TENSOR DECOMPOSITIONS WITH APPLICATIONS TO MIMO WIRELESS COMMUNICATION SYSTEMS

Gérard FAVIER

Laboratoire I3S
CNRS - Université de la Côte d'Azur
Sophia Antipolis

Ecole d'été en Traitement du Signal et des Images
Peyresq

June 2017

INTRODUCTION
FUNDAMENTAL PROBLEM IN SIGNAL PROCESSING (1)

Deconvolution (process to reverse the effects of convolution) and more generally signal estimation (recovery) from observed (received) signals

$$\mathbf{x} = \mathbb{H}[\mathbf{s}]$$

s: acoustic/seismic, sonar, radar, speech, biomedical (EEG, ECG)... signals.

$\mathbb{H}[.]$: propagation in the Earth, water, air, body...

Operator Linear/NL, Instantaneous (memoryless)/convolutive, SISO/MIMO (multiantenna system).

DECONVOLUTION IN DIGITAL COMMUNICATIONS (2)

$$\mathbf{x} = \mathbb{H}[\mathbf{s}] \Rightarrow \text{Ideally: } \hat{\mathbf{s}} = \mathbb{H}^{-1}[\mathbf{x}] \Rightarrow \text{Approximate solution.}$$

Brief history

- In the 1980s: channel equalization ⇒ adaptive (LMS/RLS) equalizers. (equalizer = device to compensate the distortion due to the communication channel)
- In the 1990s: blind deconvolution/equalization, blind source separation ⇒ High order statistics (HOS)-based methods.
- Since 2000: tensor approaches ⇒ deterministic joint semi-blind channel/symbols estimation based on multimodal/multidimensional representations of received signals.

MOBILE COMMUNICATIONS EVOLUTION (3)

- 2G systems $\Rightarrow$ SISO; GSM standard; since 1991 in Finland; 270 kbits/s.
- 3G systems $\Rightarrow$ SU-MIMO (single user); UMTS (Universal Mobile Telecom Service); EDGE (Enhanced Data Rates for GSM Evolution) standard; Internet access, video calls and mobile TV; since 2001 in Japon; $>$ 2 Mbits/s.
- 4G systems $\Rightarrow$ MU-MIMO (multi user); LTE (Long Term Evolution) standard; HD mobile TV, video conference, mobile web access...; since 2009 in Norway and Sweden; 1 Gbits/s.
- 5G systems $\Rightarrow$ Massive MIMO (very large number of antennas at the base station); from 2020; 100 Gbits/s.

A new generation of cellular standards approximately every ten years since 1G systems introduced in 1981. Each generation is characterized by new frequency bands, and higher data rates.

OBJECTIVES OF THE TALK

- To give an overview of tensor models/decompositions.
- To motivate and to illustrate the use of tensors for designing MIMO wireless communication systems.
- To present some tensor-based semi-blind receivers for joint channel/symbols estimation, in the case of point-to-point communication systems and of relaying systems.

1. Basics on MIMO wireless communication systems

2. Tensor models/decompositions
   - Background on tensors
   - Tensor models/decompositions

3. Tensor approaches for designing wireless communication systems
   - PARAFAC-CDMA system
   - PARAFAC-KRST coding system
   - PARATUCK-Tensor space-time (TST) coding system
   - Generalized PARATUCK-TSTF coding system
   - Tensor relaying communication systems
   - Nested PARAFAC two-hop relaying system
   - Tucker train two-hop relaying system

4. Conclusion and perspectives

PART 1 : MIMO WIRELESS COMMUNICATION SYSTEMS

- **1 Brief description**
- **2 Diversities and fundamental tradeoff**
- **3 Motivations for tensor modeling**

Brief description of MIMO communication systems (1)

MIMO communication systems studied since the 1990s

- Multiple antennas at the transmitter (T) and the receiver (R):

M transmit antennas; K receive antennas

$\Downarrow$

Multiple links between T and R $\Rightarrow$ Multipath-induced fading
(Random fluctuations in the received signal power)

$\Downarrow$

Propagation of information symbols through the channel $\mathbf{H} \in \mathbb{C}^{K \times M}$ or $\mathcal{H} \in \mathbb{C}^{K \times M \times F}$.

CDMA and OFDM systems (2)

CDMA and OFDM are wireless communication techniques widely used for fixed as well as for mobile applications.

CDMA (Code Division Multiple Access)
known as a spread-spectrum technique.
Used in the UMTS/3G (Universal Mobile Telecommunications System) and in GPS (Global Positioning System).
The information symbols are spread using a (spreading) pseudo-random code of length J, i.e. a sequence of chips, with values -1 and 1, or 0 and 1 $\Rightarrow$ J repetitions on chip.

OFDM (Orthogonal Frequency Division Multiplexing)
Used in ADSL/VDSL (Asymmetric/Very-high-bit rate Digital Subscriber Line), broadcast standards (DAB, DVB), and LTE/4G (Long Term Evolution) system.
Concept: multicarrier modulation technique with orthogonal subcarriers.

MIMO channels (3)

$$\mathbf{H} \in \mathbb{C}^{K \times M}$$

- $h_{k,m}$ = SISO channel gain between the kth receive antenna and the mth transmit antenna.
- $h_{k,m}$ modeled as a zero-mean circularly symmetric complex Gaussian random variable.
- Amplitude $|h_{k,m}|$ is Rayleigh distributed.

Two types of channels

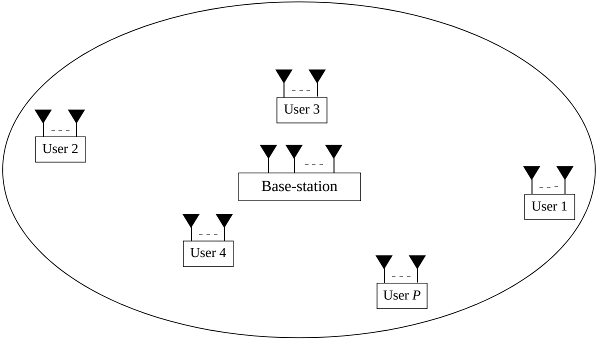
- Rich scattering $\Rightarrow$ i.i.d. frequency flat Rayleigh fading MIMO channel $\Leftrightarrow$ decorrelated channel coefficients (if adequate antenna spacing ($\geq \lambda/2$) to ensure decorrelation).
- Frequency-selective fading $\Rightarrow$ channel is frequency-dependent: $\mathcal{H} \in \mathbb{C}^{K \times M \times F}$ (increased bandwidth).

Wireless networks (4)

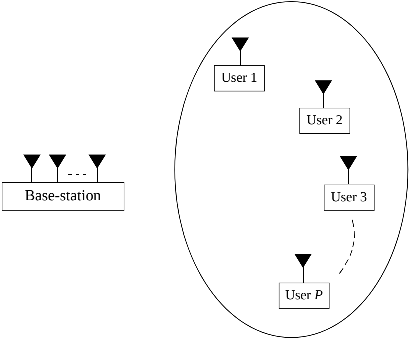
Wireless networks may be classified as cellular or ad hoc networks

- A cellular network is characterized by centralized communications: multiple users within a cell communicate with a base station that controls all transmissions and forwards data to the users $\Rightarrow$ Point-to-point communication systems.
  (Uplink/downlink: terminal$\rightarrow$ base station / station$\rightarrow$ terminal.)
- In an ad hoc network, any user can act as a sender or receiver of data, or as a relay for other users $\Rightarrow$ Relaying/Cooperative systems $\Rightarrow$ Distributed MIMO systems: multiple users cooperate to form a virtual antenna array
  $\Rightarrow$ Cooperative diversity: MIMO space diversity with single-antenna terminals (users).

Cellular systems (5)



Cooperative systems (6)



Relaying protocols (7)

Two main relaying protocols:

- Amplify-and-forward (AF) protocol: the relay amplifies/re-encodes the noisy received signals (without decoding) before forwarding.
- Decode-and-forward (DF) protocol: the relay decodes the received signals, and re-encode information signals before forwarding them.

Advantages/Drawbacks:

- AF: Simpler/Less efficient (because of noise propagation) at destination
- DF: More complex (because of decoding)/More efficient at destination

Fundamental tradeoff (8)

MIMO can be exploited to:

- Increase the rate of data transmission (transmission rate) through multiplexing.
- Improve system performance and reliability owing to space diversity.

$\Downarrow$

Fundamental tradeoff between multiplexing and diversity (i.e. transmission rate/performance).

Modulations and transmission rate (9)

Transmitted information symbols as sequences of bits depending on the used modulation

In practice, the emitter transmits data streams, each one being composed of N symbols $\Rightarrow$ Symbol matrix $\mathbf{S} \in \mathbb{C}^{N \times R}$.

Two main modulations (constellation/finite alphabet)

Quadrature Amplitude Modulation (QAM)

M-QAM finite alphabet of cardinality $M = 2^q \Rightarrow q$ bits/symbol.

16-QAM $\Rightarrow 2^4$ symbols $\Rightarrow$ 4 bits/symbol

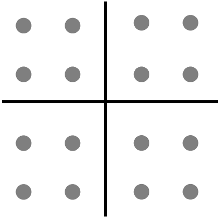
Real and imaginary parts in $\{-3, -1, 1, 3\}$.

Phase Shift Keying (PSK)

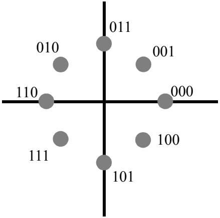
M-PSK $\Rightarrow$ symbols equally spaced on a circle with argument $\frac{2\pi}{M} m$, $m \in \{0, 1, \cdots, M-1\}$.

8-PSK $\Rightarrow 2^3$ symbols $\Rightarrow$ 3 bits/symbol.

16-QAM and 8-PSK modulations (10)



16-QAM



8-PSK

Transmission rate and performance (11)

Transmission rate

T_r = Number of bits per channel use (symbol period)

Example for TST system

Transmission of R data streams composed of N symbols ($\mathbf{S} \in \mathbb{C}^{N \times R}$)
belonging to M-QAM constellation,
spread with a spreading code of length P
(i.e. with P temporal repetitions):

$$T_r = \frac{NR}{NP} \log_2(M) = \frac{R}{P} \log_2(M) \text{ bits/channel use.}$$

Performance

in SER (Symbol Error Rate) or BER (Bit Error Rate)

Diversity gain (12)

Performance is directly linked with the diversity gain due to multiple copies of transmitted signals

$\Updownarrow$

Signal redundancy in space/time/frequency domains.

If number of copies (diversity order) $\nearrow$

then

Quality and reliability of reception $\nearrow$

Space/Time/Frequency spreading and multiplexing (13)

- **Space/Time/frequency spreading** by:
  - Transmitting the same symbols (or data streams) by means of several Tx antennas, and using several Rx antennas at the receiver.

    $\Rightarrow$ **Space diversity**

  - Repeating the same symbols during several chip periods (CDMA systems) or/and multiple time blocks

    $\Rightarrow$ **Code/Time diversities**

  - Transmitting same symbols using several subcarriers (OFDM syst.)

    $\Rightarrow$ **Frequency diversity**

    $\Downarrow$

    Performance and reliability improvement

- **Space multiplexing** by transmitting independent data streams in parallel on multiple Tx antennas $\Rightarrow$ **Transmission rate increase**

Main blocks in a MIMO system (14)

- Data streams containing information symbols to be transmitted ⇒ Symbol matrix $\mathbf{S} \in \mathbb{C}^{N \times R}$: R data streams of N symbols.
- Coding matrices/tensors.
- Resource allocation matrices/tensors.
  Resource constraints:
  - Numbers of transmit and receive antennas.
  - Limited power.
  - Frequency bandwidth.
- Channel matrices/tensors.
- Receiver

Design of receiver (15)

Objective: Best tradeoff between transmission rate and performance.

- Design of transmitter (Choice of coding).
- Choice of relaying protocol (for relaying systems).
- Design of receiver.

Three types of receiver:

- Zero-forcing (ZF): Perfect knowledge of channel (ideal performance).
- Supervised (with a training sequence): to estimate the channel and, in a second stage, the information symbols.
- Semi-blind (only a few pilot symbols are known at the receiver): to jointly estimate the channels and the symbols.

Motivations for tensor modeling of MIMO systems (1)

- MIMO systems $\Rightarrow$ Multidimensional data $\Rightarrow$ Third- to fifth-order tensors for transmitted and received signals
- Structure of tensor model results from system design
- Structure parameters (rank, mode dimensions) are design parameters (code lengths, numbers of Tx/Rx antennas, of data streams, of subcarriers, of time slots, ...)
- Tensor ST/STF coding
- Tensor of resource allocation
- Uniqueness properties of tensor models $\Rightarrow$ ambiguities eliminated with knowledge of a few pilot symbols (no training sequence is required)
- Deterministic semi-blind receivers (joint channels/symbols estimation)

Motivations for tensor modeling of MIMO systems (2)

- Aims:
  - Best tradeoff between error performance (SER or BER), transmission rate (in symbols or bits per channel use), and receiver complexity for symbol recovery.
  - Semi-blind receivers for joint channels/symbols estimation (i.e. without training sequence).
- Performance improvement by jointly exploiting several diversities.

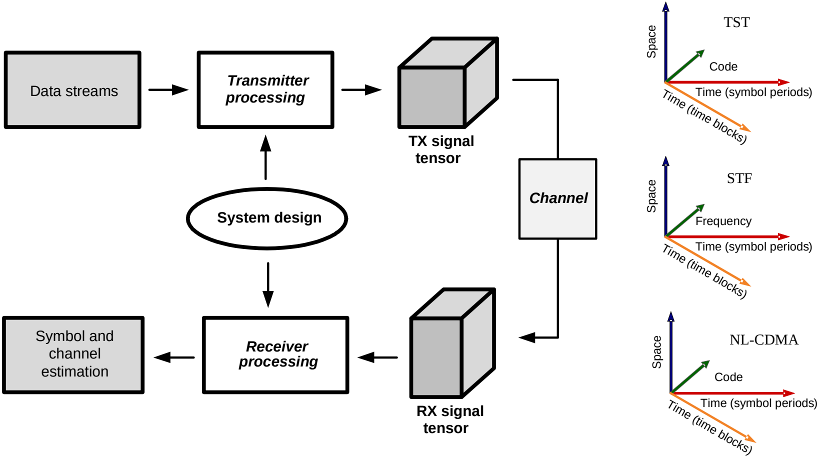
$$\Updownarrow$$

To exploit redundancy into information-bearing signals at the receiver.

$$\Updownarrow$$

Tensor spreading/Coding in space, time and/or frequency domains.

Block-diagram of tensor-based MIMO systems (3)



PART 2 : TENSOR MODELS/DECOMPOSITIONS

- **1 Brief history**
- **2 Examples and definitions**
- **3 Notations, operations and matricizations**
- **4 Tensor models/decompositions**

Brief history (1)

- From the sixties: Tensor decompositions were used for analysing collections of data matrices viewed as three-way data arrays (third-order tensors):

  - 1966: Tucker decomposition in psychometrics.

  - 1970: PARAFAC (parallel factor) decomposition introduced by Harshman in phonetics, and independently under the name CANDECOMP (canonical decomposition) by Carroll and Chang in psychometrics, also called CP (CANDECOMP/PARAFAC) by Kiers (2000). Rediscovered by Möcks (1988) under the name "topographic component model" in EEG analysis.

Brief history (2)

- From 1990: Tensor decompositions were used in:
  - Chemistry, especially in chemometrics (R. Bro's Ph.D. thesis, 1998).
  - Signal processing: blind source separation (BSS) using cumulant tensors (J.F. Cardoso, P. Comon, 1990; L. de Lathauwer's, 1997).

- Since 2000: Tensor decompositions/models are used for designing wireless communication systems (N. Sidiropoulos et al., 2000), and analysing image ensembles (Vasilescu and Terzopoulos, 2002).
- During the last decade: we developed several tensor models for designing MIMO comm. syst.: block constrained PARAFAC, CONFAC, generalized PARATUCK, nested PARAFAC, Tucker train.
- Nowadays: High order tensors, also called multi-way arrays, are used for representing and analysing multidimensional data under the form of signals, images, speech, music sequences, or texts.

Motivations for using tensor decompositions (1)

$\mapsto$ Separation of data sets into components/factors in order to extract the multimodal structure of data and useful information from noisy measurements.

$\mapsto$ Dimensionality reduction of multidimensional data:
$\Rightarrow$ Approximate low-rank tensor decompositions,
$\Rightarrow$ Tensor train decompositions.

$\mapsto$ Completion of data tensors in presence of missing data.

$\Rightarrow$ Use of a low-rank tensor decomposition for modeling the data tensor of interest.

Motivations for using tensors in SP (2)

- Moments and cumulants of RV and stochastic processes are tensors.

  ⇒ Development of tensor SP methods based on high order statistics (HOS).

- Design of MIMO wireless communication systems.

  ⇒ Semi-blind receivers for joint channel and symbols estimation.

- Modeling and analysis of multidimensional and nonlinear systems.

  ⇒ Development of Volterra-PARAFAC models, with reduced parametric complexity, by considering Volterra kernels as tensors. (Favier et al.; GRETSI'2009 and 2011, TS'2010, IJACSP'2012, SP'2012)

Some examples of tensors

- Tensors of statistics (moments and cumulants).
- Kernels of Volterra models for nonlinear system modeling.
- Tensors of transmitted and received signals in MIMO communication systems.
- Tensors of biomedical signals (EEG, ECG, MEG).
- Tensors of images and video data.
- Tensors for data analysis in phonetics, chemometrics, bioinformatics,...
- Tensors for data mining and web search.
- FaceTensors for face recognition.

Tensors of image and video data

Datasets	Modes
Color images	Spatial column $\times$ Spatial row $\times$ Color
Hyperspectral images	Spatial column $\times$ Spatial row $\times$ Spectral
Gray-level video sequences	Spatial column $\times$ Spatial row $\times$ Time
Color video sequences	Column $\times$ Row $\times$ Color $\times$ Time

Applications for compression and recognition/classification

- Medical image analysis.
- 3D object recognition.
- Surveillance: Biometrics (Face recognition); hyperspectral surveillance (military).
- Human-computer-interaction (HCI): space-time analysis of video sequences for gesture and activity recognition.
- Hyperspectral imaging used in agriculture, food industry, environment...

Notations, definitions, and tensor operations

Scalars, column vectors, matrices, and tensors of order higher than two:

$$a, \mathbf{a}, \mathbf{A}, \mathcal{A}$$

- $\circ$: vector outer product (also called tensor product).
- $\odot$: Hadamard (element-wise) product.
- $\diamond$: Khatri-Rao (column-wise Kronecker) product.
- $\otimes$: Kronecker product.
- $\times_n$: Mode-n product of a tensor $\mathcal{X}$ with a matrix $\mathbf{A}$.

Definitions

N^{th}-order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ = multidimensional array of data/measurements.

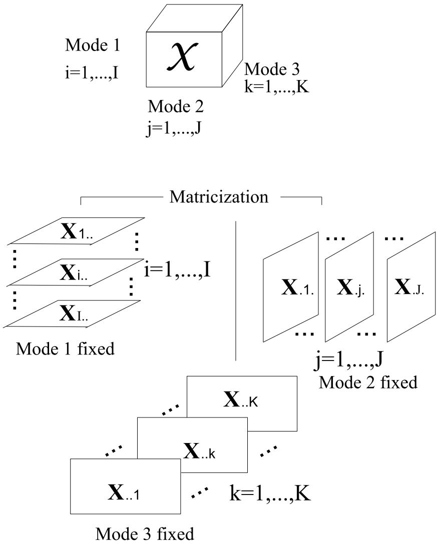
Order N = number of indices that characterize each element $x_{i_1,\cdots,i_N}$.

Each index i_n ($i_n = 1, \cdots, I_n, \text{for } n = 1, \cdots, N$) is associated with a way, also called a mode, and I_n = mode-n dimension.

Particular cases:

Cases	N	Elements	Coefficients
Vectors	1	$\mathbf{x} \in \mathbb{C}^{I \times 1}$	x_i
Matrices	2	$\mathbf{X} \in \mathbb{C}^{I \times J}$	x_{ij}
Three-way arrays	3	$\mathcal{X} \in \mathbb{C}^{I \times J \times K}$	x_{ijk}

Matrix slices of a third-order tensor (horizontal, lateral,frontal slices)



Vector slices of a third-order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$

By fixing two indices:

- Columns: j and k fixed $\Rightarrow$ JK columns $\mathbf{x}_{.jk} \in \mathbb{C}^{I \times 1}$.
- Rows: i and k fixed $\Rightarrow$ IK rows $\mathbf{x}_{i.k} \in \mathbb{C}^{J \times 1}$.
- Tubes: i and j fixed $\Rightarrow$ IJ tubes $\mathbf{x}_{ij.} \in \mathbb{C}^{K \times 1}$.

Matrix slices of a third-order tensor

Matrix slices of $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ (horizontal, lateral, frontal slices):

By fixing one index

$$\mathbf{X}_{i..} \in \mathbb{C}^{J \times K}, \mathbf{X}_{.j.} \in \mathbb{C}^{K \times I}, \mathbf{X}_{..k} \in \mathbb{C}^{I \times J}.$$

Horizontal slices

$$\mathbf{X}_{i..} = \begin{pmatrix} x_{i11} & x_{i12} & \cdots & x_{i1K} \\ x_{i21} & x_{i22} & \cdots & x_{i2K} \\ \vdots & \vdots & \ddots & \vdots \\ x_{iJ1} & x_{iJ2} & \cdots & x_{iJK} \end{pmatrix} \in \mathbb{C}^{J \times K}.$$

Matrix slices of a third-order tensor

Lateral slices

$$\mathbf{X}_{.j.} = \begin{pmatrix} x_{1j1} & x_{2j1} & \cdots & x_{Ij1} \\ x_{1j2} & x_{2j2} & \cdots & x_{Ij2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1jK} & x_{2jK} & \cdots & x_{IjK} \end{pmatrix} \in \mathbb{C}^{K \times I}.$$

Frontal slices

$$\mathbf{X}_{..k} = \begin{pmatrix} x_{11k} & x_{12k} & \cdots & x_{1Jk} \\ x_{21k} & x_{22k} & \cdots & x_{2Jk} \\ \vdots & \vdots & \ddots & \vdots \\ x_{I1k} & x_{I2k} & \cdots & x_{IJk} \end{pmatrix} \in \mathbb{C}^{I \times J}.$$

Matricization of a third-order tensor

Matricization = transformation of a tensor under the form of a matrix

Two different forms of matricization, called flat and tall matrix unfoldings:

Flat unfoldings: $\mathbf{X}_{I\times JK}, \mathbf{X}_{I\times KJ}, \mathbf{X}_{J\times KI}, \mathbf{X}_{J\times IK}, \mathbf{X}_{K\times IJ}, \mathbf{X}_{K\times JI}$

Tall unfoldings: $\mathbf{X}_{JK\times I}, \mathbf{X}_{KJ\times I}, \mathbf{X}_{KI\times J}, \mathbf{X}_{IK\times J}, \mathbf{X}_{IJ\times K}, \mathbf{X}_{JI\times K}$

$$\begin{aligned}
\mathbf{X}_{I\times KJ} &= [\mathbf{X}_{..1} \cdots \mathbf{X}_{..K}] = \mathbf{X}_{KJ\times I}^T \\
\mathbf{X}_{J\times IK} &= [\mathbf{X}_{1..} \cdots \mathbf{X}_{I..}] = \mathbf{X}_{IK\times J}^T \\
\mathbf{X}_{K\times JI} &= [\mathbf{X}_{.1.} \cdots \mathbf{X}_{.J.}] = \mathbf{X}_{JI\times K}^T
\end{aligned}$$

Convention: order of dimensions in a product IJK is linked to order of variation of the corresponding indices (i, j, k).
$\mathbf{X}_{JK\times I} \in \mathbb{C}^{JK\times I} \Rightarrow$ combination of modes (j,k) such that j varies more slowly than $k \Rightarrow x_{i,j,k} = [\mathbf{X}_{JK\times I}]_{(j-1)K+k,i} = [\mathbf{X}_{K\times IJ}]_{k,(i-1)J+j}$.

Matricization of an Nth-order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ (Favier, de Almeida; EURASIP JASP'2014)

Partitioning of $\{1, \ldots, N\}$ into two ordered subsets $\mathbb{S}_1$ and $\mathbb{S}_2$, constituted of $p \in [1, N-1]$ and $N-p$ indices, respectively.

General matricization formula

$$\mathbf{X}_{\mathbb{S}_1;\mathbb{S}_2} = \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} x_{i_1,\cdots,i_N} \left(\underset{n\in\mathbb{S}_1}{\otimes} \mathbf{e}_{i_n}^{(I_n)} \right) \left(\underset{n\in\mathbb{S}_2}{\otimes} \mathbf{e}_{i_n}^{(I_n)} \right)^T \in \mathbb{C}^{J_1 \times J_2}$$

$$\text{with } J_{n_1} = \prod_{n\in\mathbb{S}_{n_1}} I_n, \text{ for } n_1 = 1 \text{ and } 2.$$

$\mathbb{S}_1 \Leftrightarrow$ Combination of modes to form the rows of $\mathbf{X}_{\mathbb{S}_1;\mathbb{S}_2}$

$\mathbb{S}_2 \Leftrightarrow$ Combination of modes to form the columns of $\mathbf{X}_{\mathbb{S}_1;\mathbb{S}_2}$

$\mathbf{e}_{i_n}^{(I_n)} = {i_n}^{th}$ canonical vector of the Euclidean space $\mathbb{R}^{I_n}$.

Mode-n matrix unfolding of $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$

Flat matrix unfolding $\mathbf{X}_{\mathbb{S}_1;\mathbb{S}_2}$ with
$\mathbb{S}_1 = \{n\}$ and $\mathbb{S}_2 = \{n+1, \cdots, N, 1, \cdots, n-1\}$

$$\begin{aligned} \mathbf{X}_n &= \mathbf{X}_{I_n \times I_{n+1} \cdots I_N I_1 \cdots I_{n-1}} \\ &= \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} x_{i_1, \cdots, i_N} \mathbf{e}_{i_n}^{(I_n)} \left(\underset{n \in \mathbb{S}_2}{\otimes} \mathbf{e}_{i_n}^{(I_n)} \right)^T \in \mathbb{C}^{I_n \times I_{n+1} \cdots I_N I_1 \cdots I_{n-1}}. \end{aligned}$$

Column vectors of $\mathbf{X}_n$ = mode-n vectors of $\mathcal{X}$, and rank of $\mathbf{X}_n$, i.e. the dimension of the mode-n linear space spanned by the mode-n vectors, is called mode-n rank of $\mathcal{X}$, denoted by $R_n = \mathrm{rank}_n(\mathcal{X})$.

N-uplet $(R_1, \cdots, R_N)$ = multilinear rank (mrank) of $\mathcal{X}$.
In general, the mode-n ranks R_n are different, unlike the matrix case $(R_1 = R_2)$.

Mode-n product (1)

Mode-n product of a tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ with a matrix $\mathbf{A} \in \mathbb{C}^{J_n \times I_n}$ denoted $\mathcal{X} \times_n \mathbf{A}$, gives the tensor $\mathcal{Y}$ of order N and dimensions $I_1 \times \cdots \times I_{n-1} \times J_n \times I_{n+1} \times \cdots \times I_N$, such as

$$
\begin{aligned}
y_{i_1,\cdots,i_{n-1},j_n,i_{n+1},\cdots,i_N} &= \sum_{i_n=1}^{I_n} a_{j_n,i_n} x_{i_1,\cdots,i_{n-1},i_n,i_{n+1},\cdots,i_N} \\
\mathbf{Y}_n &= \mathbf{A}\mathbf{X}_n.
\end{aligned}
$$

where $\mathbf{X}_n \in \mathbb{C}^{I_n \times I_{n+1} \cdots I_N I_1 \cdots I_{n-1}}$ = Mode-n matrix unfolding of $\mathcal{X}$.

Property: for $\mathbf{A} \in \mathbb{C}^{J_n \times I_n}$ and $\mathbf{B} \in \mathbb{C}^{K_n \times J_n}$.

$$
\mathcal{X} \times_n \mathbf{A} \times_n \mathbf{B} = \mathcal{X} \times_n (\mathbf{B}\mathbf{A}) \in \mathbb{C}^{I_1 \times \cdots \times I_{n-1} \times K_n \times I_{n+1} \times \cdots \times I_N}
$$

Mode-n product (2)

$$\mathcal{Y} = \mathcal{X} \times_n \mathbf{A} \;\Leftrightarrow\; \mathbf{Y}_n = \mathbf{A}\mathbf{X}_n$$

Interpretation as a linear transformation of the mode-n space of $\mathcal{X}$, with the matrix $\mathbf{A}$

Generalization

$$\begin{aligned} \mathcal{Y} &= \mathcal{X} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \cdots \times_N \mathbf{A}^{(N)} \\ &= \mathcal{X} \times_{n=1}^{N} \mathbf{A}^{(n)} \end{aligned}$$

Multilinear (N-linear) transformation of $\mathcal{X}$

Rank-one tensor

Rank-one matrix

$$\mathbf{X} = \mathbf{u} \circ \mathbf{v} = \mathbf{u}\mathbf{v}^T \in \mathbb{C}^{I \times J} \Leftrightarrow x_{ij} = u_i v_j$$

Rank-one tensor of third-order

$$\mathcal{X} = \mathbf{u} \circ \mathbf{v} \circ \mathbf{w} \in \mathbb{C}^{I \times J \times K} \Leftrightarrow x_{ijk} = u_i v_j w_k,$$

Rank-one tensor of order N = outer product of N vectors

$$\mathcal{X} = \mathbf{u}^{(1)} \circ \mathbf{u}^{(2)} \circ \cdots \circ \mathbf{u}^{(N)} = \underset{n=1}{\overset{N}{\circ}} \mathbf{u}^{(n)} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$$

$$x_{i_1,\cdots,i_N} = \left(\underset{n=1}{\overset{N}{\circ}} \mathbf{u}^{(n)} \right)_{i_1,\cdots,i_N} = \prod_{n=1}^{N} u_{i_n}^{(n)}$$

Generalization of matrix decompositions

Matrix BD (bilinear decompos.) $\rightarrow$ PARAFAC/CANDECOMP models
also called canonical polyadic decomposition (CPD)
Harshman 1970; Carroll and Chang 1970; Hitchcock, 1927

Matrix SVD $\rightarrow$ HOSVD/Tucker models
Tucker 1966; De Lathauwer 1997

Constrained tensor models

PARALIND/CONFAC models

Bro, Harshman, Sidiropoulos, 2005; de Almeida, Favier, Motta; IEEE TSP'2008

PARATUCK / Generalized PARATUCK models

Harshman, Lundy; 1996

Favier et al.; SP'2012; Favier, de Almeida; EURASIP JASP'2014

Tensor trains (TT)

Oseledets, 2011

Special cases

Tucker trains (also called Nested Tucker (NT) models)

Favier et al., SP'2016

Nested PARAFAC models

de Almeida, Favier; IEEE SPL'2013

PARAFAC models/CPD (1)

Case of third-order tensors $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ (Harshman, 1970)

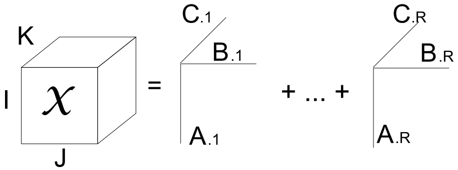
PARAFAC = Sum of R rank-one tensors (triadic decompositions)

.

$$x_{ijk} = \sum_{r=1}^{R} a_{ir} b_{jr} c_{kr}$$

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{A}_{.r} \circ \mathbf{B}_{.r} \circ \mathbf{C}_{.r} = \mathcal{I}_R \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C} = \|\mathbf{A}, \mathbf{B}, \mathbf{C}\|$$

Matrix factors: $\mathbf{A} \in \mathbb{C}^{I \times R}, \mathbf{B} \in \mathbb{C}^{J \times R}, \mathbf{C} \in \mathbb{C}^{K \times R}$



Variants of third-order PARAFAC models (2)

Models	Ref	$x_{i,j,k}$	Applications
CP	Harshman 1970	$\sum_{r=1}^{R} a_{ir} b_{jr} c_{kr}$	Many fieds
INDSCAL	Carroll et al. 1970	$\sum_{r=1}^{R} a_{ir} a_{jr} c_{kr}$	Psychometrics
Sym. CP	Comon et al. 2008	$\sum_{r=1}^{R} a_{ir} a_{jr} a_{kr}$	Volterra models
DSym CP	Favier et al. 2012	$\sum_{r=1}^{R} a_{ir} a_{jr} \bar{a}_{kr}$	NL com. chan.
ShiftCP	Morup et al. 2011	$\sum_{r=1}^{R} a_{ir} b_{j-t_k,r} c_{kr}$	Neuroimaging
	Harshman et al., 2003		
ConvCP	Morup et al. 2011	$\sum_{r=1}^{R} \sum_{t=1}^{T} a_{ir} b_{j-t,r} c_{k,r,t}$	Neuroimaging

PARAFAC models (3)

Uniqueness issue

Case of a third-order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$

PARAFAC decomposition $\|\mathbf{A}, \mathbf{B}, \mathbf{C}\|$ of rank R

Kruskal's condition (Kruskal; 1977):

$$k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2$$

where $k_{\mathbf{A}}$ denotes the k-rank of $\mathbf{A}$, i.e. the largest integer such that any set of $k_{\mathbf{A}}$ columns of $\mathbf{A}$ is linearly independent.

Remarks

- This condition is sufficient but not necessary for essential uniqueness (i.e. for column permutation and scaling ambiguities).
- This condition does not hold when $R = 1$. It is also necessary for $R = 2$ and $R = 3$ but not for $R > 3$ (ten Berge, Sidiropoulos; 2002).

PARAFAC models (4)

Uniqueness issue

Case of an Nth order PARAFAC model

$$x_{i_1,\cdots,i_N} = \sum_{r=1}^{R} \prod_{n=1}^{N} a_{i_n,r}^{(n)}$$

$$\mathcal{X} = \mathcal{I}_{N,R} \times_{n=1}^{N} \mathbf{A}^{(n)}$$

Sufficient uniqueness condition (Sidiropoulos, Bro; 2000)

$$\sum_{n=1}^{N} k_{\mathbf{A}^{(n)}} \geq 2R + N - 1$$

Generic case (full rank factor matrices; $k_{\mathbf{A}^{(n)}} = r_{\mathbf{A}^{(n)}} = \min(I_n, R)$):

$$\sum_{n=1}^{N} \min(I_n, R) \geq 2R + N - 1$$

PARAFAC models (5)

Matricization

Third-order tensors

$$\mathbf{X}_{IJ\times K} = (\mathbf{A} \diamond \mathbf{B})\mathbf{C}^T$$
$$\mathbf{X}_{JK\times I} = (\mathbf{B} \diamond \mathbf{C})\mathbf{A}^T$$
$$\mathbf{X}_{KI\times J} = (\mathbf{C} \diamond \mathbf{A})\mathbf{B}^T$$

Trilinear model w.r.t. $(\mathbf{A}, \mathbf{B}, \mathbf{C})$

Nth-order tensors

$$\mathbf{X}_{\mathbb{S}_1;\mathbb{S}_2} = \left(\underset{n\in\mathbb{S}_1}{\diamond} \mathbf{A}^{(n)} \right) \left(\underset{n\in\mathbb{S}_2}{\diamond} \mathbf{A}^{(n)} \right)^T .$$

PARAFAC model estimation (6)

Alternating Least Squares (ALS) algorithm

- Identification of a PARAFAC model = estimation of $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ from the data tensor $\mathcal{X}$, by minimizing

$$\min_{\mathbf{A},\mathbf{B},\mathbf{C}} \left\| \mathbf{X}_{JK \times I} - (\mathbf{B} \diamond \mathbf{C})\mathbf{A}^T \right\|_F^2 \Rightarrow \text{NL optimization.}$$

- Alternating minimization of three conditional LS cost functions:

$$\min_{\mathbf{A}} \left\| \mathbf{X}_{JK \times I} - (\mathbf{B}_{t-1} \diamond \mathbf{C}_{t-1})\mathbf{A}^T \right\|_F^2 \Rightarrow \mathbf{A}_t$$

$$\min_{\mathbf{B}} \left\| \mathbf{X}_{KI \times J} - (\mathbf{C}_{t-1} \diamond \mathbf{A}_t)\mathbf{B}^T \right\|_F^2 \Rightarrow \mathbf{B}_t$$

$$\min_{\mathbf{C}} \left\| \mathbf{X}_{IJ \times K} - (\mathbf{A}_t \diamond \mathbf{B}_t)\mathbf{C}^T \right\|_F^2 \Rightarrow \mathbf{C}_t.$$

  - Trilinear LS problem requiring a nonlinear optimization method transformed into three linear LS problems successively solved by means of the standard LS solution.

PARAFAC model estimation (7)

ALS algorithm

ALS algorithm

1. Initialize $\mathbf{B}_0$ and $\mathbf{C}_0$ and set $t = 0$.
2. Increment t and compute:
   - $(\mathbf{A}_t)^T = (\mathbf{B}_{t-1} \diamond \mathbf{C}_{t-1})^\dagger \mathbf{X}_{JK\times I}$.
   - $(\mathbf{B}_t)^T = (\mathbf{C}_{t-1} \diamond \mathbf{A}_t)^\dagger \mathbf{X}_{KI\times J}$.
   - $(\mathbf{C}_t)^T = (\mathbf{A}_t \diamond \mathbf{B}_t)^\dagger \mathbf{X}_{IJ\times K}$.
3. Return to step 2 until convergence.

PARAFAC model estimation (8)

ALS algorithm

- Advantages:
  - Simplicity.
  - Easy extension to higher-order PARAFAC models and other tensor models.
- Drawbacks:
  - Slow convergence (iterative algorithm).
  - Convergence towards the global minimum is not guaranteed, depending on the initialization.
- Solutions exist for improving the convergence speed: Levenberg-Marquardt, conjugate gradient, enhanced line search (ELS) methods.
- In certain applications : certain factors are known (partial estimation).

Closed form algorithm (Kibangou, Favier; EUSIPCO'2009)

Assumption: $\mathbf{C}$ known and full-column rank $\Rightarrow$ $\mathbf{C}^T$ right invertible

$$
\begin{aligned}
\mathbf{X}_{IJ\times K} &= (\mathbf{A} \diamond \mathbf{B})\mathbf{C}^T \\
&\Downarrow \\
\mathbf{W} &= \mathbf{A} \diamond \mathbf{B} = \mathbf{X}_{IJ\times K}(\mathbf{C}^T)^{\dagger} \in \mathbb{C}^{IJ\times R} \\
&\Downarrow \\
\mathbf{W}_{.r} &= \mathbf{A}_{.r} \otimes \mathbf{B}_{.r} \in \mathbb{C}^{IJ\times 1};\ r = 1, \cdots, R \\
&\Downarrow \ \left(\mathrm{vec}(\mathbf{u}\mathbf{v}^T) = \mathbf{v} \otimes \mathbf{u} \Leftrightarrow \mathrm{unvec}(\mathbf{v} \otimes \mathbf{u}) = \mathbf{u}\mathbf{v}^T\right) \\
\mathbf{F}^{(r)} &= \mathrm{unvec}_{J\times I}(\mathbf{W}_{.r}) = \mathrm{unvec}_{J\times I}(\mathbf{A}_{.r} \otimes \mathbf{B}_{.r}) = \mathbf{B}_{.r}\mathbf{A}_{.r}^T \\
&\Downarrow \ \text{SVD of a rank-one matrix} \\
\mathbf{F}^{(r)} &= \sigma_1 \mathbf{u}_1 \mathbf{v}_1^H \\
&\Downarrow \\
\mathbf{A}_{.r} &= \sqrt{\sigma_1}\mathbf{v}_1^* \ ;\ \mathbf{B}_{.r} = \sqrt{\sigma_1}\mathbf{u}_1 \\
&\Rightarrow \ \text{Computation of R SVDs to estimate } (\mathbf{A}, \mathbf{B}).
\end{aligned}
$$

Tucker models (1)

Case of third-order tensors $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ (Tucker, 1966)

$$
\begin{aligned}
x_{ijk} &= \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{s=1}^{S} g_{pqs} a_{ip} b_{jq} c_{ks} \\
\mathcal{X} &= \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{s=1}^{S} g_{pqs} \, \mathbf{A}_{.p} \circ \mathbf{B}_{.q} \circ \mathbf{C}_{.r} \\
&= \mathcal{G} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C}
\end{aligned}
$$

Core tensor $\mathcal{G} \in \mathbb{C}^{P \times Q \times S}$; Matrix factors $\mathbf{A} \in \mathbb{C}^{I \times P}, \mathbf{B} \in \mathbb{C}^{J \times Q}, \mathbf{C} \in \mathbb{C}^{K \times S}$

Special cases:

- HOSVD $\Rightarrow$ $\mathbf{A} \in \mathbb{C}^{I \times I}$, $\mathbf{B} \in \mathbb{C}^{J \times J}$ and $\mathbf{C} \in \mathbb{C}^{K \times K}$ unitary (orthog.).
- Truncated HOSVD $\Rightarrow$ $\mathbf{A} \in \mathbb{C}^{I \times P}$, $\mathbf{B} \in \mathbb{C}^{J \times Q}$ and $\mathbf{C} \in \mathbb{C}^{K \times S}$ column-orthonormal.
- PARAFAC $\Rightarrow$ $\mathcal{G} = \mathcal{I}$; $P = Q = S = R$.

Tucker model of third-order tensors $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ (2)

Matricization

Matrix representations

$$\mathbf{X}_{IJ \times K} = (\mathbf{A} \otimes \mathbf{B})\mathbf{G}_{PQ \times S}\mathbf{C}^T$$
$$\mathbf{X}_{JK \times I} = (\mathbf{B} \otimes \mathbf{C})\mathbf{G}_{QS \times P}\mathbf{A}^T$$
$$\mathbf{X}_{KI \times J} = (\mathbf{C} \otimes \mathbf{A})\mathbf{G}_{SP \times Q}\mathbf{B}^T$$

Quadrilinear model w.r.t. $(\mathcal{G}, \mathbf{A}, \mathbf{B}, \mathbf{C})$

Tucker model estimation (3)

- ALS algorithm
- Closed-form algorithm (HOSVD)
- Closed-form algorithm based on Kronecker product approximation, when the core tensor and one factor matrix are known.

Tucker model estimation (4)

For orthogonal factor matrices ($\mathbf{A}^{\dagger} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T = \mathbf{A}^T$):

$$
\begin{aligned}
\mathbf{X}_{IJ\times K} &= (\mathbf{A}\otimes\mathbf{B})\mathbf{G}_{PQ\times S}\mathbf{C}^T \\
&\Downarrow \\
\mathbf{G}_{PQ\times S} &= (\mathbf{A}\otimes\mathbf{B})^{\dagger}\mathbf{X}_{IJ\times K}(\mathbf{C}^{\mathbf{T}})^{\dagger} \\
&= (\mathbf{A}\otimes\mathbf{B})^T\mathbf{X}_{IJ\times K}\mathbf{C}
\end{aligned}
$$

HOSVD

- 1. $\mathbf{A}$ equals first P left singular vectors of $\mathbf{X}_{I\times JK}$.
- 2. $\mathbf{B}$ equals first Q left singular vectors of $\mathbf{X}_{J\times KI}$.
- 3. $\mathbf{C}$ equals first S left singular vectors of $\mathbf{X}_{K\times IJ}$.
- 4. $\mathbf{G}_{PQ\times S} = (\mathbf{A}\otimes\mathbf{B})^T\mathbf{X}_{IJ\times K}\mathbf{C}$.

Tucker models - Case of N^{th}-order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ (5)

$$x_{i_1,\cdots,i_N} = \sum_{r_1=1}^{R_1} \cdots \sum_{r_N=1}^{R_N} g_{r_1,\cdots,r_N} \prod_{n=1}^{N} a_{i_n,r_n}^{(n)}$$

Writing in terms of vector outer products:

$$\mathcal{X} = \sum_{r_1=1}^{R_1} \cdots \sum_{r_N=1}^{R_N} g_{r_1,\cdots,r_N} \mathop{\circ}_{n=1}^{N} \mathbf{A}_{.r_n}^{(n)}$$

$\Leftrightarrow$ Decompos. into a weighted sum of $\prod_{n=1}^{N} R_n$ outer products of N vectors.

Writing in terms of mode-n products:

$$\mathcal{X} \;=\; \mathcal{G} \times_{n=1}^{N} \mathbf{A}^{(n)}$$

$\Rightarrow$ Interpretation as mode-n product-based transformations of the core tensor, i.e. N linear transformations defined by the matrices $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R_n}$ applied to each mode-n vector space of $\mathcal{G} \in \mathbb{C}^{R_1 \times \cdots \times R_N}$.

Tucker models (6)

Uniqueness issue

- Generally, Tucker models are not essentially unique: their matrix factors can be only determined up to nonsingular transformations characterized by nonsingular matrices.
- Uniqueness results from the knowledge of the core tensor.
- Uniqueness can be obtained by imposing some constrained structure on the core tensor or the matrix factors.

Tucker-(N_1, N) models of an N^{th}-order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$, with $N > N_1$

(Favier, de Almeida; EURASIP JASP'2014)

$N - N_1$ factor matrices are equal to identity matrices. For instance, assuming that $\mathbf{A}^{(n)} = \mathbf{I}_{I_n}$, which implies $R_n = I_n$, for $n = N_1 + 1, \cdots, N$:

$$
\begin{aligned}
x_{i_1,\cdots,i_N} &= \sum_{r_1=1}^{R_1} \cdots \sum_{r_{N_1}=1}^{R_{N_1}} g_{r_1,\cdots,r_{N_1},i_{N_1+1},\cdots,i_N} \prod_{n=1}^{N_1} a_{i_n,r_n}^{(n)} \\
\mathcal{X} &= \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \cdots \times_{N_1} \mathbf{A}^{(N_1)} \times_{N_1+1} \mathbf{I}_{I_{N_1+1}} \cdots \times_N \mathbf{I}_{I_N} \\
&= \mathcal{G} \times_{n=1}^{N_1} \mathbf{A}^{(n)}.
\end{aligned}
$$

Tucker-(2,3) models

Case of third-order tensors

Tucker-(2,3) models, also called Tucker-2 models

Third-order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$: core tensor $\mathcal{G} \in \mathbb{C}^{P \times Q \times S}$ and matrix factors $\mathbf{A} \in \mathbb{C}^{I \times P}, \mathbf{B} \in \mathbb{C}^{J \times Q}$,$\mathbf{C} = \mathbf{I}_K$, which implies $S = K$

$$
\begin{aligned}
x_{ijk} &= \sum_{p=1}^{P} \sum_{q=1}^{Q} g_{pqk} a_{ip} b_{jq} \\
\mathcal{X} &= \mathcal{G} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{I}_K = \mathcal{G} \times_1 \mathbf{A} \times_2 \mathbf{B} \\
\mathbf{X}_{IJ \times K} &= (\mathbf{A} \otimes \mathbf{B}) \mathbf{G}_{PQ \times K}
\end{aligned}
$$

Constrained tensor models

CONFAC models (1)
(de Almeida, Favier; IEEE TSP'2008)

Tucker model with PARAFAC core tensor:

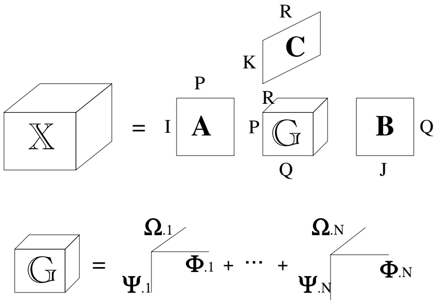
$$\begin{aligned}
\mathcal{X} &= \mathcal{G} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C} \\
\mathcal{G} &= \mathcal{I}_N \times_1 \mathbf{\Psi} \times_2 \mathbf{\Phi} \times_3 \mathbf{\Omega} \\
& \Downarrow \\
\mathcal{X} &= \mathcal{I}_N \times_1 (\mathbf{A}\mathbf{\Psi}) \times_2 (\mathbf{B}\mathbf{\Phi}) \times_3 (\mathbf{C}\mathbf{\Omega}) \\
& \Updownarrow
\end{aligned}$$

Constrained PARAFAC model (PARAFAC with Constrained Factors)

Constraint matrices $\mathbf{\Psi} \in \mathbb{R}^{P \times N}$, $\mathbf{\Phi} \in \mathbb{R}^{Q \times N}$ and $\mathbf{\Omega} \in \mathbb{R}^{R \times N}$ whose columns are chosen as canonical vectors of the Euclidean spaces $\mathbb{R}^P$, $\mathbb{R}^Q$ and $\mathbb{R}^R$, respectively, with $N \geq max(P, Q, R)$.

Third-order CONFAC models (2)

CONFAC(3) = Tucker(3) model with PARAFAC(3) core tensor:



CONFAC models (3)

In a telecommunications context, constraint matrices $(\mathbf{\Psi} \in \mathbb{R}^{P\times N}, \mathbf{\Phi} \in \mathbb{R}^{Q\times N}, \mathbf{\Omega} \in \mathbb{R}^{R\times N})$, are used for allocating (P, Q, R) resources, like data streams, codes, and transmit antennas, to the N components that form the signal to be transmitted.

$$
\begin{aligned}
x_{i,j,k} &= \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R}\Big(\sum_{n=1}^{N}\psi_{p,n}\phi_{q,n}\omega_{r,n}\Big)a_{i,p}b_{j,q}c_{k,r} \\
&= \sum_{n=1}^{N}\Big(\sum_{p=1}^{P}a_{i,p}\psi_{p,n}\Big)\Big(\sum_{q=1}^{Q}b_{j,q}\phi_{q,n}\Big)\Big(\sum_{r=1}^{R}c_{k,r}\omega_{r,n}\Big)
\end{aligned}
$$

Constrained tensor models

PARATUCK models (1)

PARATUCK-2 (or PARATUCK-(2,3)) model (Harshman, Lundy; 1996)

$$\begin{aligned} x_{i,j,k} &= \sum_{p=1}^{P}\sum_{q=1}^{Q}(w_{p,q}\psi_{p,k}\phi_{q,k})a_{i,p}b_{j,q} \\ &= \sum_{p=1}^{P}\sum_{q=1}^{Q} g_{p,q,k}a_{i,p}b_{j,q} \Leftrightarrow \text{Tucker-(2,3) model} \\ &\Downarrow \\ g_{p,q,k} &= w_{p,q}c_{p,q,k} \text{ with } c_{p,q,r} = \sum_{k=1}^{K}\psi_{p,k}\phi_{q,k}\delta_{r,k} = \psi_{p,k}\phi_{q,k} \\ &\Downarrow \\ \mathcal{C} &= \text{PARAFAC}(\|\boldsymbol{\Psi}, \boldsymbol{\Phi}, \mathbf{I}_K\|) \Rightarrow \text{PARATUCK-(2,3)}. \end{aligned}$$

Constrained tensor models

PARATUCK-2 models (2)

$$
\begin{aligned}
x_{i,j,k} &= \sum_{p=1}^{P}\sum_{q=1}^{Q} (w_{p,q}\psi_{p,k}\phi_{q,k}) a_{i,p} b_{j,q} \\
&= \sum_{p=1}^{P}\sum_{q=1}^{Q} w_{p,q} (a_{i,p}\psi_{p,k}) (b_{j,q}\phi_{q,k})
\end{aligned}
$$

Two interpretations of $\mathbf{\Psi}$ and $\mathbf{\Phi}$: Interaction or allocation matrices:

- Interactions between columns p and q of the factor matrices $\mathbf{A}$ and $\mathbf{B}$ along the mode-k of $\mathcal{X}$, with the weights $w_{p,q}$.

- Allocation of resources p and q to the mode-k of $\mathcal{X}$: allocation tensor $\mathcal{C} \in \mathbb{C}^{P\times Q\times K}$ such as $c_{p,q,k} = \psi_{p,k}\phi_{q,k}$; $\mathbf{W} \in \mathbb{C}^{P\times Q}$ = code matrix.

Applications of PARATUCK-2 models (3)

Applications in data analysis (Bro, 1998; Kiers et Smilde, 1998)

First application in SP (Kibangou, Favier; EUSIPCO'2007)
"Blind joint identification and equalization of Wiener-Hammerstein communication channels using PARATUCK-2 tensor decomposition".
$\Rightarrow$ Structured PARATUCK-2 model with Toeplitz and Vandermonde factor matrices $(\mathbf{A}, \mathbf{B})$.

First application in the context of wireless communication systems

(de Almeida, Favier; SP'2009)
"Space-time spreading-multiplexing for MIMO wireless communication systems using PARATUCK-2 tensor model":

$$x_{k,n,p} = \sum_{m=1}^{M} \sum_{r=1}^{R} \underbrace{w_{m,r}}_{\text{code}} \underbrace{h_{k,m}}_{\text{channel}} \underbrace{s_{n,r}}_{\text{symbol}} \underbrace{\phi_{p,m}\,\psi_{p,r}}_{\text{allocations}}$$

Constrained tensor models

PARATUCK-(N_1, N) models (4)

PARATUCK-(2,4) model of fourth-order tensors

TST coding system (Favier et al., EUSIPCO'2011)

$$x_{i,j,k,l} = \sum_{p=1}^{P} \sum_{q=1}^{Q} w_{p,q,l} a_{i,p} b_{j,q} \psi_{p,k} \phi_{q,k}$$

PARATUCK-(N_1, N) (Favier et al., SP'2012).

Tucker-(N_1, N) model with PARAFAC core

$\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$, with $N > N_1$

$$x_{i_1,\cdots,i_{N_1+1},\cdots,i_N} = \sum_{r_1=1}^{R_1} \cdots \sum_{r_{N_1}=1}^{R_{N_1}} c_{r_1,\cdots,r_{N_1},i_{N_1+2},\cdots,i_N} \prod_{n=1}^{N_1} a_{i_n,r_n}^{(n)} \phi_{r_n,i_{N_1+1}}^{(n)}$$

$a_{i_n,r_n}^{(n)}$, and $\phi_{r_n,i_{N_1+1}}^{(n)}$ are entries of the factor matrix $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R_n}$ and of the allocation matrix $\mathbf{\Phi}^{(n)} \in \mathbb{C}^{R_n \times I_{N_1+1}}$, $\forall n = 1, \cdots, N_1$, respectively.

Constrained tensor models

Generalized PARATUCK models (5)
(Favier, de Almeida; IEEE TSP'2014)

PARATUCK-(N_1, N) models with tensor factors $\mathcal{A}^{(n)}$, and allocation tensor $\mathcal{C}$

$$x_{i_1,\cdots,i_N} = \sum_{r_1=1}^{R_1} \cdots \sum_{r_{N_1}=1}^{R_{N_1}} w_{r_1,\cdots,r_{N_1},S} \prod_{n=1}^{N_1} a^{(n)}_{i_n,r_n,S_n} c_{r_1,\cdots,r_{N_1},T}$$

$\{r_1, \cdots, r_{N_1}\}$: input (or resource) modes,
$\{i_1, \cdots, i_N\}$: output (or diversity) modes,

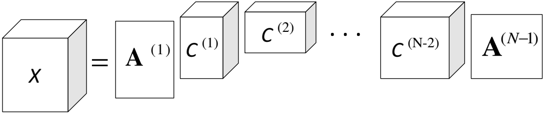
S, T, and $S_n \subseteq S \bigcup T$ (for $n = 1, \cdots, N_1$): subsets of $\{i_{N_1+1}, \cdots, i_N\}$,

$a^{(n)}_{i_n,r_n,S_n}$, $c_{r_1,\cdots,r_{N_1},T}$ (equal to 0 or 1), and $w_{r_1,\cdots,r_{N_1},S}$ are entries of $\mathcal{A}^{(n)}$, of $\mathcal{C}$, and of the core/code tensor $\mathcal{W}$, respectively.

Tensor train decompositions (TTD) (Oseledets, 2011) (1)

Case of an N^{th}-order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$

Objective: Approximation of an N^{th}-order tensor whose parametric complexity is free from exponential dependence on N.



Concatenation of third-order tensors $\mathcal{C}^{(n)} \in \mathbb{C}^{R_n \times I_{n+1} \times R_{n+1}}$, $n = 1, ..., N-2$ and two matrix factors $\mathbf{A}^{(1)} \in \mathbb{C}^{I_1 \times R_1}$, $\mathbf{A}^{(N-1)} \in \mathbb{C}^{I_N \times R_{N-1}}$

$$x_{i_1,i_2,\cdots,i_N} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \cdots \sum_{r_{N-1}=1}^{R_{N-1}} a^{(1)}_{i_1,r_1} c^{(1)}_{r_1,i_2,r_2} c^{(2)}_{r_2,i_3,r_3} \cdots c^{(N-2)}_{r_{N-2},i_{N-1},r_{N-1}} a^{(N-1)}_{i_N,r_{N-1}}$$

R_n $(n = 1, \cdots, N)$ = TT ranks, also called compression ranks.

Tensor train (TT) decompositions (2)

Parametric complexity

Other writing as a product of a row vector, $(N-2)$ matrices, and a column vector:

$$x_{i_1,i_2,\cdots,i_N} = \mathbf{A}^{(1)}_{i_1,.}\mathbf{C}^{(1)}_{.,i_2,.}\cdots\mathbf{C}^{(N-2)}_{.,i_{N-1},.}(A^{(N-1)}_{i_N,.})^T$$

$\mathbf{A}^{(1)}_{i_1,.} \in \mathbb{C}^{1\times R_1}$, $\mathbf{A}^{(N-1)}_{i_N,.} \in \mathbb{C}^{1\times R_{N-1}}$, $\mathbf{C}^{(n-1)}_{.,i_n,.} \in \mathbb{C}^{R_{n-1}\times R_n}$, $n = 2,\cdots,N-1$

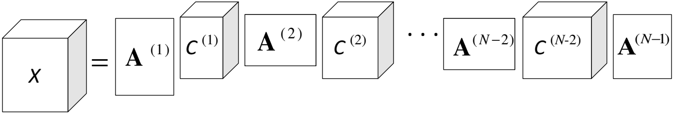
Parametric complexity of the TT representation of $\mathcal{X}$
when $I_n = I$ and $R_n = R$, $\forall n$:

Total number of entries of TT $= 2RI + (N-2)IR^2$ instead of I^N for $\mathcal{X}$.

Tucker train/Nested Tucker decompositions (NTD) (1)

(Favier et al., SP 2016)

$$\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$$

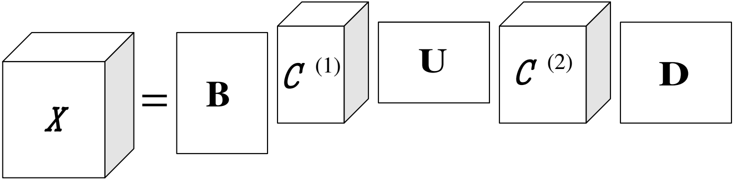


Each third-order tensor $\mathcal{C}^{(n)} \in \mathbb{C}^{R_{2n-1} \times I_{n+1} \times R_{2n}}$, $n \in [1, N-2]$, can be considered as the core tensor of a Tucker-(2,3) model having $(\mathbf{A}^{(n)}, \mathbf{I}_{I_{n+1}}, \mathbf{A}^{(n+1)})$ as matrix factors, with:

$$\mathbf{A}^{(n+1)} \in \mathbb{C}^{R_{2n} \times R_{2n+1}}, n \in [1, N-3], \mathbf{A}^{(1)} \in \mathbb{C}^{I_1 \times R_1}, \mathbf{A}^{(N-1)} \in \mathbb{C}^{I_N \times R_{2N-4}}$$

Two successive third-order Tucker-(2,3) models in the train have a matrix factor in common $\Rightarrow$ NTD

NTD(4) for a fourth-order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times I_3 \times I_4}$ (2)



$$x_{i_1,i_2,i_3,i_4} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \sum_{r_3=1}^{R_3} \sum_{r_4=1}^{R_4} \underbrace{b_{i_1,r_1} c^{(1)}_{r_1,i_2,r_2} u_{r_2,r_3}}_{\text{Tucker-(2,3)}} c^{(2)}_{r_3,i_3,r_4} d_{i_4,r_4}$$

$$= \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \sum_{r_3=1}^{R_3} \sum_{r_4=1}^{R_4} b_{i_1,r_1} c^{(1)}_{r_1,i_2,r_2} \underbrace{u_{r_2,r_3} c^{(2)}_{r_3,i_3,r_4} d_{i_4,r_4}}_{\text{Tucker-(2,3)}}$$

Nested PARAFAC model of a fourth-order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times I_3 \times I_4}$ (1)

(de Almeida, Favier; IEEE SPL'2013)

Special case of nested Tucker model with the following correspondences:

$$
\begin{aligned}
(r_1, r_2, r_3, r_4) &\leftrightarrow (r_1, r_1, r_2, r_2) \\
(\mathbf{B}, \mathcal{C}^{(1)}, \mathbf{U}, \mathcal{C}^{(2)}, \mathbf{D}) &\leftrightarrow (\mathbf{A}, \mathbf{B}, \mathbf{U}, \mathbf{C}, \mathbf{D})
\end{aligned}
$$

$$
\begin{aligned}
x_{i_1,i_2,i_3,i_4} &= \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \underbrace{a_{i_1,r_1} b_{i_2,r_1} u_{r_1,r_2}}_{\text{PARAFAC}} c_{i_3,r_2} d_{i_4,r_2} \\
&= \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} a_{i_1,r_1} b_{i_2,r_1} \underbrace{u_{r_1,r_2} c_{i_3,r_2} d_{i_4,r_2}}_{\text{PARAFAC}}
\end{aligned}
$$

Nested PARAFAC (2)

Define the third-order tensors $\mathcal{W} \in \mathbb{C}^{I_3 \times I_4 \times R_1}$ and $\mathcal{Z} \in \mathbb{C}^{I_1 \times I_2 \times R_2}$ such as

$$
\begin{aligned}
w_{i_3,i_4,r_1} &= \sum_{r_2=1}^{R_2} c_{i_3,r_2} d_{i_4,r_2} u_{r_1,r_2} \\
z_{i_1,i_2,r_2} &= \sum_{r_1=1}^{R_1} a_{i_1,r_1} b_{i_2,r_1} u_{r_1,r_2}
\end{aligned}
$$

or equivalently in terms of mode-n products

$$
\begin{aligned}
\mathcal{W} &= \mathcal{I}_{3,R_2} \times_1 \mathbf{C} \times_2 \mathbf{D} \times_3 \mathbf{U} \\
\mathcal{Z} &= \mathcal{I}_{3,R_1} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{U}^T
\end{aligned}
$$

$\Rightarrow$ $\mathcal{W}$ and $\mathcal{Z}$ satisfy two PARAFAC models.

Nested PARAFAC (3)

Combining the last two modes and the first two ones of $\mathcal{X}$, by means of $k_1 = (i_4 - 1)I_3 + i_3$ and $k_2 = (i_2 - 1)I_1 + i_1$, the 4th-order nested PARAFAC model can be rewritten as two third-order PARAFAC models of the tensors $\mathcal{X}^{(1)} \in \mathbb{C}^{I_1 \times I_2 \times K_1}$ and $\mathcal{X}^{(2)} \in \mathbb{C}^{K_2 \times I_3 \times I_4}$, where $K_1 = I_4 I_3$ and $K_2 = I_2 I_1$

$$x^{(1)}_{i_1,i_2,k_1} = \sum_{r_1=1}^{R_1} a_{i_1,r_1} b_{i_2,r_1} w_{k_1,r_1}$$

$$x^{(2)}_{k_2,i_3,i_4} = \sum_{r_2=1}^{R_2} z_{k_2,r_2} c_{i_3,r_2} d_{i_4,r_2}$$

$\mathcal{X}^{(1)}$ and $\mathcal{X}^{(2)}$ are two contracted forms of $\mathcal{X}$, which satisfy two PARAFAC models $\|\mathbf{A}, \mathbf{B}, \mathbf{W}_{K_1 \times R_1}\|$ and $\|\mathbf{Z}_{K_2 \times R_2}, \mathbf{C}, \mathbf{D}\|$ where $\mathbf{W}_{K_1 \times R_1}$ and $\mathbf{Z}_{K_2 \times R_2}$ are unfoldings of $\mathcal{W}$ and $\mathcal{Z}$ which satisfy their proper PARAFAC models.

The matrices $(\mathbf{A}, \mathbf{B}, \mathbf{U}, \mathbf{C}, \mathbf{D})$ of the nested PARAFAC model can be estimated using a five-step ALS algorithm, or two stages of BALS algo.

If some factors are known $\Rightarrow$ closed-form solution.

Parametric complexities of tensor models

- Data tensor: $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N} \Rightarrow \prod_{n=1}^{N} I_n \simeq I^N$
- PARAFAC($\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R}$) $\Rightarrow R \sum_{n=1}^{N} I_n \simeq NRI$
- Tucker($\mathcal{G} \in \mathbb{C}^{R_1 \times \cdots \times R_N}; \mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R_n}$)
  $\Rightarrow \prod_{n=1}^{N} R_n + \sum_{n=1}^{N} I_n R_n \simeq R^N + NRI$
- Tensor train $\Rightarrow I_1 R_1 + I_N R_{N-1} + \sum_{n=1}^{N-2} R_n I_{n+1} R_{n+1}$
  $\simeq 2RI + (N-2)IR^2$
- Tucker train
  $\Rightarrow I_1 R_1 + I_N R_{2N-4} + \sum_{n=1}^{N-2} R_{2n-1} I_{n+1} R_{2n} + \sum_{n=1}^{N-3} R_{2n} R_{2n+1}$
  $\simeq 2RI + (N-2)IR^2 + (N-3)R^2$

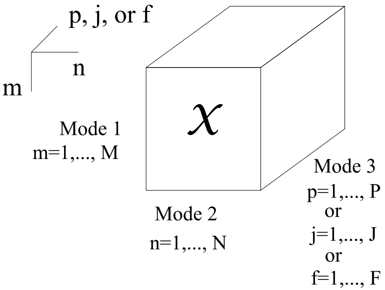
PART 3 : TENSOR-BASED MIMO COMMUNICATION SYSTEMS

- **1 Point to point communication systems**
- **2 Relaying communication systems**

Tensors of signals received by three communication systems

(CDMA, Oversampled, OFDM)

- $\mathcal{X} \in \mathbb{C}^{M \times N \times J}$ or $\in \mathbb{C}^{M \times N \times P}$ or $\in \mathbb{C}^{M \times N \times F}$: received signals tensors
  M receive antennas ; N symbol periods ; F subcarriers
  P: oversampling rate ; J: spreading code length.
- Modes: receive antenna (m), symbol period (n), oversampling period (p), chip (j), frequency (f).



Unified Block constrained PARAFAC modeling (de Almeida, Favier; SP'2007)

PARAFAC-CDMA (code division multiple access) system (1)

(Sidiropoulos, Giannakis, Bro, IEEE TSP 2000)

Q users, K Rx antennas, N symbol periods, J chips (spreading length)

n-th coded (spread) symbol of user q

$$u_{q,n,j} = s_{n,q} w_{j,q}$$

- $s_{n,q}$= symbol transmitted by the q-th user, at the n-th symbol period.
- $w_{j,q}$= j-th code used for spreading each symbol $s_{n,q}$ of the q-th user.

$\Rightarrow$ J repetitions of each symbol $s_{n,q}$.

PARAFAC-CDMA system (2)

Assumption: Multiuser (Q); Rayleigh flat fading channel

$h_{k,q}$= fading coefficient of the channel between q-th user and k-th receive antenna $\Rightarrow$ SIMO system/user.

Signals received by antenna k, during symbol period n

$$x_{k,n,j} = \sum_{q=1}^{Q} h_{k,q} u_{q,n,j} = \sum_{q=1}^{Q} h_{k,q} s_{n,q} w_{j,q} \Rightarrow \mathcal{X} \in \mathbb{C}^{K \times N \times J}$$

$$\Downarrow$$

PARAFAC model

with factors $\mathbf{H} \in \mathbb{C}^{K \times Q}, \mathbf{S} \in \mathbb{C}^{N \times Q}, \mathbf{C} \in \mathbb{C}^{J \times Q}$

Channel, Symbol, Code matrices

Three diversities: space (K), time (N), code (J).

PARAFAC-CDMA system (3)

Joint channel/symbols estimation

$$\mathcal{X} \in \mathbb{C}^{K \times N \times J}$$

Code known at the receiver $\Rightarrow$ BALS receiver

$$\begin{aligned} \mathbf{X}_{KJ \times N} &= (\mathbf{H} \diamond \mathbf{C})\mathbf{S}^T \quad \Rightarrow \quad \hat{\mathbf{S}}^T = (\mathbf{H} \diamond \mathbf{C})^\dagger \mathbf{X}_{KJ \times N} \\ \mathbf{X}_{NJ \times K} &= (\mathbf{S} \diamond \mathbf{C})\mathbf{H}^T \quad \Rightarrow \quad \hat{\mathbf{H}}^T = (\mathbf{S} \diamond \mathbf{C})^\dagger \mathbf{X}_{NJ \times K} \end{aligned}$$

PARAFAC-CDMA system (4)

Uniqueness issue

Kruskal's condition:

$$k_H + k_S + k_C \geq 2Q + 2 \quad (1)$$

Assumptions:

- User-wise independent channel gains $\Rightarrow$ **H** is full k-rank with probability one.
- **S** is full k-rank (if N is large enough).
- **C** is full k-rank (by construction).

$$\Rightarrow \ \min(K, Q) + \min(N, Q) + \min(J, Q) \geq 2Q + 2 \quad (2)$$

PARAFAC-CDMA system (5)

Practical consequences:

- If N and $J \geq Q$: $K \geq 2$ antennas are sufficient $\Rightarrow$ possibility of more users than sensors.
- If N and $K \geq Q$: $J \geq 2$ chips are sufficient.
- If some or all of $\mathbf{H}, \mathbf{S}, \mathbf{C}$ are flat instead of tall, condition (2) may still be satisfied. Example: $K = N = J = 4, Q = 5$.

PARAFAC-CDMA system (6)

Properties

- Deterministic approach for system parameters estimation using received signals only.
- Possibility to have more users (Q) than sensors (K), and/or less spreading (J) than users.
- No need of finite-alphabet, statistical independence, and constant-modulus assumptions.
- Code matrix $\mathbf{C}$ can be estimated $\Rightarrow$ Trilinear ALS algorithm.

Khatri-Rao space-time coding (1) (Sidiropoulos et al., 2002)

M transmit antennas, two coding matrices $\mathbf{\Theta} \in \mathbb{C}^{M \times M}$ and $\mathbf{W} \in \mathbb{C}^{J \times M}$, symbol matrix $\mathbf{S} \in \mathbb{C}^{N \times M}$, N symbol periods, J time blocks (temporal repetitions).

Precoded signal $v_{n,m} = \sum_{l=1}^{M} s_{n,l}\theta_{m,l}$ which combines M symbols of $\mathbf{s}_{n.}$ onto each transmit antenna m + Time spreading $\Rightarrow u_{m,n,j} = v_{n,m} w_{j,m}$.

$\Rightarrow$ Third-order tensor of coded signals: $\mathcal{U} \in \mathbb{C}^{M \times N \times J}$

$$\mathbf{U}_{NJ \times M} = \begin{bmatrix} \mathbf{U}_{.1.}^T \\ \vdots \\ \mathbf{U}_{.N.}^T \end{bmatrix} = \begin{bmatrix} \mathbf{W} D_1(\mathbf{V}) \\ \vdots \\ \mathbf{W} D_N(\mathbf{V}) \end{bmatrix} = \mathbf{V} \diamond \mathbf{W}$$

$$= \underbrace{\mathbf{S}\mathbf{\Theta}^T}_{\text{space-precoding}} \diamond \underbrace{\mathbf{W}}_{\text{time-postcoding}}$$

$\Rightarrow$ Khatri-Rao space-time (KRST) coding.

Khatri-Rao space-time coding (2)

Signal received by antenna k, during symbol period n of time block j:

$$x_{k,n,j} = \sum_{m=1}^{M} h_{k,m} u_{m,n,j} = \sum_{m=1}^{M} \sum_{l=1}^{M} h_{k,m} s_{n,l} \theta_{m,l} w_{j,m}$$
$$= \sum_{m=1}^{M} h_{k,m} v_{n,m} w_{j,m}$$

$\Rightarrow$ Third-order PARAFAC model $(\mathbf{H}, \mathbf{V}, \mathbf{W})$

Known code matrix $\mathbf{W} \Rightarrow$ Estimation of $(\mathbf{H}, \mathbf{V})$ by means of BALS algorithm.

Drawback: Decoding to estimate $\mathbf{S}$ from the estimate of $\mathbf{V} = \mathbf{S}\boldsymbol{\Theta}^T$.

PARATUCK-TST coding system (1)

(Favier et al.; SP 2012)

- MIMO communication system with M transmit antennas and K receive antennas.
- Transmission of R data streams composed of N symbols each.
- Transmission decomposed into P time blocks formed of N time slots each.

PARATUCK-TST coding system (2)

Tensor of transmitted signals (R data streams of N symbols)

ST coded signal transmitted from the transmit antenna m, during the time slot n of block p, and associated with the chip j:

$$u_{m,n,p,j} = \sum_{r=1}^{R} \underbrace{w_{m,r,j}}_{\textbf{code}} \; \underbrace{s_{n,r}}_{\textbf{symbol}} \; \underbrace{\phi_{p,m}\,\psi_{p,r}}_{\textbf{allocations}}$$

$$\textbf{code} \downarrow \mathcal{W} \in \mathbb{C}^{M\times R\times J} \qquad \textbf{symbol} \downarrow \mathbf{S} \in \mathbb{C}^{N\times R} \qquad \textbf{allocations} \downarrow \boldsymbol{\Phi} \in \mathbb{R}^{P\times M}, \boldsymbol{\Psi} \in \mathbb{R}^{P\times R}$$

$$\begin{cases} s_{n,r} = n^{\text{th}} \text{ symbol of } r^{\text{th}} \text{ data stream.} \\ \psi_{p,r} = 1 \Leftrightarrow \text{ data stream } r \text{ allocated to block } p. \\ \phi_{p,m} = 1 \Leftrightarrow \text{ transmit antenna } m \text{ allocated to block } p. \end{cases}$$

$\Rightarrow s_{n,r}$ transmitted using antenna m, during time block p.

PARATUCK-TST coding system (3)

Tensor of received signals

- **Rayleigh flat fading** propagation channel $\mathbf{H} \in \mathbb{C}^{K \times M}$ with i.i.d. CN(0,1) entries.
- Channel assumed to be **constant during at least P time blocks**.

Signal received by antenna k, associated with chip j of symbol period n of time block p:

$$x_{k,n,p,j} = \sum_{m=1}^{M} h_{k,m}\, u_{m,n,p,j}$$

$$= \sum_{m=1}^{M} \sum_{r=1}^{R} w_{m,r,j}\, h_{k,m}\, s_{n,r}\, \phi_{p,m}\, \psi_{p,r}$$

$\Rightarrow \mathcal{X} \in \mathbb{C}^{K \times N \times P \times J}$ satisfies a **PARATUCK-(2,4) model**

$$\mathbf{X}_{..pj} = \mathbf{H}\mathbf{G}_{..pj}\mathbf{S}^T \; ; \; \mathbf{G}_{..pj} = \mathbf{D}_p(\mathbf{\Phi})\mathbf{W}_{..j}\mathbf{D}_p(\mathbf{\Psi})$$

PARATUCK-TST coding system (4)

Semi-blind joint symbol and channel estimation
Matrix representations of the received signal tensor

$$\mathbf{X}_{JPK \times N} = \begin{bmatrix} \mathbf{X}_{\cdot\cdot 1,1} \\ \vdots \\ \mathbf{X}_{\cdot\cdot P,1} \\ \vdots \\ \mathbf{X}_{\cdot\cdot 1,J} \\ \vdots \\ \mathbf{X}_{\cdot\cdot P,J} \end{bmatrix} \in \mathcal{C}^{JPK \times N}, \quad \mathbf{G}_{JPM \times R} = \begin{bmatrix} \mathbf{G}_{\cdot\cdot 1,1} \\ \vdots \\ \mathbf{G}_{\cdot\cdot P,1} \\ \vdots \\ \mathbf{G}_{\cdot\cdot 1,J} \\ \vdots \\ \mathbf{G}_{\cdot\cdot P,J} \end{bmatrix} \in \mathbb{C}^{JPM \times R}, \quad \mathbf{X}_{JPN \times K} = \begin{bmatrix} \mathbf{X}^T_{\cdot\cdot 1,1} \\ \vdots \\ \mathbf{X}^T_{\cdot\cdot P,1} \\ \vdots \\ \mathbf{X}^T_{\cdot\cdot 1,J} \\ \vdots \\ \mathbf{X}^T_{\cdot\cdot P,J} \end{bmatrix} \in \mathcal{C}^{JPN \times K}, \quad \mathbf{G}_{JPR \times M} = \begin{bmatrix} \mathbf{G}^T_{\cdot\cdot 1,1} \\ \vdots \\ \mathbf{G}^T_{\cdot\cdot P,1} \\ \vdots \\ \mathbf{G}^T_{\cdot\cdot 1,J} \\ \vdots \\ \mathbf{G}^T_{\cdot\cdot P,J} \end{bmatrix} \in \mathbb{C}^{JPR \times M}$$

$$\mathbf{X}_{JPK \times N} = (\mathbf{I}_{JP} \otimes \mathbf{H})\,\mathbf{G}_{JPM \times R}\mathbf{S}^T \quad \Rightarrow \quad \mathbf{S}^T = [(\mathbf{I}_{JP} \otimes \mathbf{H})\,\mathbf{G}_{JPM \times R}]^\dagger \mathbf{X}_{JPK \times N} \qquad (3)$$

$$\mathbf{X}_{JPN \times K} = (\mathbf{I}_{JP} \otimes \mathbf{S})\,\mathbf{G}_{JPR \times M}\mathbf{H}^T \quad \Rightarrow \quad \mathbf{H}^T = [(\mathbf{I}_{JP} \otimes \mathbf{S})\,\mathbf{G}_{JPR \times M}]^\dagger \mathbf{X}_{JPN \times K} \qquad (4)$$

S and **H** are estimated **by alternately solving** (3)-(4) **in the LS sense** w.r.t. one matrix conditionally to the knowledge of previously estimated value of the other matrix (BALS algo).

PARATUCK-TST coding system (5)

Advantages

- Tensor coding and resource allocation (Tx antennas and data streams to time blocks).
- Three diversities are exploited: space (K), time (P), chip (J) $\Rightarrow$ Performance improvement w.r.t. the PARAFAC-CDMA system due to the P block repetition of each transmitted data stream, and multiple transmit antennas (M).
- Transmission rate: $\frac{R}{P} log_2(\mu)$ bits/channel use, where μ is the constellation cardinality
- Scalar scaling ambiguity $\Rightarrow$ a single pilot symbol is sufficient $\Rightarrow$ No training sequence is needed for acquiring CSI (channel state information).
- Semi-blind joint channel/symbol estimation.

PARATUCK-TST coding system (6)

Simulation results with QPSK constellation

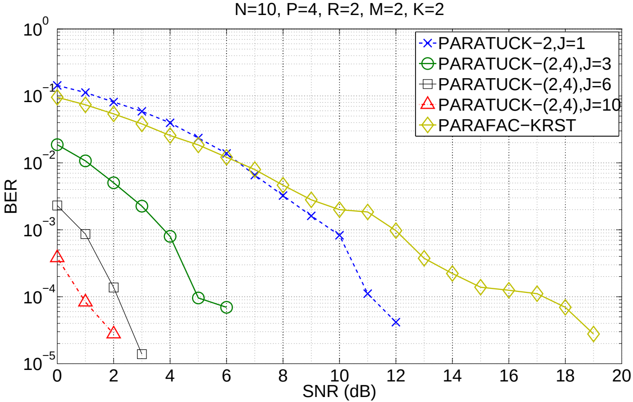


Figure: Impact of the spreading length: BER versus SNR.

BER is improved when J is increased (increase of spreading length).

PARATUCK-TST coding system (7)

Simulation results

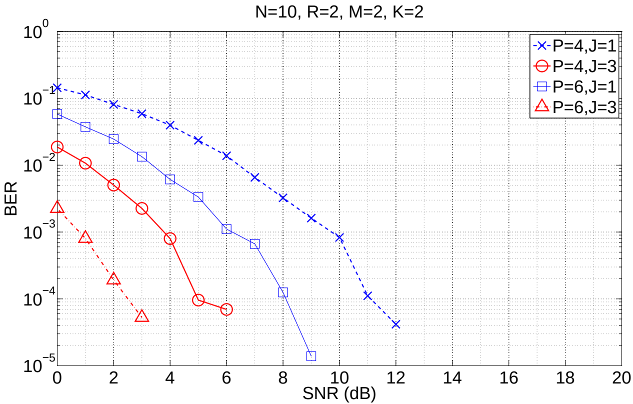


Figure: Impact of P and J: BER versus SNR.

BER ↘ when P and/or J ↗ (diversity gain proportional to KPJ).

Tensor space/time/frequency (TSTF) coding system (1)

(Favier, de Almeida; IEEE TSP'2014)

MIMO system with M transmit and K receive antennas.

Transmission decomposed into P time blocks of N symbol periods, each one being composed of J chips.

During each time block p, the transceiver uses F subcarriers to send R data streams containing N information symbols each, which form the symbol matrix $\mathbf{S} \in \mathbb{C}^{N \times R}$ with entries $s_{n,r}$, $n=1, ..., N$; $r=1, ..., R$.

$\Rightarrow$ CDMA-OFDM system.

TSTF coding system (2)

Transmitter characterized by two tensors: a fifth-order coding tensor $\mathcal{W} \in \mathbb{C}^{M \times R \times F \times P \times J}$ and a fourth-order resource allocation tensor $\mathcal{C} \in \mathbb{R}^{M \times R \times F \times P}$ composed uniquely of 1's and 0's.

$c_{m,r,f,p} = 1 \;\Rightarrow\;$ data stream r transmitted using transmit antenna m and subcarrier f, during time-block p.

Transmission of a linear combination of R coded signals:

$$u_{m,n,f,p,j} = \sum_{r=1}^{R} \underbrace{w_{m,r,f,p,j}}_{\text{code}} \; s_{n,r} \; \underbrace{c_{m,r,f,p}}_{\text{allocations}}$$

Multiplication by $w_{m,r,f,p,j} \Rightarrow$ each symbol $s_{n,r}$ is replicated four times, in the space (m), frequency (f), time (p), and chip (j) dimensions.

TSTF coding system (3)

Frequency-selective fading channel coefficients $h_{k,m,f}$ between each pair (m, k) of transmit and receive antennas, at frequency f, assumed constant during P time-blocks, independent, and circularly symmetric complex Gaussian variables, with zero-mean and unit variance.

Received signals define a fifth-order tensor $\mathcal{X} \in \mathbb{C}^{K \times N \times F \times P \times J}$:

$$
\begin{aligned}
x_{k,n,f,p,j} &= \sum_{m=1}^{M} h_{k,m,f}\, u_{m,n,f,p,j} = \sum_{m=1}^{M} \sum_{r=1}^{R} g_{m,r,f,p,j}\, h_{k,m,f}\, s_{n,r}, \\
g_{m,r,f,p,j} &= w_{m,r,f,p,j}\, c_{m,r,f,p}.
\end{aligned}
$$

Core tensor $\mathcal{G} \in \mathbb{C}^{M \times R \times F \times P \times J}$ can be interpreted as the Hadamard product of coding tensor with allocation tensor, along their common modes: $\mathcal{G} = \mathcal{W} \underset{\{m,r,f,p\}}{\odot} \mathcal{C}$.

TSTF coding system (4)

$\mathcal{X}$ satisfies a generalized PARATUCK-(2,5) model:

$$x_{i_1,i_2,i_3,i_4,i_5} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} g_{r_1,r_2,i_3,i_4,i_5}\, a^{(1)}_{i_1,r_1,i_3}\, a^{(2)}_{i_2,r_2},$$
$$g_{r_1,r_2,i_3,i_4,i_5} = w_{r_1,r_2,i_3,i_4,i_5}\, c_{r_1,r_2,i_3,i_4}.$$

with

$$\left(I_1, I_2, I_3, I_4, I_5, R_1, R_2, \mathcal{A}^{(1)}, \mathbf{A}^{(2)}\right) \leftrightarrow (K, N, F, P, J, M, R, \mathcal{H}, \mathbf{S})\,.$$

TSTF coding system (5)

Matrix unfoldings for designing semi-blind receivers

$$\mathbf{X}_{JPFK \times N} = (\mathbf{I}_{JP} \otimes bdiag(\mathbf{H}_{..f}))\mathbf{G}_{JPFM \times R}\ \mathbf{S}^T,$$
$$\mathbf{X}_{JPFN \times K} = (\mathbf{I}_{JPF} \otimes \mathbf{S})\mathbf{G}_{JPFR \times FM}\ \mathbf{H}_{FM \times K}$$

⇒ BALS semi-blind receiver (Favier, de Almeida; IEEE TSP'2016)

$$\mathbf{X}_{NK \times FPJ} = (\mathbf{S} \otimes \mathbf{H}_{K \times FM})\mathbf{G}_{RFM \times FPJ}$$

⇒ Closed form (Kronecker-based) semi-blind receiver

TSTF coding system (6)

Uniqueness issue

Generalized Tucker-(2,5) model of TSTF system

- From the unfolding $\mathbf{X}_{NK\times FPJ} = (\mathbf{S} \otimes \mathbf{H}_{K\times FM})\mathbf{G}_{RFM\times FPJ}$, it can be proved that the symbol matrix and the channel tensor are unique up to an unknown scalar factor.
- Ambiguity can be eliminated with the knowledge of a single pilot symbol at the receiver.

TSTF coding system (7)

BALS receiver

1. Initialization (it=0): randomly draw $\hat{\mathbf{S}}_{(0)}$ from the symbol alphabet.
2. $it = it + 1$.
3. Calculate the LS estimate of the channel tensor
   $(\hat{\mathbf{H}}_{FM \times K})_{(it)} = ((\mathbf{I}_{JPF} \otimes \hat{\mathbf{S}}_{(it-1)})\mathbf{G}_{JPFR \times FM})^{\dagger} \tilde{\mathbf{X}}_{JPFN \times K}$
4. Calculate the LS estimate of the symbol matrix
   $\hat{\mathbf{S}}^{T}_{(it)} = ((\mathbf{I}_{JP} \otimes bdiag(\hat{\mathbf{H}}_{..f})_{(it)})\mathbf{G}_{JPFM \times R})^{\dagger} \tilde{\mathbf{X}}_{JPFK \times N}$
5. Return to Step 2 until convergence.
6. Eliminate the scaling ambiguity
   $\hat{\mathbf{S}}_{(final)} = \hat{\mathbf{S}}_{(\infty)} \left[D_1(\hat{\mathbf{S}}_{(\infty)}) \right]^{-1}$
7. Project the estimated symbols onto the symbol alphabet.

TSTF coding system (8)

Diversity gain and transmission rate (Costa,Favier; submitted to Elsevier SP 2017)

The performance analysis is based on the pairwise error probability (PEP) of the maximum likelihood (ML) estimator of the symbol matrix $\mathbf{S}$.

The diversity gain is defined as the negative of the asymptotic slope of the plot PEP(ρ) on a log-log scale, where ρ denotes the received SNR.

TSTF coding system (9)

Diversity gain and transmission rate

Define $\alpha^{(f,p)}$ and $\beta^{(f,p)}$ as the numbers of transmit antennas used and of data streams transmitted with the subcarrier f, during the time block p.

For a full allocation strategy: $\alpha^{(f,p)} = M$, $\beta^{(f,p)} = R$, for all (f, p).

- Maximal diversity gain: $KJ \sum_{f=1}^{F} \sum_{p=1}^{P} \min\left(\alpha^{(f,p)}, \beta^{(f,p)}\right)$.
- Transmission rate (in bits per channel use): $T_r = \frac{R}{FP} \log_2(\mu)$ where μ is the cardinality of the symbol constellation.

Tensor modeling of MIMO communication systems

Comparison of tensor-based systems (Costa, Favier; submitted to SP'2017)

Simulation results

- Comparison of ZF receivers (channel perfectly known at the receiver) for ST[(1)], TST, STF[(2)], and TSTF systems, with full allocation and same product $FP = 8 \Rightarrow$ same transmission rate (1 bit/channel use).
- M=K=R=2, N=10, 16-PSK.

  - (1) de Almeida, Favier, Mota, Space-time spreading-multiplexing for MIMO wireless communication systems using the PARATUCK-2 tensor model, Signal Process. 89(11):2103-2116, Nov. 2009.
  - (2) de Almeida, Favier, Ximenes, Space-time-frequency (STF) MIMO communication systems with blind receiver based on a generalized PARATUCK2 model, IEEE TSP 61(8):1895-1909, April 2013.

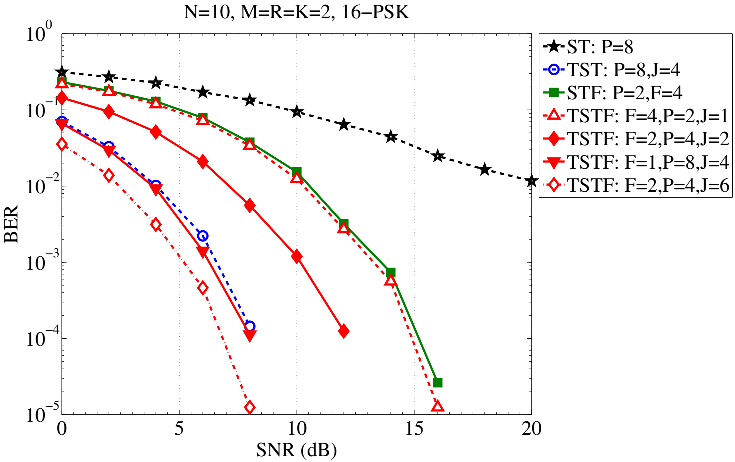


Figure: ZF receivers: Impact of the design parameters (F, P, J) on the BER.

Tensor modeling of MIMO communication systems

Comparison of tensor-based systems (7)

Comments on simulation results

- Worst BER performance with ST, due to the smallest diversity gain.
- TSTF with $(F, P, J) = (4, 2, 1)$ and STF with $(F, P) = (4, 2)$ give nearly the same BER.
- TSTF with $(F, P, J) = (1, 8, 4)$ and TST with $(P, J) = (8, 4)$ provide close BERs, with $FPJ = 32$, explaining the BER improvement.
- Impact of FPJ on the BER performance, i.e. the diversity gain: TSTF with $(F, P, J) = \{(4, 2, 1), (2, 4, 2), (1, 8, 4), (2, 4, 6)\}$, corresponding to $FPJ = \{8, 16, 32, 48\}$ $\Rightarrow$ Best performance with $(F, P, J) = (2, 4, 6)$ corresponding to $FPJ = 48$.

TSTF allows more flexibility for choosing the design parameters and best performance, due to the fifth-order coding tensor which exploits four spreading dimensions (space, frequency, time, chip).

Tensor-based relaying communication systems (1)

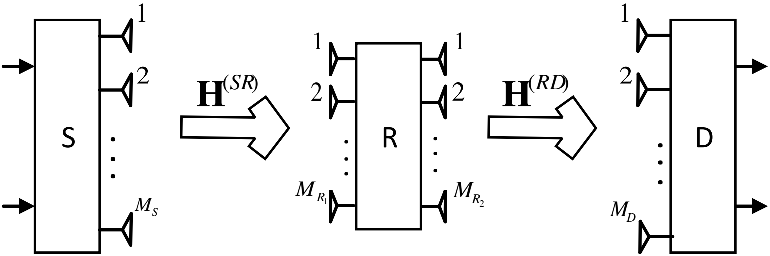


Figure: Block diagram of one-way two-hop MIMO relay system, with AF protocol.

- M_D, (M_{R_1}, M_{R_2}), and M_S antennas at destination (D), relay (R) and source (S).
- $\mathbf{H}^{(SR)}, \mathbf{H}^{(RD)}$: source-relay, and relay-destination channels.
- Symbol $\mathbf{S} \in \mathbb{C}^{N \times M_S}$ and code ($\mathbf{C} \in \mathbb{C}^{P \times M_S}$ and $\mathbf{G} \in \mathbb{C}^{J \times M_R}$) matrices.

Tensor-based relaying communication systems (2)

Three different tensor models and semi-blind receivers depending on source and relay codings:
(AF protocol)

KRST source coding ⇒ PARATUCK model (IEEE TSP 2014).

Double KRST coding ⇒ Nested PARAFAC model (IEEE TSP 2015).

Double TST coding ⇒ Nested Tucker model (Elsevier SP 2016).

Nested PARAFAC MIMO relay system (1)

(Ximenes, Favier, de Almeida; IEEE TSP'2015)

KRST coding (without precoding) at the source and the relay
$(M_{R1} = M_{R2} = M_R)$

Signals received at antenna m_R of relay

$$y_{m_R,p,n} = \sum_{m_S=1}^{M_S} h^{(SR)}_{m_R,m_S} c_{p,m_S} s_{n,m_S} \Leftrightarrow \mathrm{PARAFAC}(\mathbf{H}^{(SR)}, \mathbf{C}, \mathbf{S})$$

Signals received at antenna m_D of destination node

$$x^{(SRD)}_{m_D,j,p,n} = \underbrace{\sum_{m_R=1}^{M_R} h^{(RD)}_{m_D,m_R} g_{j,m_R} y_{m_R,p,n}}_{\mathrm{PARAFAC}(\mathbf{H}^{(RD)},\mathbf{G},\mathbf{Y}_{PN\mathrm{x}M_R})} = \underbrace{\sum_{m_R}\sum_{m_S} h^{(RD)}_{m_D,m_R} g_{j,m_R} h^{(SR)}_{m_R,m_S} c_{p,m_S} s_{n,m_S}}_{\text{Nested PARAFAC}}$$

Nested PARAFAC MIMO relay system (2)

$$x^{(SRD)}_{m_D,j,p,n} = \sum_{m_S} \underbrace{\sum_{m_R} h^{(RD)}_{m_D,m_R} g_{j,m_R} h^{(SR)}_{m_R,m_S}}_{z_{m_D,j,m_S}} c_{p,m_S} s_{n,m_S}$$

$$= \underbrace{\sum_{m_S} z_{m_D,j,m_S} c_{p,m_S} s_{n,m_S}}_{\text{PARAFAC}(\mathbf{Z}_{M_D J \times M_S},\mathbf{C},\mathbf{S})} \Rightarrow (\hat{\mathcal{Z}}, \hat{\mathbf{S}});\ \mathcal{Z} = \text{effective channel}$$

$$z_{m_D,j,m_S} = \underbrace{\sum_{m_R} h^{(RD)}_{m_D,m_R} g_{j,m_R} h^{(SR)}_{m_R,m_S}}_{\text{PARAFAC}(\mathbf{H}^{(RD)},\mathbf{G},\mathbf{H}^{(SR)^T})} \Rightarrow (\hat{\mathbf{H}}^{(RD)}, \hat{\mathbf{H}}^{(SR)})$$

Nested PARAFAC MIMO relay system (3)

Channel and symbol estimation
C and **G** assumed to be known at destination
Two solutions

Solution based on PARAFAC model of $\mathcal{Z}$

1. PARAFAC($\mathbf{Z}_{M_D JxM_S}$,**C**,**S**) of $\mathcal{X}^{(SRD)}$ $\overset{\text{Two-step ALS}}{\rightarrow}$ $(\hat{\mathcal{Z}},\hat{\mathbf{S}})$.
2. PARAFAC($\mathbf{H}^{(RD)}$,**G**,$\mathbf{H}^{(SR)^T}$) of $\mathcal{Z}$ $\overset{\text{Two-step ALS}}{\rightarrow}$ $(\hat{\mathbf{H}}^{(SR)},\hat{\mathbf{H}}^{(RD)})$.

Solution based on PARAFAC model of $\mathcal{Y}$

1. PARAFAC($\mathbf{H}^{(RD)}$,**G**,$\mathbf{Y}_{PNxM_R}$) of $\mathcal{X}^{(SRD)}$ $\overset{\text{Two-step ALS}}{\rightarrow}$ $(\hat{\mathbf{H}}^{(RD)}, \hat{\mathcal{Y}})$.
2. PARAFAC($\mathbf{H}^{(SR)}$,**C**,**S**) of $\mathcal{Y}$ $\overset{\text{Two-step ALS}}{\rightarrow}$ $(\hat{\mathbf{H}}^{(SR)},\hat{\mathbf{S}})$.

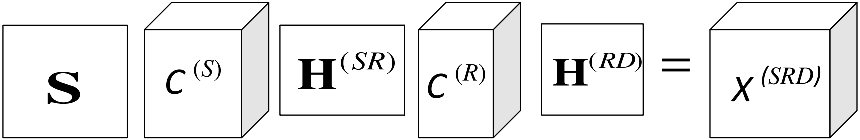
Nested PARAFAC MIMO relay system (4)

- Rewriting the nested PARAFAC model in terms of the tensors $\mathcal{Z}$ or $\mathcal{Y}$ which satisfy themselves two PARAFAC models, allows to estimate the system's parameters in two stages using a two-step ALS algorithm at each stage.
- **G** and **C** being assumed to be known at destination, each PARAFAC model contains only two unknown matrix factors ⇒ estimation can be solved using a closed-form (SVD-based) solution at each stage.

Nested Tucker MIMO relay system (1)

(Favier, Fernandes, de Almeida; SP 2016)

TST coding at the source and the destination



$$x^{(SRD)}_{m_D,j,p,n} = \sum_{m_{R_2}=1}^{M_{R_2}} \sum_{m_{R_1}=1}^{M_{R_1}} \sum_{m_S=1}^{M_S} \sum_{r=1}^{R} h^{(RD)}_{m_D,m_{R_2}} \underbrace{c^{(R)}_{m_{R_2},j,m_{R_1}}}_{\text{Relay code}} h^{(SR)}_{m_{R_1},m_S} \underbrace{c^{(S)}_{m_S,p,r}}_{\text{Source code}} s_{n,r}$$

$\Rightarrow$ Nested Tucker model $\Rightarrow$ Semi-blind ALS-based receiver for joint estimation of symbols ($\mathbf{S}$) and channels ($\mathbf{H}^{(SR)}, \mathbf{H}^{(RD)}$).

Nested Tucker MIMO relay system (2)

$$\mathcal{X}^{(SRD)} \in \mathbb{C}^{M_D \times J \times P \times N}$$

Matrix unfoldings

$$\begin{aligned}
\mathbf{X}^{(SRD)}_{JPN \times M_D} &= (\mathbf{I}_J \otimes (\mathbf{I}_P \otimes \mathbf{S})\mathbf{C}^{(S)}_{PR \times M_S} \mathbf{H}^{(SR)\,T})\mathbf{C}^{(R)}_{JM_{R_1} \times M_{R_2}} \mathbf{H}^{(RD)\,T} \\
\mathbf{X}^{(SRD)}_{PJM_D \times N} &= (\mathbf{I}_P \otimes (\mathbf{I}_J \otimes \mathbf{H}^{(RD)})\mathbf{C}^{(R)}_{JM_{R_2} \times M_{R_1}} \mathbf{H}^{(SR)})\mathbf{C}^{(S)}_{PM_S \times R} \mathbf{S}^T \\
\mathbf{x}^{(SRD)}_{PNJM_D} &= \mathrm{vec}(\mathbf{X}^{(SRD)}_{JM_D \times PN}) \\
&= ((\mathbf{I}_P \otimes \mathbf{S}) \otimes (\mathbf{I}_J \otimes \mathbf{H}^{(RD)}))(\mathbf{C}^{(S)}_{PR \times M_S} \otimes \mathbf{C}^{(R)}_{JM_{R_2} \times M_{R_1}})\mathrm{vec}(\mathbf{H}^{(SR)})
\end{aligned}$$

Define the noisy received signals tensor as: $\tilde{\mathcal{X}} = \mathcal{X} + \mathcal{N}$, where $\mathcal{N}$ is the additive noise tensor.

Nested Tucker MIMO relay system (3)

Three-step ALS receiver

1. Initialization (it=0): randomly draw $\hat{\mathbf{H}}_0^{(SR)}$ and $\hat{\mathbf{S}}_0$ from symbol alphabet
2. $it = it + 1$.
3. Calculate the LS estimate of the channel $\mathbf{H}^{(RD)}$

$$(\hat{\mathbf{H}}_{it}^{(RD)})^T = \Big((\mathbf{I}_J \otimes (\mathbf{I}_P \otimes \hat{\mathbf{S}}_{it-1})\mathbf{C}^{(S)}_{PR\times M_S}(\hat{\mathbf{H}}_{it-1}^{(SR)})^T)\mathbf{C}^{(R)}_{JM_{R_1}\times M_{R_2}} \Big)^{\dagger} \tilde{\mathbf{X}}^{(SRD)}_{JPN\times M_D}.$$

4. Calculate the LS estimate of the symbol matrix

$$\hat{\mathbf{S}}_{it}^T = \Big((\mathbf{I}_P \otimes (\mathbf{I}_J \otimes \hat{\mathbf{H}}_{it}^{(RD)})\mathbf{C}^{(R)}_{JM_{R_2}\times M_{R_1}} \hat{\mathbf{H}}_{it-1}^{(SR)})\mathbf{C}^{(S)}_{PM_S\times R} \Big)^{\dagger} \tilde{\mathbf{X}}^{(SRD)}_{PJM_D\times N}.$$

5. Calculate the LS estimate of the channel $\mathbf{H}^{(SR)}$ from $\mathrm{vec}(\tilde{\mathbf{X}}^{(SRD)}_{JM_D\times PN})$
6. Return to Step 2 until convergence.
7. Eliminate the scaling ambiguities.
8. Project the estimated symbols onto the symbol alphabet.

Nested Tucker MIMO relay system (4)

System identifiability and ambiguities elimination

Necessary identifiability conditions

$$J \geq \max(\frac{M_{R_2}}{M_{R_1}}, \frac{M_{R_1}}{M_{R_2}}), \; P \geq \max(\frac{R}{M_S}, \frac{M_S}{R}), \; M_D \geq M_{R_2}, \; N \geq R, \; PN \geq M_{R_1}$$

Equations for elimination of ambiguities

$$\hat{\mathbf{S}}_{\text{final}} = \frac{\hat{\mathbf{S}}_{\infty}}{\hat{\mathbf{s}}_{1,1}}, \; \hat{\mathbf{H}}^{(RD)}_{\text{final}} = \frac{\hat{\mathbf{H}}^{(RD)}_{\infty}}{\hat{h}^{(RD)}_{1,1}}, \; \hat{\mathbf{H}}^{(SR)}_{\text{final}} = \hat{\mathbf{s}}_{1,1} \hat{h}^{(RD)}_{1,1} \hat{\mathbf{H}}^{(SR)}_{\infty}.$$

Nested Tucker MIMO relay system (5)

Closed-form receiver (Known coding tensors)

- Two stages of Kronecker product approxim. for estimating $\mathbf{S}$ and $\mathbf{H}^{(RD)}$.
- One LS stage for estimating $\mathbf{H}^{(SR)}$.

$$\begin{aligned} \mathbf{X}^{(SRD)}_{M_D PN \times J} &= (\mathbf{H}^{(RD)} \otimes \mathbf{V})\mathbf{C}^{(R)}_{M_{R_2} M_{R_1} \times J} \Rightarrow (\hat{\mathbf{H}}^{(RD)}, \hat{\mathbf{V}}) \\ \mathbf{X}^{(SRD)}_{NJM_D \times P} &= (\mathbf{S} \otimes \mathbf{W})\mathbf{C}^{(S)}_{RM_S \times P} \Rightarrow (\hat{\mathbf{S}}, \hat{\mathbf{W}}) \end{aligned}$$

$$\begin{aligned} \mathbf{V} &= (\mathbf{I}_P \otimes \mathbf{S})\mathbf{C}^{(S)}_{PR \times M_S} \mathbf{H}^{(SR)\,T} \in \mathbb{C}^{PN \times M_{R_1}} \Rightarrow \hat{\mathbf{H}}^{(SR)} \\ \mathbf{W} &= (\mathbf{I}_J \otimes \mathbf{H}^{(RD)})\mathbf{C}^{(R)}_{JM_{R_2} \times M_{R_1}} \mathbf{H}^{(SR)} \in \mathbb{C}^{JM_D \times M_S} \Rightarrow \hat{\mathbf{H}}^{(SR)} \end{aligned}$$

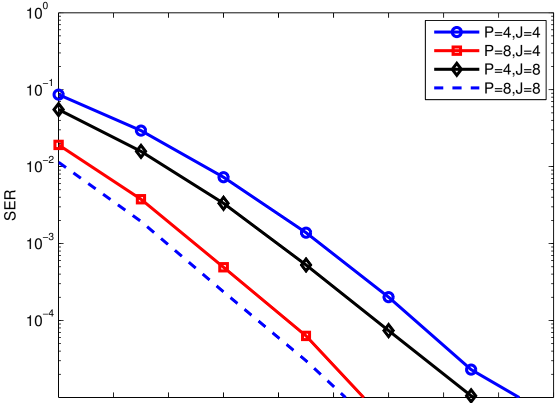
Nested Tucker MIMO relay system (6)

Simulations

ZF performance for different values of P and J

(diversity gain proportional to PJ)

BER $\searrow$ when PJ $\nearrow$

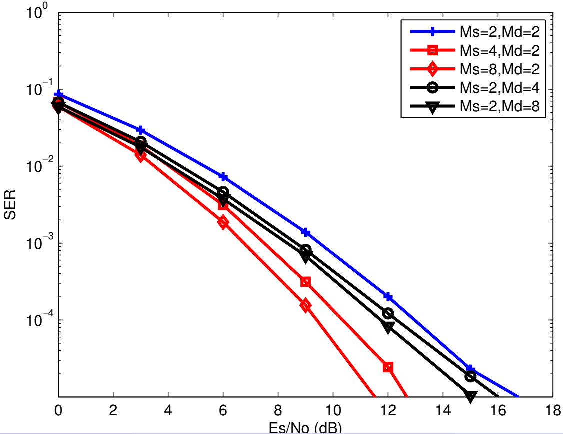


Nested Tucker MIMO relay system (7)

Simulations

ZF performance for different values of M_S and M_D

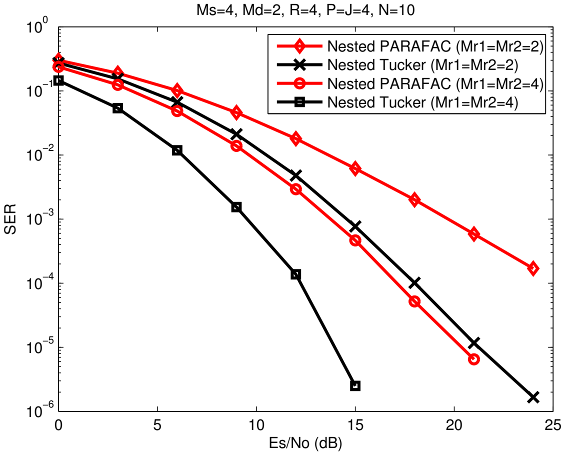
$\Rightarrow$ $M_S \nearrow$ better than $M_D \nearrow$



Nested Tucker MIMO relay system (8)

Simulations

Comparison of Nested Tucker and nested PARAFAC systems



Global design procedure (1)

- Design of the transmission system:
  Choice of coding, allocations, modulation (CDMA/OFDM/CDMA-OFDM), symbol constellation, (QAM, PSK), relaying protocol
  $\Rightarrow$ Tensor modeling of transmitted signals.
- Channel model
  $\Rightarrow$ Tensor modeling of received signals.
- Theoretical performance analysis: determination of diversity gain and transmission rate.
- Study of uniqueness of the tensor model of received signals, and determination of ambiguity relations.

Global design procedure (2)

- Determination of matrix unfoldings of the received signals tensor.
- Design of receivers:
  - ZF
  - Iterative semi-blind (ALS, Levenberg-Marquardt...)
  - Closed-form semi-blind (based on Khatri-Rao or Kronecker product)
- Study of parameter identifiability depending on the receiver.
- Experimental evaluation:
  - Test of BER performance in the case of perfect knowledge of channel (with ZF receiver)
  - Test of BER, convergence speed, computational time in the case of joint channel/symbols estimation.

Conclusion

Benefits of tensor models

Tensor models are very useful for:

- Representing, analysing and estimating multidimensional signals/data,
- Modeling and designing MIMO communication systems,
- Joint semi-blind estimation of symbols and channels in cooperative relay systems.

Tensor representations are particularly interesting when a tensor model is underlined in data as in communication systems.

Future works

- Development of new tensor models and study of their uniqueness and identifiability properties. Parameter estimation algorithms.
- Tensor-based multi-hop cooperative relay systems.
- Tensor completion (Estimation of data tensors with missing data):
  - Different solutions depending on the model used for representing the data tensor (PARAFAC, Tucker, TT...), the criterion to minimize, the choice of the modal projectors, the algorithm for optimization.
  - Applications to traffic data, sparse channel estimation for massive MIMO systems...

REFERENCES (1)

- N. D. Sidiropoulos, G. B. Giannakis, and R. Bro, "Blind PARAFAC receivers for DS-CDMA systems," *IEEE Trans. Signal Process.*, vol. 48, no. 3, pp. 810–823, March 2000.
- N. D. Sidiropoulos, and R.S. Budampati, "Khatri-Rao space-time codes," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2396–2407, 2002.
- A. L. F. de Almeida, G. Favier, and J. C. M. Mota, "PARAFAC-based unified tensor modeling for wireless communication systems with application to blind multiuser equalization," *Signal Processing*, vol. 87, pp. 337–351, 2007.
- A. L. F. de Almeida, G. Favier, and J. C. M. Mota, "A constrained factor decomposition with application to MIMO antenna system," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2429–2442, 2008.
- G. Favier, M. N. da Costa, A. L. F. de Almeida, and J. M. T. Romano, "Tensor space-time (TST) coding for MIMO wireless communication systems," *Signal Processing*, vol. 92, no. 4, pp. 1079–1092, 2012.
- A. L. F. de Almeida and G. Favier, Double Khatri-Rao space-time-frequency coding using semi-blind PARAFAC based receiver, IEEE Signal Processing Letters, 20 (5), 471-474, May 2013.

REFERENCES (2)

- G. Favier, and A.L.F. de Almeida, Tensor space-time-frequency coding with semi-blind receivers for MIMO wireless communication systems, IEEE Trans. Signal Process., 62 (22), 5987-6002, Nov. 2014.
- G. Favier, and A. L. F. de Almeida, Overview of constrained PARAFAC models, EURASIP J. on Advances in Signal Process., 62 (14), Sept. 2014.
- L.R. Ximenes, G. Favier, and A.L.F. de Almeida, Semi-blind receivers for non-regenerative cooperative MIMO communications based on nested PARAFAC modeling, IEEE Trans. Signal Process., 63 (18), 4985-4998, Sept. 2015.
- L.R. Ximenes, G. Favier, and A.L.F. de Almeida, Closed-form semi-blind receiver for MIMO relay systems using double Khatri-Rao space-time coding, IEEE Signal Process. Letters, 23 (3), 316-320, March 2016.
- G. Favier, C.A.R. Fernandes, and A.L.F. de Almeida, Nested Tucker tensor decomposition with application to MIMO relay systems using tensor space-time-frequency coding (TSTC). Signal Process., 128, 318-331, 2016.

End

Thank you for your attention

List of publications available at
http : //www.i3s.unice.fr/Gerard$_F$avier/

Copy of papers available at
https : //www.researchgate.net/profile/Gerard$_F$avier