DECOMPOSITION OF A TENSOR INTO MULTILINEAR RANK-$(M_r, N_r, \cdot)$ TERMS *

IGNAT DOMANOV†, NICO VERVLIET†, AND LIEVEN DE LATHAUWER†

Key words. multilinear algebra, third-order tensor, block term decomposition, multilinear rank

AMS subject classifications. 15A23, 15A69

1. Introduction.

1.1. Terminology and problem statement. Throughout the paper $\mathbb{F}$ denotes the field of real or complex numbers; bold-face lower case, bold-face capital case, and capital calligraphic letters refer to vectors, matrices, and tensors, respectively (e.g., $\mathbf{a}_1$, $\mathbf{B}_2$, and $\mathcal{T}$); $\mathbf{O}$ and $\mathbf{0}$ denote a zero matrix and a zero vector, respectively; $\mathbf{I}_n$ and $\mathbf{O}_{m\times n}$ denote the $n\times n$ identity matrix and the $m\times n$ zero matrix, respectively; $\mathbf{O}_n$ a is shorthand for $\mathbf{O}_{n\times n}$; $\mathrm{Null}\,(\cdot)$ denotes the null space of a matrix; "$*$" , "^{T}", and "^{H}" denote the conjugation, transpose, and hermitian transpose, respectively.

Let tensor $\mathcal{T}\in\mathbb{F}^{I\times J\times K}$ and let $\mathbf{T}_1,\ldots.\mathbf{T}_K\in\mathbb{F}^{I\times J}$ denote the frontal slices of $\mathcal{T}$. The matrices

$$\begin{aligned}\mathbf{T}_{(1)}&=[\mathbf{T}_1\ \ldots\ \mathbf{T}_K]\in\mathbb{F}^{I\times JK},\\ \mathbf{T}_{(2)}&=[\mathbf{T}_1^T\ \ldots\ \mathbf{T}_K^T]\in\mathbb{F}^{J\times IK},\\ \mathbf{T}_{(3)}&=[\mathrm{vec}(\mathbf{T}_1)\ \ldots\mathrm{vec}(\mathbf{T}_K)]^T\in\mathbb{F}^{K\times IJ},\end{aligned}$$

are called the *matrix unfoldings* of $\mathcal{T}$ and the triplet of their ranks $(r_{\mathbf{T}_{(1)}}, r_{\mathbf{T}_{(2)}}, r_{\mathbf{T}_{(3)}})$ is called the *multiLinear (ML) rank* of $\mathcal{T}$.

Tensor $\mathcal{T}$ is ML rank-$(M,N,\cdot)$ if $r_{\mathbf{T}_{(1)}}=M$, $r_{\mathbf{T}_{(2)}}=N$ and $r_{\mathbf{T}_{(3)}}$ is not specified. It can easily be shown that $\mathcal{T}$ is ML rank-$(M,N,\cdot)$ if and only if its frontal slices can be simultaneously factorized as

$$\mathbf{T}_k=\mathbf{A}\mathbf{D}_k\mathbf{B}^T,\qquad \mathbf{D}_k\in\mathbb{F}^{M\times N},\qquad k=1,\ldots,K, \tag{1}$$

where the factors $\mathbf{A}\in\mathbb{F}^{I\times M}$ and $\mathbf{B}\in\mathbb{F}^{J\times N}$ have full column rank and the matrices

$$\mathbf{D}_{(1)}:=[\mathbf{D}_1\ \ldots\ \mathbf{D}_K]\in\mathbb{F}^{M\times NK}\ \text{ and }\ \mathbf{D}_{(2)}:=[\mathbf{D}_1^T\ \ldots\ \mathbf{D}_K^T]\in\mathbb{F}^{N\times MK}$$

have full row rank. (One can, for instance, take $\mathbf{A}$ and $\mathbf{B}$ equal to the "U" factors in the compact SVDs of $\mathbf{T}_{(1)}$ and $\mathbf{T}_{(2)}$, respectively, and set $\mathbf{D}_k=\mathbf{A}^H\mathbf{T}_k\mathbf{B}^*$.) It's worth to be noted that the factors $\mathbf{A}$, $\mathbf{D}_k$, and $\mathbf{B}$ in (1) are not unique. Moreover, a generic

*Submitted to the editors DATE.

Funding: This work was funded by (1) Research Council KU Leuven: C1 project c16/15/059-nD; (2) F.W.O.: project G.0830.14N, G.0881.14N; (3) the Belgian Federal Science Policy Office: IUAP P7 (DYSCO II, Dynamical systems, control and optimization, 2012-2017); (4) EU: The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Advanced Grant: BIOTENSORS (no. 339804). This paper reflects only the authors' views and the Union is not liable for any use that may be made of the contained information TODO: Nico

† Group Science, Engineering and Technology, KU Leuven - Kulak, E. Sabbelaan 53, 8500 Kortrijk, Belgium and Dept. of Electrical Engineering ESAT/STADIUS KU Leuven, Kasteelpark Arenberg 10, bus 2446, B-3001 Leuven-Heverlee, Belgium (ignat.domanov@kuleuven.be, lieven.delathauwer@kuleuven.be).

This manuscript is for review purposes only.

ML rank-$(M, N, \cdot)$ tensor can parameterized with $(I - M)M + (J - N)N + MNK$ parameters. Indeed, if $\mathbf{X} \in \mathbb{F}^{I \times I}$ and $\mathbf{Y} \in \mathbb{F}^{J \times J}$ are nonsingular matrices, then $\mathbf{T}_1, \ldots, \mathbf{T}_K$ can also be factorized as

$$\mathbf{T}_k = (\mathbf{A}\mathbf{X}) \left(\mathbf{X}^{-1}\mathbf{D}_k\mathbf{Y}^{-T}\right) (\mathbf{B}\mathbf{Y})^T = \widehat{\mathbf{A}}\widehat{\mathbf{D}}_k\widehat{\mathbf{B}}^T, \qquad k = 1, \ldots, K. \tag{2}$$

Assuming that $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{D}_k$ in (2) are generic we can set $\mathbf{X}$ (resp. $\mathbf{Y}$) equal to the inverse of the matrix formed by the first M (resp. N) rows of $\mathbf{A}$ (resp. $\mathbf{B}$), implying that the right-hand side of the K equations

$$\mathbf{T}_k = \widehat{\mathbf{A}}\widehat{\mathbf{D}}_k\widehat{\mathbf{B}}^T = \begin{bmatrix} \mathbf{I}_M \\ * \end{bmatrix} [*] \begin{bmatrix} \mathbf{I}_N \\ * \end{bmatrix}^T, \qquad k = 1, \ldots, K.$$

can be parameterized with $(I - M)M + (J - N)N + MN \cdot K$ parameters.

We will write the K factorizations in (1) as

$$\mathcal{T} = \mathcal{D} \bullet_1 \mathbf{A} \bullet_2 \mathbf{B},$$

where $\mathcal{D}$ denotes an $M \times N \times K$ tensor with frontal slices $\mathbf{D}_1, \ldots, \mathbf{D}_K$.

In this paper we study a decomposition of $\mathcal{T}$ into a sum of ML rank-$(M_r, N_r, \cdot)$ tensors (see Figure 1)

$$\mathcal{T} = \sum_{r=1}^{R} \mathcal{D}_r \bullet_1 \mathbf{A}_r \bullet_2 \mathbf{B}_r, \qquad \mathcal{D}_r \in \mathbb{F}^{M_r \times N_r \times K}, \quad \mathbf{A}_r \in \mathbb{F}^{I \times M_r}, \quad \mathbf{B}_r \in \mathbb{F}^{J \times N_r}. \tag{3}$$

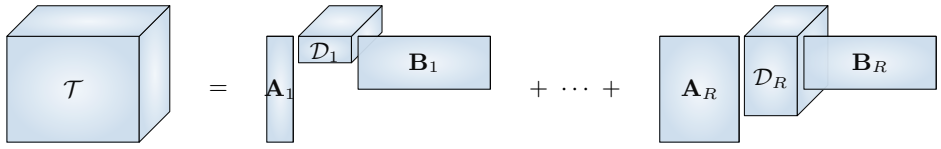


FIG. 1. *ML rank-$(M_r, N_r, \cdot)$ decomposition of a third-order tensor*

DEFINITION 1.1. *ML rank-$(M_r, N_r, \cdot)$ decomposition of tensor $\mathcal{T}$ is called* unique *if for any two decompositions of the form (3) one can be obtained from another by a permutation of summands.*

DEFINITION 1.2. *ML rank-$(M_r, N_r, \cdot)$ decomposition is called* generically unique *if*

$$\mu\{(\mathbf{A}_1, \ldots, \mathbf{A}_R, \mathbf{B}_1, \ldots, \mathbf{B}_R, \mathcal{D}_1, \ldots, \mathcal{D}_R) : \text{the decomposition of } \mathcal{T} = \sum_{r=1}^{R} \mathcal{D}_r \bullet_1 \mathbf{A}_r \bullet_2 \mathbf{B}_r \text{ is not unique}\} = 0,$$

where μ denotes a measure on $\mathbb{F}^{I\sum M_r + J\sum N_r + K\sum M_r N_r}$ absolutely continuous with respect to the Lebesgue measure.

We say that $\mathcal{T}$ is a sum of R *generic* ML rank-$(M_r, N_r, \cdot)$ terms if the entries of $\mathbf{A}_r$, $\mathbf{B}_r$, and $\mathcal{D}_r$ in (3) are randomly sampled from an absolutely continuous distribution. Thus, the generic uniqueness means uniqueness that holds with probability one.

Let

$$\mathbf{A} = [\mathbf{A}_1 \ \dots \ \mathbf{A}_R] \in \mathbb{F}^{I \times \sum M_r} \quad \text{and} \quad \mathbf{B} = [\mathbf{B}_1 \ \dots \ \mathbf{B}_R] \in \mathbb{F}^{J \times \sum N_r}. \tag{4}$$

Throughout the paper we assume that the matrices

$$\mathbf{A} \quad \text{and} \quad \mathbf{B} \quad \text{have full column rank.} \tag{5}$$

1.2. Previous results. Let $\mathbf{D}_{r,k} \in \mathbb{F}^{M_r \times N_r}$ denote the k-th frontal slice of the tensor $\mathcal{D}_r$ and let

$$\mathbf{D}_k = \operatorname{blockdiag}(\mathbf{D}_{1,k}, \dots, \mathbf{D}_{R,k}) \in \mathbb{F}^{\sum M_r \times \sum N_r}, \qquad k = 1, \dots, K \tag{6}$$

denote a block-diagonal matrix with the matrices $\mathbf{D}_{1,k}, \dots, \mathbf{D}_{R,k}$ on the diagonal. By (1), tensor decomposition (3) can be interpreted as the problem of (possibly nonsymmetric and nonsquare) *Joint Block Diagonalization* of the frontal slices of $\mathcal{T}$:

$$\mathbf{T}_k = \mathbf{A}\mathbf{D}_k\mathbf{B}^T = \mathbf{A} \operatorname{blockdiag}(\mathbf{D}_{1,k}, \dots, \mathbf{D}_{R,k})\mathbf{B}^T, \qquad k = 1, \dots, K, \tag{7}$$

where $\mathbf{A}$ and $\mathbf{B}$ are defined in (4).

Problem (7) (both uniqueness and algorithms) has been relatively well studied in the following cases.

1. JBD by congruence and $*$-congruence, i.e., the cases $\mathbf{B} = \mathbf{A}$ and $\mathbf{B} = \mathbf{A}^*$, implying the constraint $M_r = N_r$ for all r (see, for instance, [1, 2, 8] and the references therein).
2. All blocks are row vectors, i.e., $M_1 = \dots = M_R = 1$, implying that ML rank-$(M_r, N_r, \cdot)$ terms in (3) are actually ML rank-$(1, N_r, N_r)$ (see, for instance, [3, 4, 10, 9]);
3. All blocks are column vectors, i.e., $N_1 = \dots = N_R = 1$ (follows from 2. by taking transpose of the identities in (7)).

Apart from the items 1.–3. above the generic uniqueness of (7) was proved in [3] under the assumptions $M_r = N_r$ for all r and $K \geq 3$.

1.3. Our contribution. We obtain results on uniqueness and present algorithms for the computation of ML rank-$(M_r, N_r, \cdot)$ tensors in (3) in the case where $M_r \neq N_r$ for some r.

1.3.1. Exact decomposition.

THEOREM 1.3. *Let $\mathcal{T}$ admit the ML rank-$(M_r, N_r, \cdot)$ decomposition (3), the matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{D}_1, \dots, \mathbf{D}_K$ be defined in (4) and (6), respectively, and*

$$\mathbf{M}_{\mathcal{D}} = \begin{bmatrix} \mathbf{D}_1^T \otimes \mathbf{I}_I & -\mathbf{I}_J \otimes \mathbf{D}_1 \\ \vdots & \vdots \\ \mathbf{D}_K^T \otimes \mathbf{I}_I & -\mathbf{I}_J \otimes \mathbf{D}_K \end{bmatrix}. \tag{8}$$

Assume that

1. $\mathbf{A}$ *and* $\mathbf{B}$ *have full column rank and*
2. $\dim \operatorname{Null}(\mathbf{M}_{\mathcal{D}}) = R$.

Then the ML rank-$(M_r, N_r, \cdot)$ decomposition of $\mathcal{T}$ is unique and can be computed by means of Coupled Simultaneous Eigenvalue Decomposition (CS-EVD).

Proof. See section 2. ∎

Assumption 2) in Theorem 1.3 expresses the fact that the system of matrix equations

$$\mathbf{X}\mathbf{D}_k = \mathbf{D}_k\mathbf{Y}, \qquad k = 1, \dots, K \tag{9}$$

has exactly R linearly independent solutions $(\mathbf{X}, \mathbf{Y}) \in \mathbb{F}^{\sum M_r \times \sum M_r} \times \mathbb{F}^{\sum N_r \times \sum N_r}$. If

$$\begin{aligned} &\mathbf{P}_{I,r} = \text{blockdiag}(\mathbf{O}_{M_1}, \dots, \mathbf{O}_{M_{r-1}}, \mathbf{I}_{M_r}, \mathbf{O}_{M_{r+1}}, \dots, \mathbf{O}_{M_R}) \quad \text{and} \\ &\mathbf{P}_{J,r} = \text{blockdiag}(\mathbf{O}_{N_1}, \dots, \mathbf{O}_{N_{r-1}}, \mathbf{I}_{N_r}, \mathbf{O}_{N_{r+1}}, \dots, \mathbf{O}_{N_R}), \qquad r = 1, \dots, R, \end{aligned} \tag{10}$$

then $(\mathbf{X}, \mathbf{Y}) = (\mathbf{P}_{I,r}, \mathbf{P}_{J,r})$ is, obviously, a solution of (9),

$$\mathbf{P}_{I,r}\mathbf{D}_k = \mathbf{D}_k\mathbf{P}_{J,r} = \text{blockdiag}(\mathbf{O}_{M_1 \times N_1}, \dots, \mathbf{O}_{M_{r-1} \times N_{r-1}}, \mathbf{D}_{r,k}, \mathbf{O}_{M_{r+1} \times N_{r+1}}, \dots, \mathbf{O}_{M_R \times N_R}).$$

Thus, assumption 2) in Theorem 1.3 means that the dimension of the subspace of solutions of (9) is minimal.

The algebraic procedure related to Theorem 1.3 is summarized in Algorithm 1. By assumption 1) in Theorem 1.3, the number of columns of $\mathbf{A}$ and $\mathbf{B}$ cannot exceed the number of rows, i.e., $\sum M_r \le I$ and $\sum N_r \le J$. To simplify the presentation we assume in Algorithm 1 that $\sum M_r = I$ and $\sum N_r = J$, implying that $\mathbf{A}$ and $\mathbf{B}$ are square nonsingular matrices and $\mathcal{T} \in \mathbb{F}^{\sum M_r \times \sum N_r \times K}$. Since, by Lemma 2.1 below, $\mathcal{T}$ is ML rank-$(\sum M_r, \sum N_r, \cdot)$ tensor, such a compression can be done, for instance, using ML Singular Value Decomposition [5]. Note that Algorithm 1 also recovers the dimensions of the blocks, i.e., the values M_r and N_r. Indeed, in steps 1–4 of Algorithm 1 we construct matrices $\mathbf{U}_{I,1}, \dots, \mathbf{U}_{I,R} \in \mathbb{F}^{I \times I}$ and $\mathbf{U}_{J,1}, \dots, \mathbf{U}_{J,R} \in \mathbb{F}^{J \times J}$ that can be simultaneously diagonalized by $\mathbf{A}$ and $\mathbf{B}^{-T}$, respectively. In steps 5 of Algorithm 1 we use the fact that for each r the spectra of $\mathbf{U}_{I,r}$ and $\mathbf{U}_{J,r}$ relate to each other as in (12)–(13). Thus, if the values $\alpha_1, \dots, \alpha_R$ are generic, then the distinct eigenvalues of the matrices $\mathbf{F} = \alpha_1\mathbf{U}_{I,1} + \cdots + \alpha_R\mathbf{U}_{I,R}$ and $\mathbf{G} = \alpha_1\mathbf{U}_{J,1} + \cdots + \alpha_R\mathbf{U}_{J,R}$ are the same and their multiplicities are $M_1, \dots, M_R$ and $N_1, \dots, N_R$ respectively. Hence, the values M_r, N_r and the corresponding blocks $\mathbf{A}_r$, $\mathbf{B}_r$ can be recovered from the EVDs of $\mathbf{F}$ and $\mathbf{G}$.

The following corollary completes the results on uniqueness and computation of the ML rank-$(1, N_r, N_r)$ decomposition in [3, 4, 10].

COROLLARY 1.4. *Let $\mathcal{T}$ admit the ML rank-$(1, N_r, N_r)$ decomposition (3), the matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{D}_1, \dots, \mathbf{D}_K$ be defined in (4) and (6), respectively, and let*

$$\mathbf{C}_r := [\mathbf{D}_{r,1}^T \ \dots \ \mathbf{D}_{r,K}^T]^T \in \mathbb{F}^{K \times N_r}, \qquad r = 1, \dots, R. \tag{14}$$

Assume that

1. *$\mathbf{A}$ and $\mathbf{B}$ have full column rank and*
2. *$r_{[\mathbf{C}_i \ \mathbf{C}_j]} \ge \max(N_i, N_j) + 1$ for all $1 \le i < j \le R$.*

Then the ML rank-$(1, N_r, N_r)$ decomposition of $\mathcal{T}$ is unique and can be computed by means of CS-EVD.

Proof. See section 2. ∎

The following theorem contains bounds on the tensor dimensions that guarantee that assumptions 1) and 2) in Theorem 1.3 hold for generic matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{D}_1, \dots, \mathbf{D}_K$.

Algorithm 1 Computation of BTD (3) by CS-EVD (see Theorem 1.3)

Require: tensor $\mathcal{T} \in \mathbb{F}^{I \times J \times K}$ such that $I = \sum M_r$ and $J = \sum N_r$

1: Construct the IJK-by-$(I^2 + J^2)$ matrix $\mathbf{M}_\mathcal{T}$ as

$$\mathbf{M}_\mathcal{T} = \begin{bmatrix} \mathbf{T}_1^T \otimes \mathbf{I}_I & -\mathbf{I}_J \otimes \mathbf{T}_1 \\ \vdots & \vdots \\ \mathbf{T}_K^T \otimes \mathbf{I}_I & -\mathbf{I}_J \otimes \mathbf{T}_K \end{bmatrix} \tag{11}$$

2: Compute $\mathbf{u}_r = [u_{1,r} \ \dots \ u_{I^2+J^2,r}]^T$, $r = 1, \dots, R$ that form a basis of Null $(\mathbf{M}_\mathcal{T})$
3: For $r = 1, \dots, R$ reshape $[u_{1,r} \ \dots \ u_{I^2,r}]^T$ into the $I \times I$ matrix $\mathbf{U}_{I,r}$
4: For $r = 1, \dots, R$ reshape $[u_{I^2+1,r} \ \dots \ u_{I^2+J^2,r}]^T$ into the $J \times J$ matrix $\mathbf{U}_{J,r}$
5: Solve CS-EVD problem

$$\mathbf{U}_{I,r} = \mathbf{A} \,\text{blockdiag}(g_{1r}\mathbf{I}_{M_1}, \dots, g_{Rr}\mathbf{I}_{M_R})\mathbf{A}^{-1}, \tag{12}$$

$$\mathbf{U}_{J,r} = \mathbf{B}^{-T} \,\text{blockdiag}(g_{1r}\mathbf{I}_{N_1}, \dots, g_{Rr}\mathbf{I}_{N_R})\mathbf{B}^T, \quad r = 1, \dots, R \tag{13}$$

with respect to the matrices $\mathbf{A}$, $\mathbf{B}$ and the values $M_1, \dots, M_R$, $N_1, \dots, N_R$
6: Solve (7) with respect to $\mathbf{D}_1, \dots, \mathbf{D}_K$
7: For $r = 1, \dots, R$ stack $\mathbf{D}_{r,1}, \dots, \mathbf{D}_{r,K}$ into the $M_r \times N_r \times K$ tensor $\mathcal{D}_r$
8: **return** $\mathbf{A}_r \in \mathbb{F}^{I \times M_r}$, $\mathbf{B}_r \in \mathbb{F}^{J \times N_r}$ and $\mathcal{D}_r \in \mathbb{F}^{M_r \times N_r \times K}$ such that (3) holds

THEOREM 1.5. *Let $\mathcal{T}$ be a sum of R generic ML rank-$(M_r, N_r, \cdot)$ terms. Assume that*

$$I \geq \sum M_r, \qquad J \geq \sum N_r, \qquad K \geq \max\left(\max\left\lceil \frac{M_r}{N_r} \right\rceil + \max\left\lceil \frac{N_r}{M_r} \right\rceil, 3\right), \tag{15}$$

where $\lceil x \rceil$ denotes the smallest integer not less than x. Then the decomposition of $\mathcal{T}$ into a sum of ML rank-$(M_r, N_r, \cdot)$ terms is unique and can be computed by means of CS-EVD.

Proof. See section 3. ∎

The following corollaries easily follow from Theorem 1.5.

COROLLARY 1.6. *Let $\mathcal{T}$ be a sum of $R_1 \geq 1$ generic ML rank-$(1, N_r, N_r)$ terms and $R_2 \geq 1$ generic ML rank-$(M_r, 1, M_r)$ terms. Assume that*

$$I \geq R_1 + \sum M_r, \qquad J \geq R_2 + \sum N_r, \qquad \text{and} \qquad K \geq \max M_r + \max N_r.$$

Then the decomposition of $\mathcal{T}$ into a sum of R_1 ML rank-$(1, N_r, N_r)$ terms and R_2 ML rank-$(M_r, 1, M_r)$ terms is unique and can be computed by means of CS-EVD.

COROLLARY 1.7. *Let $\mathcal{T}$ be a sum of R generic ML rank-$(1, N_r, N_r)$ terms. Assume that*

$$I \geq R, \qquad J \geq \sum N_r, \qquad \text{and} \qquad K \geq \max N_r + 1. \tag{16}$$

Then the decomposition of $\mathcal{T}$ into a sum of ML rank-$(1, N_r, N_r)$ terms is unique and can be computed by means of CS-EVD.

COROLLARY 1.8. *Let $\mathcal{T}$ be a sum of R generic ML rank-$(M, N, \cdot)$ terms. Assume that*

$$I \geq RM, \qquad J \geq RN, \qquad K \geq \max(\left\lceil \frac{M}{N} \right\rceil + 1, \left\lceil \frac{N}{M} \right\rceil + 1, 3).$$

Then the decomposition of $\mathcal{T}$ into a sum of ML rank-$(M, N, \cdot)$ terms is unique and can be computed by means of CS-EVD.

The bounds in Corollary 1.6, Corollary 1.7, and Corollary 1.8 are new and guarantee that a generic ML rank $(M_r, N_r, \cdot)$ decomposition can be computed by Algorithm 1. The bound on K in (16) complements the bounds for generic uniqueness of the ML rank-$(1, N, N)$ decomposition in [3, Section 4] and is more relaxed than the bound for generic uniqueness of ML rank-$(1, N_r, N_r)$ decomposition in [9, Theorem 1.7 (E)].

In the remaining part of this section we apply Theorem 1.3 to (generic) constrained JBD problems in Table 1. Uniqueness of the constrained JBD problem

TABLE 1
Some examples of constrained JBD problems

Type of JBD problem	Constraints in (7)	$\mathbb{F}$
JBD by congruence	$\mathbf{B} = \mathbf{A}$	$\mathbb{R}$ or $\mathbb{C}$
symmetric JBD	$\mathbf{B} = \mathbf{A}$ and $\mathbf{D}_k = \mathbf{D}_k^T$, $k = 1, \ldots, K$	$\mathbb{R}$ or $\mathbb{C}$
JBD by $*$-congruence	$\mathbf{B} = \mathbf{A}^*$	$\mathbb{C}$
Hermitian JBD	$\mathbf{B} = \mathbf{A}^*$ and $\mathbf{D}_k = \mathbf{D}_k^H$, $k = 1, \ldots, K$	$\mathbb{C}$

means uniqueness of the corresponding constrained ML rank-$(N_r, N_r, 1)$ decomposition. Generic uniqueness of the constrained JBD problem can be defined in a similar way as in Definition 1.2. By way of example, we resort to the following definition of generic uniqueness of the symmetric JBD problem.

DEFINITION 1.9. *Let L_r denote a linear bijection between $\mathbb{F}^{K\frac{N_r(N_r+1)}{2}}$ and the subspace of $N_r \times N_r \times K$ tensors whose frontal slices are symmetric matrices. Let also μ be a measure on $\mathbb{F}^{I\sum N_r + K\sum \frac{N_r(N_r+1)}{2}}$ absolutely continuous with respect to the Lebesgue measure. A solution of the generic symmetric JBD problem is called unique if*

$$\mu\{(\mathbf{A}_1, \ldots, \mathbf{A}_R, \mathbf{d}_1, \ldots, \mathbf{d}_R) : \text{the constraint decomposition of } \mathcal{T} = \sum_{r=1}^{R} L_r(\mathbf{d}_r) \bullet_1 \mathbf{A}_r \bullet_2 \mathbf{A}_r \text{ is not unique}\} = 0.$$

Theorem 1.5 guarantees that if $K \geq 3$ and the diagonal blocks $\mathbf{D}_{r,k}$ in JBD problem (7) are square and generic, then the decomposition is unique and can be computed by means of CS-EVD. The following theorem states that the bound $K \geq 3$ remains valid for all constrained JBD problems in Table 1 as well and that the computation relies on S-EVD.

THEOREM 1.10. *Assume that $I \geq \sum N_r$, and $K \geq 3$. Then the solution of a generic constrained JBD problems in Table 1 is unique and can be computed by means of S-EVD.*

Steps 1–4 and modified step 5 of Algorithm 1 constitute the algebraic procedure related to Theorem 1.10. In step 5, because of the constraints $\mathbf{B} = \mathbf{A}$ and $\mathbf{B} = \mathbf{A}^*$, the CS-EVD problem (12)–(13) turns into the S-EVD problem. It can be shown that

in the case of the JBD by congruence problem the proposed reduction to S-EVD essentially coincides with the one in [1, Subsection 2.3].

Proof. See section 3. ∎

2. Proof of Theorem 1.3 and Corollary 1.4. We need the following lemmas.

LEMMA 2.1. *Let the conditions of Theorem 1.3 hold and let* $\mathbf{U}_1$ *and* $\mathbf{U}_2$ *denote the "U" factors in the compact SVDs of* $\mathbf{T}_{(1)}$ *and* $\mathbf{T}_{(2)}$*, respectively. Then*

1. *tensor* $\mathcal{T} \in \mathbb{F}^{I\times J\times K}$ *is ML rank-*$(\sum M_r, \sum N_r, \cdot)$*;*
2. *the "compressed" matrices* $\mathbf{A}^c := \mathbf{U}_1^H\mathbf{A}$ *and* $\mathbf{B}^c := \mathbf{U}_2^H\mathbf{B}$ *are square and nonsingular;*
3. $\mathbf{A} = \mathbf{U}_1\mathbf{A}^c$ *and* $\mathbf{B} = \mathbf{U}_2\mathbf{B}^c$*;*
4. *the "compressed" tensor* $\mathcal{T}^c := \mathcal{T} \bullet_1 \mathbf{U}_1^H \bullet_2 \mathbf{U}_2^H \in \mathbb{F}^{\sum M_r\times\sum N_r\times K}$ *is ML rank-*$(\sum M_r, \sum N_r, \cdot)$*;*
5. *tensor* $\mathcal{T}^c$ *admits the ML rank-*$(M_r, N_r, \cdot)$ *decomposition*

$$\mathcal{T}^c = \sum_{r=1}^{R} \mathcal{D}_r \bullet_1 (\mathbf{U}_1^H\mathbf{A}_r) \bullet_2 (\mathbf{U}_2^H\mathbf{B}_r),$$

where the tensors $\mathcal{D}_r \in \mathbb{F}^{M_r\times N_r\times K}$ *are the same as in decomposition (3).*

Proof. 1. First we show that the matrices

$$\mathbf{D}_{(1)} = [\mathbf{D}_1 \ \dots \ \mathbf{D}_K] \quad \text{and} \quad \mathbf{D}_{(2)} = [\mathbf{D}_1^T \ \dots \ \mathbf{D}_K^T] \ \text{ have full row rank.} \tag{17}$$

Assume that $\mathbf{f}^T\mathbf{D}_{(1)} = \mathbf{0}$. Then $\mathbf{1}\mathbf{f}^T\mathbf{D}_k = \mathbf{D}_k\mathbf{O}_{\sum N_r}$, where $\mathbf{1}$ denotes the $\sum M_r \times 1$ vector of ones. Hence $\begin{bmatrix}\mathbf{f}\otimes\mathbf{1}\\ \mathbf{0}\end{bmatrix} \in \operatorname{Null}(\mathbf{M}_{\mathcal{D}})$. Since $\dim\operatorname{Null}(\mathbf{M}_{\mathcal{D}}) = R$, it follows that the rank-1 matrix $\mathbf{1}\mathbf{f}^T$ is a linear combination of the matrices $\mathbf{P}_{I,1}, \dots, \mathbf{P}_{I,R}$. Hence $\mathbf{f} = \mathbf{0}$. Thus $\mathbf{D}_{(1)}$ has full row rank. In a similar way one can prove that $\mathbf{D}_{(2)}$ has also full row rank. Now, from assumption 1 in Theorem 1.3, (17), and the identities

$$\mathbf{T}_{(1)} = \mathbf{A}\mathbf{D}_{(1)}\operatorname{blockdiag}(\mathbf{B}^T, \dots, \mathbf{B}^T), \quad \mathbf{T}_{(2)} = \mathbf{B}\mathbf{D}_{(2)}\operatorname{blockdiag}(\mathbf{A}^T, \dots, \mathbf{A}^T) \tag{18}$$

it follows that

$$\operatorname{col}(\mathbf{T}_{(1)}) = \operatorname{col}(\mathbf{A}) \quad \text{and} \quad \operatorname{col}(\mathbf{T}_{(2)}) = \operatorname{col}(\mathbf{B}), \tag{19}$$

where $\operatorname{col}(\cdot)$ denotes the column space of a matrix. Hence $r_{\mathbf{T}_{(1)}} = r_{\mathbf{A}} = \sum M_r$ and $r_{\mathbf{T}_{(2)}} = r_{\mathbf{B}} = \sum N_r$, implying that the tensor $\mathcal{T} \in \mathbb{F}^{I\times J\times K}$ is ML rank-$(\sum M_r, \sum N_r, \cdot)$.

2. and 3. follow from (19) and the fact that $\mathbf{U}_1$ and $\mathbf{U}_2$ are the "U" factors in the compact SVDs of $\mathbf{T}_{(1)}$ and $\mathbf{T}_{(2)}$, respectively.

4. From the construction of $\mathbf{T}^c_{(1)}$ and (18) it follows that

$$\begin{aligned}\mathbf{T}^c_{(1)} = [\mathbf{U}_1^H\mathbf{T}_1\mathbf{U}_2^* \ \dots \ \mathbf{U}_1^H\mathbf{T}_K\mathbf{U}_2^*] = \mathbf{U}_1^H\mathbf{T}_{(1)}\operatorname{blockdiag}(\mathbf{U}_2^*, \dots, \mathbf{U}_2^*) = \\ \mathbf{U}_1^H\mathbf{A}\mathbf{D}_{(1)}\operatorname{blockdiag}((\mathbf{U}_2^H\mathbf{B})^T, \dots, (\mathbf{U}_2^H\mathbf{B})^T) = \\ \mathbf{A}^c\mathbf{D}_{(1)}\operatorname{blockdiag}((\mathbf{B}^c)^T, \dots, (\mathbf{B}^c)^T).\end{aligned} \tag{20}$$

Hence, by statement 2. and (17), $r_{\mathbf{T}^c_{(1)}} = r_{\mathbf{A}^c} = \sum M_r$. In a similar way one can prove that $r_{\mathbf{T}^c_{(2)}} = r_{\mathbf{B}^c} = \sum N_r$. Thus, $\mathcal{T}^c$ is ML rank-$(\sum M_r, \sum N_r, \cdot)$.

5. Follows from (20) and 2. ∎

LEMMA 2.2. *Let $I = \sum M_r$, $J = \sum N_r$, $\mathbf{A}$ and $\mathbf{B}$ be nonsingular, and let the matrices $\mathbf{M}_{\mathcal{D}}$ and $\mathbf{M}_{\mathcal{T}}$ be defined in (8) and (11), respectively. Then*

1) the matrices $\mathbf{M}_{\mathcal{D}}$ and $\mathbf{M}_{\mathcal{T}}$ are related by

$$\mathbf{M}_{\mathcal{T}} = \operatorname{blockdiag}(\mathbf{B} \otimes \mathbf{A}, \ldots, \mathbf{B} \otimes \mathbf{A}) \cdot \mathbf{M}_{\mathcal{D}} \cdot \begin{bmatrix} \mathbf{A}^T \otimes \mathbf{A}^{-1} & \mathbf{O} \\ \mathbf{O} & \mathbf{B}^{-1} \otimes \mathbf{B}^T \end{bmatrix};$$

2) the matrices $\mathbf{X} \in \mathbb{F}^{I \times I}$ and $\mathbf{Y} \in \mathbb{F}^{J \times J}$ satisfy the equations $\mathbf{X}\mathbf{D}_k = \mathbf{D}_k\mathbf{Y}$, $k = 1, \ldots, K$ if and only if $\mathbf{M}_{\mathcal{D}} \begin{bmatrix} \operatorname{vec}(\mathbf{X}) \\ \operatorname{vec}(\mathbf{Y}) \end{bmatrix} = \mathbf{0}$;

3) the matrices $\mathbf{X} \in \mathbb{F}^{I \times I}$ and $\mathbf{Y} \in \mathbb{F}^{J \times J}$ satisfy the equations $\mathbf{X}\mathbf{T}_k = \mathbf{T}_k\mathbf{Y}$, $k = 1, \ldots, K$ if and only if $\mathbf{M}_{\mathcal{T}} \begin{bmatrix} \operatorname{vec}(\mathbf{X}) \\ \operatorname{vec}(\mathbf{Y}) \end{bmatrix} = \mathbf{0}$;

4) $\operatorname{Null}(\mathbf{M}_{\mathcal{D}}) \supseteq \operatorname{span}\left\{ \begin{bmatrix} \operatorname{vec}(\mathbf{P}_{I,r}) \\ \operatorname{vec}(\mathbf{P}_{J,r}) \end{bmatrix} : \ r = 1, \ldots, R \right\}$, where the projectors $\mathbf{P}_{I,r}$ and $\mathbf{P}_{J,r}$ are defined in (10);

5) $\operatorname{Null}(\mathbf{M}_{\mathcal{T}}) = \begin{bmatrix} \mathbf{A}^{-T} \otimes \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{B} \otimes \mathbf{B}^{-T} \end{bmatrix} \operatorname{Null}(\mathbf{M}_{\mathcal{D}})$;

6) $\dim \operatorname{Null}(\mathbf{M}_{\mathcal{T}}) = \dim \operatorname{Null}(\mathbf{M}_{\mathcal{D}}) \geq R$;

7) if $\dim \operatorname{Null}(\mathbf{M}_{\mathcal{D}}) = R$, then the matrices

$$[\mathbf{A}\mathbf{D}_1\mathbf{B}^T \ \ldots \ \mathbf{A}\mathbf{D}_K\mathbf{B}^T] \in \mathbb{F}^{I \times JK} \quad \text{and} \quad [\mathbf{B}\mathbf{D}_1^T\mathbf{A}^T \ \ldots \ \mathbf{B}\mathbf{D}_K^T\mathbf{A}^T] \in \mathbb{F}^{J \times IK}$$

have full row rank.

Proof. 1) follows from (7), (8), and (11); 2)–4) are trivial; 5) follows from 1); 6) follows from 5) and 4); 7) follows from Lemma 2.1 1. and (18). ∎

Now we are ready to prove Theorem 1.3.

Proof of Theorem 1.3. By Lemma 2.1, WLOG, we can assume that $I = \sum M_r$ and $J = \sum N_r$.

1) First we show that decomposition (3) can be computed algebraically. From assumption 2) and Lemma 2.2 4), 6) we have that

$$\operatorname{Null}(\mathbf{M}_{\mathcal{D}}) = \operatorname{span}\left\{ \begin{bmatrix} \operatorname{vec}(\mathbf{P}_{I,r}) \\ \operatorname{vec}(\mathbf{P}_{J,r}) \end{bmatrix} : \ r = 1, \ldots, R \right\} \quad \text{and} \quad \dim \operatorname{Null}(\mathbf{M}_{\mathcal{T}}) = R.$$

Hence, by Lemma 2.2 5),

$$\operatorname{Null}(\mathbf{M}_{\mathcal{T}}) = \begin{bmatrix} \mathbf{A}^{-T} \otimes \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{B} \otimes \mathbf{B}^{-T} \end{bmatrix} \operatorname{span}\left\{ \begin{bmatrix} \operatorname{vec}(\mathbf{P}_{I,r}) \\ \operatorname{vec}(\mathbf{P}_{J,r}) \end{bmatrix} : \ r = 1, \ldots, R \right\}. \tag{21}$$

Let the columns of $\mathbf{U} \in \mathbb{F}^{(I^2+J^2) \times R}$ form a basis of $\operatorname{Null}(\mathbf{M}_{\mathcal{T}})$. We represent $\mathbf{U}$ as

$$\mathbf{U} = \begin{bmatrix} \operatorname{vec}(\mathbf{U}_{I,1}) & \ldots & \operatorname{vec}(\mathbf{U}_{I,R}) \\ \operatorname{vec}(\mathbf{U}_{J,1}) & \ldots & \operatorname{vec}(\mathbf{N}_{J,R}) \end{bmatrix}, \tag{22}$$

where $\mathbf{U}_{I,r} \in \mathbb{F}^{I \times I}$ and $\mathbf{U}_{J,r} \in \mathbb{F}^{J \times J}$, $r = 1, \ldots, R$. From (21) and (22), it follows that there exists a unique nonsingular $R \times R$ matrix $\mathbf{G}$ such that

$$\begin{aligned} \begin{bmatrix} \operatorname{vec}(\mathbf{U}_{I,1}) & \ldots & \operatorname{vec}(\mathbf{U}_{I,R}) \end{bmatrix} &= \left(\mathbf{A}^{-T} \otimes \mathbf{A}\right) \begin{bmatrix} \operatorname{vec}(\mathbf{P}_{I,1}) & \ldots & \operatorname{vec}(\mathbf{P}_{I,R}) \end{bmatrix} \mathbf{G}, \\ \begin{bmatrix} \operatorname{vec}(\mathbf{U}_{J,1}) & \ldots & \operatorname{vec}(\mathbf{U}_{J,R}) \end{bmatrix} &= \left(\mathbf{B} \otimes \mathbf{B}^{-T}\right) \begin{bmatrix} \operatorname{vec}(\mathbf{P}_{J,1}) & \ldots & \operatorname{vec}(\mathbf{P}_{J,R}) \end{bmatrix} \mathbf{G}. \end{aligned}$$

Hence, (12)–(13) hold. Since the matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{G}$ are nonsingular, it follows that the matrices $\mathbf{U}_{I,1}, \dots, \mathbf{U}_{I,R}$ are linearly independent. Thus, we can recover $\mathbf{A}$ from (12). Similarly, we can recover $\mathbf{B}$ from (13).

2) Now we show that decomposition (3) is unique. Assume that (3) holds for $\mathbf{A}_r$, $\mathbf{B}_r$, etc. replaced by $\widetilde{\mathbf{A}}_r$, $\widetilde{\mathbf{B}}_r$, etc. To prove the uniqueness it is sufficient to show that $\widetilde{\mathbf{A}}$, $\widetilde{\mathbf{B}}$ and $\mathbf{M}_{\widetilde{\mathcal{D}}}$ satisfy assumptions 1) and 2), i.e., that $\widetilde{\mathbf{A}}$ and $\widetilde{\mathbf{B}}$ are nonsingular and $\dim \operatorname{Null}\left(\mathbf{M}_{\widetilde{\mathcal{D}}}\right) = R$. By (7),

$$[\mathbf{A}\mathbf{D}_1\mathbf{B}^T \ \dots \ \mathbf{A}\mathbf{D}_K\mathbf{B}^T] = [\mathbf{T}_1 \ \dots \ \mathbf{T}_K] = [\widetilde{\mathbf{A}}\widetilde{\mathbf{D}}_1\widetilde{\mathbf{B}}^T \ \dots \ \widetilde{\mathbf{A}}\widetilde{\mathbf{D}}_K\widetilde{\mathbf{B}}^T] = \widetilde{\mathbf{A}}[\widetilde{\mathbf{D}}_1\widetilde{\mathbf{B}}^T \ \dots \ \widetilde{\mathbf{D}}_K\widetilde{\mathbf{B}}^T], \tag{23}$$

$$[\mathbf{B}\mathbf{D}_1^T\mathbf{A}^T \ \dots \ \mathbf{B}\mathbf{D}_K^T\mathbf{A}^T] = [\mathbf{T}_1^T \ \dots \ \mathbf{T}_K^T] = [\widetilde{\mathbf{B}}\widetilde{\mathbf{D}}_1^T\widetilde{\mathbf{A}}^T \ \dots \ \widetilde{\mathbf{B}}\widetilde{\mathbf{D}}_K^T\widetilde{\mathbf{A}}^T] = \widetilde{\mathbf{B}}[\widetilde{\mathbf{D}}_1^T\widetilde{\mathbf{A}}^T \ \dots \ \widetilde{\mathbf{D}}_K^T\widetilde{\mathbf{A}}^T]. \tag{24}$$

Since, by Lemma 2.2 7), the matrices in the left-hand side of (23) and (24) have full row rank, it follows that the matrices $\widetilde{\mathbf{A}}$ and $\widetilde{\mathbf{B}}$ are nonsingular. Since, by assumption 2), $\dim \operatorname{Null}(\mathbf{M}_{\mathcal{D}}) = R$, and, by Lemma 2.2 5),

$$\operatorname{Null}(\mathbf{M}_{\mathcal{T}}) = \begin{bmatrix} \mathbf{A}^{-T} \otimes \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{B} \otimes \mathbf{B}^{-T} \end{bmatrix} \operatorname{Null}(\mathbf{M}_{\mathcal{D}}) = \begin{bmatrix} \widetilde{\mathbf{A}}^{-T} \otimes \widetilde{\mathbf{A}} & \mathbf{O} \\ \mathbf{O} & \widetilde{\mathbf{B}} \otimes \widetilde{\mathbf{B}}^{-T} \end{bmatrix} \operatorname{Null}\left(\mathbf{M}_{\widetilde{\mathcal{D}}}\right),$$

it follows that $\dim \operatorname{Null}\left(\mathbf{M}_{\widetilde{\mathcal{D}}}\right) = R$. ∎

Proof of Corollary 1.4. It is sufficient to prove that assumption 2 in Corollary 1.4 implies assumption 2 in Theorem 1.3. Thus, we need to show that the subspace of solution of (9) is spanned by the trivial solutions $(\mathbf{X}, \mathbf{Y}) = (\mathbf{P}_{I,r}, \mathbf{P}_{J,r})$, $r = 1, \dots, R$, where $\mathbf{P}_{I,r}$ and $\mathbf{P}_{J,r}$ are defined in (10) and $M_1 = \cdots = M_R = 1$.

Let $\mathbf{X} = (x_{ij})_{i,j=1}^R$ and let $\mathbf{Y} = (\mathbf{Y}_{ij})_{i,j=1}^R$ consists of the blocks $\mathbf{Y}_{ij} \in \mathbb{F}^{N_i \times N_j}$. Since the matrices $\mathbf{D}_k$ are block-diagonal, we can rewrite (9) as

$$x_{ij}\mathbf{D}_{j,k} = \mathbf{D}_{i,k}\mathbf{Y}_{ij}, \qquad 1 \le i, j \le R, \qquad k = 1, \dots, K,$$

which, by (14), is equivalent to

$$x_{ij}\mathbf{C}_j = \mathbf{C}_i\mathbf{Y}_{ij}, \qquad 1 \le i, j \le R. \tag{25}$$

We need to show that

$$\mathbf{Y}_{ii} = x_{ii}\mathbf{I}_{N_i}, \qquad 1 \le i \le R, \tag{26}$$

$$\mathbf{Y}_{ij} = \mathbf{O}, \quad x_{ij} = 0, \qquad 1 \le i \ne j \le R. \tag{27}$$

(i) Let us prove (26). By (25),

$$\mathbf{C}_i(x_{ii}\mathbf{I}_{N_i} - \mathbf{Y}_{ii}) = \mathbf{O}, \qquad 1 \le i \le R. \tag{28}$$

Since the $1 \times N_i \times K$ tensor $\mathcal{D}_i$ in (3) is ML rank-$(1, N_i, N_i)$ and the second matrix unfolding of $\mathcal{D}_i$ coincides with $\mathbf{C}_i^T$, it follows that $\mathbf{C}_i^T$ has full row rank or that $\mathbf{C}_i$ has full column rank. Hence (26) follows from (28).

(ii) Let us prove (27). It is clear that (25) can be rewritten as

$$\mathbf{O} = [\mathbf{C}_i \;\; \mathbf{C}_j] \begin{bmatrix} \mathbf{Y}_{ij} \\ x_{ij}\mathbf{I}_{N_j} \end{bmatrix}. \tag{29}$$

Assume to the contrary that $x_{ij} \neq 0$. Then from (29), Sylvester's rank inequality, and assumption 2 it follows that

$$0 \geq r_{[\mathbf{C}_i \;\; \mathbf{C}_j]} + r_{[\mathbf{Y}_{ij}^T \;\; x_{ij}\mathbf{I}_{N_j}]^T} - (N_i + N_j) \geq \max(N_i, N_j) + 1 + N_j - (N_i + N_j) \geq 1,$$

which is a contradiction. Thus, $x_{ij} = 0$. Hence, by (25), $\mathbf{C}_i\mathbf{Y}_{ij} = \mathbf{O}$. Since $\mathbf{C}_i$ has full column rank, it follows that $\mathbf{Y}_{ij} = \mathbf{O}$. ∎

3. Proofs of Theorem 1.5 and Theorem 1.10 . We need the following two lemmas.

LEMMA 3.1. *Let*

$$K = \max(\left\lceil \frac{M}{N} \right\rceil + 1, \left\lceil \frac{N}{M} \right\rceil + 1, 3). \tag{30}$$

Then there exist matrices $\mathbf{E}_1, \ldots, \mathbf{E}_K \in \mathbb{F}^{M \times N}$ *such that the system of matrix equations*

$$\mathbf{X}\mathbf{E}_k = \mathbf{E}_k\mathbf{Y}, \qquad \mathbf{X} \in \mathbb{F}^{M \times M}, \;\; \mathbf{Y} \in \mathbb{F}^{N \times N} \qquad\qquad k = 1, \ldots, K \tag{31}$$

has only the trivial solutions $\mathbf{X} = c\mathbf{I}_M$ *and* $\mathbf{Y} = c\mathbf{I}_N$*, where* $c \in \mathbb{F}$*.*

Proof. It is clear that by taking the transpose of (31) we obtain an equivalent problem of the form (31) in which M and N are switched. Thus, w.l.o.g., we can always assume that $M \geq N$. We consider three cases: $M = N$, $N < M \leq 2N$, and $2N < M$.

Case $M = N$. By assumption (30), $K = 3$. Let $\mathbf{a} \in \mathbb{F}^N$ be a vector with distinct entries and let $\mathbf{D} \in \mathbb{F}^{N \times N}$ be a matrix with nonzero entries. We show that if $\mathbf{E}_1 = \mathbf{I}_N$, $\mathbf{E}_2 = \operatorname{diag}(\mathbf{a})$, and $\mathbf{E}_3 = \mathbf{D}$, then (31) has only trivial solutions.

The equation $\mathbf{X}\mathbf{E}_1 = \mathbf{E}_1\mathbf{Y}$ implies that $\mathbf{X} = \mathbf{Y}$. Hence,

$$\mathbf{Y}\operatorname{diag}(\mathbf{a}) = \mathbf{Y}\mathbf{E}_2 = \mathbf{X}\mathbf{E}_2 = \mathbf{E}_2\mathbf{Y} = \operatorname{diag}(\mathbf{a})\mathbf{Y}.$$

Since the entries of $\mathbf{a}$ are distinct, it follows that $\mathbf{Y}$ is a diagonal matrix, $\mathbf{Y} = \operatorname{diag}(y_1, \ldots, y_N)$. Finally, since the entries of $\mathbf{E}_3 = \mathbf{D}$ are nonzero, the equation

$$\operatorname{diag}(y_1, \ldots, y_N)\mathbf{D} = \mathbf{Y}\mathbf{E}_3 = \mathbf{X}\mathbf{E}_3 = \mathbf{E}_3\mathbf{Y} = \mathbf{D}\operatorname{diag}(y_1, \ldots, y_N)$$

implies that $y_1 = \cdots = y_N =: c$. Thus, $\mathbf{X} = \mathbf{Y} = c\mathbf{I}_N$.

Case $N < M \leq 2N$. By assumption (30), $K = 3$. Let $\mathbf{a} \in \mathbb{F}^N$ be a vector with distinct entries, $\mathbf{D} \in \mathbb{F}^{(M-N)\times(2N-M)}$ be a full row rank matrix with nonzero entries, and let $\widehat{\mathbf{I}}_{M-N}$ denote the matrix formed by the first $M - N$ rows of $\mathbf{I}_N$. We show that if

$$\mathbf{E}_1 = \begin{bmatrix} \mathbf{I}_N \\ \mathbf{O} \end{bmatrix}, \qquad \mathbf{E}_2 = \begin{bmatrix} \mathbf{O} \\ \mathbf{D} \end{bmatrix}, \qquad \mathbf{E}_3 = \begin{bmatrix} \operatorname{diag}(\mathbf{a}) \\ \widehat{\mathbf{I}}_{M-N} \end{bmatrix},$$

then (31) has only trivial solutions. Let $\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_3 \\ \mathbf{X}_2 & \mathbf{X}_4 \end{bmatrix}$ with $\mathbf{X}_1 \in \mathbb{F}^{N \times N}$, $\mathbf{X}_2, \mathbf{X}_3^T \in \mathbb{F}^{(M-N)\times N}$ and $\mathbf{X}_4 \in \mathbb{F}^{(M-N)\times(M-N)}$. Then the equations $\mathbf{X}\mathbf{E}_1 = \mathbf{E}_1\mathbf{Y}$, $\mathbf{X}\mathbf{E}_2 =$

$\mathbf{E}_2\mathbf{Y}$, and $\mathbf{X}\mathbf{E}_3 = \mathbf{E}_3\mathbf{Y}$ imply

$$\mathbf{X}_1 = \mathbf{Y} \quad \text{and} \quad \mathbf{X}_2 = 0, \tag{32}$$
$$\mathbf{X}_3\mathbf{D} = \mathbf{O} \quad \text{and} \quad \mathbf{X}_4\mathbf{D} = \mathbf{D}\mathbf{Y},$$

$$\mathbf{X}_1 \operatorname{diag}(\mathbf{a}) + \mathbf{X}_3\widehat{\mathbf{I}}_{M-N} = \operatorname{diag}(\mathbf{a})\mathbf{Y} \quad \text{and} \quad \mathbf{X}_2 \operatorname{diag}(\mathbf{a}) + \mathbf{X}_4\widehat{\mathbf{I}}_{M-N} = \widehat{\mathbf{I}}_{M-N}\mathbf{Y}, \tag{33}$$

respectively. Since $\mathbf{X}_3\mathbf{D} = \mathbf{O}$ and $\mathbf{D}$ has full row rank, it follows that $\mathbf{X}_3 = \mathbf{O}$. Hence, by (32) and (33),

$$\mathbf{Y} \operatorname{diag}(\mathbf{a}) = \mathbf{X}_1 \operatorname{diag}(\mathbf{a}) + \mathbf{X}_3\widehat{\mathbf{I}}_{M-N} = \operatorname{diag}(\mathbf{a})\mathbf{Y}.$$

Since all entries of $\mathbf{a}$ are distinct, it follows that $\mathbf{Y}$ is a diagonal matrix, $\mathbf{Y} = \operatorname{diag}(y_1, \dots, y_N)$. From the second identities in (32) and (33) it follows that $\mathbf{X}_4 = \operatorname{diag}(y_1, \dots, y_{M-N})$. Finally, since the entries of $\mathbf{D}$ are nonzero, it follows from $\mathbf{X}_4\mathbf{D} = \mathbf{D}\mathbf{Y}$ that $y_1 = \dots = y_N =: c$. Hence, $\mathbf{Y} = c\mathbf{I}_N$ and $\mathbf{X} = \operatorname{blockdiag}(\mathbf{X}_1, \mathbf{X}_4) = c\mathbf{I}_M$.

Case $2N < M$. By assumption (30), $K = \left\lceil \frac{M}{N} \right\rceil + 1$. Let $r = M - (K-2)N$. Then $0 < r \le N$. We define $\mathbf{E}_1, \dots, \mathbf{E}_{K-2} \in \mathbb{F}^{M\times N}$ and $\widehat{\mathbf{E}}_{K-1} \in \mathbb{F}^{M\times r}$ as blocks in the partition $\mathbf{I}_M = [\mathbf{E}_1 \ \dots \ \mathbf{E}_{K-2} \ \widehat{\mathbf{E}}_{K-1}]$ and set $\mathbf{E}_{K-1} = [\widehat{\mathbf{E}}_{K-1} \ \mathbf{O}]$, where the zero matrix has $N - r$ columns. We also set $\mathbf{E}_k = \begin{bmatrix} \operatorname{diag}(\mathbf{a}) \\ \mathbf{D} \\ \mathbf{O} \end{bmatrix}$, where $\mathbf{a} \in \mathbb{F}^N$ is a vector with distinct entries and $\mathbf{D} \in \mathbb{F}^{N\times N}$ is a matrix with nonzero entries. We show that for such choice of the matrices $\mathbf{E}_k$ system (31) has only the trivial solution. From the first $K-1$ equations in (31) it follows that $\mathbf{X} = \operatorname{blockdiag}(\mathbf{Y}, \dots, \mathbf{Y}, \widehat{\mathbf{Y}})$, where $\widehat{\mathbf{Y}}$ denotes a matrix located at the intersection of the first r rows and r columns of $\mathbf{Y}$. The remaining equation $\mathbf{X}\mathbf{E}_K = \mathbf{E}_K\mathbf{Y}$ in (31) implies that $\mathbf{Y} \operatorname{diag}(\mathbf{a}) = \operatorname{diag}(\mathbf{a})\mathbf{Y}$ and $\mathbf{Y}\mathbf{D} = \mathbf{D}\mathbf{Y}$. Now, in exactly the same way as "*Case* $M = N$" one can obtain that $\mathbf{Y} = c\mathbf{I}_N$ for some $c \in \mathbb{F}$. Hence, $\mathbf{X} = c\mathbf{I}_M$. ∎

LEMMA 3.2. *Let*

$$K = \begin{cases} 2, \text{ if } \ M_1 > N_1 \ \text{ and } \ M_2 < N_2, \\ \lceil M_2/N_2 \rceil + 1, \text{ if } \ M_1 > N_1 \ \text{ and } \ M_2 \ge N_2, \\ \lceil N_1/M_1 \rceil + 1, \text{ if } \ M_1 \le N_1 \ \text{ and } \ M_2 < N_2, \\ \lceil M_2/N_2 \rceil + \lceil N_1/M_1 \rceil, \text{ if } \ M_1 \le N_1 \ \text{ and } \ M_2 \ge N_2. \end{cases} \tag{34}$$

Then there exist matrices $\mathbf{E}_1, \dots, \mathbf{E}_K \in \mathbb{F}^{M_2\times N_2}$ *and* $\mathbf{F}_1, \dots, \mathbf{F}_K \in \mathbb{F}^{M_1\times N_1}$ *such that the system of matrix equations*

$$\mathbf{X}\mathbf{E}_k = \mathbf{F}_k\mathbf{Y}, \qquad \mathbf{X} \in \mathbb{F}^{M_1\times M_2}, \ \mathbf{Y} \in \mathbb{F}^{N_1\times N_2} \qquad k = 1, \dots, K \tag{35}$$

has only the zero solution.

Proof. We consider each of the four cases in (34) separately.

Case $M_1 > N_1$ *and* $M_2 < N_2$. By (34), $K = 2$. By definition put

$$\mathbf{E}_1 = \mathbf{O}, \qquad \mathbf{E}_2 = [\mathbf{I}_{M_2} \ \mathbf{O}], \qquad \mathbf{F}_1 = \begin{bmatrix} \mathbf{I}_{N_1} \\ \mathbf{O} \end{bmatrix}, \qquad \mathbf{E}_2 = \mathbf{O}.$$

Then the equations $\mathbf{X}\mathbf{E}_1 = \mathbf{F}_1\mathbf{Y}$ and $\mathbf{X}\mathbf{E}_2 = \mathbf{F}_2\mathbf{Y}$ imply that $\mathbf{Y} = \mathbf{O}$ and $\mathbf{X} = \mathbf{O}$, respectively.

Case $M_1 > N_1$ and $M_2 \geq N_2$. By (34), $K = \lceil M_2/N_2 \rceil + 1$. Let $\mathbf{E}_1, \ldots, \mathbf{E}_{K-1}$ be matrices such that $[\mathbf{E}_1 \ \ldots \mathbf{E}_{K-1}]$ has full row rank, let $\mathbf{E}_K$ be an arbitrary matrix, $\mathbf{F}_1, \ldots, \mathbf{F}_{K-1}$ be zero matrices, and $\mathbf{F}_K = \begin{bmatrix} \mathbf{I}_{N_1} \\ \mathbf{O} \end{bmatrix}$. Then from the first $K-1$ equations in (35) it follows that $\mathbf{X}[\mathbf{E}_1 \ \ldots \ \mathbf{E}_{K-1}] = 0$. Hence $\mathbf{X} = \mathbf{O}$. Now the remaining equation $\mathbf{X}\mathbf{E}_K = \mathbf{F}_K\mathbf{Y}$ implies that $\mathbf{Y} = \mathbf{O}$.

Case $M_1 \leq N_1$ and $M_2 < N_2$. By (34), $K = \lceil N_1/M_1 \rceil + 1$. The problem can be reduced to the setting used in the previous case by taking the transpose of the equations in (35).

Case $M_1 \leq N_1$ and $M_2 \geq N_2$. By (34), $K = \lceil M_2/N_2 \rceil + \lceil N_1/M_1 \rceil$. Let $\mathbf{E}_1, \ldots, \mathbf{E}_{\lceil M_2/N_2 \rceil}$ be matrices such that $[\mathbf{E}_1 \ \ldots \ \mathbf{E}_{\lceil M_2/N_2 \rceil}]$ has full row rank and let $\mathbf{E}_{\lceil M_2/N_2 \rceil + 1} = \mathbf{E}_K = \mathbf{O}$. Let also $\mathbf{F}_1 = \cdots = \mathbf{F}_{\lceil M_2/N_2 \rceil} = 0$ and $\mathbf{F}_{\lceil M_2/N_2 \rceil + 1}, \ldots, \mathbf{F}_K$ be matrices such that $[\mathbf{F}^T_{\lceil M_1/N_2 \rceil + 1} \ \ldots \ \mathbf{F}^T_K]^T$ has full column rank. Then from the first $\lceil M_2/N_2 \rceil$ and the last $\lceil N_1/M_1 \rceil$ equations in (35) it follows that

$$\mathbf{X}[\mathbf{E}_1 \ \ldots \ \mathbf{E}_{\lceil M_2/N_2 \rceil}] = \mathbf{O} \qquad \text{and} \qquad \mathbf{O} = [\mathbf{F}^T_{\lceil M_1/N_2 \rceil + 1} \ \ldots \ \mathbf{F}^T_K]^T\mathbf{Y},$$

respectively. Hence, $\mathbf{X} = \mathbf{O}$ and $\mathbf{Y} = \mathbf{O}$. ∎

Now we are ready to prove Theorem 1.5.

Proof of Theorem 1.5. We show that assumptions 1) and 2) in Theorem 1.3 hold for generic matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{D}_1, \ldots, \mathbf{D}_K$. The first two inequalities in (15) express the fact that the number of columns of $\mathbf{A}$ and $\mathbf{B}$ cannot exceed the number of rows. Since $\mathbf{A}$ and $\mathbf{B}$ are generic this means that assumption 1) in Theorem 1.3 holds. To prove that assumption 2) in Theorem 1.3 also holds it is sufficient to show that the subspace of solutions of (9) has dimension R. In the remaining part of the proof we show that for generic matrices $\mathbf{D}_1, \ldots, \mathbf{D}_K$ the subspace of solutions of (9) is spanned by $(\mathbf{P}_{I,1}, \mathbf{P}_{J,1}), \ldots, (\mathbf{P}_{I,R}, \mathbf{P}_{J,R})$, where the projectors $\mathbf{P}_{I,r}$, $\mathbf{P}_{J,r}$ are defined in (10).

Let $\mathbf{X} = (\mathbf{X}_{ij})_{i,j=1}^R$ and $\mathbf{Y} = (\mathbf{Y}_{ij})_{i,j=1}^R$ consist of the blocks $\mathbf{X}_{ij} \in \mathbb{F}^{M_i \times M_j}$ and $\mathbf{Y}_{ij} \in \mathbb{F}^{N_i \times N_j}$, respectively. Since the matrices $\mathbf{D}_k$ are block-diagonal, we can rewrite (9) in terms of the blocks $\mathbf{X}_{ij}$ and $\mathbf{Y}_{ij}$ as

$$\mathbf{X}_{ij}\mathbf{D}_{j,k} = \mathbf{D}_{i,k}\mathbf{Y}_{ij}, \qquad k = 1, \ldots, K, \tag{36}$$

where $i, j \in \{1, \ldots, R\}$. Let the subsets

$$\Omega_{ij}, \ \Omega_i \subseteq \left(\mathbb{F}^{M_1 \times N_1} \times \cdots \times \mathbb{F}^{M_R \times N_R}\right) \times \cdots \times \left(\mathbb{F}^{M_1 \times N_1} \times \cdots \times \mathbb{F}^{M_R \times N_R}\right)$$

consist of the elements $(\mathbf{D}_{1,1}, \ldots, \mathbf{D}_{R,1}, \ldots, \mathbf{D}_{1,K}, \ldots, \mathbf{D}_{R,K})$ and be defined as follows: the elements in Ω_{ij} $(i \neq j)$ are such that (36) has only the zero solution and the elements in Ω_i are such that (36) has only the trivial solution, i.e., $\mathbf{X}_{ii} = c_i\mathbf{I}_{M_i}$ and $\mathbf{Y}_{ii} = c_i\mathbf{I}_{N_i}$ for some $c_i \in \mathbb{F}$.

To prove that $(\mathbf{X}, \mathbf{Y})$ belongs to the linear subspace spanned by $(\mathbf{P}_{I,1}, \mathbf{P}_{J,1})$, $\ldots$, $(\mathbf{P}_{I,R}, \mathbf{P}_{J,R})$ we have to show that $(\cap\Omega_{ij}) \bigcap (\cap\Omega_i)$ is a set of full measure. Since the intersection of a finite number of sets of full measure is also a set of full measure it is sufficient to prove that all sets in the intersection have full measure. This can be done as follows.

1) Let $i \neq j$. W.l.o.g. we can assume that $i = 1$ and $j = 2$. Then, by (36),

$$\mathbf{X}_{12}\mathbf{D}_{2,k} = \mathbf{D}_{1,k}\mathbf{Y}_{12}, \qquad k = 1, \ldots, K. \tag{37}$$

It is clear that system (37) has only the zero solution if and only if the matrix

$$\mathbf{M} = \begin{bmatrix} \mathbf{D}_{2,1}^T \otimes \mathbf{I}_{M_1} & -\mathbf{I}_{N_2} \otimes \mathbf{D}_{1,1} \\ & \vdots \\ \mathbf{D}_{2,K}^T \otimes \mathbf{I}_{M_1} & -\mathbf{I}_{N_2} \otimes \mathbf{D}_{1,K} \end{bmatrix}$$

has full column rank. Hence, from Lemma 3.2 and assumption (15) it follows that there exist matrices $\mathbf{D}_{1,1}, \dots, \mathbf{D}_{1,K}$ and $\mathbf{D}_{2,1} \dots \mathbf{D}_{2,K}$ such that $\mathbf{M}$ has full column rank. Now we assume that $\mathbf{D}_{1,1}, \dots, \mathbf{D}_{1,K}$ and $\mathbf{D}_{2,1} \dots \mathbf{D}_{2,K}$ in (37) are generic. Then, by [6, Lemma 6.3], $\mathbf{M}$ has full column rank, implying that system (37) has only the zero solution. Hence, by Fubini's theorem [7, Theorem C, p.148], Ω_{12} is a set of full measure.

2) Let $i = j$. W.l.o.g. we can assume that $i = 1$. We have to prove that the system

$$\mathbf{X}_{11}\mathbf{D}_{1,k} = \mathbf{D}_{1,k}\mathbf{Y}_{11}, \qquad k = 1, \dots, K, \tag{38}$$

has only the trivial solution $\mathbf{X}_{11} = c\mathbf{I}_{M_1}$ and $\mathbf{Y}_{11} = c\mathbf{I}_{N_1}$, where $c \in \mathbb{F}$. It is clear that system (38) has only the trivial solution if and only if the dimension of the null space of the matrix

$$\mathbf{M} = \begin{bmatrix} \mathbf{D}_{1,1}^T \otimes \mathbf{I}_{M_1} & -\mathbf{I}_{N_1} \otimes \mathbf{D}_{1,1} \\ & \vdots \\ \mathbf{D}_{1,K}^T \otimes \mathbf{I}_{M_1} & -\mathbf{I}_{N_1} \otimes \mathbf{D}_{1,1} \end{bmatrix}$$

is one. Hence, from Lemma 3.1 and assumption (15) it follows that there exist matrices $\mathbf{D}_{1,1}, \dots, \mathbf{D}_{1,K}$ such that $M_1^2 + N_1^2 - 1$ columns of $\mathbf{M}$ are linearly independent. Now we assume that $\mathbf{D}_{1,1}, \dots, \mathbf{D}_{1,K}$ are generic. Then, by [6, Lemma 6.3], the same $M_1^2 + N_1^2 - 1$ columns of $\mathbf{M}$ are linearly independent, implying that system (38) has only the trivial solution. Hence, by Fubini's theorem [7, Theorem C, p.148], Ω_{11} is a set of full measure. ∎

Proof of Theorem 1.10. We show that assumptions 1) and 2) in Theorem 1.3 hold for generic matrix $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{D}_1, \dots, \mathbf{D}_K$, where $\mathbf{B} = \mathbf{A}$ or $\mathbf{B} = \mathbf{A}^*$ and $\mathbf{D}_1, \dots, \mathbf{D}_K$ may be symmetric, Hermitian or unconstrained.

1) Since $\mathbf{A}$ is generic and $I \geq \sum N_r$, it follows that $\mathbf{A}$ and $\mathbf{B}$ has full column rank.

2) WLOG we assume that $K = 3$. We proceed as in the proof of Theorem 1.5. First we show that the subspace of the solutions of (9) has dimension R for a specific choice of $\mathbf{D}_1$, $\mathbf{D}_2$, and $\mathbf{D}_3$. Let $\mathbf{D}_1 = \mathbf{I}_{\sum N_r}$, $\mathbf{D}_2$ be a diagonal matrix with distinct real values on the main diagonal and let $\mathbf{D}_3 = \operatorname{blockdiag}(\mathbf{D}_{1,3}, \dots, \mathbf{D}_{R,3})$, where the entries of $\mathbf{D}_{1,3}, \dots, \mathbf{D}_{R,3}$ are nonzero. Then $\mathbf{X}\mathbf{D}_1 = \mathbf{D}_1\mathbf{Y}$ implies that $\mathbf{X} = \mathbf{Y}$. From $\mathbf{X}\mathbf{D}_2 = \mathbf{D}_2\mathbf{Y} = \mathbf{D}_2\mathbf{X}$, it follows that $\mathbf{X} = \operatorname{diag}(x_1, \dots, x_{\sum N_r})$. Finally, from $\operatorname{diag}(x_1, \dots, x_{\sum N_r})\mathbf{D}_3 = \mathbf{X}\mathbf{D}_3 = \mathbf{D}_3\mathbf{Y} = \mathbf{D}_3 \operatorname{diag}(x_1, \dots, x_{\sum N_r})$ it follows that $\mathbf{X} = \mathbf{Y} = \operatorname{blockdiag}(c_1\mathbf{I}_{N_1}, \dots, c_R\mathbf{I}_{N_R})$ for some $c_1, \dots, c_R \in \mathbb{F}$. Thus, assumption 2) in Theorem 1.3 holds for a specific choice of $\mathbf{D}_1$, $\mathbf{D}_2$, and $\mathbf{D}_3$. Hence some $2(\sum N_r)^2 - R$ columns of $\mathbf{M}_{\mathcal{D}}$ are linearly independent. From [6, Lemma 6.3] it follows that if $\mathbf{D}_1$, $\mathbf{D}_2$, and $\mathbf{D}_3$ are generic symmetric, Hermitian or unconstrained matrices, then the same $2(\sum N_r)^2 - R$ columns of $\mathbf{M}_{\mathcal{D}}$ are linearly independent, implying that $\dim \operatorname{Null}(\mathbf{M}_{\mathcal{D}}) = R$. ∎

403 REFERENCES

404 [1] Y. Cai and C. Liu, *An algebraic approach to nonorthogonal general joint block diagonaliza-*
405 *tion*, SIAM J. Matrix Anal. Appl., 38(1) (2017), pp. 50–71.
406 [2] G. Chabriel, M. Kleinsteuber, E. Moreau, H. Shen, P. Tichavsky, and A. Yere-
407 dor, *Joint matrices decompositions and blind source separation: A survey of methods,*
408 *identification, and applications*, IEEE Signal Processing Magazine, 31 (2014), pp. 34–43.
409 [3] L. De Lathauwer, *Decompositions of a higher-order tensor in block terms — Part II: Defi-*
410 *nitions and uniqueness*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1033–1066.
411 [4] L. De Lathauwer, *Blind separation of exponential polynomials and the decomposition of a*
412 *tensor in rank-*$(L_r, L_r, 1)$ *terms*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 1451–1474.
413 [5] L. De Lathauwer, B. De Moor, and J. Vandewalle, *A multilinear singular value decom-*
414 *position*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1253–1278.
415 [6] I. Domanov and L. De Lathauwer, *On the uniqueness of the canonical polyadic decompo-*
416 *sition of third-order tensors — Part II: Overall uniqueness*, SIAM J. Matrix Anal. Appl.,
417 34 (2013), pp. 876–903.
418 [7] P. R. Halmos, *Measure theory*, Springer-Verlag, New-York, 1974.
419 [8] T. Maehara and K. Murota, *Algorithm for error-controlled simultaneous block-*
420 *diagonalization of matrices*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 605–620.
421 [9] Y. Ming, *On partial and generic uniqueness of block term tensor decompositions*, Annali
422 Dell'Universita'Di Ferrara, 60 (2014), pp. 465–493.
423 [10] M. Sørensen and L. De Lathauwer, *Coupled canonical polyadic decompositions and (cou-*
424 *pled) decompositions in multilinear rank-*$(L_{r_n}, L_{r_n}, 1)$ *terms—Part I: Uniqueness*, SIAM
425 J. Matrix Anal. Appl., 36 (2015), pp. 496–522.