UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA

PAULO RICARDO BARBOZA GOMES

TENSOR METHODS FOR ARRAY PROCESSING AND CHANNEL ESTIMATION IN WIRELESS COMMUNICATIONS SYSTEMS

FORTALEZA

2018

PAULO RICARDO BARBOZA GOMES

TENSOR METHODS FOR ARRAY PROCESSING AND CHANNEL ESTIMATION IN WIRELESS COMMUNICATIONS SYSTEMS

Tese apresentada ao Programa de Pós-graduação em Engenharia de Teleinformática do Departamento de Engenharia de Teleinformática da Universidade Federal do Ceará, como parte dos requisitos necessários para a obtenção do título de Doutor em Engenharia de Teleinformática. Área de concentração: Sinais e Sistemas. Linha de pesquisa: Processamento de Sinais e Imagens.

Orientador: Prof. Dr. André Lima Férrer de Almeida (UFC)
Coorientador: Prof. Dr. João Paulo Carvalho Lustosa da Costa (UnB)

FORTALEZA
2018

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

G616t Gomes, Paulo Ricardo Barboza.
Tensor methods for array processing and channel estimation in wireless communications systems / Paulo Ricardo Barboza Gomes. – 2018.
136 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2018.
Orientação: Prof. Dr. André Lima Férrer de Almeida.
Coorientação: Prof. Dr. João Paulo Carvalho Lustosa da Costa.

1. Tensor Modeling. 2. Array Processing. 3. Wireless Communications Systems. 4. Parameters Estimation. 5. Channel Estimation. I. Título.

CDD 621.38

PAULO RICARDO BARBOZA GOMES

TENSOR METHODS FOR ARRAY PROCESSING AND CHANNEL ESTIMATION IN WIRELESS COMMUNICATIONS SYSTEMS

Tese apresentada ao Programa de Pós-graduação em Engenharia de Teleinformática do Departamento de Engenharia de Teleinformática da Universidade Federal do Ceará, como parte dos requisitos necessários para a obtenção do título de Doutor em Engenharia de Teleinformática. Área de concentração: Sinais e Sistemas. Linha de pesquisa: Processamento de Sinais e Imagens.

Aprovada em: 21 / 09 / 2018.

BANCA EXAMINADORA

Prof. Dr. André Lima Férrer de Almeida (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. João Paulo Carvalho Lustosa da Costa (Coorientador)
Universidade de Brasília (UnB)

Prof. Dr. João Cesar Moura Mota
Universidade Federal do Ceará (UFC)

Prof. Dr. Walter da Cruz Freitas Júnior
Universidade Federal do Ceará (UFC)

Prof. Dr. Marcello Luiz Rodrigues de Campos
Universidade Federal do Rio de Janeiro (UFRJ)

Prof. Dr. Gérard Favier
Université de Nice Sophia Antipolis (UNS)
Research Director (CNRS)

This thesis is dedicated to my beloved family, specially my grandmother Nenê, my parents Paulo and Nazaré, my brothers Ruan and Matheus, my wife Flávia Ingrid, and all professors and friends. Every letter of this thesis carries a piece of you.

AGRADECIMENTOS

A Deus, pelo dom da vida, saúde e graças concedidas. Aos meus pais, Paulo e Nazaré, pelo carinho, dedicação, paciência, suporte e por todas as renúncias e esforços dedicados à minha educação. Aos meus irmãos, Ruan e Matheus, pela amizade e apoio ininterruptos. À minha avó, Nenê, e tias Expedita, Aparecida e Liduina, por cuidarem de mim como filho. À minha esposa, Flávia Ingrid, por me acompanhar desde a escola e entender durante esse longo tempo todas as renúncias com paciência, motivação e carinho. Serei eternamente grato. Amo vocês!

Ao meu orientador Prof. Dr. André Lima Férrer de Almeida, e ao meu coorientador Prof. Dr. João Paulo Carvalho Lustosa da Costa, pela excelente orientação, lições, paciência, disponibilidade, oportunidades e amizade durante esses longos anos. A humildade e entusiasmo de vocês me inspiram a ser um profissional de excelência. Muito obrigado por me guiarem nesta conquista!

À toda a equipe do Grupo de Pesquisa em Telecomunicações sem Fio (GTEL), especialmente aos professores Dr. Francisco Rodrigo Porto Cavalcanti, Dr. João Cesar Moura Mota e Dr. Tarcisio Ferreira Maciel, pelos ensinamentos e oportunidades profissionais a mim concedidas.

À todos os amigos, especialmente, Fernando Filho, Joyce, Régis Wendel, Vládia, Lucas, Karine, Bernardinho, Bruno Régis, Régis Lima, Ana Vanessa, Ana Lara e Albano. Vocês são uma extensão da minha família!

À UFC, GTEL, FUNCAP, Ericsson Telecomunicações S/A e LG Eletronics do Brasil, pela estrutura e suporte financeiro durante o período de doutorado.

À todas as pessoas que, de alguma maneira, contribuíram direta ou indiretamente para a realização deste trabalho. Os meus sinceros agradecimentos!

Paulo Ricardo Barboza Gomes
Fortaleza, 21 de setembro de 2018

Senhor, fazei de mim um instrumento da vossa paz.
Onde há ódio, que eu leve o amor.
Onde há ofensa, que eu leve o perdão.
Onde há discórdia, que eu leve a união.
Onde há dúvida, que eu leve a fé.
Onde há erro, que eu leve a verdade.
Onde há desespero, que eu leve a esperança.
Onde há tristeza, que eu leve a alegria.
Onde há trevas, que eu leve a luz.
Ó Mestre,
Fazei que eu procure mais
consolar que ser consolado
compreender que ser compreendido;
amar que ser amado.
Pois é dando que se recebe,
é perdoando que se é perdoado,
é morrendo que se vive para a vida eterna.

"Quem a tudo renuncia, tudo receberá."
São Francisco de Assis

RESUMO

Em diversas aplicações no campo de processamento digital de sinais, como por exemplo, comunicações sem-fio, sonar e radar, o sinal recebido possui natureza multidimensional que pode incluir intrinsecamente em sua estrutura dimensões como espaço, tempo, frequência, código e polarização. Em virtude disso, técnicas modernas de processamento que exploram as múltiplas dimensões do sinal podem ser desenvolvidas para melhorar o desempenho desses sistemas devido à estimativas de parâmetros mais acuradas (por exemplo: direção de partida, direção de chegada, atraso, frequência Doppler, coeficientes de canal, ruído de fase) apresentando melhores condições de identificabilidade. Nesse contexto, esta tese propõe novas modelagens tensoriais para processamento de sinais em arranjos e estimação de canal aplicada à sistemas de comunicações sem-fio. Na primeira parte desta tese, dedicada à processamento de sinais em arranjos multidimensionais de sensores e radar, propomos uma nova técnica de pré-processamento tensorial para supressão de ruído que reduz significantemente o efeito do ruído em dados matriciais e tensoriais implicando em melhores estimativas dos parâmetros desejados. Em seguida, novas modelagens tensoriais baseadas nas decomposições PARAFAC, Tucker e Nested-PARAFAC são formuladas, a partir das quais novos algoritmos para estimação conjunta de ângulo de partida e ângulo de chegada são propostos. Na segunda parte deste documento, modelagens tensoriais são desenvolvidas para resolver o problema de estimação de canal em sistemas de comunicações MIMO sem-fio. Primeiramente, propomos um esquema de codificação e retransmissão multi-frequencial que concentra o processamento associado à estimação conjunta dos canais de *downlink* e *uplink* na estação-base. Mostramos que o sinal retransmitido recebido pode ser modelado como a decomposição PARAFAC de um tensor de terceira-ordem. Em seguida, a decomposição PARAFAC é novamente explorada na modelagem de um sistema de comunicação MIMO mais realista que considera perturbações de ruídos de fase em cada antena transmissora e receptora. Algoritmos receptores para estimação de canal e ruído de fase são formulados. Resultados de simulação são apresentados para ilustrar o desempenho dos receptores propostos que são comparados ao estado-da-arte.

Palavras-chave: Modelagem tensorial, processamento em arranjos, sistemas de comunicações sem-fio, arranjos multidimensionais de sensores, radar MIMO biestático, supressão de ruído, estimação de parâmetros, estimação de canal, estimação de ruído de fase, PARAFAC, Nested-PARAFAC, decomposição de Tucker.

ABSTRACT

In several applications in the field of digital signal processing, for example, wireless communications, sonar and radar, the received signal has a multidimensional nature which can intrinsically include on its structure many dimensions such as space, time, frequency, code, and polarization. In view of this, modern processing techniques which exploit all the signal dimensions can be developed to improve the system performance due to more accurate parameter estimation (for example: direction of departure, direction of arrival, delay, Doppler frequency, channel coefficients, phase noise) with powerful identifiability conditions. In this context, this thesis proposes new tensor modeling approaches for array processing and channel estimation applied to wireless communications systems. In the first part of this thesis, devoted to multidimensional sensor array and radar processing, we propose a new tensor-based preprocessing technique for noise supression which significantly reduces the noise effect in matrix and tensor data leading to more accurate estimates of the desired parameters. Then, new tensor methods capitalizing on the PARAFAC, Tucker and Nested-PARAFAC decompositions are formulated, from which new algorithms for joint direction of departure and direction of arrival estimation are proposed. In the second part of this document, tensor modeling approaches are developed to solve channel estimation problems in MIMO wireless communications systems. Firstly, we propose a new closed-loop and multi-frequency channel training framework that concentrates the processing associated with joint downlink and uplink channel estimation at the base station. We also show that the received closed-loop signal can be modeled as the PARAFAC decomposition of a third-order tensor. Then, the PARAFAC decomposition is also exploited to modeling a more realistic MIMO communication system that considers phase noise perturbations at each transmit and receive antenna. Receiver algorithms for channel and phase noise estimation are formulated. Simulation results are presented to illustrate the performance of the proposed receivers which are compared to state-of-the-art approaches.

Keywords: Tensor modeling, array processing, wireless communications systems, multidimensional sensor arrays, bistatic MIMO radar, denoising, parameters estimation, channel estimation, phase noise estimation, PARAFAC, Nested-PARAFAC, Tucker decomposition.

LIST OF FIGURES

Figure 1 – Schematic of the parallelized procedure for the LS-KRF computation. . . 32
Figure 2 – Illustration of a rank-1 third-order tensor. 33
Figure 3 – Illustration of the horizontal, lateral and frontal slices of a third-order tensor. 34
Figure 4 – Construction process of the 1-mode, 2-mode and 3-mode unfolding matrices of a third-order tensor from its frontal, horizontal and lateral slices. 35
Figure 5 – PARAFAC decomposition of $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ into Q components. 36
Figure 6 – Tucker decomposition of a third-order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$. 42
Figure 7 – Illustration of spatial smoothing and sensors in common for $l_r = 2$ subarrays in the r-th spatial dimension. 50
Figure 8 – $\text{SNR}_{\text{output}}$ vs. $\text{SNR}_{\text{input}}$ for a third-order tensor to compare the SNR gain of MuDe and the classical single HOSVD based low-rank approximation proposed by (HAARDT, ROEMER, and DEL GALDO, 2008). In the simulated scenario: URA, $N_1 = 30$, $N_2 = 30$, $M = 6$ and $K = 10$. 54
Figure 9 – Median of the total RMSE of the spatial frequencies vs. SNR (dB) for a third-order tensor to compare the spatial parameter estimation accuracy. In the scenario: URA, $N_1 = 30$, $N_2 = 30$, $M = 6$ and $K = 10$. 55
Figure 10 – Median of the total RMSE of the spatial frequencies vs. SNR (dB) for the matrix-based approach. In the scenario: ULA, $N = 30$, $M = 6$ and $K = 30$. 56
Figure 11 – Median of the total RMSE of the spatial frequencies vs. SNR (dB) for the matrix-based approach. In the scenario: ULA, $N = 30$, $M = 1$ and $K = 30$. 57
Figure 12 – Median of the total RMSE of the spatial frequencies vs. angular spacing (in degrees) between the sources. In the scenario: ULA, $N = 30$, $M = 5$, $K = 30$ and SNR equal to 0 dB. 57
Figure 13 – Mean processing time (in seconds) vs. number of sensors. In the scenario: ULA, single-source case, $K = 30$ and SNR equal to 30 dB. 58
Figure 14 – L-shaped array configuration with $N_1 + N_2 - 1$ sensors. The distance between the sensors in the z axis is $d^{(1)}$ while the distance between the sensors in the x axis is $d^{(2)}$. The m-th wavefront has elevation and azimuth angles equal to α_m and β_m, respectively. 63
Figure 15 – Total RMSE vs. SNR for $N = 64$ sensors, $K = 10$ samples, DoAs: $\{30°, 55°\}$ and $\{45°, 60°\}$ for Hadamard sources' sequences. Uniform retangular array (URA). 67

Figure 16 –Total RMSE vs. SNR for $N = 64$ sensors, $K = 10$ samples, DoAs: $\{30°, 55°\}$ and $\{45°, 60°\}$ for BPSK sources' sequences. Uniform retangular array (URA). 68
Figure 17 –Convergence of the iterative ALS-Tucker and ALS-ProKRaft algorithms. Uniform retangular array (URA). 68
Figure 18 –Total RMSE vs. SNR for $N = 13$ sensors, $K = 500$ samples, DoAs: $\{30°, 45°\}$ and $\{50°, 55°\}$. L-Shaped array. 69
Figure 19 –Total RMSE vs. SNR (performance of the ALS-Tucker algorithm for different number os sensors). L-Shaped array. 70
Figure 20 –Total RMSE vs. number of samples. L-Shaped array. 70
Figure 21 –Bistatic MIMO radar scenario. 74
Figure 22 –Total RMSE (deg) vs. positioning error factor (α). 83
Figure 23 –Number of iterations for convergence vs. SNR (dB). 84
Figure 24 –Mean processing time (in seconds) vs. data block size. 85
Figure 25 –**Conventional training framework.** The DL and UL channel estimation problems are solved independently. The BS first transmits pilot signals. Then, the DL channel is estimated at the MSs side. The estimated DL channel is feed back to the BS via dedicated uplink resources. The UL channel is estimated at the BS side. The blue words refer to DL communication, while red words refer to UL communication. 90
Figure 26 –**Proposed closed-loop and multi-frequency based training framework.** The DL and UL channels are jointly estimated. The BS first transmits pilot signals. The MSs encode the received pilots and then feed them back to the BS. The BS jointly estimates the DL and UL channels from (120). The blue words refer to DL communication, while red words refer to UL communication. 90
Figure 27 –NMSE (Downlink Channel) vs. SNR (dB) for $N = 32$, $T = 16$, $F = 8$, $U = 8$. The SNR values in DL and UL are assumed to be equal. 98
Figure 28 –NMSE (Uplink Channel) vs. SNR (dB) for $N = 32$, $T = 16$, $F = 8$, $U = 8$. The SNR values in DL and UL are assumed to be equal. 98
Figure 29 –NMSE (DL and UL Channels) vs. number of subcarriers (F) for $N = 32$, $T = 16$, $U = 4$ and SNR = 10 dB. 99
Figure 30 –NMSE (DL and UL Channels) vs. length of the training sequence (T) for $N = 32$, $F = 4$, $U = 4$ and SNR = 10 dB. 100
Figure 31 –Number of iterations for convergence of the proposed ALS-PARAFAC receiver vs. SNR (dB) for $N = 32$, $T = 16$ and $U = 4$. 100
Figure 32 –NMSE vs. number of transmission beams (P) for the DL channel estimation, $U = 2$, $T = 16$ and $F = 25$. 107

Figure 33 – NMSE vs. number of reception beams (Q) for the UL channel estimation, $U = 2$, $T = 16$ and $F = 25$. 107
Figure 34 – NMSE vs. number of subcarriers (F) for the DL channel estimation, $U = 2$ and $T = 16$. 108
Figure 35 – NMSE vs. number of subcarriers (F) for the UL channel estimation, $U = 2$ and $T = 16$. 108
Figure 36 – NMSE vs. length of the pilot signal (T) for the DL channel estimation, $U = 2$, $P = 8$ and $F = 25$. 109
Figure 37 – NMSE vs. length of the pilot signal (T) for the UL channel estimation, $U = 2$, $P = 8$ and $F = 25$. 109
Figure 38 – Frequency-selective MIMO system equipped with different oscilators at the M transmit and N receive antennas. 113
Figure 39 – Illustration of frame and sub-frame structures. The pilot symbols part $\boldsymbol{S}_P$ is reused sub-frame to sub-frame. The PN is invariant within a sub-frame while the channel gains remain constant over the length of one frame[1]. 113
Figure 40 – NMSE of $\hat{\boldsymbol{H}}$ vs. SNR (dB) of a 4×4 MIMO system for different pilot length (L_P), $L_D = 100$ and BPSK modulation. 121
Figure 41 – NMSE of $\hat{\boldsymbol{\Phi}}^{[t]}$ vs. SNR (dB) of a 4×4 MIMO system for different pilot length (L_P), $L_D = 100$ and BPSK modulation. 121
Figure 42 – NMSE of $\hat{\boldsymbol{\Phi}}^{[r]}$ vs. SNR (dB) of a 4×4 MIMO system for different pilot length (L_P), $L_D = 100$ and BPSK modulation. 122
Figure 43 – BER vs. SNR (dB) of a 4×4 MIMO system with perfect and imperfect channel and PN knowledge for different pilot length (L_P), $L_D = 100$ and BPSK modulation. 122
Figure 44 – Number of iterations for convergence vs. SNR (dB) of a 4×4 MIMO system for different pilot length (L_P), $L_D = 100$ and BPSK modulation. . 123

LIST OF TABLES

Table 1 – Complexity analysis of different tensor-based algorithms. 82

ABBREVIATIONS AND ACRONYMS

ALS	Alternating Least Squares
AoA	Angle of Arrival
AoD	Angle of Departure
AWGN	Additive White Gaussian Noise
BALS	Bilinear Alternating Least Squares
BER	Bit Error Rate
BPSK	Binary Phase Shift Keying
BS	Base Station
CANDECOMP	CANonical DECOMPosition
CP	CANDECOMP/PARAFAC
CRB	Cramer-Rao Bound
CS	Compressed Sensing
CSI	Channel State Information
DALS	Double Alternating Least Squares
dB	Decibel
DFT	Discrete Fourier Transform
DL	Downlink
DoA	Direction of Arrival
DoD	Direction of Departure
DS-CDMA	Direct-Sequence Code-Division Multiple Access
ESPRIT	Estimation of Signal Parameters via Rotational Invariance Technique
FBA	Forward-Backward Averaging
FDD	Frequency Division Duplexing
FISTA	Fast Iterative Shrinkage-Thresholding
FLOP	Floating Point Operation
FR	Frequency Range
GNSS	Global Navigation Satellite System
HB	Hybrid Analog-Digital Beamforming
HOSVD	Higher Order Singular Value Decomposition
i.i.d.	Independent and Identically Distributed Random Variables
LS	Least-Squares
LS-KRF	Least-Squares Khatri-Rao Factorization
MIMO	Multiple-Input Multiple-Output
MMSE	Minimum Mean Square Error
mmWave	Millimeter Wave
MOS	Model Order Selection
MS	Mobile Station

MSE	Mean-Squared Error
MuDe	Multiple Denoising
MU-MIMO	Multiuser MIMO
MUSIC	Multiple Signal Classification
NMSE	Normalized Mean Square Error
NR	New Radio
NULA	Non-Uniform Linear Array
OFDM	Orthogonal Frequency Division Multiplexing
OMP	Orthogonal Matching Pursuit
OPA	Outer Product based Array
OPP	Orthogonal Procrustes Problem
PARAFAC	PARAlell FACtor
PM	Propagator Method
PN	Phase Noise
ProKRaft	Procrustes Estimation and Khatri-Rao Factorization
R-D	Multidimensional Sensor Array
RF	Radio Frequency
RMSE	Root Mean Square Error
S-CosaMP	Structured Compressive Sampling Matching Pursuit
SISO	Single-Input Single-Output
SNR	Signal-to-Noise Ratio
SS	Spatial Smoothing
STF	Space-Time-Frequency
SVD	Singular Value Decomposition
TALS	Trilinear Alternating Least Squares
TDD	Time Division Duplexing
UL	Uplink
ULA	Uniform Linear Array
URA	Uniform Retangular Array

NOTATION

Throughout this thesis the following conventions are used. Scalars are denoted by lower-case letters x, column vectors as boldface lower-case letters $\boldsymbol{x}$, matrices as boldface upper-case letters $\boldsymbol{X}$, and tensors as boldface calligraphic letters $\boldsymbol{\mathcal{X}}$.

$\mathbb{C}$	set of complex-valued numbers
$\mathbb{C}^{I}$	set of complex-valued I-dimensional vectors
$\mathbb{C}^{I\times Q}$	set of complex-valued $(I \times Q)$-matrices
$\mathbb{C}^{I_1\times I_2\times\cdots\times I_N}$	set of complex-valued $(I_1 \times I_2 \times \cdots \times I_N)$-tensors
$\{\cdot\}^{*}$	complex conjugate
$\{\cdot\}^{\mathrm{T}}$	transpose
$\{\cdot\}^{\mathrm{H}}$	conjugate transpose
$\{\cdot\}^{-1}$	inverse
$\{\cdot\}^{\dagger}$	Moore-Penrose pseudo-inverse
$\|\cdot\|_{\mathrm{F}}$	Frobenius norm
$\otimes$	Kronecker product
$\diamond$	Khatri-Rao (column-wise Kronecker) product
$\circ$	outer product
$x_{i,q}$	(i,q)-th element of $\boldsymbol{X} \in \mathbb{C}^{I\times Q}$
$x_{i_1,i_2,\cdots,i_N}$	$(i_1,i_2,\cdots,i_N)$-th element of $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1\times I_2\times\cdots\times I_N}$
$\boldsymbol{x}_q$	q-th column of $\boldsymbol{X} \in \mathbb{C}^{I\times Q}$
$\boldsymbol{X}(i,:)$	i-th row of $\boldsymbol{X} \in \mathbb{C}^{I\times Q}$
$\boldsymbol{X}(i:k,:),\ i<k$	submatrix fomed by the i-th to j-th row of $\boldsymbol{X}$
$\boldsymbol{X}_{i_1\cdot\cdot}$	i_1-th horizontal slice of $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1\times I_2\times I_3}$
$\boldsymbol{X}_{\cdot i_2\cdot}$	i_2-th lateral slice of $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1\times I_2\times I_3}$
$\boldsymbol{X}_{\cdot\cdot i_3}$	i_3-th frontal slice of $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1\times I_2\times I_3}$
$[\boldsymbol{\mathcal{X}}]_{(n)}$	n-mode unfolding matrix of $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1\times I_2\times\cdots\times I_N}$
$\times_n$	n-mode product
$\boldsymbol{I}_N$	identity matrix of size $N \times N$
$\boldsymbol{\mathcal{I}}_{N,Q}$	identity tensor of N-th order and rank-Q
$\rho_{\boldsymbol{x}}$	rank of $\boldsymbol{X}$
$\kappa_{\boldsymbol{x}}$	Kruskal-rank (k-rank) of $\boldsymbol{X}$
$\mathbb{E}\{\cdot\}$	expectation operator
$\sqcup_n$	concatenation operator along n-th mode of $\boldsymbol{\mathcal{X}}$
$\mathrm{vec}\,(\cdot)$	vectorization operator
$\mathrm{unvec}\,(\cdot)$	unvectorization operator (inverse of the vectorization operator)
$\mathrm{vecd}\,(\cdot)$	converts the main diagonal of a matrix into a column vector
$D_i\,(\boldsymbol{A})$	diagonal matrix with diagonal entries given by i-th row of $\boldsymbol{A}$

TABLE OF CONTENTS

1	**INTRODUCTION**	21
1.1	**Thesis Scope and Motivation**	21
1.2	**Main Contributions**	23
1.3	**Thesis Organization**	23
1.4	**Scientific Production**	25
1.4.1	**Participation in Research Projects**	25
1.4.2	**Journal Papers**	26
1.4.3	**Conference Papers**	26
1.4.4	**Patent**	27
2	**PREREQUISITES OF MULTILINEAR ALGEBRA**	28
2.1	**Matrix Products and Operators**	28
2.2	**Matrix Factorizations**	29
2.2.1	**Singular Value Decomposition (SVD)**	30
2.2.2	**Least-Squares Khatri-Rao Factorization (LS-KRF)**	30
2.3	**Tensor Definitions and Operations**	32
2.4	**Tensor Decompositions**	35
2.4.1	**PARAFAC Decomposition**	36
2.4.2	**Nested-PARAFAC Decomposition**	40
2.4.3	**Tucker Decomposition**	40
2.5	**Chapter Summary**	44
3	**TENSOR-BASED METHODS FOR BLIND SPATIAL SIGNATURE ESTIMATION IN MULTIDIMENSIONAL SENSOR ARRAYS**	45
3.1	**Introduction and Motivation**	45
3.1.1	**Chapter Organization**	47
3.2	**Data Model**	47
3.3	**Spatial Smoothing**	49

3.4 Proposed Multiple Denoising via Successive Spatial Smoothing, Low-Rank Approximation and Reconstruction 49
3.4.1 Tensor-based Approach (R-D sensor array case) 49
3.4.2 Computational Complexity of MuDe 52
3.4.3 Matrix-based Approach (1-D sensor array case) 53
3.5 Performance Evaluation of MuDe 53
3.6 Proposed Tensor-Based Methods for Blind Spatial Signature Estimation 57
3.6.1 Covariance Tensor Formulation 58
3.6.2 ALS-Tucker Algorithm 59
3.6.3 ALS-ProKRaft Algorithm 61
3.7 Spatial Signature Estimation in L-Shaped Sensor Arrays 63
3.7.1 Cross-Correlation Tensor for L-shaped Sensor Arrays 64
3.7.2 Estimation of the Spatial Frequencies 65
3.8 Computational Complexity 66
3.9 Advantages and Disadvantages of the Proposed Methods . . . 66
3.10 Performance Evaluation of ALS-Tucker and ALS-ProKraft Algorithms 67
3.11 Chapter Summary 71
4 A NESTED-PARAFAC BASED APPROACH FOR TARGET LOCALIZATION IN BISTATIC MIMO RADAR SYSTEMS . 72
4.1 Introduction and Motivation 72
4.1.1 Chapter Organization 73
4.2 Signal Model 74
4.3 Proposed Tensor-Based Method for Joint DoD and DoA Estimation 75
4.3.1 Proposed Nested-PARAFAC Based Modeling 77
4.3.2 First ALS Stage (Trilinear ALS) 78
4.3.3 Second ALS Stage (Bilinear ALS) 79

4.3.4 DoD and DoA Parameters Estimation . . . 79
4.3.5 Alternative Closed-Form Solution to the Second Stage . . . 80
4.4 Uniqueness Issues . . . 81
4.5 Computational Complexity . . . 82
4.6 Simulation Results . . . 83
4.7 Chapter Summary . . . 85
5 JOINT DOWNLINK AND UPLINK CHANNEL ESTIMATION USING TENSOR PROCESSING . . . 86
5.1 Introduction and Motivation . . . 86
5.1.1 Chapter Organization . . . 88
5.2 System Model . . . 88
5.2.1 Downlink Signal Model . . . 88
5.2.2 Uplink Signal Model . . . 88
5.2.3 Conventional x Proposed Channel Training Framework . . . 89
5.3 Proposed Tensor-Based Semi-Blind Receivers for Joint DL and UL Channel Estimation . . . 91
5.3.1 PARAFAC-Based Modeling . . . 91
5.3.2 ALS-PARAFAC Receiver . . . 92
5.3.3 LS-KRF Receiver . . . 93
5.4 Identifiability and Computational Complexity . . . 95
5.4.1 Identifiability Conditions . . . 95
5.4.2 Computational complexity . . . 96
5.5 Simulation Results (Part 1) . . . 97
5.6 Extension to mmWave Massive MIMO Systems . . . 101
5.6.1 Sparse Formulation for DL and UL Channel Parameters Estimation . . . 103
5.6.2 Identifiability Condition . . . 105
5.7 Simulation Results (Part 2) . . . 106
5.8 Chapter Summary . . . 110

6 TENSOR-BASED SEMI-BLIND RECEIVER FOR JOINT CHANNEL AND PHASE NOISE ESTIMATION IN FREQUENCY-SELECTIVE MIMO SYSTEMS 111
6.1 Introduction and Motivation 111
6.1.1 Chapter Organization 112
6.2 System Model 113
6.3 Proposed PARAFAC-Based Modeling 114
6.4 Proposed Semi-Blind Receiver for Joint Channel and PN Estimation 116
6.4.1 Channel Equalization 118
6.5 Identifiability and Computational Complexity 119
6.5.1 Identifiability Conditions 119
6.5.2 Computational Complexity 120
6.6 Simulation Results 120
6.7 Chapter Summary 123
7 CONCLUSIONS AND PERSPECTIVES 124
REFERENCES 126

1 INTRODUCTION

This is an introductory chapter where we present the motivation and scope of this thesis in Section 1.1. *After that, we present our main contributions in Section* 1.2, *followed by the thesis overview in Section* 1.3. *Finally, the main scientific production produced during the doctoral period are presented in Section* 1.4.

1.1 Thesis Scope and Motivation

High resolution parameter estimation and channel estimation problems play fundamental roles in several practical applications in the field of digital signal processing. For example, the first one ranging from array processing in radar, sonar, acoustics and global navigation satellite systems (GNSS), to name a few, while the latter is commonly applied to mobile communications systems. Conventional signal processing techniques usually consider only the space and time dimensions which leads to matrix-based modeling. However, often the space domain can be split into two signal dimensions (azimuth and elevation), while the time domain can be divided into other two dimensions (frames and sub-frames) so that the received signal to be processed has a multidimensional nature which can also include on its structure other dimensions (for example: frequency, code and polarization) that are not take into account since matrices are only two-dimensional data structures. In view of this, modern processing techniques that extend the existing approaches by exploiting all the signal dimensions can be developed by using tensor tools. There are significant advantages of using tensor-based signal processing instead of matrix-based signal processing. Among these advantages, we can cite the improved identifiability conditions, which generally come from the essential uniqueness property of tensor decompositions. Tensor-based methods also inherits the so-called *tensor gain*, which translates into improved accuracy due to the efficient noise rejection capability (DA COSTA *et al.*, 2011; ROEMER, 2013).

In order to deal with such a multidimensional nature of the signals, tensor decompositions have extensively been applied in recent years in array processing and wireless communications systems areas. In the array processing context, the seminal works (SIDIROPOULOS, BRO, and GIANNAKIS, 2000) and (NION and SIDIROPOULOS, 2010) linked the Parallel Factor (PARAFAC) decomposition (CARROLL and CHANG, 1970; HARSHMAN, 1970) to the problems of multiple invariance sensor array processing and targets localization in multiple-input multiple-output (MIMO) radar systems, respectively. The authors of (GUO, MIRON, and BRIE, 2008) show that a polarized source mixture on a vector sensor array can be interpreted as a PARAFAC decomposition of a third-order tensor. In (DA COSTA *et al.*, 2010), a closed-form solution to estimate the factor matrices of a PARAFAC decomposition has been proposed for spatial parameter

estimation in multidimensional sensor arrays. An interesting transmission scheme based on power variations of the transmitted signals at successive time blocks was proposed in (RONG *et al.*, 2005). It solve the problem of spatial signature estimation using two different approaches: the first one is based on the PARAFAC decomposition, while the second one relies a joint approximate diagonalization algorithm.

In the context of wireless communications systems, the authors of the seminal work (SIDIROPOULOS, GIANNAKIS, and BRO, 2000) proposed a blind multiuser separation for direct-sequence code-division multiple access (DS-CDMA) by modeling the received signal as a third-order PARAFAC decomposition. DE ALMEIDA, FAVIER, and MOTA (2007) proposes a unified PARAFAC-based modeling with application to blind multiuser equalization. In works (LIU *et al.*, 2013) and (DE ALMEIDA *et al.*, 2013), different tensor-based receivers are formulated to solving the joint symbol and channel estimation problem in space-time-frequency (STF) MIMO communication systems. In the former, the PARAFAC decomposition is exploited to derive a closed-form solution based on the factorization of the Khatri-Rao product. The latter proposed a blind receiver that exploits a generalized PARATUCK2 model (HARSHMAN and LUNDY, 1996) of the STF-MIMO transmission system.

Nowadays, millimeter wave (mmWave) MIMO communication systems has been subject of increasing interest in both academia and industry since it is a promising technology for future 5G mobile communication systems due to its potential to offer gigabit-per-second data rates by exploiting the large bandwidth available at mmWave frequencies (ZHOU *et al.*, 2016a). This has motived the development of new tensor-based approaches for mmWave MIMO channel estimation. In this way, the authors of (ZHOU *et al.*, 2016a) proposed a layered pilot transmission scheme and explored the intrinsic low-rank structure of the received signal, resulting of the sparse scattering nature of the mmWave channel, to formulate a PARAFAC-based method for the joint estimation of the uplink channels of multiple users assuming channel reciprocity. In (ZHOU *et al.*, 2017), the same authors studied the downlink channel estimation problem for frequency-selective mmWave MIMO channels. Similarly, the received signal is organized into a third-order tensor which admits a PARAFAC decomposition and the channel parameters (angles of departure and arrival, time delays and fading coefficients) are jointly estimated from the decomposed factor matrices. In contrast to (ZHOU *et al.*, 2017), the authors of (ARAÚJO and DE ALMEIDA, 2017) explore both the sparse and multidimensional structures of frequency-selective mmWave MIMO channels and recast the channel estimation problem as a multi-way compressed sensing problem. The channel parameters are estimated by solving a simpler compressive sensing problem for each channel dimension.

The key features motivating the use of tensor decompositions in the aforementioned applications come from their powerful identifiability and uniqueness properties compared with traditional matrix-based methods (HARSHMAN, 1970; KOLDA and

BADER, 2009). Additionally, one can also benefit from the multidimensional structure of tensor data to perform noise rejection/prewhitening, as shown in (DA COSTA *et al.*, 2011).

1.2 Main Contributions

In this thesis, we propose new tensor-based modeling and algorithms to solve parameter estimation problems in array processing, as well as channel and phase noise estimation problems in MIMO wireless communications systems. The main contributions can be summarized as follows:

- Development of a new denoising framework for matrix and tensor data to improve the parameter estimation accuracy in multidimensional sensor array processing;
- New tensor-based formulations for the covariance structure of the received signal in multidimensional sensor arrays and bistatic MIMO radar systems exploiting the PARAFAC, Nested-PARAFAC and Tucker decompositions;
- Development of new PARAFAC, Nested-PARAFAC and Tucker-based receiver algorithms for direction of departure and direction of arrival estimation in multidimensional sensor arrays and bistatic MIMO radar systems;
- Proposal of a new PARAFAC-based channel training framework and development of receiver algorithms which concentrate the joint downlink and uplink channel estimation at the BS side in MIMO wireless communications systems ;
- Development of a new PARAFAC-based approach for the joint channel and phase noise estimation in MIMO communication systems;
- Study of the identifiability issues and computational complexities of the proposed methods.

1.3 Thesis Organization

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

Chapter 2: *Prerequisites of Multilinear Algebra.* This chapter provides a theoretical basis for the methods developed in this thesis. It first review important definitions and operations of multilinear algebra. Then, the three most important tensor decompositions to the context of this thesis, namely PARAFAC, Nested-PARAFAC and Tucker decompositions are introduced. The contribution presented in this chapter is to unify in one place fundamental concepts spread in the literature in order to make it easily understandable for the reader.

The first part of this thesis, which comprises Chapters 3 and 4, is devoted to solve the problem of spatial parameter estimation in multidimensional sensor arrays and bistatic MIMO radar systems, respectively.

Chapter 3: *Tensor-Based Methods for Blind Spatial Signatures Estimation in Multidimensional Sensor Arrays.* This chapter tackles the denoising and spatial parameter estimation in array signal processing problems. In the first part, by exploiting the possible subarrays created by the spatial smoothing preprocessing technique and then applying successive low-rank approximations, we formulate a novel denoising framework for tensor and matrix data. In the second part, the covariance matrix of the received signal is formulated as a PARAFAC model (for uncorrelated sources) and as a Tucker model (for correlated sources). Two generalized iterative algorithms for direction of arrival estimation are proposed.

Chapter 4: *A Nested-PARAFAC Approach for Target Localization in Bistatic MIMO Radar Systems.* This chapter addresses the target localization problem in bistatic MIMO radar systems. We formulate a new tensor modeling for the cross-covariance matrix of the matched filters outputs using a Nested-PARAFAC decomposition, from which two receiver algorithms are proposed to jointly estimate the directions of departure and arrival of multiple targets located at the same range bin of interest. Identifiability issues and computational complexity of the proposed algorithms are also discussed.

The second part of this thesis, which comprises Chapters 5 and 6, is devoted to solve the problem of channel and phase noise estimation in MIMO wireless communications systems.

Chapter 5: *Joint Downlink and Uplink Channel Estimation Using Tensor Processing.* This chapter addresses the problem of channel estimation in MIMO wireless communications systems. Initially, a novel closed-loop and multi-frequency based channel training framework is presented. It allows to concentrate most of the processing burden for channel estimation at the base station, i.e., avoiding such processing at the mobile stations side. We formulate a new PARAFAC-based modeling for the received closed-loop signal. We also extend the proposed tensor signal model for millimeter-wave MIMO scenarios. Two new semi-blind receivers to perform the joint downlink and uplink channel estimation are developed, and identifiability issues and computational complexity are studied.

Chapter 6: *Tensor-Based Semi-Blind Receiver for Joint Channel and Phase Noise Estimation in Frequency-Selective MIMO Systems.* This chapter addresses the problem of channel and phase noise estimation using a PARAFAC-based modeling. By assuming a transmission scheme in which each frame is divided into small sub-frames so that the phase noise varies sub-frame to sub-frame, we show that the received signal can be modeled as a third-order PARAFAC decomposition. We propose a semi-blind receiver for

jointly recovering the channel impulse response and the phase noise at the transmitter and receiver antennas. The identifiability issues and the computational complexity of the proposed receiver are also discussed.

Chapter 7: This chapter concludes the thesis by summarizing the obtained results and listing some perspectives for future work.

Remark: This thesis has four chapters with the main contributions that cover different aspects of array signal processing and channel estimation in MIMO wireless communications systems. Every chapter is meant to be self-contained so that the reader can read them independently without loss of information.

1.4 Scientific Production

The experiences in research projects as well as the technical reports and publications produced during the doctoral period are listed as follows.

1.4.1 Participation in Research Projects

- *Speech Processing for Far Field Automatic Speech Recognition.* March/2015 - February/2016. Developed under the context of LG Electronics/UFC technical cooperation projects;
- UFC.44 *Hybrid Beamforming and Massive MIMO for 5G Wireless Systems.* October/2016 - September 2018. Developed under the context of Ericsson/UFC technical cooperation projects,

in which a number of four technical reports in UFC.44 were delivered:

- **Paulo R. B. Gomes**, and André L. F. de Almeida, *Joint DL and UL Channel Estimation Using Tensor Processing.* First Technical Report (TR01) UFC.44. GTEL-UFC, April, 2017;
- **Paulo R. B. Gomes**, and André L. F. de Almeida, *Joint DL and UL Channel Estimation for Massive MIMO Communications Using Hybrid Beamforming.* Second Technical Report (TR02) UFC.44. GTEL-UFC, October, 2017;
- **Paulo R. B. Gomes**, and André L. F. de Almeida, *Multidimensional User Grouping for Frequency-Selective MU-MIMO Systems.* Third Technical Report (TR03) UFC.44. GTEL-UFC, April, 2018.
- **Paulo R. B. Gomes**, Francisco H. C. Neto, Weskley V. F. Maurício, Daniel C. Araújo, André L. F. de Almeida, and Tarcisio F. Maciel, *Hybrid Beamforming Design Exploiting Implicit and Explicit CSI in Massive MIMO Systems.* Fourth Technical Report (TR04) UFC.44. GTEL-UFC, September, 2018.

The Chapter 5 of this thesis is related to the UFC.44 research project.

1.4.2 Journal Papers

- **Paulo R. B. Gomes**, André L. F. de Almeida, João Paulo C. L. da Costa, and G. Del Galdo, *Tensor-Based Methods for Blind Spatial Signature Estimation Under Arbitrary and Unknown Source Covariance Structure*, in Digital Signal Processing, vol. 62, pp. 197-210, 2017;
- **Paulo R. B. Gomes**, André L. F. de Almeida, João Paulo C. L. da Costa, João C. M. Mota, Daniel Valle de Lima, and G. Del Galdo, *Tensor-Based Methods for Blind Spatial Signature Estimation in Multidimensional Sensor Arrays*, in International Journal of Antennas and Propagation, vol. 2017, pp. 1-11, 2017;
- **Paulo R. B. Gomes**, João Paulo C. L. da Costa, and André L. F. de Almeida, *Tensor-Based Multiple Denoising via Successive Spatial Smoothing, Low-Rank Approximation and Reconstruction for R-D Sensor Array Processing*, in Digital Signal Processing, May 2018. *Under Major Revision*;
- **Paulo R. B. Gomes**, André L. F. de Almeida, and João Paulo C. L. da Costa, *A Nested-PARAFAC Based Approach for Target Localization in Bistatic MIMO Radar Systems*, in Digital Signal Processing, April, 2018. *Under Major Revision*;
- **Paulo R. B. Gomes**, André L. F. de Almeida, João Paulo C. L. da Costa, and Rafael T. de Sousa Jr., *Tensor-Based Semi-Blind Receiver for Joint Channel and Phase Noise Estimation in Frequency-Selective MIMO Systems*, *To be Submitted*;
- **Paulo R. B. Gomes**, and André L. F. de Almeida, *Joint DL and UL Channel Estimation for Millimeter Wave MIMO Communications Using Tensor Based Processing*, *To be Submitted*;
- **Paulo R. B. Gomes**, André L. F. de Almeida, and Tarcisio F. Maciel, *Multidimensional User Grouping for Users Scheduling in Frequency-Selective MIMO Systems*, *In Preparation.*

1.4.3 Conference Papers

- **Paulo R. B. Gomes**, André L. F. de Almeida, and João C. M. Mota, *Estimação Cega de Assinaturas Espaciais para Arranjos em Formato L Baseada em Modelagem Tensorial de Correlações Cruzadas*, in XXXIII Brazilian Symposium on Telecommunications (SBrT2015), September, 2015, Juiz de Fora, MG.
- **Paulo R. B. Gomes**, André L. F. de Almeida, João Paulo C. L. da Costa, and João C. M. Mota, *Métodos Tensoriais para Estimação Cega de Assinaturas Espaciais em Arranjos Multidimensionais de Sensores*, in XXXIV Brazilian Symposium on Telecommunications (SBrT2016), September, 2016, Santarém, PA.

1.4.4 Patent

A request for patent application related to the UFC.44 project is under evaluation at Ericsson research.

- **Paulo R. B. Gomes**, Daniel C. Araújo, André L. F. de Almeida, and Eleftherios Karipidis, *A Method for Joint Forward and Reverse CSI Acquisition.*

2 PREREQUISITES OF MULTILINEAR ALGEBRA

In this chapter, we focus on review some important basic definitions and operations of multilinear algebra that are exploited in later chapters of this thesis. Our goal here is to contribute with a compilation of fundamental concepts spread over many works in the literature in order to make it easily understandable for the reader. The chapter is divided into four parts. Firstly, we define some matrix operators, and then the Kronecker and Khatri-Rao products are introduced. Secondly, we review the matrix singular value decomposition (SVD) and least-squares Khatri-Rao factorization (LS-KRF). The third and fourth parts are particularly devoted to introduce operations involving tensors and an overview on PARAllel FACtor (PARAFAC), Nested-PARAFAC and Tucker decompositions, respectively. All the contents presented here will be extensively used throughout this thesis in the context of different applications.

2.1 Matrix Products and Operators

The matrix Kronecker and Khatri-Rao products are important operations in multilinear algebra. Often, these two products are employed to represent in a simplified way the unfolding matrices of well-used tensor decompositions. In the following, Kronecker and Khatri-Rao products are described in detail.

Definition 1. (*Kronecker product*) The Kronecker product, denoted by $\otimes$, between two matrices $\boldsymbol{A} \in \mathbb{C}^{I_1 \times Q}$ and $\boldsymbol{B} \in \mathbb{C}^{I_2 \times R}$ results in a matrix of size $I_1 I_2 \times QR$ denoted by

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{bmatrix} a_{1,1}\boldsymbol{B} & a_{1,2}\boldsymbol{B} & \cdots & a_{1,Q}\boldsymbol{B} \\ a_{2,1}\boldsymbol{B} & a_{2,2}\boldsymbol{B} & \cdots & a_{2,Q}\boldsymbol{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{I_1,1}\boldsymbol{B} & a_{I_1,2}\boldsymbol{B} & \cdots & a_{I_1,Q}\boldsymbol{B} \end{bmatrix} \in \mathbb{C}^{I_1 I_2 \times QR}. \tag{1}$$

Definition 2. (*Khatri-Rao product*) The Khatri-Rao product introduced by KHATRI and R. RAO (1968), denoted by $\diamond$, is the column-wise Kronecker product. Let $\boldsymbol{A} \in \mathbb{C}^{I_1 \times Q}$ and $\boldsymbol{B} \in \mathbb{C}^{I_2 \times Q}$ be two matrices with the same number of columns Q, the Khatri-Rao product between them results in a matrix of size $I_1 I_2 \times Q$ denoted by

$$\boldsymbol{A} \diamond \boldsymbol{B} = \begin{bmatrix} \boldsymbol{a}_1 \otimes \boldsymbol{b}_1 & \boldsymbol{a}_2 \otimes \boldsymbol{b}_2 & \cdots & \boldsymbol{a}_Q \otimes \boldsymbol{b}_Q \end{bmatrix} \in \mathbb{C}^{I_1 I_2 \times Q}, \tag{2}$$

where $\boldsymbol{a}_q \in \mathbb{C}^{I_1}$ and $\boldsymbol{b}_q \in \mathbb{C}^{I_2}$ denote the q-th column ($q = 1, \ldots, Q$) of $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively.

Definition 3. (*Vectorization operator*) Given a matrix $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_Q] \in \mathbb{C}^{I_1 \times Q}$ the vectorization operation, denoted by $\mathrm{vec}\,(\boldsymbol{A})$, converts $\boldsymbol{A}$ to a column vector $\boldsymbol{a} \in \mathbb{C}^{I_1 Q}$ by

stacking its columns $\boldsymbol{a}_q \in \mathbb{C}^{I_1}$ $(q = 1, \ldots, Q)$ on top of each other, i.e.,

$$\operatorname{vec}(\boldsymbol{A}) = \begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{a}_Q \end{bmatrix} \in \mathbb{C}^{I_1 Q}. \tag{3}$$

As an inverse operation of the vectorization, the unvectorization operator denoted by $\operatorname{unvec}_{I_1 \times Q}(\boldsymbol{a})$, reshapes the column vector $\boldsymbol{a} \in \mathbb{C}^{I_1 Q}$ into a matrix $\boldsymbol{A} \in \mathbb{C}^{I_1 \times Q}$.

Throughout this thesis, we shall make use of the following properties involving the Kronecker and Khatri-Rao products (BREWER, 1978; PETERSEN and PEDERSEN, 2008):

$$(\boldsymbol{A} \diamond \boldsymbol{B})^{\mathrm{T}} = \left[D_1(\boldsymbol{A}) \boldsymbol{B}^{\mathrm{T}} \cdots D_{I_1}(\boldsymbol{A}) \boldsymbol{B}^{\mathrm{T}} \right], \tag{4}$$

$$\boldsymbol{AC} \otimes \boldsymbol{BD} = (\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D}), \tag{5}$$

$$\boldsymbol{AC} \diamond \boldsymbol{BD} = (\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \diamond \boldsymbol{D}), \tag{6}$$

$$\operatorname{vec}\left(\boldsymbol{ABC}^{\mathrm{T}}\right) = (\boldsymbol{C} \otimes \boldsymbol{A}) \operatorname{vec}(\boldsymbol{B}), \tag{7}$$

$$\operatorname{vec}\left(\boldsymbol{ABC}^{\mathrm{T}}\right) = (\boldsymbol{C} \diamond \boldsymbol{A}) \operatorname{vecd}(\boldsymbol{B}), \tag{8}$$

$$\boldsymbol{a} \otimes \boldsymbol{b} = \operatorname{vec}(\boldsymbol{b} \circ \boldsymbol{a}), \tag{9}$$

where $\boldsymbol{B}$ is assumed to be a full matrix in (7), while it is a diagonal matrix in (8). The operator $D_{i_1}(\boldsymbol{A})$ forms a diagonal matrix holding the i_1-th row of $\boldsymbol{A} \in \mathbb{C}^{I_1 \times Q}$ $(i_1 = 1, \ldots, I_1)$ on its main diagonal. The operator $\operatorname{vecd}(\boldsymbol{B})$ converts the main diagonal elements of $\boldsymbol{B}$ into a column vector. The symbol "$\circ$"denotes the outer product operator. In each property above the vectors and matrices involved have compatible dimension.

2.2 Matrix Factorizations

In this section, we introduce the matrix singular value decomposition (SVD) and least-squares Khatri-Rao factorization (LS-KRF). The first one is an important concept to define the higher-order singular value decomposition (HOSVD) in the next sections. The LS-KRF is used when the observed data is an estimate of a Khatri-Rao product which we would like to factorize in a closed-form way by solving a set of rank-1 approximation problems.

2.2.1 Singular Value Decomposition (SVD)

The SVD, introduced by BELTRAMI (1873) and JORDAN (1874a,b), decomposes an arbitrary matrix $\boldsymbol{X} \in \mathbb{C}^{I_1 \times I_2}$ of rank-Q as

$$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{H}}, \tag{10}$$

where $\boldsymbol{U} \in \mathbb{C}^{I_1 \times I_1}$ is the unitary matrix of left singular vectors, $\boldsymbol{V} \in \mathbb{C}^{I_2 \times I_2}$ is the unitary matrix of right singular vectors, and $\boldsymbol{\Sigma} \in \mathbb{C}^{I_1 \times I_2}$ is a diagonal matrix that contains the Q non-zero singular values $\sigma_1 \geq \sigma_2 \cdots \geq \sigma_Q$ on its main diagonal and zeros elsewhere. Equation (10) can be also expressed as an economy size SVD notation. The idea is to obtain the same data matrix, without losses, by reducing the number of columns of $\boldsymbol{U}$ and $\boldsymbol{V}$. The economy size SVD of $\boldsymbol{X}$ is defined as

$$\boldsymbol{X} = \boldsymbol{U}_s\boldsymbol{\Sigma}_s\boldsymbol{V}_s^{\mathrm{H}}, \tag{11}$$

where $\boldsymbol{U}_s \in \mathbb{C}^{I_1 \times Q}$ and $\boldsymbol{V}_s \in \mathbb{C}^{I_2 \times Q}$ contain the first Q columns of $\boldsymbol{U}$ and $\boldsymbol{V}$, and $\boldsymbol{\Sigma}_s \in \mathbb{C}^{Q \times Q}$ contains only the first Q non-zero singular values on its main diagonal.

If $\boldsymbol{X}$ is a rank-1 matrix, the optimum low-rank approximation of $\boldsymbol{X}$ is obtained by truncating its SVD to a rank-1 approximation as follows:

$$\boldsymbol{X} = \sigma_1\boldsymbol{u}_1\boldsymbol{v}_1^{\mathrm{H}}, \tag{12}$$

where $\boldsymbol{u}_1 \in \mathbb{C}^{I_1}$ and $\boldsymbol{v}_1 \in \mathbb{C}^{I_2}$ are the dominant left and right singular vectors of $\boldsymbol{U}$ and $\boldsymbol{V}$, and σ_1 is the dominant singular value.

According to (GOLUB and VAN LOAN, 1996) the computational complexity of the full SVD computation of $\boldsymbol{X} \in \mathbb{C}^{I_1 \times I_2}$, in terms of floating point operations (FLOPs) counts, is given by $\mathcal{O}(I_1 \cdot I_2 \cdot \min(I_1, I_2))$. On the other hand, the computational cost of the economy size SVD (truncated to rank-Q) is $\mathcal{O}(I_1 \cdot I_2 \cdot Q)$. The solution of the rank-1 approximation problem (12) requires $\mathcal{O}(I_1 \cdot I_2)$ FLOPs.

2.2.2 Least-Squares Khatri-Rao Factorization (LS-KRF)

In some applications addressed in this thesis, we will be interested in approximating a matrix Khatri-Rao product between two factor matrices. This problem appears repeatedly over the next chapters in the context of our applications. To solve it, we make use of the well known LS-KRF algorithm proposed by KINBANGOU and FAVIER (2009); ROEMER and HAARDT (2010); ROEMER (2013). In the following, we formulate the Khatri-Rao factorization problem. In addition, the LS-KRF algorithm is summarized in the pseudo-code form.

Algorithm 1 Least-squares Khatri-Rao factorization (LS-KRF)

procedure LS-KRF
for $q = 1, \ldots, Q$
    1. *Apply the* $\text{unvec}_{I_2 \times I_1}$ *operator in the q-th column of* $\boldsymbol{X}$ *to obtain the rank-1 matrix* $\boldsymbol{\Psi}_q \in \mathbb{C}^{I_2 \times I_1}$ *in* (14).
    2. *Compute the SVD of* $\boldsymbol{\Psi}_q$ *given by* $\boldsymbol{U}_q \boldsymbol{\Sigma}_q \boldsymbol{V}_q^{\mathrm{H}}$. *Then, obtain the estimates for the q-th column of* $\hat{\boldsymbol{A}}$ *and* $\hat{\boldsymbol{B}}$ *as follows:*

$$\hat{\boldsymbol{a}}_q = \sqrt{\sigma_1}\boldsymbol{v}_1^* \quad \textit{and} \quad \hat{\boldsymbol{b}}_q = \sqrt{\sigma_1}\boldsymbol{u}_1,$$

    where $\boldsymbol{u}_1$ *and* $\boldsymbol{v}_1$ *denote the dominant left and right singular vectors of* $\boldsymbol{U}_q$ *and* $\boldsymbol{V}_q$, *and* σ_1 *represents the dominant singular value of* $\boldsymbol{\Sigma}_q$, *respectively.*
end
The final estimates are given by $\hat{\boldsymbol{A}} = [\hat{\boldsymbol{a}}_1, \cdots, \hat{\boldsymbol{a}}_Q]$ *and* $\hat{\boldsymbol{B}} = [\hat{\boldsymbol{b}}_1, \cdots, \hat{\boldsymbol{b}}_Q]$.
Remove the scaling ambiguities of $\hat{\boldsymbol{A}}$ *and* $\hat{\boldsymbol{B}}$ *through normalization procedure.*

Let $\boldsymbol{X} \in \mathbb{C}^{I_1 I_2 \times Q}$ be an observed data matrix given by the Khatri-Rao product $\boldsymbol{X} \approx \boldsymbol{A} \diamond \boldsymbol{B}$, where $\boldsymbol{A} \in \mathbb{C}^{I_1 \times Q}$ and $\boldsymbol{B} \in \mathbb{C}^{I_2 \times Q}$. We want to obtain the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ from $\boldsymbol{X}$ by solving the following optimization problem:

$$\min_{\boldsymbol{A},\boldsymbol{B}} \|\boldsymbol{X} - \boldsymbol{A} \diamond \boldsymbol{B}\|_{\mathrm{F}}^2 . \tag{13}$$

According to **Definition 2** and using the property (9), the q-th column of $\boldsymbol{X}$ can be interpreted as the vectorized form of the following rank-1 matrix

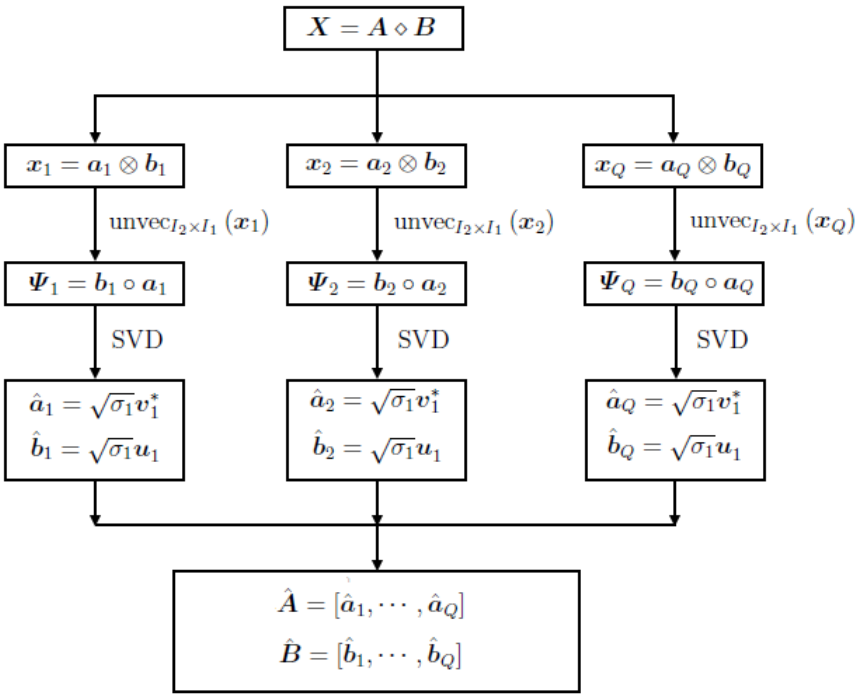
$$\boldsymbol{\Psi}_q = \boldsymbol{b}_q \circ \boldsymbol{a}_q \in \mathbb{C}^{I_2 \times I_1}, \tag{14}$$

where $\boldsymbol{a}_q \in \mathbb{C}^{I_1}$ and $\boldsymbol{b}_q \in \mathbb{C}^{I_2}$ denote the q-th column ($q = 1, \ldots, Q$) of $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively. Therefore, being $\boldsymbol{\Psi}_q \in \mathbb{C}^{I_2 \times I_1}$ a rank-1 matrix the best estimates for the vectors $\boldsymbol{a}_q$ and $\boldsymbol{b}_q$ in the least squares (LS) sense can be obtained by truncating the SVD of $\boldsymbol{\Psi}_q$, defined by $\boldsymbol{U}_q \boldsymbol{\Sigma}_q \boldsymbol{V}_q^{\mathrm{H}}$, to a rank-1 approximation, i.e.,

$$\hat{\boldsymbol{a}}_q = \sqrt{\sigma_1}\boldsymbol{v}_1^* \quad \text{and} \quad \hat{\boldsymbol{b}}_q = \sqrt{\sigma_1}\boldsymbol{u}_1, \tag{15}$$

where $\boldsymbol{u}_1$ and $\boldsymbol{v}_1$ are the dominant left and right singular vectors of $\boldsymbol{U}_q$ and $\boldsymbol{V}_q$, and σ_1 is the dominant singular value of $\boldsymbol{\Sigma}_q$, respectively. Note that the estimated vectors $\hat{\boldsymbol{a}}_q$ and $\hat{\boldsymbol{b}}_q$ are affected by non-zero complex scaling ambiguity, i.e., $\boldsymbol{a}_q \otimes \boldsymbol{b}_q = (\gamma_q \cdot \hat{\boldsymbol{a}}_q) \otimes \left(\frac{1}{\gamma_q} \cdot \hat{\boldsymbol{b}}_q\right)$. However, it can be easily removed through normalization if the first row of $\boldsymbol{A}$ or $\boldsymbol{B}$ is known (ROEMER, 2013). The final estimates $\hat{\boldsymbol{A}} = [\hat{\boldsymbol{a}}_1, \cdots, \hat{\boldsymbol{a}}_Q]$ and $\hat{\boldsymbol{B}} = [\hat{\boldsymbol{b}}_1, \cdots, \hat{\boldsymbol{b}}_Q]$ are obtained by solving this rank-1 approximation problem for each column of $\boldsymbol{X}$. Based on the aforementioned computational complexity of the rank-1 approximation problem solved via SVD, we can observe that the computational cost of the LS-KRF algorithm is

Figure 1 – Schematic of the parallelized procedure for the LS-KRF computation.



Source: Created by the author.

$\mathcal{O}(I_1 \cdot I_2 \cdot Q)$.

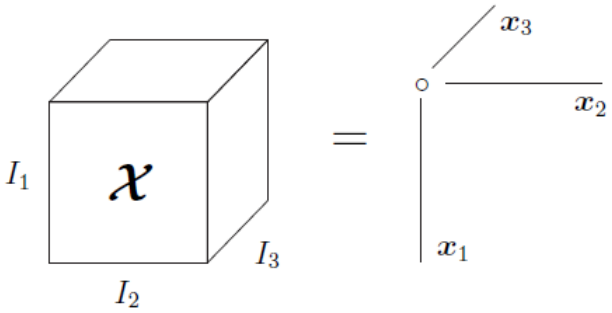
Note that the LS-KRF procedure is computed independently for each column of $\boldsymbol{X}$. Therefore, the columns of the estimated matrices $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}}$ can be obtained in a parallel way if Q processing units are available, reducing the processing time by a factor Q. The parallelized version of the LS-KRF algorithm is illustrated in Figure 1. In addition, the LS-KRF procedure is summarized in the form of pseudo-code in Algorithm 1.

2.3 Tensor Definitions and Operations

The remainder of this chapter is dedicated to presenting the theoric and algorithmic concepts of multilinear (tensor) algebra. We begin by introducing some definitions and operations of multilinear algebra. At the end, we provide an overview on the most important tensor decompositions to be used in this thesis, which are the PARAFAC, Nested-PARAFAC and Tucker decompositions.

Throughout this thesis, an N-th order tensor (or higher-order tensor) $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ with size I_n along mode (or dimension) n ($n = 1, \ldots, N$) and elements $x_{i_1, i_2, \ldots, i_N}$ is interpreted as a multidimensional array of numerical values with dimensionality N. Mathematically, it represents an element of the tensor product between N vector spaces (COMON, 2014). As special cases matrices, vectors and scalars are commoly referred as 2-order, 1-order and 0-order tensors, respectively. Other definitions and operations involving tensors are presented in the following.

Figure 2 – Illustration of a rank-1 third-order tensor.



Source: Adapted from (DE ALMEIDA, 2007).

Definition 4. (*Inner product*) The inner product between two tensors $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}$ both of N-th order, denoted by $\langle \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{Y}} \rangle$, results in a scalar value given by

$$\langle \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{Y}} \rangle = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} x_{i_1,i_2,\ldots,i_N} y_{i_1,i_2,\ldots,i_N}. \tag{16}$$

If $\langle \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{Y}} \rangle = 0$, the tensors $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}$ are mutually orthogonal (DE ALMEIDA, 2007).

Definition 5. (*Frobenius norm*) The Frobenius norm of an N-th order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ is defined as

$$\|\boldsymbol{\mathcal{X}}\|_{\mathrm{F}} = \sqrt{\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} |x_{i_1,i_2,\ldots,i_N}|^2}. \tag{17}$$

The Frobenius norm can be also represented using inner product notation as $\|\boldsymbol{\mathcal{X}}\|_{\mathrm{F}} = \sqrt{\langle \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{X}} \rangle}$. Note that (17) is similar to the matrix Frobenius norm if $N = 2$.

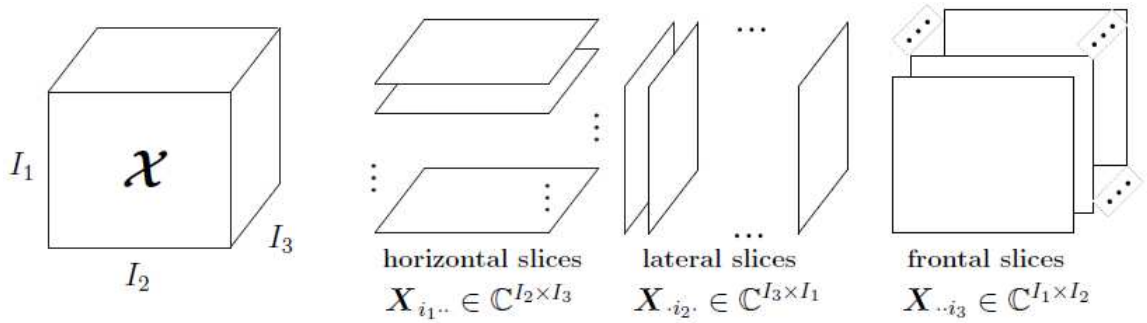
Definition 6. (*Rank-1 tensor*) An N-th order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ is said to be rank-1 when it is computed as the outer product between N vectors, i.e.,

$$\boldsymbol{\mathcal{X}} = \boldsymbol{x}_1 \circ \boldsymbol{x}_2 \circ \cdots \circ \boldsymbol{x}_N, \tag{18}$$

where $\boldsymbol{x}_n \in \mathbb{C}^{I_n}$ $(n = 1, \ldots, N)$ is called component of $\boldsymbol{\mathcal{X}}$ along the n-th mode (or dimension). As a visual example, Figure 2 illustrates a rank-1 third-order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ formed by the outer product of three vectors $\boldsymbol{x}_1 \in \mathbb{C}^{I_1}$, $\boldsymbol{x}_2 \in \mathbb{C}^{I_2}$ and $\boldsymbol{x}_3 \in \mathbb{C}^{I_3}$.

Definition 7. (*Tensor rank*) The rank of an arbitrary tensor is defined as the minimum number of rank-1 components that exactly decomposes additively a tensor (DE ALMEIDA, 2007; RIBEIRO, 2016).

Figure 3 – Illustration of the horizontal, lateral and frontal slices of a third-order tensor.



Source: Adapted from (DE ALMEIDA, 2007).

Definition 8. (*Fibers*) An n-mode fiber represents an one-dimensional section of a tensor analogue of matrix rows and columns. A fiber is obtained by fixing every index but one. An n-mode fiber has size $I_n \times 1$, where I_n denotes the size of the n-th mode of the tensor. Third-order tensors are formed by row, column and tube fibers (KOLDA and BADER, 2009).

Definition 9. (*Slices*) A slice represents a two-dimensional section of a tensor. It is obtained by fixing all but two indices. Third-order tensors are formed by horizontal, lateral and frontal slices as illustrated in Figure 3.

Definition 10. (*n-mode unfolding*) The n-mode unfolding matrix, represented by $[\boldsymbol{\mathcal{X}}]_{(n)}$, is the process of reordering the elements of an N-th order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ into a matrix without loss of information. Similar to (KOLDA and BADER, 2009) and (CICHOCKI *et al.*, 2015), we define $[\boldsymbol{\mathcal{X}}]_{(n)}$ by concatenating side-by-side the slices (or n-mode fibers) of $\boldsymbol{\mathcal{X}}$. Since many computational resources are not able to manipulate tensors of order higher than two, the unfolding matrix concept is important to facilitate the computational manipulation of a given data tensor. Figure 4 illustrates the construction process of the 1-mode $[\boldsymbol{\mathcal{X}}]_{(1)} \in \mathbb{C}^{I_1 \times I_2 I_3}$, 2-mode $[\boldsymbol{\mathcal{X}}]_{(2)} \in \mathbb{C}^{I_2 \times I_1 I_3}$ and 3-mode $[\boldsymbol{\mathcal{X}}]_{(3)} \in \mathbb{C}^{I_3 \times I_1 I_2}$ unfolding matrices of a third-order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ from its frontal, horizontal and lateral slices, respectively.

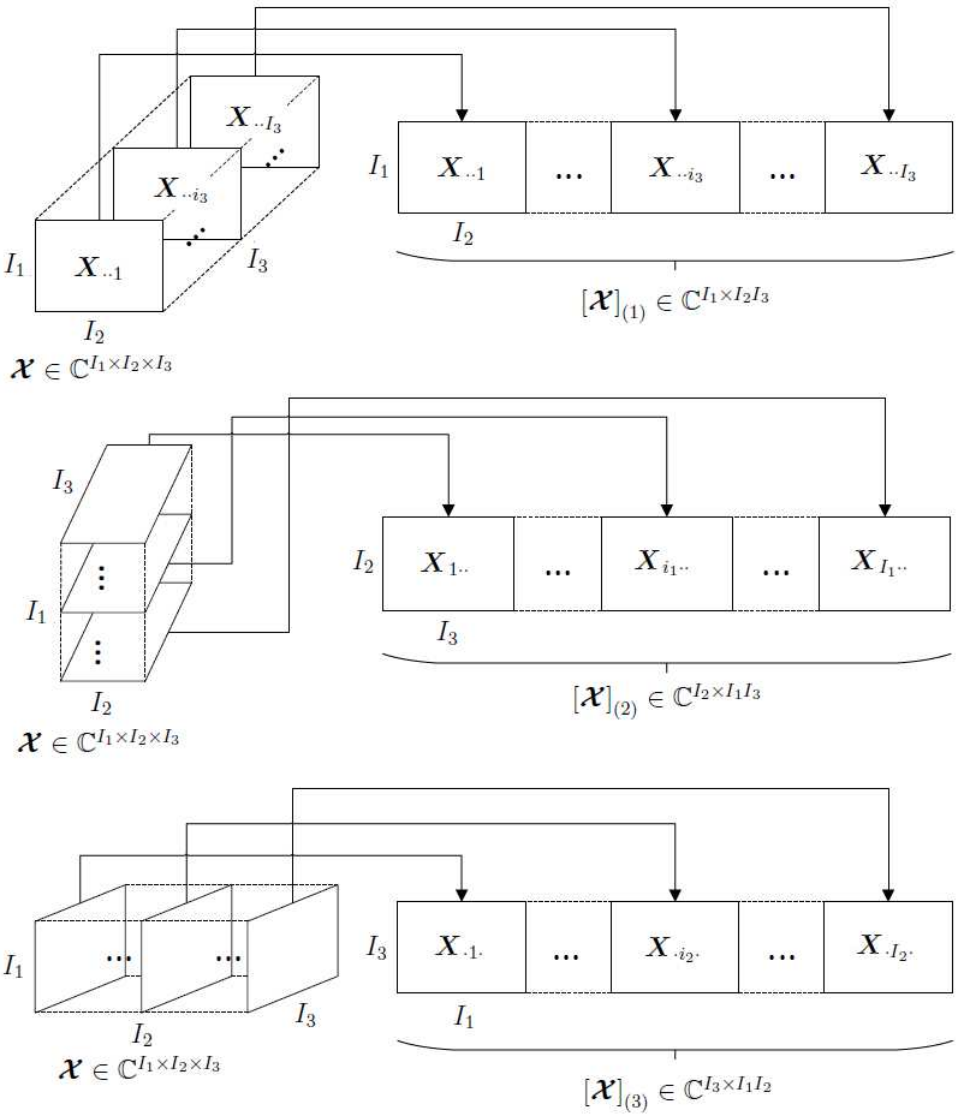
Definition 11. (*n-mode product*) The n-mode product operation, denoted by $\times_n$, consists of multiplying an N-th order tensor $\boldsymbol{X} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ by a matrix $\boldsymbol{A} \in \mathbb{C}^{Q \times I_n}$ along its n-th mode. The result of this operation is

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{X}} \times_n \boldsymbol{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_{n-1} \times Q \times I_{n+1} \times \cdots \times I_N}. \tag{19}$$

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

$$[\boldsymbol{\mathcal{Y}}]_{(n)} = \boldsymbol{A}\,[\boldsymbol{\mathcal{X}}]_{(n)}\,. \tag{20}$$

Figure 4 – Construction process of the 1-mode, 2-mode and 3-mode unfolding matrices of a third-order tensor from its frontal, horizontal and lateral slices.



Source: Adapted from (XIMENES, 2015).

In other words, the n-mode product multiplies all the n-mode fibers of $\boldsymbol{\mathcal{X}}$ with the matrix $\boldsymbol{A}$ from the left hand side. In this work, we shall make use of the following n-mode product properties (KOLDA and BADER, 2009):

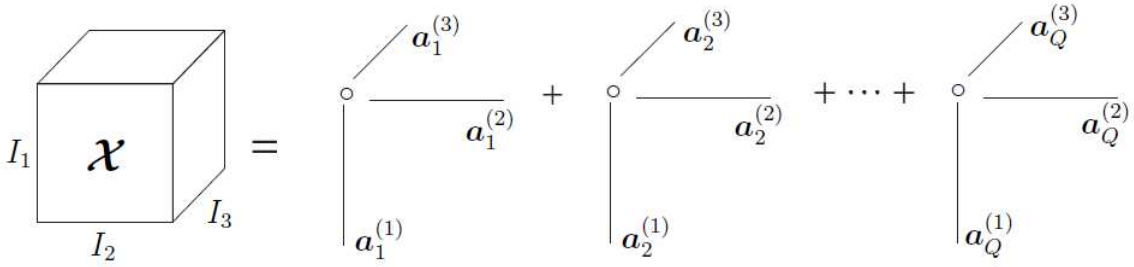
$$\boldsymbol{\mathcal{X}} \times_m \boldsymbol{A} \times_n \boldsymbol{B} = \boldsymbol{\mathcal{X}} \times_n \boldsymbol{B} \times_m \boldsymbol{A} \quad (m \neq n), \tag{21}$$

$$\boldsymbol{\mathcal{X}} \times_n \boldsymbol{A} \times_n \boldsymbol{B} = \boldsymbol{\mathcal{X}} \times_n (\boldsymbol{B}\boldsymbol{A}) . \tag{22}$$

2.4 Tensor Decompositions

This section provides an overview on the main tensor decompositions encountered in the applications investigated throughout this thesis, namely PARAFAC, Nested-PARAFAC and Tucker decompositions.

Figure 5 – PARAFAC decomposition of $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ into Q components.



Source: Adapted from (DE ALMEIDA, 2007).

2.4.1 PARAFAC Decomposition

The PARAllel FACtor (PARAFAC) decomposition or CANonical DECOMPosition (CANDECOMP) has been independently proposed by HARSHMAN (1970) in phonetics and by CARROLL and CHANG (1970) in psychometrics fields. KIERS (2000) suggests the standardization of the terminology by using the abbreviation CANDECOMP/PARAFAC (CP).

Let $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ be a third-order tensor, the PARAFAC decomposition factorizes it into a sum of Q rank-1 component tensors. i.e.,

$$\boldsymbol{\mathcal{X}} = \sum_{q=1}^{Q} \boldsymbol{a}_q^{(1)} \circ \boldsymbol{a}_q^{(2)} \circ \boldsymbol{a}_q^{(3)}, \tag{23}$$

where its (i_1, i_2, i_3)-th element is given by

$$x_{i_1,i_2,i_3} = \sum_{q=1}^{Q} a_{i_1,q}^{(1)} a_{i_2,q}^{(2)} a_{i_3,q}^{(3)}, \tag{24}$$

where Q is called number of factors or rank of $\boldsymbol{\mathcal{X}}$ and it is defined as the minimum number of rank-1 tensors for which (23) holds exactly. $a_{i_1,q}^{(1)}$, $a_{i_2,q}^{(2)}$ and $a_{i_3,q}^{(3)}$ are the elements of the factor matrices $\boldsymbol{A}^{(n)} = \left[\boldsymbol{a}_1^{(n)}, \cdots, \boldsymbol{a}_Q^{(n)}\right] \in \mathbb{C}^{I_n \times Q}$ ($i_n = 1, \ldots, I_n$ and $n = 1, 2, 3$), respectively. Figure 5 illustrates the PARAFAC decomposition of $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ into Q rank-1 components.

The PARAFAC decomposition (23) can be also alternatively represented, in terms of its factor matrices, in a more compact form by using the n-mode product notation as follows

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{I}}_{3,Q} \times_1 \boldsymbol{A}^{(1)} \times_2 \boldsymbol{A}^{(2)} \times_3 \boldsymbol{A}^{(3)}, \tag{25}$$

where $\boldsymbol{\mathcal{I}}_{3,Q}$ represents a third-order identity tensor of size $Q \times Q \times Q$. The elements of $\boldsymbol{\mathcal{I}}_{3,Q}$ are equal to 1 when all indices are equal, and 0 elsewhere.

As previously mentioned in Section (2.3), we can represent $\boldsymbol{\mathcal{X}}$ into three different ways according to **Definition 9**. The horizontal $\boldsymbol{X}_{i_1 \cdot\cdot} \in \mathbb{C}^{I_2 \times I_3}$ $(i_1 = 1, \ldots, I_1)$, lateral $\boldsymbol{X}_{\cdot i_2 \cdot} \in \mathbb{C}^{I_3 \times I_1}$ $(i_2 = 1, \ldots, I_2)$ and frontal $\boldsymbol{X}_{\cdot\cdot i_3} \in \mathbb{C}^{I_1 \times I_2}$ $(i_3 = 1, \ldots, I_3)$ slices of (25) are denoted by

$$\boldsymbol{X}_{i_1 \cdot\cdot} = \sum_{q=1}^{Q} a_{i_1,q}^{(1)} \boldsymbol{a}_q^{(2)} \boldsymbol{a}_q^{(3)^{\mathrm{T}}} = \boldsymbol{A}^{(2)} D_{i_1}\left(\boldsymbol{A}^{(1)}\right) \boldsymbol{A}^{(3)^{\mathrm{T}}}, \tag{26}$$

$$\boldsymbol{X}_{\cdot i_2 \cdot} = \sum_{q=1}^{Q} a_{i_2,q}^{(2)} \boldsymbol{a}_q^{(3)} \boldsymbol{a}_q^{(1)^{\mathrm{T}}} = \boldsymbol{A}^{(3)} D_{i_2}\left(\boldsymbol{A}^{(2)}\right) \boldsymbol{A}^{(1)^{\mathrm{T}}}, \tag{27}$$

$$\boldsymbol{X}_{\cdot\cdot i_3} = \sum_{q=1}^{Q} a_{i_3,q}^{(3)} \boldsymbol{a}_q^{(1)} \boldsymbol{a}_q^{(2)^{\mathrm{T}}} = \boldsymbol{A}^{(1)} D_{i_3}\left(\boldsymbol{A}^{(3)}\right) \boldsymbol{A}^{(2)^{\mathrm{T}}}. \tag{28}$$

According to **Definition 10** and using the property (4) and the expressions (26), (27) and (28), we can express the 1-mode $[\boldsymbol{\mathcal{X}}]_{(1)} \in \mathbb{C}^{I_1 \times I_2 I_3}$, 2-mode $[\boldsymbol{\mathcal{X}}]_{(2)} \in \mathbb{C}^{I_2 \times I_1 I_3}$ and 3-mode $[\boldsymbol{\mathcal{X}}]_{(3)} \in \mathbb{C}^{I_3 \times I_1 I_2}$ unfolding matrices of the PARAFAC decomposition as

$$\begin{aligned}
[\boldsymbol{\mathcal{X}}]_1 = [\boldsymbol{X}_{\cdot\cdot 1} \cdots \boldsymbol{X}_{\cdot\cdot I_3}] &= \boldsymbol{A}^{(1)} \left[D_1\left(\boldsymbol{A}^{(3)}\right) \boldsymbol{A}^{(2)^{\mathrm{T}}} \cdots D_{I_3}\left(\boldsymbol{A}^{(3)}\right) \boldsymbol{A}^{(2)^{\mathrm{T}}} \right] \\
&= \boldsymbol{A}^{(1)} \left(\boldsymbol{A}^{(3)} \diamond \boldsymbol{A}^{(2)} \right)^{\mathrm{T}} \in \mathbb{C}^{I_1 \times I_2 I_3},
\end{aligned} \tag{29}$$

$$\begin{aligned}
[\boldsymbol{\mathcal{X}}]_{(2)} = \left[\boldsymbol{X}_{\cdot\cdot 1}^{\mathrm{T}} \cdots \boldsymbol{X}_{\cdot\cdot I_3}^{\mathrm{T}}\right] &= \boldsymbol{A}^{(2)} \left[D_1\left(\boldsymbol{A}^{(3)}\right) \boldsymbol{A}^{(1)^{\mathrm{T}}} \cdots D_{I_3}\left(\boldsymbol{A}^{(3)}\right) \boldsymbol{A}^{(1)^{\mathrm{T}}} \right] \\
&= \boldsymbol{A}^{(2)} \left(\boldsymbol{A}^{(3)} \diamond \boldsymbol{A}^{(1)} \right)^{\mathrm{T}} \in \mathbb{C}^{I_2 \times I_1 I_3},
\end{aligned} \tag{30}$$

$$\begin{aligned}
[\boldsymbol{\mathcal{X}}]_{(3)} = [\boldsymbol{X}_{\cdot 1 \cdot} \cdots \boldsymbol{X}_{\cdot I_2 \cdot}] &= \boldsymbol{A}^{(3)} \left[D_1\left(\boldsymbol{A}^{(2)}\right) \boldsymbol{A}^{(1)^{\mathrm{T}}} \cdots D_{I_2}\left(\boldsymbol{A}^{(2)}\right) \boldsymbol{A}^{(1)^{\mathrm{T}}} \right] \\
&= \boldsymbol{A}^{(3)} \left(\boldsymbol{A}^{(2)} \diamond \boldsymbol{A}^{(1)} \right)^{\mathrm{T}} \in \mathbb{C}^{I_3 \times I_1 I_2}.
\end{aligned} \tag{31}$$

Based on the Equations (29), (30) and (31) we can formulate the classical alternating least squares (ALS) algorithm to estimate iteratively the factor matrices $\boldsymbol{A}^{(n)}$ $(n = 1, \ldots, N)$ from the measurement tensor $\boldsymbol{\mathcal{X}}$ (BRO, 1998; SMILDE, BRO, and GELADI, 2004; CICHOCKI *et al.*, 2009). Each iteration of the ALS-PARAFAC algorithm contains three LS updating steps. At each step, one factor matrix is updated while the remaining factor matrices are assumed fixed to their values obtained in the previous steps. This procedure is repeated until the convergence of the algorithm at the i-th iteration which is achieved when $|e_i - e_{i-1}| \leq \delta$, where $e_{(i)}$ denotes the residual error calculated at the i-th iteration defined by

$$e_{(i)} = \left\| \boldsymbol{\mathcal{X}} - \underbrace{\boldsymbol{\mathcal{I}}_{3,Q} \times_1 \hat{\boldsymbol{A}}_{(i)}^{(1)} \times_2 \hat{\boldsymbol{A}}_{(i)}^{(2)} \times_3 \hat{\boldsymbol{A}}_{(i)}^{(3)}}_{\hat{\boldsymbol{\mathcal{X}}}_{(i)}} \right\|_{\mathrm{F}}^2, \tag{32}$$

Algorithm 2 ALS-PARAFAC

1: **procedure** ALS-PARAFAC

2: 1. *Set* $i = 0$. *Randomly initialize* $\hat{\boldsymbol{A}}^{(2)}_{(i=0)}$ *and* $\hat{\boldsymbol{A}}^{(3)}_{(i=0)}$;

3: 2. $i \leftarrow i + 1$;

4: 3. *From* $[\boldsymbol{\mathcal{X}}]_{(1)}$ *in* (29), *obtain an LS estimate of* $\hat{\boldsymbol{A}}^{(1)}_{(i)}$:

5:
$$\hat{\boldsymbol{A}}^{(1)}_{(i)} = [\boldsymbol{\mathcal{X}}]_{(1)} \left[\left(\hat{\boldsymbol{A}}^{(3)}_{(i-1)} \diamond \hat{\boldsymbol{A}}^{(2)}_{(i-1)} \right)^{\mathrm{T}} \right]^{\dagger};$$

6: 4. *From* $[\boldsymbol{\mathcal{X}}]_{(2)}$ *in* (30), *obtain an LS estimate of* $\hat{\boldsymbol{A}}^{(2)}_{(i)}$:

7:
$$\hat{\boldsymbol{A}}^{(2)}_{(i)} = [\boldsymbol{\mathcal{X}}]_{(2)} \left[\left(\hat{\boldsymbol{A}}^{(3)}_{(i-1)} \diamond \hat{\boldsymbol{A}}^{(1)}_{(i)} \right)^{\mathrm{T}} \right]^{\dagger};$$

8: 5. *From* $[\boldsymbol{\mathcal{X}}]_{(3)}$ *in* (31), *obtain an LS estimate of* $\hat{\boldsymbol{A}}^{(3)}_{(i)}$:

9:
$$\hat{\boldsymbol{A}}^{(3)}_{(i)} = [\boldsymbol{\mathcal{X}}]_{(3)} \left[\left(\hat{\boldsymbol{A}}^{(2)}_{(i)} \diamond \hat{\boldsymbol{A}}^{(1)}_{(i)} \right)^{\mathrm{T}} \right]^{\dagger};$$

10: *Repeat steps* 2-5 *until convergence.*

where $\hat{\boldsymbol{\mathcal{X}}}_{(i)}$ is the reconstructed version of $\boldsymbol{\mathcal{X}}$ computed from the estimated factor matrices $\hat{\boldsymbol{A}}^{(n)}_{(i)}$ $(n = 1, 2, 3)$ obtained at the end of the i-th iteration. δ is a prescribed threshold value assumed to be 10^{-6} throughout this thesis. The ALS-PARAFAC algorithm is summarized in Algorithm 2. Note that the dominant cost of the ALS-PARAFAC algorithm is associated with the SVD computation used to calculate three pseudo-inverses in lines 5, 7 and 9 of Algorithm 2. Therefore, the overall computational complexity of the ALS-PARAFAC (per iteration) can be approximated to $\mathcal{O}(Q^2(I_1I_2 + I_1I_3 + I_2I_3))$ FLOPs.

Uniqueness of the PARAFAC Decomposition

The most attractive feature of the PARAFAC decomposition for signal processing applications is its simple and well defined essential uniqueness property that is guaranteed by the Kruskal's condition introduced by KRUSKAL (1977). This condition is based on the following fundamental Kruskal-rank (k-rank) definition:

Definition 12. (*k-rank*) The Kruskal-rank of $\boldsymbol{A} \in \mathbb{C}^{I_1 \times Q}$, defined as $\kappa_{\boldsymbol{A}}$, is the maximum number such that any subset of $\kappa_{\boldsymbol{A}}$ columns of $\boldsymbol{A}$ are linearly independent. As a consequence $\kappa_{\boldsymbol{A}} \leq \rho_{\boldsymbol{A}} \leq \min(I_1, Q)$, in which $\rho_{\boldsymbol{A}}$ denotes the rank of $\boldsymbol{A}$. If $\boldsymbol{A}$ is a full rank matrix it is also full k-rank, i.e., $\boldsymbol{\kappa_A} = \rho_{\boldsymbol{A}}$.

For the third-order tensor (23), if the following Kruskal's condition is satisfied

$$\kappa_{\boldsymbol{A}^{(1)}} + \kappa_{\boldsymbol{A}^{(2)}} + \kappa_{\boldsymbol{A}^{(3)}} \geq 2Q + 2, \tag{33}$$

the uniqueness of its PARAFAC decomposition is guaranteed. Basically, any triplet

$\left(\hat{\boldsymbol{A}}^{(1)}, \hat{\boldsymbol{A}}^{(2)}, \hat{\boldsymbol{A}}^{(3)}\right)$ is related to the true triplet $\left(\boldsymbol{A}^{(1)}, \boldsymbol{A}^{(2)}, \boldsymbol{A}^{(3)}\right)$ up to trivial permutation and scaling of columns. In this case, the true and estimated factor matrices are linked by the following relation:

$$\hat{\boldsymbol{A}}^{(n)} = \boldsymbol{A}^{(n)} \boldsymbol{\Pi} \boldsymbol{\Delta}^{(n)}, \quad n = 1, 2, 3, \tag{34}$$

where $\boldsymbol{\Delta}^{(n)}$ $(n = 1, 2, 3)$ are diagonal scaling matrices such as $\prod_{n=1}^{3} \boldsymbol{\Delta}^{(n)} = \boldsymbol{I}_Q$, and $\boldsymbol{\Pi}$ is an unknown permutation matrix.

Extension of the PARAFAC Decomposition to N-th Order Tensors

For simplicity of presentation, we first introduced above only the PARAFAC decomposition of a third-order tensor. However, its generalization for an N-th order tensor is straightforward. The PARAFAC decomposition of the rank-Q tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ is given by

$$\boldsymbol{\mathcal{X}} = \sum_{q=1}^{Q} \boldsymbol{a}_q^{(1)} \circ \boldsymbol{a}_q^{(2)} \circ \cdots \circ \boldsymbol{a}_q^{(N)}, \tag{35}$$

or, equivalently, by using n-mode product notation we get

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{I}}_{N,Q} \times_1 \boldsymbol{A}^{(1)} \times_2 \boldsymbol{A}^{(2)} \cdots \times_N \boldsymbol{A}^{(N)}. \tag{36}$$

The n-mode unfolding matrix of (36) is given by

$$[\boldsymbol{\mathcal{X}}]_{(n)} = \boldsymbol{A}^{(n)} \left(\boldsymbol{A}^{(N)} \diamond \cdots \diamond \boldsymbol{A}^{(n+1)} \diamond \boldsymbol{A}^{(n-1)} \diamond \cdots \diamond \boldsymbol{A}^{(1)}\right)^{\mathrm{T}}. \tag{37}$$

Remark: Throughout Chapter 3, a special attention is given to dual-symmetric tensors. The PARAFAC decomposition of a tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1 \times \cdots \times I_{2N}}$ of even order $2N$ is said to have dual-symmetry if it is defined as (WEIS, 2015):

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{I}}_{2N,Q} \times_1 \boldsymbol{A}^{(1)} \times_2 \boldsymbol{A}^{(2)} \cdots \times_N \boldsymbol{A}^{(N)} \times_{N+1} \boldsymbol{A}^{(1)^*} \times_{N+2} \boldsymbol{A}^{(2)^*} \cdots \times_{2N} \boldsymbol{A}^{(N)^*}, \tag{38}$$

where $I_{N+n} = I_n$ and $\boldsymbol{A}^{(N+n)} = \boldsymbol{A}^{(n)^*}$ $(n = 1, \ldots, N)$. This definition also applies to Tucker decomposition by simply replacing the identity tensor $\boldsymbol{\mathcal{I}}_{2N,Q}$ by an arbitrary core tensor $\boldsymbol{\mathcal{G}}$ of order $2N$.

The generalization of the Kruskal's condition (33) was formulated by SIDIROPOULOS and BRO (2000) for an N-th order tensor. For the generalized case, the Kruskal's condition becomes:

$$\sum_{n=1}^{N} \kappa_{\boldsymbol{A}^{(n)}} \geq 2Q + (N - 1). \tag{39}$$

2.4.2 Nested-PARAFAC Decomposition

The Nested-PARAFAC decomposition was introduced and also applied by DE ALMEIDA and FAVIER (2013) in the wireless communication area. The Nested-PARAFAC decomposition assumes that the n-th factor matrix $\boldsymbol{A}^{(n)} \in \mathbb{C}^{I_n \times Q}$ in (25) is itself an unfolding matrix of an additional PARAFAC decomposition. By assuming, for simplicity of presentation, that $\boldsymbol{A}^{(1)} \in \mathbb{C}^{I_1 \times Q}$ denotes the 1-mode unfolding of the third-order tensor $\boldsymbol{\mathcal{Y}} \in \mathbb{C}^{J_1 \times J_2 \times J_3}$, i.e., $\boldsymbol{A}^{(1)} = [\boldsymbol{\mathcal{Y}}]_{(1)}$ in which $J_1 = I_1$ and $J_2 J_3 = Q$, we can define the Nested-PARAFAC decomposition of $\boldsymbol{\mathcal{X}}$ in terms of the following two (linked) PARAFAC decompositions

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{I}}_{3,Q} \times_1 [\boldsymbol{\mathcal{Y}}]_{(1)} \times_2 \boldsymbol{A}^{(2)} \times_3 \boldsymbol{A}^{(3)}, \tag{40}$$

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{I}}_{3,Q} \times_1 \boldsymbol{B}^{(1)} \times_2 \boldsymbol{B}^{(2)} \times_3 \boldsymbol{B}^{(3)}, \tag{41}$$

where (40) is called outer PARAFAC part, while (41) is called inner PARAFAC part (LIU *et al.*, 2013). The matrices $\boldsymbol{B}^{(n)} \in \mathbb{C}^{J_n \times Q}$ $(n = 1, 2, 3)$ are the factor matices of the inner PARAFAC part. Note that, in a generalized representation form, the inner PARAFAC part may be associated with any mode of $\boldsymbol{\mathcal{X}}$ without loss of generality.

The estimation of the factor matrices $\boldsymbol{A}^{(n)}$ and $\boldsymbol{B}^{(n)}$ $(n = 1, 2, 3)$ can be obtained using the Algorithm 2 two times sequentially. First, the factor matrices of the outer PARAFAC part are estimated from $\boldsymbol{\mathcal{X}}$. After the convergence of the first stage, the factor matrices of the inner PARAFAC part are estimated from $\boldsymbol{\mathcal{Y}}$ previously obtained. This procedure is detailed in Algorithm 3. Since the Nested-PARAFAC decomposition can be formulated as two linked PARAFAC decompositions according to (40) and (41), its uniqueness is achieved satisfying together the following two Kruskal's conditions:

$$\kappa_{\boldsymbol{A}^{(1)}} + \kappa_{\boldsymbol{A}^{(2)}} + \kappa_{\boldsymbol{A}^{(3)}} \geq 2Q + 2 \quad \text{and} \quad \kappa_{\boldsymbol{B}^{(1)}} + \kappa_{\boldsymbol{B}^{(2)}} + \kappa_{\boldsymbol{B}^{(3)}} \geq 2Q + 2. \tag{42}$$

2.4.3 Tucker Decomposition

The Tucker decomposition was introduced by TUCKER (1966). Given a third-order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$, the Tucker decomposition factorizes it as a multilinear transformation of a third-order core tensor $\boldsymbol{\mathcal{G}} \in \mathbb{C}^{Q_1 \times Q_2 \times Q_3}$ by the factor matrices $\boldsymbol{A}^{(n)} = \left[\boldsymbol{a}_1^{(n)}, \boldsymbol{a}_2^{(n)}, \ldots, \boldsymbol{a}_{Q_n}^{(n)}\right] \in \mathbb{C}^{I_n \times Q_n}$ $(n = 1, 2, 3)$ as illustrated in Figure 6. Mathematically, the Tucker decomposition of $\boldsymbol{\mathcal{X}}$ is defined as

$$\boldsymbol{\mathcal{X}} = \sum_{q_1=1}^{Q_1} \sum_{q_2=1}^{Q_2} \sum_{q_3=1}^{Q_3} g_{q_1,q_2,q_3} \left(\boldsymbol{a}_{q_1}^{(1)} \circ \boldsymbol{a}_{q_2}^{(2)} \circ \boldsymbol{a}_{q_3}^{(3)}\right), \tag{43}$$

Algorithm 3 ALS-Nested-PARAFAC

procedure ALS-NESTED-PARAFAC

First Stage: ALS-PARAFAC in the outer PARAFAC part

1.1 *Set* $i = 0$. *Randomly initialize* $\hat{\boldsymbol{A}}_{(i=0)}^{(2)}$ *and* $\hat{\boldsymbol{A}}_{(i=0)}^{(3)}$;

1.2 $i \leftarrow i + 1$;

1.3 *From* $[\boldsymbol{\mathcal{X}}]_{(1)}$, *obtain an LS estimate of* $\hat{\boldsymbol{A}}_{(i)}^{(1)}$:

$$\hat{\boldsymbol{A}}_{(i)}^{(1)} = [\boldsymbol{\mathcal{X}}]_{(1)} \left[\left(\hat{\boldsymbol{A}}_{(i-1)}^{(3)} \diamond \hat{\boldsymbol{A}}_{(i-1)}^{(2)} \right)^{\mathrm{T}} \right]^{\dagger};$$

1.4 *From* $[\boldsymbol{\mathcal{X}}]_{(2)}$, *obtain an LS estimate of* $\hat{\boldsymbol{A}}_{(i)}^{(2)}$:

$$\hat{\boldsymbol{A}}_{(i)}^{(2)} = [\boldsymbol{\mathcal{X}}]_{(2)} \left[\left(\hat{\boldsymbol{A}}_{(i-1)}^{(3)} \diamond \hat{\boldsymbol{A}}_{(i)}^{(1)} \right)^{\mathrm{T}} \right]^{\dagger};$$

1.5 *From* $[\boldsymbol{\mathcal{X}}]_{(3)}$, *obtain an LS estimate of* $\hat{\boldsymbol{A}}_{(i)}^{(3)}$:

$$\hat{\boldsymbol{A}}_{(i)}^{(3)} = [\boldsymbol{\mathcal{X}}]_{(3)} \left[\left(\hat{\boldsymbol{A}}_{(i)}^{(2)} \diamond \hat{\boldsymbol{A}}_{(i)}^{(1)} \right)^{\mathrm{T}} \right]^{\dagger};$$

Repeat steps 1.2-1.5 *until convergence.*

Second Stage: ALS-PARAFAC in the inner PARAFAC part

2.1 *Construct the tensor* $\boldsymbol{\mathcal{Y}}$ *from* $\hat{\boldsymbol{A}}^{(1)} = [\boldsymbol{\mathcal{Y}}]_{(1)}$;

2.2 *Set* $i = 0$. *Randomly initialize* $\hat{\boldsymbol{B}}_{(i=0)}^{(2)}$ *and* $\hat{\boldsymbol{B}}_{(i=0)}^{(3)}$;

2.3 $i \leftarrow i + 1$;

2.4 *From* $[\boldsymbol{\mathcal{Y}}]_{(1)}$, *obtain an LS estimate of* $\hat{\boldsymbol{B}}_{(i)}^{(1)}$:

$$\hat{\boldsymbol{B}}_{(i)}^{(1)} = [\boldsymbol{\mathcal{Y}}]_{(1)} \left[\left(\hat{\boldsymbol{B}}_{(i-1)}^{(3)} \diamond \hat{\boldsymbol{B}}_{(i-1)}^{(2)} \right)^{\mathrm{T}} \right]^{\dagger};$$

2.5 *From* $[\boldsymbol{\mathcal{Y}}]_{(2)}$, *obtain an LS estimate of* $\hat{\boldsymbol{B}}_{(i)}^{(2)}$:

$$\hat{\boldsymbol{B}}_{(i)}^{(2)} = [\boldsymbol{\mathcal{Y}}]_{(2)} \left[\left(\hat{\boldsymbol{B}}_{(i-1)}^{(3)} \diamond \hat{\boldsymbol{B}}_{(i)}^{(1)} \right)^{\mathrm{T}} \right]^{\dagger};$$

2.6 *From* $[\boldsymbol{\mathcal{Y}}]_{(3)}$, *obtain an LS estimate of* $\hat{\boldsymbol{B}}_{(i)}^{(3)}$:

$$\hat{\boldsymbol{B}}_{(i)}^{(3)} = [\boldsymbol{\mathcal{Y}}]_{(3)} \left[\left(\hat{\boldsymbol{B}}_{(i)}^{(2)} \diamond \hat{\boldsymbol{B}}_{(i)}^{(1)} \right)^{\mathrm{T}} \right]^{\dagger};$$

Repeat steps 2.3-2.6 *until convergence.*

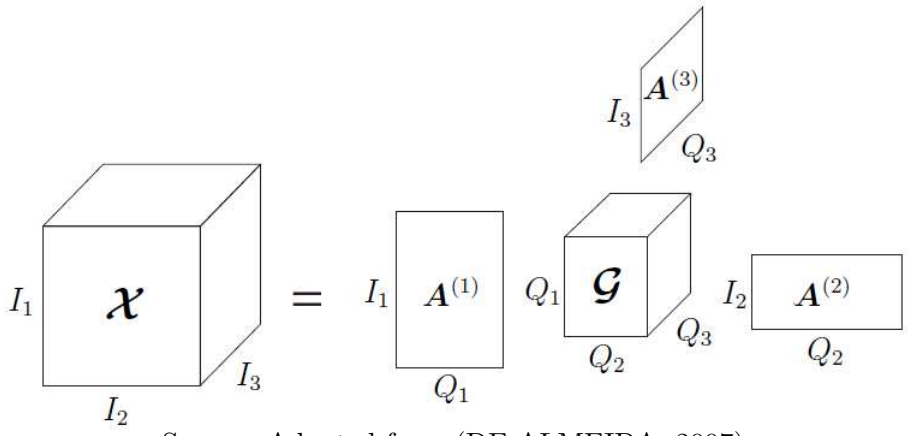
where its (i_1, i_2, i_3)-th element is given by

$$x_{i_1,i_2,i_3} = \sum_{q_1=1}^{Q_1} \sum_{q_2=1}^{Q_2} \sum_{q_3=1}^{Q_3} g_{q_1,q_2,q_3} a_{i_1,q_1}^{(1)} a_{i_2,q_2}^{(2)} a_{i_3,q_3}^{(3)}, \tag{44}$$

where Q_n is the multilinear rank of $\boldsymbol{\mathcal{X}}$, i.e., the number of components in the n-th ($n = 1, 2, 3$) mode of $\boldsymbol{\mathcal{X}}$. $a_{i_n,q_n}^{(n)}$ is the (i_n, q_n)-th element of the factor matrix $\boldsymbol{A}^{(n)} \in \mathbb{C}^{I_n \times Q_n}$ ($n = 1, 2, 3$) and g_{q_1,q_2,q_3} is the (q_1, q_2, q_3)-th element of the core tensor $\boldsymbol{\mathcal{G}} \in \mathbb{C}^{Q_1 \times Q_2 \times Q_3}$. Using the n-mode product notation, the Tucker decomposition of $\boldsymbol{\mathcal{X}}$ can be written as

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{G}} \times_1 \boldsymbol{A}^{(1)} \times_2 \boldsymbol{A}^{(2)} \times_3 \boldsymbol{A}^{(3)}. \tag{45}$$

Figure 6 – Tucker decomposition of a third-order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$.



Source: Adapted from (DE ALMEIDA, 2007).

The 1-mode $[\boldsymbol{\mathcal{X}}]_{(1)} \in \mathbb{C}^{I_1 \times I_2 I_3}$, 2-mode $[\boldsymbol{\mathcal{X}}]_{(2)} \in \mathbb{C}^{I_2 \times I_1 I_3}$ and 3-mode $[\boldsymbol{\mathcal{X}}]_{(3)} \in \mathbb{C}^{I_3 \times I_1 I_2}$ unfolding matrices of $\boldsymbol{\mathcal{X}}$ are represented, respectively, by

$$[\boldsymbol{\mathcal{X}}]_{(1)} = \boldsymbol{A}^{(1)} [\boldsymbol{\mathcal{G}}]_{(1)} \left(\boldsymbol{A}^{(3)} \otimes \boldsymbol{A}^{(2)}\right)^{\mathrm{T}}, \tag{46}$$

$$[\boldsymbol{\mathcal{X}}]_{(2)} = \boldsymbol{A}^{(2)} [\boldsymbol{\mathcal{G}}]_{(2)} \left(\boldsymbol{A}^{(3)} \otimes \boldsymbol{A}^{(1)}\right)^{\mathrm{T}}, \tag{47}$$

$$[\boldsymbol{\mathcal{X}}]_{(3)} = \boldsymbol{A}^{(3)} [\boldsymbol{\mathcal{G}}]_{(3)} \left(\boldsymbol{A}^{(2)} \otimes \boldsymbol{A}^{(1)}\right)^{\mathrm{T}}, \tag{48}$$

where $[\boldsymbol{\mathcal{G}}]_{(1)} \in \mathbb{C}^{Q_1 \times Q_2 Q_3}$, $[\boldsymbol{\mathcal{G}}]_{(2)} \in \mathbb{C}^{Q_2 \times Q_1 Q_3}$ and $[\boldsymbol{\mathcal{G}}]_{(3)} \in \mathbb{C}^{Q_3 \times Q_1 Q_2}$ denote the 1-mode, 2-mode and 3-mode unfolding matrices of $\boldsymbol{\mathcal{G}}$, respectively. The Tucker2 decomposition where $\boldsymbol{A}^{(3)} = \boldsymbol{I}_{I_3}$ and $\boldsymbol{\mathcal{G}} \in \mathbb{C}^{Q_1 \times Q_2 \times I_3}$, and the Tucker1 decomposition where $\boldsymbol{A}^{(3)} = \boldsymbol{I}_{I_3}$, $\boldsymbol{A}^{(2)} = \boldsymbol{I}_{I_2}$ and $\boldsymbol{\mathcal{G}} \in \mathbb{C}^{Q_1 \times I_2 \times I_3}$ are special cases of the Tucker decomposition (KOLDA and BADER, 2009).

Uniqueness of the Tucker Decomposition

In general, the Tucker decomposition is not unique i.e., there are infinite solutions for the factor matrices and for the core tensor that leads to the same tensor $\boldsymbol{\mathcal{X}}$.

Proof. Let $\boldsymbol{T}^{(n)}$ $(n = 1, 2, 3)$ be non-singular matrices, we can rewrite (45) as

$$\begin{aligned}
\boldsymbol{\mathcal{X}} &= \hat{\boldsymbol{\mathcal{G}}} \times_1 \hat{\boldsymbol{A}}^{(1)} \times_2 \hat{\boldsymbol{A}}^{(2)} \times_3 \hat{\boldsymbol{A}}^{(3)} \\
&= \underbrace{\boldsymbol{\mathcal{G}} \times_1 \boldsymbol{T}^{(1)^{-1}} \times_2 \boldsymbol{T}^{(2)^{-1}} \times_3 \boldsymbol{T}^{(3)^{-1}}}_{\hat{\boldsymbol{\mathcal{G}}}} \times_1 \underbrace{\boldsymbol{A}^{(1)}\boldsymbol{T}^{(1)}}_{\hat{\boldsymbol{A}}^{(1)}} \times_2 \underbrace{\boldsymbol{A}^{(2)}\boldsymbol{T}^{(2)}}_{\hat{\boldsymbol{A}}^{(2)}} \times_3 \underbrace{\boldsymbol{A}^{(3)}\boldsymbol{T}^{(3)}}_{\hat{\boldsymbol{A}}^{(3)}} \\
&= \boldsymbol{\mathcal{G}} \times_1 \boldsymbol{T}^{(1)^{-1}} \times_1 \boldsymbol{A}^{(1)}\boldsymbol{T}^{(1)} \times_2 \boldsymbol{T}^{(2)^{-1}} \times_2 \boldsymbol{A}^{(2)}\boldsymbol{T}^{(2)} \times_3 \boldsymbol{T}^{(3)^{-1}} \times_3 \boldsymbol{A}^{(3)}\boldsymbol{T}^{(3)}.
\end{aligned}$$

Making use of the property (22), we obtain

$$\begin{aligned}\boldsymbol{\mathcal{X}} &= \boldsymbol{\mathcal{G}} \times_1 \left(\boldsymbol{A}^{(1)}\boldsymbol{T}^{(1)}\boldsymbol{T}^{(1)^{-1}}\right) \times_2 \left(\boldsymbol{A}^{(2)}\boldsymbol{T}^{(2)}\boldsymbol{T}^{(2)^{-1}}\right) \times_3 \left(\boldsymbol{A}^{(3)}\boldsymbol{T}^{(3)}\boldsymbol{T}^{(3)^{-1}}\right) \\ &= \boldsymbol{\mathcal{G}} \times_1 \boldsymbol{A}^{(1)} \times_2 \boldsymbol{A}^{(2)} \times_3 \boldsymbol{A}^{(3)}. \quad \blacksquare\end{aligned}$$

From the above deduction, we can conclude that the Tucker decomposition is not unique since its factor matrices are affected by arbitrary linear transformations, while the inverse of these transformations affect the core tensor, leading to the same tensor $\boldsymbol{\mathcal{X}}$. However, in special cases in which several elements of $\boldsymbol{\mathcal{G}}$ are constrained to be equal to zero, that is, if the core tensor has some sparsity, the solutions to the factor matrices becomes unique up to trivial permutations and scaling ambiguities similar to the PARAFAC decomposition (TEN BERGE and SMILDE, 2002). The Tucker decomposition formulated throughout this thesis belong to this special category where unique solutions exist due to the sparsity of the core tensor.

Extension of the Tucker Decomposition to N-th Order Tensors

For an N-th order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$, its N-th order Tucker decomposition is given by

$$\boldsymbol{\mathcal{X}} = \sum_{q_1=1}^{Q_1} \sum_{q_2=1}^{Q_2} \cdots \sum_{q_N=1}^{Q_N} g_{q_1,q_2,\ldots,q_N} \left(\boldsymbol{a}_{q_1}^{(1)} \circ \boldsymbol{a}_{q_2}^{(2)} \circ \cdots \circ \boldsymbol{a}_{q_N}^{(N)}\right), \tag{49}$$

or, equivalently, by using n-mode product notation

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{G}} \times_1 \boldsymbol{A}^{(1)} \times_2 \boldsymbol{A}^{(2)} \cdots \times_N \boldsymbol{A}^{(N)}. \tag{50}$$

The n-mode unfolding matrix of (50) is given by

$$[\boldsymbol{\mathcal{X}}]_{(n)} = \boldsymbol{A}^{(n)} [\boldsymbol{\mathcal{G}}]_{(n)} \left(\boldsymbol{A}^{(N)} \otimes \cdots \otimes \boldsymbol{A}^{(n+1)} \otimes \boldsymbol{A}^{(n-1)} \otimes \cdots \otimes \boldsymbol{A}^{(1)}\right)^{\mathrm{T}}. \tag{51}$$

The Higher-Order Singular Value Decomposition (HOSVD)

The HOSVD, introduced by DE LATHAUWER, DE MOOR, and VANDEWALLE (2000), is a direct extension of the matrix SVD to higher-order tensors. It is a way to compute the basis for each factor matrix of the Tucker decomposition. The HOSVD of an N-th order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ is defined as

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{S}} \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \cdots \times_N \boldsymbol{U}^{(N)}, \tag{52}$$

where $\boldsymbol{U}^{(n)} \in \mathbb{C}^{I_n \times Q_n}$ $(n = 1, 2, \ldots, N)$ are unitary matrices and $\boldsymbol{\mathcal{S}} \in \mathbb{C}^{Q_1 \times Q_2 \times \cdots \times Q_N}$ is the core tensor. The set $(Q_1, Q_2, \ldots, Q_N)$ denotes the multilinear rank of $\boldsymbol{\mathcal{X}}$. The matrix

Algorithm 4 Higher-order singular value decomposition (HOSVD)

procedure HOSVD
for $n = 1, \ldots, N$
    1. *Compute* $\boldsymbol{U}^{(n)}$ *as the first* Q_n *left singular vectors of* $[\boldsymbol{\mathcal{X}}]_{(n)}$:

$$[\boldsymbol{\mathcal{X}}]_{(n)} = \boldsymbol{U}^{(n)} \boldsymbol{\Sigma}^{(n)} \boldsymbol{V}^{(n)^{\mathrm{H}}};$$

end
2. *Compute the core tensor as:*

$$\boldsymbol{\mathcal{S}} = \boldsymbol{\mathcal{X}} \times_1 \boldsymbol{U}^{(1)^{\mathrm{H}}} \times_2 \boldsymbol{U}^{(2)^{\mathrm{H}}} \cdots \times_N \boldsymbol{U}^{(N)^{\mathrm{H}}}.$$

3. *Return* $\boldsymbol{\mathcal{S}}$ *and* $\boldsymbol{U}^{(n)}$ $(n = 1, 2, \ldots, N)$.

$\boldsymbol{U}^{(n)}$ is computed as the Q_n left singular vectors of the n-th unfolding matrix of $\boldsymbol{\mathcal{X}}$, i.e.,

$$[\boldsymbol{\mathcal{X}}]_{(n)} = \boldsymbol{U}^{(n)} \boldsymbol{\Sigma}^{(n)} \boldsymbol{V}^{(n)^{\mathrm{H}}}. \tag{53}$$

From $\boldsymbol{U}^{(n)}$ $(n = 1, 2, \ldots, N)$ computed according (53), the core tensor can be obtained as

$$\boldsymbol{\mathcal{S}} = \boldsymbol{\mathcal{X}} \times_1 \boldsymbol{U}^{(1)^{\mathrm{H}}} \times_2 \boldsymbol{U}^{(2)^{\mathrm{H}}} \cdots \times_N \boldsymbol{U}^{(N)^{\mathrm{H}}}. \tag{54}$$

The HOSVD procedure is described in the pseudo-code form in Algorithm 4.

In this thesis, we will use the HOSVD concept to formulate efficient denoising strategies in Chapter 3. As aforementioned, the HOSVD provides the bases for the factor matrices of the Tucker decomposition. In contrast, in Chapter 3 an iterative algorithm based on the ALS concept is developed and applied in our context to estimate the true factor matrices and core tensor of a fourth-order Tucker decomposition.

2.5 Chapter Summary

This chapter has provided a summary of the main theoretical and algorithmic backgrounds on multilinear algebra and tensor decompositions. All the contents presented in this chapter, in the course of four parts, provide the basic material that will be exploited throughout this thesis. In the first and second parts of this chapter, we have introduced some operations, namely, the Kronecker and Khatri-Rao products, the SVD and the LS-KRF factorizations. In the third and fourth parts of this chapter, we have focused on presenting basic tensor operations and the most important tensor decompositions encountered in the application contexts of this thesis, such as array signal processing, MIMO radar and MIMO wireless communications systems.

3 TENSOR-BASED METHODS FOR BLIND SPATIAL SIGNATURE ESTIMATION IN MULTIDIMENSIONAL SENSOR ARRAYS

In this chapter, we address the spatial signatures and spatial frequencies estimation problems in multidimensional (R-D) sensor arrays. It is divided into two parts. In the first part of this chapter, we propose a tensor-based multiple denoising approach that successively applies spatial smoothing, denoising and reconstruction to the noisy data. By taking into account the knowledge of the model order and by exploiting subarrays created by the spatial smoothing, we can successively denoise the data by means of HOSVD-based and SVD-based low-rank approximation for tensor and matrix data, respectively. In the second part of this chapter, we propose two generalized iterative algorithms for spatial signatures estimation. The first tensor-based algorithm is an R-D blind spatial signature estimator that operates in scenarios where the sources' covariance matrix is nondiagonal and unknown. The second tensor-based algorithm is formulated to address the uncorrelated sources case and exploits the dual-symmetry of the covariance tensor. Additionally, a new tensor-based formulation is proposed for an L-shaped array configuration. Simulation results show that the proposed receiver, referred to as multiple denoising (MuDe), significantly reduces the noise level, yielding a more accurate estimation of the parameters without decreasing the sensor array aperture, while the algorithms proposed in the second part outperform the state-of-the-art matrix-based and tensor-based spatial parameter estimation techniques.

3.1 Introduction and Motivation

The high resolution parameter estimation from noisy signal measurements plays a key role in several practical applications in the array signal processing area ranging from MIMO communications (ZHOU *et al.*, 2016a), radar (SINGH, WANG, and CRAGÉ, 2016; NION and SIDIROPOULOS, 2010) and global navigation satellite systems (GNSS) (HAMMOUD *et al.*, 2016) to name a few. The performance of parameter estimation techniques is sensitive to the SNR and severely degrades in noisy scenarios. Therefore, a denoising preprocessing can be an effective step to improve the SNR and consequently the estimation accuracy.

The interest in denoising techniques for multidimensional data has received strong attention in seismology (MUTI and BOURENNANE, 2005) and image processing (ZHANG *et al.*, 2015; CHEN and QIAN, 2011) applications. In the array signal processing context, low-rank tensor approximation based on the HOSVD (DE LATHAUWER, DE MOOR, and VANDEWALLE, 2000) has been used as tool for denoising in multidimensional data (HAARDT, ROEMER, and DEL GALDO, 2008; GOMES *et al.*, 2017; WEIS, DEL GALDO, and HAARDT, 2007). For applying the denoising, the model order

should be previously estimated (DA COSTA *et al.*, 2011; LIU *et al.*, 2016).

From the preprocessed denoised data, several parameter estimation techniques such as the classical multiple signal classification (MUSIC) (SCHMIDT, 1986) and estimation of signal parameters via rotational invariance technique (ESPRIT) (ROY and KAILATH, 1989) can be applied to estimate the spatial frequencies of the sources. However, improved performance can be achieved by exploiting the multidimensional structure of the data by means of tensor modeling, which can include several signal dimensions such as space, time, frequency, and polarization. Tensor decompositions have been successfully employed in array signal processing for parameters estimation since they provide better identifiability conditions when compared to conventional matrix-based techniques. Another advantage of tensor-based methods is the so-called *tensor gain* which manifests itself with more precise parameter estimates due to the good noise rejection capability of tensor-based signal processing as shown in (DA COSTA *et al.*, 2011; ROEMER, 2013; DA COSTA *et al.*, 2013; DA COSTA, 2010).

In this chapter, we address the spatial signatures and spatial frequencies estimation problems in multidimensional sensor arrays. In the first part of the chapter, we show that, by taking into account the model order and by exploiting the possible subarrays created by the spatial smoothing, the data can be successively denoised by means of the HOSVD-based and SVD-based low-rank approximations, which drastically reduces the noise effect in tensor and matrix data, respectively. We initially propose a new tensor-based preprocessing method for noise reduction in noisy measurement data collected by a multidimensional (R-D) sensor array. Exploiting the spatial smoothing concept (SHAN, WAX, and KAILATH, 1985), a better signal estimation is obtained by applying successive SVD-based low-rank approximations in the output signals for subarrays of different sizes in each spatial dimension of the received signal tensor, and then reconstructing back the subarrays into the original tensor so that the procedure can be repeated. We also discuss the matrix-based approach that is a particular case of the proposed method when an one-dimensional sensor array is considered.

In the second part of the chapter, two tensor-based approaches to the estimation of spatial signatures are presented. By using the signals received on an R-D sensor array, covariance tensors are computed and solutions for correlated and uncorrelated sources are formulated, respectively. For the former scenario, in which the sources' covariance structure is nondiagonal and unknown, the covariance tensor of the received data is formulated as a Tucker decomposition of order $2R$. Such a formulation yields a generalized Tucker model based R-D sensor array processing that deals with arbitrary source covariance structures. By assuming uncorrelated sources, we then show that the problem boils down to a PARAFAC decomposition, from which a method that exploits the dual-symmetry property of the covariance tensor is derived by considering the ideas rooted in (WEIS *et al.*, 2012). For both Tucker and PARAFAC based models, the blind

estimation of the spatial signatures are achieved by means of ALS-based algorithms. Simulation results show that the MuDe achieves a significant noise reduction capability, while the proposed algorithms for spatial parameters estimation outperform the state-of-the-art matrix-based and tensor-based approaches.

3.1.1 Chapter Organization

This chapter is organized as follows. Referring to its first part, Section 3.2 presents the data model for an R-dimensional sensor array. In Section 3.3, we briefly recall the R-D spatial smoothing preprocessing scheme. The proposed multiple denoising (MuDe) via successive spatial smoothing, low-rank approximation and reconstruction is described for matrix and tensor formulations in Section 3.4. Simulation results for performance evaluation of MuDe are provided in Section 3.5. Referring to the second part of the chapter, Section 3.6 introduce a novel covariance-based tensor modeling for the received data from which the proposed algorithms for blind spatial signatures estimation are formulated. In Section 3.7, a novel tensor-based approach for parameters estimation in L-shaped sensor arrays is proposed. The computational complexity of the proposed algorithms is analyzed in Section 3.8. In Section 3.9, the advantages and the disadvantages of the proposed methods are discussed. Simulation results to the contributions contained in the second part are provided in Section 3.10. Finally, the conclusions are drawn in Section 3.11.

In the following section, we begin the first part of this chapter dedicated to formulate the proposed MuDe preprocessing method.

3.2 Data Model

Let us consider an R-dimensional sensor array of size $N_1 \times N_2 \times \cdots \times N_R$ composed by $N = \prod_{r=1}^{R} N_r$ sensor elements, where N_r $(r = 1, \ldots, R)$ denotes the number of sensors in the r-th spatial dimension, i.e., the size of the r-th array dimension. The received signal $\boldsymbol{X} \in \mathbb{C}^{N \times K}$ at the sensor array from the superposition of M far-field narrowband signal sources sampled on K subsequent time instants is given by

$$\boldsymbol{X} = \boldsymbol{A}\boldsymbol{S} + \boldsymbol{N} \in \mathbb{C}^{N \times K}, \tag{55}$$

where $\boldsymbol{S} = [\boldsymbol{s}(1), \ldots, \boldsymbol{s}(K)] \in \mathbb{C}^{M \times K}$ is the symbols matrix and $\boldsymbol{N} = [\boldsymbol{n}(1), \ldots, \boldsymbol{n}(K)] \in \mathbb{C}^{N \times K}$ is the additive white Gaussian noise (AWGN) matrix. $\boldsymbol{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_M] \in \mathbb{C}^{N \times M}$ denotes the overall array steering matrix and $\boldsymbol{a}_m$ $(m = 1, \ldots, M)$ represents the array steering vector related to the m-th signal source.

In the special case in which the received signal has a multidimensional structure, for instance, when an outer product based array (OPA) is considered (DA COSTA *et al.*, 2010), the received signal matrix in (55) can be expressed as (HAARDT, ROEMER, and DEL GALDO, 2008; STEINWANDT *et al.*, 2017)

$$\boldsymbol{X} = \left(\boldsymbol{A}^{(1)} \diamond \boldsymbol{A}^{(2)} \cdots \diamond \boldsymbol{A}^{(R)}\right) \boldsymbol{S} + \boldsymbol{N}, \tag{56}$$

where $\boldsymbol{A} = \boldsymbol{A}^{(1)} \diamond \boldsymbol{A}^{(2)} \cdots \diamond \boldsymbol{A}^{(R)} \in \mathbb{C}^{N \times M}$, $\boldsymbol{A}^{(r)} = [\boldsymbol{a}_1^{(r)}, \ldots, \boldsymbol{a}_M^{(r)}] \in \mathbb{C}^{N_r \times M}$ is the array steering matrix of the r-th array dimension. The vector $\boldsymbol{a}_m^{(r)} \in \mathbb{C}^{N_r \times 1}$ denotes the array response for the m-th signal source, $m = 1, \ldots, M$, in the r-th spatial dimension given by

$$\boldsymbol{a}_m^{(r)} = \left[1, e^{j \cdot \mu_m^{(r)}}, e^{j \cdot 2\mu_m^{(r)}}, \cdots, e^{j \cdot (N_r - 1)\mu_m^{(r)}}\right]^{\mathrm{T}}, \tag{57}$$

where $\mu_m^{(r)}$ is the spatial frequency of the m-th wavefront in the r-th array dimension.

In this case, the multidimensional structure of the data can be naturally exploited and (56) can be seen as a unimodal unfolding matrix of the noisy tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{N_1 \times N_2 \times \cdots \times N_R \times K}$ of order $R + 1$ modeled as follows

$$\begin{aligned} \boldsymbol{\mathcal{X}} &= \boldsymbol{\mathcal{X}}_o + \boldsymbol{\mathcal{N}} \\ &= \boldsymbol{\mathcal{I}}_{R+1,M} \times_1 \boldsymbol{A}^{(1)} \cdots \times_R \boldsymbol{A}^{(R)} \times_{R+1} \boldsymbol{S}^{\mathrm{T}} + \boldsymbol{\mathcal{N}}, \end{aligned} \tag{58}$$

where $\boldsymbol{\mathcal{I}}_{R+1,M}$ is the identity tensor of order $R + 1$ in which each dimension has size M. Its elements are equal to one when all indices are equal, and zero otherwise. The AWGN tensor $\boldsymbol{\mathcal{N}} \in \mathbb{C}^{N_1 \times N_2 \times \cdots \times N_R \times K}$ is obtained by arranging $\boldsymbol{N}$ as a tensor of order $R + 1$.

According to (37), the unimodal unfolding matrix $[\boldsymbol{\mathcal{X}}]_{(r)} \in \mathbb{C}^{N_r \times \prod_{j \neq r} N_j K}$ along the r-th spatial mode of $\boldsymbol{\mathcal{X}}$ ($r = 1, \ldots, R$) is expressed as

$$\begin{aligned} [\boldsymbol{\mathcal{X}}]_{(r)} &= [\boldsymbol{\mathcal{X}}_o]_{(r)} + [\boldsymbol{\mathcal{N}}]_{(r)} \\ &= \boldsymbol{A}^{(r)} \left(\boldsymbol{S}^{\mathrm{T}} \diamond \boldsymbol{A}^{(R)} \diamond \cdots \boldsymbol{A}^{(r+1)} \diamond \boldsymbol{A}^{(r-1)} \diamond \cdots \boldsymbol{A}^{(1)}\right)^{\mathrm{T}} + [\boldsymbol{\mathcal{N}}]_{(r)}. \end{aligned} \tag{59}$$

where $[\boldsymbol{\mathcal{N}}]_{(r)} \in \mathbb{C}^{N_r \times \prod_{j \neq r} N_j K}$ denotes the r-th unfolding matrix of the noise tensor $\boldsymbol{\mathcal{N}}$.

Remark: In the absence of noise, the measurement tensor $\boldsymbol{\mathcal{X}}$ in (58) has rank M. Throughout this chapter, we assume the knowledge of M at the receiver. The model order can be estimated with good accuracy by applying the techniques proposed in (DA COSTA *et al.*, 2011) and (LIU *et al.*, 2016). The model order estimation problem is beyond the scope of our study.

3.3 Spatial Smoothing

In the following, we briefly overview concepts about the R-D spatial smoothing preprocessing scheme. It will be incorporated as an important part of the proposed denoising technique formulated in the next section of this chapter.

Consider that the r-th unfolding matrix $[\boldsymbol{\mathcal{X}}]_{(r)}$, defined in (59), with N_r rows is divided into L_r subarrays of size $N_r^{(\text{sub})} = N_r - L_r + 1$ each. By applying the spatial smoothing preprocessing scheme (SHAN, WAX, and KAILATH, 1985) independently in each unfolding matrix for $r = 1, \ldots, R$, we define the r-th mode spatially smoothed matrix $\boldsymbol{X}_{\text{SS},r}^{(L_r)} \in \mathbb{C}^{N_r^{(\text{sub})} \times \prod_{j \neq r} N_j L_r K}$ which collects the output signals of the L_r subarrays as follows (THAKRE *et al.*, 2010)

$$\begin{aligned} \boldsymbol{X}_{\text{SS},r}^{(L_r)} &= \left[[\boldsymbol{\mathcal{X}}]_{(r)}^{(1)}, [\boldsymbol{\mathcal{X}}]_{(r)}^{(2)}, \ldots, [\boldsymbol{\mathcal{X}}]_{(r)}^{(L_r)} \right] \\ &= \left[\boldsymbol{J}_1^{(N_r)} [\boldsymbol{\mathcal{X}}]_{(r)}, \boldsymbol{J}_2^{(N_r)} [\boldsymbol{\mathcal{X}}]_{(r)}, \cdots, \boldsymbol{J}_{L_r}^{(N_r)} [\boldsymbol{\mathcal{X}}]_{(r)} \right], \\ & \qquad\qquad r = 1, \ldots, R \end{aligned} \tag{60}$$

where $[\boldsymbol{\mathcal{X}}]_{(r)}^{(l_r)} = \boldsymbol{J}_{l_r}^{(N_r)} [\boldsymbol{\mathcal{X}}]_{(r)} \in \mathbb{C}^{N_r^{(\text{sub})} \times \prod_{j \neq r} N_j K}$ $(l_r = 1, \ldots, L_r)$ denotes the output signal of the l_r-th subarray in the r-th dimension. The selection matrix $\boldsymbol{J}_{l_r}^{(N_r)}$ related to the l_r-th subarray in the r-th dimension is given by

$$\boldsymbol{J}_{l_r}^{(N_r)} = \left[\boldsymbol{0}_{N_r^{(\text{sub})} \times (l_r - 1)} \quad \boldsymbol{I}_{N_r^{(\text{sub})}} \quad \boldsymbol{0}_{N_r^{(\text{sub})} \times (L_r - l_r)} \right]. \tag{61}$$

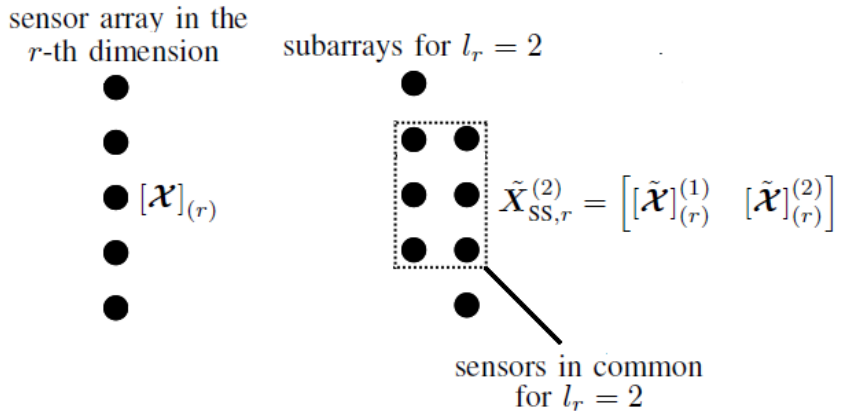
3.4 Proposed Multiple Denoising via Successive Spatial Smoothing, Low-Rank Approximation and Reconstruction

In this section, we propose the matrix and tensor based multiple denoising (MuDe) via successive spatial smoothing, low-rank approximation and reconstruction as a preprocessing step to enhance parameter estimation accuracy in R-D sensor array processing. We first start from the general case by adopting a tensor-based approach. The matrix approach will be briefly discussed as a special case.

3.4.1 Tensor-based Approach (R-D sensor array case)

The R-D spatial smoothing scheme presented in Section 3.3 provides us with subarrays for each spatial dimension of $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{N_1 \times N_2 \times \cdots \times N_R \times K}$. By fully exploiting all possible subarrays, the received tensor can be successively denoised from low-rank approximations applied to the set of spatially smoothed unfolding matrices $\left\{ \boldsymbol{X}_{\text{SS},r}^{(1)}, \ldots, \boldsymbol{X}_{\text{SS},r}^{(L_r)} \right\}$ in (60) formed from subarrays of different sizes $(l_r = 1, \ldots, L_r)$ for each spatial dimension $(r = 1, \ldots, R)$. The denoised tensor version $\tilde{\boldsymbol{\mathcal{X}}} \in \mathbb{C}^{N_1 \times N_2 \times \cdots \times N_R \times K}$ can then be obtained

Figure 7 – Illustration of spatial smoothing and sensors in common for $l_r = 2$ subarrays in the r-th spatial dimension.



Source: Created by the author.

by reconstructing back the subarrays in the r-th array dimension into the original tensor form. To this end, we formulate below the three steps of the proposed MuDe method:

• **Step 1 -** *Subarray Construction*: Initially, we construct the r-th mode spatially smoothed matrix $\boldsymbol{X}^{(l_r)}_{\mathrm{SS},r}$ according to (60) associated with the l_r-th subarray size in the r-th spatial dimension of the received data tensor.

• **Step 2 -** *Low-Rank Approximation*: By assuming the knowledge of the model order M, we compute the following low-rank approximation for the spatially smoothed matrix $\boldsymbol{X}^{(l_r)}_{\mathrm{SS},r}$ obtained in step 1:

$$\tilde{\boldsymbol{X}}^{(l_r)}_{\mathrm{SS},r} = \left[[\tilde{\mathcal{X}}]^{(1)}_{(r)} \quad \cdots \quad [\tilde{\mathcal{X}}]^{(l_r)}_{(r)}\right], \tag{62}$$

where $\tilde{\boldsymbol{X}}^{(l_r)}_{\mathrm{SS},r} = \boldsymbol{U}_s \boldsymbol{\Sigma}_s \boldsymbol{V}_s^{\mathrm{H}}$ is obtained by truncating the SVD of $\boldsymbol{X}^{(l_r)}_{\mathrm{SS},r}$ to the signal subspace. Note that this step requires that $N_r^{(\mathrm{sub})} \geq M$.

• **Step 3 -** *Reconstruction of the Denoised Tensor*: Finally, we reconstruct back the low-rank approximation of the subarrays outputs $[\tilde{\mathcal{X}}]^{(l_r)}_{(r)}$ in (62) into the original tensor form of the received data. Note that applying spatial smoothing in step 1 yields overlapping subarrays with sensors in common in the r-th spatial dimension. Thanks to this overlapping, we get a multiple denoised unfolding matrix $[\tilde{\mathcal{X}}]_{(r)} \in \mathbb{C}^{N_r \times \prod_{j \neq r} N_j K}$ associated with the r-th dimension given by

$$[\tilde{\mathcal{X}}]_{(r)} = \begin{bmatrix} [\tilde{\mathcal{X}}]_{(r)}(1,:) \\ [\tilde{\mathcal{X}}]_{(r)}(2,:) \\ \vdots \\ [\tilde{\mathcal{X}}]_{(r)}(N_r,:) \end{bmatrix} \in \mathbb{C}^{N_r \times \prod_{j \neq r} N_j K}, \tag{63}$$

which can be reconstructed from each subarray output using a mean-based reconstruction method. The n-th row of the denoised unfolding matrix $[\tilde{\boldsymbol{\mathcal{X}}}]_{(r)}$ is computed as

$$[\tilde{\boldsymbol{\mathcal{X}}}]_{(r)}(n,:) = \frac{1}{l'}\sum_{i=1}^{l_r}[\tilde{\boldsymbol{\mathcal{X}}}]_{(r)}^{(i)}(n-i+1,:). \tag{64}$$

The value of l' is equal to the number of times in which the term $[\tilde{\boldsymbol{\mathcal{X}}}]_{(r)}^{(i)}(n-i+1,:)$ is a valid output in the l_r-th subarray of the r-th dimension. This occurs when $N_r^{(\text{sub})} \geq n-i+1 > 0$. Figure 7 illustrates the use of the relation (64) for the particular case $l_r = 2$. As a numerical example, and in accordance with Figure 7, the following expression describes $[\tilde{\boldsymbol{\mathcal{X}}}]_{(r)}$ as a function of the output signal of each subarray after the subarray construction and low-rank approximation steps:

$$[\tilde{\boldsymbol{\mathcal{X}}}]_{(r)} = \begin{bmatrix} [\tilde{\boldsymbol{\mathcal{X}}}]_{(r)}^{(1)}(1,:) \\ \frac{1}{2}\left[[\tilde{\boldsymbol{\mathcal{X}}}]_{(r)}^{(1)}(2:N_r-1,:) + [\tilde{\boldsymbol{\mathcal{X}}}]_{(r)}^{(2)}(1:N_r-2,:)\right] \\ [\tilde{\boldsymbol{\mathcal{X}}}]_{(r)}^{(2)}(N_r-1,:) \end{bmatrix}.$$

Our main idea consists in achieving a higher noise reduction capability by exploiting all possible subarrays in each spatial dimension of the received data. This is possible by applying steps 1-3 successively for subarrays of sizes $l_r = 1,\ldots,L_r$, for each value $r = 1,\ldots,R$. In the end, after covering all possible subarrays sizes for each spatial dimension of the received data, the multiple denoised tensor $\tilde{\boldsymbol{\mathcal{X}}} \in \mathbb{C}^{N_1\times N_2\times\cdots\times N_R\times K}$ is obtained by arranging (63) as a tensor of order $R+1$. Then, the spatial frequencies of the M wavefronts can be estimated by resorting to R-D parameter estimation schemes to preprocessed data $\tilde{\boldsymbol{\mathcal{X}}}$ instead of the raw data $\boldsymbol{\mathcal{X}}$.

The proposed MuDe method is a general preprocessing framework for denoising, so that several R-D parameter estimation techniques such as Tensor-ESPRIT (HAARDT, ROEMER, and DEL GALDO, 2008), Tensor-MUSIC (GONG, XU, and LIU, 2006) and PARAFAC decomposition (LIU and SIDIROPOULOS, 2004) can be applied in the preprocessed data tensor to estimate the spatial frequencies of the sources with increased accuracy, due to the high noise rejection capability of MuDe. However, Tensor-MUSIC based approaches needs multi-dimensional search and have high computational complexity. Parameter estimates based on the PARAFAC decomposition are often obtained via iterative techniques such as the classical ALS-PARAFAC algorithm 2. The ALS-PARAFAC algorithm requires many iterations and do not guarantee convergence to the global optimum (RAJIH, 2006). In contrast, ESPRIT-based approaches have demonstrated to be powerful tools due to their closed-form solutions with faster processing time and performance close to the Cramér-Rao bound (CRB) (DA COSTA *et al.*, 2010).

Algorithm 5 Proposed MuDe Method

procedure MuDe
input: *received signal tensor* $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{N_1 \times N_2 \times \cdots \times N_R \times K}$, *model order* M, *maximum number of subarrays in each spatial dimension* $L_r = N_r^{(\text{sub})} - M$ $(r = 1, \ldots, R)$.
output: *denoised tensor*: $\tilde{\boldsymbol{\mathcal{X}}} \in \mathbb{C}^{N_1 \times N_2 \times \cdots \times N_R \times K}$.
for $r = 1, \ldots, R$
    for $l_r = 1, \ldots, L_r$
    1. *Compute the* r*-th mode spatially smoothed matrix* $\boldsymbol{X}_{\text{SS},r}^{(l_r)}$ *for the* l_r*-th subarray size in the* r*-th array dimension according to* (60).
    2. *Compute the low-rank approximation* $\tilde{\boldsymbol{X}}_{\text{SS},r}^{(l_r)}$ *of the spatially smoothed matrix* $\boldsymbol{X}_{SS,r}^{(l_r)}$.
    3. *Obtain the multiple denoised tensor* $\tilde{\boldsymbol{\mathcal{X}}}$ *by arranging the* r*-th unfolding matrix* (63) *as a* $(R+1)$*-th order tensor.*
    end
end

According to Algorithm 5, our approach begins with the subarray construction step by means of the spatial smoothing preprocessing scheme in the r-th array dimension for a given subarray size, as described in Section 3.3. Then, by assuming the knowledge of the model order, the low-rank approximation applied to the r-th mode spatially smoothed matrix (60) is computed. In the end, a refined estimation is obtained by reconstructing back the low-rank approximation of the subarrays outputs into the original tensor form in accordance with (63) and (64). These steps are repeated for all possible subarrays sizes $(l_r = 1, \ldots, L_r)$ in all spatial dimensions $(r = 1, \ldots, R)$ of the received data. The resulting multiple denoised tensor $\tilde{\boldsymbol{\mathcal{X}}}$ can be subsequently processed instead of the received signal $\boldsymbol{\mathcal{X}}$ in (58).

3.4.2 Computational Complexity of MuDe

The computational complexity of MuDe can be approximated as the sum of the FLOPs required to compute a total of L_r truncated SVDs for each spatial dimension of the data tensor $(r = 1, \ldots, R)$. According to (GOLUB and VAN LOAN, 1996), the computational complexity of the SVD for a matrix of size $N_r - l_r + 1 \times (\prod_{j \neq r} N_j) l_r K$ truncated to rank M is $\mathcal{O}\Big((\prod_{j \neq r} N_j)(N_r - l_r + 1) M l_r K\Big)$. From this, we deduce that the overall computational complexity of the MuDe method is $\mathcal{O}\Big(\sum_{r=1}^{R} \sum_{l_r=1}^{L_r} (\prod_{j \neq r} N_j)(N_r - l_r + 1) M l_r K\Big)$.

Remark: Note that, the higher noise reduction capability of MuDe is accompanied by an increase of computational complexity since a total of $R \cdot L_r$ truncated SVD computations are required. In contrast, the state-of-the-art approach such as (HAARDT, ROEMER, and DEL GALDO, 2008) requires only R SVD computations. The trade-off between more accurated spatial parameter estimation and computational complexity using MuDe is reinforced through computational simulations in Section 3.5.

3.4.3 Matrix-based Approach (1-D sensor array case)

A particular case of the proposed MuDe method occurs when a uniform linear array (ULA) is considered. For the 1-D array configuration, the spatial smoothing scheme is directly applied to the received data matrix in (55). Given $\boldsymbol{X} \in \mathbb{C}^{N \times K}$ as input to Algorithm 5, the steps 1-3 are repeated only L times, where L denotes the number of possible subarrays in the 1-D sensor array. Hence, in the matrix (1-D array) case, the computational complexity of the MuDe method reduces to $\mathcal{O}\Big(\sum_{l=1}^{L}\left((N-l+1)MlK\right)\Big)$.

3.5 Performance Evaluation of MuDe

In this section, we evaluate the performance of MuDe through computational experiments. We present simulation results for the proposed matrix-based and tensor-based MuDe approaches. We consider three figures of merit obtained from 50000 independent Monte Carlo runs. First, we evaluate the SNR gain when MuDe is used as a preprocessing step for noise reduction. Then, we evaluate the accuracy of the estimated spatial frequencies in terms of the total root mean square error (RMSE) for the 1-D and R-D sensor array configurations. At the end, we evaluate the mean processing time of different techniques. The spatial frequencies are uniformly distributed in $[-\pi, \pi]$ and randomly chosen in each run. The distance between neighboring sensors is assumed to be $\lambda/2$ in all array dimensions, while the source symbols are binary phase shift keying (BPSK) modulated signals. Zero-mean AWGN is assumed in all experiments.

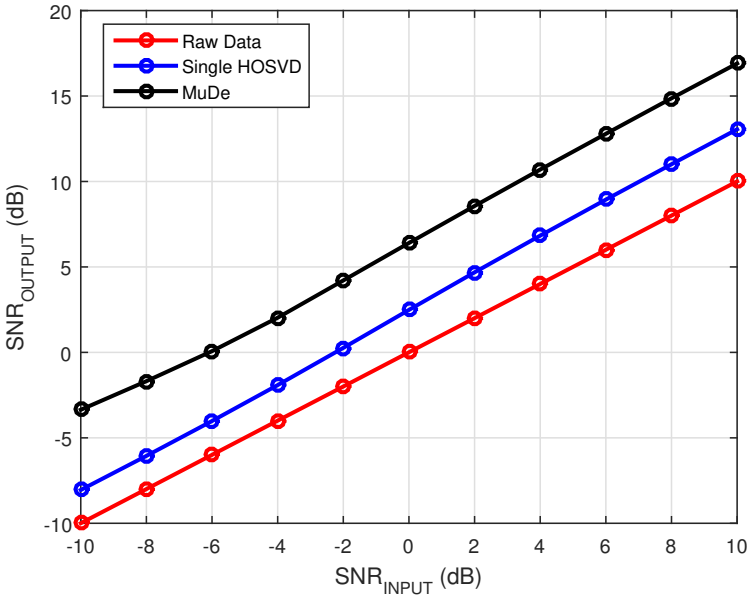
The total $\text{RMSE}_{(i)}$ computed at the i-th Monte Carlo run is defined as follows:

$$\text{RMSE}_{(i)} = \sqrt{\frac{1}{RM}\left(\sum_{r=1}^{R}\sum_{m=1}^{M}\left(\mu_{m,i}^{(r)} - \hat{\mu}_{m,i}^{(r)}\right)^2\right)}, \tag{65}$$

where $\hat{\mu}_{m,i}^{(r)}$ denotes the estimate of the m-th spatial frequency in the r-th array dimension at the i-th Monte Carlo run. The total RMSE for a given SNR value plotted in the Figures 9, 10, 11 and 12 denote the median value computed over 50000 independent Monte Carlo runs. Here, the spatial frequencies are estimated by applying the standard one-dimensional MUSIC (SCHMIDT, 1986), the one-dimensional ESPRIT (ROY and KAILATH, 1989), the standard multidimensional R-D ESPRIT (HAARDT, ROEMER, and DEL GALDO, 2008) and the ALS-PARAFAC 2 algorithms. We also compare our results to the Cramér-Rao bound (CRB) for uniform linear array (ULA) and uniform rectangular array (URA) configurations derived in (ROEMER and HAARDT, 2012).

In Figures 8 and 9, we evaluate the performance of MuDe considering a third-order received signal tensor (two spatial dimensions and one temporal dimension). In both cases, we assume a URA (i.e., $R = 2$) with N_1 and N_2 equal to 30 sensors in each spatial dimension, $M = 6$ wavefronts and $K = 10$ signal samples. The SNR gain of MuDe

Figure 8 – $\text{SNR}_{\text{output}}$ vs. $\text{SNR}_{\text{input}}$ for a third-order tensor to compare the SNR gain of MuDe and the classical single HOSVD based low-rank approximation proposed by (HAARDT, ROEMER, and DEL GALDO, 2008). In the simulated scenario: URA, $N_1 = 30$, $N_2 = 30$, $M = 6$ and $K = 10$.



Source: Created by the author.

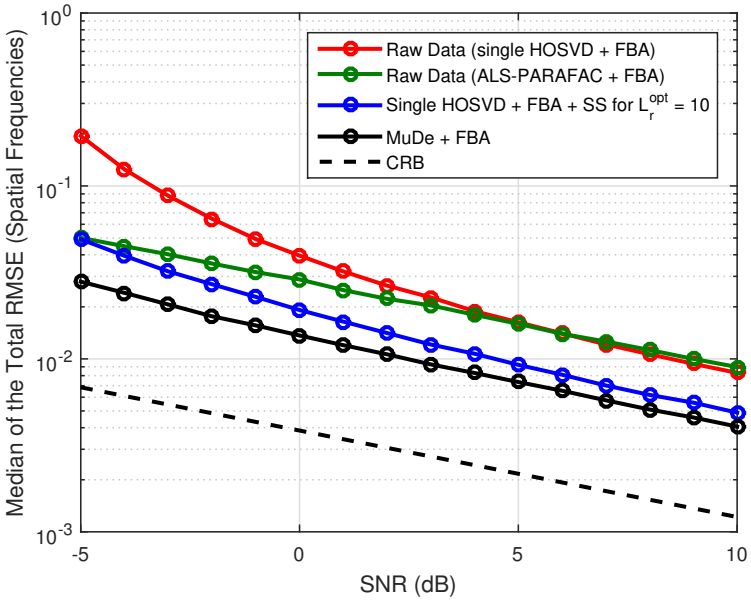
is shown in Figure 8. In this analysis, we consider a controlled scenario in which the noiseless signal term $\boldsymbol{\mathcal{X}}_o$ in (58) is available. Being $\tilde{\boldsymbol{\mathcal{X}}}$ the preprocessed data, we define the output SNR as follows:

$$\text{SNR}_{\text{output}} = 10 \cdot \log_{10} \left(\frac{\|\boldsymbol{\mathcal{X}}_o\|_{\text{F}}^2}{\|\tilde{\boldsymbol{\mathcal{X}}} - \boldsymbol{\mathcal{X}}_o\|_{\text{F}}^2} \right) \quad \text{(dB)}. \tag{66}$$

According to Figure 8, the proposed tensor-based MuDe outperforms the classical single HOSVD based low-rank approximation proposed in (HAARDT, ROEMER, and DEL GALDO, 2008) in approximatelly 4 dB. When compared to the received raw data in (58) i.e., without consider any denoising preprocessing, the tensor-based MuDe achieves a SNR gain up to 6 dB.

In Figure 9, a first preprocessing step via forward-backward averaging (FBA) technique (LINEBARGER, DE GROAT, and DOWLING, 1994; HAARDT, ROEMER, and DEL GALDO, 2008) is incorporated in all tensor-based methods. In (STEINWANDT, ROEMER, and HAARDT, 2014), the number of subarrays in the spatial smoothing is optimized by minimizing the mean-squared error (MSE) for the standard ESPRIT. In the considered scenario the optimal subarray length is $L_r^{\text{opt}} = 10$ in each spatial mode ($r = 1, 2$) for the method proposed in (STEINWANDT, ROEMER, and HAARDT, 2014) It is plotted here as a reference for comparison. For more details about the analytical derivation of the optimal number of subarrays for spatial smoothing please con-

Figure 9 – Median of the total RMSE of the spatial frequencies vs. SNR (dB) for a third-order tensor to compare the spatial parameter estimation accuracy. In the scenario: URA, $N_1 = 30$, $N_2 = 30$, $M = 6$ and $K = 10$.

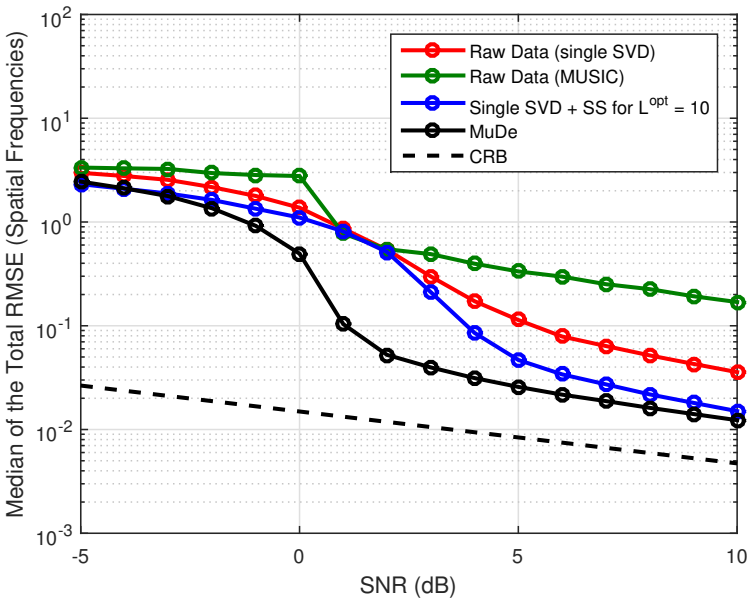


Source: Created by the author.

sult reference (STEINWANDT, ROEMER, and HAARDT, 2014). The proposed MuDe method outperforms the single HOSVD+FBA method in (HAARDT, ROEMER, and DEL GALDO, 2008) and the ALS-PARAFAC+FBA in (DA COSTA *et al.*, 2010). Compared to the HOSVD+FBA+SS method in (STEINWANDT, ROEMER, and HAARDT, 2014) for the optimal number of subarrays $L_r^{\text{opt}} = 10$, MuDe provides a lower estimation error in all the simulated SNR range, showing more accentuate gains in the low SNR regime.

In Figure 10, we assume $K = 30$ signal samples, $M = 6$ wavefronts and a ULA (i.e., $R = 1$) equipped with $N = 30$ sensors. As benchmarks to the matrix-based MuDe approach, we consider the state-of-the-art single SVD-based low-rank approximation method and its version added with spatial smoothing preprocessing scheme, where the optimal number of subarrays $L^{\text{opt}} = N/3 = 10$ is considered, as proposed in (STEINWANDT, ROEMER, and HAARDT, 2014). The standard one-dimensional MUSIC with an angular updating of 0.1° is also plotted for comparison. The proposed MuDe method outperforms the competing techniques with a more accentuated gain when compared to the classical single SVD-based low-rank approximation without spatial smoothing and MUSIC. In the single source case, the proposed method presents the same performance as that of (STEINWANDT, ROEMER, and HAARDT, 2014), as shown in Figure 11. This result corroborates the remarkable performance provided by MuDe, achieving the bound derived in (STEINWANDT, ROEMER, and HAARDT, 2014). On the other hand, the proposed method becomes preferable when a larger number of signal sources is considered,

Figure 10 – Median of the total RMSE of the spatial frequencies vs. SNR (dB) for the matrix-based approach. In the scenario: ULA, $N = 30$, $M = 6$ and $K = 30$.



Source: Created by the author.

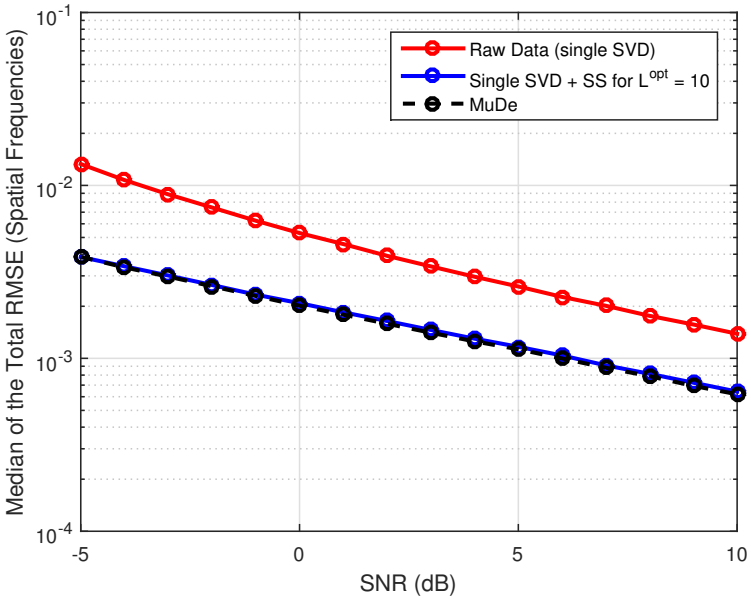
as shown in Figure 10.

In order to evaluate the spatial frequency estimation accuracy for close-sources, we present the Figure 12. In this experiment, we assume a ULA with $N = 30$ sensors, $K = 30$ and $M = 5$ closely spaced signal sources. The SNR is fixed to 0 dB. The angular spacing between the sources (in degrees) is shown on the horizontal axis. We can see that the performance of all techniques improves when the angular spacing between sources increases. As we can see from these simulation results, for very close-sources, when the angular spacing is less than 10°, the proposed MuDe is more accurate than the competitors methods. The worse performance of the approach in (STEINWANDT, ROEMER, and HAARDT, 2014) is due to the decrease of the array aperture with the use of the spatial smoothing preprocessing scheme, which decreases the array resolution for very close-sources. In contrast, the proposed MuDe significantly reduces the noise effect without decreasing the array aperture providing more accurate estimates in challeging scenarios when the sources are closely spaced.

In Figure 13, the mean processing time (in seconds) of the MuDe is compared to the state-of-the-art techniques for the single-source scenario. We can see that the MuDe is more computationally expensive than competing techniques showing an increase in the processing time when the number of sensors increases. According to Figures 9 and 10, this is the trade-off to achieve more accurate estimates for the spatial frequencies.

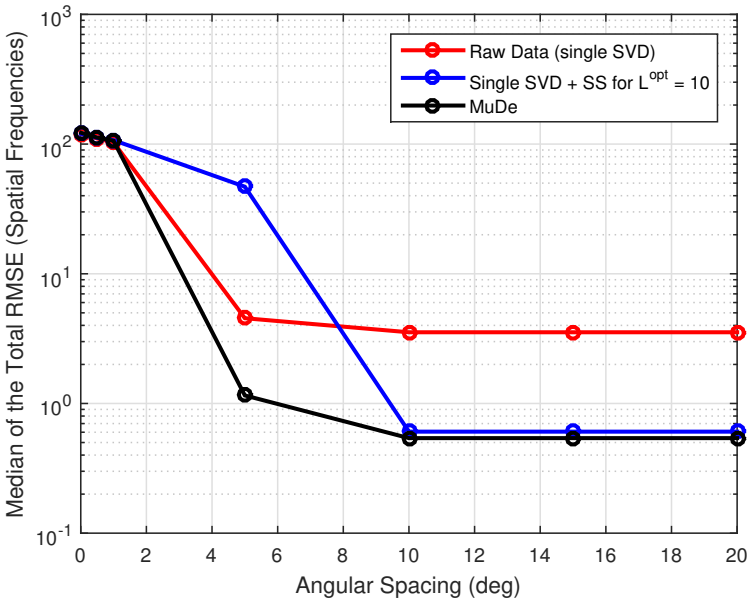
In the following section, we begin the second part of this chapter dedicated to formulate two new generalized tensor-based algorithms to solve the spatial signatures estimation problem in R-D sensor arrays.

Figure 11 – Median of the total RMSE of the spatial frequencies vs. SNR (dB) for the matrix-based approach. In the scenario: ULA, $N = 30$, $M = 1$ and $K = 30$.



Source: Created by the author.

Figure 12 – Median of the total RMSE of the spatial frequencies vs. angular spacing (in degrees) between the sources. In the scenario: ULA, $N = 30$, $M = 5$, $K = 30$ and SNR equal to 0 dB.

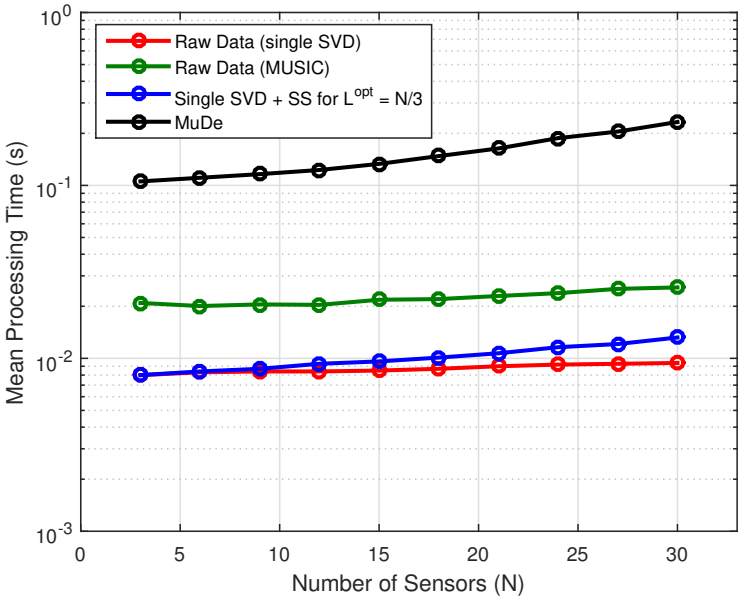


Source: Created by the author.

3.6 Proposed Tensor-Based Methods for Blind Spatial Signature Estimation

In this section, we first formulate a novel multidimensional structure from the covariance matrix of the received data. Then, an iterative ALS-based algorithm for a Tucker decomposition of order $2N$ is proposed. Finally, we derive a link between the

Figure 13 – Mean processing time (in seconds) vs. number of sensors. In the scenario: ULA, single-source case, $K = 30$ and SNR equal to 30 dB.



Source: Created by the author.

method in (WEIS *et al.*, 2012) and the covariance-based blind spatial signature estimation problem.

3.6.1 Covariance Tensor Formulation

From (56), the sample covariance matrix $\hat{\boldsymbol{R}} \in \mathbb{C}^{N \times N}$ of the signals received at the R-D sensor array can be computed as

$$\begin{aligned} \hat{\boldsymbol{R}} &\triangleq \frac{1}{K}\boldsymbol{X}\boldsymbol{X}^{\mathrm{H}} \approx \boldsymbol{A}\boldsymbol{R}_{\boldsymbol{s}}\boldsymbol{A}^{\mathrm{H}} + \sigma_n^2\boldsymbol{I} \\ &\approx \left(\boldsymbol{A}^{(1)} \diamond \cdots \diamond \boldsymbol{A}^{(R)}\right) \boldsymbol{R}_{\boldsymbol{s}} \left(\boldsymbol{A}^{(1)} \diamond \cdots \diamond \boldsymbol{A}^{(R)}\right)^{\mathrm{H}} + \sigma_n^2\boldsymbol{I}, \end{aligned} \tag{67}$$

where $\boldsymbol{R}_{\boldsymbol{s}} = \frac{1}{K}\boldsymbol{S}\boldsymbol{S}^{\mathrm{H}}$ is the sample covariance matrix of the source signals, and σ_n^2 is the noise variance.

The noiseless sample covariance matrix in (67), given by $\boldsymbol{R}_{\boldsymbol{o}} = \hat{\boldsymbol{R}} - \sigma_n^2\boldsymbol{I} \in \mathbb{C}^{N \times N}$, can be interpreted as a multimode unfolding of the noiseless covariance tensor $\boldsymbol{\mathcal{R}}_{\boldsymbol{o}} \in \mathbb{C}^{N_1 \times N_2 \times \ldots \times N_R \times N_1 \times N_2 \times \ldots \times N_R}$ of order $2N$, defined as

$$\begin{aligned} \boldsymbol{\mathcal{R}}_{\boldsymbol{o}} = \boldsymbol{\mathcal{R}}_{\boldsymbol{s}} \times_1 \boldsymbol{A}^{(1)} \times_2 \boldsymbol{A}^{(2)} \ldots \times_R \boldsymbol{A}^{(R)} \\ \times_{R+1} \boldsymbol{A}^{(1)^*} \times_{R+2} \boldsymbol{A}^{(2)^*} \ldots \times_{2R} \boldsymbol{A}^{(R)^*}, \end{aligned} \tag{68}$$

where $\boldsymbol{\mathcal{R}}_{\boldsymbol{s}}$ is the source covariance tensor, which has $2R$ dimensions, each of size M. According to (38), $\boldsymbol{\mathcal{R}}_{\boldsymbol{o}}$ is a dual-symmetric tensor, i.e. the factor matrix related to the $(R+r)$-th dimension is equal to $\boldsymbol{A}^{(r)^*}$, and $N_r = N_{R+r}$ $(r = 1, \ldots, R)$. The m-th frontal

slice of $\boldsymbol{\mathcal{R}_s}$ is a diagonal matrix whose main diagonal is given by the m-th column of the covariance matrix $\boldsymbol{R_s}$. For instance, considering $R = 2$ for the sake of notation, the following expression satisfies the relationship previously cited:

$$\boldsymbol{\mathcal{R}_s}(:,:,m,m) = \text{diag}\left(\boldsymbol{R_s}(:,m)\right), \quad m = 1,\ldots,M, \tag{69}$$

where the matrix $\boldsymbol{\mathcal{R}_s}(:,:,m,m) \in \mathbb{C}^{M\times M}$ denotes the m-th frontal slice of the covariance tensor $\boldsymbol{\mathcal{R}_s}$ obtained by fixing its last two modes. The tensor $\boldsymbol{\mathcal{R}_o}$ follows a dual-symmetric Tucker decomposition of order $2R$ with factor matrices $\boldsymbol{A}^{(r)}$ and $\boldsymbol{A}^{(r)^*}$, $r = 1,\ldots,R$, and core tensor $\boldsymbol{\mathcal{R}_s}$.

Considering the case in which the sources are uncorrelated and have unitary variance, we can rewrite (68) as

$$\begin{aligned}\boldsymbol{\mathcal{R}_o} = \boldsymbol{\mathcal{I}}_{2R,M} \times_1 \boldsymbol{A}^{(1)} \times_2 \boldsymbol{A}^{(2)} \ldots \times_R \boldsymbol{A}^{(R)} \\ \times_{R+1} \boldsymbol{A}^{(1)^*} \times_{R+2} \boldsymbol{A}^{(2)^*} \ldots \times_{2R} \boldsymbol{A}^{(R)^*},\end{aligned} \tag{70}$$

where $\boldsymbol{\mathcal{I}}_{2R,M}$ is the identity tensor of order $2R$ in which each dimension has size M. In this case, the covariance tensor $\boldsymbol{\mathcal{R}_o}$ follows a dual-symmetric PARAFAC decomposition of order $2R$.

As discussed in Section 2.4.3, the Tucker decomposition does not impose restrictions on the core tensor structure, which makes this model more flexible compared to PARAFAC decomposition. In our context, this characteristic reflects an arbitrary and unknown structure for the sources' covariance $\boldsymbol{R_s}$ which can also be estimated from (68). In contrast, the PARAFAC decomposition (70) denotes a particular case of the Tucker decomposition when the sources' signals are uncorrelated and the source covariance matrix is perfectly known (i.e. diagonal). In practice, this may not hold.

3.6.2 ALS-Tucker Algorithm

Our goal is to blindly estimate the spatial signature matrices $\boldsymbol{A}^{(r)}$ and $\boldsymbol{A}^{(r)^*}$ $(r = 1,\ldots,R)$ which refer to the different dimensions of the sensor array from the covariance tensor $\boldsymbol{\mathcal{R}_o}$. For the sake of simplicity, from this point on, we consider $\boldsymbol{A}^{(R+r)} = \boldsymbol{A}^{(r)^*}$. According to (51), the Tucker decomposition (68) allows the following factorization in terms of its factor matrices and core tensor:

$$[\boldsymbol{\mathcal{R}_o}]_{(r)} = \boldsymbol{A}^{(r)}\boldsymbol{\Delta}^{(r)}, \tag{71}$$

where

$$\boldsymbol{\Delta}^{(r)} = [\boldsymbol{\mathcal{R}}_{\boldsymbol{s}}]_{(r)} \left(\boldsymbol{A}^{(2R)} \otimes \ldots \otimes \boldsymbol{A}^{(r+1)} \otimes \boldsymbol{A}^{(r-1)} \otimes \ldots \otimes \boldsymbol{A}^{(1)}\right)^{\mathrm{T}}, \tag{72}$$

where $[\boldsymbol{\mathcal{R}}_{\boldsymbol{o}}]_{(r)}$ and $[\boldsymbol{\mathcal{R}}_{\boldsymbol{s}}]_{(r)}$ $(r = 1, \ldots, 2R)$, denote the n-mode unfolding of the covariance tensor $\boldsymbol{\mathcal{R}}_{\boldsymbol{o}}$ and the core tensor $\boldsymbol{\mathcal{R}}_{\boldsymbol{s}}$, respectively.

From the unfolding matrices of $\boldsymbol{\mathcal{R}}_{\boldsymbol{o}}$, an ALS-based algorithm can be formulated to estimate the desired factor matrices. An estimate of the spatial signature matrix $\hat{\boldsymbol{A}}^{(r)}$ $(r = 1, \ldots, 2R)$, associated with the r-th dimension of the covariance tensor, is obtained by solving the following LS problem:

$$\hat{\boldsymbol{A}}^{(r)} = \arg\min_{\boldsymbol{A}^{(r)}} \left\| [\boldsymbol{\mathcal{R}}_{\boldsymbol{o}}]_{(r)} - \boldsymbol{A}^{(r)} \boldsymbol{\Delta}^{(r)} \right\|_{\mathrm{F}}^{2}, \tag{73}$$

The solution of which is given by

$$\hat{\boldsymbol{A}}^{(r)} = [\boldsymbol{\mathcal{R}}_{\boldsymbol{o}}]_{(r)} \left(\boldsymbol{\Delta}^{(r)}\right)^{\dagger}. \tag{74}$$

Note, however, that the core tensor is also assumed to be unknown in the Tucker decomposition, and should also be estimated. Let $\boldsymbol{R}_{\boldsymbol{s}}$ be an unknown matrix of arbitrary structure, the following LS problem can be formulated from the vectorization of the sample covariance matrix $\hat{\boldsymbol{R}}$:

$$\operatorname{vec}\left(\hat{\boldsymbol{R}}_{\boldsymbol{s}}\right) = \arg\min_{\boldsymbol{R}_{\boldsymbol{s}}} \left\| \operatorname{vec}\left(\hat{\boldsymbol{R}}\right) - (\boldsymbol{A}^{*} \otimes \boldsymbol{A}) \operatorname{vec}\left(\boldsymbol{R}_{\boldsymbol{s}}\right) \right\|_{\mathrm{F}}^{2}, \tag{75}$$

from which an estimate of $\hat{\boldsymbol{R}}_{\boldsymbol{s}}$ can be obtained as

$$\operatorname{vec}\left(\hat{\boldsymbol{R}}_{\boldsymbol{s}}\right) = (\boldsymbol{A}^{*} \otimes \boldsymbol{A})^{\dagger} \operatorname{vec}\left(\hat{\boldsymbol{R}}\right), \tag{76}$$

where $\boldsymbol{A} = \boldsymbol{A}^{(1)} \diamond \boldsymbol{A}^{(2)} \ldots \diamond \boldsymbol{A}^{(R)} \in \mathbb{C}^{N \times M}$.

Since Equations (74) and (76) are nonlinear functions of the parameters to be estimated, the blind spatial signature estimation problem can be solved using a classical ALS iterative solution (BRO, 1998; SMILDE, BRO, and GELADI, 2004). The basic idea of the algorithm is to estimate one factor matrix at each step while the others remain fixed at the values obtained in previous steps. This procedure is repeated until convergence. The proposed generalized ALS-Tucker algorithm for R-D sensor arrays is summarized in Algorithm 6.

In this approach the factor matrices are treated as independent variables, i.e. the dual-symmetry property of the covariance tensor is not exploited. In this case, a final estimate of the spatial signature matrix associated with the r-th dimension of the array

Algorithm 6 Proposed ALS-Tucker algorithm

procedure ALS-TUCKER

1. *Randomly Initialize* $\hat{\boldsymbol{A}}^{(r)} \in \mathbb{C}^{N_r \times M}$ *for* $r = 2, \ldots, 2R$ *and the core tensor* $\hat{\boldsymbol{\mathcal{R}}}_{\boldsymbol{s}}$;
2. **for** $r = 1, \ldots, 2R$
According to (74), *obtain an LS estimate for the matrix* $\hat{\boldsymbol{A}}^{(r)}$*:*

$$\hat{\boldsymbol{A}}^{(r)} = [\boldsymbol{\mathcal{R}}]_{(r)} \left(\boldsymbol{\Delta}^{(r)}\right)^{\dagger};$$

• Note: *The matrix* $\boldsymbol{\Delta}^{(r)}$, *described in* (72), *is updated by fixing* $\hat{\boldsymbol{A}}^{(r)}$ *calculated* previously.
end
3. *According to* (76), *obtain an LS estimate for* $\hat{\boldsymbol{R}}_{\boldsymbol{s}}$:

$$\text{vec}\left(\hat{\boldsymbol{R}}_{\boldsymbol{s}}\right) = \left(\boldsymbol{A}^{*} \otimes \boldsymbol{A}\right)^{\dagger} \text{vec}\left(\hat{\boldsymbol{R}}\right);$$

$$\hat{\boldsymbol{R}}_{\boldsymbol{s}} = \text{unvec}_{M \times M}\left(\text{vec}\left(\hat{\boldsymbol{R}}_{\boldsymbol{s}}\right)\right);$$

4. *Using* (69), *reconstruct the core tensor* $\hat{\boldsymbol{\mathcal{R}}}_{\boldsymbol{s}}$ *from* $\hat{\boldsymbol{R}}_{\boldsymbol{s}}$;
5. *Repeat steps* 2-4 *until convergence.*

can be computed as:

$$\hat{\boldsymbol{A}}^{(r)}_{\text{final}} = \frac{\hat{\boldsymbol{A}}^{(r)} + \hat{\boldsymbol{A}}^{(R+r)^{*}}}{2}. \tag{77}$$

3.6.3 ALS-ProKRaft Algorithm

Here, a link is established between the alternating Procrustes estimation and Khatri-Rao factorization (ALS-ProKRaft) algorithm proposed by WEIS *et al.* (2012) and the blind spatial signature estimation problem in multidimensional sensor arrays. The main idea behind this algorithm is to exploit the dual-symmetry property of the PARAFAC decomposition described in (70). Next, the ALS-ProKRaft algorithm is formulated for the context of this thesis.

The multimode unfolding of the PARAFAC decomposition (70) can be rewritten as follows

$$\boldsymbol{R}_{\text{mm}} = \left(\boldsymbol{A}^{(1)} \diamond \ldots \diamond \boldsymbol{A}^{(R)}\right)\left(\boldsymbol{A}^{(1)} \diamond \ldots \diamond \boldsymbol{A}^{(R)}\right)^{\text{H}}, \tag{78}$$

which assumes the following factorization

$$\boldsymbol{R}_{\text{mm}} = \boldsymbol{R}_{\text{mm}}^{\frac{1}{2}} \cdot \left(\boldsymbol{R}_{\text{mm}}^{\frac{1}{2}}\right)^{\text{H}}, \tag{79}$$

Algorithm 7 ALS-ProKRaft algorithm

procedure ALS-ProKRaft
1. *Set* $i = 0$. *Initialize* $\hat{\boldsymbol{T}}_{(i=0)} = \boldsymbol{I}_M$;
2. *According to* (80), *obtain* $\boldsymbol{U}^{[M]}$ *and* $\boldsymbol{\Sigma}^{[M]}$ *from the SVD of the multimode unfolding matrix* $\boldsymbol{R}_{\text{mm}}$;
3. $i \leftarrow i + 1$;
4. *According to* (81), *obtain estimates for* $\hat{\boldsymbol{A}}_{(i)}^{(r)}$ $(r = 1, \ldots, R)$ *by applying the LS-KRF algorithm* 1 *on* $\boldsymbol{U}^{[M]}\boldsymbol{\Sigma}^{[M]}\hat{\boldsymbol{T}}_{(i-1)}^{\text{H}}$;
5. *According to* (84), *compute the SVD of* $\left(\hat{\boldsymbol{A}}_{(i)}^{(1)} \diamond \cdots \diamond \hat{\boldsymbol{A}}_{(i)}^{(R)}\right)^{\text{H}} \boldsymbol{U}^{[M]}\boldsymbol{\Sigma}^{[M]} = \boldsymbol{U}_P\boldsymbol{\Sigma}_P\boldsymbol{V}_P^{\text{H}}$ *and obtain* $\hat{\boldsymbol{T}}_{(i)} = \boldsymbol{U}_P\boldsymbol{V}_P^{\text{H}}$;
6. *Repeat steps* 3-5 *until convergence.*

where $\boldsymbol{R}_{\text{mm}}^{\frac{1}{2}} \in \mathbb{C}^{N\times M}$ can be obtained from the SVD of $\boldsymbol{R}_{\text{mm}}$ as follows

$$\boldsymbol{R}_{\text{mm}} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\text{H}}, \tag{80}$$

obeying the following structure

$$\boldsymbol{R}_{\text{mm}}^{\frac{1}{2}} \approx \boldsymbol{U}^{[M]}\boldsymbol{\Sigma}^{[M]}\boldsymbol{T}^{\text{H}} = \left(\boldsymbol{A}^{(1)} \diamond \ldots \diamond \boldsymbol{A}^{(R)}\right), \tag{81}$$

where $\boldsymbol{U}^{[M]} \in \mathbb{C}^{N\times M}$ is formed by the first M columns of $\boldsymbol{U}$ and $\boldsymbol{\Sigma}^{[M]} \in \mathbb{C}^{M\times M}$ is a diagonal matrix which contains the M dominant singular values of $\boldsymbol{R}_{\text{mm}}$. The matrix $\boldsymbol{T}$ represents a unitary rotation factor.

Equation (81) is an orthogonal Procrustes problem (OPP) (SCHONEMANN, 1966), in which $\boldsymbol{T}$ is a transformation matrix that maps $\boldsymbol{U}^{[M]}\boldsymbol{\Sigma}^{[M]}$ to $\left(\boldsymbol{A}^{(1)} \diamond \ldots \diamond \boldsymbol{A}^{(R)}\right)$ such that $\boldsymbol{U}^{[M]}\boldsymbol{\Sigma}^{[M]}\boldsymbol{T}^{\text{H}} = \left(\boldsymbol{A}^{(1)} \diamond \ldots \diamond \boldsymbol{A}^{(R)}\right)$. An efficient estimate for $\boldsymbol{T}$ is obtained minimizing the Frobenius norm of the residual error:

$$\arg\min_{\boldsymbol{T}} \left\| \boldsymbol{U}^{[M]}\boldsymbol{\Sigma}^{[M]}\boldsymbol{T}^{\text{H}} - \left(\boldsymbol{A}^{(1)} \diamond \ldots \diamond \boldsymbol{A}^{(R)}\right) \right\|_{\text{F}}. \tag{82}$$

This problem can be solved using a change of basis from the SVD of the matrix

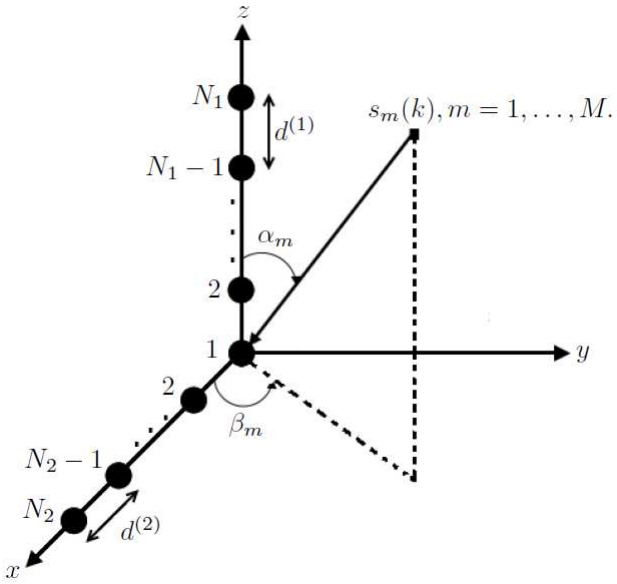
$$\left(\boldsymbol{A}^{(1)} \diamond \ldots \diamond \boldsymbol{A}^{(R)}\right)^{\text{H}} \boldsymbol{U}^{[M]}\boldsymbol{\Sigma}^{[M]} = \boldsymbol{U}_P\boldsymbol{\Sigma}_P\boldsymbol{V}_P^{\text{H}}, \tag{83}$$

which leads to the following solution according to (SCHONEMANN, 1966):

$$\hat{\boldsymbol{T}} = \boldsymbol{U}_P\boldsymbol{V}_P^{\text{H}}. \tag{84}$$

From Equations (81) and (84), an ALS-based iterative algorithm can be formulated to estimate the spatial signature matrices from the PARAFAC decomposition

Figure 14 – L-shaped array configuration with $N_1 + N_2 - 1$ sensors. The distance between the sensors in the z axis is $d^{(1)}$ while the distance between the sensors in the x axis is $d^{(2)}$. The m-th wavefront has elevation and azimuth angles equal to α_m and β_m, respectively.



Source: Created by the author.

(70). Firstly, individual estimates for each factor matrix $\hat{\boldsymbol{A}}^{(r)}$ $(r = 1, \ldots, R)$, are obtained by applying the LS-KRF algorithm 1 on the composite spatial signature matrix $\hat{\boldsymbol{A}} = \hat{\boldsymbol{A}}^{(1)} \diamond \hat{\boldsymbol{A}}^{(2)} \ldots \diamond \hat{\boldsymbol{A}}^{(R)}$. Then, the matrix $\hat{\boldsymbol{T}}$ is obtained from (84). The ALS-ProKRaft algorithm for blind spatial signature estimation in R-D sensor arrays is summarized in the pseudo-code form in Algorithm 7.

Note that compared to the conventional ALS-PARAFAC solution 2 formulated from the unimodal unfolding matrices (37), the ALS-ProKRaft algorithm becomes preferred since only half of the factors matrices needs to be estimated by exploiting the dual-symmetry property of the covariance tensor. This generally leads to a fast convergence rate of the algorithm.

3.7 Spatial Signature Estimation in L-Shaped Sensor Arrays

In this section, the blind spatial signature estimation problem is formulated for L-shaped array configuration. Considering that the receiver array is divided into smaller sub-arrays, the Tucker decomposition of a fourth-order tensor is formulated from the sample cross-correlation matrix of the data received by the different sub-arrays. From this multidimensional structure, the proposed generalized ALS-Tucker algorithm 6 can be used to estimate the sources' spatial signatures in L-shaped array scenarios.

3.7.1 Cross-Correlation Tensor for L-shaped Sensor Arrays

Consider an L-shaped sensor array equipped with N_1+N_2-1 sensors positioned in the x-z plane, as illustrated in Figure 14. Each linear array contains N_1 and N_2 sensors equally spaced at distances $d^{(1)}$ and $d^{(2)}$, respectively. We consider that each linear array is divided into P and W smaller sub-arrays, respectively. Each sub-array contains $N_1^{(\text{sub})} = N_1 - P + 1$ and $N_2^{(\text{sub})} = N_2 - W + 1$ sensors. The signal received at the p-th sub-array ($p = 1, \ldots, P$) is given by

$$\boldsymbol{X}_p^{(1)} = \boldsymbol{A_s}^{(1)} \boldsymbol{D}_p \left(\boldsymbol{\Phi}^{(1)} \right) \boldsymbol{S} + \boldsymbol{N}_p^{(1)} \in \mathbb{C}^{N_1^{(\text{sub})} \times K}, \tag{85}$$

while the signal received at the w-th sub-array ($w = 1, \ldots, W$) is given by

$$\boldsymbol{X}_w^{(2)} = \boldsymbol{A_s}^{(2)} \boldsymbol{D}_w \left(\boldsymbol{\Phi}^{(2)} \right) \boldsymbol{S} + \boldsymbol{N}_w^{(2)} \in \mathbb{C}^{N_2^{(\text{sub})} \times K}, \tag{86}$$

where $\boldsymbol{A_s}^{(r)} \in \mathbb{C}^{N_r^{(\text{sub})} \times M}$ is the spatial signature matrix of the first sub-array (or reference sub-array) for the r-th array dimension ($r = 1, 2$). $\boldsymbol{D}_p \left(\boldsymbol{\Phi}^{(1)} \right)$ and $\boldsymbol{D}_w \left(\boldsymbol{\Phi}^{(2)} \right)$ are diagonal matrices whose main diagonal is given by the p-th and w-th row of the matrices $\boldsymbol{\Phi}^{(1)} \in \mathbb{C}^{P \times M}$ and $\boldsymbol{\Phi}^{(2)} \in \mathbb{C}^{W \times M}$, respectively.

The rows of $\boldsymbol{\Phi}^{(1)}$ and $\boldsymbol{\Phi}^{(2)}$ capture the delays suffered by the signals impinging on the p-th and w-th sub-arrays with respect to the reference sub-array, which are defined based on the following spatial frequencies:

$$\mu_m^{(1)} = \frac{2\pi \cdot d^{(1)} \cdot \cos \alpha_m}{\lambda}, \tag{87}$$

$$\mu_m^{(2)} = \frac{2\pi \cdot d^{(2)} \cdot \sin \alpha_m \cdot \cos \beta_m}{\lambda}, \tag{88}$$

where α_m and β_m denote the azimuth and elevation angles of the m-th source, respectively. From (85) and (86), let us introduce now the following extended data matrices:

$$\boldsymbol{X}^{(1)} = \left[\boldsymbol{X}_1^{(1)}, \ldots, \boldsymbol{X}_P^{(1)} \right]^{\mathrm{T}} \in \mathbb{C}^{N_1^{(\text{sub})} P \times K}, \tag{89}$$

$$\boldsymbol{X}^{(2)} = \left[\boldsymbol{X}_1^{(2)}, \ldots, \boldsymbol{X}_W^{(2)} \right]^{\mathrm{T}} \in \mathbb{C}^{N_2^{(\text{sub})} W \times K}, \tag{90}$$

or, more compactly, we get

$$\boldsymbol{X}^{(r)} = \left(\boldsymbol{\Phi}^{(r)} \diamond \boldsymbol{A_s}^{(r)} \right) \boldsymbol{S} + \boldsymbol{N}^{(r)}, \quad r = 1, 2. \tag{91}$$

Similar to (67), in this approach we shall work with the following sample cross-correlation matrix:

$$\hat{\boldsymbol{R}} = \left(\boldsymbol{\Phi}^{(1)} \diamond \boldsymbol{A_s}^{(1)}\right) \boldsymbol{R_s} \left(\boldsymbol{\Phi}^{(2)} \diamond \boldsymbol{A_s}^{(2)}\right)^{\mathrm{H}} + \sigma_n^2 \boldsymbol{I}. \tag{92}$$

As aforementioned, we can see that the noiseless term in (92) denotes a multi-mode unfolding of the following cross-correlation tensor of size $N_1^{(\mathrm{sub})} \times P \times N_2^{(\mathrm{sub})} \times W$:

$$\boldsymbol{\mathcal{R}_o} = \boldsymbol{\mathcal{R}_s} \times_1 \boldsymbol{A_s}^{(1)} \times_2 \boldsymbol{\Phi}^{(1)} \times_3 \boldsymbol{A_s}^{(2)^*} \times_4 \boldsymbol{\Phi}^{(2)^*}, \tag{93}$$

where $\boldsymbol{\mathcal{R}_s} \in \mathbb{C}^{M \times M \times M \times M}$. In this modeling, the tensor $\boldsymbol{\mathcal{R}_o}$ follows a fourth-order Tucker decomposition. By analogy with (50), we can deduce the following correspondences:

$$\begin{gathered}
(I_1, I_2, I_3, I_4) \leftrightarrow \left(N_1^{(\mathrm{sub})}, P, N_2^{(\mathrm{sub})}, W\right) \\
(Q_1, Q_2, Q_3, Q_4) \leftrightarrow M \\
\left(\boldsymbol{\mathcal{G}}, \boldsymbol{A}^{(1)}, \boldsymbol{A}^{(2)}, \boldsymbol{A}^{(3)}, \boldsymbol{A}^{(4)}\right) \leftrightarrow \left(\boldsymbol{\mathcal{R}_s}, \boldsymbol{A_s}^{(1)}, \boldsymbol{\Phi}^{(1)}, \boldsymbol{A_s}^{(2)^*}, \boldsymbol{\Phi}^{(2)^*}\right).
\end{gathered} \tag{94}$$

The spatial signatures of the sources can be estimated from (93) by using the proposed generalized ALS-Tucker algorithm 6. Note that, in this approach, the ALS-Tucker algorithm is simplified to a fourth-order tensor input.

For all the previously proposed algorithms, the final estimates for the spatial signature matrices are obtained when the convergence is declared. An usually adopted criterion for convergence is defined as $|e_{(i)} - e_{(i-1)}| \leq 10^{-6}$, where $e_{(i)}$ denotes the residual error of the i-th iteration, defined as:

$$e_{(i)} = \left\| \boldsymbol{\mathcal{R}} - \hat{\boldsymbol{\mathcal{R}}}_{(i)} \right\|_{\mathrm{F}}^2, \tag{95}$$

where $\boldsymbol{\mathcal{R}} = \boldsymbol{\mathcal{R}_o} + \boldsymbol{\mathcal{N}}$ is a noisy version of the covariance tensor, $\boldsymbol{\mathcal{N}}$ is an AWGN tensor and $\hat{\boldsymbol{\mathcal{R}}}_{(i)}$ is the covariance tensor reconstructed from the estimated factor matrices and core tensor. Since the ALS-ProKRaft algorithm exploits the dual-symmetry property of the data tensor the procedure (77) is not necessary.

3.7.2 Estimation of the Spatial Frequencies

After the estimation of the spatial signatures matrices $\hat{\boldsymbol{A}}_{\mathrm{final}}^{(r)}$ $(r = 1, \ldots, R)$, the final step is to estimate the spatial frequencies of the sources $\hat{\mu}_m^{(r)}$ $(m = 1, \ldots, M)$. The final estimates can be computed from the average over the values obtained in each

row of $\hat{\boldsymbol{A}}_{\text{final}}^{(r)}$ as follows:

$$\hat{\mu}_m^{(r)} = \frac{1}{N_r - 1} \sum_{n=2}^{N_r} \frac{\arg\left\{ \hat{\boldsymbol{A}}_{\text{final}}^{(r)}(n,m) \right\}}{n-1}. \tag{96}$$

3.8 Computational Complexity

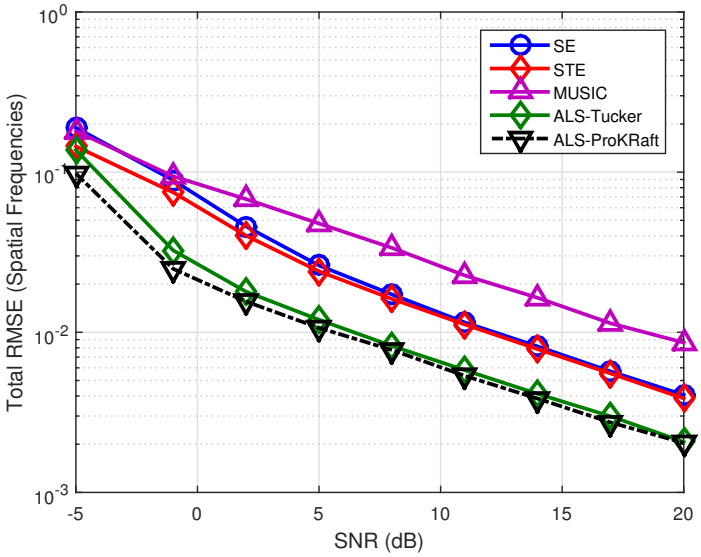
In the following, we discuss on the computational complexity of the iterative ALS-Tucker and ALS-ProKRaft algorithms. The computational complexity of one Tucker iteration refers to the cost associated with the SVD used to calculate the matrix pseudo-inverses in the least squares problems (73) and (75). The overall computational complexity per iteration of the ALS-Tucker algorithm equals the complexity of $2R$ matrix SVDs associated with each estimated factor matrix according to (74) plus the complexity of one additional matrix SVD associated with the estimated core tensor according to (76). The overall computational cost per iteration of the ALS-ProKRaft algorithm equals the complexity of $M(R-1)$ matrix SVDs associated with the application of the LS-KRF algorithm in (81) plus the complexity of one additional matrix SVD associated with the update of the unknown unitary rotation factor matrix $\boldsymbol{T}$ in (84).

3.9 Advantages and Disadvantages of the Proposed Methods

In this section, we discuss on the advantages and disadvantages of the proposed methods for blind spatial signatures estimation in R-D sensor arrays. As previously stated in Section 3.6.3, the ALS-ProKRaft algorithm works on the assumption that the sample covariance matrix of the sources $\boldsymbol{R_s}$ is perfectly known and diagonal. However, this is only true in the asymptotic case when a sufficiently large number of snapshots is assumed (i.e. $K \to \infty$), as well as when the source signals are perfectly uncorrelated. In practice, this assumption is not guaranteed. On the other hand, the ALS-Tucker algorithm previously formulated in Section 3.6.2 naturally captures any structure for the sources covariance into the core tensor $\boldsymbol{\mathcal{R}_s}$. Therefore, the assumption of uncorrelated source signals is not necessary for the ALS-Tucker algorithm, making it able to operate in scenarios where the source covariance structure is unknown and arbitrary (non-diagonal). Such scenarios occur, for instance, when the sample covariance is computed from a limited number of snapshots.

The ALS-Tucker algorithm does not exploit the dual-symmetry property of the data covariance tensor and all factor matrices and the core tensor needs to be estimated as independent variables. In contrast, in the ALS-ProKRaft algorithm, only half of the factor matrices are estimated by exploiting the dual-symmetry property of the covariance tensor. Therefore, the ALS-ProKRaft algorithm is more computationally attractive than the ALS-Tucker algorithm. When compared with the state-of-the-art matrix based algorithms such

Figure 15 – Total RMSE vs. SNR for $N = 64$ sensors, $K = 10$ samples, DoAs: $\{30°, 55°\}$ and $\{45°, 60°\}$ for Hadamard sources' sequences. Uniform retangular array (URA).



Source: Created by the author.

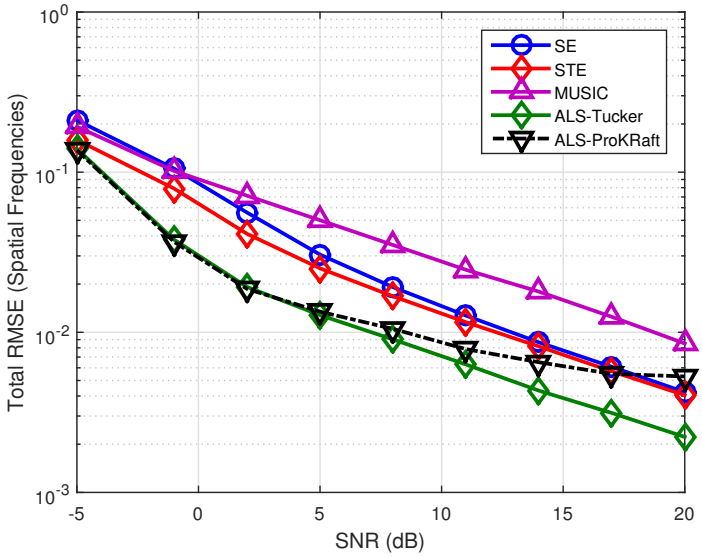
as MUSIC, ESPRIT and Propagator Method, the proposed tensor-based algorithms have the advantage of fully exploiting the multidimensional nature of the received signal in less specific scenarios, which leads to more accurate estimates. For instance, the ESPRIT algorithm was originally formulated for sensor arrays that must obey the shift invariance property. On the other hand, the MUSIC algorithm has high computational complexity due to the search of parameters in the spatial spectrum.

3.10 Performance Evaluation of ALS-Tucker and ALS-ProKraft Algorithms

In this section, the performance of the iterative ALS-Tucker and ALS-ProKRaft algorithms are investigated through computer simulations. Fisrt, the numerical results related to Section 3.6 are presented and discussed. Then, the same methodology is adopted for the L-shaped modeling developed in Section 3.7. Final results are obtained as an average of 1000 independent Monte Carlo runs. Initially, we consider a uniform rectangular array (URA) positioned on the x-z plane. The m-th wavefront has direction of arrival $\{\alpha_m, \beta_m\}_{m=1}^{M}$, where α_m and β_m are the elevation and azimuth angles, respectively.

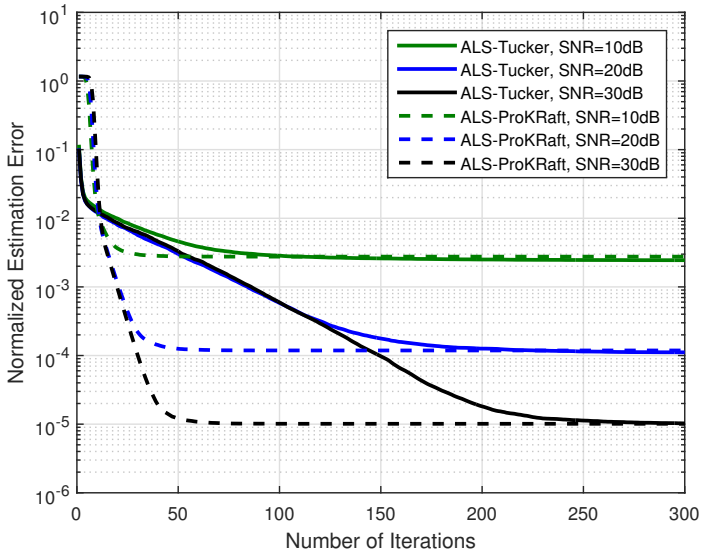
In Figures 15 and 16, the performance is measured in terms of total RMSE (65) of the estimated spatial frequencies $\hat{\mu}_m^{(r)}$ as a function of the SNR. The relations between directions of arrival and spatial frequencies are given in (87) and (88), where $d^{(r)}$ denotes the distance between sensors in the r-th array dimension, which is assumed here to be equals $\lambda/2$. In Figure 15, we consider Hadamard sequences for the sources' symbols, while in Figure 16 we consider BPSK modulated sources' symbols. In both cases, we assume $N = 64$ sensors (i.e. N_1 and N_2 equal to 8) and the sample covariance matrix of the

Figure 16 – Total RMSE vs. SNR for $N = 64$ sensors, $K = 10$ samples, DoAs: $\{30°, 55°\}$ and $\{45°, 60°\}$ for BPSK sources' sequences. Uniform retangular array (URA).



Source: Created by the author.

Figure 17 – Convergence of the iterative ALS-Tucker and ALS-ProKRaft algorithms. Uniform retangular array (URA).

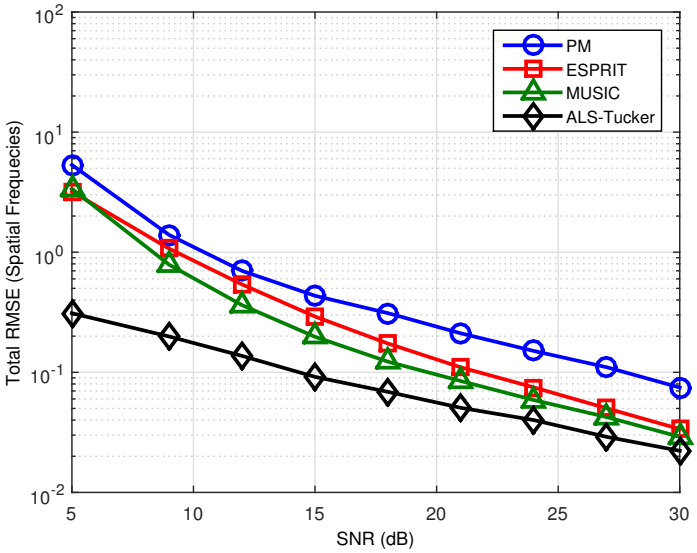


Source: Created by the author.

received data (67) computed from a reduced number of $K = 10$ samples.

From Figure 15, it can be seen that both ALS-Tucker and ALS-ProKRaft algorithms have similar performances in terms of total RMSE, when Hadamard sources' sequences are considered. On the other hand, in Figure 16, when BPSK sources' symbols are considered, a floor is exhibited by the ALS-ProKraft algorithm for high SNR values. This behavior occurs due to the modeling errors in the core tensor of the PARAFAC decomposition, which in turn arises due to the non-orthogonality of the source signals, resulting in a non-diagonal sample covariance matrix of the sources in this case. Note that

Figure 18 – Total RMSE vs. SNR for $N = 13$ sensors, $K = 500$ samples, DoAs: $\{30°, 45°\}$ and $\{50°, 55°\}$. L-Shaped array.



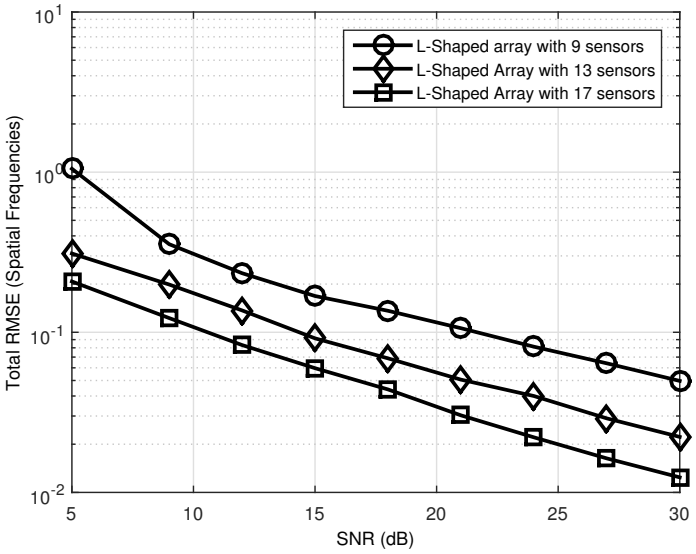
Source: Created by the author.

in the ALS-Tucker algorithm the covariance matrix of the sources possess an arbitrary and unknown structure which discards possible constraints to the source signals. This difference makes the Tucker decomposition approach more attractive in those practical scenarios in which source uncorrelatedness is not guaranteed. When compared to state-of-the-art matrix-based standard ESPRIT (SE) (ROY and KAILATH, 1989) adapted to planar arrays, matrix-based MUSIC algorithm to planar array configuration (SEKIZAWA, 1998) and standard Tensor-ESPRIT (STE) (HAARDT, ROEMER, and DEL GALDO, 2008), the proposed algorithms have improved accuracy in all the considered scenarios.

Figure 17 shows the convergence performance of the iterative algorithms. In this experiment, the median values of the normalized estimation error $e_{(i)}/N^{(2R)}$ are plotted in terms of the number of iterations for different SNR values. It is noteworthy that the ALS-ProKRaft algorithm has a very faster convergence compared to the ALS-Tucker algorithm. This behavior is expected since ALS-ProKRaft exploits the dual-symmetry property of the data tensor, which results in estimating half as many factor matrices compared to the ALS-Tucker approach.

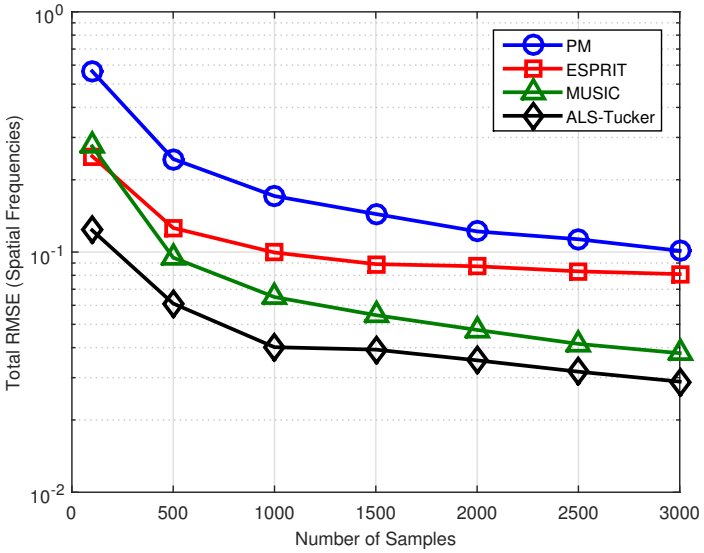
Now, we consider an L-shaped array configuration. In Figure 18, we set $N = 13$ sensors (i.e. N_1 and N_2 equal to 7) and $K = 500$ samples. Each uniform linear array is divided into $P = 2$ and $W = 2$ sub-arrays, respectively. In this experiment, the performance of the proposed ALS-Tucker algorithm is compared to the state-of-the-art matrix-based methods, namely, Propagator Method (PM) (TAYEM and KWON, 2005), MUSIC (CHANGUEL, HARABI, and GHARSALLAH, 2006) and ESPRIT (ZHANG, GAO, and CHEN, 2009), all of them originally formulated for L-shaped arrays. Note that the ALS-Tucker algorithm presents an improved performance over its competitors,

Figure 19 – Total RMSE vs. SNR (performance of the ALS-Tucker algorithm for different number os sensors). L-Shaped array.



Source: Created by the author.

Figure 20 – Total RMSE vs. number of samples. L-Shaped array.



Source: Created by the author.

with more evidenced gains in the low-to-medium SNR range. For high SNR values, the performance of the MUSIC method comes close to that of our proposal. However, the ALS-Tucker algorithm dispense any estimation procedure via bidimensional peak search as occurs with MUSIC, being the former more computationally attractive.

Figure 19 shows the performance of the ALS-Tucker by assuming different number of sensors. In this experiment, we consider $K = 500$ samples. We can observe a better performance in terms of RMSE when the number sensors of the L-shaped array is increased. This is valid for all the simulated SNR values.

In Figure 20, we analyze the influence of the number of samples K on the performance of the ALS-Tucker algorithm. This experiment considers the same parameters as the experiment of Figure 18, except the SNR value that is assumed fixed at 20 dB and the number of samples that varies between 50 and 3000. First, we can see that the performance of the algorithms improves by increasing the number of samples collected by sensor array, as expected. However, similar to Figure 18 the proposed ALS-Tucker algorithm outperforms the state-of-the-art PM, ESPRIT and MUSIC methods.

3.11 Chapter Summary

In this chapter, we have addressed the spatial frequencies estimation problem in multidimensional sensors arrays. Initially, we have proposed a tensor-based multiple denoising (MuDe) method that enhances parameter estimation in R-D sensor array processing. The MuDe exploits all possible subarrays created by an R-D spatial smoothing and achieves a higher noise reduction capability with successive low-rank approximations in tensor and matrix data. For a reduced number of sources close to one, the MuDe has a remarkable performance achieving the bound derived in (STEINWANDT, ROEMER, and HAARDT, 2014). Our proposed approach considerably outperforms state-of-the-art solutions for scenarios with a high number of sources.

In addition, we also have proposed two tensor-based approaches based on the Tucker and PARAFAC decompositions to solve the blind spatial signatures estimation problem. One is a covariance-based generalization of the Tucker decomposition for R-D sensor arrays. The other one, is a link between the ALS-ProKRaft algorithm and covariance-based array processing for blind spatial signatures estimation. As another contribution, we have formulated a cross-correlation-based fourth-order Tucker decomposition which makes the proposed ALS-Tucker algorithm applicable in scenarios composed by L-Shaped array configurations. The two proposed tensor methods differ in the structure assumed for the source covariance. It is worth pointing out that in realistic scenarios, when the received covariance matrix is calculated from a reduced number of samples, or snapshots, the ALS-Tucker algorithm becomes preferable since it operates with an arbitrary and unknown structure for the covariance of the source signals. On the other hand, when the sources can be assumed to be uncorrelated, we can achieve improved performance by exploiting the dual-symmetry property of the covariance tensor, which makes the ALS-ProKRaft algorithm preferable since it provides good estimation accuracy with a smaller number of iterations. When compared with other state-of-the-art matrix-based and tensor-based techniques, the proposed tensor-based iterative algorithms have shown their effectiveness with remarkable gains in terms of estimation accuracy.

4 A NESTED-PARAFAC BASED APPROACH FOR TARGET LOCALIZATION IN BISTATIC MIMO RADAR SYSTEMS

In this chapter, we make use of the Nested-PARAFAC decomposition to formulate a tensor-based method for joint direction of departure (DoD) and direction of arrival (DoA) estimation in bistatic MIMO radar systems. By assuming that the transmit array is divided into two maximally overlapping subarrays, we initially model the cross-covariance matrix of the matched filters outputs as a Nested-PARAFAC decomposition of a fourth-order covariance tensor. Then, exploiting the algebraic structure of this decomposition, we first propose a two stage algorithm for joint DoD and DoA estimation of multiple targets based on double alternating least squares (DALS). In addition, for scenarios in which the number of receive antennas exceeds the number of targets, we propose a closed-form solution to the second stage of the proposed method based on the LS-KRF concept. Simulation results show that the proposed method offers a highly-accurate localization of multiple targets in real-world scenarios where the antenna elements at the transmit and receive arrays have positioning errors as well as less complexity compared to competing state-of-the-art tensor-based solutions.

4.1 Introduction and Motivation

In recent years, research on target detection and localization techniques for MIMO radar systems has drawn great attention in the signal processing community (ZHENG, ZHANG, and ZHANG, 2012; GU, HE, and ZHU, 2013; JIANG *et al.*, 2013). In contrast to conventional phased-array radars that emits coherent waveforms, MIMO radar systems simultaneously transmit orthogonal waveforms using multiple antennas, increasing the performance of parameter estimation compared to previous ones (WANG *et al.*, 2014). Regarding the configuration of transmit and receive antennas, MIMO radars are divided in statistical (HAIMOVICH, BLUM, and CIMINI, 2008) and colocated (LI and STOICA, 2007) MIMO radars. The former assumes that transmit and receive antennas are widely spaced, while the latter assumes antennas closely spaced, resulting in monostatic (TAN, NIE, and WEN, 2017) or bistatic (LIU and LIAO, 2010) MIMO radar schemes. Throughout this chapter, we focus on the bistatic MIMO radar case, where the direction of departure (DoD) and direction of arrival (DoA) of the targets need to be jointly estimated.

Many techniques to solve the multiple targets localization problem in MIMO radar are based on traditional eigenspace algorithms, such as MUSIC (XIE, LIU, and WU, 2011; LI *et al.*, 2013) and ESPRIT (CHEN, GU, and SU, 2010; JIN, LIAO, and LI, 2009). However, MUSIC-based methods need a bidimensional search in the spatial spectrum to find the DoDs and DoAs of the targets. In this case, the accuracy of the estimated pa-

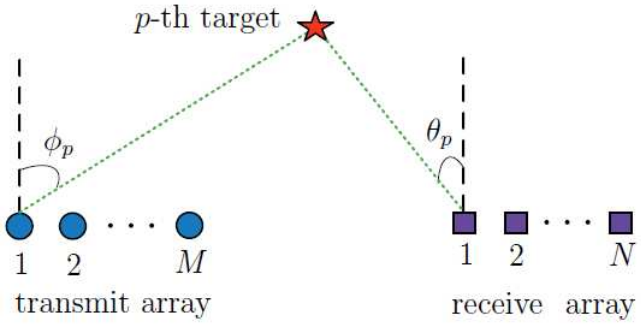
rameters is directly related to an increase in the computational cost. On the other hand, ESPRIT-based methods avoid the exhaustive peak search by exploiting the shift invariance property of the antenna array. However, this approach implies constraints on the geometry of the transmit and receive arrays, and a performance degradation is observed when the antenna elements have positioning errors. In (NION and SIDIROPOULOS, 2010), a direct-data tensor formulation based on the PARAFAC decomposition for joint DoD and DoD estimation in MIMO radar has been proposed. However, this approach becomes prohibitive when a large number of snapshots are processed at the receiver, motivating in this way, a tensor formulation based on second-order statistics which are long-term parameters that can be assumed to be constant over a larger time-scale. In addition, (NION and SIDIROPOULOS, 2010) also provides a performance analysis of tensor-based and matrix-based approaches, but it does not consider real-world scenarios where calibration errors are present at the transmit and receive arrays.

In this chapter, we make use of the Nested-PARAFAC decomposition (see Section 2.4.2) to formulate a tensor-based method for joint DoD and DoA estimation in bistatic MIMO radar systems. By assuming that the transmit array is divided into two maximally overlapping subarrays, we show that the cross-correlation matrix of the received signals after the matched filters can be modeled as a fourth-order tensor following a Nested-PARAFAC decomposition. Exploiting the multilinear structure of this decomposition, a two stage double alternating least squares (DALS) estimator is formulated from the inner and outer third-order PARAFAC parts of the Nested-PARAFAC model to solve the multi-target localization problem. In contrast to classical matrix-based and tensor-based ESPRIT versions (ROY and KAILATH, 1989; HAARDT, ROEMER, and DEL GALDO, 2008), the proposed method does not impose any constraints on the geometry of the transmit and receive arrays. Moreover, it divides a high complexity fourth-order estimation problem into two third-order subproblems leading to a reduction in the processing time and complexity for parameter estimation when compared to state-of-the-art direct data (NION and SIDIROPOULOS, 2010) and covariance-based (GOMES *et al.*, 2017; WANG *et al.*, 2018) solutions. Our simulation results also show that the proposed tensor-based receiver is capable of jointly estimating the spatial parameters of multiple targets with good accuracy even when the antenna arrays are vulnerable to calibration errors or have arbitrary geometries.

4.1.1 Chapter Organization

This chapter is organized as follows. In Section 4.2, the basic signal model for a bistatic MIMO radar system is presented. In Section 4.3, the cross-covariance tensor of the matched filters outputs for two transmit subarrays is modeled as a Nested-PARAFAC decomposition. Then, the proposed two-stage tensor-based algorithm for joint DoD and

Figure 21 – Bistatic MIMO radar scenario.



Source: Adapted from (WANG *et al.*, 2018).

DoA estimation is formulated. Uniqueness issues of the proposed Nested-PARAFAC based estimator are discussed in Section 4.4. In Section 4.5, the computational complexities of proposed and competing state-of-the-art algorithms are discussed. Simulation results are provided in Section 4.6 and the chapter is concluded in Section 4.7.

4.2 Signal Model

We consider a narrowband bistatic MIMO radar system in which the transmit and receive arrays have M and N antenna elements, respectively. This scenario is shown in Figure 21. At the transmit side, each antenna element emits one waveform. The M transmitted waveforms have identical bandwidth and center frequency but are temporally orthogonal. The waveform vector transmitted by the m-th array element within one repetition interval is denoted by $\boldsymbol{s}_m \in \mathbb{C}^{1\times K}$, $m = 1, \ldots, M$, where K is the number of samples per pulse period and $\boldsymbol{s}_m \boldsymbol{s}_m^{\mathrm{H}} = K$. We also assume that P targets in the far-field of the transmit and receive arrays, with different Doppler frequencies are located at the same range bin of interest. The location of the p-th target ($p = 1, \ldots, P$) is denoted by (ϕ_p, θ_p), where ϕ_p and θ_p are the DoD and DoA with respect to the transmit and receive arrays, respectively. The received signal $\boldsymbol{X}^{(l)} \in \mathbb{C}^{N\times K}$ at the output of the receive array in the l-th pulse period, $l = 1, \ldots, L$, from reflections of P targets is given by (CHEN, GU, and SU, 2010)

$$\boldsymbol{X}^{(l)} = \sum_{p=1}^{P} \boldsymbol{a}_{\mathrm{r}}(\theta_p)\beta_p^{(l)}\boldsymbol{a}_{\mathrm{t}}^{\mathrm{T}}(\phi_p) \begin{bmatrix} \boldsymbol{s}_1 \\ \vdots \\ \boldsymbol{s}_M \end{bmatrix} e^{j2\pi f_{\mathrm{D}p} t_l} + \boldsymbol{N}^{(l)}, \qquad l = 1, \ldots, L \tag{97}$$

where $\beta_p^{(l)}$ denotes the complex-valued reflection coefficient of the p-th target assumed constant during a pulse period, $f_{\mathrm{D}p}$ is the Doppler frequency of the p-th target, t_l is the slow time index of the l-th pulse period and $\boldsymbol{N}^{(l)} \in \mathbb{C}^{N\times K}$ denotes the AWGN contribution.

The steering vectors of the transmit and receive arrays with respect to the p-th target, $\boldsymbol{a}_{\mathrm{t}}(\phi_p) \in \mathbb{C}^{M\times 1}$ and $\boldsymbol{a}_{\mathrm{r}}(\theta_p) \in \mathbb{C}^{N\times 1}$, assume arbitrary geometries, and are computed as functions of the angular parameters ϕ_p and θ_p, respectively. For instance, considering a linear array with positioning errors between the antenna elements, the m-th and n-th entries of $\boldsymbol{a}_{\mathrm{t}}(\phi_p)$ and $\boldsymbol{a}_{\mathrm{r}}(\theta_p)$ are modeled as

$$[\boldsymbol{a}_{\mathrm{t}}(\phi_p)]_m = e^{j\frac{2\pi}{\lambda}[(m-1)d_t+\alpha\epsilon_m]\sin\phi_p} \quad \text{and} \quad [\boldsymbol{a}_{\mathrm{r}}(\theta_p)]_n = e^{j\frac{2\pi}{\lambda}[(n-1)d_r+\alpha\epsilon_n]\sin\theta_p}, \tag{98}$$

where λ denotes the wavelength of the transmitted signal, d_t and d_r are the inter-element spacing of the transmit and receive arrays, α denotes a positioning error factor, while ϵ_m and ϵ_n are zero mean real Gaussian random variables assumed different for each value of $m = 1,\ldots,M$ and $n = 1\ldots,N$. In contrast to real-world scenarios in which positioning errors are present in both transmit and receive arrays, the ESPRIT-based algorithms (ROY and KAILATH, 1989; HAARDT, ROEMER, and DEL GALDO, 2008) assume antenna arrays that satisfy the shift invariance property and have good accuracy (only valid in the ideal case, when $\alpha \to 0$).

Due to the orthogonality of the transmitted waveforms, the received signal (97) can be matched to the m-th waveform $\boldsymbol{s}_m$, $m = 1,\ldots,M$. The output of the matched filters w.r.t. the m-th transmitted waveform is given by

$$\boldsymbol{y}_m^{(l)} = \frac{1}{K}\boldsymbol{X}^{(l)}\boldsymbol{s}_m^{\mathrm{H}} = \boldsymbol{A}_{\mathrm{r}}\boldsymbol{D}_m\left(\boldsymbol{A}_{\mathrm{t}}\right)\boldsymbol{\gamma}^{(l)} + \boldsymbol{w}_m^{(l)} \in \mathbb{C}^{N\times 1}, \tag{99}$$

where $\boldsymbol{A}_{\mathrm{t}} = [\boldsymbol{a}_{\mathrm{t}}(\phi_1),\ldots,\boldsymbol{a}_{\mathrm{t}}(\phi_P)] \in \mathbb{C}^{M\times P}$ denotes the transmit steering matrix, $\boldsymbol{A}_{\mathrm{r}} = [\boldsymbol{a}_{\mathrm{r}}(\theta_1),\ldots,\boldsymbol{a}_{\mathrm{r}}(\theta_P)] \in \mathbb{C}^{N\times P}$ is the receive steering matrix, the vector gains is denoted by $\boldsymbol{\gamma}^{(l)} = \left[\beta_1^{(l)}e^{j2\pi f_{\mathrm{D}1}t_l},\ldots,\beta_P^{(l)}e^{j2\pi f_{\mathrm{D}P}t_l}\right]^{\mathrm{T}} \in \mathbb{C}^{P\times 1}$ and $\boldsymbol{w}_m^{(l)} = \frac{1}{K}\boldsymbol{N}^{(l)}\boldsymbol{s}_m^{\mathrm{H}} \in \mathbb{C}^{N\times 1}$ is the noise contribution at the matched filter output associated with the m-th transmitted waveform.

Our aim in this chapter is to locate the P targets in the same range bin of interest by estimating their DoDs and DoAs. In the following, by starting from the output signals of the matched filters in (99), we formulate the proposed two-stage tensor-based method for the joint estimation of the spatial target parameters.

4.3 Proposed Tensor-Based Method for Joint DoD and DoA Estimation

Let us consider that the transmit array is divided into two smaller subarrays. We assume maximally overlapping subarrays, i.e., the first subarray contains the first $M-1$ antenna elements of the transmit array, while the second subarray contains its last $M-1$ antenna elements. The received signal at the output of the matched filters

associated with the first subarray is given by

$$\boldsymbol{z}_1^{(l)} = \begin{bmatrix} \boldsymbol{y}_1^{(l)} \\ \vdots \\ \boldsymbol{y}_{M-1}^{(l)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_{\mathrm{r}} \boldsymbol{D}_1 \left(\boldsymbol{A}_{\mathrm{t}} \right) \boldsymbol{\gamma}^{(l)} \\ \vdots \\ \boldsymbol{A}_{\mathrm{r}} \boldsymbol{D}_{M-1} \left(\boldsymbol{A}_{\mathrm{t}} \right) \boldsymbol{\gamma}^{(l)} \end{bmatrix} + \begin{bmatrix} \boldsymbol{w}_1^{(l)} \\ \vdots \\ \boldsymbol{w}_{M-1}^{(l)} \end{bmatrix}$$

$$= \left(\boldsymbol{A}_{\mathrm{t},1} \diamond \boldsymbol{A}_{\mathrm{r}} \right) \boldsymbol{\gamma}^{(l)} + \boldsymbol{w}_{\mathrm{MF1}}^{(l)} \in \mathbb{C}^{(M-1)N \times 1}, \tag{100}$$

where $\boldsymbol{A}_{\mathrm{t},1} = [\boldsymbol{a}_{\mathrm{t},1}(\phi_1), \ldots, \boldsymbol{a}_{\mathrm{t},1}(\phi_P)] \in \mathbb{C}^{M-1 \times P}$ denotes the transmit steering matrix of the first subarray. In a similar way, the output signals associated with the second subarray can be written as

$$\boldsymbol{z}_2^{(l)} = \begin{bmatrix} \boldsymbol{y}_2^{(l)} \\ \vdots \\ \boldsymbol{y}_{M}^{(l)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_{\mathrm{r}} \boldsymbol{D}_2 \left(\boldsymbol{A}_{\mathrm{t}} \right) \boldsymbol{\gamma}^{(l)} \\ \vdots \\ \boldsymbol{A}_{\mathrm{r}} \boldsymbol{D}_{M} \left(\boldsymbol{A}_{\mathrm{t}} \right) \boldsymbol{\gamma}^{(l)} \end{bmatrix} + \begin{bmatrix} \boldsymbol{w}_2^{(l)} \\ \vdots \\ \boldsymbol{w}_{M}^{(l)} \end{bmatrix}$$

$$= \left(\boldsymbol{A}_{\mathrm{t},2} \diamond \boldsymbol{A}_{\mathrm{r}} \right) \boldsymbol{\gamma}^{(l)} + \boldsymbol{w}_{\mathrm{MF2}}^{(l)} \in \mathbb{C}^{(M-1)N \times 1}, \tag{101}$$

where $\boldsymbol{A}_{\mathrm{t},2} = [\boldsymbol{a}_{\mathrm{t},2}(\phi_1), \ldots, \boldsymbol{a}_{\mathrm{t},2}(\phi_P)] \in \mathbb{C}^{M-1 \times P}$ denotes the transmit steering matrix of the second subarray. The vector $\boldsymbol{w}_{\mathrm{MF}i}^{(l)} \in \mathbb{C}^{(M-1)N \times 1}$ is the noise term at the matched filter output with respect to the i-th transmit subarray ($i = 1, 2$).

From (100) and (101), the cross-covariance matrix $\boldsymbol{R} \in \mathbb{C}^{(M-1)N \times (M-1)N}$ between the subarrays output signals is given by

$$\boldsymbol{R} = E\left\{ \boldsymbol{z}_2^{(l)} \boldsymbol{z}_1^{(l)\mathrm{H}} \right\}, \quad l = 1, \ldots, L$$

$$= \left(\boldsymbol{A}_{\mathrm{t},2} \diamond \boldsymbol{A}_{\mathrm{r}} \right) \boldsymbol{R}_{\gamma} \left(\boldsymbol{A}_{\mathrm{t},1} \diamond \boldsymbol{A}_{\mathrm{r}} \right)^{\mathrm{H}} + \boldsymbol{R}_{w}, \tag{102}$$

where $\boldsymbol{R}_{\gamma} = E\left\{ \boldsymbol{\gamma}^{(l)} \boldsymbol{\gamma}^{(l)\mathrm{H}} \right\} \in \mathbb{C}^{P \times P}$. The term $\boldsymbol{R}_{w} = E\left\{ \boldsymbol{w}_{\mathrm{MF2}}^{(l)} \boldsymbol{w}_{\mathrm{MF1}}^{(l)\mathrm{H}} \right\} \in \mathbb{C}^{(M-1)N \times (M-1)N}$ denotes the cross-covariance matrix of the unknown noise contribution.

Note that the cross-covariance matrix $\boldsymbol{R}$ in (102) can be viewed as a multimode unfolding of the following fourth-order cross-covariance tensor $\boldsymbol{\mathcal{R}} \in \mathbb{C}^{N \times M-1 \times N \times M-1}$

$$\boldsymbol{\mathcal{R}} = \boldsymbol{\mathcal{R}}_{\gamma} \times_1 \boldsymbol{A}_{\mathrm{r}} \times_2 \boldsymbol{A}_{\mathrm{t},2} \times_3 \boldsymbol{A}_{\mathrm{r}}^{*} \times_4 \boldsymbol{A}_{\mathrm{t},1}^{*} + \boldsymbol{\mathcal{R}}_{w}, \tag{103}$$

where $\boldsymbol{\mathcal{R}}_{\gamma} \in \mathbb{C}^{P \times P \times P \times P}$ and $\boldsymbol{\mathcal{R}}_{w} \in \mathbb{C}^{N \times M-1 \times N \times M-1}$ are obtained by "tensorizing" $\boldsymbol{R}_{\gamma}$ and $\boldsymbol{R}_{w}$ as fourth-order tensors of sizes $P \times P \times P \times P$ and $N \times M-1 \times N \times M-1$, respectively. Indeed, the noiseless part of (103) satisfies a fourth-order Tucker decomposition. By comparing (103) with (50), the following correspondences can be deduced:

$$\left(I_1, I_2, I_3, I_4, Q_{n=1,2,3,4} \right) \leftrightarrow \left(N, M-1, N, M-1, P \right),$$

$$\left(\boldsymbol{\mathcal{G}}, \boldsymbol{A}^{(1)}, \boldsymbol{A}^{(2)}, \boldsymbol{A}^{(3)}, \boldsymbol{A}^{(4)} \right) \leftrightarrow \left(\boldsymbol{\mathcal{R}}_{\gamma}, \boldsymbol{A}_{r}, \boldsymbol{A}_{t,2}, \boldsymbol{A}_{r}^{*}, \boldsymbol{A}_{t,1}^{*} \right).$$

The estimation of the targets' spatial parameters (DoDs and DoAs) can be obtained using different approaches such as ESPRIT-based algorithms (ROY and KAILATH, 1989; HAARDT, ROEMER, and DEL GALDO, 2008), PARAFAC-based direct-data algorithms (NION and SIDIROPOULOS, 2010), and ALS-Tucker4 algorithms (GOMES *et al.*, 2017; WANG *et al.*, 2018). The first approach (ROY and KAILATH, 1989; HAARDT, ROEMER, and DEL GALDO, 2008) assumes perfect array calibration. Array imperfections in real-world scenarios may severely degrade the estimation accuracy of these methods. Although the direct-data algorithm in (NION and SIDIROPOULOS, 2010) provide good estimation accuracy by applying the ALS-PARAFAC algorithm 2 directly to the received data tensor, its computational complexity may be high when the data block size is large. In the ALS-Tucker4 algorithm proposed by GOMES *et al.* (2017), five LS estimation steps are necessary to estimate the factor matrices and the core tensor from the cross-correlation tensor (103). However, this approach also has high computational cost and may present slow convergence due the wide search space to alternately minimize a cost function with five unknown variables, represented by the core tensor and four factor matrices that model the covariance tensor. In the following, we formulate a new algorithm for joint DoD and DoA estimation that is based on a two-stage Nested-PARAFAC modeling approach. As will be clear in the sequel, the Nested-PARAFAC approach breaks down the complex fourth-order tensor problem (103) into two smaller third-order tensor subproblems, without affecting the accuracy of the parameters estimation.

4.3.1 Proposed Nested-PARAFAC Based Modeling

The cross-covariance matrix $\boldsymbol{R}$ in (102) can now be viewed as an unfolding matrix of a third-order cross-covariance tensor $\boldsymbol{\mathcal{R}} \in \mathbb{C}^{(M-1)N \times N \times M-1}$ expressed using the n-mode product notation as

$$\boldsymbol{\mathcal{R}} = \boldsymbol{\mathcal{I}}_{3,P} \times_1 [\boldsymbol{\mathcal{A}}]^{\mathrm{T}}_{(1)} \times_2 \boldsymbol{A}^{*}_{\mathrm{r}} \times_3 \boldsymbol{A}^{*}_{\mathrm{t},1} + \boldsymbol{\mathcal{R}}_w, \tag{104}$$

where $\boldsymbol{\mathcal{R}}_w \in \mathbb{C}^{(M-1)N \times N \times M-1}$ is now a third-order cross-covariance noise tensor. The factor matrix $[\boldsymbol{\mathcal{A}}]^{\mathrm{T}}_{(1)} \in \mathbb{C}^{(M-1)N \times P}$ in (104) denotes the transpose of the 1-mode unfolding of the following third-order PARAFAC decomposition

$$\boldsymbol{\mathcal{A}} = \boldsymbol{\mathcal{I}}_{3,P} \times_1 \boldsymbol{R}^{\mathrm{T}}_{\gamma} \times_2 \boldsymbol{A}_{\mathrm{r}} \times_3 \boldsymbol{A}_{\mathrm{t},2}. \tag{105}$$

Indeed, the noiseless term of the cross-covariance tensor (104) corresponds to the Nested-PARAFAC decomposition of a fourth-order tensor, according to Section 2.4.2.

By analogy with (40) and (41), we can deduce the following correspondences:

$$\underbrace{\left(\boldsymbol{\mathcal{X}}, [\boldsymbol{\mathcal{Y}}]_{(1)}, \boldsymbol{A}^{(2)}, \boldsymbol{A}^{(3)}\right) \leftrightarrow \left(\boldsymbol{\mathcal{R}}, [\boldsymbol{\mathcal{A}}]_{(1)}^{\mathrm{T}}, \boldsymbol{A}_r^*, \boldsymbol{A}_{t,1}^*\right)}_{\text{outer PARAFAC part}},$$

$$\underbrace{\left(\boldsymbol{\mathcal{Y}}, \boldsymbol{B}^{(1)}, \boldsymbol{B}^{(2)}, \boldsymbol{B}^{(3)}\right) \leftrightarrow \left(\boldsymbol{\mathcal{A}}, \boldsymbol{R}_\gamma^{\mathrm{T}}, \boldsymbol{A}_r, \boldsymbol{A}_{t,2}\right)}_{\text{inner PARAFAC part}}.$$

4.3.2 First ALS Stage (Trilinear ALS)

According to (29), (30) and (31), the outer PARAFAC part (104) admits the following representations in terms of its 1-mode, 2-mode and 3-mode unfolding matrices

$$[\boldsymbol{\mathcal{R}}]_{(1)} = [\boldsymbol{\mathcal{A}}]_{(1)}^{\mathrm{T}} \left(\boldsymbol{A}_{\mathrm{t},1}^* \diamond \boldsymbol{A}_{\mathrm{r}}^*\right)^{\mathrm{T}} + [\boldsymbol{\mathcal{R}}_w]_{(1)}, \tag{106}$$

$$[\boldsymbol{\mathcal{R}}]_{(2)} = \boldsymbol{A}_{\mathrm{r}}^* \left(\boldsymbol{A}_{\mathrm{t},1}^* \diamond [\boldsymbol{\mathcal{A}}]_{(1)}^{\mathrm{T}}\right)^{\mathrm{T}} + [\boldsymbol{\mathcal{R}}_w]_{(2)}, \tag{107}$$

$$[\boldsymbol{\mathcal{R}}]_{(3)} = \boldsymbol{A}_{\mathrm{t},1}^* \left(\boldsymbol{A}_{\mathrm{r}}^* \diamond [\boldsymbol{\mathcal{A}}]_{(1)}^{\mathrm{T}}\right)^{\mathrm{T}} + [\boldsymbol{\mathcal{R}}_w]_{(3)}. \tag{108}$$

Assuming that the number of targets is known or has been previously estimated using, e.g., the methods proposed in (DA COSTA *et al.*, 2011; LIU *et al.*, 2016), estimates of the factor matrices $[\boldsymbol{\mathcal{A}}]_{(1)}^{\mathrm{T}}$, $\boldsymbol{A}_{\mathrm{r}}^*$ and $\boldsymbol{A}_{\mathrm{t},1}^*$ can be obtained from the outer PARAFAC part using the ALS-PARAFAC algorithm 2. The first trilinear ALS (TALS) stage of the proposed method consists of estimating the matrices of interest in an alternating way from the unfolding matrices $[\boldsymbol{\mathcal{R}}]_{(n=1,2,3)}$ by solving three linear LS problems from (106), (107) and (108), whose analytic solutions are given by $[\hat{\boldsymbol{\mathcal{A}}}]_{(1)}^{\mathrm{T}} = [\boldsymbol{\mathcal{R}}]_{(1)} \left[\left(\boldsymbol{A}_{\mathrm{t},1}^* \diamond \boldsymbol{A}_{\mathrm{r}}^*\right)^{\mathrm{T}}\right]^{\dagger}$, $\hat{\boldsymbol{A}}_{\mathrm{r}}^* = [\boldsymbol{\mathcal{R}}]_{(2)} \left[\left(\boldsymbol{A}_{\mathrm{t},1}^* \diamond [\boldsymbol{\mathcal{A}}]_{(1)}^{\mathrm{T}}\right)^{\mathrm{T}}\right]^{\dagger}$ and $\hat{\boldsymbol{A}}_{\mathrm{t},1}^* = [\boldsymbol{\mathcal{R}}]_{(3)} \left[\left(\boldsymbol{A}_{\mathrm{r}}^* \diamond [\boldsymbol{\mathcal{A}}]_{(1)}^{\mathrm{T}}\right)^{\mathrm{T}}\right]^{\dagger}$, respectively.

Each iteration of the first TALS stage contains three LS estimation steps to update $[\hat{\boldsymbol{\mathcal{A}}}]_{(1)}^{\mathrm{T}}$, $\hat{\boldsymbol{A}}_{\mathrm{r}}^*$ and $\hat{\boldsymbol{A}}_{\mathrm{t},1}^*$ in an iterative way. At each step, a given factor matrix is updated by fixing the other two to their estimates obtained at previous updating steps. This procedure is repeated until the convergence of the first stage. Convergence of the first TALS stage at the i-th iteration is declared when $|e_{(i)} - e_{(i-1)}| \leq 10^{-6}$, where $e_{(i)}$ denotes the residual error at the i-th iteration defined as

$$e_{(i)} = \left\| \boldsymbol{\mathcal{R}} - \hat{\boldsymbol{\mathcal{R}}}_{(i)} \right\|_{\mathrm{F}}^2, \tag{109}$$

where $\hat{\boldsymbol{\mathcal{R}}}_{(i)}$ is the reconstructed version of the cross-covariance tensor obtained from the estimated factor matrices at the end of the i-th iteration.

4.3.3 Second ALS Stage (Bilinear ALS)

After the convergence of the first TALS stage, a second bilinear ALS (BALS) stage can be formulated to estimate the factor matrices $\boldsymbol{A}_{t,2}$ and $\boldsymbol{R}_{\gamma}$ from the outputs $\hat{\boldsymbol{A}}_r$ and $[\hat{\boldsymbol{\mathcal{A}}}]_{(1)}$ of the first stage. Thanks to the Nested-PARAFAC structure of $\boldsymbol{\mathcal{R}}$, the estimated factor matrix $[\hat{\boldsymbol{\mathcal{A}}}]_{(1)}$ can be recast as a third-order PARAFAC decomposition in unfolded form. Thus, in a similar way, and assuming now that $\boldsymbol{A}_r$ is known in the second stage, the matrices of interest can be obtained from the unfolding matrices

$$[\boldsymbol{\mathcal{A}}]_{(1)} = \boldsymbol{R}_{\gamma}^{\mathrm{T}} \left(\boldsymbol{A}_{t,2} \diamond \boldsymbol{A}_r\right)^{\mathrm{T}}, \tag{110}$$

$$[\boldsymbol{\mathcal{A}}]_{(3)} = \boldsymbol{A}_{t,2} \left(\boldsymbol{A}_r \diamond \boldsymbol{R}_{\gamma}^{\mathrm{T}}\right)^{\mathrm{T}}, \tag{111}$$

by solving two linear LS problems whose analytic solutions are $\hat{\boldsymbol{R}}_{\gamma}^{\mathrm{T}} = [\boldsymbol{\mathcal{A}}]_{(1)} \left[\left(\boldsymbol{A}_{t,2} \diamond \boldsymbol{A}_r\right)^{\mathrm{T}}\right]^{\dagger}$ and $\hat{\boldsymbol{A}}_{t,2} = [\boldsymbol{\mathcal{A}}]_{(3)} \left[\left(\boldsymbol{A}_r \diamond \boldsymbol{R}_{\gamma}^{\mathrm{T}}\right)^{\mathrm{T}}\right]^{\dagger}$, respectively. Note that, this second BALS stage has a faster convergence since it estimates only two factor matrices due to the knowledge of $\boldsymbol{A}_r$ previously obtained in the first TALS stage.

4.3.4 DoD and DoA Parameters Estimation

By exploiting the redundancy inserted by the maximally overlapping subarrays, a final estimate of the transmit steering matrix is computed from $\hat{\boldsymbol{A}}_{\mathrm{t},1}$ obtained in the first TALS stage and from $\hat{\boldsymbol{A}}_{\mathrm{t},2}$ obtained in the second BALS stage using the following relation:

$$\hat{\boldsymbol{A}}_{\mathrm{t}} = \begin{bmatrix} \hat{\boldsymbol{A}}_{\mathrm{t},1}\left(1,:\right) \\ \frac{1}{2}\left[\hat{\boldsymbol{A}}_{\mathrm{t},1}\left(2:M-1,:\right) + \hat{\boldsymbol{A}}_{\mathrm{t},2}\left(1:M-2,:\right)\right] \\ \hat{\boldsymbol{A}}_{\mathrm{t},2}\left(M-1,:\right) \end{bmatrix}. \tag{112}$$

The final step of the proposed method is to estimate the DoDs and DoAs of P targets. For the p-th target, which corresponds to the p-th column of the estimated transmit and receive steering matrices, estimates of its angular parameters $\hat{\phi}_p$ and $\hat{\theta}_p$ can be obtained from a one-dimensional peak search in the following spatial spectra (DA COSTA *et al.*, 2010):

$$\hat{\phi}_p = \underset{\phi_p}{\operatorname{argmax}} \left| \boldsymbol{a}_{\mathrm{t}}^{\mathrm{H}}(\phi_p) \cdot \hat{\boldsymbol{a}}_{\mathrm{t}}(\hat{\phi}_p) \right|, \tag{113}$$

$$\hat{\theta}_p = \underset{\theta_p}{\operatorname{argmax}} \left| \boldsymbol{a}_{\mathrm{r}}^{\mathrm{H}}(\theta_p) \cdot \hat{\boldsymbol{a}}_{\mathrm{r}}(\hat{\theta}_p) \right|, \tag{114}$$

where $\boldsymbol{a}_{\mathrm{t}}(\phi_p)$ and $\boldsymbol{a}_{\mathrm{r}}(\theta_p)$ are known array monifolds of the transmit and receive arrays for a given localization (ϕ_p, θ_p), respectively.

Algorithm 8 Proposed DALS algorithm for joint DoD and DoA estimation

procedure
First Trilinear ALS Stage
1.1 *Set* $i = 0$ *and initialize randomly* $\hat{\boldsymbol{A}}_{\mathrm{t},1(i=0)}$ *and* $\hat{\boldsymbol{A}}_{\mathrm{r}(i=0)}$;
1.2 $i \leftarrow i + 1$;
1.3 *From* $[\boldsymbol{\mathcal{R}}]_{(1)}$, *obtain an LS estimate of* $[\boldsymbol{\mathcal{A}}]_{(1)}$:

$$[\hat{\boldsymbol{\mathcal{A}}}]^{\mathrm{T}}_{(1)(i)} = [\boldsymbol{\mathcal{R}}]_{(1)} \left[\left(\hat{\boldsymbol{A}}_{\mathrm{t},1(i-1)} \diamond \hat{\boldsymbol{A}}_{\mathrm{r}(i-1)} \right)^{\mathrm{H}} \right]^{\dagger};$$

1.4 *From* $[\boldsymbol{\mathcal{R}}]_{(2)}$, *obtain an LS estimate of* $\boldsymbol{A}^{*}_{\mathrm{r}}$:

$$\hat{\boldsymbol{A}}^{*}_{\mathrm{r}(i)} = [\boldsymbol{\mathcal{R}}]_{(2)} \left[\left(\boldsymbol{A}^{*}_{\mathrm{t},1(i-1)} \diamond [\hat{\boldsymbol{\mathcal{A}}}]^{\mathrm{T}}_{(1)(i)} \right)^{\mathrm{T}} \right]^{\dagger};$$

1.5 *From* $[\boldsymbol{\mathcal{R}}]_{(3)}$, *obtain an LS estimate of* $\boldsymbol{A}^{*}_{\mathrm{t},1}$:

$$\hat{\boldsymbol{A}}^{*}_{\mathrm{t},1(i)} = [\boldsymbol{\mathcal{R}}]_{(3)} \left[\left(\boldsymbol{A}^{*}_{\mathrm{r}(i)} \diamond [\hat{\boldsymbol{\mathcal{A}}}]^{\mathrm{T}}_{(1)(i)} \right)^{\mathrm{T}} \right]^{\dagger};$$

1.6 *Repeat steps* 1.2-1.5 *until convergence of the first stage.*
Second Bilinear ALS Stage
2.1 *Obtain* $\hat{\boldsymbol{\mathcal{A}}} \in \mathbb{C}^{P \times N \times M-1}$ *by reshaping the estimated 1-mode unfolding matrix* $[\hat{\boldsymbol{\mathcal{A}}}]_{(1)(i)}$ *obtained in the first stage as a third-order tensor in accordance with* (105);
2.2 *Set* $j = 0$ *and initialize randomly* $\hat{\boldsymbol{R}}_{\gamma(j=0)}$ *and* $\hat{\boldsymbol{A}}_{\mathrm{t},2(j=0)}$;
2.3 $j \leftarrow j + 1$;
2.4 *From* $[\hat{\boldsymbol{\mathcal{A}}}]_{(1)}$ *and* $\hat{\boldsymbol{A}}_{\mathrm{r}(i)}$, *obtain an LS estimate of* $\boldsymbol{R}^{\mathrm{T}}_{\gamma}$:

$$\hat{\boldsymbol{R}}^{\mathrm{T}}_{\gamma(j)} = [\hat{\boldsymbol{\mathcal{A}}}]_{(1)} \left[\left(\hat{\boldsymbol{A}}_{\mathrm{t},2(j-1)} \diamond \hat{\boldsymbol{A}}_{\mathrm{r}(i)} \right)^{\mathrm{T}} \right]^{\dagger};$$

2.5 *From* $[\hat{\boldsymbol{\mathcal{A}}}]_{(3)}$ *and* $\hat{\boldsymbol{A}}_{\mathrm{r}(i)}$, *obtain an LS estimate of* $\boldsymbol{A}_{\mathrm{t},2}$:

$$\hat{\boldsymbol{A}}_{\mathrm{t},2(j)} = [\hat{\boldsymbol{\mathcal{A}}}]_{(3)} \left[\left(\hat{\boldsymbol{A}}_{\mathrm{r}(i)} \diamond \hat{\boldsymbol{R}}^{\mathrm{T}}_{\gamma(j)} \right)^{\mathrm{T}} \right]^{\dagger};$$

2.6 *Repeat steps* 2.3-2.5 *until convergence of the second stage.*
DoD and DoA Parameters Estimation
3.1 *From* $\hat{A}_{\mathrm{t},1(i)}$ *and* $\hat{A}_{\mathrm{t},2(j)}$, *obtain a final estimate of* $\hat{\boldsymbol{A}}_{\mathrm{t}}$ *using the relation* (112);
3.2 *From the p-th column of* $\hat{\boldsymbol{A}}_{\mathrm{t}}$ *and* $\hat{\boldsymbol{A}}_{\mathrm{r}}$, *find the estimates of* $\hat{\phi}_p$ *and* $\hat{\theta}_p$ $(p = 1, \ldots, P)$ *performing the one-dimensional peak search in the spatial spectrum* (113) *and* (114), *respectively.*

The proposed tensor-based double alternating least squares (DALS) algorithm for multi-target localization is summarized in the form of pseudo-code in Algorithm 8.

4.3.5 Alternative Closed-Form Solution to the Second Stage

When $N \geq P$ an estimate to $\boldsymbol{A}_{\mathrm{t},2}$ can be obtained from the first stage output $[\hat{\boldsymbol{\mathcal{A}}}]_{(1)}$ in closed-form by means of the LS-KRF algorithm 1 presented in Section 2.2.2. Multiplying both sides of $[\hat{\boldsymbol{\mathcal{A}}}]^{\mathrm{T}}_{(2)}$ by the pseudo-inverse of $\hat{\boldsymbol{A}}^{\mathrm{T}}_{\mathrm{r}}$ obtained in the first step, the p-th column of the resulting matrix $\hat{\boldsymbol{A}}_{\mathrm{t},2} \diamond \hat{\boldsymbol{R}}^{\mathrm{T}}_{\gamma}$ denotes the vectorization operation of the rank-1 matrix $\boldsymbol{\Psi}_p = \boldsymbol{r}^{*}_p \circ \boldsymbol{a}_{\mathrm{t},2}(\phi_p) \in \mathbb{C}^{P \times M-1}$, where $\boldsymbol{r}_p \in \mathbb{C}^{P \times 1}$ is the p-th column of the covariance matrix $\hat{\boldsymbol{R}}_{\gamma}$. Defining $\boldsymbol{U}_p \boldsymbol{\Sigma}_p \boldsymbol{V}^{\mathrm{H}}_p$ as the SVD of $\boldsymbol{\Psi}_p$, estimates of $\boldsymbol{r}^{*}_p$ and $\boldsymbol{a}_{\mathrm{t},2}(\phi_p)$ $(p = 1, \ldots, P)$ can be obtained by truncating the SVD to a rank-1 approximation as follows:

$$\hat{\boldsymbol{r}}^{*}_p = \sqrt{\sigma_1}\boldsymbol{u}_1 \quad \text{and} \quad \hat{\boldsymbol{a}}_{\mathrm{t},2}(\phi_p) = \sqrt{\sigma_1}\boldsymbol{v}^{*}_1, \tag{115}$$

Algorithm 9 Algorithm with alternative closed-form solution to the second stage

procedure

1. *From the covariance tensor* $\boldsymbol{\mathcal{R}}$ *in* (104)*, obtain the estimates of* $\hat{\boldsymbol{\mathcal{A}}}$*,* $\hat{\boldsymbol{A}}_{\mathrm{t},1}$ *and* $\hat{\boldsymbol{A}}_{\mathrm{r}}$ *using the first trilinear ALS stage in Algorithm* 8;
2. *From* $[\hat{\boldsymbol{\mathcal{A}}}]_{(2)}^{\mathrm{T}} \cdot \left(\hat{\boldsymbol{A}}_{\mathrm{r}}^{\mathrm{T}}\right)^{\dagger} = \hat{\boldsymbol{A}}_{\mathrm{t},2} \diamond \hat{\boldsymbol{R}}_{\gamma}^{\mathrm{T}}$*, obtain the estimates of* $\hat{\boldsymbol{A}}_{\mathrm{t},2}$ *and* $\hat{\boldsymbol{R}}_{\gamma}^{\mathrm{T}}$ *using the LS-KRF procedure in Algorithm* 1;
3. *From the p-th column of* $\hat{\boldsymbol{A}}_{\mathrm{t}}$ *and* $\hat{\boldsymbol{A}}_{\mathrm{r}}$*, find the estimates of* $\hat{\phi}_p$ *and* $\hat{\theta}_p$ $(p = 1, \ldots, P)$ *performing the one-dimensional peak search in the spatial spectrum* (113) *and* (114)*, respectively.*

where $\boldsymbol{u}_1 \in \mathbb{C}^{P\times 1}$ and $\boldsymbol{v}_1 \in \mathbb{C}^{M-1\times 1}$ are the first left and right singular vectors of $\boldsymbol{U}_p$ and $\boldsymbol{V}_p$, respectively, and σ_1 is the largest singular value. The estimates of the matrices $\hat{\boldsymbol{A}}_{\mathrm{t},2}$ and $\hat{\boldsymbol{R}}_{\gamma}$ are obtained by repeating this procedure P times according to Algorithm 1.

Remark: Note that the LS-KRF and peak search steps can be computed independently for each column of the estimated matrices. Thus, the estimates of DoDs and DoAs can be obtained in a parallel way if P processors are available, reducing the processing time by a factor of P. A version of the proposed tensor-based algorithm using the closed-form LS-KRF solution in the second stage is summarized in Algorithm 9.

4.4 Uniqueness Issues

We assume that each factor matrix in the Nested-PARAFAC model (104) has full-rank, i.e., all targets are uncorrelated with each other and have different DoDs and DoAs. Based on this assumption, and applying the Kruskal's condition (33) to the inner and outer PARAFAC parts of the Nested-PARAFAC decomposition, we can deduce that the estimated matrices $\hat{\boldsymbol{A}}_{\mathrm{t}}$ and $\hat{\boldsymbol{A}}_{\mathrm{r}}$ are unique up to permutation and scaling columns if

$$\min\left(N(M-1), P\right) + \min\left(N, P\right) + \min\left(M-1, P\right) \geq 2P + 2 \tag{116}$$

$$\min\left(N, P\right) + \min\left(M-1, P\right) \geq P + 2. \tag{117}$$

The combination of conditions (116) and (117) yields the following two corollaries:

- If $N \geq P$, then $M \geq 3$ transmit antennas are necessary to uniquely recover the spatial parameters of P targets;
- If $M - 1 \geq P$, then $N \geq 2$ receive antennas are necessary to uniquely recover the spatial parameters of P targets.

Since positioning errors can be measured as a function of the first antenna element (being the reference one), the first row of each matrix $\boldsymbol{A}_{\mathrm{t}}$ and $\boldsymbol{A}_{\mathrm{r}}$ is assumed to have unity entries. Hence, at the output of the first TALS stage, the scaling ambiguity in $\hat{\boldsymbol{A}}_{\mathrm{t},1}$ and $\hat{\boldsymbol{A}}_{\mathrm{r}}$ can be removed from this *a priori* information through normalization. By its turn, the scaling ambiguity affecting the estimate $\hat{\boldsymbol{A}}_{\mathrm{t},2}$ obtained in the second BALS stage

Table 1 – Complexity analysis of different tensor-based algorithms.

Algorithm	**Number of FLOPs**
ALS-PARAFAC (NION and SIDIROPOULOS, 2010)	$LMP^2 + LNP^2 + MNP^2$
ALS-Tucker4 (GOMES *et al.*, 2017)	$2N(M-1)^2P^2 + 2N^2(M-1)P^2 + N^2(M-1)^2P^4$
Nested-PARAFAC	
TALS (1st stage)	$N(M-1)^2P^2 + N^2(M-1)P^2 + N(M-1)P^2$
BALS (2nd stage)	$N(M-1)P^2 + NP^3$
LS-KRF (2nd stage)	$(M-1)P^2 + NP^2$

Source: Created by the author.

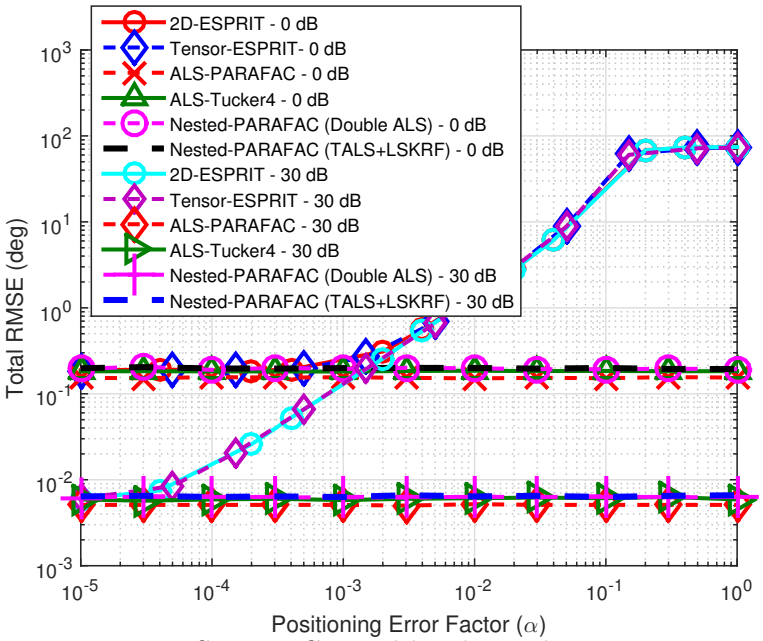
can be removed from the second row of $\hat{\boldsymbol{A}}_{\mathrm{t},1}$ already estimated in the first stage. Note that, the column permutation is the same for all matrices which means that the proposed method provides automatically paired estimates of the targets' DoDs and DoAs.

4.5 Computational Complexity

In the following, we evaluate the computational complexity of our proposed Nested-PARAFAC based estimator with respect to competing tensor-based algorithms recently proposed in the literature. The complexity of each algorithm in terms of FLOPs counts at each iteration is summarized in Table 1. In these expressions, the number of FLOPs refers to the dominant cost of the algorithms, which are the computations of the SVDs required to calculate pseudo-inverses at each iteration of the iterative tensor-based algorithms. Similar to (GOLUB and VAN LOAN, 1996), we assume here that the computational cost to calculate the SVD of a matrix of size $I_1 \times I_2$ is $\mathcal{O}\left(I_1 \cdot I_2 \cdot \min(I_1, I_2)\right)$.

As previously mentioned, the complexity of the ALS-PARAFAC algorithm is a function of the data block size L to be processed at the receiver. In Table 1, we can see that the first two terms of the ALS-Tucker4 algorithm has double the complexity than the first two terms in the TALS stage of the Nested-PARAFAC algorithm, while the last term has quadratic complexity compared to the last term in the TALS stage. Moreover, the knowledge of $\boldsymbol{A}_r$ in the BALS stage results in a faster convergence and lower complexity estimation, in comparison with the first TALS stage. Thus, the dominant part of the complexity of the Nested-PARAFAC refers to its first TALS stage. This becomes more evident when the alternative closed-form solution based on the LS-KRF algorithm is used in the second stage. Based on this analysis, we can conclude that the ALS-Tucker4 is computationally more expansive than the proposed Nested-PARAFAC algorithm. This discussion will be reinforced by the numerical results of Section 4.6, showing that the proposed approach yields similar performance as (NION and SIDIROPOULOS, 2010) and (GOMES *et al.*, 2017) while being more computationally attractive.

Figure 22 – Total RMSE (deg) vs. positioning error factor (α).



Source: Created by the author.

4.6 Simulation Results

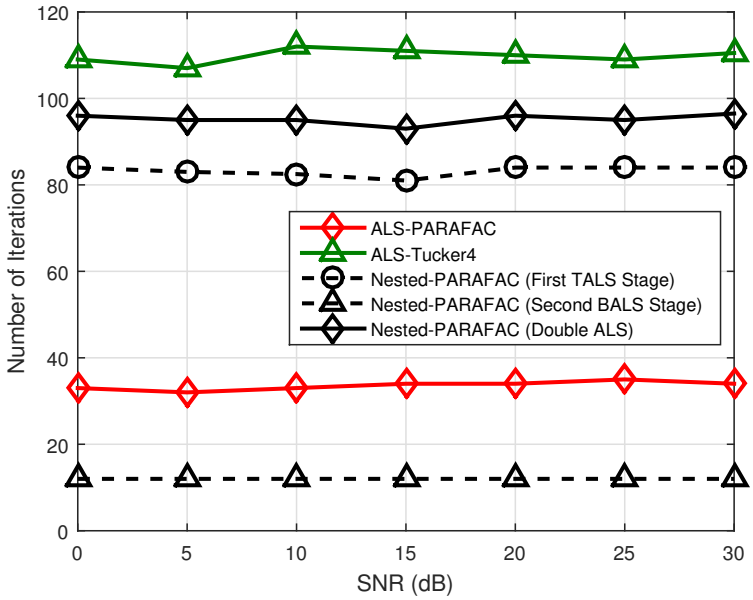
In this section, we evaluate the performance of the proposed Nested-PARAFAC estimator from computer simulations. We consider $M = 6$ and $N = 5$ antennas at the transmit and receive arrays, respectively. The antenna arrays are assumed to be non-uniform linear arrays (NULAs) with known positioning errors, i.e., the array manifolds are known. We assume $K = 128$ samples per pulse period, a pulse duration of 5 μs, and data block of size $L = 100$. Our simulated scenario has three targets with Doppler frequencies $\{f_{\mathrm{D}p}\}_{p=1}^{3} = \{300, 400, 400\}$ Hz and reflection coefficients $\{|\beta_p|\}_{p=1}^{3} = 1$, which are located at $(\phi_1, \theta_1) = (30°, 20°)$, $(\phi_2, \theta_2) = (50°, 40°)$ and $(\phi_3, \theta_3) = (70°, 60°)$. The numerical results represent an average of 1000 independent Monte Carlo runs.

In Figure 22, we show as a figure of merit the total root mean square error (RMSE) of the estimated angles as a function of the positioning error factor α. The total RMSE is defined as:

$$\mathrm{RMSE} = \sqrt{E\left\{\sum_{p=1}^{P}(\phi_p - \hat{\phi}_p)^2 + (\theta_p - \hat{\theta}_p)^2\right\}}, \tag{118}$$

where $\hat{\phi}_p$ and $\hat{\theta}_p$ are the estimated DoD and DoA of the p-th target, respectively. We compare the estimation accuracy of the proposed method with matrix-based 2D-ESPRIT (CHEN, GU, and SU, 2008) and Tensor-ESPRIT (HAARDT, ROEMER, and DEL GALDO, 2008). We also include in our evaluation, the tensor-based competitors ALS-Tucker4 (GOMES *et al.*, 2017) and ALS-PARAFAC (NION and SIDIROPOULOS, 2010), which

Figure 23 – Number of iterations for convergence vs. SNR (dB).

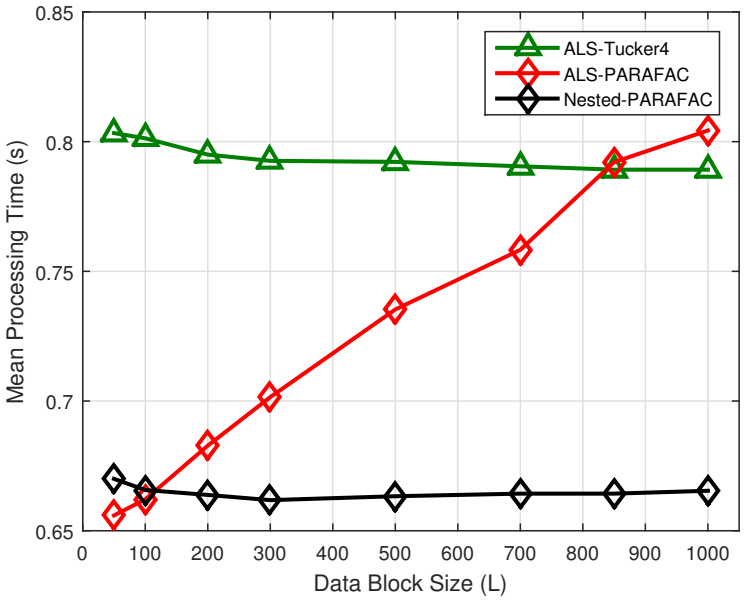


Source: Created by the author.

solve the same problem. We call attention that in the ALS-Tucker4 algorithm, the covariance matrix (102) is interpreted as a multimode unfolding of a fourth-order Tucker decomposition, and a five-step ALS-based algorithm is proposed to estimate the DoDs and DoAs of the targets from (103). The ALS-PARAFAC is a direct-data approach in which a trilinear ALS-based fitting algorithm is applied directly to received signal tensor (97), which is of size $N \times M \times L$. According to Figure 22, the iterative ALS-based algorithms are robust to real-world scenarios with array calibration errors and have no performance degradation when the error factor α increases. On the other hand, the accuracy of matrix-based and tensor-based ESPRIT techniques is degraded due to the violation of the shift-invariance property that only holds in the ideal scenario with $\alpha \to 0$. Note that all tensor-based iterative algorithms have similar estimation accuracy. Thus, the best choice within the simulated tensor-based methods can be guided mainly by the computational complexity. This aspect is evaluated in our next experiment.

In Figure 23, we evaluate the convergence rate of the iterative ALS-based algorithms. We plot as a figure of merit the average number of iterations required for convergence as a function of signal-to-noise ratio (SNR) in dB. In this experiment, the fast convergence of the ALS-PARAFAC algorithm does not expose its high computational cost per iteration when L is large, as shown in Table 1. Despite the same performance in terms of RMSE, the proposed Nested-PARAFAC estimator requires less iterations than the ALS-Tucker4 one. Indeed, as discussed in Section 4.5, most of the processing burden of the Nested-PARAFAC estimator is concentrated in its first TALS stage which corresponds to approximately 90% of the overall number of iterations. Figure 23 together with the values in Table 1 reveal that the proposed Nested-PARAFAC algorithm is preferable than

Figure 24 – Mean processing time (in seconds) vs. data block size.



Source: Created by the author.

the ALS-Tucker4 in terms of computational complexity.

Figure 24 depicts the mean processing time (in seconds) as a function of the data block size L processed at the receiver. In this experiment, we set the SNR to 30 dB. In contrast to Figure 23, we can observe more clearly the dependence between the computational complexity of the ALS-PARAFAC and the data block size. More specifically, for a reasonable value of L, the Nested-PARAFAC approach outperforms the ALS-PARAFAC one since the latter operates directly on the received raw data samples and thus has a higher computational complexity when L is large, as shown earlier in Table 1. These numerical results validate the merits of the proposed tensor-based method compared to state-of-the-art tensor-based ones.

4.7 Chapter Summary

In this chapter, we have proposed a new Nested-PARAFAC based approach for solving the multi-target localization problem in bistatic MIMO radar systems without imposing constraints on the geometry of the transmit and receive arrays. The proposed Nested-PARAFAC estimator offers the same performance in terms of DoD and DoA accuracy estimation as that of competing tensor-based methods in (NION and SIDIROPOULOS, 2010) and (GOMES *et al.*, 2017), while being less computationally complex. At the same time, the proposed solution is robust to positioning errors of the antenna elements, in contrast to competing ESPRIT and Tensor-ESPRIT techniques in (CHEN, GU, and SU, 2008) and (HAARDT, ROEMER, and DEL GALDO, 2008), respectively.

5 JOINT DOWNLINK AND UPLINK CHANNEL ESTIMATION USING TENSOR PROCESSING

In this chapter, we address the problem of joint downlink (DL) and uplink (UL) channel estimation in multiuser wireless communications systems. We first propose a closed-loop and multi-frequency based channel training framework in which pilot signals received by multiple mobile stations (MSs) are coded, spread in the frequency-domain via multiple adjacent subcarriers, and then reported back to the BS. Then, we show that the received closed-loop signal can be modeled as a third-order PARAFAC decomposition from which two semi-blind receivers to perform the joint estimation of the DL and UL channels are formulated. In addition, as another contribution, we extend the proposed framework for millimeter wave (mmWave) MIMO scenarios in which the BS is assumed to be equipped with a hybrid analog-digital beamforming (HB) architecture. In contrast to the classical approach, in which the DL and UL channel estimation problems are usually considered as two separate problems, our idea is to jointly estimate both the DL and UL channels as a single problem by concentrating most of the processing burden for channel estimation at the BS side. The proposed framework alleviates the computational overhead for channel estimation at the MSs side with limited processing capability, while relaxing channel reciprocity assumptions. Our identifiability study shows that the proposed receivers can achieve a reduced training overhead as well as require few frequency resources. Simulation results demonstrate that the proposed receivers achieve a performance close to the classical approach, with the advantage of avoiding complex computations for channel estimation at the MSs side as well as dedicated feedback channels for each MS.

5.1 Introduction and Motivation

The MIMO systems has been applied to many wireless standards since it can increase the capacity and reliability in wireless communications systems (LU *et al.*, 2014). In recent years, great attention has been given to multiuser MIMO systems (MU-MIMO) (GESBERT *et al.*, 2007), where a base station (BS) equipped whith multiple antennas simultaneously serves a set of (single or multiple antennas) mobile stations (MSs). The performance of MU-MIMO systems strongly depends on the efficient measurement of the channel state information (CSI) by the BS. This is possible from a training phase in which the BS transmits known training signals to the MSs. According to (CHOI, LOVE, and BIDIGARE, 2014), the BS can learn the CSI from limited feedback in frequency division duplexing (FDD) (LOVE *et al.*, 2008) or assuming channel reciprocity in time division duplexing (TDD) (MARZETA, 2010). However, in practice, the downlink (DL) channel estimated by the uplink (UL) channel considering channel reciprocity may not be accurate (GUEY and LARSSON, 2004). Futhermore, the channel acquisition becomes a challenge

if the BS has to transmit long DL training sequences, and the MS has to report back its large channel matrix estimates. The amount of overhead in the DL channel estimation due to the long training sequences and in the UL channel due to the large matrices to be reported can severely decrease the spectral efficiency of the system, calling for novel solutions that enable us to reduce the unnecessary overhead.

Alternatively to the classical MIMO systems, the millimeter wave (mmWave) massive MIMO technology has been subject of increasing interest in both academia and industry for future wireless standards, due to its great potential to provide substancial energy efficiency and data rate gains. The first one is due to the reduced number of radio frequency (RF) chains, assumed to be smaller that the number of antennas. The last one is due to the large chunck of available frequencies in the mmWave band. However, due to the severe path loss over the mmWave frequency bands, large antenna arrays should be deployed at the BS and MSs to provide sufficient beamforming gains using hybrid analog-digital beamforming (HB) architectures (ALKHATEEB *et al.*, 2014b; RIAL *et al.*, 2016). To fully benefit from the beamforming gains in mmWave MIMO systems, accurate CSI estimation is also crucial to realize the hybrid precoding designs in which the analog part can be used to improve the signal power, while the digital part is designed to supress inter-user interferences (LIANG, XU, and DONG, 2014; ALKHATEEB, HEATH, and LEUS, 2015; GALLACHER and RAHMAN, 2015).

Motivated by the strong importance of accurate CSI estimates in the MIMO systems performance, several channel estimation techniques have been proposed (ALKHATEEB *et al.*, 2014a; VENUGOPAL *et al.*, 2017; ZHOU *et al.*, 2016b, 2017). However, they assume conventional channel training framework, in which the DL and UL channel estimation problems are trated as two decoupled procedures at the MSs and BS, respectively. In contrast, in this chapter we study the problem of joint DL and UL channel estimation. Initially, as a first contribution, we propose a novel closed-loop and multi-frequency based channel training framework in which the pilot signals received by multiple MSs are coded and spread in the frequency-domain, and then feed back to the BS over the same UL resources. Making use of this framework, the received closed-loop signal can be modeled as a third-order PARAFAC decomposition. As a second contribution, by capitalizing the multidimensional structure of the received signal, we propose two tensor-based semi-blind receivers for joint DL and UL channel estimation. The first one is an iterative solution based on the ALS concept. The second one is a closed-form solution based on the LS-KRF algorithm. As a third contribution, we extend the proposed techniques to opperate in mmWave MIMO scenarios. The proposed framework allows concentrating most of the processing burden for channel estimation at the BS side, i.e., avoiding computational overhead for channel estimation at the MSs with limited processing capability. Simulation results reveal that the proposed receivers achieve a performance close to the classical framework that treat the estimation of DL and UL channels as separate problems.

5.1.1 Chapter Organization

This chapter is organized as follows. In Section 5.2, we present the proposed closed-loop channel training framework and the proposed system model. Section 5.3 formulates the proposed tensor-based semi-blind receivers for joint DL and UL channel estimation. In Section 5.4, we analyze the uniqueness issues and computational complexity of the proposed receivers. The first set of simulation results are provided in Section 5.5. The extension of the proposed techniques for mmWave massive MIMO scenarios is provided in Section 5.6. The set of simulation results related to the proposed extension are provided in Section 5.7. Finally, the chapter conclusions are drawn in Section 5.8.

5.2 System Model

In this section, we introduce the proposed closed-loop and multi-frequency channel training framework. Based on it, we formulate the corresponding DL and UL signal models. Our focus here is to show that the proposed framework allows joint DL and UL channel estimation by concentrating the processing burden associated with channel estimation at the BS, while relaxing channel reciprocity assumption.

5.2.1 Downlink Signal Model

Consider a multiuser wireless communication system consisting of a BS and U MSs. The BS is equiped with a ULA of N elements while each MS is a single-antenna device. Let $\boldsymbol{S} \in \mathbb{C}^{T\times N}$ be a known pilot signal sent by the BS to all MSs in which T denotes the length of the training sequence. The signal $\boldsymbol{y}_u \in \mathbb{C}^T$ received by the u-th MS is given by

$$\boldsymbol{y}_u = \boldsymbol{S}\boldsymbol{h}_u + \boldsymbol{v}_u^{(\mathrm{DL})}, \tag{119}$$

where $\boldsymbol{h}_u \in \mathbb{C}^N$ denotes the DL channel vector and $\boldsymbol{v}_u^{(\mathrm{DL})} \in \mathbb{C}^T$ is the AWGN vector, both related to the u-th MS.

5.2.2 Uplink Signal Model

The pilot signal (119) received by the u-th MS is feed back to the BS after a multi-frequency coding operation (i.e., no channel estimation is done at the MS side). More specifically, we assume a preprocessing step in which $\boldsymbol{y}_u$ $(u = 1, \ldots, U)$ is coded and spread in the frequency-domain across F adjacent subcarriers over which the UL channel is assumed to be constant. The u-th MS loads the received pilot into the f-th subcarrier using the code factor $c_{f,u}$. The BS then receives the sum of U co-channel signals. The

received closed-loop signal at the BS associated with the f-th subcarrier is given by

$$\begin{aligned} \boldsymbol{X}_f &= \left(\sum_{u=1}^{U} c_{f,u}\boldsymbol{g}_u\boldsymbol{y}_u^{\mathrm{T}}\right) + \boldsymbol{V}_f^{(\mathrm{UL})} \\ &= \boldsymbol{G}D_f\left(\boldsymbol{C}\right)\boldsymbol{Y}^{\mathrm{T}} + \boldsymbol{V}_f^{(\mathrm{UL})} \in \mathbb{C}^{N\times T}, \quad f = 1,\ldots,F, \end{aligned} \tag{120}$$

where $c_{f,u}$ denotes the (f,u)-th entry of the so called frequency spreading matrix $\boldsymbol{C} = [\boldsymbol{c}_1,\ldots,\boldsymbol{c}_U] \in \mathbb{C}^{F\times U}$. The u-th column of $\boldsymbol{C}$ contains the set of code coefficients used by the u-th MS over F frequencies. The matrix $\boldsymbol{Y} = \boldsymbol{S}\boldsymbol{H} + \boldsymbol{V}^{(\mathrm{DL})} \in \mathbb{C}^{T\times U}$ collects in its u-th column the pilot signal (119) received by the u-th MS. The matrices $\boldsymbol{H} = [\boldsymbol{h}_1,\ldots,\boldsymbol{h}_U] \in \mathbb{C}^{N\times U}$ and $\boldsymbol{G} = [\boldsymbol{g}_1,\ldots,\boldsymbol{g}_U] \in \mathbb{C}^{N\times U}$ denote the DL and UL channel matrices, respectively. $\boldsymbol{V}^{(\mathrm{DL})} \in \mathbb{C}^{T\times U}$ and $\boldsymbol{V}_f^{(\mathrm{UL})} \in \mathbb{C}^{N\times T}$ denote the overall DL and UL noise contributions. Throughout this chapter, we assume that $\boldsymbol{H}$ and $\boldsymbol{G}$ are possibly different, which means channel reciprocity may not hold.

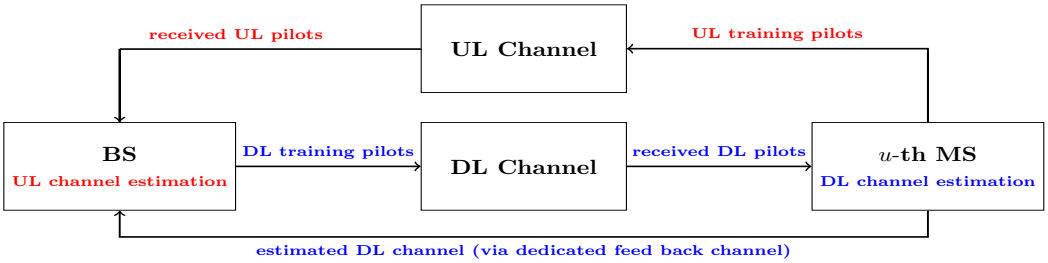
Remark: It is worth noting that the coding vectors used by the different MSs do not need to be orthogonal. In practice, this means that these codes can be locally generated at each MS as pseudo-random sequences, i.e., no prior signaling and coordination between MSs is necessary.

5.2.3 Conventional x Proposed Channel Training Framework

The conventional channel training framework, illustrated in Figure 25, assumes channel reciprocity in TDD or treat the DL and UL channel estimation as two separated problems in FDD, i.e., the DL and UL channel estimation procedures are performed independently at the MS and BS, respectively. In the DL channel estimation the BS (first) sends pilot signals, and the MS use them to obtain the DL channel estimation by means of state-of-the-art LS, minimum mean square error (MMSE) or compressed sensing (CS) based estimators. Then, the estimated DL channel is feedback to the BS via dedicated UL resources (SHEN *et al.*, 2016). There are two ways of feeding back the DL channel information to the BS. Explicit feedback consists in reporting the estimated channel vector $\hat{\boldsymbol{h}}_u$ computed by the u-th MS. However, its report may produce important overhead when the number of antennas is large. On the other hand, implicit feedback means transmitting quantized versions of the channel, which reduces the overhead to the cost of channel accuracy (ARAÚJO, 2016). To solve the UL channel estimation problem, the pilot signal is sent by the MS to the BS. Finally, the UL channel estimation procedure is performed by the BS.

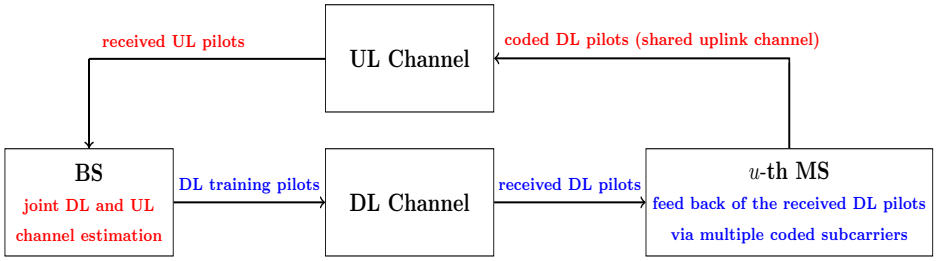
Reciprocity-based channel estimation assume that DL and UL channels are reciprocal, i.e., one is a transposed version of another one. The channel estimation procedure initiates at the MS by sending pilot signals toward BS. It collects the pilot signals

Figure 25 – **Conventional training framework.** The DL and UL channel estimation problems are solved independently. The BS first transmits pilot signals. Then, the DL channel is estimated at the MSs side. The estimated DL channel is feed back to the BS via dedicated uplink resources. The UL channel is estimated at the BS side. The blue words refer to DL communication, while red words refer to UL communication.



Source: Created by the author.

Figure 26 – **Proposed closed-loop and multi-frequency based training framework.** The DL and UL channels are jointly estimated. The BS first transmits pilot signals. The MSs encode the received pilots and then feed them back to the BS. The BS jointly estimates the DL and UL channels from (120). The blue words refer to DL communication, while red words refer to UL communication.



Source: Created by the author.

and estimates the UL channel by using state-of-the-art techniques, after that, MS uses the reciprocity assumption to obtain an estimate of the DL channel (ARAÚJO, 2016). In practice, the reciprocity assumption may not be hold due to radio frequency distortions, mutual coupling effects between antenna elements and different carrier frequency such as in FDD (SHEPARD *et al.*). Furthermore, the conventional framework may imply a high computational complexity at the MS side, especially when devices with limited processing capability are utilized by the users.

By making use of the proposed closed-loop and multi-frequency channel training framework, illustrated in Figure 26, no processing for channel estimation is performed at the MSs side. In contrast to the conventional approach in Figure 25, the pilot signal received by each MS is feedback to the BS after a multi-frequency coding operation across adjacent subcarriers. After this closed-loop transmission scheme, the joint DL and UL channel estimation can be performed at the BS from the received signal (120). The proposed framework alleviates computational overhead due to processing for channel estimation at the MS side, by shiffiting channel estimation processing to the BS side. Furthermore, it also relaxes channel reciprocity assumptions, typical in TDD systems, since the DL and UL channels can be estimated as independent variables from (120).

5.3 Proposed Tensor-Based Semi-Blind Receivers for Joint DL and UL Channel Estimation

Our aim in this chapter is to jointly estimate the DL and UL channel matrices $\boldsymbol{H}$ and $\boldsymbol{G}$ by solving a multiuser channel estimation problem at the BS. For this purpose, we first recall the received closed-loop signal (120) using tensor formalism. Then, by capitalizing its multidimensional structure, we formulate two tensor-based semi-blind receivers to estimate the channel matrices by means of the well-known ALS and LS-KRF algorithms.

5.3.1 PARAFAC-Based Modeling

Similar to (28), the noiseless term in the closed-loop received signal (120) represents the f-th frontal slice of the third-order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{N \times T \times F}$ obtained by concatenating the F signal matrices $\{\boldsymbol{X}_f\}_{f=1}^{F}$ associated to the different adjacent subcarriers along the third mode of $\boldsymbol{\mathcal{X}}$, i.e.,

$$\boldsymbol{\mathcal{X}} = \boldsymbol{X}_1 \sqcup_3 \boldsymbol{X}_2 \sqcup_3 \ldots \sqcup_3 \boldsymbol{X}_F. \tag{121}$$

The tensor $\boldsymbol{\mathcal{X}}$ assumes the following PARAFAC decomposition

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{I}}_{3,U} \times_1 \boldsymbol{G} \times_2 \boldsymbol{Y} \times_3 \boldsymbol{C}. \tag{122}$$

The following correspondences can be stated by comparing (122) with (25):

$$(I_1, I_2, I_3, Q) \leftrightarrow (N, T, F, U)$$
$$\left(\boldsymbol{A}^{(1)}, \boldsymbol{A}^{(2)}, \boldsymbol{A}^{(3)}\right) \leftrightarrow (\boldsymbol{G}, \boldsymbol{Y}, \boldsymbol{C})\,.$$

According to (29), (30) and (31), the tensor $\boldsymbol{\mathcal{X}}$ admits the following factorizations in terms of its unfolding and factor matrices:

$$[\boldsymbol{\mathcal{X}}]_{(1)} = \boldsymbol{G}\,(\boldsymbol{C} \diamond \boldsymbol{Y})^{\mathrm{T}} \in \mathbb{C}^{N \times TF}, \tag{123}$$
$$[\boldsymbol{\mathcal{X}}]_{(2)} = \boldsymbol{Y}\,(\boldsymbol{C} \diamond \boldsymbol{G})^{\mathrm{T}} \in \mathbb{C}^{T \times NF}, \tag{124}$$
$$[\boldsymbol{\mathcal{X}}]_{(3)} = \boldsymbol{C}\,(\boldsymbol{Y} \diamond \boldsymbol{G})^{\mathrm{T}} \in \mathbb{C}^{F \times NT}. \tag{125}$$

In order to jointly estimate the channel matrices $\boldsymbol{H}$ and $\boldsymbol{G}$ from the received signal tensor $\boldsymbol{\mathcal{X}}$ in (122), we formulate in the following the two proposed tensor-based semi-blind receivers which are based on the ALS and LS-KRF algoritms, respectively.

5.3.2 ALS-PARAFAC Receiver

From (122), estimates of $\boldsymbol{G}$, $\boldsymbol{Y}$ and $\boldsymbol{C}$ can be obtained by solving the following quadratic optimization problem:

$$\min_{\hat{\boldsymbol{G}}, \hat{\boldsymbol{Y}}, \hat{\boldsymbol{C}}} \left\| \boldsymbol{\mathcal{X}} - \sum_{u=1}^{U} \hat{\boldsymbol{g}}_u \circ \hat{\boldsymbol{y}}_u \circ \hat{\boldsymbol{c}}_u \right\|_{\mathrm{F}}^2 . \tag{126}$$

This problem can be solved by means of the ALS-PARAFAC algorithm 2. It consists of estimating in an alternating way the factor matrices from the unfolding matrices $[\boldsymbol{\mathcal{X}}]_{(n)}$ $(n = 1, \ldots, 3)$ by solving the following three linear LS problems:

$$\underset{\boldsymbol{G}}{\operatorname{argmin}} \left\| [\boldsymbol{\mathcal{X}}]_{(1)} - \boldsymbol{G}\,(\boldsymbol{C} \diamond \boldsymbol{Y})^{\mathrm{T}} \right\|_{\mathrm{F}}^2, \tag{127}$$
$$\underset{\boldsymbol{Y}}{\operatorname{argmin}} \left\| [\boldsymbol{\mathcal{X}}]_{(2)} - \boldsymbol{Y}\,(\boldsymbol{C} \diamond \boldsymbol{G})^{\mathrm{T}} \right\|_{\mathrm{F}}^2, \tag{128}$$
$$\underset{\boldsymbol{C}}{\operatorname{argmin}} \left\| [\boldsymbol{\mathcal{X}}]_{(3)} - \boldsymbol{C}\,(\boldsymbol{Y} \diamond \boldsymbol{G})^{\mathrm{T}} \right\|_{\mathrm{F}}^2 . \tag{129}$$

The solutions of which are given by $\hat{\boldsymbol{G}} = [\boldsymbol{\mathcal{X}}]_{(1)} \left[(\boldsymbol{C} \diamond \boldsymbol{Y})^{\mathrm{T}}\right]^{\dagger}$, $\hat{\boldsymbol{Y}} = [\boldsymbol{\mathcal{X}}]_{(2)} \left[(\boldsymbol{C} \diamond \boldsymbol{G})^{\mathrm{T}}\right]^{\dagger}$ and $\hat{\boldsymbol{C}} = [\boldsymbol{\mathcal{X}}]_{(3)} \left[(\boldsymbol{Y} \diamond \boldsymbol{G})^{\mathrm{T}}\right]^{\dagger}$, respectively.

Each iteration of the ALS-PARAFAC algorithm has three LS updating steps. At each step, one factor matrix is updated while the remaining factor matrices are assumed fixed to their values obtained in the previous steps. This procedure is repeated until the

convergence of the algorithm. Denoting by

$$e_{(i)} = \left\| [\boldsymbol{\mathcal{X}}]_{(1)} - [\hat{\boldsymbol{\mathcal{X}}}]_{(1)} \right\|_{\mathrm{F}}^{2} \tag{130}$$

the residual error between the received signal and the reconstructed signal at the i-th iteration, defined as $[\hat{\boldsymbol{\mathcal{X}}}]_{(1)} = \hat{\boldsymbol{G}} \left(\hat{\boldsymbol{C}} \diamond \hat{\boldsymbol{Y}} \right)^{\mathrm{T}}$, the convergence at the i-th iteration is declared when $|e_{(i)} - e_{(i-1)}| \leqslant 10^{-6}$.

By assuming a coordinated scenario in which the frequency spreading matrix $\boldsymbol{C}$ is known at the BS, only column scaling ambiguity in the estimated matrices $\hat{\boldsymbol{G}}$ and $\hat{\boldsymbol{Y}}$ exists. The scaling ambiguity matrix $\boldsymbol{\Delta_G}$ can be determined by assuming knowledge of the first row of $\boldsymbol{G}$. In turn, the ambiguity matrix $\boldsymbol{\Delta_Y}$ can be obtained by means of the following relation:

$$\boldsymbol{\Delta_G} \boldsymbol{\Delta_Y} = \boldsymbol{I}_U. \tag{131}$$

In practice, the knowledge of the first row of $\boldsymbol{G}$ can be obtained using a simple supervised procedure in which a known pilot symbol is sent to the BS by each MS (FERNANDES, FAVIER, and MOTA, 2011; XIMENES, FAVIER, and DE ALMEIDA, 2015). We assume that before transmission, each MS send a known pilot sequence to estimate its link between the first receive antenna at the BS. This "pre-phase" is essential for the receiver to remove the scaling ambiguity in the estimated matrices.

Downlink Channel Estimation: After the convergence of the ALS-PARAFAC algorithm, an additional step becomes necessary to estimate the DL channel matrix $\hat{\boldsymbol{H}}$ from the output matrix $\hat{\boldsymbol{Y}} = \boldsymbol{S}\hat{\boldsymbol{H}} + \hat{\boldsymbol{V}}^{(\mathrm{DL})}$. To this end, we incorporate the classical training-based MMSE channel estimator as final step of the proposed receiver. The MMSE estimate of $\hat{\boldsymbol{H}}$ is given by (KAY, 2000; CHO *et al.*, 2010)

$$\hat{\boldsymbol{H}} = \left[\boldsymbol{S}^{\mathrm{H}} \boldsymbol{R}_{vv}^{(\mathrm{DL})^{-1}} \boldsymbol{S} + \boldsymbol{R}_{hh}^{(\mathrm{DL})^{-1}} \right]^{-1} \boldsymbol{S}^{\mathrm{H}} \boldsymbol{R}_{vv}^{(\mathrm{DL})^{-1}} \hat{\boldsymbol{Y}}, \tag{132}$$

where $\boldsymbol{R}_{vv}^{(\mathrm{DL})} \in \mathbb{C}^{T \times T}$ and $\boldsymbol{R}_{hh}^{(\mathrm{DL})} \in \mathbb{C}^{N \times N}$ denote the noise covariance matrix and the covariance matrix of the DL channel, respectively. The proposed ALS-PARAFAC receiver for joint estimation of the DL and UL channels is summarized in Algorithm 10.

5.3.3 LS-KRF Receiver

Alternatively to the iterative ALS-PARAFAC receiver, estimates of the DL and UL channel matrices can be obtained in a closed-form way by exploiting the knowledge of the frequency spreading matrix $\boldsymbol{C}$, and then solving a set of rank-1 approximation problems by means of the LS-KRF algorithm formulated in Section 2.2.2.

Algorithm 10 Proposed ALS-PARAFC receiver for joint DL and UL channel estimation

1: **procedure**
2: 1. *Set* $i = 0$. *Randomly initialize* $\hat{\boldsymbol{Y}}_{(i=0)}$;
3: 2. $i \leftarrow i + 1$;
4: 3. *Using* $[\boldsymbol{\mathcal{X}}]_{(1)}$ in (123), *find an LS estimate of* $\hat{\boldsymbol{G}}_{(i)}$:
5:
$$\hat{\boldsymbol{G}}_{(i)} = [\boldsymbol{\mathcal{X}}]_{(1)} \left[\left(\boldsymbol{C} \diamond \hat{\boldsymbol{Y}}_{(i-1)} \right)^{\mathrm{T}} \right]^{\dagger};$$
6: 4. *Using* $[\boldsymbol{\mathcal{X}}]_{(2)}$ in (124), *find an LS estimate of* $\hat{\boldsymbol{Y}}_{(i)}$:
7:
$$\hat{\boldsymbol{Y}}_{(i)} = [\boldsymbol{\mathcal{X}}]_{(2)} \left[\left(\boldsymbol{C} \diamond \hat{\boldsymbol{G}}_{(i)} \right)^{\mathrm{T}} \right]^{\dagger};$$
8: 5. *Repeat steps* 2-4 *until convergence;*
9: 6. *Obtain the MMSE estimate for* $\hat{\boldsymbol{H}}$ using (132).

By assuming $F \geq U$, and multiplying both sides of $[\boldsymbol{\mathcal{X}}]_{(3)}^{\mathrm{T}}$ in (125) by the pseudo-inverse of $\boldsymbol{C}^{\mathrm{T}}$ from the right hand side, we get

$$\begin{aligned} \boldsymbol{Y} \diamond \boldsymbol{G} &= [\boldsymbol{\mathcal{X}}]_{(3)}^{\mathrm{T}} \left(\boldsymbol{C}^{\mathrm{T}} \right)^{\dagger} && (133) \\ &= [\boldsymbol{y}_1 \otimes \boldsymbol{g}_1, \ldots, \boldsymbol{y}_U \otimes \boldsymbol{g}_U] \in \mathbb{C}^{NT \times U}. && (134) \end{aligned}$$

According to property (9), the u-th column of (134) can be rewritten as

$$\boldsymbol{y}_u \otimes \boldsymbol{g}_u = \mathrm{vec}\left(\boldsymbol{g}_u \circ \boldsymbol{y}_u \right), \tag{135}$$

that represents the vectorization operation of the rank-1 matrix $\boldsymbol{W}_u = \boldsymbol{g}_u \circ \boldsymbol{y}_u \in \mathbb{C}^{N \times T}$. Being $\boldsymbol{U}_u \boldsymbol{\Sigma}_u \boldsymbol{V}_u^{\mathrm{H}}$ the SVD of $\boldsymbol{W}_u$, estimates for $\boldsymbol{g}_u$ and $\boldsymbol{y}_u$ $(u = 1, \ldots, U)$ can be obtained by truncating the SVD for a rank-1 approximation, i.e.,

$$\hat{\boldsymbol{g}}_u = \sqrt{\sigma_1} \boldsymbol{u}_1 \quad \text{and} \quad \hat{\boldsymbol{y}}_u = \sqrt{\sigma_1} \boldsymbol{v}_1^*, \tag{136}$$

where $\boldsymbol{u}_1 \in \mathbb{C}^{N \times 1}$ and $\boldsymbol{v}_1 \in \mathbb{C}^{T \times 1}$ are the first left and right singular vectors of $\boldsymbol{U}_u$ and $\boldsymbol{V}_u$, respectively, and σ_1 is the largest singular value. Final estimates for the matrices $\boldsymbol{Y}$ and $\boldsymbol{G}$ are obtained by repeating this process for $u = 1, \ldots, U$ as in Algorithm 1. Analogous to the proposed ALS-PARAFAC receiver, the column scaling ambiguity in $\hat{\boldsymbol{G}}$ and $\hat{\boldsymbol{Y}}$ can be removed when the first row of $\boldsymbol{G}$ is known. In the same way, the MMSE estimate for $\hat{\boldsymbol{H}}$ can be obtained from (132). The proposed LS-KRF receiver is shown in Algorithm 11.

Note that in the proposed LS-KRF receiver, a single SVD is computed independently for each column of the Khatri-Rao product $\boldsymbol{Y} \diamond \boldsymbol{G}$. Therefore, the U columns of the estimated matrices $\hat{\boldsymbol{Y}}$ and $\hat{\boldsymbol{G}}$ can be computed in a parallel way when U proces-

Algorithm 11 Proposed LS-KRF receiver for joint DL and UL channel estimation

1: **procedure**
2: 1. *From* $\boldsymbol{Y} \diamond \boldsymbol{G} = [\boldsymbol{\mathcal{X}}]_{(3)}^{\mathrm{T}} \left(\boldsymbol{C}^{\mathrm{T}}\right)^{\dagger}$, *obtain the estimates of* $\hat{\boldsymbol{G}}$ *and* $\hat{\boldsymbol{Y}}$ *using the*
3: *LS-KRF algorithm* 1;
4: 2. *Obtain the MMSE estimate for* $\hat{\boldsymbol{H}}$ using (132).

sors are available at the BS. In other words, the processing delay associated with channel estimation and symbol detection at the BS can be kept constant (i.e. it does not increase with the number of MSs), as long as the BS is equipped with multiple (at least U) DSP units. Note also that the complexity per unit is that of a rank-1 approximation problem.

5.4 Identifiability and Computational Complexity

In this section, we study the identifiability issues under which the downlink $\boldsymbol{H}$ and uplink $\boldsymbol{G}$ channel matrices can be jointly and uniquely recovered using the proposed ALS-PARAFAC and LS-KRF receivers. In addition, we also present a brief discussion about the computational complexity of the proposed methods.

5.4.1 Identifiability Conditions

1. *ALS-PARAFAC receiver:* Based on the received signal tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{N \times T \times F}$ in (120), the Kruskal's condition (33) can be rewritten to our signal model as:

$$\kappa_{\boldsymbol{G}} + \kappa_{\boldsymbol{Y}} + \kappa_{\boldsymbol{C}} \geq 2U + 2. \tag{137}$$

We assume the following: (i) The DL and UL channels of the u-th MS are modeled as a combination of L_u paths, each one of them being characterized by differents AoDs, AoAs and fading coefficients (DE ALMEIDA, FAVIER, and MOTA, 2006); (ii) The antenna array response has a Vandermonde structure (we assume a ULA at the BS) and (iii) The factor matrix $\boldsymbol{S}$ is randomly generated and follows a uniform distribution while $\boldsymbol{C}$ is a DFT matrix. Under the assumptions (i) and (ii) the DL and UL channel matrices $\boldsymbol{H}$ and $\boldsymbol{G}$ have full rank with probability one. From **Definition 12**, this implies $\kappa_{\boldsymbol{G}} = \min(N, U)$. Assumption (iii) implies that the matrices $\boldsymbol{S}$ and $\boldsymbol{C}$ have full rank, and consequently $\kappa_{\boldsymbol{Y}} = \min(T, U)$ and $\kappa_{\boldsymbol{C}} = \min(F, U)$. From this analysis, the Kruskal's condition (137) can be equivalently written as follows:

$$\min(N, U) + \min(T, U) + \min(F, U) \geq 2U + 2. \tag{138}$$

Since the assumption that the number of antennas at the BS is greater than the number of MSs is reasonable, the identifiability condition (138) simplifies to

$$\min(T, U) + \min(F, U) \geq U + 2. \tag{139}$$

From condition (139), we can analyze the following scenarios:

1) Considering $T \geq U$, only $F = 2$ subcarriers are required to estimate the DL and UL channels of U MSs. Otherwise stated, the ALS-PARAFAC receiver requires a reduced number of frequency resources (subcarriers).

2) Considering $F \geq U$, a training sequence of length $T = 2$ pilots is enough to estimate the DL and UL channels of U MSs. Due to the small size of the training sequence, the ALS-PARAFAC receiver has a reduced training overhead.

2. *LS-KRF receiver:* The LS-KRF receiver requires that the following necessary and sufficient uniqueness condition be satisfied:

$$F \geq U. \tag{140}$$

Note that this condition represents a particular case of (139), which indicates that the application of the LS-KRF receiver requires a more restricted scenario compared to the ALS-PARAFAC receiver, since the number of used frequency resources (subcarriers) increases with the number of active MSs. On the other hand, the LS-KRF receiver is a closed-form solution that allows parallel (user-wise) channel estimation and symbol detection, in contrast to the ALS-PARAFAC one, where all users are processed jointly.

5.4.2 Computational complexity

1. *ALS-PARAFAC receiver:* We approximate the computational complexity of the proposed ALS-PARAFAC receiver, in terms of FLOPs, considering only the cost associated with the SVD used to compute the matrix pseudo-inverses in the LS solutions of (127) and (128). According to (GOLUB and VAN LOAN, 1996), each iteration of the ALS-PARAFAC receiver requires approximately $\mathcal{O}(U^2FT + U^2FN)$ FLOPs.

2. *LS-KRF receiver:* The computational complexity of the LS-KRF receiver can be approximated as the cost to obtain $\boldsymbol{Y} \diamond \boldsymbol{G}$ from (125) added with the cost to calculate U SVD-based rank-1 approximations to estimate $\hat{\boldsymbol{G}}$ and $\hat{\boldsymbol{Y}}$ from the previous Khatri-Rao product. Thus, the LS-KRF requires approximately $\mathcal{O}(U^2F + UNT)$ FLOPs.

Therefore, the overall complexities of the proposed semi-blind receivers are equal to the sum of the the above complexities and the computational cost $\mathcal{O}(T^3 + N^3)$ associated with the MMSE estimate of $\hat{\boldsymbol{H}}$ from $\hat{\boldsymbol{Y}}$ according to (132).

5.5 Simulation Results (Part 1)

The first set of simulation results presented here is devoted to evaluate the performance of the proposed ALS-PARAFAC and LS-KRF semi-blind receivers in some selected system configurations. We evaluate the performance in DL and UL channel estimations in terms of normalized mean square error (NMSE) of the estimated matrices compared to the actual channel matrices. We consider a ULA composed by $N = 32$ antenna elements at the BS. The DL and UL channels of the u-th MS are generated assuming a multipath channel model (DE ALMEIDA, FAVIER, and MOTA, 2006) composed by L_u paths being each one characterized by a particular AoD, AoA and path gain. At each run, L_u is set randomly between one and five for each MS. The AoDs and AoAs to the l-th path of the u-th MS are randomly distributed in the interval $[0, 2\pi]$ and the path gains follow a complex-valued Gaussian distribution with zero-mean and unit variance. We also assume that the known pilot signal $\boldsymbol{S}$ is a BPSK modulated sequence while the frequency spreading matrix $\boldsymbol{C}$ is designed to be DFT. The results are averaged over 2000 independent Monte Carlo runs.

The NMSE of the estimated DL and UL channel matrices are defined as:

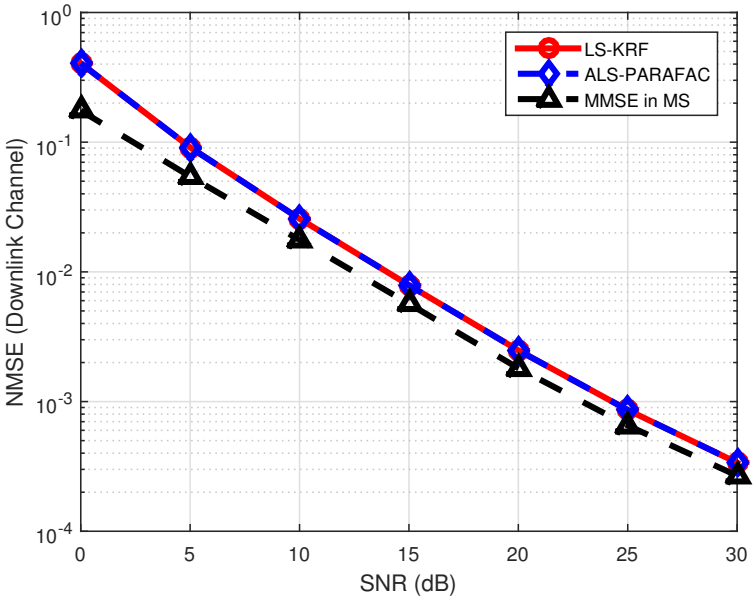
$$\text{NMSE}\left(\hat{\boldsymbol{H}}\right) = \frac{1}{2000}\sum_{\tau=1}^{2000}\frac{\|\boldsymbol{H}_{(\tau)} - \hat{\boldsymbol{H}}_{(\tau)}\|_{\text{F}}^2}{\|\boldsymbol{H}_{(\tau)}\|_{\text{F}}^2}, \tag{141}$$

$$\text{NMSE}\left(\hat{\boldsymbol{G}}\right) = \frac{1}{2000}\sum_{\tau=1}^{2000}\frac{\|\boldsymbol{G}_{(\tau)} - \hat{\boldsymbol{G}}_{(\tau)}\|_{\text{F}}^2}{\|\boldsymbol{G}_{(\tau)}\|_{\text{F}}^2}, \tag{142}$$

where $\boldsymbol{H}_{(\tau)}$ and $\boldsymbol{G}_{(\tau)}$ denote the true DL and UL channel matrices while $\hat{\boldsymbol{H}}_{(\tau)}$ and $\hat{\boldsymbol{G}}_{(\tau)}$ denote the DL and UL channel estimates, both with respect to the τ-th Monte Carlo run.

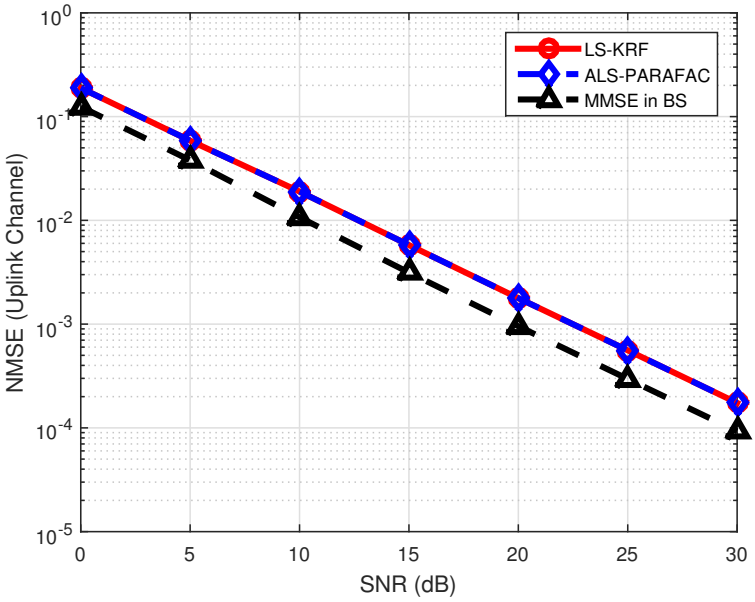
In Figures 27 and 28, the NMSE is evaluated as a function of the SNR. In these experiments, we consider a pilot signal of length $T = 16$, $U = 8$ MSs and $F = 8$ subcarriers for a "symmetric" scenario in which the SNR in the DL and UL are equal. As references for comparison, we also evaluate the performance of the MMSE-based channel estimators. These two approaches are referred in our plots as MMSE in MS and MMSE in BS. They represent the solutions to the conventional channel training framework in which the estimates for the DL and UL channels are computed at the MS and BS, respectively. The proposed ALS-PARAFAC and LS-KRF receivers present performance close to the competing MMSE estimators. However, the latter requires two independent estimation steps at the MSs and BS side. The proposed receivers concentrate the processing burden for channel estimation at the BS, avoiding such previously mentioned some practical limitations such as channel reciprocity as well as high complexity processing for channel estimation at each MS. Despite the same performance between the proposed tensor-based receivers, the LS-KRF receiver is more limited due to its more restrictive identifiability

Figure 27 – NMSE (Downlink Channel) vs. SNR (dB) for $N = 32$, $T = 16$, $F = 8$, $U = 8$. The SNR values in DL and UL are assumed to be equal.



Source: Created by the author.

Figure 28 – NMSE (Uplink Channel) vs. SNR (dB) for $N = 32$, $T = 16$, $F = 8$, $U = 8$. The SNR values in DL and UL are assumed to be equal.

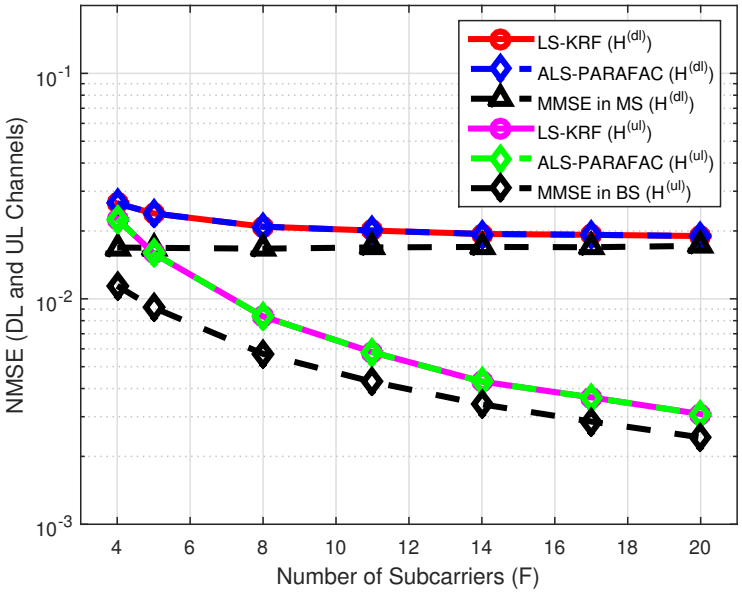


Source: Created by the author.

condition in (140).

In Figure 29, the performance is evaluated as a function of the number of subcarriers F. The parameters T and U are fixed to $T = 16$ and $U = 4$, respectively. The SNR in the DL and UL communications are set equal to 10 dB. This experiment shows the accuracy in the DL channel estimate is not affected by the number of subcarriers used. In contrast, when the number of subcarriers increases, the proposed receivers get better

Figure 29 – NMSE (DL and UL Channels) vs. number of subcarriers (F) for $N = 32$, $T = 16$, $U = 4$ and SNR = 10 dB.



Source: Created by the author.

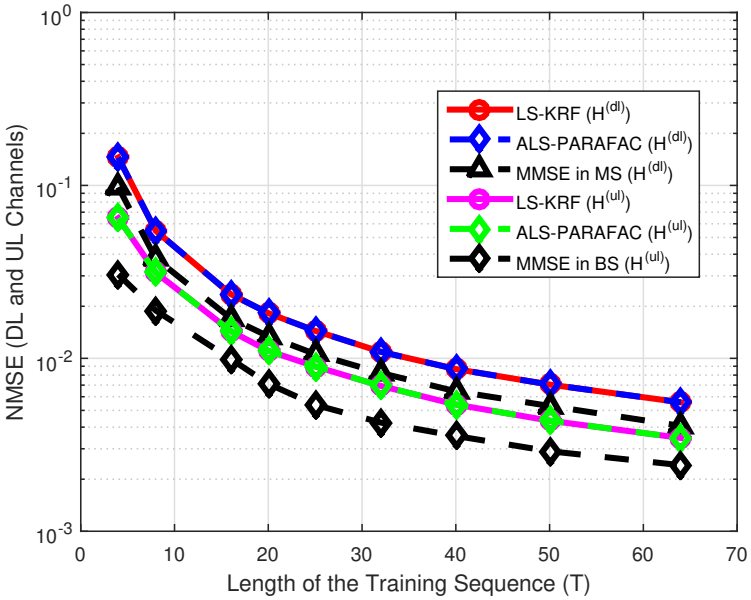
performance in the UL channel estimation. We also observe a strong proximity between the proposed methods and the MMSE-based approaches. The performance gap between the curves presents low variability when the number of subcarriers is modified.

In Figure 30, we plot the NMSE as a function of the length of the training sequence T. We consider $F = 4$, $U = 4$ and the same SNR value as the previous experiment. We can see that the performance of all methods improves when T increases. In the important case of $T < N$ we can notice the proposed LS-KRF and ALS-PARAFAC receivers present performance very close to the MMSE estimators. For a massive scenario, this result implies in a satisfactory performance of the proposed methods even when a training sequence of reduced length is used. In this case, the proposed receivers achieve a substancial training overhead reduction.

In Figure 31, we analyze the convergence rate of the proposed ALS-PARAFAC receiver. We plot the number of iterations for convergence as a function of the SNR by assuming different subcarrier numbers. In general, the ALS convergence depends on the factor matrices initialization and can be very slow when all matrices are unknown. However, we can see that the convergence of the ALS-PARAFAC receiver is achieved with few iterations due to the knowledge of the frequency spreading matrix $\boldsymbol{C}$ at the BS. We can also notice that the number of iterations for convergence is sensitive to the number of subcarriers and the difference is more pronounced when a reduced number of subcarriers is assumed in the UL transmission.

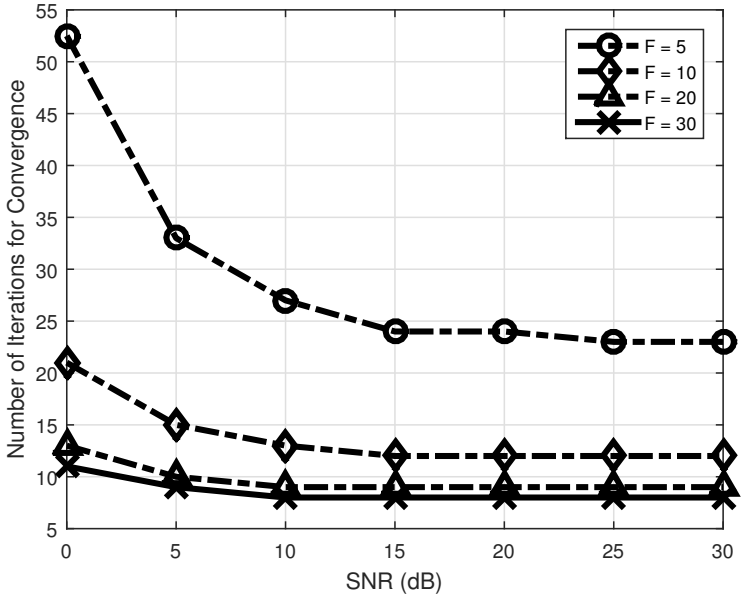
In the following section, we formulate an extension of the proposed methods able to opperate in mmWave MIMO scenarios. In this new approach, we consider that

Figure 30 – NMSE (DL and UL Channels) vs. length of the training sequence (T) for $N = 32$, $F = 4$, $U = 4$ and SNR = 10 dB.



Source: Created by the author.

Figure 31 – Number of iterations for convergence of the proposed ALS-PARAFAC receiver vs. SNR (dB) for $N = 32$, $T = 16$ and $U = 4$.



Source: Created by the author.

the BS is equipped with a HB architecture and the DL and UL channels assume sparse representations typical of mmWave propagation environments. Additional simulation results for performance evaluation are also provided in the sequence.

5.6 Extension to mmWave Massive MIMO Systems

Consider now a BS equipped with N_{BS} antennas serving simultaneously U MSs equipped with N_{MS} antennas. BS employs a HB architecture using M_{RF} RF chains. Due to the different instants of time dedicated to DL and UL communications, the beamforming matrix associated with DL transmission is denoted by $\boldsymbol{W} = \boldsymbol{W}_{\mathrm{RF}}\boldsymbol{W}_{\mathrm{BB}} \in \mathbb{C}^{N_{\mathrm{BS}}\times M_{\mathrm{RF}}}$, while the beamforming matrix associated with UL reception is denoted by $\boldsymbol{F} = \boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{BB}} \in \mathbb{C}^{N_{\mathrm{BS}}\times M_{\mathrm{RF}}}$. Note that equal beamforming matrices can also be considered in the transmission and reception phases without loss of generality. The BS transmits a length-T pilot sequence $\boldsymbol{s}_p \in \mathbb{C}^T$ over the p-th spatial direction using the beamforming vector $\boldsymbol{w}_p \in \mathbb{C}^{N_{\mathrm{BS}}}$ ($p = 1, \ldots, P$ and $P \leq M_{\mathrm{RF}}$). The received signal at the u-th MS over P different directions is given by

$$\boldsymbol{Y}_u = \boldsymbol{H}_u\boldsymbol{W}\boldsymbol{S} + \boldsymbol{V}_u^{(\mathrm{DL})} \in \mathbb{C}^{N_{\mathrm{MS}}\times T}, \tag{143}$$

where $\boldsymbol{H}_u \in \mathbb{C}^{N_{\mathrm{MS}}\times N_{\mathrm{BS}}}$ denotes the DL channel matrix associated to the u-th MS, $\boldsymbol{W} = [\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_P] \in \mathbb{C}^{N_{\mathrm{BS}}\times P}$ denotes the transmission beamforming matrix, $\boldsymbol{S} = [\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_P]^{\mathrm{T}} \in \mathbb{C}^{P\times T}$ concatenates the pilot signals to be sent by each transmission beam. The matrix $\boldsymbol{V}_u^{(\mathrm{DL})} \in \mathbb{C}^{N_{\mathrm{MS}}\times T}$ is the AWGN term at the u-th MS.

During the channel training phase, we assume identity matrices for the digital beamforming matrices, while the analog beamforming matrices have constant unit modulus entries with random phases. Therefore, the entries of $\boldsymbol{W}$ and $\boldsymbol{F}$ are chosen uniformly from a unit circle scaled by a constant $1/\sqrt{N_{\mathrm{BS}}}$, i.e.,

$$[\boldsymbol{W}]_{i,j} = \frac{1}{\sqrt{N_{\mathrm{BS}}}}e^{j\vartheta_{i,j}} \quad \text{and} \quad [\boldsymbol{F}]_{i,j} = \frac{1}{\sqrt{N_{\mathrm{BS}}}}e^{j\varphi_{i,j}}, \tag{144}$$

where $\vartheta_{i,j}$ and $\varphi_{i,j} \in [-\pi, \pi]$ follow a uniform distribution. Since this work deals with the channel estimation problem, the optimum design of the beamforming matrices is not adressed.

Making use of the proposed channel training framework, the coded version of (143) feed back to the BS at the f-th subcarrier is denoted by

$$\boldsymbol{Y}_{f,u} = \mathrm{diag}(\boldsymbol{c}_{f,u})\boldsymbol{Y}_u \in \mathbb{C}^{N_{\mathrm{MS}}\times T}, \tag{145}$$

where $\boldsymbol{c}_{f,u} \in \mathbb{C}^{N_{\mathrm{MS}}}$ denotes the known code vector associated with the f-th subcarrier used by the u-th MS.

In the UL communication, the BS employs Q beamforming vectors $\boldsymbol{f}_q \in \mathbb{C}^{N_{\mathrm{BS}}}$ ($q = 1, \ldots, Q$ and $Q \leq M_{\mathrm{RF}}$) to receive the coded UL pilot signals over a set of Q different spatial directions. The received closed-loop signal at the BS associated with the f-th subcarrier is given by

$$\begin{aligned} \boldsymbol{X}_f &= \boldsymbol{F}^{\mathrm{H}} \left(\sum_{u=1}^{U} \boldsymbol{G}_u \boldsymbol{Y}_{f,u} + \boldsymbol{V}^{(\mathrm{UL})} \right) \\ &= \left(\boldsymbol{F}^{\mathrm{H}} \boldsymbol{G}_{\mathrm{e}} \right) \mathrm{diag}(\boldsymbol{c}_f) \boldsymbol{Y}_{\mathrm{e}}^{\mathrm{T}} + \boldsymbol{F}^{\mathrm{H}} \boldsymbol{V}^{(\mathrm{UL})} \in \mathbb{C}^{Q \times T}, \end{aligned} \tag{146}$$

where $\boldsymbol{G}_{\mathrm{e}} = [\boldsymbol{G}_1, \boldsymbol{G}_2, \ldots, \boldsymbol{G}_U] \in \mathbb{C}^{N_{\mathrm{BS}} \times U N_{\mathrm{MS}}}$ denotes an extended version of the UL channel matrix that concatenates the U UL channel matrices $\boldsymbol{G}_u \in \mathbb{C}^{N_{\mathrm{BS}} \times N_{\mathrm{MS}}}$ $(u = 1, \ldots, U)$ of all MSs, $\boldsymbol{Y}_{\mathrm{e}} = \left[\boldsymbol{Y}_1^{\mathrm{T}}, \boldsymbol{Y}_2^{\mathrm{T}}, \ldots, \boldsymbol{Y}_U^{\mathrm{T}} \right] \in \mathbb{C}^{T \times U N_{\mathrm{MS}}}$ denotes an extended matrix that concatenates the feed back signals send by all MSs, $\boldsymbol{c}_f = \left[\boldsymbol{c}_{f,1}^{\mathrm{T}}, \boldsymbol{c}_{f,2}^{\mathrm{T}}, \ldots, \boldsymbol{c}_{f,U}^{\mathrm{T}} \right]^{\mathrm{T}} \in \mathbb{C}^{U N_{\mathrm{MS}}}$ is an extended code vector that contains the coding vectors of all MSs at the f-th subcarrier, and $\boldsymbol{F}^{\mathrm{H}} \boldsymbol{V}^{(\mathrm{UL})}$ represents the filtered noise version at the RF chains output.

Remark: For this approach, we particularly consider a mmWave propagation environment with limited scattering. According (ALKHATEEB *et al.*, 2014a), mmWave channels can be modeled by a clustered channel model with few L_u dominant paths between the u-th MS and the BS. Therefore, the DL channel matrix $\boldsymbol{H}_u \in \mathbb{C}^{N_{\mathrm{MS}} \times N_{\mathrm{BS}}}$ associated with the u-th MS can be expressed as

$$\boldsymbol{H}_u = \sum_{l=1}^{L_u} \alpha_{u,l} \boldsymbol{a}_{\mathrm{MS}}(\theta_{u,l}) \boldsymbol{a}_{\mathrm{BS}}(\phi_{u,l})^{\mathrm{T}}, \tag{147}$$

where $\alpha_{u,l}$ denotes the complex path gain of the u-th MS related to the l-th path. The path gains are modeled as circular symmetric Gaussian random variables with zero mean and unit variance. $\boldsymbol{a}_{\mathrm{MS}}(\theta_{u,l}) \in \mathbb{C}^{N_{\mathrm{MS}}}$ and $\boldsymbol{a}_{\mathrm{BS}}(\phi_{u,l}) \in \mathbb{C}^{N_{\mathrm{BS}}}$ are the antenna array response vectors evaluated at angle of arrival $\theta_{u,l}$ and angle of departure $\phi_{u,l}$ uniformly distributed in the interval $[0, 2\pi]$. We assume ULA at the BS and MSs. However, the proposed method can be applied to arbitrary array geometries without loss of generality. For ULA configurations with inter-antennas spacing equals to $d = \frac{\lambda}{2}$, where λ denotes the wavelength of the signal, the array response vectors at the MS and BS can be modeled as follows

$$\boldsymbol{a}_{\mathrm{MS}}(\theta_{u,l}) = \frac{1}{\sqrt{N_{\mathrm{MS}}}} \left[1, e^{j\pi \cos\theta_{u,l}}, \cdots, e^{j\pi (N_{\mathrm{MS}}-1) \cos\theta_{u,l}} \right]^{\mathrm{T}}, \tag{148}$$

$$\boldsymbol{a}_{\mathrm{BS}}(\phi_{u,l}) = \frac{1}{\sqrt{N_{\mathrm{BS}}}} \left[1, e^{j\pi \cos\phi_{u,l}}, \cdots, e^{j\pi (N_{\mathrm{BS}}-1) \cos\phi_{u,l}} \right]^{\mathrm{T}}. \tag{149}$$

Using a matrix fashion notation, $\boldsymbol{H}_u$ can be rewritten as

$$\boldsymbol{H}_u = \boldsymbol{A}_{\mathrm{MS}} \mathrm{diag}(\boldsymbol{\alpha}) \boldsymbol{A}_{\mathrm{BS}}^{\mathrm{T}}, \tag{150}$$

where $\boldsymbol{\alpha} = \sqrt{N_{\mathrm{MS}} N_{\mathrm{BS}} / L_u} \left[\alpha_{u,1}, \alpha_{u,2}, \ldots, \alpha_{u,L_u} \right]^{\mathrm{T}} \in \mathbb{C}^{L_u}$ denotes the vector that contains the L_u path gains in the DL. The array response matrices $\boldsymbol{A}_{\mathrm{MS}} \in \mathbb{C}^{N_{\mathrm{MS}} \times L_u}$ and $\boldsymbol{A}_{\mathrm{BS}} \in$

$\mathbb{C}^{N_{\rm BS}\times L_u}$ at the MS and BS are expressed by

$$\boldsymbol{A}_{\rm MS} = [\boldsymbol{a}_{\rm MS}(\theta_{u,1}), \boldsymbol{a}_{\rm MS}(\theta_{u,2}), \cdots, \boldsymbol{a}_{\rm MS}(\theta_{u,L_u})], \quad (151)$$

$$\boldsymbol{A}_{\rm BS} = [\boldsymbol{a}_{\rm BS}(\phi_{u,1}), \boldsymbol{a}_{\rm BS}(\phi_{u,2}), \cdots, \boldsymbol{a}_{\rm BS}(\phi_{u,L_u})]. \quad (152)$$

The UL channel matrix $\boldsymbol{G}_u \in \mathbb{C}^{N_{\rm BS}\times N_{\rm MS}}$ from the u-th MS to the BS can be represented in a similar way. We define $\boldsymbol{G}_u$ as follows

$$\boldsymbol{G}_u = \boldsymbol{A}_{\rm BS}{\rm diag}(\boldsymbol{\beta})\boldsymbol{A}_{\rm MS}^{\rm T}, \quad (153)$$

where $\boldsymbol{A}_{\rm BS} \in \mathbb{C}^{N_{\rm BS}\times M_u}$ and $\boldsymbol{A}_{\rm MS} \in \mathbb{C}^{N_{\rm MS}\times M_u}$ are now function of the spatial parameters in the UL, while $\boldsymbol{\beta} = \sqrt{N_{\rm MS}N_{\rm BS}/M_u}\,[\beta_{u,1}, \beta_{u,2}, \ldots, \beta_{u,M_u}]^{\rm T} \in \mathbb{C}^{M_u}$ denotes the vector that contains the M_u path gains of the UL channel.

Note that, similar to (120) the noiseless signal term in (146) denotes the f-th frontal slice of the PARAFAC tensor

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{I}}_{3,U\cdot N_{\rm MS}} \times_1 \left(\boldsymbol{F}^{\rm H}\boldsymbol{G}_{\rm e}\right) \times_2 \boldsymbol{Y}_{\rm e} \times_3 \boldsymbol{C} \in \mathbb{C}^{Q\times T\times F}, \quad (154)$$

where

$$\left(\boldsymbol{A}^{(1)}, \boldsymbol{A}^{(2)}, \boldsymbol{A}^{(3)}\right) \leftrightarrow \left(\boldsymbol{F}^{\rm H}\boldsymbol{G}_{\rm e}, \boldsymbol{Y}_{\rm e}, \boldsymbol{C}\right)$$
$$(I_1, I_2, I_3, Q) \leftrightarrow (Q, T, F, UN_{\rm MS}).$$

The three dimensions, or modes, of $\boldsymbol{\mathcal{X}}$ stands for the number of receive beams, pilot signal length and number of subcarriers. The matrix $\boldsymbol{F}$ can be seen as a compression matrix associated with the first mode of $\boldsymbol{\mathcal{X}}$ which reduces the size of the first mode from $N_{\rm BS}$ to Q RF chains (i.e., number of beams). The f-th row of the frequency spreading matrix $\boldsymbol{C} \in \mathbb{C}^{F\times UN_{\rm MS}}$ contains the code coefficients used by the U MSs at the f-th subcarrier, i.e.,

$$\boldsymbol{C} = [\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_K]^{\rm T} \in \mathbb{C}^{F\times UN_{\rm MS}}. \quad (155)$$

Remark: Note that, the use of the proposed ALS-PARAFAC 10 and LS-KRF 11 receivers to estimate the compressed DL and UL channel matrices $\boldsymbol{F}^{\rm H}\boldsymbol{G}_{\rm e} \in \mathbb{C}^{Q\times UN_{\rm MS}}$ and $\boldsymbol{Y}_{\rm e} \in \mathbb{C}^{T\times UN_{\rm MS}}$ from (154) is straightforward.

5.6.1 Sparse Formulation for DL and UL Channel Parameters Estimation

A second processing step consists of estimating the channel parameters (AoDs, AoAs and path gains) to reconstruct the channel matrices $\hat{\boldsymbol{H}}_u$ and $\hat{\boldsymbol{G}}_u$ ($u = 1, \ldots, U$).

Let us introduce the following short block representation:

$$\boldsymbol{F}^{\mathrm{H}}\hat{\boldsymbol{G}}_{\mathrm{e}} = \left[\hat{\boldsymbol{\Phi}}_1, \hat{\boldsymbol{\Phi}}_2, \cdots, \hat{\boldsymbol{\Phi}}_U\right] \in \mathbb{C}^{Q\times UN_{\mathrm{MS}}}, \tag{156}$$

where

$$\hat{\boldsymbol{\Phi}}_u = \boldsymbol{F}^{\mathrm{H}}\hat{\boldsymbol{G}}_u \in \mathbb{C}^{Q\times N_{\mathrm{MS}}}, \quad u = 1, \ldots, U. \tag{157}$$

By replacing $\hat{\boldsymbol{G}}_u$ for (153), and then vectorizing (157) according to property (8), we get

$$\hat{\boldsymbol{\varphi}}_u = \mathrm{vec}\left(\hat{\boldsymbol{\Phi}}_u\right) = \left(\hat{\boldsymbol{A}}_{\mathrm{MS}} \diamond \boldsymbol{F}^{\mathrm{H}}\hat{\boldsymbol{A}}_{\mathrm{BS}}\right)\hat{\boldsymbol{\beta}}. \tag{158}$$

Using the property (6), we straightforwardly obtain

$$\hat{\boldsymbol{\varphi}}_u = \left(\boldsymbol{I}_{N_{\mathrm{MS}}} \otimes \boldsymbol{F}^{\mathrm{H}}\right)\left(\hat{\boldsymbol{A}}_{\mathrm{MS}} \diamond \hat{\boldsymbol{A}}_{\mathrm{BS}}\right)\hat{\boldsymbol{\beta}} \in \mathbb{C}^{QN_{\mathrm{MS}}}, \tag{159}$$

where $\boldsymbol{I}_{N_{\mathrm{MS}}}$ denotes an identity matrix of size $N_{\mathrm{MS}} \times N_{\mathrm{MS}}$.

The same procedure can be directly performed in the u-th block $\hat{\boldsymbol{Y}}_u^{\mathrm{T}}$ of the estimated factor matrix

$$\hat{\boldsymbol{Y}}_{\mathrm{e}} = \left[\hat{\boldsymbol{Y}}_1^{\mathrm{T}}, \hat{\boldsymbol{Y}}_2^{\mathrm{T}}, \cdots, \hat{\boldsymbol{Y}}_U^{\mathrm{T}}\right] \in \mathbb{C}^{T\times UN_{\mathrm{MS}}}, \tag{160}$$

where

$$\hat{\boldsymbol{Y}}_u^{\mathrm{T}} = \boldsymbol{S}^{\mathrm{T}}\boldsymbol{W}^{\mathrm{T}}\hat{\boldsymbol{H}}_u^{\mathrm{T}} + \boldsymbol{V}_u^{(\mathrm{DL})^{\mathrm{T}}} \in \mathbb{C}^{T\times N_{\mathrm{MS}}}. \tag{161}$$

Similar to (159), we can obtain the following vector formulation

$$\hat{\boldsymbol{y}}_u = \left(\boldsymbol{I}_{N_{\mathrm{MS}}} \otimes \boldsymbol{S}^{\mathrm{T}}\boldsymbol{W}^{\mathrm{T}}\right)\left(\hat{\boldsymbol{A}}_{\mathrm{MS}} \diamond \hat{\boldsymbol{A}}_{\mathrm{BS}}\right)\hat{\boldsymbol{\alpha}} + \boldsymbol{v}_u^{(\mathrm{DL})} \in \mathbb{C}^{TN_{\mathrm{MS}}}, \tag{162}$$

where $\hat{\boldsymbol{y}}_u = \mathrm{vec}\left(\hat{\boldsymbol{Y}}_u^{\mathrm{T}}\right)$ and $\boldsymbol{v}_u^{(\mathrm{DL})} = \mathrm{vec}\left(\boldsymbol{V}_u^{(\mathrm{DL})^{\mathrm{T}}}\right)$.

From (159) and (162), two independent CS problems can be formulated to estimate the channel parameters of the u-th MS. We assume that grid quantization errors are neglected, i.e., the AoDs and AoAs are drawn from a uniform angle grid of N_g points contained in the set $\left\{0, \frac{2\pi}{N_g}, \ldots, \frac{2\pi(N_g-1)}{N_g}\right\}$, with $N_g \gg L_u$ and $N_g \gg M_u$. Based on this assumption, we can obtain the following sparse formulations for the vectors $\hat{\boldsymbol{y}}_u$ and $\hat{\boldsymbol{\varphi}}_u$:

$$\hat{\boldsymbol{y}}_u = \left(\boldsymbol{I}_{N_{\mathrm{MS}}} \otimes \boldsymbol{S}^{\mathrm{T}}\boldsymbol{W}^{\mathrm{T}}\right)\boldsymbol{\Sigma}_{\mathrm{D}}\bar{\boldsymbol{\alpha}}, \tag{163}$$

$$\hat{\boldsymbol{\varphi}}_u = \left(\boldsymbol{I}_{N_{\mathrm{MS}}} \otimes \boldsymbol{F}^{\mathrm{H}}\right)\boldsymbol{\Sigma}_{\mathrm{D}}\bar{\boldsymbol{\beta}}, \tag{164}$$

where $\boldsymbol{\Sigma}_{\mathrm{D}} = \bar{\boldsymbol{A}}_{\mathrm{MS}} \otimes \bar{\boldsymbol{A}}_{\mathrm{BS}} \in \mathbb{C}^{N_{\mathrm{MS}}N_{\mathrm{BS}}\times N_g^2}$ denotes the known dictionary matrix used to solve the sparse signal recovery problem. The matrices $\bar{\boldsymbol{A}}_{\mathrm{MS}} \in \mathbb{C}^{N_{\mathrm{MS}}\times N_g}$ and $\bar{\boldsymbol{A}}_{\mathrm{BS}} \in \mathbb{C}^{N_{\mathrm{BS}}\times N_g}$

that build the dictionary are denoted by

$$\bar{\boldsymbol{A}}_{\mathrm{MS}} = \left[\boldsymbol{a}_{\mathrm{MS}}\left(0\right), \boldsymbol{a}_{\mathrm{MS}}\left(\frac{2\pi}{N_g}\right), \ldots, \boldsymbol{a}_{\mathrm{MS}}\left(\frac{2\pi(N_g-1)}{N_g}\right)\right], \tag{165}$$

$$\bar{\boldsymbol{A}}_{\mathrm{BS}} = \left[\boldsymbol{a}_{\mathrm{BS}}\left(0\right), \boldsymbol{a}_{\mathrm{BS}}\left(\frac{2\pi}{N_g}\right), \ldots, \boldsymbol{a}_{\mathrm{BS}}\left(\frac{2\pi(N_g-1)}{N_g}\right)\right], \tag{166}$$

and contains all points of the angle grid. The left hand side matrices $\left(\boldsymbol{I}_{N_{\mathrm{MS}}} \otimes \boldsymbol{S}^{\mathrm{T}}\boldsymbol{W}^{\mathrm{T}}\right)$ and $\left(\boldsymbol{I}_{N_{\mathrm{MS}}} \otimes \boldsymbol{F}^{\mathrm{H}}\right)$ are called measurement matrices of the sparse problems. $\bar{\boldsymbol{\alpha}} \in \mathbb{C}^{N_g^2}$ and $\bar{\boldsymbol{\beta}} \in \mathbb{C}^{N_g^2}$ are sparse vectors obtained by augmenting the vector gains $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ with zero elements, respectively.

Estimates for the channel parameters can be obtained by applying state-of-the-art CS algorithms in the estimated sparse vectors (163) and (164). Many efficient algorithms such as orthogonal matching pursuit (OMP) (PATI, REZAIIFAR, and KRISHNAPRASAD, 1993), structured compressive sampling matching pursuit (S-CoSaMP) (SHEN *et al.*, 2016) and fast iterative shrinkage-thresholding (FISTA) (BECK and TEBOULLE, 2009), to name a few, can be used to solve these two sparse signal recovery problems. In a simplified view, the estimates for the path gains $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ correspond to non-zero entries of the estimated sparse vectors $\bar{\boldsymbol{\alpha}}$ and $\bar{\boldsymbol{\beta}}$, while estimation for the AoDs and AoAs are obtained by selecting the columns of the dictionary matrix $\boldsymbol{\Sigma}_{\mathrm{D}}$ related to the positions of the estimated path gains in the sparse vector.

Finally, from the estimated channel parameters the BS can construct the estimated DL and UL channel matrices $\hat{\boldsymbol{H}}_u$ and $\hat{\boldsymbol{G}}_u$ of the u-th MS according to relations (150) and (153) as follows:

$$\hat{\boldsymbol{H}}_u = \hat{\boldsymbol{A}}_{\mathrm{MS}}\mathrm{diag}(\hat{\boldsymbol{\alpha}})\hat{\boldsymbol{A}}_{\mathrm{BS}}^{\mathrm{T}}, \tag{167}$$

$$\hat{\boldsymbol{G}}_u = \hat{\boldsymbol{A}}_{\mathrm{BS}}\mathrm{diag}(\hat{\boldsymbol{\beta}})\hat{\boldsymbol{A}}_{\mathrm{MS}}^{\mathrm{T}}. \tag{168}$$

Remark: Compared to (159), the sparse signal recovery problem formulated from (162) naturally incorporates the DL noise contribution in its structure. For this reason, the proposed framework can lead to some performance degradation in the DL channel estimation compared to UL channel estimation. Therefore, we can observe a trade-off between DL channel estimation accuracy performed by the BS and reduction of the processing cost for channel estimation at the users' side. This discussion is later reinforced by means of numerical simulations in the next section.

5.6.2 Identifiability Condition

Unique LS solutions for the compressed DL and UL channel matrices $\boldsymbol{F}^{\mathrm{H}}\boldsymbol{G}_{\mathrm{e}}$ and $\boldsymbol{Y}_{\mathrm{e}}$ obtained from the 1-mode and 2-mode unfoldings of (154) requires that $\left(\boldsymbol{C} \diamond \boldsymbol{Y}_{\mathrm{e}}\right)^{\mathrm{T}} \in \mathbb{C}^{UN_{\mathrm{MS}} \times FT}$ and $\left(\boldsymbol{C} \diamond \boldsymbol{F}^{\mathrm{H}}\boldsymbol{G}_{\mathrm{e}}\right)^{\mathrm{T}} \in \mathbb{C}^{UN_{\mathrm{MS}} \times FQ}$ be full row-rank to be right-invertible. For

this, the following two conditions must be satisfied:

$$FT \geqslant UN_{\mathrm{MS}} \quad \text{and} \quad FQ \geqslant UN_{\mathrm{MS}}. \tag{169}$$

Combining these conditions yields the following lower bound on the number of subcarriers:

$$F \geq \max\left(\left\lceil \frac{UN_{\mathrm{MS}}}{T}, \frac{UN_{\mathrm{MS}}}{Q} \right\rceil\right), \tag{170}$$

where $\lceil x \rceil$ denotes the smallest integer number that is greater or equal to x.

5.7 Simulation Results (Part 2)

Here, we provide a second set of simulation results referring to Section 5.6.1. We compare the proposed method with the conventional approach illustrated in Figure 25. The state-of-the-art compressed sensing OMP algorithm (PATI, REZAIIFAR, and KRISHNAPRASAD, 1993) is used in both MSs and BS as a channel estimation technique. It is also considered as the second stage of our algorithm. Both MSs and BS employ ULA with half wavelength spaced antennas. We set $N_{\mathrm{BS}} = 32$, $N_{\mathrm{MS}} = 16$, $U = 2$, $N_g = 64$ and equal SNR values for the DL and UL communications in all experiments. The obtained results are averaged over 1000 independent Monte Carlo runs. In each run, the DL and UL channel matrices with $L_u = 3$ and $M_u = 3$ paths per user and HB matrices are generated in accordance with Equations (150), (153) and (144), respectively. The pilot signal $\boldsymbol{S}$ is a BPSK modulated matrix, and the frequency spreading matrix $\boldsymbol{C}$ is a random matrix that follows a uniform distribution.

The algorithm performance is evaluated in terms of the NMSE between the estimated and true channel matrices, given by

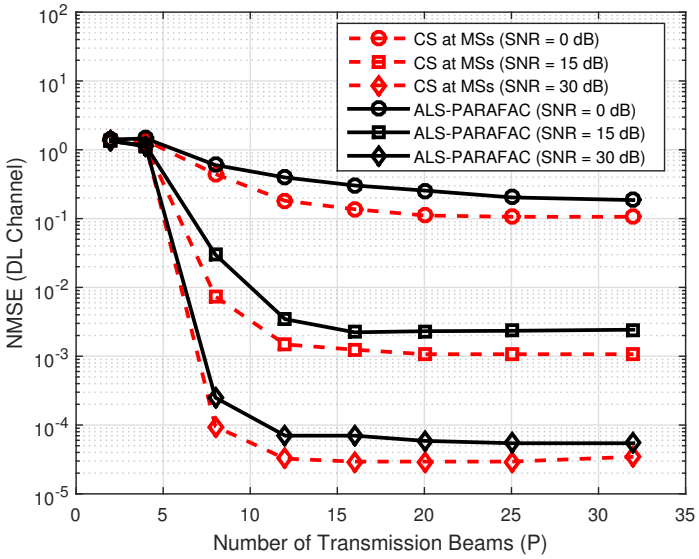
$$\mathrm{NMSE}\left(\hat{\boldsymbol{H}}\right) = \frac{\sum_{u=1}^{U} \|\boldsymbol{H}_u - \hat{\boldsymbol{H}}_u\|_{\mathrm{F}}^2}{\sum_{u=1}^{U} \|\boldsymbol{H}_u\|_{\mathrm{F}}^2}, \tag{171}$$

$$\mathrm{NMSE}\left(\hat{\boldsymbol{G}}\right) = \frac{\sum_{u=1}^{U} \|\boldsymbol{G}_u - \hat{\boldsymbol{G}}_u\|_{\mathrm{F}}^2}{\sum_{u=1}^{U} \|\boldsymbol{G}_u\|_{\mathrm{F}}^2}. \tag{172}$$

In our experiments, we selected some system configurations by varying the SNR values, number of transmission (P) and reception (Q) beams, number of subcarriers (F) and the length of the pilot signal (T).

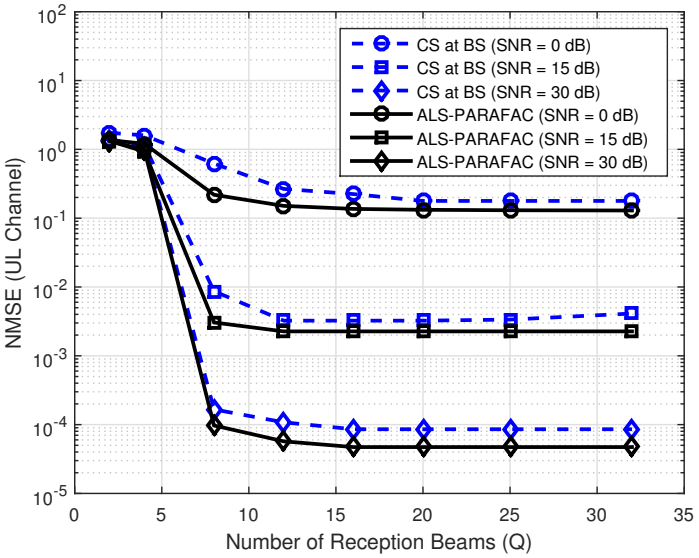
Figures 32 and 33 show the NMSE as a function of the number of transmission (P) and reception (Q) beams. The proposed method ouperforms the classical framework to the UL channel estimation, while the DL performance is worse in the all SNR range considered. From this experiment, we observe a trade-off between DL channel estimation accuracy and computational complexity for channel estimation. In other words, the pro-

Figure 32 – NMSE vs. number of transmission beams (P) for the DL channel estimation, $U = 2$, $T = 16$ and $F = 25$.



Source: Created by the author.

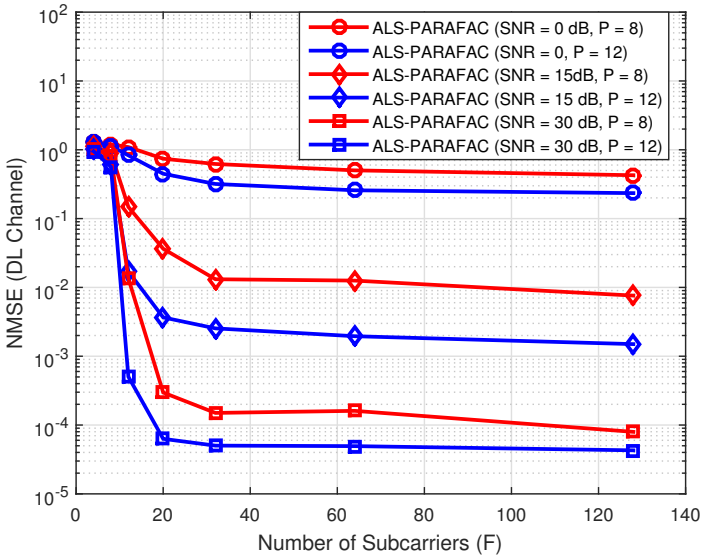
Figure 33 – NMSE vs. number of reception beams (Q) for the UL channel estimation, $U = 2$, $T = 16$ and $F = 25$.



Source: Created by the author.

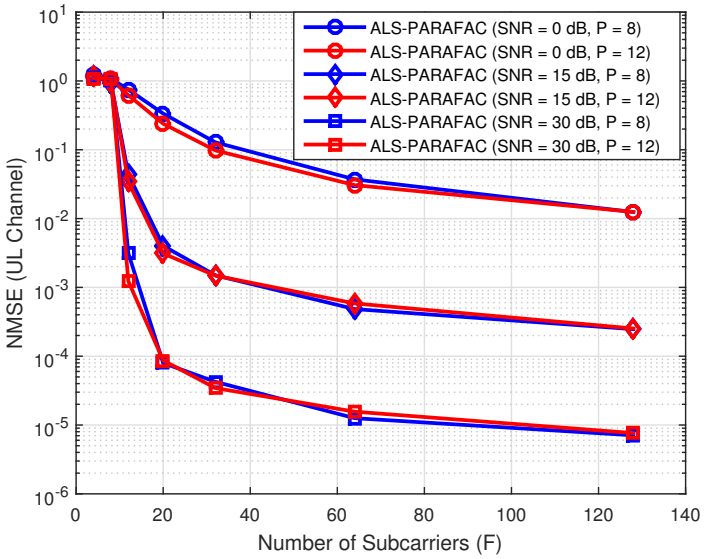
posed method concentrate most of the processing burden for channel estimation at the BS side, while a better performance of DL channel estimation is accompanied of a high computational cost to complex channel estimation processing at the MSs side when the conventional framework is considered. On the other hand, the performance loss at DL is compensated with more accurate estimations at UL. In addition, the NMSE performance is not influenced by the number of RF chains when P and Q are greater than 12. This result reveals that the proposed method provides good channel estimation accuracy even when the BS is equipped with few number of RF chains.

Figure 34 – NMSE vs. number of subcarriers (F) for the DL channel estimation, $U = 2$ and $T = 16$.



Source: Created by the author.

Figure 35 – NMSE vs. number of subcarriers (F) for the UL channel estimation, $U = 2$ and $T = 16$.

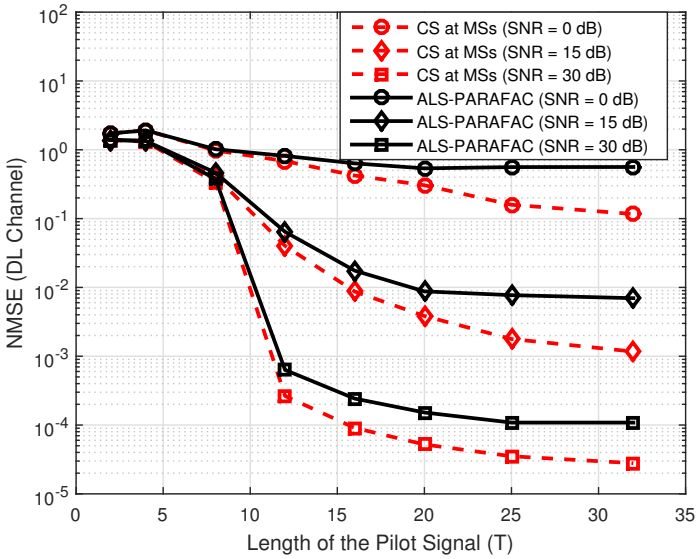


Source: Created by the author.

In Figures 34 and 35, the NMSE performance is avaluated as a function of the number of subcarriers (F). An increase of F leads to an improved performance only until $F = 32$ subcarriers in the DL channel estimation, while for the UL channel estimation this value is approximately equal to $F = 64$ subcarriers. This result show that the proposed method can operate with few frequency resources to estimate jointly the MSs channels with high accuracy.

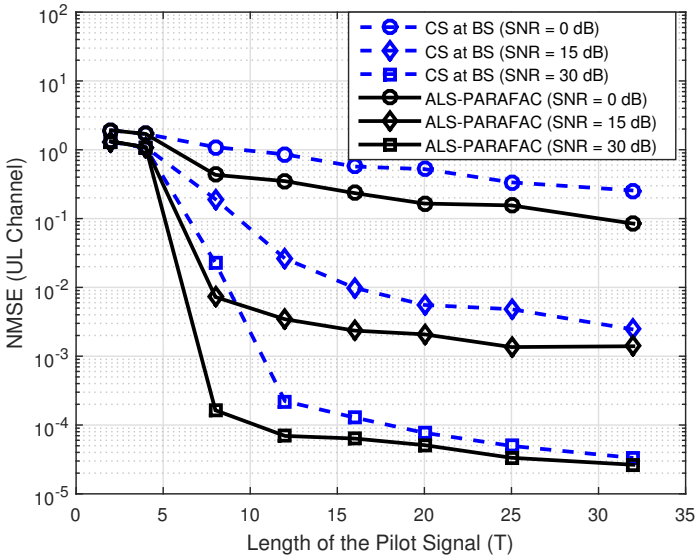
In Figures 36 and 37, we evaluate the NMSE performance in terms of the pilot signal length (T). In this experiment we also set $U = 2$, $P = 8$ and $F = 25$. Here we

Figure 36 – NMSE vs. length of the pilot signal (T) for the DL channel estimation, $U = 2$, $P = 8$ and $F = 25$.



Source: Created by the author.

Figure 37 – NMSE vs. length of the pilot signal (T) for the UL channel estimation, $U = 2$, $P = 8$ and $F = 25$.



Source: Created by the author.

conclude that short pilots signals are necessary to estimate the DL and UL channels by means of the proposed method. Low variability in the NMSE values is observed when $T > 15$. For a massive MIMO scenario, this result implies in a substancial reduction in the pilot overhead to training.

5.8 Chapter Summary

. In this chapter, we have addressed the joint DL and UL channel estimation problem in MU-MIMO wireless communications systems. We first proposed a novel multi-carrier based training scheme able to concentrate the processing burden for channel estimation at the BS. Then, we have formulated two tensor-based receivers for joint DL and UL channel estimation. In addition, we have also proposed an extension of these methods able to operate in mmWave propagation environments. Simulation results showed that the proposed receivers achieve a performance close to the classical channel estimation framework, with the advantage of avoiding complex processing for channel estimation at the MSs, which in many cases are power-limited devices.

6 TENSOR-BASED SEMI-BLIND RECEIVER FOR JOINT CHANNEL AND PHASE NOISE ESTIMATION IN FREQUENCY-SELECTIVE MIMO SYSTEMS

The MIMO technologies significantly enhance the spectral efficiency of wireless communications systems. However, its performance can be severely degrated due to oscillator imperfections that lead to an unknown phase noise (PN) per transmit and receive antenna which needs to be compensated. In this chapter, we propose a new method to solve the joint channel and PN estimation problem in frequency-selective MIMO systems. By assuming a transmission scheme in which each frame is divided into small sub-frames so that the PN varies sub-frame to sub-frame, the received signal can be formulated using tensor modeling concepts. From the obtained tensor model, we propose an iterative semi-blind receiver for channel estimation in the presence of PN perturbations. The proposed receiver consists of two stages. In the first stage, the frequency-selective MIMO channel is directly estimated through a tensor-based ALS-PARAFAC algorithm that fits a PARAFAC model to the noisy received signal. In the second one, the LS-KRF algorithm is used to extract the PN components at the transmitter and receiver. The identifiability conditions of the proposed tensor model and the computational complexity of the our receiver algorithm are also analyzed. Our approach dispenses idealized assumptions such as perfect knowledge of the channel and statistical distribution for the PN process, which makes it applicable to more challenging scenarios compared to other solutions in the literature.

6.1 Introduction and Motivation

Modern wireless communications systems are expected to support high data rates by increasing the spectral efficiency of 5G New Radio (NR) with Frequency Range (FG) 1 up to 6 GHz and by mitigating the losses in the 5G NR with FR 2 from 24 to 86 GHz. MIMO technologies achieve such improved spectral efficiency by employing multiple antennas at both the transmitter and the receiver to exploit the four typical array signal processing gains, namely, array diversity, multiplexing and reduced interference gains (FOSCHINI and GANS, 1998; TAROKH, SESHADRI, and CALDERBANK, 1998). However, for the practical implementation of MIMO systems, there are hardware impairments due to non-ideal radio frequency (RF) front-ends, such as oscillator imperfections that result in unknown phase noise (PN) per antenna. Therefore, distortions can be introduced in the transmitted and received signal leading to a severe performance loss. Moreover, the instantaneous channel acquisition that can be used to improve system performance through precoding/beamforming techniques becomes a challenging problem since the effective channel is a distorted version of the true channel, and thus the PN has to be compensated in the channel acquisition step. Therefore, new approaches that can

jointly estimate the channel and PN at the transmitter and receiver have great practical appeal.

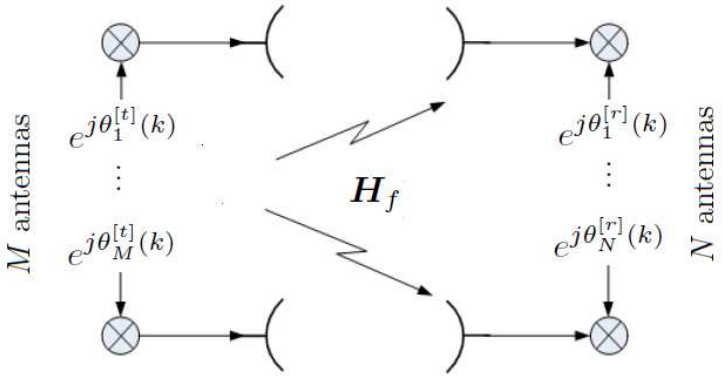
The PN compensation problem has been extensively studied in the past years. For instance, in (PETROVIC, RAVE, and FETTWEIS, 2007) and (KIM *et al.*, 2008), techniques for PN compensation are proposed for single-input single-output (SISO) systems. However, they assume perfect CSI knowledge at the receiver, which is not feasible in practice. In MIMO systems, (NGEBANI *et al.*, 2017) proposes a novel placement of pilot carriers in the preamble and data portion of the MIMO-OFDM frame for joint channel and PN estimation. The authors in (HADASCHIK *et al.*, 2005) and (ISHAQUE and ASCHEID, 2012) propose compensation schemes based on the knowledge of the statistical modelling of PN process. Note that, in contrast to (HADASCHIK *et al.*, 2005; ISHAQUE and ASCHEID, 2012), our goal is to devise a method that can provide accurate channel and PN estimation by avoiding perfect CSI knowledge and at the same time being robust to PN model variations, i.e., without requiring also explicit knowledge of the PN model.

In this chapter, we propose a new tensor-based method for frequency-selective MIMO channel estimation in the presence of PN impairments. By assuming that each transmitted frame is divided into multiple sub-frames and that the PN perturbations vary between sub-frames, we show that the received signal can be modeled as a third-order PARAFAC decomposition. Motivated by the multidimensional structure of the received signal, we propose a two-stage iterative semi-blind receiver for the joint estimation of the channel and PN. In the first stage, estimates of the channel gains for different frequencies are obtained by means of an ALS algorithm that fits a PARAFAC model to the noisy received signal, while the second one obtain closed-form estimates of the PN per antenna port using a LS-KRF algorithm. The identifiability conditions of the proposed tensor model and the computational complexity of our two-stage receiver algorithm are also discussed. The proposed receiver does not require perfect knowledge of the channel and PN model. Therefore, our solution becomes more attractive compared to other works in the literature with non-realistic assumptions. Simulation results show the effectiveness and high accuracy of the proposed receiver for joint estimation of the channel and PN impairments.

6.1.1 Chapter Organization

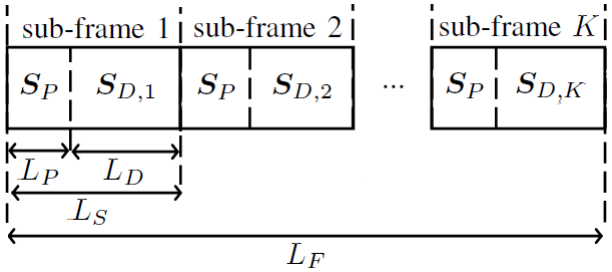
This chapter is organized as follows. The data model and main assumptions for the considered frequency-selective MIMO system with PN are presented in Section 6.2. In Section 6.3, the joint channel and PN estimation problem is recast using tensor-based modeling. Then, the proposed semi-blind receiver algorithm is formulated in Section 6.4. The uniqueness issues of the proposed tensor-based system model and the computational complexity of the proposed receiver algorithm are analyzed in Section 6.5. Simulation

Figure 38 – Frequency-selective MIMO system equipped with different oscilators at the M transmit and N receive antennas.



Source: Adapted from (HUANG, WANG, and HE, 2015).

Figure 39 – Illustration of frame and sub-frame structures. The pilot symbols part $\boldsymbol{S}_P$ is reused sub-frame to sub-frame. The PN is invariant within a sub-frame while the channel gains remain constant over the length of one frame[1].



Source: Created by the author.

results are shown in Section 6.6. Finally, Section 6.7 brings some concluding remarks.

6.2 System Model

Let us consider a point to point frequency-selective MIMO system with transmitter and receiver having M and N antennas, respectively. Each transmit and receive antenna is equipped with its own independent oscillator so that the PN is assumed to be different between the antennas, as illustrated in Figure 38. Each transmitted frame of length $L_F = K \cdot L_S$ is composed by K sub-frames, each having a length of L_S samples. Each sub-frame consists of a known pilot symbols part $\boldsymbol{S}_P \in \mathbb{C}^{M \times L_P}$ of length L_P and an unknown data part $\boldsymbol{S}_{D,k} \in \mathbb{C}^{M \times L_D}$ of length L_D. Thus, each transmitted sub-frame has size $L_S = L_P + L_D$. The pilot symbols part is reused sub-frame to sub-frame. Figure 39 illustrates the considered frame structure. We assume that the system operates with high sampling rate and the PN is time-invariant[1] within a small sub-frame, but varying from sub-frame to sub-frame. The channel is frequency-selective with L taps assumed to remain constant over the length of one frame, i.e., the channel vary more slowly than the

[1]For most practical oscillators the temporal innovation variance of the PN at each transmit and receive antenna is very small in order of $(10^{-3}, 10^{-5})$ rad^2 (MEHRPOUYAN *et al.*, 2012). Therefore, it is a valid assumption for sub-frames with small size.

PN process. In the frequency domain, the received signal $\boldsymbol{X}_{k,f} \in \mathbb{C}^{N \times L_S}$ associated with the k-th sub-frame at the f-th frequency is denoted by

$$\boldsymbol{X}_{k,f} = \boldsymbol{W} D_k \left(\boldsymbol{\Phi}^{[r]} \right) \boldsymbol{H}_f D_k \left(\boldsymbol{\Phi}^{[t]} \right) \boldsymbol{S}_k + \boldsymbol{W} \boldsymbol{V}_{k,f}, \tag{173}$$

where $\boldsymbol{W} \in \mathbb{C}^{N \times N}$ denotes the combining matrix assumed fixed to all sub-frames and frequencies, $\boldsymbol{H}_f \in \mathbb{C}^{N \times M}$ is the MIMO channel matrix associated with the f-th frequency, $\boldsymbol{S}_k = [\boldsymbol{S}_P | \boldsymbol{S}_{D,k}] \in \mathbb{C}^{M \times L_S}$ denotes the k-th transmitted sub-frame, and $\boldsymbol{V}_{k,f} \in \mathbb{C}^{N \times L_S}$ is the cicularly symmetric complex AWGN matrix. In (173), $\boldsymbol{H}_f$ represents the f-th frontal slice of the tensor $\boldsymbol{\mathcal{H}} = \tilde{\boldsymbol{\mathcal{H}}} \times_3 \boldsymbol{F}_L \in \mathbb{C}^{N \times M \times F}$ obtained after the multiplication of the channel impulse response tensor $\tilde{\boldsymbol{\mathcal{H}}} \in \mathbb{C}^{N \times M \times L}$ with a DFT matrix $\boldsymbol{F}_L \in \mathbb{C}^{F \times L}$ along the 3-mode. The n-th and m-th row of the matrices $\boldsymbol{\Phi}^{[r]} \in \mathbb{C}^{K \times N}$ and $\boldsymbol{\Phi}^{[t]} \in \mathbb{C}^{K \times M}$ can be expressed as:

$$\boldsymbol{\Phi}^{[r]}(k,:) = \left[e^{j\theta_1^{[r]}(k)}, \cdots, e^{j\theta_N^{[r]}(k)} \right] \in \mathbb{C}^{1 \times N} \tag{174}$$

and

$$\boldsymbol{\Phi}^{[t]}(k,:) = \left[e^{j\theta_1^{[t]}(k)}, \cdots, e^{j\theta_M^{[t]}(k)} \right] \in \mathbb{C}^{1 \times M}. \tag{175}$$

These row vectors contain the unknown PN $\theta_n^{[r]}(k)$ and $\theta_m^{[t]}(k)$ at the n-th ($n = 1, \ldots, N$) receive antenna and m-th ($m = 1, \ldots, M$) transmit antenna within the k-th sub-frame, respectively. Note that the data model in (173), (174) and (175) accepts any statistical distribution for the PN process. We do not make any assumption on the statistical distribution of the PN.

6.3 Proposed PARAFAC-Based Modeling

At the receiver, the frame shown in Figure 39 is processed in two sequential ways. Firstly, a training-based processing step for joint channel and PN estimation is performed, which is followed by a processing step for data information decoding. In the first step, the receiver extracts only the pilot preamble of each sub-frame. From (173), the contribution of the pilot part $\boldsymbol{X}_{k,f}^{(P)} \in \mathbb{C}^{N \times L_P}$ associated to the k-th sub-frame at the f-th frequency is represented by

$$\boldsymbol{X}_{k,f}^{(P)} = \boldsymbol{W} D_k \left(\boldsymbol{\Phi}^{[r]} \right) \boldsymbol{H}_f D_k \left(\boldsymbol{\Phi}^{[t]} \right) \boldsymbol{S}_P + \boldsymbol{W} \boldsymbol{V}_{k,f}^{(P)}, \tag{176}$$

where $\boldsymbol{W} \boldsymbol{V}_{k,f}^{(P)} \in \mathbb{C}^{N \times L_P}$ denotes the filtered noise contribution associated to the pilot part of the k-th sub-frame. In a similar way, the contribution of the information symbols part $\boldsymbol{X}_{k,f}^{(D)} \in \mathbb{C}^{N \times L_D}$ associated to the k-th sub-frame at the f-th frequency is given by

$$\boldsymbol{X}_{k,f}^{(D)} = \boldsymbol{W} D_k \left(\boldsymbol{\Phi}^{[r]} \right) \boldsymbol{H}_f D_k \left(\boldsymbol{\Phi}^{[t]} \right) \boldsymbol{S}_{D,k} + \boldsymbol{W} \boldsymbol{V}_{k,f}^{(D)}, \tag{177}$$

where $\boldsymbol{WV}_{k,f}^{(D)} \in \mathbb{C}^{N\times L_D}$ denotes the filtered noise contribution associated to the information symbols part of the k-th sub-frame. Note that $\boldsymbol{X}_{k,f} = \left[\boldsymbol{X}_{k,f}^{(P)} | \boldsymbol{X}_{k,f}^{(D)}\right] \in \mathbb{C}^{N\times L_S}$.

According to property (7), by vectorizing the pilot part $\boldsymbol{X}_{k,f}^{(P)}$ of the k-th sub-frame (176) we obtain

$$\boldsymbol{x}_{k,f}^{(P)} = \left(\boldsymbol{S}_P^{\mathrm{T}} \otimes \boldsymbol{W}\right) \mathrm{vec}\left(D_k\left(\boldsymbol{\Phi}^{[r]}\right) \boldsymbol{H}_f D_k\left(\boldsymbol{\Phi}^{[t]}\right)\right) + \tilde{\boldsymbol{v}}_{k,f}^{(P)}, \tag{178}$$

where $\boldsymbol{x}_{k,f}^{(P)} = \mathrm{vec}\left(\boldsymbol{X}_{k,f}^{(P)}\right) \in \mathbb{C}^{NL_P}$ and $\tilde{\boldsymbol{v}}_{k,f}^{(P)} = \mathrm{vec}\left(\boldsymbol{WV}_{k,f}^{(P)}\right) \in \mathbb{C}^{NL_P}$ for simplicity of notation.

By applying again the property (7) to the second term in the right hand side of (178), we get

$$\boldsymbol{x}_{k,f}^{(P)} = \left(\boldsymbol{S}_P^{\mathrm{T}} \otimes \boldsymbol{W}\right) \left(D_k\left(\boldsymbol{\Phi}^{[t]}\right) \otimes D_k\left(\boldsymbol{\Phi}^{[r]}\right)\right) \boldsymbol{h}_f + \tilde{\boldsymbol{v}}_{k,f}^{(P)}, \tag{179}$$

where $\boldsymbol{h}_f = \mathrm{vec}\left(\boldsymbol{H}_f\right) \in \mathbb{C}^{NM}$. Now using the property $\mathrm{diag}(\boldsymbol{a})\boldsymbol{b} = \mathrm{diag}(\boldsymbol{b})\boldsymbol{a}$, we can rewrite (179) as follows

$$\boldsymbol{x}_{k,f}^{(P)} = \left(\boldsymbol{S}_P^{\mathrm{T}} \otimes \boldsymbol{W}\right) \mathrm{diag}\left(\boldsymbol{h}_f\right) \left(\boldsymbol{\Phi}^{[t]^{\mathrm{T}}}(k,:) \otimes \boldsymbol{\Phi}^{[r]^{\mathrm{T}}}(k,:)\right) + \tilde{\boldsymbol{v}}_{k,f}^{(P)}. \tag{180}$$

Collecting the vectorized received pilots $\boldsymbol{x}_{k,f}^{(P)}$ for all the $k = 1,\ldots,K$ sub-frames that form a given received frame at the f-th frequency as the columns of the resulting matrix $\boldsymbol{X}_f^{(P)} = \left[\boldsymbol{x}_{1,f}^{(P)},\ldots,\boldsymbol{x}_{K,f}^{(P)}\right] \in \mathbb{C}^{NL_P\times K}$ and according to the definition of the Khatri-Rao product in (2), which corresponds to the column-wise Kronecker product, we have

$$\boldsymbol{X}_f^{(P)} = \left(\boldsymbol{S}_P^{\mathrm{T}} \otimes \boldsymbol{W}\right) \mathrm{diag}\left(\boldsymbol{h}_f\right) \left(\boldsymbol{\Phi}^{[t]^{\mathrm{T}}} \diamond \boldsymbol{\Phi}^{[r]^{\mathrm{T}}}\right) + \tilde{\boldsymbol{V}}_f^{(P)}, \tag{181}$$

where $\tilde{\boldsymbol{V}}_f^{(P)} = \left[\tilde{\boldsymbol{v}}_{1,f}^{(P)},\ldots,\tilde{\boldsymbol{v}}_{K,f}^{(P)}\right] \in \mathbb{C}^{NL_P\times K}$.

Note that, by comparing (181) with (28) is easy to see that the noiseless term of $\boldsymbol{X}_f^{(P)}$ can be interpreted as the f-th frontal slice of the third-order tensor $\boldsymbol{\mathcal{X}}^{(P)} \in \mathbb{C}^{NL_P\times K\times F}$ which corresponds to the following PARAFAC decomposition

$$\boldsymbol{\mathcal{X}}^{(P)} = \boldsymbol{\mathcal{I}}_{3,MN} \times_1 \left(\boldsymbol{S}_P^{\mathrm{T}} \otimes \boldsymbol{W}\right) \times_2 \left(\boldsymbol{\Phi}^{[t]^{\mathrm{T}}} \diamond \boldsymbol{\Phi}^{[r]^{\mathrm{T}}}\right)^{\mathrm{T}} \times_3 \boldsymbol{H}, \tag{182}$$

where $\boldsymbol{H}(f,:) = \boldsymbol{h}_f^{\mathrm{T}} \in \mathbb{C}^{1\times MN}$ $(f = 1,\ldots,F)$. By analogy with (25), the correspondences below are valid:

$$\left(\boldsymbol{A}^{(1)}, \boldsymbol{A}^{(2)}, \boldsymbol{A}^{(3)}\right) \leftrightarrow \left(\left(\boldsymbol{S}_P^{\mathrm{T}} \otimes \boldsymbol{W}\right), \left(\boldsymbol{\Phi}^{[t]^{\mathrm{T}}} \diamond \boldsymbol{\Phi}^{[r]^{\mathrm{T}}}\right)^{\mathrm{T}}, \boldsymbol{H}\right),$$
$$\left(I_1, I_2, I_3, Q\right) \leftrightarrow \left(NL_P, K, F, MN\right).$$

According to (29), (30) and (31), the 1-mode, 2-mode and 3-mode unfolding matrices $\left[\boldsymbol{\mathcal{X}}^{(P)}\right]_{(1)} \in \mathbb{C}^{N_{L_P} \times KF}$, $\left[\boldsymbol{\mathcal{X}}^{(P)}\right]_{(2)} \in \mathbb{C}^{K \times FNL_P}$ and $\left[\boldsymbol{\mathcal{X}}^{(P)}\right]_{(3)} \in \mathbb{C}^{F \times KNL_P}$ of $\boldsymbol{\mathcal{X}}^{(P)}$ take the following forms:

$$\left[\boldsymbol{\mathcal{X}}^{(P)}\right]_{(1)} = \left(\boldsymbol{S}_P^{\mathrm{T}} \otimes \boldsymbol{W}\right)\left(\boldsymbol{H} \diamond \left(\boldsymbol{\Phi}^{[t]^{\mathrm{T}}} \diamond \boldsymbol{\Phi}^{[r]^{\mathrm{T}}}\right)^{\mathrm{T}}\right)^{\mathrm{T}}, \tag{183}$$

$$\left[\boldsymbol{\mathcal{X}}^{(P)}\right]_{(2)} = \left(\boldsymbol{\Phi}^{[t]^{\mathrm{T}}} \diamond \boldsymbol{\Phi}^{[r]^{\mathrm{T}}}\right)^{\mathrm{T}}\left(\boldsymbol{H} \diamond \left(\boldsymbol{S}_P^{\mathrm{T}} \otimes \boldsymbol{W}\right)\right)^{\mathrm{T}}, \tag{184}$$

$$\left[\boldsymbol{\mathcal{X}}^{(P)}\right]_{(3)} = \boldsymbol{H}\left(\left(\boldsymbol{\Phi}^{[t]^{\mathrm{T}}} \diamond \boldsymbol{\Phi}^{[r]^{\mathrm{T}}}\right)^{\mathrm{T}} \diamond \left(\boldsymbol{S}_P^{\mathrm{T}} \otimes \boldsymbol{W}\right)\right)^{\mathrm{T}}. \tag{185}$$

It is worth mentioning that the estimates of the frequency-selective MIMO channel and PN impairments can be obtained from the above equations by means of a two-stage receiver algorithm, as it will be explained in the sequel.

6.4 Proposed Semi-Blind Receiver for Joint Channel and PN Estimation

Our goal is to jointly estimate from the noisy version of $\boldsymbol{\mathcal{X}}^{(P)}$ the frequency-selective channel and the PN per sub-frame without *a priori* knowledge of the CSI and PN model. Note that the combining filter $\boldsymbol{W}$ is fixed (and known) at the receiver. The proposed semi-blind receiver algorithm consists of two stages. In the first stage, estimates of $\boldsymbol{H}_f$ for different frequencies are obtained using the state-of-the-art ALS-PARAFAC algorithm 2. In the second one, the individual estimates of $\boldsymbol{\Phi}^{[t]}$ and $\boldsymbol{\Phi}^{[r]}$ are obtained by applying the LS-KRF algorithm 1 in the estimated factor matrix (related to the second mode of $\boldsymbol{\mathcal{X}}^{(P)}$) previously computed in the first ALS stage. In the following, we formulate in detail each stage of the proposed tensor-based semi-blind receiver.

The first stage of the proposed algorithm, called BALS, consists of estimating the factor matrices $\boldsymbol{\Phi} = \left(\boldsymbol{\Phi}^{[t]^{\mathrm{T}}} \diamond \boldsymbol{\Phi}^{[r]^{\mathrm{T}}}\right)^{\mathrm{T}}$ and $\boldsymbol{H}$ in an alternating way from the unfolding matrices $\left[\boldsymbol{\mathcal{X}}^{(P)}\right]_{(2)}$ and $\left[\boldsymbol{\mathcal{X}}^{(P)}\right]_{(3)}$ in (184) and (185), respectively. This can be done by optimizing, respectively, the following two LS criteria:

$$\hat{\boldsymbol{\Phi}} = \underset{\boldsymbol{\Phi}}{\operatorname{argmin}} \left\| \left[\boldsymbol{\mathcal{X}}^{(P)}\right]_{(2)} - \boldsymbol{\Phi}\left(\boldsymbol{H} \diamond \left(\boldsymbol{S}_P^{\mathrm{T}} \otimes \boldsymbol{W}\right)\right)^{\mathrm{T}} \right\|_{\mathrm{F}}^2, \tag{186}$$

$$\hat{\boldsymbol{H}} = \underset{\boldsymbol{H}}{\operatorname{argmin}} \left\| \left[\boldsymbol{\mathcal{X}}^{(P)}\right]_{(3)} - \boldsymbol{H}\left(\boldsymbol{\Phi} \diamond \left(\boldsymbol{S}_P^{\mathrm{T}} \otimes \boldsymbol{W}\right)\right)^{\mathrm{T}} \right\|_{\mathrm{F}}^2. \tag{187}$$

The solutions of which are given by

$$\hat{\boldsymbol{\Phi}} = \left[\boldsymbol{\mathcal{X}}^{(P)}\right]_{(2)} \left[\left(\boldsymbol{H} \diamond \left(\boldsymbol{S}_P^{\mathrm{T}} \otimes \boldsymbol{W}\right)\right)^{\mathrm{T}}\right]^{\dagger}, \tag{188}$$

$$\hat{\boldsymbol{H}} = \left[\boldsymbol{\mathcal{X}}^{(P)}\right]_{(3)} \left[\left(\boldsymbol{\Phi} \diamond \left(\boldsymbol{S}_P^{\mathrm{T}} \otimes \boldsymbol{W}\right)\right)^{\mathrm{T}}\right]^{\dagger}, \tag{189}$$

Algorithm 12 Proposed two-stage semi-blind-receiver for joint channel and PN estimation

procedure
First Stage (BALS)
1.1 *Set* $i = 0$. *Initialize randomly the factor matrix* $\hat{\boldsymbol{H}}_{(i=0)}$;
1.2 $i \leftarrow i + 1$;
1.3 *According to* (188), *obtain an LS estimate of* $\boldsymbol{\Phi}$:

$$\hat{\boldsymbol{\Phi}}_{(i)} = \left[\boldsymbol{\mathcal{X}}^{(P)}\right]_{(2)} \left[\left(\hat{\boldsymbol{H}}_{(i-1)} \diamond \left(\boldsymbol{S}_P^{\mathrm{T}} \otimes \boldsymbol{W}\right)\right)^{\mathrm{T}}\right]^{\dagger};$$

1.4 *According to* (189), *obtain an LS estimate of* $\boldsymbol{H}$:

$$\hat{\boldsymbol{H}}_{(i)} = \left[\boldsymbol{\mathcal{X}}^{(P)}\right]_{(3)} \left[\left(\hat{\boldsymbol{\Phi}}_{(i)} \diamond \left(\boldsymbol{S}_P^{\mathrm{T}} \otimes \boldsymbol{W}\right)\right)^{\mathrm{T}}\right]^{\dagger};$$

1.5 *Repeat steps* 1.2-1.4 *until convergence.*
Second Stage (LS-KRF)
2.1 *From* $\hat{\boldsymbol{\Phi}}$, *obtained in the BALS stage, obtain the estimates of* $\boldsymbol{\Phi}^{[t]}$ *and* $\boldsymbol{\Phi}^{[r]}$ *using the LS-KRF algorithm* 1.

respectively.

Due to the knowledge of $\boldsymbol{W}$ and $\boldsymbol{S}_P$ at the receiver, each iteration of the BALS stage contains only two updating steps. At each step, the fitting error is minimized with respect to one given factor matrix by fixing the other to its value obtained at previous updating step. This procedure is repeated until the convergence of the BALS stage at the i-th iteration. The convergence is declared when $|e_{(i)} - e_{(i-1)}| \leq 10^{-6}$, where $e_{(i)}$ denotes the residual error calculated at the i-th iteration defined as

$$e_{(i)} = \left\| \boldsymbol{\mathcal{X}}^{(P)} - \hat{\boldsymbol{\mathcal{X}}}_{(i)}^{(P)} \right\|_{\mathrm{F}}^2, \tag{190}$$

where $\hat{\boldsymbol{\mathcal{X}}}_{(i)}^{(P)}$ represents the reconstructed version of $\boldsymbol{\mathcal{X}}^{(P)}$ computed from the estimated factor matrices at the end of the i-th iteration. It is important to mention that the convergence of the BALS stage to the global minimum is always achieved within a few iterations (between 8 and 28 in our simulation). The permutation ambiguity does not exist due to the knowledge of 1-mode factor matrix. Therefore, the column scaling ambiguity affecting the estimated factor matrices can be removed with a simple normalization procedure.

The second stage, called LS-KRF, consists of estimating the PN matrices $\boldsymbol{\Phi}^{[t]}$ and $\boldsymbol{\Phi}^{[r]}$ from the estimated factor matrix $\hat{\boldsymbol{\Phi}}^{\mathrm{T}}$ previously computed in the BALS stage. For the sake of convenience, let us define $\boldsymbol{\varphi}_k^{[t]} \in \mathbb{C}^M$ and $\boldsymbol{\varphi}_k^{[r]} \in \mathbb{C}^N$ as the k-th column of $\boldsymbol{\Phi}^{[t]^{\mathrm{T}}}$ and $\boldsymbol{\Phi}^{[r]^{\mathrm{T}}}$, respectively. According to definition (2), we have

$$\hat{\boldsymbol{\Phi}}^{\mathrm{T}} = \left[\boldsymbol{\varphi}_1^{[t]} \otimes \boldsymbol{\varphi}_1^{[r]}, \ldots, \boldsymbol{\varphi}_K^{[t]} \otimes \boldsymbol{\varphi}_K^{[r]}\right] \in \mathbb{C}^{MN \times K}. \tag{191}$$

According to property (9), the k-th column of (191) can be interpreted as the vectorized form of the following rank-1 matrix

$$\boldsymbol{\varphi}_k^{[t]} \otimes \boldsymbol{\varphi}_k^{[r]} = \operatorname{vec}\left(\boldsymbol{\Psi}_k\right), \tag{192}$$

where $\boldsymbol{\Psi}_k = \boldsymbol{\varphi}_k^{[r]} \circ \boldsymbol{\varphi}_k^{[t]} \in \mathbb{C}^{N \times M}$. Therefore, estimates for the vectors $\boldsymbol{\varphi}_k^{[r]}$ and $\boldsymbol{\varphi}_k^{[t]}$ can be obtained by truncating the SVD of $\boldsymbol{\Psi}_k$ defined by $\boldsymbol{U}_k \boldsymbol{\Sigma}_k \boldsymbol{V}_k^{\mathrm{H}}$ to a rank-1 approximation as follows:

$$\hat{\boldsymbol{\varphi}}_k^{[r]} = \sqrt{\sigma_1}\boldsymbol{u}_1 \quad \text{and} \quad \hat{\boldsymbol{\varphi}}_k^{[t]} = \sqrt{\sigma_1}\boldsymbol{v}_1^*, \tag{193}$$

where $\boldsymbol{u}_1 \in \mathbb{C}^N$ and $\boldsymbol{v}_1 \in \mathbb{C}^M$ are the dominant left and right singular vectors of $\boldsymbol{U}_k$ and $\boldsymbol{V}_k$, and σ_1 is the dominant singular value, respectively. Note that, the final estimates of $\hat{\boldsymbol{\Phi}}^{[r]^{\mathrm{T}}}$ and $\hat{\boldsymbol{\Phi}}^{[t]^{\mathrm{T}}}$ are obtained by solving this rank-1 approximation problem for $k = 1, \ldots, K$ according to Algorithm 1. The proposed two-stage tensor-based semi-blind receiver for joint channel and PN estimation is summarized in Algorithm 12.

6.4.1 Channel Equalization

Since the proposed semi-blind receiver algorithm provides estimates for the frequency-selective channel and PN impairments for both transmit and receive antennas, this information can be used to equalize the channel and decode the transmitted symbols. From (177), by collecting the data part of the k-th sub-frame for $f = 1, \ldots, F$ frequencies, we obtain

$$\underbrace{\begin{bmatrix} \boldsymbol{X}_{k,1}^{(D)} \\ \vdots \\ \boldsymbol{X}_{k,F}^{(D)} \end{bmatrix}}_{\boldsymbol{X}_k^{(D)}} = \underbrace{\begin{bmatrix} \boldsymbol{W} D_k\left(\boldsymbol{\Phi}^{[r]}\right) \boldsymbol{H}_1 D_k\left(\boldsymbol{\Phi}^{[t]}\right) \\ \vdots \\ \boldsymbol{W} D_k\left(\boldsymbol{\Phi}^{[r]}\right) \boldsymbol{H}_F D_k\left(\boldsymbol{\Phi}^{[t]}\right) \end{bmatrix}}_{\boldsymbol{H}_k^{(\mathrm{eff})}} \boldsymbol{S}_{D,k} + \underbrace{\begin{bmatrix} \boldsymbol{W}\boldsymbol{V}_{k,1}^{(D)} \\ \vdots \\ \boldsymbol{W}\boldsymbol{V}_{k,F}^{(D)} \end{bmatrix}}_{\tilde{\boldsymbol{V}}_k^{(D)}}$$

or, equivalently,

$$\boldsymbol{X}_k^{(D)} = \boldsymbol{H}_k^{(\mathrm{eff})} \boldsymbol{S}_{D,k} + \tilde{\boldsymbol{V}}_k^{(D)} \in \mathbb{C}^{FN \times L_D}. \tag{194}$$

Therefore, an LS estimation for the data part $\boldsymbol{S}_{D,k}$ in the k-th sub-frame can be obtained from the estimated matrices $\hat{\boldsymbol{H}}$, $\hat{\boldsymbol{\Phi}}^{[t]}$ and $\hat{\boldsymbol{\Phi}}^{[r]}$ as follows:

$$\hat{\boldsymbol{S}}_{D,k} = \left(\hat{\boldsymbol{H}}_k^{(\mathrm{eff})}\right)^{\dagger} \boldsymbol{X}_k^{(D)} \in \mathbb{C}^{M \times L_D}, \tag{195}$$

where

$$\hat{\boldsymbol{H}}_k^{(\mathrm{eff})} = \begin{bmatrix} \boldsymbol{W} D_k\left(\hat{\boldsymbol{\Phi}}^{[r]}\right) \hat{\boldsymbol{H}}_1 D_k\left(\hat{\boldsymbol{\Phi}}^{[t]}\right) \\ \vdots \\ \boldsymbol{W} D_k\left(\hat{\boldsymbol{\Phi}}^{[r]}\right) \hat{\boldsymbol{H}}_F D_k\left(\hat{\boldsymbol{\Phi}}^{[t]}\right) \end{bmatrix} \in \mathbb{C}^{FN \times M}. \tag{196}$$

Remark: Note that according to (176), an LS estimate of the effective channel matrix $\hat{\boldsymbol{H}}_{k,f}^{(\text{eff})} = \boldsymbol{X}_{k,f}^{(P)} \boldsymbol{S}_P^{\dagger}$ can be directly obtained from the pilot preamble for each sub-frame if $L_P \geq M$. Then, channel equalization can be performed using the estimated effective channel to obtain an estimate of the transmmited symbols matrix $\hat{\boldsymbol{S}}_{D,k} = \left(\hat{\boldsymbol{H}}_{k,f}^{(\text{eff})}\right)^{\dagger} \boldsymbol{X}_{k,f}^{(D)}$. However, this standard approach provides an estimate of the PN-distorted channel instead of the true frequency-selective MIMO channel. In contrast, our tensor-based semi-blind receiver provides an estimate of the true channel matrix $\hat{\boldsymbol{H}}$ which can be used, for instance, to design efficient beamforming strategies in order to improve system performance. Furthermore, our approach can reduce pilot overhead since it relaxes the identifiability condition of the standard LS estimator (which requires $L_P \geq M$), as will be analyzed in the following.

6.5 Identifiability and Computational Complexity

In the following, we analyze under which conditions the channel and PN matrices can be uniquely estimated. In addition, we also study the computational complexity of the proposed receiver algorithm.

6.5.1 Identifiability Conditions

According to (188) and (189), unique LS solutions for $\boldsymbol{\Phi}$ and $\boldsymbol{H}$ obtained from the unfolding matrices $\left[\boldsymbol{\mathcal{X}}^{(P)}\right]_{(2)}$ and $\left[\boldsymbol{\mathcal{X}}^{(P)}\right]_{(3)}$ requires that $\left(\boldsymbol{H} \diamond \left(\boldsymbol{S}_P^{\mathrm{T}} \otimes \boldsymbol{W}\right)\right)^{\mathrm{T}} \in \mathbb{C}^{MN \times FNL_P}$ and $\left(\boldsymbol{\Phi} \diamond \left(\boldsymbol{S}_P^{\mathrm{T}} \otimes \boldsymbol{W}\right)\right)^{\mathrm{T}} \in \mathbb{C}^{MN \times KNL_P}$ be full row-rank to be right-invertible. Therefore, the following conditions must be satisfied:

$$FL_P \geq M \quad \text{and} \quad KL_P \geq M. \tag{197}$$

By combining these two conditions, the lower bound on the length of the pilot preamble L_P (per sub-frame) is given by

$$L_P \geq \max\left(\left\lceil \frac{M}{F} \right\rceil, \left\lceil \frac{M}{K} \right\rceil\right), \tag{198}$$

where $\lceil x \rceil$ denotes the smallest integer number that is greater or equal to x. From (198), we can conclude that the proposed approach requires less pilot overhead when compared to the standard pilot-based LS approach, which requires $L_P \geq M$. Note that, this feature has great relevance in scenarios with a high number of transmit antennas.

6.5.2 Computational Complexity

We evaluate the computational complexity of the proposed receiver in terms of FLOPs. By observing (188), (189) and (193) we can note that the dominant cost of the proposed algorithm is associated with two matrix inverses per iteration with sizes $MN \times FNL_P$ and $MN \times KNL_P$ plus K rank-1 approximation problems of matrices with sizes $N \times M$ computed via SVD. According to (GOLUB and VAN LOAN, 1996), the matrix inverses in (188) and (189) have complexities $\mathcal{O}(N^3M^2FL_P)$ and $\mathcal{O}(N^3M^2KL_P)$, respectively. We can conclude that the BALS stage has complexity $\mathcal{O}(N^3M^2L_P(F+K))$ at each iteration, while the LS-KRF stage has complexity $\mathcal{O}(MNK)$. Therefore, we can approximate the complexity of our receiver algorithm as $\mathcal{O}((N^3M^2L_P(F+K)I) + KNM)$, where I denotes the number of iterations until the convergence of the BALS stage.

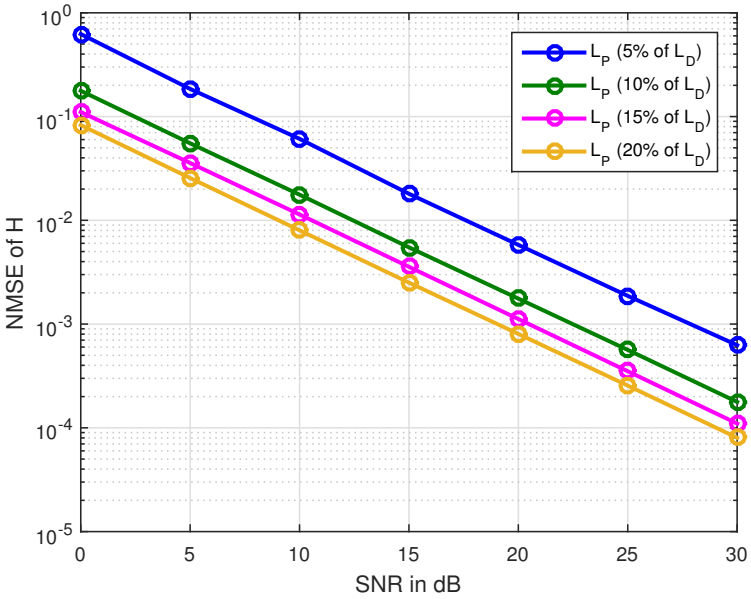
6.6 Simulation Results

In this section, we evaluate the performance of the proposed receiver algorithm in terms of the normalized mean square error (NMSE) between the estimated and true frequency-selective MIMO channel and PN matrices, the bit error rate (BER) and the number of iterations for convergence. The simulation results represent an average over 3000 independent Monte Carlo runs. Each run corresponds to an independent realization of the channel, PN matrices, pilots, data and additive noise. In all experiments we set a 4×4 frequency-selective MIMO system, $L_D = 100$ snapshots, $K = 10$ sub-frames and $F = 10$ subcarriers. The pilot and data symbols are BPSK modulated. The PN perturbations are normally distributed random variables, and independently generated for each transmit and receive antenna. The channel impulse response tensor $\tilde{\boldsymbol{\mathcal{H}}}$ is assumed to have independent and identically distributed (i.i.d.) complex Gaussian entries with zero-mean and unit variance. The matrices $\boldsymbol{W}$ and $\boldsymbol{F}_L$ are DFT matrices. The NMSE of the estimated channel and PN matrices are defined as follows:

$$\text{NMSE}\left(\hat{\boldsymbol{H}}\right) = \frac{\left\|\boldsymbol{H} - \hat{\boldsymbol{H}}\right\|_{\text{F}}^2}{\|\boldsymbol{H}\|_{\text{F}}^2}, \quad \text{NMSE}\left(\hat{\boldsymbol{\Phi}}^{[t]}\right) = \frac{\left\|\boldsymbol{\Phi}^{[t]} - \hat{\boldsymbol{\Phi}}^{[t]}\right\|_{\text{F}}^2}{\left\|\boldsymbol{\Phi}^{[t]}\right\|_{\text{F}}^2} \quad \text{and} \quad \text{NMSE}\left(\hat{\boldsymbol{\Phi}}^{[r]}\right) = \frac{\left\|\boldsymbol{\Phi}^{[r]} - \hat{\boldsymbol{\Phi}}^{[r]}\right\|_{\text{F}}^2}{\left\|\boldsymbol{\Phi}^{[r]}\right\|_{\text{F}}^2}.$$

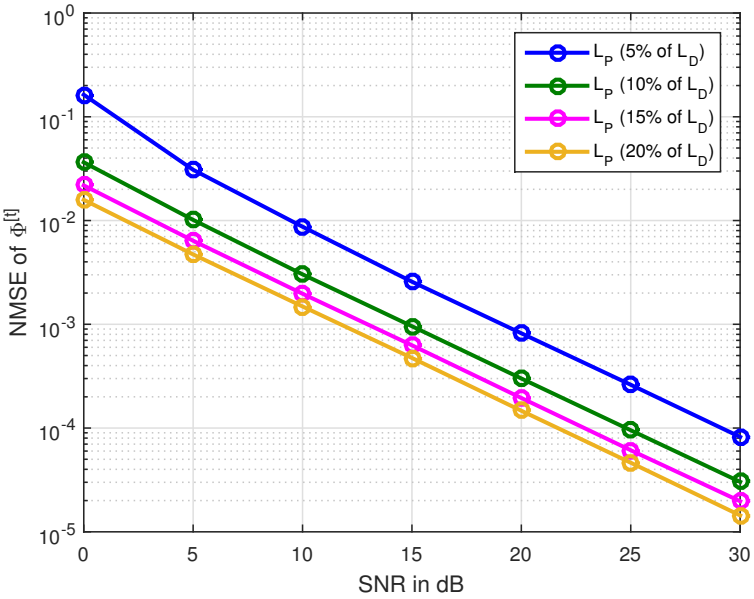
Figures 40, 41 and 42 show the estimation accuracy of the channel and PN matrices *versus* SNR when the length of the pilot preamble (L_P) is $5\%, 10\%, 15\%$ and 20% of the length of the data part (L_D). As expected, the estimation accuracy improves when the length of the pilot preamble increases. However, good performance results are achieved with fewer pilots symbols approximately 20% of the data part length for all simulated SNR range. These three experiments demonstrate the effectiveness of the proposed tensor-based semi-blind receiver in estimating jointly the frequency-selective MIMO channel and PN matrices.

Figure 40 – NMSE of $\hat{\boldsymbol{H}}$ vs. SNR (dB) of a 4×4 MIMO system for different pilot length (L_P), $L_D = 100$ and BPSK modulation.



Source: Created by the author.

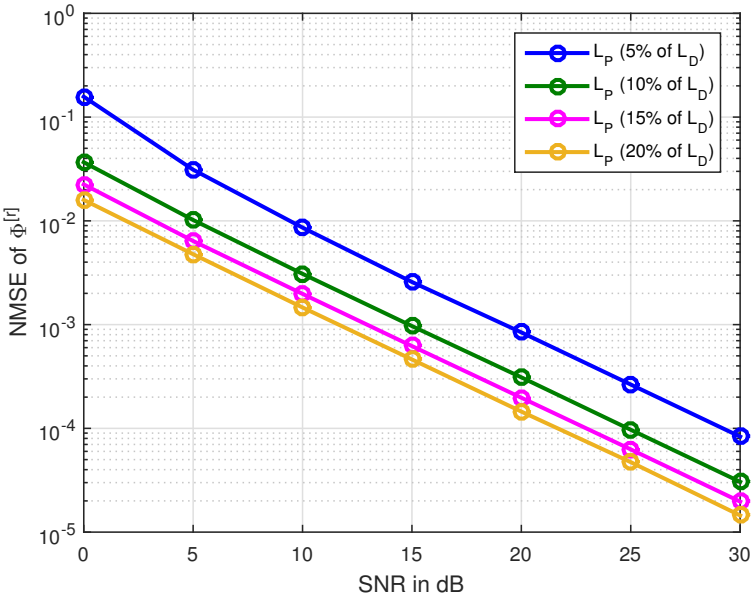
Figure 41 – NMSE of $\hat{\boldsymbol{\Phi}}^{[t]}$ vs. SNR (dB) of a 4×4 MIMO system for different pilot length (L_P), $L_D = 100$ and BPSK modulation.



Source: Created by the author.

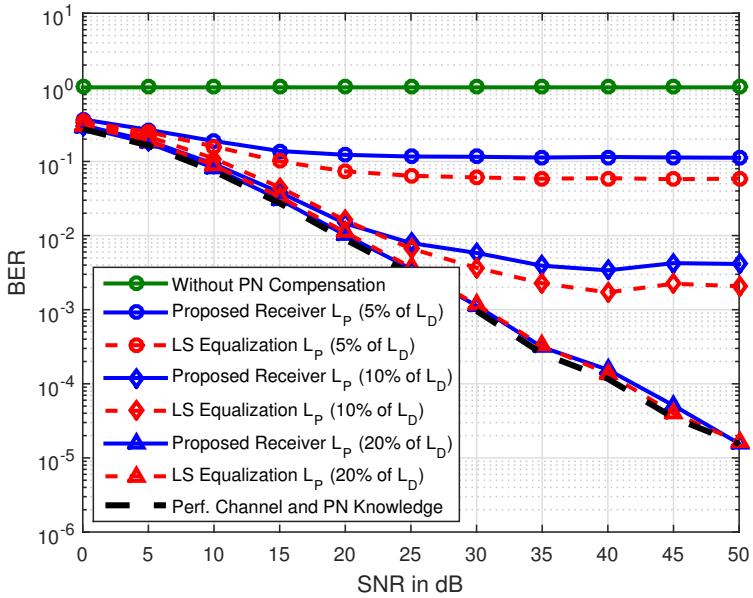
Figure 43 despicts the BER as a function of the SNR. In this experiment, the performance of the proposed receiver is compared with that of the standard pilot-aidded LS equalizer. The ideal case with perfect channel and PN knowledge is also plotted as a lower bound. As can be seen, without PN compensation at the receiver, the system performance severely deteriorates as expected. The standard LS equalizer outperforms the proposed receiver when L_P is less than 20% of the data length. However, as a disad-

Figure 42 – NMSE of $\hat{\boldsymbol{\Phi}}^{[r]}$ vs. SNR (dB) of a 4×4 MIMO system for different pilot length (L_P), $L_D = 100$ and BPSK modulation.



Source: Created by the author.

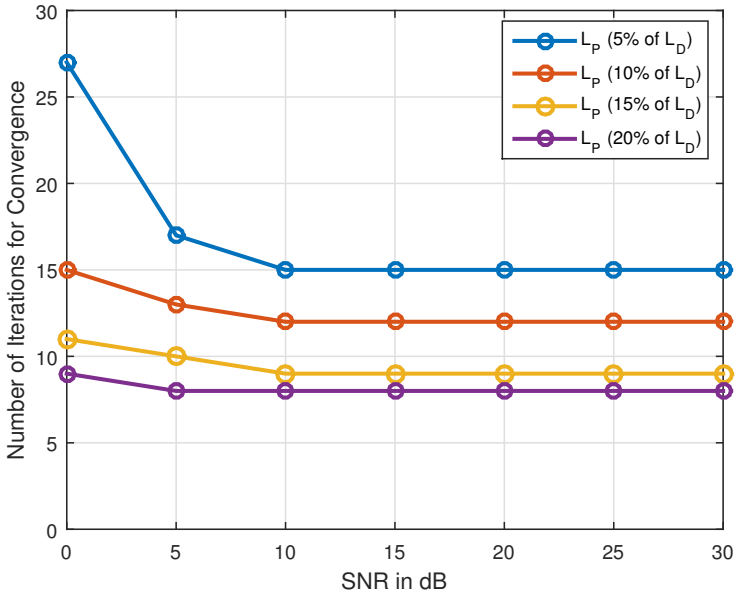
Figure 43 – BER vs. SNR (dB) of a 4×4 MIMO system with perfect and imperfect channel and PN knowledge for different pilot length (L_P), $L_D = 100$ and BPSK modulation.



Source: Created by the author.

vantage, the LS receiver provides an estimate of the effective channel matrix instead true frequency-selective MIMO channel. More importantly, the system performance is very close to the ideal case of perfect channel and PN knowledge with few pilot overhead, e.g., when L_P is equal to 20% of the data length.

Figure 44 – Number of iterations for convergence vs. SNR (dB) of a 4×4 MIMO system for different pilot length (L_P), $L_D = 100$ and BPSK modulation.



Source: Created by the author.

Figure 44 provides the mean number of iterations required for the convergence of the BALS stage. This result confirms experimentally the rapid convergence of the proposed receiver (between 8 and 28 iterations).

6.7 Chapter Summary

In this chapter, we have showed that the received signal in a frequency-selective MIMO system in which both the transmit and the receive antennas are equipped with independent oscillators can be modeled as a third-order PARAFAC model. Furthermore, we have proposed a two-stage semi-blind receiver to solve the joint channel and PN estimation problem using the state-of-the-art ALS-PARAFAC and LS-KRF algorithms. The proposed approach avoids idealized assumptions such as perfect channel and PN model knowledge, which makes it applicable to more challenging scenarios compared to other solutions in the literature. Simulation results indicate that the proposed receiver can achieve a performance very close to the ideal case of perfect channel and PN knowledge with fewer pilot symbols, rapid convergence and providing an accurate estimate of the true frequency-selective MIMO channel matrix which can be used a *posteriori*, for instance, to design efficient beamforming strategies in order to improve system performance.

7 CONCLUSIONS AND PERSPECTIVES

In this thesis, we have developed new tensor modeling approaches and signal processing algorithms for channel and signal parameters estimation in multidimensional sensor arrays, bistatic MIMO radar and MIMO wireless communications systems. The main conclusions for each chapter are summarized below:

- In **Chapter 2**, we have contributed with a review on tensor algebra and tensor decompositions concepts needed throughout this thesis. The PARAFAC, Nested-PARAFAC and Tucker decompositions have been presented in different (scalar, slices, unfoldings and n-mode product) notations. The state-of-the-art ALS algorithms to estimate the factor matrices, and the uniqueness conditions of these tensor models have also been presented.
- In **Chapter 3**, the spatial parameters estimation problem in multidimensional sensor arrays was addressed. We developed a preprocessing framework for denoising called MuDe. We found that MuDe provides more accurate estimation of the spatial parameters compared to state-of-the-art techniques, while achieving the MSE bound derived in (STEINWANDT, ROEMER, and HAARDT, 2014) for the single source scenario. We also exploited the multidimensional structure of the covariance matrix of the signal received at an R-D sensor array to formulate two generalized Tucker-based and PARAFAC-based algorithms for spatial signature estimation. The first one has a higher computational complexity, but can effectively handle unknown structures for the covariance of the source signals. On the other hand, the lower computational complexity of the second one is accompanied by the assumption of uncorrelated sources, which may not hold in more realistic scenarios. Simulation results showed that the proposed tensor methods outperform the state-of-the-art matrix-based and tensor-based techniques in terms of parameters estimation accuracy.
- In **Chapter 4**, the multiple targets localization problem in bistatic MIMO radar systems was addressed. We have capitalized on the Nested-PARAFAC decomposition to formulate new solutions for joint DoD and DoA estimation. The Nested-PARAFAC decomposition allows to decouple a complex fourth-order tensor problem into two smaller third-order tensor subproblems without affecting the accuracy of the parameter estimates. The proposed method has a lower complexity compared to state-of-the-art solutions, being robust to positioning errors between the antenna elements.
- In **Chapter 5**, the impact of the joint DL and UL channel estimation at the BS was evaluated. We have proposed a new closed-loop and multi-frequency channel training framework that allows to concentrate the processing burden for channel estimation at the BS, thus avoiding, additional processing for channel estimation

at the MSs as well as channel reciprocity assumptions. The proposed framework yields the formulation of the received closed-loop signal as a third-order PARAFAC decomposition. A new PARAFAC modeling is also proposed for MIMO wireless communications systems equipped with hybrid analog-digital beamforming architectures and operating in mmWave propagation environments. This approach capitalizes on the sparse representation of the channel from which sparse signal recovery problems are solved by the BS to estimate the channel parameters. Numerical results showed a trade-off between DL channel estimation accuracy performed at the BS and reduction of the processing cost for channel estimation at the MS side.

- In **Chapter 6**, the problem of joint channel and phase noise estimation was addressed for frequency-selective MIMO systems. The PARAFAC decomposition was exploited for modeling a more realistic MIMO communication system that assumes real-world hardware impairments that result in unknown phase noise perturbations per antenna. The proposed PARAFAC-based receiver relies on the ALS and LS-KRF algorithms to jointly estimate the channel and phase noise. Simulation results showed that the proposed method provides an efficient phase noise compensation achieving performance very close to that obtained with perfect CSI knowledge and without requiring the knowledge of the channel and statistical distribution of the phase noise process.

Several perspectives may be raised, as listed below:

- With respect to **Chapter 3**, we shall develop a full framework that takes into account a first step for model order selection (MOS) and incorporate the preprocessing multidimensional prewhitening stage in (DA COSTA *et al.*, 2013) to evaluate the MuDe performance in the spatial colored noise environments. Additionally, we also intend to extend the MuDe to operate in other sensor array geometries, e.g., those considered in (PAL and VAIDYANATHAN, 2010) and (P.PAL and VAIDYANATHAN, 2012).
- Perspectives for the **Chapter 5** include its extension to frequency-selective channels, a performance evaluation assuming different HB designs, grid mismatch and realistic channel models.
- For **Chapter 6**, we shall incorporate a more realistic model for the temporal variation for the phase noise perturbations and investigate its impact on the performance of the proposed method. A new tensor modeling that consider the same phase noise perturbations for a group of antennas is also an interesting research topic for distributed MIMO systems.

REFERENCES

ALKHATEEB, A.; EL AYACH, O.; LEUS, G.; HEATH, R. W. Channel estimation and hybrid precoding for millimeter wave cellular systems. *IEEE Journal of Selected Topics in Signal Processing*, v. 8, n. 5, p. 831–846, 2014a.

ALKHATEEB, A.; HEATH, R. W.; LEUS, G. Achievable rates of multiuser millimeter wave systems with hybrid precoding. *IEEE International Conference on Communication Workshop (ICCW)*. London, UK, 2015.

ALKHATEEB, A.; MO, J.; PRELCIC, N. G.; HEATH, R. W. MIMO precoding and combining solutions for millimeter-wave systems. *IEEE Communications Magazine*, v. 52, n. 12, p. 122–131, 2014b.

ARAÚJO, D. C. *Channel estimation techniques applied to massive MIMO systems using sparsity and statistical approaches*. Doctoral Thesis, 2016.

ARAÚJO, D. C.; DE ALMEIDA, A. L. F. Tensor-based compressed estimation of frequency-selective mmwave MIMO channels. *IEEE International Workshop on Computational Advances in Multi-sensor Adaptive Processing (CAMSAP)*. Curacao, Netherlands Antilles, 2017.

BECK, A.; TEBOULLE, M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, v. 2, n. 1, p. 183–202, 2009.

BELTRAMI, E. Sulle funzioni bilineari. *Giornale di matematiche ad uso degli studenti delle universita*, v. 11, p. 98–106, 1873.

BREWER, J. W. Kronecker products and matrix calculus in system theory. *IEEE Transactions on Circuits and Systems*, v. CAS-25, n. 9, p. 772–781, 1978.

BRO, R. *Multi-way analysis in the food industry: models, algorithms and applications*. Doctoral Thesis, 1998.

CARROLL, J. D.; CHANG, J. Analysis of individual differences in multidimensional scaling via an N-way generalization of Eckart-Young decomposition. *Psychometrika*, v. 35, n. 3, p. 283–319, 1970.

CHANGUEL, H.; HARABI, F.; GHARSALLAH, A. 2-L-shape two-dimensional arrival angle estimation with a classical subspace algorithm. *IEEE International Symposium on Industrial Eletronics (ISIE)*. Montreal, Canada, 2006.

CHEN, G.; QIAN, S. E. Denoising of hyperspectral imagery using principal component analysis and wavelet shrinkage. *IEEE Transactions on Geoscience and Remote Sensoring*, v. 49, n. 3, p. 973–980, 2011.

CHEN, J.; GU, H.; SU, W. Angle estimation using ESPRIT without pairing in MIMO radar. *Eletronics Letters*, v. 44, n. 24, p. 1422–1423, 2008.

CHEN, J.; GU, H.; SU, W. A new method for joint DOD and DOA estimation in bistatic MIMO radar. *Signal Processing*, v. 90, p. 714–718, 2010.

CHO, Y. S.; KIM, J.; WANG, W. Y.; KANG, C. G. *MIMO-OFDM wireless communications with MATLAB*. Wiley Publishing, 2010.

CHOI, J.; LOVE, D. J.; BIDIGARE, P. Downlink training techniques for FDD massive MIMO systems: open-loop and closed-loop training with memory. *IEEE Journal of Selected Topics in Signal Processing*, v. 8, n. 5, p. 802–814, 2014.

CICHOCKI, A.; PHAN, A. H.; CAIAFA, C. F.; ZHOU, G.; ZHAO, Q.; DE LATHAUWER, L. Tensor decompositions for signal processing applications: from two-way to multiway component analysis. *IEEE Signal Processing Magazine*, v. 32, n. 2, p. 145–163, 2015.

CICHOCKI, A.; ZDUNEK, R.; PHAN, A. H.; AMARI, S. *Nonnegative matrix and tensor factorizations*. Wiley, 2009.

COMON, P. Tensors: a brief introduction. *IEEE Signal Processing Magazine*, v. 31, n. 3, p. 44–53, 2014.

DA COSTA, J. P. C. L. *Parameter estimation techniques for multidimensional array signal processing*. Shaker Publisher, 2010.

DA COSTA, J. P. C. L.; LIU, K.; CHEUNG, H.; SCHWARS, S.; HAARDT, M.; ROEMER, F. Multidimensional prewhitening for enhanced signal reconstruction and parameter estimation in colored noise with Kronecker correlation structure. *Signal Processing*, v. 93, n. 11, p. 3209–3226, 2013.

DA COSTA, J. P. C. L.; ROEMER, F.; HAARDT, M.; DE SOUSA JR., R. T. Multi-dimensional model order selection. *EURASIP Journal on Advances in Signal Processing*, v. 26, n. 20, 2011.

DA COSTA, J. P. C. L.; ROEMER, F.; WEIS, M.; HAARDT, M. Robust R-D parameter estimation via closed-form PARAFAC. *Proc. ITG Workshop on Smart Antennas (WSA10)*, 2010.

DE ALMEIDA, A. L. F. *Tensor modeling and signal processing for wireless communication systems*. Doctoral Thesis, 2007.

DE ALMEIDA, A. L. F.; FAVIER, G.; ; XIMENES, L. R. Space-time-frequency (STF) MIMO communication systems with blind receiver based on a generalized PARATUCK2 model. *IEEE Transactions on Signal Processing*, v. 61, n. 8, p. 337–351, 2013.

DE ALMEIDA, A. L. F.; FAVIER, G. Double Khatri-Rao space-time-frequency coding using semi-blind PARAFAC based receiver. *IEEE Signal Processing Letters*, v. 20, n. 5, p. 471–474, 2013.

DE ALMEIDA, A. L. F.; FAVIER, G.; MOTA, J. C. M. Multipath parameter estimation of time-varying space-time communication channles using parallel factor analysis. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Tolouse, Fance, 2006.

DE ALMEIDA, A. L. F.; FAVIER, G.; MOTA, J. C. M. PARAFAC-based unified tensor modeling for wireless communication systems with application to blind multiuser equalization. *Signal Processing*, v. 87, p. 337–351, 2007.

DE LATHAUWER, L.; DE MOOR, B.; VANDEWALLE, J. A multilinear singular value decomposition. *SIAM Journal on Matrix Analysis and Applications*, v. 21, p. 1253–1278, 2000.

FERNANDES, C. A. R.; FAVIER, G.; MOTA, J. C. M. PARAFAC-based channel estimation and data recovery in nonlinear MIMO spread spectrum communication systems. *Signal Processing*, v. 91, n. 2, p. 311–322, 2011.

FOSCHINI, G. J.; GANS, M. J. On limits of wireless communications when using multiple antennas. *Wireless Personal Communications*, v. 6, n. 3, p. 311–335, 1998.

GALLACHER, R. A. S.; RAHMAN, M. S. Multi-user MIMO strategies for a millimeter wave communication system using hybrid beamforming. *IEEE International Conference on Communications (ICC)*. London, UK, 2015.

GESBERT, D.; KOUNTOURIS, M.; HEATH, R. W.; CHAE, C.; SALZER, T. From single user to multiuser communications: shifting the MIMO paradigm. *IEEE Signal Processing Magazine*, v. 24, n. 5, p. 36–46, 2007.

GOLUB, G. H.; VAN LOAN, C. F. *Matrix computations*. Johns Hopkins University Press., 1996.

GOMES, P. R. B.; DE ALMEIDA, A. L. F.; DA COSTA, J. P. C. L.; DEL GALDO, G. Tensor-based methods for blind spatial signature estimation under arbitrary and unknown source covariance structure. *Digital Signal Processing*, v. 62, p. 197–210, 2017.

GONG, X.; XU, Y.; LIU, Z. On the equivalence of tensor-MUSIC and matrix-MUSIC. *International Symposium on Antennas, Propagation and EM Theory (ISAPE)*. Guilin, China, 2006.

GU, C.; HE, J.; ZHU, X. Target localization using MIMO eletromagnetic vector array systems. *Signal Processing*, v. 93, p. 2103–2107, 2013.

GUEY, J.; LARSSON, L. D. Modeling and evaluation of MIMO systems exploiting channel reciprocity in TDD systems. *IEEE Vehicular Technology Conference (VTC)*. 2004.

GUO, X.; MIRON, S.; BRIE, D. Identificability of the PARAFAC model for polarized source mixture on a vector sensor array. *IEEE International Conference on Acoustic, Speech and Signal PRocessing (ICASSP)*. Las Vegas, USA, 2008.

HAARDT, M.; ROEMER, F.; DEL GALDO, G. Higher-order SVD-based subspace estimation to improve the parameter estimation accuracy in multidimensional harmonic retrieval problems. *IEEE Transactions on Signal Processing*, v. 56, n. 7, p. 3198–3213, 2008.

HADASCHIK, N.; DORPINGHAUS, M.; SENST, A.; HARMJANZ, O.; KAUFER, U.; ASCHEID, G.; MEYER, H. Improving MIMO phase noise estimation by exploiting spatial correlations. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. Philadelphia, USA, 2005, p. 833–836.

HAIMOVICH, A. M.; BLUM, R.; CIMINI, L. MIMO radar with widely separated antennas. *IEEE Signal Processing Magazine*, v. 25, n. 1, p. 116–129, 2008.

HAMMOUD, B.; ANTREICH, F.; NOSSEK, J. A.; DA COSTA, J. P. C. L.; DE ALMEIDA, A. L. F. Tensor-based approach for time-delay estimation. *International ITG Workshop on Smart Antennas (WSA)*. Munich, Germany, 2016.

HARSHMAN, R. A. Foundations of the PARAFAC procedure: model and conditions for an explanatory multi-mode factor analysis. *UCLA Working Papers in Phonetics*, v. 16, p. 1–84, 1970.

HARSHMAN, R. A.; LUNDY, M. E. Uniqueness proof for a family of models sharing features of Tucker's three-mode factor analysis and PARAFAC/ CANDECOMP. *Psychometrika*, v. 61, p. 133–154, 1996.

HUANG, H.; WANG, W. G. J.; HE, J. Phase noise and frequency offset compensation in high frequency MIMO-OFDM systems. *IEEE International Conference on Communications (ICC)*. London, UK, 2015, p. 1280–1285.

ISHAQUE, A.; ASCHEID, G. On the efficient mitigation of phase noise in MIMO-OFDM receivers. *International Symposium on Signals, Systems, and Eletronics*. Potsdam, Germany, 2012, p. 1–6.

JIANG, H.; ZHANG, Y.; LI, J.; CUI, H. PARAFAC-based algorithm for multidimensional parameter estimation in polarimetric bistatic MIMO radar. *EURASIP Journal on Advances in Signal Processing*, p. 1–14, 2013.

JIN, M.; LIAO, G.; LI, J. Joint DOD and DOA estimation for bistatic MIMO radar.

Signal Processing, v. 89, p. 244–251, 2009.

JORDAN, C. Mémoire sur les bilinéaires. *J. Math Pures Appl.*, v. 19, p. 35–54, 1874a.

JORDAN, C. Sur la réduction des formes bilinéaires. *Comptes Rendus de l'Academic Sciences*, v. 78, p. 614–617, 1874b.

KAY, S. M. *Fundamentals of statistical signal processing: estimation and theory.* Prentice-Hall, 2000.

KHATRI, C. G.; R. RAO, C. Solutions to some functional equations and their applications to characterization of probability distributions. *Sankhya*, v. 30, p. 167–180, 1968.

KIERS, H. A. L. Towards a standardized notation and terminology in multiway analysis. *Journal of Chemometrics*, v. 14, p. 105–122, 2000.

KIM, K. Y.; ZHOU, Q.; CHOI, H. J.; SAYED, A. H. An efficient carrier phase synchronization technique for high-order M-QAMCOFDM. *IEEE Transactions on Signal Processing*, v. 56, n. 8, p. 3789–3794, 2008.

KINBANGOU, A. Y.; FAVIER, G. Non-iterative solution for PARAFAC with a Toeplitz matrix factor. *European Signal Processing Conference (EUSIPCO)*. Glasglow, Scotland, 2009.

KOLDA, T. G.; BADER, B. W. Tensor decompositions and applications. *SIAM Review*, v. 51, n. 3, p. 455–500, 2009.

KRUSKAL, J. B. Three-way arrays: rank and uniqueness or trilinear decompositions, with applications to arithmetic complexity and statistics. *Linear Algebra and Its Applications*, v. 18, p. 95–138, 1977.

LI, J.; STOICA, P. MIMO radar with colocated antennas. *IEEE Signal Processing Magazine*, v. 24, n. 5, p. 106–114, 2007.

LI, J.; ZHANG, X.; CAO, R.; ZHOU, M. Reduced-dimension MUSIC for angle and array gain-phase error estimation in bistatic MIMO radar. *IEEE Communications Letters*, v. 17, n. 3, p. 443–446, 2013.

LIANG, L.; XU, W.; DONG, X. Low-complexity hybrid precoding in massive multiuser MIMO systems. *IEEE Wireless Communications Letters*, v. 3, n. 6, p. 653–656, 2014.

LINEBARGER, D. A.; DE GROAT, R. D.; DOWLING, E. M. Efficient direction finding methods employing forward/backward averaging. *IEEE Transactions on Signal Processing*, v. 42, n. 8, p. 2136–2145, 1994.

LIU, K.; DA COSTA, J. P. C. L.; SO, C.; HUANG, L.; YE, J. Detection of number of components in CANDECOMP/PARAFAC models via minimum description length. *Digital Signal Processing*, v. 51, p. 110–123, 2016.

LIU, K.; DA COSTA, J. P. C. L.; SO, H. C.; DE ALMEIDA, A. L. F. Semi-blind receivers for joint symbol and channel estimation in space-time-frequency MIMO-OFDM Systems. *IEEE Transactions on Signal Processing*, v. 61, n. 21, p. 5444–5457, 2013.

LIU, X.; SIDIROPOULOS, N. *High-resolution and robust signal processing.* Marcel Dekker, Chapter 3, 2004.

LIU, X. L.; LIAO, G. S. Multi-target localisation in bistatic MIMO radar. *Eletronics Letters*, v. 46, n. 13, p. 945–946, 2010.

LOVE, D. J.; HEATH, R. W.; LAU, V. K. N.; GESBERT, D.; RAO, B. D.; ANDREWS, M. An overview of limited feedback in wireless communication systems. *IEEE Journal of Selected Topics in Signal Processing*, v. 26, n. 8, p. 1341–1365, 2008.

LU, L.; LI, G. Y.; SWINDLEHURST, A. L.; ASHIKHMIN, A.; ZHANG, R. An overview of massive MIMO: benefits and challenges. *IEEE Journal of Selected Topics in Signal Processing*, v. 8, n. 5, p. 742–758, 2014.

MARZETA, T. L. Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Transactions on Wireless Communication*, v. 9, n. 11, p. 3590–3600, 2010.

MEHRPOUYAN, H.; NASIR, A. A.; ERIKSSON, T.; BLOSTEIN, S. D.; KARAGIANNIDIS, G. K.; SVENSSON, T. Time-varying phase noise and channel estimation in MIMO systems. *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC).* Cesme, Turkey, 2012.

MUTI, D.; BOURENNANE, S. Multidimensional filtering based on tensor approach. *Signal Processing*, v. 85, n. 12, p. 2338–2353, 2005.

NGEBANI, I. M.; CHUMA, J. M.; ZIBANI, I.; MATLOTSE, E.; TSAMASSE, K. Joint channel and phase noise estimation in MIMO-OFDM systems. *IEEE Radio and Antenna Days of the Indian Ocean*, 2017.

NION, D.; SIDIROPOULOS, N. D. Tensor algebra and multidimensional harmonic retrieval in signal processing for MIMO radar. *IEEE Transactions on Signal Processing*, v. 58, n. 11, p. 5693–5705, 2010.

PAL, P.; VAIDYANATHAN, P. P. Nested arrays: a novel approach to array processing with enhanced degrees of freedom. *IEEE Transactions on Signal Processing*, v. 58, n. 8, p. 4167–4181, 2010.

PATI, Y. C.; REZAIIFAR, R.; KRISHNAPRASAD, P. S. Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. *Asilomar Conference on Signals, Systems, and Computers*. Pacific Grove, CA, 1993.

PETERSEN, K. B.; PEDERSEN, M. S. *The matrix cookbook*. http://matrixcookbook.com, 2008.

PETROVIC, D.; RAVE, W.; FETTWEIS, G. Effects of phase noise on OFDM systems with and without PLL: characterization and compensation. *IEEE Transactions on Communications*, v. 55, n. 8, p. 1607–1616, 2007.

P.PAL; VAIDYANATHAN, P. P. Nested arrays in two dimensions, part I: geometrical considerations. *IEEE Transactions on Signal Processing*, v. 60, n. 9, p. 4694–4705, 2012.

RAJIH, M. *Blind identification of underdetermined mixtures based on the characteristic function*. Doctoral Thesis, 2006.

RIAL, R. M.; RUSU, C.; PRELCIC, N. G.; ALKHATEEB, A.; HEATH, R. W. Hybrid MIMO architecture for millimeter wave communications: phased shifters or switches? *IEEE Access*, v. 4, n. 4, p. 247–267, 2016.

RIBEIRO, L. N. *On supervised multilinear filtering: applications to system identification and antenna beamforming*. Master of Science Thesis, 2016.

ROEMER, F. *Advanced algebraic concepts for efficient multi-channel signal processing*. Doctoral Thesis, 2013.

ROEMER, F.; HAARDT, M. Tensor-based channel estimation and iterative refinements for two-way relaying with multiple antennas and spatial reuse. *IEEE Transactions on Signal Processing*, v. 58, n. 11, p. 5720–5735, 2010.

ROEMER, F.; HAARDT, M. A framework for the analytical performance assessment of matrix and tensor-based ESPRIT-type algorithms. *CoRR*, v. 1209.3253, p. 1–24, 2012.

RONG, Y.; VOROBYOV, S. A.; GERSHMAN, A. B.; SIDIROPOULOS, N. D. Blind spatial signature estimation via time-varying user power loading and parallel factor analysis. *IEEE Transactions on Signal Processing*, v. 53, n. 8, p. 1697–1710, 2005.

ROY, R.; KAILATH, T. ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech and Signal Processing*, v. 37, n. 7, p. 984–995, 1989.

SCHMIDT, R. O. Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, v. 34, n. 3, p. 276–280, 1986.

SCHONEMANN, P. H. A generalized solution of the orthogonal Procrustes problem. *Psychometrika*, v. 31, n. 1, p. 1–10, 1966.

SEKIZAWA, S. Estimation of arrival directions using MUSIC algorithm with planar arrays. *IEEE International Conference on Universal Personal Communications (ICUPC)*. Florence, Italy, 1998.

SHAN, T.J.; WAX, M.; KAILATH, T. On spatial smoothing for direction-of-arrival estimation of coherent signals. *IEEE Transactions on Acoustics, Speech and Signal Processing*, v. 33, p. 806–811, 1985.

SHEN, W.; DAI, L.; SHI, Y.; SHIM, B.; WANG, Z. Joint channel training and feedback for FDD massive MIMO systems. *IEEE on Vehicular Technology*, v. 65, n. 10, p. 8762–8767, 2016.

SHEPARD, C.; YU, H.; AN, N.; LI, L. E.; MARZETTA, T.; YANG, R.; ZHONG, L. Argos: practical many-antenna base stations. *IEEE on Vehicular Technology*, ????

SIDIROPOULOS, N. D.; BRO, R. On the uniqueness of multilinear decomposition of N-way arrays. *Journal of Chemometrics*, v. 14, p. 229–239, 2000.

SIDIROPOULOS, N. D.; BRO, R.; GIANNAKIS, G. B. Parallel factor analysis in sensor array processing. *IEEE Transactions on Signal Processing*, v. 48, n. 8, p. 2377–2388, 2000.

SIDIROPOULOS, N. D.; GIANNAKIS, G. B.; BRO, R. Blind PARAFAC receivers for DS-CDMA systems. *IEEE Transactions on Signal Processing*, v. 48, n. 3, p. 810–823, 2000.

SINGH, P. R.; WANG, Y.; CRAGÉ, P. Bistatic MIMO radar for near field source localisation using PARAFAC. *Eletronics Letters*, v. 52, n. 12, p. 1060–1061, 2016.

SMILDE, A.; BRO, R.; GELADI, P. *Multi-way analysis with applications in the chemical sciences*. Wiley, 2004.

STEINWANDT, J.; ROEMER, F.; HAARDT, M. Asymptotic performance analysis of ESPRIT-type algorithms for circular and strictly non-circular sources with spatial smoothing. *IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)*. Florence, Italy, 2014.

STEINWANDT, J.; ROEMER, F.; HAARDT, M.; DEL GALDO, G. Performance analysis of multi-dimensional ESPRIT-type algorithms for arbitrary and strictly non-circular sources with spatial smoothing. *IEEE Transactions on Signal Processing*, v. 65, n. 9, p. 2262–2276, 2017.

TAN, J.; NIE, Z.; WEN, D. Low complexity MUSIC-based direction-of-arrival algorithm

for monostatic MIMO radar. *Eletronics Letters*, v. 53, n. 4, p. 275–277, 2017.

TAROKH, V.; SESHADRI, N.; CALDERBANK, A. R. Space-time codes for high data rates wireless communications: performance criterion and code construction. *IEEE Transactions on Information Theory*, v. 44, n. 2, p. 744–765, 1998.

TAYEM, N.; KWON, H. M. L-shaped 2-dimensional arrival angle estimation with propagator method. *IEEE Transactions on Antennas and Propagation*, v. 53, n. 5, p. 1622–1630, 2005.

TEN BERGE, J. M. F.; SMILDE, A. K. Non-triviality and identification of a constrained Tucker3 analysis. *Journal of Chemometrics*, v. 16, p. 609–612, 2002.

THAKRE, A.; HAARDT, M.; ROEMER, F.; GIRIDHAR, K. Tensor-based spatial smoothing (TB-SS) using multiple snapshots. *IEEE Transactions on Signal Processing*, v. 58, n. 5, p. 2715–2728, 2010.

TUCKER, L. R. Some mathematical notes on three-mode factor analysis. *Psychometrika*, v. 31, p. 279–311, 1966.

VENUGOPAL, K.; ALKHATEEB, A.; HEATH, R. W.; PRELCIC, N. G. Time-domain channel estimation for wideband millimeter wave systems with hybrid architecture. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. New Orleans, USA, 2017.

WANG, X.; WANG, W.; LI, X.; WANG, J. A tensor-based subspace approach for bistatic MIMO radar in spatial colored noise. *Sensors*, v. 14, p. 3897–3907, 2014.

WANG, Z.; CAI, C.; WEN, F.; HUANG, D. A quadrilinear decomposition method for direction estimation in bistatic MIMO radar. *IEEE Access*, v. 6, p. 13766–13772, 2018.

WEIS, M. *Multi-dimensional signal decomposition techniques for the analysis of EEG data*. Doctoral Thesis, 2015.

WEIS, M.; DEL GALDO, G.; HAARDT, M. A correlation tensor-based model for time variant frequency selective MIMO channels. *International ITG/IEEE Workshop on Smart Antennas (WSA)*. Viena, Austria, 2007.

WEIS, M.; ROEMER, F.; HAARDT, M.; HUSAR, P. Dual-symmetric parallel factor analysis using procrustes estimation and Khatri-Rao factorization. *European Signal Processing Conference (EUSIPCO)*. Bucharest, Romania, 2012.

XIE, R.; LIU, Z.; WU, J. Direction finding with automatic pairing for bistatic MIMO radar. *Signal Processing*, v. 92, p. 198–203, 2011.

XIMENES, L R. *Tensor-based MIMO relaying communication systems*. Doctoral Thesis, 2015.

XIMENES, L. R.; FAVIER, G.; DE ALMEIDA, A. L. F. Semi-blind receivers for non-regenerative cooperative MIMO communications based on nested PARAFAC modeling. *IEEE Transactions on Signal Processing*, v. 63, n. 18, p. 4985–4998, 2015.

ZHANG, X.; GAO, X.; CHEN, W. Improved blind 2D-direction of arrival estimation with L-shaped array using shift invariance property. *Journal of Eletromagnetics Waves and Applications*, v. 23, n. 5-6, p. 593–606, 2009.

ZHANG, X.; XU, Z.; JIA, N.; YANG, W.; FENG, Q.; CHEN, W.; FENG, Y. Denoising of 3D magnetic resonance images by using higher-order singular value decomposition. *Medical Image Analysis*, v. 19, n. 1, p. 75–86, 2015.

ZHENG, Z.; ZHANG, J.; ZHANG, J. Joint DOD and DOA estimation for bistatic MIMO radar in the presence of unknown mutual coupling. *Signal Processing*, v. 92, p. 3039–3048, 2012.

ZHOU, Z.; FANG, J.; YANG, L.; LI, H.; CHEN, Z.; BLUM, R. S. Low-rank tensor decomposition-aided channel estimation for millimeter wave MIMO-OFDM systems. *IEEE Journal of Selected Areas in Communications*, v. 35, n. 7, p. 1524–1538, 2017.

ZHOU, Z.; FANG, J.; YANG, L.; LI, H.; CHEN, Z.; LI, S. Channel estimation for millimeter wave multiuser MIMO systems via PARAFAC decomposition. *IEEE Transactions on Wireless Communications*, v. 15, n. 11, p. 7501–7516, 2016a.

ZHOU, Z.; FANG, J.; YANG, L.; LI, H.; CHEN, Z.; LI, S. Channel estimation for millimeter wave multiuser MIMO systems via PARAFAC decomposition. *IEEE Transactions on Wireless Communications*, v. 15, n. 11, p. 7501–7516, 2016b.