Universidade de Brasília
Faculdade de Tecnologia
Departamento de Engenharia Mecânica

Decomposição tensorial aninhada acoplada aplicada a sistemas de comunicação MIMO duplamente polarizado

Maria de Fátima Kallynna Bezerra Couras

TESE DE DOUTORADO

PROGRAMA DE PÓS-GRADUAÇÃO EM SISTEMAS MECATRÔNICOS

Brasília

2023

Universidade de Brasília
Faculdade de Tecnologia
Departamento de Engenharia Mecânica

Coupled Nested Tensor Decomposition applied to Dual-Polarized MIMO Communication Systems

Maria de Fátima Kallynna Bezerra Couras

Tese de Doutorado submetida ao Departamento de Engenharia Mecânica da Faculdade de Tecnologia da Universidade Brasília como parte dos requisitos necessários para a obtenção do grau de Doutor.

Orientador: Prof. Dr.-Ing. João Paulo Javidi da Costa

Brasília
2023

Bezerra Couras, Maria de Fátima Kallynna.

B769d Decomposição tensorial aninhada acoplada aplicada a sistemas de comunicação MIMO duplamente polarizado / Maria de Fátima Kallynna Bezerra Couras; orientador João Paulo Javidi da Costa . -- Brasília, 2023.

140 p.

Tese de Doutorado (Programa de Pós-Graduação em Sistemas Mecatrônicos) -- Universidade de Brasília, 2023.

1. Decomposição Tensorial. 2. Produto de Kronecker. 3. Modelagem do canal. 4. Receptores semi-cegos. I. , João Paulo Javidi da Costa, orient. II. Título

Universidade de Brasília
Faculdade de Tecnologia
Departamento de Engenharia Mecânica

Decomposição tensorial aninhada acoplada aplicada a sistemas de comunicação MIMO duplamente polarizado

Maria de Fátima Kallynna Bezerra Couras

Tese de Doutorado submetida ao Departamento de Engenharia Mecânica da Faculdade de Tecnologia da Universidade Brasília como parte dos requisitos necessários para a obtenção do grau de Doutor.

Trabalho aprovado. Brasília, 27 de junho de 2023:

Prof. Dr-Ing. João Paulo Javidi da Costa
Orientador

Prof. Dr. Ricardo Zelenovsky
Examinador interno

Prof. Dr. José Alfredo Ruiz Vargas
Examinador interno

Prof. Dr. Walter da Cruz Freitas Júnior, UFC/PPGETI
Examinador externo

Prof. Dr. Tarcisio F. Maciel, UFC/PPGETI
Examinador externo

Brasília
2023

To my husband Pablo and my parents Pastoura and Francisco (in memorian).

Acknowledgements

I would like to thank all the people who contributed to the development of this Thesis, especially:

To my Advisors Prof. Dr. Gérard Favier, Prof. Dr. Vicente Zarzoso, Prof. Dr. André Lima Ferrer de Almeida and Prof. Dr. João Paulo Javidi da Costa.

Especially my thanks to prof. Gérard Favier, for his mentorship and never-ending supply of fascinating tasks. His humble approach to research and love to tensors is an inspiration. This approach is evident in his simple but obvious writing style, which I aspire to emulate throughout my career. Thank you so much for everything that taught me about tensors, because everything I learned I owe to you.

My dear husband Pablo Henrique Ursulino de Pinho all my thanks and gratitude. For staying with me in this dream of becoming PhDs and for all the support at crucial moments in our personal and academic life.

To my parents Francisco Bezerra Couras (in memorian) and Pastoura Gomes Bezerra for granting me the gift of life together with God. To my brothers Vicente Bezerra Couras and Carlos Alberto Bezerra Couras for their support at different times.

My dear University of Brasilia - UnB and Université Côte D'Azur and in particular to the I3S laboratory for the technical support given to the development of this work.

To my friends from Brasilia especially Jacinto Bezerra Gomes, Sílvia Martins, Gustavo Martins, João Batista Martins, Gileno Santos, Italene Aparecida, Pedro Lucas and João Paulo for a beautiful and true friendship.

To my friends in Brazil, Layne Leal, Luciano Monteiro, Evandson Dantas, Alisson Calvante, Ellen Castro, Valmir Silvino, and Aldenir Sudario, for the true feeling of friendship.

To a great friend and one of the greatest supporters of my doctorate and my husband's at the University of Brasília-UnB, Dr.-Ing. Ricardo Kehrle Miranda.

I am also grateful to the LASP laboratories, the University of Brasília-UnB and GTEL, Federal University of Ceará-UFC for their technical support and to my graduating program in Mechatronic systems - PPMEC / UnB.

To the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-CAPES and Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq for the scholarship granted.

To my Parents-in-law, Edilson Ursulino de Morais (in memorian) Leny Campelo de Pinho. And to my brother-in-law Edilson Ursulino de Morais Filho and my sisters-in-law

Fernanda Brígida Ursulino de Pinho. To my nephews and nieces, Ianca Lavigne, Carlos Daniel, Karla Vitorya, Davi Pinho, Igo Alexandre and Sophia Morais.

To my grandparents (in memorian) Vicente Gomes da Silva, Raimunda Bezerra Feitosa, Ricarte Bezerra Couras and Expedita Bezerra Pereira.

To my dear professors Dr. João César Mota and Dr. Walter Freitas for having been the main collaborators of this doctorate in a thesis cotutela regime with Université Côte D'Azur, in Nice, France.

“Logic will get you from A to Z.
Imagination will get you anywhere.”
(Albert Einstein)

Abstract

In recent years, massive Multiple-Input-Multiple-Output (MIMO) systems have been the subject of intense research due to their great potential to provide energy efficiency, data rate gains and diversity. Semi-blind receivers based on tensor decompositions for MIMO massive systems have been extensively studied in this context. These receivers allow us a better estimate of the channel and symbols without any information about the channel. This thesis presents a received signal model based on tensor decompositions that combine an extension of the multiple Kronecker space-time (MKronST) coding and fifth-order channel tensor to transmit the symbols. The coding extension is based on combining the tensor space-time (TST) coding and the multiple symbol matrices Kronecker product (MSMKron), called TST-MSMKron coding. The channel assumes a uniform rectangular array (URA) at both transmitter and receiver, allowing us to model the channel as a tensor. More specifically, the theoretical contributions of this thesis are the proposal of new semi-blind receivers to jointly estimate the symbol matrices, channel and channel parameters without prior knowledge. In the first part of this thesis, a decomposition based on column selection, called multidimensional CX decomposition, for tensors is proposed and one algorithm is presented to estimate and reconstruct the data tensor. In the second part, the TST-MSMKron coding is presented for massive MIMO systems, where a model of the received signal is proposed that combines a fifth-order channel with the TST-MSMKron coding. This system allows us to model the received signal as a coupled-nested-Tucker-PARAFAC (parallel factor analysis). In addition, two-step semi-blind receivers are proposed to jointly estimate the symbols, the channel and the channel parameters. In each part of the thesis, Monte Carlo simulations are provided to evaluate the performance of the proposed algorithms. The results show the efficiency of algorithms in the reconstruction of the data and joint estimation of the symbols, channel and channel parameters of the proposed system, respectively.

Keywords: Tensor Decomposition. Kronecker product. Channel modeling. Semi-blind receivers.

Resumo

Nos últimos anos, os sistemas MIMO (do inglês, *Multiple-Input-Multiple-Output*) massivos têm sido objeto de intensa pesquisa devido a seu grande potencial para fornecer eficiência energética, ganhos de taxa de dados e diversidade. Neste contexto, receptores semi-cegos baseados em tensores têm sido extensivamente estudados para sistemas MIMO. Esses receptores nos permitem uma estimativa melhor do canal e dos símbolos sem qualquer informação a respeito do canal. Esta tese apresenta um modelo de sinais recebidos com base em decomposições tensoriais que combinam uma extensão da codificação MKronST (do inglês, *multiple Kronecker space-time*) e um tensor do canal de ordem 5 para transmitir os símbolos. A extensão da codificação é baseada na combinação da codificação TST e codificação baseada nos múltiplos produtos de Kronecker das matrizes de símbolos (MSMKron), chamada de codificação TST-MSMKron. O canal é modelado com um arranjo uniforme retangular (do inglês, *uniform rectangular array*, URA) no transmissor e no receptor o que permite modelar o canal como um tensor. Mais especificamente, as contribuições teóricas desta tese estão em torno da proposição de novos receptores semi-cegos para estimar conjuntamente as matrizes de símbolo, os parâmetros de canal e o canal sem conhecimento a *priori* sobre eles. Na primeira parte desta tese é proposta uma decomposição baseada na seleção de colunas, chamada multidimensional CX, para tensores e um algoritmo é apresentado para estimar e reconstruir o tensor de dados. Na segunda parte desta tese, a codificação TST-MSMKron é apresentada para sistemas MIMO massivos, onde é proposto um modelo do sinal recebido que combina um canal de ordem 5 com a codificação TST-MSMKron. Este sistema nos permite modelar o sinal recebido como uma decomposição Nested-Tucker-PARAFAC (do ingês, *parallel factor analysis*). Além disso, os receptores semi-cegos em duas etapas são propostos para estimar em conjunto os símbolos, o canal e os parâmetros do canal. Em cada parte são fornecidas simulações de Monte Carlo para avaliar o desempenho dos algoritmos propostos. Os resultados mostram a eficiência dos algoritmos na reconstrução dos conjuntos de dados e estimativa conjunta dos símbolos, do canal e parâmetros do canal do sistema proposto, respectivamente.

Palavras-chave: Decomposição Tensorial. Produto de Kronecker. Modelagem do canal. Receptores semi-cegos.

List of Figures

List of Tables

List of abbreviations and acronyms

6G	sixth-generation	22
ALM	alternating least mean squares	88
ALS	alternating least squares	24
ATALS	accelerated trilinear alternating least squares	88
BCI	brain-computer interface	22
Bi-ALS	bi-alternating least-squares	26
BS	base station	76
CPD	canonical polyadic decomposition	34
CS	compressed/compressive sensing	62
CSI	channel state information	22
D-KRSTF	double-KRST	75
DD	double directional	23
DFT	discrete Fourier transform	90
DoA	direction-of-arrival	24
DoD	direction-of-departure	24
DP	double polarized	24
DS-CDMA	direct-sequence code division multiple access	74
ESPRIT	estimation of signal parameters via rotational invariance thecnique	24
FDD	frequency division duplex	74
GHz	gigahertz	23
GKRF	generalized KRF	45
GKronF	generalized KronF	45
H_R	horizontally	25
HOSVD	higher order singular value decomposition	45
IoE	internet of everything	22
IRS	intelligent reflecting surface	76
ITS-ALS	improved two-stage ALS	88

JIRAFE	joint dimensionality reduction and factor retrieval	24
KALMS	Kronecker alternating least mean squares	88
KALS	Kronecker ALS	88
KLS	Kronecker least squares	88
KPLS	Kronecker product least square	88
KRF	Khatri-Rao factorization	45
KRF	eigenvalue decomposition	58
KronF	Kronecker factorization	26
KRST	Khatri-Rao space time	74
KSVD	Kronecker SVD	88
LS	least square	49
MC	matrix completion	63
MIMO	multiple input multiple output	22
MKRF	multiple KRF	88
MKronST	multiple Kronecker space-time	24
mmWave	millimeter-wave	22
MSE	mean square error	28
MSMKron	multiple symbol matrices Kronecker coding	27
MUSIC	multiple signal classification	24
NMSE	normalized mean square error	28
OSTBC	orthogonal space-time block codes	76
PARAFAC	parallel factor analysis	33
QAM	quadrature amplitude modulation	99
RectALS	rectified alternating least squares	57
RIS	reconfigurable intelligent surface	111
SER	symbol error rate	28
SIMO	single input multiple output	74
SIP	shift invariance principle	58
SISO	single input single output	74
SKALS	supervised ALS	88

STF	space-time-frequency	23
SVD	singular value decomposition	34
TAS	transmit antenna selection	76
TC	tensor completion	63
TenRICE	tensor-based RIS channel estimation	88
THOSVD	truncated higher-order singular value decomposition	26
THz	terahertz	22
TR$_1$A	toeplitz rank-one approximation	58
TR$_1$A$_{imp}$	improved TR$_1$A	58
TST	tensor space-time	23
TSTF	tensor space-time-frequency	23
TT	tensor train	34
TT-HSVD	tensor train hierarchical singular value decomposition	45
TT-SVD	tensor train singular value decomposition	54
UAV	unmanned aerial vehicle	76
ULA	uniform linear array	24
UPA	uniform planar array	74
URA	uniform rectangular array	24
V$_R$	vertically	25
ZF	zero-forcing	93

List of symbols

$\circ$	outer product	38
$\delta_{r_1,\ldots,r_N}$	generalized Kronecker delta	33
$\diamond$	Khatri-Rao product	30
$\hat{\hat{\mathbf{X}}}$	matrix $\hat{\mathbf{X}}$ after the correction of ambiguities	29
$\hat{\mathbf{X}}$	estimate of $\mathbf{X}$	29
$\Lambda^{(n)}$	scaling matrix n	36
$\lambda^{(n)}$	n-th scaling element	46
$\mathbb{C}$	set of complex-valued numbers	29
$\mathbb{C}^{I \times J}$	set of complex-valued of size $I \times J$	29
$\mathbb{C}^{I_1 \times \ldots \times I_N}$	set of complex-valued of size $I_1 \times \ldots \times I_N$	29
$\mathbb{R}$	set of real-valued numbers	29
$\mathbb{R}^I$	set of real-valued I-dimensional vectors	29
$\mathbf{\Pi}$	permutation matrix	36
$\mathcal{I}_{R,N}$	identity tensor of N-th order and R-rank	29
$\mathcal{X}$	calligraphic letters to represent tensors	29
$\mathrm{diag}(\cdot)$	diagonal matrix from its vector argument	29
$\mathrm{unvec}(\cdot)$	unvectorization operator (inverse of the vectorization operator)	29
$\mathrm{vec}(\cdot)$	vectorization operator	29
$\min_{x \in \mathbb{C} f(x)}$	minimum argument $x \in \mathbb{C}$ of a function $f(x)$	46
$\odot$	Hadamard product	29
$\otimes$	Kronecker product	30
$\lVert \cdot \rVert_*$	nuclear norm	29
$\lVert \cdot \rVert_2$	l_2-norm	29
$\lVert \cdot \rVert_F$	Frobenius norm	29
$\mathbf{1}_R$	column vector with length R with all elements equal to 1	29
$\mathbf{e}_n^{(N)}$	n-th canonical basis vector of the Euclidian space $\mathbb{R}^N$	48
$\mathbf{I}_R$	identity matrix of size $R \times R$	29

$\mathbf{X}$	boldface upper-case to represent matrices	29
$\mathbf{x}$	boldface lower-case to represent vectors	29
$\mathbf{X}^*$	complex conjugate matrix	29
$\mathbf{X}^{-1}$	inverse matrix	29
$\mathbf{X}^\dagger$	Moore-Penrose pseudo-inverse matrix	29
$\mathbf{X}^H$	Hermitian matrix operator	29
$\mathbf{X}^T$	transpose matrix	29
$\mathbf{X}_{..k}$	k-th frontal slice of $\mathbf{X} \in \mathbb{C}^{I \times J \times K}$	29
$\mathbf{X}_{.j.}$	j-th lateral slice of $\mathbf{X} \in \mathbb{C}^{I \times J \times K}$	29
$\mathbf{X}_{.j}$	j-th column of $\mathbf{X} \in \mathbb{C}^{I \times J}$	29
$\mathbf{X}_{i..}$	i-th horizontal slice of $\mathbf{X} \in \mathbb{C}^{I \times J \times K}$	29
$\mathbf{X}_{i.}$	i-th row of $\mathbf{X} \in \mathbb{C}^{I \times J}$	29
$\mathbf{X}_{I_n \times I_{n+1} \ldots I_N I_1 \ldots I_{n-1}}$	n-mode unfolding of $\mathcal{X} \in \mathbb{C}^{I_1 \times \ldots \times I_N}$	29
$\times_n$	n-mode product	33
$\times_n^m$	contraction operator	34
$D_i(\mathbf{X})$	diagonal matrix with diagonal entries given by i-th row of $\mathbf{X}$	29
$k_{\mathbf{X}}$	Kruskal-rank (k-rank) of $\mathbf{X}$	33
O	big O notation for complexity order	66
$r_{\mathbf{X}}$	rank of $\mathbf{X}$	29
x	lower-case to represent scalars	29
$x_{i,j}$	(i,j)-th element of $\mathbf{X} \in \mathbb{C}^{I \times J}$	29
$x_{i_1,\ldots,i_N}$	$(i_1,\ldots,i_N)$-th element of $\mathcal{X} \in \mathbb{C}^{I_1 \times \ldots \times I_N}$	29

Contents

1 Introduction

1.1 Thesis scope and motivation

The number of devices connected to the internet is exponentially increasing, and emerging technologies such as smart cities smart homes and internet of everything (IoE) are already a reality. Emerging applications such as streaming TV, extended reality services, telemedicine, wireless brain-computer interface (BCI) services, autonomous vehicles and systems, holographic and tactile communications, videoconferencing, intelligent transportation systems, among many others applications demand high throughput and broader coverage services. To withstand the challenges of new technologies, a sixth-generation (6G) wireless system needs to be developed with new attractive features such as high performance in terms of data rate, latency, reliability, coverage, massive and mobile connectivity, and energy efficiency [1, 2, 3].

6G wireless systems and networks must be smart and open to adapt to constantly changing services, which requires collaborative cognitive communication services, defined as the 3Cs [4]. 6G networks will be seen as a distributed, cell-less multiple input multiple output (MIMO) system, where each smart device will be served by more than one access point or a base station using multiplexing or transmission coordination that will be connected to the cloud to access cloud services. They will be able to use frequencies higher than in 5G networks, on the order of terahertz (THz), providing substantially higher capacity and speed with much lower latency, in addition to the combination of sub-millimeter-wave (mmWave) frequencies and the possibility of frequency selectivity in network deployment, leading to significant advances in user experience [4, 5]. The low latency of 6G networks will allow everything to be faster and more accessible, also enabling large data transmission, augmented reality, and autonomous operation of various equipment including cars. See [1, 4, 5] for more details about 6G networks. For the development of 6G wireless systems, many challenges must be faced.

In the case of wireless communication systems, satisfying these requirements greatly depends on an accurate channel state information (CSI) to compensate for the perturbations caused by wireless propagation on the transmitted signals. These perturbations include multipath propagation, reflections and diffractions from large obstacles, and fading effects. Therefore, the study of channel behavior is fundamental, where mathematical models that aim to describe the behavior of the communication channels and, consequently, their influence on the transmitted information are constantly exploited by the scientific community for designing precoders, directional and adaptive beamformers, and receivers adapted to each situation. In general, the receivers can be divided into two classes: supervised (or pilot-assisted) and unsupervised (or semi-blind) receivers. In the first case, training sequences known by the receiver are used to estimate the channel, which generally implies a significant training overhead due to pilots' transmission. Statistical models with time-varying random variables are used to characterize the channels. Besides the study of channel behavior, the development of techniques that avoid the deep fading of the signal before reaching its destination is needed. An alternative to deal with these effects is to exploit signal diversity, such as space, time, frequency, coding, and cooperation diversity [6].

During the last decades, tensor decompositions have been used to model received signals as well as to derive receivers to estimate the channel and symbols exploiting various forms of diversity. The use of tensor decompositions combined with new codings to transmit signals has been extensively proposed to introduce extra diversities for both point-to-point and multi-hop MIMO systems. Some examples are double Khatri-Rao space-time-frequency (STF) , tensor space-time (TST), tensor space-time-frequency (TSTF) codings [7, 8, 9], and coding based on multiple Khatri-Rao and Kronecker products of symbol matrices [10, 11, 12]. The use of these codings with tensor approaches for the design of the MIMO wireless communication systems has led to the development of new tensor models [8, 9, 13, 12, 14], and allows the proposition of semi-blind receivers for jointly estimating the transmitted symbols and channel parameters, in addition, to providing an increase in the transmission rate compared to the others codings [11]. In Chapter 4 a review of codings and MIMO communication system can be found, respectively.

Massive MIMO systems are part of current wireless systems [15]. They are very important in the development of the 6G wireless networks because they need to deal with massive volumes of data and very high data rate connectivity per device [2, 1]. In the last years, massive MIMO systems combined with mmWave communication systems have been the subject of intense research due to their great potential to improve system capacity and spectral efficiency. Such systems operate in the 30-300 gigahertz (GHz) spectrum with large antennas arrays at both the transmitter and receiver, and can achieve gigabit-per-second data rates [16, 17, 18]. Note that the sparsity of mmWave channels can be exploited in formulating the channel estimation as a sparse signal recovery problem which can be solved using compressed sensing methods, resulting in a significant reduction of the training overhead [19, 20, 21]. Some works have recently proposed supervised solutions for channel parameters' estimation based on a low-rank PARAFAC decomposition [22], as briefly reviewed hereafter [23, 24, 25].

3D channel modeling has been the subject of intensive research works [26], because it allows exploring the channel's degrees of freedom in the elevation direction, particularly useful for massive MIMO systems characterized by a large number of antennas. 3D channels models are in effect double directional (DD) MIMO channels [27], in sense that they include consideration of directions-of-arrival (DoAs) of the multipath components at both the base and mobile stations [28]. 3D MIMO channels also can consider using a double polarization leads to double polarizarized (DP) MIMO channels. DP antenna arrays, using horizontal and vertical polarizations, may offer large multiplexing (i.e. capacity) gains to MIMO systems by combining space and polarization diversities, as showed in [29]. However, estimating the key parameters of a DD-DP channel is a challenge due to the great number of channel parameters. The application of standard array processing algorithms (e.g., multiple signal classification (MUSIC) [30] and estimation of signal parameters via rotational invariance thecnique (ESPRIT) [31] algorithms) is a hard task due to restrictive assumptions about the system configuration in terms of numbers of multipath and transmit/receive antennas [25]. Tensor-based signal processing has shown to be a promising approach for future wireless communications systems due to the increase in dimensions, including space (antennas), time (snapshots and periods), code, frequency and polarization. Tensor-based signal processing techniques can naturally exploit the multidimensional structure of the data and its multiple forms of diversity, allowing to devise receiver algorithms with good performance-complexity tradeoffs while operating under more relaxed parameter choices than conventional matrix-based receivers [32].

Qian *et al.* [25] proposed a tensor-based approach for DP MIMO channel estimation by recasting the MIMO channel as a fourth-order tensor. The authors assumed a MIMO system with a uniform rectangular array (URA) at the transmitter and a (ULA) at the receiver. The identifiability of the channel parameters is discussed and a channel estimation algorithm based on alternating least squares (ALS) solution for estimating the steering and path-losses matrices combined with a closed-form method for the direction-of-departure (DoD) and direction-of-arrival (DoA) estimation is presented taking into account the Vandermonde structure of the steering matrices. Zniyed *et al.* [33] extended the fourth-order channel presented in [25] to a fifth-order channel tensor model. The authors assume a MIMO system with a URA at both the transmitter and the receiver and it is proposed the joint dimensionality reduction and factor retrieval (JIRAFE) algorithm for channel estimation by exploiting an equivalence between a PARAFAC model and a tensor train decomposition [33]. This method makes it possible to reduce the computational complexity and avoids ill-convergence problems, as highly time consuming and slowly convergence, linked to the application of the ALS algorithm to a N-order PARAFAC decomposition [34].

In this thesis, we address novel semi-blind receivers to jointly estimate symbols, channel, and channel parameters in DD-DP MIMO wireless communication systems. In particular, one of the main contributions of this thesis relies on the proposition of a new DD-DP MIMO system, equipped with URAs at both ends of the link, and combining multiple Kronecker products of symbol matrices with a TST coding, which we will call TST-MSMKron coding. The MSMKron coding is a particular case of the multiple Kronecker space-time (MKronST) coding presented in [10] for multi-hop MIMO relay systems. The channel tensor is represented by a fifth-order PARAFAC decomposition separated into two parts associated with the vertically (V_R) and horizontally (H_R) polarized receive antennas, respectively. Then, it is established that the tensors of received signals satisfy a new coupled nested Tucker-PARAFAC model, whose core tensor is the coding tensor. The coupling is due to the coding that is common to the V_R- and H_R-polarized receive antennas. The nested structure results from the fact that a matrix unfolding of the channel tensor constitutes one matrix factor of the Tucker model, the other factors being the steering and path-loss matrices.

Capitalizing on this new tensor model with the tensor coding knowledge, semi-blind receivers are derived for estimating the transmitted symbols, and channel parameters. These receivers are composed of two stages. In the first one, the information symbols and the channel are jointly estimated and in the second stage, from the estimated channel tensor, the channel parameters' estimation is performed. Due to the semi-blind nature of the proposed receivers, this works extends the approach of [25] in the sense that no pilot sequence is used for channel estimation, which allows to increase the transmission rate. In comparison with [33] which assumes that the channel is known a priori, the proposed receivers first jointly and semi-blindly estimate the channel and the information symbols before applying the JIRAFE method combined with a new rectification method exploiting the Vandermonde structure of the steering matrices, to estimate the channel parameters.

There is another contribution of this thesis. We propose a multidimensional column-space decomposition to perform a low-rank approximation of tensors based on the CX decomposition for matrices. By exploiting this new tensor model, an algorithm based on the l_2-norm to perform the approximation of the tensor data is proposed. To illustrate the performance of the CX-tensor algorithm, Monte Carlo simulations are performed.

1.2 Thesis organization

This thesis is divided into six chapters, including this introductory chapter. In the following, we briefly describe the content of the five remaining chapters.

Chapter 2: *Tensor Preliminaries*: This chapter presents a theoretical basis for the methods developed in this thesis. An important review of the definitions and operations of multilinear algebra is presented, where the notations and main matrix and tensor operations that will be used throughout the thesis are reviewed. Then, the main tensor decompositions are summarized, where the new coupled nested Tucker-PARAFAC model will be presented, as well as, its uniqueness property, thus introducing the first original contributions of the thesis. In addition to the tensor models, the algorithms used to estimate parameters in matrix and tensor decompositions are also presented.

Chapter 3: *CX decomposition for tensors*: This chapter presents the second contribution of this thesis. First, a multidimensional column-space decomposition to perform a low-rank approximation of tensors based on the CX decomposition for matrices is presented. Then in the second part, an algorithm based on probabilities of the columns of each n-mode matrix unfolding is presented to reconstruct the data tensor. The parameter identifiability conditions and computational complexity are analyzed. Finally, Monte Carlo simulation results are provided to illustrate the effectiveness of the proposed algorithm compared with the state of the art.

Chapter 4: *Dual-polarized MIMO communication systems using combined TST-MSMKron coding*: This chapter presents a bibliography review of existing MIMO communication systems based on tensor decompositions, as well as coding schemes and receivers. Furthermore, in the second part, the new coding scheme based on multiple Kronecker products of symbol matrices combined with TST coding is proposed to transmit the symbols and for this combined coding, a new system is provided based on the new tensor model. We extend the DD-DP channel model presented in [33] and combine it with TST-MSMKron coding to propose a new communication system.

Chapter 5: *Tensor-based semi-blind receivers for MIMO communication system using a combined TST-MSMKron coding*: In this chapter, the proposed DP MIMO system presented in Chapter 4 is explored to derive semi-blind receivers in two stages that jointly estimate the transmitted symbols and channel parameters (DoD and DoA angles, path-loss coefficients). In the first stage, the information symbols and the channel are jointly estimated using either a bi-alternating least-squares (Bi-ALS) algorithm followed by the Kronecker factorization (KronF) algorithm to separate the symbol matrices or the truncated higher-order singular value decomposition (THOSVD) method [35], which allows to directly and jointly estimate the symbol matrices and the channel in closed-form. In the second stage, from the estimated channel tensor, the channel parameters' estimation is performed using the JIRAFE algorithm combined with a new rectification method to take into account the Vandermonde structure of the steering matrices. Parameter identifiability and computational complexity of each receiver are analyzed. Monte Carlo simulation results are provided to illustrate the effectiveness of the proposed coding schemes and semi-blind receivers.

Chapter 6: *Conclusions and Perspectives*: In this chapter, the main conclusions about the contributions of this work are provided. Some advantages and limitations of the proposed systems and receivers are presented. In addition, some perspectives for future research are outlined.

1.3 Main original contributions

The different contributions of this thesis are distributed among into three chapters. Briefly, the main contributions of this thesis can be summarized as follows:

Chapter 2

- Review of basic operations of tensor decomposition and study of algorithms for matrix and tensor estimation;
- Proposition of the new tensor model named coupled nested Tucker-PARAFAC, which combines the existing Tucker-(N_1, N), and the PARAFAC model for higher-order tensors;
- Demonstration of the uniqueness of the new tensor model under certain conditions, based on the uniqueness of Tucker and PARAFAC models.

Chapter 3

- Proposition of the new tensor model named multidimensional CX decomposition, which is an extension of the CX decomposition of matrices for higher-order tensors and demonstration of the uniqueness of the new tensor model;
- The proposition of a new algorithm that performs the approximation of the data tensor based on the l_2-norm;
- Study of the performance of the proposed CX algorithm and the impact of the design parameters by means of extensive Monte Carlo simulations.

Chapter 4

- Proposition of the channel tensor model proposed in [33] and the combined TST-MSMKron (multiple symbol matrices Kronecker) coding used to encode the signals to be transmitted;
- Proposition of the new DD-DP MIMO communication systems, which uses the combined TST-MSMKron and the DD-DP channel tensor. It is established that the tensors of received signals satisfy a new coupled nested Tucker-PARAFAC model;
- Study of the uniqueness issues of the new system model.

Chapter 5

- The proposition of new semi-blind receivers composed of two stages to jointly estimate the symbols and the channel parameters (DoD and DoA angles, path-loss coefficients);
- Study of the identifiability conditions and computational complexity of each proposed semi-blind receiver;
- Study of the performance of the combined TST-MSMKron coding and the impact of the design parameters under the assumption of perfect channel knowledge, by means of extensive Monte Carlo simulations;

- Study of the proposed semi-blind receivers in terms of symbol error rate (SER) for symbols estimation, normalized mean square error (NMSE) for channel estimation, and mean square error (MSE) for channel parameters estimation.

1.4 Scientific production

This thesis has originated two journal papers and three publications at conferences. In the following, a list of publications is presented:

JOURNAL PAPERS:

- **M. F. K. B. Couras**, P. H. de Pinho, G. Favier, V. Zarzoso, de A. L. F. Almeida, and J. P. J. da Costa, Semi-blind receivers based on a coupled nested Tucker-PARAFAC model for dual-polarized MIMO systems using combined TST and MSMKron codings. **Digital Signal Processing - Elsevier**, p. 104043, 2023.
- P. H. de Pinho, **M. F. K. B. Couras**, G. Favier, de A. L. F. Almeida, J. P. J. da Costa, Semi-blind receivers for two-hop MIMO relay systems with a combined TSTF-MSMKron coding. **Sensors**, 2023, 23(13), 5963.

CONFERENCE PAPERS:

- **M. F. K. B. Couras**, P. H. de Pinho, G. Favier, J. P. J. da Costa, V. Zarzoso, and de A. L. F. Almeida, Multidimensional CX decomposition of tensors. In: **2019 Workshop on Communication Networks and Power Systems (WCNPS). IEEE**, p. 1-4, 2019.
- P. H. de Pinho, **M. F. K. B. Couras**, G. Favier, J. P. J. da Costa, de A. L. F. Almeida, J. P. A. Maranhão, Semi-supervised receivers for MIMO systems with multiple Khatri-Rao coding. In: **13th International Conference on Signal Processing and Communication Systems (ICSPCS)**. **IEEE**, p. 1-7, 2019.
- F. G. Constancio, **M. F. K. B. Couras**, D. X. Nogueira, J. P. J. da Costa, M. da R. Zanatta, R. T. de Sousa, and N. T. da Mota, Extended ADDIE model for improved distance learning courses. In: **2018 IEEE Frontiers in Education Conference (FIE)**, p. 1-5, 2018.

2 Tensor Preliminaries

In this chapter, we present a review of basic definitions and operations of multilinear algebra applied in this thesis. This chapter is divided into five sections. In Section 2.1, we define the notations. In Section 2.2 basic operations involving matrices are introduced as Khatri-Rao and Kronecker products. In Section 2.3, the basic operations involving tensors are defined, while the main tensor decompositions such as PARAFAC, Tucker, and their generalizations for N-order tensors, coupled-PARAFAC, HOSVD, and TT are presented in Section 2.4. These decompositions are frequently explored in the literature. The first contribution of this thesis, the coupled-nested Tucker-PARAFAC decomposition, is also introduced in Section 2.4, as well as its uniqueness property. Section 2.5 introduces additional algorithms for matrices and parameter estimation for matrix and tensor decompositions used throughout this thesis.

2.1 Notation

$\mathbb{R}$ and $\mathbb{C}$ denote the sets of real and complex numbers, respectively. Scalars, vectors, matrices and tensors are denoted by lower-case, boldface lower-case, boldface upper-case and calligraphic letters, e.g., x, $\mathbf{x}$, $\mathbf{X}$, $\mathcal{X}$, respectively. The operator diag$(\cdot)$ forms a diagonal matrix from its vector argument, while $D_k(\mathbf{C})$ forms a diagonal matrix holding the k-th row of $\mathbf{C} \in \mathbb{C}^{K\times K}$ on the diagonal. The transpose, complex conjugate, conjugate transpose and Moore-Penrose pseudo-inverse operators are represented by $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$ and $(\cdot)^\dagger$, respectively.

$\mathbf{e}_n^{(N)}$ represents the n-th canonical basis vector of the Euclidian space $\mathbb{R}^N$. Considering the matrix $\mathbf{A} \in \mathbb{C}^{I\times J}$, the vec$(\cdot)$ and unvec$(\cdot)$ operators are defined so that $\mathbf{a} = \text{vec}(\mathbf{A}) \in \mathbb{C}^{JI} \leftrightarrow \mathbf{A} = \text{unvec}(\mathbf{a}) \in \mathbb{C}^{I\times J}$. $\mathbf{A}_{.j}$ (respectively $\mathbf{A}_{i.}$) denotes the j-th column (respectively i-th row) of the matrix $\mathbf{A}$, while $\hat{\mathbf{A}}$ denotes an estimate of $\mathbf{A}$ and $\hat{\hat{\mathbf{A}}}$ represents the matrix $\hat{\mathbf{A}}$ after the correction of ambiguities. $\|\cdot\|_F$, $\|\cdot\|_*$ and $\|\cdot\|_2$ are used to represent the Frobenius norm, nuclear norm and the l_2-norm, respectively. The outer, Kronecker, Khatri-Rao, Hadamard and n-mode products are denoted by $\circ$, $\otimes$, $\diamond$, $\odot$ and $\times_n$, respectively.

$\mathcal{I}_{R,N}$, $\mathbf{I}_R$, $\mathbf{1}_R^T$, and $r_\mathbf{A}$ represent the identity tensor of order N and of size $R\times \ldots \times R$, the identity matrix of size $R\times R$, the all-ones row vector of dimensions $1\times R$, and the rank of the matrix $\mathbf{A}$, respectively. The i-th element of $\mathbf{a} \in \mathbb{C}^I$ is denoted by a_i, the (i,j)-th element of $\mathbf{A}$ is represented by $a_{i,j}$, and the $(i_1,\ldots,i_N)$-th element of $\mathcal{A} \in \mathbb{C}^{I_1\times\ldots\times I_N}$ is given by $a_{i_1,\ldots,i_N}$.

2.2 Matrix products and operations

In this section, the Kronecker and Khatri-Rao products are presented. These operations are very important in multilinear algebra.

Definition 1. (Kronecker product) *The Kronecker product of* $\boldsymbol{A} \in \mathbb{C}^{I\times M}$ *and* $\boldsymbol{B} \in \mathbb{C}^{J\times N}$ *is defined as:*

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{1,1}\mathbf{B} & a_{1,2}\mathbf{B} & \dots & a_{1,M}\mathbf{B} \\ a_{2,1}\mathbf{B} & a_{2,2}\mathbf{B} & \dots & a_{2,M}\mathbf{B} \\ \vdots & \vdots & \vdots & \vdots \\ a_{I,1}\mathbf{B} & a_{I,2}\mathbf{B} & \dots & a_{I,M}\mathbf{B} \end{bmatrix} \in \mathbb{C}^{IJ\times MN}. \tag{2.1}$$

Note that the Kronecker product between **A** and **B** has all possible combinations of products of the elements of **A** and **B**. To calculate the Kronecker product, the number of columns of the matrices does not need to be equal.

Definition 2. (Khatri-Rao product) *The Khatri-Rao product of* $\boldsymbol{A} \in \mathbb{C}^{I\times M}$ *and* $\boldsymbol{C} \in \mathbb{C}^{J\times M}$ *is defined as:*

$$\mathbf{A} \diamond \mathbf{C} = \begin{bmatrix} \mathbf{A}_{.1} \otimes \mathbf{C}_{.1} & \mathbf{A}_{.2} \otimes \mathbf{C}_{.2} & \dots & \mathbf{A}_{.M} \otimes \mathbf{C}_{.M} \end{bmatrix} \in \mathbb{C}^{IJ\times M}. \tag{2.2}$$

The Khatri-Rao product also can be seen as a column-wise Kronecker product. Note that the Khatri-Rao product of two matrices exists only if these matrices have the same number of columns. Another way to compute the Khatri-Rao product is given as:

$$\mathbf{A} \diamond \mathbf{C} = \begin{bmatrix} \mathbf{C}D_1(\mathbf{A}) \\ \vdots \\ \mathbf{C}D_I(\mathbf{A}) \end{bmatrix}. \tag{2.3}$$

Throughout this thesis, we shall make use of the following properties involving the Kronecker and Khatri-Rao products. For this, we consider the matrices $\mathbf{A} \in \mathbb{C}^{I\times M}$, $\mathbf{B} \in \mathbb{C}^{J\times N}$, $\mathbf{C} \in \mathbb{C}^{J\times M}$, $\mathbf{D} \in \mathbb{C}^{I\times N}$ and $\mathbf{E} \in \mathbb{C}^{M\times N}$.

Property 1.

$$(\mathbf{A} \otimes \mathbf{B})^H = \mathbf{A}^H \otimes \mathbf{B}^H, \tag{2.4}$$

$$(\mathbf{B} \diamond \mathbf{E})^H = \begin{bmatrix} D_1^*(\mathbf{B})\mathbf{E}^H & \dots & D_J^*(\mathbf{B})\mathbf{E}^H \end{bmatrix} \in \mathbb{C}^{N\times JM}, \tag{2.5}$$

where $D_j^*(\mathbf{B}) = \operatorname{diag}\begin{pmatrix} b_{j1}^* & \dots & b_{jN}^* \end{pmatrix}$.

Property 2.

$$(\mathbf{A} \diamond \mathbf{C})^H (\mathbf{D} \diamond \mathbf{B}) = \mathbf{A}^H\mathbf{D} \odot \mathbf{C}^H\mathbf{B}, \tag{2.6}$$

Property 3.

$$\operatorname{vec}\left(\mathbf{A}\mathbf{E}\mathbf{B}^T\right) = (\mathbf{B} \otimes \mathbf{A}) \operatorname{vec}(\mathbf{E}), \tag{2.7}$$

Property 4.

$$\operatorname{vec}\left(\mathbf{a}^{(N)} \circ \dots \circ \mathbf{a}^{(1)}\right) = \mathbf{a}^{(1)} \otimes \dots \otimes \mathbf{a}^{(N)}. \tag{2.8}$$

Property 5. *Given $\boldsymbol{C} = \boldsymbol{A} \otimes \boldsymbol{B}$, $\boldsymbol{C}$ is full column rank if and only if $\boldsymbol{A}$ and $\boldsymbol{B}$ are full column rank, such that, $r_{\boldsymbol{C}} = r_{\boldsymbol{A}} r_{\boldsymbol{B}}$ with $r_{\boldsymbol{A}}$, $r_{\boldsymbol{B}}$ and $r_{\boldsymbol{C}}$ being ranks of $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$, respectively.*

2.3 Basics of tensor algebra

This section presents definitions and operations involving tensors. An N-order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \ldots \times I_N}$ with elements $x_{i_1,\ldots,i_N}$ is interpreted as a multidimensional array of numerical values. For demonstrations and discussions on these definitions and operations, see [36, 37, 38].

Definition 3. (Fiber) *Fibers are vectors obtained by fixing every index of the modes of a tensor, except for one. The third-order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ is formed by column fibers ($\boldsymbol{X}_{.i_2 i_3} \in \mathbb{C}^{I_1}$), row fibers ($\boldsymbol{X}_{i_1 . i_3} \in \mathbb{C}^{I_2}$) and tubes fibers ($\boldsymbol{X}_{i_1 i_2 .} \in \mathbb{C}^{I_3}$) as illustrated in Figure 1.*

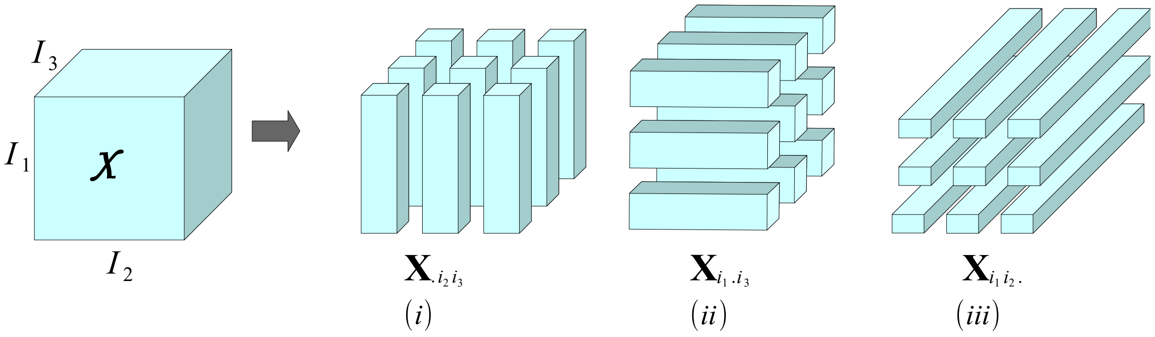


Figure 1 – (i) column fibers; (ii) row fibers; (iii) tube fibers.

Definition 4. (Matrix slice) *A matrix slice represents a two-dimensional section of a tensor. It is obtained by fixing all indices except two. A third-order tensor $\mathcal{X}$ is formed by frontal slices ($\boldsymbol{X}_{..i_3} \in \mathbb{C}^{I_1 \times I_2}$), lateral slices ($\boldsymbol{X}_{.i_2.} \in \mathbb{C}^{I_1 \times I_3}$) and horizontal slices ($\boldsymbol{X}_{i_1..} \in \mathbb{C}^{I_2 \times I_3}$). Figure 2 illustrates the three types of slices of the third-order tensor $\mathcal{X}$.*

Definition 5. (Matrix unfoldings) *The matrix unfolding is the process of reordering the elements of an N-order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \ldots \times I_N}$ into a matrix without loss of information [36]. There are different ways of matricizing the tensor $\mathcal{X}$ according to the partitioning of the set $\{1,\ldots,N\}$ into two ordered subsets $\mathbb{S}_1$ and $\mathbb{S}_2$, constituted of p and $N-p$ indices, respectively. A formula for the matricization for $p \in [1, N-1]$ is* [38]:

$$\mathbf{X}_{\mathbb{S}_1 \times \mathbb{S}_2} = \sum_{i_1}^{I_1} \ldots \sum_{i_N}^{I_N} y_{i_1,\ldots,i_N} \left(\bigotimes_{n \in \mathbb{S}_1} \mathbf{e}_{i_n}^{(I_n)} \right) \left(\bigotimes_{n \in \mathbb{S}_2} \mathbf{e}_{i_n}^{(I_n)} \right)^T \in \mathbb{C}^{J_1 \times J_2}, \tag{2.9}$$

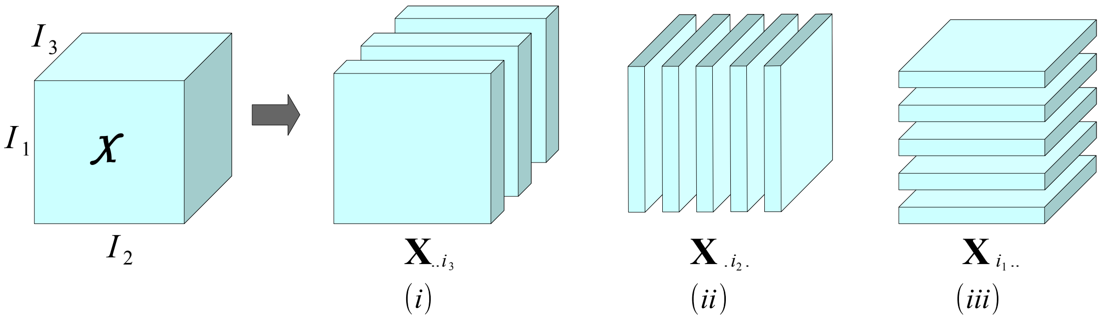


Figure 2 – (i) frontal slice; (ii) lateral slice; (iii) horizontal slice.

with $J_{n_1} = \prod_{n\in\mathbb{S}_{n_1}} I_n$, for $n_1 \in [1,2]$. Note that an n-mode unfolding is a rearrangement of the elements of $\mathcal{X}$ obtained by varying the index i_n and keeping the others indices fixed, in such a way that the fibers of the n-mode are placed along the rows (flat unfolding) or columns (tall unfolding). Considering a third-order tensor $\mathcal{X}$, the matrix unfolding $\mathbf{X}_{I_1\times I_3 I_2}$ is defined as Eq. (2.10). This unfolding is obtained by stacking the slices of a given tensor mode:

$$\mathbf{X}_{I_1\times I_3 I_2} = \sum_{i_1=1}^{I_1}\sum_{i_2=1}^{I_2}\sum_{i_3=1}^{I_3} x_{i_1,i_2,i_3}\left(\mathbf{e}_{i_1}^{(I_1)}\right)\left(\mathbf{e}_{i_3}^{(I_3)}\otimes \mathbf{e}_{i_2}^{(I_2)}\right)^T \in \mathbb{C}^{I_1\times I_3 I_2}. \tag{2.10}$$

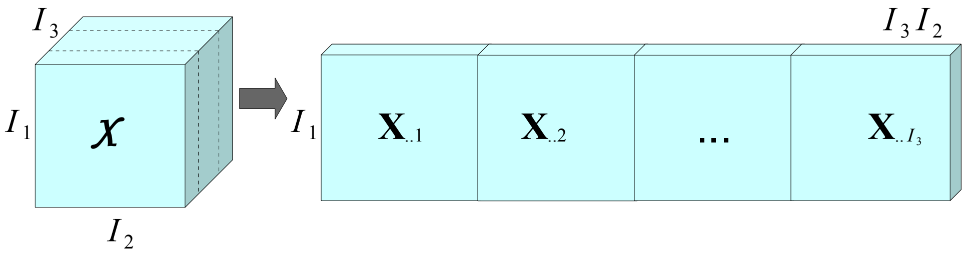


Figure 3 – Matrix unfolding representation $\mathbf{X}_{I_1\times I_3 I_2}$ of the third-order tensor $\mathcal{X}$.

Definition 6. (Frobenius norm) *Frobenius norm of $\mathcal{X}$ is defined as the square root of the inner product of the tensor with itself, i.e.:*

$$\| \mathcal{X} \|_F = \sqrt{\langle \mathcal{X}\mathcal{X}\rangle} = \left(\sum_{i_1=1}^{I_1}\cdots\sum_{i_N=1}^{I_N} |x_{i_1,\dots,i_N}|^2\right)^{1/2}. \tag{2.11}$$

Definition 7. (Rank-one tensor) *A tensor $\mathcal{X}$ is said to be a rank-one tensor if we can write $\mathcal{X}$ as the outer product of N vectors $\mathbf{a}^{(n)} \in \mathbb{C}^{I_n}$, with $n \in [1, N]$, as follows:*

$$\mathcal{X} = \mathbf{a}^{(1)} \circ ... \circ \mathbf{a}^{(N)}. \tag{2.12}$$

Parallel factor analysis (PARAFAC) [39] decomposition expresses the tensor as linear combinations of rank-one tensors. This definition is a generalized concept of the rank-one matrix, where a matrix $\mathbf{X} \in \mathbb{C}^{I_1 \times I_2}$ has rank one if there are two vectors $\mathbf{a}^{(1)} \in \mathbb{C}^{I_1}$ and $\mathbf{a}^{(2)} \in \mathbb{C}^{I_2}$, such that, $\mathbf{X} = \mathbf{a}^{(1)} \circ \mathbf{a}^{(2)} = \mathbf{a}^{(1)}\mathbf{a}^{(2)T}$.

Definition 8. (Kruskal rank) *The Kruskal rank (k-rank) of a matrix $\boldsymbol{A} \in \mathbb{C}^{M \times N}$ is the largest integer such that every set of $k_{\boldsymbol{A}}$ columns of $\boldsymbol{A}$ is linearly independent. Note that the k-rank is always $k_{\boldsymbol{A}} \leq r_{\boldsymbol{A}} \leq \min(M, N)$.*

Definition 9. (Identity tensor) *The identity tensor $\mathcal{I}_{N,R} \in \mathbb{R}^{R \times R \times ... \times R}$ is a diagonal hypercubic tensor whose elements $\delta_{r_1,...,r_N}$ are defined by means of the generalized Kronecker delta, i.e.:*

$$\delta_{r_1,...,r_N} = \begin{cases} 1, & \text{if} \quad r_1 = ... = r_N \\ 0, & \text{otherwise} \end{cases}, \tag{2.13}$$

and $R_n = R$, $\forall_n \in [1, N]$.

Definition 10. (n-mode product) *The n-mode product consists of multiplying an N-order tensor $\mathcal{X}$ by a matrix (or a vector) along its n-th mode. For the matrix case, the n-mode product of a N-order tensor $\mathcal{X}$ with a matrix $\boldsymbol{U} \in \mathbb{C}^{J \times I_n}$ is denoted by $\mathcal{X} \times_n \boldsymbol{U}$ and is of size $I_1 \times ... I_{n-1} \times J \times I_{n+1} \times ... \times I_N$. Elementwise, we have:*

$$b_{i_1,...,i_{n-1},j,i_{n+1},...,i_N} = \sum_{i_n=1}^{I_n} x_{i_1,...,i_{n-1},i_n,i_{n+1},...,i_N} u_{j,i_n}. \tag{2.14}$$

The n-mode product can be also represented in terms of the n-mode unfolding matrices of $\mathcal{X}$ as follows:

$$\mathbf{B}^{(n)} = \mathbf{U}\mathbf{X}^{(n)}, \tag{2.15}$$

with $\boldsymbol{X}^{(n)} \in \mathbb{C}^{I_n \times I_{n+1} ... I_N I_1 ... I_{n-1}}$ and $\boldsymbol{B}^{(n)} \in \mathbb{C}^{J \times I_{n+1} ... I_N I_1 ... I_{n-1}}$. We have the following properties for the n-mode product:

Property 6. *For any permutation $\pi(\cdot)$ of P distinct indices $m_p \in [1, N]$ such as $q_p = \pi(m_p)$, $p \in [1, P]$, with $P \leq N$, we have [38]:*

$$\mathcal{X} \times_{q=q_1}^{q_p} \mathbf{A}^{(q)} = \mathcal{X} \times_{m=m_1}^{m_p} \mathbf{A}^{(m)}, \tag{2.16}$$

which means that the order of the m_p-mode products is irrelevant when the indices m_p are all distinct.

Property 7. *Consider $\boldsymbol{A} \in \mathbb{C}^{K\times I_n}$ and $\mathbf{C} \in \mathbb{C}^{L\times K}$. For two products of $\mathcal{X} \in \mathbb{C}^{I_1\times\ldots\times I_N}$ along the same n-mode we have:*

$$\mathcal{X} \times_n \mathbf{A} \times_n \mathbf{C} = \mathcal{X} \times_n (\mathbf{CA}), \tag{2.17}$$

for $n \in [1, N]$. For the same n-mode, the order of the multiplication is relevant.

Definition 11. (Contraction operation) *The contraction between two tensors $\mathcal{Z} \in \mathbb{C}^{I_1\times\ldots\times I_N}$ and $\mathcal{X} \in \mathbb{C}^{J_1\times\ldots\times J_M}$, with $I_n = J_m$ is a tensor of order $(N + M - 2)$ such that:*

$$\mathcal{Z} \times_n^m \mathcal{X} = \sum_{k=1}^{I_n} z_{i_1,\ldots,i_{n-1},k,i_{n+1},\ldots,i_N} x_{j_1,\ldots,j_{m-1},k,j_{m+1},\ldots,j_M}, \tag{2.18}$$

where $(\mathcal{Z} \times_n^m \mathcal{X}) \in \mathbb{C}^{I_1\times\ldots\times I_{n-1}\times I_{n+1}\times\ldots\times I_N\times J_1\times\ldots\times J_{m-1}\times J_{m+1}\times\ldots\times J_M}$.

Definition 12. (Multilinear rank) *Considering an N-order tensor $\mathcal{X} \in \mathbb{C}^{I_1\times\ldots\times I_N}$, the N-tuplet $(R_1,\ldots,R_N)$ is called multilinear rank of $\mathcal{X}$, where the column vectors of the matricized form $\boldsymbol{X}^{(n)} \in \mathbb{C}^{I_n\times I_{n+1}\ldots I_N I_1\ldots I_{n-1}}$ are the n-mode vectors and $R_n = r\left(\boldsymbol{X}^{(n)}\right) = r_n(\mathcal{X}) \leq I_n$ is the n-mode rank of $\mathcal{X}$ [35].*

2.4 Background on tensor decompositions

In the last years, the use of multilinear algebra for communication systems has been growing. Tensor decompositions were used to model received signals and derive receivers to estimate the channel and symbol matrices, exploring various forms of diversity. Tensor decompositions can be seen as extensions of matrix decompositions to higher-order matrices (matrices with more than two dimensions) that allow recovering information from multivariate datasets decomposing tensors into elementary factors. The two most famous tensor decompositions are PARAFAC [39] and Tucker [40] models, as well as their variations. PARAFAC decomposition, also called canonical polyadic decomposition (CPD), was introduced independently by R. Harshman *et al.* [39] and J. Carroll & J. Chang [41], both in 1970, for applications in psychometrics and phonetics, respectively. Tucker decomposition was introduced by L. Tucker [40] in 1966 and was presented as an extension of the singular value decomposition (SVD) to three-way matrices, which gave rise to generalization as HOSVD [35]. PARAFAC and Tucker decompositions influenced the development of new tensor decompositions for specific cases such as nested PARAFAC [7, 42, 43], nested Tucker [44], generalized nested PARAFAC [12], coupled nested Tucker [14] and tensor train (TT) [45, 46]. This section provides an overview of the main tensor decompositions used in this thesis. These models are namely PARAFAC, nested PARAFAC, Tucker, nested coupled Tucker-PARAFAC and TT.

2.4.1 PARAFAC decomposition

PARAFAC model decomposes a tensor into a sum of rank-one tensors. Then, let $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ be a third-order tensor, the PARAFAC decomposition of $\mathcal{X}$ can be expressed as:

$$x_{i,j,k} = \sum_{r=1}^{R} a_{i,r} b_{j,r} c_{k,r}, \tag{2.19}$$

where $a_{i,r}$, $b_{j,r}$ and $c_{k,r}$ are the elements of $\mathbf{A} \in \mathbb{C}^{I \times R}$, $\mathbf{B} \in \mathbb{C}^{J \times R}$ and $\mathbf{C} \in \mathbb{C}^{K \times R}$, respectively. PARAFAC decomposition also can be represented as a sum of R rank-one component tensors, i.e.:

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{A}_{.r} \circ \mathbf{B}_{.r} \circ \mathbf{C}_{.r} \in \mathbb{C}^{I \times J \times K}, \tag{2.20}$$

where $\mathbf{A}_{.r} \in \mathbb{C}^{I}$, $\mathbf{B}_{.r} \in \mathbb{C}^{J}$ and $\mathbf{C}_{.r} \in \mathbb{C}^{K}$ are the r-th column vectors of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, respectively, with $r \in [1, R]$ and R known as the *rank* of the model or tensor rank. The number of factors is the dimension of the tensor. In this case, there are three factors $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$. Therefore, we have a third-dimensional tensor. Figure 4 illustrates the PARAFAC decomposition of a third-order tensor $\mathcal{X}$ into R components.

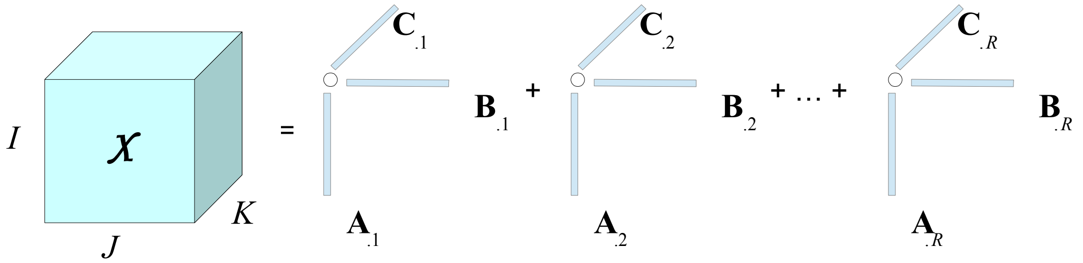


Figure 4 – PARAFAC decomposition of a third-order tensor $\mathcal{X}$ into R components.

Definition 13. (Tensor rank) *Tensor rank corresponds to the minimal number of rank-one tensors needed to write this tensor as a linear combination. It can also be called canonical rank.* The previous definition implies that an arbitrary tensor of rank $R \geq 1$ can be written as a sum of R rank-one tensors as Eq. (2.20).

PARAFAC decomposition can be also represented as a n-mode product, such that:

$$\mathcal{X} = \mathcal{I}_{3,R} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C} \in \mathbb{C}^{I \times J \times K}, \tag{2.21}$$

where $\mathcal{I}_{3,R} \in \mathbb{R}^{R \times R \times R}$ is the third-order identity tensor. The main characteristic of the PARAFAC model is its intrinsic uniqueness. In fact, PARAFAC decomposition is essentially unique, i.e., the matrices can be estimated up to scaling and permutation ambiguities under mild conditions. The 1-mode, 2-mode and 3-mode matrix unfoldings of the PARAFAC decomposition can be expressed as:

$$\mathbf{X}_{JK\times I} = (\mathbf{B} \diamond \mathbf{C})\,\mathbf{A}^T, \tag{2.22}$$

$$\mathbf{X}_{KI\times J} = (\mathbf{C} \diamond \mathbf{A})\,\mathbf{B}^T, \tag{2.23}$$

$$\mathbf{X}_{IJ\times K} = (\mathbf{A} \diamond \mathbf{B})\,\mathbf{C}^T. \tag{2.24}$$

Now, let us consider an N-order tensor $\mathcal{Y} \in \mathbb{C}^{I_1\times\ldots\times I_N}$ that satisfies the PARAFAC decomposition with rank R expressed as:

$$y_{i_1,\ldots,i_N} = \sum_{r=1}^{R}\prod_{n=1}^{N} a^{(n)}_{i_n,r}, \tag{2.25}$$

where $a^{(n)}_{i_n,r}$ are the elements of $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n\times R}$, for $n \in [1, N]$. PARAFAC model is also written in terms of the n-mode product:

$$\mathcal{Y} = \mathcal{I}_{N,R} \times_1 \mathbf{A}^{(1)} \times_2 \ldots \times_N \mathbf{A}^{(N)}. \tag{2.26}$$

The generic formulation of a tall n-mode unfolding of the PARAFAC decomposition is given by:

$$\mathbf{Y}_{I_{n+1}\ldots I_N I_1\ldots I_{n-1}\times I_n} = \left(\mathbf{A}^{(n+1)} \diamond \ldots \diamond \mathbf{A}^{(N)} \diamond \mathbf{A}^{(1)} \diamond \ldots \diamond \mathbf{A}^{(n-1)}\right)\mathbf{A}^{(n)T}. \tag{2.27}$$

2.4.1.1 Uniqueness

The uniqueness of the PARAFAC model has been discussed in several works [47, 48, 49, 50, 51, 52, 53]. Kruskal derived sufficient conditions for the uniqueness of third-order PARAFAC decompositions of real-valued tensors [50]. Around two decades later, Sidiropoulos *et al.* [47] extended the Kruskal condition to complex-valued tensors. Sidiropoulos & Bro [49] further generalized Kruskal's uniqueness condition to N-order tensors. Since any arbitrary tensor can be written as a PARAFAC decomposition, the following theorem holds for any tensor that satisfies the following condition:

Theorem 1. *Consider an N-order tensor $\mathcal{Y} \in \mathbb{C}^{I_1\times\ldots\times I_N}$ that satisfies a PARAFAC decomposition. If the condition*

$$\sum_{n=1}^{N} k_{\mathbf{A}^{(n)}} \geq 2R + (N-1), \tag{2.28}$$

is satisfied, the factor matrices $\mathbf{A}^{(n)}$, for $n \in [1, N]$, are unique up to permutation and scaling ambiguities, so that $\hat{\mathbf{A}}^{(n)} = \mathbf{A}^{(n)}\mathbf{\Pi}\mathbf{\Lambda}^{(n)}$, where $\mathbf{\Pi} \in \mathbb{C}^{R\times R}$ is the permutation matrix and $\mathbf{\Lambda}^{(n)} \in \mathbb{C}^{R\times R}$ are diagonal scaling matrices, with $\prod_{n=1}^{N}\mathbf{\Lambda}^{(n)} = \mathbf{I}_R$.

The condition (2.28) is sufficient but not necessary to guarantee essential uniqueness [54].

2.4.2 Coupled PARAFAC

Let us consider the following decomposition for the fourth-order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times J_1 \times I_2 \times J_2}$ [55]:

$$x_{i_1,j_1,i_2,j_2} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} a^{(1)}_{i_1,r_1} b^{(1)}_{j_1,r_1} a^{(2)}_{i_2,r_2} b^{(2)}_{j_2,r_2} g_{r_1,r_2}, \tag{2.29}$$

where $a^{(1)}_{i_1,r_1}$, $a^{(2)}_{i_2,r_2}$, $b^{(1)}_{i_1,r_1}$, $b^{(2)}_{i_2,r_2}$ and g_{r_1,r_2} are the entries of the factor matrices $\mathbf{A}^{(1)} \in \mathbb{C}^{I_1 \times R_1}$, $\mathbf{A}^{(2)} \in \mathbb{C}^{I_2 \times R_2}$, $\mathbf{B}^{(1)} \in \mathbb{C}^{J_1 \times R_1}$, $\mathbf{B}^{(2)} \in \mathbb{C}^{J_2 \times R_2}$ and $\mathbf{G} \in \mathbb{C}^{R_1 \times R_2}$, respectively. This model can be interpreted as two coupled third-order PARAFAC models sharing $\mathbf{G}$ as a common factor matrix. Indeed, let us define the third-order tensors $\mathcal{Z}^{(1)} \in \mathbb{C}^{I_1 \times J_1 \times R_2}$ and $\mathcal{Z}^{(2)} \in \mathbb{C}^{I_2 \times J_2 \times R_1}$ as:

$$z^{(1)}_{i_1,j_1,r_2} = \sum_{r_1}^{R_1} a^{(1)}_{i_1,r_1} b^{(1)}_{j_1,r_1} g_{r_1,r_2}, \tag{2.30}$$

$$z^{(2)}_{i_2,j_2,r_1} = \sum_{r_2}^{R_2} a^{(2)}_{i_2,r_2} b^{(2)}_{j_2,r_2} g_{r_1,r_2}. \tag{2.31}$$

Eqs. (2.30)-(2.31) correspond to PARAFAC decompositions of the tensors $\mathcal{Z}^{(1)}$ and $\mathcal{Z}^{(2)}$, with factor matrices ($\mathbf{A}^{(1)}$,$\mathbf{B}^{(1)}$,$\mathbf{G}^{T}$) and ($\mathbf{A}^{(2)}$,$\mathbf{B}^{(2)}$,$\mathbf{G}$), respectively. These tensors admit the following matrix unfolding forms:

$$\mathbf{Z}^{(n)} = \left(\mathbf{A}^{(n)} \diamond \mathbf{B}^{(n)}\right) \mathbf{C}^{(n)T} \in \mathbb{C}^{K_n \times R_{n_1}}, \tag{2.32}$$

with

$$\mathbf{C}^{(n)} = \begin{cases} \mathbf{G}^{T}, & \text{for} \quad n = 1, n_1 = 2, \\ \mathbf{G}, & \text{for} \quad n = 2, n_1 = 1 \end{cases}. \tag{2.33}$$

and $K_n = I_n J_n$. These matrix representations of $\mathcal{Z}^{(1)}$ and $\mathcal{Z}^{(2)}$ are associated with a contraction of the first two modes ($k_n = (i_n - 1)J_n + j_n$, for $n \in [1,2]$). We can rewrite Eq. (2.29) as two coupled PARAFAC decompositions:

$$x^{(1)}_{i_1,j_1,k_2} = \sum_{r_1=1}^{R_1} a^{(1)}_{i_1,r_1} b^{(1)}_{j_1,r_1} z^{(2)}_{k_2,r_1}, \tag{2.34}$$

$$x^{(2)}_{i_2,j_2,k_1} = \sum_{r_2=1}^{R_2} a^{(2)}_{i_2,r_2} b^{(2)}_{j_2,r_2} z^{(1)}_{k_1,r_2}, \tag{2.35}$$

with respective factor matrices ($\mathbf{A}^{(1)}$,$\mathbf{B}^{(1)}$,$\mathbf{Z}^{(2)}$) and ($\mathbf{A}^{(2)}$,$\mathbf{B}^{(2)}$,$\mathbf{Z}^{(1)}$), where $\mathbf{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$ are defined as:

$$z^{(1)}_{k_1,r_2} = z^{(1)}_{i_1,j_1,r_2}, \quad \text{and} \quad z^{(2)}_{k_2,r_1} = z^{(2)}_{i_2,j_2,r_1}. \tag{2.36}$$

It is worth noting that Eqs. (2.34)-(2.35) are different contracted representations of the same fourth-order tensor $\mathcal{X}$ defined in Eq. (2.29), corresponding to two different ways of coupling the third-order PARAFAC decompositions in Eqs. (2.30)-(2.31) containing the full information of the original tensor model in Eq. (2.29). In terms of n-mode unfolding, Eqs. (2.30)-(2.31) can be represented as:

$$\mathcal{X}^{(1)} = \mathcal{I}_{3,R_1} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{B}^{(1)} \times_3 \mathbf{Z}^{(2)} \in \mathbb{C}^{I_1 \times J_1 \times K_2}, \tag{2.37}$$

$$\mathcal{X}^{(2)} = \mathcal{I}_{3,R_2} \times_1 \mathbf{A}^{(2)} \times_2 \mathbf{B}^{(2)} \times_3 \mathbf{Z}^{(1)} \in \mathbb{C}^{I_2 \times J_2 \times K_1}. \tag{2.38}$$

The decomposition in Eqs. (2.37)-(2.38) admits the following matrix unfoldings representations:

$$\mathbf{X}^{(1)}_{J_1K_2 \times I_1} = \left(\mathbf{B}^{(1)} \diamond \mathbf{Z}^{(2)}\right) \mathbf{A}^{(1)T} \in \mathbb{C}^{J_1K_2 \times I_1}, \tag{2.39}$$

$$\mathbf{X}^{(2)}_{J_2K_1 \times I_2} = \left(\mathbf{B}^{(2)} \diamond \mathbf{Z}^{(1)}\right) \mathbf{A}^{(2)T} \in \mathbb{C}^{J_2K_1 \times I_2}. \tag{2.40}$$

Coupled PARAFAC decompositions (2.37)-(2.38) are essentially unique, i.e., their factor matrices are unique up to column permutation and scaling ambiguities, if [55]:

$$k_{\mathbf{A}^{(n)}} + k_{\mathbf{B}^{(n)}} + k_{\mathbf{Z}^{(n_1)}} \leq 2R_n + 2, \tag{2.41}$$

for $(n, n_1) \in [(1,2),(2,1)]$.

2.4.3 Tucker decomposition

Given a third-order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$, Tucker model decomposes it as a multilinear transformation of a third-order tensor $\mathcal{G} \in \mathbb{C}^{P \times Q \times R}$ by the factor matrices $\mathbf{A} \in \mathbb{C}^{I \times P}$, $\mathbf{B} \in \mathbb{C}^{J \times Q}$ and $\mathbf{C} \in \mathbb{C}^{K \times R}$ as illustrated in Figure 5. Tucker decomposition of $\mathcal{X}$ can be defined as:

$$x_{i,j,k} = \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} g_{p,q,r} a_{i,p} b_{j,q} c_{k,r}, \tag{2.42}$$

where $a_{i,p}$, $b_{j,q}$ and $c_{k,r}$ are the (i,p)-th, (j,q)-th and (k,r)-th elements of $\mathbf{A} \in \mathbb{C}^{I \times P}$, $\mathbf{B} \in \mathbb{C}^{J \times Q}$ and $\mathbf{C} \in \mathbb{C}^{K \times R}$, respectively, and $g_{p,q,r}$ is the (p,q,r)-th element of the core tensor $\mathcal{G} \in \mathbb{C}^{P \times Q \times R}$. In terms of the outer product, Tucker decomposition is represented by:

$$\mathcal{X} = \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} g_{p,q,r} \left(\mathbf{A}_{.p} \circ \mathbf{B}_{.q} \circ \mathbf{C}_{.r}\right) \in \mathbb{C}^{I \times J \times K}, \tag{2.43}$$

where $\mathbf{A}_{.p} \in \mathbb{C}^{I}$, $\mathbf{B}_{.q} \in \mathbb{C}^{J}$ and $\mathbf{C}_{.r} \in \mathbb{C}^{K}$ are the column vectors of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, respectively. Using the n-mode product, the Tucker decomposition of $\mathcal{X}$ can be written as:

$$\mathcal{X} = \mathcal{G} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C}. \tag{2.44}$$

This decomposition can also be represented in a compact form as $(\mathcal{G}, \mathbf{A}, \mathbf{B}, \mathbf{C})$. The n-mode matrix unfoldings of $\mathcal{X}$ are represented by:

$$\mathbf{X}_{JK \times I} = (\mathbf{B} \otimes \mathbf{C})\, \mathbf{G}_{QR \times P} \mathbf{A}^{T}, \tag{2.45}$$

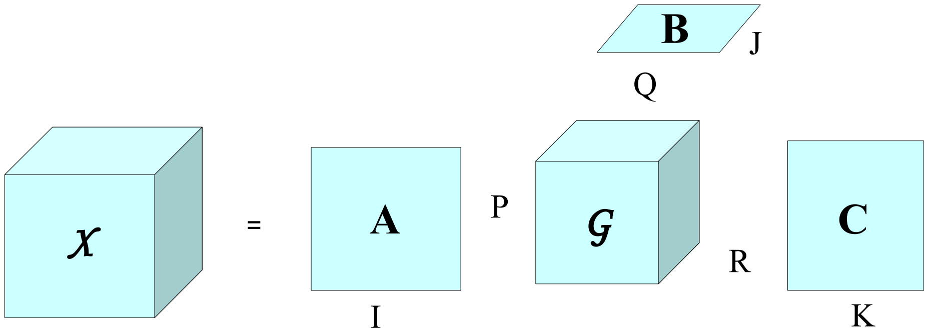


Figure 5 – Tucker decomposition of a third-order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$.

$$\mathbf{X}_{KI \times J} = (\mathbf{C} \otimes \mathbf{A})\, \mathbf{G}_{RP \times Q} \mathbf{B}^T, \tag{2.46}$$

$$\mathbf{X}_{IJ \times K} = (\mathbf{A} \otimes \mathbf{B})\, \mathbf{G}_{PQ \times R} \mathbf{C}^T, \tag{2.47}$$

where $\mathbf{G}_{QR \times P}$, $\mathbf{G}_{RP \times Q}$ and $\mathbf{G}_{PQ \times R}$ denote 1-mode, 2-mode and 3-mode matrix unfoldings of $\mathcal{G}$, respectively.

For an N-order tensor $\mathcal{Y} \in \mathbb{C}^{I_1 \times \ldots \times I_N}$, its Tucker decomposition is given by:

$$y_{i_1,\ldots,i_N} = \sum_{r_1=1}^{R_1} \ldots \sum_{r_N=1}^{R_N} g_{r_1,\ldots,r_N} \prod_{n=1}^{N} a^{(n)}_{i_n,r_n}, \tag{2.48}$$

where $g_{r_1,\ldots,r_N}$ are the elements of $\mathcal{G} \in \mathbb{C}^{R_1 \times \ldots \times R_N}$ and $a^{(n)}_{i_n,r_n}$ are the elements of $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R_n}$, for $n \in [1, N]$. Tucker decomposition can be also expressed as:

$$\mathcal{Y} = \sum_{r_1=1}^{R_1} \ldots \sum_{r_N=1}^{R_N} g_{r_1,\ldots,r_N} \left(\mathbf{A}^{(1)}_{.r_1} \circ \ldots \circ \mathbf{A}^{(N)}_{.r_N} \right), \tag{2.49}$$

where $\mathbf{A}^{(n)}_{.r_n} \in \mathbb{C}^{I_n}$ are the column vectors of the factor matrices $\mathbf{A}^{(n)}$. The N-order Tucker decomposition can also be represented by the n-mode product as:

$$\mathcal{Y} = \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \ldots \times_N \mathbf{A}^{(N)}. \tag{2.50}$$

The n-mode generic matrix unfolding of Eq. (2.50) is given by:

$$\mathbf{Y}_{I_{n+1} \ldots I_N I_1 \ldots I_{n-1} \times I_n} = \left(\mathbf{A}^{(n+1)} \otimes \ldots \otimes \mathbf{A}^{(N)} \otimes \mathbf{A}^{(1)} \otimes \ldots \otimes \mathbf{A}^{(n-1)} \right) \mathbf{G}_{R_{n+1} \ldots R_N R_1 \ldots R_{n-1} \times R_n} \mathbf{A}^{(n)T}, \tag{2.51}$$

where $\mathbf{G}_{R_{n+1} \ldots R_N R_1 \ldots R_{n-1} \times R_n} \in \mathbb{C}^{R_{n+1} \ldots R_N R_1 \ldots R_{n-1} \times R_n}$ is the corresponding n-mode unfolding of $\mathcal{G}$.

Special cases of the Tucker decomposition were introduced in [38]. Then, let us consider the decomposition of an N-order tensor that has N_1 factor matrices, such that

$N > N_1$ and $N - N_1$ denotes the number of factor matrices equal to the identity matrix. Tucker-(N_1, N) decomposition can be written to N-order tensors as [38]:

$$\mathcal{Y} = \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \ldots \times_{N_1} \mathbf{A}^{(N_1)} \times_{N_1+1} \mathbf{I}_{I_{N_1+1}} \times_{N_1+2} \ldots \times_N \mathbf{I}_{I_N}, \tag{2.52}$$

where $\mathbf{A}^{(n)} = \mathbf{I}_{I_n}$ for $n \in [N_1 + 1, N]$, which implies: $\mathcal{G} \in \mathbb{C}^{R_1 \times \ldots \times R_{N_1} \times I_{N_1+1} \times \ldots \times I_N}$ and $R_n = I_n$. Or simply:

$$\mathcal{Y} = \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \ldots \times_{N_1} \mathbf{A}^{(N_1)} = \mathcal{G} \times_{n=1}^{N_1} \mathbf{A}^{(n)}. \tag{2.53}$$

For example, considering a third-order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ and the third factor matrix equal to the identity matrix, i.e., $\mathbf{C} = \mathbf{I}_K \in \mathbb{C}^{K \times K}$. The Tucker-(2,3) or Tucker-2 is given by:

$$\mathcal{X} = \mathcal{G} \times_1 \mathbf{A} \times_2 \mathbf{B} \in \mathbb{C}^{I \times J \times K}, \tag{2.54}$$

where $\mathcal{G} \in \mathbb{C}^{P \times Q \times K}$. Similarly, considering the second and the third factor matrices equal to identity matrices, i.e., $\mathbf{B} = \mathbf{I}_J \in \mathbb{C}^{J \times J}$ and $\mathbf{C} = \mathbf{I}_K$. The Tucker-1 or Tucker-(1,3) decomposition is defined as:

$$\mathcal{X} = \mathcal{G} \times_1 \mathbf{A} \in \mathbb{C}^{I \times J \times K}, \tag{2.55}$$

where $\mathcal{G} \in \mathbb{C}^{P \times J \times K}$.

2.4.3.1 Uniqueness

Tucker model is not essentially unique, since the factor matrices **A**, **B** and **C** and the core tensor $\mathcal{G}$ are not identifiable in a unique way. The singularity can be obtained by imposing some constraints on $\mathcal{G}$ or on **A**, **B** and **C**. To demonstrate the non-uniqueness of the Tucker model, we have the third-order tensor $\mathcal{X}$ decomposed as:

$$\mathcal{X} = \hat{\mathcal{G}} \times_1 \hat{\mathbf{A}} \times_2 \hat{\mathbf{B}} \times_3 \hat{\mathbf{C}}, \tag{2.56}$$

where $\hat{\mathcal{G}} \in \mathbb{C}^{P \times Q \times R}$, $\hat{\mathbf{A}} \in \mathbb{C}^{I \times P}$, $\hat{\mathbf{B}} \in \mathbb{C}^{J \times Q}$ and $\hat{\mathbf{C}} \in \mathbb{C}^{K \times R}$. This decomposition is not unique because it has freedom of rotation. For example, defining:

$$\begin{cases} \hat{\mathcal{G}} = \mathcal{G} \times_1 \mathbf{T}^{(1)} \times_2 \mathbf{T}^{(2)} \times_3 \mathbf{T}^{(3)}, \\ \hat{\mathbf{A}} = \mathbf{A}[\mathbf{T}^{(1)}]^{-1}, \\ \hat{\mathbf{B}} = \mathbf{B}[\mathbf{T}^{(2)}]^{-1}, \\ \hat{\mathbf{C}} = \mathbf{C}[\mathbf{T}^{(3)}]^{-1}, \end{cases} \tag{2.57}$$

with $\mathbf{T}^{(1)} \in \mathbb{C}^{P \times P}$, $\mathbf{T}^{(2)} \in \mathbb{C}^{Q \times Q}$ and $\mathbf{T}^{(3)} \in \mathbb{C}^{R \times R}$. Substituting (2.57) in (2.56) and applying the property 7 we have:

$$\begin{aligned} \mathcal{X} &= \mathcal{G} \times_1 \mathbf{T}^{(1)} \times_2 \mathbf{T}^{(2)} \times_3 \mathbf{T}^{(3)} \times_1 \mathbf{A}[\mathbf{T}^{(1)}]^{-1} \times_2 \mathbf{B}[\mathbf{T}^{(2)}]^{-1} \times_3 \mathbf{C}[\mathbf{T}^{(3)}]^{-1} \\ &= \mathcal{G} \times_1 \mathbf{A}[\mathbf{T}^{(1)}]^{-1}\mathbf{T}^{(1)} \times_2 \mathbf{B}[\mathbf{T}^{(2)}]^{-1}\mathbf{T}^{(2)} \times_3 \mathbf{C}[\mathbf{T}^{(3)}]^{-1}\mathbf{T}^{(3)} \\ &= \mathcal{G} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C}, \end{aligned} \tag{2.58}$$

where $[\mathbf{T}^{(1)}]^{-1}\mathbf{T}^{(1)} = \mathbf{I}_P \in \mathbb{R}^{P\times P}$, $[\mathbf{T}^{(2)}]^{-1}\mathbf{T}^{(2)} = \mathbf{I}_Q \in \mathbb{R}^{Q\times Q}$ and $[\mathbf{T}^{(3)}]^{-1}\mathbf{T}^{(3)} = \mathbf{I}_R \in \mathbb{R}^{R\times R}$. From the above deduction, we can conclude that there are alternatives to the factor matrices and core tensor that satisfy the same Tucker decomposition. The above deduction can be generalized to an N-order Tucker model.

2.4.4 Coupled nested Tucker-PARAFAC decomposition

This section addresses one of the main contributions of this thesis, the proposed coupled nested Tucker-PARAFAC model. This decomposition is a combination of the PARAFAC and Tucker decompositions. Let us consider a fifth-order tensor $\mathcal{X} \in \mathbb{C}^{I_n\times J\times K\times L\times M}$ modeled as a Tucker-(4, 5) decomposition:

$$\mathcal{X} = \mathcal{G} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C} \times_4 \mathbf{D} \times_5 \mathbf{I}_M \in \mathbb{C}^{I_n\times J\times K\times L\times M}, \tag{2.59}$$

where $\mathcal{G} \in \mathbb{C}^{R_1\times R_2\times R_3\times R_4\times M}$ is the core tensor and $\mathbf{A} \in \mathbb{C}^{I_n\times R_1}$, $\mathbf{B} \in \mathbb{C}^{J\times R_2}$, $\mathbf{C} \in \mathbb{C}^{K\times R_3}$ and $\mathbf{D} \in \mathbb{C}^{L\times R_4}$ are the factor matrices and $\mathbf{I}_M \in \mathbb{R}^{M\times M}$ is the identity matrix. The factor matrix $\mathbf{A}$ corresponds to a n-mode matrix unfolding of the N-order tensor $\mathcal{A} \in \mathbb{C}^{I_1\times\ldots\times I_N}$ modeled as a PARAFAC decomposition:

$$\mathcal{A} = \mathcal{I}_{N,P} \times_1 \mathbf{A}^{(1)} \times_2 \ldots \times_N \mathbf{A}^{(N)} \in \mathbb{C}^{I_1\times\ldots\times I_N}, \tag{2.60}$$

where $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n\times P}$ are the factor matrices of the PARAFAC model, for $n \in [1, N]$. The n-mode matrix unfolding $\mathbf{A}$ can be expressed as:

$$\mathbf{A} = \mathbf{A}^{(n)} \left(\mathbf{A}^{(n+1)} \diamond \ldots \diamond \mathbf{A}^{(N)} \diamond \mathbf{A}^{(1)} \diamond \ldots \diamond \mathbf{A}^{(n-1)}\right)^T \in \mathbb{C}^{I_n\times I_{n+1}\ldots I_N I_1\ldots I_{n-1}}. \tag{2.61}$$

From Eq. (2.61), note that $R_1 = I_{n+1}\ldots I_N I_1\ldots I_{n-1}$. For the PARAFAC model (2.60), consider the factor matrix $\mathbf{A}^{(N)}$ formed by the concatenation of two other matrices $\mathbf{A}^{(N)'} \in \mathbb{C}^{I'_N\times P}$ and $\mathbf{A}^{(N)''} \in \mathbb{C}^{I''_N\times P}$, with $I_N = I'_N + I''_N$, such that:

$$\mathbf{A}^{(N)} = \begin{bmatrix} \mathbf{A}^{(N)'} \\ \mathbf{A}^{(N)''} \end{bmatrix} \in \mathbb{C}^{I_N\times P}. \tag{2.62}$$

Substituting $\mathbf{A}^{(N)}$ in Eq. (2.60), we have:

$$\mathcal{A} = \begin{bmatrix} \mathcal{A}' \\ \mathcal{A}'' \end{bmatrix} = \mathcal{I}_{N,P} \times_1 \mathbf{A}^{(1)} \times_2 \ldots \times_N \begin{bmatrix} \mathbf{A}^{(N)'} \\ \mathbf{A}^{(N)''} \end{bmatrix}, \tag{2.63}$$

where $\mathcal{A}' \in \mathbb{C}^{I_1\times\ldots\times I'_N}$ and $\mathcal{A}'' \in \mathbb{C}^{I_1\times\ldots\times I''_N}$. Then, substituting $\mathcal{A}$ in Eq. (2.59), we have:

$$\mathcal{X} = \begin{bmatrix} \mathcal{X}' \\ \mathcal{X}'' \end{bmatrix} = \mathcal{G} \times_1 \begin{bmatrix} \mathbf{A}' \\ \mathbf{A}'' \end{bmatrix} \times_2 \mathbf{B} \times_3 \mathbf{C} \times_4 \mathbf{D} \times_5 \mathbf{I}_M. \tag{2.64}$$

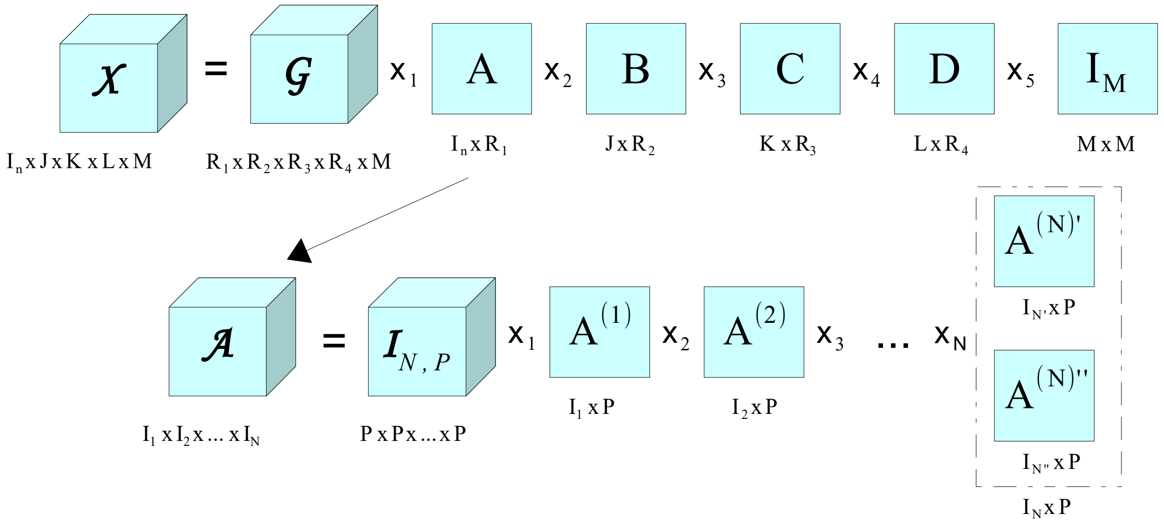


Figure 6 – Coupled nested Tucker-PARAFAC decomposition of a (4,5)-order $\mathcal{X} \in \mathbb{C}^{I_n \times J \times K \times L \times M}$.

From Eqs. (2.59)-(2.64), we can conclude that the tensor $\mathcal{X}$ satisfies a (4,5)-order coupled nested Tucker-PARAFAC model, as illustrated in Figure 6. The PARAFAC part of this model is associated with the tensor $\mathcal{A}$ represented by the PARAFAC model (2.60), whose unfolding $\mathbf{A}$ is the factor matrix along the first mode of $\mathcal{X}$. That explains the nested Tucker-PARAFAC structure. The coupled structure is due to the core tensor $\mathcal{G}$ that is common to the two components $(\mathcal{X}', \mathcal{X}'')$ of $\mathcal{X}$. Based on the contracted fifth-order Tucker model (2.59), it is easy to deduce the following matrix unfoldings of tensor $\mathcal{X}$:

$$\mathbf{X}_{JKLM \times I_n} = (\mathbf{B} \otimes \mathbf{C} \otimes \mathbf{D} \otimes \mathbf{I}_M)\, \mathbf{G}_{R_2R_3R_4M \times R_1} \mathbf{A}^T \in \mathbb{C}^{JKLM \times I_n}, \tag{2.65}$$

$$\mathbf{X}_{KLMI_n \times J} = (\mathbf{C} \otimes \mathbf{D} \otimes \mathbf{I}_M \otimes \mathbf{A})\, \mathbf{G}_{R_3R_4MR_1 \times R_2} \mathbf{B}^T \in \mathbb{C}^{KLMI_n \times J}, \tag{2.66}$$

$$\mathbf{X}_{LMI_nJ \times K} = (\mathbf{D} \otimes \mathbf{I}_M \otimes \mathbf{A} \otimes \mathbf{B})\, \mathbf{G}_{R_4MR_1R_2 \times R_3} \mathbf{C}^T \in \mathbb{C}^{LMI_nJ \times K}, \tag{2.67}$$

$$\mathbf{X}_{MI_nJK \times L} = (\mathbf{I}_M \otimes \mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C})\, \mathbf{G}_{MR_1R_2R_3 \times R_4} \mathbf{D}^T \in \mathbb{C}^{MI_nJK \times L}, \tag{2.68}$$

$$\mathbf{X}_{I_nJKL \times M} = (\mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C} \otimes \mathbf{D})\, \mathbf{G}_{R_1R_2R_3R_4 \times M} \in \mathbb{C}^{I_nJKL \times M}. \tag{2.69}$$

In MIMO communication systems, $\mathcal{A}$ can be replaced by a channel tensor $\mathcal{H}$ modeled by the PARAFAC decomposition, the tensor $\mathcal{G}$ would represent a coding tensor (e.g. TS, TST, TSTF, etc.) and $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ would represent the transmitted symbol matrices using the MSMKron coding proposed in this thesis. By exploiting this coupled nested Tucker-PARAFAC decomposition presented above, a wireless MIMO communication system is proposed.

2.4.4.1 Uniqueness

Coupled nested Tucker PARAFAC decomposition uniqueness results from the uniqueness of the Tucker model (2.59) of $\mathcal{X}$ combined with the uniqueness of the PARAFAC decomposition (2.60) of $\mathcal{A}$.

As it is well known, Tucker models are generally not essentially unique [40]. Uniqueness can be obtained by imposing some constraints on the core tensor or the factor matrices. For the Tucker model (2.59), the core tensor will be assumed to be known, which implies the uniqueness of the factor matrices up to scaling factors such as:

$$\hat{\mathbf{A}} = \mathbf{A}\lambda^{(A)}, \quad \hat{\mathbf{B}} = \mathbf{B}\lambda^{(B)}, \quad \hat{\mathbf{C}} = \mathbf{C}\lambda^{(C)}, \quad \hat{\mathbf{D}} = \mathbf{D}\lambda^{(D)}, \quad \lambda^{(A)}\lambda^{(B)}\lambda^{(C)}\lambda^{(D)} = 1 \ . \tag{2.70}$$

The tensor $\mathcal{A}$ (2.60) is modeled by a rank-P PARAFAC decomposition that is essentially unique if the following sufficient condition is satisfied [51]:

$$\sum_{n=1}^{N} k_{\mathbf{A}^{(n)}} \geq 2P + 2, \tag{2.71}$$

where $k_{\mathbf{A}^{(n)}}$ is the Kruskal rank of $\mathbf{A}^{(n)}$, for $n \in [1, N]$. If the condition (2.71) is satisfied, any set of matrices $(\hat{\mathbf{A}}^{(1)}, ..., \hat{\mathbf{A}}^{(N)})$ that satisfy Eq. (2.63) are related with the original matrices by:

$$\hat{\mathbf{A}}^{(n)} = \mathbf{A}^{(n)}\mathbf{\Pi}\mathbf{\Lambda}^{(n)}, \quad \prod_{n=1}^{N} \mathbf{\Lambda}^{(n)} = \mathbf{I}_P, \tag{2.72}$$

where $\mathbf{\Pi} \in \mathbb{C}^{P \times P}$ is a permutation matrix and $\mathbf{\Lambda}^{(n)}$ are diagonal matrices. The matrices $\mathbf{A}^{(n)}$ are full rank. Then, in the generic case, condition (2.71) becomes:

$$\sum_{n=1}^{N} \min\left(I_n, P\right) \geq 2P + 2. \tag{2.73}$$

If the matrices $\mathbf{A}^{(n)}$ are full rank, then the condition (2.73) is always satisfied, which ensures essential uniqueness of the PARAFAC decomposition of the tensor $\mathcal{A}$.

2.4.5 Tensor train (TT) decomposition

The TT decomposition was introduced in [45, 46] and was created to solve the curse of dimensionality, being used to perform approximation and compression of the higher-order tensor. TT decomposition can be calculated using standard decomposition as SVD [45]. Considering a fourth-order tensor $\mathcal{Y} \in \mathbb{C}^{I_1 \times I_2 \times I_3 \times I_4}$, the TT decomposition in scalar form can be written as:

$$y_{i_1,i_2,i_3,i_4} = \sum_{r_1}^{R_1}\sum_{r_2}^{R_2}\sum_{r_3}^{R_3}\sum_{r_4}^{R_4} g^{(1)}_{i_1,r_1} g^{(2)}_{r_1,i_2,r_2} g^{(3)}_{r_2,i_3,r_3} g^{(4)}_{r_3,i_4}, \tag{2.74}$$

where $g^{(1)}_{i_1,r_1}$, $g^{(2)}_{r_1,i_2,r_2}$, $g^{(3)}_{r_2,i_3,r_3}$ and $g^{(4)}_{r_3,i_4}$ are the elements of the TT-cores $\mathbf{G}^{(1)} \in \mathbb{C}^{I_1 \times R_1}$, $\mathcal{G}^{(2)} \in \mathbb{C}^{R_1 \times I_2 \times R_2}$, $\mathcal{G}^{(3)} \in \mathbb{C}^{R_2 \times I_3 \times R_3}$ and $\mathbf{G}^{(4)} \in \mathbb{C}^{R_3 \times I_4}$, respectively. In this decomposition, two successive TT-cores have a common index (r_n) that is referred to as auxiliary indices, in contrast to the indices i_n which are called spatial indices. The quantities R_n are the TT ranks. The TT-cores $\mathcal{G}^{(n)}$ have size $R_{n-1} \times I_n \times R_n$ except for $n = 1$ and $n = 4$, where the sizes are $I_1 \times R_1$ and $R_3 \times I_4$ with $R_0 = R_4 = 1$ respectively [46]. The TT decomposition represented in Figure 7 can be expressed using the contraction operation (Definition 11) as:

$$\mathcal{Y} = \mathbf{G}^{(1)} \times^1_2 \mathcal{G}^{(2)} \times^1_3 \mathcal{G}^{(3)} \times^1_4 \mathbf{G}^{(4)}. \tag{2.75}$$

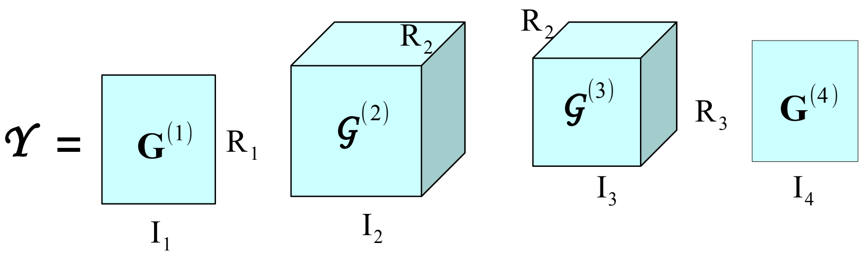


Figure 7 – TT decomposition of a fourth-order tensor $\mathcal{Y}$.

The matrix unfolding of $\mathcal{Y}$ in MATLAB can be obtained by the reshape function as:

$$\mathbf{Y}^{(n)} = \text{reshape}\left(\mathcal{Y}, \left[\prod_{s=1}^{n} I_s, \prod_{k=n+1}^{N} I_k\right]\right), \tag{2.76}$$

where the reshape function rearranges $\mathcal{Y}$ in terms of the dimensions $\left(\prod_{s=1}^{n} I_s\right) \times \left(\prod_{k=n+1}^{N} I_k\right)$. From Eq. (2.76), we have the following matrix unfoldings for $\mathcal{Y}$:

$$\mathbf{Y}^{(1)} = \text{reshape}\left(\mathcal{Y}, \left[I_1, \prod_{k=2}^{4} I_k\right]\right) \in \mathbb{C}^{I_1 \times I_2 I_3 I_4}, \tag{2.77}$$

$$\mathbf{Y}^{(2)} = \text{reshape}\left(\mathcal{Y}, \left[\prod_{s=1}^{2} I_s, \prod_{k=3}^{4} I_k\right]\right) \in \mathbb{C}^{I_1 I_2 \times I_3 I_4}, \tag{2.78}$$

$$\mathbf{Y}^{(3)} = \text{reshape}\left(\mathcal{Y}, \left[\prod_{s=1}^{3} I_s, I_4\right]\right) \in \mathbb{C}^{I_1 I_2 I_3 \times I_4}. \tag{2.79}$$

2.5 Algorithms

In this section, some algorithms used in the rest of the thesis are introduced to estimate the factor matrices in tensor and matrix decompositions. In the case of matrix decompositions, in Section 2.5.1, the algorithms are the SVD, Khatri-Rao factorization (KRF), Kronecker factorization (KronF) and their generalizations, (GKRF) and (GKronF), respectively. To estimate the factor matrices in tensor decompositions, Section 2.5.2 presents some algorithms such as the ALS, HOSVD, tensor train hierarchical singular value decomposition (TT-HSVD) and JIRAFE, among others. These algorithms can be used to estimate factor matrices in a closed-form or iterative way.

2.5.1 Algorithms for matrix factorizations

2.5.1.1 Singular value decomposition (SVD)

The SVD of a matrix $\mathbf{X} \in \mathbb{C}^{m \times n}$ with rank-R can be written as:

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{H}, \tag{2.80}$$

where $\mathbf{U} \in \mathbb{C}^{m \times m}$ and $\mathbf{V} \in \mathbb{C}^{n \times n}$ are unitary matrices and $\mathbf{\Sigma} \in \mathbb{C}^{m \times n}$ contains the r non-zero singular values $[\sigma_1, \sigma_2, ..., \sigma_r]$ on its main diagonal and zeros elsewhere, such that the singular values are ordered by magnitude, i.e., $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r > 0$, with $r \leq \min(m,n)$.

Considering $\mathbf{X}$, the products $\mathbf{X}\mathbf{X}^{H}$ and $\mathbf{X}^{H}\mathbf{X}$ (or $\mathbf{X}\mathbf{X}^{T}$ and $\mathbf{X}^{T}\mathbf{X}$ for real matrix) are Hermitian matrices (or real symmetric matrices) and they are therefore diagonalizable by means of their eigendecompositions as [54]:

$$\mathbf{X}\mathbf{X}^{H} = \mathbf{U}\mathbf{D}_1\mathbf{U}^{H}, \tag{2.81}$$

$$\mathbf{X}^{H}\mathbf{X} = \mathbf{V}\mathbf{D}_2\mathbf{V}^{H}, \tag{2.82}$$

where $\mathbf{U}$ is the matrix of eigenvectors of $\mathbf{X}\mathbf{X}^{H}$ and $\mathbf{V}$ is the matrix of eigenvectors of $\mathbf{X}^{H}\mathbf{X}$, whose columns form two orthonormal bases, which implies $\mathbf{U}\mathbf{U}^{H} = \mathbf{U}^{H}\mathbf{U} = \mathbf{I}_m$ and $\mathbf{V}\mathbf{V}^{H} = \mathbf{V}^{H}\mathbf{V} = \mathbf{I}_n$. The columns of $\mathbf{U} \in \mathbb{C}^{m \times m}$ represent the left singular vectors of $\mathbf{X}$ and the columns of $\mathbf{V} \in \mathbb{C}^{n \times n}$ are the right singular vectors of $\mathbf{X}$. The non-zero eigenvalues of $\mathbf{X}\mathbf{X}^{H}$ and $\mathbf{X}^{H}\mathbf{X}$ are equal, non-negative and ordered by the magnitude, i.e, $\lambda_1 \geq \lambda_2 \geq ... \geq ...\lambda_r > 0$.

Now, considering only the first r rows and columns of $\mathbf{\Sigma}$ as non-zero, Eq. (2.80) can alternatively be expressed in more compact form as:

$$\mathbf{X} = \mathbf{U}_r\mathbf{\Sigma}_r\mathbf{V}_r^{H}, \tag{2.83}$$

where $\mathbf{U}_r \in \mathbb{C}^{m \times r}$ and $\mathbf{V}_r \in \mathbb{C}^{n \times r}$ contain the first r columns of $\mathbf{U}$ and $\mathbf{V}$, respectively and $\mathbf{\Sigma}_r \in \mathbb{C}^{r \times r}$ is the upper-left $(r \times r)$ block of $\mathbf{\Sigma}$.

Let us consider $\mathbf{X}$ as a rank-one matrix, its low-rank approximation is calculated by truncating its SVD to a rank-one approximation as follows:

$$\mathbf{X} = \sigma_1\mathbf{u}_1\mathbf{v}_1^{H}, \tag{2.84}$$

where $\mathbf{u}_1 \in \mathbb{C}^{m}$ and $\mathbf{v}_1 \in \mathbb{C}^{n}$ are the left and right dominant singular vectors of $\mathbf{U}$ and $\mathbf{V}$, respectively and σ_1 is the dominant singular value.

2.5.1.2 Kronecker factorization (KronF)

In this section, the KronF algorithm is presented [56, 57, 58, 59]. This algorithm estimates the matrices using a rank-one approximation of the Kronecker product. We assume a matrix $\mathbf{C} = \mathbf{A} \otimes \mathbf{B} \in \mathbb{C}^{IJ \times KQ}$ and the following minimization problem:

$$\min_{\mathbf{A},\mathbf{B}} \parallel \mathbf{C} - \mathbf{A} \otimes \mathbf{B} \parallel_F^2, \tag{2.85}$$

where $\mathbf{A} \in \mathbb{C}^{I \times K}$ and $\mathbf{B} \in \mathbb{C}^{J \times Q}$. The factor matrices of the Kronecker product can be estimated using a rank-one approximation of the matrix such that:

$$\min_{\mathbf{a},\mathbf{b}} \parallel \tilde{\mathbf{C}} - \mathbf{b}\mathbf{a}^T \parallel_F^2, \tag{2.86}$$

with $\tilde{\mathbf{C}} \in \mathbb{C}^{QJ \times KI}$, $\mathbf{a} = \text{vec}\,(\mathbf{A}) \in \mathbb{C}^{KI}$ and $\mathbf{b} = \text{vec}\,(\mathbf{B}) \in \mathbb{C}^{QJ}$. Defining the SVD of $\tilde{\mathbf{C}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H$, $\mathbf{a}$ and $\mathbf{b}$ are calculated as [58]:

$$\hat{\mathbf{a}} = \sqrt{\sigma_1}\mathbf{V}^*_{.1}, \quad \hat{\mathbf{b}} = \sqrt{\sigma_1}\mathbf{U}_{.1} \ , \tag{2.87}$$

where $\mathbf{U}_{.1} \in \mathbb{C}^{QJ}$ is the first column of the left singular vector matrix, $\mathbf{V}_{.1} \in \mathbb{C}^{KI}$ is the first column of the right singular vector matrix and σ_1 is the dominant singular value. To find $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$, we must unvectorize the vectors $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$, respectively, as:

$$\hat{\mathbf{A}} = \text{unvec}(\hat{\mathbf{a}}) \in \mathbb{C}^{I \times K}, \quad \hat{\mathbf{B}} = \text{unvec}\left(\hat{\mathbf{b}}\right) \in \mathbb{C}^{J \times Q}. \tag{2.88}$$

Note that the estimated matrices are affected by non-zero complex scaling ambiguity [58]. To solve this, we have the following relation:

$$\hat{\mathbf{A}} = \mathbf{A}\lambda^{(1)}, \quad \hat{\mathbf{B}} = \mathbf{B}\lambda^{(2)}, \quad \lambda^{(1)}\lambda^{(2)} = 1. \tag{2.89}$$

To find the scaling element, it is necessary to know one element of $\mathbf{A}$ or $\mathbf{B}$ [58, 10]. Considering the element $a_{1,1}$ of $\mathbf{A}$ known. We have for $\lambda^{(1)}$:

$$\lambda^{(1)} = \frac{\hat{a}_{1,1}}{a_{1,1}}, \tag{2.90}$$

where $\hat{a}_{1,1}$ is the element (1,1) of $\hat{\mathbf{A}}$. Now, for the matrix $\mathbf{B}$, we have:

$$\lambda^{(2)} = \left(\lambda^{(1)}\right)^{-1}. \tag{2.91}$$

By the relation (2.89), we can find the estimated matrices as:

$$\hat{\hat{\mathbf{A}}} = \hat{\mathbf{A}}\left(\lambda^{(1)}\right)^{-1}, \quad \hat{\hat{\mathbf{B}}} = \hat{\mathbf{B}}\lambda^{(1)} \ . \tag{2.92}$$

The KronF algorithm is described in Table 1.

KronF algorithm for estimation of $\hat{\hat{\mathbf{A}}}$ and $\hat{\hat{\mathbf{B}}}$

Input: matrix $\mathbf{C}$, I, J, K and Q

Output: Estimated factor matrices $\hat{\hat{\mathbf{A}}}$ and $\hat{\hat{\mathbf{B}}}$

1) Reshape $\mathbf{C} \in \mathbb{C}^{IJ \times KQ}$ as:

$$\tilde{\mathbf{C}} = \text{reshape}\left(\mathbf{C}, [QJ, KI]\right).$$

2) Calculate the SVD for $\tilde{\mathbf{C}}$:

$$\tilde{\mathbf{C}} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H.$$

3) Calculate $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ using Eq. (2.87).
4) Unvectorize $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ using Eq. (2.88).
5) Eliminate scaling ambiguities using Eqs. (2.89)-(2.91).
6) Adjust the factor matrices using Eq. (2.92).

Table 1 – Kronecker Factorization algorithm.

2.5.1.3 Generalized Kronecker factorization (GKronF)

In this subsection, the generalized KronF algorithm is summarized according to [59]. Let us consider a Kronecker product of N matrices as:

$$\mathbf{A} = \otimes_{n=1}^{N} \mathbf{A}^{(n)} \in \mathbb{C}^{I \times R}. \tag{2.93}$$

The minimization problem becomes:

$$\min_{\mathbf{A}^{(n)}, n \in [1,\ldots,N]} \parallel \mathbf{A} - \otimes_{n=1}^{N} \mathbf{A}^{(n)} \parallel_F^2, \tag{2.94}$$

where $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R_n}$ are the matrices to be estimated such that, $I = \prod_{n=1}^{N} I_n$ and $R = \prod_{n=1}^{N} R_n$. The problem (2.94) can be solved iteratively by determining the matrices $\mathbf{A}^{(n)}$ with a two-by-two search. To illustrate the GKronF algorithm, consider the case where $N = 3$, such that: $\mathbf{A} = \mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)} \otimes \mathbf{A}^{(3)}$. Then, the minimization problem becomes:

$$\min_{\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{A}^{(3)}} \parallel \mathbf{A} - \mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)} \otimes \mathbf{A}^{(3)} \parallel_F^2 . \tag{2.95}$$

The matrices $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(3)}$ in (2.95) can be estimated using the KronF algorithm presented in Table 1 to the following two decompositions of $\mathbf{A}$:

$$\mathbf{A} = \mathbf{A}^{(1)} \otimes \mathbf{A}^{(2,3)} \in \mathbb{C}^{I_1 I_2 I_3 \times R_1 R_2 R_3}, \tag{2.96}$$

$$\mathbf{A} = \mathbf{A}^{(1,2)} \otimes \mathbf{A}^{(3)} \in \mathbb{C}^{I_1 I_2 I_3 \times R_1 R_2 R_3}, \tag{2.97}$$

where $\mathbf{A}^{(2,3)} = \mathbf{A}^{(2)} \otimes \mathbf{A}^{(3)}$ and $\mathbf{A}^{(1,2)} = \mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)}$. To estimate $\mathbf{A}^{(2)}$, we use the following equation obtained by permuting the matrices $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ as:

$$\mathbf{A}_{\pi}^{(2)} = \left(\mathbf{\Pi}_{I_2, I_1} \otimes \mathbf{I}_{I_3}\right) \mathbf{A} \left(\mathbf{\Pi}_{R_2, R_1} \otimes \mathbf{I}_{R_3}\right) = \mathbf{A}^{(2)} \otimes \mathbf{A}^{(1)} \otimes \mathbf{A}^{(3)} \in \mathbb{C}^{I_2 I_1 I_3 \times R_2 R_1 R_3}, \tag{2.98}$$

$$\mathbf{A}_{\pi}^{(2)} = \mathbf{A}^{(2)} \otimes \mathbf{A}^{(1,3)} \in \mathbb{C}^{I_2 I_1 I_3 \times R_2 R_1 R_3}, \tag{2.99}$$

where $\mathbf{A}^{(1,3)} = \mathbf{A}^{(1)} \otimes \mathbf{A}^{(3)}$, $\mathbf{A}_{\pi}^{(2)}$ indicates the permutation between $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ and $\mathbf{\Pi}_{I_2,I_1} \in \mathbb{C}^{I_2 I_1 \times I_1 I_2}$, $\mathbf{\Pi}_{R_2,R_1} \in \mathbb{C}^{R_2 R_1 \times R_1 R_2}$ are defined as:

$$\mathbf{\Pi}_{I_2,I_1} = \sum_{i_2} \sum_{i_1} \left(\mathbf{e}_{i_2}^{(I_2)} \mathbf{e}_{i_1}^{(I_1)T} \right) \otimes \left(\mathbf{e}_{i_1}^{(I_1)} \mathbf{e}_{i_2}^{(I_2)T} \right) \in \mathbb{C}^{I_2 I_1 \times I_1 I_2}, \tag{2.100}$$

$$\mathbf{\Pi}_{R_2,R_1} = \sum_{r_2} \sum_{r_1} \left(\mathbf{e}_{r_2}^{(R_2)} \mathbf{e}_{r_1}^{(R_1)T} \right) \otimes \left(\mathbf{e}_{r_1}^{(R_1)} \mathbf{e}_{r_2}^{(R_2)T} \right) \in \mathbb{C}^{R_2 R_1 \times R_1 R_2}, \tag{2.101}$$

where $\mathbf{e}_{i_1}^{(I_1)}$ is the i_1-th canonical basis vector of the Euclidean space $\mathbb{R}^{I_1}$, similarly for $\mathbf{e}_{i_2}^{(I_2)}$, $\mathbf{e}_{r_1}^{(R_1)}$ and $\mathbf{e}_{r_2}^{(R_2)}$. From Eq. (2.99), we apply the KronF algorithm in Table 1 that allows to estimate $\mathbf{A}^{(2)}$. A disadvantage of this closed-form algorithm is that its performance degrades when N increases [59].

2.5.1.4 Kronecker product approximation to rank-one tensors

In this section, the Kronecker product approximation to estimate N matrices is presented. The Kronecker product of N matrices can be reorganized as a tensor and the matrices can be estimated using the THOSVD algorithm [35]. Then, considering $\mathbf{A}$ as:

$$\mathbf{A} = \mathbf{A}^{(1)} \otimes \ldots \otimes \mathbf{A}^{(N)} \in \mathbb{C}^{I_1 \ldots I_N \times R_1 \ldots R_N}, \tag{2.102}$$

where $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R_n}$ are the matrices to be estimated for $n \in [1,N]$. This operation was defined in [60] for the Kronecker product of multiple matrices. The problem (2.102) becomes:

$$\min_{\mathbf{A}^{(n)}, n \in [1,N]} \parallel \mathbf{A} - \mathbf{A}^{(1)} \otimes \ldots \otimes \mathbf{A}^{(N)} \parallel_F^2, \tag{2.103}$$

$\mathbf{A}$ can be expressed as a rank-one tensor rearranging the Kronecker product into an outer product. The problem (2.103) now becomes:

$$\min_{\mathbf{A}^{(n)}} \parallel \mathbf{A} - \mathbf{A}^{(1)} \otimes .. \otimes \mathbf{A}^{(N)} \parallel_F^2 \Leftrightarrow \min_{\mathbf{a}^{(n)}} \parallel \mathcal{A} - \mathbf{a}^{(1)} \circ \ldots \circ \mathbf{a}^{(N)} \parallel_F^2, \tag{2.104}$$

where $\mathbf{a}^{(n)}$ =vec($\mathbf{A}^{(n)}$) $\in \mathbb{C}^{R_n I_n}$ and $\mathcal{A} \in \mathbb{C}^{R_1 I_1 \times \ldots \times R_N I_N}$ is the rank-one tensor obtained by reshaping the multiple Kronecker product:

$$\mathcal{A} = \text{reshape}\left(\mathbf{A}, [R_1 I_1, \ldots, R_N I_N]\right). \tag{2.105}$$

The vectors $\mathbf{a}^{(n)}$ are estimated using the THOSVD algorithm and the estimated factor matrix $\hat{\mathbf{A}}^{(n)}$ is deduced using the unvec operator [60, 61, 54]. Each estimate $\hat{\mathbf{A}}^{(n)}$ is subject to a scalar scaling ambiguity $\lambda^{(n)}$ such as the corrected estimate is given by:

$$\hat{\hat{\mathbf{A}}}^{(n)} = \hat{\mathbf{A}}^{(n)} \lambda^{(n)} \quad \text{with} \quad \prod_{n=1}^{N} \lambda^{(n)} = 1, \tag{2.106}$$

$\lambda^{(n)}$ is determined from the knowledge of one element of $\mathbf{A}^{(n)}$.

2.5.1.5 Khatri-Rao factorization (KRF)

In some applications covered in this thesis, we will be interested in approximating a Khatri-Rao product between two factor matrices. To solve it, we make use of the Khatri-Rao factorization (KRF) proposed in [62]. This factorization estimates the factor matrices through a rank-one approximation. We assume a matrix $\mathbf{A} \in \mathbb{C}^{IJ \times K}$ given by:

$$\mathbf{A} = \mathbf{B} \diamond \mathbf{C} \in \mathbb{C}^{IJ \times K}, \tag{2.107}$$

where $\mathbf{B} \in \mathbb{C}^{I \times K}$ and $\mathbf{C} \in \mathbb{C}^{J \times K}$. The minimization problem becomes:

$$\min_{\mathbf{B},\mathbf{C}} \| \mathbf{A} - \mathbf{B} \diamond \mathbf{C} \|_F^2 \,. \tag{2.108}$$

The Khatri-Rao product can be seen as a column-wise Kronecker product and we can calculate a rank-one approximation of the matrix defined for each column ($k \in [1,K]$) as:

$$\mathbf{Y}_k = \text{unvec}\left(\mathbf{A}_{.k}\right) = \mathbf{C}_{.k}\mathbf{B}_{.k}^T, \tag{2.109}$$

where $\mathbf{B}_{.k} \in \mathbb{C}^I$ and $\mathbf{C}_{.k} \in \mathbb{C}^J$ denote the k-th column of $\mathbf{B}$ and $\mathbf{C}$, respectively. Therefore, since $\mathbf{Y}_k \in \mathbb{C}^{J \times I}$ is a rank-one matrix, the best estimate of $\mathbf{B}_{.k}$ and $\mathbf{C}_{.k}$ is found by solving the least squares (LS) problem by the SVD of $\mathbf{Y}_k$ defined as $\mathbf{Y}_k = \mathbf{U}_k\mathbf{\Sigma}_k\mathbf{V}_k^H$. The k-th column of $\mathbf{B}$ and $\mathbf{C}$ are given by [62]:

$$\hat{\mathbf{B}}_{.k} = \sqrt{\sigma_{1,k}}\left(\mathbf{V}_k^*\right)_{.1}, \quad \hat{\mathbf{C}}_{.k} = \sqrt{\sigma_{1,k}}\left(\mathbf{U}_k\right)_{.1}, \tag{2.110}$$

where $\left(\mathbf{U}_k\right)_{.1} \in \mathbb{C}^J$ and $\left(\mathbf{V}_k\right)_{.1} \in \mathbb{C}^I$ are the dominant left and right singular vectors of $\mathbf{U}_k$ and $\mathbf{V}_k$, respectively and σ_{1k} is the dominant singular value of $\mathbf{\Sigma}_k$. The estimated vectors are affected by a non-zero complex scaling ambiguity [62]. For the scaling ambiguities, we have the following relation:

$$\hat{\mathbf{B}} = \mathbf{B}\mathbf{\Lambda}^{(1)}, \quad \hat{\mathbf{C}} = \mathbf{C}\mathbf{\Lambda}^{(2)}, \quad \mathbf{\Lambda}^{(1)}\mathbf{\Lambda}^{(2)} = \mathbf{I}_K \,. \tag{2.111}$$

It is necessary to know one row of the original factor matrices $\mathbf{B}$ or $\mathbf{C}$ to find the scaling matrices [62]. Let us consider the first row of the matrix $\mathbf{B}$ known. We have for each element of $\mathbf{\Lambda}^{(1)}$:

$$\lambda_k^{(1)} = \frac{\hat{b}_{1,k}}{b_{1,k}}, \tag{2.112}$$

where $\hat{b}_{1,k}$ and $b_{1,k}$ are the elements of the first row of $\hat{\mathbf{B}}$ and $\mathbf{B}$, respectively for $k \in [1, K]$. $\mathbf{\Lambda}^{(1)}$ corresponds to a scaling matrix with the elements $\lambda_k^{(1)}$ on the main diagonal. By the relation (2.111), we can find the estimated matrices as:

$$\hat{\hat{\mathbf{B}}} = \hat{\mathbf{B}}\left(\mathbf{\Lambda}^{(1)}\right)^{-1}, \quad \hat{\hat{\mathbf{C}}} = \hat{\mathbf{C}}\mathbf{\Lambda}^{(1)} \,. \tag{2.113}$$

The KRF algorithm for matrices estimation is summarized in Table 2.

```
KRF algorithm for estimation of $\hat{\hat{\mathbf{B}}}$ and $\hat{\hat{\mathbf{C}}}$
Input: matrix $\mathbf{A}$, $I$, $J$ and $K$
Output: Estimated factor matrices $\hat{\hat{\mathbf{B}}}$ and $\hat{\hat{\mathbf{C}}}$
1) Unvectorize the matrix $\mathbf{A} \in \mathbb{C}^{IJ \times K}$:
for $k = 1 : K$
    $\mathbf{Y}_k = \text{unvec}\left(\mathbf{A}_{.k}\right),$
    2) Calculate the SVD for each column of $\mathbf{Y}$:
    $\mathbf{Y}_k = \mathbf{U}_k \mathbf{\Sigma}_k \mathbf{V}_k^H.$
    3) Calculate $\hat{\mathbf{B}}$ and $\hat{\mathbf{C}}$ using Eq. (2.110).
end
4) Store in the matrices as:
    $\hat{\mathbf{B}} = \left[\ \sqrt{\sigma_{1,1}}(\mathbf{V}_1)_{.1} \ \ \dots \ \ \sqrt{\sigma_{1,K}}(\mathbf{V}_K)_{.1}\ \right]^* \in \mathbb{C}^{I \times K},$
    $\hat{\mathbf{C}} = \left[\ \sqrt{\sigma_{1,1}}(\mathbf{U}_1)_{.1} \ \ \dots \ \ \sqrt{\sigma_{1,K}}(\mathbf{U}_K)_{.1}\ \right] \in \mathbb{C}^{J \times K}.$
5) Eliminate scaling ambiguities using Eqs. (2.111)-(2.112).
6) Adjust the factor matrices using Eq. (2.113).
```

Table 2 – Khatri-Rao Factorization algorithm.

2.5.2 Higher-order singular value decomposition (HOSVD)

The HOSVD of an N-order tensor $\mathcal{Y} \in \mathbb{C}^{I_1 \times \dots \times I_N}$ with multilinear-rank-$(R_1, \dots, R_N)$ is defined as [35]:

$$\mathcal{Y} = \mathcal{S} \times_1 \mathbf{U}^{(1)} \times_2 \dots \times_N \mathbf{U}^{(N)} \in \mathbb{C}^{I_1 \times \dots \times I_N}, \tag{2.114}$$

where $\mathbf{U}^{(n)} \in \mathbb{C}^{I_n \times R_n}$ are the left singular vector matrices of each n-mode unfolding, with $n \in [1, N]$ and $\mathcal{S} \in \mathbb{C}^{R_1 \times \dots \times R_N}$ being the core tensor. The matrix $\mathbf{U}^{(n)}$ is computed from the n-th matrix unfolding of $\mathcal{Y}$, i.e.:

$$\mathbf{Y}_{I_n \times I_{n+1} \dots I_N I_1 \dots I_{n-1}} = \mathbf{U}^{(n)} \mathbf{\Sigma}^{(n)} \mathbf{V}^{(n)H}. \tag{2.115}$$

From $\mathbf{U}^{(n)}$ computed according to Eq. (2.115), the core tensor can be obtained as:

$$\mathcal{S} = \mathcal{Y} \times_1 \mathbf{U}^{(1)H} \times_2 \dots \times_N \mathbf{U}^{(N)H}. \tag{2.116}$$

Now, considering the third-order Tucker decomposition of $\mathcal{Y} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ in Eq. (2.44), the HOSVD algorithm computes a base for each factor matrix via the SVD of each n-mode matrix unfolding of $\mathcal{Y}$ as:

$$\mathbf{Y}_{I_1 \times I_2 I_3} = \mathbf{U}^{(1)} \mathbf{\Sigma}^{(1)} \mathbf{V}^{(1)H}, \tag{2.117}$$

$$\mathbf{Y}_{I_2 \times I_3 I_1} = \mathbf{U}^{(2)} \mathbf{\Sigma}^{(2)} \mathbf{V}^{(2)H}, \tag{2.118}$$

$$\mathbf{Y}_{I_3 \times I_1 I_2} = \mathbf{U}^{(3)} \mathbf{\Sigma}^{(3)} \mathbf{V}^{(3)H}, \tag{2.119}$$

where $\mathbf{U}^{(n)}$ is a matrix that spans the subspace of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, for $n \in [1, 3]$, respectively. The HOSVD procedure is described in Table 3.

HOSVD algorithm for estimation of $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, $\hat{\mathbf{C}}$ and $\hat{\mathcal{G}}$

Input: tensor $\mathcal{Y}$

Output: Estimated factor matrices $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, $\hat{\mathbf{C}}$ and $\hat{\mathcal{G}}$

1) Compute the SVD for n-mode unfolding of $\mathcal{Y}$:

$$\mathbf{Y}_{I_1 \times I_2 I_3} = \mathbf{U}^{(1)}\mathbf{\Sigma}^{(1)}\mathbf{V}^{(1)H},$$

$$\mathbf{Y}_{I_2 \times I_3 I_1} = \mathbf{U}^{(2)}\mathbf{\Sigma}^{(2)}\mathbf{V}^{(2)H},$$

$$\mathbf{Y}_{I_3 \times I_1 I_2} = \mathbf{U}^{(3)}\mathbf{\Sigma}^{(3)}\mathbf{V}^{(3)H}.$$

2) Compute $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$ and $\hat{\mathbf{C}}$ as:

$$\hat{\mathbf{A}} = \mathbf{U}^{(1)},$$

$$\hat{\mathbf{B}} = \mathbf{U}^{(2)},$$

$$\hat{\mathbf{C}} = \mathbf{U}^{(3)}.$$

3) Compute the core tensor $\hat{\mathcal{G}}$:

$$\hat{\mathcal{G}} = \mathcal{Y} \times_1 \hat{\mathbf{A}}^H \times_2 \hat{\mathbf{B}}^H \times_3 \hat{\mathbf{C}}^H.$$

Table 3 – HOSVD algorithm.

2.5.3 Alternating least squares (ALS) algorithm

The ALS algorithm is an iterative algorithm applied to estimate the factor matrices. Based on Eqs. (2.22)-(2.24) for PARAFAC decomposition, we can formulate the classic ALS algorithm for iterative estimation of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ from the tensor $\mathcal{X}$ [39]. The ALS algorithm results from the following optimization problem deduced from Eq. (2.21):

$$\min_{\mathbf{A},\mathbf{B},\mathbf{C}} \| \mathcal{X} - \mathcal{I}_{3,R} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C} \|_F^2 . \tag{2.120}$$

Fixing one of the matrix factors ($\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$) with its value estimated at the previous iteration, the ALS method replaces the optimization problem (2.120) by three LS sub-problems deduced from the matrix unfoldings (2.22)-(2.24), leading to the alternate minimization of the following LS criteria:

$$\min_{\mathbf{A}} \| \mathbf{X}_{JK \times I} - \left(\hat{\mathbf{B}}_{[l-1]} \diamond \hat{\mathbf{C}}_{[l-1]}\right) \mathbf{A}^T \|_F^2 \longrightarrow \hat{\mathbf{A}}_{[l]}, \tag{2.121}$$

$$\min_{\mathbf{B}} \| \mathbf{X}_{KI \times J} - \left(\hat{\mathbf{C}}_{[l-1]} \diamond \hat{\mathbf{A}}_{[l]}\right) \mathbf{B}^T \|_F^2 \longrightarrow \hat{\mathbf{B}}_{[l]}, \tag{2.122}$$

$$\min_{\mathbf{C}} \| \mathbf{X}_{IJ \times K} - \left(\hat{\mathbf{A}}_{[l]} \diamond \hat{\mathbf{B}}_{[l]}\right) \mathbf{C}^T \|_F^2 \longrightarrow \hat{\mathbf{C}}_{[l]}. \tag{2.123}$$

The matrices $(\hat{\mathbf{B}} \diamond \hat{\mathbf{C}})$, $(\hat{\mathbf{C}} \diamond \hat{\mathbf{A}})$ and $(\hat{\mathbf{A}} \diamond \hat{\mathbf{B}})$ must have full column rank for ensuring uniqueness of the LS estimates, which implies the following necessary condition: $R \leq \min(JK, KI, IJ)$. The update equations at iteration $[l]$ are given by:

$$\hat{\mathbf{A}}_{[l]} = \left[\left(\hat{\mathbf{B}}_{[l-1]} \diamond \hat{\mathbf{C}}_{[l-1]} \right)^{\dagger} \mathbf{X}_{JK \times I} \right]^{T}, \tag{2.124}$$

$$\hat{\mathbf{B}}_{[l]} = \left[\left(\hat{\mathbf{C}}_{[l-1]} \diamond \hat{\mathbf{A}}_{[l]} \right)^{\dagger} \mathbf{X}_{KI \times J} \right]^{T}, \tag{2.125}$$

$$\hat{\mathbf{C}}_{[l]} = \left[\left(\hat{\mathbf{A}}_{[l]} \diamond \hat{\mathbf{B}}_{[l]} \right)^{\dagger} \mathbf{X}_{K \times IJ} \right]^{T}, \tag{2.126}$$

where $\left(\hat{\mathbf{B}} \diamond \hat{\mathbf{C}}\right)^{\dagger}$, $\left(\hat{\mathbf{C}} \diamond \hat{\mathbf{A}}\right)^{\dagger}$ and $\left(\hat{\mathbf{A}} \diamond \hat{\mathbf{B}}\right)^{\dagger}$ denote the pseudo-inverses of $\left(\hat{\mathbf{B}} \diamond \hat{\mathbf{C}}\right)$, $\left(\hat{\mathbf{C}} \diamond \hat{\mathbf{A}}\right)$ and $\left(\hat{\mathbf{A}} \diamond \hat{\mathbf{B}}\right)$, respectively. Taking into account the property 2 of the Khatri-Rao product, the computation of the pseudo-inverse in the ALS algorithm can be simplified as follows:

$$\begin{aligned} \left(\hat{\mathbf{B}}_{[l-1]} \diamond \hat{\mathbf{C}}_{[l-1]}\right)^{\dagger} &= \left[(\hat{\mathbf{B}}_{[l-1]} \diamond \hat{\mathbf{C}}_{[l-1]})^{H} (\hat{\mathbf{B}}_{[l-1]} \diamond \hat{\mathbf{C}}_{[l-1]}) \right]^{-1} \left(\hat{\mathbf{B}}_{[l-1]} \diamond \hat{\mathbf{C}}_{[l-1]}\right)^{H} \\ &= \left(\hat{\mathbf{B}}_{[l-1]}^{H} \hat{\mathbf{B}}_{[l-1]} \odot \hat{\mathbf{C}}_{[l-1]}^{H} \hat{\mathbf{C}}_{[l-1]} \right)^{-1} \left(\hat{\mathbf{B}}_{[l-1]} \diamond \hat{\mathbf{C}}_{[l-1]}\right)^{H}, \end{aligned} \tag{2.127}$$

$$\begin{aligned} \left(\hat{\mathbf{C}}_{[l]} \diamond \hat{\mathbf{A}}_{[l-1]}\right)^{\dagger} &= \left[(\hat{\mathbf{C}}_{[l]} \diamond \hat{\mathbf{A}}_{[l-1]})^{H} (\hat{\mathbf{C}}_{[l]} \diamond \hat{\mathbf{A}}_{[l-1]}) \right]^{-1} \left(\hat{\mathbf{C}}_{[l]} \diamond \hat{\mathbf{A}}_{[l-1]}\right)^{H} \\ &= \left(\hat{\mathbf{C}}_{[l]}^{H} \hat{\mathbf{C}}_{[l]} \odot \hat{\mathbf{A}}_{[l-1]}^{H} \hat{\mathbf{A}}_{[l-1]} \right)^{-1} \left(\hat{\mathbf{C}}_{[l]} \diamond \hat{\mathbf{A}}_{[l-1]}\right)^{H}, \end{aligned} \tag{2.128}$$

$$\begin{aligned} \left(\hat{\mathbf{A}}_{[l]} \diamond \hat{\mathbf{B}}_{[l]}\right)^{\dagger} &= \left[(\hat{\mathbf{A}}_{[l]} \diamond \hat{\mathbf{B}}_{[l]})^{H} (\hat{\mathbf{A}}_{[l]} \diamond \hat{\mathbf{B}}_{[l]}) \right]^{-1} \left(\hat{\mathbf{A}}_{[l]} \diamond \hat{\mathbf{B}}_{[l]}\right)^{H} \\ &= \left(\hat{\mathbf{A}}_{[l]}^{H} \hat{\mathbf{A}}_{[l]} \odot \hat{\mathbf{B}}_{[l]}^{H} \hat{\mathbf{B}}_{[l]} \right)^{-1} \left(\hat{\mathbf{B}}_{[l]} \diamond \hat{\mathbf{A}}_{[l]}\right)^{H}. \end{aligned} \tag{2.129}$$

This amounts to replacing the computation of the pseudo-inverses of matrices of size $KJ \times R$, $IK \times R$ and $JI \times R$ by the computation of the inverses of three matrices of size $R \times R$. For deciding the convergence of the ALS algorithm, we consider the error at the $[l]$-th iteration deduced from (2.24) as:

$$\epsilon_{[l]} = \| \mathbf{X}_{IJ \times K} - \left(\hat{\mathbf{A}}_{[l]} \diamond \hat{\mathbf{B}}_{[l]}\right) \hat{\mathbf{C}}_{[l]}^{T} \|_{F}^{2} \, . \tag{2.130}$$

Convergence at the $[l]$-th iteration is declared when this error does not significantly change between two successive iterations, i.e., $|\epsilon_{[l-1]} - \epsilon_{[l]}| \leq \delta$, where $\delta > 0$ is a predefined threshold assumed to be 10^{-6} throughout this thesis. ALS algorithm is summarized in Table 4.

2.5.4 Tensor train-singular value decomposition (TT-SVD) algorithm

The TT decomposition of an arbitrary tensor can be calculated using successive SVDs [46]. Given a fourth-order tensor $\mathcal{Y} \in \mathbb{C}^{I_1 \times I_2 \times I_3 \times I_4}$ decomposed in terms of the TT model as

```
**ALS algorithm for estimation of $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$ and $\hat{\mathbf{C}}$**
**Input:** tensor $\mathcal{X}$
**Output:** Estimated matrices $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$ and $\hat{\mathbf{C}}$
$l = 0$
1) Randomly initialize $\mathbf{B}_{[0]}$ and $\mathbf{C}_{[0]}$.
2) Update the estimates of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ as:
    $\hat{\mathbf{A}}_{[l]} = \left[\left(\hat{\mathbf{B}}_{[l-1]} \diamond \hat{\mathbf{C}}_{[l-1]}\right)^{\dagger} \mathbf{X}_{JK \times I}\right]^{T},$
    $\hat{\mathbf{B}}_{[l]} = \left[\left(\hat{\mathbf{C}}_{[l-1]} \diamond \hat{\mathbf{A}}_{[l]}\right)^{\dagger} \mathbf{X}_{KI \times J}\right]^{T},$
    $\hat{\mathbf{C}}_{[l]} = \left[\left(\hat{\mathbf{A}}_{[l]} \diamond \hat{\mathbf{B}}_{[l]}\right)^{\dagger} \mathbf{X}_{K \times IJ}\right]^{T}.$
3) Calculate the error function (2.130) and $err = |\epsilon_{[l-1]} - \epsilon_{[l]}|$.
- **if** $|\epsilon_{[l-1]} - \epsilon_{[l]}| < \delta$ or $l$ =maximum number of iterations
- **stop**
- **else** $l \rightarrow l+1$;
$\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$ and $\hat{\mathbf{C}}$
**end**
```

Table 4 – Alternating least square algorithm.

in Eq. (2.74), with $\mathbf{G}^{(1)} \in \mathbb{C}^{I_1 \times R_1}$, $\mathcal{G}^{(2)} \in \mathbb{C}^{R_1 \times I_2 \times R_2}$, $\mathcal{G}^{(3)} \in \mathbb{C}^{R_2 \times I_3 \times R_3}$ and $\mathbf{G}^{(4)} \in \mathbb{C}^{R_3 \times I_4}$. The TT decomposition can be computed by a sequence of SVD decompositions, in which the SVDs of each unfolding of $\mathcal{Y}$ are calculated [45, 46]. So, first we compute the unfolding $\mathbf{Y}_{I_1 \times I_2 I_3 I_4}$ as in (2.77) such that:

$$\mathbf{Y}^{(1)} = \text{reshape}\,(\mathcal{Y}, [I_1, I_2 I_3 I_4]) \in \mathbb{C}^{I_1 \times I_2 I_3 I_4}. \tag{2.131}$$

Subsequently the SVD of the unfolding $\mathbf{Y}^{(1)}$ is calculated as:

$$\mathbf{Y}^{(1)} = \mathbf{U}^{(1)} \mathbf{\Sigma}^{(1)} \mathbf{V}^{(1)H} = \mathbf{U}^{(1)} \mathbf{Z}^{(1)}, \tag{2.132}$$

where $\mathbf{Z}^{(1)} = \mathbf{\Sigma}^{(1)} \mathbf{V}^{(1)H} \in \mathbb{C}^{R_1 \times I_2 I_3 I_4}$ and $\mathbf{U}^{(1)} \in \mathbb{C}^{I_1 \times R_1}$, with R_1 =rank($\mathbf{Y}^{(1)}$). The matrix $\mathbf{G}^{(1)}$ is given by:

$$\mathbf{G}^{(1)} = \mathbf{U}^{(1)} \in \mathbb{C}^{I_1 \times R_1}. \tag{2.133}$$

The unfolding $\mathbf{Y}^{(2)}$ is calculated as follow:

$$\mathbf{Y}^{(2)} = \text{reshape}\left(\mathbf{Z}^{(1)}, [R_1 I_2, I_3 I_4]\right) \in \mathbb{C}^{R_1 I_2 \times I_3 I_4}. \tag{2.134}$$

The SVD of the unfolding $\mathbf{Y}^{(2)}$ is calculated as:

$$\mathbf{Y}^{(2)} = \mathbf{U}^{(2)} \mathbf{\Sigma}^{(2)} \mathbf{V}^{(2)H} = \mathbf{U}^{(2)} \mathbf{Z}^{(2)}, \tag{2.135}$$

where $\mathbf{Z}^{(2)} = \mathbf{\Sigma}^{(2)}\mathbf{V}^{(2)H} \in \mathbb{C}^{R_2 \times I_3 I_4}$ and $\mathbf{U}^{(2)} \in \mathbb{C}^{R_1 I_2 \times R_2}$, with R_2 =rank($\mathbf{Y}^{(2)}$). The tensor $\mathcal{G}^{(2)}$ is given by:

$$\mathcal{G}^{(2)} = \text{reshape}\left(\mathbf{U}^{(2)}, [R_1, I_2, R_2]\right) \in \mathbb{C}^{R_1 \times I_2 \times R_2}. \tag{2.136}$$

For the unfolding $\mathbf{Y}^{(3)}$ we have:

$$\mathbf{Y}^{(3)} = \text{reshape}\left(\mathbf{Z}^{(2)}, [R_2 I_3, I_4]\right) \in \mathbb{C}^{R_2 I_3 \times I_4}. \tag{2.137}$$

The SVD of $\mathbf{Y}^{(3)}$ is given by:

$$\mathbf{Y}^{(3)} = \mathbf{U}^{(3)}\mathbf{\Sigma}^{(3)}\mathbf{V}^{(3)H} = \mathbf{U}^{(3)}\mathbf{Z}^{(3)}, \tag{2.138}$$

where $\mathbf{Z}^{(3)} = \mathbf{\Sigma}^{(3)}\mathbf{V}^{(3)H} \in \mathbb{C}^{R_3 \times I_4}$ and $\mathbf{U}^{(3)} \in \mathbb{C}^{R_2 I_3 \times R_3}$, with R_3 =rank($\mathbf{Y}^{(3)}$). The tensor $\mathcal{G}^{(3)}$ is given by:

$$\mathcal{G}^{(3)} = \text{reshape}\left(\mathbf{U}^{(3)}, [R_2, I_3, R_3]\right) \in \mathbb{C}^{R_2 \times I_3 \times R_3}. \tag{2.139}$$

To find the matrix $\mathbf{G}^{(4)}$ we have:

$$\mathbf{G}^{(4)} = \mathbf{Z}^{(3)} \in \mathbb{C}^{R_3 \times I_4}. \tag{2.140}$$

Figure 8 illustrates the TT-SVD algorithm to calculate the TT-cores of the fourth-order $\mathcal{Y} \in \mathbb{C}^{I_1 \times I_2 \times I_3 \times I_4}$.

2.5.5 Tensor train-hierarchical singular value decomposition (TT-HSVD) algorithm

The TT-HSVD algorithm was presented by Y. Zniyed *et al.* [63] and this algorithm derives the TT-cores in a parallel hierarchical way. The main difference between the tensor train singular value decomposition (TT-SVD) and TT-HSVD algorithms lies in the initial matrix unfolding to be processed by the SVD and the reshaping strategy. In Figure 9, TT-HSVD algorithm is presented to calculate the TT-cores of the fourth-order tensor $\mathcal{Y} \in \mathbb{C}^{I_1 \times I_2 \times I_3 \times I_4}$.

For the TT-HSVD algorithm, considering an a priori chosen index $\overline{N} \in [1, N]$, the first matrix unfolding $\mathbf{Y}^{(\overline{N})}$ is of size $(I_1 ... I_{\overline{N}}) \times (I_{\overline{N}+1} ... I_N)$, which leads to a more rectangular matrix in relation to the matrix unfolding of TT-SVD. Its $R_{\overline{N}}$-truncated SVD provides two factors $\mathbf{U}^{(n)}_{\overline{N}} \in \mathbb{C}^{I_1 ... I_{\overline{N}} \times R_{\overline{N}}}$ and $\mathbf{V}^{(n)}_{\overline{N}} \in \mathbb{C}^{R_{\overline{N}} \times I_{\overline{N}+1} ... I_N}$. These two factors are now reshaped in parallel, which constitutes the main difference with the TT-SVD algorithm for which only a single reshaping operation is applied to $\mathbf{V}^{(1)}_2$ [63]. This process is repeated after each SVD computation.

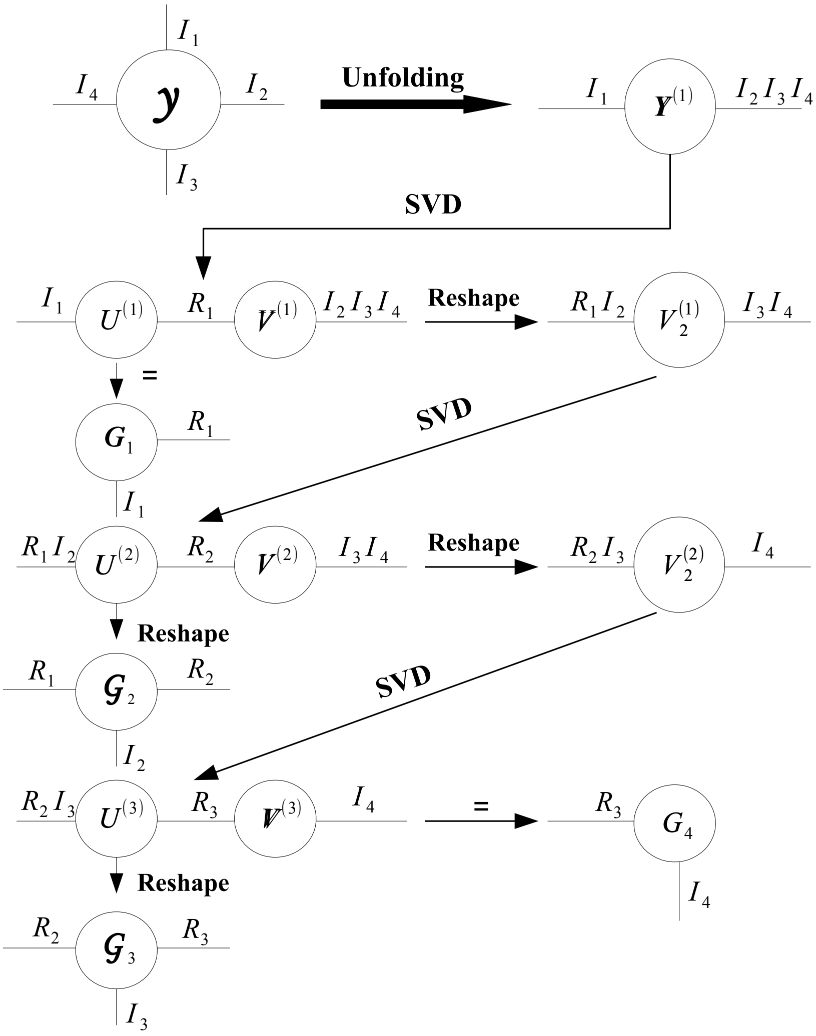


Figure 8 – TT-SVD algorithm applied on the fourth-order tensor $\mathcal{Y} \in \mathbb{C}^{I_1 \times I_2 \times I_3 \times I_4}$ [34].

2.5.6 PARAFAC and TT decompositions equivalence and JIRAFE algorithm

In this section, we present the equivalence between an N-order PARAFAC decomposition and an N-order TT decomposition for dimensionality reduction. The equivalence between these two models was presented in [34, 64]. Given an N-order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \ldots \times I_N}$ that admits a rank-R PARAFAC decomposition with factor matrices $\mathbf{A}^{(n)}$ of full column rank defined as:

$$\mathcal{X} = \mathcal{I}_{N,R} \times_1 \mathbf{A}^{(1)} \times_2 \ldots \times_N \mathbf{A}^{(N)} \in \mathbb{C}^{I_1 \times \ldots \times I_N}, \tag{2.141}$$

where $\mathcal{I}_{N,R} \in \mathbb{R}^{R \times \ldots \times R}$ and $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R}$ are the factor matrices. Then its TT decomposition involves a train of two matrices and $(N-2)$ tensors, each modeled as a third-order PARAFAC decomposition with the same TT-rank $R_1 = R_2 = \ldots = R_{N-1} = R$. The factor matrices in Eq. (2.141) can then be directly derived from the TT-cores up to $(N-1)$ basis-change matrices such as [34]:

$$\mathbf{G}^{(1)} = \mathbf{A}^{(1)} \mathbf{M}_1^{-1} \in \mathbb{C}^{I_1 \times R}, \tag{2.142}$$

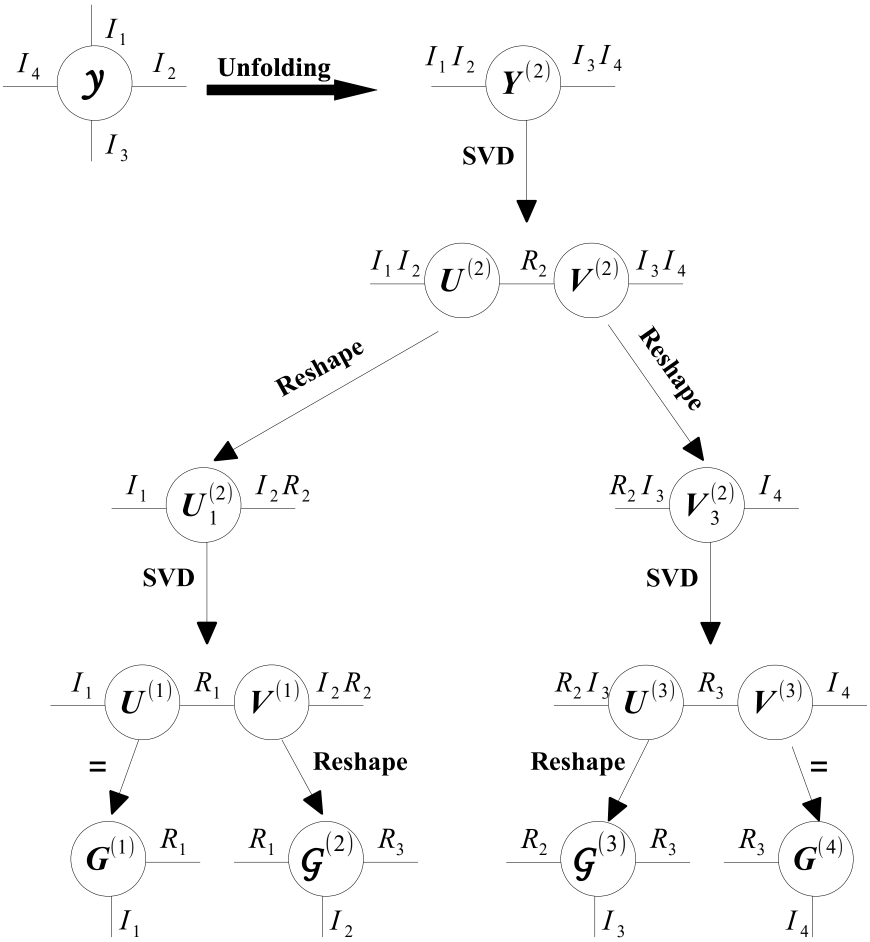


Figure 9 – TT-HSVD algorithm applied on the fourth-order tensor $\mathcal{Y} \in \mathbb{C}^{I_1 \times I_2 \times I_3 \times I_4}$ [63].

$$\mathcal{G}^{(k)} = \mathcal{I}_{3,R} \times_1 \mathbf{M}_{k-1} \times_2 \mathbf{A}^{(k)} \times_k \mathbf{M}_k^{-T} \in \mathbb{C}^{R \times I_k \times R}, 2 \leq k \leq N-1, \tag{2.143}$$

$$\mathbf{G}^{(N)} = \mathbf{M}_{N-1}[\mathbf{A}^{(N)}]^T \in \mathbb{C}^{R \times I_N}, \tag{2.144}$$

where $\mathbf{M}_k \in \mathbb{C}^{R \times R}$ for $k \in [1, N-1]$ are nonsingular basis-change matrices. Note that two neighboring core tensors $\mathcal{G}^{(k)}$ and $\mathcal{G}^{(k+1)}$ satisfy coupled third-order PARAFACs which share the common factor $\mathbf{M}_k$. To estimate the factor matrices of the PARAFAC decomposition, the JIRAFE algorithm summarized in Table 5, can be used. This algorithm has been successfully applied in the context of multidimensional harmonic retrieval [64]. The TT-SVD or TT-HSVD algorithms, described in Sections 2.5.4 and 2.5.5, respectively, is used to determine the TT-core estimates $(\mathbf{G}^{(1)}, \mathcal{G}^{(2)}, ..., \mathcal{G}^{(N-1)}, \mathbf{G}^{(N)})$. Second, the factors $(\mathbf{M}_1, \mathbf{A}^{(2)}, \mathbf{M}_2)$ are estimated by means of the ALS method. Then, the other factor matrices are estimated using the KRF method, recalled in Subsection 2.5.1.4. Thus, for $3 \leq k \leq N-1$, the KRF algorithm allows to estimate $\mathbf{A}^{(k)}$ and $\mathbf{M}_k^{-T}$ from the LS estimate of their Khatri-Rao product:

$$\widehat{\mathbf{A}^{(k)} \diamond \mathbf{M}_k^{-T}} = \left(\mathbf{M}_{k-1}^{\dagger} \hat{\mathbf{G}}^{(k)}_{R \times I_{k-1} R} \right)^T . \tag{2.145}$$

The combination of the TTHSVD, ALS and KRF methods to estimate the PARAFAC factor matrices results in the JIRAFE algorithm, summarized in Table 5, where $[\hat{\mathbf{M}}_1, \hat{\mathbf{A}}^{(2)}, \hat{\mathbf{M}}_2^{-T}]$ $= \text{ALS}(\hat{\mathcal{G}}^{(2)}, R)$ means that the ALS algorithm is applied to the third-order tensor PARAFAC decomposition (2.143) for $k = 2$. Then, the other factor matrices are estimated using the KRF method applied to (2.143) for $3 \leq k \leq N-1$, with the matrix $\mathbf{M}_{k-1}$ estimated in the previous step. Thus, the KRF algorithm allows to estimate $\hat{\mathbf{A}}^{(2)}$ and $\mathbf{M}_k^{-T}$ from the LS estimate of their Khatri-Rao product as:

$$\widehat{\mathbf{A}^{(k)} \diamond \mathbf{M}_k^{-T}} = \mathbf{G}^{(k)}_{I_kR\times R}\hat{\mathbf{M}}_{k-1}^{-T}, \tag{2.146}$$

$\mathbf{A}^{(1)}$ and $\mathbf{A}^{(N)}$ are estimated using the LS algorithm applied to Eqs. (2.142). JIRAFE algorithm has been successfully applied in the context of multidimensional harmonic retrieval [64]. It allows us to reduce the computational complexity while avoiding ill-convergence problems encountered with the standard ALS algorithms applied to a higher-order PARAFAC, as slow convergence and possible convergence towards local minima. With JIRAFE, the ALS algorithm is applied to a third-order PARAFAC instead of an n-order one.

JIRAFE algorithm for the estimation of $\hat{\mathbf{A}}^{(1)}, \ldots, \hat{\mathbf{A}}^{(N)}$.

Input: N-order tensor $\mathcal{X}$ of rank-R

Output: Estimated matrices $\hat{\mathbf{A}}^{(1)}, \ldots, \hat{\mathbf{A}}^{(N)}$

1) **Dimensionality reduction:** estimate the TT-cores by means of the TT-SVD/TT-HSVD algorithm applied to $\mathcal{X}$:

$$[\hat{\mathbf{G}}^{(1)}, \hat{\mathcal{G}}^{(2)}, \ldots, \hat{\mathcal{G}}^{(N-1)}, \hat{\mathbf{G}}^{(N)}] \leftarrow \text{TT}-\text{SVD}/\text{TT}-\text{HSVD}\left(\mathcal{X}, R\right).$$

2) **PARAFAC factors retrieval:** estimate the factor matrices by means of the ALS and KRF algorithms:

$$[\hat{\mathbf{M}}_1, \hat{\mathbf{A}}^{(2)}, \hat{\mathbf{M}}_2^{-T}] \leftarrow \text{ALS}\left(\hat{\mathcal{G}}^{(2)}, R\right).$$

for $k \in [3, N-1]$

$$[\hat{\mathbf{A}}^{(k)}, \hat{\mathbf{M}}_k^{-T}] \leftarrow \text{KRF}\left(\hat{\mathcal{G}}^{(k)}, \hat{\mathbf{M}}_{k-1}, R\right),$$

end

$$\hat{\mathbf{A}}^{(1)} = \hat{\mathbf{G}}^{(1)}\hat{\mathbf{M}}_1, \quad \hat{\mathbf{A}}^{(N)} = [\hat{\mathbf{G}}^{(N)}]^T\hat{\mathbf{M}}_{N-1}^{-T} \; .$$

Table 5 – JIRAFE algorithm.

2.5.7 Rectified alternating least squares (RectALS) algorithm

Several variants of the ALS algorithm have been proposed in the literature. Here, the RectALS is summarized in Table 6. For more information see [65, 64]. The RectALS algorithm takes into account the Vandermonde structure of the factor matrices, which is based on iterative projections and LS criteria. Considering the TT-core $\mathcal{G}^{(2)}$ estimated by JIRAFE, the RectALS algorithm is applied to the third-order TT-core $\mathcal{G}^{(2)}$ to retrieve the matrices $\mathbf{M}_1$, $\mathbf{A}^{(2)}$ and $\mathbf{M}_2$, where one of the matrices of this model has a Vandermonde structure. The Vandermonde matrix $\mathbf{A}^{(2)} \in \mathbb{C}^{I_2 \times R}$ of rank R is defined as [64]:

$$\mathbf{A}^{(2)} = [\ \mathbf{a}^{(2)}(z_{I_2,1}) \ \ \dots \ \ \mathbf{a}^{(2)}(z_{I_2,R})] \in \mathbb{C}^{I_2 \times R}, \tag{2.147}$$

where each column $\mathbf{a}^{(2)}(z_{I_2,r})$ depends on a single parameter $(z_{I_2,r})$ of unit modulus, $z_{I_2,r} = e^{i\omega_{I_2,r}}$ is the pole and $\omega_{I_2,r}$ is the r-th angular frequency along the i_2-th dimension. $\mathbf{a}^{(2)}(z_{I_2,r})$ is defined as:

$$\mathbf{a}^{(2)}(z_{I_2,r}) = [\ z^0_{I_2,r} \ \ z^1_{I_2,r} \ \ z^2_{I_2,r} \ \ \dots \ \ z^{I_2-1}_{I_2,r} \ ]^T. \tag{2.148}$$

The estimation of the factor matrices is similar to the ALS algorithm presented before, the main difference is the estimation of the angular frequencies $(\omega_{I_2,r})$ of the Vandermonde matrix $\mathbf{A}^{(2)}$. We first estimate $\mathbf{A}^{(2)}$ from Eq. (2.149) and after estimating the angular frequencies using the rectification strategies presented in the next subsections. And finally, we update $\mathbf{M}_2$. The RectALS algorithm is summarized in Table 6.

2.5.7.1 Rectification strategy

In this section, we present rectification strategies for estimating a Vandermonde matrix (2.147), with $\omega_{I_2,r}$ replaced by ω_r, combined with the RectALS algorithm. The first one is the shift invariance principle (SIP) introduced in [64] and summarized in Table 7. SIP strategy is inspired from the notion of pencil matrices. The second strategy is the toeplitz rank-one approximation (TR$_1$A) method proposed in [65], used in [33] and summarized in Table 8. And the third strategy is (TR$_1$A$_{imp}$) method which is an improved version of the TR$_1$A method, proposed in [32] and presented in this thesis in Table 9. TR$_1$A$_{imp}$ strategy is proposed to enforce the Vandermonde structure of the steering matrices estimated in presence of noise, whose elements of the first row are normalized to 1 [32].

TR$_1$A rectification method consists in constructing a rank-one hermitian Toeplitz matrix $\mathbf{T}_r = \hat{\mathbf{A}}^{(2)}_{.r} \hat{\mathbf{A}}^{(2)H}_{.r}$, of size $I_2 \times I_2$, from each column r of an estimated $I_2 \times R$ Vandermonde matrix, and computing its SVD $\mathbf{T}_k = \mathbf{U}_r \mathbf{\Sigma}_r \mathbf{V}^*_r$ to estimate the r-th angular frequency.

Our proposed rectification method consists in constructing a rank-one hermitian Toeplitz matrix $\mathbf{T}_r = \hat{\mathbf{A}}^{(2)}_{.r} \hat{\mathbf{A}}^{(2)H}_{.r}$, of size $I_2 \times I_2$, from each column r of an estimated $I_2 \times R$ Vandermonde matrix, and computing its eigenvalue decomposition (EVD) $\mathbf{T}_r = \mathbf{P}_r \mathbf{D}_r \mathbf{P}^H_r$ to estimate the r-th angular frequency.

Unlike the original TR$_1$A method which identifies only one element of the Toeplitz matrix with its corresponding term in the EVD of the constructed rank-one Toeplitz matrix, our method exploits the identification of $I - 1$ terms as:

$$e^{-j\omega_r(i)} = d_r (p_r)_{1,1} (p_r)^*_{i+1,1} \ \Rightarrow \ \hat{\omega}_r = \frac{1}{i} \angle \Big((p_r)_{1,1} (p_r)^*_{i+1,1} \Big), \tag{2.149}$$

```
RectALS algorithm for estimation of $\mathbf{M}_1$, $\mathbf{A}^{(2)}$and $\mathbf{M}_2$
Input: tensor $\mathcal{G}^{(2)}$
Output: Estimated $\mathbf{M}_1$, $\mathbf{A}^{(2)}$, $\mathbf{M}_2$ and angular frequencies $\{\omega_{I_2,1},...,\omega_{I_2,R}\}$
$l = 0$
1) Randomly initialize $\mathbf{A}^{(2)}_{[0]}$, $M^{-T}_{2[0]}$ with random inputs
2) Update the estimtes as:
    $\hat{\mathbf{M}}_{1[l]} = \mathbf{G}^{(2)}_{R\times I_2 R}[(\hat{\mathbf{A}}^{(2)}_{[l-1]} \diamond \hat{\mathbf{M}}^{-T}_{2[l-1]})^T]^{\dagger}$,
    $\hat{\mathbf{A}}^{(2)}_{[l]} = \mathbf{G}^{(2)}_{I_2\times R^2}[(\hat{\mathbf{M}}^{-T}_{2[l-1]} \diamond \hat{\mathbf{M}}_{1[l]})^T]^{\dagger}$.
3) Apply the rectification strategy
    $[\hat{\mathbf{A}}^{(2)}_{[l]}] = \text{rectification} \quad \text{strategy}\left(\hat{\mathbf{A}}^{(2)}_{[l]}\right)$.
    $\hat{\mathbf{M}}^{-T}_{2[l]} = \mathbf{G}^{(2)}_{R\times RI_2}\left[\left(\hat{\mathbf{M}}_{1[l]} \diamond \hat{\mathbf{A}}^{(2)}_{[l]}\right)^T\right]^{\dagger}$,
end
3) Calculate the error function as:
    $\epsilon_{[l]} = \| \mathbf{G}^{(2)}_{I_2\times R^2} - \hat{\mathbf{A}}^{(2)}_{[l]}\left(\hat{\mathbf{M}}^{-T}_{2[l]} \diamond \hat{\mathbf{M}}_{1[l]}\right)^T \|^2_F$ .
- if $|\epsilon_{[l-1]} - \epsilon_{[l]}| \leq \delta$
- stop
- else $l \to l + 1$;
$\hat{\mathbf{M}}_1$, $\hat{\mathbf{A}}^{(2)}$ and $\hat{\mathbf{M}}_2$
end
```

Table 6 – RectALS algorithm.

for $i \in [1, I_2 - 1]$, where $(p_r)_{i,1}$ denotes the $(i,1)$ element of the eigenvector matrix $\mathbf{P}_r$, and the operator $\angle$ stands for the angle of its complex scalar argument. Then, the estimated angular frequency is calculated as an average of these $I_2 - 1$ estimates (2.149) as:

$$\hat{\omega}_r = \frac{1}{I_2 - 1}\sum_{i=1}^{I_2-1}\frac{1}{i}\angle\left((p_r)_{1,1}(p_r)^*_{i+1,1}\right). \tag{2.150}$$

2.6 Chapter summary

In this chapter, multilinear algebra operations and definitions, tensor decompositions and algorithms were provided. These concepts serve as the basis for the next chapters. In the first part of this chapter, some notations extensively used in this thesis were introduced. In the second part of this chapter, some operations involving matrices, namely Kronecker and

SIP strategy to rectify an estimated Vandermonde matrix $\hat{\mathbf{A}}^{(2)}$

```
Input: Estimated matrix $\hat{\mathbf{A}}^{(2)}$
Output: Rectified Vandermonde matrix $\hat{\mathbf{A}}^{(2)}$.
for $r \in [1, R]$
    $y_r = \overline{\hat{\mathbf{A}}}_{.r}^{(2)H} \underline{\hat{\mathbf{A}}}_{.r}^{(2)}$,
    $\hat{\omega}_{I_2,r} = \angle\left(\frac{1}{y_r}\right)$,
    for $i \in [1,I_2]$
        $\hat{a}_{i,r}^{(2)} = e^{-j\hat{\omega}_r(i-1)}$,
    end
    $\mathbf{A}_{.r}^{(2)} = [\ \hat{a}_{1,r}^{(2)} \ \dots \ \hat{a}_{I_2,r}^{(2)}\ ]$.
end
```

Table 7 – SIP strategy.

TR$_1$A strategy to rectify an estimated Vandermonde matrix $\hat{\mathbf{A}}^{(2)}$

```
Input: Estimated matrix $\hat{\mathbf{A}}^{(2)}$
Output: Rectified Vandermonde matrix $\hat{\mathbf{A}}^{(2)}$.
for $r \in [1,R]$
    1) Form a rank-one hermitian Toeplitz matrix $\mathbf{T}_r$ from the $r$-th column $\hat{\mathbf{A}}_{.r}^{(2)}$ as:
        $\mathbf{T}_r = \hat{\mathbf{A}}_{.r}^{(2)} \hat{\mathbf{A}}_{.r}^{(2)H}$,
    2) Calculate the SVD of $\mathbf{T}_r = \mathbf{U}_r \Sigma_r \mathbf{V}_r^*$.
    for $i \in [1,I_2 - 1]$
        3) Estimate the angular frequency $\hat{\omega}_r$ associated with the $r$-th column of the Vandermonde matrix as:
            $\hat{\omega}_r = \angle\left((u_r)_{1,1}((u_r)^*_{(i+1),1}\right)$,
            $\hat{a}_{i,r}^{(2)} = e^{-j\hat{\omega}_r(i)}$.
    end
    4) Build the rectified estimated $r$-th column $\hat{\hat{\mathbf{A}}}_{.r}$ as:
        $\mathbf{A}_{.r}^{(2)} = [\ \hat{a}_{1,r}^{(2)} \ \dots \ \hat{a}_{I_2,r}^{(2)}\ ]$.
end
```

Table 8 – TR$_1$A strategy.

```
TR$_1$A$_{imp}$ strategy to rectify an estimated Vandermonde matrix $\hat{\mathbf{A}}^{(2)}$
Input: Estimated matrix $\hat{\mathbf{A}}^{(2)}$
Output: Rectified Vandermonde matrix $\hat{\hat{\mathbf{A}}}^{(2)}$.
for $r \in [1, R]$
    1) Form a rank-one hermitian Toeplitz matrix $\mathbf{T}_r$ from the $r$-th column $\hat{\mathbf{A}}^{(2)}_{.r}$ as:
        $$\mathbf{T}_r = \hat{\mathbf{A}}^{(2)}_{.r}\hat{\mathbf{A}}^{(2)H}_{.r},$$
    2) Calculate the EVD of $\mathbf{T}_r = \mathbf{P}_r\mathbf{D}_r\mathbf{P}_r^H$.
    for $i \in [1, I_2 - 1]$
        3) Estimate the angular frequency $\omega_r$ associated with the $r$-th column of the Vandermonde matrix as:
            $$\hat{\omega}_r = \frac{1}{I_2 - 1}\left(\sum_{i=1}^{I_2-1}\frac{1}{i}\angle\Big((p_r)_{1,1}((p_r)^*_{i+1,1})\Big)\right),$$
            $$\hat{a}^{(2)}_{i,r} = e^{-j\hat{\omega}_r(i)},$$
    end
    4) Build the rectified estimated $r$-th column $\hat{\hat{\mathbf{A}}}_{.r}$ as:
        $$\mathbf{A}^{(2)}_{.r} = [\ \hat{a}^{(2)}_{1,r}\ \ \dots\ \ \hat{a}^{(2)}_{I_2,r}\ ].$$
end
```

Table 9 – TR$_1$A$_{imp}$ strategy.

Khatri-Rao products were introduced, as well as, their properties. Basic tensor algebra definitions and a background in tensor decompositions were explored in the third and fourth parts of this chapter, where the new coupled-nested Tucker PARAFAC model and its uniqueness property were presented. The new tensor decomposition combines Tucker and PARAFAC models in a unique decomposition where the core tensor must be known, being one of the main contributions of this thesis. And in the fifth part, some algorithms were presented to estimate the factor matrices as SVD, KronF, and KRF algorithms. The link between TT and PARAFAC decomposition was recalled and some algorithms were presented to estimate the factor matrices in tensor models such as ALS, TTHSVD, JIRAFE, RectALS, and others, as well as algorithms for estimating angular frequencies in Vandermonde matrices.

3 Multidimensional CX decomposition of tensors

In this chapter, the second original contribution of this thesis is presented. First, a multidimensional column-space decomposition to perform a low-rank approximation of tensors based on the CX decomposition for matrices is presented. CX decomposition is a low-rank tensor decomposition that is explicitly expressed in terms of a small number of actual columns of the data tensor. CX decomposition for tensors is used to reduce complexity and preserve the initial information contained in the data tensor. For this new tensor model is proposed an algorithm that performs the approximation of the tensor based on the l_2*-norm.*

The computational complexity of the algorithm is analyzed and compared to the state-of-the-art THOSVD algorithm. Monte Carlo simulation results are provided to illustrate the performance of the proposed CX-tensor decomposition and the associated algorithm to tensor reconstruction and the impact of design parameters on the system performance in terms of NMSE.

3.1 Introduction

The large-scale data is characterized not only by its huge volume but also by its heterogeneity, precision and incompleteness. In recent years, tensors have been commonly used to designate generalized multidimensional data of large matrices, such as big data, reducing the problem of dimensionality [66]. Various tensor models are applied because they are well adapted to represent great data, such as Tucker hierarchical models [67], tensor networks [68] and TT decomposition [69].

As an attractive approach employing sparse representations, the compressed/compressive sensing (CS) technique is proposed to reduce data purchasing costs, allowing Nyquist sub-sampling, to be widely applied in many areas [70]. The key idea of the CS is to allow the exact reconstruction of the signal of much less samples than required by the Nyquist–Shannon sampling theorem since the signal admits a sparse representation in a given domain [70, 71]. Sparsity is the inherent property of those signals for which the whole of the information contained in the signal can be represented only with the help of a few significant components, as compared to the total length of the signal [72]. In the CS method, the compressed samples are acquired by a small set of linear and usually random measures, where the recovery of the signal is generally formulated as a problem of minimizing the l_0-norm [73]. As the minimization of l_0-norm is an NP-Hard problem, most existing research solves this problem by bringing it closer to a convex problem of minimization of l_1-norm [70].

CS and matrix completion (MC) have some similarities, where the goal is to recover a low-rank matrix by using a small number of matrix measurements [70, 74, 75, 76]. The matrix reconstruction methods recover the absent inputs of a matrix $\mathbf{T} \in \mathbb{R}^{M \times N}$ of its partially known entries given by a random subset Ω of size $M \times N$. This problem is known as a low-rank matrix optimization problem [69]:

$$\min_{\mathbf{X}} \operatorname{rank}(\mathbf{X}) \quad \text{subject} \quad \text{to} \quad \mathbf{X}_\Omega = \mathbf{T}_\Omega. \tag{3.1}$$

The missing inputs of $\mathbf{X}$ are as small as possible. Due to the nature of the rank estimation, problem (3.1) is considered NP-hard and can be fexibilized to a convex optimization problem l_1 as:

$$\min_{\mathbf{X}} \parallel \mathbf{X} \parallel_* \quad \text{subject} \quad \text{to} \quad \mathbf{X}_\Omega = \mathbf{T}_\Omega, \tag{3.2}$$

where $\parallel \mathbf{X} \parallel_*$ corresponds to the matrix nuclear norm. CS and MC were extended to the tensor completion (TC). TC is a problem of filling out the missing inputs of partially observed low-rank tensors [77]. This problem can be written as a generalization of the MC problem as follows. Let us consider the N-order tensor $\mathcal{T} \in \mathbb{R}^{I_1 \times \ldots \times I_N}$ and we must reconstruct it from its known entries given by a set (Ω) of indexes. The following tensor rank minimization problem is proposed as [69, 78]:

$$\min_{\mathbf{X}^{(n)}} \sum_{n=1}^{N} \alpha_n \operatorname{rank}(\mathbf{X}^{(n)}) \quad \text{subject} \quad \text{to} \quad \mathcal{X}_\Omega = \mathcal{T}_\Omega, \tag{3.3}$$

where α_n are the weights that satisfy $\sum_{n=1}^{N} \alpha_n = 1$. Eq. (3.3) is addressed by the following L_1 optimization problem:

$$\min_{\mathbf{X}^{(n)}} \sum_{n=1}^{N} \alpha_n \parallel \mathbf{X}^{(n)} \parallel_* \quad \text{subject} \quad \text{to} \quad \mathcal{X}_\Omega = \mathcal{T}_\Omega, \tag{3.4}$$

where $\parallel \mathbf{X}^{(n)} \parallel_*$ corresponds to the matrix nuclear norm of $\mathbf{X}^{(n)}$ for $n \in [1,N]$. In most cases, the tensor reconstruction is performed by reorganizing tensors as n-mode matrix unfoldings and then applying matrix reconstruction techniques [77, 75]. Tensor reconstruction algorithms and their applications are receiving a lot of attention in some areas, such as data mining and computational vision [79], collaborative filtering [80], medical images [78], signal processing [81], neuroscience [82], telecommunications [83] and traffic data estimation [84].

The CX decomposition is particularly attractive for large-scale [85] matrices because it provides a representation of data as a linear combination of a few columns of the original data matrix [86, 87], which does not happen when we consider SVD. SVD provides the best low-rank approximation of a matrix but, the new “dimensions” (the so-called eigencolumns and eigenrows) of the matrix are linear combinations of (up to all) the original dimensions, being difficult to interpret in terms of the underlying data and processes generating that data [88].

Few works addressed the CX decompositions for data tensors [87, 89]. In [87], two different generalizations of the CUR matrix decomposition are presented to tensors. These generalizations are based only on some selected entries of the original tensor. The authors also developed an adaptive algorithm for selecting the n-mode fibers from the data tensor. In addition to providing a low-rank approximation, such CX tensor decompositions have the property of preserving the initial information contained in the tensor, while keeping the sparsity of the initial data tensor, which is not the case of the HOSVD [89]. In [89], an algorithm for the selection of fibers of the tensor is introduced. The authors applied the approximation for rows and columns to a particular matrix unfolding providing an approximation based on a few fibers extracted from the frontal slices of a third-order tensor.

Based on related works, we propose a low-rank tensor decomposition obtained by applying the matrix CX decomposition to each n-mode unfolding of the tensor. The proposed decomposition not only preserves the initial information contained in the data tensor but also reduces the storage requirements, which makes easier the tasks of tensor reconstruction. An algorithm based on the l_2-norm of columns of each unfolding is derived to perform the CX-tensor decomposition. The proposed algorithm can be seen as an extension of the algorithm proposed in [90] to higher-order tensors.

3.2 CX decomposition for matrices

First, we need to understand CX decomposition for matrices. Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ be a matrix and $\mathbf{C} \in \mathbb{R}^{m\times c}$ be a matrix whose columns consist of c columns of $\mathbf{A}$. The matrix $\hat{\mathbf{A}} = \mathbf{CX} \in \mathbb{R}^{m\times n}$ is a column-based low-rank matrix approximation of $\mathbf{A}$, for some matrix $\mathbf{X} \in \mathbb{R}^{c\times n}$ [86, 88]. CX decomposition provides a low-rank approximation expressed in a small number of columns c of the original matrix $\mathbf{A}$. Since the columns in $\mathbf{A}$ are selected to form $\mathbf{C}$, the optimization problem for the matrix $\mathbf{X}$ to be solved is:

$$\min_{\mathbf{X}\in\mathbb{R}^{c\times n}} \parallel \mathbf{A} - \mathbf{CX} \parallel_F^2, \tag{3.5}$$

where $\mathbf{X} = \mathbf{C}^{\dagger}\mathbf{A}$. Solving the LS problem (3.5) for $\mathbf{X}$, we provide the “best” approximation of $\mathbf{A}$ as:

$$\hat{\mathbf{A}} = \mathbf{CC}^{\dagger}\mathbf{A} = P_{\mathbf{C}}\mathbf{A}, \tag{3.6}$$

where $P_{\mathbf{C}}\mathbf{A}$ is the projection of $\mathbf{A}$ onto the subspace spanned by the columns of $\mathbf{C}$ with $P_{\mathbf{C}} = \mathbf{CC}^{\dagger}$. The quality of a CX matrix approximation depends on the choice of $\mathbf{C}$ and $\mathbf{X}$ with the following upper bound for the approximation error [86, 88]:

$$\parallel \mathbf{A} - \mathbf{CX} \parallel_F^2 \leq (1+\epsilon) \parallel \mathbf{A} - \mathbf{A}_k \parallel_F^2, \tag{3.7}$$

where $\mathbf{A}_k$ is the best rank-k approximation of $\mathbf{A}$ and ϵ is the error parameter.

A particular case of CX decomposition for a second-order tensor is the CUR decomposition in which the subsets of columns and rows of the matrix $\mathbf{A}$ are extracted. The subset of c columns and r rows selected fills the matrices $\mathbf{C} \in \mathbb{R}^{m \times c}$ and $\mathbf{R} \in \mathbb{R}^{r \times n}$, respectively. Next we calculate a suitable intersection matrix $\mathbf{U} \approx \mathbf{C}^{\dagger}\mathbf{A}\mathbf{R}^{\dagger} \in \mathbb{R}^{c \times r}$, and estimate a low-rank approximation $\hat{\mathbf{A}}$ as [90]:

$$\mathbf{A} \approx \hat{\mathbf{A}} = \mathbf{CUR}. \tag{3.8}$$

3.3 CX decomposition for N-order tensors

Based on the CX decomposition previously presented, we developed a generalized CX decomposition for the tensor case as presented in [91]. Let $\mathcal{Y} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$ be an N-order tensor decomposed by means of the Tucker decomposition with multilinear rank-$(R_1, R_2, \ldots, R_N)$ as:

$$\mathcal{Y} = \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \ldots \times_N \mathbf{A}^{(N)}, \tag{3.9}$$

where $\mathcal{G} \in \mathbb{R}^{R_1 \times R_2 \times \ldots \times R_N}$ is the core tensor and $\mathbf{A}^{(n)} \in \mathbb{R}^{I_n \times R_n}$ are the factor matrices, with $n \in [1, N]$. To perform the CX decomposition for tensors it is necessary to select P_n columns of each n-mode unfolding $\mathbf{Y}^{(n)} \in \mathbb{C}^{I_n \times I_{n+1} \ldots I_N I_1 \ldots I_{n-1}}$ of $\mathcal{Y}$ in order to reduce the dimensionality of the tensor $\mathcal{Y}$. Thus, we can rewrite the Tucker decomposition as the CX model, such that:

$$\hat{\mathcal{Y}} \approx \mathcal{W} \times_1 \mathbf{C}^{(1)} \times_2 \ldots \times_N \mathbf{C}^{(N)}, \tag{3.10}$$

where $\mathcal{W} \in \mathbb{R}^{P_1 \times P_2 \times \ldots \times P_N}$ is the core tensor and $\mathbf{C}^{(n)} \in \mathbb{R}^{I_n \times P_n}$ contains the selected columns from each n-mode unfolding. CX decomposition for each n-mode matrix unfolding is given by:

$$\mathbf{Y}^{(n)} \approx \mathbf{C}^{(n)}\mathbf{X}^{(n)}, \tag{3.11}$$

with $\mathbf{X}^{(n)} \in \mathbb{R}^{P_n \times I_{n+1} \ldots I_N I_1 \ldots I_{n-1}}$. The following optimization problem can be solved as:

$$\min_{\mathcal{W}} \| \, \mathcal{Y} - \mathcal{W} \times_1 \mathbf{C}^{(1)} \times_2 \ldots \times_N \mathbf{C}^{(N)} \, \|_F^2 \, . \tag{3.12}$$

By minimizing the cost function (3.12) for the core tensor $\mathcal{W} \in \mathbb{R}^{P_1 \times P_2 \times \ldots \times P_N}$ yields:

$$\hat{\mathcal{W}} = \mathcal{Y} \times_1 \mathbf{C}^{(1)\dagger} \times_2 \ldots \times_N \mathbf{C}^{(N)\dagger}, \tag{3.13}$$

$\mathbf{C}^{(n)}$ must be full column rank to ensure the uniqueness of the pseudo-inverses $\mathbf{C}^{(n)\dagger}$, i.e.,

$$P_n \leq I_n, \tag{3.14}$$

with $n \in [1, N]$. Then, rank$(\mathbf{C}^{(n)}) = P_n$ and $\mathbf{C}^{(N)\dagger}\mathbf{C}^{(N)} = \mathbf{I}_{P_n}$. The n-mode matrix unfoldings of (3.10) are given by:

$$\hat{\mathbf{Y}}^{(n)} \approx \mathbf{C}^{(n)}\hat{\mathbf{W}}^{(n)} \left(\mathbf{C}^{(n+1)} \otimes \ldots \otimes \mathbf{C}^{(N)} \otimes \mathbf{C}^{(1)} \otimes \ldots \otimes \mathbf{C}^{(n-1)} \right)^T, \tag{3.15}$$

with $\hat{\mathbf{Y}}^{(n)} \in \mathbb{R}^{I_n \times I_{n+1} \ldots I_N I_1 \ldots I_{n-1}}$ and $\hat{\mathbf{W}}^{(n)} \in \mathbb{R}^{P_n \times P_{n+1} \ldots P_N P_1 \ldots P_{n-1}}$. As for the matrix case, the quality of CX tensor approximation depends on the choice of matrices $\mathbf{C}^{(n)}$ and P_n.

3.4 Algorithm for CX decomposition of N-order tensors

In this section, an algorithm is proposed to select P_n columns of each n-mode matrix unfolding of the N-order tensor $\mathcal{Y}$ based on the column selection algorithm for matrices, described in [90]. This algorithm is based on l_2-norm to select the columns of each n-mode unfolding $\mathbf{Y}^{(n)}$, where a selection probability is calculated for each column of $\mathbf{Y}^{(n)}$ derived from your l_2-norm. The underlying idea is that if the column of the matrix unfolding $\mathbf{Y}^{(n)}$ has relevant information, this column will have a high probability to be selected for the matrix $\mathbf{C}^{(n)}$. The probabilities $p_j^{(n)}$ are defined for each column j of the matrix unfolding $\mathbf{Y}^{(n)}$ as [91]:

$$p_j^{(n)} = \frac{\parallel \mathbf{Y}_{.j}^{(n)} \parallel_2^2}{\parallel \mathbf{Y}^{(n)} \parallel_F^2}, \tag{3.16}$$

where $\mathbf{Y}_{.j}^{(n)}$ is the j-th column of $\mathbf{Y}^{(n)}$, $j \in [1, J]$ and $J = I_{n+1}...I_N I_1...I_{n-1}$. The probability is calculated for each column of the n-mode matrix unfolding $\mathbf{Y}^{(n)}$, where the P_n columns (in Eq. (3.16)) are selected, with $P_n \leq I_n$. After calculating the estimated probabilities $p_j^{(n)}$ of the j-th column of $\mathbf{Y}^{(n)}$, the matrix $\mathbf{C}^{(n)}$ is built with the P_n columns that have the highest probabilities. The selected columns are scaled as follows:

$$\mathbf{C}_{\cdot j_{p_n}}^{(n)} = \frac{\mathbf{Y}_{\cdot j_{p_n}}^{(n)}}{\sqrt{P_n p_{j_{p_n}}^{(n)}}}, \tag{3.17}$$

where j_{p_n} represents the selected columns of the n-mode matrix unfolding $\mathbf{Y}^{(n)}$, $\mathbf{C}_{\cdot j_{p_n}}^{(n)} \in \mathbb{R}^{I_n}$ is the j_{p_n}-th selected column scaled by $p_n \in [1, P_n]$ and $p_{j_{p_n}}^{(n)}$ is the probability of the j_{p_n}-th column. The P_n scaled columns fill the matrix $\mathbf{C}^{(n)} \in \mathbb{R}^{I_n \times P_n}$ in the same order in which they were taken from the original unfolding $\mathbf{Y}^{(n)}$. And the core sub-tensor $\hat{\mathcal{W}} \in \mathbb{R}^{P_1 \times P_2 \times ... \times P_N}$ is calculated as Eq. (3.13). The CX algorithm for N-order tensors is summarized in Table 10.

3.5 Computational complexity

Considering a N-order hyper cubic tensor $\mathcal{Y} \in \mathbb{R}^{I \times I \times ... \times I}$ with multilinear-rank $(R, R, ..., R)$, the complexity of CX algorithm is:

$$O\left(NI^N + NPI^2 + \sum_{k=1}^{N} P^k I^{N-k+1}\right), \tag{3.18}$$

flops to get all the factors, where P is the number of columns selected. Here, the same number of columns selected are considered for all matrix unfoldings. The first term in the CX algorithm complexity is due to the calculation of the probability $p_j^{(n)}$ for I^{N-1} columns of dimension I of each matrix unfolding, the second term is related to the pseudo-inverse of the matrices $\mathbf{C}^{(n)}$ to compute the core tensor and the third term is required to compute the core tensor by means of N matrix multiplications required in (3.13).

CX algorithm to choose the column matrices $\mathbf{C}^{(n)}$ and core sub-tensor $\mathcal{W}$

```
Input: Tensor $\mathcal{Y}$ , number of columns $P_n$.
Output: Estimated $\mathbf{C}^{(n)}$, $\mathbf{X}^{(n)}$, $\hat{\mathcal{W}}$ and $\hat{\mathcal{Y}}$.
for $n = 1 : N$
  for $j = 1 : J$
    1) Calculate the probability for each column as:
       $$p_j^{(n)} = \frac{\| \mathbf{Y}_{.j}^{(n)} \|_2^2}{\| \mathbf{Y}^{(n)} \|_F^2},$$
  end
  2) For the matrix $\mathbf{C}^{(n)}$:
  - Select $P_n$ columns of each matrix unfolding with highest probabilities;
  - Store the column and position;
  - Scale $P_n$ selected columns and store them in matrix $\mathbf{C}^{(n)}$ as:
  for $j_{p_n} = 1 : P_n$
     $$\mathbf{C}_{.j_{pn}}^{(n)} = \frac{\mathbf{Y}_{.j_{pn}}^{(n)}}{\sqrt{P_n p_{j_{pn}}^{(n)}}},$$
  end
end
3) Calculate the core sub-tensor $\hat{\mathcal{W}}$ as:
   $$\hat{\mathcal{W}} = \mathcal{Y} \times_1 \mathbf{C}^{(1)\dagger} \times_2 \ldots \times_N \mathbf{C}^{(N)\dagger},$$
4) Construct $\hat{\mathcal{Y}}$ as:
   $$\hat{\mathcal{Y}} = \hat{\mathcal{W}} \times_1 \mathbf{C}^{(1)} \times_2 \ldots \times_N \mathbf{C}^{(N)}.$$
```

Table 10 – CX algorithm.

3.6 Simulation results

In this section, we evaluate the performance of the proposed CX algorithm. First, in Section 3.6.1, the simulations and the considered performance criterion are described. In Section 3.6.2, the impact of design parameters is studied in terms of reconstructed signal normalized mean square (NMSE). The results presented in Section 3.6.2 are compared with the state-of-the-art THOSVD algorithm for the approximation of a third-order tensor.

3.6.1 General description of the simulations

In this section, the description of the simulations is considered. The data tensor is simulated as in [92]. First, a cubic data tensor $\mathcal{Y} \in \mathbb{R}^{I \times I \times I}$ with multilinear rank-(R, R, R) is constructed. The core tensor $\mathcal{G} \in \mathbb{R}^{R\times R\times R}$ is generated in the entry drawn from the Gaussian distribution. The factor matrices $\mathbf{U}^{(n)}$ are generated as:

$$\mathbf{U}^{(n)} = \mathbf{Q}_n \mathrm{diag}\left(1, 2^{-\theta}, ..., I^{-\theta}\right) \in \mathbb{R}^{I\times R}, \tag{3.19}$$

with $\mathbf{Q}_n$ being a random orthogonal matrix and θ controlling the rate of decay of the singular values. The higher the value of θ, the closer to zero are the elements of the matrix $\mathbf{U}^{(n)}$, ensuring the sparsity of the data matrices. With the generated factor matrices ($\mathbf{U}^{(1)}$, $\mathbf{U}^{(2)}$, $\mathbf{U}^{(3)}$), and core tensor ($\mathcal{G}$), the data tensor $\mathcal{X}$ is simulated as:

$$\mathcal{X} = \mathcal{G} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)} \in \mathbb{R}^{I\times I\times I}, \tag{3.20}$$

which is then normalized as:

$$\mathcal{X}_{norm} = \frac{\mathcal{X}}{\| \mathcal{X} \|_F}. \tag{3.21}$$

The Gaussian noise tensor $\mathcal{V} \in \mathbb{R}^{I\times I\times I}$ is generated and normalized as $\mathcal{X}$, such that $\mathcal{V}_{norm} = \dfrac{\mathcal{V}}{\| \mathcal{V} \|_F}$ and then added to the tensor $\mathcal{X}_{norm}$, generating the data tensor $\mathcal{Y} \in \mathbb{R}^{I\times I\times I}$ as:

$$\mathcal{Y} = \mathcal{X}_{norm} + \alpha \mathcal{V}_{norm}, \tag{3.22}$$

where α allows fixing the signal to noise ratio (SNR) calculated as:

$$\mathrm{SNR} = 20\log\frac{\| \mathcal{X}_{norm} \|_F}{\| \mathcal{V}_{norm} \|_F}, \tag{3.23}$$

which gives $\alpha = \dfrac{\| \mathcal{X}_{norm} \|_F}{\| \mathcal{V}_{norm} \|_F} 10^{-\mathrm{SNR}/20}$. The dimension and the order of tensor $\mathcal{Y}$ are fixed for $I_n = I = 20$ and $N = 3$, respectively, to evaluate the approximation. Subsequently, the estimation of the tensor $\hat{\mathcal{Y}}$ is performed using the THOSVD and CX algorithms. The number of selected columns P_n is equal to the truncated rank T_n determined for each unfolding, to compare the results , i.e, $P_n = T_n = P = T$. T and P are varied such that $(P, T) < R$, to evaluate which number of columns selected/truncated rank provides the best reconstruction. The design parameters are the SNR $\in [10, 30, 50]$ dB and $\theta \in [1, 3, 5, 10]$. The estimated model is evaluated by the NMSE, given by:

$$NMSE = \frac{1}{M} \sum_{m=1}^{M} \frac{\| \mathcal{Y}_m - \hat{\mathcal{Y}}_m \|}{\| \mathcal{Y}_m \|}, \tag{3.24}$$

where $\hat{\mathcal{Y}}_m$ is the tensor $\mathcal{Y}_m$ estimated at the m-th Monte Carlo run. The NMSE is calculated by averaging the results over $M = 1000$ Monte Carlo runs, after truncating the 5% worse and 5% better values to eliminate the influence of ill-convergence and outliers.

3.6.2 Impact of design parameters

In this section, we evaluate the performance of the proposed algorithm compared with the state-of-the-art THOSVD algorithm. The results obtained are shown in Figures 10 and 11 for a tensor of third-order $\mathcal{Y} \in \mathbb{R}^{20\times20\times20}$. The SNR is varied to observe the impact on the choice of the number of columns or ranks. In Figures 10 and 11 it is possible to verify that, as expected, the NMSE performance improves as SNR increases.

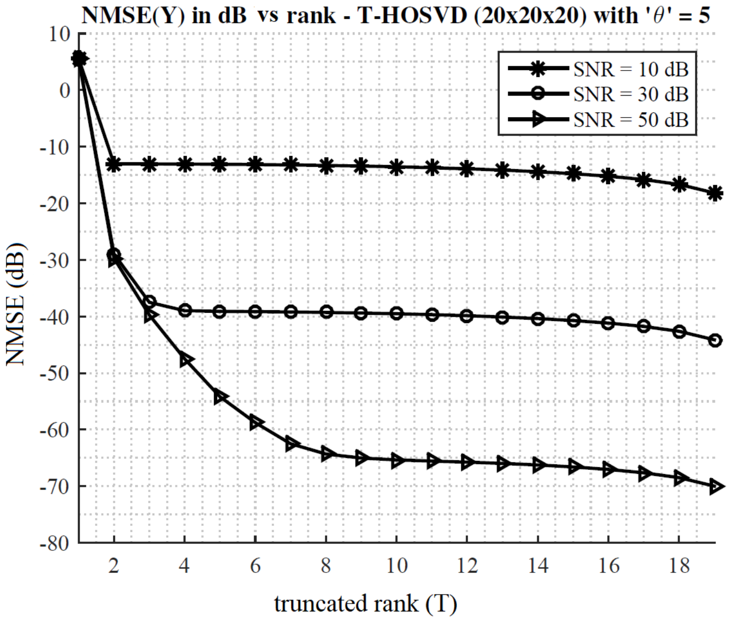


Figure 10 – NMSE vs. Rank for fixed θ - THOSVD algorithm.

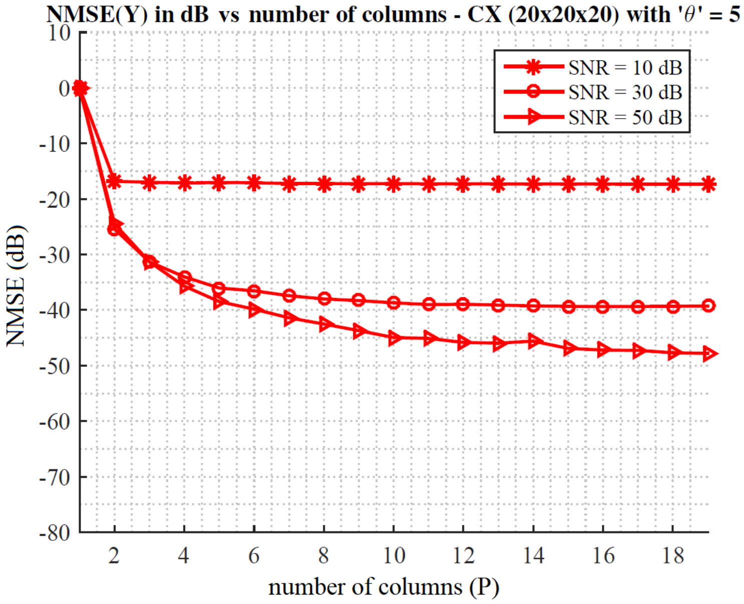


Figure 11 – NMSE vs. the number of columns selected for fixed θ - THOSVD algorithm.

Figure 12 shows the comparison between the THOSVD and CX algorithms considering a fixed SNR to 50 dB and varying θ.

In Figure 12, the performance of the THOSVD algorithm is better than that of the proposed CX algorithm since THOSVD provides the best low-rank approximation for tensor reconstruction. With the curves in Figure 12, it is also possible to observe the impact of

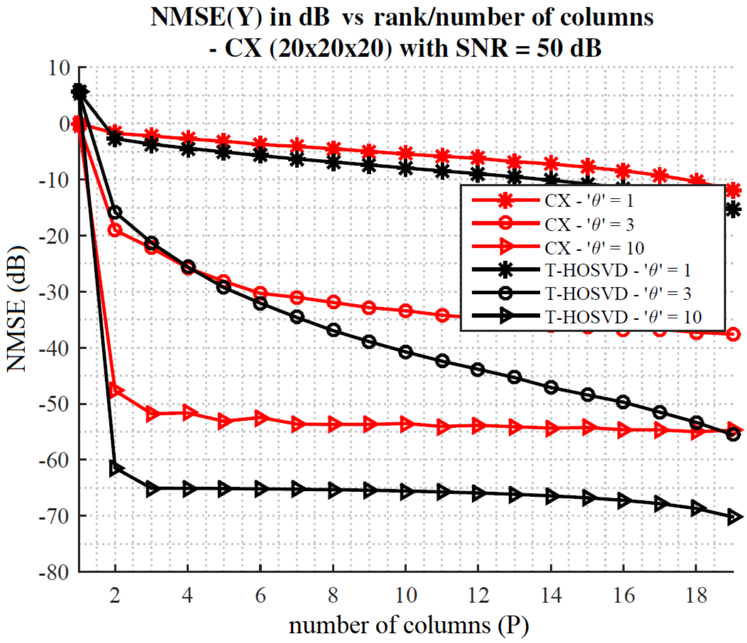


Figure 12 – NMSE vs. the number of columns selected for different values of angular decay θ - THOSVD and CX algorithms.

Algorithms	**Computational Complexity**
CX	$O\left(NI^N + NPI^2 + \sum_{k=1}^{N} P^k I^{N-k+1}\right)$
THOSVD	$O\left(NI^{N+1} + NTI^2 + \sum_{k=1}^{N} T^k I^{N-k+1}\right)$

Table 11 – Computational complexity for the THOSVD and CX algorithms.

parameter θ on the estimation of tensor $\hat{\mathcal{Y}}$. As θ increases the singular values of the tensor decrease rapidly and it is possible to reconstruct the tensor with a small truncated rank. We can also conclude that when the truncated rank T and the number of selected columns P are closer to the full tensor rank R, the reconstruction improves.

Note also that the simulation results show a degradation of the approximation performance of the CX algorithm compared with experiments with the THOSVD algorithm, which is the price to pay for keeping the initial information of the data tensor. However, keeping the initial information contained in the tensor, as in the proposed solution, allows us to respect some characteristics of the data like their sparsity and to simplify the interpretation of results after a classification task, which is not the case with the THOSVD algorithm.

In Table 11, a comparison of the complexities of THOSVD and CX algorithms is provided by evaluating the cost in relation to the construction of the column matrices, left singular vector matrices and core tensors. Defining the ratio $O_1 = O_{THOSVD}/O_{CX}$, which expresses how many times THOSVD is more computationally demanding than CX for data reconstruction. We have

$$O_1 = \frac{NI^{N+1} + NTI^2 + \sum_{k=1}^{N} T^k I^{N-k+1}}{NI^N + NPI^2 + \sum_{k=1}^{N} P^k I^{N-k+1}}, \tag{3.25}$$

where we are considering $T = P$. Note that the complexities of the algorithms are similar, but typically the CX algorithm requires more columns than the THOSVD algorithm to perform a better estimate of the tensor, i.e., $T \leq P$.

Figure 13 shows the complexity ratio O_1 considering the variation of the number of selected columns P and $N \in [3, 4, 5]$, i.e. the order of the data tensor with I remaining fixed and equal to 20. Regarding the variation of the number of selected columns note that increasing P, the computational complexity of THOSVD compared to CX are similar, as they differ only by the factor I in the first element that is constant. Regarding the variation of the order of the tensor, increasing N, the closer are the complexities of THOSVD and CX. And when P is close to the tensor rank, the complexity rates are close independent of the tensor order.

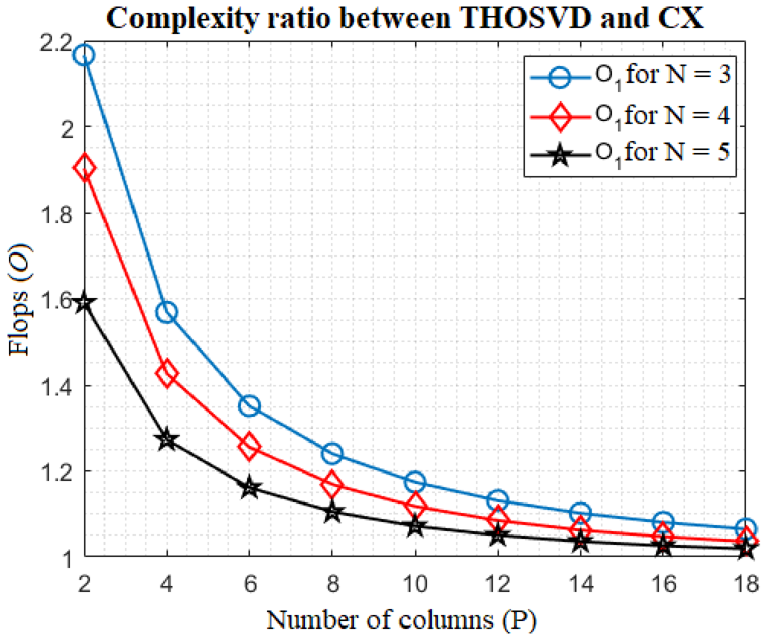


Figure 13 – Complexity ratio between THOSVD and CX algorithms.

3.7 Chapter summary

In this chapter, the second contribution of this thesis was presented. The multidimensional CX decomposition has been proposed to perform a low-rank approximation of an N-order tensor preserving initial information. For the multidimensional CX decomposition, an algorithm based on the computation of the probabilities of the columns of each matrix unfolding is proposed to perform the reconstruction of the data tensor. This decomposition allows us to preserve the data tensor information.

Monte Carlo simulation results have illustrated the effectiveness of the proposed multidimensional CX decomposition compared to the state-of-the-art THOSVD method, with a similar computation time, and with the advantage of keeping the information of the original data as needed for TC and tensor reconstruction methods.

4 Dual-polarized MIMO communication systems using combined TST-MSM-Kron coding

In this chapter, some contributions of this research are presented. First, a bibliographic overview of MIMO communication systems is presented in terms of modeling and coding schemes. In the second part of the chapter, we present a received signal model. We propose a new DD-DP MIMO system, which combines a multiple Kronecker product of symbol matrices with a TST coding, called TST-MSMKron coding. These codings allow us to jointly estimate transmitted symbols, channel, and channel parameters (steering matrices, and path losses) through the closed-form algorithms. The channel model is first detailed, where we consider a URA at both transmitter and receiver. It is represented by a fifth-order PARAFAC decomposition separated into two parts associated with the vertically (V_R)- and horizontally (H_R)- polarized receive antennas, respectively. Then, it is established that the tensor of received signals satisfy a new coupled nested Tucker-PARAFAC model, whose core tensor is the coding tensor. The identifiability conditions are established.

4.1 Bibliographic overview of MIMO systems

In the last years, massive MIMO systems have been the subject of intense research due to their great potential to provide energy efficiency and data rate gains [34]. However, high-speed wireless transmission has fundamental physical limitations that prevent the system from providing the best performance. Basic signal processing techniques normally consider only two dimensions (space and time) which leads to models based on a matrix. However, often the space domain can be split into more signal dimensions as azimuth and elevation, while the time domain can be divided into other dimensions, such as frames and sub-frames or we can also include on its structure other dimensions, for example, frequency, code and polarization that are not taken into account when matrix decompositions are used [93].

During the last decade, the use of multilinear algebra for communications systems has been growing [94]. Tensor decompositions have been extensively used to model received signals as well as to derive receivers to estimate the channel and symbols by exploiting various forms of diversity. The two most common tensor decompositions are PARAFAC [39] and Tucker [40]. PARAFAC decomposition is the most popular one and is extensively used in wireless communication systems because it presents a conceptual simplicity and a simple uniqueness property [50]. There are many systems based on the PARAFAC decomposition as the single input multiple output (SIMO)-OFDM and the MIMO-OFDM systems in [95]; the direct-sequence code division multiple access (DS-CDMA) uplink system in [96]; the channel estimation method for millimeter waves in [21]; the single input single output (SISO)-OFDM system proposed in [97]; MIMO-OFDM systems in [98, 99, 100]; and the channel tensor with a URA at both, transmitter and receiver proposed in [33]. Regarding Tucker decomposition, a massive MIMO-OFDM system is proposed in [101]; and a massive frequency division duplex (FDD) downlink MIMO system with a single BS equipped with a uniform planar array (UPA) and a single-antenna user, such that the system is presented in [102], where the channel is modeled by PARAFAC. For all the above works, the received signals are modeled as PARAFAC or Tucker decompositions that allow the proposition of semi-blind or supervised receivers to estimate channel and symbol matrices.

When the diversity in MIMO systems is exploited, it means that are created redundancies in the signal by exploiting the random nature of the radio propagation in such a way that different and independent versions of the same signal reach the destination. There are many ways to obtain diversity. For example, spatial diversity can be found in wireless MIMO systems through the use of multiple antennas at the transmit and receive nodes. The benefit from spatial diversity comes from the redundancies in the transmitted signal, leading the receive antennas to possibly obtain uncorrelated faded versions of the same signal increasing the probability of effective reception of the transmitted information. Time diversity can be obtained via coding where the information is coded and dispersed in the time domain in different periods so that different parts of the codewords experience roughly uncorrelated fading. The frequency diversity can be exploited when the channel is frequency-selective, where the signals are replayed across multiple subcarriers. Since diversity is such an important resource, a wireless system can use several types of diversity simultaneously [6].

In recent years, the application of codings with tensor approaches for the design of the MIMO wireless communication systems has led to the development of new tensor models such as the PARATUCK-(N_1,N) [8], generalized PARATUCK [9, 13], nested PARAFAC [12], and coupled nested Tucker [14]. Codes based on the Khatri-Rao and Kronecker products were incorporated into point-to-point MIMO and multi-hop MIMO systems to introduce extra diversity to the systems [10, 11]. This combination allows the proposition of new received tensor signals and associated semi-blind receivers. In Table 12 some tensor approaches for wireless communication systems with their received signals tensor models and codings are summarized. Sidiropoulos *et al.* (2000) [47] was the first that combined a DS-CDMA MIMO system with a coding matrix, where the received signals were modeled by the PARAFAC decomposition. Sidiropoulos & Budampati (2002) [103] presented the Khatri-Rao space time (KRST) coding, where a multi-antenna system modeled by the PARAFAC was proposed. The TST coding was considered in [8] to propose a point-to-point MIMO system using the

Reference	Received signals tensor model	MIMO system	Coding
Sidiropoulos *et al.* (2000)[47]	PARAFAC	DS-CDMA	Coding matrix
Sidiropoulos & Budampati (2002) [103]	PARAFAC	multi-antenna	KRST
Favier *et al.* (2012) [8]	(2,4)-PARATUCK	point-to-point	TST
Almeida & Favier (2013) [7]	nested PARAFAC	OFDM	D-KRSTF
Da Costa (2014) [104]	(2,4)-PARATUCK	multicarrier	TST and allocation resources
Favier & Almeida (2014) [9]	generalized PARATUCK	OFDM-CDMA	TSTF and allocation tensor
Da Costa *et al.* (2018) [58]	generalized PARATUCK	OFDM-CDMA	TSTF
De Pinho *et al.* (2019) [11]	PARAFAC	point-to-point	MSMKR
Randriambelonoro *et al.* (2021) [12]	nested PARAFAC	multipath	KRST and MKRSM

Table 12 – Tensor approaches for wireless communication systems.

(2,4)-PARATUCK decomposition. Almeida & Favier (2013) [7] combined the STF and KRST codings, called double-KRST (D-KRSTF), to propose a MIMO-OFDM system modeled by the nested PARAFAC decomposition. Da Costa (2014) [104] combined TST coding with resource allocation, where the received signals were modeled as (2,4)-PARATUCK. Also, it considered an existing multicarrier MIMO wireless system with STF coding to compare against the proposed system. A MIMO OFDM-CDMA system is proposed by Favier & Almeida (2014) [9] where a tensor space-time-frequency (TSTF) coding is combined with an allocation tensor. The received signals are modeled by the generalized PARATUCK decomposition. A MIMO OFDM-CDMA system combined with TSTF was presented in [58]. The received signal was modeled by the generalized PARATUCK decomposition. De Pinho *et al.* (2019) [11] introduced the MSMKR coding as a particular case of MKRST coding and proposes a point-to-point MIMO system modeled by PARAFAC. A MIMO system in a time-varying multipath environment combined with KRST and MKRSM codings was proposed by Randriambelonoro *et al.* (2021) [12]. The system is equipped with ULAs at both, transmit and receiver. The received signal is modeled by a $(N + 2)$-order nested PARAFAC decomposition. The use of these codings with tensor approaches for the design of MIMO wireless communication systems has led to the development of new tensor models. Overviews of such MIMO systems can be found in [58].

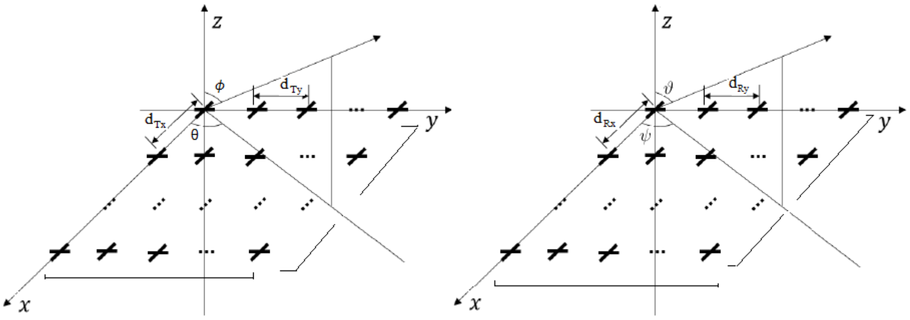
Millimeter-wave (mmWave) communication systems are systems with large antennas

Reference	Received signals tensor model	mmWave MIMO system	Coding
Zhou *et al.* (2017) [23]	-	FD-MIMO	precoding matrices
Du *et al.* (2021) [105]	optimized nested PARAFAC	massive	KRST
Elganimi *et al.* (2022) [106]	-	IRS-assited	OSTBC
Chang *et al.* (2021) [107]	PARAFAC	OFDM	hybrid precoding
Gherekhloo *et al.* (2021) [108]	PARAFAC	point-to-point RIS-aided	precoding and decoding matrix

Table 13 – Tensor approaches for mmWave MIMO systems.

arrays at both the transmitter and receiver, they operate in the 30-300 GHz spectrum and can achieve gigabit-per-second data rates [16, 17, 18]. Massive MIMO system is combined with mmWave communications to improve system capacity and spectral efficiency. Table 13 summarizes mmWave MIMO systems with their received signals tensor models and codings. Zhou *et al.* (2017) [23] considered mmWave FD-MIMO based unmanned aerial vehicle (UAV) base station (BS) combined with a precoding matrix. Du *et al.* (2021) [105] proposed a massive mmWave MIMO system combined with KRST coding and modeled by the PARAFAC decomposition. An intelligent reflecting surface (IRS)-assisted mmWave massive MIMO system was proposed by Elganimi *et al.* (2022) [106] with transmit antenna selection (TAS) using orthogonal space-time block codes (OSTBC) scheme. Chang *et al.* (2021) [107] considered a multiuser OFDM massive mmWave and adopted a hybrid precoding structure to model the received signals as a low-rank PARAFAC decomposition. A RIS-aided mmWave MIMO communication system was presented in [108], where the low-rank nature of mmWave MIMO channels was considered and the received training signals were written as a low-rank multi-way tensor admitting a PARAFAC decomposition. Systems presented in Table 13 generally are modeled by the PARAFAC decomposition and consider pilot-assisted symbols to be transmitted. MmWave systems combined with matrix/tensor coding schemes still need to be studied.

In this chapter, we propose a new DD-DP MIMO system presented in [109]. This system is equipped with URAs at both ends of the link, where a new coding based on the Kronecker product of symbol matrices is combined with a TST coding to transmit the symbols. Our proposed system differs from the systems in Table 13 by the proposed TST-MSMKron coding which induces an increase in the diversity gain than the codings used by the systems presented above, and by the proposition of receivers that can estimate the symbol matrices, channel and channel parameters due to the signal tensor model proposed, whereas all the systems in Table 13 requires knowledge of the symbols transmitted (pilot symbols). In the



(a) URA at the transmitter

(b) URA at the receiver

Figure 14 – Illustration of an URA at the transmitter (a) and receiver (b).

next sections, the system model will be presented, in terms of the channel model, coding and received signals tensor.

4.2 System model

Consider a DD-DP MIMO wireless communication system, equipped with uniform rectangular arrays (URAs) both at the transmitter and receiver, as considered in [33]. The objective of this section is to present the system model under the standard form, in the noise-free case:

$$\mathbf{X} = \mathbf{H}\mathbf{V}, \tag{4.1}$$

where $\mathbf{H}$, $\mathbf{V}$ and $\mathbf{X}$ represent respectively matrix unfoldings of the channel tensor, and of tensors containing the coded signals to be transmitted and the received signals. This result will be demonstrated via the theoretical developments in the next three subsections, leading to the system model (4.38). The tensor model of the channel is first detailed. Then, in Sections 4.2.2 and 4.2.3, new coding and the tensor model of received signals will be described, respectively.

4.2.1 Channel tensor model

We assume that the receiver and transmitter employ URAs, with cross-polarized transmit and receive antenna pairs, respectively. Each antenna pair consists of a vertical (V)- and an horizontal (H)-polarized antenna, as illustrated in Figures 14 (a) and (b), respectively. Each cross represents a cross-polarized antenna pair [25].

Defining (M_{R_x}, M_{R_y}) and (M_{T_x}, M_{T_y}) as the numbers of receive and transmit antennas, in directions x and y, respectively, and $M_T = M_{T_x}M_{T_y}$ and $M_R = M_{R_x}M_{R_y}$ as the total numbers of transmit and receive antennas, respectively, the DD-DP channel is represented as a matrix $\mathbf{H} \in \mathbb{C}^{2M_RM_T}$, partitioned as follows [25]:

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}^{(V_R,V_T)} & \mathbf{H}^{(V_R,H_T)} \\ \mathbf{H}^{(H_R,V_T)} & \mathbf{H}^{(H_R,H_T)} \end{bmatrix} = \begin{bmatrix} \mathbf{H}^{(V_R)} \\ \mathbf{H}^{(H_R)} \end{bmatrix}, \tag{4.2}$$

where the row-blocks $\mathbf{H}^{(V_R)}$ and $\mathbf{H}^{(H_R)}$ represent the channels of each polarization at the receiver, $\mathbf{H}^{(V_R,V_T)} \in \mathbb{C}^{M_R \times M_T}$ is the sub-channel matrix between V-polarized transmit and receive antennas, and $\mathbf{H}^{(V_R,H_T)} \in \mathbb{C}^{M_R \times M_T}$ is the sub-channel matrix between H-polarized transmit and V-polarized receive antennas. The other two blocks in (4.2) are defined similarly. Each sub-channel $\mathbf{H}^{(r,q)}$, with $r \in \{V_R, H_R\}$ and $q \in \{V_T, H_T\}$ is modeled as:

$$\mathbf{H}^{(r,q)} = \mathbf{A}_R \text{diag}\left(\mathbf{b}^{(r,q)}\right) \mathbf{A}_T^T \in \mathbb{C}^{M_R \times M_T}, \tag{4.3}$$

where $b_k^{(r,q)}$ is the complex path-loss vector defined as:

$$\mathbf{b}^{(r,q)} = \begin{bmatrix} b_1^{(r,q)} & \dots & b_K^{(r,q)} \end{bmatrix}^T \in \mathbb{C}^K, \tag{4.4}$$

with $b_k^{(r,q)}$ the complex path-loss for the k-th path and the (r,q)-th sub-channel, containing the polarization and fading effects, and K is the number of path between the transmitter and receiver.

The steering matrices are given by:

$$\mathbf{A}_T = \ [\mathbf{a}_T(\theta_1,\phi_1) \ \dots \ \mathbf{a}_T(\theta_K,\phi_K)] \ = \mathbf{A}_{T_x} \diamond \mathbf{A}_{T_y} \in \mathbb{C}^{M_T \times K}, \tag{4.5}$$

$$\mathbf{A}_R = \ [\mathbf{a}_R(\psi_1,\vartheta_1) \ \dots \ \mathbf{a}_R(\psi_K,\vartheta_K)] \ = \mathbf{A}_{R_x} \diamond \mathbf{A}_{R_y} \in \mathbb{C}^{M_R \times K}, \tag{4.6}$$

where $\mathbf{A}_{T_x} \in \mathbb{C}^{M_{T_x} \times K}$, $\mathbf{A}_{T_y} \in \mathbb{C}^{M_{T_y} \times K}$, $\mathbf{A}_{R_x} \in \mathbb{C}^{M_{R_x} \times K}$ and $\mathbf{A}_{R_y} \in \mathbb{C}^{M_{R_y} \times K}$ are the steering matrices in directions x and y, at the transmitter and receiver.

The array steering vectors for the k-th path are defined as:

$$\mathbf{a}_T(\theta_k,\phi_k) = \mathbf{a}_{T_x,k} \otimes \mathbf{a}_{T_y,k} \in \mathbb{C}^{M_{T_x} M_{T_y}}, \tag{4.7}$$

$$\mathbf{a}_R(\psi_k,\vartheta_k) = \mathbf{a}_{R_x,k} \otimes \mathbf{a}_{R_y,k} \in \mathbb{C}^{M_{R_x} M_{R_y}}, \tag{4.8}$$

Writing the steering matrices as:

$$\mathbf{A}_P = \ [\mathbf{a}_{P,1} \ \dots \ \mathbf{a}_{P,K}] \ \in \mathbb{C}^{M_P \times K}, \tag{4.9}$$

with $P \in \{T_x, T_y, R_x, R_y\}$, each steering vector is a Vandermonde vector, function of both azimuth and elevation angles, such as:

$$\mathbf{a}_{T_x,k} = \begin{bmatrix} 1 & e^{-j\omega_{T_x,k}} & \dots & e^{-j\omega_{T_x,k}(M_{T_x}-1)} \end{bmatrix}^T, \tag{4.10}$$

$$\mathbf{a}_{T_y,k} = \begin{bmatrix} 1 & e^{-j\omega_{T_y,k}} & \dots & e^{-j\omega_{T_y,k}(M_{T_y}-1)} \end{bmatrix}^T, \tag{4.11}$$

$$\mathbf{a}_{R_x,k} = \begin{bmatrix} 1 & e^{-j\omega_{R_x,k}} & \dots & e^{-j\omega_{R_x,k}(M_{R_x}-1)} \end{bmatrix}^T, \tag{4.12}$$

$$\mathbf{a}_{R_y,k} = \begin{bmatrix} 1 & e^{-j\omega_{R_y,k}} & \dots & e^{-j\omega_{R_y,k}(M_{R_y}-1)} \end{bmatrix}^T, \tag{4.13}$$

which assumes that the reference antennas are placed at the origin. The angular frequencies are defined as:

$$\omega_{T_x,k} = \frac{2\pi(m_{T_x} - 1)d_{T_x}\sin(\phi_k)\cos(\theta_k)}{\nu}, \tag{4.14}$$

$$\omega_{T_y,k} = \frac{2\pi(m_{T_y} - 1)d_{T_y}\sin(\phi_k)\sin(\theta_k)}{\nu}, \tag{4.15}$$

$$\omega_{R_x,k} = \frac{2\pi(m_{R_x} - 1)d_{R_x}\sin(\vartheta_k)\cos(\psi_k)}{\nu}, \tag{4.16}$$

$$\omega_{R_y,k} = \frac{2\pi(m_{R_y} - 1)d_{R_y}\sin(\vartheta_k)\sin(\psi_k)}{\nu}, \tag{4.17}$$

where (ϕ_k, θ_k) and (ϑ_k, ψ_k) are the DoD and DoA azimuth and elevation angles for the k-th path, respectively, ν is the wavelength, (d_{T_x}, d_{T_y}) and (d_{R_x}, d_{R_y}) are the antenna spacings for horizontal and vertical units at the transmitter and receiver, respectively.

The structure (4.13) of the steering vectors leads to the following Vandermonde structure for the steering matrices:

$$\mathbf{A}_P = \begin{bmatrix} 1 & \cdots & 1 \\ e^{-j\omega_{P,1}} & \cdots & e^{-j\omega_{P,K}} \\ \vdots & \vdots & \ddots \\ e^{-j\omega_{P,1}(M_P-1)} & \cdots & e^{j\omega_{P,K}(M_P-1)} \end{bmatrix} \in \mathbb{C}^{M_P \times K}, \tag{4.18}$$

where $\omega_{P,k}$ is the P-th angular frequency along the k-th path.

Substituting (4.3) into (4.2) gives:

$$\mathbf{H} = \begin{bmatrix} \mathbf{A}_R & 0 \\ 0 & \mathbf{A}_R \end{bmatrix} \begin{bmatrix} D_1\left(\mathbf{B}^{(V_R)}\right) & D_2\left(\mathbf{B}^{(V_R)}\right) \\ D_1\left(\mathbf{B}^{(H_R)}\right) & D_2\left(\mathbf{B}^{(H_R)}\right) \end{bmatrix} \begin{bmatrix} \mathbf{A}_T & 0 \\ 0 & \mathbf{A}_T \end{bmatrix}^T, \tag{4.19}$$

with:

$$\mathbf{B}^{(V_R)} = \begin{bmatrix} \mathbf{b}^{(V_R,V_T)} & \mathbf{b}^{(V_R,H_T)} \end{bmatrix}^T \in \mathbb{C}^{2\times K}, \tag{4.20}$$

$$\mathbf{B}^{(H_R)} = \begin{bmatrix} \mathbf{b}^{(H_R,V_T)} & \mathbf{b}^{(H_R,H_T)} \end{bmatrix}^T \in \mathbb{C}^{2\times K}, \tag{4.21}$$

Using property (2.1), the channels $\mathbf{H}^{(V_R)}$ and $\mathbf{H}^{(H_R)}$ can be written as:

$$\mathbf{H}^{(V_R)} = \mathbf{A}_R \left(\mathbf{B}^{(V_R)} \diamond \mathbf{A}_T\right)^T, \tag{4.22}$$

$$\mathbf{H}^{(H_R)} = \mathbf{A}_R \left(\mathbf{B}^{(H_R)} \diamond \mathbf{A}_T\right)^T. \tag{4.23}$$

Replacing $\mathbf{A}_T$ and $\mathbf{A}_R$ by their expressions (4.5) and (4.6), respectively, gives:

$$\mathbf{H}^{(V_R)} = \left(\mathbf{A}_{R_x} \diamond \mathbf{A}_{R_y}\right)\left(\mathbf{B}^{(V_R)} \diamond \mathbf{A}_{T_x} \diamond \mathbf{A}_{T_y}\right)^T, \tag{4.24}$$

$$\mathbf{H}^{(H_R)} = \left(\mathbf{A}_{R_x} \diamond \mathbf{A}_{R_y}\right)\left(\mathbf{B}^{(H_R)} \diamond \mathbf{A}_{T_x} \diamond \mathbf{A}_{T_y}\right)^T. \tag{4.25}$$

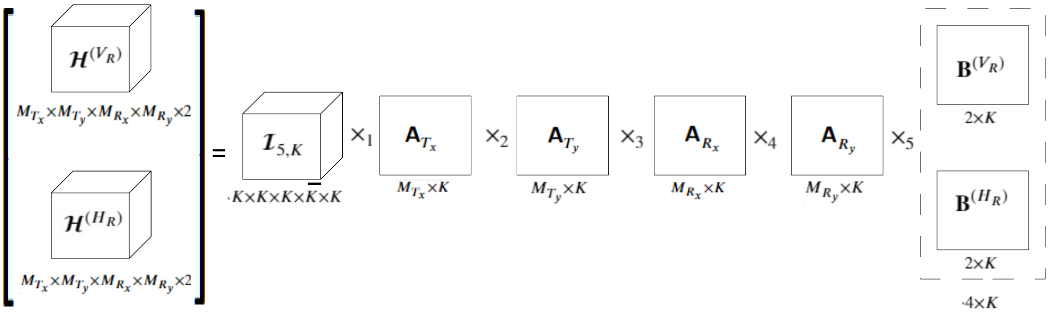


Figure 15 – Global channel tensor $\mathcal{H}$ modeled as a PARAFAC decomposition.

These equations can be interpreted as matrix unfoldings of two fifth-order tensors $\mathcal{H}^{(V_R)}$ and $\mathcal{H}^{(H_R)}$ satisfying the PARAFAC decompositions $[\![\mathbf{A}_{T_x}, \mathbf{A}_{T_y}, \mathbf{A}_{R_x}, \mathbf{A}_{R_y}, \mathbf{B}^{(V_R)}; K]\!]$ and $[\![\mathbf{A}_{T_x}, \mathbf{A}_{T_y}, \mathbf{A}_{R_x}, \mathbf{A}_{R_y}, \mathbf{B}^{(H_R)}; K]\!]$, respectively, i.e.:

$$\mathcal{H}^{(V_R)} = \mathcal{I}_{5,K} \times_1 \mathbf{A}_{T_x} \times_2 \mathbf{A}_{T_y} \times_3 \mathbf{A}_{R_x} \times_4 \mathbf{A}_{R_y} \times_5 \mathbf{B}^{(V_R)} \in \mathbb{C}^{M_{T_x}\times M_{T_y}\times M_{R_x}\times M_{R_y}\times 2}, \tag{4.26}$$

$$\mathcal{H}^{(H_R)} = \mathcal{I}_{5,K} \times_1 \mathbf{A}_{T_x} \times_2 \mathbf{A}_{T_y} \times_3 \mathbf{A}_{R_x} \times_4 \mathbf{A}_{R_y} \times_5 \mathbf{B}^{(H_R)} \in \mathbb{C}^{M_{T_x}\times M_{T_y}\times M_{R_x}\times M_{R_y}\times 2}. \tag{4.27}$$

The model (4.26) can also be written as the sum of K rank-one tensors:

$$\mathcal{H}^{(V_R)} = \sum_{k=1}^{K} \mathbf{a}_{T_x,k} \circ \mathbf{a}_{T_y,k} \circ \mathbf{a}_{R_x,k} \circ \mathbf{a}_{R_y,k} \circ \mathbf{b}^{(V_R)}_{.k}, \tag{4.28}$$

where $\mathbf{b}^{(V_R)}_{.k} \in \mathbb{C}^2$, with $\mathbf{b}^{(V_R)}_{1.} = \left[\mathbf{b}^{(V_R,V_T)}\right]^T$, and $\mathbf{b}^{(V_R)}_{2.} = \left[\mathbf{b}^{(V_R,H_T)}\right]^T$. This writing of the channel $\mathcal{H}^{(V_R)}$ highlights the transmission through K paths, between the transmitter and the V-polarized receive antennas. Similarly, the PARAFAC model (4.27) can be written as Eq. (4.28), replacing $\mathbf{b}^{(V_R)}_{.k}$ by $\mathbf{b}^{(H_R)}_{.k}$, with $\mathbf{b}^{(H_R)}_{1.} = \left[\mathbf{b}^{(H_R,V_T)}\right]^T$, and $\mathbf{b}^{(H_R)}_{2.} = \left[\mathbf{b}^{(H_R,H_T)}\right]^T$.

Concatenating the channel tensors $\mathcal{H}^{(V_R)}$ and $\mathcal{H}^{(H_R)}$ along the fifth mode and defining $\mathbf{B} \in \mathbb{C}^{4\times K}$ as the row-block concatenation of the matrices $\mathbf{B}^{(V_R)}$ and $\mathbf{B}^{(H_R)}$:

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}^{(V_R)} \\ \mathbf{B}^{(H_R)} \end{bmatrix}, \tag{4.29}$$

the global channel tensor $\mathcal{H} \in \mathbb{C}^{M_{T_x}\times M_{T_y}\times M_{R_x}\times M_{R_y}\times 4}$ can be written as:

$$\mathcal{H} = \begin{bmatrix} \mathcal{H}^{(V_R)} \\ \mathcal{H}^{(H_R)} \end{bmatrix} = \mathcal{I}_{5,K} \times_1 \mathbf{A}_{T_x} \times_2 \mathbf{A}_{T_y} \times_3 \mathbf{A}_{R_x} \times_4 \mathbf{A}_{R_y} \times_5 \mathbf{B} \in \mathbb{C}^{M_{T_x}\times M_{T_y}\times M_{R_x}\times M_{R_y}\times 4}, \tag{4.30}$$

which is a fifth-order rank-K PARAFAC decomposition illustrated in Figure 15.

The resulting channel tensor is an extension of the fourth-order one presented in [25], with a URA at the transmitter and a uniform linear array (ULA) at the receiver. Note that the considered channel modeling does not take delays and Doppler shifts into account, assuming the channel bandwidth is sufficiently small and the channel is sufficiently slowly varying over the transmission duration. In Table 14, the definitions of the channel parameters are summarized.

Channel parameters	Definitions
M_{T_x}, M_{T_y}	numbers of transmit antennas in directions x and y, respectively
$M_T = M_{T_x} M_{T_y}$	total number of transmit antennas
M_{R_x}, M_{R_y}	numbers of receive antennas in directions x and y, respectively
$M_R = M_{R_x} M_{R_y}$	total number of receive antennas
K	number of paths
ϕ_k, θ_k	azimuth and elevation DoD angles for the k-th path
ϑ_k, ψ_k	azimuth and elevation DoA angles for the k-th path

Table 14 – Channel parameters.

4.2.2 TST-MSMKron coding

In this section, we present the coding used at the transmitter which combines TST and MSMKRon codings presented in [109]. This last one can be viewed as the MKronST coding [10], without prior knowledge of one symbol matrix. It consists in a multiple Kronecker product of N symbol matrices $\mathbf{S}^{(n)} \in \mathbb{C}^{J_n \times R_n}$, $n \in [1, N]$, each one being formed of R_n data streams composed of J_n symbols each:

$$\mathbf{S} = \otimes_{n=1}^{N} \mathbf{S}^{(n)} \triangleq \mathbf{S}^{(1)} \otimes \ldots \otimes \mathbf{S}^{(N)} \in \mathbb{C}^{J \times R}, \tag{4.31}$$

where $R = \prod_{n=1}^{N} R_n$ and $J = \prod_{n=1}^{N} J_n$. The scalar writing of (4.31) is as follows:

$$s_{j,r} = \prod_{n=1}^{N} s_{j_n, r_n}^{(n)}, \quad j \in [1, J], \quad r \in [1, R], \tag{4.32}$$

with $j = j_N^{(N)} + \left(j_{N-1}^{(N-1)} - 1\right) J_N + \cdots + \left(j_1^{(1)} - 1\right) \prod_{n=2}^{N} J_n$, and $r = r_N^{(N)} + \left(r_{N-1}^{(N-1)} - 1\right) R_N + \cdots + \left(r_1^{(1)} - 1\right) \prod_{n=2}^{N} R_n$, where $j_n^{(n)} \in [1, J_n]$ and $r_n^{(n)} \in [1, R_n]$ denote the indices j_n and r_n in $s_{j_n, r_n}^{(n)}$. This writing (4.32) highlights the way to compute each element of the multiple Kronecker product matrix S.

The MSMKron coding is combined with a coding tensor $\mathcal{G} \in \mathbb{C}^{2M_T \times R_1 \times \ldots \times R_N \times P}$ in such a way that the coded signals form an $(N+2)$-order tensor which satisfies the Tucker model given by:

$$\mathcal{V} = \mathcal{G} \times_1 \mathbf{I}_{2M_T} \times_2 \mathbf{S}^{(1)} \times_3 \ldots \times_{N+1} \mathbf{S}^{(N)} \times_{N+2} \mathbf{I}_P \in \mathbb{C}^{2M_T \times J_1 \times \ldots \times J_N \times P}. \tag{4.33}$$

Note that the core tensor of this decomposition is the coding tensor $\mathcal{G}$. The coded signals to be transmitted by the m_t-th transmit antenna, during the p-th time block, and the j_n-th symbol period, with $n \in [1, N]$, can be written in scalar form as:

$$v_{m_t, j_1, \cdots, j_N, p} = \sum_{r_1=1}^{R_1} \cdots \sum_{r_N=1}^{R_N} g_{m_t, r_1, \cdots, r_N, p} \prod_{n=1}^{N} s_{j_n, r_n}, \tag{4.34}$$

Steering matrices
$\mathbf{A}_P = \begin{bmatrix} \mathbf{a}_{P,1} & \dots & \mathbf{a}_{P,K} \end{bmatrix} \in \mathbb{C}^{M_P \times K}$, for $P \in [T_x, T_y, R_x, R_y]$
$\mathbf{A}_T = \mathbf{A}_{T_x} \diamond \mathbf{A}_{T_y} \in \mathbb{C}^{M_T \times K}$, with $M_T = M_{T_x} M_{T_y}$
$\mathbf{A}_R = \mathbf{A}_{R_x} \diamond \mathbf{A}_{R_y} \in \mathbb{C}^{M_R \times K}$, with $M_R = M_{R_x} M_{R_y}$
Path-loss matrices
$\mathbf{B}^{(r)} \in \mathbb{C}^{2 \times K}$, with $r \in [V_R, H_R]$
$\mathbf{B} = \begin{bmatrix} \mathbf{B}^{(V_R)} \\ \mathbf{B}^{(H_R)} \end{bmatrix} \in \mathbb{C}^{4 \times K}$
Channel matrices and tensors
$\mathbf{H}^{(r,q)} \in \mathbb{C}^{M_R \times M_T}$, with $r \in [V_R, H_R]$ and $q \in [V_T, H_T]$
$\mathbf{H}^{(r)} \in \mathbb{C}^{M_R \times 2M_T}$
$\mathbf{H} = \begin{bmatrix} \mathbf{H}^{(V_R)} \\ \mathbf{H}^{(H_R)} \end{bmatrix} \in \mathbb{C}^{2M_R \times 2M_T}$
$\mathcal{H}^{(r)} \in \mathbb{C}^{M_{T_x} \times M_{T_y} \times M_{R_x} \times M_{R_y} \times 2}$
$\mathcal{H} \in \mathbb{C}^{M_{T_x} \times M_{T_y} \times M_{R_x} \times M_{R_y} \times 4}$
Symbol matrices
$\mathbf{S}^{(n)} \in \mathbb{C}^{J_n \times R_n}$, for $n \in [1,N]$
$\mathbf{S} = \mathbf{S}^{(1)} \otimes \dots \otimes \mathbf{S}^{(N)} \in \mathbb{C}^{J \times R}$
$R = \prod_{n=1}^{N} R_n,\ J = \prod_{n=1}^{N} J_n$
Tensor space-time coding
$\mathcal{G} \in \mathbb{C}^{2M_T \times R_1 \times \dots \times R_N \times P}$
Received signals tensor
$\mathcal{X} \in \mathbb{C}^{2M_R \times J_1 \times \dots \times J_N \times P}$

Table 15 – System matrices and tensors.

with $m_t \in [1, 2M_T]$. The following matrix unfolding can be easily deduced:

$$\mathbf{V}_{2M_T \times PJ} = \mathbf{G}_{2M_T \times PR} \left(\mathbf{I}_P \otimes \mathbf{S} \right)^T \in \mathbb{C}^{2M_T \times PJ}, \tag{4.35}$$

where $\mathbf{G}_{2M_T \times PR} \in \mathbb{C}^{2M_T \times PR}$ is an unfolding of the coding tensor $\mathcal{G}$, and $\mathbf{S}$ is defined in (4.31). The TST-MSMKron coding increases space-time diversity, as it will be illustrated in the simulations. Note that the identity matrix $\mathbf{I}_P \in \mathbb{R}^{P \times P}$ in (4.33), which gives the Kronecker product in (4.35), is associated with P repetitions of the symbol matrices via transmission in P blocks. The definitions of the system matrices and tensors are summarized in Table 15.

4.2.3 Received signals model

In this section, the tensor model of the received signals is described. In the noise-free case, the received signals are obtained by transmitting the coded signals tensor (4.33) through the channel $\mathbf{H}$ defined in (4.2), which gives [109]:

$$\mathcal{X} = \mathcal{V} \times_1 \mathbf{H} = \mathcal{G} \times_1 \mathbf{H} \times_2 \mathbf{S}^{(1)} \times_3 \dots \times_{N+1} \mathbf{S}^{(N)} \times_{N+2} \mathbf{I}_P. \tag{4.36}$$

Regarding expression (4.36), we conclude that $\mathcal{X}$ represents the $(N+2)$-order Tucker model whose core tensor is the coding tensor $\mathcal{G}$ and the matrix factors are the channel $\mathbf{H}$ and the symbol matrices $\mathbf{S}^{(n)}$, $n \in [1, N]$ and the identity matrix $\mathbf{I}_P$. The signal received by the m_R-th receive antenna, during the p-th time block, and the j_n-th symbol period, with $n \in [1, N]$, can be written as:

$$x_{m_R, j_1, \cdots, j_N, p} = \sum_{m_T=1}^{2M_T} \sum_{r_1=1}^{R_1} \cdots \sum_{r_N=1}^{R_N} g_{m_t, r_1, \cdots, r_N, p} \, h_{m_R, m_T} \prod_{n=1}^{N} s_{j_n, r_n}^{(n)}, \tag{4.37}$$

with $m_R \in [1, 2M_R]$. The received signal in (4.36) can also be expressed in terms of mode-1 unfolding:

$$\mathbf{X}_{2M_R \times PJ} = \mathbf{H}\mathbf{G}_{2M_T \times PR} \left(\mathbf{I}_P \otimes \mathbf{S}\right)^T. \tag{4.38}$$

Note that the transmission via the channel $\mathbf{H}$ can be interpreted as a mode-1 linear transformation, with matrix $\mathbf{H}$, applied to the tensor $\mathcal{V}$ of coded signals. Considering the channel blocks defined by (4.24) and (4.25), the received signals for each polarization can be separated from (4.38) as follows:

$$\mathbf{X}_{M_R \times PJ}^{(V_R)} = \mathbf{H}^{(V_R)} \mathbf{G}_{2M_T \times PR} \left(\mathbf{I}_P \otimes \mathbf{S}\right)^T \in \mathbb{C}^{M_R \times PJ}, \tag{4.39}$$

$$\mathbf{X}_{M_R \times PJ}^{(H_R)} = \mathbf{H}^{(H_R)} \mathbf{G}_{2M_T \times PR} \left(\mathbf{I}_P \otimes \mathbf{S}\right)^T \in \mathbb{C}^{M_R \times PJ}. \tag{4.40}$$

From (4.36)-(4.40), we can conclude that the tensor of received signals satisfies an $(N+2)$-order coupled nested Tucker-PARAFAC model, as illustrated by means of the blocks diagram in Figure 16. The PARAFAC part of this model is associated with the channel tensor $\mathcal{H}$ represented by the PARAFAC model (4.30), whose unfolding $\mathbf{H}$ is the factor matrix along the first mode of $\mathcal{X}$. That explains the nested Tucker-PARAFAC structure. The coupled structure is due to the coding that is common to the two components $(\mathcal{X}^{(V_R)}, \mathcal{X}^{(H_R)})$ of $\mathcal{X}$, associated with the V-polarized $(\mathcal{X}^{(V_R)})$ and H-polarized $(\mathcal{X}^{(H_R)})$ receive antennas, respectively. This coupling is composed of the MSMKron and TST coding blocks, as shown on Figure 16.

Combining modes 2 to $N+1$ of tensors $\mathcal{G}$ and $\mathcal{X}$ results in contracted forms $\mathcal{G}_c \in \mathbb{C}^{2M_T \times R \times P}$ and $\mathcal{X}_c \in \mathbb{C}^{2M_R \times J \times P}$, and the expression (4.36) can be rewritten as:

$$\mathcal{X}_c = \mathcal{G}_c \times_1 \mathbf{H} \times_2 \mathbf{S} \times_3 \mathbf{I}_P. \tag{4.41}$$

Based on the contracted third-order Tucker model (4.41), it is easy to deduce the following matrix unfoldings of tensor $\mathcal{X}$:

$$\mathbf{X}_{PJ \times 2M_R} = \left(\mathbf{I}_P \otimes \mathbf{S}\right) \mathbf{G}_{PR \times 2M_T} \mathbf{H}^T \in \mathbb{C}^{PJ \times 2M_R}, \tag{4.42}$$

$$\mathbf{X}_{2PM_R \times J} = \left(\mathbf{I}_P \otimes \mathbf{H}\right) \mathbf{G}_{2PM_T \times R} \mathbf{S}^T \in \mathbb{C}^{2PM_R \times J}, \tag{4.43}$$

$$\mathbf{X}_{P \times 2M_R J} = \mathbf{G}_{P \times 2M_T R} \left(\mathbf{H} \otimes \mathbf{S}\right)^T \in \mathbb{C}^{P \times 2M_R J}. \tag{4.44}$$

The advantages of this model are:

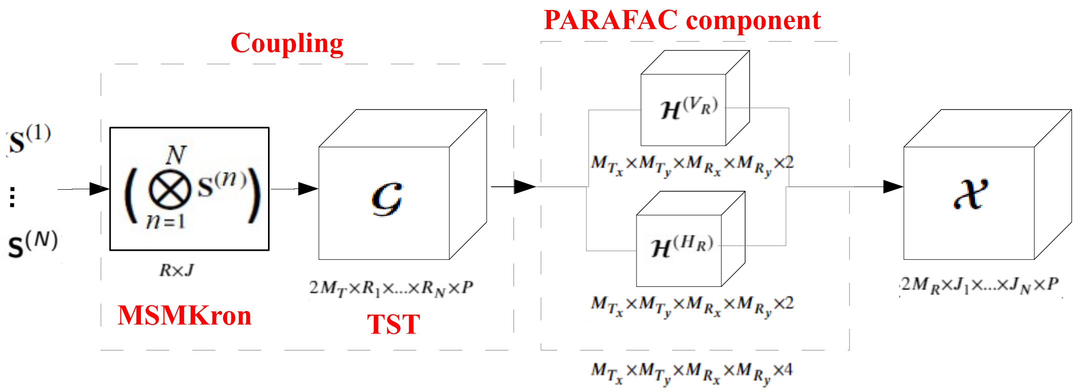

Coupling
PARAFAC component
$\mathbf{S}^{(1)}$
⋮
$\mathbf{S}^{(N)}$
$\left(\bigotimes_{n=1}^{N} \mathbf{S}^{(n)}\right)$
$R{\times}J$
MSMKron
$\mathcal{G}$
$2M_T{\times}R_1{\times}...{\times}R_N{\times}P$
TST
$\mathcal{H}^{(V_R)}$
$M_{T_x}{\times}M_{T_y}{\times}M_{R_x}{\times}M_{R_y}{\times}2$
$\mathcal{H}^{(H_R)}$
$M_{T_x}{\times}M_{T_y}{\times}M_{R_x}{\times}M_{R_y}{\times}2$
$M_{T_x}{\times}M_{T_y}{\times}M_{R_x}{\times}M_{R_y}{\times}4$
$\mathcal{X}$
$2M_R{\times}J_1{\times}...{\times}J_N{\times}P$

Figure 16 – Block diagram of the system represented by a coupled nested Tucker-PARAFAC model.

- The path-losses can be obtained separately for each polarization in the reception, consequently it is possible to identify the received signal and the channel for each polarization in the reception;
- The TST-MSMKron coding allows for an increase in the space-time diversity that it is imposed by multiple Kronecker products of symbol matrices combined with the TST coding;
- The symbol matrices transmitted, the channel, the steering matrices and the path-losses can be estimated by semi-blind receivers without prior knowledge of the channel, as long as the TST coding and one symbol of each symbol matrix are known.

4.2.3.1 Uniqueness analysis

System model uniqueness results from uniqueness of the Tucker model (4.36) of the received signals tensor $\mathcal{X}$ combined with uniqueness of the PARAFAC decomposition (4.30) of the channel tensor $\mathcal{H}$ [109].

As it is well known, Tucker models are generally not essentially unique [40]. Uniqueness can be obtained by imposing some constraints on the core tensor or the factor matrices. For the Tucker model (4.41), the core tensor will be assumed to be known at the reception, which implies uniqueness of the factor matrices up to scaling factors such as [109]:

$$\hat{\mathbf{S}} = \mathbf{S}\lambda^{(S)}, \quad \hat{\mathbf{H}} = \mathbf{H}\lambda^{(H)}, \quad \lambda^{(S)}\lambda^{(H)} = 1\ . \tag{4.45}$$

The channel tensor (4.30) is modeled by a rank-K PARAFAC decomposition that is essentially unique if the following sufficient condition is satisfied [51]:

$$k_{\mathbf{A}_{T_x}} + k_{\mathbf{A}_{T_y}} + k_{\mathbf{A}_{R_x}} + k_{\mathbf{A}_{R_y}} + k_{\mathbf{B}} \geq 2K + 2, \tag{4.46}$$

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

If the condition (4.46) is satisfied, any set of matrices ($\hat{\mathbf{A}}_{T_x}$, $\hat{\mathbf{A}}_{T_y}$, $\hat{\mathbf{A}}_{R_x}$, $\hat{\mathbf{A}}_{R_y}$, $\hat{\mathbf{B}}$) that satisfy (4.30) are related with the original matrices by:

$$\begin{aligned} \hat{\mathbf{A}}_{T_y} &= \mathbf{A}_{T_y}\mathbf{\Pi}\mathbf{\Lambda}^{(T_y)}, & \hat{\mathbf{A}}_{T_x} &= \mathbf{A}_{T_x}\mathbf{\Pi}\mathbf{\Lambda}^{(T_x)}, \\ \hat{\mathbf{A}}_{R_y} &= \mathbf{A}_{R_y}\mathbf{\Pi}\mathbf{\Lambda}^{(R_y)}, & \hat{\mathbf{A}}_{R_x} &= \mathbf{A}_{R_x}\mathbf{\Pi}\mathbf{\Lambda}^{(R_x)}, \\ \hat{\mathbf{B}} &= \mathbf{B}\mathbf{\Pi}\mathbf{\Lambda}^{(B)}, & \mathbf{\Lambda}^{(T_y)}\mathbf{\Lambda}^{(T_x)}\mathbf{\Lambda}^{(R_y)}\mathbf{\Lambda}^{(R_x)}\mathbf{\Lambda}^{(B)} &= \mathbf{I}_K, \end{aligned} \tag{4.47}$$

where $\mathbf{\Pi} \in \mathbb{C}^{K\times K}$ is a permutation matrix and $\mathbf{\Lambda}^{(f)}$ are diagonal matrices, with $f \in \{T_y, T_x, R_y, R_x, B\}$. Note that the permutation ambiguity is irrelevant from a practical point of view, since it corresponds to a permutation of paths. Due to the random nature of the channel parameters, the matrices $\mathbf{A}_{T_x}$, $\mathbf{A}_{T_y}$, $\mathbf{A}_{R_x}$, $\mathbf{A}_{R_y}$ and $\mathbf{B}$ are full rank. Then, in the generic case, condition (4.46) becomes:

$$\min\left(M_{T_y},K\right) + \min\left(M_{T_x},K\right) + \min\left(M_{R_y},K\right) + \min\left(M_{R_x},K\right) + \min\left(4,K\right) \geq 2K+2. \tag{4.48}$$

If we assume that all steering and path-loss matrices are full column rank, which implies $K \leq \min(M_{T_y}, M_{T_x}, M_{R_y}, M_{R_x}, 4)$, then the above condition is always satisfied, which ensures essential uniqueness of the PARAFAC decomposition of the channel tensor. Note that a small value of K results from the sparse nature of the channel. In this work the tensor rank is assumed to be known. When K is unknown, a supplementary step is needed to estimate the number of paths. The scaling ambiguities can be easily eliminated by exploiting the Vandermonde structure of the steering matrices (4.18) whose first row is composed of ones.

4.3 Chapter summary

In this chapter, a DD-DP MIMO wireless system has been proposed with the TST coding combined with the particular case of the MKronST coding, so-called TST-MSMKron coding. This coding provides an increase in space-time diversity by multiple Kronecker products of symbol matrices combined with the TST coding. We have shown that the combination of TST-MSMKron coding with the generalized channel modeled as a 5-order PARAFAC decomposition to transmit the symbols provides a signal received modeled by the $(N+2)$-order coupled nested Tucker-PARAFAC model, where the core tensor is the TST coding. The coupling is due to the coding that is common to the VR- and HR- polarized receive antennas. The nested structure results from the fact that a matrix unfolding of the channel tensor constitutes one matrix factor of the Tucker model, the other factors being the steering and path-loss matrices. The essential uniqueness of the tensor model is insured under mild conditions.

The contributions in this chapter extend previous works in different ways, either by using a particular case of MKronST coding or by extending the channel tensor combined with the new coding schemes for building tensors of received signals. By exploiting the new received signal model, semi-blind receivers are derived in Chapter 5 to estimate the symbol matrices, channel and channel parameters.

5 Tensor-based semi-blind receivers for Dual-polarized MIMO communication systems using combined TST-MSMKron coding

In this chapter, by exploiting the received signals tensor presented in Chapter 4 and with the tensor coding knowledge, we propose two-semi blind receivers composed of two steps to joint estimate symbol matrices, channel and channel parameters (DoA and DoD angles, path-loss coefficients). These receivers are composed of two stages. In the first one, the information symbols and the channel are jointly estimated, using either a Bi-ALS algorithm, followed by the KronF algorithm to separate the symbol matrices, or the THOSVD method, which allows to directly and jointly estimate the symbol matrices and the channel in closed-form. In the second stage, from the estimated channel, the channel parameters estimation is performed using the JIRAFE algorithm including a new rectification strategy for taking the Vandermonde structure into account. Parameter identifiability and computational complexity are analyzed for each receiver.

Monte Carlo simulation results are provided to illustrate the performance of the TST-MSMKron coding and the impact of design parameters on the system performance and the behavior of the proposed receivers in terms of SER, channel NMSE and angle MSE.

5.1 Bibliographic overview of receivers for MIMO systems

In the last years, the proposition of receivers to accurately estimate channels on massive MIMO systems has increased. The CSI estimation needs to be accurate as it may affect the performance of transmit beamforming at the transmitters and decoding accuracy at the receiver [110]. For this, it is necessary to propose receivers that allow estimating the channel and its parameters with better precision. Receivers can be divided into two classes: supervised (or pilot-assisted) and unsupervised (or semi-blind). In the first case, training sequences known by the receiver are used to estimate the channel, which generally implies a significant training overhead due to pilots transmission [32].

Table 16 summarizes the wireless communication systems presented in Chapter 4, in Tables 12-13, with their MIMO system type, and receivers. Sidiropoulos *et al.* (2000) [47] proposed a blind receiver by combining compression methods and the ALS algorithm to estimate the symbols and the channel. A blind CSI recovery receiver was proposed based on the ALS algorithm by Sidiropoulos & Budampati (2002) [103]. Favier *et al.* (2012) [8] proposed a blind receiver based on the ALS algorithm. Almeida & Favier (2013) [7] proposed a recursive semi-blind receiver to estimate the symbols and channel. This receiver considers the ALS algorithm in two steps. Da Costa (2014) [104] proposes three semi-blind receivers based on the ALS algorithm and Kronecker product to estimate the parameters of a multicarrier MIMO wireless system with TST/STF coding. Favier & Almeida (2014) [9] proposed three receivers based on the ALS and Kronecker product least square (KPLS) methods. Of these receivers, two are semi-blind and one is supervised. Da Costa *et al.* (2018) [58] proposed five receivers: three semi-blind and two supervised. The receivers are based on the ALS and SVD methods. Two semi-supervised receivers in two stages based on the KRF and ALS algorithms were proposed by De Pinho *et al.* (2019) [11] to estimate the symbol matrices and the channel. Randriambelonoro *et al.* (2021) [12] proposed closed-form semi-blind receivers based on ALS and KRF methods. Note that the majority of the proposed receivers are blind or semi-blind, where one matrix, row, or one element of the symbols matrices needs to be known to estimate the symbols and channel without channel knowledge. The algorithms considered are based on classical methods such as ALS and SVD.

In terms of mmWave MIMO systems, we can cite some works. Zhou *et al.* (2017) [23] proposed a receiver based on the ALS algorithm. The channel parameters, such as the DoAs and DoDs, time delays and fading coefficients, were extracted based on the estimated factor matrices. Du *et al.* (2021) [105] proposed semi-blind receivers based on the ALS and SVD methods to estimate the compound channel matrix, detect the information symbols and extract the channel parameters considering the sparse scattering nature of mmWave channels. Chang *et al.* (2021) [107] presented an accelerated trilinear alternating least squares (ATALS) to jointly estimate the multiuser channel parameters via a one-dimensional search. A tensor-based RIS channel estimation method (TenRICE) was proposed by Gherekhloo *et al.* (2021) [108] to estimate the matrices using an ALS algorithm. MmWave channels combined with coding and tensor techniques can improve parameter estimation techniques.

The researches presented above often consider pilot-assisted symbols because they are interesting only in parameter channel estimation, not considering the complete system model to be estimated, i.e., the symbols, channel and channel parameters, being a hard task to solve. An important difference between the systems presented in Table 16 and our proposed system is related to the a priori information needed to eliminate scaling ambiguities. The systems presented require knowledge of one element or one row of the channel matrix, while our system only requires a priori knowledge of one symbol of each symbol matrix, where the receivers are semi-blind.

In this chapter, we consider the DD-DP MIMO system introduced in [109] and presented in Chapter 4 to derive two-semi-blind receivers for estimating the transmitted symbols

Reference	**MIMO system**	**Receivers**
Sidiropoulos *et al.* (2000) [47]	DS-CDMA	Based on ALS
Sidiropoulos & Budampati (2002) [103]	multi-antenna	ALS
Favier *et al.* (2012) [8]	point-to-point	ALS
Almeida & Favier (2013) [7]	OFDM	ALS
Da Costa (2014) [104]	multicarrier	ALS, ALM and KLS
Favier & Almeida (2014) [9]	OFDM-CDMA	ALS and KPLS
Da Costa *et al.* (2018) [58]	OFDM-CDMA	KALS, SKALS, KPLS, KSVD and KALMS
De Pinho *et al.* (2019) [11]	point-to-point	KRF-KRF, ALS-KRF and TALS
Randriambelonoro *et al.* (2021) [12]	multipath	MKRF/LS/ALS and MKRF/ALS
Zhou *et al.* (2017) [23]	OFDM-MIMO massive mmWave	ALS
Du *et al.* (2021) [105]	mm-Wave	ITS-ALS, TALS and SVD-ALS
Chang *et al.* (2021) [107]	OFDM mmWave	ATALS
Gherekhloo *et al.* (2021) [108]	point-to-point RIS-aided mmWave	TenRICE

Table 16 – Tensor approaches for wireless communication systems.

and channel parameters (DoD and DoA angles, path-loss coefficients). These receivers are composed of two stages. In the first one, the information symbols and the channel are jointly estimated, using either a Bi-ALS algorithm, followed by the KronF algorithm to separate the symbol matrices, or the THOSVD method, which allows directly and jointly estimate the symbol matrices and the channel in closed-form. In the second stage, from the estimated channel, the channel parameters estimation is performed using the JIRAFE algorithm. Parameter identifiability and computational complexity are analyzed for each receiver. In the next sections, the receivers composed of two stages, computational complexity analysis and results will be presented.

5.2 Semi-blind receivers for DP MIMO systems using TST-MSMKron codings

By exploiting the matrix unfoldings (4.42)-(4.44) of the received signals tensor, two semi-blind receivers are now developed for estimating the symbol matrices, the channel and channel parameters (DoA and DoD angles, path-losses). These receivers are composed of two stages. In the first one, the symbol matrices and the channel are estimated, while in the second stage, the channel parameters are derived from the estimated channel [109].

5.2.1 Stage 1 - channel and symbol estimation

The Bi-ALS algorithm is used to jointly estimate the MSMKron product **S** and the channel matrix unfolding **H**. Then, the KronF algorithm is applied to separate the symbol matrices. The Bi-ALS algorithm results from the following optimization problem deduced from Eq. (4.41) [109]:

$$\min_{\mathbf{S},\mathbf{H}} \parallel \mathcal{X}_c - \mathcal{G}_c \times_1 \mathbf{H} \times_2 \mathbf{S} \times_3 \mathbf{I}_P \parallel_F^2, \tag{5.1}$$

where $\| \,.\, \|_F$ is the Frobenius norm. Fixing one of the matrix factors ($\mathbf{H}$, $\mathbf{S}$) with its value estimated at previous iteration, the Bi-ALS method replaces the optimization problem (5.1) by two LS sub-problems deduced from the matrix unfoldings (4.42) and (4.43), leading to the alternate minimization of the following LS criteria:

$$\min_{\mathbf{H}} \parallel \mathbf{X}_{PJ\times 2M_R} - \left(\mathbf{I}_P \otimes \hat{\mathbf{S}}_{[it-1]}\right) \mathbf{G}_{PR\times 2M_T} \mathbf{H}^T \parallel_F^2 \longrightarrow \hat{\mathbf{H}}^T_{[it]}, \tag{5.2}$$

$$\min_{\mathbf{S}} \parallel \mathbf{X}_{2PM_R\times J} - \left(\mathbf{I}_P \otimes \hat{\mathbf{H}}_{[it]}\right) \mathbf{G}_{2PM_T\times R} \mathbf{S}^T \parallel_F^2 \longrightarrow \hat{\mathbf{S}}^T_{[it]}. \tag{5.3}$$

The update equations at iteration $[it]$ are given by:

$$\hat{\mathbf{H}}^T_{[it]} = \left[\left(\mathbf{I}_P \otimes \hat{\mathbf{S}}_{[it-1]}\right) \mathbf{G}_{PR\times 2M_T}\right]^{\dagger} \mathbf{X}_{PJ\times 2M_R} \in \mathbb{C}^{2M_T\times 2M_R}, \tag{5.4}$$

$$\hat{\mathbf{S}}^T_{[it]} = \left[\left(\mathbf{I}_P \otimes \hat{\mathbf{H}}_{[it]}\right) \mathbf{G}_{2PM_T\times R}\right]^{\dagger} \mathbf{X}_{2PM_R\times J} \in \mathbb{C}^{R\times J}. \tag{5.5}$$

The matrices $\left[\left(\mathbf{I}_P \otimes \hat{\mathbf{S}}_{[it-1]}\right) \mathbf{G}_{PR\times 2M_T}\right]$ and $\left[\left(\mathbf{I}_P \otimes \hat{\mathbf{H}}_{[it]}\right) \mathbf{G}_{2PM_T\times R}\right]$ must have full column rank for ensuring uniqueness of the LS estimates, which implies the following necessary conditions: $2M_T \leq PJ$ and $R \leq 2PM_R$. To simplify the computation of the estimate $\hat{\mathbf{H}}$ in Eq. (5.4), we assume that the matrices $\mathbf{G}_{PR\times 2M_T}$ and $\mathbf{S}$ have full column rank which implies: $2M_T \leq PR$ and $R \leq J$, respectively. Moreover, to simplify the computation of $\hat{\mathbf{S}}$ in Eq. (5.5), we assume that the unfolding $\mathbf{G}_{2PM_T\times R}$ is chosen as a full column rank truncated discrete Fourier transform (DFT) matrix, which allows us to replace its pseudo-inverse by its transconjugate, implying the necessary condition $R \leq 2PM_T$. We also assume that $\mathbf{H}$ has full column rank, implying $M_T \leq M_R$. Exploiting these assumptions simplify the LS estimates (5.4)-(5.5) as:

$$\hat{\mathbf{H}}^T_{[it]} = \left(\mathbf{G}_{PR\times 2M_T}\right)^{\dagger} \left(\mathbf{I}_P \otimes \hat{\mathbf{S}}^{\dagger}_{[it-1]}\right) \mathbf{X}_{PJ\times 2M_R}, \tag{5.6}$$

$$\hat{\mathbf{S}}^T_{[it]} = \left(\mathbf{G}_{2PM_T\times R}\right)^{H} \left(\mathbf{I}_P \otimes \hat{\mathbf{H}}^{\dagger}_{[it]}\right) \mathbf{X}_{2PM_R\times J}. \tag{5.7}$$

The Bi-ALS algorithm (5.6)-(5.7) is a simplified version of (5.4)-(5.5) in terms of pseudo-inverses computation. However, this simplification is at the price of additional constraints on the design parameters.

For deciding the convergence of the Bi-ALS algorithm, we consider the error at the $[it]$-th iteration deduced from (4.42) as:

$$err_{[it]} = \| \mathbf{X}_{PJ\times 2M_R} - \left(\mathbf{I}_P \otimes \hat{\mathbf{S}}_{[it]}\right) \mathbf{G}_{PR\times 2M_T} \hat{\mathbf{H}}_{[it]}^{T} \|_F^2 . \tag{5.8}$$

Convergence at the $[it]$-th iteration is declared when this error does not significantly change between two successive iterations, i.e., $|err_{[it-1]} - err_{[it]}| \leq \epsilon$, where ϵ is a predefined threshold.

As already mentioned, the Tucker model (4.41) is unique up to scalar scaling ambiguities defined by Eqs. (4.45). Noting $\hat{\mathbf{H}}$ and $\hat{\mathbf{S}}$ the LS estimates (5.4) and (5.5), or (5.6) and (5.7), at convergence, the estimates after correcting the ambiguities are given by:

$$\hat{\hat{\mathbf{S}}} = \hat{\mathbf{S}}\left(\lambda^{(S)}\right)^{-1}, \quad \hat{\hat{\mathbf{H}}} = \hat{\mathbf{H}}\lambda^{(S)}. \tag{5.9}$$

For eliminating these scaling ambiguities, it is sufficient to assume that one element of $\mathbf{S}$ is known a priori, e.g., $s_{11} = 1$. Under this assumption, $\lambda^{(S)}$ is calculated as: $\lambda^{(S)} = \hat{s}_{11}$. The symbol matrices $\mathbf{S}^{(n)}$ are then estimated using the KronF algorithm which minimizes the following LS cost function:

$$\min_{\mathbf{S}^{(n)},\, n\in\{1,\dots,N\}} \| \hat{\hat{\mathbf{S}}} - \mathbf{S}^{(1)} \otimes \dots \otimes \mathbf{S}^{(N)} \|_F^2 . \tag{5.10}$$

This problem is solved by reshaping $\hat{\hat{\mathbf{S}}}$ as a rank-one tensor [60, 61, 54]:

$$\hat{\hat{\mathcal{S}}} = \mathbf{s}^{(1)} \circ \dots \circ \mathbf{s}^{(N)} \in \mathbb{C}^{R_1J_1\times\dots\times R_NJ_N}, \tag{5.11}$$

where $\mathbf{s}^{(n)}$ =vec($\mathbf{S}^{(n)}$) $\in \mathbb{C}^{R_nJ_n}$, and replacing the cost function (5.10) by:

$$\min_{\mathbf{s}^{(n)},\, n\in[1,N]} \| \hat{\hat{\mathcal{S}}} - \mathbf{s}^{(1)} \circ \cdots \circ \mathbf{s}^{(N)} \|_F^2 . \tag{5.12}$$

This rewriting of the cost function as a rank-one approximation of an N-order tensor allows to estimate each vectorized form $\mathbf{s}^{(n)}$ by calculating the SVD of the mode-n unfolding of the tensor $\hat{\hat{\mathcal{S}}}$:

$$\hat{\hat{\mathbf{S}}}_{R_nJ_n\times R_{n+1}J_{n+1}\dots R_NJ_NR_1J_1\dots R_{n-1}J_{n-1}} = \mathbf{U}^{(n)}\boldsymbol{\Sigma}^{(n)}\mathbf{V}^{(n)H}, \tag{5.13}$$

with $\hat{\mathbf{s}}^{(n)}$ chosen as the first left singular vector associated with the dominant singular value of this mode-n unfolding, i.e.:

$$\hat{\mathbf{s}}^{(n)} = \mathbf{U}^{(n)}_{.1} \in \mathbb{C}^{R_nJ_n}. \tag{5.14}$$

The estimated symbol matrix is then obtained by unvectorizing $\hat{\mathbf{s}}^{(n)}$ as:

$$\hat{\mathbf{S}}^{(n)} = \text{unvec}\left(\hat{\mathbf{s}}^{(n)}\right) \in \mathbb{C}^{J_n\times R_n}. \tag{5.15}$$

```
Bi-ALS-KronF receiver for estimating the symbol matrices $\mathbf{S}^{(n)}$ and the channel $\mathbf{H}$
Input: tensors $\mathcal{X}$, $\mathcal{G}$
Output: Estimated symbol matrices $\hat{\hat{\mathbf{S}}}^{(n)}$ and estimated channel $\hat{\hat{\mathbf{H}}}$.
Step 1 - BiALS algorithm
$it = 0$
1) Randomly initialize $\mathbf{S}^{(n)}_{[0]}$ with symbols drawn from the alphabet and $s^{(n)}_{11} =$ 1.
2) Update the estimates of $\mathbf{S}$ and $\mathbf{H}$ using (5.4)-(5.5) or (5.6)-(5.7).
3) Calculate the error (5.8) and $|err_{[it-1]} - err_{[it]}|$.
-if $|err_{[it-1]} - err_{[it]}| \leq \epsilon$ or $it =$ maximum number of iterations
- stop
-else $it \rightarrow it + 1$;
4) Eliminate the scaling ambiguities.

Step 2 - KronF algorithm
5) Build the rank-one tensor as: $\hat{\hat{\mathcal{S}}} = \text{reshape}\left(\hat{\hat{\mathbf{S}}}, [R_1 J_1, \ldots, R_N J_N]\right)$.
6) Estimate each vector $\hat{\mathbf{s}}^{(n)}$ as the first left singular vector using (5.14), and unvectorize it using (5.15).
7) Eliminate the scaling ambiguities using (5.16).
8) Project the estimated symbols onto the finite alphabet.
```

Table 17 – Bi-ALS-KronF receiver.

Assuming $s^{(n)}_{11} = 1$, the scalar ambiguity is corrected by:

$$\hat{\hat{\mathbf{S}}}^{(n)} = \hat{\mathbf{S}}^{(n)} \left(\lambda^{(n)}\right)^{-1}, \text{with } \lambda^{(n)} = \hat{s}^{(n)}_{11}. \tag{5.16}$$

The final estimated symbols are obtained after a projection onto the finite alphabet. The Bi-ALS-KronF algorithm is summarized in Table 17.

5.2.1.1 THOSVD receiver

A second semi-blind receiver is proposed to directly and jointly estimate the channel and symbol matrices, using a closed-form solution based on the THOSVD algorithm [35]. From the matrix unfolding (4.44), with $\mathbf{S}$ replaced by its expression (4.31), the LS estimate of the multiple Kronecker product is given by [109]:

$$\mathbf{Y} \triangleq \mathbf{H} \otimes \widehat{\mathbf{S}^{(1)} \otimes} \ldots \otimes \mathbf{S}^{(N)} = \left[(\mathbf{G}_{P \times 2M_T R})^{\dagger} \mathbf{X}_{P \times 2M_R J}\right]^T \in \mathbb{C}^{2M_R J \times 2M_T R}. \tag{5.17}$$

The unfolding $\mathbf{G}_{P \times 2M_T R}$ must have full column rank for ensuring uniqueness of this LS estimate, which induces the necessary condition: $2M_T R \leq P$. The symbol matrices $\mathbf{S}^{(n)}$ and channel $\mathbf{H}$ are then jointly estimated by means of the KronF algorithm, as in step 2 of the Bi-ALS-KronF receiver, with the difference that we can now simultaneously estimate all the matrices $(\mathbf{H}, \mathbf{S}^{(1)}, \cdots, \mathbf{S}^{(N)})$. The THOSVD receiver is summarized in Table 18.

THOSVD receiver for estimating the symbol matrices $\mathbf{S}^{(n)}$ and the channel $\mathbf{H}$.

Input: tensors $\mathcal{X}$, $\mathcal{G}$

Output: Estimated symbol matrices $\hat{\hat{\mathbf{S}}}^{(n)}$ and estimated channel $\hat{\hat{\mathbf{H}}}$.

1) Calculate the LS estimate $\mathbf{Y}$ defined in (5.17).

2) Build the rank-one tensor $\mathcal{Y}$ of size $4M_TM_R \times R_1J_1 \times ... \times R_NJ_N$ from $\mathbf{Y}$.

3) Compute the SVD of each mode-n unfolding of $\mathcal{Y}$, and calculate the estimates $\hat{\mathbf{s}}^{(n)} = \text{vec}(\hat{\mathbf{S}}^{(n)})$ and $\hat{\mathbf{h}} = \text{vec}(\hat{\mathbf{H}})$ as the first left singular vector of each mode-n unfolding.

4) Unvectorize $\hat{\mathbf{s}}^{(n)}$ and $\hat{\mathbf{h}}$ to obtain the estimates $\hat{\mathbf{S}}^{(n)}$ and $\hat{\mathbf{H}}$.

5) Eliminate the scaling ambiguities.

6) Project the estimated symbols onto the finite alphabet.

Table 18 – THOSVD receiver.

5.2.1.2 Zero-forcing (ZF)-KronF receiver

To evaluate the impact of design parameters on the system performance, we use the zero-forcing (ZF)-KronF receiver which assumes a perfect channel knowledge. The LS estimate $\hat{\mathbf{S}}_{ZF}$ of $\mathbf{S}$ is then calculated using (5.5) or (5.7), with $\mathbf{H}_{[it]}$ replaced by the true channel $\mathbf{H}$, which gives:

$$\hat{\mathbf{S}}_{ZF}^{T} = \left[(\mathbf{I}_P \otimes \mathbf{H})\,\mathbf{G}_{2PM_T\times R}\right]^{\dagger} \mathbf{X}_{2PM_R\times J}, \tag{5.18}$$

or

$$\hat{\mathbf{S}}_{ZF}^{T} = \left(\mathbf{G}_{2PM_T\times R}\right)^{H} \left(\mathbf{I}_P \otimes \mathbf{H}^{\dagger}\right) \mathbf{X}_{2PM_R\times J}. \tag{5.19}$$

As for the Bi-ALS algorithm, use of (5.5) or (5.7) implies the following necessary conditions: $R \leq 2PM_R$ or $R \leq 2PM_T$ and $M_T \leq M_R$, for the ZF-KronF receiver (5.18) and its simplified version (5.19), respectively. Then, the symbol matrices $\mathbf{S}^{(n)}$ are estimated using the KronF algorithm as in the second step of the Bi-ALS-KronF receiver.

In Table 19, the identifiability conditions for the proposed receivers are summarized. By comparing the conditions for the Bi-ALS algorithm (5.6)-(5.7) with the ones for the Bi-ALS algorithm (5.4)-(5.5), we can deduce some implications. Indeed, for the estimate (5.6), the conditions $2M_T \leq PR$ and $R \leq J$ imply $2M_T \leq PJ$, i.e., the identifiability condition for the LS solution (5.4). For the estimate (5.7), the conditions $R \leq 2PM_T$ and $M_T \leq M_R$ imply $R \leq 2PM_R$, i.e., the identifiability condition for the LS solution (5.5). In conclusion, if the identifiability conditions for (5.6)-(5.7) are satisfied, then the ones for the Bi-ALS algorithm (5.4)-(5.5) are automatically satisfied. Note also that $R \leq 2PM_T$ and $2M_T \leq PR$ imply $R \leq P^2R$, which is always satisfied. Therefore, the condition $2M_T \leq PR$ can be discarded.

We can also conclude that the THOSVD receiver is more restrictive than the Bi-ALS receivers in the sense that a higher value of P is required, which implies a reduction in the

Receiver	**Identifiability conditions**
Bi-ALS-KronF Eqs. (5.4)-(5.5)	$2M_T \leq PJ$; $R \leq 2PM_R$
Bi-ALS-KronF Eqs. (5.6)-(5.7)	$R \leq \min(2PM_T, J)$; $2M_T \leq PR$; $M_T \leq M_R$
THOSVD	$2M_TR \leq P$;
ZF-KronF	$R \leq 2PM_R$ or $R \leq 2PM_T$, $M_T \leq M_R$

Table 19 – Identifiability conditions for the receivers.

transmission rate. Finally, as the ZF-KronF receiver only estimates the symbol matrices using the LS solutions (5.18) or (5.19), identifiability conditions are subsets of the ones for the Bi-ALS-KronF receivers.

5.2.2 Stage 2 - channel parameters estimation using JIRAFE

In this section, the JIRAFE algorithm is presented to estimate the steering vectors and path-losses from the estimated channel matrix $\hat{\hat{\mathbf{H}}}$, reshaped as the tensor $\hat{\hat{\mathcal{H}}}$ satisfying the PARAFAC decomposition (4.30):

$$\hat{\hat{\mathcal{H}}} = \text{reshape}\left(\hat{\hat{\mathbf{H}}}, \left[M_{T_x}, M_{T_y}, M_{R_x}, M_{R_y}, 4\right]\right) \in \mathbb{C}^{M_{T_x} \times M_{T_y} \times M_{R_x} \times M_{R_y} \times 4}. \tag{5.20}$$

Exploiting the equivalence between TT and PARAFAC decompositions recalled in section 2.5.6, the fifth-order estimated channel tensor (5.20) can be rewritten as the following TT decomposition deduced from (2.75), for $N = 5$:

$$\hat{\hat{\mathcal{H}}} = \mathbf{G}^{(1)} \times_2^1 \mathcal{G}^{(2)} \times_3^1 \mathcal{G}^{(3)} \times_4^1 \mathcal{G}^{(4)} \times_5^1 G^{(5)}, \tag{5.21}$$

where $\mathbf{G}^{(1)} \in \mathbb{C}^{M_{T_x} \times K}$, $\mathcal{G}^{(2)} \in \mathbb{C}^{K \times M_{T_y} \times K}$, $\mathcal{G}^{(3)} \in \mathbb{C}^{K \times M_{R_x} \times K}$, $\mathcal{G}^{(4)} \in \mathbb{C}^{K \times M_{R_y} \times K}$ and $\mathbf{G}^{(5)} \in \mathbb{C}^{K \times 4}$ are the TT-cores, linked with the PARAFAC factors by means of the equivalence relations (2.142)-(2.143) which become [34, 33]:

$$\mathbf{G}^{(1)} = \mathbf{A}_{T_x}\mathbf{M}_1^{-1}, \tag{5.22}$$

$$\mathcal{G}^{(2)} = \mathcal{I}_{3,K} \times_1 \mathbf{M}_1 \times_2 \mathbf{A}_{T_y} \times_3 \mathbf{M}_2^{-T}, \tag{5.23}$$

$$\mathcal{G}^{(3)} = \mathcal{I}_{3,K} \times_1 \mathbf{M}_2 \times_2 \mathbf{A}_{R_x} \times_3 \mathbf{M}_3^{-T}, \tag{5.24}$$

$$\mathcal{G}^{(4)} = \mathcal{I}_{3,K} \times_1 \mathbf{M}_3 \times_2 \mathbf{A}_{R_y} \times_3 \mathbf{M}_4^{-T}, \tag{5.25}$$

$$\mathbf{G}^{(5)} = \mathbf{M}_4\mathbf{B}^T, \tag{5.26}$$

where $\mathbf{M}_n \in \mathbb{C}^{K \times K}$, for $n \in [1, 4]$, are nonsingular basis-change matrices. The TTHSVD algorithm, described in section 2.5.5, is used to determine the TT core estimates $\hat{\mathbf{G}}^{(k)}$, for $k \in [1, 5]$, and $\hat{\mathcal{G}}^{(k)}$, for $k \in [2, 3, 4]$.

A new rectification strategy, denoted $\text{TR}_1\text{A}_{imp}$ and presented in Table 20, is proposed to enforce the Vandermonde structure of the steering matrices estimated in presence of

```
TR$_1$A$_{imp}$ method to rectify an estimated Vandermonde matrix $\hat{\mathbf{A}} \in \mathbb{C}^{I\times K}$

Input: Estimated matrix $\hat{\mathbf{A}}$
Output: Rectified Vandermonde matrix $\hat{\hat{\mathbf{A}}}$.
for $k \in [1, K]$
    1) Form a rank-one hermitian Toeplitz matrix $\mathbf{T}_k$ from the $k$-th column $\hat{\mathbf{A}}_{.k}$
       as: $\mathbf{T}_k = \hat{\mathbf{A}}_{.k}\hat{\mathbf{A}}^H_{.k}$.
    2) Calculate the EVD of $\mathbf{T}_k = \mathbf{P}_k\mathbf{D}_k\mathbf{P}^H_k$.
    for $i \in [1, I-1]$
        3) Estimate the angular frequency $\omega_k$ associated with the $k$-th column of he
           Vandermonde matrix as:
           $$\hat{\omega}_k = \frac{1}{I-1}\left(\sum_{i=1}^{I-1}\frac{1}{i}\angle\Big((p_k)_{1,1}(p_k)^*_{i+1,1}\Big)\right).$$
        4) Build the rectified estimated $k$-th column $\hat{\hat{\mathbf{A}}}_{.k}$ as:
           $$\hat{\hat{\mathbf{A}}}_{.k} = [\ \hat{a}_{1,k} \quad \hat{a}_{2,k} \quad \cdots \quad \hat{a}_{I_1,k}\ ]^T \ \text{with}\ \hat{a}_{i,k} = e^{-j\hat{\omega}_k(i-1)}.$$
    end
end
```

Table 20 – TR$_1$A$_{imp}$ rectification method.

noise, whose elements of the first row are normalized to 1, as in (4.18). This strategy is an improved version of the Toeplitz rank-one approximation method, proposed in [65] and denoted TR$_1$A.

Our proposed rectification method consists in constructing a rank-one hermitian Toeplitz matrix $\mathbf{T}_k = \hat{\mathbf{A}}_{.k}\hat{\mathbf{A}}^H_{.k}$, of size $I \times I$, from each column k of an estimated $I \times K$ Vandermonde matrix, and computing its eigenvalue decomposition (EVD) $\mathbf{T}_k = \mathbf{P}_k\mathbf{D}_k\mathbf{P}^H_k$ to estimate the k-th angular frequency.

Unlike the original TR$_1$A method which identifies only one element of the Toeplitz matrix with its corresponding term in the EVD of the constructed rank-one Toeplitz matrix, our method exploits the identification of $I-1$ terms as:

$$e^{-j\omega_k i} = d_k(p_k)_{1,1}(p_k)^*_{i+1,1} \;\Rightarrow\; \hat{\omega}_k = \frac{1}{i}\angle\Big((p_k)_{1,1}(p_k)^*_{i+1,1}\Big), \tag{5.27}$$

for $i \in [1, I-1]$, where $(p_k)_{i,1}$ denotes the $(i,1)$ element of the eigenvector matrix $\mathbf{P}_k$, and the operator $\angle$ stands for the angle of its complex scalar argument. Then, the estimated angular frequency is calculated as an average of these $I-1$ estimates (A.48) as:

$$\hat{\omega}_k = \frac{1}{I-1}\sum_{i=1}^{I-1}\frac{1}{i}\angle\Big((p_k)_{1,1}(p_k)^*_{i+1,1}\Big). \tag{5.28}$$

JIRAFE algorithm for the estimation of $\hat{\mathbf{A}}_{T_x}$, $\hat{\mathbf{A}}_{T_y}$, $\hat{\mathbf{A}}_{R_x}$, $\hat{\mathbf{A}}_{R_y}$ and $\hat{\mathbf{B}}$.

Input: Estimated channel $\hat{\hat{\mathbf{H}}}$
Output: Estimated matrices $\hat{\mathbf{A}}_{T_x}$, $\hat{\mathbf{A}}_{T_y}$, $\hat{\mathbf{A}}_{R_x}$, $\hat{\mathbf{A}}_{R_y}$ and $\hat{\mathbf{B}}$.

1) Reshape the estimated channel matrix $\hat{\hat{\mathbf{H}}}$ as a tensor $\hat{\hat{\mathcal{H}}}$,

$$\hat{\hat{\mathcal{H}}} = \text{reshape}\left(\hat{\hat{\mathbf{H}}}, \left[M_{T_x}, M_{T_y}, M_{R_x}, M_{R_y}, 4\right]\right).$$

2) Estimate the TT cores by means of the TTHSVD algorithm applied to $\hat{\hat{\mathcal{H}}}$:

$$[\hat{\mathbf{G}}^{(1)}, \hat{\mathcal{G}}^{(2)}, \hat{\mathcal{G}}^{(3)}, \hat{\mathcal{G}}^{(4)}, \hat{\mathbf{G}}^{(5)}] \leftarrow \text{TTHSVD}\ (\hat{\hat{\mathcal{H}}}, K).$$

3) Estimate the factor matrices by means of the RectALS and RectKRF algorithms:

$$\left[\hat{\mathbf{M}}_1, \hat{\mathbf{A}}_{T_y}, \hat{\mathbf{M}}_2^{-T}\right] \leftarrow \text{RectALS}\left(\hat{\mathcal{G}}^{(2)}, K\right),$$

$$\left[\hat{\mathbf{A}}_{R_x}, \hat{\mathbf{M}}_3^{-T}\right] \leftarrow \text{RectKRF}\left(\hat{\mathcal{G}}^{(3)}, \hat{\mathbf{M}}_2, K\right),$$

$$\left[\hat{\mathbf{A}}_{R_y}, \hat{\mathbf{M}}_4^{-T}\right] \leftarrow \text{RectKRF}\left(\hat{\mathcal{G}}^{(4)}, \hat{\mathbf{M}}_3, K\right),$$

$$\hat{\mathbf{A}}_{T_x} = \hat{\mathbf{G}}^{(1)}\hat{\mathbf{M}}_1, \quad \hat{\mathbf{B}} = \hat{\mathbf{G}}^{(5)T}\hat{\mathbf{M}}_4^{-T}.$$

4) Rectify $\hat{\mathbf{A}}_{T_x}$ using TR$_1$A$_{imp}$ method in Table 20.

Table 21 – JIRAFE algorithm.

This strategy is applied with the ALS algorithm for estimating $\mathbf{A}_{T_y}$ from the estimated core $\hat{\mathcal{G}}^{(2)}$, and with the KRF algorithm to estimate $\mathbf{A}_{R_x}$ and $\mathbf{A}_{R_y}$, from the estimated cores $\hat{\mathcal{G}}^{(3)}$ and $\hat{\mathcal{G}}^{(4)}$, respectively, which leads to the RectALS and RectKRF algorithms. Finally, LS estimates of $\mathbf{A}_{T_x}$ and $\mathbf{B}$ are deduced from Eqs. (5.22) and (5.26), respectively, with a rectification of $\hat{\mathbf{A}}_{T_x}$. The JIRAFE algorithm is summarized in Table 21.

The azimuth and elevation angles ϕ_k, θ_k, ϑ_k and ψ_k are then obtained by means of the following equations [25]:

$$\hat{\phi}_k = \sin^{-1}\left(\sqrt{\left(\frac{\nu}{2\pi d_{T_x}}\hat{\omega}_{T_x,k}\right)^2 + \left(\frac{\nu}{2\pi d_{T_y}}\hat{\omega}_{T_y,k}\right)^2}\right), \quad \hat{\theta}_k = \tan^{-1}\left(\frac{d_{T_x}\hat{\omega}_{T_y,k}}{d_{T_y}\hat{\omega}_{T_x,k}}\right), \tag{5.29}$$

$$\hat{\vartheta}_k = \sin^{-1}\left(\sqrt{\left(\frac{\nu}{2\pi d_{R_x}}\hat{\omega}_{R_x,k}\right)^2 + \left(\frac{\nu}{2\pi d_{R_y}}\hat{\omega}_{R_y,k}\right)^2}\right), \quad \hat{\psi}_k = \tan^{-1}\left(\frac{d_{R_x}\hat{\omega}_{R_y,k}}{d_{R_y}\hat{\omega}_{R_x,k}}\right). \tag{5.30}$$

The Bi-ALS/KronF-JIRAFE and THOSVD-JIRAFE receivers, composed of two stages, are summarized in Table 22, and illustrated by means of Figure 17.

Bi-ALS-KronF-JIRAFE and THOSVD-JIRAFE receivers

Input: Tensors $\mathcal{X}$, $\mathcal{G}$

Output: $\hat{\hat{\mathcal{H}}}$, $\hat{\hat{\mathbf{S}}}^{(n)}$, for $n \in [1, N]$, $\hat{\mathbf{A}}_{T_x}$, $\hat{\mathbf{A}}_{T_y}$, $\hat{\mathbf{A}}_{R_x}$, $\hat{\mathbf{A}}_{R_y}$ and $\hat{\mathbf{B}}$.

Stage 1 - Estimation of the channel and symbol matrices

$[\hat{\hat{\mathbf{H}}}, \hat{\hat{\mathbf{S}}}^{(n)}] \leftarrow$ Bi-ALS-KronF or THOSVD, using the Algorithms in Tables 17 and 18.

Stage 2 - Estimation of the channel parameters

$[\hat{\mathbf{A}}_{T_x}, \hat{\mathbf{A}}_{T_y}, \hat{\mathbf{A}}_{R_x}, \hat{\mathbf{A}}_{R_y}, \hat{\mathbf{B}}] \leftarrow$ JIRAFE $\left(\hat{\hat{\mathcal{H}}}, K\right)$, usisng the Algorithm in Table 21.

- Angle estimation

$\left[\hat{\phi}_k, \hat{\theta}_k, \hat{\vartheta}_k, \hat{\psi}_k\right] \leftarrow$ Eqs. (5.29)-(5.30).

Table 22 – Bi-ALS/KronF-JIRAFE and THOSVD-JIRAFE receivers.

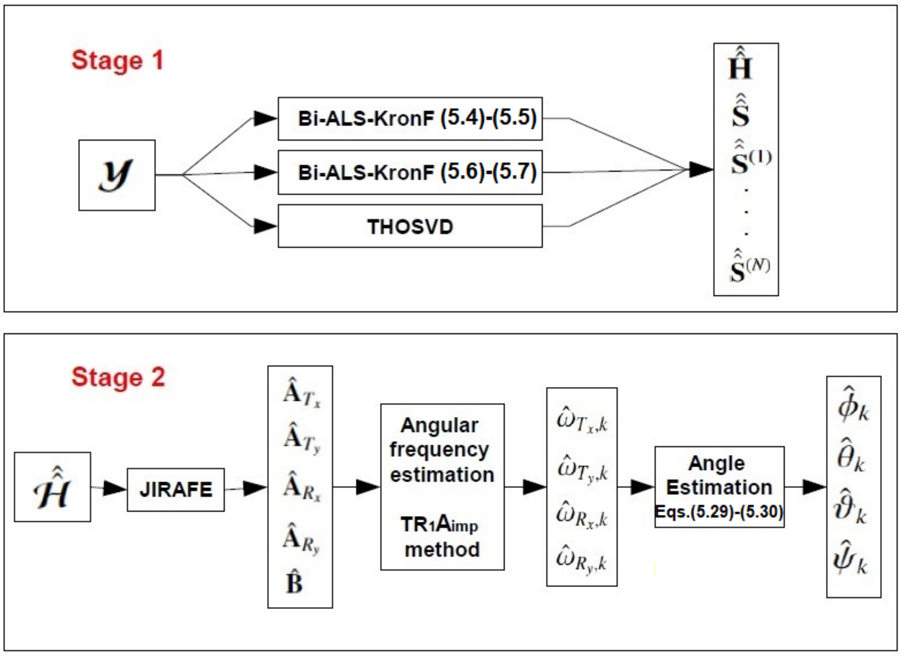


Figure 17 – Block-diagram of the proposed receivers.

5.3 Complexity analysis

In this section, we compare the computational complexity of the proposed semi-blind receivers, considering that the most expensive matrix operations are the SVD and EVD whose complexities are $O(IJ \min(I,J))$ for an $I \times J$ matrix and $O(I^3)$ for an $I \times I$ matrix, respectively. The complexities are evaluated taking the identifiability conditions into account.

Computing the HOSVD of an N-th-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ requires N SVDs of $I_n \times I_{n+1} ... I_N I_1 ... I_{n-1}$ matrices, for $n \in [1,N]$, inducing the following overall computational complexity $O\left((\sum_{n=1}^{N} I_n) \prod_{q=1}^{N} I_q\right)$ if $I_n \leq \prod_{q \neq n}^{N} I_q$. In particular, the complexity of the KronF algorithm described in section 2.5.1.2 for estimating the symbol matrices is $O\left((\sum_{n=1}^{N} R_n J_n) \prod_{q=1}^{N} R_q J_q\right)$.

Algorithms	Computational Complexity
Bi-ALS-KonF (5.4) and (5.5)	$O\left(4M_T^2PJ\right) + O\left(2R^2PM_R\right) +$ $O\left(\left(\sum_{n=1}^N R_nJ_n\right)\prod_{q=1}^N R_qJ_q\right)$
Bi-ALS-KonF (5.6) and (5.7)	$O\left(4M_T^2PR\right) + O\left(R^2J\right) + O\left(8M_T^2M_R\right) +$ $O\left(\left(\sum_{n=1}^N R_nJ_n\right)\prod_{q=1}^N R_qJ_q\right)$
THOSVD	$O\left(4M_T^2R^2P\right) \quad + \quad O\left(4M_TM_RR^2J^2\right) \quad +$ $O\left(4M_RM_T(\sum_{n=1}^N R_nJ_n)\prod_{q=1}^N R_qJ_q\right)$

Table 23 – Computational complexity of the Bi-AKS-KronF and THOSVD algorithms.

Algorithms	Computational Complexity
ALS	$O\left(K^2(16M^3 + M^4)\right)$
RectALS	$O\left(K^2(16M^3 + M^4)\right) + O\left(4KM^3\right) \approx O\left(K^2(16M^3 + M^4)\right)$
JIRAFE without rectification	$O\left(4KM^4 + K^2M^4 + 2K^3M + K^4 + 2K^5M\right)$
JIRAFE with rectification	$O\left(4KM^4 + K^2M^4 + 2K^3M + K^4 + 2K^5M\right) + O\left(4KM^3\right)$ $\approx O\left(4KM^4 + K^2M^4 + 2K^3M + K^4 + 2K^5M\right)$

Table 24 – Computational complexity of the ALS/RectALS/JIRAFE algorithms in the second stage.

To compute the PARAFAC decomposition of a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ assumed to be of rank R, using the ALS algorithm, requires to compute N LS estimates, which needs to pseudo-inverse $\prod_{q \neq n}^N I_q \times R$ matrices, for $n \in [1,N]$, and induces the overall computational complexity $O\left(R^2 \sum_{n=1}^N (\prod_{q \neq n}^N I_q)\right)$, at each iteration. See [111] for more details.

In Table 23, the computational complexities of the algorithms of stage 1 in Figure 17 are compared, i.e., the Bi-ALS-KronF and THOSVD algorithms.

The complexity of the Bi-ALS-KronF (5.6) and (5.7) is lower than that of the Bi-ALS-KronF (5.4) and (5.5) due to the simplification of the pseudo-inverses in (5.4) and (5.5). Note that, unlike the complexity of HOSVD, the one for the Bi-ALS algorithms must be multiplied by the number of iterations needed for convergence, which explains why the computation time with the closed-form solution (HOSVD algorithm) is generally lower than with the iterative Bi-ALS algorithms.

Table 24 provides the computational complexity of the ALS, RectALS and JIRAFE algorithms of stage 2, assuming $M_{T_x} = M_{T_y} = M_{R_x} = M_{R_y} = M$, and with the complexity of the rectification methods given by: $K\left(M_{T_x}^3 + M_{T_y}^3 + M_{R_x}^3 + M_{R_y}^3\right) = 4KM^3$. For the complexity of JIRAFE, see [63]. As previously concluded for the comparison of the Bi-ALS and HOSVD algorithms, the computation time with the closed-form JIRAFE method is generally much lower than with the ALS and RectALS iterative algorithms due to the great number of iterations needed for their convergence. The computational complexities of the rectification algorithms TR$_1$A and TR$_1$A$_{imp}$ are the same, equal to $O\left(4KM^3\right)$, which is negligible with respect to the ones of the standard ALS and JIRAFE algorithms.

5.4 Simulation results

In this section, we evaluate the performance of the proposed DD-DP MIMO system and associated receivers. First, in Section 5.4.1, we describe the simulations and present the considered performance criteria. In Section 5.4.2, we study the impact of design parameters on the symbol error rate (SER), using the ZF-KronF receiver. In Section 5.4.3, the proposed semi-blind receivers are compared in terms of SER, channel normalized mean square error (NMSE) and reconstructed signal NMSE. Finally, in Section 5.4.4, the results of the multipath parameters estimation are presented, and a comparison with state-of-the-art methods is carried out.

5.4.1 General description of the simulations

The noisy received signal tensor $\mathcal{Y}$ is simulated as:

$$\mathcal{Y} = \mathcal{X} + \alpha\mathcal{N}, \tag{5.31}$$

where $\mathcal{X}$ contains the noise-free received signals obtained by means of Eq. (4.36), and $\mathcal{N}$ is the additive noise tensor whose entries are zero-mean circularly-symmetric complex-valued Gaussian random variables, with unit variance, and α allows to fix the SNR calculated as:

$$\text{SNR} = 20\log\frac{\| \mathcal{X} \|_F}{\alpha \| \mathcal{N} \|_F}, \tag{5.32}$$

which gives $\alpha = \frac{\|\mathcal{X}\|_F}{\|\mathcal{N}\|_F}10^{-\text{SNR}/20}$. The steering matrices $(\mathbf{A}_{T_x}, \mathbf{A}_{T_y}, \mathbf{A}_{R_x}, \mathbf{A}_{R_y})$ are simulated as Vandermonde matrices, with random realizations of DoD and DoA angles $(\phi_k, \theta_k, \vartheta_k, \psi_k)$ following a uniform distribution in $[0, \pi]$ rad. At each run, the coefficients of $\mathbf{B} \in \mathbb{C}^{4\times K}$ are randomly drawn from a complex Gaussian distribution with zero mean and unit variance. The channels $\mathcal{H}^{(V_R)}$ and $\mathcal{H}^{(H_R)}$ are simulated using Eqs. (4.26)-(4.27). The symbols to be transmitted are randomly generated from the 16-QAM (quadrature amplitude modulation) alphabet, with a uniform distribution. As mentioned before, the coding tensor is designed for each Monte Carlo run, such as its unfolding $\mathbf{G}_{2PM_T\times R}$ be a truncated DFT matrix.

The performance criteria, plotted versus the SNR, are calculated as:

$$\text{NMSE}_{\mathcal{Z}} = \frac{1}{L}\sum_{l=1}^{L}\frac{\| \hat{\mathcal{Z}}_l - \mathcal{Z}_l \|_F^2}{\| \mathcal{Z}_l \|_F^2}, \tag{5.33}$$

where $\hat{\mathcal{Z}}_l$ is the tensor $\mathcal{Z}_l$ estimated at the l-th run, with $\mathcal{Z}_l \in \{\mathcal{H}_l, \mathcal{X}_l\}$. For DoD and DoA angles estimation, the mean square error (MSE) criterion is calculated as:

$$\text{MSE}_{DoD} = \frac{1}{K}\sum_{k=1}^{K}\left(\phi_k - \hat{\phi}_k\right)^2 + \left(\theta_k - \hat{\theta}_k\right)^2, \tag{5.34}$$

$$\mathrm{MSE}_{DoA} = \frac{1}{K}\sum_{k=1}^{K}\left(\vartheta_k - \hat{\vartheta}_k\right)^2 + \left(\psi_k - \hat{\psi}_k\right)^2. \quad (5.35)$$

where the MSE is calculated for the estimated angles of transmission (MSE_{DoD}) and of reception (MSE_{DoA}), respectively. The SER, NMSE and MSE are calculated by averaging the results over $L = 5.10^4$ Monte Carlo runs, after truncating the 5% worse and 5% better values to eliminate the influence of ill-convergence and outliers.

The transmission rate T_R is given by:

$$T_R = \frac{\sum_{n=1}^{N} J_n R_n - N}{P\prod_{n=1}^{N} J_n}\log_2(\mu), \quad (5.36)$$

where $\sum_{n=1} J_n R_n$ corresponds to the total number of transmitted symbols, N is the number of symbols assumed to be a priori known for ambiguity suppression, and μ denotes the cardinality of the symbol alphabet, i.e., the number of constellation points. Note that increasing the number J_n of symbols in the symbol matrix $\mathbf{S}^{(n)}$ induces an increase of coding diversity and a lower transmission rate. On the other hand, an increase of the number R_n of data streams leads to an increase of T_R, while an increase of the number P of repetitions implies a decrease of T_R, as shown in Table 25.

5.4.2 Impact of design parameters

First, we evaluate the performance of the proposed system under perfect channel knowledge, as illustrated in Figures 18(a)-21(b). In this case, the ZF receiver (5.19) is used to estimate the matrix $\mathbf{S}$, combined with the KronF algorithm to separate the symbol matrices. The design parameters considered in the simulations are given in Table 25.

Figure 18(a) compares the SER for three different data stream numbers: $R_1 = R_2 \in [2, 4, 6]$. From this figure, it can be concluded that increasing R_1 and R_2 induces a degradation of symbol estimation, while the transmission rate increases (see Table 25). As expected, this happens because an increase of R_1 and R_2 implies an increase of the number of symbols to be estimated, without increasing the number of data in the tensor $\mathcal{Y}$ for performing the symbols estimation.

Figure 18(b) shows the impact on the SER for different numbers of symbols per data stream: $J_1 = J_2 \in [2, 4, 6]$. From these simulation results, it can be concluded that the SER is improved when the numbers of symbols increase, which implies an increase of coding diversity at reception, since J_1 and J_2 are dimensions of the data tensor, which is not the case for R_1 and R_2. See Eq. (4.39). On the other hand, the transmission rate decreases as shown in Table 25.

Figures 19(a) and 19(b) illustrate the impact of numbers of transmit and receive antennas, with: $(M_{T_x}, M_{T_y}) \in [(2,2), (4,4), (5,5)]$ and $(M_{R_x}, M_{R_y}) \in [(3,3), (4,4), (5,5)]$, respectively.

Impact of	Design Parameters	Transmission rate	Figures
Number of data streams	$(M_{T_x};M_{T_y}) = (4{,}4); (M_{R_x},M_{R_y}) = (5{,}5);$ $K = 2; J_1 = J_2 = 2; P = 3; R_1 = R_2 \in [2,4,6]$	$T_R =$ $2; 4.66; 7.33$	Figure 18(a)
Number of symbols per data stream	$(M_{T_x},M_{T_y}) = (4{,}4); (M_{R_x},M_{R_y}) = (5{,}5);$ $K = 2; R_1 = R_2 = 2; P = 2; J_1 = J_2 \in [2,4,6]$	$T_R = 3;$ $1.75; 1.22$	Figure 18(b)
Number of transmit antennas	$(M_{R_x},M_{R_y}) = (6{,}6); K = 2; R_1 = R_2 = 2;$ $J_1 = J_2 = 4; P = 2; (M_{T_x},M_{T_y}) \in$ $[(2{,}2),(4{,}4),(5{,}5)]$	$T_R = 1.75$	Figure 19(a)
Number of receive antennas	$(M_{T_x},M_{T_y}) = (2{,}2); K = 2; R_1 = R_2 = 2;$ $J_1 = J_2 = 4; P = 2; (M_{R_x},M_{R_y}) \in$ $[(3{,}3),(4{,}4),(5{,}5)]$	$T_R = 1.75$	Figure 19(b)
Number of time blocks	$(M_{T_x},M_{T_y}) = (3{,}3); (M_{R_x},M_{R_y}) = (4{,}4);$ $K = 2; R_1 = R_2 = 4; J_1 = J_2 = 2; P \in [2,4,6]$	$T_R = 7;$ $3.5; 2.33$	Figure 20(a)
m-QAM	$(M_{T_x},M_{T_y}) = (3{,}3); (M_{R_x},M_{R_y}) = (4{,}4);$ $K = 2; R_1 = R_2 = 2; J_1 = J_2 = 4; P = 2;$ $m \in [4,8,16,32]$	$T_R = 0.87;$ $1.31; 1.75;$ 2.18	Figure 20(b)
Number of symbol matrices	$\mathbf{N=1}: (M_{T_x},M_{T_y}) = (7{,}7); (M_{R_x},M_{R_y}) = (8{,}8); K = 2, P = 16, J_1 = 4; R_1 = 9$ $\mathbf{N=2}: J_1 = J_2 = 4; R_1 = R_2 = 4$ $\mathbf{N=3}: J_1 = J_2 = 4, J_3 = 1; R_1 = 4, R_2 = 2, R_3 = 9;$ $\mathbf{N=5}: J_1 = J_2 = J_3 = J_4 = 2, J_5 = 1; R_1 = R_2 = R_3 = R_4 = 4, R_5 = 3$	$T_R = 0.46$	Figure 21(a)
SER$_{global}$, **SER of** $\mathbf{S}^{(1)}$ **and** $\mathbf{S}^{(2)}$	$(M_{T_x},M_{T_y}) = (4{,}4), (M_{R_x},M_{R_y}) = (5{,}5), K = P = 2, R_1 = R_2 = 2, J_1 = 2, J_2 = 8$	$T_R = 2.25$	Figure 21(b)
Comparison of the TSTF-MSMKron and OSTBC codings	$(M_{T_x},M_{T_y}) = (3,3); (M_{R_x},M_{R_y}) = (4,4);$ $R_1 = R_2 = 2; J_1 = J_2 = 4; P = 2; J = 2; R = 7$	$T_R = 1.75;$ $T_{OSTBC} =$ 6.22	Figure 22
Comparison of the receivers	$(M_{T_x},M_{T_y}) = (2{,}2), (M_{R_x},M_{R_y}) = (3{,}3), K = 3, R_1 = R_2 = 2, J_1 = J_2 = 4, P = 34$	$T_R = 0.109$	Figures 23 - 27

Table 25 – Design parameters for the simulations.

When the number of receive antennas is increased, the quality of symbol estimation is improved, due to an increase of space diversity, via an augmentation of the number of received signals. The impact of the number of transmit antennas is less important than the one of receive antennas. Note that the SER is null for SNR > 0 dB. Note also that the transmission rate does not vary with the antenna numbers.

Figure 20(a) presents the SER for different numbers of time blocks: $P \in [2,4,6]$. When this number is increased, time diversity of the system is increased and consequently

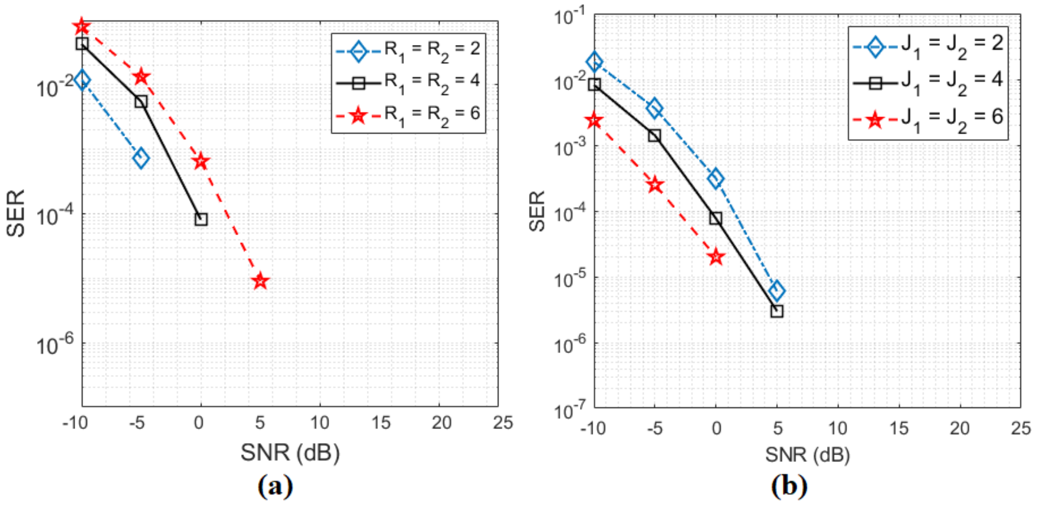


Figure 18 – Impact of: (a) data stream numbers and (b) numbers of symbols per data stream.

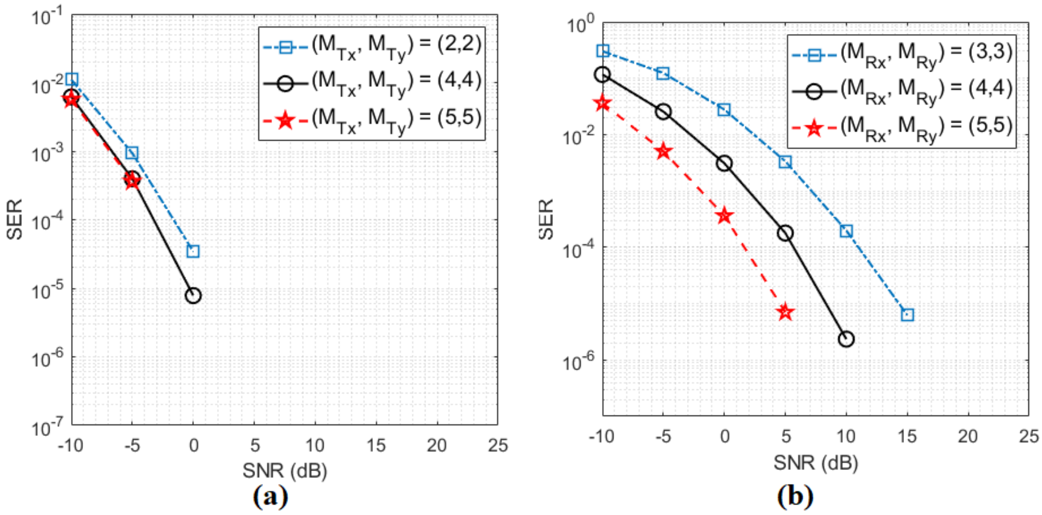


Figure 19 – Impact of: (a) transmit antennas number and (b) receive antennas number.

the SER is improved. On the other hand, the transmission rate decreases as we can see in Table 25.

Figure 20(b) compares the SER obtained with four different modulations m-QAM, $m \in [4,8,16,32]$. As expected, the SER performance is better when 4-QAM is used, because the decoding with 4-QAM is easier than with the other modulations, while inducing a lower transmission rate (see Table 25).

In Figure 21(a), we compare the MSMKron coding for different numbers of symbol matrices: $N \in [1,2,3,5]$. In the case where $N = 1$, i.e., when only a single symbol matrix is transmitted, then the MSMKron coding simplifies as a TST coding. From this figure, we conclude that an increase of N implies a significant improvement of the SER. The best performance is obtained with $N = 5$, which provides a gain of 10dB for a SER of 10^{-3}, in comparison with $N = 1$. These results corroborate the coding gain provided by the MSMKron coding owing the multiple Kronecker product of symbol matrices, which induces

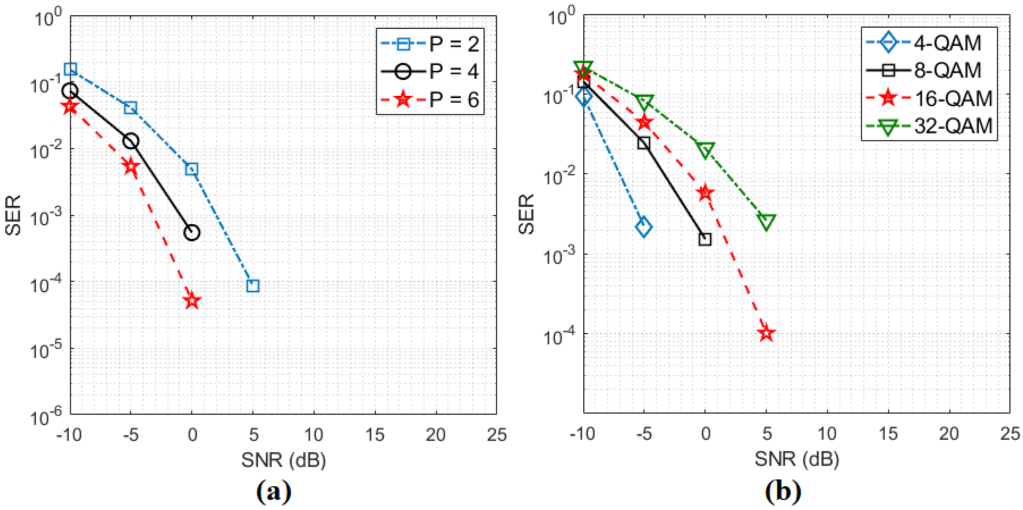


Figure 20 – Impact of: (a) time blocks number and (b) modulation (m-QAM).

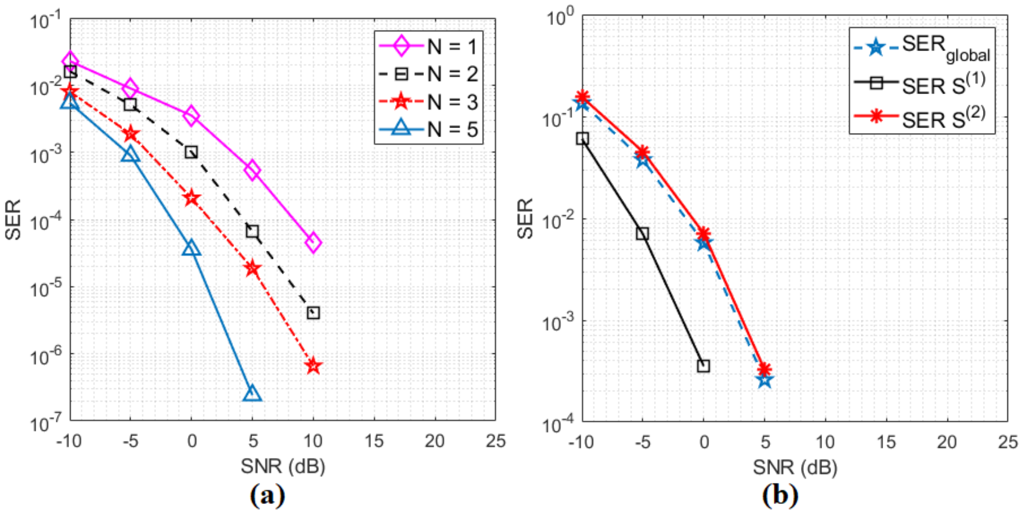


Figure 21 – (a) Impact of N on the SER and (b) SER of individual symbol matrices.

a redundancy augmentation when N is increased.

In Figure 21(b), the SER_{global} is compared with the individual SERs of the symbol matrices $\mathbf{S}^{(1)}$ and $\mathbf{S}^{(2)}$, considering $J_1 = 2$ and $J_2 = 8$. The Kronecker product of $\mathbf{S}^{(1)}$ and $\mathbf{S}^{(2)}$ induces a greater redundancy for $\mathbf{S}^{(1)}$ than for $\mathbf{S}^{(2)}$, since each symbol of $\mathbf{S}^{(1)}$ is repeated $8R_2$ times while each symbol of $\mathbf{S}^{(2)}$ is repeated only $2R_1$ times. As expected, the simulations results show that the best individual SER is obtained for $\mathbf{S}^{(1)}$ with the smallest dimension J_1 with respect to J_2, due to a greater redundancy provided by $\mathbf{S}^{(2)}$ in the Kronecker product. The global SER is close to the individual SER of $\mathbf{S}^{(2)}$.

In Figure 22, the proposed TST-MSMKron coding is compared with the OSTBC coding matrix [112] using a single symbol matrix $\mathbf{S} \in \mathbb{C}^{J \times R}$ instead of a multiple Kronecker product of symbol matrices. With the OSTBC coding, the symbol matrix is estimated through

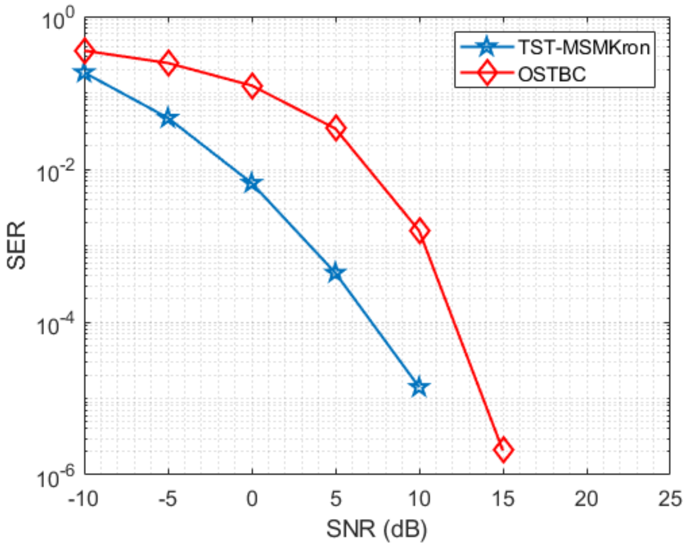


Figure 22 – Comparison of the TST-MSMKron and OSTBC codings.

the ZF receiver as:

$$\hat{\mathbf{S}}_{ZF}^{T} = (\mathbf{HW})^{\dagger}\,\mathbf{X}, \tag{5.37}$$

where $\mathbf{W} \in \mathbb{C}^{2M_T \times R}$ is the OSTBC coding matrix and $\mathbf{X} \in \mathbb{C}^{2M_R \times J}$. The transmission rate is given by:

$$T_{OSTBC} = \frac{JR}{M_T}\log_2(\mu). \tag{5.38}$$

where JR is the total number of transmitted symbols. For both codings, the number of transmitted symbols is the same. See the design parameters in Table 25. As expected, from Figure 22, we conclude that the TST-MSMKron coding gives a better SER than the OSTBC coding thanks to a greater coding diversity brought by the Kronecker product of symbol matrices. As a counterpart, the transmission rate with the TSTF-MSMKron coding is smaller than the one with the OSTBC coding. See Table 25.

Analyzing the above presented numerical results allows us to draw the following conclusions about the new TST-MSMKron coding:

- space diversity is provided by both the receive antennas (M_R) and the coding tensor $\mathcal{G}$ via the transmit antennas, implying an increase of the dimensions of the received signals tensor, and a greater redundancy in the transmitted symbols, respectively. See Eqs. (4.34) and (4.37);

- time diversity is provided by the transmission in P blocks which implies a repetition of transmitted symbols and an increase of the number of received signals;

- coding diversity is introduced by the TST-MSMKron coding which creates a redundancy of each transmitted symbol, while increasing the number of received signals via the dimensions J_n of the received signals tensor.

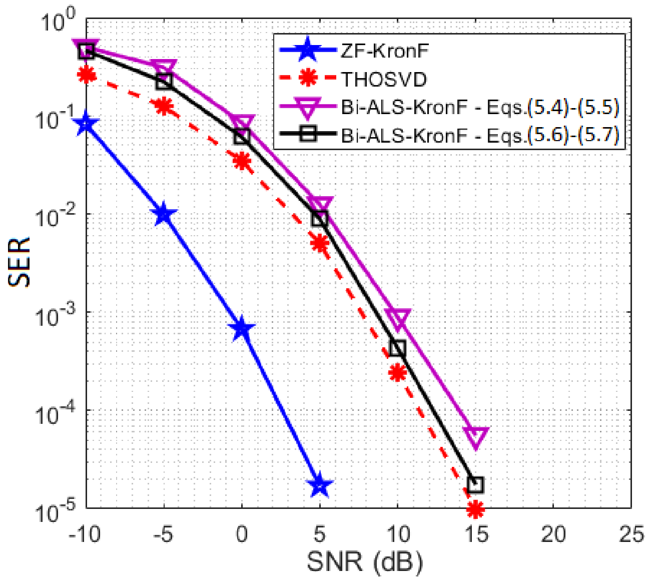


Figure 23 – SER comparison with THOSVD, Bi-ALS-KronF and ZF receivers.

Space, time and coding diversities are highlighted in the expression (4.39) of received signals, owing the sums on transmit antennas due to the tensor coding, and on the numbers of symbols contained in each data stream, as resulting from cross-multiplications between symbols provided by multiple Kronecker products which define the MSMKron coding.

5.4.3 Comparison of the proposed semi-blind receivers

In the next experiments, we first compare the SERs obtained with the proposed semi-blind and ZF-KronF receivers (using Eq. (5.19)), as shown in stage 1 of Figure 17. For the Bi-ALS-KronF receiver, both versions corresponding to Eqs. (5.4)-(5.5) and Eqs. (5.6)-(5.7) are considered. Then, we compare the performance of semi-blind receivers, in terms of channel and reconstructed signal NMSEs. For these simulations, the design parameters have the following values: $(M_{T_x}, M_{T_y}) = (2{,}2)$, $(M_{R_x}, M_{R_y}) = (3{,}3)$, $K = 3$, $J_1 = J_2 = 4$, $R_1 = R_2 = 2$, $P = 34$.

From Figures 23 and 24, we can conclude that the THOSVD receiver provides better performance than the Bi-ALS-KronF ones. That is due to the closed-form of THOSVD allowing to jointly estimate the channel and symbol matrices, while the Bi-ALS-KronF receivers are composed of two steps, one iterative to estimate $(\mathbf{H}, \mathbf{S})$, and one in closed-form to separate the symbol matrices from $\mathbf{S}$ estimated in the first step. Note also that the simplified Bi-ALS-KronF receiver (5.6)-(5.7) performs a little better than the other version (5.4)-(5.5) due to the simplification of the pseudo-inverse calculation. As expected, the ZF-KronF receiver provides the better SER due to a priori knowledge of the channel.

In Table 23, a comparison of the complexities of Bi-ALS-KronF receiver, both versions corresponding to Eqs. (5.4)-(5.5) and Eqs. (5.6)-(5.7) and THOSVD is provided. Based

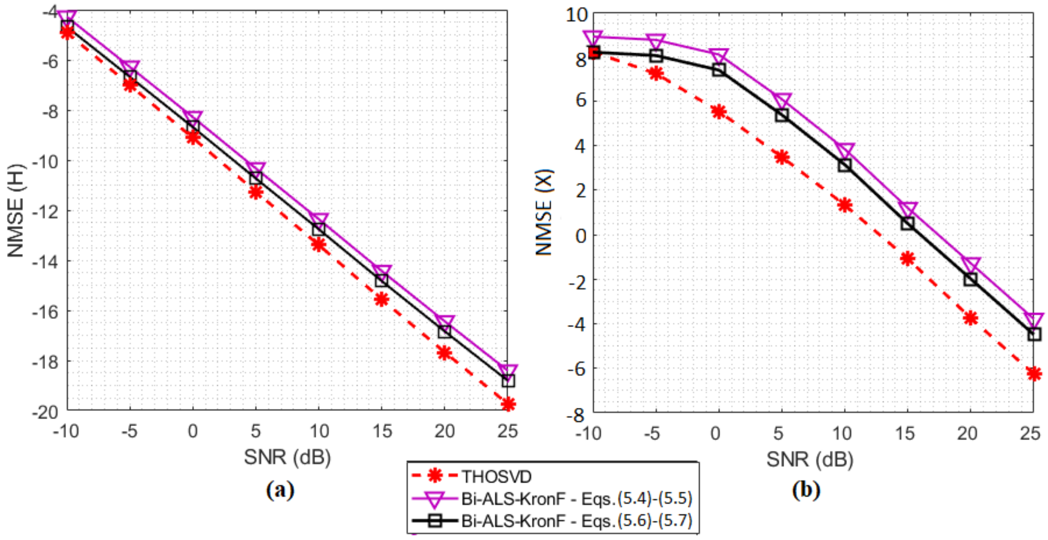


Figure 24 – (a) Channel NMSE comparison and (b) Reconstructed signal NMSE comparison.

on this table we can define the ratios $O_1 = O_{Bi-ALS-KronF(5.4)-(5.5)}/O_{Bi-ALS-KronF(5.6)-(5.5)_1}$, $O_2 = O_{THOSVD}/O_{Bi-ALS-KronF(5.4)-(5.5)}$ and $O_3 = O_{THOSVD}/O_{Bi-ALS-KronF(5.6)-(5.7)}$, which expresses how many times Bi-ALS-KronF (5.4)-(5.5) is more computational complexity than Bi-ALS-KronF (5.6)-(5.7) and how many times THOSVD algorithm is more computationally demanding than Bi-ALS-KronF algorithms. We have

$$O_1 = \frac{i_1(4M_T^2PJ + 2R^2PM_R) + \left(\sum_{n=1}^{N} R_nJ_n\right)\prod_{q=1}^{N} R_qJ_q}{i_2(4M_T^2PR + R^2J + 8M_T^2M_R) + \left(\sum_{n=1}^{N} R_nJ_n\right)\prod_{q=1}^{N} R_qJ_q} \tag{5.39}$$

$$O_2 = \frac{4M_T^2R^2P + 4M_TM_RR^2J^2 + 4M_RM_T(\sum_{n=1}^{N} R_nJ_n)\prod_{q=1}^{N} R_qJ_q}{i_1(4M_T^2PJ + 2R^2PM_R) + \left(\sum_{n=1}^{N} R_nJ_n\right)\prod_{q=1}^{N} R_qJ_q} \tag{5.40}$$

$$O_3 = \frac{4M_T^2R^2P + 4M_TM_RR^2J^2 + 4M_RM_T(\sum_{n=1}^{N} R_nJ_n)\prod_{q=1}^{N} R_qJ_q}{i_2(4M_T^2PR + R^2J + 8M_T^2M_R) + \left(\sum_{n=1}^{N} R_nJ_n\right)\prod_{q=1}^{N} R_qJ_q} \tag{5.41}$$

where i_1 and i_2 are the average numbers of iterations for convergence of the Bi-ALS-KronF (5.4)-(5.5) and Bi-ALS-KronF (5.6)-(5.7) algorithms, respectively. Figure 25 shows the complexity ratios O_1, O_2 and O_3 calculated using average values for i_1 and i_2 obtained from all the Monte Carlo runs and considering the variation of the number of time blocks $P \in [34, 40]$. From this figure, we can note that even as the number of time blocks increases, the complexities for the three receivers are linear. The Bi-ALS-KronF (5.6)-(5.7) is much less computationally demanding than Bi-ALS-KronF (5.4)-(5.5) and THOSVD algorithms due the simplification of the pseudo-inverse where this algorithm needs less number of iterations to converge, such that $O_{Bi-ALS-KronF(5.4)-(5.5)} >> O_{THOSVD} >> O_{Bi-ALS-KronF(5.6)-(5.7)}$.

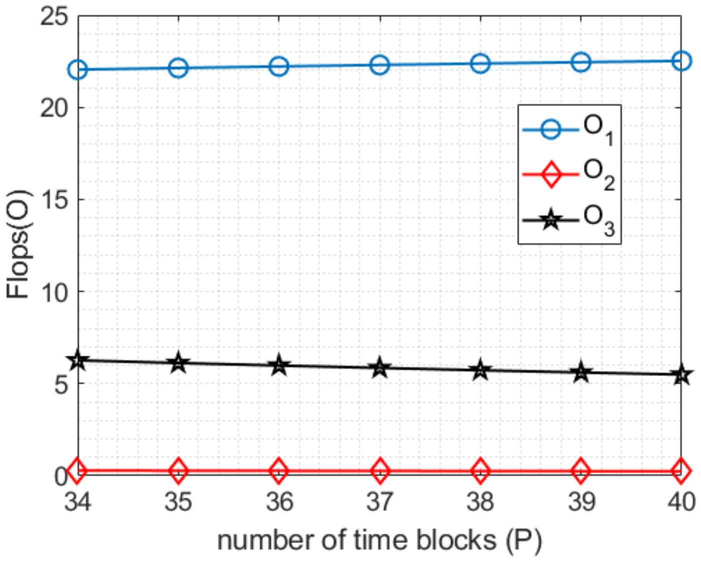


Figure 25 – Complexity ratio of THOSVD, Bi-ALS-KronF - Eqs. (5.4) and (5.7) and Bi-ALS-KronF - Eqs. (5.6)-(5.7) receivers.

5.4.4 Multipath parameters estimation

In this section, we evaluate the performance of the second stage of the proposed receivers in Figure 17, to estimate the multipath channel parameters whose values considered in the simulations are given in Table 25. The MSE performance for angles estimation at transmission and reception, respectively, obtained with the JIRAFE method is compared with the one provided by the RectALS algorithm which results from the following optimization problem:

$$\min_{\mathbf{A}_{T_x},\mathbf{A}_{T_y},\mathbf{A}_{R_x},\mathbf{A}_{R_y},\mathbf{B}} \| \mathcal{H} - \mathcal{I}_{5,K} \times_1 \mathbf{A}_{T_x} \times_2 \mathbf{A}_{T_y} \times_3 \mathbf{A}_{R_x} \times_4 \mathbf{A}_{R_y} \times_5 \mathbf{B} \|_F^2 \ . \tag{5.42}$$

A comparison is also made of the JIRAFE method without rectification and with the TR$_1$A and TR$_1$A$_{imp}$ rectification methods. Note that the estimated channel used in stage 2 is the one obtained with the best algorithm in stage 1, i.e., with the THOSVD receiver.

Figures 26(a) and 26(b) present the estimated angles MSE of transmission and reception, respectively, while Figure 27 plot the path-loss matrix **B** and reconstructed channel NMSEs. From Figures 26(a) and 26(b), we can conclude that the JIRAFE method combined with the TR1A$_{imp}$ rectification strategy outperforms the RectALS algorithm. This can be explained by the noise reduction property of the truncated SVD in the TTHSVD steps, which makes the JIRAFE method more robust to noise. Another cause is that the RectALS method is based on a five-step iterative algorithm, while the JIRAFE algorithm uses only a two-step ALS algorithm.

For the same reasons, Figures 27 (a) and (b) illustrate the superiority of the JIRAFE algorithm over the RectALS algorithm in terms of both the NMSE of the path-loss matrix

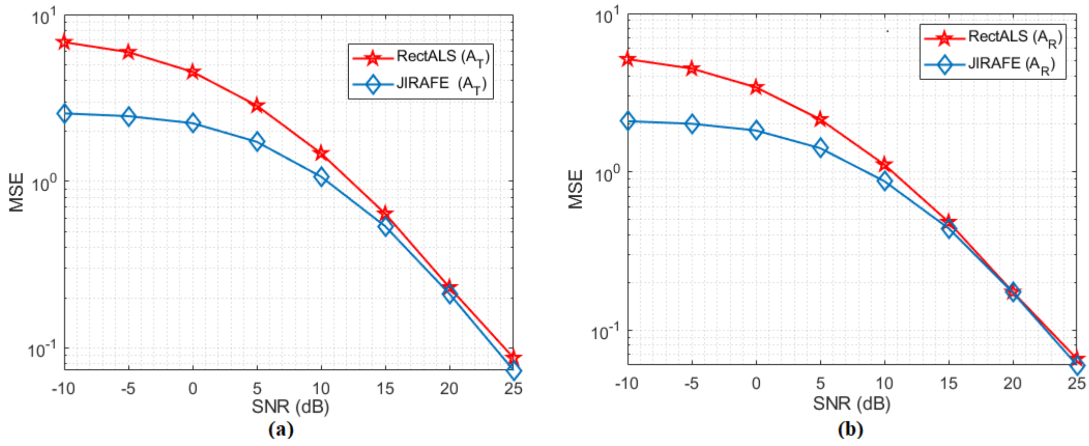


Figure 26 – Estimated angles MSE comparison of transmission (a) and reception (b).

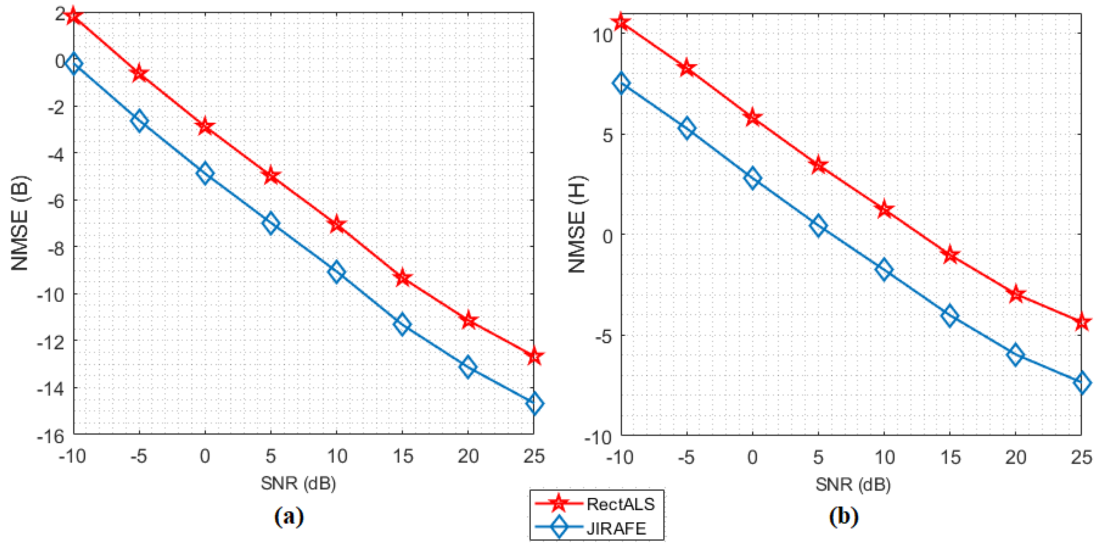


Figure 27 – (a) Estimated path-losses NMSEs and (b) reconstructed channel NMSEs.

and the reconstructed channel, with the best performance provided by the JIRAFE method combined with the TR1A$_{imp}$ rectification strategy.

5.5 Chapter summary

In this chapter, by exploiting the received signals tensor presented in Chapter 4 and assuming the TST coding tensor known, two semi-blind receivers in two stages have been derived to joint estimate the transmitted information symbols, the channel and channel parameters (DoD and DoA angles, path-loss coefficients). At the first stage, the first receiver is composed of two steps. The first one is based on the iterative ALS algorithm, while the second one considers the closed-form KronF algorithm to separate the symbol matrices. The second receiver is based on the closed-form THOSVD algorithm which can directly and jointly estimate the channel and symbol matrices. In comparison with the supervised system in [25] which requires using a pilot sequence to estimate the channel, the proposed receivers only need a priori knowledge of one symbol per symbol matrix. Necessary conditions for system identifiability have been established for each receiver, showing that the THOSVD receiver is more constraining than the Bi-ALS-KronF one. In the second stage, the estimated channel is used to estimate the channel parameters by means of the JIRAFE algorithm. A new rectification method has been proposed for ensuring the Vandermonde structure of the steering matrices, which allows to improve the accuracy of the DoD and DoA angles estimation.

Extensive Monte Carlo simulations have allowed to illustrate the impact of all the design parameters on the SER performance. And the performances of the proposed semi-blind receivers have been compared in terms of SER, channel NMSE and angles MSE. As expected, the closed-form receiver outperforms the iterative receiver. Simulation results also have illustrated the great flexibility of the TST-MSMKron coding, and corroborated the very good SER performance and of angles estimation obtained with the closed-form THOSVD algorithm combined with the JIRAFE method.

6 Conclusions and Perspectives

6.1 Conclusions

This thesis has addressed the study of tensor decompositions and algorithms/semi-blind receivers applied to data reconstruction and DD-DP FDD Massive MIMO system. In particular, new tensor models for data compression and for received signals were presented. The second one combines a particular case of MKronST coding with the TST coding and a 5th-order channel tensor, modeled by PARAFAC decomposition to transmit the symbol matrices. Based on this model, new semi-blind receivers composed of two steps were proposed that perform the joint estimation of the symbols, the channel, and the channel parameters without channel knowledge. Performance analysis of each proposed system showed their behavior, and effectiveness, evaluating the improvements of the techniques addressed. In the sequel, we provide a brief conclusion of each chapter that has proposed original contributions.

In **Chapter 3**, we presented the multidimensional CX decomposition to perform a low-rank approximation of an N-order tensor preserving initial information. Based on the proposed tensor model, an algorithm was presented based on the computation of the probabilities of the columns of each matrix unfolding to perform the reconstruction of the data that preserves some data tensor information. The simulation results illustrated the effectiveness of the proposed algorithm in terms of reconstruction compared to the standard THOSVD method. The proposed algorithm corroborates to have a similar computation time to the THOSVD and with the advantage of keeping the information of the original data.

In **Chapter 4**, we presented a new received signal that uses the particular case of MKronST coding [10] combined with a TST coding and the 5-order channel tensor model to transmit the symbols. The coding used is called tensor space-time (TST)-Multiple Symbol Matrices Kronecker (MSMKron) coding, this coding does not require a pre-coding matrix and allows us to propose semi-blind receivers in two steps to joint estimate symbol matrix, channel, and channel parameters. Based on the TST-MSMKron coding and the channel tensor, the signals received were modeled as a new decomposition called $(N + 2)$-order coupled nested Tucker-PARAFAC decomposition, where the core tensor is the TST coding. The coupling is due to the coding that is common to the V_R- and H_R- polarized receive antennas. The nested structure results from the fact that a matrix unfolding of the channel tensor constitutes one matrix factor of the Tucker model, the other factors being the steering and path-loss matrices. Identifiability conditions were established for the new model. Under the assumption that the tensor coding is known at the reception, this system model is essentially unique under mild conditions.

In **Chapter 5**, by exploiting the tensor model of the received signals presented in Chapter 4 and the knowledge of the coding tensor, we present semi-blind receivers in two steps to joint estimate symbol matrix, channel, and channel parameters (DoD and DoA angles, path-loss coefficients). In comparison with the supervised system in [25] which requires using a pilot sequence to estimate the channel, the proposed receivers only need a priori knowledge of one symbol per symbol matrix. Parameter identifiability conditions and computational complexity for each receiver were established and compared. In the first step, the symbol matrices and channel are estimated by means of the Bi-ALS-KronF or THOSVD algorithms, and in the second step, the estimated channel is used to estimate the channel parameters (DoD and DoA angles, and path-loss coefficients) by means of the JIRAFE algorithm which exploits an equivalence between PARAFAC and tensor train decompositions. A new rectification method was proposed for ensuring the Vandermonde structure of the steering matrices. Simulation results showed that the semi-blind receiver is efficient to estimate the symbols, the channel, and the channel parameters and the great flexibility of the TST-MSMKron coding.

6.2 Perspectives

In the following, we list some perspectives from this thesis:

- Based on the tensor decomposition proposed in Chapter 3, one perspective is the development of more competitive CX tensor decomposition algorithms in terms of computational cost, as well as their combination with tensor completion algorithms related to applications as compressive/compressed sensing [113, 114] and data reconstruction for Big Data [115];

- Perspectives of Chapter 4 include the extension of this new tensor model for cooperative (multi-relay) massive MIMO mmWave cellular systems, with the objective to improve coverage. More generally, we need to exploit new technologies like reconfigurable intelligent surfaces (RIS), massive reconfigurable MIMO antennas, and mmWave/THz technologies, which opens new multidisciplinary research issues at the intersection of wireless communications, computer science, artificial intelligence, physics, and mathematics [116];

- We should propose a new coding based on the Kronecker or Khatri-Rao product of the symbol matrices, and propose new semi-blind receivers to estimate the parameters based on this new signals received model. This can be considered in order to increase the transmission rate;

- The proposed FDD MIMO system must be compared to the FDD MIMO systems presented in the literature to analyze the performance. We must consider the system presented in [25] to compare, as it considers the pilot symbols to perform the supervised estimation of the channel;

- We should consider different scenarios, such as the number of paths $K > 4$, to propose new semi-blind receivers to estimate the symbol matrices, channel and channel parameters.

References

[1] M.Z. Chowdhury et al. “6G wireless communication systems: Applications, requirements, technologies, challenges, and research directions”. In: *IEEE Open Journal of the Communications Society* 1 (2020), pp. 957–975.

[2] S. Mumtaz, J.M. Jornet, et al. “Terahertz communication for vehicular networks”. In: *IEEE Transactions on Vehicular Technology* 66.7 (2017).

[3] W. Saad, M. Bennis, and M. Chen. “A vision of 6G wireless systems: Applications, trends, technologies, and open research problems”. In: *IEEE network* 34.3 (2019), pp. 134–142.

[4] S. Shahzadi, M. Iqbal, and N.R. Chaudhry. “6g vision: Toward future collaborative cognitive communication (3c) systems”. In: *IEEE Communications Standards Magazine* 5.2 (2021), pp. 60–67.

[5] T. Nakamura. “5G Evolution and 6G”. In: *IEEE symposium on VLSI technology*. IEEE. 2020, pp. 1–5.

[6] D.S. Rocha. “Nested tensor decomposition applied to cooperative MIMO communication systems”. PhD thesis. Brazil: University of Ceara, 2019.

[7] A.L.F. de Almeida and G. Favier. “Double Khatri-Rao space-time-frequency coding using semi-blind PARAFAC based receiver”. In: *IEEE Signal Processing Letters* 20.5 (2013), pp. 471–474.

[8] G. Favier, M.N. da Costa, et al. “Tensor space-time (TST) coding for MIMO wireless communication systems”. In: *Signal Processing* 92.4 (2012), pp. 1079–1092.

[9] G. Favier and A.L.F. de Almeida. “Tensor space-time-frequency coding with semiblind receivers for MIMO wireless communication systems”. In: *IEEE Transactions on Signal Processing* 62.22 (2014), pp. 5987–6002.

[10] W. Freitas Jr, G. Favier, and A.L.F. de Almeida. “Generalized Khatri-Rao and Kronecker space-time coding for MIMO relay systems with closed-form semi-blind receivers”. In: *Signal Processing* 151 (2018), pp. 19–31.

[11] P.H.U. de Pinho, M.F.K.B. Couras, et al. “Semi-supervised receivers for MIMO systems with multiple Khatri-Rao coding”. In: *13th International Conference on Signal Processing and Communication Systems (ICSPCS)*. IEEE. 2019, pp. 1–7.

[12] S.V. N. Randriambelonoro, G. Favier, and R. Boyer. “Semi-blind joint symbols and multipath parameters estimation of MIMO systems using KRST/MKRSM coding”. In: *Digital Signal Processing* 109 (2021), p. 102908.

[13] 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”. In: *IEEE Transactions on Signal Processing* 61.8 (2013), pp. 1895–1909.

[14] D.S. Rocha, C.A.R. Fernandes, and G. Favier. “MIMO multi-relay systems with tensor space-time coding based on coupled nested Tucker decomposition”. In: *Digital Signal Processing* 89 (2019), pp. 170–185.

[15] A. Mchangama, J. Ayadi, et al. “MmWave massive MIMO small cells for 5G and beyond mobile networks: An overview”. In: *12th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP)*. IEEE. 2020, pp. 1–6.

[16] Z. Pi and F. Khan. “An introduction to millimeter-wave mobile broadband systems”. In: *IEEE communications magazine* 49.6 (2011), pp. 101–107.

[17] A.L. Swindlehurst, E. Ayanoglu, et al. “Millimeter-wave massive MIMO: The next wireless revolution?” In: *IEEE Communications Magazine* 52.9 (2014), pp. 56–62.

[18] S. Rangan, T.S Rappaport, and E. Erkip. “Millimeter-wave cellular wireless networks: Potentials and challenges”. In: *Proceedings of the IEEE* 102.3 (2014), pp. 366–385.

[19] A. Alkhateeb, O. El Ayach, et al. “Channel estimation and hybrid precoding for millimeter wave cellular systems”. In: *IEEE journal of selected topics in signal processing* 8.5 (2014), pp. 831–846.

[20] P. Schniter and A. Sayeed. “Channel estimation and precoder design for millimeter-wave communications: The sparse way”. In: *48th Asilomar conference on signals, systems and computers*. IEEE. 2014, pp. 273–277.

[21] Z. Zhou, J. Fang, et al. “Channel estimation for millimeter-wave multiuser MIMO systems via PARAFAC decomposition”. In: *IEEE Transactions on Wireless Communications* 15.11 (2016), pp. 7501–7516.

[22] R.W. Heath and N. others Gonzalez-Prelcic. “An overview of signal processing techniques for millimeter wave MIMO systems”. In: *IEEE journal of selected topics in signal processing* 10.3 (2016), pp. 436–453.

[23] Z. Zhou, J. Fang, et al. “Low-rank tensor decomposition-aided channel estimation for millimeter wave MIMO-OFDM systems”. In: *IEEE Journal on Selected Areas in Communications* 35.7 (2017), pp. 1524–1538.

[24] J. Rodriguez-Fernandez, N. Gonzalez-Prelcic, et al. “Frequency-domain compressive channel estimation for frequency-selective hybrid millimeter wave MIMO systems”. In: *IEEE Transactions on Wireless Communications* 17.5 (2018), pp. 2946–2960.

[25] C. Qian, X. Fu, et al. “Tensor-based channel estimation for dual-polarized massive MIMO systems”. In: *IEEE Transactions on Signal Processing* 66.24 (2018), pp. 6390–6403.

[26] A. Kammoun, H. Khanfir, et al. "Preliminary results on 3D channel modeling: From theory to standardization". In: *IEEE Journal on Selected Areas in Communications* 32.6 (2014), pp. 1219–1229.

[27] M. Steinbauer et al. "Array measurement of the double-directional mobile radio channel". In: *VTC2000-Spring. 2000 IEEE 51st Vehicular Technology Conference Proceedings (Cat. No. 00CH37026).* Vol. 3. IEEE. 2000, pp. 1656–1662.

[28] K. Kalliola et al. "3-D double-directional radio channel characterization for urban macrocellular applications". In: *IEEE Transactions on Antennas and Propagation* 51.11 (2003), pp. 3122–3133.

[29] C. Oestges, B. Clerckx, et al. "Dual-polarized wireless communications: from propagation models to system performance evaluation". In: *IEEE Transactions on Wireless Communications* 7.10 (2008), pp. 4019–4031.

[30] R. Schmidt. "Multiple emitter location and signal parameter estimation". In: *IEEE transactions on antennas and propagation* 34.3 (1986), pp. 276–280.

[31] R. Roy and T. Kailath. "ESPRIT-estimation of signal parameters via rotational invariance techniques". In: *IEEE Transactions on acoustics, speech, and signal processing* 37.7 (1989), pp. 984–995.

[32] M.F.K.B. Couras, P.H.U. de Pinho, et al. "Semi-blind receivers based on a coupled nested Tucker-PARAFAC model for dual-polarized MIMO systems using combined TST and MSMKron codings". In: *Digital Signal Processing* (2023), p. 104043.

[33] Y. Zniyed, R. Boyer, et al. "Tensor train representation of MIMO channels using the JIRAFE method". In: *Signal Processing* 171 (2020), p. 107479.

[34] Y. Zniyed, R. Boyer, et al. "High-order CPD estimation with dimensionality reduction using a tensor train model". In: *26th European Signal Processing Conference (EUSIPCO).* IEEE. 2018, pp. 2613–2617.

[35] L. De Lathauwer, B. De Moor, and J. Vandewalle. "A multilinear singular value decomposition". In: *SIAM journal on Matrix Analysis and Applications* 21.4 (2000), pp. 1253–1278.

[36] T.G. Kolda and B.W. Bader. "Tensor decompositions and applications". In: *SIAM review* 51.3 (2009), pp. 455–500.

[37] S. Liu, G. Trenkler, et al. "Hadamard, Khatri-Rao, Kronecker and other matrix products". In: *International Journal of Information and Systems Sciences* 4.1 (2008), pp. 160–177.

[38] G. Favier and A.L.F. de Almeida. "Overview of constrained PARAFAC models". In: *EURASIP Journal on Advances in Signal Processing* 2014.1 (2014), pp. 1–25.

[39] R.A. Harshman et al. “Foundations of the PARAFAC procedure: Models and conditions for an ”explanatory” multimodal factor analysis”. In: (1970).

[40] L.R. Tucker. “Some mathematical notes on three-mode factor analysis”. In: *Psychometrika* 31.3 (1966), pp. 279–311.

[41] J.D. Carroll and J. Chang. “Analysis of individual differences in multidimensional scaling via an N-way generalization of ”Eckart-Young” decomposition”. In: *Psychometrika* 35.3 (1970), pp. 283–319.

[42] 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”. In: *IEEE transactions on signal processing* 63.18 (2015), pp. 4985–4998.

[43] 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”. In: *IEEE Signal Processing Letters* 23.3 (2016), pp. 316–320.

[44] 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 coding (TSTC)”. In: *Signal Processing* 128 (2016), pp. 318–331.

[45] I. Oseledets and E. Tyrtyshnikov. “TT-cross approximation for multidimensional arrays”. In: *Linear Algebra and its Applications* 432.1 (2010), pp. 70–88.

[46] I.V. Oseledets. “Tensor-train decomposition”. In: *SIAM Journal on Scientific Computing* 33.5 (2011), pp. 2295–2317.

[47] N.D. Sidiropoulos, G.B. Giannakis, and R. Bro. “Blind PARAFAC receivers for DS-CDMA systems”. In: *IEEE Transactions on Signal Processing* 48.3 (2000), pp. 810–823.

[48] 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”. In: *Psychometrika* 61 (1996), pp. 133–154.

[49] N.D. Sidiropoulos and R. Bro. “On the uniqueness of multilinear decomposition of N-way arrays”. In: *Journal of Chemometrics: A Journal of the Chemometrics Society* 14.3 (2000), pp. 229–239.

[50] J.B. Kruskal. “Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics”. In: *Linear algebra and its applications* 18.2 (1977), pp. 95–138.

[51] J.B. Kruskal. “Rank, decomposition, and uniqueness for 3-way and N-way arrays”. In: *Multiway data analysis* (1989), pp. 7–18.

[52] R.A. Harshman. “Determination and proof of minimum uniqueness conditions for PARAFAC1”. In: *UCLA working papers in phonetics* 22.111-117 (1972), p. 3.

[53] L. De Lathauwer. “A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization”. In: *SIAM journal on Matrix Analysis and Applications* 28.3 (2006), pp. 642–666.

[54] G. Favier. *Matrix and Tensor Decompositions in Signal Processing, Volume 2*. John Wiley & Sons, 2021.

[55] A.L.F. de Almeida and G. Favier. “Double Khatri–Rao space-time-frequency coding using semi-blind PARAFAC based receiver”. In: *IEEE Signal Processing Letters* 20.5 (2013), pp. 471–474.

[56] N. Pitsianis and C.F. Van Loan. *Approximation with Kronecker Products. Linear Algebra for Large Scale and Real Time application*. 1993.

[57] F. Roemer. “Advanced algebraic concepts for efficient multi-channel signal processing”. PhD thesis. Germany: Universitats bibliothek Ilmenau, 2013.

[58] M.N. da Costa, G. Favier, and J.M.T. Romano. “Tensor modelling of MIMO communication systems with performance analysis and Kronecker receivers”. In: *Signal Processing* 145 (2018), pp. 304–316.

[59] W.C. Freitas Jr, G. Favier, and A.L.F. de Almeida. “Sequential closed-form semiblind receiver for space-time coded multihop relaying systems”. In: *IEEE Signal Processing Letters* 24.12 (2017), pp. 1773–1777.

[60] K.K. Wu, Y. Yam, et al. “Kronecker product approximation with multiple factor matrices via the tensor product algorithm”. In: *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE. 2016, pp. 004277–004282.

[61] B. Sokal, A.L.F. de Almeida, and M. Haardt. “Semi-blind receivers for MIMO multi-relaying systems via rank-one tensor approximations”. In: *Signal Processing* 166 (2020), p. 107254.

[62] A.Y. Kibangou and G. Favier. “Non-iterative solution for PARAFAC with a Toeplitz matrix factor”. In: *17th European Signal Processing Conference*. IEEE. 2009, pp. 691–695.

[63] Y. Zniyed, R. Boyer, et al. “A TT-based hierarchical framework for decomposing high-order tensors”. In: *SIAM Journal on Scientific Computing* 42.2 (2020), A822–A848.

[64] Y. Zniyed, R. Boyer, et al. “Multidimensional harmonic retrieval based on Vandermonde tensor train”. In: *Signal Processing* 163 (2019), pp. 75–86.

[65] R. Boyer and P. Comon. “Rectified ALS algorithm for multidimensional harmonic retrieval”. In: *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*. IEEE. 2016, pp. 1–5.

[66] A.K. Smilde, P. Geladi, and R. Bro. *Multi-way analysis: applications in the chemical sciences*. John Wiley & Sons, 2005.

[67] C. Da Silva and F.J. Herrmann. "Optimization on the hierarchical tucker manifold–applications to tensor completion". In: *Linear Algebra and its Applications* 481 (2015), pp. 131–173.

[68] R. Socher, D. Chen, et al. "Reasoning with neural tensor networks for knowledge base completion". In: *Advances in neural information processing systems* 26 (2013).

[69] H.N. Phien, H.D. Tuan, et al. "Efficient tensor completion: Low-rank tensor train". In: *arXiv preprint arXiv:1601.01083* (2016).

[70] E. Candes and J. Romberg. "Sparsity and incoherence in compressive sampling". In: *Inverse problems* 23.3 (2007), p. 969.

[71] D.L. Donoho. "Compressed sensing". In: *IEEE Transactions on information theory* 52.4 (2006), pp. 1289–1306.

[72] M. Rani, S.B. Dhok, and R.B. Deshmukh. "A systematic review of compressive sensing: Concepts, implementations and applications". In: *IEEE access* 6 (2018), pp. 4875–4894.

[73] Z. Qin, J. Fan, et al. "Sparse representation for wireless communications: A compressive sensing approach". In: *IEEE Signal Processing Magazine* 35.3 (2018), pp. 40–58.

[74] E.J. Candes and B. Recht. "Exact low-rank matrix completion via convex optimization". In: *46th Annual Allerton Conference on Communication, Control, and Computing*. IEEE. 2008, pp. 806–812.

[75] B. Recht, M. Fazel, and P.A. Parrilo. "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization". In: *SIAM review* 52.3 (2010), pp. 471–501.

[76] M. Fazel. "Matrix rank minimization with applications". PhD thesis. Washigton: PhD thesis, Stanford University, 2002.

[77] M. Yuan and C. Zhang. "On tensor completion via nuclear norm minimization". In: *Foundations of Computational Mathematics* 16.4 (2016), pp. 1031–1068.

[78] S. Gandy, B. Recht, and I. Yamada. "Tensor completion and low-n-rank tensor recovery via convex optimization". In: *Inverse problems* 27.2 (2011), p. 025010.

[79] J. Liu, P. Musialski, et al. "Tensor completion for estimating missing values in visual data". In: *IEEE transactions on pattern analysis and machine intelligence* 35.1 (2012), pp. 208–220.

[80] A. Karatzoglou, X. Amatriain, et al. “Multiverse recommendation: n-dimensional tensor factorization for context-aware collaborative filtering”. In: *Proceedings of the fourth ACM conference on Recommender systems*. 2010, pp. 79–86.

[81] Q. Song, H. Ge, et al. “Tensor completion algorithms in big data analytics”. In: *ACM Transactions on Knowledge Discovery from Data (TKDD)* 13.1 (2019), pp. 1–48.

[82] H. Lee, Y. Kim, et al. “Nonnegative tensor factorization for continuous EEG classification”. In: *International journal of neural systems* 17.04 (2007), pp. 305–317.

[83] L. Huang, A.L.F. de Almeida, and H.C. So. “Target estimation in bistatic MIMO radar via tensor completion”. In: *Signal Processing* 120 (2016), pp. 654–659.

[84] J.H.M. Goulart, A.Y. Kibangou, and G. Favier. “Traffic data imputation via tensor completion based on soft thresholding of Tucker core”. In: *Transportation Research Part C: Emerging Technologies* 85 (2017), pp. 348–362.

[85] A. Gittens, J. Kottalam, et al. “A multi-platform evaluation of the randomized CX low-rank matrix factorization in Spark”. In: *IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*. IEEE. 2016, pp. 1403–1412.

[86] P. Drineas, M.W. Mahoney, and S. Muthukrishnan. “Relative-error CUR matrix decompositions”. In: *SIAM Journal on Matrix Analysis and Applications* 30.2 (2008), pp. 844–881.

[87] C.F. Caiafa and A. Cichocki. “Generalizing the column–row matrix decomposition to multi-way arrays”. In: *Linear Algebra and its Applications* 433.3 (2010), pp. 557–573.

[88] P. Drineas, M.W. Mahoney, and S. Muthukrishnan. “Subspace sampling and relative-error matrix approximation: Column-row-based methods”. In: *14th Annual European Symposium, Zurich, Switzerland, September 11-13, 2006. Proceedings 14*. Springer. 2006, pp. 304–314.

[89] M.W. Mahoney, M. Maggioni, and P. Drineas. “Tensor-CUR decompositions for tensor-based data”. In: *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*. 2006, pp. 327–336.

[90] N. Mitrovic, M.T. Asif, et al. “CUR decomposition for compression and compressed sensing of large-scale traffic data”. In: *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*. IEEE. 2013, pp. 1475–1480.

[91] M.F.K.B. Couras et al. “Multidimensional CX Decomposition of Tensors”. In: *2019 Workshop on Communication Networks and Power Systems (WCNPS)*. IEEE. 2019, pp. 1–4.

[92] J.H.M. Goulart and G. Favier. “A general iterative imputation scheme with feedback control for tensor completion (IFCTC)”. In: *Colloque GRETSI*. 2017.

[93] P.R.B. Gomes. “Tensor methods for array processing and channel estimation in wireless communications systems”. PhD thesis. Brazil: University of Ceara, 2018.

[94] B. Sokal. “Semi-blind receivers for multi-relaying mimo systems using rank-one tensor factorizations”. PhD thesis. Brazil: University of Ceara, 2017.

[95] T. Jiang and N.D. Sidiropoulos. “A direct blind receiver for SIMO and MIMO OFDM systems subject to unknown frequency offset and multipath”. In: *4th IEEE Workshop on Signal Processing Advances in Wireless Communications-SPAWC 2003 (IEEE Cat. No. 03EX689)*. IEEE. 2003, pp. 358–362.

[96] M. Sorensen and L. De Lathauwer. “Tensor decompositions with Vandermonde factor and applications in signal processing”. In: *2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*. IEEE. 2012, pp. 890–894.

[97] Z. Luo, L. Zhu, and C. Li. “Vandermonde constrained tensor decomposition based blind carrier frequency synchronization for OFDM transmissions”. In: *Wireless Personal Communications* 95 (2017), pp. 3459–3475.

[98] R. Yang, W. Zhang, and S. Lou. “Join adaptive channel estimation and data detection for MIMO-OFDM systems”. In: *Wireless Communications and Mobile Computing* 2020 (2020), pp. 1–9.

[99] L. Chen, X. Jiang, et al. “Tensor-Based Downlink Channel Reconstruction for FDD Massive MIMO”. In: *55th Asilomar Conference on Signals, Systems, and Computers*. IEEE. 2021, pp. 1453–1458.

[100] Y. Gong, H. Zhao, et al. “A multipath-aided localization method for MIMO-OFDM systems via tensor decomposition”. In: *IEEE Wireless Communications Letters* 11.6 (2022), pp. 1225–1228.

[101] D.C. Araujo, A.L.F. de Almeida, et al. “Tensor-based channel estimation for massive MIMO-OFDM systems”. In: *IEEE Access* 7 (2019), pp. 42133–42147.

[102] M.S. Ibrahim, C.I. Kanatsoulis, and N.D. Sidiropoulos. “Downlink channel feedback for FDD massive MIMO systems via tensor compression and sampling”. In: *54th Asilomar Conference on Signals, Systems, and Computers*. IEEE. 2020, pp. 27–31.

[103] N.D. Sidiropoulos and R.S. Budampati. “Khatri-Rao space-time codes”. In: *IEEE Transactions on Signal Processing* 50.10 (2002), pp. 2396–2407.

[104] M.N. da Costa et al. “Tensor space-time coding for MIMO wireless communication systems”. PhD thesis. Brazil: Universite Nice Sophia Antipolis; Universidade estadual de Campinas, 2014.

[105] J. Du, M. Han, et al. “Tensor-based joint channel estimation and symbol detection for time-varying mmWave massive MIMO systems”. In: *IEEE Transactions on Signal Processing* 69 (2021), pp. 6251–6266.

[106] T.Y. Elganimi, K.M. Rabie, and G. Nauryzbayev. “IRS-Assisted Millimeter-wave Massive MIMO with Transmit Antenna Selection for IoT Networks”. In: *arXiv preprint arXiv:2212.05854* (2022).

[107] R. Chang, C. Yuan, and J. Du. “PARAFAC-Based Multiuser Channel Parameter Estimation for MmWave Massive MIMO Systems over Frequency Selective Fading Channels”. In: *Electronics* 10.23 (2021), p. 2983.

[108] S. Gherekhloo, K. Ardah, et al. “Tensor-based channel estimation and reflection design for RIS-aided millimeter-wave MIMO communication systems”. In: *55th Asilomar Conference on Signals, Systems, and Computers*. IEEE. 2021, pp. 1683–1689.

[109] M.F.K.B. Couras et al. “Semi-blind receivers based on a coupled nested Tucker-PARAFAC model for dual-polarized MIMO systems using combined TST and MSMKron codings”. In: *Digital Signal Processing* 137 (2023), p. 104043.

[110] C. Qian, X. Fu, and N.D. Sidiropoulos. “Algebraic channel estimation algorithms for FDD massive MIMO systems”. In: *IEEE Journal of Selected Topics in Signal Processing* 13.5 (2019), pp. 961–973.

[111] P. Comon, X. Luciani, and A.L.F. de Almeida. “Tensor decompositions, alternating least squares and other tales”. In: *Journal of Chemometrics: A Journal of the Chemometrics Society* 23.7–8 (2009), pp. 393–405.

[112] C. M. Lau et al. “Performance of MIMO Systems Using Space Time Block Codes (STBC)”. In: *Open Journal of Applied Sciences* 11.03 (2021), p. 273.

[113] S. Fowler, G.G. Baravdish, and G. Baravdish. “Compressed sensing of wireless signals for image tensor reconstruction”. In: *2019 IEEE Global Communications Conference (GLOBECOM)*. IEEE. 2019, pp. 1–6.

[114] M. Nourazar and B. Goossens. “Accelerating iterative CT reconstruction algorithms using Tensor Cores”. In: *Journal of Real-Time Image Processing* 18.6 (2021), pp. 1979–1991.

[115] J.B. Ong. “Tensor computing for big data analytic”. PhD thesis. Singapore, 2020.

[116] E. Basar, M. Di Renzo, et al. “Wireless communications through reconfigurable intelligent surfaces”. In: *IEEE access* 7 (2019), pp. 116753–116773.

Appendix

APPENDIX A – Resumo Estendido em Língua Portuguesa

Título: Decomposição tensorial aninhada acoplada aplicada a sistemas de telecomunicação MIMO duplamente polarizado
Autor: Maria de Fátima Kallynna Bezerra Couras
Orientador: João Paulo Javidi da Costa
Coorientador: Vicente Zarzoso
Programa de Pós-Graduação em Sistemas Mecatrônicos
Brasília, 27 de junho de 2023

Palavras-chave: Decomposição Tensorial 1. Produto de Kronecker 2. Modelagem do canal 3. Receptores semi-cegos 4.

Introdução

O número de dispositivos conectados à Internet está aumentando exponencialmente e tecnologias emergentes, como cidades e casas inteligentes já são uma realidade. Aplicações emergentes, como streaming de TV, serviços de realidade virtual, entre muitas outras exigem transmissão de dados de alto rendimento e serviços de cobertura mais amplos. Para suportar os desafios das novas tecnologias, é necessário que sistemas sem fio de sexta geração (6G) sejam desenvolvidos com novos recursos atraentes, como alto desempenho em termos de taxa de dados, latência, confiabilidade, cobertura, conectividade maciça e móvel e eficiência energética [1, 2, 3]. Para o desenvolvimento dos sistemas de rede sem fio 6G, muitos desafios devem ser enfrentados. A geração de sinal em THz contínuo é difícil porque possui requisitos mais rigorosos em relação ao tamanho, é um sinal de comunicação de curta distância e tem mais complexidade no projeto das antena/transmissor, além de altas perdas de caminho de transmissão e alto custo [5, 6].

No caso de sistemas de comunicação sem fio, a satisfação desses requisitos depende muito de uma informação precisa do conhecimento do estado do canal (CSI) para compensar as perturbações causadas pela propagação sem fio que induz interações de sinais transmitidos com o ambiente. Os canais seguem modelos estatísticos caracterizados por meio de variáveis aleatórias que variam no tempo. Além do estudo do comportamento do canal, é necessário o desenvolvimento de técnicas que evitem o desvanecimento profundo do sinal antes de chegar ao seu destino. Uma alternativa para lidar com esses efeitos presentes nos canais sem fio é explorar as diversidades de sinal, como espaço, tempo, frequência, codificação e diversidades de cooperação. Quando a diversidade nos sistemas MIMO é explorada, significa que são criadas redundâncias no sinal, explorando a natureza aleatória da propagação de rádio de tal maneira que as versões diferentes e independentes do mesmo sinal atingem o destino. Como a diversidade é um recurso tão importante, um sistema sem fio pode usar vários tipos de diversidade simultaneamente [7].

Durante as últimas décadas, as decomposições tensoriais têm sido utilizadas para modelar sinais recebidos, bem como derivar receptores para estimar o canal e os símbolos que exploram várias formas de diversidade. O uso de decomposições tensoriais combinadas com novas codificações para transmitir sinais são amplamente propostas para introduzir diversidades extras aos sistemas MIMO ponto a ponto e relé. Alguns exemplos são as codificações espaço-tempo-frequência (STF), tensor espaço-tempo (TST) e tensor espaço-tempo-frequência (TSTF) [8, 9, 10] e codificações com base em múltiplos produtos de Khatri-Rao e Kronecker das matrizes de símbolos [11, 12, 13]. O uso dessas codificações com abordagens tensoriais para modelar o design do sistema MIMO levou ao desenvolvimento de novos modelos tensoriais [9, 10, 14, 13, 15] e permite a proposição de receptores semi-cegos para estimar conjuntamente os símbolos transmitidos e parâmetros do canal, fornecendo um aumento na taxa de transmissão em comparação com as outras codificações [12].

Os sistemas MIMO massivos fazem parte dos sistemas sem fio atuais [16]. Eles são muito importantes no desenvolvimento das redes sem fio 6G, porque precisam lidar com o manuseio de volumes de dados massivos e conectividade com alta taxa de dados por dispositivo [1, 2]. Os sistemas MIMO massivos também fornecem ganhos de desempenho em termos de confiabilidade e segurança sobre os sistemas de comunicação MIMO existentes [27, 23]. No entanto, usufruir dessas vantagens na prática depende da estimativa precisa do CSI, que afeta o desempenho da forma de transmissão de feixe nos transmissores e precisão da decodificação nos receptores [28]. Para isso, é necessário propor modelos de sinal recebidos e receptores que permitam estimar o canal e seus parâmetros com melhor precisão. A modelagem de canais polarizados 3D tem sido objeto de estudo de intensa pesquisa [27], porque permite explorar a componente do canal na direção de elevação, particularmente útil para sistemas MIMO massivos caracterizados por um grande número de antenas. Isso leva a canais MIMO direcionais duplos (DD), enquanto o uso de uma polarização dupla (DP) leva a canais MIMO duplo polarizados. Matrizes de antena DP, usando polarizações horizontais e verticais, podem oferecer grandes ganhos de multiplexação (isto é, capacidade) para os sistemas MIMO, combinando diversidades de espaço e polarização, como mostrado em [28]. No entanto, estimar os principais parâmetros de um canal DD-DP é um desafio devido ao grande número de parâmetros de canal a serem estimados. As técnicas de processamento de sinais baseadas em tensores podem explorar naturalmente a estrutura multidimensional dos dados e suas múltiplas formas de diversidade, permitindo elaborar algoritmos para receptor com boas compensações de complexidade de desempenho enquanto operam sobre escolhas mais relaxadas de parâmetros do que os receptores convencionais baseados em matrizes

[29].

Nesta tese, abordamos novos receptores semi-cegos para estimar conjuntamente símbolos, canal e parâmetros do canal em sistemas de comunicação sem fio MIMO DD-DP. Em particular, uma das principais contribuições desta tese está em torno da proposição de um novo sistema MIMO DD-DP, equipado com URAs nas duas extremidades do link e combinando com as codificações TST e do produto Kronecker de múltiplas matrizes de símbolos, que chamamos codificação TST-MSMKron. A codificação MSMKron é um caso particular da codificação MKronST (do inglês, múltiplo Kronecker Space-time) apresentada em [11]. O tensor do canal é representado por uma decomposição PARAFAC de quinta ordem separado em duas partes associadas às antenas de recepção verticalmente (VR) e horizontalmente (HR) polarizadas, respectivamente. Em seguida, é estabelecido que os tensores de sinais recebidos satisfazem um novo modelo Tucker-PARAFAC aninhado acoplado, cujo tensor core é o tensor de codificação TST.

Capitalizando esse novo modelo do tensor e o conhecimento do tensor de codificação, receptores semi-cegos são derivados para estimar os símbolos transmitidos e os parâmetros do canal (ângulos DOD e DOA e coeficientes de perda de caminho). Esses receptores são compostos por dois estágios. No primeiro, os símbolos da informação e o canal são estimados em conjunto, usando um algoritmo de mínimos quadrados alternados (Bi-ALS), seguido pelo algoritmo KronF para separar as matrizes de símbolo ou aplicando o algoritmo THOSVD, que estima direta e conjuntamente as matrizes de símbolo e o canal de forma fechada. No segundo estágio, a partir do tensor de canal estimado, os parâmetros do canal são estimados utilizando o algoritmo JIRAFE combinado com um novo método de retificação que leva em consideração a estrutura de Vandermonde das matrizes de direção. As condiçoes de identificabilidade são analisadas para cada receptor. Devido à natureza semi-cega dos receptores propostos, este trabalho estende a abordagem de [26], que usa um método supervisionado para estimativa do canal. Para ilustrar o comportamento e a eficácia dos esquemas propostos, são realizadas extensas simulações de Monte Carlo.

Materiais e Métodos

Nas próximas seções, o modelo do sistema é apresentado, em termos do modelo do canal, codificação e tensor de sinais recebidos.

Modelo do tensor do canal

Assumimos que o receptor e o transmissor empregam URAs, com pares de antenas polarizado cruzado de recepção e transmissão, respectivamente. Cada par de antenas consiste em uma antena polarizada vertical (V)- e uma polarizada horizontal (H). Definindo (M_{R_x}, M_{R_y}) and (M_{T_x}, M_{T_y}) como o número de antenas de recepção e transmissão, nas direções

x e y, respectivamente, e $M_T = M_{T_x} M_{T_y}$ e $M_R = M_{R_x} M_{R_y}$ como o número total de antenas de transmissão e recepção, o canal DD-DP é representado como uma matriz $\mathbf{H} \in \mathbb{C}^{2M_R M_T}$, particionado da seguinte maneira [26]:

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}^{(V_R,V_T)} & \mathbf{H}^{(V_R,H_T)} \\ \mathbf{H}^{(H_R,V_T)} & \mathbf{H}^{(H_R,H_T)} \end{bmatrix} = \begin{bmatrix} \mathbf{H}^{(V_R)} \\ \mathbf{H}^{(H_R)} \end{bmatrix}, \tag{A.1}$$

onde os blocos $\mathbf{H}^{(V_R)}$ e $\mathbf{H}^{(H_R)}$ representam os canais de cada polarização no receptor, $\mathbf{H}^{(V_R,V_T)} \in \mathbb{C}^{M_R \times M_T}$ é a matriz do sub-canal entre as antenas de transmissão e recepção V-polarizadas, e $\mathbf{H}^{(V_R,H_T)} \in \mathbb{C}^{M_R \times M_T}$ é a matriz do sub-canal entre as antenas de transmissão H-polarizada e as antenas de recepção V-polarizadas. Os outros dois blocos (A.1) são definidos da mesma forma. Cada sub-canal $\mathbf{H}^{(r,q)}$, com $r \in \{V_R, H_R\}$ e $q \in \{V_T, H_T\}$ é modelado como:

$$\mathbf{H}^{(r,q)} = \mathbf{A}_R \mathrm{diag}\left(\mathbf{b}^{(r,q)}\right) \mathbf{A}_T^T \in \mathbb{C}^{M_R \times M_T}, \tag{A.2}$$

com $b_k^{(r,q)}$ sendo a perda de caminho complexa para o k-ésimo caminho e o (r,q)-ésimo sub-canal, contendo os efeitos de polarização e desbotamento, e K é o número de caminho entre o transmissor e o receptor. As matrizes de direção são dadas por:

$$\mathbf{A}_T = \ [\mathbf{a}_T(\theta_1,\phi_1) \ \ \ldots \ \ \mathbf{a}_T(\theta_K,\phi_K)] \ = \mathbf{A}_{T_x} \diamond \mathbf{A}_{T_y} \in \mathbb{C}^{M_T \times K}, \tag{A.3}$$

$$\mathbf{A}_R = \ [\mathbf{a}_R(\psi_1,\vartheta_1) \ \ \ldots \ \ \mathbf{a}_R(\psi_K,\vartheta_K)] \ = \mathbf{A}_{R_x} \diamond \mathbf{A}_{R_y} \in \mathbb{C}^{M_R \times K}, \tag{A.4}$$

onde $\mathbf{A}_{T_x} \in \mathbb{C}^{M_{T_x} \times K}$, $\mathbf{A}_{T_y} \in \mathbb{C}^{M_{T_y} \times K}$, $\mathbf{A}_{R_x} \in \mathbb{C}^{M_{R_x} \times K}$ e $\mathbf{A}_{R_y} \in \mathbb{C}^{M_{R_y} \times K}$ são as matrizes de direção nas direções x e y, no transmissor e no receptor. Os vetores de direção para o k-ésimo caminho são definidos como:

$$\mathbf{a}_T(\theta_k,\phi_k) = \mathbf{a}_{T_x,k} \otimes \mathbf{a}_{T_y,k} \in \mathbb{C}^{M_{T_x} M_{T_y}}, \quad \mathbf{a}_R(\psi_k,\vartheta_k) = \mathbf{a}_{R_x,k} \otimes \mathbf{a}_{R_y,k} \in \mathbb{C}^{M_{R_x} M_{R_y}}, \tag{A.5}$$

cada vetor de direção é um vetor de Vandermonde em função dos ângulos de azimute e elevação. Substituindo (A.2) em (A.1) temos:

$$\mathbf{H} = \begin{bmatrix} \mathbf{A}_R & 0 \\ 0 & \mathbf{A}_R \end{bmatrix} \begin{bmatrix} D_1\left(\mathbf{B}^{(V_R)}\right) & D_2\left(\mathbf{B}^{(V_R)}\right) \\ D_1\left(\mathbf{B}^{(H_R)}\right) & D_2\left(\mathbf{B}^{(H_R)}\right) \end{bmatrix} \begin{bmatrix} \mathbf{A}_T & 0 \\ 0 & \mathbf{A}_T \end{bmatrix}^T, \tag{A.6}$$

com $\mathbf{B}^{(V_R)} = \begin{bmatrix} \mathbf{b}^{(V_R,V_T)} & \mathbf{b}^{(V_R,H_T)} \end{bmatrix}^T \in \mathbb{C}^{2 \times K}$, $\mathbf{B}^{(H_R)} = \begin{bmatrix} \mathbf{b}^{(H_R,V_T)} & \mathbf{b}^{(H_R,H_T)} \end{bmatrix}^T \in \mathbb{C}^{2 \times K}$, . Usando a propriedade (2.1) e substituindo $\mathbf{A}_T$ e $\mathbf{A}_R$ por suas expressões (A.3) e (A.4), respectivamente, temos:

$$\mathbf{H}^{(V_R)} = \left(\mathbf{A}_{R_x} \diamond \mathbf{A}_{R_y}\right)\left(\mathbf{B}^{(V_R)} \diamond \mathbf{A}_{T_x} \diamond \mathbf{A}_{T_y}\right)^T, \tag{A.7}$$

$$\mathbf{H}^{(H_R)} = \left(\mathbf{A}_{R_x} \diamond \mathbf{A}_{R_y}\right)\left(\mathbf{B}^{(H_R)} \diamond \mathbf{A}_{T_x} \diamond \mathbf{A}_{T_y}\right)^T. \tag{A.8}$$

Essas equações podem ser interpretadas como o desdobramento da matriz de dois tensores de quinta ordem $\mathcal{H}^{(V_R)}$ e $\mathcal{H}^{(H_R)}$ satisfazendo a decomposição PARAFAC $[\![\mathbf{A}_{T_x}, \mathbf{A}_{T_y}, \mathbf{A}_{R_x}, \mathbf{A}_{R_y}, \mathbf{B}^{(V_R)}; K]\!]$ e $[\![\mathbf{A}_{T_x}, \mathbf{A}_{T_y}, \mathbf{A}_{R_x}, \mathbf{A}_{R_y}, \mathbf{B}^{(H_R)}; K]\!]$, respectivamente:

$$\mathcal{H}^{(V_R)} = \mathcal{I}_{5,K} \times_1 \mathbf{A}_{T_x} \times_2 \mathbf{A}_{T_y} \times_3 \mathbf{A}_{R_x} \times_4 \mathbf{A}_{R_y} \times_5 \mathbf{B}^{(V_R)} \in \mathbb{C}^{M_{T_x} \times M_{T_y} \times M_{R_x} \times M_{R_y} \times 2}, \tag{A.9}$$

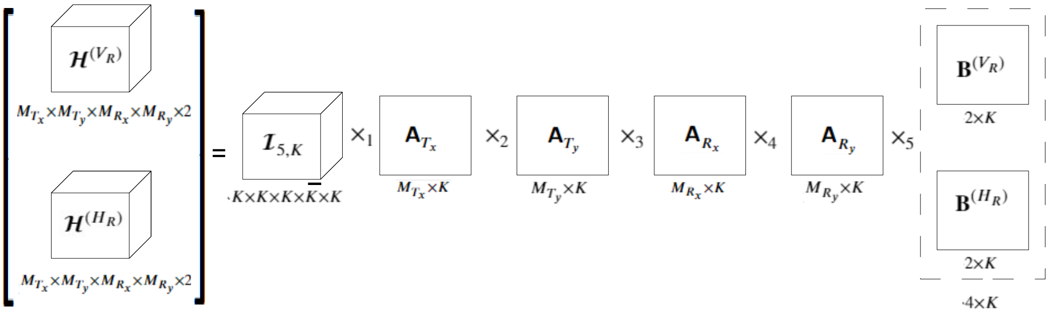


Figure 28 – Tensor do canal $\mathcal{H}$ modelado como uma decomposição PARAFAC.

$$\mathcal{H}^{(H_R)} = \mathcal{I}_{5,K} \times_1 \mathbf{A}_{T_x} \times_2 \mathbf{A}_{T_y} \times_3 \mathbf{A}_{R_x} \times_4 \mathbf{A}_{R_y} \times_5 \mathbf{B}^{(H_R)} \in \mathbb{C}^{M_{T_x}\times M_{T_y}\times M_{R_x}\times M_{R_y}\times 2}. \tag{A.10}$$

Concatenando os tensores do canal $\mathcal{H}^{(V_R)}$ e $\mathcal{H}^{(H_R)}$ ao longo do quinto modo e definindo $\mathbf{B} \in \mathbb{C}^{4\times K}$ como a concatenação das matrizes $\mathbf{B}^{(V_R)}$ e $\mathbf{B}^{(H_R)}$:

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}^{(V_R)} \\ \mathbf{B}^{(H_R)} \end{bmatrix}, \tag{A.11}$$

o canal do tensor global $\mathcal{H} \in \mathbb{C}^{M_{T_x}\times M_{T_y}\times M_{R_x}\times M_{R_y}\times 4}$ pode ser escrito como:

$$\mathcal{H} = \begin{bmatrix} \mathcal{H}^{(V_R)} \\ \mathcal{H}^{(H_R)} \end{bmatrix} = \mathcal{I}_{5,K} \times_1 \mathbf{A}_{T_x} \times_2 \mathbf{A}_{T_y} \times_3 \mathbf{A}_{R_x} \times_4 \mathbf{A}_{R_y} \times_5 \mathbf{B} \in \mathbb{C}^{M_{T_x}\times M_{T_y}\times M_{R_x}\times M_{R_y}\times 4}, \tag{A.12}$$

que é uma decomposição PARAFAC de quinta ordem e rank K ilustrada na Figura 28.

O tensor de canal resultante é uma extensão do tensor do canal de quarta ordem, apresentado em [26], com uma URA no transmissor e uma ULA no receptor. Observe que a modelagem do canal considerada não leva atrasos e mudanças de Doppler em consideração, assumindo que a largura de banda do canal é suficientemente pequena e o canal varia suficientemente lentamente em relação à duração da transmissão.

Codificação TST-MSMKron

Nesta seção, apresentamos a codificação usada no transmissor que combina as codificações TST e MSMKRon. Este último pode ser visto como a codificação MKronST [11], sem o conhecimento prévio de uma matriz de símbolo. Consiste em um produto de Kronecker de N símbolos $\mathbf{S}^{(n)} \in \mathbb{C}^{J_n\times R_n}$, com $n \in [1, N]$, cada uma sendo formado por R_n fluxos de dados compostos por J_n símbolos cada:

$$\mathbf{S} = \otimes_{n=1}^{N} \mathbf{S}^{(n)} \triangleq \mathbf{S}^{(1)} \otimes \ldots \otimes \mathbf{S}^{(N)} \in \mathbb{C}^{J\times R}, \tag{A.13}$$

onde $R = \prod_{n=1}^{N} R_n$ e $J = \prod_{n=1}^{N} J_n$. A codificação MSMKron é combinada com um tensor de codificação $\mathcal{G} \in \mathbb{C}^{2M_T\times R_1\times \ldots \times R_N \times P}$ de tal maneira que os sinais codificados formam um tensor de $(N+2)$-ordem que satisfaz o modelo Tucker dado por:

$$\mathcal{V} = \mathcal{G} \times_1 \mathbf{I}_{2M_T} \times_2 \mathbf{S}^{(1)} \times_3 \ldots \times_{N+1} \mathbf{S}^{(N)} \times_{N+2} \mathbf{I}_P \in \mathbb{C}^{2M_T\times J_1\times \ldots \times J_N \times P}. \tag{A.14}$$

Observe que o tensor do núcleo desta decomposição é o tensor de codificação $\mathcal{G}$. O seguinte desdobramento pode ser deduzido:

$$\mathbf{V}_{2M_T \times PJ} = \mathbf{G}_{2M_T \times PR} \left(\mathbf{I}_P \otimes \mathbf{S}\right)^T \in \mathbb{C}^{2M_T \times PJ}, \tag{A.15}$$

onde $\mathbf{G}_{2M_T \times PR} \in \mathbb{C}^{2M_T \times PR}$ é um desdobramento do tensor de codificação $\mathcal{G}$, e $\mathbf{S}$ é definido em (A.13). Observe que a matriz de identidade $\mathbf{I}_P \in \mathbb{R}^{P \times P}$ in (A.14), que fornece ao produto Kronecker em (A.15), está associado a P repetições das matrizes de símbolos por transmissão em blocos P.

Modelo do sinal recebido

Nesta seção, o modelo tensorial dos sinais recebidos é descrito. No caso sem ruído, os sinais recebidos são obtidos transmitindo a matriz de sinais codificados (A.15) através do canal $\mathbf{H}$ definido em (A.1), que fornece um modelo do sistema padrão como (4.1):

$$\mathbf{X}_{2M_R \times PJ} = \mathbf{H}\mathbf{V}_{2M_T \times PJ} \in \mathbb{C}^{2M_R \times PJ}. \tag{A.16}$$

Substituindo (A.15) em (A.16) leva à equação:

$$\mathbf{X}_{2M_R \times PJ} = \mathbf{H}\mathbf{G}_{2M_T \times PR} \left(\mathbf{I}_P \otimes \mathbf{S}\right)^T. \tag{A.17}$$

Levando em consideração a definição (A.13) de $\mathbf{S}$, concluímos que esta equação representa um desdobramento de um modelo Tucker de $(N+2)$-ordem cujo o tensor do núcleo is é a codificação $\mathcal{G}$ e as matrizes de fatores são canal $\mathbf{H}$, as matrizes de símbolos $\mathbf{S}^{(n)}, n \in [1, N]$ e a matriz identidade $\mathbf{I}_P$. Este modelo de Tucker para o tensor de sinais recebidos $\mathcal{X} \in \mathbb{C}^{2M_R \times J_1 \times \ldots \times J_N \times P}$ pode ser deduzido apartir de (A.14) e (A.16) como:

$$\mathcal{X} = \mathcal{V} \times_1 \mathbf{H} = \mathcal{G} \times_1 \mathbf{H} \times_2 \mathbf{S}^{(1)} \times_3 \ldots \times_{N+1} \mathbf{S}^{(N)} \times_{N+2} \mathbf{I}_P. \tag{A.18}$$

Combinando os modos 2 a $N+1$ dos tensores $\mathcal{G}$ e $\mathcal{X}$ resulta na forma contraída $\mathcal{G}_c \in \mathbb{C}^{2M_T \times R \times P}$ e $\mathcal{X}_c \in \mathbb{C}^{2M_R \times J \times P}$, e a expressão (A.18) pode ser escrita como:

$$\mathcal{X}_c = \mathcal{G}_c \times_1 \mathbf{H} \times_2 \mathbf{S} \times_3 \mathbf{I}_P. \tag{A.19}$$

Com base no modelo Tucker contraído de terceira ordem (A.19), é fácil deduzir os seguintes desdobramentos para $\mathcal{X}$:

$$\mathbf{X}_{PJ \times 2M_R} = \left(\mathbf{I}_P \otimes \mathbf{S}\right) \mathbf{G}_{PR \times 2M_T} \mathbf{H}^T \in \mathbb{C}^{PJ \times 2M_R}, \tag{A.20}$$

$$\mathbf{X}_{2PM_R \times J} = \left(\mathbf{I}_P \otimes \mathbf{H}\right) \mathbf{G}_{2PM_T \times R} \mathbf{S}^T \in \mathbb{C}^{2PM_R \times J}, \tag{A.21}$$

$$\mathbf{X}_{P \times 2M_R J} = \mathbf{G}_{P \times 2M_T R} \left(\mathbf{H} \otimes \mathbf{S}\right)^T \in \mathbb{C}^{P \times 2M_R J}. \tag{A.22}$$

Unicidade

A unicidade do modelo do sistema resulta unicidade do modelo Tucker (A.18) do tensor de sinal recebido $\mathcal{X}$ combinado com a unicidade da decomposição PARAFAC (A.12) do tensor do canal $\mathcal{H}$. Para o modelo Tucker (A.19),o tensor do núcleo será considerado conhecido na recepção, o que implica unicidade das matrizes até fatores de escala como:

$$\hat{\mathbf{S}} = \mathbf{S}\lambda^{(S)}, \quad \hat{\mathbf{H}} = \mathbf{H}\lambda^{(H)}, \quad \lambda^{(S)}\lambda^{(H)} = 1 \ . \tag{A.23}$$

O tensor do canal (A.12) é modelado pela decomposição PARAFAC de posto K que é essencialmente única se a seguinte condição for satisfeita [47]:

$$k_{\mathbf{A}_{T_x}} + k_{\mathbf{A}_{T_y}} + k_{\mathbf{A}_{R_x}} + k_{\mathbf{A}_{R_y}} + k_{\mathbf{B}} \geq 2K + 2, \tag{A.24}$$

onde $k_{\mathbf{A}}$ é o posto Kruskal de $\mathbf{A}$, ou seja, o maior número inteiro de tal forma que todo conjunto de $k_{\mathbf{A}}$ colunas de $\mathbf{A}$ seja linearmente independente. Se (A.24) é satisfeita, qualquer conjunto de matrizes $(\hat{\mathbf{A}}_{T_x}, \hat{\mathbf{A}}_{T_y}, \hat{\mathbf{A}}_{R_x}, \hat{\mathbf{A}}_{R_y}, \hat{\mathbf{B}})$ que satisfaça (A.12) estão relacionados com as matrizes originais por:

$$\begin{aligned} \hat{\mathbf{A}}_{T_y} &= \mathbf{A}_{T_y}\mathbf{\Pi}\mathbf{\Lambda}^{(T_y)}, & \hat{\mathbf{A}}_{T_x} &= \mathbf{A}_{T_x}\mathbf{\Pi}\mathbf{\Lambda}^{(T_x)}, \\ \hat{\mathbf{A}}_{R_y} &= \mathbf{A}_{R_y}\mathbf{\Pi}\mathbf{\Lambda}^{(R_y)}, & \hat{\mathbf{A}}_{R_x} &= \mathbf{A}_{R_x}\mathbf{\Pi}\mathbf{\Lambda}^{(R_x)}, \\ \hat{\mathbf{B}} &= \mathbf{B}\mathbf{\Pi}\mathbf{\Lambda}^{(B)}, & \mathbf{\Lambda}^{(T_y)}\mathbf{\Lambda}^{(T_x)}\mathbf{\Lambda}^{(R_y)}\mathbf{\Lambda}^{(R_x)}\mathbf{\Lambda}^{(B)} &= \mathbf{I}_K, \end{aligned} \tag{A.25}$$

onde $\mathbf{\Pi} \in \mathbb{C}^{K\times K}$ é a matriz de permutação e $\mathbf{\Lambda}^{(f)}$ são matrizes diagonais, com $f \in \{T_y, T_x, R_y, R_x, B\}$. Observe que a ambiguidade da permutação é irrelevante do ponto de vista prático, pois corresponde a uma permutação de caminhos. Devido à natureza aleatória dos parâmetros do canal, as matrizes $\mathbf{A}_{T_x}$, $\mathbf{A}_{T_y}$, $\mathbf{A}_{R_x}$, $\mathbf{A}_{R_y}$ e $\mathbf{B}$ são de posto completo. Então, no caso genérico, a condição (A.24) se torna:

$$\min\left(M_{T_y},K\right) + \min\left(M_{T_x},K\right) + \min\left(M_{R_y},K\right) + \min\left(M_{R_x},K\right) + \min\left(4,K\right) \geq 2K + 2. \tag{A.26}$$

Se assumirmos que todas as matrizes de direção e perda de caminho são posto de coluna completa, o que implica $K \leq \min(M_{T_y}, M_{T_x}, M_{R_y}, M_{R_x}, 4)$, então a condição acima é sempre satisfeita, o que garante a singularidade essencial da decomposição PARAFAC do tensor do canal. Observe que um pequeno valor de K resulta da natureza esparsa do canal. Neste trabalho, supõe-se que o posto do tensor seja conhecido. Quando K é desconhecido, é necessária uma etapa suplementar para estimar o número de caminhos. As ambiguidades de escala podem ser facilmente eliminadas, explorando a estrutura de Vandermonde das matrizes de direção (4.18) cuja primeira linha é composta por uns.

Receptores semi-cegos para systema MIMO DP usando a codificação TST-MSMKron

Ao explorar os desdobramentos (A.20) e (A.22) o tensor de sinais recebidos, dois receptores semi-cegos são desenvolvidos para estimar as matrizes de símbolo, o canal e os parâmetros do canal (ângulos DoA e DoD, perdas de caminho). Esses receptores são compostos por dois estágios. No primeiro, as matrizes de símbolo e o canal são estimadas, enquanto no segundo estágio, os parâmetros do canal são derivados do canal estimado.

Estágio 1 - Estimação do canal e símbolos

Receptor Bi-ALS-KronF

O algoritmo Bi-ALS é usado para estimar conjuntamente o produto MSMKron **S** e o desdobramento do canal **H**. Em seguida, o algoritmo KronF é aplicado para separar as matrizes de símbolo. O algoritmo Bi-ALS resulta do seguinte problema de otimização deduzido da Eq.(A.19):

$$\min_{\mathbf{S},\mathbf{H}} \| \mathcal{X}_c - \mathcal{G}_c \times_1 \mathbf{H} \times_2 \mathbf{S} \times_3 \mathbf{I}_P \|_F^2, \tag{A.27}$$

Fixando uma das matrizes $(\mathbf{H}, \mathbf{S})$ om seu valor estimado na iteração anterior, o BI-ALS substitui o problema de otimização (A.27) por dois sub-problemas de mínimos quadrados deduzidos dos desdobramentos (A.20) e (A.21), tal que:

$$\min_{\mathbf{H}} \| \mathbf{X}_{PJ\times 2M_R} - \left(\mathbf{I}_P \otimes \hat{\mathbf{S}}_{[it-1]}\right) \mathbf{G}_{PR\times 2M_T} \mathbf{H}^T \|_F^2 \longrightarrow \hat{\mathbf{H}}^T_{[it]}, \tag{A.28}$$

$$\min_{\mathbf{S}} \| \mathbf{X}_{2PM_R\times J} - \left(\mathbf{I}_P \otimes \hat{\mathbf{H}}_{[it]}\right) \mathbf{G}_{2PM_T\times R} \mathbf{S}^T \|_F^2 \longrightarrow \hat{\mathbf{S}}^T_{[it]}. \tag{A.29}$$

Para a convergência do algoritmo Bi-ALS, consideramos o erro na $[it]$-ésima iteração deduzido apartir de (A.20) como:

$$err_{[it]} = \| \mathbf{X}_{PJ\times 2M_R} - \left(\mathbf{I}_P \otimes \hat{\mathbf{S}}_{[it]}\right) \mathbf{G}_{PR\times 2M_T} \hat{\mathbf{H}}^T_{[it]} \|_F^2 \, . \tag{A.30}$$

Convergência na $[it]$-th iteração é declarada quando $|err_{[it-1]} - err_{[it]}| \leq \epsilon$, onde ϵ é um limiar pré-definido. Considerando as ambiguidade de escala para $\hat{\mathbf{H}}$ e $\hat{\mathbf{S}}$, as estimativas após corrigir as ambiguidades são dadas por:

$$\hat{\hat{\mathbf{S}}} = \hat{\mathbf{S}} \left(\lambda^{(S)}\right)^{-1}, \quad \hat{\hat{\mathbf{H}}} = \hat{\mathbf{H}} \lambda^{(S)}. \tag{A.31}$$

Para eliminar as ambiguidades de escala, é suficiente assumir que um elemento de **S** é conhecido a priori, por exemplo, $s_{11} = 1$. Sob essa suposição, $\lambda^{(S)}$ é calculado como $\lambda^{(S)} = \hat{s}_{11}$. As matrizes de símbolos $\mathbf{S}^{(n)}$ são então estimados usando o algoritmo KronF que minimiza a seguinte equação:

$$\min_{\mathbf{S}^{(n)},\, n\in\{1,\ldots,N\}} \| \hat{\hat{\mathbf{S}}} - \mathbf{S}^{(1)} \otimes \ldots \otimes \mathbf{S}^{(N)} \|_F^2 \, . \tag{A.32}$$

Este problema é resolvido reorganizando $\hat{\hat{\mathbf{S}}}$ como um tensor de posto um [56, 57, 50]:

$$\hat{\hat{\mathcal{S}}} = \mathbf{s}^{(1)} \circ \ldots \circ \mathbf{s}^{(N)} \in \mathbb{C}^{R_1 J_1 \times \ldots \times R_N J_N}, \tag{A.33}$$

onde $\mathbf{s}^{(n)}$ =vec($\mathbf{S}^{(n)}$) $\in \mathbb{C}^{R_n J_n}$, e substituindo (A.32) por:

$$\min_{\mathbf{s}^{(n)},\, n \in [1,N]} \parallel \hat{\hat{\mathcal{S}}} - \mathbf{s}^{(1)} \circ \cdots \circ \mathbf{s}^{(N)} \parallel_F^2 . \tag{A.34}$$

Cada forma vetorizada $\mathbf{s}^{(n)}$ pode ser estimada calculado a SVD de cada desdobramento de $\hat{\hat{\mathcal{S}}}$:

$$\hat{\hat{\mathbf{S}}}_{R_n J_n \times R_{n+1} J_{n+1} \ldots R_N J_N R_1 J_1 \ldots R_{n-1} J_{n-1}} = \mathbf{U}^{(n)} \mathbf{\Sigma}^{(n)} \mathbf{V}^{(n)H}, \tag{A.35}$$

com $\hat{\mathbf{s}}^{(n)}$ escolhido como o primeiro vetor singular esquerdo associado ao valor singular dominante do desdobramento de modo-n, ou seja,:

$$\hat{\mathbf{s}}^{(n)} = \mathbf{U}^{(n)}_{.1} \in \mathbb{C}^{R_n J_n}. \tag{A.36}$$

A matriz de símbolo estimada é obtida "desvetorizando" $\hat{\mathbf{s}}^{(n)}$ como:

$$\hat{\mathbf{S}}^{(n)} = \text{unvec}\left(\hat{\mathbf{s}}^{(n)}\right) \in \mathbb{C}^{J_n \times R_n}. \tag{A.37}$$

Assumindo $s^{(n)}_{11} = 1$,á ambiguidade escalar é corrigida como:

$$\hat{\hat{\mathbf{S}}}^{(n)} = \hat{\mathbf{S}}^{(n)} \left(\lambda^{(n)}\right)^{-1}, \text{with}, \lambda^{(n)} = \hat{s}^{(n)}_{11}. \tag{A.38}$$

Receptor THOSVD

É proposto um segundo receptor semi-cego para estimar direta e conjuntamente as matrizes de símbolos e canal, usando uma solução de forma fechada baseada no algoritmo THOSVD [31]. A partir do desdobramento (A.22), com $\mathbf{S}$ substituído por (A.13), a estimação de mínimos quadrados do produto de Kronecker é dado por:

$$\mathbf{Y} \triangleq \mathbf{H} \otimes \mathbf{S}^{\widehat{(1)} \otimes \ldots} \otimes \mathbf{S}^{(N)} = \left[(\mathbf{G}_{P \times 2M_T R})^{\dagger} \mathbf{X}_{P \times 2M_R J}\right]^T \in \mathbb{C}^{2M_R J \times 2M_T R}. \tag{A.39}$$

O desdobramento $\mathbf{G}_{P \times 2M_T R}$ deve ter um posto de coluna completa para garantir a unicidade da estimativa de mínimos quadrados, que induz a condição necessária: $2M_T R \leq P$. As matrizes de símbolos $\mathbf{S}^{(n)}$ e canal $\mathbf{H}$ são estimados conjuntamente através do algoritmo KronF, como na etapa 2 do receptor Bi-ALS-KronF, com a diferença que nós podemos simultaneamente estimar todas as matrizes $(\mathbf{H}, \mathbf{S}^{(1)}, \cdots, \mathbf{S}^{(N)})$.

Receptor Zero-Forcing (ZF)-KronF

Para avaliar o impacto dos parâmetros de design no desempenho do sistema, usamos o receptor zero-forcing (ZF)-KronF que assume um conhecimento perfeito do canal. $\hat{\mathbf{S}}_{ZF}$ de $\mathbf{S}$ é calculado usando (5.5) or (5.7), com $\mathbf{H}_{[it]}$ substituído pelo verdadeiro canal $\mathbf{H}$, tal que:

$$\hat{\mathbf{S}}^T_{ZF} = \left[(\mathbf{I}_P \otimes \mathbf{H})\, \mathbf{G}_{2PM_T \times R}\right]^{\dagger} \mathbf{X}_{2PM_R \times J}. \tag{A.40}$$

ou

$$\hat{\mathbf{S}}^T_{ZF} = \left(\mathbf{G}_{2PM_T \times R}\right)^H \left(\mathbf{I}_P \otimes \mathbf{H}^{\dagger}\right) \mathbf{X}_{2PM_R \times J}. \tag{A.41}$$

Em seguida, as matrizes de símbolo $\mathbf{S}^{(n)}$ são estimadas usando o algoritmo KronF, como na segunda etapa do receptor Bi-ALS-KronF.

Estágio 2 - Estimação dos parâmetros do canal usando JIRAFE

Nesta seção o algoritmo JIRAFE é apresentado para estimar os vetores de direção e as perdas de caminho a partir do desdobramento do canal estimado $\hat{\hat{\mathbf{H}}}$, reorganizando-o como o tensor $\hat{\hat{\mathcal{H}}}$ satisfazendo a decomposição PARAFAC (A.12):

$$\hat{\hat{\mathcal{H}}} = \text{reshape}\left(\hat{\hat{\mathbf{H}}}, \left[M_{T_x}, M_{T_y}, M_{R_x}, M_{R_y}, 4\right]\right) \in \mathbb{C}^{M_{T_x}\times M_{T_y}\times M_{R_x}\times M_{R_y}\times 4}. \quad \text{(A.42)}$$

Explorar a equivalência entre as decomposições de TT e PARAFAC, o tensor do canal estimado de quinta ordem (A.42) ode ser reescrito como a seguinte decomposição de TT deduzida de (2.75), para $N = 5$:

$$\hat{\hat{\mathcal{H}}} = \mathbf{G}^{(1)} \times_2^1 \mathcal{G}^{(2)} \times_3^1 \mathcal{G}^{(3)} \times_4^1 \mathcal{G}^{(4)} \times_5^1 G^{(5)}, \quad \text{(A.43)}$$

onde $\mathbf{G}^{(1)} \in \mathbb{C}^{M_{T_x}\times K}$, $\mathcal{G}^{(2)} \in \mathbb{C}^{K\times M_{T_y}\times K}$, $\mathcal{G}^{(3)} \in \mathbb{C}^{K\times M_{R_x}\times K}$, $\mathcal{G}^{(4)} \in \mathbb{C}^{K\times M_{R_y}\times K}$ e $\mathbf{G}^{(5)} \in \mathbb{C}^{K\times 4}$ são os núcleos TT, ligado aos fatores PARAFAC por meio das relações de equivalência (2.142)-(2.143) que se tornam [59, 30]:

$$\mathbf{G}^{(1)} = \mathbf{A}_{T_x}\mathbf{M}_1^{-1}, \quad \mathbf{G}^{(5)} = \mathbf{M}_4\mathbf{B}^T, \quad \text{(A.44)}$$

$$\mathcal{G}^{(2)} = \mathcal{I}_{3,K} \times_1 \mathbf{M}_1 \times_2 \mathbf{A}_{T_y} \times_3 \mathbf{M}_2^{-T}, \quad \text{(A.45)}$$

$$\mathcal{G}^{(3)} = \mathcal{I}_{3,K} \times_1 \mathbf{M}_2 \times_2 \mathbf{A}_{R_x} \times_3 \mathbf{M}_3^{-T}, \quad \text{(A.46)}$$

$$\mathcal{G}^{(4)} = \mathcal{I}_{3,K} \times_1 \mathbf{M}_3 \times_2 \mathbf{A}_{R_y} \times_3 \mathbf{M}_4^{-T}, \quad \text{(A.47)}$$

onde $\mathbf{M}_n \in \mathbb{C}^{K\times K}$, para $n \in [1,4]$, são são matrizes de mudança de base não singulares. O algoritmo TTHSVD, descrito na Seção 2.5.5, é usado para determinar as estimativas do núcleo TT $\hat{\mathbf{G}}^{(k)}$, para $k \in [1,5]$, e $\hat{\mathcal{G}}^{(k)}$, para $k \in [2,3,4]$.

Uma nova estratégia de retificação TR$_1$A$_{imp}$ é proposta para garantir a estrutura de Vandermonde das matrizes de direção estimadas na presença de ruído, cujos elementos da primeira linha são normalizados para 1, como em (4.18). Essa estratégia é uma versão aprimorada do método de aproximação de posto um Toeplitz (TR$_1$A) proposto em [62]. Nosso método de retificação proposto consiste na construção de uma matriz hermitiana Toeplitz de posto um $\mathbf{T}_k = \hat{\mathbf{A}}_{.k}\hat{\mathbf{A}}_{.k}^H$, de tamanho $I \times I$, aparti de cada coluna k de uma matriz de Vandermonde estimada, calculando sua decomposição de autovalores (EVD) $\mathbf{T}_k = \mathbf{P}_k\mathbf{D}_k\mathbf{P}_k^H$ para estimar a k-ésima frequência angular. Ao contrário do método TR$_1$A e identifica apenas um elemento da matriz Toeplitz com seu termo correspondente no EVD da matriz de posto 1 Toeplitz construída, nosso método explora a identificação de termos de $I - 1$ como:

$$e^{-j\omega_k i} = d_k(p_k)_{1,1}(p_k)^*_{i+1,1} \Rightarrow \hat{\omega}_k = \frac{1}{i}\angle\left((p_k)_{1,1}(p_k)^*_{i+1,1}\right), \quad \text{(A.48)}$$

para $i \in [1, I-1]$, onde $(p_k)_{i,1}$ denota o $(i,1)$ elemento da matriz de autovetor $\mathbf{P}_k$, e o operador $\angle$ significa o ângulo de seu complexo argumento escalar. Em seguida, a frequência angular estimada é calculada como uma média dessas estimativas de $I - 1$ (A.48) como:

$$\hat{\omega}_k = \frac{1}{I-1}\sum_{i=1}^{I-1}\frac{1}{i}\angle\left((p_k)_{1,1}(p_k)^*_{i+1,1}\right). \quad \text{(A.49)}$$

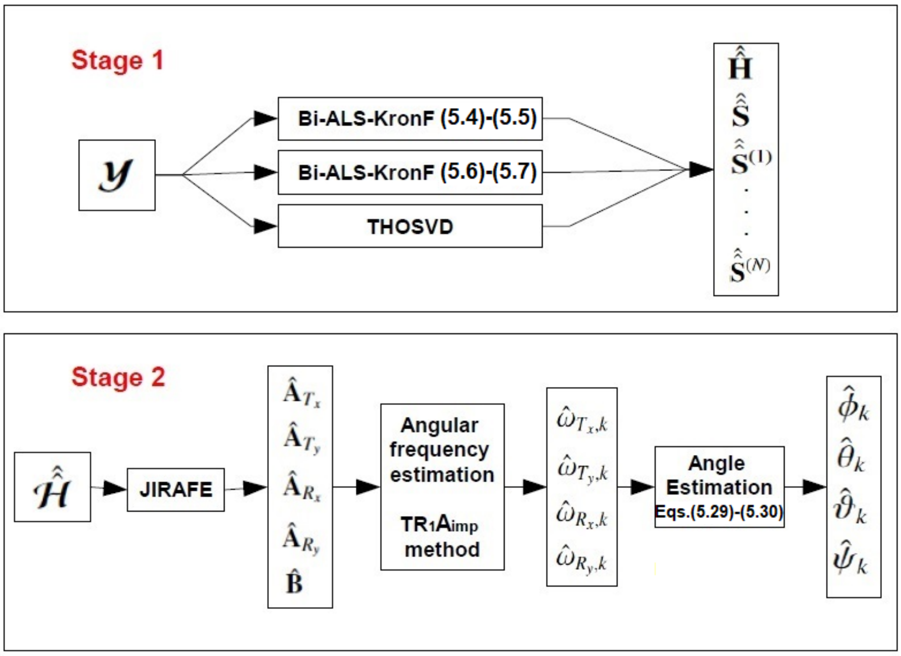


Figure 29 – Diagrama em blocos dos receptores propostos.

Essa estratégia é aplicada com o algoritmo ALS para estimar $\mathbf{A}_{T_y}$ a partir de $\hat{\mathcal{G}}^{(2)}$, e com o algoritmoKRF algorithm para estimar $\mathbf{A}_{R_x}$ e $\mathbf{A}_{R_y}$, a partir de $\hat{\mathcal{G}}^{(3)}$ e $\hat{\mathcal{G}}^{(4)}$, respectivamente, o que leva aos algoritmos RectALS e RectKRF. Finalmente, a estimação de $\mathbf{A}_{T_x}$ e $\mathbf{B}$ são deduzidas aparttir da Eq.(A.44) com a retificação de $\hat{\mathbf{A}}_{T_x}$. os receptores Bi-ALS/KronF-JIRAFE e THOSVD-JIRAFE, compostos de ois estágios, são ilustrados na Figura 29.

Resultados e Discussões

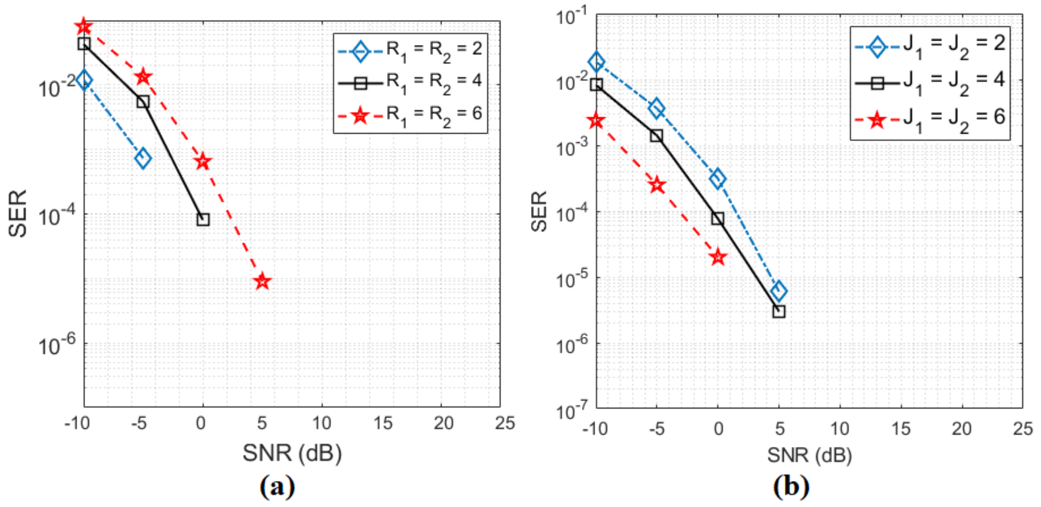
Impacto dos parâmetros de design

Primeiro, avaliamos o desempenho do sistema proposto sob o conhecimento perfeito do canal, conforme ilustrado nas Figuras 30 (a)-33 (b). Nesse caso, o receptor ZF (A.41)é usado para estimar $\mathbf{S}$, combinado com o algoritmo KronF para separar as matrizes de símbolos. Os parâmetros de design considerados nas simulações são fornecidos na Tabela 26.

Figura 30 (a) compara a SER para três números diferentes de fluxo de dados: $R_1 = R_2 \in [2, 4, 6]$. A partir desta figura, pode -se concluir que o aumento de R_1 e R_2 induz uma degradação da estimativa de símbolos, enquanto a taxa de transmissão aumenta (veja a Tabela 26). Como esperado, isso acontece porque um aumento de R_1 e R_2 implica um aumento do número de símbolos a serem estimados, sem aumentar o número de dados no tensor $\mathcal{Y}$ para realizar a estimação dos símbolos. Figura 30 (b) mostra o impacto na SER para diferentes números de símbolos por fluxo de dados: $J_1 = J_2 \in [2, 4, 6]$. A partir desses resultados da simulação, pode -se concluir que a SER melhora quando o número de símbolos aumenta, o que implica um aumento da diversidade da codificação na recepção, uma vez

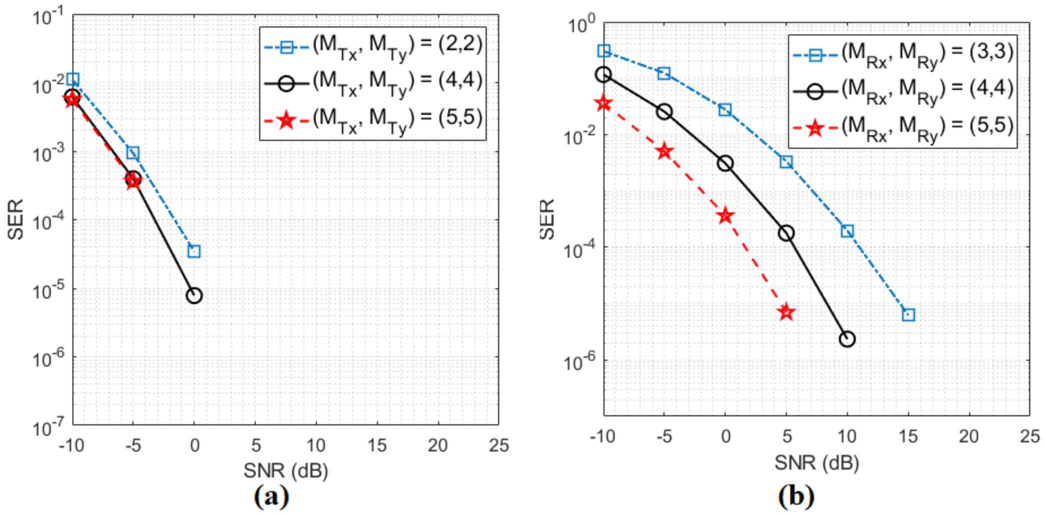
Impacto dos	**Parametros de design**	**Taxa de transmissão**	**Figuras**
Número de fluxos de dados	$(M_{T_x}, M_{T_y}) = (4{,}4), (M_{R_x}, M_{R_y}) = (5{,}5), K = 2,$ $J_1 = J_2 = 2, P = 3, R_1 = R_2 \in [2, 4, 6]$	$T_R = 2; 4.66; 7.33$	Figure 30(a)
Número de símbolos por fluxo de dados	$(M_{T_x}, M_{T_y}) = (4{,}4), (M_{R_x}, M_{R_y}) = (5{,}5), K = 2,$ $R_1 = R_2 = 2, P = 2, J_1 = J_2 \in [2, 4, 6]$	$T_R = 3; 1.75; 1.22$	Figure 30(b)
Número de antenas de transmissão	$(M_{R_x}, M_{R_y}) = (6{,}6)$, $K = 2$, $R_1 = R_2 = 2$, $J_1 = J_2 = 4$, $P = 2, (M_{T_x}, M_{T_y}) \in [(2{,}2), (4{,}4), (5{,}5)]$	$T_R = 1.75$	Figure 31(a)
Número de antenas de recepção	$(M_{T_x}, M_{T_y}) = (2{,}2)$, $K = 2$, $R_1 = R_2 = 2$, $J_1 = J_2 = 4$, $P = 2, (M_{R_x}, M_{R_y}) \in [(3{,}3), (4{,}4), (5{,}5)]$	$T_R = 1.75$	Figure 31(b)
Número de blocos de tempo	$(M_{T_x}, M_{T_y}) = (3{,}3), (M_{R_x}, M_{R_y}) = (4{,}4), K = 2,$ $R_1 = R_2 = 4, J_1 = J_2 = 2, P \in [2, 4, 6];$	$T_R = 7; 3.5; 2.33$	Figure 32(a)
m-QAM	$(M_{T_x}, M_{T_y}) = (3{,}3), (M_{R_x}, M_{R_y}) = (4{,}4), K = 2,$ $R_1 = R_2 = 2, J_1 = J_2 = 4, P = 2, m \in [4, 8, 16, 32]$	$T_R = 0.87; 1.31; 1.75; 2.18$	Figure 32(b)
Número de matrizes de símbolos	$\mathbf{N = 1}$: $(M_{T_x}, M_{T_y}) = (7{,}7); (M_{R_x}, M_{R_y}) = (8{,}8); K = 2, P = 16, J_1 = 4; R_1 = 9$ $\mathbf{N = 2}$: $J_1 = J_2 = 4; R_1 = R_2 = 4$ $\mathbf{N = 3}$: $J_1 = J_2 = 4, J_3 = 1; R_1 = 4, R_2 = 2, R_3 = 9;$ $\mathbf{N = 5}$: $J_1 = J_2 = J_3 = J_4 = 2, J_5 = 1; R_1 = R_2 = R_3 = R_4 = 4, R_5 = 3$	$T_R = 0.46$	Figure 33(a)
SER$_{global}$, **SER of S**$^{(1)}$ **and S**$^{(2)}$	$(M_{T_x}, M_{T_y}) = (4{,}4), (M_{R_x}, M_{R_y}) = (5{,}5), K =$ $P = 2, R_1 = R_2 = 2, J_1 = 2, J_2 = 8$	$T_R = 2.25$	Figure 33(b)
comparação dos receptores	$(M_{T_x}, M_{T_y}) = (2{,}2), (M_{R_x}, M_{R_y}) = (3{,}3), K = 3, R_1 = R_2 = 2, J_1 = J_2 = 4, P = 34$	$T_R = 0.109$	Figures 34 - 37

Table 26 – Parametros de design para as simulações.



(a) (b)

Figure 30 – Impacto do: (a) número de fluxos de dados e (b) número de símbolos por fluxo de dados.



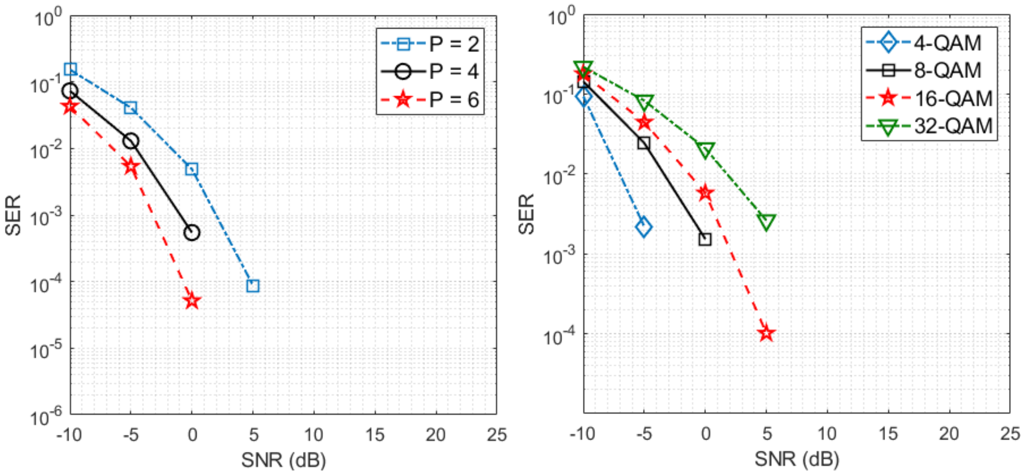
(a) (b)

Figure 31 – Impacto do: (a) número de antenas de transmissão e (b) número de antenas de recepção.

que J_1 e J_2 ão dimensões do tensor de dados, o que não é o caso de R_1 e R_2. Por outro lado, a taxa de transmissão diminui conforme mostrado na Tabela 26.

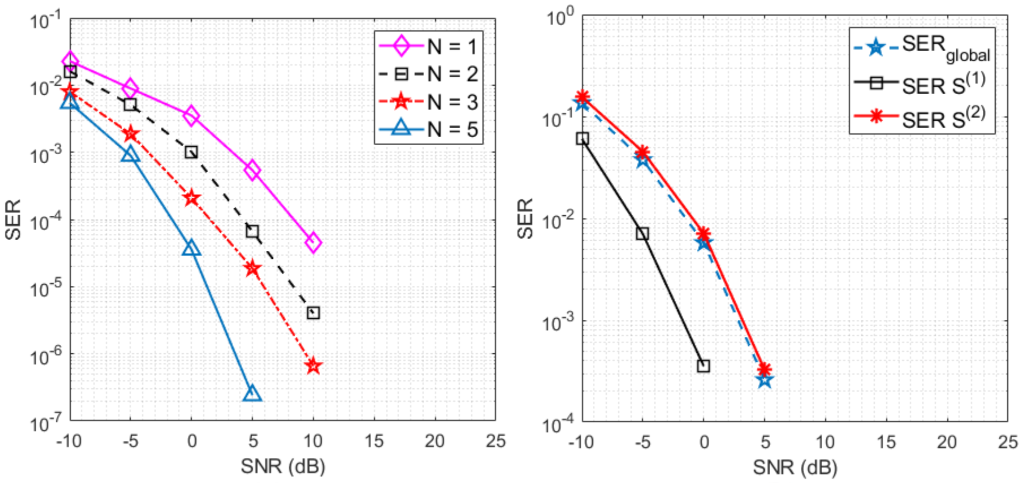
Figuras 31 (a) e 31 (b) illustra o impacto do número de antenas de transmissão e recepçao, com: $(M_{T_x}, M_{T_y}) \in [(2{,}2),(4{,}4),(5{,}5)]$ e $(M_{R_x}, M_{R_y}) \in [(3{,}3),(4{,}4),(5{,}5)]$, respectivamente. Quando o número de antenas de recepção aumenta, a qualidade da estimatição dos símbolos melhora, devido a um aumento da diversidade espacial, através do aumento do número de sinais recebidos. O impacto do número de antenas de transmissão é menos importante que o das antenas de recepção. Observe que a SER é nula para SNR $>$ 0 dB. Observe também que a taxa de transmissão não varia com os números da antena.

Figura 32 (a) apresenta a SER para diferentes números de blocos de tempo: $P \in [2,4,6]$. Quando esse número aumenta, a diversidade de tempo do sistema aumenta e, consequentemente, a SER melhora. Por outro lado, a taxa de transmissão diminui. Figura 32 (b) compara a SER obtida com quatro modulações m-QAM diferentes, $m \in [4{,}8{,}16{,}32]$.



(a) **(b)**

Figure 32 – Impacto do: (a) número de blocos de tempo e (b) modulação (m-QAM).



(a) **(b)**

Figure 33 – (a) Impacto de N na SER e (b) SER das matrizes de símbolos individuais.

Como esperado, o desempenho da SER é melhor quando a modulação 4-QAM é usada, porque 4-QAM é mais fácil de demodular do que as outras modulações, enquanto induz uma taxa de transmissão mais baixa.

Na Figura 33 (a), comparamos a codificação MSMKron para diferentes números de matrizes de símbolos: $N \in [1,2,3,5]$. No caso em que $N = 1$, ou seja, quando somente uma matriz de símbolos é transmitida , então a codificação MSMKron se reduz a codificação TST. A partir desta figura, concluímos que um aumento de N implica uma melhoria significativa da SER. O melhor desempenho é obtido com $N = 5$, que fornece um ganho de 10dB para uma SER de 10^{-3}, em comparação com $N = 1$. Esses resultados corroboram o ganho de codificação fornecido pela codificação MSMKron devido aos múltiplos produtos de kKonecker das matrizes de símbolos, o que induz um aumento de redundância quando N aumenta.

Na Figura 33 (b), a SER$_{global}$ é comparado com as SERs individuais das matrizes de símbolos $\mathbf{S}^{(1)}$ e $\mathbf{S}^{(2)}$, considerando $J_1 = 2$ e $J_2 = 8$. O produto de Kronecker de $\mathbf{S}^{(1)}$ e $\mathbf{S}^{(2)}$ nduz uma redundância maior para $\mathbf{S}^{(1)}$ do que para $\mathbf{S}^{(2)}$, uma vez que cada símbolo de $\mathbf{S}^{(1)}$é repetido $8R_2$ vezes enquanto cada símbolo de $\mathbf{S}^{(2)}$ é repetido somente $2R_1$ vezes. Como esperado, os resultados das simulações mostram que a melhor SER individual é obtido para $\mathbf{S}^{(1)}$ com a menor dimensão J_1 em relação a J_2, devido a uma maior redundância fornecida por $\mathbf{S}^{(2)}$ no produto de Kronecker. A SER global está próxima da SER individual de $\mathbf{S}^{(2)}$.

Comparação dos receptores semi-cegos propostos

Nos próximos experimentos, comparamos primeiro as SERs obtidas com os receptores semi-cegos propostos e os receptores ZF-KronF (usando Eq. (A.41)), como mostrado no estágio 1 da Figura 29. ambas as versões correspondentes às Eqs.(5.4)-(5.5) e Eqs.(5.6)-(5.7) são consideradas. Em seguida, comparamos a performance dos receptores semi-cegos , em termos de estimação de canal e reconstrução do sinal através da NMSE. Para essas simulações, os parâmetros de design têm os seguintes valores: $(M_{T_x}, M_{T_y}) = (2{,}2)$, $(M_{R_x}, M_{R_y}) = (3{,}3)$, $K = 3, J_1 = J_2 = 4, R_1 = R_2 = 2, P = 34$.

A partir das Figuras 34 e 35,podemos concluir que o receptor THOSVD fornece melhor desempenho que os receptorers Bi-ALS-KronF. Isso se deve à forma fechada do THOSVD, permitindo estimar conjuntamente as matrizes de canal e símbolo, enquanto os receptores Bi-ALS-KronF são compostos de duas etapas, sendo uma iterativa para estimar $(\mathbf{H}, \mathbf{S})$, e uma de forma fechada para separar as matrizes de símbolosa partir de $\mathbf{S}$ estimado na primeira etapa. Observe também que o receptor simplificado Bi-ALS-KronF (5.6)-(5.7) estima as matrizes um pouco melhor que a outra versão (5.4)-(5.5) devido à simplificação do cálculo da pseudo-inversa. Como esperado, o receptor ZF-KronF fornece a melhor SER devido ao conhecimento a priori do canal.

Estimativa de parâmetros de multi-percurso

Nesta seção, avaliamos o desempenho do segundo estágio dos receptores propostos na Figura 29, para estimar os parâmetros de multi-percurso do canal cujos valores considerados nas simulações são fornecidos na Tabela 26. O desempenho do MSE para a estimativa de ângulos obtidos com o método JIRAFE é comparado com o fornecido pelo algoritmo ALS retificado, que resulta do seguinte problema de otimização:

$$\min_{\mathbf{A}_{T_x}, \mathbf{A}_{T_y}, \mathbf{A}_{R_x}, \mathbf{A}_{R_y}, \mathbf{B}} \| \mathcal{H} - \mathcal{I}_{5,K} \times_1 \mathbf{A}_{T_x} \times_2 \mathbf{A}_{T_y} \times_3 \mathbf{A}_{R_x} \times_4 \mathbf{A}_{R_y} \times_5 \mathbf{B} \|_F^2 \,. \tag{A.50}$$

Também é feita uma comparação do método JIRAFE sem retificação e com os métodos de retificação TR$_1$A and TR$_1$A$_{imp}$. Observe que o canal estimado usado no estágio 2 é o obtido com o melhor algoritmo no estágio 1, isto é, com o receptor THOSVD.

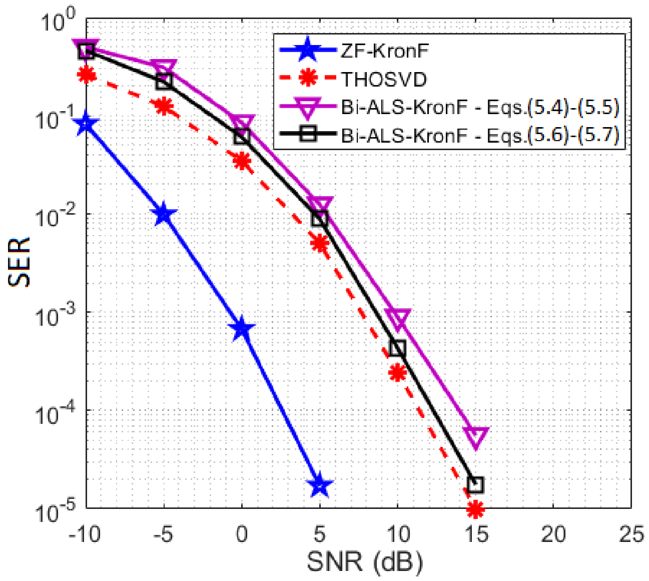


Figure 34 – Comparação da SER com receptores THOSVD, Bi-ALS-KronF e ZF-KronF.

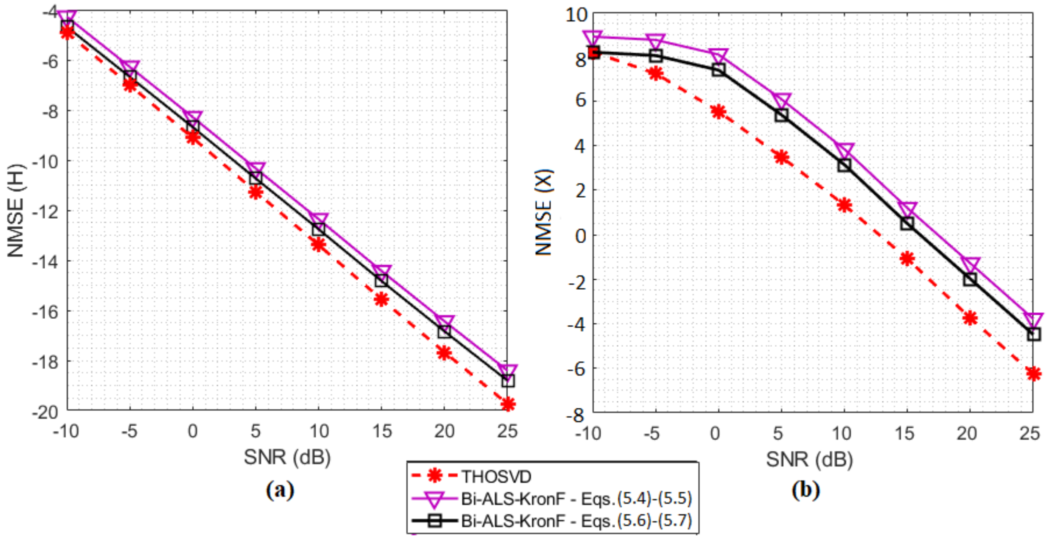


Figure 35 – (a) Comparação de NMSE do canal e (b) Comparação de NMSE de sinal reconstruído.

Figura 36 apresenta a MSE para os ângulos estimados, enquanto Figuras 37 plotam a NMSE para reconstrução da matriz de perdas de caminho **B** e o canal reconstruído. Para a Figura 36, podemos concluir que o método JiIRAFE combinado com a estratégia de retificação TR1A$_{imp}$ supera os outros algoritmos. Além disso, as três variantes JIRAFE (com e sem retificação) fornecem MSEs melhores do que os algoritmos ALS e Rectals. Isso pode ser explicado pela propriedade de redução de ruído do SVD truncado nas etapas do TTHSVD, o que torna o método JIRAFE mais robusto ao ruído. Outra causa é que os métodos ALS e rectALS são baseados em um algoritmo iterativo em cinco etapas, enquanto o algoritmo JIRAFE usa apenas um algoritmo ALS de duas etapas.

Pelas mesmas razões, as Figuras 37 ilustram a superioridade das três variantes do JIRAFE sobre os algoritmos ALS e rectALS em termos da NMSE da matriz de perda de

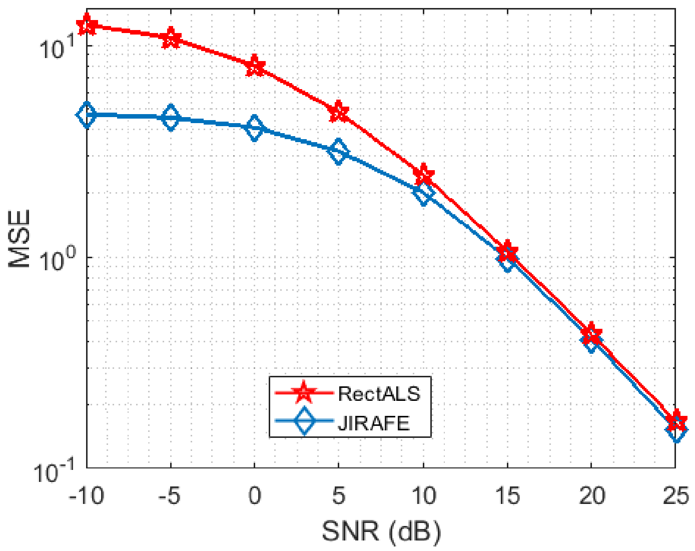


Figure 36 – Comparação da MSE para ângulos estimados.

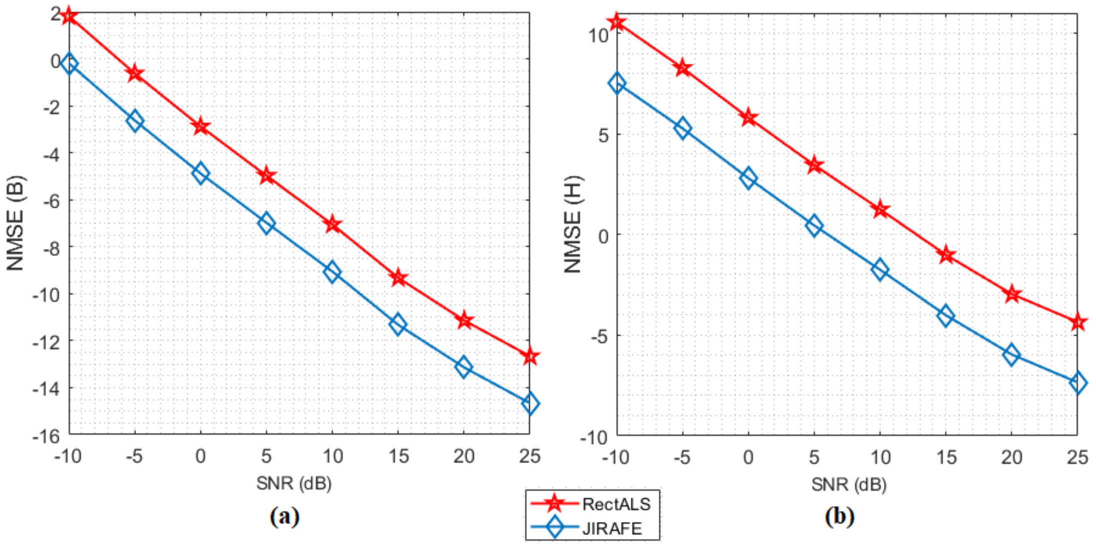


Figure 37 – (a) NMSEs estimados de perda de caminho e (b) NMSEs do canal reconstruído.

caminho e do canal reconstruído, com a melhor performance fornecida por o método JIRAFE combinado com a estratégia de retificação TR1A$_{imp}$. Para Figura 36, também podemos concluir que levar em consideração a estrutura de Vandermonde na estimativa das matrizes de direção (o método rectALS e JIRAFE com retificação) permite melhorar significativamente sua estimativa em comparação com a ALS padrão e o método JIRAFE sem retificação, induzindo uma melhor estimativa do canal como mostrado na Figura 37 (b).

Conclusão

Nesta tese, apresentamos um novo modelo do sinal recebido que considera um caso específico da codificação MKronST combinada com a codificação TST e o modelo do tensor de canal de quinta ordem para transmitir os símbolos, chamada TST-MSMkron.

Essa codificação não requer uma matriz de pré-codificação e nos permite propor receptores semi-cegos em duas etapas para estimar conjuntamente as matrizes de símbolos, canal e parâmetros do canal. Com base na codificação TST-MSMKron e no tensor do canal, os sinais recebidos foram modelados como uma nova decomposição chamada $(N+2)$-ordem decomposição Tucker-PARAFAC aninhada acoplada, onde o tensor do núcleo é a codificação TST. O acoplamento é devido à codificação que é comum aos V_r- e H_r- antenas de recepção polarizadas. A estrutura aninhada resulta do fato de que uma matriz desdobra o tensor do canal constitui um fator matricial do modelo Tucker, os outros fatores sendo as matrizes de direção e perda de caminho. As condições de identificação foram estabelecidas para o novo modelo. Sob a suposição de que a codificação do tensor é conhecida na recepção, esse modelo do sistema é essencialmente único sob condições suaves.

Ao explorar o modelo tensorial dos sinais recebidos e o conhecimento da codificação TST, apresentamos receptores semi-cegos em duas etapas para a estimativa conjunta das matrizes de símbolos, canal e parâmetros do canal (ângulos DOD e DOA, coeficientes de perda de caminho). Em comparação com o sistema supervisionado em [28], que requer o uso de uma sequência piloto para estimar o canal, os receptores propostos precisam apenas de um conhecimento a priori de um símbolo de cada matriz de símbolos. As condições de identificabilidade dos parâmetros e a complexidade computacional para cada receptor foram estabelecidas e comparadas. Na primeira etapa, as matrizes e o canal de símbolos são estimados por meio dos algoritmos Bi-ALS-KronF ou THOSVD e, na segunda etapa, o canal estimado é usado para estimar os parâmetros do canal (ângulos DOD e DOA e coeficientes das perda de caminho) por meio do algoritmo JIRAFE, que explora uma equivalência entre pARAFAC e decomposições TT. Um novo método de retificação foi proposto para garantir a estrutura de Vandermonde das matrizes de direção. Os resultados da simulação mostraram que o receptor semi-cego é eficiente para estimar os símbolos, o canal e os parâmetros do canal e a grande flexibilidade da codificação TST-MSMKron.