# Double Contractions of Tensors to Represent Slice-wise Multiplications of Tensors – with Applications to PARAFAC2 and Multi-Carrier MIMO Systems in Wireless Communications


***Martin Haardt***

Ilmenau University of Technology
Communications Research Laboratory
98684 Ilmenau, Germany

E-Mail: Martin.Haardt@tu-ilmenau.de

# Acknowledgments

- Joint work with
  - ⇒ Kristina Naskovska, TU Ilmenau
  - ⇒ Damir Rakhimov, TU Ilmenau
  - ⇒ André de Almeida, Federal University of Ceará (UFC), Fortaleza, Brazil

# Motivation

- Slice-wise multiplication of two tensors
  - $\Rightarrow$ is required in a variety of tensor decompositions
    - PARAFAC2
    - PARATUCK2
  - $\Rightarrow$ and is encountered in many applications
    - biomedical data analysis (EEG, MEG, etc.)
    - multi-carrier MIMO systems
  - $\Rightarrow$ provide a new tensor representation
    - that is not based on a slice-wise (matrix) description
  - $\Rightarrow$ can be represented by a **double contraction** of two tensors
    - efficiently calculated via generalized unfoldings
    - leads to new tensor models that do not depend on the chosen unfolding
    - **reveal the tensor structure** of the data model (**constrained CP**)

# Outline

- Tensor Algebra and Notation
- Computation of Tensor Contractions via Generalized Unfoldings
- Double Contractions of Tensors to Represent Slice-wise Multiplications
  - $\Rightarrow$ Element-wise Multiplication of Vectors and Matrices using Contraction
  - $\Rightarrow$ Slice-wise Multiplication of Tensors using Contraction
- Constrained CP model of PARAFAC2 derived via Double Contractions
- Applications to Multi-Carrier MIMO Systems in Wireless Communications
  - $\Rightarrow$ Khatri-Rao Coded MIMO OFDM Systems
  - $\Rightarrow$ MIMO OFDM Systems
- Conclusions

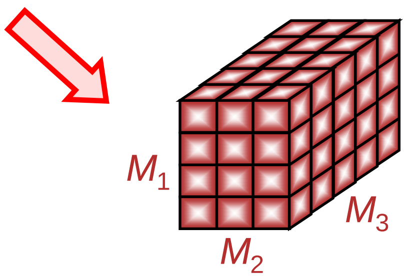
# Outline

- Tensor Algebra and Notation
- Computation of Tensor Contractions via Generalized Unfoldings
- Double Contractions of Tensors to Represent Slice-wise Multiplications
  - $\Rightarrow$ Element-wise Multiplication of Vectors and Matrices using Contraction
  - $\Rightarrow$ Slice-wise Multiplication of Tensors using Contraction
- Constrained CP model of PARAFAC2 derived via Double Contractions
- Applications to Multi-Carrier MIMO Systems in Wireless Communications
  - $\Rightarrow$ Khatri-Rao Coded MIMO OFDM Systems
  - $\Rightarrow$ MIMO OFDM Systems
- Conclusions

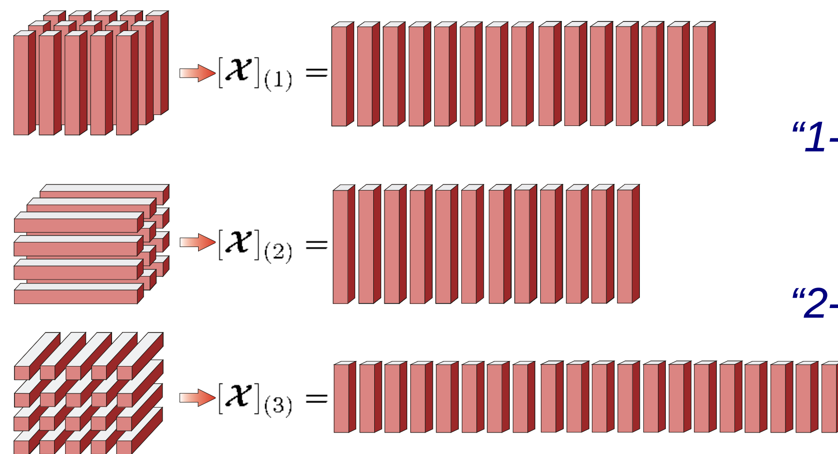
# Tensor algebra

- 3-D tensor = 3-way array

$\mathcal{X} \in \mathbb{C}^{M_1 \times M_2 \times M_3}$




- Unfoldings $[\mathcal{X}]_{(n)}$ $\quad n\text{-rank}(\mathcal{X}) = \text{rank}([\mathcal{X}]_{(n)})$



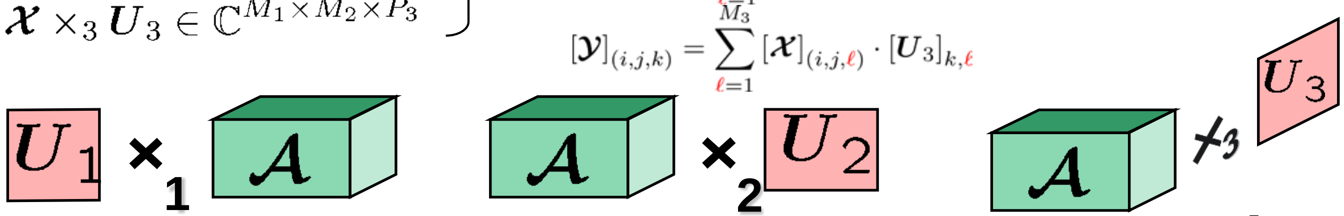

$\in \mathbb{C}^{M_1 \times M_2 \cdot M_3}$ *"1-mode vectors"*

$\in \mathbb{C}^{M_2 \times M_3 \cdot M_1}$ *"2-mode vectors"*

$\in \mathbb{C}^{M_3 \times M_1 \cdot M_2}$ *"3-mode vectors"*

- *n*-mode products between $\mathcal{X} \in \mathbb{C}^{M_1 \times M_2 \times M_3}$ and $\boldsymbol{U}_n \in \mathbb{C}^{P_n \times M_n}$

$$\left.\begin{aligned} \mathcal{Y} &= \mathcal{X} \times_1 \boldsymbol{U}_1 \in \mathbb{C}^{P_1 \times M_2 \times M_3} \\ \mathcal{Y} &= \mathcal{X} \times_2 \boldsymbol{U}_2 \in \mathbb{C}^{M_1 \times P_2 \times M_3} \\ \mathcal{Y} &= \mathcal{X} \times_3 \boldsymbol{U}_3 \in \mathbb{C}^{M_1 \times M_2 \times P_3} \end{aligned}\right\} \Leftrightarrow \begin{aligned} [\mathcal{Y}]_{(i,j,k)} &= \sum_{\ell=1}^{M_1} [\mathcal{X}]_{(\ell,j,k)} \cdot [\boldsymbol{U}_1]_{i,\ell} \\ [\mathcal{Y}]_{(i,j,k)} &= \sum_{\ell=1}^{M_2} [\mathcal{X}]_{(i,\ell,k)} \cdot [\boldsymbol{U}_2]_{j,\ell} \\ [\mathcal{Y}]_{(i,j,k)} &= \sum_{\ell=1}^{M_3} [\mathcal{X}]_{(i,j,\ell)} \cdot [\boldsymbol{U}_3]_{k,\ell} \end{aligned}$$

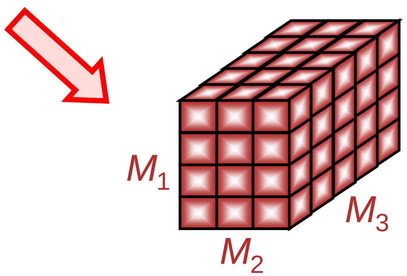



$n = 1$ $\qquad n = 2$ $\qquad n = 3$

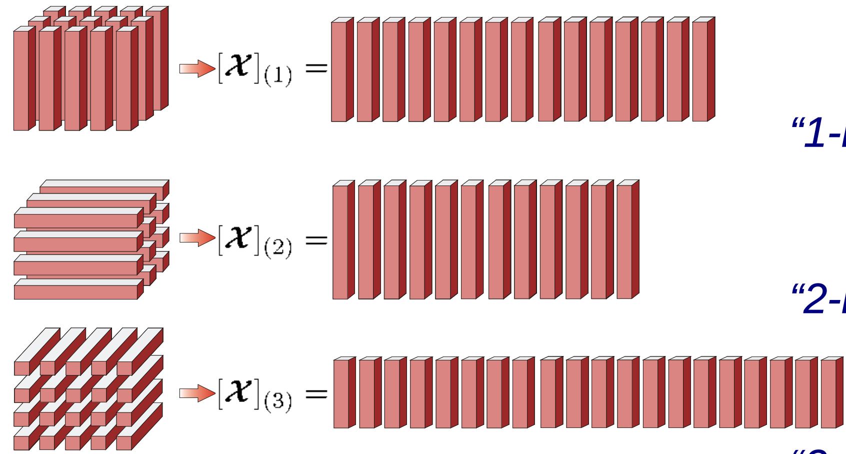
# Tensor algebra

- 3-D tensor = 3-way array

$\boldsymbol{\mathcal{X}} \in \mathbb{C}^{M_1 \times M_2 \times M_3}$




- Unfoldings $[\boldsymbol{\mathcal{X}}]_{(n)}$ $\quad n\text{-rank}(\boldsymbol{\mathcal{X}}) = \text{rank}([\boldsymbol{\mathcal{X}}]_{(n)})$




$\in \mathbb{C}^{M_1 \times M_2 \cdot M_3}$ *"1-mode vectors"*

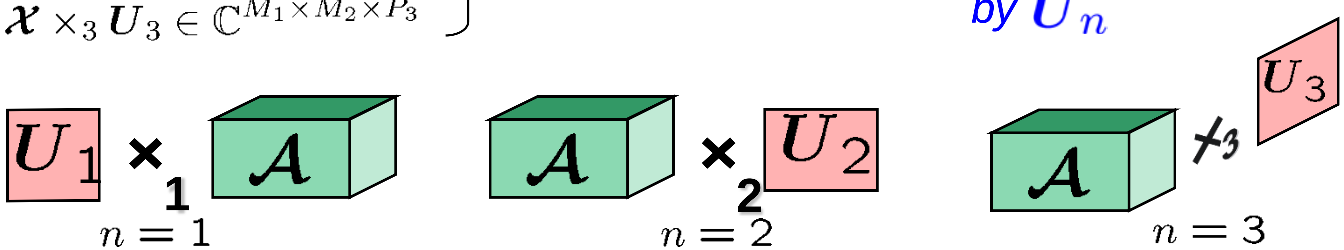
$\in \mathbb{C}^{M_2 \times M_3 \cdot M_1}$ *"2-mode vectors"*

$\in \mathbb{C}^{M_3 \times M_1 \cdot M_2}$ *"3-mode vectors"*

- *n*-mode products between $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{M_1 \times M_2 \times M_3}$ and $\boldsymbol{U}_n \in \mathbb{C}^{P_n \times M_n}$

$$\left.\begin{aligned} \boldsymbol{\mathcal{Y}} &= \boldsymbol{\mathcal{X}} \times_1 \boldsymbol{U}_1 \in \mathbb{C}^{P_1 \times M_2 \times M_3} \\ \boldsymbol{\mathcal{Y}} &= \boldsymbol{\mathcal{X}} \times_2 \boldsymbol{U}_2 \in \mathbb{C}^{M_1 \times P_2 \times M_3} \\ \boldsymbol{\mathcal{Y}} &= \boldsymbol{\mathcal{X}} \times_3 \boldsymbol{U}_3 \in \mathbb{C}^{M_1 \times M_2 \times P_3} \end{aligned}\right\} \Leftrightarrow [\boldsymbol{\mathcal{Y}}]_{(n)} = \boldsymbol{U}_n \cdot [\boldsymbol{\mathcal{X}}]_{(n)}$$

*i.e., all the n-mode vectors multiplied from the left-hand-side by $\boldsymbol{U}_n$*

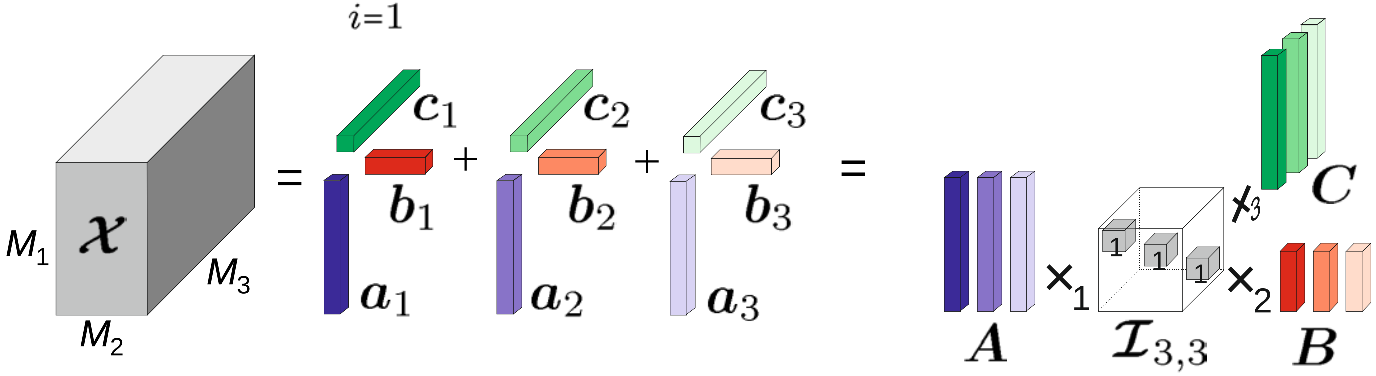


# Canonical Polyadic (CP) Decomposition

- Decomposes a given *R*-way array (tensor) into a sum of (the minimum number of) rank-one components
  - ⇒ **C**anonical **P**olyadic (**CP**) **decomposition**, also known as **Para**llel **Fac**tor (**PARAFAC**) analysis or **Can**onical **Decomp**osition (**CANDECOMP**)

$$\boldsymbol{\mathcal{X}} = \sum_{i=1}^{d} \boldsymbol{a}_i \circ \boldsymbol{b}_i \circ \boldsymbol{c}_i = \boldsymbol{\mathcal{I}}_{3,d} \times_1 \boldsymbol{A} \times_2 \boldsymbol{B} \times_3 \boldsymbol{C}$$

$d$: tensor rank$(\boldsymbol{\mathcal{X}})$




  - ⇒ applications in psychometrics, chemometrics, array signal processing, communications, biomedical signal processing, data mining, image/video processing, etc.

# $n$-mode unfoldings of a 4-way CP tensor

- 4-way CP: $\mathcal{X} = \mathcal{I}_{4,d} \times_1 \boldsymbol{F}^{(1)} \times_2 \boldsymbol{F}^{(2)} \times_3 \boldsymbol{F}^{(3)} \times_4 \boldsymbol{F}^{(4)} \in \mathbb{C}^{M_1 \times M_2 \times M_3 \times M_4}$

$\mathcal{X} \in \mathbb{C}^{M_1 \times M_2 \times M_3 \times M_4}$

$$[\mathcal{X}]_{(1)} \in \mathbb{C}^{M_1 \times M_2 \cdot M_3 \cdot M_4}$$

$$\boldsymbol{F}^{(1)} \cdot \left(\boldsymbol{F}^{(2)} \diamond \boldsymbol{F}^{(3)} \diamond \boldsymbol{F}^{(4)}\right)^{\mathrm{T}}$$

$\mathcal{X} \in \mathbb{C}^{M_1 \times M_2 \times M_3 \times M_4}$

$$[\mathcal{X}]_{(3)} \in \mathbb{C}^{M_3 \times M_4 \cdot M_1 \cdot M_2}$$

$$\boldsymbol{F}^{(3)} \cdot \left(\boldsymbol{F}^{(4)} \diamond \boldsymbol{F}^{(1)} \diamond \boldsymbol{F}^{(2)}\right)^{\mathrm{T}}$$

$\mathcal{X} \in \mathbb{C}^{M_1 \times M_2 \times M_3 \times M_4}$

$$[\mathcal{X}]_{(2)} \in \mathbb{C}^{M_2 \times M_3 \cdot M_4 \cdot M_1}$$

$$\boldsymbol{F}^{(2)} \cdot \left(\boldsymbol{F}^{(3)} \diamond \boldsymbol{F}^{(4)} \diamond \boldsymbol{F}^{(1)}\right)^{\mathrm{T}}$$

$\mathcal{X} \in \mathbb{C}^{M_1 \times M_2 \times M_3 \times M_4}$

$$[\mathcal{X}]_{(4)} \in \mathbb{C}^{M_4 \times M_1 \cdot M_2 \cdot M_3}$$

$$\boldsymbol{F}^{(4)} \cdot \left(\boldsymbol{F}^{(1)} \diamond \boldsymbol{F}^{(2)} \diamond \boldsymbol{F}^{(3)}\right)^{\mathrm{T}}$$

$\Rightarrow$ **one** index in the rows, **all others** in the columns

# “Generalized” unfoldings of a 4-way CP tensor

- 4-way CP: $\mathcal{X} = \mathcal{I}_{4,d} \times_1 \boldsymbol{F}^{(1)} \times_2 \boldsymbol{F}^{(2)} \times_3 \boldsymbol{F}^{(3)} \times_4 \boldsymbol{F}^{(4)} \in \mathbb{C}^{M_1 \times M_2 \times M_3 \times M_4}$

$\mathcal{X} \in \mathbb{C}^{M_1 \times M_2 \times M_3 \times M_4}$

$$\boxed{[\mathcal{X}]_{([1,2],[3,4])}} \in \mathbb{C}^{M_1 \cdot M_2 \times M_3 \cdot M_4}$$

$$\left(\boldsymbol{F}^{(2)} \diamond \boldsymbol{F}^{(1)}\right) \cdot \left(\boldsymbol{F}^{(4)} \diamond \boldsymbol{F}^{(3)}\right)^{\mathrm{T}}$$

$\mathcal{X} \in \mathbb{C}^{M_1 \times M_2 \times M_3 \times M_4}$

$$\boxed{[\mathcal{X}]_{([1,4],[2,3])}} \in \mathbb{C}^{M_1 \cdot M_4 \times M_2 \cdot M_3}$$

$$\left(\boldsymbol{F}^{(4)} \diamond \boldsymbol{F}^{(1)}\right) \cdot \left(\boldsymbol{F}^{(3)} \diamond \boldsymbol{F}^{(2)}\right)^{\mathrm{T}}$$

$\mathcal{X} \in \mathbb{C}^{M_1 \times M_2 \times M_3 \times M_4}$

$$\boxed{[\mathcal{X}]_{([1,3],[2,4])}} \in \mathbb{C}^{M_1 \cdot M_3 \times M_2 \cdot M_4}$$

$$\left(\boldsymbol{F}^{(3)} \diamond \boldsymbol{F}^{(1)}\right) \cdot \left(\boldsymbol{F}^{(4)} \diamond \boldsymbol{F}^{(2)}\right)^{\mathrm{T}}$$

(etc.)

- In general: $[\mathcal{X}]_{([\alpha_P,\ldots,\alpha_1],[\alpha_R,\ldots,\alpha_{P+1}])} = \left(\boldsymbol{F}^{(\alpha_1)} \diamond \ldots \diamond \boldsymbol{F}^{(\alpha_P)}\right) \cdot \left(\boldsymbol{F}^{(\alpha_{P+1})} \diamond \ldots \diamond \boldsymbol{F}^{(\alpha_R)}\right)^{\mathrm{T}}$

$\Rightarrow$ an arbitrary subset of indices in the rows, all other indices in the columns

[RSH12] F. Roemer, C. Schroeter, and M. Haardt, “A semi-algebraic framework for approximate CP decompositions via joint matrix diagonalization and generalized unfoldings,” *in Proc. of the 46th Asilomar Conference on Signals, Systems, and Computers*, pp. 2023–2027, November 2012.

# Outline

- Tensor Algebra and Notation
- Computation of Tensor Contractions via Generalized Unfoldings
- Double Contractions of Tensors to Represent Slice-wise Multiplications
  - ⇒ Element-wise Multiplication of Vectors and Matrices using Contraction
  - ⇒ Slice-wise Multiplication of Tensors using Contraction
- Constrained CP model of PARAFAC2 derived via Double Contractions
- Applications to Multi-Carrier MIMO Systems in Wireless Communications
  - ⇒ Khatri-Rao Coded MIMO OFDM Systems
  - ⇒ MIMO OFDM Systems
- Conclusions

# Motivation: Tensor Contraction (1)

- **Slice-wise multiplication of two tensors** is required in many signal processing applications and tensor decompositions, such as
  - ⇒ PARAFAC2 tensor decomposition


[H72] R. A. Harshman, PARAFAC2: "Mathematical and technical notes," UCLA Working Papers in Phonetics, vol. 22, pp. 30–47, 1972.

[K93] H. A. L. Kiers, "An alternating least squares algorithm for PARAFAC2 and three-way DEDICOM," Comput. Stat. Data. An., vol. 16, pp. 103–118, 1993.

[NCAH18] K. Naskovska, Y. Cheng, A. L. F. de Almeida, and M. Haardt, "Efficient computation of the PARAFAC2 decomposition via generalized tensor contractions," in *Proc. of 52nd Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, Oct. 2018.


  - ⇒ data analysis based on the PARAFAC2 tensor decomposition


[WJR+10] M. Weis, D. Jannek, F. Roemer, T. Guenther, M. Haardt, and P. Husar, "Multi-dimensional PARAFAC2 component analysis of multi-channel EEG data including temporal tracking," in *Proc. 32-th Inter. Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Sept. 2010.

[WJG+10] M. Weis, D. Jannek, T. Guenther, P. Husar, F. Roemer, and M. Haardt, "Temporally resolved multi-way component analysis of dynamic sources in event-related EEG data using PARAFAC2," in *Proc. 18-th European Signal Processing Conference (EUSIPCO 2010),* pp. 696-700, Aug. 2010.

# Motivation: Tensor Contraction (2)

⇒ PARATUCK2 tensor decomposition

| | |
|---|---|
| [HL96] | R. A. Harshman and M. E. Lundy, Uniqueness proof for a family of models sharing features of Tucker's three-mode factor analysis and PARAFAC/CANDECOMP, Psychometrika, vol. 61, pp. 133–154, 1996. |

⇒ PARATUCK2 is used to model space-time-frequency tensors for multi-carrier MIMO systems in wireless communication

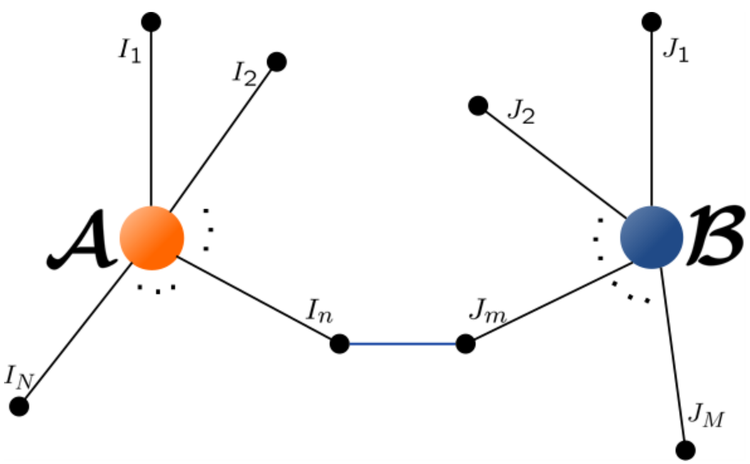
| | |
|---|---|
| [AFX13] | A. L. F. de Almeida, G. Favier, and L. R. Ximenes, "Space-time-frequency (STF) MIMO communication systems with blind receiver based on a generalized PARATUCK2 model," *IEEE Transactions on Signal Processing*, vol. 61, no. 8, pp. 1895–1909, 2013. |
| [AF13] | A. L. F. de Almeida and G. Favier, "Unified tensor model for space-frequency spreading-multiplexing (SFSM) MIMO communication systems," *EURASIP Journal on Advances in Signal Processing*, vol. 48, 2013 |
| [FA14] | G. Favier and A. L. F. de Almeida, "Tensor space-time-frequency coding with semi-blind receivers for MIMO wireless communication systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 22, pp. 5987– 6002, 2014. |
| [NCH+17] | K. Naskovska, S. A. Cheema, M. Haardt, B. Valeev, and Y. Evdokimov, "Iterative GFDM receiver based on the PARATUCK2 tensor decomposition," *in Proc. 21-st International ITG Workshop on Smart Antennas*, 2017. |

# Tensor Contraction

- $\mathcal{A} \bullet_n^m \mathcal{B}$ represents an **inner product** of the $n$-th mode of $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \ldots \times I_N}$ with the $m$-th mode of $\mathcal{B} \in \mathbb{C}^{J_1 \times J_2 \ldots \times J_M}$, provided $I_n = J_m$

$$(\mathcal{A} \bullet_n^m \mathcal{B})_{(i_1,\ldots,i_{n-1},i_{n+1},\ldots,i_N,j_1,\ldots,j_{m-1},j_{m+1},\ldots,j_M)} = \mathcal{T}_{(i_1,\ldots,i_{n-1},i_{n+1},\ldots,i_N,j_1,\ldots,j_{m-1},j_{m+1},\ldots,j_M)}$$

$$\mathcal{T}_{(i_1,\ldots,i_{n-1},i_{n+1},\ldots,i_N,j_1,\ldots,j_{m-1},j_{m+1},\ldots,j_M)} = \sum_{i=1}^{I_N} \mathcal{A}_{(i_1,\ldots,i_{n-1},i,i_{n+1},\ldots,i_N)} \cdot \mathcal{B}_{(j_1,\ldots,j_{m-1},i,j_{m+1},\ldots,j_M)}$$




$$\mathcal{T} \in \mathbb{C}^{I_1 \times \ldots \times I_{n-1} \times I_{n+1} \times \ldots \times I_N \times J_1 \times \ldots \times J_{m-1} \times J_{m+1} \times \ldots \times J_M}$$

- Let us now consider some simpler examples

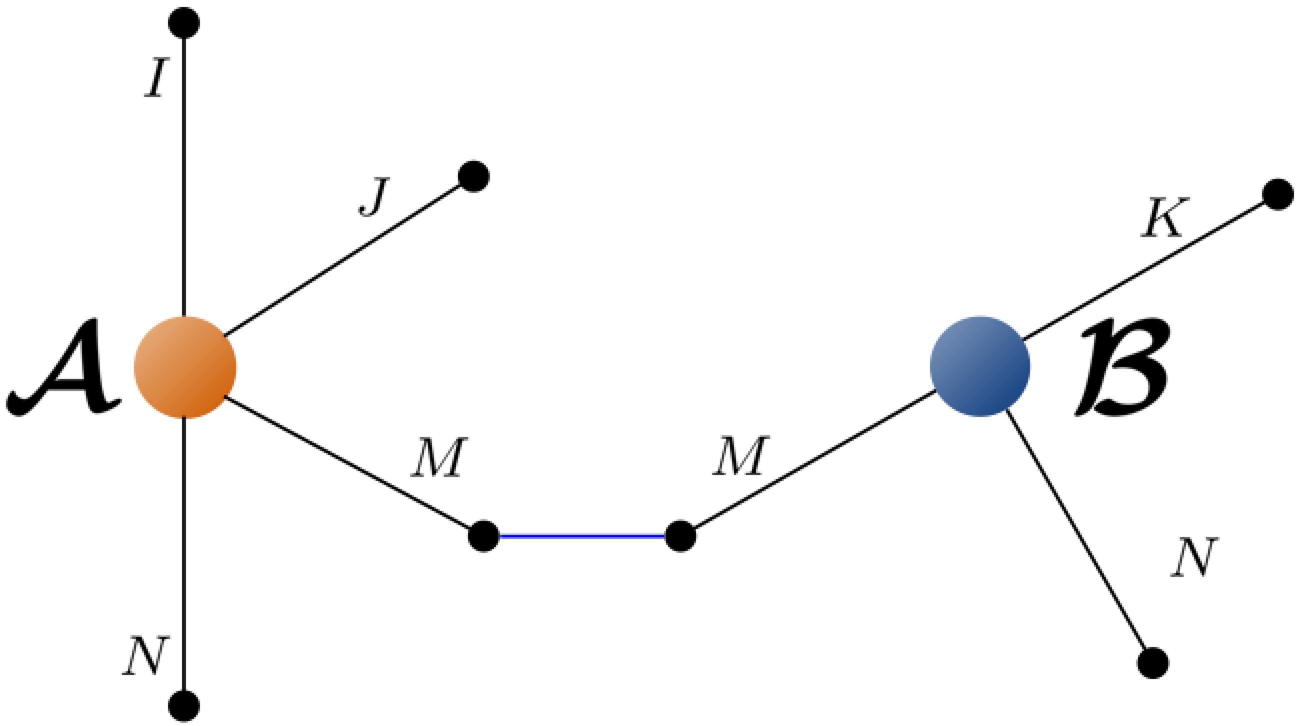

[C14] A. Cichocki, “Era of big data processing: A new approach via tensor networks and tensor decompositions,” *arXiv:1403.2048 [cs.ET]*, 2014.

[HRD08] M. Haardt, F. Roemer, and G. Del Galdo, “Higher-order SVD based subspace estimation to improve the parameter estimation accuracy in multi-dimensional harmonic retrieval problems,” *IEEE Transactions on Signal Processing*, vol. 56, pp. 3198–3213, 2008.

# "Generalized" unfoldings and Contraction

- Let us take $\mathcal{A} \in \mathbb{C}^{I \times J \times M \times N}$ and $\mathcal{B} \in \mathbb{C}^{M \times N \times K}$ as an example

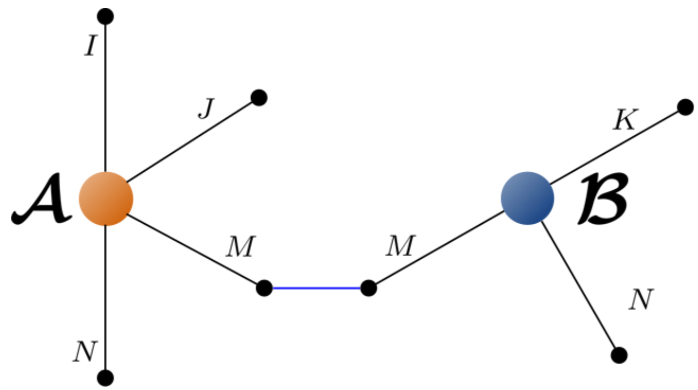
$$\Rightarrow \quad \mathcal{T} = \mathcal{A} \bullet_3^1 \mathcal{B} \in \mathbb{C}^{I \times J \times N \times N \times K}$$



# "Generalized" unfoldings and Contraction

- Let us take $\mathcal{A} \in \mathbb{C}^{I \times J \times M \times N}$ and $\mathcal{B} \in \mathbb{C}^{M \times N \times K}$ as an example

$$\Rightarrow \mathcal{T} = \mathcal{A} \bullet_3^1 \mathcal{B} \in \mathbb{C}^{I \times J \times N \times N \times K}$$




$$\mathcal{A} \in \mathbb{C}^{I \times J \times M \times N} \qquad \mathcal{B} \in \mathbb{C}^{M \times N \times K}$$

$$\boxed{[\mathcal{A}]_{([1,2,4],[3])}} \in \mathbb{C}^{IJN \times M} \qquad \boxed{[\mathcal{B}]_{([1],[2,3])}} \in \mathbb{C}^{M \times NK}$$

$$\boxed{[\mathcal{T}]_{([1,2,3],[4,5])} = [\mathcal{A}]_{([1,2,4],[3])} \cdot [\mathcal{B}]_{([1],[2,3])} \in \mathbb{C}^{IJN \times NK}}$$

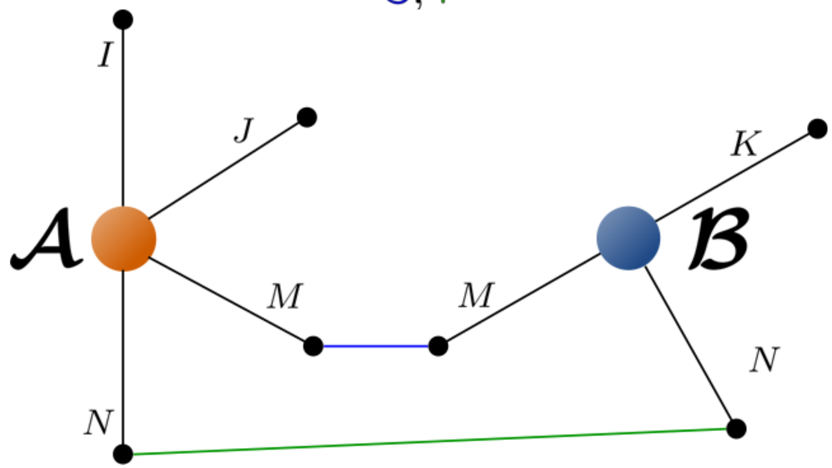
# "Generalized" unfoldings and Double Contraction

- Moreover, more than one mode can be contracted

$\mathcal{A} \in \mathbb{C}^{I \times J \times M \times N}$

$\mathcal{B} \in \mathbb{C}^{M \times N \times K}$

$\Rightarrow$ double contraction $\mathcal{T} = \mathcal{A} \bullet_{3,4}^{1,2} \mathcal{B} \in \mathbb{C}^{I \times J \times K}$




$\mathcal{A} \in \mathbb{C}^{I \times J \times M \times N}$

$\mathcal{B} \in \mathbb{C}^{M \times N \times K}$

$\boxed{[\mathcal{A}]_{([1,2],[3,4])}} \in \mathbb{C}^{IJ \times MN}$

$\boxed{[\mathcal{B}]_{([1,2],[3])}} \in \mathbb{C}^{MN \times K}$

$$\boxed{[\mathcal{T}]_{([1,2],[3])} = [\mathcal{A}]_{([1,2],[3,4])} \cdot [\mathcal{B}]_{([1,2],[3])} \in \mathbb{C}^{IJ \times K}}$$

# Matrix and *n*-mode Product using Contraction

- The matrix product $\boldsymbol{A} \cdot \boldsymbol{X}$ between a matrix $\boldsymbol{A} \in \mathbb{C}^{L \times L}$ and $\boldsymbol{X} \in \mathbb{C}^{L \times K}$ can be expressed as contraction.

$$\boldsymbol{A} \cdot \boldsymbol{X} \Leftrightarrow \boxed{\boldsymbol{A} \bullet_2^1 \boldsymbol{X}}$$

$\Rightarrow$ similar, $\boldsymbol{A}^\top \cdot \boldsymbol{X} \Leftrightarrow \boxed{\boldsymbol{A} \bullet_1^1 \boldsymbol{X}}$

- The $n$-mode product between a tensor $\mathcal{B} \in \mathbb{C}^{I_1 \times I_2 \times \ldots I_N}$ and a matrix $\boldsymbol{X} \in \mathbb{C}^{J \times I_n}$ can also be expressed as contraction.

$$\mathcal{T}_1 = \mathcal{B} \times_n \boldsymbol{X} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_{n-1} \times J \times I_{n+1} \times \ldots \times I_N}$$

$$\Updownarrow$$

$$\boxed{\mathcal{T}_2 = \mathcal{B} \bullet_n^2 \boldsymbol{X}} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_{n-1} \times I_{n-1} \times \ldots \times I_N \times J}$$

$\Rightarrow$ Both tensors contain the same values, only their dimensions are permuted.

# Outline

- Tensor Algebra and Notation
- Computation of Tensor Contractions via Generalized Unfoldings
- Double Contractions of Tensors to Represent Slice-wise Multiplications
  - $\Rightarrow$ Element-wise Multiplication of Vectors and Matrices using Contraction
  - $\Rightarrow$ Slice-wise Multiplication of Tensors using Contraction
- Constrained CP model of PARAFAC2 derived via Double Contractions
- Applications to Multi-Carrier MIMO Systems in Wireless Communications
  - $\Rightarrow$ Khatri-Rao Coded MIMO OFDM Systems
  - $\Rightarrow$ MIMO OFDM Systems
- Conclusions

# Element-wise Multiplication of Vectors using Contraction

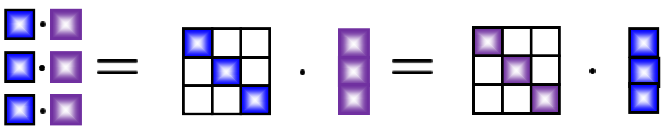
- Let us define two vectors $\boldsymbol{a} \in \mathbb{C}^{M \times 1}$ and $\boldsymbol{b} \in \mathbb{C}^{M \times 1}$

  $\Rightarrow$ the element-wise multiplication (Schur-Hadamard product)

  $$\boldsymbol{c} = \boldsymbol{a} \odot \boldsymbol{b} \quad \text{or} \quad c_{(m)} = a_{(m)} b_{(m)} \quad \forall m = 1, \ldots, M$$

  $\Rightarrow$ is equivalent to

  $$\boldsymbol{c} = \operatorname{diag}(\boldsymbol{a})\, \boldsymbol{b} = \operatorname{diag}(\boldsymbol{b})\, \boldsymbol{a}$$



  $\Rightarrow$ using the contraction operator

  $$\boxed{\boldsymbol{c} = \operatorname{diag}(\boldsymbol{a}) \bullet_2^1 \boldsymbol{b} = \operatorname{diag}(\boldsymbol{b}) \bullet_2^1 \boldsymbol{a}}$$

  $\Rightarrow$ where

  $$\operatorname{diag}(\boldsymbol{a}) = \boldsymbol{D}^{(\boldsymbol{a})} = \boldsymbol{I}_M \diamond \boldsymbol{a}^\top \qquad \text{(Khatri-Rao product)}$$

  $$\operatorname{diag}(\boldsymbol{b}) = \boldsymbol{D}^{(\boldsymbol{b})} = \boldsymbol{I}_M \diamond \boldsymbol{b}^\top$$

# Element-wise Multiplication of Matrices using Contraction (1)

- Next, let us assume two matrices $\boldsymbol{A} \in \mathbb{C}^{M \times N}$ and $\boldsymbol{B} \in \mathbb{C}^{M \times N}$
  - $\Rightarrow$ the element-wise multiplication (in two dimensions)

$$\boldsymbol{C} = \boldsymbol{A} \odot \boldsymbol{B} \quad \text{or} \quad C_{(m,n)} = A_{(m,n)} B_{(m,n)} \quad \forall m = 1, \ldots, M, n = 1, \ldots, N$$

  - $\Rightarrow$ using the contraction operator, it is equivalent to

$$\boxed{\boldsymbol{C} = \boldsymbol{\mathcal{D}}_A \bullet_{2,4}^{1,2} \boldsymbol{B} = \boldsymbol{\mathcal{D}}_B \bullet_{2,4}^{1,2} \boldsymbol{A}}$$

$$\mathcal{D}_{A(m,m,n,n)} = A_{(m,n)} \qquad \mathcal{D}_{B(m,m,n,n)} = B_{(m,n)}$$

$$\boldsymbol{\mathcal{D}}_A \in \mathbb{C}^{M \times M \times N \times N} \qquad \boldsymbol{\mathcal{D}}_B \in \mathbb{C}^{M \times M \times N \times N}$$

  - $\Rightarrow$ where

$$[\boldsymbol{\mathcal{D}}_A]_{([1,3],[2,4])} = \boldsymbol{I}_{MN} \diamond \text{vec}\,(\boldsymbol{A})^T \qquad \text{(diagonal matrix)}$$

$$[\boldsymbol{\mathcal{D}}_B]_{([1,3],[2,4])} = \boldsymbol{I}_{MN} \diamond \text{vec}\,(\boldsymbol{B})^T \qquad \text{(diagonal matrix)}$$

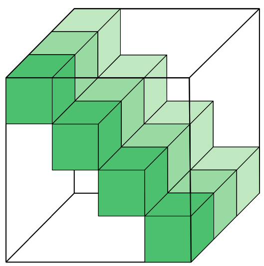
# Element-wise Multiplication of Matrices using Contraction (2)

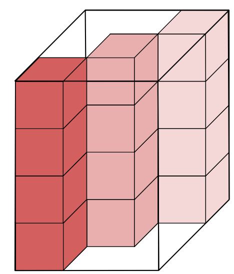
$\Rightarrow$ or, for the same two matrices $\boldsymbol{A} \in \mathbb{C}^{M\times N}$ and $\boldsymbol{B} \in \mathbb{C}^{M\times N}$

$$\boxed{\boldsymbol{C} = \boldsymbol{\mathcal{D}}^{(A)} \bullet_{2,3}^{1,3} \boldsymbol{\mathcal{D}}^{(B)}}$$

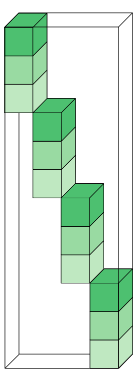
$$\boldsymbol{\mathcal{D}}^{(A)}{}_{(m,m,n)} = \boldsymbol{A}_{(m,n)} \qquad \boldsymbol{\mathcal{D}}^{(B)}{}_{(m,n,n)} = \boldsymbol{B}_{(m,n)}$$

$$\boldsymbol{\mathcal{D}}^{(A)} = \boldsymbol{\mathcal{I}}_{3,M} \times_3 \boldsymbol{A}^T \in \mathbb{C}^{M\times M\times N} \qquad \boldsymbol{\mathcal{D}}^{(B)} = \boldsymbol{\mathcal{I}}_{3,N} \times_1 \boldsymbol{B} \in \mathbb{C}^{M\times N\times N}$$

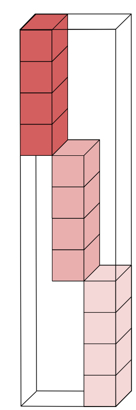




$\Rightarrow$ where



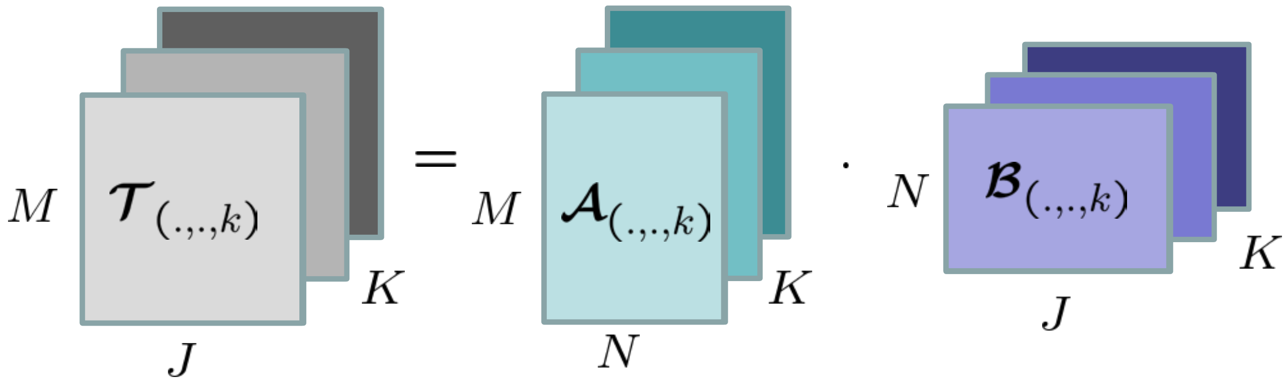
$$\left[\boldsymbol{\mathcal{D}}^{(A)}\right]_{([3,2],[1])} = \boldsymbol{I}_M \diamond \boldsymbol{A}^T \qquad \left[\boldsymbol{\mathcal{D}}^{(B)}\right]_{([1,3],[2])} = \boldsymbol{I}_N \diamond \boldsymbol{B}$$



# Slice-wise Multiplication using Contraction (1)

- A slice wise multiplication of two tensors $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{M \times N \times K}$ and $\boldsymbol{\mathcal{B}} \in \mathbb{C}^{N \times J \times K}$

$$\boldsymbol{\mathcal{T}}_{(.,.,k)} = \boldsymbol{\mathcal{A}}_{(.,.,k)} \boldsymbol{\mathcal{B}}_{(.,.,k)} \quad \forall k = 1, \ldots, K$$




- is also equal to

$$\boxed{\boldsymbol{\mathcal{T}} = \boldsymbol{\mathcal{A}} \bullet_{2,3}^{1,4} \boldsymbol{\mathcal{D}}_B \in \mathbb{C}^{M \times J \times K},}$$

$$[\boldsymbol{\mathcal{D}}_B]_{([1,2,4],[3])} = \boldsymbol{I}_K \diamond [\boldsymbol{\mathcal{B}}]_{([1,2],[3])} \qquad \boldsymbol{\mathcal{D}}_B \in \mathbb{C}^{N \times J \times K \times K}$$

# Slice-wise Multiplication using Contraction (2)

- A slice wise multiplication of two tensors $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{M \times N \times K}$ and $\boldsymbol{\mathcal{B}} \in \mathbb{C}^{N \times J \times K}$

$$\boldsymbol{\mathcal{T}}_{(.,.,k)} = \boldsymbol{\mathcal{A}}_{(.,.,k)} \boldsymbol{\mathcal{B}}_{(.,.,k)} \quad \forall k = 1, \ldots, K$$

- is also equal to

$$\boxed{\boldsymbol{\mathcal{T}} = \boldsymbol{\mathcal{A}} \bullet_{2,3}^{1,4} \boldsymbol{\mathcal{D}}_B \in \mathbb{C}^{M \times J \times K},}$$

$$[\boldsymbol{\mathcal{D}}_B]_{([1,2,4],[3])} = \boldsymbol{I}_K \diamond [\boldsymbol{\mathcal{B}}]_{([1,2],[3])} \qquad \boldsymbol{\mathcal{D}}_B \in \mathbb{C}^{N \times J \times K \times K}$$

- Further combination are also possible leading to the same result, such as

$\Rightarrow$ $\boxed{\boldsymbol{\mathcal{T}}_1 = \boldsymbol{\mathcal{D}}_B \bullet_{1,4}^{2,3} \boldsymbol{\mathcal{A}} \in \mathbb{C}^{J \times K \times M}}$

$\Rightarrow$ or, $\boxed{\boldsymbol{\mathcal{T}}_2 = \boldsymbol{\mathcal{D}}_A \bullet_{2,4}^{1,3} \boldsymbol{\mathcal{B}} \in \mathbb{C}^{M \times K \times J}}$ with $[\boldsymbol{\mathcal{D}}_A]_{([1,2,4],[3])} = \boldsymbol{I}_K \diamond [\boldsymbol{\mathcal{A}}]_{([1,2],[3])}$

$$\boldsymbol{\mathcal{D}}_A \in \mathbb{C}^{M \times N \times K \times K}$$

$\Rightarrow$ however, the ordering of the dimensions is permuted

# Outline

- Tensor Algebra and Notation
- Computation of Tensor Contractions via Generalized Unfoldings
- Double Contractions of Tensors to Represent Slice-wise Multiplications
  - ⇒ Element-wise Multiplication of Vectors and Matrices using Contraction
  - ⇒ Slice-wise Multiplication of Tensors using Contraction
- Constrained CP model of PARAFAC2 derived via Double Contractions
- Applications to Multi-Carrier MIMO Systems in Wireless Communications
  - ⇒ Khatri-Rao Coded MIMO OFDM Systems
  - ⇒ MIMO OFDM Systems
- Conclusions

# Computation of the PARAFAC2 decomposition

- PARAFAC2 was proposed in

| | |
|---|---|
| [H72] | R.A. Harshman, "PARAFAC2: Mathematical and technical notes," *CLA Working Papers in Phonetics* (University Microfilms, Ann Arbor, Michigan, No. 10,085), vol. *22*, pp. 30-44, 1972. |

- The computation is based on ALS (Alternating Least Squares)
  - $\Rightarrow$ **indirect fitting approach** (via PARATUCK2)

| | |
|---|---|
| [K93] | H. A. L. Kiers, "An alternating least squares algorithm for PARAFAC2 and three-way DEDICOM," in *Computational Statistics & Data Analysis* vol. 16, number 1, pp. 103–118, 1993. |

  - $\Rightarrow$ **direct fitting approach with two loops**
    - outer loop - Orthogonal Procrustes Problem (OPP)
    - inner loop - ALS

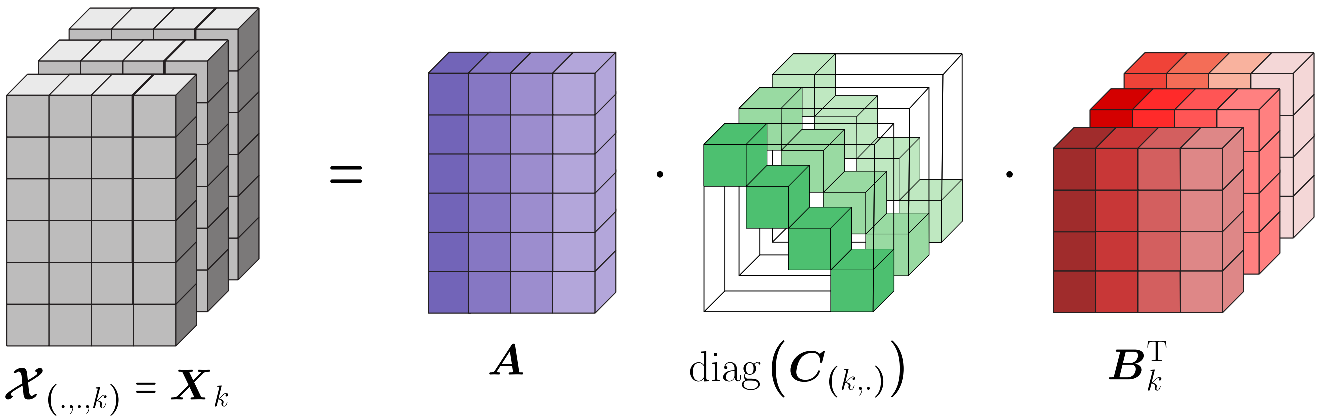
| | |
|---|---|
| [KBB99] | H. A. L. Kiers, J. M. F. T. Berge, and R. Bro, „PARAFAC2 - Part I. A direct fitting algorithm for the PARAFAC2 model," in *J. Chemometrics* vol.13, no. 3, issue 4, pp. 275–294, 1999. |

  - $\Rightarrow$ also another direct fitting approach **with two loops**
    - outer loop - Orthogonal Procrustes Problem (OPP)
    - inner loop – Simultaneous Matrix Diagonalization (SMD)

# PARAFAC2 decomposition

- The PARAFAC2 decomposition can be seen as coupled matrix decomposition, coupled in one mode

$$\boldsymbol{X}_k = \boldsymbol{A} \cdot \operatorname{diag}\left(\boldsymbol{C}_{(k,.)}\right) \boldsymbol{B}_k^{\top} \quad \forall k = 1, \ldots, K$$




**Harshman constraint:**

$$\boldsymbol{B}_k^{\top} = \boldsymbol{F}^{\top} \cdot \boldsymbol{V}_k$$

$$\boldsymbol{V}_k \cdot \boldsymbol{V}_k^{\top} = \boldsymbol{I}_R$$

$$\boldsymbol{B}_k^{\top} \cdot \boldsymbol{B}_k = \boldsymbol{F}^{\top} \cdot \boldsymbol{F}$$

- or, the slice-wise multiplication of tensors

$$\boldsymbol{\mathcal{X}}_{(.,.,k)} = \boldsymbol{A} \cdot \boldsymbol{\mathcal{C}}_{(.,.,k)} \cdot \boldsymbol{\mathcal{B}}_{(.,.,k)}$$

$\Rightarrow$ Slice-wise multiplication can be expressed in terms of contraction

$\Rightarrow$ Tensor structure enables simultaneous view of all dimensions

  - Efficient computation of the PARAFAC2 decomposition

# Slice-wise multiplication in terms of contraction

- $\mathcal{X}_{(.,.,k)} = \boldsymbol{A} \cdot \mathcal{C}_{(.,.,k)} \cdot \mathcal{B}_{(.,.,k)} \qquad (I \times J \times K)$

$\boldsymbol{A} \in \mathbb{R}^{I \times R}$

$\mathcal{C} = \mathcal{I}_{3,R} \times_3 \boldsymbol{C} \in \mathbb{R}^{R \times R \times K}$

$\mathcal{B} \in \mathbb{R}^{R \times J \times K}$

$\mathcal{V} \in \mathbb{R}^{R \times J \times K}$

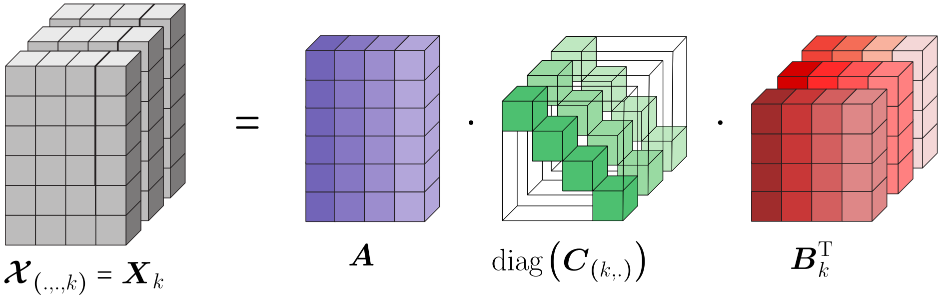
$\boldsymbol{F} \in \mathbb{R}^{R \times R}$

- Using the Harshman constraint, we have

$\Rightarrow \ \mathcal{B} = \mathcal{V} \times_1 \boldsymbol{F}^{\top}$

$\Rightarrow \ \mathcal{B} \bullet_{2,3}^{2,3} \mathcal{B} = K \boldsymbol{F}^{\top} \boldsymbol{F} \in \mathbb{C}^{R \times R}$

$\Rightarrow \ \mathcal{V} \bullet_{2,3}^{2,3} \mathcal{V} = K \boldsymbol{I}_R \in \mathbb{C}^{R \times R}$




- In terms of contraction

$\mathcal{X} = (\mathcal{D}_C \times_1 \boldsymbol{A}) \bullet_{2,4}^{1,3} \mathcal{B} \qquad (I \times K \times J)$ ***note***: permuted dimensions

$[\mathcal{D}_C]_{([1,2,4],[3])} = \boldsymbol{I}_K \diamond [\mathcal{C}]_{([1,2],[3])} \qquad \mathcal{D}_C \in \mathbb{R}^{R \times R \times K \times K}$

# Contraction based on generalized unfoldings (1)

- It can be shown that $\mathcal{D}_C \in \mathbb{R}^{R\times R\times K\times K}$

$$\mathcal{D}_C = \underbrace{\left(\mathcal{I}_{4,K} \otimes \mathcal{I}_{3,R}\right) \times_1 \left(\mathbf{1}_K^{\mathsf{T}} \otimes \boldsymbol{I}_R\right) \times_2 \left(\mathbf{1}_K^{\mathsf{T}} \otimes \boldsymbol{I}_R\right) \times_3 \left((\boldsymbol{I}_K \otimes \mathbf{1}_R^{\mathsf{T}}) \diamond \mathrm{vec}\left(\boldsymbol{C}^{\mathsf{T}}\right)^{\mathsf{T}}\right)}$$

$$\mathcal{X} = (\mathcal{D}_C \times_1 \boldsymbol{A}) \bullet_{2,4}^{1,3} \mathcal{B}$$

- Using the generalized unfoldings, we get

$$[\mathcal{X}]_{([1,2],3)} = [\mathcal{D}_C \times_1 \boldsymbol{A}]_{([1,3],[2,4])} \, [\mathcal{B}]_{([1,3],2)}$$

---

Kronecker product between two tensors $\mathcal{A} \in \mathbb{C}^{M\times N\times L}$ and $\mathcal{B} \in \mathbb{C}^{P\times Q\times R}$ yields

$\mathcal{K} = \mathcal{A} \otimes \mathcal{B} \in \mathbb{C}^{MP\times NQ\times LR}$

# Contraction based on generalized unfoldings (2)

- $[\mathcal{X}]_{([1,2],3)} = [\mathcal{D}_C \times_1 \boldsymbol{A}]_{([1,3],[2,4])} \, [\mathcal{B}]_{([1,3],2)}$ $\qquad \mathcal{B} = \mathcal{V} \times_1 \boldsymbol{F}^\top$

$$[\mathcal{X}]_{([1,2],3)} = \left(\left(\left(\boldsymbol{I}_K \otimes \mathbf{1}_R^\top\right) \diamond \mathrm{vec}\left(\boldsymbol{C}^\top\right)^\top\right) \otimes \boldsymbol{A}\left(\mathbf{1}_K^\top \otimes \boldsymbol{I}_R\right)\right) \cdot \boxed{\left[\mathcal{I}_{4,K} \otimes \mathcal{I}_{3,R}\right]_{([1,3],[2,4])} \cdot \left(\boldsymbol{I}_K \otimes \mathbf{1}_K^\top \otimes \boldsymbol{I}_R\right)^\top} \cdot \left(\boldsymbol{I}_K \otimes \boldsymbol{F}^\top\right) \cdot [\mathcal{V}]_{([1,3],2)}$$

It is a selection matrix that converts the Kronecker into a Khatri-Rao product

$$\boxed{\boldsymbol{I}_{RK} \diamond \boldsymbol{I}_{RK}}$$

- $$[\mathcal{X}]_{([1,2],3)} = \left(\left(\left(\boldsymbol{I}_K \otimes \mathbf{1}_R^\top\right) \diamond \mathrm{vec}\left(\boldsymbol{C}^\top\right)^\top\right) \diamond \boldsymbol{A}\left(\mathbf{1}_K^\top \otimes \boldsymbol{I}_R\right)\right) \cdot \left(\boldsymbol{I}_K \otimes \boldsymbol{F}^\top\right) \cdot [\mathcal{V}]_{([1,3],2)}$$

$\Rightarrow$ This is an unfolding of **constrained CP decomposition**.

# PARAFAC2 as a Constrained CP Decomposition

$$\mathcal{X} = \mathcal{I}_{3,RK} \times_1 \underbrace{\boldsymbol{A}\left(\mathbf{1}_K^\top \otimes \boldsymbol{I}_R\right)}_{\bar{\boldsymbol{A}} \in \mathbb{R}^{I \times RK}} \times_2 \underbrace{\left((\boldsymbol{I}_K \otimes \mathbf{1}_R^\top) \diamond \mathrm{vec}\left(\boldsymbol{C}^\top\right)^\top\right)}_{\bar{\boldsymbol{C}} \in \mathbb{R}^{K \times RK}} \times_3 \underbrace{\left(\left(\boldsymbol{I}_K \otimes \boldsymbol{F}^\top\right) \cdot [\mathcal{V}]_{([1,3],2)}\right)^\top}_{\bar{\boldsymbol{B}} \in \mathbb{R}^{J \times RK}}$$

$\Rightarrow$ where, $\bar{\boldsymbol{B}} = [\mathcal{B}]_{(2,[1,3])}$

- degenerate CP model in all three modes
- in general not unique
- but, the factor matrices are constrained

$\Rightarrow$ exploiting this matrix structure leads to better identifiability

[dAFM08] A. L. F. de Almeida, G. Favier, and J. C. M. Mota, “A Constrained Factor Decomposition with Application to MIMO Antenna Systems,” *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2429–2442, 2008.

# LS estimates of $\mathcal{V}$

- Assuming that $\boldsymbol{A}$, $\boldsymbol{C}$, $\boldsymbol{F}$, and known

$$[\mathcal{V}]_{([1,3],2)} = \cancel{\left((\bar{\boldsymbol{C}} \diamond \bar{\boldsymbol{A}}) \cdot (\boldsymbol{I}_K \otimes \boldsymbol{F}^\top)\right)^+} \cdot [\mathcal{X}]_{([1,2],3)}$$

  - does not take into account the orthogonality constraints
  - but, applicable if the Harshman constraint is not considered,

- alternatively

$$\cancel{\bar{\boldsymbol{B}} = [\mathcal{B}]_{(2,[1,3])} = [\mathcal{X}]_{(3,[1,2])}} \cdot \left((\bar{\boldsymbol{C}} \diamond \bar{\boldsymbol{A}})^\top\right)^+ \qquad \mathcal{B} = \mathcal{V} \times_1 \boldsymbol{F}^\top$$

$$\cancel{[\mathcal{B}]_{(1,[2,3])} = \boldsymbol{F}^\top \cdot [\mathcal{V}]_{(1,[2,3])}}$$

$\Rightarrow$ Orthogonal Procrustes Problem (OPP)

  - less accurate in the noisy case than solving directly the OPP

- $\boxed{\text{or solving directly the OPP on } \mathcal{X} = (\mathcal{D}_C \times_1 \boldsymbol{A}) \bullet_{2,4}^{1,3} (\mathcal{V} \times_1 \boldsymbol{F}^\top)}$


[S66] P. H. Schoenemann, “A generalized solution of the orthogonal Procrustes problem,” *Psychometrika*, vol. 31, no. 1, pp. 1–10, 1966.

[KBK99] H. A. L. Kiers, J. M. F. Ten Berge, and R. Bro, “PARAFAC2 — Part I. A direct fitting algorithm for the PARAFAC2 model,” *J. Chemometrics*, vol. 13, pp. 275–294, 1999.

# LS estimates of $A$, $F$, $\mathcal{B}$, and $C$

- Estimate $A$,

$$\boxed{A = [\mathcal{X}]_{(1,[2,3])} \cdot \left( (\mathbf{1}_K^\top \otimes I_R) \cdot (\bar{B} \diamond \bar{C})^\top \right)^+}$$

- utilizing the orthogonality of $\mathcal{V}$, estimate $F$,

$$\tilde{\mathcal{X}} = \mathcal{X} \bullet_{2,3}^{2,3} \mathcal{D}_{\mathcal{V}} = \mathcal{I}_{3,R} \times_1 A \times_2 F^\top \times_3 C$$

$$\left[\mathcal{D}_{\mathcal{V}}\right]_{([1,2,4],[3])} = I_K \diamond [\mathcal{V}]_{([1,2],[3])}$$

$$\boxed{F = \left[\tilde{\mathcal{X}}\right]_{(2,[1,3])} \cdot \left( (C \diamond A)^\top \right)^+}$$

- compute $\mathcal{B}$,

$$\boxed{\mathcal{B} = \mathcal{V} \times_1 F^\top}$$

- and estimate $C$

$$\boxed{\text{vec}\left(C^\top\right) = \left( \bar{B} \diamond (I_K \otimes \mathbf{1}_R^\top) \diamond \bar{A} \right)^+ \cdot \text{vec}\left(\mathcal{X}\right)}$$

Estimate $\mathcal{V}$ by solving the OPP

$$\mathcal{X} = (\mathcal{D}_C \times_1 A) \bullet_{2,4}^{1,3} (\mathcal{V} \times_1 F^\top)$$

**Initialization**

$A$ based on SVD

$C$ random

$F = I_R$

# ALS algorithm for PARAFAC2

- Based on these estimates we propose a **direct fitting ALS** algorithm that only has a **single loop**.

- **Initialization**
  - $\Rightarrow$ $\boldsymbol{A}$ based on SVD
  - $\Rightarrow$ $\boldsymbol{C}$ random
  - $\Rightarrow$ $\boldsymbol{F} = \boldsymbol{I}_R$

- The algorithm is stopped if it
  - $\Rightarrow$ exceeds the predefined maximum number of iterations: 2000
  - $\Rightarrow$ or reaches a predefined minimum of the cost function (reconstruction error): $10^{-7}$

# Simulation Results

- Synthetic data:

$$\mathcal{X} = \mathcal{X}_0 + \mathcal{N}$$

  $\Rightarrow$ random factors with elements drawn from Gaussian distribution

  $\Rightarrow$ $\mathcal{N}$: i.i.d. zero mean Gaussian noise (SNR = $\sigma_{\mathrm{n}}^{-2}$)

- Accuracy measure

  $\Rightarrow$ **S**quared **R**econstruction **E**rror (**SRE**)

$$\boxed{\mathrm{SRE} = \frac{\|\hat{\mathcal{X}} - \mathcal{X}_0\|_{\mathrm{H}}^2}{\|\mathcal{X}_0\|_{\mathrm{H}}^2}}$$

- Comparison

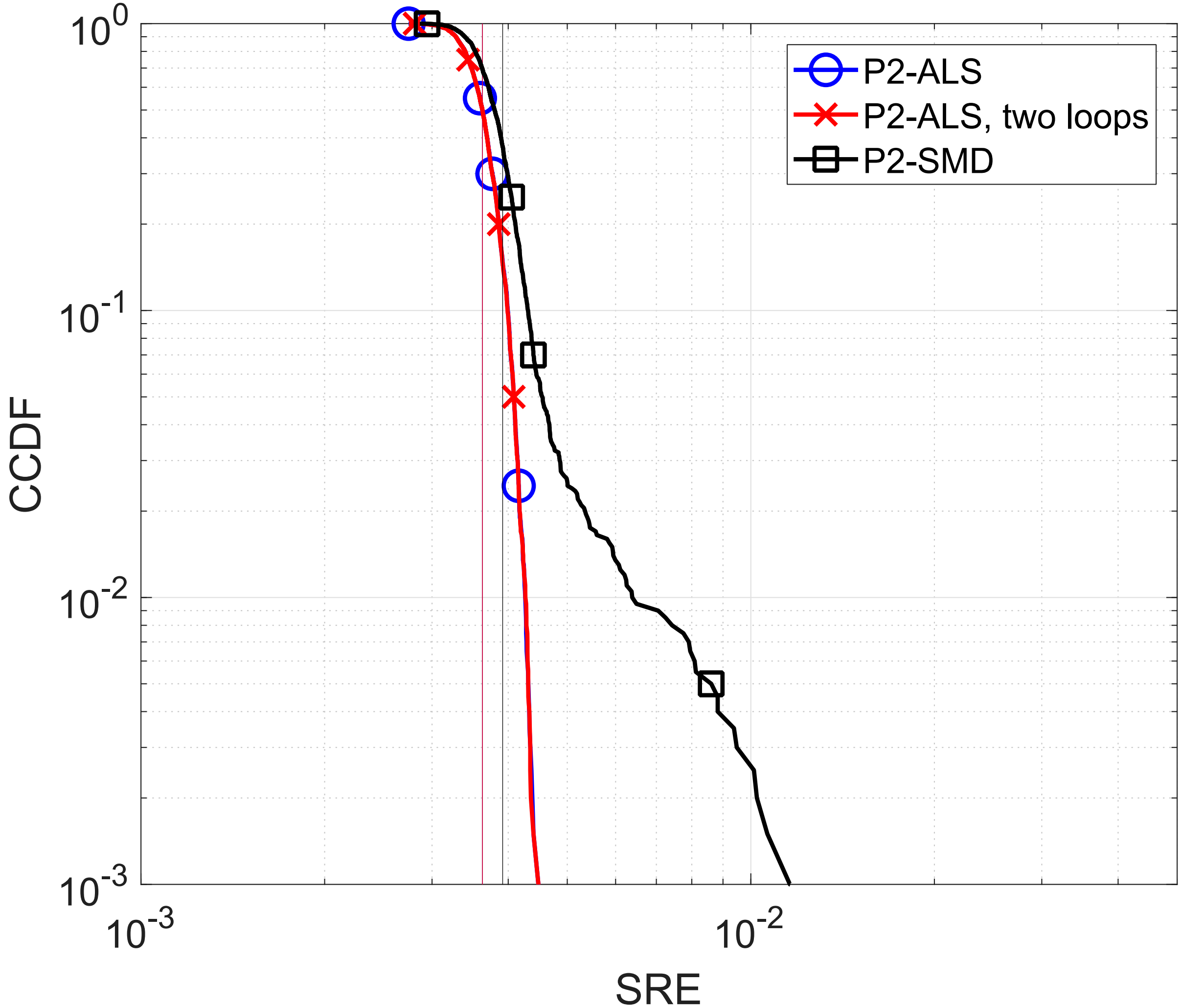
  $\Rightarrow$ **ALS** - proposed ALS algorithm with a single loop

  $\Rightarrow$ **2 ALS loops** - ALS algorithm, outer loop OPP, inner loop CP-ALS [KBK99]

  $\Rightarrow$ **ALS + SMD** - outer loop OPP, inner loop CP-SMD

[KBK99] H. A. L. Kiers, J. M. F. Ten Berge, and R. Bro, “PARAFAC2 — Part I. A direct fitting algorithm for the PARAFAC2 model”, J. Chemometrics,vol. 13, pp. 275–294, 1999.

# CCDF of SRE, real-valued tensor, 8 x 10 x 12




$8 \times 10 \times 12$

$R = 3$

$\mathrm{SNR} = 20$ dB

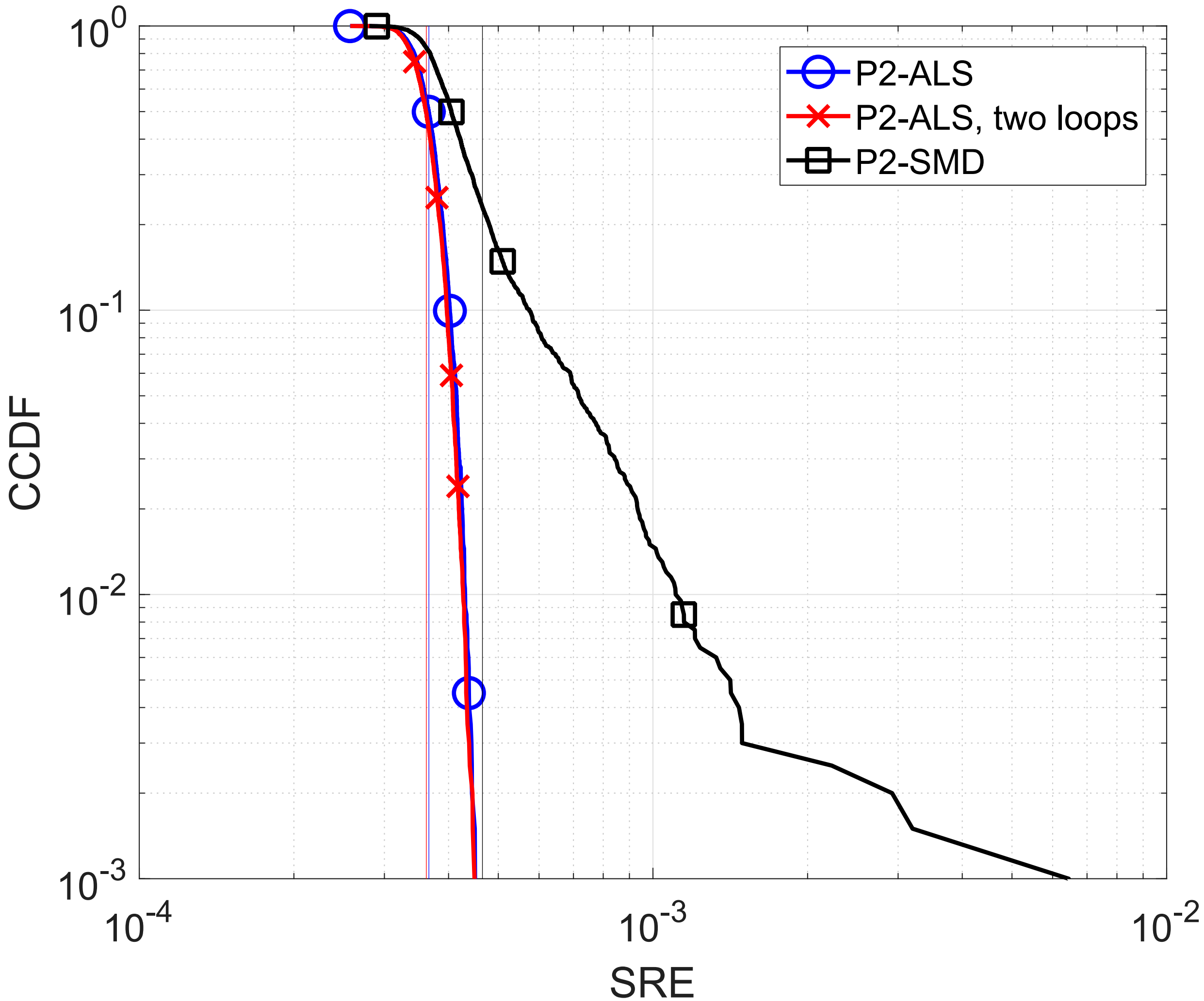
2000 realizations

Initialization based on SVD
All algorithms are initialized with the same matrices

**Stopping criteria**

- maximum number of iterations (2000)
- $10^{-7}$ minimum of the cost function
- **inner ALS loop:** maximum 5 iterations
- **inner SMD loop**: maximum 50 iterations

# CCDF of SRE, correlated real-valued tensor, 8 x 10 x 12




$8 \times 10 \times 12$

$R = 3$

$\mathrm{SNR} = 30$ dB

2000 realizations

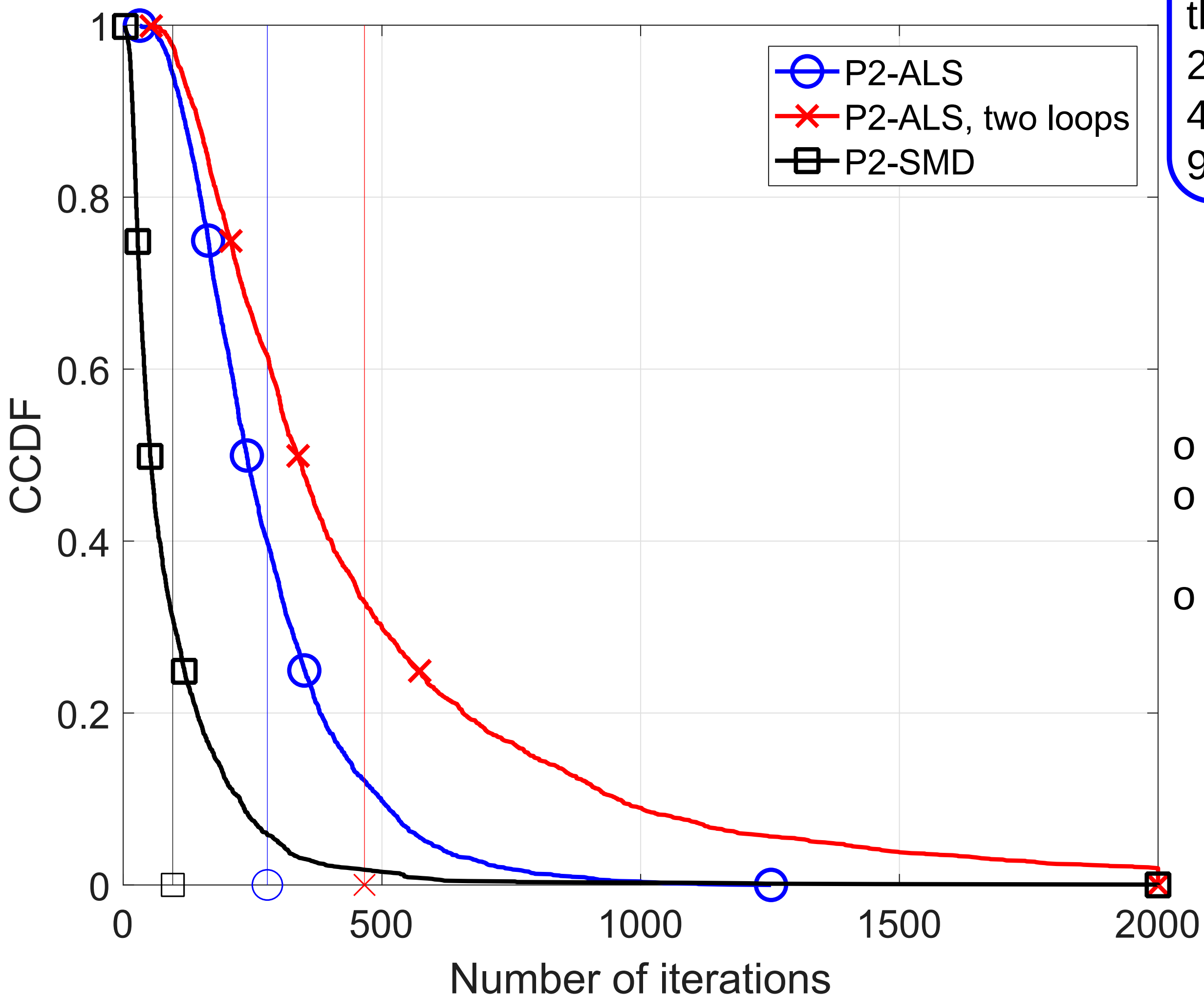
$\rho_C = 0.8$

Initialization based on SVD
All algorithms are initialized with the same matrices

**Stopping criteria**

- maximum number of iterations (2000)
- $10^{-7}$ minimum of the cost function
- **inner ALS loop:** maximum 5 iterations
- **inner SMD loop:** maximum 50 iterations

# Number of iterations, correlated real-valued tensor, 8 x 10 x 12




Average number over the 2000 realizations:
278
466 (outer loop only)
95 (outer loop only)

SNR = 30 dB

2000 realizations

$\rho_C = 0.8$

- Initialization based on SVD
- All algorithms are initialized with the same matrices
- **Stopping criteria**
  - maximum number of iterations (2000)
  - $10^{-7}$ minimum of the cost function
  - **inner ALS loop:** maximum 5 iterations
  - **inner SMD loop:** maximum 50 iterations

# Conclusions: PARAFAC2

- Using this new (general) approach we have shown
  - $\Rightarrow$ PARAFAC2 is equivalent to a **constrained, degenerate CP model**
  - $\Rightarrow$ leads to a direct fitting ALS algorithm
    - with only a **single loop**
    - resulting in less iterations as compared to state-of-the-art algorithms
    - **reduced computational complexity**

[NCAH18] K. Naskovska, Y. Cheng, A. L. F. de Almeida, and M. Haardt, “Efficient computation of the PARAFAC2 decomposition via generalized tensor contractions,” in *Proc. of 52nd Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, Oct. 2018.

# Outline

- Tensor Algebra and Notation
- Computation of Tensor Contractions via Generalized Unfoldings
- Double Contractions of Tensors to Represent Slice-wise Multiplications
  - ⇒ Element-wise Multiplication of Vectors and Matrices using Contraction
  - ⇒ Slice-wise Multiplication of Tensors using Contraction
- Constrained CP model of PARAFAC2 derived via Double Contractions
- Applications to Multi-Carrier MIMO Systems in Wireless Communications
  - ⇒ Khatri-Rao Coded MIMO OFDM Systems
  - ⇒ MIMO OFDM Systems
- Conclusions

# Introduction and Motivation: Khatri-Rao Coded MIMO OFDM Systems (1)

- Tensor-based signal processing has a very broad range of applications for **multi-dimensional data** such as
  - $\Rightarrow$ compressed sensing,
  - $\Rightarrow$ processing of multidimensional big data,
  - $\Rightarrow$ blind source separation using antenna arrays,
  - $\Rightarrow$ modeling communications systems.
- A multi-user communication system was modeled in terms of multilinear algebra for the design of a blind receiver in

| | |
|---|---|
| [SGB00] | N. D. Sidiropoulos, G. B. Giannakis, and R. Bro, "Blind PARAFAC receivers for DS-CDMA systems," *IEEE Transactions on Signal Processing*, vol. 48, no. 3, pp. 810–823, 2000. |

- An iterative receiver for MIMO-**GFDM** (**G**eneralized **F**requency **D**ivision **M**ultiplexing) systems was presented in

| | |
|---|---|
| [NCH+17] | K. Naskovska, S. A. Cheema, M. Haardt, B. Valeev, and Y. Evdokimov, "Iterative GFDM receiver based on the PARATUCK2 tensor decomposition," *in Proc. 21-st International ITG Workshop on Smart Antennas*, 2017. |

# Introduction and Motivation: Khatri-Rao Coded MIMO OFDM Systems (2)

- Space-time-frequency models for MIMO communication systems leading to semi-blind receivers were proposed in

[AFX13] A. L. F. de Almeida, G. Favier, and L. R. Ximenes, "Space-time-frequency (STF) MIMO communication systems with blind receiver based on a generalized PARATUCK2 model," *IEEE Transactions on Signal Processing*, vol. 61, no. 8, pp. 1895–1909, 2013.

[FA14] G. Favier and A. L. F. de Almeida, "Tensor space-time-frequency coding with semi-blind receivers for MIMO wireless communication systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 22, pp. 5987– 6002, 2014.

- We present a **tensor model** for MIMO OFDM systems based on a **double contraction** between a channel tensor and a signal tensor. This tensor model
  - ⇒ provides a **compact** and **flexible** formulation of the MIMO OFDM system model
  - ⇒ exploiting it at the receiver leads to a **tensor gain**
  - ⇒ exploits the channel **correlation** to reduce the number required pilot symbols (as compared to other tensor models)
  - ⇒ can be easily **extended** to any other multicarrier system, such as GFDM or FBMC

# System Model

- **MIMO-OFDM** communication system in the frequency domain
  - $\Rightarrow M_T$ transmit antennas
  - $\Rightarrow M_R$ receive antennas
  - $\Rightarrow N$ subcarriers
  - $\Rightarrow P$ frames, $P = K \cdot Q$, $K$ groups of $Q$ blocks
  - $\Rightarrow \tilde{\boldsymbol{\mathcal{X}}} \in \mathbb{C}^{N \times M_T \times K \times Q}$ is the transmitted signal tensor
  - $\Rightarrow \tilde{\boldsymbol{\mathcal{H}}} \in \mathbb{C}^{N \times N \times M_R \times M_T}$ is the channel tensor in the frequency domain
    - $\tilde{\boldsymbol{\mathcal{H}}}_{(.,.,m_R,m_T)} = \text{diag}\left(\boldsymbol{F}_L \cdot \boldsymbol{h}_L^{(m_R,m_T)}\right) = \text{diag}\left(\tilde{\boldsymbol{h}}_L^{(m_R,m_T)}\right)$
- $$\boxed{\tilde{\boldsymbol{\mathcal{Y}}} = \tilde{\boldsymbol{\mathcal{H}}} \bullet_{2,4}^{1,2} \tilde{\boldsymbol{\mathcal{X}}} + \tilde{\boldsymbol{\mathcal{N}}} = \tilde{\boldsymbol{\mathcal{Y}}}_0 + \tilde{\boldsymbol{\mathcal{N}}} \in \mathbb{C}^{N \times M_R \times K \times Q}}$$
  - $\Rightarrow$ compact and flexible formulation of the MIMO OFDM system
  - $\Rightarrow$ can be easily extended to any other multi-carrier system

---

$\boldsymbol{F}_L$ contains the first $L$ columns of a DFT matrix of size $N \times N$

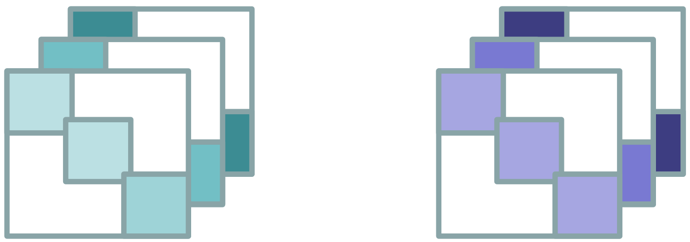
$L$ channel taps $\quad \boldsymbol{h}_L^{(m_R,m_T)}$ channel impulse response

[NHdA17] K. Naskovska, M. Haardt, and A. L. F. de Almeida, “Generalized tensor contraction with application to Khatri-Rao Coded MIMO OFDM systems,“ in *Proc. IEEE 7th Int. Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pp. 286 - 290, Dec. 2017.

# Channel Tensor (1)

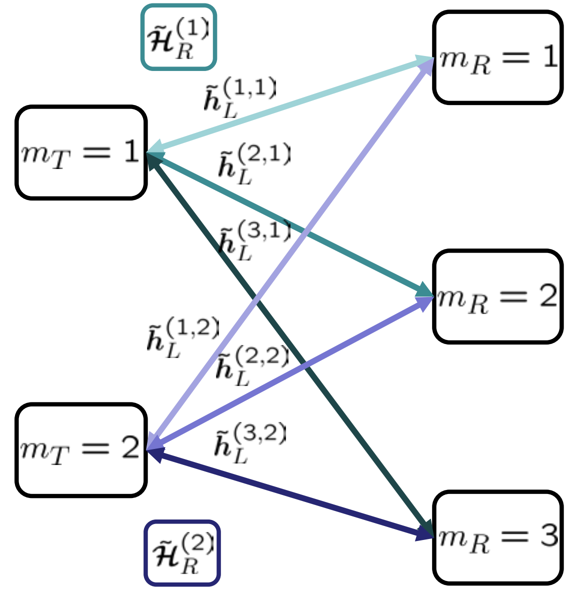
- Let us collect all channel coefficients for the $m_T$-th transmit and all receive antennas

$$\tilde{\boldsymbol{\mathcal{H}}}^{(m_T)}_{R(.,.,m_R)} = \text{diag}\left(\tilde{\boldsymbol{h}}_L^{(m_R,m_T)}\right) \qquad m_R = 1,\ldots,M_R$$



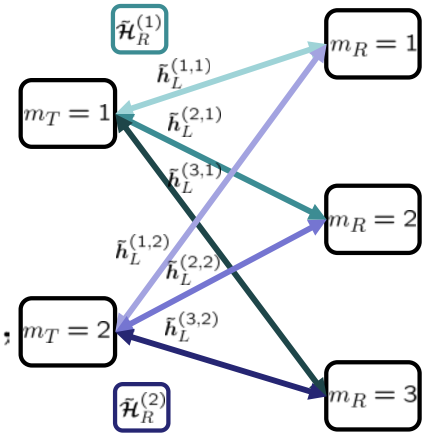
$$\tilde{\boldsymbol{H}}_R^{(m_T)} = \left[\tilde{\boldsymbol{h}}_L^{(1,m_T)} \quad \tilde{\boldsymbol{h}}_L^{(2,m_T)} \quad \ldots \quad \tilde{\boldsymbol{h}}_L^{(M_R,m_T)}\right]^T \in \mathbb{C}^{M_R \times N}$$

$$\boxed{\tilde{\boldsymbol{\mathcal{H}}}_R^{(m_T)} = \boldsymbol{\mathcal{I}}_{3,N} \times_1 \boldsymbol{I}_N \times_2 \boldsymbol{I}_N \times_3 \tilde{\boldsymbol{H}}_R^{(m_T)} \in \mathbb{C}^{N \times N \times M_R}}$$



# Channel Tensor (2)

- and now collecting all of the transmit antennas




$$\begin{aligned}
\tilde{\mathcal{H}} &= \left[\tilde{\mathcal{H}}_R^{(1)} \sqcup_4 \tilde{\mathcal{H}}_R^{(2)} \sqcup_4 \dots \tilde{\mathcal{H}}_R^{(M_T)}\right] \\
&= \left[\tilde{\mathcal{H}}_R^{(1)} \sqcup_4 \tilde{\mathcal{H}}_R^{(2)} \sqcup_4 \dots \tilde{\mathcal{H}}_R^{(M_T)}\right] \times_4 \boldsymbol{I}_{M_T} \\
&= \sum_{m_T=1}^{M_T} \tilde{\mathcal{H}}_R^{(m_T)} \circ \boldsymbol{e}_{m_T} = \sum_{m_T=1}^{M_T} \boldsymbol{\mathcal{D}} \times_1 \boldsymbol{I}_N \times_2 \boldsymbol{I}_N \times_3 \tilde{\boldsymbol{H}}_R^{(m_T)} \times_4 \boldsymbol{e}_{m_T};
\end{aligned}$$

$\Rightarrow$ where

- $\boldsymbol{e}_{m_T}$ is a pivoting vector of length $M_T$ containing all zeros and one at the $m_T$-th position
- $\boldsymbol{\mathcal{D}} \in \mathbb{R}^{N \times N \times N \times 1}$

  $\boldsymbol{\mathcal{D}} = \boldsymbol{\mathcal{I}}_{4,1} \otimes \boldsymbol{\mathcal{I}}_{3,N}$

---

Kronecker product between two tensors $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{M \times N \times L}$ and $\boldsymbol{\mathcal{B}} \in \mathbb{C}^{P \times Q \times R}$ yields

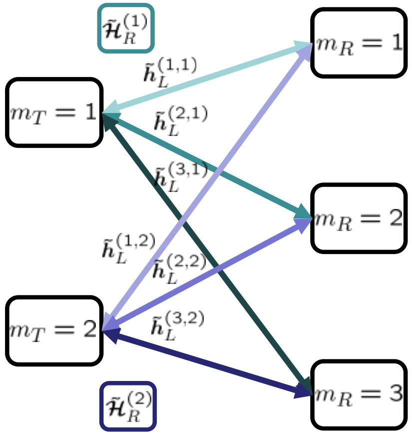
$\boldsymbol{\mathcal{K}} = \boldsymbol{\mathcal{A}} \otimes \boldsymbol{\mathcal{B}} \in \mathbb{C}^{MP \times NQ \times LR}$

# Channel Tensor (3)

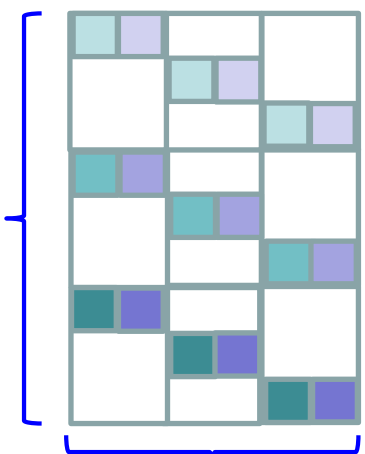
- The $([1,3],[4,2])$ generalized unfolding

$$\boxed{[\tilde{\mathcal{H}}]_{([1,3],[4,2])} = \bar{\boldsymbol{H}} \diamond (\boldsymbol{I}_N \otimes \mathbf{1}_{M_T}^T)}$$

$$\Rightarrow \quad \bar{\boldsymbol{H}} = \left[\tilde{\boldsymbol{H}}_R^{(1)} \quad \dots \quad \tilde{\boldsymbol{H}}_R^{(M_T)}\right] \cdot \boldsymbol{P} \in \mathbb{C}^{M_R \times M_T \cdot N}$$




$[\tilde{\mathcal{H}}]_{([1,3],[4,2])} =$ $N M_R$



$M_T N$

$\boldsymbol{e}_i \in \mathbb{R}^{M_T \times 1}$ $\quad \boldsymbol{\mathcal{D}}_{(.,.,.,1)} = \boldsymbol{\mathcal{I}}_{3,N} \in \mathbb{R}^{N \times N \times N \times 1}$ $\quad \tilde{\boldsymbol{\mathcal{H}}} \in \mathbb{C}^{N \times N \times M_R \times M_T}$

$\boldsymbol{P} \in \mathbb{R}^{M_T \cdot N \times M_T \cdot N}$ reorders the columns such that the faster increasing index is $M_T$ instead of $N$

# Data Transmission

- Data transmission using **K**hatri-**R**ao **s**pace-**t**ime (**KRST**) coding

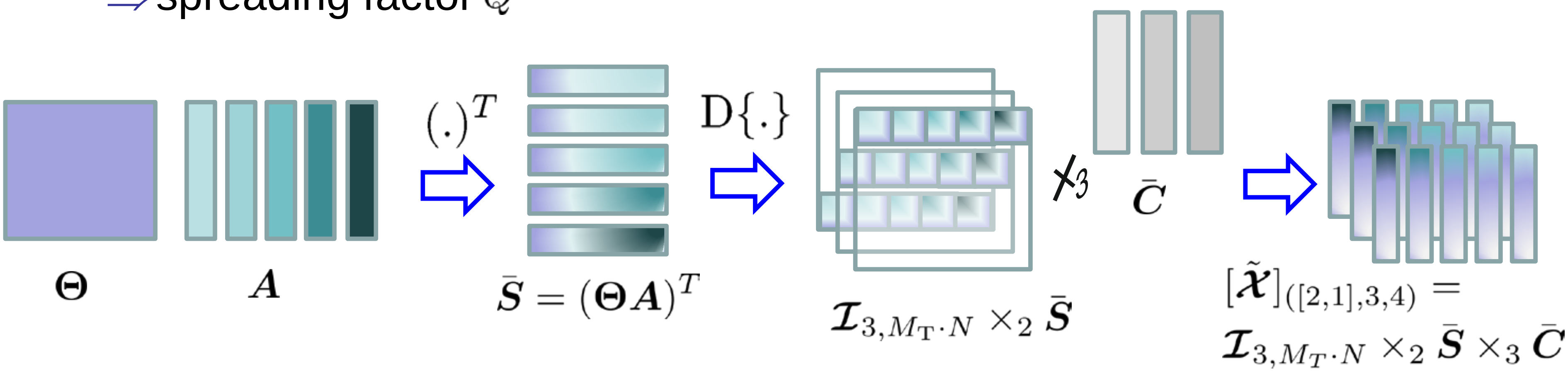
[SB02] N. D. Sidiropoulos and R. S. Budampati, “Khatri-Rao Space-Time codes,” *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2396–2407, 2002.

$\Rightarrow$ the $k$-th transmitted block:

$$\boxed{\tilde{\boldsymbol{X}}_k = \mathrm{D}_k\{\bar{\boldsymbol{S}}\} \cdot \bar{\boldsymbol{C}}^T \in \mathbb{C}^{M_\mathrm{T} N \times Q}, k = 1, \ldots, K}$$

$\mathrm{D}_k\{.\}$ converts the $k$-th row into diagonal matrix

$\Rightarrow$ spreading factor $Q$




| | |
|---|---|
| $\boldsymbol{\Theta} \in \mathbb{C}^{M_\mathrm{T}N \times M_\mathrm{T}N}$ | Unitary precoding matrix |
| $\boldsymbol{A} \in \mathbb{C}^{M_\mathrm{T}N \times K}$ | Symbol matrix |
| $\bar{\boldsymbol{S}} = \begin{bmatrix}\boldsymbol{S}_1 & \ldots & \boldsymbol{S}_N\end{bmatrix} \in \mathbb{C}^{K \times M_T \cdot N}$ | Overall data to be transmitted in $M_\mathrm{T}N$ blocks |
| $\bar{\boldsymbol{C}} = \begin{bmatrix}\boldsymbol{C}_1 & \ldots & \boldsymbol{C}_N\end{bmatrix} \in \mathbb{C}^{Q \times M_T \cdot N}$ | KRST coding matrix, spatial code rate $M_\mathrm{T}N/Q$ |

# Khatri-Rao coded signal tensor

- $\tilde{\boldsymbol{\mathcal{X}}} \in \mathbb{C}^{N \times M_T \times K \times Q}$ contains Khatri-Rao coded symbols
  - $\Rightarrow$ for each subcarrier $n = 1, \ldots, N$
  - $\Rightarrow$ the coding is proportional to the number of transmit antennas, $Q = M_{\mathrm{T}}$
  - $\Rightarrow$ the generalized unfolding of the signal tensor is

$$\boxed{\begin{aligned}[\tilde{\boldsymbol{\mathcal{X}}}]_{([2,1],[4,3])} &= \begin{bmatrix} \boldsymbol{S}_1 \diamond \boldsymbol{C}_1 & \boldsymbol{S}_2 \diamond \boldsymbol{C}_2 & \ldots & \boldsymbol{S}_N \diamond \boldsymbol{C}_N \end{bmatrix}^T \in \mathbb{C}^{M_T N \times QK} \\ &= \boldsymbol{I}_{M_T \cdot N} (\bar{\boldsymbol{S}} \diamond \bar{\boldsymbol{C}})^T\end{aligned}}$$

---

$$\boldsymbol{S}_n \in \mathbb{C}^{K \times M_T}$$

$$\boldsymbol{C}_n \in \mathbb{C}^{Q \times M_T}$$

$$\bar{\boldsymbol{S}} = \begin{bmatrix} \boldsymbol{S}_1 & \ldots & \boldsymbol{S}_N \end{bmatrix} \in \mathbb{C}^{K \times M_T \cdot N}$$

$$\bar{\boldsymbol{C}} = \begin{bmatrix} \boldsymbol{C}_1 & \ldots & \boldsymbol{C}_N \end{bmatrix} \in \mathbb{C}^{Q \times M_T \cdot N}$$

# Noiseless Received Signal

$$\tilde{\mathcal{Y}}_0 = \tilde{\mathcal{H}} \bullet_{2,4}^{1,2} \tilde{\mathcal{X}} \in \mathbb{C}^{N \times M_R \times K \times Q}$$

- can be expressed as

$$[\tilde{\mathcal{Y}}_0]_{([1,2],[4,3])} = [\tilde{\mathcal{H}}]_{([1,3],[4,2])} \cdot [\tilde{\mathcal{X}}]_{([2,1],[4,3])}$$

$\Rightarrow$ by inserting the corresponding unfoldings of the channel and the signal tensor, we get

$$[\tilde{\mathcal{Y}}_0]_{([1,2],[4,3])} = \left(\bar{\boldsymbol{H}} \diamond (\boldsymbol{I}_N \otimes \boldsymbol{1}_{M_T}^T)\right) \cdot (\bar{\boldsymbol{S}} \diamond \bar{\boldsymbol{C}})^T$$

"generalized" unfolding of a 4-way **constrained CP tensor**

- $$\boxed{\tilde{\mathcal{Y}}_0 = \mathcal{I}_{4,M_T \cdot N} \times_1 (\boldsymbol{I}_N \otimes \boldsymbol{1}_{M_T}^T) \times_2 \bar{\boldsymbol{H}} \times_3 \bar{\boldsymbol{S}} \times_4 \bar{\boldsymbol{C}}}$$

$\Rightarrow$ depending on the available a priori knowledge at the receiver,

- channel estimation, symbol estimation,
  or joint channel and symbol estimation can be performed
- channel correlation on adjacent subcarriers can be exploited

# Khatri-Rao Receiver

- From the tensor model, we get

$$[\tilde{\mathcal{Y}}]_{([1,4],[3,2])} \approx \left(\bar{\boldsymbol{C}} \diamond (\boldsymbol{I}_N \otimes \mathbf{1}_{M_T}^T)\right) \cdot (\bar{\boldsymbol{H}} \diamond \bar{\boldsymbol{S}})^T$$

$$\Rightarrow Q = M_T$$

$$\Rightarrow \left(\bar{\boldsymbol{C}} \diamond (\boldsymbol{I}_N \otimes \mathbf{1}_{M_T}^T)\right) \in \mathbb{C}^{N \cdot Q \times M_T \cdot N}$$

  - block diagonal
  - left invertible
  - known at the receiver

$$\Rightarrow \boldsymbol{C}_n^H \boldsymbol{C}_n = M_T \cdot \boldsymbol{I}_{M_T}$$

$$\boxed{\bar{\boldsymbol{Y}} = \frac{1}{M_T}\left(\bar{\boldsymbol{C}} \diamond (\boldsymbol{I}_N \otimes \mathbf{1}_{M_T}^T)\right)^H \cdot [\tilde{\mathcal{Y}}]_{([1,4],[3,2])} \approx (\bar{\boldsymbol{H}} \diamond \bar{\boldsymbol{S}})^T}$$

- Using the **Least Squares – Khatri-Rao Factorization** we can estimate

$$\hat{\bar{\boldsymbol{H}}} = \bar{\boldsymbol{H}} \cdot \boldsymbol{\Lambda},$$
$$\hat{\bar{\boldsymbol{S}}} = \bar{\boldsymbol{S}} \cdot \boldsymbol{\Lambda}^{-1}$$

$\boldsymbol{\Lambda} \in \mathbb{C}^{M_T \cdot N \times M_T \cdot N}$ is a diagonal scaling matrix

[CCH+17] Y. Cheng, S. A. Cheema, M. Haardt, A. Weiss, and A. Yeredor, "Performance analysis of least-squares Khatri-Rao factorization," in *Proc. IEEE 7th Int. Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Curacao, Dutch Antilles, pp. 447 - 486, Dec. 2017.

# Resolving the Scaling Ambiguity

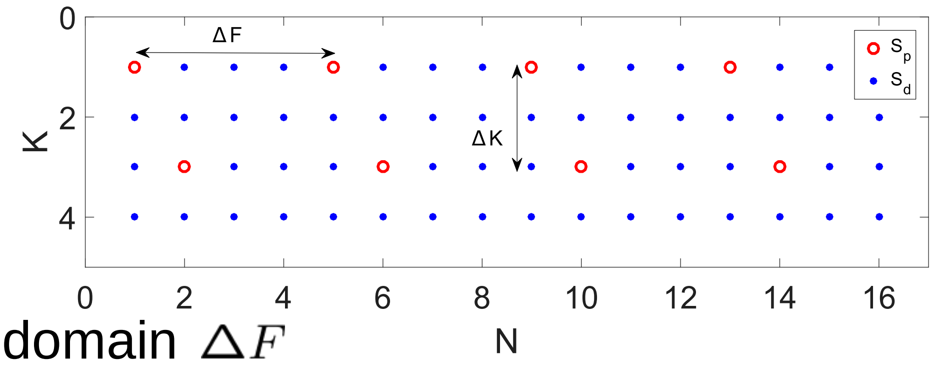
- $\boldsymbol{\Lambda} \in \mathbb{C}^{M_T \cdot N \times M_T \cdot N}$

  $\Rightarrow$ with the knowledge of one row of the matrix $\bar{\boldsymbol{S}}$

  - equivalent to $M_T N$ pilot symbols

  $\Rightarrow$ using less pilot symbols and applying channel interpolation techniques for OFDM

$$\bar{\boldsymbol{S}} = \underbrace{\bar{\boldsymbol{S}}_d}_{\text{data}} + \underbrace{\bar{\boldsymbol{S}}_p}_{\text{pilots}}$$




  - spacing in the frequency domain $\Delta F$
  - spacing in the time domain $\Delta K$
  - with the prior knowledge of the pilot symbols and their positions, we obtain a **pilot based** channel estimate $\hat{\boldsymbol{H}}_p$ by exploiting the channel **correlation**
  - $\hat{\boldsymbol{\Lambda}} = \text{diag}\left(\frac{1}{M_R}\sum_{i=1}^{M_R} \hat{\boldsymbol{H}}(i,.) \oslash \hat{\boldsymbol{H}}_p(i,.)\right)$

$\oslash$ element-wise division

# Enhancement via Least Squares

- After the scaling ambiguity that affects the columns of $\bar{\boldsymbol{H}}$ and $\bar{\boldsymbol{S}}$ is resolved we can **enhance** these estimates based on least squares
  - ⇒ using the decoded symbols $Q(\hat{\bar{\boldsymbol{S}}})$

$$\Rightarrow \hat{\bar{\boldsymbol{H}}}^T_{\mathrm{LS}} = \left( (\boldsymbol{I}_N \otimes \mathbf{1}^T_{M_T}) \diamond \bar{\boldsymbol{C}} \diamond Q(\hat{\bar{\boldsymbol{S}}}) \right)^+ \cdot [\tilde{\mathcal{Y}}]_{([2,4,1],[3])}$$

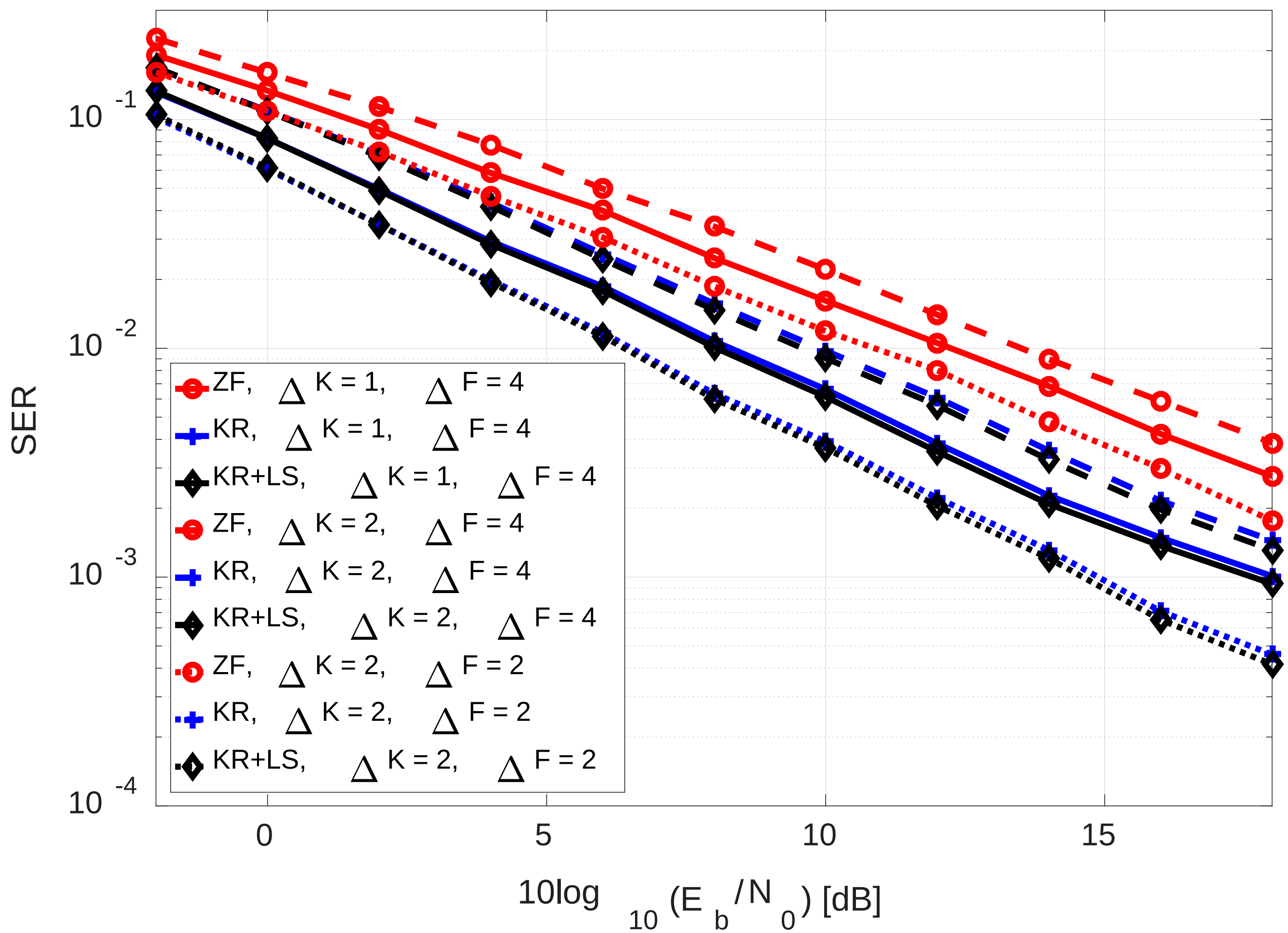
  - ⇒ improve the estimate of the diagonal scaling matrix

$$\hat{\boldsymbol{\Lambda}}_{\mathrm{LS}} = \mathrm{diag}\left( \frac{1}{M_R} \sum_{i=1}^{M_R} \hat{\bar{\boldsymbol{H}}}(i,.) \oslash \hat{\bar{\boldsymbol{H}}}_{\mathrm{LS}}(i,.) \right)$$

  - ⇒ improve the estimate of the symbols

$$\hat{\bar{\boldsymbol{S}}}_{\mathrm{LS}} = \hat{\bar{\boldsymbol{S}}} \cdot \hat{\boldsymbol{\Lambda}}_{\mathrm{LS}}$$

# Comparison of the SER for different spacings of the pilot positions




$N = 128$

$K = 2$

$Q = 2$

$M_T = 2$

$M_R = 2$

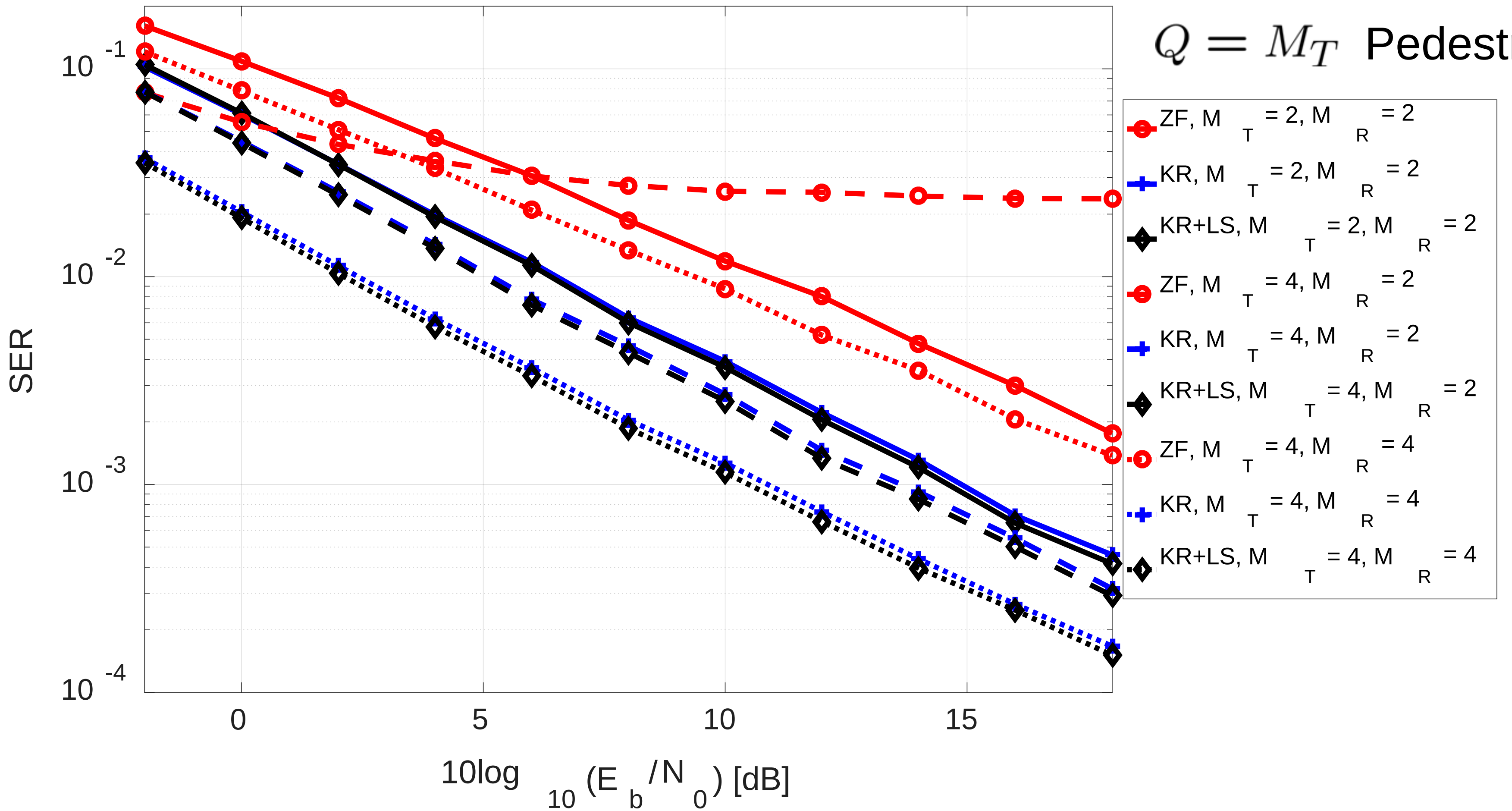
Pedestrian A

# Comparison of the SER for different numbers of transmit and receive antennas

$N = 128 \quad \Delta K = 2$

$K = 2 \quad \Delta F = 2$

$Q = M_T$ Pedestrian A



# Conclusions: Khatri-Rao Coded MIMO OFDM Systems

- Tensor model for MIMO OFDM systems based on a **double contraction** between a channel tensor and a signal tensor
  - $\Rightarrow$ provides a **compact** and **flexible** formulation of this Khatri-Rao coded MIMO OFDM system
  - $\Rightarrow$ can be easily **extended** to any other multi-carrier schemes, such as GFDM or FBMC
  - $\Rightarrow$ exploiting it at the receiver side leads to a **tensor gain**
- The proposed **Khatri-Rao receiver**
  - $\Rightarrow$ has an **improved** performance in terms of the SER
  - $\Rightarrow$ **exploits the channel correlation** between adjacent subcarriers
  - $\Rightarrow$ requires the same amount of training symbols as traditional OFDM techniques
  - $\Rightarrow$ can be extended to an iterative receiver if $Q \leq M_T$ for an improved performance
  - $\Rightarrow$ can be extended to a system with random coding matrices

# Outline

- Tensor Algebra and Notation
- Computation of Tensor Contractions via Generalized Unfoldings
- Double Contractions of Tensors to Represent Slice-wise Multiplications
  - ⇒ Element-wise Multiplication of Vectors and Matrices using Contraction
  - ⇒ Slice-wise Multiplication of Tensors using Contraction
- Constrained CP model of PARAFAC2 derived via Double Contractions
- Applications to Multi-Carrier MIMO Systems in Wireless Communications
  - ⇒ Khatri-Rao Coded MIMO OFDM Systems
  - ⇒ MIMO OFDM Systems
- Conclusions

# Introduction and Motivation: MIMO OFDM Systems

- We propose a **tensor model** for MIMO OFDM system
  - $\Rightarrow$ based on the **double contraction** between a channel tensor and a signal tensor
    - no additional spreading
- This tensor model
  - $\Rightarrow$ provides a **compact** and **flexible** formulation of the MIMO OFDM system
  - $\Rightarrow$ exploiting it at the receiver side leads to a **tensor gain**
  - $\Rightarrow$ exploits the channel **correlation** to reduce the number required pilot symbols (as compared to other tensor models)
  - $\Rightarrow$ can be easily **extended** to any other multi-carrier system, such as GFDM or FBMC

[NHdA18] K. Naskovska, M. Haardt, and A. L. F. de Almeida, “Generalized Tensor Contractions for an Improved Receiver design in MIMO-OFDM Systems,“ in *Proc. IEEE Int. Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Calgary, Alberta, Canada, pp. 3186 - 3190, Apr. 2018.

# System Model

- **MIMO-OFDM** communication system in the frequency domain
  - $\Rightarrow M_T$ transmit antennas
  - $\Rightarrow M_R$ receive antennas
  - $\Rightarrow N$ subcarriers
  - $\Rightarrow K$ frames
  - $\Rightarrow \tilde{\boldsymbol{\mathcal{S}}} \in \mathbb{C}^{N \times M_T \times K}$ is the signal tensor
  - $\Rightarrow \tilde{\boldsymbol{\mathcal{H}}} \in \mathbb{C}^{N \times N \times M_R \times M_T}$ is the channel tensor in the frequency domain
    - $\tilde{\boldsymbol{\mathcal{H}}}_{(.,.,m_R,m_T)} = \text{diag}\left(\boldsymbol{F}_L \cdot \boldsymbol{h}_L^{(m_R,m_T)}\right) = \text{diag}\left(\tilde{\boldsymbol{h}}_L^{(m_R,m_T)}\right)$
- $$\boxed{\tilde{\boldsymbol{\mathcal{Y}}} = \tilde{\boldsymbol{\mathcal{H}}} \bullet_{2,4}^{1,2} \tilde{\boldsymbol{\mathcal{S}}} + \tilde{\boldsymbol{\mathcal{N}}} = \tilde{\boldsymbol{\mathcal{Y}}}_0 + \tilde{\boldsymbol{\mathcal{N}}} \in \mathbb{C}^{N \times M_R \times K}}$$
  - $\Rightarrow$ compact and flexible formulation of the MIMO OFDM system
  - $\Rightarrow$ can be easily extended to any other multi-carrier system

---

$\boldsymbol{F}_L$ contains the first $L$ columns of DFT matrix of size $N \times N$

$L$ channel taps

$\boldsymbol{h}_L^{(m_R,m_T)}$ channel impulse response

# Data transmission

- The transpose of the 3-mode unfolding of uncoded signal tensor $\tilde{\boldsymbol{\mathcal{S}}} \in \mathbb{C}^{N \times M_T \times K}$ is

$$\bar{\boldsymbol{S}} = \tilde{\boldsymbol{\mathcal{S}}}^T_{([1,2],[3])} = \left[ \; \tilde{\boldsymbol{S}}^{(1)} \quad \tilde{\boldsymbol{S}}^{(2)} \quad \ldots \quad \tilde{\boldsymbol{S}}^{(M_T)} \; \right] \in \mathbb{C}^{K \times N \cdot M_T}$$

$\tilde{\boldsymbol{S}}^{(m_T)} \in \mathbb{C}^{K \times N}$ contains the symbols transmitted via the $m_T$-th antenna

- Moreover, we assume that the symbol matrix consists of data and pilot symbols

$$\bar{\boldsymbol{S}} = \underbrace{\bar{\boldsymbol{S}}_d}_{\text{data}} + \underbrace{\bar{\boldsymbol{S}}_p}_{\text{pilots}}$$

$\Rightarrow$ The matrix $\bar{\boldsymbol{S}}_d$ contains zeros at the positions of the pilot symbols.

$\Rightarrow$ The matrix $\bar{\boldsymbol{S}}_p$ contains zeros at the positions of the data symbols.

$\Rightarrow$ Pilots are sent **only** in the **first frame** with a subcarrier spacing of $\Delta F$ between two pilot symbols

# OFDM Receivers

- Received signal

$$\tilde{\mathcal{Y}} = \tilde{\mathcal{H}} \bullet_{2,4}^{1,2} \tilde{\mathcal{S}} + \tilde{\mathcal{N}} = \tilde{\mathcal{Y}}_0 + \tilde{\mathcal{N}} \in \mathbb{C}^{N \times M_R \times K}$$

$$[\tilde{\mathcal{Y}}]_{([1,2],[3])} = \underbrace{[\tilde{\mathcal{H}}]_{([1,3],[2,4])}}\underbrace{\tilde{\mathcal{S}}_{([1,2],[3])}} + [\tilde{\mathcal{N}}]_{([1,2],[3])} \in \mathbb{C}^{N \cdot M_R \times K}$$

$$[\tilde{\mathcal{Y}}]_{([1,2],[3])} = \boxed{\left(\bar{\boldsymbol{H}} \diamond (\mathbf{1}_{M_T}^T \otimes \boldsymbol{I}_N)\right)} \cdot \boxed{\bar{\boldsymbol{S}}^T} + [\tilde{\mathcal{N}}]_{([1,2],[3])}$$

$$\tilde{\mathcal{Y}} = \mathcal{I}_{3,N \cdot M_T} \times_1 (\mathbf{1}_{M_T}^T \otimes \boldsymbol{I}_N) \times_2 \bar{\boldsymbol{H}} \times_3 \bar{\boldsymbol{S}} + \tilde{\mathcal{N}}$$

$\Rightarrow$ **constrained CP model**

$\Rightarrow$ degenerate CP model in all three modes (more challenging)

- **Goal**

$\Rightarrow$ jointly estimate the channel and the symbols

# ZF Receiver

- With the prior knowledge of the pilot symbols and their positions, we obtain a **pilot based** channel estimate $\hat{\hat{\boldsymbol{H}}}_p(\hat{\hat{\boldsymbol{\mathcal{H}}}}_p)$ by exploiting the channel **correlation**

- **Zero Forcing** (ZF) Receiver

initialization $\tilde{\boldsymbol{\mathcal{H}}}_{\mathrm{p}}$
**for** $n = 1 : N$ **do**
$\quad \hat{\tilde{\boldsymbol{\mathcal{S}}}}_{(n,.,.)} \approx \hat{\tilde{\boldsymbol{\mathcal{H}}}}^{+}_{\mathrm{p}(n,n,.,.)} \tilde{\boldsymbol{\mathcal{Y}}}_{(n,.,.)}$
**end**

# Improved Receiver Design

- Alternatively, the channel and the symbols on each subcarrier can be estimated from

$$[\tilde{\mathcal{Y}}]^T_{([1],[2,3])} = \sum_{m_T=1}^{M_T} \left( \tilde{\boldsymbol{S}}^{(m_T)} \diamond \tilde{\boldsymbol{H}}_R^{(m_T)} \right) + [\tilde{\mathcal{N}}]^T_{([1],[3,2])} \in \mathbb{C}^{KM_R \times N}$$

$\Rightarrow$ This sum can be resolved by imposing orthogonality constraints

- Khatri-Rao Coding

[NHA17] K. Naskovska, M. Haardt, and A. L. F. de Almeida, "Generalized tensor contraction with application to Khatri-Rao coded MIMO OFDM systems," *in Proc. IEEE 7th Int. Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pp. 286 – 290, 2017.

$\Rightarrow$ or in column-wise fashion

$$\tilde{\boldsymbol{Y}}_n \approx \tilde{\boldsymbol{H}}_n \cdot \tilde{\boldsymbol{S}}_n \in \mathbb{C}^{M_R \times K}$$

$\tilde{\boldsymbol{y}}_n = \mathrm{vec}(\tilde{\boldsymbol{Y}}_n)$ where $\tilde{\boldsymbol{y}}_n$ denotes the $n$-th column of $[\tilde{\mathcal{Y}}]^T_{([1],[2,3])}$

$\tilde{\boldsymbol{H}}_n = \tilde{\mathcal{H}}_{(n,n,.,.)}$ $\quad \tilde{\boldsymbol{S}}_n = \tilde{\mathcal{S}}_{(n,.,.)}$

[TVP94] S. Talwar, M. Viberg, and A. Paulraj, "Blind estimation of multiple cochannel digital signals using an antenna array," *IEEE Signal Processing Letters*, vol. 1, no. 2, pp. 29–31, 1994.

[TVP96] ——, "Blind separation of synchronous co-channel digital signals using an antenna array. I. Algorithms," *IEEE Trans. Signal Process.*, vol. 44, no. 5, pp. 1184–1197, 1996.

# Iterative Least Squares with Projection (ILSP)

initialization $\tilde{\mathcal{H}}_{\mathrm{p}}$, maxIteration, minErr

**for** $n = 1 : N$ **do**

set $i = 1$, $e = \infty$

**while** $i < maxIteration$ *or* $e > minErr$ **do**

$$\bar{\tilde{\boldsymbol{S}}}_n^{(i)} = (\tilde{\boldsymbol{H}}_n^{(i-1)H} \tilde{\boldsymbol{H}}_n^{(i-1)})^{-1} \tilde{\boldsymbol{H}}_n^{(i-1)H} \tilde{\boldsymbol{Y}}_n$$

$$\tilde{\boldsymbol{S}}_n^{(i)} = \mathrm{proj}\left(\bar{\tilde{\boldsymbol{S}}}_n^{(i)}\right)$$

**if** rank $\left(\tilde{\boldsymbol{S}}_n^{(i)}\right) = M_T$ **then**

$$\tilde{\boldsymbol{H}}_n^{(i)} = \tilde{\boldsymbol{Y}}_n \tilde{\boldsymbol{S}}_n^{(i)H} (\tilde{\boldsymbol{S}}_n^{(i)H} \tilde{\boldsymbol{S}}_n^{(i)})^{-1}$$

**else**

$$\tilde{\boldsymbol{H}}_n^{(i)} = \tilde{\boldsymbol{H}}_n^{(i-1)}$$

**end**

$i = i + 1$, $e = \|\tilde{\boldsymbol{H}}_n^{(i-1)} - \tilde{\boldsymbol{H}}_n^{(i)}\|_{\mathrm{F}}^2$

**end**

**end**

for each subcarrier

in each iteration

Symbols are estimated via LS and projected onto the finite alphabet

If the estimated symbol matrix has rank $M_T$, an update of the channel estimate will be computed in a LS sense

# Iterative Least Squares with Enumeration (ILSE)

initialization $\tilde{\mathcal{H}}_{\mathrm{p}}$, maxIteration, minErr

**for** $n = 1 : N$ **do**

set $i = 1$, $e = \infty$

**while** $i < maxIteration \text{ or } e > minErr$ **do**

**for** $k = 1 : K$ **do**

$$\hat{\boldsymbol{s}} = \arg \min_{\boldsymbol{s}^{(j)} \in \Omega} \| \tilde{\boldsymbol{Y}}_{n(.,k)} - \tilde{\boldsymbol{H}}_n^{(i-1)} \boldsymbol{s}^{(j)} \|$$

$$j = 1, \ldots M^{M_T}$$

$$\tilde{\boldsymbol{S}}_{n(.,k)}^{(i)} = \hat{\boldsymbol{s}}$$

**end**

**if** rank $\left( \tilde{\boldsymbol{S}}_n^{(i)} \right) = M_T$ **then**

$$\tilde{\boldsymbol{H}}_n^{(i)} = \tilde{\boldsymbol{Y}}_n \tilde{\boldsymbol{S}}_n^{(i)H} (\tilde{\boldsymbol{S}}_n^{(i)H} \tilde{\boldsymbol{S}}_n^{(i)})^{-1}$$

**else**

$$\tilde{\boldsymbol{H}}_n^{(i)} = \tilde{\boldsymbol{H}}_n^{(i-1)}$$

**end**

$i = i + 1$, $e = \| \tilde{\boldsymbol{H}}_n^{(i-1)} - \tilde{\boldsymbol{H}}_n^{(i)} \|_{\mathrm{F}}^2$

**end**

**end**

for each subcarrier

in each iteration

Symbols are estimated based on enumeration (exhaustive search)

If the estimated symbol matrix has rank $M_T$, an update of the channel estimate will be computed in a LS sense

# Recursive Least Squares with Projection (RLSP)

initialization $\tilde{\mathcal{H}}_{\mathrm{p}}$, $0 \leq \alpha \leq 1$

**for** $n = 1 : N$ **do**

$$\bar{\tilde{\boldsymbol{S}}}_n = (\tilde{\boldsymbol{H}}_n^H \tilde{\boldsymbol{H}}_n)^{-1} \tilde{\boldsymbol{H}}_n^H \tilde{\boldsymbol{Y}}_n$$

$$\tilde{\boldsymbol{S}}_n = \mathrm{proj}\left(\bar{\tilde{\boldsymbol{S}}}_n\right)$$

set $\boldsymbol{P}^{(0)} = \boldsymbol{I}_{M_T}$, $\tilde{\boldsymbol{H}}_n^{(0)} = \tilde{\boldsymbol{H}}_n$

**for** $k = 1 : K$ **do**

$$\boldsymbol{s} = \tilde{\boldsymbol{S}}_{n(.,k)}$$

$$\tilde{\boldsymbol{H}}_n^{(k)} = \tilde{\boldsymbol{H}}_n^{(k-1)} + \frac{\left(\tilde{\boldsymbol{Y}}_{n(.,k)} - \tilde{\boldsymbol{H}}_n^{(k-1)} \boldsymbol{s}\right)}{\alpha + \boldsymbol{s}^H \boldsymbol{P}^{(k-1)} \boldsymbol{s}} \boldsymbol{s}^H \boldsymbol{P}^{(k-1)}$$

$$\boldsymbol{P}^{(k)} = \frac{1}{\alpha}\left(\boldsymbol{P}^{(k-1)} - \frac{\boldsymbol{P}^{(k-1)} \boldsymbol{s} \boldsymbol{s}^H \boldsymbol{P}^{(k-1)}}{\alpha + \boldsymbol{s}^H \boldsymbol{P}^{(k-1)} \boldsymbol{s}}\right)$$

**end**

**end**

for each subcarrier

Symbols are estimated via LS and projected onto the finite alphabet

Recursive least squares estimate of the channel

# Recursive Least Squares with Enumeration (RLSE)

initialization $\tilde{\mathcal{H}}_{\mathrm{p}},\ 0 \le \alpha \le 1$

**for** $n = 1 : N$ **do**

set $\boldsymbol{P}^{(0)} = \boldsymbol{I}_{M_T},\ \tilde{\boldsymbol{H}}_n^{(0)} = \tilde{\boldsymbol{H}}_n$

**for** $k = 1 : K$ **do**

$$\hat{\boldsymbol{s}} = \arg \min_{\boldsymbol{s}^{(j)} \in \Omega} \|\tilde{\boldsymbol{Y}}_{n(.,k)} - \tilde{\boldsymbol{H}}_n^{(k-1)} \boldsymbol{s}^{(j)}\|$$

$j = 1, \ldots M^{M_T}$

$\tilde{\boldsymbol{S}}_{n(.,k)} = \hat{\boldsymbol{s}}$

$$\tilde{\boldsymbol{H}}_n^{(k)} = \tilde{\boldsymbol{H}}_n^{(k-1)} + \frac{\left(\tilde{\boldsymbol{Y}}_{n(.,k)} - \tilde{\boldsymbol{H}}_n^{(k-1)} \hat{\boldsymbol{s}}\right)}{\alpha + \hat{\boldsymbol{s}}^H \boldsymbol{P}^{(k-1)} \hat{\boldsymbol{s}}} \hat{\boldsymbol{s}}^H \boldsymbol{P}^{(k-1)}$$

$$\boldsymbol{P}^{(k)} = \frac{1}{\alpha} \left(\boldsymbol{P}^{(k-1)} - \frac{\boldsymbol{P}^{(k-1)} \hat{\boldsymbol{s}} \hat{\boldsymbol{s}}^H \boldsymbol{P}^{(k-1)}}{\alpha + \hat{\boldsymbol{s}}^H \boldsymbol{P}^{(k-1)} \hat{\boldsymbol{s}}}\right)$$

**end**

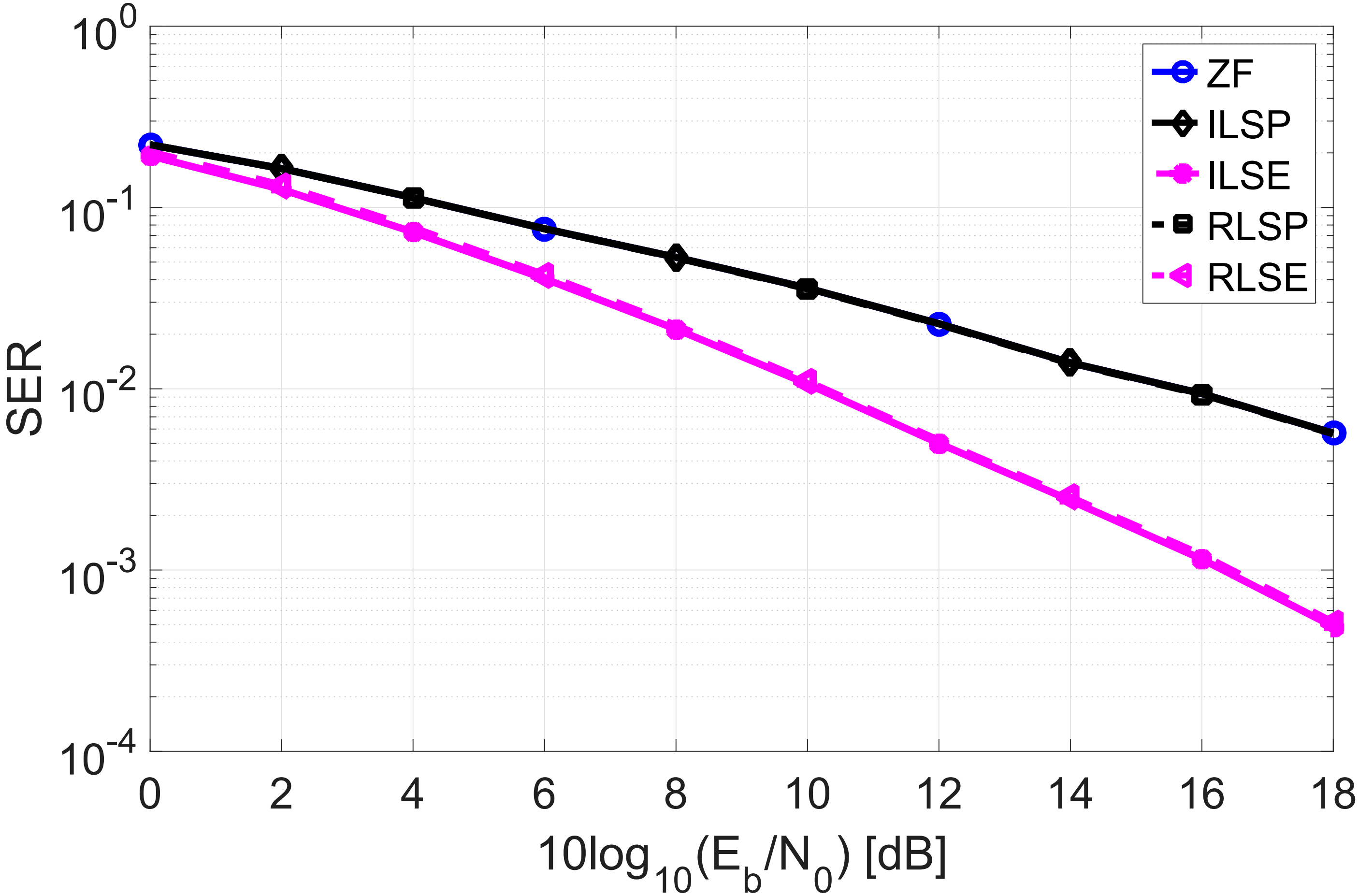
**end**

for each subcarrier

Symbols are estimated based on enumeration (exhaustive search)

Recursive least squares estimate of the channel

# ZF, ILSP, ILSE, RLSP, and RLSE for *K* = 2




$2 \times 2$

4QAM

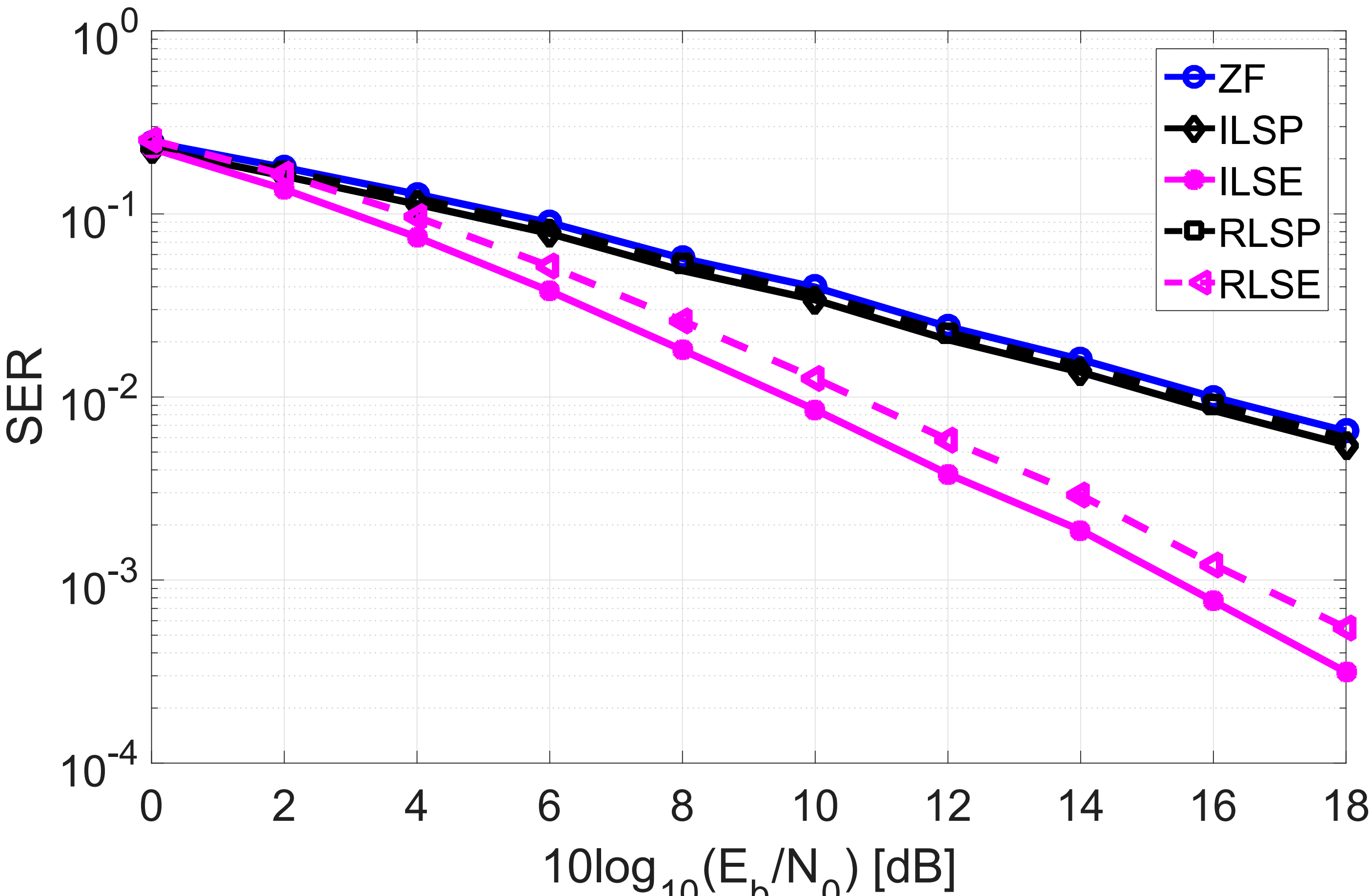
$K = 2$

$N = 128$

$\Delta F = 3$

Pedestrian A

# ZF, ILSP, ILSE, RLSP, and RLSE for $K$ = 8




$2 \times 2$

4QAM

$K = 8$

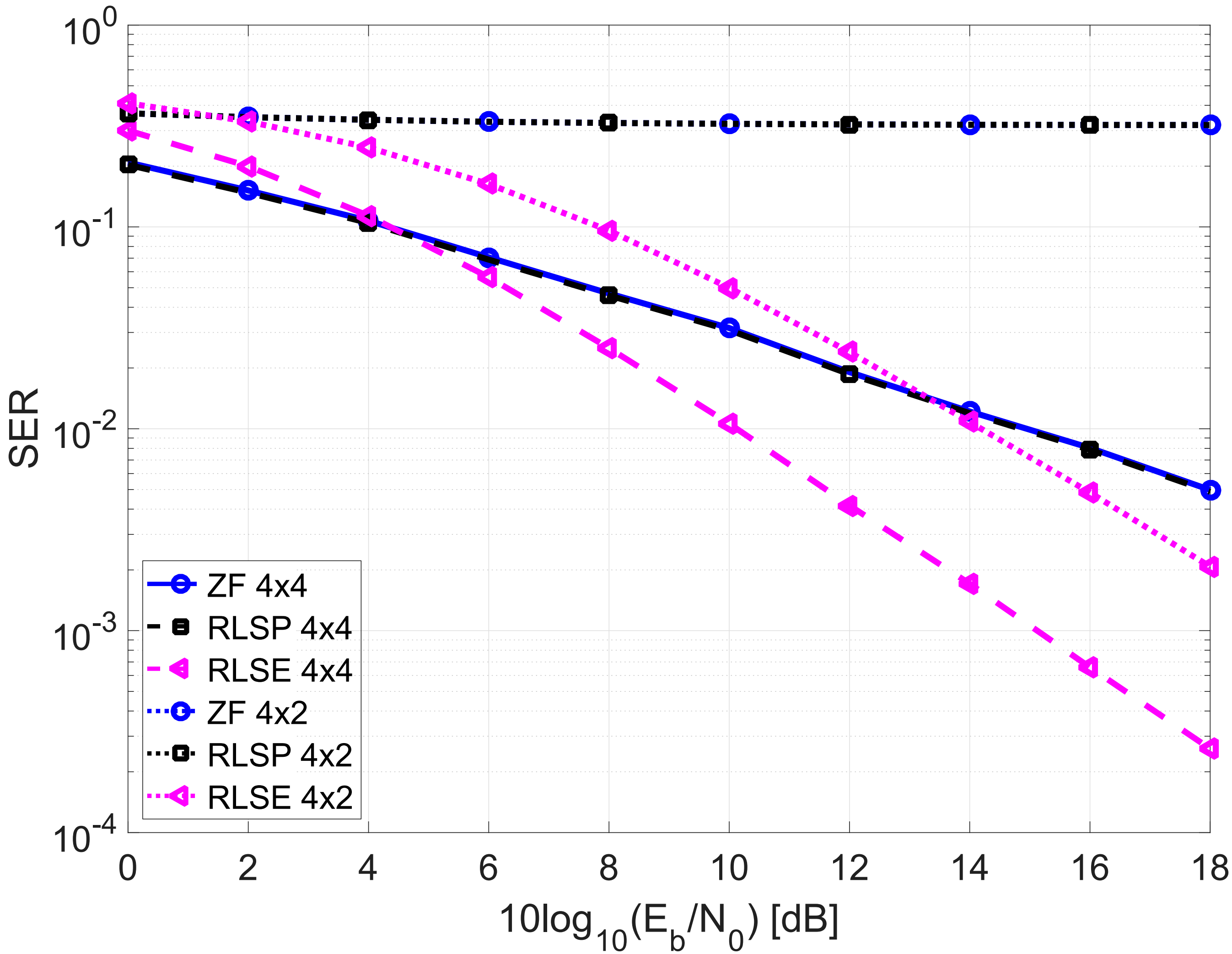
$N = 128$

$\Delta F = 3$

Pedestrian A

| Algorithm | ILSP | ILSE | RLSP | RLSE |
|---|---|---|---|---|
| Total Time [s] | 2.815 | 6.910 | 7.962 | 4.810 |

# Comparison of ZF, RLSP, and RLSE for Different Antenna Configurations




$M_R \times M_T$

4QAM

$K = 8$

$N = 512$

$\Delta F = 4$

Pedestrian A

# Conclusions: MIMO OFDM Systems (1)

- We present a tensor model for MIMO OFDM systems based on a **double contraction** between a channel tensor and a signal tensor
  - ⇒ provides a **compact** and **flexible** formulation of the MIMO OFDM system
  - ⇒ can be easily **extended** to other multi-carrier schemes, such as GFDM or FBMC
- This model facilitates the design of **several types of receivers** based on iterative LS and recursive LS (RLS) algorithms
  - ⇒ ILSP and RLSP show a similar performance as the ZF algorithm
  - ⇒ ISLE and RLSE based on enumeration, outperform the rest of the algorithms at the cost of increased complexity
  - ⇒ both recursive algorithms have less computational complexity
  - ⇒ the RLSE algorithm does not need matrix inversion
    - suitable for any configuration setup, even time-varying channels

# Conclusions: MIMO OFDM Systems (2)

- In the future,
  - $\Rightarrow$ recursive algorithms can be used to exploit the correlation of the channel tensor
  - $\Rightarrow$ multiple unfoldings can be exploited sequentially to capture the tensor structure in the receiver design
- Exploitation of the **sparse tensor structure** of **mmWave channels**
  - $\Rightarrow$ significant reduction of the pilot overhead (semi-blind channel and symbol estimation)
  - $\Rightarrow$ increase of the spectral efficiency
  - $\Rightarrow$ reduction of the latency
  - $\Rightarrow$ facilitates fast tracking of rapidly varying channels

# Conclusions

- Slice-wise multiplication of two tensors
  - $\Rightarrow$ is required in a variety of tensor decompositions
    - PARAFAC2, PARATUCK2, …
  - $\Rightarrow$ and is encountered in many applications
    - biomedical data (EEG, MEG, etc.)
    - multi-carrier MIMO systems
  - $\Rightarrow$ provide a new tensor representation
    - that is not based on a slice-wise (matrix) description
  - $\Rightarrow$ can be represented by a **double contraction**
    - efficiently calculated via generalized unfoldings
    - leads to new tensor models that do not depend on the chosen unfolding
    - **reveal the constrained CP tensor structure** of the data model
- Can be exploited to derive **improved receivers** and/or **improved** (blind or semi-blind) **model identification algorithms**

# THANK YOU!