# Tensor and coupled decompositions in block terms: uniqueness and irreducibility

#### Dana Lahat<sup>1,2</sup> Christian Jutten<sup>2</sup>

<sup>1</sup>Institut de Recherche en Informatique de Toulouse (IRIT), CNRS, France <sup>2</sup>Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, 38000 Grenoble, France

June 12, 2018





## A simple model for the data



- $\xi_{dt}$  is the (d, t)th data sample (e.g., observation)
- Each observation is a sum of contributions from  $R \ge 2$  signals
- dth sensor and ith signal are related by a coefficient a<sub>di</sub>
- We ignore noise (any contribution not explained by this model)

### A not-so-simple model for data analysis

$$D \ge 2 \text{ sensors} \left\{ \underbrace{\Xi}_{T \ge 2 \text{ samples } R} = \underbrace{\mathbf{A}}_{R} \underbrace{\mathbf{\Sigma}^{\top}}_{T \ge 2 \text{ samples } R} = \underbrace{\begin{bmatrix} \mathbf{a}_{1} \boldsymbol{\sigma}_{1}^{\top} \\ \vdots \\ \mathbf{\sigma}_{1}^{\top} \\ \mathbf{a}_{1} \\ \text{sum of } R > 2 \text{ rank-1 terms} \end{bmatrix} \right\}$$

Desired uniqueness: up to arbitrary scaling  $(\mathbf{a}_i \lambda_i^{-1})(\lambda_i \boldsymbol{\sigma}_i^{\top})$  and ordering

This factorization is generally not unique for  $R \ge 2$ :

 $\boldsymbol{\Xi} = \boldsymbol{A}\boldsymbol{\Sigma}^{\top} = \boldsymbol{A}\boldsymbol{Z}^{-1}\,\boldsymbol{Z}\boldsymbol{\Sigma}^{\top} \text{ for any nonsingular } \boldsymbol{R} \times \boldsymbol{R} \text{ matrix } \boldsymbol{Z}$ 

#### Why do we need a unique decomposition?

Uniqueness is necessary to achieve interpretability, i.e., attach physical meaning to the output [Harshman, 1970], [Cattell, 1944]

### A not-so-simple model for data analysis

$$D \ge 2 \text{ sensors} \left\{ \underbrace{\Xi}_{T \ge 2 \text{ samples } R} = \underbrace{\mathbf{A}}_{R} \underbrace{\mathbf{\Sigma}^{\top}}_{T \ge 2 \text{ samples } R} = \underbrace{\begin{bmatrix} \mathbf{a}_{1} \boldsymbol{\sigma}_{1}^{\top} \\ \vdots \\ \mathbf{\sigma}_{1}^{\top} \\ \mathbf{a}_{1} \\ \text{sum of } R > 2 \text{ rank-1 terms} \end{bmatrix} \right\}$$

Desired uniqueness: up to arbitrary scaling  $(\mathbf{a}_i \lambda_i^{-1})(\lambda_i \boldsymbol{\sigma}_i^{\top})$  and ordering

This factorization is generally not unique for  $R \ge 2$ :

 $\boldsymbol{\Xi} = \boldsymbol{A}\boldsymbol{\Sigma}^\top = \boldsymbol{A}\boldsymbol{Z}^{-1}\,\boldsymbol{Z}\boldsymbol{\Sigma}^\top \ \text{ for any nonsingular } \boldsymbol{R}\times\boldsymbol{R} \ \text{matrix } \boldsymbol{Z}$ 

#### Why do we need a unique decomposition?

Uniqueness is necessary to achieve interpretability, i.e., attach physical meaning to the output [Harshman, 1970], [Cattell, 1944]

 $\implies$  Diversity

### Data with three-way diversity

A polyadic decomposition of  $\mathcal{T} \in \mathbb{F}^{I \times J \times K}$  in sum of R rank-1 tensors:



$$\underbrace{\mathbf{A}}_{I \times R} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_R \end{bmatrix}, \underbrace{\mathbf{B}}_{J \times R} = \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_R \end{bmatrix}, \underbrace{\mathbf{C}}_{K \times R} = \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_R \end{bmatrix}$$

- When R is minimal, this decomposition is called the Canonical polyadic decomposition (CPD), and R is the rank of T, denoted r<sub>T</sub>
- The CPD of a tensor is unique if it is subject only to trivial indeterminacies: arbitrary permuting of rank-1 terms, and arbitrary scaling of vectors within each rank-1 term

# Collinearity and Kruskal rank

The concept of collinearity, and linear dependence among columns of factor matrices, is fundamental to the uniqueness of tensor decompositions

- If two or more columns of a factor matrix of a tensor  ${\cal T}$  have collinear columns, then the overall CPD is not unique
- For more than two columns, we have

#### Kruskal Rank or k-rank [Kruskal, 1977]

 $k_A$  is the largest number such that any set of  $k_A$  columns (vectors) of **A** is linearly independent

Generically,  $k_{\mathbf{A}} = \min(I, R)$  if  $\mathbf{A} \in \mathbb{F}^{I \times R}$ .

#### Kruskal's condition [Kruskal, 1977]

Let  $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$  and  $k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \ge 2R + 2$ . Then  $r_{\mathcal{T}} = R$  and the CPD of  $\mathcal{T}$  is unique.

# Collinearity and Kruskal rank

The concept of collinearity, and linear dependence among columns of factor matrices, is fundamental to the uniqueness of tensor decompositions

- If two or more columns of a factor matrix of a tensor  ${\cal T}$  have collinear columns, then the overall CPD is not unique
- For more than two columns, we have

#### Kruskal Rank or k-rank [Kruskal, 1977]

 $k_A$  is the largest number such that any set of  $k_A$  columns (vectors) of **A** is linearly independent

Generically,  $k_{\mathbf{A}} = \min(I, R)$  if  $\mathbf{A} \in \mathbb{F}^{I \times R}$ .

#### Kruskal's condition [Kruskal, 1977]

Let  $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$  and  $k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \ge 2R + 2$ . Then  $r_{\mathcal{T}} = R$  and the CPD of  $\mathcal{T}$  is unique.

#### In this talk, we suggest generalizations to these concepts

## From rank-1 to rank- $L_i$ , $L_i \ge 1$

- $\checkmark\,$  Many types of latent phenomena are not rank-1
  - Fetal electrocardiogram (FECG) [De Lathauwer et al., 1995]
  - Cosmic microwave background radiation (CMB) [Cardoso et al., 2008]
- Higher accuracy, better interpretability
- ✓ Computational advantages: no need to separate within subspaces

$$D \ge 2\left\{ \underbrace{\Xi}_{T \ge 2} = \underbrace{\mathbf{A}}_{\sum_{i=1}^{R} L_{i}} \underbrace{\mathbf{\Sigma}^{\top}}_{\mathbf{A}_{1}} = \underbrace{\Box}_{\mathbf{\Sigma}_{1}^{\top}} \underbrace{\mathbf{\Sigma}_{1}^{\top}}_{\mathbf{A}_{1}} \underbrace{\mathbf{L}_{1}}_{\mathbf{A}_{R}} \underbrace{\mathbf{L}_{R}}_{\mathbf{A}_{R}} \underbrace{\mathbf{L}_{R}} \underbrace{\mathbf{L}_{R}} \underbrace{\mathbf{L}_{R}}$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_R \end{bmatrix}, \ \mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_1 & \mathbf{\Sigma}_2 & \cdots & \mathbf{\Sigma}_R \end{bmatrix}$$

Desired uniqueness: up to arbitrary  $L_i \times L_i$  nonsingular  $Z_i$ :  $(\mathbf{A}_i \mathbf{Z}_i^{-1})(\mathbf{Z}_i \mathbf{\Sigma}_i^{\top}) \forall r$ , and ordering  $\rightarrow$  Only span $(\mathbf{A}_i)$  can be uniquely identified This factorization is generally not unique for  $R \ge 2$ 

### Do the rank- $L_i$ terms exist?

Let us look again at the decomposition

$$\boxed{\Xi} = \boxed{\mathbf{A}} \boxed{\boldsymbol{\Sigma}^{\top}} = \boxed{\underbrace{\boldsymbol{\Sigma}^{\top}}_{\mathbf{A}_{1}} \mathbf{\Sigma}_{1}^{\top}} + \dots + \boxed{\underbrace{\boldsymbol{\Sigma}^{\top}}_{\mathbf{A}_{R}} \mathbf{\Sigma}_{R}^{\top}} \mathbf{\Sigma}_{R}^{\top}$$

with  $L_i \ge 2$  for at least one i = 1, ..., R. Why (or when) can't we write  $\Xi$  as a sum of  $(\sum_{i=1} L_i)$  rank-1 terms?

#### Reducibility and irreducibility

When a term of rank  $\geq 2$  can be factorized into several terms of smaller rank, we say that it is **reducible**. Otherwise, it is **irreducible**.

- Reducible terms may cause non-uniqueness and non-identifiability
- Reducibility and irreducibility depend on
  - data
  - model

Decomposition of  $K \ge 2$  datasets in sum of rank- $L_i$  terms



 $\mathbf{A}^{[k]}_{i} = [\mathbf{A}^{[k]}_{1}|\cdots|\mathbf{A}^{[k]}_{R}], \ \mathbf{A}^{[k]}_{i} \in \mathbb{R}^{D \times L_{i}}, \ k = 1, \dots, K$  $\mathbf{A}^{[k]}_{i} \mathbf{\Sigma}^{[k]\top}_{i} = \mathbf{A}^{[k]}_{i} \mathbf{Z}^{-[k]}_{i} \mathbf{Z}^{[k]}_{i} \mathbf{\Sigma}^{[k]\top}_{i}, \ \mathbf{Z}^{[k]}_{i} \in \mathbb{R}^{L_{i} \times L_{i}} \text{ nonsingular } \forall k, i, t$ 

Potential applications: Data fusion, multiset data analysis, frequency domain analysis of convolutive mixtures, and more

D. Lahat, C. Jutten

TRICAP 2018 8 / 23

# Decomposition of $K \ge 2$ datasets in sum of rank- $L_i$ terms

Independent multivariate random processes



For a specific dataset (mixture) k (or K = 1)

- Independent subspace analysis (ISA) [Comon, 1995], [De Lathauwer et al., 1995] [Cardoso, 1998]
- If  $L_i = 1 \ \forall i$ , independent component analysis (ICA) [Comon, 1994]

Therefore, this model subsumes and generalizes tensor-related ICA results

# Decomposition of $K \ge 2$ datasets in sum of rank- $L_i$ terms

Independent multivariate random processes



- Soft links among datasets [Lahat, Adalı, Jutten, 2015]
- Uniqueness: up to  $(\mathbf{A}_i^{[k]} \mathbf{Z}_i^{-[k]}) (\mathbf{Z}_i^{[k]} \mathbf{\Sigma}_i^{[k]\top})$  and global ordering
- Joint independent subspace analysis (JISA) [Lahat and Jutten, 2014]

•  $L_i = 1 \; orall k, i 
ightarrow$  Independent vector analysis (IVA) [Kim et al., 2006]

### A simple JISA model based on second-order statistics

- Simplest JISA model, no diversity among samples within dataset
- Each dataset not unique and not identifiable individually
- Uniqueness, identifiability, and irreducibility, only due to the link among the datasets

Cross-correlation between *i*th and *j*th elements in datasets k and  $\ell$ :

$$\underbrace{\mathbf{S}_{ij}^{[k,\ell]}}_{L_i \times L_j} = \begin{cases} E\{\mathbf{\Sigma}_i^{[k]\top} \mathbf{\Sigma}_i^{[\ell]}\} & \text{nonsingular} \quad i = j \\ E\{\mathbf{\Sigma}_i^{[k]\top} \mathbf{\Sigma}_j^{[\ell]}\} & \mathbf{0} & i \neq j \end{cases}$$

Cross-correlation between all the elements in datasets k and  $\ell$ :

$$\underbrace{\mathbf{S}_{RL_i \times RL_i}^{[k,\ell]}}_{RL_i \times RL_i} = E\{\mathbf{\Sigma}^{[k]\top}\mathbf{\Sigma}^{[\ell]}\} = \begin{bmatrix} \mathbf{S}_{11}^{[k,\ell]} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_{RR}^{[k,\ell]} \end{bmatrix} = \underbrace{\mathbf{S}_{11}^{[k,\ell]} \oplus \cdots \oplus \mathbf{S}_{RR}^{[k,\ell]}}_{\text{block-diagonal}}$$

### JISA as a coupled block decomposition (or diagonalization)

Cross-correlation between observations in datasets k and  $\ell$ :

$$\mathbf{X}^{[k,\ell]} = E\{\mathbf{\Xi}^{[k]}\mathbf{\Xi}^{[\ell]\top}\} = \mathbf{A}^{[k]} \underbrace{\mathbf{S}^{[k,\ell]}}_{\text{block-diagonal}} \mathbf{A}^{[\ell]\top} = \sum_{i=1}^{R} \mathbf{A}^{[k]}_{i} \mathbf{S}^{[k,\ell]}_{ii} \mathbf{A}^{[\ell]\top}_{i} \quad \forall k, \ell$$
  
Coupled block decomposition (CBD) of  $\{\mathbf{X}^{[k,\ell]}\}_{k,\ell=1}^{K}$ 

When  $\mathbf{A}^{[k]}$  is nonsingular  $\forall k$ :

$$\mathbf{A}^{-[k]}\mathbf{X}^{[k,\ell]}\mathbf{A}^{-[\ell]\top} = \mathbf{S}^{[k,\ell]} = \begin{bmatrix} \mathbf{S}_{11}^{[k,\ell]} & \mathbf{0} \\ & \ddots \\ \mathbf{0} & \mathbf{S}_{RR}^{[k,\ell]} \end{bmatrix} \quad \forall k,k$$

Coupled block diagonalization (CBD) of  $\{\mathbf{X}^{[k,\ell]}\}_{k,\ell=1}^{K}$ 

Here, the data cannot be stacked in a single tensor

# Coupled block diagonalization - example



- R = 3 low-rank terms in each dataset
- Dimensions of low-rank terms in each dataset:  $L_1 = 2$ ,  $L_2 = 1$ ,  $L_3 = 3$
- K ≥ 2 datasets

### Coupled block diagonalization - example



#### Reducibility and irreducibility in CBD

- If, for fixed *i*, all blocks S<sup>[k,ℓ]</sup><sub>ii</sub> can be further block-diagonalized with the same block pattern, the CBD of {X<sup>[k,ℓ]</sup>}<sup>K</sup><sub>k,ℓ=1</sub> is said to be reducible. Otherwise, it is irreducible.
- Generally irreducible as soon as  $K \ge 3$

# CBD uniqueness through JISA identifiability

- Since our statistical model is simple, we can derive the Fisher information matrix (FIM) in closed form
- The inverse of the FIM is a lower bound on the covariance of the parameters
- A singular FIM means that the model is not identifiable
- When  $\{\mathbf{X}^{[k,\ell]}\}_{k,\ell=1}^{K}$  represents the sufficient statistics of JISA, CBD uniqueness  $\Leftrightarrow$  JISA identifiability
- We assume irreducibility (i.e., reducible solutions excluded)

# CBD uniqueness and JISA identifiability

#### Theorem ([Lahat & Jutten, 2015, 2018])

A coupled block diagonalization of  $\{\mathbf{X}^{[k,\ell]}\}_{k,\ell=1}^{K}$  with irreducible block terms is not unique iff there exists at least one pair (i,j),  $i \neq j$ , for which  $L_j = L_i$  and

$$\mathbf{S}_{jj}^{[k,\ell]} = \mathbf{\Psi}^{[k]} \mathbf{S}_{jj}^{[k,\ell]} \mathbf{\Psi}^{[\ell]\top} \quad \forall k, \ell$$

where  $\{\Psi^{[k]}\}_{k=1}^{K}$  are nonsingular  $L_i \times L_i$  matrices.

If  $\{\mathbf{X}^{[k,\ell]}\}_{k,\ell=1}^{K}$  are the sufficient statistics of the JISA model, then this theorem equally characterizes the necessary and sufficient conditions for JISA identifiability.

In this example, R = 3,  $\{\mathbf{S}_{ii}^{[k,\ell]}\}_{k,\ell=1}^{K}$  irreducible and with no zero values.





- In this example, R = 3,  $\{\mathbf{S}_{ii}^{[k,\ell]}\}_{k,\ell=1}^{K}$  irreducible and with no zero values.
  - Collect all *r*th blocks of all diagonals in one matrix, whose (k, ℓ)th element is S<sup>[k,ℓ]</sup><sub>rr</sub>



In this example, R = 3,  $\{\mathbf{S}_{ii}^{[k,\ell]}\}_{k,\ell=1}^{K}$  irreducible and with no zero values.

 Collect all *r*th blocks of all diagonals in one matrix, whose (k, ℓ)th element is S<sup>[k,ℓ]</sup><sub>rr</sub>







In this example, R = 3,  $\{\mathbf{S}_{ii}^{[k,\ell]}\}_{k,\ell=1}^{K}$  irreducible and with no zero values.

 Collect all *r*th blocks of all diagonals in one matrix, whose (k, ℓ)th element is S<sup>[k,ℓ]</sup><sub>rr</sub>



where  $\mathbf{\Psi}^{[k]}$  are nonsingular  $L_i \times L_i$  matrices [Lahat et al., 2015]

 $\Rightarrow\,$  Signals with different block sizes are always identifiable

# How is CBD related to tensor decompositions? Let $\mathbf{A}^{[k]} \mapsto \mathbf{A} \ \forall k$ , $(k, \ell) \mapsto q$ , q = 1 : Q. Then

$$\mathbf{X}^{(q)} = \mathbf{A}\mathbf{S}^{(q)}\mathbf{A}^{\top} = \sum_{i=1}^{R} \mathbf{A}_{i}\mathbf{S}^{(q)}_{ii}\mathbf{A}_{i}^{\top} \forall q \quad \Leftrightarrow \quad \mathcal{X} = \sum_{i=1}^{R} \mathcal{C}_{i} \times_{1} \mathbf{A}_{i} \times_{2} \mathbf{A}_{i}$$



- $\mathcal{X} \in \mathbb{R}^{D \times D \times Q}$ ,  $\mathbf{A} = [\mathbf{A}_1, \dots, \mathbf{A}_R]$ ,  $\mathbf{A}_i \in \mathbb{R}^{D \times L_i}$ ,  $\mathbf{S}_{ii}^{(q)} \in \mathbb{R}^{L_i \times L_i}$ ,  $\mathbf{S}^{(q)} = \mathbf{S}_{ii}^{(1)} \oplus \dots \oplus \mathbf{S}_{ii}^{(Q)}$ ,  $\mathcal{C}_i \in \mathbb{R}^{L_i \times L_i \times Q}$
- Each core tensor  $C_i$  is irreducible
- A special case of rank-( $L_r, M_r, \cdot$ ) block term decomposition (BTD) [De Lathauwer, 2008]

# How is CBD related to tensor decompositions? Let $\mathbf{A}^{[k]} \mapsto \mathbf{A} \ \forall k$ , $(k, \ell) \mapsto q$ , q = 1 : Q. Then

$$\mathbf{X}^{(q)} = \mathbf{A}\mathbf{S}^{(q)}\mathbf{A}^{\top} = \sum_{i=1}^{R} \mathbf{A}_{i}\mathbf{S}^{(q)}_{ii}\mathbf{A}_{i}^{\top} \forall q \quad \Leftrightarrow \quad \mathbf{\mathcal{X}} = \sum_{i=1}^{R} \mathbf{C}_{i} \times_{1} \mathbf{A}_{i} \times_{2} \mathbf{A}_{i}$$



- $\mathcal{X} \in \mathbb{R}^{D \times D \times Q}$ ,  $\mathbf{A} = [\mathbf{A}_1, \dots, \mathbf{A}_R]$ ,  $\mathbf{A}_i \in \mathbb{R}^{D \times L_i}$ ,  $\mathbf{S}_{ii}^{(q)} \in \mathbb{R}^{L_i \times L_i}$ ,  $\mathbf{S}^{(q)} = \mathbf{S}_{ii}^{(1)} \oplus \dots \oplus \mathbf{S}_{ii}^{(Q)}$ ,  $\mathcal{C}_i \in \mathbb{R}^{L_i \times L_i \times Q}$
- Each core tensor  $C_i$  is irreducible
- A special case of rank-(L<sub>r</sub>, M<sub>r</sub>, ·) block term decomposition (BTD) [De Lathauwer, 2008]
- If **A** is nonsingular: joint block diagonalization (JBD) of  $\mathcal{X}$  by  $\mathbf{A}^{-1}$

$$\mathbf{A}^{-1}\mathbf{X}^{(q)}\mathbf{A}^{- op} = \mathbf{S}^{(q)} \ \forall q \quad \Leftrightarrow \quad \mathcal{X} imes_1 \mathbf{A}^{-1} imes_2 \mathbf{A}^{-1} =$$

D. Lahat, C. Jutten

### CBD uniqueness versus JBD uniqueness

Recall the equivalence relation for CBD non-uniqueness:

$$\mathbf{S}_{ii}^{[k,\ell]} = \mathbf{\Psi}^{[k]} \mathbf{S}_{ii}^{[k,\ell]} \mathbf{\Psi}^{[\ell]\top} \quad \forall k, \ell$$

where  $\{\Psi^{[k]}\}_{k=1}^{K}$  are nonsingular  $L_i \times L_i$  matrices

Applying the same type of simplification,  $(k, \ell) \mapsto q$ , to the CBD uniqueness results, and setting  $\Psi^{[k]} \mapsto \Phi$ , we obtain

$$\mathbf{S}_{ii}^{(q)} = \mathbf{\Phi} \mathbf{S}_{ii}^{(q)} \mathbf{\Phi}^{\top} \forall q \quad \Leftrightarrow \quad \mathcal{C}_j = \mathcal{C}_i \times_1 \mathbf{\Phi} \times_2 \mathbf{\Phi}$$

where  $\mathbf{\Phi}$  is a nonsingular  $L_i \times L_i$  matrix.

### CBD uniqueness versus JBD uniqueness

Recall the equivalence relation for CBD non-uniqueness:

$$\mathbf{S}_{ii}^{[k,\ell]} = \mathbf{\Psi}^{[k]} \mathbf{S}_{ii}^{[k,\ell]} \mathbf{\Psi}^{[\ell]\top} \quad \forall k, \ell$$

where  $\{\Psi^{[k]}\}_{k=1}^{K}$  are nonsingular  $L_i \times L_i$  matrices

Applying the same type of simplification,  $(k, \ell) \mapsto q$ , to the CBD uniqueness results, and setting  $\Psi^{[k]} \mapsto \Phi$ , we obtain

$$\mathbf{S}_{ii}^{(q)} = \mathbf{\Phi} \mathbf{S}_{ii}^{(q)} \mathbf{\Phi}^{ op} \ \forall q \quad \Leftrightarrow \quad \mathcal{C}_j = \mathcal{C}_i \times_1 \mathbf{\Phi} \times_2 \mathbf{\Phi}$$

where  $\mathbf{\Phi}$  is a nonsingular  $L_i \times L_i$  matrix.

These are the necessary and sufficient conditions for uniqueness of JBD [Lahat, Messer, Cardoso, 2012]

### CBD uniqueness versus JBD uniqueness



where  $\mathbf{\Psi}^{[k]}$  are nonsingular  $L_i \times L_i$  matrices [Lahat et al., 2015]

- A pair (*i*, *j*) of block terms is not identifiable iff its covariance profiles satisfy the given equivalence relation
- Only pairs of sources with the same size can cause non-identifiability

### Implications

We now argue that

Our results are useful also for more general types of data, e.g.,

- Complex-valued
- Singular factor matrices
- Not necessarily positive-definite covariance matrices
- Our results can be regarded as generalizations to the concept of Kruskal's rank

# Non-uniqueness of rank- $(L, M, \cdot)$ BTD

Consider the rank- $(L, M, \cdot)$  block decomposition with R = 2



where  $\mathcal{T} \in \mathbb{C}^{I \times J \times K}$ ,  $\mathcal{C}_i \in \mathbb{C}^{L \times M \times K}$ ,  $\mathbf{A}_i \in \mathbb{C}^{I \times L}$ ,  $\mathbf{B}_i \in \mathbb{C}^{J \times M}$ , i = 1, 2. Consider an alternative rank- $(L, M, \cdot)$  BTD of  $\mathcal{T}$  with

$$\overline{\mathbf{A}} = \sqrt{2}\mathbf{A} \begin{bmatrix} \mathbf{I} & -\mathbf{\Psi}^{-1} \\ \mathbf{I} & \mathbf{\Psi}^{-1} \end{bmatrix}^{-1} , \ \overline{\mathbf{B}} = \sqrt{2}\mathbf{B} \begin{bmatrix} \mathbf{I} & -\mathbf{\Phi}^{-1} \\ \mathbf{I} & \mathbf{\Phi}^{-1} \end{bmatrix}^{-1}$$

where  $\Psi \in \mathbb{F}^{L \times L}$  and  $\Phi \in \mathbb{F}^{M \times M}$  are nonsingular matrices. If  $C_2 = C_1 \times_1 \Psi \times_2 \Phi$ , then

$$\frac{1}{2}\mathcal{C} \times_1 \begin{bmatrix} \mathbf{I} & -\mathbf{\Psi}^{-1} \\ \mathbf{I} & \mathbf{\Psi}^{-1} \end{bmatrix} \times_2 \begin{bmatrix} \mathbf{I} & -\mathbf{\Phi}^{-1} \\ \mathbf{I} & \mathbf{\Phi}^{-1} \end{bmatrix} = \begin{array}{c} \mathcal{C}_1 \\ \mathcal{C}_1 \\ \mathcal{C}_1 \end{array}$$

Rank-(L<sub>r</sub>, M<sub>r</sub>, ·) BTD not unique if C<sub>j</sub> = C<sub>i</sub> ×<sub>i</sub> Ψ ×<sub>2</sub> Φ for some (i, j)
An analogous result can be shown for the CBD

# Generalizing k-rank to core tensors in rank- $(L_r, M_r, \cdot)$ BTD

#### Definition: k-rank for core tensors in rank- $(L_r, M_r, \cdot)$ BTD

- $k_{\mathcal{C}}^{\text{core}} = 1$  if for at least one pair (i, j),  $i \neq j$ ,  $\mathcal{C}_i = \mathcal{C}_j \times_1 \Psi \times_2 \Phi$  with  $\Psi$  and  $\Phi$  nonsingular
- 2  $k_{\mathcal{C}}^{\text{core}} \geq 2 \text{ if } k_{\mathcal{C}}^{\text{core}} \neq 1 \text{ and } \mathcal{C}_r \neq \mathcal{O} \ \forall r$
- ◎  $k_{C}^{core} = 2$  if  $k_{C}^{core} \ge 2$ , and there is a triplet (i, j, k),  $i \ne j \ne k$  satisfying

$$\boldsymbol{\mathcal{C}}_{k} = \boldsymbol{\mathcal{C}}_{j} \times_{1} \boldsymbol{\Psi} \times_{2} \boldsymbol{\Phi} + \boldsymbol{\mathcal{C}}_{j} \times_{1} \boldsymbol{\Psi}' \times_{2} \boldsymbol{\Phi}'$$

with  $\Psi$ ,  $\Psi'$ ,  $\Phi$ , and  $\Phi'$  nonsingular matrices

•  $k_{\mathcal{C}}^{\text{core}} = n \ n \ge 3$ : defined analogously

- **()** Kruskal's rank is a special case when  $L_r = 1 = M_r \ \forall r$
- k<sup>core</sup> may be larger than 1 even if some, but not all, columns in the core tensors are proportional
- Senerically,  $k_{\mathcal{C}}^{\text{core}} = \min(K, R)$  if  $\mathcal{C}_i \in \mathbb{C}^{L \times L \times K}$  and  $L_i = L \forall i$

# Previous generalizations to Kruskal's rank and condition

Motivation: uniqueness of different types of BTD

Generalizing k-rank to partitioned matrices [De Lathauwer, 2008]

The k'-rank of a (not necessarily uniformly) partitioned matrix **A**, denoted by  $k'_{\mathbf{A}}$ , is the maximal number r such that any set of r submatrices of **A** yields a set of linearly independent columns.

Let  $\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \cdots & \mathbf{A}_R \end{bmatrix} \in \mathbb{F}^{l \times LR}$  be uniformly partitioned in R matrices  $\mathbf{A}_r \in {}^{l \times L}$ . Generically,  $k'_{\mathbf{A}} = \min(\lfloor \frac{l}{L} \rfloor, R)$ .

• The k'-rank still depends on individual columns of A

The uniqueness results for BTD in [De Lathauwer, 2008]

- Only for generic block terms  $\Rightarrow$  in particular, generic core tensors
- In the nongeneric case, considers lack of uniqueness only due to reducibility
- No Kruskal-like conditions for the rank- $(L_r, M_r, \cdot)$  BTD

# A suggested generalization to Kruskal's condition

#### Conjecture

The rank- $(L_r, M_r, \cdot)$  BTD of a tensor  $\mathcal{T}$  is unique if each rank- $(L_r, M_r, \cdot)$  term is irreducible, and

$$k'_{\mathbf{A}} + k'_{\mathbf{B}} + k^{\mathsf{core}}_{\mathcal{C}} \ge 2R + 2$$

- **1** *R* must reflect the number of irreducible terms
- **2** Kruskal's condition is a special case when  $L_r = 1 = M_r \ \forall r$
- Section 2008 Sect

# Concluding remarks

- k-rank for CBD can be defined analogously.
- e However, extending Kruskal's condition for CBD is less obvious, because of the multiple datasets
- We presented results on the uniqueness of rank-(L<sub>r</sub>, M<sub>r</sub>, ·) BTD and CBD that indicate that the concept of Kruskal's rank may −and sould− be generalized in new directions, to accommodate more elaborate structures in the data
- Validation and/or proof needed for conjectures
- Potentially useful for new uniqueness results on BTD and coupled decompositions
- **o** Irreducibility of the block terms is a prerequisite for uniqueness

This work was supported by the project CHESS, 2012-ERC-AdG-320684. GIPSA-Lab is a partner of the LabEx PERSYVAL-Lab (ANR-11-LABX-0025).

chess