# CPD COMPUTATION VIA RECURSIVE EIGENSPACE DECOMPOSITIONS

*Eric Evert, Michiel Vandecappelle, and Lieven De Lathauwer*

KU Leuven – Kulak

## ABSTRACT

The Canonical Polyadic Decomposition (CPD) is a fundamental tensor decomposition which has widespread use in signal processing due to its ability to extract component information. A popular for algorithm CPD is the generalized eigenvalue decomposition (GEVD) which is based on the generalized eigenvectors of a subpencil of a tensor. GEVD plays an important role in applications as it provides strong algebraic initializations for optimization routines for CPD computation. In fact, using GEVD initializations can improve final accuracy and reduce computation time. However, despite GEVD's success, the algorithm underperforms in some settings and exhibits pencil-based instability.

We present a recursive generalized eigenspace decomposition (GESD) for CPD computation. Rather using one subpencil, GESD combines generalized eigenspace information from many subpencils to compute a CPD. GESD is more accurate than GEVD and thereby improves the reliability of the components extracted by CPD. We also give a Cramér-Rao based analysis for the accuracy of GESD.

***Index Terms*—** tensors, canonical polyadic decomposition, eigenvalue decomposition, multilinear algebra

## 1. INTRODUCTION

Tensors, or multiindexed numerical arrays, play an important role in fields such as machine learning and signal processing [1, 2]. These higher-order generalizations of matrices allow for preservation of higher-order structure present in data and low rank decompositions of tensors allow for compression of data and recovery of underlying information [1–3]. A fundamental decomposition for tensors is the canonical polyadic decomposition (CPD) which expresses a tensor as a sum of rank one tensors. A key feature of the CPD is that, with mild assumptions [4–6], the CPD of a low rank tensor is unique. For a tensor with a unique CPD, the CPD can often be found algebraically in the noiseless setting. Algebraic solutions can be obtained with limited computation time and are often used as initializations for optimization based methods.

One highly popular algorithm for algebraic computation of a CPD is the generalized eigenvalue decomposition (GEVD) [7–9]. GEVD exploits the fact that a factor matrix of a tensor may be revealed by computing the generalized eigenvectors of a subpencil of the tensor. However, it has been observed that GEVD can underperform when tensor rank is large or when the angle between factor matrix columns is small. Also, it has recently been shown that pencil based algorithms such as GEVD are not stable [10].

The accuracy of generalized eigenvectors computed for a matrix pencil depends heavily on the separation among the generalized eigenvalues and eigenvectors of the pencil [11–13]. If the generalized eigenvalues and eigenvectors are well separated, a GEVD may be accurately computed. However, when a pair of generalized eigenvalues or eigenvectors is near parallel, the accuracy of GEVD deteriorates. As such, GEVD runs into challenges if one is unable to find a subpencil in which all generalized eigenvalues are well separated.

We present an extension of GEVD which combines information from many subpencils. The resulting generalized eigenspace decomposition (GESD) is more accurate than GEVD. The increase in accuracy comes from only computing generalized eigenvectors and eigenspaces which correspond to well separated generalized eigenvalues. The computed generalized eigenspaces are used to express the original tensor as a sum of tensors of reduced rank. This process is recursively applied until arriving at a collection of rank one tensors. Also see [14] for a different type of recursive CPD algorithm which uses successive optimization subject to a multilinear rank constraint on the residual.

Section 2 introduces the GESD algorithm. In Section 3 we examine the impact of noise on GESD. The performance of GESD on synthetic data is illustrated in Section 4.

### 1.1. Notation and definitions

Let $\mathbb{K}$ denote either $\mathbb{R}$ or $\mathbb{C}$. We denote scalars, vectors, matrices, and tensors by lower case $(a)$, bold lower case $(\mathbf{a})$, bold upper case $(\mathbf{A})$, and calligraphic script $(\mathcal{A})$, respectively. For a matrix $\mathbf{A} \in \mathbb{K}^{I \times J}$, we let $\mathbf{A}^{\mathrm{T}}$ denote the transpose of $\mathbf{A}$

**Funding:** (1) Flemish Government: This work was supported by the Fonds de la Recherche Scientifique–FNRS and the Fonds Wetenschappelijk Onderzoek–Vlaanderen under EOS Project no 30468160 (SeLMA); (2) KU Leuven Internal Funds C16/15/059 and ID-N project no 3E190402. (3) This research received funding from the Flemish Government under the "Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen" programme. (4) Work supported by Leuven Institute for Artificial Intelligence (Leuven.ai). (5) In addition, Michiel Vandecappelle was partially supported by an SB Grant (1S66617N) from the Research Foundation—Flanders (FWO).

while $\mathbf{A}^{\mathrm{H}}$ denotes the conjugate transpose of $\mathbf{A}$. If $\mathbf{A}$ is invertible, then we let $\mathbf{A}^{-\mathrm{T}}$ denote the inverse of $\mathbf{A}^{\mathrm{T}}$.

A **tensor** is a multiindexed array with entries in $\mathbb{K}$. For the sake of exposition, we restrict our attention to tensors of order three. Given nonzero vectors $\mathbf{a} \in \mathbb{K}^{I_1}, \mathbf{b} \in \mathbb{K}^{I_2}, \mathbf{c} \in \mathbb{K}^{I_3}$, let

$$\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c} \in \mathbb{K}^{I_1 \times I_2 \times I_3}$$

denote the $I_1 \times I_2 \times I_3$ tensor with $i, j, k$ entry equal to $a_i b_j c_k$. A tensor of this form is called a **rank one tensor**. The minimal integer $R$ such that $\mathcal{T} \in \mathbb{K}^{I_1 \times I_2 \times I_3}$ may be written as

$$\mathcal{T} = \sum_{r=1}^{R} \mathbf{a}_r \otimes \mathbf{b}_r \otimes \mathbf{c}_r$$

where the vectors $\mathbf{a}_r, \mathbf{b}_r, \mathbf{c}_r$ have entries in $\mathbb{K}$ is called the $\mathbb{K}$-rank of the tensor $\mathcal{T}$, and a decomposition of this form is called a **canonical polyadic decomposition** (CPD) of $\mathcal{T}$. Compactly we write $\mathcal{T} = [\![\mathbf{A}, \mathbf{B}, \mathbf{C}]\!]$ where the matrices $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$ have $\mathbf{a}_r, \mathbf{b}_r$ and $\mathbf{c}_r$ as their $r$th columns, respectively. The matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ are called the **factor matrices** for $\mathcal{T}$.

A **mode-$\ell$ fiber** of a tensor is a vector obtained by fixing all indices but the $\ell$th. The **mode-$\ell$ unfolding** $\mathbf{T}_{[\ell;j,i]}$ of $\mathcal{T}$ is the matrix obtained by stacking all mode-$\ell$ fibers of $\mathcal{T}$ as columns of a matrix, where the mode $i$ indices increment faster than the mode $j$ indices. The **$\ell$-mode** product $\mathcal{T} \cdot_\ell \mathbf{A}$ between a matrix $\mathbf{A}$ and a tensor $\mathcal{T}$ is the tensor with mode-$\ell$ unfolding equal to $\mathbf{A}\mathbf{T}_{[\ell;j,i]}$. We call the matrices $\mathbf{T}(:,:,k)$ the **frontal slices** of $\mathcal{T}$ and use $\mathbf{T}_k$ to denote $\mathbf{T}(:,:,k)$.

Given a matrix pencil $(\mathbf{T}_1, \mathbf{T}_2) \in \mathbb{K}^{R \times R \times 2}$, we say $\mathbf{z} \neq \mathbf{0} \in \mathbb{K}^R$ is a **generalized eigenvector** of $(\mathbf{T}_1, \mathbf{T}_2)$ if there exists a $\boldsymbol{\lambda} \in \mathbb{K}^2$ and a $\mathbf{q} \neq \mathbf{0} \in \mathbb{K}^R$ such that

$$\mathbf{T}_1\mathbf{z} = \lambda_1\mathbf{q} \qquad \text{and} \qquad \mathbf{T}_2\mathbf{z} = \lambda_2\mathbf{q}. \tag{1}$$

We call the subspace $\operatorname{span}((\lambda_1, \lambda_2)) \subset \mathbb{K}^2$ a **generalized eigenvalue** of $\mathcal{T}$. Given a subspace $\mathfrak{Z} \subset \mathbb{K}^R$ of dimension $\ell$ and a tensor $\mathcal{T} \in \mathbb{K}^{R \times R \times K}$, we say $\mathfrak{Z}$ is a **generalized eigenspace** of $\mathcal{T}$ if there is a subspace $\mathfrak{Q} \subset \mathbb{K}^R$ of dimension less than or equal to $\ell$ such that

$$\mathbf{T}_k\mathfrak{Z} \subseteq \mathfrak{Q} \qquad \text{for all } k = 1, \ldots, K.$$

See [11–13] for further discussion.

## 2. GESD DESCRIPTION

GESD uses generalized eigenspaces to recursively split a given tensor into subtensors of reduced rank until the subtensors all have rank one (up to noise). We compute CPD of each (noisy) rank-1 tensor using rank-1 approximations.

Before proceeding we note that generalized eigenvalue based algorithms such as GEVD and GESD require the tensor $\mathcal{T}$ has at least two factor matrices with full column rank. In particular, letting $R$ denote the rank of the signal portion of $\mathcal{T} \in \mathbb{K}^{I_1 \times I_2 \times I_3}$, one has $R \leq \min\{I_1, I_2\}$ up to a permutation of indices. For this reason, by computing a (truncated) orthogonal compression of $\mathcal{T}$ such as a multilinear singular value decomposition, see e.g. [15], we can restrict to the case where $\mathcal{T}$ has rank $R$ and size $R \times R \times K$ with $R \geq K$.

Now let $\mathcal{T}$ be an $R \times R \times K$ tensor of rank $R$ with CPD $\mathcal{T} = [\![\mathbf{A}, \mathbf{B}, \mathbf{C}]\!]$ and assume $\mathbf{A}$ and $\mathbf{B}$ are invertible. Let $\mathbf{V} \in \mathbb{K}^{K \times 2}$ be a column-wise orthonormal matrix of size $K \times 2$, and set $\mathcal{S} = \mathcal{T} \cdot_3 \mathbf{V}^{\mathrm{T}}$. The GESD algorithm begins by computing a collection of generalized eigenspaces $\mathfrak{Z}_1, \ldots, \mathfrak{Z}_L$ for the generalized eigenvalue problem

$$\mathbf{S}_i\mathbf{z} = \boldsymbol{\lambda}_i\mathbf{q} \text{ for } i = 1, 2 \qquad \operatorname{span}(\{\mathfrak{Z}_\ell\}_{\ell=1}^{L}) = \mathbb{K}^R. \tag{2}$$

To simplify exposition we take $L = 2$.

Say the generalized eigenvalues of $\mathcal{S}$ are split into two distinct groups of size $j_1$ and $j_2$ where $j_1 + j_2 = R$. The corresponding eigenspaces may then for instance be computed with a $\mathbf{QZ}$ decomposition[1] of the matrix pencil $(\mathbf{S}_1, \mathbf{S}_2)$ [16]. That is, we find unitary matrices $\mathbf{Q}, \mathbf{Z} \in \mathbb{K}^{R \times R}$ such that

$$\mathbf{S}_1 = \mathbf{Q}\mathbf{X}_1\mathbf{Z}^{\mathrm{H}} \qquad \text{and} \qquad \mathbf{S}_2 = \mathbf{Q}\mathbf{X}_2\mathbf{Z}^{\mathrm{H}} \tag{3}$$

where $\mathbf{X}_1$ and $\mathbf{X}_2$ are upper triangular matrices with the generalized eigenvalues of the pencil on the diagonal.

Using [17, Theorem 2.3 (1)], one may show that $\mathbf{Z}_1$ and $\mathbf{Z}_2$ have the form $\mathbf{Z}_i = \mathbf{B}^{-\mathrm{T}}\mathbf{E}_i$ where the $j$th row of $\mathbf{E}_1$ is nonzero if and only if the $j$th row of $\mathbf{E}_2$ is a zero row, and vice versa. Define the tensors $\mathcal{T}^1$ and $\mathcal{T}^2$ by

$$\mathcal{T}^i = \mathcal{T} \cdot_2 \mathbf{Z}_i^{\mathrm{T}} = [\![\mathbf{A}, \mathbf{Z}_i^{\mathrm{T}}\mathbf{B}, \mathbf{C}]\!] = [\![\mathbf{A}, \mathbf{E}_i^{\mathrm{T}}, \mathbf{C}]\!]$$

and for $i = 1, 2$ set $R_i = \operatorname{rank}(\mathcal{T}^i)$. Since the $\mathbf{E}_i^{\mathrm{T}}$ have zero columns corresponding to the zero rows of $\mathbf{E}_i$, we have that $R_1 + R_2 = \operatorname{rank}(\mathcal{T})$. Moreover, suppose $\mathcal{T}^i$ has CPD $[\![\mathbf{A}_i, \check{\mathbf{E}}_i, \mathbf{C}_i]\!]$ for $i = 1, 2$. Here the matrix $\check{\mathbf{E}}_i \in \mathbb{K}^{R \times R_i}$ is the matrix obtained by removing the zero columns of $\mathbf{E}_i^{\mathrm{T}}$. Then up to permutation and scaling, one has $\mathbf{A} = (\mathbf{A}_1 \ \mathbf{A}_2)$. Similarly, the factor matrix $\mathbf{C}$ of $\mathcal{T}$ has the form $\mathbf{C} = (\mathbf{C}_1 \ \mathbf{C}_2)$ up to scaling and the same permutation as $\mathbf{A}$.

This procedure partitions the rank one terms of $\mathcal{T}$ into two distinct collections so that one collection gives a decomposition for the tensor $\mathcal{T}_1$ while the other collection gives a decomposition for $\mathcal{T}_2$. The number of rank one terms in each collection, i.e., the rank of $\mathcal{T}_1$ and of $\mathcal{T}_2$, is determined by the number of generalized eigenvalues in each cluster.

This method may be recursively applied to each $\mathcal{T}^i$ to reduce their ranks until one obtains a collection of tensors all having rank one (up to noise). The columns of the factor matrices $\mathbf{A}$ and $\mathbf{C}$ may then be approximated by computing rank one approximations for each of these (approximately) rank one tensors, as discussed below. Having computed $\mathbf{A}$ and $\mathbf{C}$, we may compute $\mathbf{B}$ by solving the set of linear equations

$$\mathbf{T}_{[2;3,1]} = \mathbf{B}(\mathbf{C} \odot \mathbf{A})^{\mathrm{T}}.$$

[1] The $\mathbf{QZ}$ decomposition is the generalization of the Schur decomposition to matrix pencils, e.g., see [11, Chapter VI, Theorem 1.9].

Now consider the case where $\mathbf{Z}_1 \in \mathbb{K}^{R\times 1}$ spans an eigenspace of dimension one. In this case $\mathbf{Z}_1$ has the form

$$\mathbf{Z}_1 = \beta \mathbf{B}^{-\mathrm{T}} \mathbf{e}_\ell^{\mathrm{T}}$$

where $\mathbf{e}_\ell$ is a standard basis vector and $\beta \in \mathbb{K}$. Thus we have

$$\mathcal{T} \cdot_2 \mathbf{Z}_1^{\mathrm{T}} = [\![\mathbf{A}, \beta \mathbf{e}_\ell, \mathbf{C}]\!] = [\![\mathbf{a}_\ell, \beta, \mathbf{c}_\ell]\!] .$$

We may compute $\mathbf{a}_\ell$ and $\mathbf{c}_\ell$ by computing a rank one approximation of the matrix $\mathcal{T} \cdot_2 \mathbf{Z}_1^{\mathrm{T}}$. One may verify that $\mathcal{T} \cdot_2 \mathbf{Z}_1^{\mathrm{T}}$ has rank one (up to noise) by computing its singular values.

## 3. A PROBABILISTIC BOUND FOR GESD

We now examine the impact of noise on GESD. Our goal is to provide a probabilistic bound for the existence of a subpencil of the tensor of interest which has two distinct clusters of generalized eigenvalues that have not coalesced.

Suppose $\mathcal{T} \in \mathbb{K}^{R\times R\times 2}$ is a matrix pencil with distinct generalized eigenvalues. For sufficiently small perturbations $\mathcal{N}$, the generalized eigenvalues and generalized eigenvectors of $\mathcal{T} + \mathcal{N}$ are continuous in $\mathcal{N}$. However, if the perturbation $\mathcal{N}$ causes the matrix pencil $\mathcal{T} + \mathcal{N}$ to have a repeated generalized eigenvalue, i.e., causes at least two generalized eigenvalues of $\mathcal{T} + \mathcal{N}$ to coalesce, then the path of generalized eigenvectors can no longer be unambiguously continued [18].

Despite this, the path of the corresponding generalized eigenspaces is still continuous and unambiguous in $\mathcal{N}$ so long as the underlying generalized eigenvalues do not coalesce with other generalized eigenvalues. Thus guaranteeing the existence of two clusters of generalized eigenvalues which cannot have coalesced guarantees that $\mathcal{T}+\mathcal{N}$ has two generalized eigenspaces which are still continuous in the perturbation $\mathcal{N}$ and are therefore closely related to the original generalized eigenspaces of $\mathcal{T}$. See [13, 18] for further discussion.

### 3.1. A Cramér-Rao based bound for eigenvalue clusters

Our basic strategy to obtain a probabilistic bound is to use the Cramér-Rao induced bound (CRIB) for the mean-square angular deviation of the vectors of a factor matrix as described in [19] to build confidence intervals for the generalized eigenvalues of a matrix pencil. These confidence intervals can be used to say that clusters of generalized eigenvalues have not coalesced with high probability, i.e., that the separation of the corresponding eigenspaces is well justified in GESD.

Let $\mathcal{T} = [\![\mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \in \mathbb{K}^{R\times R\times K}$ be a rank $R$ tensor and let $\mathcal{N}(\sigma)$ be a random Gaussian noise tensor whose independent identically distributed (i.i.d.) entries have variance $\sigma^2$. Assume that $\mathcal{T} + \mathcal{N}(\sigma)$ has a best rank $R$ approximation denoted by $\hat{\mathcal{T}} = [\![\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}]\!]$ where $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ are invertible. Let $\mathbf{c}_k$ and $\hat{\mathbf{c}}_k$ denote the $k$th columns of the factor matrices $\mathbf{C}$ and $\hat{\mathbf{C}}$, respectively. The CRIB for a column $\mathbf{c}_k$ of the factor matrix $\mathbf{C}$ computes a lower bound for the variance of the angle between $\mathbf{c}_k$ and $\hat{\mathbf{c}}_k$. For details on theory and computation of the CRIB we refer the reader to [19, 20]. For general discussion of Cramér-Rao bounds for tensors, also see [21].

If $\mathcal{T}$ has size $R \times R \times 2$, then the generalized eigenvalues of the pencil $(\mathbf{T}_1, \mathbf{T}_2)$ are given by the spans of the columns of the factor matrix $\mathbf{C}$. Therefore, we may directly apply the CRIB to a matrix pencil to get a lower bound on the variance of the angles of the pencil's generalized eigenvalues.

Let $\mathfrak{L}_1, \ldots, \mathfrak{L}_R$ and $\hat{\mathfrak{L}}_1, \ldots, \hat{\mathfrak{L}}_R$ be the generalized eigenvalues of the pencils $\mathcal{T}$ and $\mathcal{T} + \mathcal{N}(\sigma) \in \mathbb{K}^{R\times R\times 2}$, respectively. Assuming that the angular deviation between $\mathfrak{L}_r$ and $\hat{\mathfrak{L}}_r$ is normally distributed for each $r$, for a given variance $\sigma^2$ and probability $p$ we may then use the CRIB to determine intervals $I_r(\sigma)$ around each $\mathfrak{L}_r$ such that $\hat{\mathfrak{L}}_r \in I_r(\sigma)$ for all $r$ with probability bounded above by $p$.

Let $\sigma_*$ be the supremum over $\sigma$ such that the set $\cup_r I_r(\sigma_*)$ is not connected. Then $\sigma_*^2$ gives an approximation for the variance of the entries of $\mathcal{N}$ such that with probability $p$ the generalized eigenvalues of the pencil $(\hat{\mathbf{T}}_1, \hat{\mathbf{T}}_2)$ may be divided into two distinct clusters which have not coalesced. Thus, in practice, if $\sigma \leq \sigma_*$, then the factors GESD recovers from $\mathcal{T} + \mathcal{N}(\sigma)$ are closely related to the factors of $\mathcal{T}$.

For use in upcoming experiments, given $\sigma_*$ we may approximate the expected value of the Frobenius norm of the noise tensor $\mathcal{N}(\sigma_*) \in \mathbb{K}^{R\times R\times K}$ as $\mathbb{E}(\|\mathcal{N}(\sigma_*)\|_{\mathrm{F}}) \approx \sigma_* \sqrt{R^2 K}$. We call the bound $\mathbb{E}(\|\mathcal{N}(\sigma_*)\|_{\mathrm{F}})$ as computed for a given $p$ the **Cramér-Rao existence bound (CREB)**.

## 4. NUMERICAL EXPERIMENTS

We now compare GESD to GEVD in a direction-of-arrival (DOA) retrieval [22] experiment that considers the case of line-of-sight signals impinging on a uniform rectangular array (URA) as the CPD can be applied to find the DOAs in this case [23, 24]. We consider a URA with $M \times M$ sensors, with $M = 20$, collecting $K = 20$ samples from $R = 8$ sources in the far field with as azimuths [5 20 25 30 35 40 55 60] and elevations [5 10 20 25 30 35 40 45] degrees, respectively. The omnidirectional sensors are evenly spaced with $\Delta$ as inter-sensor spacing. The collected azimuths $\mathbf{Z} \in \mathbb{C}^{M\times R}$ and elevations $\mathbf{L} \in \mathbb{C}^{M\times R}$ of the $R$ sources at each sensor in the array generate an observed tensor $\mathcal{T} \in \mathbb{C}^{M\times M\times K}$, for which the $k$th frontal slice admits a low rank decomposition $\mathcal{T} = [\![\mathbf{A}^{(k)}, \mathbf{E}^{(k)}, \mathbf{s}^{(k)}]\!]$. The matrix $\mathbf{A}^{(k)} \in \mathbb{C}^{M\times R}$ has entries $a_{mr}^{(k)} = \exp((m-1)2\pi/\lambda \sin(z_{rk}\pi/180)\Delta i)$. Meanwhile, the matrix $\mathbf{E}^{(k)} \in \mathbb{C}^{M\times R}$ has entries $e_{mr}^{(k)} = \exp((m-1)2\pi/\lambda \sin(l_{rk}\pi/180)\Delta i)$. The vector $\mathbf{s}^{(k)} \in \mathbb{C}^R$ contains the sources and $\lambda$ is the signal wavelength. A rank-$R$ CPD of $\mathcal{T}$ can be computed and the azimuths and elevations of the sources can be recovered from the mode-1 and mode-2 factor matrices of the CPD, respectively.

We perturb the tensor $\mathcal{T}$ with Gaussian distributed noise in the range $[-20, 40]$ dB and compute CPDs with both the

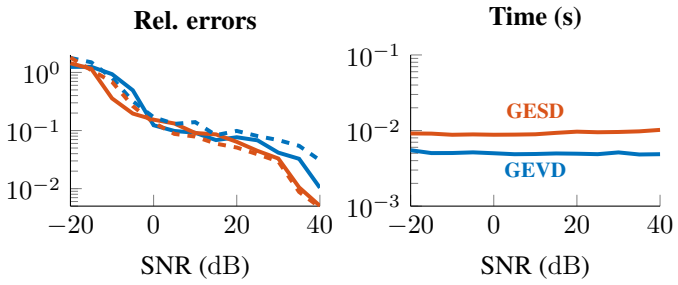


**Fig. 1**. GESD estimates the azimuths and elevations of the sources more accurately than GEVD, and is only slightly slower. Left: mean relative errors over the eight sources of the estimated azimuths (——) and elevations (- - -) for GEVD and the estimated azimuths (——) and elevations (- - -) for GESD. Right: computation time for both methods.

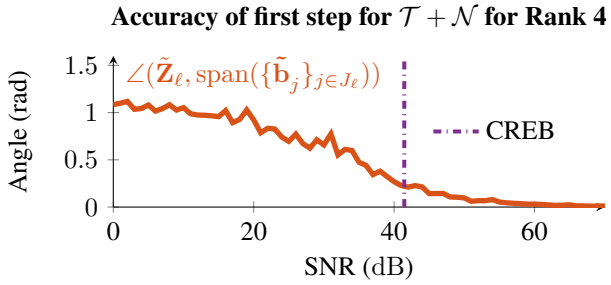


**Fig. 2**. Angle between $\tilde{\mathbf{Z}}_\ell$ and columns of $\mathbf{B}^{-\mathrm{T}}$ vs. the CREB of Section 3 (computed for $p = 0.9$). Here $\mathcal{T}$ has rank-4 and size $4 \times 4 \times 4$. The SNR is varied from 0 dB to 70 dB.

GEVD and the GESD methods. In Figure 1, the median error (left) and computation time (right) over 50 trials is shown for the estimation of the azimuths and elevations of the eight sources. The GESD method is more accurate than the GEVD method, especially for the estimation of the elevations of the sources, while being only slightly slower.

### 4.1. Accuracy of GESD versus the CREB

We now compare the CREB against the accuracy of the first step of GESD. Let $\mathcal{T} \in \mathbb{R}^{R\times R\times K}$ be a real rank $R$ tensor and let $\mathcal{N}$ be a noise tensor. As previously discussed, given $\mathcal{T}+\mathcal{N}$ as an input, the goal of the first step of GESD is to decompose $\mathbb{R}^R$ into (at least) two disjoint JGE spaces of $\mathcal{T}$.

Let $\tilde{\mathbf{Z}}_1$ and $\tilde{\mathbf{Z}}_2$ be the matrices computed by the first step of GESD. Then there should exist disjoint index sets $J_1, J_2 \subset \{1,\dots,R\}$ such that $J_1 \cup J_2 = \{1,\dots,R\}$ and such that

$$\operatorname{ran} \tilde{\mathbf{Z}}_\ell \approx \operatorname{span}(\{\tilde{\mathbf{b}}_j\}_{j\in J_\ell}) \qquad \text{for } \ell = 1, 2.$$

Here the $\tilde{\mathbf{b}}_\mathbf{j}$ are columns of the matrix $\mathbf{B}^{-\mathrm{T}}$.

In Figure 2, we plot the median (over 50 trials) of the aforementioned angle as a function of the SNR of $\mathcal{T}$ to $\mathcal{N}$ where $\mathcal{N}$ is randomly generated Gaussian noise. In each trial we compute subspaces $\tilde{\mathbf{Z}}_1$ and $\tilde{\mathbf{Z}}_2$ for the tensor $\mathcal{T}+\mathcal{N}$ using GESD. We check accuracy by plotting median of the CREB and of the angle described above. Here, $\mathcal{T}$ was generated by independently sampling all entries of its factor matrices uniformly on $[0, 1]$ and forming the associated low-rank tensor.

Our experiment shows that the CREB performs well in predicting errors in GESD. For emphasis, this experiment is performed on the core of the tensor $\mathcal{T}$, a setting which is particularly challenging. The SNR for the CREB depends heavily on the dimensions of $\mathcal{T}$, and in the setting of rank less than dimensions common in signal processing applications, the SNR of the CREB is expected to be much lower. Thus, the focus of this experiment is that the CREB is predictive of errors rather than the specific SNR of the CREB.

Both of our experiments are performed on a machine with an Intel Core i7-6820HQ CPU at 2.70GHz and 16GB of RAM using MATLAB R2020b and Tensorlab 3.0 [25].

## 5. CONCLUSION

We presented the algebraic GESD algorithm for CPD computation which uses multiple subpencils of a given tensor to compute a CPD. GESD combines eigenspace computations from many subpencils of a tensor to compute a CPD, thus allowing one to only perform trustworthy computations in each pencil. In numerical experiments we showed that GESD outperforms GEVD for direction-of-arrival estimation in terms of accuracy and that GESD takes only slightly longer to compute these estimates. In addition, we performed a Cramér-Rao based analysis for the accuracy of GESD.

A direction for future work is that for a very high rank tensor, one cannot in general expect to find a subpencil which has clusters of well-separated generalized eigenvalues. Since the number of generalized eigenvalues of a subpencil of a tensor is equal to the rank of a tensor and since generalized eigenvalues are subspaces of $\mathbb{K}^2$, the expected separation between generalized eigenvalues becomes small as $R$ becomes large.

In this case, one can weaken our assumptions to require only partial separation of eigenvalue clusters. E.g., one can form four clusters of generalized eigenvalues $C_1, C_2, C_3, C_4$ where the clusters are chosen so that all eigenvalues in $C_1$ are well separated from eigenvalues of $C_3$, i.e., eigenvalues in $C_1$ can only be close to eigenvalues of $C_2$ or $C_4$, and this can only happen near the border between $C_1$ and $C_2$ (and $C_1$ and $C_4$, respectively). If one computes corresponding generalized eigenspaces $\mathfrak{E}_1, \mathfrak{E}_2, \mathfrak{E}_3, \mathfrak{E}_4$, then there can be no interference between $\mathfrak{E}_1$ and $\mathfrak{E}_3$ since the eigenpairs corresponding to $C_1$ have been well separated from eigenpairs that correspond to $C_3$. In addition, interference between $\mathfrak{E}_1$ and $\mathfrak{E}_2$ or $\mathfrak{E}_1$ and $\mathfrak{E}_4$ can only occur at parts of these eigenspaces corresponding to eigenvalues near the boundaries of the clusters.

## 6. REFERENCES


[1] A. Cichocki, D. P. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. F. Caiafa, and A.-H. Phan, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 145–163, 2015.

[2] N. 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, 2017.

[3] P. Comon and C. Jutten, *Handbook of blind source separation: Independent component analysis and applications*, Academic press, 2009.

[4] I. Domanov and L. De Lathauwer, "On the uniqueness of the canonical polyadic decomposition of third-order tensors — Part I: Basic results and uniqueness of one factor matrix," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 3, pp. 855–875, 2013.

[5] J. B. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Linear Algebra Appl.*, vol. 18, pp. 95–138, 1977.

[6] N. D. Sidiropoulos and R. Bro, "On the uniqueness of multilinear decomposition of N-way arrays," *J. Chemometrics*, vol. 14, no. 3, pp. 229–239, 2000.

[7] I. Domanov and L. De Lathauwer, "Canonical polyadic decomposition of third-order tensors: Relaxed uniqueness conditions and algebraic algorithm," *Linear Algebra and its Applications*, vol. 513, pp. 342–375, 2017.

[8] E. Sanchez and B.R. Kowalski, "Tensorial resolution: A direct trilinear decomposition," *J. Chemometrics*, vol. 4, no. 1, pp. 29–45, 1990.

[9] S. E. Leurgans, R. T. Ross, and R. B. Abel, "A decomposition for three-way arrays," *SIAM J. Matrix Anal. Appl.*, vol. 14, no. 4, pp. 1064–1083, 1993.

[10] C. Beltrán, P. Breiding, and N. Vannieuwenhoven, "Pencil-based algorithms for tensor rank decomposition are not stable," *SIAM J. Matrix Anal. Appl.*, vol. 40, no. 2, pp. 739–773, 2019.

[11] G. W. Stewart and J.G. Sun, *Matrix Perturbation Theory*, Academic Press, 1990.

[12] G.H. Golub and C.F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, 1996.

[13] J. Demmel and B. Kågström, "Computing stable eigendecompositions of matrix pencils," *Linear Algebra and its Applications*, vol. 88-89, pp. 139 – 186, 1987.

[14] A. Phan, P. Tichavský, and A. Cichocki, "Tensor deflation for CANDECOMP/PARAFAC— Part I: Alternating subspace update algorithm," *IEEE Trans. Signal Process.*, vol. 63, no. 22, pp. 5924–5938, 2015.

[15] 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.

[16] D. Camps, *Pole Swapping Methods for the Eigenvalue Problem—Rational QR Algorithms*, Ph.D. thesis, KU Leuven, 2019.

[17] E. Evert and L. De Lathauwer, "Guarantees for existence of a best canonical polyadic approximation of a noisy low-rank tensor," *Technical Report 19–71, ESAT-STADIUS, KU Leuven, Leuven, Belgium*, 2021, (To appear in SIAM J. Matrix Anal. Appl.).

[18] T. Kato, *Perturbation Theory for Linear Operators*, Springer, 1995.

[19] P. Tichavský and Z. Koldovský, "Weight adjusted tensor method for blind seperation of underdetermined mixtures of nonstationary sources," *IEEE Trans. on Signal Process.*, vol. 59, pp. 1037–1047, 2011.

[20] P. Tichavský, A.-H. Phan, and Z. Koldovský, "Cramér–Rao-induced bounds for CANDECOMP/PARAFAC tensor decomposition," *IEEE Trans. Signal Process.*, vol. 61, no. 8, pp. 1986–1997, 2013.

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

[22] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE signal processing magazine*, vol. 13, no. 4, pp. 67–94, 1996.

[23] N. D. Sidiropoulos, R. Bro, and G. B. Giannakis, "Parallel factor analysis in sensor array processing," *IEEE transactions on Signal Processing*, vol. 48, no. 8, pp. 2377–2388, 2000.

[24] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 984–995, 1989.

[25] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer, "Tensorlab 3.0," Mar. 2016.