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

Dana Lahat ${ }^{1,2}$ Christian Jutten ${ }^{2}$<br>${ }^{1}$ Institut de Recherche en Informatique de Toulouse (IRIT), CNRS, France<br>${ }^{2}$ Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, 38000 Grenoble, France

June 12, 2018



## A simple model for the data

$$
D \geq 2 \text { sensors }\{\underbrace{\text { Observations }}_{\underbrace{\square=}}=\underbrace{\boxed{A}}_{R} \boldsymbol{\Sigma}^{\top}=\overbrace{\underbrace{\overbrace{\mathbf{a}_{1}}^{\sigma_{1}^{\top}}}_{\text {sum of } R \geq 2 \text { rank-1 terms }}+\cdots+\overbrace{\mathbf{a}_{R}}^{\boldsymbol{\sigma}_{R}^{\top}}}^{\mathbf{a}_{1} \boldsymbol{\sigma}_{1}^{\top}}
$$

$$
\begin{aligned}
& \overbrace{\mathbf{A}}^{D \times R}=\left[\begin{array}{llll}
\mathbf{a}_{1} & \mathbf{a}_{2} & \cdots & \mathbf{a}_{R}
\end{array}\right] \quad \text { mixing matrix } \\
& \overbrace{\boldsymbol{\Sigma}}^{R \times T}=\left[\begin{array}{llll}
\boldsymbol{\sigma}_{1} & \sigma_{2} & \cdots & \boldsymbol{\sigma}_{R}
\end{array}\right] \quad \text { (latent) signals }
\end{aligned}
$$

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


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

$$
D \geq 2 \text { sensors }\{\underbrace{\underbrace{\equiv}_{R}}_{T \geq 2 \text { samples }}=\underbrace{\boldsymbol{\Sigma}^{\top}}_{\underbrace{\mathbf{A}}_{\text {sum of } R \geq 2 \text { rank-1 terms }}}=\overbrace{\overbrace{\mathbf{a}_{1}}^{\boldsymbol{\sigma}_{1}^{\top}}}^{\mathbf{a}_{1} \boldsymbol{\sigma}_{1}^{\top}}+\cdots+\rrbracket_{\mathbf{a}_{R}}^{\boldsymbol{\sigma}_{R}^{\top}}
$$

Desired uniqueness: up to arbitrary scaling $\left(\mathbf{a}_{i} \lambda_{i}^{-1}\right)\left(\lambda_{i} \boldsymbol{\sigma}_{i}^{\top}\right)$ and ordering This factorization is generally not unique for $R \geq 2$ :

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

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

$$
D \geq 2 \text { sensors }\{\underbrace{\underbrace{\equiv}_{R}}_{T \geq 2 \text { samples }}=\underbrace{\boldsymbol{\Sigma}^{\top}}_{\underbrace{\mathbf{A}}_{\text {sum of } R \geq 2 \text { rank-1 terms }}}=\overbrace{\overbrace{\mathbf{a}_{1}}^{\boldsymbol{\sigma}_{1}^{\top}}}^{\mathbf{a}_{1} \boldsymbol{\sigma}_{1}^{\top}}+\cdots+\rrbracket_{\mathbf{a}_{R}}^{\boldsymbol{\sigma}_{R}^{\top}}
$$

Desired uniqueness: up to arbitrary scaling $\left(\mathbf{a}_{i} \lambda_{i}^{-1}\right)\left(\lambda_{i} \boldsymbol{\sigma}_{i}^{\top}\right)$ and ordering This factorization is generally not unique for $R \geq 2$ :

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

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]

$$
\Longrightarrow \text { Diversity }
$$

## 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:

$\underbrace{\mathbf{A}}_{I \times R}=\left[\begin{array}{lll}\mathbf{a}_{1} & \cdots & \mathbf{a}_{R}\end{array}\right], \underbrace{\mathbf{B}}_{J \times R}=\left[\begin{array}{lll}\mathbf{b}_{1} & \cdots & \mathbf{b}_{R}\end{array}\right], \underbrace{\mathbf{C}}_{K \times R}=\left[\begin{array}{lll}\mathbf{c}_{1} & \cdots & \mathbf{c}_{R}\end{array}\right]$

- When $R$ is minimal, this decomposition is called the Canonical polyadic decomposition (CPD), and $R$ is the rank of $\mathcal{T}$, denoted $r_{\mathcal{T}}$
- 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 $\boldsymbol{\mathcal { T }}$ have collinear columns, then the overall CPD is not unique
- For more than two columns, we have


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

$k_{\mathbf{A}}$ is the largest number such that any set of $k_{\mathbf{A}}$ columns (vectors) of $\mathbf{A}$ 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 $k_{\mathbf{A}}+k_{\mathbf{B}}+k_{\mathbf{C}} \geq 2 R+2$. 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 $\boldsymbol{\mathcal { T }}$ have collinear columns, then the overall CPD is not unique
- For more than two columns, we have


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

$k_{\mathbf{A}}$ is the largest number such that any set of $k_{\mathbf{A}}$ columns (vectors) of $\mathbf{A}$ 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 $k_{\mathbf{A}}+k_{\mathbf{B}}+k_{\mathbf{C}} \geq 2 R+2$. 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- $L_{i}, L_{i} \geq 1$

$\checkmark$ 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]
$\checkmark$ Higher accuracy, better interpretability
Computational advantages: no need to separate within subspaces

$$
\left.\left.\begin{array}{rl}
D \geq 2\{ & \underbrace{\square \Xi}_{T \geq 2}=\underbrace{\boxed{A}}_{\sum_{i=1}^{R} L_{i}} \boldsymbol{\Sigma}^{\top}
\end{array}=\square_{\mathbf{A}_{1}}^{\square \boldsymbol{\Sigma}_{1}^{\top}}{ }^{\square}+\cdots+\bigsqcup_{\mathbf{A}_{R}}^{\boldsymbol{\Sigma}_{R}^{\top}}\right\} L_{R}\right\}
$$

Desired uniqueness: up to arbitrary $L_{i} \times L_{i}$ nonsingular $\mathbf{Z}_{i}$ : $\left(\mathbf{A}_{i} \mathbf{Z}_{i}^{-1}\right)\left(\mathbf{Z}_{i} \boldsymbol{\Sigma}_{i}^{\top}\right) \forall r$, and ordering
$\rightarrow$ Only $\operatorname{span}\left(\mathbf{A}_{i}\right)$ can be uniquely identified
This factorization is generally not unique for $R \geq 2$

## Do the rank- $L_{i}$ terms exist?

Let us look again at the decomposition

$$
\left.\left.\bar{\equiv}=\boxed{\mathbf{A}} \boldsymbol{\Sigma}^{\top}=\square_{\mathbf{A}_{1}}^{\square \boldsymbol{\Sigma}_{1}^{\top}}\right\}{L_{1}}^{\square}+\cdots+\prod_{\mathbf{A}_{R}}^{\boldsymbol{\Sigma}_{R}^{\top}}\right\} L_{R}
$$

with $L_{i} \geq 2$ for at least one $i=1, \ldots, R$.
Why (or when) can't we write $\equiv$ 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 $K \geq 2$ datasets in sum of rank- $L_{i}$ terms

$$
\begin{aligned}
& D\{\underbrace{\boxed{\Xi^{[K]}}}_{\tilde{T}}=\underbrace{\boldsymbol{A}^{[K]}}_{\sum_{i=1}^{R} L_{i}} \boldsymbol{\Sigma}^{[K] \top} \mid=\mathbf{A}_{1}^{[K]} \boldsymbol{\Sigma}_{1}^{[K] \top}\}+\cdots+\underbrace{\square \boldsymbol{\Sigma}_{R}^{[K] \top}}_{\mathbf{A}_{R}^{[K]}}\} L_{R} \\
& \mathbf{A}^{[k]}=\left[\mathbf{A}_{1}^{[k]}|\cdots| \mathbf{A}_{R}^{[k]}\right], \mathbf{A}_{i}^{[k]} \in \mathbb{R}^{D \times L_{i}}, k=1, \ldots, K \\
& \mathbf{A}_{i}^{[k]} \boldsymbol{\Sigma}_{i}^{[k] \top}=\mathbf{A}_{i}^{[k]} \mathbf{Z}_{i}^{-[k]} \mathbf{Z}_{i}^{[k]} \boldsymbol{\Sigma}_{i}^{[k] \top}, \mathbf{Z}_{i}^{[k]} \in \mathbb{R}^{L_{i} \times L_{i}} \text { nonsingular } \forall k, i, t
\end{aligned}
$$

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

## Decomposition of $K \geq 2$ 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 $K \geq 2$ datasets in sum of rank- $L_{i}$ terms

Independent multivariate random processes


- Soft links among datasets [Lahat, Adall, Jutten, 2015]
- Uniqueness: up to $\left(\mathbf{A}_{i}^{[k]} \mathbf{Z}_{i}^{-[k]}\right)\left(\mathbf{Z}_{i}^{[k]} \boldsymbol{\Sigma}_{i}^{[k] \top}\right)$ and global ordering
- Joint independent subspace analysis (JISA) [Lahat and Jutten, 2014]
- $L_{i}=1 \forall k, i \rightarrow$ 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 $\ell$ :

$$
\underbrace{\mathbf{S}_{i j}^{[k, \ell]}}_{L_{i} \times L_{j}}=\left\{\begin{array}{lll}
E\left\{\boldsymbol{\Sigma}_{i}^{[k] \top} \boldsymbol{\Sigma}_{i}^{[\ell]}\right\} & \text { nonsingular } & i=j \\
E\left\{\boldsymbol{\Sigma}_{i}^{\left.[k] \boldsymbol{\Sigma}_{j}^{[\ell]}\right\}}\right. & \mathbf{0} & i \neq j
\end{array}\right.
$$

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

$$
\underbrace{\mathbf{S}^{[k, \ell]}}_{R L_{i} \times R L_{i}}=E\left\{\boldsymbol{\Sigma}^{[k] \top} \boldsymbol{\Sigma}^{[l]}\right\}=\left[\begin{array}{ccc}
\mathbf{S}_{11}^{[k, \ell]} & \mathbf{0} \\
\mathbf{0} & \ddots & \mathbf{S}_{R R}^{[k, \ell]}
\end{array}\right]=\underbrace{\mathbf{S}_{11}^{[k, \ell]} \oplus \cdots \oplus \mathbf{S}_{R R}^{[k, \ell]}}_{\text {block-diagonal }}
$$

JISA as a coupled block decomposition (or diagonalization)
Cross-correlation between observations in datasets $k$ and $\ell$ :

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

When $\mathbf{A}^{[k]}$ is nonsingular $\forall k$ :
$\mathbf{A}^{-[k]} \mathbf{X}^{[k, \ell]} \mathbf{A}^{-[\ell] \top}=\mathbf{S}^{[k, \ell]}=\left[\begin{array}{ccc}\mathbf{S}_{11}^{[k, \ell]} & & \mathbf{0} \\ 0 & \ddots & \mathbf{S}_{R R}^{[k, \ell]}\end{array}\right] \quad \forall k, \ell$
Coupled block diagonalization (CBD) of $\left\{\mathbf{X}^{[k, \ell]}\right\}_{k, \ell=1}^{K}$
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: $L_{1}=2, L_{2}=1, L_{3}=3$
- $K \geq 2$ datasets


## Coupled block diagonalization - example

$$
\mathbf{A}^{-[1]} \quad \ldots \quad \mathbf{A}^{-[K]}
$$



## Reducibility and irreducibility in CBD

- If, for fixed $i$, all blocks $\mathbf{S}_{i i}^{[k, \ell]}$ can be further block-diagonalized with the same block pattern, the CBD of $\left\{\mathbf{X}^{[k, \ell]}\right\}_{k, \ell=1}^{K}$ is said to be reducible. Otherwise, it is irreducible.
- Generally irreducible as soon as $K \geq 3$


## 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 $\left\{\mathbf{X}^{[k, \ell]}\right\}_{k, \ell=1}^{K}$ 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 $\left\{\mathbf{X}^{[k, \ell]}\right\}_{k, \ell=1}^{K}$ 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

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

where $\left\{\boldsymbol{\Psi}^{[k]}\right\}_{k=1}^{K}$ are nonsingular $L_{i} \times L_{i}$ matrices.

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

JISA uniqueness and non-identifiability [Lahat \& Jutten, 2015]
In this example, $R=3,\left\{\mathbf{S}_{i i}^{[k, \ell]}\right\}_{k, \ell=1}^{K}$ irreducible and with no zero values.


## JISA uniqueness and non-identifiability [Lahat \& Jutten, 2015]

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


- Collect all $r$ th blocks of all diagonals in one matrix, whose $(k, \ell)$ th element is $\mathbf{S}_{r r}^{[k, \ell]}$



## JISA uniqueness and non-identifiability [Lahat \& Jutten, 2015]

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


- Collect all $r$ th blocks of all diagonals in one matrix, whose $(k, \ell)$ th element is $\mathbf{S}_{r r}^{[k, \ell]}$



## JISA uniqueness and non-identifiability [Lahat \& Jutten, 2015]

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


- Collect all $r$ th blocks of all diagonals in one matrix, whose $(k, \ell)$ th element is $\mathbf{S}_{r r}^{[k, \ell]}$

This CBD is not unique iff for some pair $(i, j), L_{i}=L_{j}$ and $\mathbf{S}_{j j}^{[k, \ell]}=\boldsymbol{\Psi}^{[k]} \mathbf{S}_{i i}^{[k, \ell]} \boldsymbol{\Psi}^{[\ell] \top} \forall k, \ell$

$$
\left[\begin{array}{cc}
\boldsymbol{\square}_{1} & \cdots \\
\vdots & \vdots \\
\square & \cdots
\end{array}\right]=\left[\begin{array}{ccc}
\boldsymbol{\Psi}^{[1]} & & \mathbf{0} \\
& \ddots & \\
\mathbf{0} & & \boldsymbol{\Psi}^{[K]}
\end{array}\right]\left[\begin{array}{ccc}
\square & \cdots & \square \\
\vdots & \vdots \\
\square & \cdots
\end{array}\right]\left[\begin{array}{ccc}
\boldsymbol{\Psi}^{[1]} & & \mathbf{0} \\
& \ddots & \\
\mathbf{0} & & \boldsymbol{\Psi}^{[K]}
\end{array}\right]^{\top}
$$

where $\boldsymbol{\Psi}^{[k]}$ 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

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



- $\mathcal{X} \in \mathbb{R}^{D \times D \times Q}, \mathbf{A}=\left[\mathbf{A}_{1}, \ldots, \mathbf{A}_{R}\right], \mathbf{A}_{i} \in \mathbb{R}^{D \times L_{i}}, \mathbf{S}_{i i}^{(q)} \in \mathbb{R}^{L_{i} \times L_{i}}$, $\mathbf{S}^{(q)}=\mathbf{S}_{i i}^{(1)} \oplus \cdots \oplus \mathbf{S}_{i i}^{(Q)}, \mathcal{C}_{i} \in \mathbb{R}^{L_{i} \times L_{i} \times Q}$
- Each core tensor $\mathcal{C}_{i}$ is irreducible
- A special case of rank- $\left(L_{r}, M_{r}, \cdot\right)$ 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

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



- $\mathcal{X} \in \mathbb{R}^{D \times D \times Q}, \mathbf{A}=\left[\mathbf{A}_{1}, \ldots, \mathbf{A}_{R}\right], \mathbf{A}_{i} \in \mathbb{R}^{D \times L_{i}}, \mathbf{S}_{i i}^{(q)} \in \mathbb{R}^{L_{i} \times L_{i}}$, $\mathbf{S}^{(q)}=\mathbf{S}_{i i}^{(1)} \oplus \cdots \oplus \mathbf{S}_{i i}^{(Q)}, \mathcal{C}_{i} \in \mathbb{R}^{L_{i} \times L_{i} \times Q}$
- Each core tensor $\mathcal{C}_{i}$ is irreducible
- A special case of rank- $\left(L_{r}, M_{r}, \cdot\right)$ block term decomposition (BTD) [De Lathauwer, 2008]
- If $\mathbf{A}$ is nonsingular: joint block diagonalization (JBD) of $\mathcal{X}$ by $\mathbf{A}^{-1}$

$$
\mathbf{A}^{-1} \mathbf{X}^{(q)} \mathbf{A}^{-\top}=\mathbf{S}^{(q)} \forall \boldsymbol{q} \quad \Leftrightarrow \quad \mathcal{X} \times{ }_{1} \mathbf{A}^{-1} \times_{2} \mathbf{A}^{-1}=\mathcal{C}
$$

## CBD uniqueness versus JBD uniqueness

Recall the equivalence relation for CBD non-uniqueness:

$$
\begin{gathered}
\mathbf{S}_{j j}^{[k, \ell]}=\boldsymbol{\Psi}^{[k]} \mathbf{S}_{i i}^{[k, \ell]} \boldsymbol{\Psi}^{[\ell] \top} \quad \forall k, \ell \\
\text { where }\left\{\boldsymbol{\Psi}^{[k]}\right\}_{k=1}^{K} \text { are nonsingular } L_{i} \times L_{i} \text { matrices }
\end{gathered}
$$

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

$$
\begin{gathered}
\mathbf{S}_{i j}^{(q)}=\boldsymbol{\Phi} \mathbf{S}_{i i}^{(q)} \boldsymbol{\Phi}^{\top} \forall q \Leftrightarrow \mathcal{C}_{j}=\mathcal{C}_{i} \times{ }_{1} \boldsymbol{\Phi} \times{ }_{2} \boldsymbol{\Phi} \\
\text { where } \boldsymbol{\Phi} \text { is a nonsingular } L_{i} \times L_{i} \text { matrix. }
\end{gathered}
$$

## CBD uniqueness versus JBD uniqueness

Recall the equivalence relation for CBD non-uniqueness:

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

where $\left\{\boldsymbol{\Psi}^{[k]}\right\}_{k=1}^{K}$ 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 $\boldsymbol{\Psi}^{[k]} \mapsto \boldsymbol{\Phi}$, we obtain

$$
\begin{gathered}
\mathbf{S}_{j j}^{(q)}=\boldsymbol{\Phi} \mathbf{S}_{i i}^{(q)} \boldsymbol{\Phi}^{\top} \forall q \Leftrightarrow \mathcal{C}_{j}=\mathcal{C}_{i} \times_{1} \boldsymbol{\Phi} \times_{2} \boldsymbol{\Phi} \\
\text { where } \boldsymbol{\Phi} \text { is a nonsingular } L_{i} \times L_{i} \text { matrix. }
\end{gathered}
$$

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

## CBD uniqueness versus JBD uniqueness

Compare with JBD:


$$
\mathcal{C}_{j}=\mathcal{C}_{i} \times_{1} \boldsymbol{\Phi} \times_{2} \boldsymbol{\Phi}
$$

Not unique iff for some pair $(i, j), \mathbf{S}_{j j}^{[k, \ell]}=\boldsymbol{\Psi}^{[k]} \mathbf{S}_{i i}^{[k, \ell]} \boldsymbol{\Psi}^{[\ell] \top} \forall k, \ell$

$$
\left.\left[\begin{array}{ccc}
\square & \cdots & \square \\
\vdots & \vdots \\
\square & \cdots & \square
\end{array}\right]=\left[\begin{array}{ccc}
\boldsymbol{\Psi}^{[1]} & & \mathbf{0} \\
& \ddots & \\
\mathbf{0} & & \boldsymbol{\Psi}^{[K]}
\end{array}\right]\left[\begin{array}{cc}
\square & \cdots \\
\vdots & \\
\square & \cdots
\end{array}\right]\left[\begin{array}{ccc}
\boldsymbol{\Psi}^{[1]} & & \mathbf{0} \\
& \ddots & \\
\mathbf{0} & & \boldsymbol{\Psi}
\end{array}\right]^{[K]}\right]^{\top}
$$

where $\boldsymbol{\Psi}^{[k]}$ 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
(1) Our results are useful also for more general types of data, e.g.,

- Complex-valued
- Singular factor matrices
- Not necessarily positive-definite covariance matrices
(2) 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

$$
\overline{\mathbf{A}}=\sqrt{2} \mathbf{A}\left[\begin{array}{cc}
\mathbf{I} & -\boldsymbol{\Psi}^{-1} \\
\mathbf{I} & \boldsymbol{\Psi}^{-1}
\end{array}\right]^{-1}, \overline{\mathbf{B}}=\sqrt{2} \mathbf{B}\left[\begin{array}{cc}
\mathbf{I} & -\boldsymbol{\Phi}^{-1} \\
\mathbf{I} & \boldsymbol{\Phi}^{-1}
\end{array}\right]^{-1}
$$

where $\boldsymbol{\Psi} \in \mathbb{F}^{L \times L}$ and $\boldsymbol{\Phi} \in \mathbb{F}^{M \times M}$ are nonsingular matrices. If $\mathbf{C}_{2}=\mathcal{C}_{1} \times_{1} \boldsymbol{\Psi} \times{ }_{2} \boldsymbol{\Phi}$, then

$$
\frac{1}{2} \mathcal{C} \times{ }_{1}\left[\begin{array}{cc}
\mathbf{I} & -\boldsymbol{\Psi}^{-1} \\
\mathbf{I} & \boldsymbol{\Psi}^{-1}
\end{array}\right] \times \times_{2}\left[\begin{array}{cc}
\mathbf{I} & -\boldsymbol{\Phi}^{-1} \\
\mathbf{I} & \boldsymbol{\Phi}^{-1}
\end{array}\right]=\mathcal{C}_{1} / \mathcal{C}_{1}
$$

- Rank- $\left(L_{r}, M_{r}, \cdot\right)$ BTD not unique if $\mathbf{C}_{j}=\mathcal{C}_{i} \times{ }_{i} \boldsymbol{\Psi} \times{ }_{2} \boldsymbol{\Phi}$ for some $(i, j)$
- An analogous result can be shown for the CBD


## Generalizing k-rank to core tensors in rank- $\left(L_{r}, M_{r}, \cdot\right)$ BTD

Definition: k-rank for core tensors in rank- $\left(L_{r}, M_{r}, \cdot\right)$ BTD
(1) $k_{\mathcal{C}}^{\text {core }}=1$ if for at least one pair $(i, j), i \neq j, \mathcal{C}_{i}=\mathcal{C}_{j} \times_{1} \boldsymbol{\Psi} \times{ }_{2} \boldsymbol{\Phi}$ with $\boldsymbol{\Psi}$ and $\boldsymbol{\Phi}$ nonsingular
(2) $k_{\mathcal{C}}^{\text {core }} \geq 2$ if $k_{\mathcal{C}}^{\text {core }} \neq 1$ and $\mathcal{C}_{r} \neq \mathcal{O} \forall r$
(3) $k_{\mathcal{C}}^{\text {core }}=2$ if $k_{\mathcal{C}}^{\text {core }} \geq 2$, and there is a triplet $(i, j, k), i \neq j \neq k$ satisfying

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

with $\boldsymbol{\Psi}, \boldsymbol{\Psi}^{\prime}, \boldsymbol{\Phi}$, and $\boldsymbol{\Phi}^{\prime}$ nonsingular matrices
(9) $k_{\mathcal{C}}^{\text {core }}=n n \geq 3$ : defined analogously
(1) Kruskal's rank is a special case when $L_{r}=1=M_{r} \forall r$
(2) $k_{\mathcal{C}}^{\text {core }}$ may be larger than 1 even if some, but not all, columns in the core tensors are proportional
(3) Generically, $k_{\mathcal{C}}^{\text {core }}=\min (K, R)$ 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 BTDGeneralizing k-rank to partitioned matrices [De Lathauwer, 2008]
The $\mathrm{k}^{\prime}$-rank of a (not necessarily uniformly) partitioned matrix A, denoted by $k_{\mathbf{A}}^{\prime}$, is the maximal number $r$ such that any set of $r$ submatrices of $\mathbf{A}$ yields a set of linearly independent columns.

Let $\mathbf{A}=\left[\begin{array}{lll}\mathbf{A}_{1} & \cdots & \mathbf{A}_{R}\end{array}\right] \in \mathbb{F}^{I \times L R}$ be uniformly partitioned in $R$ matrices $\mathbf{A}_{r} \epsilon^{I \times L}$. Generically, $k_{\mathbf{A}}^{\prime}=\min \left(\left\lfloor\frac{I}{L}\right\rfloor, R\right)$.

- The k'-rank still depends on individual columns of $\mathbf{A}$ 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- $\left(L_{r}, M_{r}, \cdot\right)$ BTD


## A suggested generalization to Kruskal's condition

## Conjecture

The rank- $\left(L_{r}, M_{r}, \cdot\right)$ BTD of a tensor $\mathcal{T}$ is unique if each rank- $\left(L_{r}, M_{r}, \cdot\right)$ term is irreducible, and

$$
k_{\mathbf{A}}^{\prime}+k_{\mathbf{B}}^{\prime}+k_{\mathcal{C}}^{\text {core }} \geq 2 R+2
$$

(1) $R$ must reflect the number of irreducible terms
(2) Kruskal's condition is a special case when $L_{r}=1=M_{r} \forall r$
(3) Explains all rank- $\left(L_{r}, M_{r}, \cdot\right)$ BTD cases in [De Lathauwer, 2008]

## Concluding remarks

(1) k-rank for CBD can be defined analogously.
(2) However, extending Kruskal's condition for CBD is less obvious, because of the multiple datasets
(3) We presented results on the uniqueness of rank- $\left(L_{r}, M_{r}, \cdot\right) \mathrm{BTD}$ 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
(9) Validation and/or proof needed for conjectures
(0) Potentially useful for new uniqueness results on BTD and coupled decompositions
(0) Irreducibility of the block terms is a prerequisite for uniqueness

## chess

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).

