Tensor Decompositions with Banded Matrix Factors

Mikael Sorensen, Pierre Comon

Tensor Decompositions with Banded Matrix Factors

Mikael Sørensen[a], Pierre Comon[b]

[a]K.U.Leuven - E.E. Dept. (ESAT) - SCD-SISTA, Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, and Group Science, Engineering and Technology, K.U.Leuven Campus Kortrijk, E. Sabbelaan 53, 8500 Kortrijk, Belgium, Mikael.Sorensen@kuleuven-kortrijk.be.
[b]I3S, CNRS, University of Nice, B.P. 121, 06903, Sophia Antipolis, and INRIA, GALAAD, B.P. 93, 06902, Sophia Antipolis, France, p.comon@ieee.org.

Abstract

The computation of the model parameters of a Canonical Polyadic Decomposition (CPD), also known as the parallel factor (PARAFAC) or canonical decomposition (CANDECOMP) or CP decomposition, is typically done by resorting to iterative algorithms, e.g. either iterative alternating least squares type or descent methods. In many practical problems involving tensor decompositions such as signal processing, some of the matrix factors are banded. First, we develop methods for the computation of CPDs with one banded matrix factor. It results in best rank-1 tensor approximation problems. Second, we propose methods to compute CPDs with more than one banded matrix factor. Third, we extend the developed methods to also handle banded and structured matrix factors such as Hankel or Toeplitz. Computer results are also reported.

Keywords: Tensors, Polyadic decomposition, canonical decomposition (CANDECOMP), parallel factor (PARAFAC), CP decomposition, Banded matrices, Hankel matrices, Toeplitz matrices.

1. Introduction

Tensor decompositions with banded matrix factors and possibly also (block-) Toeplitz/Hankel structures have found application in signal pro-

Email addresses: `Mikael.Sorensen@kuleuven-kortrijk.be` (Mikael Sørensen), `p.comon@ieee.org` (Pierre Comon)

Preprint submitted to Linear Algebra and Its Applications

cessing, in particular in cumulant based blind identification of convolutive mixtures [5], [22], [2], [3], [12], [13] and in tensor based blind equalization of wireless communication systems [10], [20], [1]. In this case the banded structure of the matrix factors are directly related to the filter orders of the given system we attempt to identify.

It also occurs in blind separation of communication signals [25] and in computational multilinear algebra [9] involving triangular structured matrix factors. Hence, the development of methods to deal with CPDs with banded and possibly also structured matrix factors such as Hankel or Toeplitz must be considered important.

Recently an approach to address structured CPDs was proposed in [7]. However, it was mainly concerned with deriving necessary uniqueness conditions. In this paper we limit the discussion to tensors with banded matrix factors, but also provide sufficient uniqueness conditions for CPDs with banded and possibly also structured matrix factors such as Hankel or Toeplitz. Second, we develop numerical methods for the computation of CPDs with banded and possibly also structured matrix factors.

The results presented in this paper are also valid for more general tensor decompositions with banded matrix factors such as the Tucker decomposition [24] or the family of block tensor decompositions [11].

The paper is organised as follows. In the rest of the introduction, the applied notation is presented followed by a brief review of the CPD. Section 3 provides uniqueness results for CPDs with banded matrix factors while section 4 present numerical methods to compute them. Next, in section 5 we provide uniqueness results for CPDs with banded and structured matrix factors while section 6 describes numerical methods to compute them. In section 7, some numerical experiments are reported. A concluding section eventually ends the paper.

2. Notation

Vectors, matrices and tensors are denoted by lower case boldface, upper case boldface and upper case calligraphic letters, respectively. The symbol $\otimes$ denotes the Kronecker product

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots \\ \vdots & \vdots & \ddots \end{bmatrix},$$

where $(\mathbf{A})_{mn} = a_{mn}$. The symbol $\odot$ denotes the Khatri-Rao product

$$\mathbf{A} \odot \mathbf{B} = \left[\begin{array}{ccc} \mathbf{a}_1 \otimes \mathbf{b}_1 & \mathbf{a}_2 \otimes \mathbf{b}_2 & \dots \end{array} \right],$$

where $\mathbf{a}_r$ and $\mathbf{b}_r$ denote the rth column vector of $\mathbf{A}$ and $\mathbf{B}$, respectively. Let $\circ$ denote the outer product of N vectors $\mathbf{a}^{(n)} \in \mathbb{C}^{I_n}$ such that

$$\left(\mathbf{a}^{(1)} \circ \mathbf{a}^{(2)} \circ \cdots \circ \mathbf{a}^{(N)}\right)_{i_1 i_2 \dots i_N} = a^{(1)}_{i_1} a^{(2)}_{i_2} \cdots a^{(N)}_{i_N}.$$

Further, $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$, $(\cdot)^\dagger$, $\|\cdot\|_F$, $\mathrm{Col}(\cdot)$ and $\mathrm{Row}(\cdot)$ denote the transpose, conjugate, conjugate-transpose, one-sided inverse, Frobenius norm, column space and row space of a matrix, respectively.

Let $\mathbf{A} \in \mathbb{C}^{I \times J}$, then $\mathrm{vec}(\mathbf{A}) \in \mathbb{C}^{IJ}$ denotes the column vector defined by $(\mathrm{vec}(\mathbf{A}))_{i+(j-1)I} = (\mathbf{A})_{ij}$. Given $\mathbf{a} \in \mathbb{C}^{IJ}$, then the reverse operation is $\mathrm{Unvec}(\mathbf{a}) = \mathbf{A} \in \mathbb{C}^{I \times J}$ such that $(\mathbf{a})_{i+(j-1)I} = (\mathbf{A})_{ij}$.

Matlab index notation will be used to denote submatrices of a given matrix. For example, $\mathbf{A}(1:k,:)$ denotes the submatrix of $\mathbf{A}$ consisting of the rows from 1 to k. The anti-identity matrix is denoted by $\mathbf{J}_R \in \mathbb{C}^{R \times R}$ and it is equal to

$$\mathbf{J}_R = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ \vdots & \iddots & \iddots & 0 \\ 0 & 1 & \iddots & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{C}^{R \times R}.$$

Let $\mathbf{a}$ be a column vector, then we say that $\mathbf{a}$ is lower bounded by $\mathrm{b}_L(\mathbf{a})$ if its entries satisfy $a_{\mathrm{b}_L(\mathbf{a})} \neq 0$ and $a_{\mathrm{b}_L(\mathbf{a})+n} = 0$, $\forall n > 0$. Similarly, we say that $\mathbf{a}$ is upper bounded by $\mathrm{b}_U(\mathbf{a})$ if its entries satisfy $a_{\mathrm{b}_U(\mathbf{a})} \neq 0$ and $a_{\mathrm{b}_U(\mathbf{a})-n} = 0$, $\forall n > 0$.

We say that a matrix $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_R] \in \mathbb{C}^{I \times R}$ is banded if $\mathrm{b}_U(\mathbf{a}_m) \neq \mathrm{b}_U(\mathbf{a}_n)$ and $\mathrm{b}_U(\mathbf{a}_m) > 1$, or $\mathrm{b}_L(\mathbf{a}_m) \neq \mathrm{b}_L(\mathbf{a}_n)$ and $\mathrm{b}_L(\mathbf{a}_m) < I$, for some $m, n \in [1, R]$. Furthermore, we say that $\mathbf{A}$ is lower banded if $\mathrm{b}_U(\mathbf{a}_1) < \cdots < \mathrm{b}_U(\mathbf{a}_R)$ and anti-lower banded if $\mathrm{b}_U(\mathbf{a}_1) > \cdots > \mathrm{b}_U(\mathbf{a}_R)$. Similarly, we say that $\mathbf{A}$ is upper banded if $\mathrm{b}_L(\mathbf{a}_1) < \cdots < \mathrm{b}_L(\mathbf{a}_R)$ and anti-upper banded if $\mathrm{b}_L(\mathbf{a}_1) > \cdots > \mathrm{b}_L(\mathbf{a}_R)$. Note that this definition of banded matrices is less restrictive than the standard definitions such as in [14].

2.1. Canonical Polyadic Decomposition (CPD)

An Nth order rank-1 tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ is defined as the outer product of some non-zero vectors $\mathbf{a}^{(n)} \in \mathbb{C}^{I_n}$, $n \in [1, N]$, such that $\mathcal{X}_{i_1 \ldots i_N} = \prod_{n=1}^{N} a_{i_n}^{(n)}$. The rank of a tensor $\mathcal{X}$ is equal to the minimal number of rank-1 tensors that yield $\mathcal{X}$ in a linear combination. Assume that the rank of $\mathcal{X}$ is R, then it can be written as

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{a}_r^{(1)} \circ \cdots \circ \mathbf{a}_r^{(N)}, \tag{1}$$

where $\mathbf{a}_r^{(n)} \in \mathbb{C}^{I_n}$. The decomposition (1) is sometimes referred to as the polyadic decomposition [16], PARAFAC decomposition [15], CANDECOMP decomposition [4] or CP decomposition [17], [23], [18], [6]. In this paper it will be referred to as the Canonical Polyadic Decomposition (CPD) of $\mathcal{X}$. Let us stack the vectors $\{\mathbf{a}_r^{(n)}\}$ into the matrices

$$\mathbf{A}^{(n)} = \begin{bmatrix} \mathbf{a}_1^{(n)}, & \cdots & , \mathbf{a}_R^{(n)} \end{bmatrix} \in \mathbb{C}^{I_n \times R}, \quad n \in [1, N]. \tag{2}$$

The matrices $\mathbf{A}^{(n)}$ in (2) will be referred to as the matrix factors of the CPD of the tensor $\mathcal{X}$ in (1). Furthermore, we say that the tensor $\mathcal{X}$ is partially symmetric if $\mathbf{A}^{(m)} = \mathbf{A}^{(n)}$ for some $m \neq n$. Similarly, we say that the tensor $\mathcal{X}$ is partially Hermitian symmetric if $\mathbf{A}^{(m)} = \mathbf{A}^{(n)*}$ for some $m \neq n$.

The following two matrix representations of the CPD of an Nth order tensor will be used throughout the paper. First, consider $\mathcal{T} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ with rank R and let the row vectors $\mathbf{t}^{(i_1, \ldots, i_{p-1}, i_{p+1}, \ldots, i_N)} \in \mathbb{C}^{1 \times I_p}$, where $p \in [1, N]$, be constructed such that $\mathbf{t}_{i_p}^{(i_1, \ldots, i_{p-1}, i_{p+1}, \ldots, i_N)} = \mathcal{T}_{i_1, \ldots, i_N}$, then

$$\mathbf{t}^{(i_1, \ldots, i_{P-1}, i_{P+1}, \ldots, i_N)} = \sum_{i_P=1}^{I_P} \mathcal{T}_{i_1 \ldots i_N} \mathbf{e}_{i_P}^{(I_P)T} = \sum_{r=1}^{R} \prod_{\substack{n=1 \\ n \neq P}}^{N} a_{i_n r}^{(n)} \mathbf{a}_r^{(P)T}$$

where $a_{i_n r}^{(n)} = \mathbf{A}^{(n)}(i_n, r)$ and $\mathbf{e}_{i_P}^{(I_P)} \in \mathbb{C}^{I_P}$ is a unit vector with unit element at entry i_P and zero elsewhere. Stack the vectors $\{\mathbf{t}^{(i_1, \ldots, i_{P-1}, i_{P+1}, \ldots, i_N)}\}$ into the matrix

$$\mathbb{C}^{\prod_{n=1, n \neq P}^{N} I_n \times I_P} \ni \mathbf{T}_{[P]} \triangleq \begin{bmatrix} \mathbf{t}^{(1, \ldots, 1, 1, \ldots, 1)} \\ \mathbf{t}^{(1, \ldots, 1, 1, \ldots, 2)} \\ \vdots \\ \mathbf{t}^{(I_1, \ldots, I_{P-1}, I_{P+1}, \ldots, I_N)} \end{bmatrix} = \left(\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(P-1)} \odot \mathbf{A}^{(P+1)} \odot \cdots \odot \mathbf{A}^{(N)} \right) \mathbf{A}^{(P)T}.$$

Second, we will also apply the matrix representation

$$\mathbb{C}^{\prod_{m=1}^{P} I_m \times \prod_{n=P+1}^{N} I_n} \ni \mathbf{T}^{[P]} \triangleq \begin{bmatrix} \mathcal{T}_{1,\ldots,1,1,\ldots,1} & \mathcal{T}_{1,\ldots,1,1,\ldots,2} & \cdots & \mathcal{T}_{1,\ldots,1,I_{P+1},\ldots,I_N} \\ \mathcal{T}_{1,\ldots,2,1,\ldots,1} & \mathcal{T}_{1,\ldots,2,1,\ldots,2} & \cdots & \mathcal{T}_{1,\ldots,2,I_{P+1},\ldots,I_N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{T}_{I_1,\ldots,I_P,1,\ldots,1} & \mathcal{T}_{I_1,\ldots,I_P,1,\ldots,2} & \cdots & \mathcal{T}_{I_1,\ldots,I_P,I_{P+1},\ldots,I_N} \end{bmatrix}$$
$$= \left(\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(P)}\right)\left(\mathbf{A}^{(P+1)} \odot \cdots \odot \mathbf{A}^{(N)}\right)^T.$$

The CPD of $\mathcal{X}$ in (1) is said to be essentially unique if all the N-tuplets $\left(\overline{\mathbf{A}}^{(1)}, \ldots, \overline{\mathbf{A}}^{(N)}\right)$ satisfying (1) are related via

$$\overline{\mathbf{A}}^{(n)} = \mathbf{A}^{(n)} \Pi \Delta_{\mathbf{A}^{(n)}}, \quad \forall n \in [1, N],$$

where Π is a permutation matrix and $\{\Delta_{\mathbf{A}^{(n)}}\}$ are diagonal matrices satisfying $\prod_{n=1}^{N} \Delta_{\mathbf{A}^{(n)}} = \mathbf{I}_R$ [19].

If the matrix factor $\mathbf{A}^{(P)}$ is known up to a permutation and scaling of its column vectors and has full column rank, then the other matrix factors follow in an essentially unique manner from the decoupled rank-1 tensor approximation problems [7]:

$$\mathbf{F} = \mathbf{T}_{[P]} \left(\mathbf{A}^{(P)T}\right)^{\dagger} = \mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(P-1)} \odot \mathbf{A}^{(P+1)} \odot \cdots \odot \mathbf{A}^{(N)}. \tag{3}$$

Indeed, the Khatri-Rao product structure of (3) indicates that each column vector of $\mathbf{F}$ corresponds to a vectorized version of a rank-1 tensor.

Similarly, assume that the matrix factors $\{\mathbf{A}^{(n)}\}_{n=1}^{P}$ are known up to a permutation and scaling of their column vectors and the matrix $\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(P)}$ has full column rank, then the remaining matrix factors follow in an essentially unique manner from the decoupled rank-1 tensor approximation problems

$$\mathbf{G} = \left(\left(\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(P)}\right)^{\dagger} \mathbf{T}^{[P]}\right)^T = \mathbf{A}^{(P+1)} \odot \cdots \odot \mathbf{A}^{(N)}. \tag{4}$$

Again, the Khatri-Rao product structure of (4) indicates that each column vector of $\mathbf{G}$ corresponds to a vectorized version of a rank-1 tensor.

To summarize, if some of the matrix factors of a CPD are known and the matrix of Khatri-Rao products of them is a full column rank, then the remaining matrix factors for this CPD follow from decoupled rank-1 tensor approximation problems. This is the foundation of the numerical methods presented in this paper.

The problem of solving best rank-1 tensor approximation problems is well-posed. Furthermore, when $N = P + 1$ or $N = P + 2$, then the problem reduces to linear system solving and best rank-1 matrix approximation problems, respectively, which can be solved by standard numerical linear algebra methods. This is in contrast to iterative methods such as Alternating Least Squares (ALS) which can be prone to local minima and slow convergence.

3. Uniqueness of CPDs with Banded Matrix Factors

In this section we present a uniqueness result for Nth order CPDs exploiting that $P < N$ of the involved matrix factors are banded matrices. More precisely, it allows us to identify the banded matrix factors from $\mathrm{Col}\left(\mathbf{T}^{[P]}\right)$.

Proposition 3.1. *Consider $\mathcal{T} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ with rank R and matrix representation $\boldsymbol{T}^{[P]} = \left(\boldsymbol{A}^{(1)} \odot \cdots \odot \boldsymbol{A}^{(P)}\right)\left(\boldsymbol{A}^{(P+1)} \odot \cdots \boldsymbol{A}^{(N)}\right)^T$ where $P < N$. Assume that the matrices $\boldsymbol{A}^{(1)} \odot \cdots \odot \boldsymbol{A}^{(P)}$ and $\boldsymbol{A}^{(P+1)} \odot \cdots \odot \boldsymbol{A}^{(N)}$ have full column rank and the matrix $\boldsymbol{A}^{(1)} \odot \cdots \odot \boldsymbol{A}^{(P)}$ satisfies*

$$1 \leq b_U\left(\boldsymbol{a}_1^{(1)} \otimes \cdots \otimes \boldsymbol{a}_1^{(P)}\right) < \cdots < b_U\left(\boldsymbol{a}_R^{(1)} \otimes \cdots \otimes \boldsymbol{a}_R^{(P)}\right), \tag{5}$$

$$b_L\left(\boldsymbol{a}_1^{(1)} \otimes \cdots \otimes \boldsymbol{a}_1^{(P)}\right) < \cdots < b_L\left(\boldsymbol{a}_R^{(1)} \otimes \cdots \otimes \boldsymbol{a}_R^{(P)}\right) \leq \prod_{n=1}^{P} I_n \tag{6}$$

up to a permutation of its rows and columns, i.e., there exist permutation matrices $\Pi_1 \in \mathbb{C}^{\prod_{n=1}^{P} \times \prod_{n=1}^{P}}$ and $\Pi_2 \in \mathbb{C}^{R \times R}$ such that $\Pi_1\left(\boldsymbol{A}^{(1)} \odot \cdots \odot \boldsymbol{A}^{(P)}\right)\Pi_2$ satisfies the inequalities (5) *and* (6)*. Let $\widehat{\boldsymbol{A}}^{(1)} \odot \cdots \odot \widehat{\boldsymbol{A}}^{(P)}$ be another matrix with same structure as $\boldsymbol{A}^{(1)} \odot \cdots \odot \boldsymbol{A}^{(P)}$, then $Col\left(\widehat{\boldsymbol{A}}^{(1)} \odot \cdots \odot \widehat{\boldsymbol{A}}^{(P)}\right) \subseteq Col\left(\boldsymbol{A}^{(1)} \odot \cdots \odot \boldsymbol{A}^{(P)}\right) = Col\left(\boldsymbol{T}^{[P]}\right)$ if and only if $\widehat{\boldsymbol{A}}^{(1)} \odot \cdots \odot \widehat{\boldsymbol{A}}^{(P)} = \left(\boldsymbol{A}^{(1)} \odot \cdots \odot \boldsymbol{A}^{(P)}\right)\boldsymbol{D}$, where $\boldsymbol{D}$ is a nonsingular diagonal matrix.*

Proof. *Assume that* $\widehat{\boldsymbol{A}}^{(1)} \odot \cdots \odot \widehat{\boldsymbol{A}}^{(P)} = \left(\boldsymbol{A}^{(1)} \odot \cdots \odot \boldsymbol{A}^{(P)}\right)\boldsymbol{D}$*, where and* $\boldsymbol{D}$ *is a nonsingular diagonal matrix, then obviously* $Col\left(\widehat{\boldsymbol{A}}^{(1)} \odot \cdots \odot \widehat{\boldsymbol{A}}^{(P)}\right) \subseteq Col\left(\boldsymbol{A}^{(1)} \odot \cdots \odot \boldsymbol{A}^{(P)}\right) = Col\left(\boldsymbol{T}^{[P]}\right)$*. Conversely, assume that* $Col\left(\widehat{\boldsymbol{A}}^{(1)} \odot \cdots \odot \widehat{\boldsymbol{A}}^{(P)}\right) \subseteq Col\left(\boldsymbol{A}^{(1)} \odot \cdots \odot \boldsymbol{A}^{(P)}\right) = Col\left(\boldsymbol{T}^{[P]}\right)$*, then there exists a nonsingular matrix* $\boldsymbol{M} \in \mathbb{C}^{R\times R}$ *such that*

$$\left(\widehat{\boldsymbol{A}}^{(1)} \odot \cdots \odot \widehat{\boldsymbol{A}}^{(P)}\right)\boldsymbol{M} \quad = \quad \boldsymbol{A}^{(1)} \odot \cdots \odot \boldsymbol{A}^{(P)}. \qquad (7)$$

By assumption we also assume that there exist permutation matrices Π_1 *and* Π_2 *such that* $\Pi_1\left(\boldsymbol{A}^{(1)} \odot \cdots \odot \boldsymbol{A}^{(P)}\right)\Pi_2$ *satisfies the inequalities* (5) *and* (6)*. Hence, from* (7) *we get*

$$\Pi_1\left(\widehat{\boldsymbol{A}}^{(1)} \odot \cdots \odot \widehat{\boldsymbol{A}}^{(P)}\right)\Pi_2\overline{\boldsymbol{M}} = \Pi_1\left(\boldsymbol{A}^{(1)} \odot \cdots \odot \boldsymbol{A}^{(P)}\right)\Pi_2,$$

where $\Pi_1\left(\widehat{\boldsymbol{A}}^{(1)} \odot \cdots \odot \widehat{\boldsymbol{A}}^{(P)}\right)\Pi_2$ *and* $\Pi_1\left(\boldsymbol{A}^{(1)} \odot \cdots \odot \boldsymbol{A}^{(P)}\right)\Pi_2$ *are banded matrices satisfying* (5) *and* (6) *and* $\overline{\boldsymbol{M}} = \Pi_2^T\boldsymbol{M}\Pi_2$ *is a nonsingular matrix. Due to property* (5) *the matrix* $\overline{\boldsymbol{M}}$ *must be lower triangular and due to property* (6) *the matrix* $\overline{\boldsymbol{M}}$ *must also be upper triangular. Hence, the matrix* $\overline{\boldsymbol{M}}$ *must be diagonal. Hence* $\boldsymbol{M} = \boldsymbol{D}$.

We notice from proposition 3.1 that if only one banded matrix factor is exploited, then it is required to be upper and lower banded. However, if several of the matrix factors are banded, then from proposition 3.1 we only require that the Khatri-Rao product of the banded matrix factors is upper and lower banded. This means, by exploiting the structure of several banded matrix factors, we obtain more relaxed uniqueness results.

4. Computation of CPDs with Banded Matrix Factors

In this section we develop numerical methods to compute CPDs containing banded matrix factors. To simplify the discussion, let us only consider computational methods exploiting that one or two of the matrix factors are banded. The presented method can be generalized in such a way that it is capable of jointly exploiting that three or more of the matrix factors are banded. Subsection 4.1 and 4.2 provide numerical methods to compute a CPD exploiting that one and two of the matrix factors are banded, respectively.

4.1. Exploiting one Banded Matrix Factor

The method outlined in this section will be referred to as CPBAND1. Let $\mathcal{T} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ with rank R and let $\mathbf{T}_{[N]} = \left(\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(N-1)}\right) \mathbf{A}^{(N)T}$ be its matrix representation. Assume that the matrices $\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(N-1)}$ and $\mathbf{A}^{(N)}$ have full column rank. Let $\mathbf{T}_{[N]} = \mathbf{U}\Sigma\mathbf{V}^H$ denote the compact Singular Value Decomposition (SVD) of $\mathbf{T}_{[N]}$, then there exists a nonsingular matrix $\mathbf{N} \in \mathbb{C}^{R \times R}$ such that

$$\mathbf{U}\Sigma\mathbf{N}^{-1} = \mathbf{A}^{(1)} \odot \mathbf{A}^{(2)} \odot \cdots \odot \mathbf{A}^{(N-1)} \tag{8}$$

$$\mathbf{V}^*\mathbf{N} = \mathbf{A}^{(N)} \tag{9}$$

Assume also that $\mathbf{A}^{(N)}$ is a banded matrix with $1 \leq \mathrm{b}_U\left(\mathbf{a}_1^{(N)}\right) < \cdots < \mathrm{b}_U\left(\mathbf{a}_R^{(N)}\right)$ and $\mathrm{b}_L\left(\mathbf{a}_1^{(N)}\right) < \cdots < \mathrm{b}_L\left(\mathbf{a}_R^{(N)}\right) \leq I_N$ up to a permutation of its rows and columns, then according to proposition 3.1, we can find $\mathbf{N}$ and $\mathbf{A}^{(N)}$ from $\mathrm{Row}\left(\mathbf{T}_{[N]}\right)$.

Let the rth column vector of $\mathbf{A}^{(N)}$ be parameterized as

$$\mathbf{a}_r^{(N)} = \sum_{m=\mathrm{b}_U\left(\mathbf{a}_r^{(N)}\right)}^{\mathrm{b}_L\left(\mathbf{a}_r^{(N)}\right)} a_m \mathbf{e}_m^{(I_N)},$$

where $\mathbf{e}_m^{(I_N)} \in \mathbb{C}^{I_N}$ is the unit vector which is equal to one at entry m and zero elsewhere. The linear system (9), then decouples into the R linear systems

$$\begin{aligned} \mathbf{V}^*\mathbf{n}_r &= \sum_{m=\mathrm{b}_U\left(\mathbf{a}_r^{(N)}\right)}^{\mathrm{b}_L\left(\mathbf{a}_r^{(N)}\right)} a_m \mathbf{e}_m^{(I_N)} \\ &= \left[\mathbf{e}^{(I_N)}_{\mathrm{b}_U\left(\mathbf{a}_r^{(N)}\right)}, \ldots, \mathbf{e}^{(I_N)}_{\mathrm{b}_L\left(\mathbf{a}_r^{(N)}\right)}\right] \mathbf{a}_r^{(N)}, \quad r \in [1, R], \end{aligned} \tag{10}$$

where $\mathbf{n}_r$ is the rth column vector of $\mathbf{N}$ and the vector $\mathbf{a}_r^{(N)} \in \mathbb{C}^{\mathrm{b}_L\left(\mathbf{a}_r^{(N)}\right) - \mathrm{b}_U\left(\mathbf{a}_r^{(N)}\right) + 1}$ contains the coefficients used to construct the rth column of $\mathbf{A}^{(N)}$. Let

$$\mathbf{E}^{(r)} = \left[\mathbf{e}^{(I_N)}_{\mathrm{b}_U\left(\mathbf{a}_r^{(N)}\right)}, \ldots, \mathbf{e}^{(I_N)}_{\mathrm{b}_L\left(\mathbf{a}_r^{(N)}\right)}\right] \in \mathbb{C}^{I_N \times \left(\mathrm{b}_L\left(\mathbf{a}_r^{(N)}\right) - \mathrm{b}_U\left(\mathbf{a}_r^{(N)}\right) + 1\right)}, \quad r \in [1, R],$$

then equation (10) can be written as

$$\mathbf{V}^*\mathbf{n}_r = \mathbf{E}^{(r)}\mathbf{a}_r^{(N)} \Leftrightarrow \left[\mathbf{V}^*, -\mathbf{E}^{(r)}\right]\begin{bmatrix} \mathbf{n}_r \\ \mathbf{a}_r^{(N)} \end{bmatrix} = \mathbf{0}\,, \quad r \in [1, R]\,. \tag{11}$$

From proposition 3.1 we know that the solution to (11) is essentially unique. The Least Squares (LS) solution to (11) is given by the right singular vector of the matrix $\left[\mathbf{V}^*, -\mathbf{E}^{(r)}\right]$ associated with its smallest singular value. Let $\mathbf{x} \in \mathbb{C}^{R+\mathrm{b}_L\left(\mathbf{a}_r^{(N)}\right)-\mathrm{b}_U\left(\mathbf{a}_r^{(N)}\right)+1}$ be the solution to the system (11), then

$$\begin{aligned} \mathbf{n}_r &= \mathbf{x}(1:R)\,, & r \in [1, R] \\ \mathbf{a}_r^{(N)} &= \mathbf{x}\left(R+1 : R+1+\mathrm{b}_L\left(\mathbf{a}_r^{(N)}\right) - \mathrm{b}_U\left(\mathbf{a}_r^{(N)}\right)\right)\,, & r \in [1, R] \end{aligned}$$

Once $\mathbf{A}^{(N)}$ and $\mathbf{N}$ have been obtained, then from (8) we have the relation $\mathbf{F} = \mathbf{U}\Sigma\mathbf{N}^{-1} = \mathbf{A}^{(1)} \odot \mathbf{A}^{(2)} \odot \cdots \odot \mathbf{A}^{(N-1)}$. Thus, we can find the remaining unknown matrix factors from the R decoupled best rank-1 tensor approximation problems

$$\min_{\mathbf{a}_r^{(1)},\ldots,\mathbf{a}_r^{(N-1)}} \left\|\mathbf{f}_r - \mathbf{a}_r^{(1)} \otimes \cdots \otimes \mathbf{a}_r^{(N-1)}\right\|_F^2, \quad r \in [1, R], \tag{12}$$

where $\mathbf{f}_r$ and $\mathbf{a}_r^{(n)}$ denote the rth column vector of $\mathbf{F}$ and $\mathbf{A}^{(n)}$, respectively. To numerically solve the best rank-1 tensor approximation problems (12) the higher-order power method [8] could for instance be used.

Remark that if $N = 2$, then $\mathbf{T}_{[2]} = \mathbf{A}^{(1)}\mathbf{A}^{(2)T}$ and we directly obtain $\mathbf{A}^{(1)}$ from $\mathbf{A}^{(1)} = \mathbf{T}_{[2]}\left(\mathbf{A}^{(2)T}\right)^{\dagger}$ or alternatively $\mathbf{A}^{(1)} = \mathbf{U}\Sigma\mathbf{N}^{-1}$. Furthermore, if $N = 3$, then $\mathbf{T}_{[3]} = \left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right)\mathbf{A}^{(3)T}$ and $\mathbf{F} = \mathbf{U}\Sigma\mathbf{N}^{-1} = \mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}$. We have $\mathbf{f}_r = \mathbf{a}_r^{(1)} \otimes \mathbf{a}_r^{(2)}$ and $\mathbf{F}_r = \mathrm{Unvec}\left(\mathbf{f}_r\right) = \mathbf{a}_r^{(2)}\mathbf{a}_r^{(1)T}$, where $\mathbf{f}_r$, $\mathbf{a}_r^{(1)}$ and $\mathbf{a}_r^{(2)}$ denotes the rth column vector of $\mathbf{F}$, $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$, respectively. The matrices $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ can therefore be found from the R decoupled best rank-1 *matrix* approximation problems

$$\min_{\mathbf{a}_r^{(1)},\mathbf{a}_r^{(2)}} \left\|\mathbf{F}_r - \mathbf{a}_r^{(2)}\mathbf{a}_r^{(1)T}\right\|_F^2, \quad r \in [1, R],$$

which can be solved by standard numerical linear algebra methods.

Assume that the matrices $\mathbf{A}^{(p)}$, $1 \leq p \leq P \leq N$ are lower and upper banded matrix factors. If the matrix representations $\mathbf{T}_{[p]}$ only consist of full column rank matrices $\forall p \in [1, P]$, then we can successively compute the banded matrix factor $\mathbf{A}^{(p)}$ from $\mathrm{Row}\left(\mathbf{T}_{[p]}\right)$ in a non-iterative way $\forall p \in [1, P]$. This is done by repeating the procedure just outlined above in this subsection.

Let us discuss this approach in more detail when $P = 2$. Assume that $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ are banded matrices satisfying the conditions stated in proposition 3.1. Assume also that the matrices $\mathbf{A}^{(1)}$, $\mathbf{A}^{(2)}$, $\left(\mathbf{A}^{(2)} \odot \mathbf{A}^{(3)} \odot \cdots \odot \mathbf{A}^{(N)}\right)$, $\left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(3)} \odot \cdots \odot \mathbf{A}^{(N)}\right)$ and $\left(\mathbf{A}^{(3)} \odot \mathbf{A}^{(4)} \odot \cdots \odot \mathbf{A}^{(N)}\right)$ have full column rank.

We first compute $\mathbf{A}^{(1)}$. Let $\mathbf{T}_{[1]} = \mathbf{U}\Sigma\mathbf{V}^{H}$ denote the compact SVD of $\mathbf{T}_{[1]}$, then there exists a nonsingular matrix $\mathbf{N} \in \mathbb{C}^{R\times R}$ such that $\mathbf{V}^{*}\mathbf{N} = \mathbf{A}^{(1)}$. By applying the procedure described in this subsection we obtain $\mathbf{A}^{(1)}$ from $\mathrm{Row}\left(\mathbf{T}_{[1]}\right)$.

Next, we compute $\mathbf{A}^{(2)}$. Let $\mathbf{T}_{[2]} = \mathbf{U}\Sigma\mathbf{V}^{H}$ denote the compact SVD of $\mathbf{T}_{[2]}$, then there exists a nonsingular matrix $\mathbf{N} \in \mathbb{C}^{R\times R}$ such that $\mathbf{V}^{*}\mathbf{N} = \mathbf{A}^{(2)}$. Again, by applying the procedure described in this subsection we obtain $\mathbf{A}^{(2)}$ from $\mathrm{Row}\left(\mathbf{T}_{[2]}\right)$.

In the final step we find the remaining matrix factors. Let

$$\mathbf{T}^{[2]} = \left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right)\left(\mathbf{A}^{(3)} \odot \cdots \odot \mathbf{A}^{(N)}\right)^{T} = \overline{\mathbf{U}}\,\overline{\Sigma}\,\overline{\mathbf{V}}^{H}$$

denote the compact SVD of $\mathbf{T}^{[2]}$, then there exists a nonsingular matrix $\mathbf{M} \in \mathbb{C}^{R\times R}$ such that $\overline{\mathbf{U}}\mathbf{M} = \left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right) \Leftrightarrow \mathbf{M} = \overline{\mathbf{U}}^{H}\left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right)$. Let $\mathbf{F} = \overline{\mathbf{V}}^{*}\Sigma\mathbf{M}^{-T} = \mathbf{A}^{(3)} \odot \cdots \odot \mathbf{A}^{(N)}$, then the remaining matrix factors follow from the R decoupled best rank-1 tensor approximation problems

$$\min_{\mathbf{a}_r^{(3)},\ldots,\mathbf{a}_r^{(N)}} \left\|\mathbf{f}_r - \mathbf{a}_r^{(3)} \otimes \cdots \otimes \mathbf{a}_r^{(N)}\right\|_F^2, \quad r \in [1, R],$$

where $\mathbf{f}_r$ and $\mathbf{a}_r^{(n)}$ denote the rth column vector of $\mathbf{F}$ and $\mathbf{A}^{(n)}$, respectively.

Remark that if $N = 3$, then $\mathbf{T}_{[3]} = \left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right)\mathbf{A}^{(3)T}$ and we *directly* obtain $\mathbf{A}^{(3)T} = \left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right)^{\dagger}\mathbf{T}_{[3]}$ or alternatively $\mathbf{A}^{(3)} = \mathbf{V}^{*}\Sigma\mathbf{M}^{-T}$. Furthermore, if $N = 4$, then $\mathbf{T}^{[2]} = \left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right)\left(\mathbf{A}^{(3)} \odot \mathbf{A}^{(4)}\right)^{T}$ and $\mathbf{F} = \mathbf{V}^{*}\Sigma\mathbf{M}^{-T} = \mathbf{A}^{(3)} \odot \mathbf{A}^{(4)}$. We have $\mathbf{f}_r = \mathbf{a}_r^{(3)} \otimes \mathbf{a}_r^{(4)}$ and $\mathbf{F}_r = \mathrm{Unvec}\left(\mathbf{f}_r\right) = \mathbf{a}_r^{(4)}\mathbf{a}_r^{(3)T}$, where $\mathbf{f}_r$, $\mathbf{a}_r^{(3)}$ and $\mathbf{a}_r^{(4)}$ denote the rth column vector of $\mathbf{F}$, $\mathbf{A}^{(3)}$ and $\mathbf{A}^{(4)}$, respectively. The matrices $\mathbf{A}^{(3)}$ and $\mathbf{A}^{(4)}$ can then be found from the R decoupled best rank-1 *matrix* approximation problems

$$\min_{\mathbf{a}_r^{(3)},\mathbf{a}_r^{(4)}} \left\|\mathbf{F}_r - \mathbf{a}_r^{(4)}\mathbf{a}_r^{(3)T}\right\|_F^2, \quad r \in [1, R],$$

which can be solved by standard numerical linear algebra methods.

4.2. Exploiting Jointly two Banded Matrix Factors

In this subsection we present a method for computing a CPD which jointly exploits that two of the matrix factors are banded. It is referred to as CPBAND2.

Let $\mathcal{T} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ with rank R and matrix representation

$$\mathbf{T}^{[2]} = \left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right)\left(\mathbf{A}^{(3)} \odot \mathbf{A}^{(4)} \odot \cdots \odot \mathbf{A}^{(N)}\right)^T.$$

Assume that the matrices $\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}$ and $\mathbf{A}^{(3)} \odot \mathbf{A}^{(4)} \odot \cdots \odot \mathbf{A}^{(N)}$ have full column rank. Note that this condition does not require that the matrix factors $\{\mathbf{A}^{(n)}\}$ to have full column rank. Furthermore, assume that $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ are banded matrices satisfying the conditions stated in proposition 3.1. Let $\mathbf{T}^{[2]} = \mathbf{U}\Sigma\mathbf{V}^H$ denote the compact SVD of $\mathbf{T}^{[2]}$, then there exists a nonsingular matrix $\mathbf{M} \in \mathbb{C}^{R \times R}$ such that

$$\mathbf{U}\mathbf{M} = \mathbf{A}^{(1)} \odot \mathbf{A}^{(2)} \tag{13}$$

$$\mathbf{V}^*\Sigma\mathbf{M}^{-T} = \mathbf{A}^{(3)} \odot \mathbf{A}^{(4)} \odot \cdots \odot \mathbf{A}^{(N)} \tag{14}$$

According to proposition 3.1, we can find $\mathbf{M}$, $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ from $\mathrm{Col}\left(\mathbf{T}^{[2]}\right)$. Let the rth column vector of $\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}$ be parameterized as

$$\mathbf{a}_r^{(1)} \otimes \mathbf{a}_r^{(2)} = \sum_{m=\mathrm{b}_U\left(\mathbf{a}_r^{(1)}\right)}^{\mathrm{b}_L\left(\mathbf{a}_r^{(1)}\right)} \sum_{n=\mathrm{b}_U\left(\mathbf{a}_r^{(2)}\right)}^{\mathrm{b}_L\left(\mathbf{a}_r^{(2)}\right)} a_{mr}^{(1)} a_{nr}^{(2)} \mathbf{e}_m^{(I_1)} \otimes \mathbf{e}_n^{(I_2)},$$

where $a_{mr}^{(p)} = \mathbf{A}^{(p)}(m, r)$ and $\mathbf{e}_m^{(I_n)} \in \mathbb{C}^{I_n}$ is the unit vector which is equal to one at entry m and zero elsewhere. The system of equations (13), then decouples into R independent systems of equations

$$\begin{aligned}\mathbf{U}\mathbf{m}_r &= \sum_{m=\mathrm{b}_U\left(\mathbf{a}_r^{(1)}\right)}^{\mathrm{b}_L\left(\mathbf{a}_r^{(1)}\right)} \sum_{n=\mathrm{b}_U\left(\mathbf{a}_r^{(2)}\right)}^{\mathrm{b}_L\left(\mathbf{a}_r^{(2)}\right)} a_{mr}^{(1)} a_{nr}^{(2)} \mathbf{e}_m^{(I_1)} \otimes \mathbf{e}_n^{(I_2)} \\ &= \left[\mathbf{e}_{\mathrm{b}_U\left(\mathbf{a}_r^{(1)}\right)}^{(I_1)} \otimes \mathbf{e}_{\mathrm{b}_U\left(\mathbf{a}_r^{(2)}\right)}^{(I_2)}, \ldots, \mathbf{e}_{\mathrm{b}_L\left(\mathbf{a}_r^{(1)}\right)}^{(I_1)} \otimes \mathbf{e}_{\mathrm{b}_L\left(\mathbf{a}_r^{(2)}\right)}^{(I_2)}\right] \mathbf{a}_r^{(1)} \otimes \mathbf{a}_r^{(2)}, \quad r \in [1, R], \end{aligned} \tag{15}$$

where $\mathbf{m}_r$ is the rth column vector of $\mathbf{M}$ and the vector $\mathbf{a}_r^{(n)} \in \mathbb{C}^{\mathrm{b}_L\left(\mathbf{a}_r^{(n)}\right)-\mathrm{b}_U\left(\mathbf{a}_r^{(n)}\right)+1}$ contains the coefficients for the basis vectors used to construct the rth column vector of $\mathbf{A}^{(n)}$. Let

$$\mathbf{E}^{(r)} = \left[\mathbf{e}^{(I_1)}_{\mathrm{b}_U\left(\mathbf{a}_r^{(1)}\right)} \otimes \mathbf{e}^{(I_2)}_{\mathrm{b}_U\left(\mathbf{a}_r^{(2)}\right)}, \dots, \mathbf{e}^{(I_1)}_{\mathrm{b}_L\left(\mathbf{a}_r^{(1)}\right)} \otimes \mathbf{e}^{(I_2)}_{\mathrm{b}_L\left(\mathbf{a}_r^{(2)}\right)}\right] \in \mathbb{C}^{I_1 I_2 \times \left(\mathrm{b}_L\left(\mathbf{a}_r^{(1)}\right)-\mathrm{b}_U\left(\mathbf{a}_r^{(1)}\right)+1\right)\left(\mathrm{b}_L\left(\mathbf{a}_r^{(2)}\right)-\mathrm{b}_U\left(\mathbf{a}_r^{(2)}\right)+1\right)},$$

where $r \in [1, R]$, then equation (15) can be written as

$$\mathbf{U}\mathbf{m}_r = \mathbf{E}^{(r)}\left(\mathbf{a}_r^{(1)} \otimes \mathbf{a}_r^{(2)}\right) \Leftrightarrow \left[\mathbf{U}, -\mathbf{E}^{(r)}\right]\begin{bmatrix} \mathbf{m}_r \\ \mathbf{a}_r^{(1)} \otimes \mathbf{a}_r^{(2)} \end{bmatrix} = \mathbf{0}, \quad r \in [1, R]. \tag{16}$$

From proposition 3.1 we know that the solution to (16) is essentially unique. In order to solve the system (16) we ignore the Kronecker product structure $\mathbf{a}_r^{(1)} \otimes \mathbf{a}_r^{(2)}$. Remark this relaxation may affect uniqueness of the problem. However, numerical experiments indicate that this relaxation step does not seem to affect the uniqueness of the solution. The LS solution to the relaxed version of (16) is given by the right singular vector of the matrix $\left[\mathbf{U}, -\mathbf{E}^{(r)}\right]$ associated with its smallest singular value. Let $\mathbf{x} \in \mathbb{C}^{R+\left(\mathrm{b}_L\left(\mathbf{a}_r^{(1)}\right)-\mathrm{b}_U\left(\mathbf{a}_r^{(1)}\right)+1\right)\left(\mathrm{b}_L\left(\mathbf{a}_r^{(2)}\right)-\mathrm{b}_U\left(\mathbf{a}_r^{(2)}\right)+1\right)}$ be the LS solution to the relaxed version of the system (16), then we set

$$\begin{aligned}
\mathbf{m}_r &= \mathbf{x}(1:R), \\
\mathbf{y}_r &= \mathbf{a}_r^{(1)} \otimes \mathbf{a}_r^{(2)} \\
&= \mathbf{x}\left(R+1 : R + \left(\mathrm{b}_L\left(\mathbf{a}_r^{(1)}\right) - \mathrm{b}_U\left(\mathbf{a}_r^{(1)}\right) + 1\right)\left(\mathrm{b}_L\left(\mathbf{a}_r^{(2)}\right) - \mathrm{b}_U\left(\mathbf{a}_r^{(2)}\right) + 1\right)\right),
\end{aligned}$$

where $r \in [1, R]$. Let $\mathbf{Y}_r = \mathrm{Unvec}\left(\mathbf{y}_r\right) = \mathbf{a}_r^{(2)}\mathbf{a}_r^{(1)T}$, then $\{\mathbf{a}_r^{(1)}\}$ and $\{\mathbf{a}_r^{(2)}\}$ follow from the R decoupled best rank-1 matrix approximation problems

$$\min_{\mathbf{a}_r^{(1)}, \mathbf{a}_r^{(2)}} \left\|\mathbf{Y}_r - \mathbf{a}_r^{(2)}\mathbf{a}_r^{(1)T}\right\|_F^2, \quad r \in [1, R].$$

Remark that this method is able to take a partial symmetry $\mathbf{a}_r^{(1)} = \mathbf{a}_r^{(2)}$ or a partial Hermitian symmetry $\mathbf{a}_r^{(1)} = \mathbf{a}_r^{(2)*}$ into account.

Once the matrices $\mathbf{M}$, $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ have been found, then from (14) we have the relation $\mathbf{F} = \mathbf{V}^{*}\Sigma\mathbf{M}^{-T} = \mathbf{A}^{(3)} \odot \mathbf{A}^{(4)} \odot \cdots \odot \mathbf{A}^{(N)}$. Hence, the remaining matrix factors follow from the R decoupled best rank-1 tensor approximation problems

$$\min_{\mathbf{a}_r^{(3)},\ldots,\mathbf{a}_r^{(N)}} \left\| \mathbf{f}_r - \mathbf{a}_r^{(3)} \otimes \cdots \otimes \mathbf{a}_r^{(N)} \right\|_F^2, \quad r \in [1, R],$$

where $\mathbf{f}_r$ and $\mathbf{a}_r^{(n)}$ denote the rth column vector of $\mathbf{F}$ and $\mathbf{A}^{(n)}$, respectively.

Again, remark that if $N = 3$, then we *directly* obtain $\mathbf{A}^{(3)T} = \left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right)^{\dagger} \mathbf{T}_{(1)}$ or alternatively $\mathbf{A}^{(3)} = \mathbf{V}^{*}\Sigma\mathbf{M}^{-T}$. Furthermore, if $N = 4$, then $\mathbf{T}^{[2]} = \left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right)\left(\mathbf{A}^{(3)} \odot \mathbf{A}^{(4)}\right)^{T}$ and $\mathbf{F} = \mathbf{V}^{*}\Sigma\mathbf{M}^{-T} = \mathbf{A}^{(3)} \odot \mathbf{A}^{(4)}$. We have $\mathbf{f}_r = \mathbf{a}_r^{(3)} \otimes \mathbf{a}_r^{(4)}$ and $\mathbb{C}^{I_4 \times I_3} \ni \mathbf{F}_r = \mathrm{Unvec}\left(\mathbf{f}_r\right) = \mathbf{a}_r^{(4)}\mathbf{a}_r^{(3)T}$, where $\mathbf{f}_r$, $\mathbf{a}_r^{(3)}$ and $\mathbf{a}_r^{(4)}$ denotes the rth column vector of $\mathbf{F}$, $\mathbf{A}^{(3)}$ and $\mathbf{A}^{(4)}$, respectively. Thus, the matrices $\mathbf{A}^{(3)}$ and $\mathbf{A}^{(4)}$ follow from the solutions of the R decoupled best rank-1 *matrix* approximation problems

$$\min_{\mathbf{a}_r^{(3)},\mathbf{a}_r^{(4)}} \left\| \mathbf{F}_r - \mathbf{a}_r^{(4)}\mathbf{a}_r^{(3)T} \right\|_F^2, \quad r \in [1, R],$$

which can be solved by standard numerical linear algebra methods.

5. Uniqueness of CPDs with Banded and Structured Matrix Factors

We say that a matrix $\mathbf{A} \in \mathbb{C}^{I \times R}$ is structured if it can be written as $\mathbf{A} = \sum_{l=1}^{L} a_l \mathbf{E}_l$ where the matrices $\mathbf{E}_l$ are given, and $L < IR$ [7]. Examples of such matrices are banded Hankel or Toeplitz matrices which will be discussed in more details. Let $\mathbf{A}$ be a structured matrix, then $\dim\left(\mathbf{A}\right)$ denotes the dimension of the subspace of structured matrices consisting of matrices with the same structure as $\mathbf{A}$.

For instance, let $\mathbf{A} \in \mathbb{C}^{I \times R}$ be a banded Hankel matrix with $\mathrm{b}_U\left(\mathbf{a}_r\right) = \mathrm{b}_U\left(\mathbf{a}_1\right) + r - 1$ and $\mathrm{b}_L\left(\mathbf{a}_r\right) = \mathrm{b}_L\left(\mathbf{a}_1\right) + r - 1$. Let $\{\mathbf{E}_l\}$ be a basis for the banded Hankel matrix $\mathbf{A}$ with predefined values for $\mathrm{b}_U\left(\mathbf{a}_1\right)$ and $\mathrm{b}_L\left(\mathbf{a}_1\right)$, then $\mathbf{A} = \sum_{l=1}^{L} a_l \mathbf{E}_l$, where $L = \dim\left(\mathbf{A}\right) = \mathrm{b}_L\left(\mathbf{a}_1\right) - \mathrm{b}_U\left(\mathbf{a}_1\right) + 1$. In particular, basis matrices of the form

$$\left(\mathbf{E}_l\right)_{ij} = \begin{cases} 1, & i = \mathrm{b}_U\left(\mathbf{a}_1\right) + l - j \\ 0, & \text{otherwise} \end{cases}$$

will be used in section 7.

As another example, let $\mathbf{A} \in \mathbb{C}^{I \times R}$ be a banded Toeplitz matrix with $\mathrm{b}_U(\mathbf{a}_r) = \mathrm{b}_U(\mathbf{a}_1) - r + 1$ and $\mathrm{b}_L(\mathbf{a}_r) = \mathrm{b}_L(\mathbf{a}_1) - r + 1$. Let $\{\mathbf{E}_l\}$ be a basis for the banded Toeplitz matrix $\mathbf{A}$ with predefined values for $\mathrm{b}_U(\mathbf{a}_1)$ and $\mathrm{b}_L(\mathbf{a}_1)$, then $\mathbf{A} = \sum_{l=1}^{L} a_l \mathbf{E}_l$, where $L = \dim(\mathbf{A}) = \mathrm{b}_L(\mathbf{a}_1) - \mathrm{b}_U(\mathbf{a}_1) + 1$. Again, basis matrices of the form

$$(\mathbf{E}_l)_{ij} = \begin{cases} 1, & i = \mathrm{b}_U(\mathbf{a}_1) + l + j - 2 \\ 0, & \text{otherwise} \end{cases}$$

will be used in section 7.

In subsection 5.1 we extend the uniqueness result from section 3 concerning tensor decompositions with banded matrix factors to the case of tensor decompositions with banded and structured matrix factors. Next, in subsection 5.2 we provide some uniqueness results valid for tensor decompositions with a banded Hankel or Toeplitz matrix factor.

5.1. Banded and Structured Matrix Factors

The following proposition 5.2 presented in this subsection is an extension of proposition 3.1 to the case when the matrix factors under consideration are banded and structured. In order to prove proposition 5.2 we will make use of lemma 5.1.

Lemma 5.1. *Let $\boldsymbol{A}^{(n)} \in \mathbb{C}^{I_n \times R}$ belong to a subspace of structured matrices of dimension L_n and $n \in [1, N]$. Let $\widehat{\boldsymbol{A}}^{(n)}$ have the same structure as $\boldsymbol{A}^{(n)}$ and assume that*

$$\widehat{\boldsymbol{A}}^{(1)} \odot \cdots \odot \widehat{\boldsymbol{A}}^{(N)} = \left(\boldsymbol{A}^{(1)} \odot \cdots \odot \boldsymbol{A}^{(N)}\right)\boldsymbol{D},$$

where $\boldsymbol{D} \in \mathbb{C}^{R \times R}$ is a nonsingular diagonal matrix up to a possible permutation of its rows and columns. If there exist orthogonal bases $\{\boldsymbol{E}_{l_n}^{(n)}\}$ such that $\boldsymbol{A}^{(n)} = \sum_{l_n=1}^{L_n} a_{l_n}^{(n)} \boldsymbol{E}_{l_n}^{(n)}$ with properties $\left(\boldsymbol{E}_{l_n}^{(n)}\right)_{pq} \in \{0, 1\}$, $\forall l_n, n, p, q$, and $\forall n \in [1, N]$ $\exists l_n \in [1, L_n]$ such that $\forall r \in [1, R]$, $\exists i_r \in [1, I_n]$, $\left(\boldsymbol{E}_{l_n}^{(n)}\right)_{i_r r} = 1$, i.e. for every n, there exists at least one basis matrix $\boldsymbol{E}_{l_n}^{(n)}$ having a 1 in each column, then $\boldsymbol{D} = \alpha \boldsymbol{I}_R$, $\alpha \in \mathbb{C}$.

Proof. *Assume that*

$$\sum_{n=1}^{N} \sum_{l_n=1}^{L_n} \hat{a}_{l_1}^{(1)} \cdots \hat{a}_{l_N}^{(N)} \left(\boldsymbol{E}_{l_1}^{(1)} \odot \cdots \odot \boldsymbol{E}_{l_N}^{(N)}\right) = \sum_{n=1}^{N} \sum_{l_n=1}^{L_n} a_{l_1}^{(1)} \cdots a_{l_N}^{(N)} \left(\boldsymbol{E}_{l_1}^{(1)} \odot \cdots \odot \boldsymbol{E}_{l_N}^{(N)}\right)\boldsymbol{D},$$

where $\mathbf{D} \in \mathbb{C}^{R\times R}$ *is a nonsingular diagonal matrix up to a possible permutation of its rows and columns. Since there exist orthogonal bases* $\{\mathbf{E}_{l_n}^{(n)}\}$ *with properties* $\left(\mathbf{E}_{l_n}^{(n)}\right)_{pq} \in \{0,1\}$, $\forall l_n, n, p, q$, *and* $\forall n \in [1,N]$ $\exists l_n \in [1,L_n]$ *such that* $\forall r \in [1,R]$, $\exists i_r \in [1,I_n]$, $\left(\mathbf{E}_{l_n}^{(n)}\right)_{i_r r} = 1$, *we first notice that there exists a basis matrix* $\mathbf{E}_{l_1}^{(1)} \odot \cdots \odot \mathbf{E}_{l_N}^{(N)}$ *containing a unit entry in each of its column vectors. This means that* $a_{l_1}^{(1)} \cdots a_{l_N}^{(N)}$ *must appear in each column of this particular matrix* $a_{l_1}^{(1)} \cdots a_{l_N}^{(N)} \left(\mathbf{E}_{l_1}^{(1)} \odot \cdots \odot \mathbf{E}_{l_N}^{(N)}\right)$. *Second, since* $\{\mathbf{E}_{l_n}^{(n)}\}$ *are orthogonal bases, then due to the property of the Khatri-Rao product the set* $\{\mathbf{E}_{l_1}^{(1)} \odot \cdots \odot \mathbf{E}_{l_N}^{(N)}\}$ *is also an orthogonal basis. Third, the basis matrices* $\{\mathbf{E}_{l_n}^{(n)}\}$ *are constructed such that* $\left(\mathbf{E}_{l_n}^{(n)}\right)_{pq} \in \{0,1\}$, $\forall l_n, n, p, q$ *and therefore* $\left(\mathbf{E}_{l_1}^{(1)} \odot \cdots \odot \mathbf{E}_{l_N}^{(N)}\right)_{pq} \in \{0,1\}$, $\forall l_1, \ldots, l_N, p, q$. *Due to orthogonality,* $vec\left(\mathbf{E}_{l_1}^{(1)} \odot \cdots \odot \mathbf{E}_{l_N}^{(N)}\right)^H vec\left(\mathbf{E}_{l'_1}^{(1)} \odot \cdots \odot \mathbf{E}_{l'_N}^{(N)}\right) = 0$ *when* $l_n \neq l'_n$ *for some* n. *The last two facts imply that if* $\left(\mathbf{E}_{l_1}^{(1)} \odot \cdots \odot \mathbf{E}_{l_N}^{(N)}\right)_{pq} = 1$ *then* $\left(\mathbf{E}_{l'_1}^{(1)} \odot \cdots \odot \mathbf{E}_{l'_N}^{(N)}\right)_{pq} = 0$ *when* $l_n \neq l'_n$ *for some* n. *Hence, in order to preserve the Khatri-Rao product structure of the linearly structured matrices* $\{\mathbf{E}_{l_n}^{(n)}\}$ *the diagonal matrix must be* $\mathbf{D} = \alpha \mathbf{I}_R$, *where* $\alpha \in \mathbb{C}$.

Proposition 5.2. *Consider* $\mathcal{T} \in \mathbb{C}^{I_1\times\cdots\times I_N}$ *with rank R and matrix representation* $\mathbf{T}^{[P]} = \left(\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(P)}\right)\left(\mathbf{A}^{(P+1)} \odot \cdots \odot \mathbf{A}^{(N)}\right)^T$ *where* $P < N$. *Assume that the matrices* $\left(\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(P)}\right)$ *and* $\left(\mathbf{A}^{(P+1)} \odot \cdots \odot \mathbf{A}^{(N)}\right)$ *have full column rank. Assume also that* $\mathbf{A}^{(n)}$, $n \in [1,P]$, *belong to a subspace of structured matrices of dimension* L_n *and satisfying the assumptions made in proposition 3.1. Moreover, assume that there exist orthogonal bases* $\{\mathbf{E}_{l_n}^{(n)}\}$ *such that* $\mathbf{A}^{(n)} = \sum_{l_n=1}^{L_n} a_{l_n}^{(n)} \mathbf{E}_{l_n}^{(n)}$ $\forall n \in [1,P]$ *with properties* $\left(\mathbf{E}_{l_n}^{(n)}\right)_{pq} \in \{0,1\}$, $\forall p, q, l_n, n$ *and* $\forall n \in [1,P]$ $\exists l_n \in [1,L_n]$ *such that* $\forall r \in [1,R]$, $\exists i_r \in [1,I_n]$, $\left(\mathbf{E}_{l_n}^{(n)}\right)_{i_r r} = 1$. *Let* $\widehat{\mathbf{A}}^{(n)}$ *be matrices with same structures as* $\mathbf{A}^{(n)}$, $\forall n \in [1,P]$, *then* $Col\left(\widehat{\mathbf{A}}^{(1)} \odot \cdots \odot \widehat{\mathbf{A}}^{(P)}\right) \subseteq Col\left(\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(P)}\right) = Col\left(\mathbf{T}^{[P]}\right)$ *if and only if* $\widehat{\mathbf{A}}^{(n)} = \mathbf{A}^{(n)}\alpha_n$, *where* $\alpha_n \in \mathbb{C}$, $\forall n \in [1,P]$.

Proof. *Let* $\widehat{\mathbf{A}}^{(n)} = \mathbf{A}^{(n)}\alpha_n$, *where* $\alpha_n \in \mathbb{C}$, $\forall n \in [1,P]$, *then obviously* $Col\left(\widehat{\mathbf{A}}^{(1)} \odot \cdots \odot \widehat{\mathbf{A}}^{(P)}\right) \subseteq$

$Col\left(\boldsymbol{A}^{(1)} \odot \cdots \odot \boldsymbol{A}^{(P)}\right) = Col\left(\boldsymbol{T}^{[P]}\right)$. Conversely, assume that $Col\left(\widehat{\boldsymbol{A}}^{(1)} \odot \cdots \odot \widehat{\boldsymbol{A}}^{(P)}\right) \subseteq Col\left(\boldsymbol{A}^{(1)} \odot \cdots \odot \boldsymbol{A}^{(P)}\right) = Col\left(\boldsymbol{T}^{[P]}\right)$, then there exists a nonsingular matrix $\boldsymbol{M} \in \mathbb{C}^{R\times R}$ such that

$$\widehat{\boldsymbol{A}}^{(1)} \odot \cdots \odot \widehat{\boldsymbol{A}}^{(P)} \;=\; \left(\boldsymbol{A}^{(1)} \odot \cdots \odot \boldsymbol{A}^{(P)}\right)\boldsymbol{M}.$$

Due to proposition 3.1 we know that $\boldsymbol{M}$ must be equal to a diagonal matrix $\boldsymbol{D}$ up to a possible permutation of its rows and columns. From lemma 5.1 we know that $\boldsymbol{D} = \alpha \boldsymbol{I}_R$, where $\alpha \in \mathbb{C}$.

5.2. Banded Toeplitz or Hankel Matrix Factor

The following proposition 5.3 shows that we can identify an upper banded Toeplitz matrix factor or a lower banded Toeplitz matrix factor $\mathbf{A}^{(P)}$ from Row $(\mathbf{T}_{[P]})$. Similarly, proposition 5.4 explains that we can identify an anti-upper banded Hankel matrix factor or an anti-lower banded Hankel matrix factor $\mathbf{A}^{(P)}$ from Row $(\mathbf{T}_{[P]})$.

Proposition 5.3. *Consider $\mathcal{T} \in \mathbb{C}^{I_1\times\cdots\times I_N}$ with rank R and matrix representation $\boldsymbol{T}_{[P]} = \left(\boldsymbol{A}^{(1)} \odot \cdots \odot \boldsymbol{A}^{(P-1)} \odot \boldsymbol{A}^{(P+1)} \odot \cdots \odot \boldsymbol{A}^{(N)}\right)\boldsymbol{A}^{(P)T}$. Assume that the matrices $\boldsymbol{A}^{(P)}$ and $\left(\boldsymbol{A}^{(1)} \odot \cdots \odot \boldsymbol{A}^{(P-1)} \odot \boldsymbol{A}^{(P+1)} \odot \cdots \odot \boldsymbol{A}^{(N)}\right)$ have full column rank and $\boldsymbol{A}^{(P)}$ is an upper or lower banded Toeplitz matrix with*

$$b_L\left(\boldsymbol{a}_R^{(P)}\right) \leq I_P \text{ and } b_L\left(\boldsymbol{a}_R^{(P)}\right) - R \geq 1 \tag{17}$$

or

$$b_U\left(\boldsymbol{a}_1^{(P)}\right) \geq 1 \text{ and } b_U\left(\boldsymbol{a}_1^{(P)}\right) + R \leq I_P \tag{18}$$

in the respective cases. Let $\widehat{\boldsymbol{A}}^{(P)}$ be another matrix with same structure as $\boldsymbol{A}^{(P)}$, then $Row\left(\widehat{\boldsymbol{A}}^{(P)T}\right) \subseteq Row\left(\boldsymbol{A}^{(P)T}\right) = Row\left(\boldsymbol{T}_{[P]}\right)$ if and only if $\widehat{\boldsymbol{A}}^{(P)} = \boldsymbol{A}^{(P)}\alpha$, where $\alpha \in \mathbb{C}$.

Proof. *Assume that $\widehat{\boldsymbol{A}}^{(P)} = \boldsymbol{A}^{(P)}\alpha$ for some $\alpha \in \mathbb{C}$, then obviously $Row\left(\widehat{\boldsymbol{A}}^{(P)T}\right) \subseteq Row\left(\boldsymbol{A}^{(P)T}\right) = Row\left(\boldsymbol{T}_{[P]}\right)$. Conversely, assume that $b_U\left(\boldsymbol{a}_1^{(P)}\right) \geq 1$, $b_U\left(\boldsymbol{a}_1^{(P)}\right) + R \leq I_P$ and $Row\left(\widehat{\boldsymbol{A}}^{(P)T}\right) \subseteq Row\left(\boldsymbol{A}^{(P)T}\right) = Row\left(\boldsymbol{T}_{[P]}\right)$. Then there exists a nonsingular matrix $\boldsymbol{M} \in \mathbb{C}^{R\times R}$ such that*

$$\widehat{\boldsymbol{A}}^{(P)} \;=\; \boldsymbol{A}^{(P)}\boldsymbol{M}.$$

According to (18) *the matrices* $\widehat{\mathbf{A}}^{(P)}$ *and* $\mathbf{A}^{(P)}$ *are lower banded and therefore* $\mathbf{M}$ *must be lower triangular.*

We have that $\hat{a}_{b_U(\mathbf{a}_1^{(P)})+i-1,i} = a_{b_U(\mathbf{a}_1^{(P)})+i-1,i} m_{i,i}$, $\forall i \in [1, R]$ *which together with the relations* $\hat{a}_{b_U(\mathbf{a}_1^{(P)})+i-1,i} = \hat{a}_{b_U(\mathbf{a}_1^{(P)}),1}$, $a_{b_U(\mathbf{a}_1^{(P)})+i-1,i} = a_{b_U(\mathbf{a}_1^{(P)}),1}$, $\forall i \in [1, R]$ *implies that* $m_{i,i} = m_{R,R}$, $\forall i \in [1, R]$. *Since the lower triangular matrix* $\mathbf{M}$ *is nonsingular we also have that* $m_{R,R} \neq 0$ *and hence* $m_{i,i} \neq 0$, $\forall i \in [1, R]$.

Next, we have that $\hat{a}_{b_U(\mathbf{a}_1^{(P)})+i,i} = \hat{a}_{b_U(\mathbf{a}_1^{(P)})+1,1}$ *and* $a_{b_U(\mathbf{a}_1^{(P)})+i,i} = a_{b_U(\mathbf{a}_1^{(P)})+1,1}$, $\forall i \in [1, R]$, *and*

$$\hat{a}_{b_U(\mathbf{a}_1^{(P)})+i,i} = \begin{cases} a_{b_U(\mathbf{a}_1^{(P)})+i,i} m_{i,i} + a_{b_U(\mathbf{a}_1^{(P)})+i,i+1} m_{i+1,i} & , i \in [1, R-1] \\ a_{b_U(\mathbf{a}_1^{(P)})+i,i} m_{i,i} & , i = R \end{cases}$$

The fact that $m_{i,i} = m_{1,1}$, $\forall i \in [1, R]$ *and the assumption* $a_{b_U(\mathbf{a}_1^{(P)})+i,i+1} \neq 0$, $\forall i \in [1, R-1]$ *implies that* $m_{i+1,i} = 0$, $\forall i \in [1, R-1]$.

Furthermore, we have that that $\hat{a}_{b_U(\mathbf{a}_1^{(P)})+i+1,i} = \hat{a}_{b_U(\mathbf{a}_1^{(P)})+2,1}$ *and* $a_{b_U(\mathbf{a}_1^{(P)})+i+1,i} = a_{b_U(\mathbf{a}_1^{(P)})+2,1}$, $\forall i \in [1, R-1]$, *and*

$$\hat{a}_{b_U(\mathbf{a}_1^{(P)})+i+1,i} = \begin{cases} a_{b_U(\mathbf{a}_1^{(P)})+i+1,i} m_{i,i} + a_{b_U(\mathbf{a}_1^{(P)})+i+1,i+2} m_{i+2,i} & , i \in [1, R-2] \\ a_{b_U(\mathbf{a}_1^{(P)})+i+1,i} m_{i,i} & , i = R-1 \end{cases}$$

The fact that $m_{i,i} = m_{1,1}$, $\forall i \in [1, R]$ *and the assumption* $a_{b_U(\mathbf{a}_1^{(P)})+i+1,i+2} \neq 0$, $\forall i \in [1, R-2]$ *implies that* $m_{i+2,i} = 0$, $\forall i \in [1, R-2]$. *By repeating this procedure we can conclude that* $\mathbf{M}$ *must be upper triangular. Thus,* $\widehat{\mathbf{A}}^{(P)} = \mathbf{A}^{(P)}\alpha$, *where* $\alpha \in \mathbb{C}$.

The proof for the case when $\mathbf{A}^{(P)}$ *is an upper banded Toeplitz matrix satisfying the assumptions* (17) *is analogue. Notice that* $\mathbf{J}_{I_P}\mathbf{A}^{(P)}\mathbf{J}_R$ *is a lower banded Toeplitz matrix satisfying the assumptions* (18). *Hence the result also follows directly from the above reasoning.*

Let us now state a similar result for Hankel matrix factors.

Proposition 5.4. *Consider* $\mathcal{T} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ *with rank R and matrix representation* $\mathbf{T}_{[P]} = \left(\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(P-1)} \odot \mathbf{A}^{(P+1)} \odot \cdots \odot \mathbf{A}^{(N)}\right) \mathbf{A}^{(P)T}$. *Assume that the matrices* $\mathbf{A}^{(P)}$ *and* $\left(\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(P-1)} \odot \mathbf{A}^{(P+1)} \odot \cdots \odot \mathbf{A}^{(N)}\right)$ *have full column rank and* $\mathbf{A}^{(P)}$

is an anti-upper or anti-lower banded Hankel matrix with

$$b_L\left(\boldsymbol{a}_1^{(P)}\right) \leq I_P \text{ and } b_L\left(\boldsymbol{a}_1^{(P)}\right) - R \geq 1 \tag{19}$$

or

$$b_U\left(\boldsymbol{a}_R^{(P)}\right) \geq 1 \text{ and } b_U\left(\boldsymbol{a}_R^{(P)}\right) + R \leq I_P \tag{20}$$

in the respective cases. Let $\widehat{\boldsymbol{A}}^{(P)}$ *be another matrix with same structure as* $\boldsymbol{A}^{(P)}$*, then* $Row\left(\widehat{\boldsymbol{A}}^{(P)T}\right) \subseteq Row\left(\boldsymbol{A}^{(P)T}\right) = Row\left(\boldsymbol{T}_{[P]}\right)$ *if and only if* $\widehat{\boldsymbol{A}}^{(P)} = \boldsymbol{A}^{(P)}\alpha$*, where* $\alpha \in \mathbb{C}$*.*

Proof. *Assume that* $\widehat{\boldsymbol{A}}^{(P)} = \boldsymbol{A}^{(P)}\alpha$ *for some* $\alpha \in \mathbb{C}$*, then obviously* $Row\left(\widehat{\boldsymbol{A}}^{(P)T}\right) \subseteq Row\left(\boldsymbol{A}^{(P)T}\right) = Row\left(\boldsymbol{T}_{[P]}\right)$*. Conversely, assume that* $b_L\left(\boldsymbol{a}_1^{(P)}\right) \leq I_P$*,* $b_L\left(\boldsymbol{a}_1^{(P)}\right) - R \geq 1$ *and* $Row\left(\widehat{\boldsymbol{A}}^{(P)T}\right) \subseteq Row\left(\boldsymbol{A}^{(P)T}\right) = Row\left(\boldsymbol{T}_{[P]}\right)$*. Then there exists a nonsingular matrix* $\boldsymbol{M} \in \mathbb{C}^{R\times R}$ *such that*

$$\widehat{\boldsymbol{A}}^{(P)} = \boldsymbol{A}^{(P)}\boldsymbol{M}.$$

We have

$$\widehat{\boldsymbol{A}}^{(P)}\boldsymbol{J}_R = \boldsymbol{A}^{(P)}\boldsymbol{J}_R\boldsymbol{J}_R^T\boldsymbol{M}\boldsymbol{J}_R = \boldsymbol{A}^{(P)}\boldsymbol{J}_R\overline{\boldsymbol{M}},$$

where $\overline{\boldsymbol{M}} = \boldsymbol{J}_R^T\boldsymbol{M}\boldsymbol{J}_R$ *is a nonsingular matrix. Since* $\widehat{\boldsymbol{A}}^{(P)}\boldsymbol{J}_R$ *and* $\boldsymbol{A}^{(P)}\boldsymbol{J}_R$ *are upper banded Toeplitz matrices satisfying the inequalities* (17) *the result follows from proposition 5.3.*

The proof for the case when $\boldsymbol{A}^{(P)}$ *is an anti-lower banded Hankel matrix satisfying the assumptions* (20) *is analogue. In fact, since* $\boldsymbol{A}^{(P)}\boldsymbol{J}_R$ *is a lower banded Toeplitz matrix satisfying the inequalities* (18) *the result follows from proposition 5.3.*

6. Computation of CPDs with Banded and Structured Matrix Factors

Subsection 6.1 and 6.2 explains how to compute a Nth order CPD by exploiting that one or two of the matrix factors are banded and structured matrix factors, respectively. The presented method can be generalized in such a way that it is capable of jointly exploiting that three or more of the matrix factors are banded and structured.

6.1. Exploiting one Banded and Structured Matrix Factor

In this subsection we propose a method which exploits that one of the matrix factors is banded and structured. When the matrix factor is Toeplitz or Hankel structured, then the method will be referred to as CPTOEP1 and CPHANK1, respectively. Let $\mathcal{T} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ with rank R and let $\mathbf{T}_{[N]} = \left(\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(N-1)}\right) \mathbf{A}^{(N)T}$ be its matrix representation. Assume that the matrices $\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(N-1)}$ and $\mathbf{A}^{(N)}$ have full column rank and that $\mathbf{A}^{(N)}$ is a banded and structured matrix with $\dim\left(\mathbf{A}^{(N)}\right) = L_N$. Let $\mathbf{T}_{[N]} = \mathbf{U}\Sigma\mathbf{V}^H$ denote the compact SVD of $\mathbf{T}_{[N]}$, then there exists a nonsingular matrix $\mathbf{M} \in \mathbb{C}^{R \times R}$ such that

$$
\begin{aligned}
\mathbf{U}\Sigma\mathbf{M}^{-1} &= \mathbf{A}^{(1)} \odot \mathbf{A}^{(2)} \odot \cdots \odot \mathbf{A}^{(N-1)} && (21)\\
\mathbf{M}\mathbf{V}^H &= \mathbf{A}^{(N)T}. && (22)
\end{aligned}
$$

If $\mathbf{A}^{(N)}$ is a banded and structured matrix satisfying the assumptions stated in proposition 5.2, 5.3 or 5.4, then we can find $\mathbf{M}$ and $\mathbf{A}^{(N)}$ from $\text{Row}\left(\mathbf{T}_{[N]}\right)$. Equation (22) can be written as

$$
\mathbf{M}\mathbf{V}^H = \mathbf{A}^{(N)T} = \sum_{l=1}^{L_N} a_l^{(N)} \mathbf{E}_l^{(N)T}, \tag{23}
$$

where $\mathbf{V} \in \mathbb{C}^{I_N \times R}$ and $\{\mathbf{E}_l^{(N)}\}$ is an orthonormal basis for the matrix $\mathbf{A}^{(N)}$ and $\mathbf{a}^{(N)} = \left[a_1^{(N)}, \ldots, a_{L_N}^{(N)}\right]^T \in \mathbb{C}^{L_N}$ is its coefficient vector. Equation (23) is equivalent to

$$
(\mathbf{V}^* \otimes \mathbf{I}_R)\,\text{vec}\,(\mathbf{M}) = \left[\text{vec}\left(\mathbf{E}_1^{(N)T}\right), \ldots, \text{vec}\left(\mathbf{E}_{L_N}^{(N)T}\right)\right] \mathbf{a}^{(N)}
$$

and

$$
\mathbf{G}\begin{bmatrix} \text{vec}\,(\mathbf{M}) \\ \mathbf{a}^{(N)} \end{bmatrix} = \mathbf{0}, \tag{24}
$$

where

$$
\mathbf{G} = \left[\mathbf{V}^* \otimes \mathbf{I}_R, -\text{vec}\left(\mathbf{E}_1^{(N)T}\right), \ldots, -\text{vec}\left(\mathbf{E}_{L_N}^{(N)T}\right)\right] \in \mathbb{C}^{I_N R \times (R^2 + L_N)}.
$$

From proposition 5.2, 5.3 or 5.4 we know that the solution to (24) is unique up to a scaling ambiguity. The LS solution to (24) is given by the right singular vector of the matrix $\mathbf{G}$ associated with its smallest singular value. Let $\mathbf{x} \in \mathbb{C}^{R^2+L_N}$ be the LS solution to the system (24), then

$$\operatorname{vec}(\mathbf{M}) = \mathbf{x}\left(1 : R^2\right)$$
$$\mathbf{a}^{(N)} = \mathbf{x}\left(R^2 + 1 : R^2 + 1 + L_N\right).$$

Let $\mathbf{F} = \mathbf{U}\Sigma\mathbf{M}^{-1}$, then from (21) the remaining matrix factors follow from the R decoupled best rank-1 tensor approximation problems

$$\min_{\mathbf{a}_r^{(1)},\ldots,\mathbf{a}_r^{(N-1)}} \left\|\mathbf{f}_r - \mathbf{a}_r^{(1)} \otimes \cdots \otimes \mathbf{a}_r^{(N-1)}\right\|_F^2, \quad r \in [1, R], \tag{25}$$

where $\mathbf{f}_r$ and $\mathbf{a}_r^{(n)}$ denotes the rth column vector of $\mathbf{F}$ and $\mathbf{A}^{(n)}$, respectively. When $N = 2$, then $\mathbf{T}_{[2]} = \mathbf{A}^{(1)}\mathbf{A}^{(2)T}$ and we get $\mathbf{A}^{(1)} = \mathbf{T}_{[2]}\left(\mathbf{A}^{(2)T}\right)^{\dagger}$ or alternatively $\mathbf{A}^{(1)} = \mathbf{U}\Sigma\mathbf{M}^{-1}$. When $N = 3$, then (25) reduces to solving the R decoupled best rank-1 matrix approximation problems

$$\min_{\mathbf{a}_r^{(1)},\mathbf{a}_r^{(2)}} \left\|\mathbf{f}_r - \mathbf{a}_r^{(1)} \otimes \mathbf{a}_r^{(2)}\right\|_F^2, \quad r \in [1, R],$$

which can be solved by standard numerical linear algebra methods.

If the CPD contains several banded and structured matrix factors, then under certain conditions we can compute them in a successive manner as follows. Consider $\mathcal{T} \in \mathbb{C}^{I_1\times\cdots\times I_N}$ with rank R. Assume that the matrix factors $\mathbf{A}^{(p)}$, $1 \leq p \leq P \leq N$ are banded and structured and they satisfy the assumptions stated in proposition 5.2, 5.3 or 5.4. Furthermore, assume that the involved matrices of the matrix representations $\mathbf{T}_{[p]}$ only consist of full column rank matrices $\forall p \in [1, P]$. By applying the procedure described in this subsection we can obtain $\mathbf{A}^{(p)}$ from $\operatorname{Row}\left(\mathbf{T}_{[p]}\right)$. By repeating this procedure the remaining banded and structured matrix factors can be found. Once the banded and structured matrix factors have been found, then the remaining unstructured matrix factors follow from R decoupled best rank-1 tensor approximation problems as earlier explained.

Let us be more specific for the case $P = 2$. Assume that $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ are banded and structured and satisfying the assumptions stated in proposition 5.2, 5.3 or 5.4. Assume also that the matrices $\mathbf{A}^{(1)}$, $\mathbf{A}^{(2)}$, $\left(\mathbf{A}^{(2)} \odot \mathbf{A}^{(3)} \odot \cdots \odot \mathbf{A}^{(N)}\right)$, $\left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(3)} \odot \cdots \odot \mathbf{A}^{(N)}\right)$ and $\left(\mathbf{A}^{(3)} \odot \mathbf{A}^{(4)} \odot \cdots \odot \mathbf{A}^{(N)}\right)$ have full column rank.

We first compute $\mathbf{A}^{(1)}$. Let $\mathbf{T}_{[1]} = \mathbf{U}\Sigma\mathbf{V}^H$ denote the compact SVD of $\mathbf{T}_{[1]}$, then there exists a nonsingular matrix $\mathbf{N} \in \mathbb{C}^{R\times R}$ such that $\mathbf{V}^*\mathbf{N} = \mathbf{A}^{(1)}$. By applying the procedure described in this subsection we obtain $\mathbf{A}^{(1)}$ from $\text{Row}(\mathbf{T}_{[1]})$.

Next, we compute $\mathbf{A}^{(2)}$. Let $\mathbf{T}_{[2]} = \mathbf{U}\Sigma\mathbf{V}^H$ denote the compact SVD of $\mathbf{T}_{[2]}$, then there exists a nonsingular matrix $\mathbf{N} \in \mathbb{C}^{R\times R}$ such that $\mathbf{V}^*\mathbf{N} = \mathbf{A}^{(2)}$. Again, by applying the procedure described in this subsection we obtain $\mathbf{A}^{(2)}$ from $\text{Row}(\mathbf{T}_{[2]})$.

In the final step we find the remaining matrix factors. Let

$$\mathbf{T}^{[2]} = \left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right)\left(\mathbf{A}^{(3)} \odot \cdots \odot \mathbf{A}^{(N)}\right)^T = \overline{\mathbf{U}}\,\Sigma\,\overline{\mathbf{V}}^H$$

denote the compact SVD of $\mathbf{T}^{[2]}$, then there exists a nonsingular matrix $\mathbf{M} \in \mathbb{C}^{R\times R}$ such that $\overline{\mathbf{U}}\mathbf{M} = \left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right) \Leftrightarrow \mathbf{M} = \overline{\mathbf{U}}^H\left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right)$. We have $\mathbf{F} = \overline{\mathbf{V}}^*\overline{\Sigma}\mathbf{M}^{-T} = \mathbf{A}^{(3)} \odot \cdots \odot \mathbf{A}^{(N)}$ and the remaining matrix factors follow from the R decoupled rank-1 tensor approximation problems

$$\min_{\mathbf{a}_r^{(3)},\ldots,\mathbf{a}_r^{(N)}} \left\|\mathbf{f}_r - \mathbf{a}_r^{(3)} \otimes \cdots \otimes \mathbf{a}_r^{(N)}\right\|_F^2, \quad r \in [1, R],$$

where $\mathbf{f}_r$ and $\mathbf{a}_r^{(n)}$ denotes the rth column vector of $\mathbf{F}$ and $\mathbf{A}^{(n)}$, respectively. Remark that when $N = 3$, then $\mathbf{A}^{(3)}$ follows *directly* from $\mathbf{A}^{(3)T} = \left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right)^\dagger \mathbf{T}^{[2]}$ or alternatively $\mathbf{A}^{(3)} = \overline{\mathbf{V}}^*\overline{\Sigma}\mathbf{M}^{-T}$. When $N = 4$, then $\mathbf{F} = \overline{\mathbf{V}}^*\overline{\Sigma}\mathbf{M}^{-T} = \mathbf{A}^{(3)} \odot \mathbf{A}^{(4)}$ which means that $\mathbf{A}^{(3)}$ and $\mathbf{A}^{(4)}$ follows from solving the R decoupled rank-1 *matrix* approximation problems

$$\min_{\mathbf{a}_r^{(3)},\mathbf{a}_r^{(4)}} \left\|\mathbf{f}_r - \mathbf{a}_r^{(3)} \otimes \mathbf{a}_r^{(4)}\right\|_F^2, \quad r \in [1, R],$$

which can be solved by standard numerical linear algebra methods.

6.2. Exploiting Jointly two Banded and Structured Matrix Factors

In this subsection we propose a method which exploits that two of the matrix factors are banded and structured. When both the structured matrix factors are Toeplitz or Hankel structured, then the method will be referred to as CPTOEP2 and CPHANK2, respectively.

Let $\mathcal{T} \in \mathbb{C}^{I_1\times\cdots\times I_N}$ with rank R and let $\mathbf{T}^{[2]} = \left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right)\left(\mathbf{A}^{(3)} \odot \cdots \odot \mathbf{A}^{(N)}\right)^T$ be its matrix representation. Assume that the matrices $\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}$ and

$\mathbf{A}^{(3)} \odot \cdots \odot \mathbf{A}^{(N)}$ have full column rank. Assume also that $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ are banded and structured matrices with $\dim\left(\mathbf{A}^{(1)}\right) = L_1$, $\dim\left(\mathbf{A}^{(2)}\right) = L_2$ and satisfying the assumptions stated in proposition 5.2. According to proposition 5.2 we can find the matrix factors $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ from $\mathrm{Col}\left(\mathbf{T}^{[2]}\right)$. Let $\mathbf{T}^{[2]} = \mathbf{U}\Sigma\mathbf{V}^H$ denote the compact SVD of $\mathbf{T}^{[2]}$, then there exists a nonsingular matrix $\mathbf{M} \in \mathbb{C}^{R\times R}$ such that

$$\begin{aligned} \mathbf{UM} &= \mathbf{A}^{(1)} \odot \mathbf{A}^{(2)} && (26)\\ \mathbf{V}^*\Sigma\mathbf{M}^{-T} &= \mathbf{A}^{(3)} \odot \cdots \odot \mathbf{A}^{(N)}. && (27) \end{aligned}$$

Let $\{\mathbf{E}_{l_p}^{(p)}\}$ be an orthonormal basis for $\mathbf{A}^{(p)}$ and $\mathbf{a}^{(p)} = \left[a_1^{(p)}, \ldots, a_{L_p}^{(p)}\right]^T \in \mathbb{C}^{L_p}$ be its coefficient vector such that $\mathbf{A}^{(p)} = \sum_{l_p=1}^{L_p} a_{l_p}^{(p)} \mathbf{E}_{l_p}^{(p)}$, where $p \in \{1, 2\}$. Equation (26) can now be written as

$$(\mathbf{I}_R \otimes \mathbf{U})\,\mathrm{vec}\,(\mathbf{M}) = \left[\mathrm{vec}\left(\mathbf{E}_1^{(1)} \odot \mathbf{E}_1^{(2)}\right), \ldots, \mathrm{vec}\left(\mathbf{E}_{L_1}^{(1)} \odot \mathbf{E}_{L_2}^{(2)}\right)\right]\left(\mathbf{a}^{(1)} \otimes \mathbf{a}^{(2)}\right)$$

and

$$\mathbf{G}\begin{bmatrix} \mathrm{vec}\,(\mathbf{M}) \\ \mathbf{a}^{(1)} \otimes \mathbf{a}^{(2)} \end{bmatrix} = \mathbf{0}, \qquad (28)$$

where

$$\mathbf{G} = \left[\mathbf{I}_R \otimes \mathbf{U}, -\mathrm{vec}\left(\mathbf{E}_1^{(1)} \odot \mathbf{E}_1^{(2)}\right), \ldots, -\mathrm{vec}\left(\mathbf{E}_{L_1}^{(1)} \odot \mathbf{E}_{L_2}^{(2)}\right)\right] \in \mathbb{C}^{I_1 I_2 R \times (R^2 + L_1 L_2)}.$$

From proposition 5.2 we know that solution to (28) is unique up to a scalar ambiguity. In order to solve the system (28) we will ignore the Kronecker product structure $\mathbf{a}^{(1)} \otimes \mathbf{a}^{(2)}$. Remark this relaxation may affect the uniqueness of the problem. However, numerical experiments indicate that this relaxation step does not seem to affect the uniqueness of the solution. The LS solution to the relaxed version of (28) is given by the right singular vector of the matrix $\mathbf{G}$ associated with its smallest singular value. Let $\mathbf{x} \in \mathbb{C}^{R^2 + L_1 L_2}$ be the LS solution to the relaxed version of the system (28), then we set

$$\begin{aligned} \mathrm{vec}\,(\mathbf{M}) &= \mathbf{x}\left(1 : R^2\right) \\ \mathbf{y} &= \mathbf{x}\left(R^2 + 1 : R^2 + L_1 L_2\right) = \mathbf{a}^{(1)} \otimes \mathbf{a}^{(2)}. \end{aligned}$$

Let $\mathbb{C}^{L_2 \times L_1} \ni \mathbf{Y} = \text{Unvec}\left(\mathbf{x}\left(R^2+1 : R^2 + L_1 L_2\right)\right)$, then we find the unknown coefficient vectors $\mathbf{a}^{(1)}$ and $\mathbf{a}^{(2)}$ up to scale factor ambiguities from the rank-1 matrix approximation problem

$$\min_{\mathbf{a}^{(1)},\mathbf{a}^{(2)}} \left\| \mathbf{Y} - \mathbf{a}^{(2)}\mathbf{a}^{(1)T} \right\|_F^2 ,$$

which can be solved by standard numerical linear algebra methods. Remark that this method is able to take a partial symmetry $\mathbf{a}^{(1)} = \mathbf{a}^{(2)}$ or a partial Hermitian symmetry $\mathbf{a}^{(1)} = \mathbf{a}^{(2)*}$ into account.

Once the matrices $\mathbf{M}$, $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ have been found, then from (27) we have $\mathbf{F} = \mathbf{V}^* \Sigma \mathbf{M}^{-T} = \mathbf{A}^{(3)} \odot \cdots \odot \mathbf{A}^{(N)}$. This implies that the remaining matrix factors follow from the R decoupled rank-1 tensor approximation problems

$$\min_{\mathbf{a}_r^{(3)},\ldots,\mathbf{a}_r^{(N)}} \left\| \mathbf{f}_r - \mathbf{a}_r^{(3)} \otimes \cdots \otimes \mathbf{a}_r^{(N)} \right\|_F^2 , \quad r \in [1, R] ,$$

where $\mathbf{f}_r$ and $\mathbf{a}_r^{(n)}$ denotes the rth column vector of $\mathbf{F}$ and $\mathbf{A}^{(n)}$, respectively.

Again, remark that if $N = 3$, then $\mathbf{A}^{(3)}$ follows from $\mathbf{A}^{(3)T} = \left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right)^{\dagger} \mathbf{T}^{[2]}$ or alternatively $\mathbf{A}^{(3)} = \mathbf{V}^* \Sigma \mathbf{M}^{-T}$. When $N = 4$, then $\mathbf{F} = \mathbf{V}^* \Sigma \mathbf{M}^{-T} = \mathbf{A}^{(3)} \odot \mathbf{A}^{(4)}$ which means that $\mathbf{A}^{(3)}$ and $\mathbf{A}^{(4)}$ follow from solving the R decoupled rank-1 matrix approximation problems

$$\min_{\mathbf{a}_r^{(3)},\mathbf{a}_r^{(4)}} \left\| \mathbf{f}_r - \mathbf{a}_r^{(3)} \otimes \mathbf{a}_r^{(4)} \right\|_F^2 , \quad r \in [1, R] ,$$

which can be solved by standard numerical linear algebra methods.

7. Numerical Experiments

Let us restrict the simulation study to tensors of order $N = 3$. The real and imaginary entries of all the involved matrix factors and tensors are randomly drawn from a uniform distribution with support $[-\frac{1}{2}, \frac{1}{2}]$. Let $\mathcal{T} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ with rank R denote the structured tensor we attempt to estimate from the observed tensor $\mathcal{X} = \mathcal{T} + \beta \mathcal{N}$, where $\mathcal{N} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ is a perturbation tensor and $\beta \in \mathbb{R}$. The following Signal-to-Noise Ratio (SNR) measure will be used

$$\text{SNR [dB]} = 10 \log \left(\frac{\left\| \mathbf{T}_{[1]} \right\|_F^2}{\left\| \beta \mathbf{N}_{[1]} \right\|_F^2} \right) .$$

Furthermore, when the matrix $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R}$ is banded, the following performance measure will also be used

$$\mathrm{P}_{\Lambda}\left(\mathbf{A}^{(n)}\right) = \min_{\Lambda} \frac{\left\|\mathbf{A}^{(n)} - \widehat{\mathbf{A}}^{(n)}\Lambda\right\|_F}{\left\|\mathbf{A}^{(n)}\right\|_F},$$

where $\widehat{\mathbf{A}}^{(n)}$ denotes the estimated matrix factor and Λ denotes a diagonal matrix. When the matrix $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R}$ is banded and Hankel or Toeplitz structured, we will use the following function

$$\mathrm{P}_{\alpha}\left(\mathbf{A}^{(n)}\right) = \min_{\alpha \in \mathbb{C}} \frac{\left\|\mathbf{A}^{(n)} - \widehat{\mathbf{A}}^{(n)}\alpha\right\|_F}{\left\|\mathbf{A}^{(n)}\right\|_F},$$

as a performance measure. To measure the elapsed time in second used to execute the algorithms in MATLAB, the built-in functions *tic*$(\cdot)$ and *toc*$(\cdot)$ is used.

We compare the presented methods with the popular ALS method. The ALS method is randomly initialized and we decide that the ALS method has converged when the applied cost function at iteration k and $k+1$ has changed less than $\epsilon_{\mathrm{ALS}} = 10^{-8}$ or the number of iterations has exceeded 2000. For the ALS method we will use the following performance measure

$$\mathrm{P}_{\Pi\Lambda}\left(\mathbf{A}^{(n)}\right) = \min_{\Pi\Lambda} \frac{\left\|\mathbf{A}^{(n)} - \widehat{\mathbf{A}}^{(n)}\Pi\Lambda\right\|_F}{\left\|\mathbf{A}^{(n)}\right\|_F},$$

where $\widehat{\mathbf{A}}^{(n)}$ denotes the estimated matrix factor, Π a permutation matrix and Λ denotes a diagonal matrix. In order to find Π and Λ the greedy LS column matching algorithm between $\mathbf{A}^{(n)}$ and $\widehat{\mathbf{A}}^{(n)}$ proposed in [21] is used. We note in passing that other methods for determining the permutation ambiguity have been proposed in [6].

7.1. Banded Matrix Factors

In all the simulations in this subsection we set $I_1 = 10$, $I_2 = 7$, $I_3 = 5$ and $R = 5$, the data is complex, i.e., $\mathcal{T}, \mathcal{N} \in \mathbb{C}^{10\times 7\times 5}$, and the matrix factor $\mathbf{A}^{(3)} \in \mathbb{C}^{5\times 5}$ is an unstructured matrix. When the CPBAND1 and CPBAND2 methods are followed by an ALS refinement step of at most 200 ALS iterations, then the refined CPBAND1 and CPBAND2 methods will be referred to as CPBAND1-ALS and CPBAND2-ALS, respectively.

Case 1. Let $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ be narrow lower and upper banded matrices of the form

$$\mathbf{A}^{(1)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ x & 0 & 0 & 0 & 0 \\ x & x & 0 & 0 & 0 \\ x & x & x & 0 & 0 \\ x & x & x & x & 0 \\ 0 & x & x & x & x \\ 0 & 0 & x & x & x \\ 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & x \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \in \mathbb{C}^{10\times 5}, \quad \mathbf{A}^{(2)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ x & 0 & 0 & 0 & 0 \\ x & x & 0 & 0 & 0 \\ 0 & x & x & 0 & 0 \\ 0 & 0 & x & x & 0 \\ 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & x \end{bmatrix} \in \mathbb{C}^{7\times 5},$$

where x denotes a non-zero entry. Remark that since the matrix factors are lower and upper banded both CPBAND1 and CPBAND2 are valid methods. The mean and standard deviation (std) $\mathrm{P}_{\Pi\Lambda}\left(\mathbf{A}^{(n)}\right)$, $\mathrm{P}_{\Lambda}\left(\mathbf{A}^{(n)}\right)$ and time values over 100 trials as a function of SNR can be seen in the left column of figure 1. We notice that the CPBAND2 method works better than the CPBAND1 method. At low SNR the CPBAND2 method also works better than the ALS method while at high SNR the ALS method performs better than the CPBAND2 method. A reason that the ALS method performs better than the CPBAND2 at high SNR could be due to the relaxation step of the Kronecker product structure in the CPBAND2 method. However, the ALS method is also more costly than the CPBAND2 method. We also notice that the CPBAND1-ALS and CPBAND2-ALS methods yield a similar performance as the ALS method but at a lower computational cost.

Case 2. Let the matrix factor $\mathbf{A}^{(1)}$ be a upper banded matrix and let the $\mathbf{A}^{(2)}$ be a lower banded matrix of the forms

$$\mathbf{A}^{(1)} = \begin{bmatrix} x & x & 0 & 0 & 0 \\ x & x & x & 0 & 0 \\ x & x & x & x & 0 \\ x & x & x & x & x \\ x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & 0 & x & x & x \\ 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & x \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \in \mathbb{C}^{10\times 5}, \quad \mathbf{A}^{(2)} = \begin{bmatrix} x & 0 & 0 & 0 & 0 \\ x & x & 0 & 0 & 0 \\ x & x & x & 0 & 0 \\ x & x & x & x & 0 \\ x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \end{bmatrix} \in \mathbb{C}^{7\times 5},$$

where x denotes a non-zero entry. Remark that the matrix factors are not lower and upper banded, but the Khatri-Rao product of them is still lower and upper banded. This means that the CPBAND1 method is not valid while the CPBAND2 method still is. The mean and std $\mathrm{P}_{\Pi\Lambda}\left(\mathbf{A}^{(n)}\right)$, $\mathrm{P}_{\Lambda}\left(\mathbf{A}^{(n)}\right)$ and time values over 100 trials as a function of SNR can be seen in the right column of figure 1. As expected, we first notice that the CPBAND2 method works while the CPBAND1 method does not. The CPBAND2-ALS and the ALS methods yield a similar performance. However, the CPBAND2-ALS method is less costly than the ALS method.

7.2. Banded Toeplitz Matrix Factors

In all the simulations in this subsection we set $I_1 = 8$, $I_2 = 9$, $I_3 = 10$ and $R = 6$, the data is complex, i.e., $\mathcal{T}, \mathcal{N} \in \mathbb{C}^{8\times 9\times 10}$, and the matrix factor $\mathbf{A}^{(3)} \in \mathbb{C}^{10\times 6}$ is an unstructured matrix.

Case 1. Let $\mathbf{A}^{(1)}$ be a narrow lower and upper banded Toeplitz matrix and let $\mathbf{A}^{(2)}$ be a lower banded Toeplitz matrix of the forms

$$\mathbf{A}^{(1)} = \begin{bmatrix} a_1^{(1)} & 0 & 0 & 0 & 0 & 0 \\ a_2^{(1)} & a_1^{(1)} & 0 & 0 & 0 & 0 \\ 0 & a_2^{(1)} & a_1^{(1)} & 0 & 0 & 0 \\ 0 & 0 & a_2^{(1)} & a_1^{(1)} & 0 & 0 \\ 0 & 0 & 0 & a_2^{(1)} & a_1^{(1)} & 0 \\ 0 & 0 & 0 & 0 & a_2^{(1)} & a_1^{(1)} \\ 0 & 0 & 0 & 0 & 0 & a_2^{(1)} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \in \mathbb{C}^{8\times 6}, \quad \mathbf{A}^{(2)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ a_1^{(1)} & 0 & 0 & 0 & 0 & 0 \\ a_2^{(1)} & a_1^{(1)} & 0 & 0 & 0 & 0 \\ a_3^{(1)} & a_2^{(1)} & a_1^{(1)} & 0 & 0 & 0 \\ 0 & a_3^{(1)} & a_2^{(1)} & a_1^{(1)} & 0 & 0 \\ 0 & 0 & a_3^{(1)} & a_2^{(1)} & a_1^{(1)} & 0 \\ 0 & 0 & 0 & a_3^{(1)} & a_2^{(1)} & a_1^{(1)} \\ 0 & 0 & 0 & 0 & a_3^{(1)} & a_2^{(1)} \end{bmatrix} \in \mathbb{C}^{9\times 6}.$$

Remark that the conditions stated in propositions 5.2 and 5.3 are satisfied. The mean and std $\mathrm{P}_{\Pi\Lambda}\left(\mathbf{A}^{(1)}\right)$, $\mathrm{P}_{\Lambda}\left(\mathbf{A}^{(1)}\right)$ and time values over 100 trials as a function of SNR can be seen in the left column of figure 2. We notice that CPTOEP2 method performs better than the CPTOEP1 method which in turn performs better than the ALS method. We also notice that the ALS method is more costly than the CPTOEP1 and CPTOEP2 methods.

Case 2. Let $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ be lower banded Toeplitz matrices of the form

$$\mathbf{A}^{(1)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ a_1^{(1)} & 0 & 0 & 0 & 0 & 0 \\ a_2^{(1)} & a_1^{(1)} & 0 & 0 & 0 & 0 \\ a_3^{(1)} & a_2^{(1)} & a_1^{(1)} & 0 & 0 & 0 \\ a_4^{(1)} & a_3^{(1)} & a_2^{(1)} & a_1^{(1)} & 0 & 0 \\ a_5^{(1)} & a_4^{(1)} & a_3^{(1)} & a_2^{(1)} & a_1^{(1)} & 0 \\ 0 & a_5^{(1)} & a_4^{(1)} & a_3^{(1)} & a_2^{(1)} & a_1^{(1)} \\ 0 & 0 & a_5^{(1)} & a_4^{(1)} & a_3^{(1)} & a_2^{(1)} \end{bmatrix} \in \mathbb{C}^{8\times 6}, \quad \mathbf{A}^{(2)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ a_1^{(2)} & 0 & 0 & 0 & 0 & 0 \\ a_2^{(2)} & a_1^{(2)} & 0 & 0 & 0 & 0 \\ a_3^{(2)} & a_2^{(2)} & a_1^{(2)} & 0 & 0 & 0 \\ a_4^{(2)} & a_3^{(2)} & a_2^{(2)} & a_1^{(2)} & 0 & 0 \\ a_5^{(2)} & a_4^{(2)} & a_3^{(2)} & a_2^{(2)} & a_1^{(2)} & 0 \\ 0 & a_5^{(2)} & a_4^{(2)} & a_3^{(2)} & a_2^{(2)} & a_1^{(2)} \\ 0 & 0 & a_5^{(2)} & a_4^{(2)} & a_3^{(2)} & a_2^{(2)} \end{bmatrix} \in \mathbb{C}^{9\times 6}.$$

Remark that the condition stated in proposition 5.2 and 5.3 are satisfied. The mean and std $\mathrm{P}_{\Pi\Lambda}\left(\mathbf{A}^{(1)}\right)$, $\mathrm{P}_{\Lambda}\left(\mathbf{A}^{(1)}\right)$ and time values over 100 trials as a function of SNR can be seen in the right column of figure 2. We notice that CPTOEP2 method performs better than the ALS method which in turn performs better than the CPTOEP1 method. Again, we notice that the ALS method is more costly than the CPTOEP1 and CPTOEP2 methods.

7.3. Banded Hankel Matrix Factors and a Partial Symmetry

In all the simulations in this subsection we set $I_1 = 8$, $I_2 = 8$, $I_3 = 10$ and $R = 6$, the data is real, i.e., $\mathcal{T}, \mathcal{N} \in \mathbb{R}^{8\times 8\times 10}$, and the matrix factor $\mathbf{A}^{(3)} \in \mathbb{R}^{10\times 6}$ is an unstructured matrix.

Case 1. Let the $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ be anti-lower and anti-upper banded Hankel matrices of the form

$$\mathbf{A}^{(1)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & a_1^{(1)} \\ 0 & 0 & 0 & 0 & a_1^{(1)} & a_2^{(1)} \\ 0 & 0 & 0 & a_1^{(1)} & a_2^{(1)} & a_3^{(1)} \\ 0 & 0 & a_1^{(1)} & a_2^{(1)} & a_3^{(1)} & 0 \\ 0 & a_1^{(1)} & a_2^{(1)} & a_3^{(1)} & 0 & 0 \\ a_1^{(1)} & a_2^{(1)} & a_3^{(1)} & 0 & 0 & 0 \\ a_2^{(1)} & a_3^{(1)} & 0 & 0 & 0 & 0 \\ a_3^{(1)} & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{8\times 6}, \quad \mathbf{A}^{(2)} = \mathbf{A}^{(1)}.$$

Note that the conditions stated in propositions 5.2 and 5.4 are satisfied. The mean and std $\mathrm{P}_{\Pi\Lambda}\left(\mathbf{A}^{(1)}\right)$, $\mathrm{P}_{\Lambda}\left(\mathbf{A}^{(1)}\right)$ and time values over 100 trials as a function of SNR can be seen in the left column of figure 3. We notice that the CPHANK2 method performs better than CPHANK1 method which in turn performs better than the ALS method. We also notice that the ALS method is more costly than the CPHANK1 and CPHANK2 methods.

Case 2. Let the $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ be anti-upper banded Hankel matrices of the form

$$\mathbf{A}^{(1)} = \begin{bmatrix} 0 & a_1^{(1)} & a_2^{(1)} & a_3^{(1)} & a_4^{(1)} & a_5^{(1)} \\ a_1^{(1)} & a_2^{(1)} & a_3^{(1)} & a_4^{(1)} & a_5^{(1)} & 0 \\ a_2^{(1)} & a_3^{(1)} & a_4^{(1)} & a_5^{(1)} & 0 & 0 \\ a_3^{(1)} & a_4^{(1)} & a_5^{(1)} & 0 & 0 & 0 \\ a_4^{(1)} & a_5^{(1)} & 0 & 0 & 0 & 0 \\ a_5^{(1)} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{8\times 6}, \quad \mathbf{A}^{(2)} = \mathbf{A}^{(1)}.$$

Remark that the conditions stated in propositions 5.2 and 5.4 are not satisfied. The mean and std $\mathrm{P}_{\Pi\Lambda}\left(\mathbf{A}^{(1)}\right)$, $\mathrm{P}_{\Lambda}\left(\mathbf{A}^{(1)}\right)$ and time values over 100 trials as a function of SNR can be seen in the right column of figure 3. We notice that CPHANK2 performs better than ALS method while the CPHANK1 fails. Again, we notice that the ALS method is more costly than the CPHANK1 and CPHANK2 methods.

8. Conclusion

We first presented uniqueness results for CPDs with a banded and possibly also structured matrix factor. Furthermore, procedures for the computation of a CPD containing a banded and possibly also structured matrix factors were provided. It resulted in best rank-1 tensor approximation problems.

Next, we presented uniqueness results for CPDs with several banded and possibly also structured matrix factors. More relaxed uniqueness results were obtained when jointly taking the structure of several of the matrix factors into account. Procedures for the computation of a CPD containing several banded and possibly also structured matrix factors were also provided.

The results presented in this paper are also valid for more general tensor decompositions with banded matrix factors such as the Tucker decomposition [24] or the family of block tensor decompositions [11].

Numerical experiments showed that often an increase in performance can be expected when several of the banded and possibly also structured matrix factors are taken into account in the computation of the structured CPDs. In the case of banded matrix factors the numerical experiments indicated that the proposed methods can be used to speed up the popular ALS method. Numerical experiments performed in the presence of Toeplitz or Hankel matrix factors indicated that the proposed methods performed well, compared to ALS.

Directions of future research include the impact of the ignorance of the Kronecker structure when at least two matrix factors are banded or structured, and the generalization of our results to the case of non banded matrix factors.

Acknowledgments

The authors thank the reviewers for their careful reading and relevant and helpful comments.

References

[1] A. L. F. De Almeida and G. Favier and J. C. M. Mota, *PARAFAC-based Unified Tensor Modeling for Wireless Communication Systems with Application to Blind Multiuser Equalization*, Signal Processing, 87(2007), pp. 337-351.

[2] A. Belouchrani and B. Derras, *An Efficient Fourth Order System Identification (FOSI) Algorithm utilizing the Joint Diagonalization Procedure*, Proceedings of the Tenth IEEE Workshop on Statistical Signal and Array Processing, (2000), pp. 621-625.

[3] H. Bousbiah-Salah and A. Belouchrani and K. Abed-Meraim, *Jacobi-like algorithm for blind separation of convolutive mixtures*, Electronics Letter, 37(2001), pp. 1049–1050.

[4] J. Carroll and J. Chang, *Analysis of Individual Differences in Multidimensional Scaling via an N-way Generalization of "Eckart-Young" Decomposition*, Psychometrika, 9(1970), pp. 267-283.

[5] P. Comon, *MA Identification using Fourth Order Cumulants*, Signal Processing, 26(1992), pp. 381–388

[6] P. Comon and X. Luciani and A.L.F. De Almeida, *Tensor Decompositions, Alternating Least Squares and other Tales*, J. Chemometrics, 23(2009), pp. 393–405

[7] P. Comon and M. Sørensen and E. Tsigaridas, *Decomposing Tensors with Structured Matrix Factors Reduces to Rank-1 Approximations*, Proc. ICASSP, March 14-19, Dallas, USA, 2010

[8] L. De Lathauwer and 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., 21(2000), pp. 1324-1342.

[9] L. De Lathauwer and B. De Moor and J. Vandewalle, *Computation of the Canonical Decomposition by means of a Simultaneous Generalized Schur Decomposition*, SIAM J. Matrix Anal. Appl., 26(2004), pp. 295–327.

[10] L. De Lathauwer and A. de Baynast, *Blind Deconvolution of DS-CDMA Signals by Means of Decomposition in Rank-(1,L,L) Terms*, IEEE Trans. Signal Processing, 56(2008), pp. 1562–1571.

[11] L. De Lathauwer, *Decomposition of a Higher-Order Tensor in Block Terms –Part II: Definitions and Uniqueness*, SIAM J. Matrix Anal. Appl., 30(2008), pp. 1033–1066.

[12] C.E.R. Fernandes and G. Favier and J.M.M. Mota , *Blind Channel Identification Algorithms Based on the PARAFAC Decomposition of Cumulant Tensors: The Single User and Multiuser Cases*, Signal Processing, 88(2008), pp. 1382–1401.

[13] C.E.R. Fernandes and P. Comon and G. Favier, *Blind Identification of MISO-FIR Channels*, Signal Processing, 90(2010), pp. 490-503.

[14] G.H. Golub and C.F. Van Loan, *Matrix Computations*, The John Hopkins University Press, 3rd ed., 1996.

[15] R. A. Harshman, *Foundations of the Parafac Procedure: Models and Conditions for an Explanatory Multimodal Factor Analysis*, UCLA Working Papers in Phonetics, 16(1970), pp. 1–84.

[16] F. L. Hitchcock, *Multiple Invariants and Generalized Rank of a P-way Matrix or Tensor*, J. Math. and Phys., 7(1927), pp. 39–79.

[17] H. A. L. Kiers, *Towards a Standardized Notation and Terminology in Multiway Analysis*, J. Chemometrics, 14(2000), pp. 105–122

[18] T. G. Kolda and B. W. Bader, *Tensor Decompositions and Applications*, SIAM Review, 51(2009), pp. 455–500

[19] J. B. Kruskal, *Three-way Arrays: Rank and Uniqueness of Trilinear Decompositions, with Applications to Arithmetic Complexity and Statistics*, Linear Algebra and its Applications, 18(1977), pp. 95-138.

[20] D. Nion and L. De Lathauwer, *A Block Component Model-Based Blind DS-CDMA Receiver*, IEEE Trans. Signal Processing, 56(2008), pp. 5567–5579.

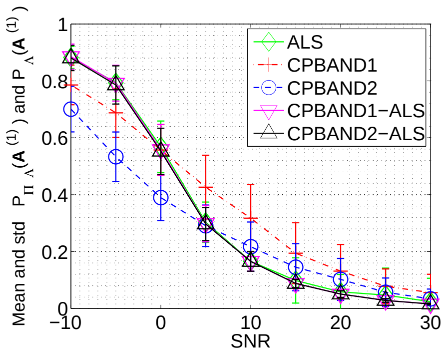
[21] N. D. Sidiropoulos and G. B. Giannakis and R. Bro, *Blind PARAFAC Receivers for DS-CDMA Systems*, IEEE Trans. Signal Process., 48(2000), pp. 810–823.

[22] L. Tong, *Identification of Multichannel MA Parameters using Higher-order Statistics*, Signal Processing, 53(1996), pp. 195–209.

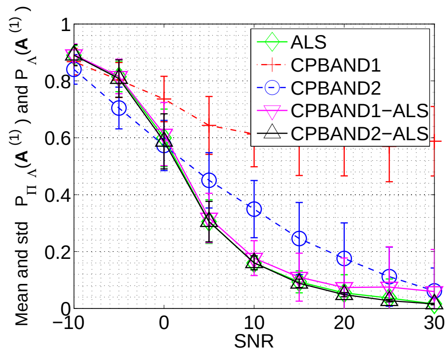
[23] J. M. F. Ten Berge and N. D. Sidiropoulos, *On Uniqueness in CANDECOMP/PARAFAC*, Psychometrika, 67(2002), pp. 399-409.

[24] L. R. Tucker, *Some Mathematical Notes on the Three-mode Factor Analysis*, Psychometrika, 31(1966), pp. 279-311.

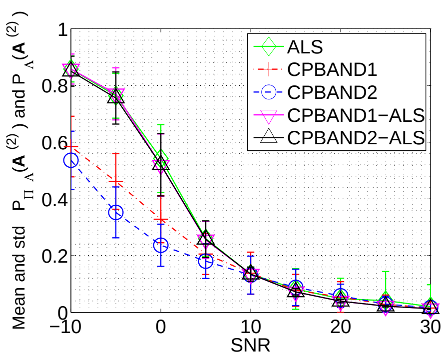
[25] A.-J. Van Der Veen and A. Paulraj, *Analytical Constant Modulus Algorithm*, IEEE Trans. Signal Process., 44(1996), pp. 1136–1155.



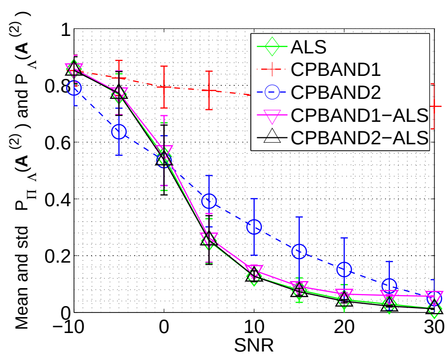
(a) Mean and std $P_{\Pi\Lambda}\left(\mathbf{A}^{(1)}\right)$ and $P_{\Lambda}\left(\mathbf{A}^{(1)}\right)$.



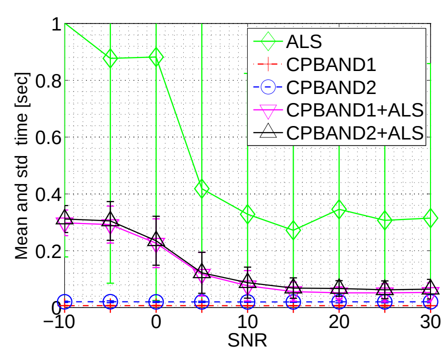
(b) Mean and std $P_{\Pi\Lambda}\left(\mathbf{A}^{(1)}\right)$ and $P_{\Lambda}\left(\mathbf{A}^{(1)}\right)$.



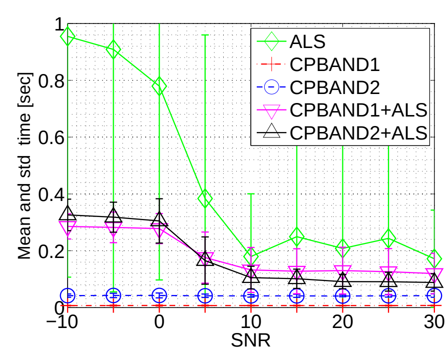
(c) Mean and std $P_{\Pi\Lambda}\left(\mathbf{A}^{(2)}\right)$ and $P_{\Lambda}\left(\mathbf{A}^{(2)}\right)$.



(d) Mean and std $P_{\Pi\Lambda}\left(\mathbf{A}^{(2)}\right)$ and $P_{\Lambda}\left(\mathbf{A}^{(2)}\right)$.

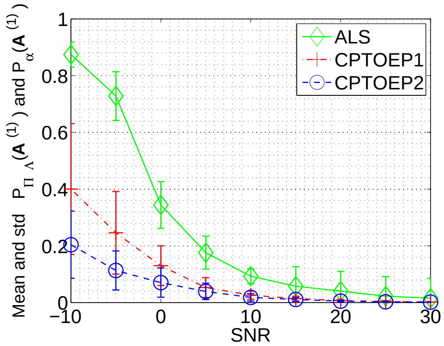


(e) Mean and std time.

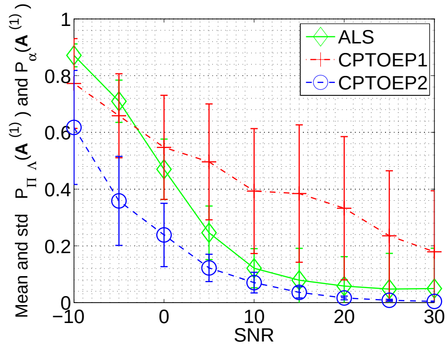


(f) Mean and std time.

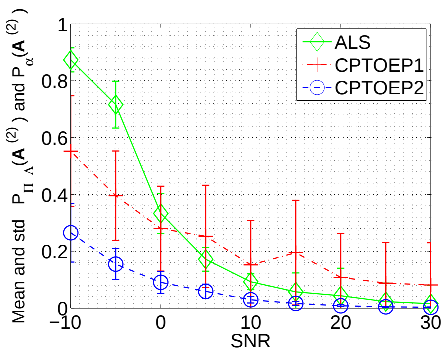
Figure 1: Mean and std $P_{\Pi\Lambda}\left(\mathbf{A}^{(n)}\right)$ and $P_{\Lambda}\left(\mathbf{A}^{(n)}\right)$ and time values over 100 trials while SNR is varying from -10 to 30 dB for the banded simulation cases. Case 1 on the left column and case 2 on the right column of the figure.



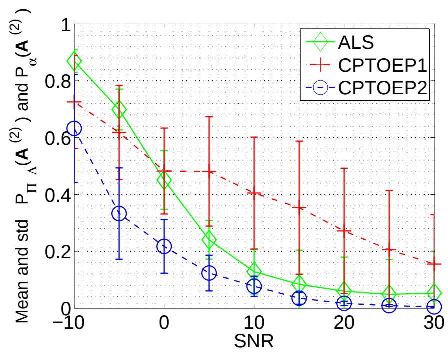
(a) Mean and std $P_{\Pi\Lambda}\left(\mathbf{A}^{(1)}\right)$ and $P_\alpha\left(\mathbf{A}^{(1)}\right)$.



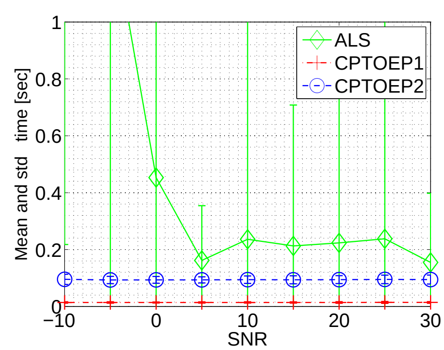
(b) Mean and std $P_{\Pi\Lambda}\left(\mathbf{A}^{(2)}\right)$ and $P_\alpha\left(\mathbf{A}^{(2)}\right)$.



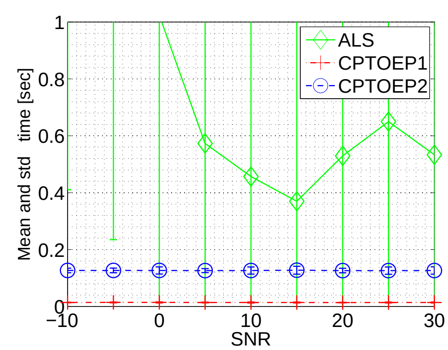
(c) Mean and std $P_{\Pi\Lambda}\left(\mathbf{A}^{(2)}\right)$ and $P_\alpha\left(\mathbf{A}^{(2)}\right)$.



(d) Mean and std $P_{\Pi\Lambda}\left(\mathbf{A}^{(2)}\right)$ and $P_\alpha\left(\mathbf{A}^{(2)}\right)$.

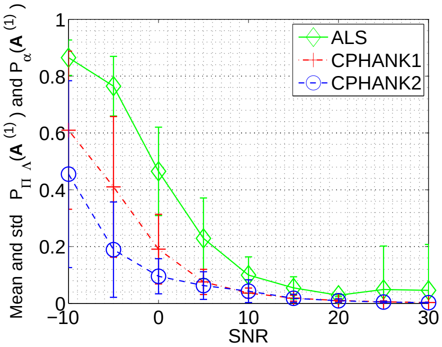


(e) Mean and std time.

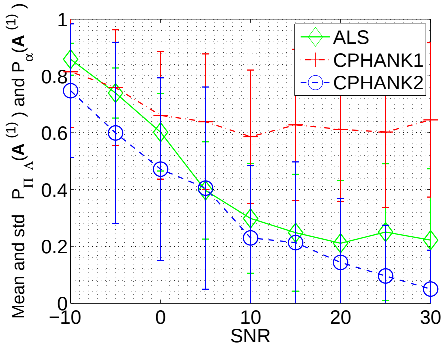


(f) Mean and std time.

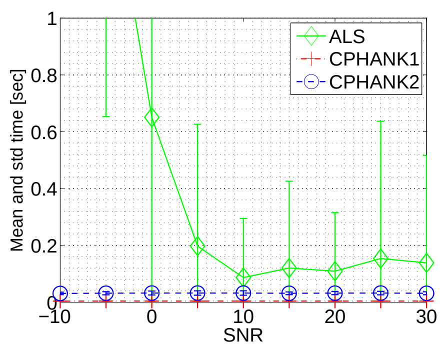
Figure 2: Mean and std $P_{\Pi\Lambda}\left(\mathbf{A}^{(n)}\right)$ and $P_\alpha\left(\mathbf{A}^{(n)}\right)$ and time values over 100 trials while SNR is varying from -10 to 30 dB for the banded Toepliz simulation cases. Case 1 on the left column and case 2 on the right column of the figure.



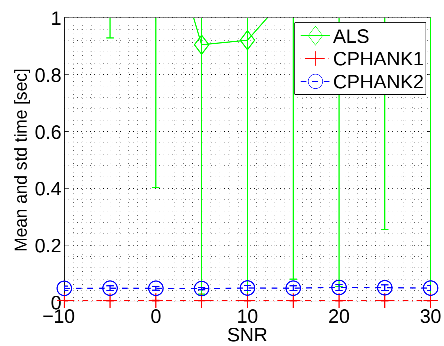
(a) Mean and std $P_{\Pi\Lambda}\left(\mathbf{A}^{(1)}\right)$ and $P_{\alpha}\left(\mathbf{A}^{(1)}\right)$.



(b) Mean and std $P_{\Pi\Lambda}\left(\mathbf{A}^{(2)}\right)$ and $P_{\alpha}\left(\mathbf{A}^{(2)}\right)$.



(c) Mean and std time.



(d) Mean and std time.

Figure 3: Mean and std $P_{\Pi\Lambda}\left(\mathbf{A}^{(n)}\right)$ and $P_{\alpha}\left(\mathbf{A}^{(n)}\right)$ and time values over 100 trials while SNR is varying from −10 to 30 dB for the banded Hankel simulation cases. Case 1 on the left column and case 2 on the right column of the figure.