Nice-Sophia Antipolis University
Doctoral School of Information and Communication Sciences
Information and Communication Sciences & Technologies

and

University of Campinas
School of Electrical and Computer Engineering
Telecommunications and Telematics

Michele NAZARETH DA COSTA

TENSOR SPACE-TIME CODING FOR MIMO WIRELESS COMMUNICATION SYSTEMS

Doctorate thesis in co-tutorship presented to the School of Electrical and Computer Engineering in partial fulfillment of the requirements for the degree of Doctor in Sciences, mention: Automatic control, Signal and Image Processing from UNS and in Electrical Engineering, concentration area: Wireless Communication and Telematics from UNICAMP.

Advisors: **Gérard FAVIER** (UNS) and **João Marcos T. ROMANO** (UNICAMP)
Co-supervisor: **André L. F. DE ALMEIDA** (UFC)

Campinas, SP - Brazil
2014

Abstract

Since the growing success of mobile systems in the 1990s, new wireless technologies have been developed in order to support a growing demand for high-quality multimedia services while still being flexible to accommodate new services with low error rates. An interesting way to improve the error performance and to achieve better transmission rates is to combine the use of various diversities and multiplexing access techniques in the MIMO system context. The incorporation of *oversampling*, *spreading* and *multiplexing* operations and additional diversities on wireless systems lead to multidimensional received signals which naturally satisfy tensor models. This thesis proposes a new tensorial approach based on a tensor space-time (TST) coding for MIMO wireless communication systems. The signals received by multiple antennas form a fourth-order tensor that satisfies a new tensor model, referred to as PARATUCK-(2,4) model. A performance analysis is carried out for the proposed TST system and a recent space-time-frequency (STF) system, which allows to derive expressions for the maximum diversity gain over a flat fading channel. An uplink processing based on the TST coding with allocation resources is proposed. A new tensor decomposition is introduced, the so-called PARATUCK-(N_1, N), which generalizes the standard PARATUCK-2 and our PARATUCK-(2,4) model. This thesis establishes uniqueness conditions for the PARATUCK-(N_1, N) model. From these results, joint symbol and channel estimation is ensured for the TST and STF systems. Semi-blind receivers are proposed based on the well-known Alternating Least Squares (ALS) algorithm and the Levenberg-Marquardt (LM) method. A semi-blind receiver based on the Kronecker Least Squares (KLS) is also proposed for both systems. Simulation results are presented to illustrate the efficiency of the proposed receivers in terms of symbol recovery and convergence speed when compared to other methods from the literature.

Keywords: MIMO wireless communication systems, symbol estimation, space-time code, PARATUCK model, CANDECOMP/PARAFAC model, tensor modeling, CDMA, OFDM.

Resumo

Desde o crescente sucesso de sistemas móveis na década de 90, novas tecnologias sem fio têm sido desenvolvidas a fim de suportar a crescente demanda de serviços de multimídia de alta qualidade e ainda flexível para implantar novos serviços com baixas taxas de erro. Uma forma interessante de melhorar o desempenho de erro e de obter melhores taxas de transmissão consiste em combinar o emprego de várias diversidades com técnicas de múltiplo acesso no contexto de sistemas MIMO. A incorporação de operações de *sobreamostragem*, *espalhamento* e *multiplexação*, e diversidades adicionais em sistemas sem fio levam a sinais recebidos multidimensionais que, naturalmente, satisfazem modelos tensoriais. Esta tese propõe uma nova abordagem tensorial baseada em uma codificação tensorial espaço-temporal (TST) para sistemas de comunicação sem fio MIMO. Os sinais recebidos por múltiplas antenas formam um tensor de quarta ordem que satisfaz um novo modelo tensorial, referido como PARATUCK-(2,4). A análise de desempenho é realizada para o sistema proposto TST e um recente sistema espaço-tempo-frequencial (STF), a qual permite derivar expressões para o ganho máximo de diversidade através de um canal com desvanecimento plano. Propõe-se um sistema de transmissão baseado em codificação TST com recursos de alocação de antenas para sistemas MIMO com múltiplos usuários. Uma nova decomposição tensorial é introduzida, denominada PARATUCK-(N_1, N), e esta generaliza o modelo padrão PARATUCK-2 e nosso modelo PARATUCK-(2,4). A presente tese estabelece as condições de unicidade para o modelo PARATUCK-(N_1, N). A partir desses resultados, a estimativa conjunta do símbolo e canal é assegurada para os sistemas TST e STF. Os receptores semi-cegos propostos para os dois sistemas baseiam-se no algoritmo do tipo mínimos quadrados alternados ("Alternanting Least Squares", ALS) e no método de otimização Levenberg-Marquardt (LM). Um receptor baseado na estrutura do produto de Kronecker, denominado "Kronecker Least Squares" (KLS), também é proposto para ambos os sistemas. Resultados de simulações são apresentados para ilustrar a eficiência dos receptores propostos em termos de recuperação de símbolo e a velocidade de convergência quando comparados com outros métodos da literatura.

Palavras-chave: sistemas de comunicação sem fio MIMO, estimação de símbolo, codificação espaço-temporal, modelo PARATUCK, modelo CANDECOMP/PARAFAC, modelagem tensorial, CDMA, OFDM.

Résumé

Depuis le succès croissant des systèmes mobiles au cours des années 1990, les nouvelles technologies sans fil ont été développées afin de répondre à la demande croissante de services multimédias de haute qualité et une plus grande flexibilité pour déployer de nouveaux services avec des taux d'erreur les plus faibles possibles. Un moyen intéressant pour améliorer les performances et obtenir de meilleurs taux de transmission consiste à combiner l'utilisation de plusieurs diversités avec un accès de multiplexage dans le cadre des systèmes MIMO. L'utilisation de techniques de *sur-échantillonnage*, d'*étalement* et de *multiplexage*, et de diversités supplémantaires conduit à des signaux multidimensionnels, au niveau de la réception, qui satisfont des modèles tensoriels. Cette thèse propose une nouvelle approche tensorielle basée sur un codage spatio-temporal tensoriel (TST) pour les systèmes de communication sans fil MIMO. Les signaux reçus par plusieurs antennes forment un tenseur d'ordre quatre qui satisfait un nouveau modèle tensoriel, dénommé modèle PARATUCK-(2,4). Une analyse de performance est réalisée pour le système TST ainsi que pour un système spatio-temporel-fréquentiel (STF) récemment proposé dans la littérature, avec l'obtention du gain maximum de diversité dans le cas d'un canal à évanouissement plat. Un système de transmission basé sur le codage TST est proposé pour les systèmes MIMO avec plusieurs utilisateurs. Une nouvelle décomposition tensorielle est introduite, appelée PARATUCK-(N_1, N), qui généralise le modèle standard PARATUCK-2 et notre modèle PARATUCK-(2,4). Cette thèse établit les conditions d'unicité du modèle PARATUCK-(N_1, N). À partir de ces résultats, différents récepteurs semi-aveugles sont proposés pour une estimation conjointe des symboles transmis et du canal, pour les systèmes TST et STF. Cette approche tensorielle ne nécessite pas de supposer l'indépendance statistique des symboles transmis. Les récepteurs proposés pour les deux systèmes font appel soit à un algorithme du type moindres carrés alternés "Alternating Least Squares": (ALS), soit à la méthode d'optimisation de Levenberg -Marquardt (LM). Un récepteur basé sur la structure du produit de Kronecker, appelé méthode "Kronecker Least Squares" (KLS), est aussi proposé. Des résultats de simulations sont présentés pour illustrer l'efficacité des récepteurs proposés en termes de récupération de symboles et de vitesse de convergence par rapport à d'autres méthodes de la littérature.

Mots-clés: systèmes de communication sans fil MIMO, estimation de symbole, codage spatio-temporel, modèle PARATUCK, modèle CANDECOMP/PARAFAC, codage tensoriel, CDMA, OFDM.

List of Figures

3.1 Transmit processing based on the TST coding with resource allocations. 28
3.2 Visualization of the tensor slice $\mathcal{X}_{..p.}$ of the PARATUCK-(2,4) model. 29
3.3 Uplink processing based on the TST coding with resource allocations. 44

4.1 Computational complexity: Influence of the number P of data blocks. 80
4.2 Computational complexity: Influence of the number K of receive antennas. . . . 81
4.3 Computational complexity: Influence of the spreading length J and of the number F of subcarriers. 81
4.4 Computational complexity: Influence of the number R of data streams. 82
4.5 ZF-TST receiver: Influence of the spreading code length. 86
4.6 ZF-TST receiver: Influence of the number of phases of PSK modulation. 86
4.7 ZF-TST receiver: Influence of the number of data blocks. 87
4.8 ZF-TST receiver: Influence of the number of data streams. 87
4.9 ZF-TST receiver: Influence of the number of receive antennas. 88
4.10 ZF-STF receiver: Influence of the number of subcarriers. 88
4.11 ZF-STF receiver: Influence of the number of phases of PSK modulation. 89
4.12 ZF-STF receiver: Influence of the number of data blocks. 89
4.13 ZF-STF receiver: Influence of the number of data streams. 90
4.14 ZF-STF receiver: Influence of the number of receive antennas. 90
4.15 TST $\times$ STF systems: Influence of the diversity gain by J and F. 91
4.16 Influence of the number of receive antennas: BER versus SNR. 91
4.17 Influence of the number of receive antennas: Channel NMSE versus SNR. 92
4.18 Influence of the number of data blocks: BER versus SNR. 92
4.19 Influence of the number of data blocks: Channel NMSE versus SNR. 93
4.20 KLS-TST Receiver: Influence of a priori information in the symbol recovery. . . 93
4.21 KLS-TST Receiver: Influence of a priori information in the channel estimation. . 94
4.22 KLS-STF Receiver: Influence of a priori information in the symbol recovery. . . 94
4.23 KLS-STF Receiver: Influence of a priori information in the channel estimation. . 95
4.24 TST system: Performance of all receivers for two values of J. 95
4.25 STF system: Performance of all receivers for two values of F. 96
4.26 TST system: Performance of all receivers for two values of R. 96
4.27 STF system: Performance of all receivers for two values of R. 96

4.28 TST system: Convergence speed of all algorithms. 97
4.29 STF system: Convergence speed of all algorithms. 97
4.30 TST system: Convergence of all algorithms with initialization based on the KLS method. 98
4.31 STF system: Convergence of all algorithms with initialization based on the KLS method. 99
4.32 Comparison with KRST and STM for QPSK. 100
4.33 Comparison with KRST and STM for 16-PSK. 100
4.34 Multiuser TST system: Performance of all receivers for *Case 1*. For better visualization, the BER curves for each user are separately plotted for each algorithm in Fig. 4.35. 102
4.35 Multiuser TST system: Performance of all receivers for *Case 1*, (–): averaged over all users and (–): each user. The red curves are also plotted in Fig. 4.34 with the purpose of comparing all algorithms. 102
4.36 ALS-TST receiver: Condition on the code tensor for *Case 1*. 103
4.37 ALS-TST receiver: Condition on the allocation matrices for *Case 2*. 104
4.38 ALS-TST receiver: Influence of the number of receive antennas for *Case 1*. . . . 104

List of Tables

2.1 Summary of the uniqueness results for PT-(2,4) model. 26

3.1 Summary of the uniqueness results for the TST and STF systems. 50

4.1 Minimum value of J for LS identifiability of $\mathbf{S}$ and $\mathbf{H}$ 54
4.2 ALS algorithm for semi-blind joint symbol and channel estimation. 65
4.3 LM algorithm for semi-blind joint symbol and channel estimation. 73
4.4 ALM algorithm for semi-blind joint symbol and channel estimation. 75
4.5 KLS algorithm for semi-blind joint symbol and channel estimation. 79
4.6 Computational complexity per iteration. 79
4.7 Summary of the identifiability conditions with $N \geq R$ and $K \geq M$ (or, $N \geq QR$ and $K \geq QM$). 82

Abbreviations, mathematical notations and operations

List of abbreviations

ALM	Alternating Levenberg-Marquardt
ALS	Alternating Least Squares
BER	Bit-Error-Rate
CDMA	Code Division Multiple Access
CONFAC	Constrained Factor
CP	CANDECOMP/PARAFAC
DS-CDMA	Direct-Sequence Code Division Multiple Access
FDMA	Frequency Division Multiplexing Access
KLS	Kronecker Least Squares
KRST	Khatri-Rao Space-Time
LM	Levenberg-Marquardt
MIMO	Multiple-Input Multiple-Output
ML	Maximum Likelihood
OFDM	Orthogonal Frequency Division Multiplexing
PEP	Pairwise Error Probability
PT	PARATUCK
QPSK	Quadrature Phase Shift Keying
SDMA	Space Division Multiple Access
SER	Symbol-Error-Rate
SF	Space-Frequency
SNR	Signal-to-Noise Ratio
ST	Space-Time
STF	Space-Time-Frequency
TDMA	Time Division Multiple Access
TST	Tensor Space-Time
W-CDMA	Wideband Code Division Multiple Access

List of notations and operations

a	scalars, lower-case
$\mathbf{a}$	column vectors, boldface lower-case
$\mathbf{A}$	matrices, boldface upper-case
$\mathcal{A}$	higher-order tensors, calligraphic letters
$\mathbb{R}$	set of real-valued numbers
$\mathbb{C}$	set of complex-valued numbers
$\mathbb{C}^I$	set of complex-valued I-dimensional vectors
$\mathbb{C}^{I\times J}$	set of complex-valued $(I \times J)$-matrices
$\mathbb{C}^{I_1\times\ldots\times I_N}$	set of complex-valued $(I_1 \times \ldots \times I_N)$-tensors
$[\mathbf{A}]_{i_1,i_2} = a_{i_1,i_2}$	(i_1, i_2)-th element of matrix $\mathbf{A} \in \mathbb{C}^{I_1\times I_2}$
$[\mathcal{A}]_{i_1,\cdots,i_N} = a_{i_1,\ldots,i_N}$	$(i_1, \ldots, i_N)$-th element of tensor $\mathcal{A} \in \mathbb{C}^{I_1\times\ldots\times I_N}$
$\mathcal{A}_{i_1\cdot\cdot} \in \mathbb{C}^{I_2\times I_3}$	i_1-th first-mode matrix-slice of tensor $\mathcal{A} \in \mathbb{C}^{I_1\times I_2\times I_3}$
$\mathcal{A}_{\cdot i_2\cdot} \in \mathbb{C}^{I_1\times I_3}$	i_2-th second-mode matrix-slice of tensor $\mathcal{A} \in \mathbb{C}^{I_1\times I_2\times I_3}$
$\mathcal{A}_{\cdot\cdot i_3} \in \mathbb{C}^{I_1\times I_2}$	i_3-th third-mode matrix-slice of tensor $\mathcal{A} \in \mathbb{C}^{I_1\times I_2\times I_3}$
$(\cdot)^*$	complex conjugate of matrix or vector
$(\cdot)^{\mathrm{T}}$	transpose of matrix or vector
$(\cdot)^{\mathrm{H}}$	conjugate transpose of matrix or vector
$\mathbf{A}^{-1}$	inverse of $\mathbf{A}$
$\mathbf{A}^{\dagger}$	Moore-Penrose pseudoinverse of $\mathbf{A}$
$\mathbf{A}_{i\cdot}$ (resp. $\mathbf{A}_{\cdot j}$)	i-th row (resp. j-th column) of $\mathbf{A}$
$\mathbf{I}_N$	Identity matrix of order N
$\mathbf{1}_N$	all-ones column vector of dimension N
$\mathbf{0}_{M\times N}$	all-zeros $(M \times N)$-matrix
$\mathbf{0}_N$	all-zeros column vector of dimension N
$\mathbf{e}_n^{(N)}$	n-th canonical vector of dimension N
$\lceil a \rceil$	the smallest integer number greater than or equal to a
$\lvert a \rvert$	absolute value (magnitude) of a
$\lVert\cdot\rVert_F$	Frobenius norm
$\mathrm{vec}(\mathbf{A})$	stacks $\mathbf{A}$ into a column vector (vectorization operator)
$\mathrm{diag}(\mathbf{a})$	diagonal matrix built from $\mathbf{a}$
$\mathrm{bdiag}(\mathbf{A}_1, \ldots, \mathbf{A}_{\mathrm{N}})$	block-diagonal matrix built from the N matrices $\{\mathbf{A}_1, \ldots, \mathbf{A}_N\}$
$\mathrm{rank}(\mathbf{A})$	rank of $\mathbf{A}$
$k(\mathbf{A})$ or $k_{\mathbf{A}}$	Kruskal rank (k-rank) of $\mathbf{A}$
$\mathrm{tr}(\mathbf{A})$	trace of $\mathbf{A}$
$\exp(a)$ or e^a	exponential function of a
$P(\cdot)$	probability density function
$Q(\cdot)$	Q-function, defined as $Q(x) = (1/\sqrt{2\pi}) \int_x^\infty e^{-t^2/2}\mathrm{d}t$
$\mathcal{A} \times_n \mathbf{U}$	n-mode product of $\mathcal{A}$ by $\mathbf{U}$
$\circ$	outer product
$\otimes$	Kronecker product
$\diamond$	Khatri-Rao (column-wise Kronecker) product
$\lvert\otimes\rvert$	partition-wise Kronecker product
$\odot$	Hadamard product

Contents

1 Introduction **1**

1.1 General overview of wireless communications 1

1.2 Chapter contents and contributions 6

1.3 Publications 8

2 Tensor decompositions: background and new contributions **9**

2.1 Matrix operations 9

2.2 Basic tensor operations 11

2.3 Background on tensor decompositions and new contributions 14

2.3.1 CANDECOMP/PARAFAC decomposition 14

2.3.2 PARATUCK-(N_1, N) decomposition 19

3 MIMO systems with joint multiplexing and spreading operations **27**

3.1 Tensor Space-Time (TST) system 27

3.2 Space-Time-Frequency (STF) system 32

3.3 Performance analysis 34

3.3.1 TST system 35

3.3.2 STF system 39

3.4 Uniqueness analysis 40

3.4.1 TST system 40

3.4.2 STF system 43

3.5 Generalization of TST systems to multiuser case 44

3.6 Resume of uniqueness results 50

4 Semi-blind receivers **51**

4.1 Alternating Least Squares (ALS) 52

4.1.1 TST system 52

4.1.2 STF system 60

4.1.3 ALS receivers 64

4.1.4 Zero-Forcing (ZF) receivers 64

4.2 Levenberg-Marquardt (LM) 65

4.2.1 Gradient and Jacobian expressions 66
4.2.2 LM and ALM receivers 69
4.3 Kronecker based Least Squares (KLS) 74
4.3.1 TST system 74
4.3.2 STF system 78
4.4 Complexity analysis of algorithms 79
4.5 General discussion 81
4.6 Simulation results 83
4.6.1 Performance analysis of the TST system 85
4.6.2 Performance analysis of the STF system 88
4.6.3 Comparison between TST and STF systems 90
4.6.4 KLS algorithm: Influence of a priori information 93
4.6.5 Comparison between different algorithms 94
4.6.6 Comparison with KRST and STM 99
4.6.7 Generalization of TST systems to multiuser case 101

5 Conclusion and Perspectives **105**

A Basic properties **109**

B Generalization of STF system to multiuser case **111**

Bibliography **115**

Chapter 1

Introduction

1.1 General overview of wireless communications

Radio broadcasting, television broadcasting, satellite communications, and mobile systems are applications of wireless communication systems. A recent interest in wireless communication come from the growing success of cellular systems, mainly since the 1990s with the second generation (known as 2G) technology replacing analogy technology by digital communication. In the past few decades, mobile wireless technologies have been classified according to their generations. Such classifications are distinguished by the type of service (expansion of communication to other sources like images, video, and data), improved privacy, data transfer speeds, improved spectral efficiency, volume data broadcast, and quality of service.

In every new generation of mobile wireless communication, such as the current 3G and 4G technologies, the systems must be designed to support a growing demand for high-quality multimedia services while still being flexible to accommodate new services. These systems must take into account the best tradeoffs between error performance (in terms of symbol or bit error rates, abbreviated as SER and BER, respectively), transmission rate (in symbols or bits per channel use), power and spectral efficiency, and receiver complexity for symbol recovery.

A characteristic of the wireless channels is the existence of many different paths between transmitter and receiver, which leads to different versions of the transmitted signal at the receiver. These resulting signals can widely vary in amplitude and phase. The recovery of the information data may be impaired by the interference from other sources of electromagnetic waves, or even between two or more versions of each transmitted signal (from propagation effects). Moreover, another major limitation of wireless system performance comes from the fading in wireless link.

A key idea for improving the error performance is to exploit the propagation of several paths jointly, which means incurring redundancy into the information-bearing signals available at the receiver. The principle of diversity techniques is to provide different versions of the same signal at receiver, in which each version is ideally affected by independent fading channel. As a consequence, the probability of all signals fading in at the same time is drastically reduced [1, 2]. This enables the mitigation of fading in wireless link and hence, an increase of the reception reliability, which leads to a reduction of error rate.

This redundancy can be provided by channels as it is the case of frequency-selective and time-selective channels, leading to what are called frequency diversity and Doppler diversity, respectively. Redundant information can also be obtained by spreading operations at the transmitter, in space, time and/or frequency domains [3]. It is important to notice that redundancy does not always reflect on the information diversity. In general, the diversity is only achieved when the different versions of the same signal are independently affected by different fading channels.

There are many forms to achieve diversity. They can be classified in four classes, as follows: *space*, *time*, *frequency*, and *polarization* [4, 5]. Generally speaking, space diversity results from the use of multiple antennas at both transmitter and/or receiver ends, which leads to multiple-antenna communication systems with multiple-input multiple-output (MIMO) channels. It is well-known that the deployment of multiple antennas in wireless systems allows improving the transmission rate and reliability over single-transmit antenna systems, while keeping the same transmission bandwidth and power [6, 7, 8]. The works [9, 10, 11] have independently studied and derived expressions for the capacity of MIMO fading channels, concluding that the use of more antennas drastically improves the channel capacity.

An important aspect of the space diversity is that it can be achieved in two cases. The first case is when the transmit and/or receive antennas are properly separated, i.e. with more than half wavelength (exactly 0.38λ according to [4]), which consequently imposes a minimal physical separation preventing its use on small devices. The second case, also known as *angular diversity*, is obtained from the use of directional antennas, which enables its application in small devices.

When the space diversity is obtained from the use of multiple transmit antennas, denoted by *transmitter space diversity*, the transmit power must be divided along all transmit antennas. Contrarily, the space diversity generated from several receive antennas, known as *receiver space diversity*, does not require additional transmit power.

Temporal diversity is derived from the transmission of multiple coded versions of the same signal at different time instants. This diversity takes advantage of variations in the channel in the sense that the coherence time of the channel is small in fast fading channels. The *interleaving* of the coded symbols before the transmission ensures independent fading channels even in the case of slow channels. Thus, robustness of the communication to a temporary deep fade is increased. Nevertheless, the transmission rate is reduced in the sense that an entire block must be received before the initialization of the decoding process.

Frequency diversity employs different carrier frequencies which must be separated by more than the coherence bandwidth of the channel in order to induce different multipath. Another frequency diversity is achieved by frequency-selective channels [2], since the signal bandwidth suffers spreading and the channel gain varies across it. Both time and frequency diversities are not bandwidth efficiency and just the second one requires additional transmit power.

The last diversity refers to the transmission and the reception of signals employing antennas with different polarization, in which the main types are linear, circular and elliptical polarizations. The main disadvantages of this type of diversity are its limitations with respect to the number of polarization types and that it incurs power loss at the transmitter and/or receiver

whereas the power is divided between the polarized antennas. Yet this thesis is focused on the first three types of diversity.

Multiple versions of the transmitted signal obtained from different diversity techniques can be combined with the purpose of improving the system performance [1, 5]. Note that a diversity can not always be available inasmuch as depends on the system feature. Antenna diversity is harder to implement in small devices, such as cellular phones, than in base-stations, because the diversity is impaired as much the channels corresponding to different antennas are correlated.

Temporal diversity in fast fading channels is not recommended for delay-sensitive and stationary applications [6]. In the first application, real-time information is required to be received within a certain time deadline. In stationary applications, the coherence time of channels is infinite which leads to correlated signal versions. Frequency diversity can not be achieved when the delay spread is small, because the frequency components will correspond to correlated fading channels.

In general, the use of several transmit antennas for transmitting the same data signal, known as *space spreading*, can improve the system reliability as well as the transmission of the same symbol during several time periods, known as *time spreading*. On the other hand, the signals can be independently transmitted in parallel by several antennas, denoted by *space multiplexing*, leading to an increase of the transmission rate.

As discussed previously, when multiple antennas are employed in the transmission and/or reception of the same information, the average error probability drastically decreases which leads to an increase of *diversity gain*. At the same time, multiple antennas can be used to transmit in parallel different information data which results in an increase of transmission data rate and thus a *multiplexing gain*. Note that both gains can be simultaneously obtained, but generally the maximal diversity and the maximal multiplexing gains can not be simultaneously achievable [12]. A tradeoff between spatial-multiplexing and diversity gain was proposed in [7]. Another gain similar to diversity gain, the *array gain* results from average power of combining of multiple received signals, which leads to an increase in average received SNR relative to the single-branch average SNR [4, 1].

Alamouti proposed in [13] a transmit diversity scheme using two transmit antennas and one receive antenna with the purpose of fading channel mitigation. Although the first space-time (ST) coding was developed by Tarok *et al.* [6], proposing the construction criteria for ST codes. Since then many works have proposed a variety of ST transmission schemes in order to attain a good compromise between error performance and transmission rate in different system contexts [14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24]. In order to exploit the spatial and frequency diversities over frequency-selective MIMO channels, the space-frequency (SF) coding was proposed in [25]. Initially, ST codes were used directly as SF codes, just replacing the time domain by the frequency domain [25, 26, 27]. However, they showed in [28] that the frequency diversity available in frequency-selective MIMO channels, in general, can not be exploited by using directly ST schemes, requiring a mapping from ST codes to SF codes proposed later in [29, 30]. The proposed mapping in [30] allows that existing ST schemes can be employed to attain full diversity in MIMO-OFDM systems. They have also proposed a tradeoff between error performance and transmission rate for MIMO-OFDM systems. The space-time-frequency

(STF) coding was introduced with the purpose of exploiting jointly the space, time and frequency diversities over frequency-selective channels [31, 32, 33].

In wireless communication, *multiple-access* (MA) techniques permit that the communication resources of the channel to be shared properly by multiple users or local stations. These methods are based on frequency-division (FD), code-division (CD), time-division (TD), space-division (SD) or also combinations of these techniques. The choice of the multiple-access method to be employed depends on the application characteristics and performance requirements of the system.

The combination of ST, SF or STF codes and multiplexing access techniques in the MIMO systems has received much attention in the last years. In the MIMO-OFDM system, SF and STF codings have been employed for providing high data rates and/or reliability through the system diversities available [25, 28, 32, 34, 35, 36, 37, 38, 39]. For the same purpose, the MIMO-CDMA and MIMO-SDMA systems have exploited the ST diversities [40, 41, 42, 17, 43, 12, 44].

Many technical challenges required in the data transmission have driven additional signal processing complexity at the receiver. Channel identification and equalization traditionally use a training period to estimate the propagation channel for a transmitted symbols recovery. As the sources transmit periodically a training sequence known at the receiver, the transmission rate is inevitably affected and for some applications with fast fading channels for example, the training becomes even not effective. For those, and other reasons, blind methods are more appropriate, mainly when the communication requires higher transmission rate.

Blind techniques exploit temporal properties of the signals, channel features or spatial properties of the receiver, such as finite alphabet, orthogonality of the sources, stationarity channel, constant-modulus of the signal constellation, cyclostationarity or statistical independence of the sources [45, 46, 47, 48, 49], for performing channel identification, equalization and source separation.

The use of tensor tools has aroused interest in signal processing applications for wireless communication systems since the pioneer work [50] in 2000. They proposed a blind multiuser separation-equalization-detection for *direct-sequence code-division multiple access* (or simply, DS-CDMA) systems from the modeling of received signals by the parallel factors (CANDECOMP/PARAFAC or shortly, CP) [51, 52]. An interesting advantage is that the CP DS-CDMA receiver does not require knowledge of spreading codes and of channel coefficients, finite alphabet/constant modulus or statistic independence to recover the transmitted signals.

One common feature of all tensor approaches is to perform a jointly blind symbol and channel estimation without a priori channel state information (CSI) at the receiver under identifiability conditions more relaxed than those based on conventional matrix models and without requiring statistical independence between the signals transmitted. Furthermore, the signal processing can be approached in a deterministic way and can directly exploit special features of the system.

Signal processing usually considers space and time dimensions which leads to matrix models. The incorporation of *oversampling*, *spreading* and *multiplexing* operations, and additional diversity on wireless communication systems can be represented as new dimensions that suggests naturally the use of tensors to represent the system model. Another advantage is that tensor structure exploits jointly all available information into signals at the receiver for a signal

recovery purpose.

The tensor decompositions are very useful in modeling signal received, a direct association is established between the parameters of tensor decomposition and the physical parameters of the link communications such as transmitted symbols, channel attenuation coefficients and coding coefficients. As any tensor decomposition, the estimation of certain parameters is desired and its unique determination is ensured by uniqueness conditions for this particular model. The advantages of tensor approach are direct consequence of the essential uniqueness property.

During the last decade, tensorial approaches have been employed to exploit diversities of MIMO wireless systems providing a more reliability to recover the transmitted symbols with blind detection [50, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62]. Sidiropoulos and Budampati in [53] proposed a ST coding based on the Khatri-Rao matrix product, named Khatri-Rao Space-Time (KRST) code, which combines a spatial precoding with a temporal spreading. This code allows to attain a desired performance and provides since full transmission rate to full diversity gain by varying the length of temporal spreading code.

A block tensor model for multiple-access MIMO systems with multiple transmit antennas per user was proposed in [54]. They combined *direct sequence code division* multiple access technique with spatial multiplexing. The received signal is decomposed in a generalized CP, which can be rewritten as a sum of rank-$(L, L, 1)$ terms with tensor rank denoting the number of users [63]. In [60], they presented a bound on the number of users under which blind separation and deconvolution is guaranteed. This system was extended in [61] for multiple users at reception resulting in a generalization of the model, the so-called block component model (BCM).

A three-dimensional tensor coding is constructed from the spatial and temporal spreading and spatial multiplexing in [55]. The blind receiver is based on a constrained CP model with fixed constrained structure. The extension for multiuser can be found in [56]. Differently from [53], the approach of [55] introduces some flexibility at the transceiver by allowing choosing a number of data streams different from the number of transmit antennas.

The works [57], [58] and [59] have introduced the ST coding with allocation structures represented by two or three matrices in order to control the design transmit schemes. In [57], two allocation matrices define the allocation of users' data streams and spreading codes to their transmit antennas. De Almeida *et al.* in [58] generalize [57] by including a third matrix which defines the mapping of the precoded signals to the transmit antennas. The received signals in [57, 58] satisfy a third-order constrained CP model, named as CONFAC model. Differently to [57, 58], two allocation matrices in [59] jointly control the spatial and temporal allocations, i.e. the allocations of data streams to transmit antennas and time-slots. The received signals in [59] satisfy a PARATUCK-2 (or shortly, PT-2) model [64].

Observe that all these last works are restricted to single-carrier transmissions. In the context of multicarrier systems, [62] considers a MIMO system with STF spreading-multiplexing. It is employed two third-order interaction tensors that define a joint temporal and frequency allocation of the data streams to the transmit antennas, thus allowing to adjust the multiplexing degree and spreading redundancy in space (transmit antennas), time (blocks) and frequency (subcarriers). Besides the difference in relation to third-order allocation structures and allocation matrices in [56, 57, 58, 59], the MIMO channel for the STF system links each transmit

antenna with each receive antenna for each subcarrier, which leads to a MIMO channel tensor instead of a matrix.

This work is based on the improvement of system performance given in [59] by the introduction of temporal and frequency diversities. The increment of diversities in the PT-2 model allows to reach a new tensor decomposition, named as PARATUCK-(N_1, N) (or PT-(N_1, N)) [65, 66] which is an extension of PT-2 model. We have investigated the uniqueness conditions of this new model employing CP and PT-2 uniqueness results. We also analyze the performance of our proposed system and the STF system [62] in terms of diversity gain and derive the maximum diversity gain over a at fading channel (per each subcarrier). Conditions for identifiability and uniqueness for both systems are also established.

1.2 Chapter contents and contributions

In addition to this introduction and the concluding remarks, the thesis is organized into three main chapters, as follows:

Chapter 2. We give an overview of matrix operations involving Kronecker and Khatri-Rao products and the concept of k-rank that will be necessary along of this work. Two lemmas concerning the Khatri-Rao product between N matrices are deduced, resulting in generalizations of the lemmas presented respectively in [67, 68]. Both lemmas are important to analyze the uniqueness and identifiability conditions developed in the next chapters of this thesis. Some basic operations and matrix representations of higher-order tensors are introduced.

We present a historical overview of the well-known CP decomposition [52, 51] containing main uniqueness results with a discussion and comparison of these conditions. We also propose a sufficient condition and two extensions for some results in the literature. In the sequence, our proposed decomposition, PT-(N_1, N), is introduced. We show that this model generalizes the PT-2 model [64] and takes advantages of its properties. The uniqueness conditions for PT-2 model have been analyzed and extended. Finally, we deduce the uniqueness results for PT-(N_1, N) decomposition.

Chapter 3. We introduce our proposed space-time coding , named as Tensor Space-Time (TST), and derive the expressions of transmit and receive signals [65, 66]. A performance analysis of the TST coding is deduced with the purpose of evaluating the diversity of information transmitted, which allows us to express a maximum diversity gain in terms of some system parameters and taking into account the structures of antenna and stream allocations per time-block. We propose an uplink processing based on the TST coding with allocation resources for multiple users at the transmission.

The Space-Time-Frequency (STF) system in [62] is described and we derive a performance analysis based upon the one done previously for TST system. The uniqueness conditions for both systems are investigated in an unified way using the results presented in the previous chapter and taking into consideration the structure of each system.

Chapter 4. In this chapter, we present some semi-blind receivers for joint channel estimation and symbol recovery. We apply in both systems the well-known Alternating Least Squares (ALS) and the Levenberg-Marquardt (LM) algorithms, and propose two new algorithms: the

Alternating Levenberg-Marquardt (ALM) and the Kronecker based Least Square (KLS). The ALM is a simplified version of the LM algorithm and the KLS is a non-iterative algorithm based on the structure of the Kronecker product. All algorithms are compared in terms of identifiability conditions, complexity and convergence speed. Finally, some simulation results are provided to evaluate the performance of these receivers for both systems and to compare the TST coding with other ST codings based on tensor approaches.

Main contributions

We highlight clearly all contributions developed in this thesis per chapter as follows.

Chapter 2.

- Development of two lemmas regarding the k-rank of Khatri-Rao product between several matrices, which generalize two lemmas presented respectively in [67, 68] (Section 2.1).
- Proposition of a sufficient uniqueness condition for the N-th order CP model based on the results in [69] and generalization of two theorems proposed in [70] concerning partial uniqueness conditions for third-order CP model extended to an N-th order CP model (Subsection 2.3.1).
- Proposition of the PT-(N_1, N) model, which generalizes the PT-2 model [64] (Subsection 2.3.2).
- Development of the uniqueness conditions for our proposed model based on the uniqueness results in [59], [69] and [70] (Subsection 2.3.2).

Chapter 3.

- Proposition of the TST coding for MIMO wireless communication systems modeled by the PT-(2,4) and a constrained CP models (Section 3.1).
- Proposition of an uplink processing based on the TST coding with allocation resources.
- Development of a performance analysis based on the diversity of transmitted information and derivation of the maximum diversity gain for TST [66] and STF [62] systems (Section 3.3).
- Proposition of the uniqueness conditions based on the results in Chapter 2 for both systems in an unified way (Section 3.4).

Chapter 4.

- Development of semi-blind receivers for joint channel estimation and symbol recovery applying the ALS and LM algorithms (Sections 4.1 and 4.2).
- Proposition of new algorithms: the ALM, a simplified version of the LM algorithm, and the KLS, a non-iterative algorithm based on the structure of the Kronecker product (Sections 4.3 and 4.2).

- A complete comparison between all algorithms in terms of identifiability conditions, complexity per iteration, convergence speed for both systems.

1.3 Publications

- G. Favier, M. N. da Costa, A. L. F. de Almeida, J. M. T. Romano. "Tensor coding for CDMA-MIMO wireless communication systems", 19th European Signal Processing Conference (EUSIPCO), 2011, Barcelone.

- G. Favier, M. N. da Costa, A. L. F. de Almeida, J. M. T. Romano. "Tensor Space-time (TST) coding for MIMO wireless communication systems", Signal Processing 92 (4) 2012, 1079-1092.

- M. N. da Costa, G. Favier, A. L. F. de Almeida, J. M. T. Romano. "A comparative study of TST and STF codings for tensor based MIMO communication systems with semi-blind receivers", submitted.

- M. N. da Costa, G. Favier, A. L. F. de Almeida, J. M. T. Romano. "Semi-blind receivers for TST-MIMO multiuser systems", in preparation.

Chapter 2

Tensor decompositions: background and new contributions

In this chapter, we introduce the basic definitions and operations of multilinear algebra which are used in this thesis. Firstly, let us show some matrix operations and in the sequence, two lemmas derived from works of Sidiropoulos *et al.* [67, 71] which correspond to contributions of this thesis and are important to study the uniqueness and identifiability conditions. In the second part, we bring up some basic tensor operations. Finally, in the third part, we present a historical overview of the CANDECOMP/PARAFAC (or in abbreviation, CP) model and introduce the proposed PARATUCK-(N_1,N) (or in abbreviation, PT-(N_1,N)) model giving the main results concerning the uniqueness of both decompositions. Another contribution of this chapter concerns the proposition of a sufficient uniqueness condition based on the results in [69] for the N-th order CP model, the extension of the uniqueness conditions proposed in [70] for any N-th order CP model, and the proposition of uniqueness conditions for our proposed PT-(N_1, N) model, based on the CP and PT-2 results.

2.1 Matrix operations

There are two principal products widely employed in the tensor approaches, known as *Kronecker* and *Khatri-Rao* products. We will see in the next section that higher-order tensors can be represented by *unfolded* matrices and these products are frequently used in the context of tensor decompositions to simplify expressions and to explore their properties [72, 73].

The Khatri-Rao product, which is equivalent to a column-wise Kronecker product of two matrices, was introduced by Khatri and Rao [74]. Before to transcribe the definition of the Khatri-Rao product, recall that the Kronecker product between two matrices $\mathbf{A} \in \mathbb{C}^{I\times J}$ and $\mathbf{B} \in \mathbb{C}^{M\times N}$ results in

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

Definition 2.1. (Khatri-Rao product).[74] *The Khatri-Rao product, denoted by $\diamond$, of two matrices $\mathbf{A} \in \mathbb{C}^{I\times R}$ and $\mathbf{B} \in \mathbb{C}^{J\times R}$ results in a $(IJ \times R)$-matrix given by*

$$\mathbf{A} \diamond \mathbf{B} = \left[\begin{array}{ccc} \mathbf{A}_{\cdot 1} \otimes \mathbf{B}_{\cdot 1} & \ldots & \mathbf{A}_{\cdot R} \otimes \mathbf{B}_{\cdot R} \end{array} \right] \in \mathbb{C}^{IJ\times R}. \tag{2.2}$$

Let us also define the partition-wise Kronecker product of two matrices as follows

Definition 2.2. (partition-wise Kronecker product).[75] *Let $\mathbf{A} = \left[\mathbf{A}^{(1)} \cdots \mathbf{A}^{(R)}\right] \in \mathbb{C}^{I\times RJ}$ and $\mathbf{B} = \left[\mathbf{B}^{(1)} \cdots \mathbf{B}^{(R)}\right] \in \mathbb{C}^{M\times RN}$ be two partitioned matrices. The partition-wise Kronecker product, denoted by $|\otimes|$, of two matrices $\mathbf{A} \in \mathbb{C}^{I\times RJ}$ and $\mathbf{B} \in \mathbb{C}^{M\times RN}$ results in a $(IM \times RJN)$-matrix given by*

$$\mathbf{A}|\otimes|\mathbf{B} = \left[\begin{array}{ccc} \mathbf{A}^{(1)} \otimes \mathbf{B}^{(1)} & \ldots & \mathbf{A}^{(R)} \otimes \mathbf{B}^{(R)} \end{array} \right] \in \mathbb{C}^{IM\times RJN}. \tag{2.3}$$

In order to deduce two new lemmas proposed in this thesis that will be useful for the uniqueness and identifiability analysis, we need to give a brief overview of the definition of the Kruskal rank of a matrix. The concept of k-rank was introduced by Kruskal in 1977 [76] which cames from his studies of the uniqueness for the CP decomposition but the term *k-rank* for *Kruskal rank* was just later named by Harshman and Lundy [77]. This concept is important because the most general results on uniqueness involve or depend on the k-rank of a matrix.

Definition 2.3. (Kruskal rank or k-rank).[76] *The Kruskal rank or k-rank of a matrix $\mathbf{A}$, denoted by $k(\mathbf{A})$ or $k_{\mathbf{A}}$, is the maximal number $k_{\mathbf{A}}$ such that any set of $k_{\mathbf{A}}$ columns of $\mathbf{A}$ is linearly independent.*

The distinction between k-rank and rank is important. Verify that $k(\mathbf{A}) = x$ implies that every x columns of $\mathbf{A}$ are linearly independent, whereas $\text{rank}(\mathbf{A}) = x$ requires that at least x columns are linearly independent. Thus, the k-rank is more constrained than the rank of a matrix $\mathbf{A}$ implying $k(\mathbf{A}) \leq \text{rank}(\mathbf{A})$.

Now, let us introduce a lemma that gives a lower-bound on the k-rank of Khatri-Rao product between N matrices, this lemma is an extension of the lemma for a Khatri-Rao product (i.e. $N = 2$) proved in [67, 68] and posteriorly proved in a different way in [78, 79].

Lemma 2.1. (k-rank of Khatri-Rao products). *Consider the Khatri-Rao product of N matrices*

$$\mathbf{A}^{(1)} \diamond \cdots \diamond \mathbf{A}^{(N)} = \mathop{\diamond}_{n=1}^{N} \mathbf{A}^{(n)} = \left[\begin{array}{ccc} \mathop{\otimes}_{n=1}^{N} \mathbf{A}_{\cdot 1}^{(n)} & \cdots & \mathop{\otimes}_{n=1}^{N} \mathbf{A}_{\cdot R}^{(n)} \end{array} \right] \in \mathbb{C}^{I_1\cdots I_N\times R} \tag{2.4}$$

and define $k_{\mathbf{A}^{(n)}} \triangleq k\left(\mathbf{A}^{(n)}\right)$ as the k-rank of $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n\times R}$ for $n \in \{1, ..., N\}$.

(i) If $k_{\mathbf{A}^{(n)}} = 0$ for a given n, then $k_{\mathbf{A}^{(1)}\diamond\cdots\diamond\mathbf{A}^{(N)}} = 0$.

(ii) If $k_{\mathbf{A}^{(n)}} \geq 1$ for all n, then $k_{\mathbf{A}^{(1)}\diamond\cdots\diamond\mathbf{A}^{(N)}} \geq \min\left(k_{\mathbf{A}^{(1)}} + \cdots + k_{\mathbf{A}^{(N)}} - (N-1), R\right)$.

Proof: First, let us prove (i). When $k_{\mathbf{A}^{(n)}} = 0$ for a given n, $\mathbf{A}^{(n)}$ has at least one zero column. Thereby, supposing $\mathbf{A}_{\cdot r^*}^{(n^*)} = \mathbf{0}_{I_{n^*}}$ for a given r^* and n^*, it gets from (2.4): $\mathop{\otimes}_{n=1}^{N} \mathbf{A}_{\cdot r^*}^{(n)} = \mathbf{0}_{I_1\ldots I_N}$, i.e. $\mathbf{A}^{(1)} \diamond \cdots \diamond \mathbf{A}^{(N)}$ has one zero column which implies $k_{\mathbf{A}^{(1)}\diamond\cdots\diamond\mathbf{A}^{(N)}} = 0$. We can prove (ii) applying

the Lemma 1 in [67, 68] to the Khatri-Rao product between two matrices. Suppose $k_{\mathbf{A}^{(n)}} \geq 1$ for all n, we obtain

$$k_{\mathbf{A}^{(1)} \diamond \mathbf{A}^{(2)}} \geq \min\left(k_{\mathbf{A}^{(1)}} + k_{\mathbf{A}^{(2)}} - 1, R\right), \tag{2.5}$$

$$\begin{aligned} k_{\mathbf{A}^{(1)} \diamond \mathbf{A}^{(2)} \diamond \mathbf{A}^{(3)}} &\geq \min\left(k_{\mathbf{A}^{(1)} \diamond \mathbf{A}^{(2)}} + k_{\mathbf{A}^{(3)}} - 1, R\right) \\ &\geq \min\left(k_{\mathbf{A}^{(1)}} + k_{\mathbf{A}^{(2)}} + k_{\mathbf{A}^{(3)}} - 2, R\right). \end{aligned} \tag{2.6}$$

Both expressions can be easily extended to the general case of Khatri-Rao product between N matrices, i.e. $N-1$ Khatri-Rao products, by:

$$k_{\mathbf{A}^{(1)} \diamond \cdots \diamond \mathbf{A}^{(N)}} \geq \min\left(k_{\mathbf{A}^{(1)}} + \cdots + k_{\mathbf{A}^{(N)}} - (N-1), R\right). \tag{2.7}$$

■

This contribution allows us to derive the next lemma that leads to an extension of the lemma for full rank of one Khatri-Rao product proposed in [71].

Lemma 2.2. (Full rank of Khatri-Rao products). *Consider* $\mathbf{A}^{(1)} \diamond \cdots \diamond \mathbf{A}^{(N)}$ *with* $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R}$, $n \in \{1, ..., N\}$. *If* $k_{\mathbf{A}^{(n)}} \geq 1$ *for all* n *and* $\sum_{n=1}^{N} k_{\mathbf{A}^{(n)}} \geq R + N - 1$, *then* $\mathbf{A}^{(1)} \diamond \cdots \diamond \mathbf{A}^{(N)}$ *is full column-rank, implying* $I_1 \cdots I_N \geq R$.

Proof: According to Lemma 2.1 for $k_{\mathbf{A}^{(n)}} \geq 1$, $n \in \{1, \ldots, N\}$, we have

$$k_{\mathbf{A}^{(1)} \diamond \cdots \diamond \mathbf{A}^{(N)}} \geq \min\left(k_{\mathbf{A}^{(1)}} + \cdots + k_{\mathbf{A}^{(N)}} - (N-1), R\right). \tag{2.8}$$

If $\sum_{n=1}^{N} k_{\mathbf{A}^{(n)}} \geq R + N - 1$, then $\min\left(k_{\mathbf{A}^{(1)}} + \cdots + k_{\mathbf{A}^{(N)}} - (N-1), R\right) = R$. By the k-rank definition $k_{\mathbf{A}^{(1)} \diamond \cdots \diamond \mathbf{A}^{(N)}} \leq \min\left(I_1 \ldots I_N, R\right)$, it leads to $k_{\mathbf{A}^{(1)} \diamond \cdots \diamond \mathbf{A}^{(N)}} = R$ implying that $\mathbf{A}^{(1)} \diamond \cdots \diamond \mathbf{A}^{(N)}$ is full column-rank. ■

2.2 Basic tensor operations

A tensor is a multi-dimensional array of numerical values. The order of a tensor is the dimensionality of array, or equivalently, the number of indices. Thereby, a matrix and a vector can be respectively represented by a 2-dimensional and a 1-dimensional array and therefore, are respectively a second-order and a first-order tensor. A scalar is a single number and thus, zero-order tensor.

Definition 2.4. (Scalar notation). *Each element of an* N*-th order tensor,* $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$, *can be denoted by*

$$a_{i_1, i_2, \ldots, i_N} \triangleq [\mathcal{A}]_{i_1, i_2, \ldots, i_N}, \tag{2.9}$$

$i_n \in \{1, \ldots, I_n\}$ *with* $n \in \{1, \ldots, N\}$ *is an* n*-th dimension or also called as an* n*-th mode of* $\mathcal{A}$.

Definition 2.5. *The outer product between a tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_M}$ of M-th order and $\mathcal{B} \in \mathbb{C}^{J_1 \times J_2 \times \cdots \times J_N}$ of N-th order results in a tensor of $(M+N)$-th order defined as*

$$[\mathcal{A} \circ \mathcal{B}]_{i_1,i_2,\ldots,i_M,j_1,j_2,\ldots,j_N} \triangleq a_{i_1,i_2,\ldots,i_M} b_{j_1,j_2,\ldots,j_N}, \tag{2.10}$$

for all index values.

The outer product of two tensors results in another tensor and the order of this new tensor is given by the sum of the order of both tensors. Observe that (2.10) is a generalization of the outer product of two vectors, which results in a matrix.

Definition 2.6. (Rank-one tensor). *An N-th order tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ has rank one if it can be written as the outer product of N vectors $\mathbf{u}^{(n)} \in \mathbb{C}^{I_n}$, $n \in \{1, \ldots, N\}$:*

$$\mathcal{A} = \mathbf{u}^{(1)} \circ \mathbf{u}^{(2)} \circ \cdots \circ \mathbf{u}^{(N)} \quad \Longleftrightarrow \quad a_{i_1,\ldots,i_N} = u^{(1)}_{i_1} u^{(2)}_{i_2} \ldots u^{(N)}_{i_N}, \tag{2.11}$$

for all index values.

This definition is a generalization of a rank-one matrix: a matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$ has rank-one if and only if it can be written as the outer product of two vectors $\mathbf{u} \in \mathbb{C}^M$ and $\mathbf{v} \in \mathbb{C}^N$, i.e. $\mathbf{A} = \mathbf{u} \circ \mathbf{v} = \mathbf{u}\mathbf{v}^{\mathrm{T}} \Longleftrightarrow a_{m,n} = u_m v_n$ for all index values.

Definition 2.7. *The Frobenius norm of a tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ is defined as*

$$\|\mathcal{A}\|_F \triangleq \sqrt{\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} |a_{i_1,i_2,\ldots,i_N}|^2}. \tag{2.12}$$

This is analogous to the matrix Frobenius norm.

Definition 2.8. *The n-mode product of a tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_n \times \cdots \times I_N}$ and a matrix $\mathbf{U} \in \mathbb{C}^{J_n \times I_n}$ is a $(I_1 \times \cdots \times I_{n-1} \times J_n \times I_{n+1} \times \cdots \times I_N)$-tensor given by*

$$[\mathcal{A} \times_n \mathbf{U}]_{i_1,\ldots,i_{n-1},j_n,i_{n+1},\ldots,i_N} \triangleq \sum_{i_n=1}^{I_n} a_{i_1,\ldots,i_n,\ldots,i_N} u_{j_n,i_n}, \tag{2.13}$$

for all index values.

The n-mode product is a compact form to represent linear transformations involving tensors. Rewriting (2.13) as $\mathcal{B} = \mathcal{A} \times_n \mathbf{U}$, it gets

$$\begin{bmatrix} b_{i_1,\ldots,i_{n-1},1,i_{n+1},\ldots,i_N} \\ \vdots \\ b_{i_1,\ldots,i_{n-1},J_n,i_{n+1},\ldots,i_N} \end{bmatrix} = \mathbf{U} \begin{bmatrix} a_{i_1,\ldots,i_{n-1},1,i_{n+1},\ldots,i_N} \\ \vdots \\ a_{i_1,\ldots,i_{n-1},I_n,i_{n+1},\ldots,i_N} \end{bmatrix}, \tag{2.14}$$

which represents a linear transformation mapping $\mathbb{C}^{I_n}$ to $\mathbb{C}^{J_n}$ on the n-th dimension of $\mathcal{A}$.

Corollary 2.1. *Given a tensor* $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ *and two matrices* $\mathbf{F} \in \mathbb{C}^{J_n \times I_n}$ *and* $\mathbf{G} \in \mathbb{C}^{J_m \times I_m}$ *with* $m \neq n$*, one has:*

$$(\mathcal{A} \times_n \mathbf{F}) \times_m \mathbf{G} = (\mathcal{A} \times_m \mathbf{G}) \times_n \mathbf{F} = \mathcal{A} \times_n \mathbf{F} \times_m \mathbf{G}. \tag{2.15}$$

And two n-mode products satisfy the identity

$$(\mathcal{A} \times_n \mathbf{F}) \times_n \mathbf{G} = \mathcal{A} \times_n (\mathbf{GF}) \tag{2.16}$$

with $\mathbf{F} \in \mathbb{C}^{L_n \times I_n}$ *and* $\mathbf{G} \in \mathbb{C}^{J_n \times L_n}$*.*

From this notation, the matrix product $\mathbf{UAV}^{\mathrm{H}}$, with $\mathbf{A} \in \mathbb{C}^{I_1 \times I_2}$, $\mathbf{U} \in \mathbb{C}^{J_1 \times I_1}$ and $\mathbf{V} \in \mathbb{C}^{J_2 \times I_2}$, can be equivalently rewritten as $\mathbf{A} \times_1 \mathbf{U} \times_2 \mathbf{V}^*$.

Matrix representation of a higher-order tensor

The process of reordering the elements of an N-th order tensor into a matrix is known as *unfolding* or *matricization*. In our notation, we explicit the size of the unfolded matrix in terms of the sizes of the tensor dimensions in order to convert the matrix back to the original tensor.

Definition 2.9. (Matrix unfolding). *Assume an N-th order tensor* $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ *and let the ordered sets* $\mathbb{A} = \{a_1, \ldots, a_K\}$ *and* $\mathbb{B} = \{b_1, \ldots, b_L\}$ *be a partitioning of its modes* $\mathbb{D} = \{1, \ldots, N\}$*. The unfolded matrix* $\mathbf{X}_{P \times Q} \in \mathbb{C}^{P \times Q}$ *of this tensor can be denoted by*

$$[\mathbf{X}_{P \times Q}]_{p,q} = x_{i_1, i_2, \ldots, i_N}, \text{ with } P = \prod_{p \in \mathbb{A}} I_p \text{ and } Q = \prod_{q \in \mathbb{B}} I_q, \tag{2.17}$$

both indices p and q are associated with several modes.

Let us illustrate the above definition by the following example: For $\mathbb{A} = \{1, 2\}$ and $\mathbb{B} = \{3, \ldots, N\}$, then we can write $\mathbf{X}_{I_1 I_2 \times I_3 \ldots I_N}$. The name *matricization* comes from the analogy to the more common term *vectorization*, the vectorization of a matrix is just a special case of the matricization of a tensor, which all modes become row modes, i.e. $\mathrm{vec}(\mathbf{X}) = \begin{bmatrix} \mathbf{X}_{\cdot 1}^{\mathrm{T}} & \ldots & \mathbf{X}_{\cdot I_2}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \in \mathbb{C}^{I_2 I_1}$ for $\mathbf{X} \in \mathbb{C}^{I_1 \times I_2}$. The vectorization of a higher-order tensor can be derived from the vectorization of a matrix unfolding of this tensor. Observe that the order of the dimensions is relevant because denotes the speed at each mode changes, it will be clear along this work.

Any higher-order tensor can be represented by matrix unfoldings. Verify that there are many different ways to write an unfolded matrix of a higher-order tensor by rearranging its modes into a matrix and by permutation of its indices. Consider for example a tensor with three dimensions $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$, we can define three different matrices: $\mathbf{A}_{I_1 \times I_2 I_3} \triangleq \begin{bmatrix} \mathbf{A}_{\cdot 1 \cdot} & \ldots & \mathbf{A}_{\cdot I_2 \cdot} \end{bmatrix}$, $\mathbf{A}_{I_2 \times I_3 I_1} \triangleq \begin{bmatrix} \mathbf{A}_{\cdot \cdot 1}^{\mathrm{T}} & \ldots & \mathbf{A}_{\cdot \cdot I_3}^{\mathrm{T}} \end{bmatrix}$ and $\mathbf{A}_{I_3 \times I_1 I_2} \triangleq \begin{bmatrix} \mathbf{A}_{1 \cdot \cdot}^{\mathrm{T}} & \ldots & \mathbf{A}_{I_1 \cdot \cdot}^{\mathrm{T}} \end{bmatrix}$ to represent the same tensor with

$$a_{i_1, i_2, i_3} = [\mathbf{A}_{I_1 \times I_2 I_3}]_{i_1, (i_2 - 1) I_3 + i_3} = [\mathbf{A}_{I_2 \times I_3 I_1}]_{i_2, (i_3 - 1) I_1 + i_1} = [\mathbf{A}_{I_3 \times I_1 I_2}]_{i_3, (i_1 - 1) I_2 + i_2}, \tag{2.18}$$

for all index values and $\mathbf{A}_{i_1 \cdot \cdot} \in \mathbb{C}^{I_2 \times I_3}$, $\mathbf{A}_{\cdot i_2 \cdot} \in \mathbb{C}^{I_1 \times I_3}$ and $\mathbf{A}_{\cdot \cdot i_3} \in \mathbb{C}^{I_1 \times I_2}$ denoting the *slice* of $\mathcal{A}$, constructed by fixing the modes 1, 2 and 3, respectively. Remark from (2.18) that the order of the indices is important to construct the unfolded matrices from the matrix *slicings*, the indices placed more to the left vary slower and the ones placed more to the right vary faster.

2.3 Background on tensor decompositions and new contributions

This section gives an overview of the most known tensor model, called CANDECOMP/ PARAFAC [52, 51] and presents the main results concerning the uniqueness conditions of this model. In the sequence, we introduce our proposed model named as PARATUCK-(N_1, N), or simply PT-(N_1, N), which generalizes the standard PARATUCK-2 (or in abbreviation, PT-2) model [64]. We recall the uniqueness results for the PT-2 model and analyze the uniqueness conditions for a special structure of the PT-(N_1, N), for $N_1 = 2$ and $N = 4$, which will be useful to our specific problem. Finally, we extend the uniqueness conditions to the PT-(N_1, N) model from the results deduced from PT-(2,4) and CP decompositions.

2.3.1 CANDECOMP/PARAFAC decomposition

The idea of expressing a tensor as a sum of rank-one tensors was originally introduced by Hitchcock in 1927 [80]. The decomposition of three-way arrays has been developed in an independent way by Carrol and Chang [52] in psychometrics, named as CANDECOMP (canonical decomposition), and by Harshman [51] in phonetics, named as PARAFAC (parallel factors). Both names report to different features of this model, in this thesis we use the abbreviation CP to refer to the CANDECOMP/PARAFAC decomposition.

The CP model decomposes an N-th order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ into a sum of rank-one tensors so it can be expressed using the outer product notation (2.11) and also in a concise form as

$$\begin{aligned} \mathcal{X} &= \left[\!\left[\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \ldots, \mathbf{A}^{(N)} \right]\!\right] \\ &= \sum_{r=1}^{R} \mathbf{A}^{(1)}_{\cdot r} \circ \mathbf{A}^{(2)}_{\cdot r} \circ \cdots \circ \mathbf{A}^{(N)}_{\cdot r}, \end{aligned} \tag{2.19}$$

where its elements are given by

$$x_{i_1,\ldots,i_N} = \sum_{r=1}^{R} a^{(1)}_{i_1,r}\, a^{(2)}_{i_2,r} \ldots a^{(N)}_{i_N,r} = \sum_{r=1}^{R} \prod_{n=1}^{N} a^{(n)}_{i_n,r}, \tag{2.20}$$

$\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R}$ for $n \in \{1, \ldots, N\}$ and $R > 0$ denote, respectively, the matrix factors and the rank of $\mathcal{X}$.

The usual definition of tensor rank, proposed in [80] and independently later in [76], comes from the minimum number R of rank-one tensors sufficient to decompose a tensor. Therefore, the decomposition in (2.19) is irreducible in the sense that it can not be represented using less than R components of rank-one.

Remark that any higher-order tensor can be decomposed in the form of (2.19), it leads to the rank definition: Any matrix can be decomposed in a sum of rank-one matrices, i.e. $\mathbf{X} = \sum_{r=1}^{R} \mathbf{a}_r \mathbf{b}_r^{\mathrm{T}} \in \mathbb{C}^{M \times N}$ for $\mathbf{a}_r \in \mathbb{C}^M$ and $\mathbf{b}_r \in \mathbb{C}^N$, where the minimum number of R defines the matrix rank. Contrary to matrices, the rank of higher-order tensors is not bounded by the

tensor dimensions. The determination of tensor rank is not easy and can be found in some special cases [81, 82, 83].

As we have presented in Section 2.2, any higher-order tensor can be represented in terms of matrix unfoldings. Thus, we can express the CP model in a matricized form using the Khatri-Rao product by

$$\mathbf{X}_{I_n \times I_1 \ldots I_{n-1} I_{n+1} \ldots I_N} = \mathbf{A}^{(n)} \left(\mathbf{A}^{(1)} \diamond \cdots \diamond \mathbf{A}^{(n-1)} \diamond \mathbf{A}^{(n+1)} \diamond \cdots \diamond \mathbf{A}^{(N)} \right)^{\mathrm{T}} \tag{2.21}$$

for a given $n \in \{1, \ldots, N\}$. Remark that $\mathbf{X}_{I_n \times I_1 \ldots I_{n-1} I_{n+1} \ldots I_N}$ is just one way to represent the tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \ldots I_N}$ isolating an n-th matrix factor $\mathbf{A}^{(n)}$ on the left side of the expression (2.21).

The CP decomposition of an N-th order tensor, defined in (2.19), consists in determining N matrix factors $\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(N)}$ from which the original tensor was constructed. It is important to investigate the conditions that can guarantee the uniqueness. The uniqueness of the CP decomposition has been explored since early in the 70's and the main advantage of this decomposition is due to the well-known essential uniqueness of this model, which is crucial in many applications.

Uniqueness results

The first published uniqueness result is due to Harshman [51], giving credit to Robert Jennrich of the UCLA Department of Mathematics for the uniqueness proof in this work. Several results regarding uniqueness of the CP model have been developed ever since [84, 76, 85, 86, 50, 67, 87, 88, 89, 90, 91, 78, 92, 69, 70], but the most general sufficient condition and well-known result on uniqueness is attributed to Kruskal [76].

Kruskal has proposed a sufficient condition in [76] for the essential uniqueness of real third-order CP decomposition that was better discussed in [85] and alternative proofs were presented in [78] and [93]. This uniqueness condition was extended for complex case by Sidiropoulos, Giannakis and Bro in [50] and generalized for an N-th order tensor, i.e. for $N > 3$, by Sidiropoulos and Bro in [67]. This result is reproduced as follows:

Theorem 2.1. [67] *Consider the N-th order tensor $\mathcal{X} = \left[\left[\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \ldots, \mathbf{A}^{(N)}\right]\right]$ with $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R}$, $n \in \{1, \ldots, N\}$ and suppose R is its rank. Thus, $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R}$, $n \in \{1, \ldots, N\}$ are unique up to permutation and scaling of columns provided that*

$$\sum_{n=1}^{N} k_{\mathbf{A}^{(n)}} \geq 2R + (N - 1). \tag{2.22}$$

The proof of this theorem exploits Lemma 1 in [67] which gives a lower-bound on the k-rank of Khatri-Rao product and is based on rewriting of a N-th order CP model as another CP model of third order, implying the concatenation of $N - 2$ loading matrices in only one matrix. Thus, the principle of lemma 2.1 proposed and formulated in this thesis is implicitly suggested in this Lemma proof.

Remark from (2.22) that each dimension of the tensor increments just one in the minimum value of the total sum of k-rank, then the increase of tensor dimensions N decreases the k-rank required per dimension. The minimum k-rank required per each matrix factor is two for three

dimensions ($N = 3$), i.e. $k_{\mathbf{A}^{(n)}} \geq 2$ for $n \in \{1, 2, 3\}$, and for more than 3 dimensions ($N > 3$), it is possible to guarantee uniqueness with one matrix of k-rank equals to one. Therefore, the uniqueness condition (2.22) becomes less restrictive with rising of the tensor order N.

When the loading matrices $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R}$ for $n = 1, \ldots, N$ are drawn independently from absolutely continuous distributions, these matrices are full k-rank, i.e. $k_{\mathbf{A}^{(n)}} = \min(I_n, R)$, $n = 1, \ldots, N$ and the condition (2.22) becomes [67]

$$\sum_{n=1}^{N} \min(I_n, R) \geq 2R + (N - 1). \tag{2.23}$$

Jiang and Sidiropoulos in [89] and De Lathauwer in [90] have independently derived an alternative uniqueness condition for third-order CP decomposition where one of the factor matrix is full column-rank. For the third-order CP model $\mathcal{X} = \left[\!\left[\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{A}^{(3)}\right]\!\right]$ with $\mathbf{A}^{(3)}$ being full column-rank ($k_{\mathbf{A}^{(3)}} = R \leq I_3$), these conditions lead to $k_{\mathbf{A}^{(1)}} + k_{\mathbf{A}^{(2)}} \geq R + 2$, which is a special case of Kruskal's condition [76] and according to Lemma 2.2, $k_{\mathbf{A}^{(1)}} + k_{\mathbf{A}^{(2)}} \geq R + 2$ implies that $\mathbf{A}^{(1)} \diamond \mathbf{A}^{(2)}$ is full column-rank ($I_1 I_2 \geq R$). Differently to Jiang and Sidiropoulos, De Lathauwer proposed a deterministic condition for the same case, considering a third-order tensor with two *generic* component matrices which are randomly sampled from a continuous distribution. This condition was derived in the form of a dimensionality constraint for third and also fourth-order tensors.

In [91], a link is established between the uniqueness conditions of Jiang and Sidiropoulos and of De Lathauwer, and it is proved that both conditions are more relaxed than the classical Kruskal's condition for the special case in which one factor matrix is full column-rank. These results were posteriorly extended to arbitrary order $N > 3$ tensors by Stegeman in [69]. Although both conditions are equivalent, the condition deduced by De Lathauwer is easer to be checked. Let us enunciated it as follows for $N \geq 3$. Consider an N-th order CP model $\mathcal{X} = \left[\!\left[\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(N)}\right]\!\right]$, with generic $\left(\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(N-1)}\right)$, $\mathbf{A}^{(N)}$ full column-rank and $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R}$, $n \in \{1, \ldots, N\}$. This decomposition is unique if

$$\frac{R(R-1)}{2} \leq \sum_{n=2}^{N-1} \left(2^{n-1} - 1\right) Q_{(n,N)}, \tag{2.24}$$

with

$$Q_{(n,N)} = \sum_{\mathbb{S}_n} \prod_{j \in \mathbb{S}_n} \frac{I_j(I_j - 1)}{2} \prod_{j \notin \mathbb{S}_n} I_j \tag{2.25}$$

where the summation is over all subsets $\mathbb{S}_n$ of $n \in \{1, \ldots, N-1\}$ containing n distinct elements. If $n = N - 1$, then we set $\prod_{j \notin \mathbb{S}_n} I_j = 1$.

Although these conditions are more relaxed than Kruskal's condition, they are still sufficient conditions and furthermore, for a restricted CP model. The condition based on Jiang and Sidiropoulos approach is not practical and the one based on De Lathauwer is not guaranteed for matrices no randomly constructed.

Since the proposed condition for uniqueness by Kruskal [76], attempts to prove that it is also necessary have been investigated. In [88], Ten Berge and Sidiropoulos proved that Kruskal's

condition is not only sufficient but also necessary for tensors of rank $R = 2$ and $R = 3$, but not necessary when $R > 3$.

More general necessary uniqueness condition of a third-order CP model $\mathcal{X} = \left[\left[\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{A}^{(3)}\right]\right]$ was proposed by Liu and Sidiropoulos in [87] and checked in [78], this condition is given by

$$\min\left(\operatorname{rank}\left(\mathbf{A}^{(1)} \diamond \mathbf{A}^{(2)}\right), \operatorname{rank}\left(\mathbf{A}^{(1)} \diamond \mathbf{A}^{(3)}\right), \operatorname{rank}\left(\mathbf{A}^{(2)} \diamond \mathbf{A}^{(3)}\right)\right) = R. \tag{2.26}$$

In another way, the Khatri-Rao product of any two matrix factors must be full column-rank in order to guarantee the CP decomposition uniqueness. This work gave also an idea for its extension to arbitrary dimension N that was appropriately proved by Stegeman in [69]. For higher-order CP model, the Khatri-Rao product between any $N-1$ matrices, i.e. any *leave-one-out* selection of matrices $\mathbf{A}^{(n)}$ with $n \in \{1, ..., N\}$, must be full column-rank, i.e.

$$\min_{n=1,\ldots,N}\left(\operatorname{rank}\left(\mathbf{A}^{(1)} \diamond \ldots \diamond \mathbf{A}^{(n-1)} \diamond \mathbf{A}^{(n+1)} \diamond \ldots \diamond \mathbf{A}^{(N)}\right)\right) = R \tag{2.27}$$

for the model to be unique. Note that the rank of Khatri-Rao products is not affected by order in which the multiplications are carried out.

The generalized Kruskal condition (2.22) is a sufficient condition for uniqueness, whereas (2.27) is necessary for any value of R. Hence, the sufficient condition (2.22) implies the necessary condition (2.27) and the reverse is not true. Another advantage is due to the facility of testing the condition (2.27) in terms of computational complexity [87].

The condition (2.27) can be rewritten in an equivalent way as

$$\operatorname{rank}\left(\bigdiamond_{\substack{n=1\\n\neq x}}^{N} \mathbf{A}^{(n)}\right) = R, \qquad \text{for all } x \in \{1, \ldots, N\} \tag{2.28}$$

and it is implicit $\prod_{n=1,n\neq x}^{N} I_n \geq R$ for all $x \in \{1, \ldots, N\}$. Since $\operatorname{rank}(\mathbf{A} \diamond \mathbf{B}) \leq \operatorname{rank}(\mathbf{A} \otimes \mathbf{B}) = \operatorname{rank}(\mathbf{A}) \operatorname{rank}(\mathbf{B})$, it can be generalized for several Khatri-Rao products which results in: $\operatorname{rank}\left(\mathbf{A}^{(1)} \diamond \ldots \diamond \mathbf{A}^{(N)}\right) \leq \operatorname{rank}\left(\mathbf{A}^{(1)}\right) \ldots \operatorname{rank}\left(\mathbf{A}^{(N)}\right)$. Thus, (2.28) leads to

$$\prod_{\substack{n=1\\n\neq x}}^{N} \operatorname{rank}\left(\mathbf{A}^{(n)}\right) \geq R, \qquad \text{for all } x \in \{1, \ldots, N\}. \tag{2.29}$$

Note that (2.28) implies (2.29) and the reverse statement is not true.

From (2.27)-(2.29), it is clear that an all-zero column in one of the matrix factors implies nonuniqueness of the decomposition and in this sense, $k_{\mathbf{A}^{(n)}} \geq 1$ for all $n \in \{1, \ldots, N\}$ is essential for the CP uniqueness [78]. As the condition in (2.27) corresponds to Khatri-Rao products between $N-1$ matrices be full column-rank, implying $R \leq I_1 \ldots I_{n-1} I_{n+1} \ldots I_N$, and by the k-rank definition, we have that $k\left(\mathbf{A}^{(1)} \diamond \ldots \diamond \mathbf{A}^{(n-1)} \diamond \mathbf{A}^{(n+1)} \diamond \ldots \diamond \mathbf{A}^{(N)}\right) = R$ for any n. From Lemmas 2.1 and 2.2, we propose the theorem below which leads to satisfy both conditions (2.27) and (2.28).

Theorem 2.2. *Recall the CP model of an N-th order array $\mathcal{X} = \left[\left[\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(N)}\right]\right] \in \mathbb{C}^{I_1 \times \ldots \times I_N}$, with $\mathbf{A}^{(n)} \in \mathcal{C}^{I_n \times R}$ for $n = \{1, \ldots, N\}$. If $\mathbf{A}^{(n)}$ has no zero columns for $n \in \{1, \ldots, N\}$ and the*

condition

$$\sum_{\substack{n=1 \\ n\neq x}}^{N} k_{\mathbf{A}^{(n)}} \geq R + (N-2), \qquad \textit{for all } x \in \{1,\ldots,N\}. \tag{2.30}$$

holds, the loading matrices $\mathbf{A}^{(n)}$ *for* $n \in \{1,\ldots,N\}$ *are unique up to permutation and scaling.*

All uniqueness results discussed until now were based on the *essential* uniqueness of the CP model, which means that all matrix factors of the model are uniquely determined up to permutation and scaling of columns. According to (2.30), when a third-order CP model has at least two collinear loading matrices, i.e. k-rank of each matrix equals to 1, the (essential) uniqueness model is no longer guaranteed. Generalizing to an N-th order tensor, when the N-th order CP model has $N-1$ component matrices with collinear columns, the essential uniqueness of this model is not achieved. However, the uniqueness for some loading matrices can be still ensured, which is known as *partial* uniqueness.

The partial uniqueness has been discussed since the standard CP decomposition by Harshman in [51, 84] and has received more attention for CP with a fixed pattern of linear dependencies in the loading vectors [94, 95, 58, 96, 97, 98, 70]. In some practical applications, we are interested in the uniqueness of one particular component matrix [94, 58, 70] which is called *uni-mode* uniqueness [70].

Now, let us present two following theorems concerning the uni-mode uniqueness of the third-order CP model derived in [70]. Guo *et al.* have proposed sufficient conditions to ensure the essential uniqueness of one loading matrix, i.e. one matrix factor can be uniquely identified.

Theorem 2.3. [70] *Recall the CP model of a three-way array* $\mathcal{X} = \big[\big[\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{A}^{(3)}\big]\big] \in \mathbb{C}^{I_1 \times I_2 \times I_3}$, *with* $\mathbf{A}^{(n)} \in \mathcal{C}^{I_n \times R}$ *for* $n = \{1,2,3\}$. *If* $\mathbf{A}^{(1)}$ *has no zero columns and the condition*

$$\operatorname{rank}\big(\mathbf{A}^{(1)}\big) + k_{\mathbf{A}^{(2)}} + k_{\mathbf{A}^{(3)}} \geq 2R + 2 \tag{2.31}$$

holds, the first mode loading $\mathbf{A}^{(1)}$ *is unique up to permutation and scaling of the columns.*

If $\mathbf{A}^{(1)}$ is full column-rank implying $k_{\mathbf{A}^{(1)}} = \operatorname{rank}\big(\mathbf{A}^{(1)}\big)$, the condition in (2.31) becomes identical to Kruskal's condition which leads to the essential uniqueness of all the loading matrices $\mathbf{A}^{(1)}$, $\mathbf{A}^{(2)}$ and $\mathbf{A}^{(3)}$ and not only for $\mathbf{A}^{(1)}$ when $k_{\mathbf{A}^{(1)}} < \operatorname{rank}\big(\mathbf{A}^{(1)}\big)$. Remark that the rank instead of the k-rank of $\mathbf{A}^{(1)}$ makes the condition (2.31) less restrictive than Kruskal's condition in the sense that even when $\mathbf{A}^{(1)}$ has collinear columns implying $k_{\mathbf{A}^{(1)}} = 1$, $\operatorname{rank}\big(\mathbf{A}^{(1)}\big)$ can be greater than 1.

It is noticed in [70] that the condition (2.31) becomes more relaxed when $\mathbf{A}^{(2)}$ and $\mathbf{A}^{(3)}$ are not full rank, i.e. when $k_{\mathbf{A}^{(2)}} < \operatorname{rank}\big(\mathbf{A}^{(2)}\big)$ and $k_{\mathbf{A}^{(3)}} < \operatorname{rank}\big(\mathbf{A}^{(3)}\big)$ respectively. It is presented in the next theorem as follows.

Theorem 2.4. [70] *Recall the CP model of a three-way array* $\mathcal{X} = \big[\big[\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{A}^{(3)}\big]\big] \in \mathbb{C}^{I_1 \times I_2 \times I_3}$, *with* $\mathbf{A}^{(n)} \in \mathcal{C}^{I_n \times R}$ *for* $n = \{1,2,3\}$. *If* $\mathbf{A}^{(1)}$ *has no zero columns,* $\mathbf{A}^{(2)}$ *and* $\mathbf{A}^{(3)}$ *are not full rank and the condition*

$$\operatorname{rank}\big(\mathbf{A}^{(1)}\big) + k_{\mathbf{A}^{(2)}} + k_{\mathbf{A}^{(3)}} \geq 2R + 1 \tag{2.32}$$

is satisfied, then the first mode loading $\mathbf{A}^{(1)}$ *is unique up to permutation and scaling of the columns.*

Note that both conditions (2.31) and (2.32) lead to $k_{\mathbf{A}^{(n)}} \geq 2$ for $n = \{1, 2, 3\}$ since $\min(I_n, R) \geq \text{rank}\big(\mathbf{A}^{(n)}\big) \geq k_{\mathbf{A}^{(n)}}$. It implies that all loading matrices do not have collinear columns. From the same idea employed in [67] which has been previously discussed, we can generalize the last two theorems proposed in [70] for an N-th order CP decomposition by the concatenation of $N-2$ component matrices in one matrix and using Lemma 2.1. We obtain

$$\text{rank}\big(\mathbf{A}^{(1)}\big) + \sum_{n=2}^{N} k_{\mathbf{A}^{(n)}} \geq 2R + (N-1) \tag{2.33}$$

and when $\mathbf{A}^{(n)}$ for $n = \{2, \ldots, N\}$ are not full column-rank

$$\text{rank}\big(\mathbf{A}^{(1)}\big) + \sum_{n=2}^{N} k_{\mathbf{A}^{(n)}} \geq 2R + (N-2). \tag{2.34}$$

Observe that these extensions of uniqueness condition imply only the essential uniqueness of the first mode $\mathbf{A}^{(1)}$.

2.3.2 PARATUCK-(N_1, N) decomposition

First, let us give an overview about the standard PARATUCK-2 model and present our proposed model in the sequence. The PT-2 model was introduced by Harshman and Lundy in psychometrics [64], this model can be viewed as a general version of the well-known CP model which incorporates interacting dimensions. This model combines some properties of both CP and TUCKER-2 [99] models, hence the name PARATUCK-2.

Recalling the TUCKER-2 and PT-2 models respectively by

$$x_{i_1,i_2,i_3} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} g_{r_1,r_2,i_3}\, a^{(1)}_{i_1,r_1}\, a^{(2)}_{i_2,r_2}, \tag{2.35}$$

$$x_{i_1,i_2,i_3} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} c_{r_1,r_2}\, a^{(1)}_{i_1,r_1}\, a^{(2)}_{i_2,r_2}\, \phi^{(1)}_{r_1,i_3}\, \phi^{(2)}_{r_2,i_3}, \tag{2.36}$$

note that the PT-2 model can be rewritten as a constrained TUCKER-2 model with the core tensor given by: $g_{r_1,r_2,i_3} = c_{r_1,r_2}\phi^{(1)}_{r_1,i_3}\phi^{(2)}_{r_2,i_3}$. The restricted structure of the core compared to TUCKER-2 retains uniqueness properties [64, 100]. The Tucker likeness due to the insertion of $c_{r_1 r_2}$, the second mode of both factor matrices $\mathbf{A}^{(1)} \in \mathbb{C}^{I_1 \times R_1}$ and $\mathbf{A}^{(2)} \in \mathbb{C}^{I_2 \times R_2}$ does not require to have the same size [64, 100].

Taking into account theses advantages of the PT-2 model, we proposed a new constrained tensor model that generalizes the PT-2 model, named as PT-(N_1, N). Given an N-th order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \ldots \times I_N}$, the PT-$(N_1, N)$ model of $\mathcal{X}$, with $N > N_1$, is defined in scalar form by the following expression [66]:

$$x_{i_1,\ldots,i_{N_1},\ldots,i_N} = \sum_{r_1=1}^{R_1} \cdots \sum_{r_{N_1}=1}^{R_{N_1}} c_{r_1,\ldots,r_{N_1},i_{N_1+2},\ldots,i_N} \prod_{n=1}^{N_1} a^{(n)}_{i_n,r_n}\, \phi^{(n)}_{r_n,i_{N_1+1}}, \tag{2.37}$$

where $a^{(n)}_{i_n,r_n}$ and $\phi^{(n)}_{r_n,i_{N_1+1}}$ are, respectively, the entries of the factor matrix $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R_n}$ and the weighting matrix $\mathbf{\Phi}^{(n)} \in \mathbb{C}^{R_n \times I_{N_1+1}}$ for $n = 1, \ldots, N_1$.

This model can be interpreted as a transformation of the input tensor $\mathcal{C} \in \mathbb{C}^{R_1 \times \cdots \times R_{N_1} \times I_{N_1+2} \times \cdots \times I_N}$ via its multiplication by the factor matrices $\mathbf{A}^{(n)}$, $n = 1, \ldots, N_1$, along its first N_1 modes, combined with an n-th weighting matrix $\mathbf{\Phi}^{(n)}$ $(n = 1, \ldots, N_1)$ relatively to the mode-(N_1+1) of the transformed tensor $\mathcal{X}$.

From the notation of Kronecker and Khatri-Rao products, we can also represent the PT-(N_1, N) model by the following matrix unfoldings

$$\begin{aligned}
\mathbf{X}_{I_1 \ldots I_{N_1} \times I_{N_1+1} \ldots I_N} &= \left(\mathbf{A}^{(1)} \otimes \ldots \otimes \mathbf{A}^{(N_1)}\right) \left(\left(\mathbf{\Phi}^{(1)} \diamond \ldots \diamond \mathbf{\Phi}^{(N_1)}\right)^{\mathrm{T}} \diamond \mathbf{C}_{I_{N_1+2} \ldots I_N \times R_1 \ldots R_{N_1}}\right)^{\mathrm{T}} \\
\mathbf{X}_{I_{N_1+1} \times I_1 \ldots I_{N_1} I_{N_1+2} \ldots I_N} &= \left(\mathbf{\Phi}^{(1)} \diamond \ldots \diamond \mathbf{\Phi}^{(N_1)}\right)^{\mathrm{T}} \left(\left(\mathbf{A}^{(1)} \otimes \ldots \otimes \mathbf{A}^{(N_1)}\right) \diamond \mathbf{C}_{I_{N_1+2} \ldots I_N \times R_1 \ldots R_{N_1}}\right)^{\mathrm{T}} \\
\mathbf{X}_{I_{N_1+2} \ldots I_N \times I_1 \ldots I_{N_1+1}} &= \mathbf{C}_{I_{N_1+2} \ldots I_N \times R_1 \ldots R_{N_1}} \left(\left(\mathbf{A}^{(1)} \otimes \ldots \otimes \mathbf{A}^{(N_1)}\right) \diamond \left(\mathbf{\Phi}^{(1)} \diamond \ldots \diamond \mathbf{\Phi}^{(N_1)}\right)^{\mathrm{T}}\right)^{\mathrm{T}},
\end{aligned} \tag{2.38}$$

where $\mathbf{C}_{I_{N_1+2} \ldots I_N \times R_1 \ldots R_{N_1}}$ denotes a $(I_{N_1+2} \ldots I_N \times R_1 \ldots R_{N_1})$- matrix unfolding of the tensor $\mathcal{C}$. Remark that there are many forms to write the matrix unfoldings and each expression above isolates as the left factor, respectively, the factor matrices $\mathbf{A}^{(n)}$ $(n = 1, \ldots, N_1)$, the weighting matrices $\mathbf{\Phi}^{(n)}$ $(n = 1, \ldots, N_1)$ and the core tensor $\mathcal{C}$ represented by a matrix unfolding.

Special cases:

We can list two cases of interest derived from the PT-(N_1, N) model: the standard PT-2 and the PT-(2,4) models, the last one will be useful to represent a particular application. From (2.37) then we obtain:

- The standard PT-2 model for $N_1 = 2$ and $N = 3$:

$$x_{i_1,i_2,i_3} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} c_{r_1,r_2} a^{(1)}_{i_1,r_1} a^{(2)}_{i_2,r_2} \phi^{(1)}_{r_1,i_3} \phi^{(2)}_{r_2,i_3}, \tag{2.39}$$

- The following PT-(2,4) model for $N_1 = 2$ and $N = 4$:

$$x_{i_1,i_2,i_3,i_4} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} c_{r_1,r_2,i_4} a^{(1)}_{i_1,r_1} a^{(2)}_{i_2,r_2} \phi^{(1)}_{r_1,i_3} \phi^{(2)}_{r_2,i_3}. \tag{2.40}$$

The uniqueness results for the PT-2 model have not been widely investigated because the PT-2 model has a more complex structure and has been less employed in applications. Now, we show the uniqueness results for the PT-2 model, analyze the conditions for the PT-(2,4) model and its generalization for the PT-(N_1, N) model.

Uniqueness results of PARATUCK-2 model

In 1996, Harshman and Lundy have provided uniqueness analysis for the PT-2 model subject to $R_1 = R_2$ and $c_{r_1,r_2} \neq 0$ for all $r_1 \in \{1, \ldots, R_1\}$, $r_2 \in \{1, \ldots, R_2\}$ considering the two following cases: the general PT-2 model and the symmetrically weighted version (i.e., when $\mathbf{\Phi}^{(1)} = \mathbf{\Phi}^{(2)}$) [64]. The purpose of these restrictions is explained by the simplification of the proof in this work.

Despite having considered $R_1 = R_2$ in this proof, they have presented as well an experimental evidence of uniqueness being obtainable even when $R_1 \neq R_2$. Let us formulate appropriately the sufficient uniqueness condition proved in [64] as follows.

Theorem 2.5. [64] *Consider the PT-2 model of a three-way array given by* (2.39) *with* $R_1 = R_2$ *and* $c_{r_1,r_2} \neq 0$ *for all index values. If* $\mathbf{A}^{(1)} \in \mathbb{C}^{I_1 \times R_1}$, $\mathbf{A}^{(2)} \in \mathbb{C}^{I_2 \times R_2}$ *and* $\mathbf{C} \in \mathbb{C}^{R_1 \times R_2}$ *are full column-rank, implying* $I_1, I_2 \geq R_1 = R_2$, *and there is an "adequate"* I_3, *then* $\mathbf{A}^{(1)}$ *and* $\mathbf{A}^{(2)}$ *are unique up to permutation and/or scaling of the columns.*

The term *adequate* was employed in [64] and its definition is not simple, they did not specify any rule to set I_3 needed for uniqueness of this model. They have proved the uniqueness for a minimum value of I_3 and have given examples for some values of $R_1 = R_2$. When $R_1 = R_2 = 2$ and $R_1 = R_2 = 3$ for a general PT-2 model, it is respectively required at least $I_3 = 9$ and $I_3 = 36$. When $R_1 = R_2 = 2$, $R_1 = R_2 = 3$ and $R_1 = R_2 = 4$ for the special case of PT-2 with $\mathbf{\Phi}^{(1)} = \mathbf{\Phi}^{(2)}$, at least $I_3 = 5$, $I_3 = 15$ and $I_3 = 35$, respectively.

Posteriorly, the PT-2 model has been exploited in chemometrics: by Bro [100] and in signal processing: by Kibangou and Favier [101], and by De Almeida *et al.* [59]. But new contributions on uniqueness of the PT-2 model were only derived in the last two works, both have investigated uniqueness in the context of their applications considering particular structures for the parameters of this decomposition.

In [101], they have written the output signal tensor as a constrained PT-2 model and thus, the uniqueness condition was analyzed by considering structural constraints such as Toeplitz and Vandermonde forms for some of its matrix factors. Thanks to several assumptions in the structure of the parameters was possible to ensure the uniqueness of a constrained version of PT-2 model. However, as theses assumptions are not easily exploited for a more general case, we restrict to present just the results proposed in [59].

Theorem 2.6. [59] *Consider the PT-2 model of a three-way array given by* (2.39) *and assume the perfect knowledge of* $\mathbf{\Phi}^{(1)}$, $\mathbf{\Phi}^{(2)}$ *and* $\mathbf{C}$. *If* $\mathbf{A}^{(1)}$, $\mathbf{A}^{(2)}$ *and* $\left(\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)}\right)^{\mathrm{T}}$ *are full column-rank, implying respectively* $I_1 \geq R_1$, $I_2 \geq R_2$ *and* $I_3 \geq R_1 R_2$, *then both factor matrices* $\mathbf{A}^{(1)}$ *and* $\mathbf{A}^{(2)}$ *are unique up to a scalar factor.*

Observe that Theorem 2.6 provides a sufficient condition to ensure the uniqueness of both matrix factors $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ imposing a building restriction over the weighting matrices $\mathbf{\Phi}^{(1)}$ and $\mathbf{\Phi}^{(2)}$ ($I_3 \geq R_1 R_2$). Contrarily, the sufficient condition proposed by Theorem 2.5 imposes mainly that $\mathbf{C}$ is full column-rank ($R_1 \geq R_2$) and also requires an enough value of I_3 not defined. In this sense, Theorem 2.6 becomes more advantageous. The assumption of knowledge of some parameters in both theorems is explained by the practical applications.

Uniqueness results of PARATUCK-(2,4) model

From (2.38), let us write the matrix unfoldings of PT-(2,4) model with $N_1 = 2$ and $N = 4$ as follows.

$$\begin{aligned}
\mathbf{X}_{I_1 I_2 \times I_3 I_4} &= \left(\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)}\right) \left(\left(\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)}\right)^{\mathrm{T}} \diamond \mathbf{C}_{I_4 \times R_1 R_2}\right)^{\mathrm{T}} \\
\mathbf{X}_{I_3 \times I_1 I_2 I_4} &= \left(\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)}\right)^{\mathrm{T}} \left(\left(\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)}\right) \diamond \mathbf{C}_{I_4 \times R_1 R_2}\right)^{\mathrm{T}} \\
\mathbf{X}_{I_4 \times I_1 I_2 I_3} &= \mathbf{C}_{I_4 \times R_1 R_2} \left(\left(\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)}\right) \diamond \left(\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)}\right)^{\mathrm{T}}\right)^{\mathrm{T}}.
\end{aligned} \tag{2.41}$$

One way to analyze the uniqueness conditions for this model is by means its matrix unfoldings. In our analysis, we assume the knowledge of both weighting matrices $\mathbf{\Phi}^{(1)} \in \mathbb{C}^{R_1 \times I_3}$ and $\mathbf{\Phi}^{(2)} \in \mathbb{C}^{R_2 \times I_3}$, and also of the core tensor $\mathcal{C} \in \mathbb{C}^{R_1 \times R_2 \times I_4}$. In our practical context, we are interested to guarantee the model uniqueness for the estimation of the factor matrices. As we want to determine only both matrices $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$, let us consider the first matrix unfolding of (2.41).

The study of uniqueness for the PT-(2,4) model proposed in this thesis is based on the one for the PT-2 model in [59]. Considering $\hat{\mathbf{A}}^{(1)}$ and $\hat{\mathbf{A}}^{(2)}$ as alternative solutions that satisfy (2.41), we can write $\hat{\mathbf{A}}^{(n)} = \mathbf{A}^{(n)}\,\mathbf{U}^{(n)}$ with $\mathbf{U}^{(n)} \in \mathbb{C}^{R_n \times R_n}$, $n = 1, 2$, non-singular matrices. Thus, the unfolded matrix $\mathbf{X}_{I_1 I_2 \times I_3 I_4}$ can be rewritten using the Kronecker property (A.4) as

$$\begin{aligned}
\left(\hat{\mathbf{A}}^{(1)} \otimes \hat{\mathbf{A}}^{(2)}\right) \left(\left(\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)}\right)^{\mathrm{T}} \diamond \mathbf{C}_{I_4 \times R_1 R_2}\right)^{\mathrm{T}} &= \left(\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)}\right) \left(\mathbf{U}^{(1)} \otimes \mathbf{U}^{(2)}\right) \\
&\quad \left(\left(\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)}\right)^{\mathrm{T}} \diamond \mathbf{C}_{I_4 \times R_1 R_2}\right)^{\mathrm{T}}.
\end{aligned} \tag{2.42}$$

From (2.42) the uniqueness of the PT-(2,4) model can be proved. It is necessary to remove any ambiguity caused by the nonsingular transformation matrices $\mathbf{U}^{(1)}$ and $\mathbf{U}^{(2)}$ to recover both matrices $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$. The next theorem shows a sufficient condition which ensures the uniqueness of $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$.

Theorem 2.7. *Consider the PT-(2,4) model of a four-way array given by* (2.40)*, suppose the perfect knowledge of* $\mathbf{\Phi}^{(1)}$*,* $\mathbf{\Phi}^{(2)}$ *and* $\mathcal{C}$*, and that* ${\mathbf{\Phi}^{(1)}}^{\mathrm{T}}$*,* ${\mathbf{\Phi}^{(2)}}^{\mathrm{T}}$ *and* $\mathbf{C}_{I_4 \times R_1 R_2}$ *have no zero-columns. If* $\mathbf{A}^{(1)}$ *and* $\mathbf{A}^{(2)}$ *are full column-rank implying* $I_1 \geq R_1$*,* $I_2 \geq R_2$*, and* $\mathbf{\Phi}^{(1)}$*,* $\mathbf{\Phi}^{(2)}$ *and* $\mathcal{C}$ *are chosen such that*

$$k\left(\left(\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)}\right)^{\mathrm{T}}\right) + k(\mathbf{C}_{I_4 \times R_1 R_2}) \geq R_1 R_2 + 1, \tag{2.43}$$

then both factor matrices $\mathbf{A}^{(1)}$ *and* $\mathbf{A}^{(2)}$ *are unique up to a factor scaling.*

Proof: If $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ are full column-rank, then the left-inverse of $\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)}$ exits and is unique. Assuming that ${\mathbf{\Phi}^{(1)}}^{\mathrm{T}}$, ${\mathbf{\Phi}^{(2)}}^{\mathrm{T}}$ and $\mathbf{C}_{I_4 \times R_1 R_2}$ have no zero-columns imply that the k-rank of these matrices is greater than or equal to 1. According to Lemma 2.2, if $\mathbf{\Phi}^{(1)}$, $\mathbf{\Phi}^{(2)}$ and $\mathcal{C}$ satisfy the inequality in (2.43), then $\left(\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)}\right)^{\mathrm{T}} \diamond \mathbf{C}_{I_4 \times R_1 R_2}$ is full column-rank, which leads to the existence of the right-inverse of $\left(\left(\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)}\right)^{\mathrm{T}} \diamond \mathbf{C}_{I_4 \times R_1 R_2}\right)^{\mathrm{T}}$. Finally, (2.42) gives

$$\mathbf{U}^{(1)} \otimes \mathbf{U}^{(2)} = \mathbf{I}_{R_1 R_2}. \tag{2.44}$$

From (2.44), the only solution happens when both matrices $\mathbf{U}^{(1)}$ and $\mathbf{U}^{(2)}$ are identity matrices up to scalar factors that compensate each other, i.e. $\mathbf{U}^{(1)} = \alpha\,\mathbf{I}_{R_1}$ and $\mathbf{U}^{(2)} = 1/\alpha\,\mathbf{I}_{R_2}$, which leads to

$$\hat{\mathbf{A}}^{(1)} = \alpha\,\mathbf{A}^{(1)}, \qquad \hat{\mathbf{A}}^{(2)} = 1/\alpha\,\mathbf{A}^{(2)} \tag{2.45}$$

and concludes the proof. ∎

Remark 2.1. *The uniqueness of PT-(2,4) model can be achieved even when* $\left(\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)}\right)^{\mathrm{T}}$ *is not full column-rank, if* $\mathcal{C}$ *is appropriately chosen satisfying* $\mathbf{C}_{I_4 \times R_1 R_2}$ *full column-rank. Thus, the extension of PT-2 model to PT-(2,4) model provides a flexibility of the uniqueness condition given by Theorem* (2.6)*.*

The last result is further generalized for any N_1 and N with $N_1 < N$, which leads to a sufficient uniqueness condition for the PT-(N_1, N) model. The uniqueness analyze of this model can be deduced, analogously to the previous theorems, from the first matrix unfolding $\mathbf{X}_{I_1 \ldots I_{N_1} \times I_{N_1+1} \ldots I_N}$ in (2.38).

Theorem 2.8. *Consider the PT-*(N_1, N) *model of a* N*-way array given by* (2.37) *with* $N_1 < N$*, suppose the perfect knowledge of all weighting matrices* $\mathbf{\Phi}^{(n)}$ $(n = 1, \ldots, N_1)$ *and core tensor* $\mathcal{C}$*, and* $\mathbf{\Phi}^{(n)^{\mathrm{T}}}$ $(n = 1, \ldots, N_1)$ *and* $\mathbf{C}_{I_{N_1+2} \ldots I_N \times R_1 \ldots R_{N_1}}$ *have no zero-columns. If all factor matrices* $\mathbf{A}^{(n)}$ $(n = 1, \ldots, N_1)$ *are full column-rank implying* $I_n \geq R_n$ *for* $n = 1, \ldots, N_1$*, and* $\mathbf{\Phi}^{(n)}$ $(n = 1, \ldots, N_1)$ *and* $\mathcal{C}$ *are chosen such that*

$$k\Big(\left(\mathbf{\Phi}^{(1)} \diamond \ldots \diamond \mathbf{\Phi}^{(N_1)}\right)^{\mathrm{T}}\Big) + k\big(\mathbf{C}_{I_{N_1+2} \ldots I_N \times R_1 \ldots R_{N_1}}\big) \geq \prod_{n=1}^{N_1} R_n + 1, \tag{2.46}$$

then all factor matrices $\mathbf{A}^{(n)}$ $(n = 1, \ldots, N_1)$ *are unique up to a scaling factor.*

Uniqueness results derived for the PT-(N_1, N) model correspond to sufficient conditions and consider the knowledge of all weighting matrices and of the core tensor. In order to exploit the uniqueness results concerning the CP model, let us rewrite the PT-(N_1, N) model as a constrained CP as follows.

Let us consider a third order CP decomposition $\mathcal{X} = \big[\big[\mathbf{B}^{(1)}, \mathbf{B}^{(2)}, \mathbf{B}^{(3)}\big]\big] \in \mathbb{C}^{J_1 \times J_2 \times J_3}$, where each loading matrices are given by

$$\begin{aligned} \mathbf{B}^{(1)} &= \mathbf{A}^{(1)} \otimes \ldots \otimes \mathbf{A}^{(N_1)} \in \mathbb{C}^{J_1 \times R} \\ \mathbf{B}^{(2)} &= \left(\mathbf{\Phi}^{(1)} \diamond \ldots \diamond \mathbf{\Phi}^{(N_1)}\right)^{\mathrm{T}} \in \mathbb{C}^{J_2 \times R}, \\ \mathbf{B}^{(3)} &= \mathbf{C}_{I_{N_1+2} \ldots I_N \times R_1 \ldots R_{N_1}} \in \mathbb{C}^{J_3 \times R} \end{aligned} \tag{2.47}$$

and the following correspondences $J_1 = I_1 \ldots I_{N_1}$, $J_2 = I_{N_1+1}$, $J_3 = I_{N_1+2} \ldots I_N$ and $R = R_1 \ldots R_{N_1}$. According to the Kronecker product, note that $\operatorname{rank}(\mathbf{A}^{(1)} \otimes \ldots \otimes \mathbf{A}^{(N_1)}) = \prod_{n=1}^{N_1} \operatorname{rank}(\mathbf{A}^{(n)})$. If we only consider the uniqueness of all matrix factors for practical reasons, we can directly apply Theorem 2.3 and deduce a wider condition by the next theorem.

Theorem 2.9. *Consider the PT-(N_1, N) model given by* (2.37) *and assume that all matrix factors $\mathbf{A}^{(n)}$ for $n = 1, \ldots, N_1$ have no zero-columns. If the following condition*

$$\prod_{n=1}^{N_1} \operatorname{rank}\left(\mathbf{A}^{(n)}\right) + k\left(\left(\mathbf{\Phi}^{(1)} \diamond \ldots \diamond \mathbf{\Phi}^{(N_1)}\right)^{\mathrm{T}}\right) + k\left(\mathbf{C}_{I_{N_1+2}\ldots I_N \times R_1 \ldots R_{N_1}}\right) \geq 2\left(\prod_{n=1}^{N_1} R_n\right) + 2 \tag{2.48}$$

holds, $\mathbf{A}^{(1)} \otimes \ldots \otimes \mathbf{A}^{(N_1)}$ is unique up to permutation and scaling of the columns.

Remark 2.2.

- *As discussed above for the CP results, the condition* (2.48) *leads to $k(\mathbf{C}_{I_{N_1+2}\ldots I_N \times R_1 \ldots R_{N_1}}) \geq 2$, $\operatorname{rank}(\mathbf{A}^{(n)}) \geq k(\mathbf{A}^{(n)}) \geq 2$ for a given $n \in \{1, \ldots, N_1\}$ and $k\left(\left(\mathbf{\Phi}^{(1)} \diamond \ldots \diamond \mathbf{\Phi}^{(N_1)}\right)^{\mathrm{T}}\right) \geq 2$. Therefore, $\mathbf{A}^{(n)}$ for a given $n \in \{1, \ldots, N_1\}$, $\left(\mathbf{\Phi}^{(1)} \diamond \ldots \diamond \mathbf{\Phi}^{(N_1)}\right)^{\mathrm{T}}$ and $\mathbf{C}_{I_{N_1+2}\ldots I_N \times R_1 \ldots R_{N_1}}$ do not have collinear columns.*
- *The sufficient condition* (2.48) *is more restrictive than the one given in Theorem 2.8 when all matrix factors $\mathbf{A}^{(n)}$, $n = 1, \ldots, N_1$, are full column-rank.*

If we apply Theorem 2.9 to the PT-(2,4) model ($N_1 = 2$ and $N = 4$), the condition (2.48) becomes

$$\operatorname{rank}\left(\mathbf{A}^{(1)}\right) \operatorname{rank}\left(\mathbf{A}^{(2)}\right) + k\left(\left(\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)}\right)^{\mathrm{T}}\right) + k(\mathbf{C}_{I_4 \times R_1 R_2}) \geq 2R_1R_2 + 2, \tag{2.49}$$

which represents an alternative condition when the matrix factors $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ are not full column-rank.

We can also propose an uniqueness condition for the PT-(N_1, N) model applying the condition (2.26) to the constrained 3-CP model with factor matrices defined in (2.47). If the Khatri-Rao product of any two matrix factors is full column-rank, i.e. the following matrices

$$\begin{aligned}
&\left(\mathbf{A}^{(1)} \otimes \ldots \otimes \mathbf{A}^{(N_1)}\right) \diamond \left(\mathbf{\Phi}^{(1)} \diamond \ldots \diamond \mathbf{\Phi}^{(N_1)}\right)^{\mathrm{T}} \in \mathbb{C}^{I_1 \ldots I_{N_1+1} \times R_1 \ldots R_{N_1}} \\
&\left(\mathbf{A}^{(1)} \otimes \ldots \otimes \mathbf{A}^{(N_1)}\right) \diamond \mathbf{C}_{I_{N_1+2}\ldots I_N \times R_1 \ldots R_{N_1}} \in \mathbb{C}^{I_1 \ldots I_{N_1} I_{N_1+2} \ldots I_N \times R_1 \ldots R_{N_1}} \\
&\left(\mathbf{\Phi}^{(1)} \diamond \ldots \diamond \mathbf{\Phi}^{(N_1)}\right)^{\mathrm{T}} \diamond \mathbf{C}_{I_{N_1+2}\ldots I_N \times R_1 \ldots R_{N_1}} \in \mathbb{C}^{I_{N_1+1} \ldots I_N \times R_1 \ldots R_{N_1}}
\end{aligned} \tag{2.50}$$

are full column-rank, then the essential uniqueness of the third-order CP model with the matrix factors given in (2.47) is ensured. Now, let us enunciate the next theorem applying the equivalent condition (2.30).

Theorem 2.10. *Consider the PT-(N_1, N) model given by* (2.37). *If all following conditions*

$$\begin{aligned}
k\left(\mathbf{A}^{(1)} \otimes \ldots \otimes \mathbf{A}^{(N_1)}\right) + k\left(\left(\mathbf{\Phi}^{(1)} \diamond \ldots \diamond \mathbf{\Phi}^{(N_1)}\right)^{\mathrm{T}}\right) &\geq \prod_{n=1}^{N_1} R_n + 1 \\
k\left(\mathbf{A}^{(1)} \otimes \ldots \otimes \mathbf{A}^{(N_1)}\right) + k\left(\mathbf{C}_{I_{N_1+2}\ldots I_N \times R_1 \ldots R_{N_1}}\right) &\geq \prod_{n=1}^{N_1} R_n + 1 \\
k\left(\left(\mathbf{\Phi}^{(1)} \diamond \ldots \diamond \mathbf{\Phi}^{(N_1)}\right)^{\mathrm{T}}\right) + k\left(\mathbf{C}_{I_{N_1+2}\ldots I_N \times R_1 \ldots R_{N_1}}\right) &\geq \prod_{n=1}^{N_1} R_n + 1
\end{aligned} \tag{2.51}$$

hold, then $\mathbf{A}^{(1)} \otimes \ldots \otimes \mathbf{A}^{(N_1)}$, $\left(\mathbf{\Phi}^{(1)} \diamond \ldots \diamond \mathbf{\Phi}^{(N_1)}\right)^{\mathrm{T}}$ and $\mathbf{C}_{I_{N_1+2}\ldots I_N \times R_1 \ldots R_{N_1}}$ are unique up to permutation and scaling of the columns.

Remark 2.3.

- *The conditions* (2.51) *imply that* $\mathbf{C}_{I_{N_1+2}\ldots I_N \times R_1 \ldots R_{N_1}}$, *all matrix factors* $\mathbf{A}^{(n)}$ *and* $\left(\mathbf{\Phi}^{(1)} \diamond \ldots \diamond \mathbf{\Phi}^{(N_1)}\right)^{\mathrm{T}}$ *have no zero-columns.*
- *If any two matrices of* (2.47) *are full column-rank, the condition* (2.51) *given in Theorem 2.10 is more relaxed than the one* (2.48) *given in Theorem 2.9.*

Now, let us resume the uniqueness results obtained from Theorems 2.8, 2.9 and 2.10 for $N_1 = 2$ and $N = 4$ in Table 3.1. According to Table 3.1, assuming that $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ are full column-rank, the second condition becomes more restrictive than the first condition and the third condition becomes equal to the first condition. If both matrices $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ are not full column-rank, the first condition is not achieved. By the k-rank definition, we obtain $\operatorname{rank}\left(\mathbf{A}^{(1)}\right)\operatorname{rank}\left(\mathbf{A}^{(2)}\right) = \operatorname{rank}\left(\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)}\right) \geq k\left(\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)}\right)$ thus the third condition is more flexible than the second condition.

Observe that the condition of Theorem 2.8 is derived taking into account one unfolded matrix $\mathbf{X}_{I_1 I_2 \times I_3 I_4}$. If we analyze the uniqueness of this model using the same reasoning for the other unfolded matrices given in (2.41), i.e. $\mathbf{X}_{I_3 \times I_1 I_2 I_4}$ and $\mathbf{X}_{I_4 \times I_1 I_2 I_3}$, then the two inequalities of the third condition will be also achieved.

Table 2.1: Summary of the uniqueness results for PT-(2,4) model.

Theorem	**Uniqueness condition**	**Tensor model**	**Base reference**
2.8	$k\Big(\big(\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)}\big)^{\mathrm{T}}\Big) + k(\mathbf{C}_{I_4 \times R_1 R_2}) \geq R_1 R_2 + 1$ $\mathbf{A}^{(1)}$, $\mathbf{A}^{(2)}$ full column-rank	PT-2 model	[59]
2.9	$\operatorname{rank}\big(\mathbf{A}^{(1)}\big) \operatorname{rank}\big(\mathbf{A}^{(2)}\big) + k\Big(\big(\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)}\big)^{\mathrm{T}}\Big) + k(\mathbf{C}_{I_4 \times R_1 R_2}) \geq 2R_1 R_2 + 2$	CP model	[70]
2.10	$k\big(\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)}\big) + k\Big(\big(\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)}\big)^{\mathrm{T}}\Big) \geq R_1 R_2 + 1$ $k\big(\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)}\big) + k(\mathbf{C}_{I_4 \times R_1 R_2}) \geq R_1 R_2 + 1$ $k\Big(\big(\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)}\big)^{\mathrm{T}}\Big) + k(\mathbf{C}_{I_4 \times R_1 R_2}) \geq R_1 R_2 + 1$	CP model	[69]

Chapter 3

MIMO systems with joint multiplexing and spreading operations

A tensor space-time (TST) coding for MIMO wireless communication systems [65, 66] is introduced in this chapter. In the context of multicarrier systems, we present a MIMO wireless communication system with space-time-frequency (STF) spreading-multiplexing proposed in [62]. The expressions of the transmit and receive signals are established. We analyze the performance of both systems regarding the diversity of transmitted information, allowing to deduce the maximum gain for each system. From the results derived in Chapter 2, we also propose uniqueness conditions for both systems.

3.1 Tensor Space-Time (TST) system

Consider a MIMO wireless communication system with M transmit antennas and K receive antennas, and we denote by $s_{n,r}$ the n-th symbol of the r-th data stream, each data stream $(r = 1, ..., R)$ being composed of N information symbols.

The TST coding allows spreading and multiplexing the transmitted symbols, belonging to R data streams, in both space (by employing multiple antennas) and time (by transmission over time blocks and time spreading during several chip periods) domains, through the employment of a third-order code tensor admitting transmit antenna, data stream and chip as modes, $\mathcal{W} \in \mathbb{C}^{M \times R \times J}$, and two allocation matrices that allocate transmit antennas and data streams to each block, $\mathbf{C}^{(\mathbf{H})} \in \mathbb{R}^{P \times M}$ and $\mathbf{C}^{(\mathbf{S})} \in \mathbb{R}^{P \times R}$ respectively.

The transmission is assumed to be decomposed into P data blocks, each block being formed of N symbol periods. At each symbol period n of the p-th block, the transceiver transmits a linear combination of the n-th symbols of the data streams determined by the *stream-to-block* allocation matrix $\mathbf{C}^{(\mathbf{S})}$, across a set of transmit antennas fixed by the *antenna-to-block* allocation matrix $\mathbf{C}^{(\mathbf{H})}$, both matrices are composed uniquely of 1's and 0's. It is important to notice that, during each block p, a different set of data streams can be sent using a different set of transmit antennas, these two sets depending on the row vectors $\mathbf{C}^{(\mathbf{S})}_{p.}$ and $\mathbf{C}^{(\mathbf{H})}_{p.}$ of the two allocation matrices, respectively.

Each symbol $s_{n,r}$ is replicated several times after multiplication by a three-dimensional

spreading code $w_{m,r,j}$, in such a way that the signal transmitted from the m-th antenna during the n-th symbol period of the p-th block, and associated with the j-th chip, is given by

$$u_{m,n,p,j} = \sum_{r=1}^{R} w_{m,r,j}\, s_{n,r}\, c_{p,m}^{(\mathbf{H})}\, c_{p,r}^{(\mathbf{S})} = \sum_{r=1}^{R} g_{m,r,p,j}\, s_{n,r} \tag{3.1}$$

with

$$g_{m,r,p,j} \stackrel{\Delta}{=} w_{m,r,j}\, c_{p,m}^{(\mathbf{H})}\, c_{p,r}^{(\mathbf{S})}. \tag{3.2}$$

Equation (3.1) defines the transmit processing composed of three blocks carrying the following operations: data stream allocation, code tensor and transmit antenna allocation.

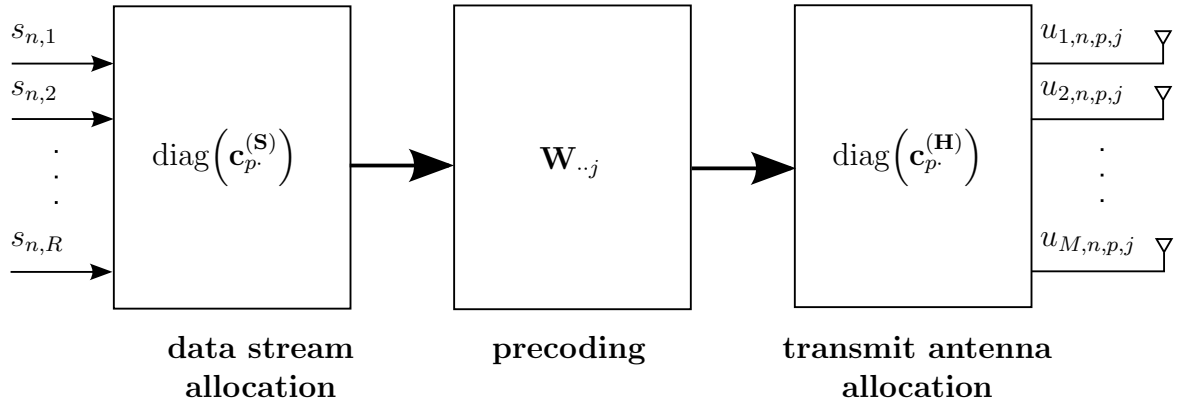


Figure 3.1: Transmit processing based on the TST coding with resource allocations.

In Figure 3.1, the functioning of this transceiver is illustrated for the p-th block and the j-th chip. The first and third black boxes $\left(\mathrm{diag}\left(\mathbf{c}_{p\cdot}^{(\mathbf{H})}\right)\right.$ and $\left.\mathrm{diag}\left(\mathbf{c}_{p\cdot}^{(\mathbf{S})}\right)\right)$ select the data streams to be sent and the transmit antennas to be used for transmission during the p-th block, respectively, whereas the second black box ($\mathbf{W}_{\cdot\cdot j}$) spreads the selected data streams on the selected antennas to deliver the following matrix of coded signals:

$$\mathbf{U}_{\cdot\cdot p,j} = \mathbf{G}_{\cdot\cdot p,j}\, \mathbf{S}^{\mathrm{T}} \in \mathbb{C}^{M\times N}, \tag{3.3}$$

where $\mathbf{S} \in \mathbb{C}^{N\times R}$ and $\mathbf{G}_{\cdot\cdot p,j} \in \mathbb{C}^{M\times R}$ can be deduced from (3.2)

$$\mathbf{G}_{\cdot\cdot p,j} = \mathrm{diag}\left(\mathbf{c}_{p\cdot}^{(\mathbf{H})}\right)\, \mathbf{W}_{\cdot\cdot j}\, \mathrm{diag}\left(\mathbf{c}_{p\cdot}^{(\mathbf{S})}\right). \tag{3.4}$$

In the noiseless case of flat Rayleigh fading channel, the signal received at the k-th antenna during the j-th chip period of the n-th symbol period of the p-th block, is given by

$$x_{k,n,p,j} = \sum_{m=1}^{M} h_{k,m}\, u_{m,n,p,j} = \sum_{m=1}^{M}\sum_{r=1}^{R} g_{m,r,p,j}\, h_{k,m}\, s_{n,r}. \tag{3.5}$$

The fading coefficients $h_{k,m}$ between transmit antenna (m) and receive antenna (k) are assumed to be independent and identically distributed (i.i.d.) zero-mean complex Gaussian random variables. They are also assumed to be constant during at least P blocks.

Comparing (3.2) and (3.5) with (2.40), we have the following correspondences

$$\begin{aligned}(I_1, I_2, I_3, I_4, R_1, R_2) &\longleftrightarrow (K, N, P, J, M, R)\,,\\ \left(\mathcal{C}, \mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{\Phi}^{(1)}, \mathbf{\Phi}^{(2)}\right) &\longleftrightarrow \left(\mathcal{W}, \mathbf{H}, \mathbf{S}, \mathbf{C}^{(\mathbf{H})^{\mathrm{T}}}, \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right).\end{aligned} \tag{3.6}$$

Therefore, the received signal $x_{k,n,p,j}$ satisfies the PT-(2,4) model given by (2.40).

From the n-mode product definition (2.13), (3.1) and (3.5), we can write the following fourth-order tensor $\mathcal{X} \in \mathbb{C}^{K\times N\times P\times J}$ as

$$\mathcal{X} = \mathcal{U} \times_1 \mathbf{H} = (\mathcal{G} \times_2 \mathbf{S}) \times_1 \mathbf{H} = \mathcal{G} \times_1 \mathbf{H} \times_2 \mathbf{S}, \tag{3.7}$$

where $\mathcal{U} = \mathcal{G} \times_2 \mathbf{S} \in \mathbb{C}^{M\times N\times P\times J}$ denotes the tensor of transmitted signals and $\mathcal{G} \in \mathbb{C}^{M\times R\times P\times J}$ represent the tensor deduced from the combination between allocation matrices, $\mathbf{C}^{(\mathbf{H})}$ and $\mathbf{C}^{(\mathbf{H})}$, and code tensor $\mathcal{W}$. Analogously, the tensorial slice $\mathcal{X}_{..p.} \in \mathbb{C}^{K\times N\times J}$ of $\mathcal{X}$, containing all the signals received during the p-th block, can be written from (2.13), (2.16), (3.3) and (3.4) as

$$\begin{aligned}\mathcal{X}_{..p.} &= \mathcal{W} \times_1 \left(\mathbf{H}\,\mathrm{diag}\big(\mathbf{c}_{p.}^{(\mathbf{H})}\big)\right) \times_2 \left(\mathbf{S}\,\mathrm{diag}\big(\mathbf{c}_{p.}^{(\mathbf{S})}\big)\right)\\ &= \mathcal{G}_{..p.} \times_1 \mathbf{H} \times_2 \mathbf{S}\end{aligned} \tag{3.8}$$

with

$$\mathcal{G}_{..p.} = \mathcal{W} \times_1 \mathrm{diag}\big(\mathbf{c}_{p.}^{(\mathbf{H})}\big) \times_2 \mathrm{diag}\big(\mathbf{c}_{p.}^{(\mathbf{S})}\big) \tag{3.9}$$

and also visualized in Fig. 3.2.

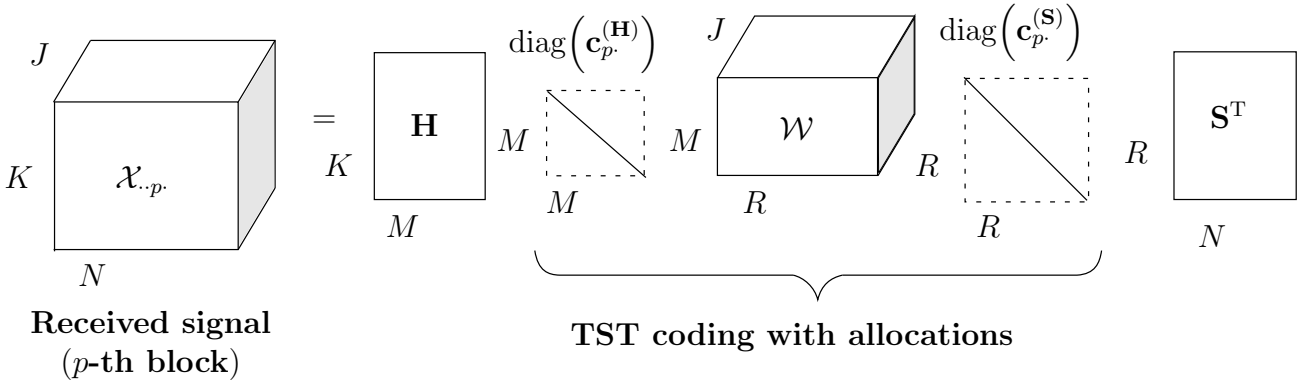


Figure 3.2: Visualization of the tensor slice $\mathcal{X}_{..p.}$ of the PARATUCK-(2,4) model.

Let us define $\mathbf{X}_{..p,j} \in \mathbb{C}^{K\times N}$ as a matrix slice of the received signal tensor $\mathcal{X}$. Using (3.5), it leads to the following factorization

$$\mathbf{X}_{..p,j} = \mathbf{H}\,\mathbf{U}_{..p,j} = \mathbf{H}\,\mathbf{G}_{..p,j}\,\mathbf{S}^{\mathrm{T}}, \tag{3.10}$$

with $\mathbf{G}_{..p,j} \in \mathbb{C}^{M\times R}$ given by (3.4). By stacking column-wise the set of matrix slices $\{\mathbf{X}_{..1,1}, \ldots, \mathbf{X}_{..P,J}\}$ and $\{\mathbf{X}_{..1,1}^{\mathrm{T}}, \ldots, \mathbf{X}_{..P,J}^{\mathrm{T}}\}$, and using (3.10), we can deduce, respectively, the following two

matrix unfoldings of $\mathcal{X}$ as

$$\begin{aligned}\mathbf{X}_{PJK\times N} &\triangleq \begin{bmatrix}\mathbf{X}_{\cdot\cdot 1,1}\\ \vdots\\ \mathbf{X}_{\cdot\cdot 1,J}\\ \vdots\\ \mathbf{X}_{\cdot\cdot P,1}\\ \vdots\\ \mathbf{X}_{\cdot\cdot P,J}\end{bmatrix} \in \mathbb{C}^{PJK\times N} \qquad \mathbf{X}_{PJN\times K} \triangleq \begin{bmatrix}\mathbf{X}^{\mathrm{T}}_{\cdot\cdot 1,1}\\ \vdots\\ \mathbf{X}^{\mathrm{T}}_{\cdot\cdot 1,J}\\ \vdots\\ \mathbf{X}^{\mathrm{T}}_{\cdot\cdot P,1}\\ \vdots\\ \mathbf{X}^{\mathrm{T}}_{\cdot\cdot P,J}\end{bmatrix} \in \mathbb{C}^{PJN\times K}\\ &= (\mathbf{I}_{PJ}\otimes\mathbf{H})\,\mathbf{G}_{PJM\times R}\mathbf{S}^{\mathrm{T}}, \qquad = (\mathbf{I}_{PJ}\otimes\mathbf{S})\,\mathbf{G}_{PJR\times M}\mathbf{H}^{\mathrm{T}},\end{aligned} \tag{3.11}$$

where

$$\mathbf{G}_{PJM\times R} \triangleq \begin{bmatrix}\mathbf{G}_{\cdot\cdot 1,1}\\ \vdots\\ \mathbf{G}_{\cdot\cdot 1,J}\\ \vdots\\ \mathbf{G}_{\cdot\cdot P,1}\\ \vdots\\ \mathbf{G}_{\cdot\cdot P,J}\end{bmatrix} \in \mathbb{C}^{PJM\times R}, \quad \mathbf{G}_{PJR\times M} \triangleq \begin{bmatrix}\mathbf{G}^{\mathrm{T}}_{\cdot\cdot 1,1}\\ \vdots\\ \mathbf{G}^{\mathrm{T}}_{\cdot\cdot 1,J}\\ \vdots\\ \mathbf{G}^{\mathrm{T}}_{\cdot\cdot P,1}\\ \vdots\\ \mathbf{G}^{\mathrm{T}}_{\cdot\cdot P,J}\end{bmatrix} \in \mathbb{C}^{PJR\times M} \tag{3.12}$$

represent two matrix unfoldings of $\mathcal{G} \in \mathbb{C}^{M\times R\times P\times J}$.

Applying property (A.3) to (3.10) and (3.4), denoting $\mathrm{vec}(\cdot)$ as the vectorization operator, we have

$$\mathrm{vec}(\mathbf{X}_{\cdot\cdot p,j}) = (\mathbf{S}\otimes\mathbf{H})\,\mathrm{vec}(\mathbf{G}_{\cdot\cdot p,j}) \in \mathbb{C}^{NK\times 1} \tag{3.13}$$

and

$$\begin{aligned}\mathrm{vec}(\mathbf{G}_{\cdot\cdot p,j}) &= \left(\mathrm{diag}\left(\mathbf{c}^{(\mathbf{S})}_{\mathrm{p}\cdot}\right)\otimes\mathrm{diag}\left(\mathbf{c}^{(\mathbf{H})}_{\mathrm{p}\cdot}\right)\right)\mathrm{vec}(\mathbf{W}_{\cdot\cdot j}) \in \mathbb{C}^{RM\times 1}\\ &= \mathrm{diag}(\mathrm{vec}(\mathbf{W}_{\cdot\cdot j}))\left(\mathbf{c}^{(\mathbf{S})^{\mathrm{T}}}_{p\cdot}\otimes\mathbf{c}^{(\mathbf{H})^{\mathrm{T}}}_{p\cdot}\right),\end{aligned} \tag{3.14}$$

which gives

$$\mathrm{vec}(\mathbf{X}_{\cdot\cdot p,j}) = (\mathbf{S}\otimes\mathbf{H})\,\mathrm{diag}(\mathrm{vec}(\mathbf{W}_{\cdot\cdot j}))\left(\mathbf{c}^{(\mathbf{S})^{\mathrm{T}}}_{p\cdot}\otimes\mathbf{c}^{(\mathbf{H})^{\mathrm{T}}}_{p\cdot}\right). \tag{3.15}$$

Using (3.13) and (3.14), we can deduce a third matrix unfolding of $\mathcal{X}$ as

$$\begin{aligned}\mathbf{X}_{KN\times JP} &\triangleq \left[\mathrm{vec}\left(\mathbf{X}^{\mathrm{T}}_{\cdot\cdot 1,1}\right)\ \cdots\ \mathrm{vec}\left(\mathbf{X}^{\mathrm{T}}_{\cdot\cdot P,1}\right)\ \cdots\ \mathrm{vec}\left(\mathbf{X}^{\mathrm{T}}_{\cdot\cdot 1,J}\right)\ \cdots\ \mathrm{vec}\left(\mathbf{X}^{\mathrm{T}}_{\cdot\cdot P,J}\right)\right]\\ &= (\mathbf{H}\otimes\mathbf{S})\,\mathbf{G}_{MR\times JP} \in \mathbb{C}^{KN\times JP}\end{aligned} \tag{3.16}$$

with a matrix unfoldings of $\mathcal{G}$ given by

$$\begin{aligned}\mathbf{G}_{MR\times JP} &\triangleq \left[\mathrm{vec}\left(\mathbf{G}^{\mathrm{T}}_{\cdot\cdot 1,1}\right)\ \cdots\ \mathrm{vec}\left(\mathbf{G}^{\mathrm{T}}_{\cdot\cdot P,1}\right)\ \cdots\ \mathrm{vec}\left(\mathbf{G}^{\mathrm{T}}_{\cdot\cdot 1,J}\right)\ \cdots\ \mathrm{vec}\left(\mathbf{G}^{\mathrm{T}}_{\cdot\cdot P,J}\right)\right] \in \mathbb{C}^{MR\times JP}\\ &= \left[\mathrm{diag}\left(\mathrm{vec}\left(\mathbf{W}^{\mathrm{T}}_{\cdot\cdot 1}\right)\right)\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}}{}_{\diamond}\,\mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)\ \cdots\ \mathrm{diag}\left(\mathrm{vec}\left(\mathbf{W}^{\mathrm{T}}_{\cdot\cdot J}\right)\right)\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}}{}_{\diamond}\,\mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)\right].\end{aligned} \tag{3.17}$$

The tensor $\mathbf{G}_{MR\times JP}$ can be also expressed applying property (A.1) to (3.17) in the form

$$\mathbf{G}_{MR\times JP} = \left(\mathbf{W}_{J\times MR} \diamond \left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}}\right)^{\mathrm{T}} \tag{3.18}$$

where

$$\mathbf{W}_{J\times MR} \triangleq \left[\mathrm{vec}\left(\mathbf{W}_{::1}^{\mathrm{T}}\right) \ \cdots \ \mathrm{vec}\left(\mathbf{W}_{::J}^{\mathrm{T}}\right)\right]^{\mathrm{T}} \in \mathbb{C}^{J\times MR}. \tag{3.19}$$

The received signal given by (3.5) can be also rewritten as a constrained CP model of third-order instead of the PT-(2,4) model which will allow us to exploit the CP results. Firstly, we rewrite the PT-(2,4) as a constrained CP-3 model aiming to relate both models and then restrict them to our problem.

TST system modeled by a constrained CP model

Now, let us rewrite the PT-(2,4) model as the following third-order constrained CP model $\mathcal{X} \in \mathbb{C}^{I_1 I_2\times I_3\times I_4}$

$$x_{j_1,i_3,i_4} = \sum_{r=1}^{R_1R_2} \overline{a}_{j_1,r}\, \overline{b}_{i_3,r}\, \overline{c}_{i_4,r}, \qquad j_1 \in \{1,...,I_1I_2\}, \tag{3.20}$$

whose matrix factors $\overline{\mathbf{A}} \in \mathbb{C}^{I_1I_2\times R_1R_2}$, $\overline{\mathbf{B}} \in \mathbb{C}^{I_3\times R_1R_2}$, $\overline{\mathbf{C}} \in \mathbb{C}^{I_4\times R_1R_2}$ are given by

$$\begin{aligned}
\overline{\mathbf{A}} &= \mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)},\\
\overline{\mathbf{B}} &= \left(\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)}\right)^{\mathrm{T}},\\
\overline{\mathbf{C}} &= \mathbf{C}_{I_4\times R_1R_2}.
\end{aligned} \tag{3.21}$$

From (3.21) and the correspondences in (3.6), the received signal can be expressed by a constrained CP-3 model $\mathcal{X} \in \mathbb{C}^{KN\times P\times J}$, where

$$\begin{aligned}
\overline{\mathbf{A}} &= \mathbf{H} \otimes \mathbf{S} \in \mathbb{C}^{KN\times MR},\\
\overline{\mathbf{B}} &= \left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}} \in \mathbb{C}^{P\times MR},\\
\overline{\mathbf{C}} &= \mathbf{W}_{J\times MR} \in \mathbb{C}^{J\times MR},
\end{aligned} \tag{3.22}$$

are its loading matrices.

From (2.21) and (3.22), let us write the following unfolded matrix

$$\mathbf{X}_{KN\times JP} = (\mathbf{H} \otimes \mathbf{S}) \left(\mathbf{W}_{J\times MR} \diamond \left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}}\right)^{\mathrm{T}}. \tag{3.23}$$

Verify that this unfolded matrix constructed by the factors $\overline{\mathbf{A}}$, $\overline{\mathbf{B}}$ and $\overline{\mathbf{C}}$ given in (3.22) results exactly in the same expression for the unfolded matrix $\mathbf{X}_{KN\times JP}$ given in (3.16) with (3.18).

Special case: TST system without allocation structures

In order to simplify the system, we can consider the fourth-order code tensor denoted by $\mathcal{W} \in \mathbb{C}^{M \times R \times P \times J}$ instead of the combination of allocation structures with precoding. Thus, the received signal can be rewritten as

$$x_{k,n,p,j} = \sum_{m=1}^{M} \sum_{r=1}^{R} w_{m,r,p,j}\, h_{k,m}\, s_{n,r}. \tag{3.24}$$

The employment of a fourth-order code tensor simplifies the analyze of identifiability and uniqueness conditions, and also allows an increase of information diversity because the model becomes more flexible in terms of structure, we can fix an optimal code which will satisfy appropriately both conditions and can provide better performance. It will be clear later.

3.2 Space-Time-Frequency (STF) system

Consider a multicarrier MIMO wireless communication system using M transmit antennas, K receive antennas and F subcarriers. We assume a transmission consisting of P data blocks, each one composed of N symbol periods. For a fixed symbol period and subcarrier, the (m, p, f)-th space-time-frequency (STF) coded signal, associated with the m-th transmit antenna, p-th block and f-th subcarrier is generated by a tensor coding operating on R data streams of N information symbols each one.

The STF coding structure employs two allocation tensors: the stream allocation tensor $\mathcal{C}^{(\mathbf{S})} \in \mathbb{R}^{F \times P \times R}$ and the antenna allocation tensor $\mathcal{C}^{(\mathcal{H})} \in \mathbb{R}^{F \times P \times M}$, which are composed uniquely of 1's and 0's. The first tensor determines the time-frequency mapping of the R data streams across P blocks and F subcarriers, and the second one determines the time-frequency mapping of the M transmit antennas.

The (f, m, n, p)-th element of the coded signal tensor $\mathcal{U} \in \mathbb{C}^{F \times M \times N \times P}$ associated with the f-th subcarrier, m-th transmit antenna, n-th symbol period and p-th data block, is given by [62]

$$u_{f,m,n,p} = \sum_{r=1}^{R} w_{m,r}\, s_{n,r}\, c^{(\mathcal{H})}_{f,p,m}\, c^{(\mathbf{S})}_{f,p,r} = \sum_{r=1}^{R} t_{f,m,r,p}\, s_{n,r} \tag{3.25}$$

with

$$t_{f,m,r,p} \triangleq w_{m,r}\, c^{(\mathcal{H})}_{f,p,m}\, c^{(\mathbf{S})}_{f,p,r}, \tag{3.26}$$

where $w_{m,r}$ is (m, r)-th entry of the code matrix $\mathbf{W}$ and $s_{n,r}$ denotes the n-th transmitted symbol associated with the r-th data stream.

Let us define the MIMO-OFDM channel given by the tensor $\mathcal{H} \in \mathbb{C}^{F \times K \times M}$, where $h_{f,k,m}$ is a complex coefficient of the channel linking the m-th transmit antenna with the k-th receive antenna for the f-th subcarrier. The fading coefficients are assumed to be constant during at least P blocks.

In the noiseless case of scattering-rich multipath fading channel, the received signal tensor $\mathcal{X} \in \mathbb{C}^{F \times K \times N \times P}$ associated with the f-th subcarrier and received at the k-th antenna during

the n-th symbol period of the p-th data block, is given by [62]

$$x_{f,k,n,p} = \sum_{m=1}^{M} h_{f,k,m}\, u_{f,m,n,p} = \sum_{m=1}^{M}\sum_{r=1}^{R} t_{f,m,r,p}\, h_{f,k,m}\, s_{n,r}. \tag{3.27}$$

According to [62], (3.27) follows a generalized PT-2 model. Using the n-mode product definition (2.13) and (3.27), we can write the f-th tensorial slice of $\mathcal{X} \in \mathbb{C}^{F\times K\times N\times P}$, containing all received signals associated with the f-th subcarrier, as

$$\mathcal{X}^{(f)} = \mathcal{T}^{(f)} \times_1 \mathbf{H}^{(f)} \times_2 \mathbf{S} \tag{3.28}$$

where $\mathcal{X}^{(f)} \triangleq \mathcal{X}_{f\cdots} \in \mathbb{C}^{K\times N\times P}$, $\mathcal{T}^{(f)} \triangleq \mathcal{T}_{f\cdots} \in \mathbb{C}^{M\times R\times P}$ and $\mathbf{H}^{(f)} \triangleq \mathbf{H}_{f\cdot\cdot} \in \mathbb{C}^{K\times M}$.

Let us define $\mathbf{X}_{f\cdot\cdot p} \in \mathbb{C}^{K\times N}$ as the matrix slice obtained by fixing f and p indices of $\mathcal{X} \in \mathbb{C}^{F\times K\times N\times P}$.

$$\mathbf{X}_{f\cdot\cdot p} = \mathbf{H}_{f\cdot\cdot}\, \mathbf{T}_{f\cdot\cdot p}\, \mathbf{S}^{\mathrm{T}} = \mathbf{H}_{f\cdot\cdot}\, \mathbf{U}_{f\cdot\cdot p}, \tag{3.29}$$

with

$$\mathbf{T}_{f\cdot\cdot p} = \operatorname{diag}\Big(\mathbf{c}^{(\mathcal{H})}_{f,p\cdot}\Big)\ \mathbf{W}\ \operatorname{diag}\Big(\mathbf{c}^{(\mathbf{S})}_{f,p\cdot}\Big) \in \mathbb{C}^{M\times R} \tag{3.30}$$

and

$$\mathbf{U}_{f\cdot\cdot p} = \mathbf{T}_{f\cdot\cdot p}\, \mathbf{S}^{\mathrm{T}} \in \mathbb{C}^{M\times N}. \tag{3.31}$$

Analogously to (3.11) and using (3.29), we can build two matrix unfoldings of $\mathcal{X}$ as

$$\begin{aligned}
\mathbf{X}_{PFK\times N} &\triangleq \begin{bmatrix} \mathbf{X}_{1\cdot\cdot 1} \\ \vdots \\ \mathbf{X}_{F\cdot\cdot 1} \\ \vdots \\ \mathbf{X}_{1\cdot\cdot P} \\ \vdots \\ \mathbf{X}_{F\cdot\cdot P} \end{bmatrix} \in \mathbb{C}^{PFK\times N} \\
&= (\mathbf{I}_P \otimes \operatorname{bdiag}(\mathbf{H}_{1\cdot\cdot}, \ldots, \mathbf{H}_{F\cdot\cdot}))\, \mathbf{T}_{PFM\times R}\, \mathbf{S}^{\mathrm{T}},
\end{aligned} \qquad
\begin{aligned}
\mathbf{X}_{PFN\times K} &\triangleq \begin{bmatrix} \mathbf{X}^{\mathrm{T}}_{1\cdot\cdot 1} \\ \vdots \\ \mathbf{X}^{\mathrm{T}}_{F\cdot\cdot 1} \\ \vdots \\ \mathbf{X}^{\mathrm{T}}_{1\cdot\cdot P} \\ \vdots \\ \mathbf{X}^{\mathrm{T}}_{F\cdot\cdot P} \end{bmatrix} \in \mathbb{C}^{PFN\times K} \\
&= (\mathbf{I}_{PF} \otimes \mathbf{S})\, \mathbf{T}_{PFR\times FM}\, \mathbf{H}_{FM\times K},
\end{aligned} \tag{3.32}$$

where

$$\mathbf{T}_{PFM\times R} \triangleq \begin{bmatrix} \mathbf{T}_{1\cdot\cdot 1} \\ \vdots \\ \mathbf{T}_{F\cdot\cdot 1} \\ \vdots \\ \mathbf{T}_{1\cdot\cdot P} \\ \vdots \\ \mathbf{T}_{F\cdot\cdot P} \end{bmatrix} \in \mathbb{C}^{PFM\times R}, \quad \mathbf{T}_{PFR\times FM} \triangleq \begin{bmatrix} \mathbf{T}^{\mathrm{T}}_{1\cdot\cdot 1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{T}^{\mathrm{T}}_{F\cdot\cdot 1} \\ & \vdots & \\ \mathbf{T}^{\mathrm{T}}_{1\cdot\cdot P} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{T}^{\mathrm{T}}_{F\cdot\cdot P} \end{bmatrix} \in \mathbb{C}^{PFR\times FM}, \tag{3.33}$$

with $\mathbf{H}_{FM\times K}$ denotes an unfolded matrix of $\mathcal{H}\in\mathbb{C}^{F\times K\times M}$ which can be expressed in terms of its matrix slices $\mathbf{H}_{f\cdot\cdot}$ by

$$\mathbf{H}_{FM\times K}\triangleq\left[\mathbf{H}_{1\cdot\cdot}\ \cdots\ \mathbf{H}_{F\cdot\cdot}\right]^{\mathrm{T}}\in\mathbb{C}^{FM\times K}. \tag{3.34}$$

Applying property (A.3) to (3.29) and (3.30), we have

$$\mathrm{vec}\left(\mathbf{X}_{f\cdot\cdot p}^{\mathrm{T}}\right)=\left(\mathbf{H}_{f\cdot\cdot}\otimes\mathbf{S}\right)\mathrm{vec}\left(\mathbf{T}_{f\cdot\cdot p}^{\mathrm{T}}\right)\in\mathbb{C}^{KN\times 1} \tag{3.35}$$

and

$$\begin{aligned}\mathrm{vec}\left(\mathbf{T}_{f\cdot\cdot p}^{\mathrm{T}}\right)&=\left(\mathrm{diag}\left(\mathbf{c}_{\mathrm{f,p}\cdot}^{(\mathcal{H})}\right)\otimes\mathrm{diag}\left(\mathbf{c}_{\mathrm{f,p}\cdot}^{(\mathbf{S})}\right)\right)\mathrm{vec}\left(\mathbf{W}^{\mathrm{T}}\right)\\&=\mathrm{diag}\left(\mathbf{c}_{f,p\cdot}^{(\mathcal{H})}\otimes\mathbf{c}_{f,p\cdot}^{(\mathbf{S})}\right)\mathrm{vec}\left(\mathbf{W}^{\mathrm{T}}\right)\in\mathbb{C}^{MR\times 1},\end{aligned} \tag{3.36}$$

which gives

$$\mathrm{vec}\left(\mathbf{X}_{f\cdot\cdot p}^{\mathrm{T}}\right)=\left(\mathbf{H}_{f\cdot\cdot}\otimes\mathbf{S}\right)\mathrm{diag}\left(\mathbf{c}_{f,p\cdot}^{(\mathcal{H})}\otimes\mathbf{c}_{f,p\cdot}^{(\mathbf{S})}\right)\mathrm{vec}\left(\mathbf{W}^{\mathrm{T}}\right). \tag{3.37}$$

For convenience, let us define a third unfolded matrix analogously to (3.16) as follows

$$\begin{aligned}\mathbf{X}_{KN\times FP}&\triangleq\left[\mathrm{vec}\left(\mathbf{X}_{1\cdot\cdot 1}^{\mathrm{T}}\right)\ \cdots\ \mathrm{vec}\left(\mathbf{X}_{1\cdot\cdot P}^{\mathrm{T}}\right)\ \cdots\ \mathrm{vec}\left(\mathbf{X}_{F\cdot\cdot 1}^{\mathrm{T}}\right)\ \cdots\ \mathrm{vec}\left(\mathbf{X}_{F\cdot\cdot P}^{\mathrm{T}}\right)\right]\\&=\left(\mathbf{H}_{K\times FM}\otimes\mathbf{S}\right)\mathbf{T}_{FMR\times FP}\in\mathbb{C}^{KN\times FP},\end{aligned} \tag{3.38}$$

where

$$\begin{aligned}\mathbf{T}_{FMR\times FP}&\triangleq\begin{bmatrix}\left[\mathrm{vec}\left(\mathbf{T}_{1\cdot\cdot 1}^{\mathrm{T}}\right)\cdots\mathrm{vec}\left(\mathbf{T}_{1\cdot\cdot P}^{\mathrm{T}}\right)\right] & & \mathbf{0}\\ & \ddots & \\ \mathbf{0} & & \left[\mathrm{vec}\left(\mathbf{T}_{F\cdot\cdot 1}^{\mathrm{T}}\right)\cdots\mathrm{vec}\left(\mathbf{T}_{F\cdot\cdot P}^{\mathrm{T}}\right)\right]\end{bmatrix}\\&\in\mathbb{C}^{FMR\times FP}\end{aligned} \tag{3.39}$$

and from (3.36) and the Khatri-Rao definition (A.1), we can write

$$\begin{aligned}\left[\mathrm{vec}\left(\mathbf{T}_{f\cdot\cdot 1}^{\mathrm{T}}\right)\cdots\mathrm{vec}\left(\mathbf{T}_{f\cdot\cdot P}^{\mathrm{T}}\right)\right]^{\mathrm{T}}&=\begin{bmatrix}\mathrm{vec}\left(\mathbf{W}^{\mathrm{T}}\right)^{\mathrm{T}}\mathrm{diag}\left(\mathbf{c}_{f,1\cdot}^{(\mathcal{H})}\otimes\mathbf{c}_{f,1\cdot}^{(\mathbf{S})}\right)\\ \vdots\\ \mathrm{vec}\left(\mathbf{W}^{\mathrm{T}}\right)^{\mathrm{T}}\mathrm{diag}\left(\mathbf{c}_{f,P\cdot}^{(\mathcal{H})}\otimes\mathbf{c}_{f,P\cdot}^{(\mathbf{S})}\right)\end{bmatrix}\\&=\mathrm{vec}\left(\mathbf{W}^{\mathrm{T}}\right)^{\mathrm{T}}\diamond\left(\mathbf{C}_{f\cdot\cdot}^{(\mathcal{H})^{\mathrm{T}}}\diamond\mathbf{C}_{f\cdot\cdot}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}}\in\mathbb{C}^{P\times MR}.\end{aligned} \tag{3.40}$$

3.3 Performance analysis

In this section, we analyze the performance of the TST coding focusing on the diversity of information transmitted and derive the maximum diversity gain over a flat fading channel, and in the sequence, extend this analysis for the STF system.

3.3.1 TST system

Consider that the channel matrix $\mathbf{H}$ has independent entries following a circular symmetric complex Gaussian random distribution, i.e. $h_{k,m} \sim \mathcal{CN}(0,1)$, or equivalently its real and imaginary components are i.i.d. and distributed as $\mathcal{N}(0,1/2)$, which corresponds to the assumption of flat Rayleigh fading. We also assume that the receiver has perfect knowledge of $\mathbf{H}$ and of the TST coding parameter set $\{\mathcal{W}, \mathbf{C}^{(\mathbf{S})}, \mathbf{C}^{(\mathbf{H})}\}$.

For matrix ST coding, the transmitted ST code matrix, also called ST codeword, is defined as the matrix associated with the coding mapping:

$$\mathbf{s}_n \in \mathbb{C}^{M\times 1} \longrightarrow \mathbf{C}_n \in \mathbb{C}^{M\times T} \tag{3.41}$$

where M and T denote the space and time spreading lengths. In the case of the proposed tensor ST coding, the ST codeword is the fourth-order tensor associated with the coding mapping:

$$\mathbf{S}^{\mathrm{T}} \in \mathbb{C}^{R\times N} \longrightarrow \mathcal{U} \in \mathbb{C}^{M\times N\times P\times J} \tag{3.42}$$

whose dimensions are the lengths of space, time, block and chip spreadings.

As already mentioned, matrix ST coding approaches are based on codeword estimation, followed by a decoding step for estimating the transmitted symbols. So, the performance analysis is generally based on the pairwise error probability (PEP) of the maximum likelihood (ML) estimator of the codeword matrix, defined as the probability that the ML estimator estimates $\hat{\mathbf{C}}_n$ when $\mathbf{C}_n$ is actually sent. In the case of TST coding, it is possible to directly estimate the symbol matrix instead of the codeword tensor, which explains why our performance analysis is based on the PEP of the ML estimator of $\mathbf{S}$ instead of $\mathcal{U}$.

The diversity gain d is defined as the negative of the asymptotic slope of the plot PEP(ρ) on a log-log scale, where ρ denotes the received signal-to-noise ratio (SNR), and PEP is hereafter the probability that the ML estimator estimates $\hat{\mathbf{S}}$ when $\mathbf{S}$ is actually transmitted.

In the sequel, we first determine the function PEP(ρ), and then we deduce the diversity gain for TST coding.

The conditional PEP between $\mathbf{S}$ and $\hat{\mathbf{S}}$ can be approximated by [6],[102]:

$$P\Big(\mathbf{S} \to \hat{\mathbf{S}}|\mathbf{H}\Big) = Q\left(\sqrt{\frac{1}{2N_0}\|\mathcal{X} - \hat{\mathcal{X}}\|_F^2}\right), \tag{3.43}$$

where N_0 is the noise variance per (real and imaginary) dimension and $Q(\cdot)$ is the complementary cumulative distribution function of a Gaussian variable defined as

$$Q(x) \triangleq P(x \geq y) = \frac{1}{\sqrt{2\pi}} \int_y^{\infty} \exp\left(-\frac{y^2}{2}\right) \mathrm{d}y. \tag{3.44}$$

The Q-function can be rewritten in an alternative form referred to as Craig's formula [1]

$$Q(x) = \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \exp\left(-\frac{x^2}{2\sin^2(\beta)}\right) \mathrm{d}\beta, \tag{3.45}$$

applying this definition to (3.43), one gets

$$P\Big(\mathbf{S}\to\hat{\mathbf{S}}|\mathbf{H}\Big)=\frac{1}{\pi}\int_0^{\frac{\pi}{2}}\exp\left(-\frac{\|\mathcal{X}-\hat{\mathcal{X}}\|_F^2}{4N_0\sin^2(\beta)}\right)\mathrm{d}\beta. \tag{3.46}$$

Defining the difference between the matrix slices (p,j) of the codeword tensors $\mathcal{U}$ and $\hat{\mathcal{U}}$ as

$$\mathbf{E}^{(p,j)}=\mathbf{U}_{\cdot\cdot p,j}-\hat{\mathbf{U}}_{\cdot\cdot p,j}\ \in\mathbb{C}^{M\times N} \tag{3.47}$$

and using (3.5) and (3.10), we have

$$\begin{aligned}\|\mathcal{X}-\hat{\mathcal{X}}\|_F^2&=\sum_{k=1}^{K}\sum_{n=1}^{N}\sum_{p=1}^{P}\sum_{j=1}^{J}|x_{k,n,p,j}-\hat{x}_{k,n,p,j}|^2\\&=\sum_{p=1}^{P}\sum_{j=1}^{J}\|\mathbf{H}\,\mathbf{E}^{(p,j)}\|_F^2=\sum_{p=1}^{P}\sum_{j=1}^{J}\mathrm{tr}\big(\mathbf{H}\mathbf{A}^{(p,j)}\mathbf{H}^{\mathrm{H}}\big)\end{aligned} \tag{3.48}$$

where $\mathbf{A}^{(p,j)}\triangleq\mathbf{E}^{(p,j)}\left(\mathbf{E}^{(p,j)}\right)^{\mathrm{H}}$ is Hermitian and nonnegative definite. Observe that the trace of a product can also be written in the following form:

$$y^{(p,j)}\triangleq\mathrm{tr}\big(\mathbf{H}\mathbf{A}^{(p,j)}\mathbf{H}^{\mathrm{H}}\big)=\mathrm{vec}\left(\mathbf{H}^{\mathrm{T}}\right)^{\mathrm{T}}\left(\mathbf{I}_K\otimes\mathbf{A}^{(p,j)}\right)\mathrm{vec}\left(\mathbf{H}^{\mathrm{T}}\right)^{*}. \tag{3.49}$$

From (3.46) and (3.49), we can derive the average PEP as follows

$$\begin{aligned}P\Big(\mathbf{S}\to\hat{\mathbf{S}}\Big)&=\int_0^{\infty}\left(\frac{1}{\pi}\int_0^{\frac{\pi}{2}}\prod_{p=1}^{P}\prod_{j=1}^{J}\exp\left(-\frac{y^{(p,j)}}{4N_0\sin^2(\beta)}\right)\mathrm{d}\beta\right)\ p_{y^{(p,j)}}\left(y^{(p,j)}\right)\mathrm{d}y^{(p,j)}\\&=\frac{1}{\pi}\int_0^{\frac{\pi}{2}}\prod_{p=1}^{P}\prod_{j=1}^{J}\left(\int_0^{\infty}\exp\left(-\frac{y^{(p,j)}}{4N_0\sin^2(\beta)}\right)p_{y^{(p,j)}}\left(y^{(p,j)}\right)\mathrm{d}y^{(p,j)}\right)\mathrm{d}\beta\\&=\frac{1}{\pi}\int_0^{\frac{\pi}{2}}\prod_{p=1}^{P}\prod_{j=1}^{J}M_{y^{(p,j)}}\left(-\frac{1}{4N_0\sin^2(\beta)}\right)\mathrm{d}\beta,\end{aligned} \tag{3.50}$$

where

$$M_y(\gamma)=\int_0^{\infty}\exp(\gamma\,y)\,p_y(y)\mathrm{d}y \tag{3.51}$$

denotes the moment generating function of y. It is important to attend that the probability density function of $y^{(p,j)}$ denoted by $p_{y^{(p,j)}}(\cdot)$ does not depend on the variables p and j (i.e. $\mathbf{A}^{(p,j)}$).

Knowing that the channel coefficients $h_{k,m}$ are i.i.d and have a circular symmetric complex Gaussian random distribution with zero-mean and unit variance, let us introduce an interesting theorem [103].

Theorem 3.1. [103] *The moment generating function of a Hermitian quadratic form in a complex Gaussian random variable $y=\mathbf{z}^{\mathrm{H}}\mathbf{F}\mathbf{z}$, where $\mathbf{z}$ is a circularly symmetric complex Gaussian vector with mean $\bar{\mathbf{z}}$ and covariance matrix $\mathbf{R}_{\mathbf{z}}$ and $\mathbf{F}$ is a Hermitian matrix, is given by*

$$M_y(s)\triangleq\int_0^{\infty}\exp(sy)\,p_y\left(y\right)\mathrm{d}y=\frac{\exp\left(s\bar{\mathbf{z}}^{\mathrm{H}}\mathbf{F}\left(\mathbf{I}-s\mathbf{R}_{\mathbf{z}}\mathbf{F}\right)^{-1}\bar{\mathbf{z}}\right)}{\det\left(\mathbf{I}-\mathrm{s}\mathbf{R}_{\mathbf{z}}\mathbf{F}\right)}. \tag{3.52}$$

Considering the correspondences $\mathbf{F} = \mathbf{I}_K \otimes \mathbf{A}^{(p,j)}$, $\mathbf{R}_{\mathbf{z}} = \mathbf{I}_{KM}$, $\mathbf{z} = \mathrm{vec}\left(\mathbf{H}^{\mathrm{T}}\right)$, $\bar{\mathbf{z}} = \mathbf{0}_{KM}$ and $y^{(p,j)}$ defined in (3.49), we can apply (3.52) to (3.50) getting

$$\begin{aligned} P\Big(\mathbf{S} \to \hat{\mathbf{S}}\Big) &= \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \prod_{p=1}^{P} \prod_{j=1}^{J} \det\left(\mathbf{I}_{KM} + \frac{1}{4N_0 \sin^2(\beta)} \left(\mathbf{I}_K \otimes \mathbf{A}^{(p,j)}\right)\right)^{-1} \mathrm{d}\beta \\ &= \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \prod_{p=1}^{P} \prod_{j=1}^{J} \det\left(\mathbf{I}_M + \frac{1}{4N_0 \sin^2(\beta)} \mathbf{A}^{(p,j)}\right)^{-K} \mathrm{d}\beta. \end{aligned} \tag{3.53}$$

To solve the integration over β in the expression (3.53) is not simple, hence we can employ the Chernoff bound [102, 1] in order to eliminate the integral and to give an upper bound of PEP

$$P\Big(\mathbf{S} \to \hat{\mathbf{S}}\Big) \leq \prod_{p=1}^{P} \prod_{j=1}^{J} \det\left(\mathbf{I}_M + \frac{1}{4N_0} \mathbf{A}^{(p,j)}\right)^{-K}. \tag{3.54}$$

By definition, the Chernoff bound is obtained by taking $\sin^2(\beta) = 1$ in (3.45), and so rewriting (3.53) with $\sin^2(\beta) = 1$.

Since $\det(\mathbf{I} + \alpha\mathbf{A}) = \prod_{i=1}^{\mathrm{rank}(\mathbf{A})} (1 + \alpha\lambda_i(\mathbf{A}))$ with $\lambda_i(\mathbf{A})$ eigenvalue of $\mathbf{A}$, we can rewrite (3.54) as

$$P\Big(\mathbf{S} \to \hat{\mathbf{S}}\Big) \leq \prod_{p=1}^{P} \prod_{j=1}^{J} \prod_{i=1}^{r^{(p,j)}} \left(1 + \frac{1}{4N_0} \lambda_i^{(p,j)}\right)^{-K}, \tag{3.55}$$

where $\lambda_i^{(p,j)}$ and $r^{(p,j)} \triangleq \mathrm{rank}\big(\mathbf{A}^{(\mathrm{p,j})}\big)$ denote the non-zero eigenvalues and the rank of $\mathbf{A}^{(p,j)}$, respectively.

At high SNR, i.e. for small values of N_0, the above upper bound on the PEP becomes

$$P\Big(\mathbf{S} \to \hat{\mathbf{S}}\Big) \leq \prod_{p=1}^{P} \prod_{j=1}^{J} \prod_{i=1}^{r^{(p,j)}} \left(\lambda_i^{(p,j)}\right)^{-K} \left(\frac{1}{4N_0}\right)^{-K \sum\limits_{j,p} r^{(p,j)}} \tag{3.56}$$

which gives the following diversity gain

$$d^{\mathrm{TST}} = K \sum_{j=1}^{J} \sum_{p=1}^{P} r^{(p,j)}. \tag{3.57}$$

Recalling that $\mathbf{A}^{(p,j)} = \mathbf{E}^{(p,j)} \left(\mathbf{E}^{(p,j)}\right)^{\mathrm{H}}$, we have

$$r^{(p,j)} = \mathrm{rank}\big(\mathbf{A}^{(p,j)}\big) = \mathrm{rank}\big(\mathbf{E}^{(p,j)}\big). \tag{3.58}$$

Using (3.3), the difference (3.47) of the codeword matrix slices can be rewritten as

$$\mathbf{E}^{(p,j)} = \mathbf{G}_{\cdot p,j} \left(\mathbf{S} - \hat{\mathbf{S}}\right)^{\mathrm{T}}. \tag{3.59}$$

As the symbol matrix $\mathbf{S}$ has independent entries following a random distribution, generically, we have $\text{rank}(\mathbf{S}) = \min(N, R)$. Applying the well-known property $\text{rank}(\mathbf{AB}) \leq \min(\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B}))$ for arbitrary matrices $\mathbf{A}$ and $\mathbf{B}$, and assuming $N \geq R$, we deduce that

$$\begin{aligned} r^{(p,j)} &= \text{rank}\big(\mathbf{E}^{(p,j)}\big) \leq \min\Big(\text{rank}(\mathbf{G}_{..p,j}), \text{rank}\Big(\mathbf{S} - \hat{\mathbf{S}}\Big)\Big) \\ &\leq \min(\min(M, R), \min(N, R)) \\ &\leq \min(M, R), \quad \forall\, p, j. \end{aligned} \tag{3.60}$$

It is interesting to observe that the maximum rank of $\mathbf{E}^{(p,j)}$ and consequently, of $\mathbf{G}_{..p,j}$ is depending on the structure of the allocation matrices $\mathbf{C}^{(\mathbf{S})}$ and $\mathbf{C}^{(\mathbf{H})}$. Let us assume that $\beta_p^{(\mathbf{S})}$ and $\beta_p^{(\mathbf{H})}$ denote the number of zero elements of $\mathbf{c}_{p\cdot}^{(\mathbf{S})}$ and $\mathbf{c}_{p\cdot}^{(\mathbf{H})}$, respectively. From (3.4), $\mathbf{G}_{..p,j}$ has $\beta_p^{(\mathbf{S})}$ zero columns and $\beta_p^{(\mathbf{H})}$ zero rows, then $\text{rank}(\mathbf{G}_{..p,j}) \leq \min\Big(M - \beta_p^{(\mathbf{H})}, R - \beta_p^{(\mathbf{S})}\Big)$. Thus, the maximum diversity gain is deduced from (3.57) in replacing $r^{(p,j)}$ by its upper bound given in (3.60) for all the values of p and j.

Theorem 3.2. (Maximum diversity gain for the TST system). *Assuming that $N \geq R$, then the TST system characterized by the design parameter set $\{P, J, M, R, K\}$ provides a maximum diversity gain equal to*

$$d_{\max}^{\text{TST}} = KJ \sum_{p=1}^{P} \min\big(M - \beta_p^{(\mathbf{H})}, R - \beta_p^{(\mathbf{S})}\big), \tag{3.61}$$

with $\beta_p^{(\mathbf{S})}$ and $\beta_p^{(\mathbf{H})}$ denoting the number of zero elements of $\mathbf{c}_{p\cdot}^{(\mathbf{S})}$ and $\mathbf{c}_{p\cdot}^{(\mathbf{H})}$, respectively.

So, we can conclude from (3.61) that the maximum diversity gain depends on the allocation matrices and therefore, we may have different performances for each particular allocation structure.

From (3.61), we can observe in the best case $d_{\max}^{\text{TST}} = KJP \min(M, R)$ and therefore, the TST coding provides a better diversity than standard matrix ST coding schemes that ensure a maximum diversity gain of KM. Moreover, for fixed numbers (K and M) of receive and transmit antennas, the maximum diversity gain can be increased by independently increasing the design parameters P, J and R (up to $R = M$).

However, we have to recall that an increase of P decreases the transmission rate*, while an increase of R increases the transmission rate. Moreover, for a fixed R, i.e. a fixed number of data streams to be estimated, an increase of P or J implies an increase of the number of received signals to be used for channel and symbol estimation, and thus an improvement of the estimation quality, while for fixed P and J, an increase of R implies an increase of the number of parameters (symbols) to be estimated, which degrades the quality of estimation. From these considerations, we see that the design parameters (P, J, M, R) must be chosen in such a way that the best tradeoff between transmission rate and BER performance be satisfied.

*The transmission rate (in bits per channel use) is given by $\frac{R}{P}\log_2(\mu)$, where μ is the cardinality of the information symbol constellation.

3.3.2 STF system

Let us derive the average PEP and the maximum diversity gain over a flat fading channel for the STF system. Analogously to (3.47), we can define the difference between the matrix slices (p, f) of the codeword tensors $\mathcal{U}$ and $\hat{\mathcal{U}}$ as

$$\mathbf{E}^{(p,f)} = \mathbf{U}_{f\cdot\cdot p} - \hat{\mathbf{U}}_{f\cdot\cdot p} \ \in \mathbb{C}^{M\times N} \tag{3.62}$$

and using (3.27) and (3.29), we have

$$\begin{aligned}\|\mathcal{X} - \hat{\mathcal{X}}\|_F^2 &= \sum_{k=1}^{K}\sum_{n=1}^{N}\sum_{p=1}^{P}\sum_{f=1}^{F} |x_{f,k,n,p} - \hat{x}_{f,k,n,p}|^2 \\ &= \sum_{p=1}^{P}\sum_{f=1}^{F} \|\mathbf{H}_{f\cdot\cdot}\,\mathbf{E}^{(p,f)}\|_F^2 = \sum_{p=1}^{P}\sum_{f=1}^{F} \mathrm{tr}\big(\mathbf{H}_{f\cdot\cdot}\mathbf{A}^{(p,f)}\mathbf{H}_{f\cdot\cdot}^{\mathrm{H}}\big)\,,\end{aligned} \tag{3.63}$$

where $\mathbf{A}^{(p,f)} \triangleq \mathbf{E}^{(p,f)}\left(\mathbf{E}^{(p,f)}\right)^{\mathrm{H}}$ is Hermitian and nonnegative definite.

Observe that both systems, TST and STF, show close expressions and we can employ the same reasoning to develop each step for the STF. For simplicity and in order to avoid repetitions, let us restrict to the main expressions.

We assume that the channel tensor $\mathcal{H}$ has independent entries following a circular symmetric complex Gaussian random distribution, i.e. $h_{f,k,m} \sim \mathcal{CN}(0, 1)$. The receiver has also perfect knowledge of $\mathcal{H}$ and of the STF parameter set $\{\mathbf{W}, \mathcal{C}^{(\mathbf{S})}, \mathcal{C}^{(\mathcal{H})}\}$.

Comparing (3.63) with (3.48), we can easily obtain the above upper bound on the PEP at high SNR for the STF, analogously to (3.56),

$$P\Big(\mathbf{S} \to \hat{\mathbf{S}}\Big) \leq \prod_{p=1}^{P}\prod_{f=1}^{F}\prod_{i=1}^{r^{(p,f)}} \left(\lambda_i^{(p,f)}\right)^{-K} \left(\frac{1}{4N_0}\right)^{-K\sum\limits_{f,p} r^{(p,f)}}, \tag{3.64}$$

where $\lambda_i^{(p,f)}$ and $r^{(p,f)} \triangleq \mathrm{rank}\big(\mathbf{A}^{(p,f)}\big)$ denote, respectively, the non-zero eigenvalues and the rank of $\mathbf{A}^{(p,f)}$, the diversity gain is given by

$$d^{\mathrm{STF}} = K\sum_{f=1}^{F}\sum_{p=1}^{P} r^{(p,f)}. \tag{3.65}$$

From (3.57) and (3.65), we note that both systems achieve the same diversity gain when $N \geq R$ and, $\mathbf{G}_{\cdot\cdot p,j}$ and $\mathbf{T}_{f\cdot\cdot p}$ for all p, j, f are chosen such that

$$\sum_{j=1}^{J}\sum_{p=1}^{P} \mathrm{rank}(\mathbf{G}_{\cdot\cdot p,j}) = \sum_{f=1}^{F}\sum_{p=1}^{P} \mathrm{rank}(\mathbf{T}_{f\cdot\cdot p})\,.$$

Analogously to (3.61) in Theorem 3.2 and from (3.65), we proposed a maximum diversity gain for the STF system given in the following theorem.

Theorem 3.3. (Maximum diversity gain for the STF system). *Assuming that $N \geq R$, then the STF system characterized by the design parameter set $\{P, F, M, R, K\}$ provides a maximum diversity gain equal to*

$$d_{\max}^{\mathrm{STF}} = K \sum_{f=1}^{F} \sum_{p=1}^{P} \min\Big(M - \beta_{p,f}^{(\mathcal{H})}, R - \beta_{p,f}^{(\mathbf{S})}\Big), \tag{3.66}$$

where $\beta_{p,f}^{(\mathbf{S})}$ and $\beta_{p,f}^{(\mathcal{H})}$ denote the number of zero elements of $\mathbf{c}_{f,p\cdot}^{(\mathbf{S})}$ and $\mathbf{c}_{f,p\cdot}^{(\mathcal{H})}$, respectively.

According to (3.61) and (3.66), we can verify that the allocation structures limit the maximum diversity gain for both systems. Thus, systems with different allocation structures may provide different performances.

Remark also that the allocation tensors for the STF system arrange the data transmission across the subcarriers and therefore, setting the same design parameters for both systems $\{K, P, M, R\}$ including $J = F$, we have $d_{\max}^{\mathrm{STF}} \leq d_{\max}^{\mathrm{TST}}$ from (3.61) and (3.66). It indicates that the extra time diversity J presented in the TST system can improve more the reliability of symbol recovery compared to the frequency diversity F in the STF system. In this sense, we can conclude that both diversities are not equivalent.

3.4 Uniqueness analysis

In this section, we employ the uniqueness results developed in Chapter 2 to derive practical conditions with the purpose of ensuring the estimation of symbol and channel for the TST and STF systems. We provide a more practical uniqueness condition to be satisfied than the one proposed in [62] for the STF system.

3.4.1 TST system

Analogously the uniqueness analysis for PT-(2,4) model in last chapter, let us consider the unfolded matrix defined in (3.23) to analyze the uniqueness conditions of this model in the TST context. Due to the special structure of this unfolded matrix, we can isolate the matrices to be estimated, i.e. the channel and symbol matrices. This analysis takes into account the knowledge of the allocation matrices and the code tensor at both transceiver and receiver sides.

We have to ensure the uniqueness of the PT-(2,4) model in order to guarantee the unique estimation of the symbol $\mathbf{S}$ and channel $\mathbf{H}$ matrices, hence both estimates of $\mathbf{S}$ and $\mathbf{H}$ have to satisfy the unfolded matrices in (3.11) and (3.16). Let $\hat{\mathbf{S}}$ and $\hat{\mathbf{H}}$ be alternative solutions which can write as $\hat{\mathbf{S}} = \mathbf{S}\,\mathbf{U}$ and $\hat{\mathbf{H}} = \mathbf{H}\,\mathbf{V}$, with $\mathbf{U} \in \mathbb{C}^{R\times R}$ and $\mathbf{V} \in \mathbb{C}^{M\times M}$ non-singular matrices. Choosing for simplicity (3.16), the unfolded matrix $\mathbf{X}_{KN\times JP}$ can be rewritten using the Kronecker property (A.4) as

$$\begin{aligned}(\mathbf{H}\otimes\mathbf{S})\,\mathbf{G}_{MR\times JP} &= \left(\hat{\mathbf{H}}\otimes\hat{\mathbf{S}}\right)\mathbf{G}_{MR\times JP} \\ &= (\mathbf{H}\otimes\mathbf{S})\,(\mathbf{V}\otimes\mathbf{U})\,\mathbf{G}_{MR\times JP},\end{aligned} \tag{3.67}$$

with $\mathbf{G}_{MR\times JP}$ given by (3.18).

Thus, we can directly apply Theorem 2.7 to derive a sufficient condition for TST system taking account the system structures as it is enunciated in the next theorem.

Theorem 3.4. *Suppose that* $\mathbf{S}$ *and* $\mathbf{H}$ *are full column-rank, and the perfect knowledge of the code tensor* $\mathcal{W}$ *and the allocation matrices* $\mathbf{C}^{(\mathbf{H})}$ *and* $\mathbf{C}^{(\mathbf{S})}$*. If the code tensor is chosen such that* $\mathbf{W}_{J\times MR}$ *is full column-rank and* ${\mathbf{C}^{(\mathbf{H})}}^{\mathrm{T}} \diamond {\mathbf{C}^{(\mathbf{S})}}^{\mathrm{T}}$ *has no zero-rows, then the estimates of* $\mathbf{S}$ *and* $\mathbf{H}$ *are unique up to a scaling factor* α*, i.e.*

$$\mathbf{S} = \alpha\,\hat{\mathbf{S}}, \quad \mathbf{H} = \frac{1}{\alpha}\,\hat{\mathbf{H}}. \tag{3.68}$$

Proof: As $\mathbf{S}$ and $\mathbf{H}$ are full column-rank, the left-inverse of $\mathbf{H} \otimes \mathbf{S}$ exists and is unique. By k-rank definition, if ${\mathbf{C}^{(\mathbf{H})}}^{\mathrm{T}} \diamond {\mathbf{C}^{(\mathbf{S})}}^{\mathrm{T}}$ has no zero-rows then $k\left(\left({\mathbf{C}^{(\mathbf{H})}}^{\mathrm{T}} \diamond {\mathbf{C}^{(\mathbf{S})}}^{\mathrm{T}}\right)^{\mathrm{T}}\right) \geq 1$. According to Theorem 2.7, if $\mathbf{W}_{J\times MR}$ is full column-rank, implying that $\mathbf{G}_{MR\times JP}$ is full row-rank, then both symbol and channel matrices are unique up to a scaling factor, i.e. (3.68), which concludes the proof. ∎

Remark 3.1.

- *As the elements of* $\mathbf{H}$ *are drawn from continuous probability density function,* $\mathbf{H}$ *is almost surely full column-rank when* $K \geq M$*.*
- *It is important to emphasize that the model does not present the permutation ambiguity and only the scalar ambiguity which can be removed by the knowledge of only one symbol at the receiver.*
- *According to Theorem 2.7, even when* $\mathbf{W}_{J\times MR}$ *is not full column-rank it is possible to achieve the uniqueness condition satisfying* $k(\mathbf{W}_{J\times MR}) + k\left(\left({\mathbf{C}^{(\mathbf{H})}}^{\mathrm{T}} \diamond {\mathbf{C}^{(\mathbf{S})}}^{\mathrm{T}}\right)^{\mathrm{T}}\right) \geq 1$*.*

According to the theorem above, the allocation structures have to satisfy $k\left(\left({\mathbf{C}^{(\mathbf{H})}}^{\mathrm{T}} \diamond {\mathbf{C}^{(\mathbf{S})}}^{\mathrm{T}}\right)^{\mathrm{T}}\right) \geq 1$, which is equivalent to the next equality:

$$\left\{\forall m, r,\ \exists p \,\middle|\, c^{(\mathbf{H})}_{p,m}\, c^{(\mathbf{S})}_{p,r} = 1\right\}. \tag{3.69}$$

Verify that (3.69) means that each r-th data stream has to be transmitted by each m-th antenna during at least one time block p. Consequently, (3.69) has a physical interpretation which allows to construct the allocation matrices.

It is important to emphasize that the characteristics of wireless communication system interfere or can even restrict the choice of allocation matrices. As for example, the amount of available antennas per each time block, restriction of the transmission rate and required reliability of communication directly affect the dimensions and structure of allocation matrices. For this reason, it is convenient to relax the construction rule of allocation matrices by appropriately fixing the code tensor.

We can exploit the uniqueness results for CP model in order to deduce another conditions to ensure channel and symbol estimation for the TST coding. Thus, let us consider the received signal tensor decomposed as a constrained CP model which was developed in Section 3.1 and recalled by convenience hereafter

$$\mathcal{X} = \left[\!\left[\overline{\mathbf{A}}^{(1)}, \overline{\mathbf{A}}^{(2)}, \overline{\mathbf{A}}^{(3)}\right]\!\right], \tag{3.70}$$

with

$$\begin{aligned}\overline{\mathbf{A}} &= \mathbf{H} \otimes \mathbf{S} \in \mathbb{C}^{KN \times MR},\\ \overline{\mathbf{B}} &= \left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}} \in \mathbb{C}^{P \times MR},\\ \overline{\mathbf{C}} &= \mathbf{W}_{J \times MR} \in \mathbb{C}^{J \times MR}.\end{aligned} \tag{3.71}$$

Theorems 2.9 and 2.10 proposed in Chapter 2 can be applied to extend the uniqueness conditions through the constrained CP model given in (3.70) and (3.71). Let us firstly employ Theorem 2.9 to derive the next theorem for the TST system.

Theorem 3.5. *Suppose the perfect knowledge of the code tensor $\mathcal{W}$ and the allocation matrices $\mathbf{C}^{(\mathbf{H})}$ and $\mathbf{C}^{(\mathbf{S})}$. Assuming that the code tensor is chosen such that $\mathbf{W}_{J \times MR}$ is full column-rank, if the following condition is satisfied*

$$\operatorname{rank}(\mathbf{H}) \operatorname{rank}(\mathbf{S}) + k\left(\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}}\right) \geq MR + 2, \tag{3.72}$$

we can guarantee the uniqueness of $\mathbf{S}$ and $\mathbf{H}$ assuming the knowledge of one symbol.

Proof: Observe that $\operatorname{rank}(\mathbf{H} \otimes \mathbf{S}) = \operatorname{rank}(\mathbf{H}) \operatorname{rank}(\mathbf{S})$. Applying directly (2.48), we have that if the condition (3.72) holds then $\mathbf{H} \otimes \mathbf{S}$ is unique up to permutation and scaling of the columns. The knowledge of $\mathcal{W}$, $\mathbf{C}^{(\mathbf{H})}$ and $\mathbf{C}^{(\mathbf{S})}$ allow to eliminate ambiguities of the estimation of $\mathbf{H} \otimes \mathbf{S}$ and it is possible to separate $\mathbf{S}$ and $\mathbf{H}$ using the knowledge of just one symbol, thanks to the special structure of the Kronecker product. ∎

Remark 3.2.

- *It is implicity in* (3.72) *the following restriction $k\left(\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}}\right) \geq 2$, which means that $\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}$ does not have collinear rows. Theorem 3.5 does not impose that $\mathbf{S}$ and $\mathbf{H}$ are full column-rank, but the condition on the allocation matrices is stronger than the one given in Theorem 3.4.*
- *The sufficient condition* (3.72) *is more restrictive than the one given in Theorem 3.4 when both $\mathbf{S}$ and $\mathbf{H}$ are full column-rank, implying $N \geq R$ and $K \geq M$. Thus, the most important aspect of the Theorem 3.5 lies in the fact that it allows the use of more transmit antennas (M) than receiver antennas (K), i.e. $M > K$.*

We can obtain another uniqueness condition for the TST system applying Theorem 2.10. Let us enunciate the following theorem.

Theorem 3.6. *Suppose the perfect knowledge of the code tensor $\mathcal{W}$ and the allocation matrices $\mathbf{C}^{(\mathbf{H})}$ and $\mathbf{C}^{(\mathbf{S})}$. Assuming that the code tensor is chosen such that $\mathbf{W}_{J \times MR}$ is full column-rank, if the following condition is satisfied*

$$k(\mathbf{H} \otimes \mathbf{S}) + k\left(\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}}\right) \geq MR + 1, \tag{3.73}$$

we can guarantee the uniqueness of $\mathbf{S}$ and $\mathbf{H}$ assuming the knowledge of one symbol.

Proof: According to Theorem 2.10 for $\mathbf{W}_{J\times MR}$ is full column-rank, we directly obtain the inequality (3.73). Finally, $\mathbf{S}$ and $\mathbf{H}$ are correctly recover thanks to the structure of Kronecker product and the knowledge of $\mathcal{W}$, $\mathbf{C}^{(\mathbf{H})}$, $\mathbf{C}^{(\mathbf{S})}$ and at least one symbol. ∎

Remark 3.3.

- *When* $\mathbf{S}$ *and* $\mathbf{H}$ *are full column-rank implying* $N \geq R$ *and* $K \geq M$*, Theorem 3.6 leads to Theorem 3.4.*
- *As* $k(\mathbf{H} \otimes \mathbf{S}) \leq \operatorname{rank}(\mathbf{H} \otimes \mathbf{S}) = \operatorname{rank}(\mathbf{H}) \operatorname{rank}(\mathbf{S}) \leq \min(N, R) \min(K, M)$*, hence the condition* (3.73) *of Theorem 3.6 is less constrained than the condition* (3.72) *of Theorem 3.5.*
- *If* $\mathbf{S}$ *has collinear columns, then* $k(\mathbf{H} \otimes \mathbf{S}) = 1$ *and as consequence,* $\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}$ *has to be full row-rank in order to satisfy* (3.73)*.*

According to the last remark, Theorem 3.5 is unnecessary and Theorem 3.4 is a particular case of Theorem 3.6. The interesting advantage of Theorem 3.6 is that $\mathbf{S}$ and $\mathbf{H}$ do not need to be full column-rank, which allows the use of more transmit antennas (M) than receiver antennas (K), i.e. $M > K$.

3.4.2 STF system

Analogously to the uniqueness analysis for the TST coding, we study the uniqueness condition from the convenient structure of the unfolded matrix $\mathbf{X}_{KN\times FP}$ given in (3.38).

Let us consider $\hat{\mathbf{S}}$ and $\hat{\mathbf{H}}_{K\times FM}$ as alternative solutions that satisfy (3.38) and assume $\hat{\mathbf{S}} = \mathbf{S}\mathbf{U}$ and $\hat{\mathbf{H}}_{K\times FM} = \mathbf{H}_{K\times FM}\mathbf{V}$ with $\mathbf{U} \in \mathbb{C}^{R\times R}$ and $\mathbf{V} \in \mathbb{C}^{FM\times FM}$ nonsingular matrices. Thus, we can rewrite (3.38) as

$$\mathbf{X}_{KN\times FP} = \left(\hat{\mathbf{H}}_{K\times FM} \otimes \hat{\mathbf{S}}\right) \mathbf{T}_{FMR\times FP} = \left(\mathbf{H}_{K\times FM} \otimes \mathbf{S}\right) \left(\mathbf{V} \otimes \mathbf{U}\right) \mathbf{T}_{FMR\times FP}. \tag{3.74}$$

We can prove the uniqueness model of (3.27) from (3.74) and the next theorem shows a sufficient condition that ensure the uniqueness of this model.

Theorem 3.7. *Suppose that* $\mathbf{S}$ *and* $\mathbf{H}_{K\times FM}$ *are full column-rank, and the perfect knowledge of the code matrix* $\mathbf{W}$ *and the allocation tensors* $\mathcal{C}^{(\mathbf{S})}$ *and* $\mathcal{C}^{(\mathcal{H})}$*. If we choose* $\mathbf{W}$ *and* $\mathcal{C}^{(\mathbf{S})}$ *and* $\mathcal{C}^{(\mathcal{H})}$ *such that* $w_{m,r} \neq 0$ *for all* $m \in \{1, ..., M\}$ *and* $r \in \{1, ..., R\}$*, and* $\mathbf{C}_{f..}^{(\mathcal{H})^{\mathrm{T}}} \diamond \mathbf{C}_{f..}^{(\mathbf{S})^{\mathrm{T}}}$ *full row-rank for all* $f \in \{1, ..., F\}$ *implying* $MR \leq P$*, then we can uniquely estimate* $\mathbf{S}$ *and* $\mathbf{H}_{K\times FM}$ *up to a scalar factor* α*, i.e.*

$$\mathbf{S} = \alpha\, \hat{\mathbf{S}}, \qquad \mathbf{H}_{K\times FM} = \frac{1}{\alpha}\, \hat{\mathbf{H}}_{K\times FM}. \tag{3.75}$$

Proof: If $\mathbf{S}$, $\mathbf{H}_{K\times FM}$ and $\mathbf{T}_{FMR\times FP}^{\mathrm{T}}$ are full column-rank, then (3.74) can be rewritten as

$$\mathbf{V} \otimes \mathbf{U} = \mathbf{I}_{FMR}. \tag{3.76}$$

The only solution for (3.76) happens when both matrices $\mathbf{U}$ and $\mathbf{V}$ are identity matrices up to scalar factors that compensate each other, which leads to (3.75). From (3.39) and (3.40), if the elements of the code matrix are nonzero and $\mathbf{C}_{f..}^{(\mathcal{H})^{\mathrm{T}}} \diamond \mathbf{C}_{f..}^{(\mathbf{S})^{\mathrm{T}}}$ is full row-rank for all f, then $\mathbf{T}_{FMR\times FP}$ will be full row-rank as well. It concludes the proof. ∎

Remark 3.4.

- *Analogously to the TST system, assuming that the channel coefficients have entries i.i.d. and a continuous distribution,* $\mathbf{H}_{K\times FM}$ *is almost surely full column-rank if* $K \geq FM$*. Thus,* $K \geq FM$ *is a restrictive condition because we have always to use* F *times more receive antennas than transmit antennas.*
- *Theorem 3.7 provides a more practical uniqueness condition for a generalized PT-2 model than the conditions in [62].*
- *Differently to Theorem 3.7 for the STF systems, the uniqueness results for the TST systems permit a more relaxed condition over the allocation structures by appropriately choosing the code and it is not required a large amount of data blocks* P *as for the STF uniqueness* $(MR \leq P)$*.*

3.5 Generalization of TST systems to multiuser case

Let us generalize the TST coding for Q users as illustrated in Figure 3.3. Each user transmits R input data streams using M different antennas and each data stream is composed of N information symbols. We consider two allocation matrices $\mathbf{C}^{(\mathbf{H},q)} \in \mathbb{R}^{P\times M}$ and $\mathbf{C}^{(\mathbf{S},q)} \in \mathbb{R}^{P\times R}$ for each user q, which allocate transmit antennas and data streams to each block p.

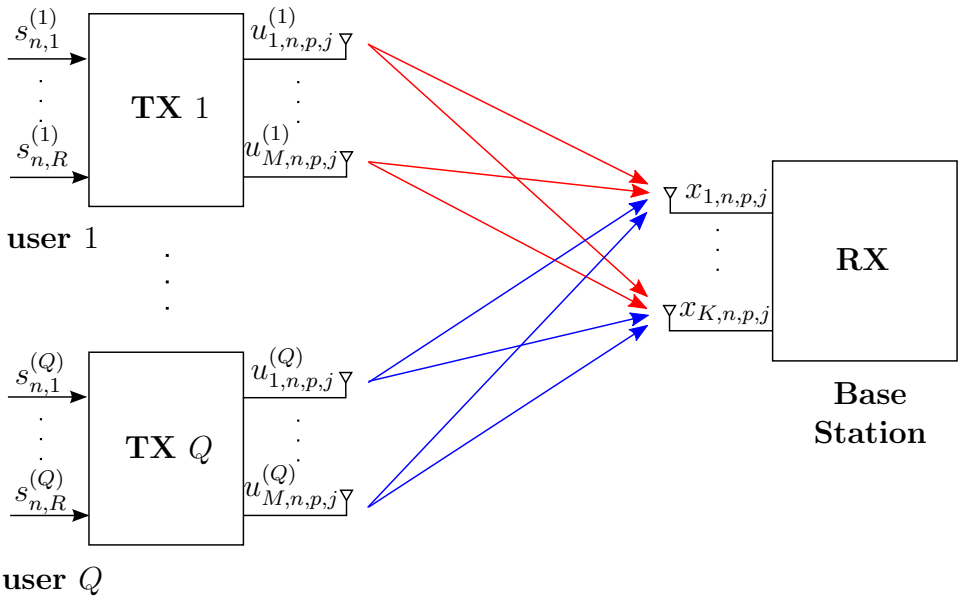


Figure 3.3: Uplink processing based on the TST coding with resource allocations.

From (3.1), the signal associated with the q-th user and the j-th chip transmitted from the m-th antenna during the n-th symbol period of the p-th block is given by

$$u^{(q)}_{m,n,p,j} = \sum_{r=1}^{R} g^{(q)}_{m,r,p,j}\, s^{(q)}_{n,r} \tag{3.77}$$

with

$$g^{(q)}_{m,r,p,j} \triangleq w^{(q)}_{m,r,j}\, c^{(\mathbf{H},q)}_{p,m}\, c^{(\mathbf{S},q)}_{p,r}. \tag{3.78}$$

In the absence of noise, the received signal associated with the q-th user can be written using (3.5) as

$$x_{k,n,p,j}^{(q)} = \sum_{m=1}^{M} h_{k,m}^{(q)}\, u_{m,n,p,j}^{(q)} = \sum_{m=1}^{M}\sum_{r=1}^{R} g_{m,r,p,j}^{(q)}\, h_{k,m}^{(q)}\, s_{n,r}^{(q)}. \tag{3.79}$$

We obtain the tensor of the overall received signal $\mathcal{X} \in \mathbb{C}^{K\times N\times P\times J}$ by summing the Q contributions and using the n-mode product definition (2.13)

$$\mathcal{X} \triangleq \sum_{q=1}^{Q} \mathcal{X}^{(q)} = \sum_{q=1}^{Q} \mathcal{G}^{(q)} \times_1 \mathbf{H}^{(q)} \times_2 \mathbf{S}^{(q)}. \tag{3.80}$$

Remark that each received signal tensor $\mathcal{X}^{(q)} \in \mathbb{C}^{K\times N\times P\times J}$ satisfies the PARATUCK-(2,4) model.

From (3.11) and (3.16), we can express three matrix unfoldings of $\mathcal{X}$ as follows

$$\begin{aligned}\mathbf{X}_{PJK\times N} &\triangleq \sum_{q=1}^{Q} \mathbf{X}_{PJK\times N}^{(q)} = \left[\mathbf{I}_{PJ}\otimes\mathbf{H}^{(1)} \;\cdots\; \mathbf{I}_{PJ}\otimes\mathbf{H}^{(Q)}\right]\mathbf{G}_{QPJM\times QR}\,\mathbf{S}^{\mathrm{T}} \\ &= (\mathbf{\Omega}\,|\otimes|\,\mathbf{H}_{K\times QM})\,\mathbf{G}_{QPJM\times QR}\,\mathbf{S}_{QR\times N} \in \mathbb{C}^{PJK\times N},\end{aligned} \tag{3.81}$$

$$\begin{aligned}\mathbf{X}_{PJN\times K} &\triangleq \sum_{q=1}^{Q} \mathbf{X}_{PJN\times K}^{(q)} = \left[\mathbf{I}_{PJ}\otimes\mathbf{S}^{(1)} \;\cdots\; \mathbf{I}_{PJ}\otimes\mathbf{S}^{(Q)}\right]\mathbf{G}_{QPJR\times QM}\,\mathbf{H}^{\mathrm{T}} \\ &= (\mathbf{\Omega}\,|\otimes|\,\mathbf{S}_{N\times QR})\,\mathbf{G}_{QPJR\times QM}\,\mathbf{H}_{QM\times K} \in \mathbb{C}^{PJN\times K},\end{aligned} \tag{3.82}$$

$$\begin{aligned}\mathbf{X}_{KN\times JP} &\triangleq \sum_{q=1}^{Q} \mathbf{X}_{KN\times JP}^{(q)} = \left[\mathbf{H}^{(1)}\otimes\mathbf{S}^{(1)} \;\cdots\; \mathbf{H}^{(Q)}\otimes\mathbf{S}^{(Q)}\right]\mathbf{G}_{QMR\times JP} \\ &= (\mathbf{H}_{K\times QM}\,|\otimes|\,\mathbf{S}_{N\times QR})\,\mathbf{G}_{QMR\times JP} \in \mathbb{C}^{KN\times JP},\end{aligned} \tag{3.83}$$

with $\mathbf{S}_{N\times QR} \triangleq \left[\mathbf{S}^{(1)} \;\cdots\; \mathbf{S}^{(Q)}\right] \in \mathbb{C}^{N\times QR}$, $\mathbf{H}_{K\times QM} \triangleq \left[\mathbf{H}^{(1)} \;\cdots\; \mathbf{H}^{(Q)}\right] \in \mathbb{C}^{K\times QM}$, $\mathbf{\Omega} \triangleq \mathbf{1}_Q^{\mathrm{T}}\otimes\mathbf{I}_{PJ} \in \mathbb{C}^{PJ\times QPJ}$,

$$\mathbf{G}_{QPJM\times QR} \triangleq \underset{\in\mathbb{C}^{QPJM\times QR}}{\begin{bmatrix}\mathbf{G}_{PJM\times R}^{(1)} & & \mathbf{0}\\ & \ddots & \\ \mathbf{0} & & \mathbf{G}_{PJM\times R}^{(Q)}\end{bmatrix}}, \quad \mathbf{G}_{QPJR\times QM} \triangleq \underset{\in\mathbb{C}^{QPJR\times QM}}{\begin{bmatrix}\mathbf{G}_{PJR\times M}^{(1)} & & \mathbf{0}\\ & \ddots & \\ \mathbf{0} & & \mathbf{G}_{PJR\times M}^{(Q)}\end{bmatrix}}, \tag{3.84}$$

$$\mathbf{G}_{QMR\times JP} \triangleq \left[{\mathbf{G}_{MR\times JP}^{(1)}}^{\mathrm{T}} \;\cdots\; {\mathbf{G}_{MR\times JP}^{(Q)}}^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{C}^{QMR\times JP}. \tag{3.85}$$

Each matrix $\mathbf{G}_{MR\times JP}^{(q)}$ can be written using (3.18) and (3.19) as

$$\mathbf{G}_{MR\times JP}^{(q)} = \left(\mathbf{W}_{J\times MR}^{(q)} \diamond \left({\mathbf{C}^{(\mathbf{H},q)}}^{\mathrm{T}} \diamond {\mathbf{C}^{(\mathbf{S},q)}}^{\mathrm{T}}\right)^{\mathrm{T}}\right)^{\mathrm{T}} \tag{3.86}$$

with

$$\mathbf{W}_{J\times MR}^{(q)} \triangleq \left[\operatorname{vec}\left(\mathbf{W}_{\cdot\cdot 1}^{(q)^{\mathrm{T}}}\right) \ \cdots \ \operatorname{vec}\left(\mathbf{W}_{\cdot\cdot J}^{(q)^{\mathrm{T}}}\right)\right]^{\mathrm{T}} \in \mathbb{C}^{J\times MR}. \tag{3.87}$$

Applying (3.86) to (3.85) and using the definition of the Khatri-Rao product (A.1), we have

$$\begin{aligned}\mathbf{G}_{QMR\times JP} &= \left[\mathbf{W}_{J\times MR}^{(1)} \diamond \left(\mathbf{C}^{(\mathbf{H},1)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S},1)^{\mathrm{T}}}\right)^{\mathrm{T}} \ \cdots \ \mathbf{W}_{J\times MR}^{(Q)} \diamond \left(\mathbf{C}^{(\mathbf{H},Q)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S},Q)^{\mathrm{T}}}\right)^{\mathrm{T}}\right]^{\mathrm{T}} \\ &= \left(\mathbf{W}_{J\times QMR} \diamond \left[\left(\mathbf{C}^{(\mathbf{H},1)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S},1)^{\mathrm{T}}}\right)^{\mathrm{T}} \ \cdots \ \left(\mathbf{C}^{(\mathbf{H},Q)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S},Q)^{\mathrm{T}}}\right)^{\mathrm{T}}\right]\right)^{\mathrm{T}},\end{aligned} \tag{3.88}$$

where

$$\mathbf{W}_{J\times QMR} \triangleq \left[\mathbf{W}_{J\times MR}^{(1)} \ \cdots \ \mathbf{W}_{J\times MR}^{(Q)}\right] \in \mathbb{C}^{J\times QMR}. \tag{3.89}$$

Let us define the selection matrix

$$\tilde{\mathbf{\Phi}}_{QMR} \triangleq \left[\begin{array}{cccc} \mathbf{\Phi}_1 & \mathbf{\Phi}_2 & \cdots & \mathbf{\Phi}_Q \end{array}\right] \in \mathbb{C}^{QMQR\times QMR}, \tag{3.90}$$

where

$$\mathbf{\Phi}_q \triangleq \left[\begin{array}{cccc} \mathbf{E}_{(q-1)MQ+q} & \mathbf{E}_{(q-1)MQ+Q+q} & \cdots & \mathbf{E}_{(q-1)MQ+(M-1)Q+q} \end{array}\right] \in \mathbb{C}^{QMQR\times MR} \tag{3.91}$$

and $\mathbf{E}_l$ is a matrix $QMQR\times R$ with a identity matrix in the l-th block $R\times R$ and zeros elsewhere, i.e.

$$\mathbf{E}_l \triangleq \left[\underbrace{\mathbf{0}_R \ \cdots \ \mathbf{0}_R}_{R\times(l-1)R} \quad \mathbf{I}_R \quad \underbrace{\mathbf{0}_R \ \cdots \ \mathbf{0}_R}_{R\times(QMQ-l)R}\right]^{\mathrm{T}}. \tag{3.92}$$

From the selection matrix $\tilde{\mathbf{\Phi}}_{QMR}$ defined in (3.90), we can express the following equality

$$\begin{aligned}\left[\left(\mathbf{C}^{(\mathbf{H},1)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S},1)^{\mathrm{T}}}\right)^{\mathrm{T}} \ \cdots \ \left(\mathbf{C}^{(\mathbf{H},Q)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{H},Q)^{\mathrm{T}}}\right)^{\mathrm{T}}\right] &= \left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}} \tilde{\mathbf{\Phi}}_{QMR} \\ &\in \mathbb{C}^{P\times QMR}\end{aligned} \tag{3.93}$$

where

$$\mathbf{C}^{(\mathbf{H})} \triangleq \left[\mathbf{C}^{(\mathbf{H},1)} \ \cdots \ \mathbf{C}^{(\mathbf{H},Q)}\right] \in \mathbb{C}^{P\times QM}, \tag{3.94}$$

$$\mathbf{C}^{(\mathbf{S})} \triangleq \left[\mathbf{C}^{(\mathbf{S},1)} \ \cdots \ \mathbf{C}^{(\mathbf{S},Q)}\right] \in \mathbb{C}^{P\times QR} \tag{3.95}$$

represent the global allocation matrices which concatenate the antenna and stream allocation matrices for all users. Observe that $\tilde{\mathbf{\Phi}}_{QMR}$ selects only QMR columns of $\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}} \in \mathbb{C}^{P\times QMQR}$.

Applying (3.93) to (3.88), we can rewrite (3.88) as

$$\mathbf{G}_{QMR\times JP} = \left(\mathbf{W}_{J\times QMR} \diamond \left(\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}} \tilde{\mathbf{\Phi}}_{QMR}\right)\right)^{\mathrm{T}}. \tag{3.96}$$

Observe that the overall received signal tensor (3.80) can be rewritten as a constrained CP decomposition with the purpose of applying the uniqueness results of the CP model.

TST coding with multiple users modeled by a constrained CP model

Consider a third-order CP model $\mathcal{X} \in \mathbb{C}^{KN\times P\times J}$ whose the loading matrices are given by

$$\begin{aligned}
\overline{\mathbf{A}} &= \mathbf{H}_{K\times QM} \,|\otimes|\, \mathbf{S}_{N\times QR} \in \mathbb{C}^{KN\times QMR},\\
\overline{\mathbf{B}} &= \left[\left(\mathbf{C}^{(\mathbf{H},1)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S},1)^{\mathrm{T}}}\right)^{\mathrm{T}} \;\cdots\; \left(\mathbf{C}^{(\mathbf{H},Q)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{H},Q)^{\mathrm{T}}}\right)^{\mathrm{T}}\right]\\
&= \left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}} \tilde{\mathbf{\Phi}}_{QMR} \in \mathbb{C}^{P\times QMR},\\
\overline{\mathbf{C}} &= \left[\mathbf{W}^{(1)}_{J\times MR} \;\cdots\; \mathbf{W}^{(Q)}_{J\times MR}\right]\\
&= \mathbf{W}_{J\times QMR} \in \mathbb{C}^{J\times MR},
\end{aligned} \tag{3.97}$$

From (2.21) and (3.97), we obtain the same unfolded matrix given in (3.83) by the matrix factors $\overline{\mathbf{A}}$, $\overline{\mathbf{B}}$ and $\overline{\mathbf{C}}$.

Uniqueness analysis

We study the uniqueness conditions for the TST system considering Q users analogously to the development of Theorem 3.4. The unfolded matrix $\mathbf{X}_{KN\times JP}$ given in (3.83) is used to analyze the uniqueness condition of this new model.

From the selection matrix defined in (3.90), we can rewrite the unfolded matrix $\mathbf{X}_{KN\times JP}$ using the following expression which relates the Kronecker and partition-wise Kronecker products

$$\mathbf{H}_{K\times QM} \,|\otimes|\, \mathbf{S}_{N\times QR} = (\mathbf{H}_{K\times QM} \otimes \mathbf{S}_{N\times QR})\, \tilde{\mathbf{\Phi}}_{QMR}. \tag{3.98}$$

Note that $\tilde{\mathbf{\Phi}}_{QMR}$ selects only QMR columns of $\mathbf{H}_{K\times QM}\otimes\mathbf{S}_{N\times QR} \in \mathbb{C}^{KN\times QMQR}$ and the resultant matrix is represented by $\mathbf{H}_{K\times QM} \,|\otimes|\, \mathbf{S}_{N\times QR} \in \mathbb{C}^{KN\times QMR}$. Applying (3.98) to (3.83) gives

$$\mathbf{X}_{KN\times JP} = (\mathbf{H}_{K\times QM} \,|\otimes|\, \mathbf{S}_{N\times QR})\, \mathbf{G}_{QMR\times JP} = (\mathbf{H}_{K\times QM} \otimes \mathbf{S}_{N\times QR})\, \tilde{\mathbf{\Phi}}_{QMR}\, \mathbf{G}_{QMR\times JP}. \tag{3.99}$$

Considering $\hat{\mathbf{S}}_{N\times QR} = \left[\hat{\mathbf{S}}^{(1)} \;\cdots\; \hat{\mathbf{S}}^{(Q)}\right]$ and $\hat{\mathbf{H}}_{K\times QM} = \left[\hat{\mathbf{H}}^{(1)} \;\cdots\; \hat{\mathbf{H}}^{(Q)}\right]$ as alternative solutions that satisfy (3.83), we can write $\hat{\mathbf{S}}_{N\times QR} = \mathbf{S}_{N\times QR}\,\mathbf{U}$ and $\hat{\mathbf{H}}_{K\times QM} = \mathbf{H}_{K\times QM}\,\mathbf{V}$, with

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}^{(1,1)} & \cdots & \mathbf{U}^{(1,Q)} \\ \vdots & \ddots & \vdots \\ \mathbf{U}^{(Q,1)} & \cdots & \mathbf{U}^{(Q,Q)} \end{bmatrix} \in \mathbb{C}^{QR\times QR}, \qquad \mathbf{V} = \begin{bmatrix} \mathbf{V}^{(1,1)} & \cdots & \mathbf{V}^{(1,Q)} \\ \vdots & \ddots & \vdots \\ \mathbf{V}^{(Q,1)} & \cdots & \mathbf{V}^{(Q,Q)} \end{bmatrix} \in \mathbb{C}^{QM\times QM} \tag{3.100}$$

non-singular. Thus, the unfolded matrix $\mathbf{X}_{KN\times JP}$ can be rewritten using the Kronecker property (A.5) as

$$(\mathbf{H}_{K\times QM} \otimes \mathbf{S}_{N\times QR})\, (\mathbf{V} \,|\otimes|\, \mathbf{U})\, \mathbf{G}_{QMR\times JP} = (\mathbf{H}_{K\times QM} \otimes \mathbf{S}_{N\times QR})\, \tilde{\mathbf{\Phi}}_{QMR}\, \mathbf{G}_{QMR\times JP}. \tag{3.101}$$

Theorem 3.8. *Suppose that $\mathbf{S}_{N\times QR}$ and $\mathbf{H}_{K\times QM}$ are full column-rank and the perfect knowledge of the code tensor $\mathcal{W}^{(q)}$ and the allocation matrices $\mathbf{C}^{(\mathbf{H},q)}$ and $\mathbf{C}^{(\mathbf{S},q)}$ for all user. If the code tensor for each user is chosen such that $\mathbf{W}_{J\times QMR}$ is full column-rank, implying $J \geq QRM$, and*

$\mathbf{C}^{(\mathbf{H},q)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S},q)^{\mathrm{T}}}$ *has no zero-rows for all* q*, then the estimates of* $\mathbf{S}^{(q)}$ *and* $\mathbf{H}^{(q)}$*, for* $q = 1, ..., Q$*, are unique up to a scalar factor, i.e.*

$$\mathbf{S}^{(q)} = \alpha_q \hat{\mathbf{S}}^{(q)}, \quad \mathbf{H}^{(q)} = \frac{1}{\alpha_q} \hat{\mathbf{H}}^{(q)}. \tag{3.102}$$

Proof: If $\mathbf{S}_{N\times QR}$ and $\mathbf{H}_{K\times QM}$ are full column-rank and $\mathbf{G}_{QMR\times JP}$ is full row-rank, then (3.101) can be rewritten as

$$\begin{aligned} \mathbf{V} \,|\otimes|\, \mathbf{U} = \tilde{\mathbf{\Phi}}_{QMR} \qquad &\Longrightarrow \\ \begin{bmatrix} \mathbf{V}^{(1,q)} \\ \vdots \\ \mathbf{V}^{(Q,q)} \end{bmatrix} \otimes \begin{bmatrix} \mathbf{U}^{(1,q)} \\ \vdots \\ \mathbf{U}^{(Q,q)} \end{bmatrix} &= \mathbf{\Phi}_q \in \mathbb{C}^{QMQR\times MR} \\ &= \begin{bmatrix} \mathbf{E}_{(q-1)MQ+q} & \mathbf{E}_{(q-1)MQ+Q+q} & \cdots & \mathbf{E}_{(q-1)MQ+(M-1)Q+q} \end{bmatrix}, \end{aligned} \tag{3.103}$$

which, from the definition given in (3.92), results in

$$\begin{aligned} & v_{m,m}^{(q,q)} \mathbf{U}^{(q,q)} = \mathbf{I}_R, \quad \forall m = 1, ..., M, \forall q = 1, ..., Q, \\ \Longrightarrow \quad & \mathbf{V}^{(q,q)} \otimes \mathbf{U}^{(q,q)} = \mathbf{I}_{MR}, \quad \forall q = 1, ..., Q, \end{aligned} \tag{3.104}$$

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}^{(1,1)} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{V}^{(Q,Q)} \end{bmatrix}, \qquad \mathbf{U} = \begin{bmatrix} \mathbf{U}^{(1,1)} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{U}^{(Q,Q)} \end{bmatrix} \tag{3.105}$$

and consequently, it gets

$$\mathbf{S}^{(q)} = \hat{\mathbf{S}}^{(q)} \mathbf{U}^{(q,q)}, \qquad \mathbf{H}^{(q)} = \hat{\mathbf{H}}^{(q)} \mathbf{V}^{(q,q)}. \tag{3.106}$$

The only solution for $\mathbf{V}^{(q,q)} \otimes \mathbf{U}^{(q,q)} = \mathbf{I}_{MR}$, for $q = 1, ..., Q$, happens when both matrices $\mathbf{U}^{(q,q)}$ and $\mathbf{V}^{(q,q)}$ are identity matrices up to scalar factors that compensate each other, i.e. $\mathbf{U}^{(q,q)} = \alpha_q \mathbf{I}_R$ and $\mathbf{V}^{(q,q)} = 1/\alpha_q \mathbf{I}_M$, which leads to (3.102).

In order to guaranty the full row-rank property of $\mathbf{G}_{QMR\times JP}$, let us consider the matrix unfolding given in (3.96).

According to Lemma 2.2 and from (3.96), the matrix $\mathbf{G}_{JP\times QMR} = \mathbf{G}_{QMR\times JP}^{\mathrm{T}}$ is full column-rank if $\mathbf{W}_{J\times QMR}$, $\mathbf{C}^{(\mathbf{S})} \in \mathbb{C}^{P\times QR}$ and $\mathbf{C}^{(\mathbf{H})} \in \mathbb{C}^{P\times QM}$ satisfy the following condition

$$k(\mathbf{W}_{J\times QMR}) + k\left(\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}} \tilde{\mathbf{\Phi}}_{QMR}\right) \geq QMR + 1. \tag{3.107}$$

Assuming that $\mathbf{C}^{(\mathbf{H},q)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S},q)^{\mathrm{T}}}$ has non zero-rows for all q and from (3.93), we obtain $k\left(\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}} \tilde{\mathbf{\Phi}}_{QMR}\right) \geq 1$. Finally, if the code tensors for each users are chosen such that $\mathbf{W}_{J\times QMR}$ defined in (3.89) is full column-rank, then (3.107) is satisfied and $\mathbf{G}_{JP\times QMR}$ is full column-rank. ∎

Remark 3.5.

- *According to Theorem 3.8, the channel and symbol estimates do not present permutation ambiguity and only scalar ambiguity which can be eliminated by the knowledge of one symbol for each user q.*
- *As the channel coefficients $h_{k,m}^{(q)}$ are i.i.d and have a continuous distribution, $\mathbf{H}_{K\times QM}$ is full column-rank with probability 1 when $K \geq QM$.*
- *Choosing the code tensor for each user such that $\mathbf{W}_{J\times QMR}$ is full column-rank, we obtain the same condition on the allocation matrices given by Theorem 3.4 for each user.*

When all users employ the same code tensor, i.e. $\mathcal{W}^{(1)} = \ldots = \mathcal{W}^{(Q)}$, with non-zero entries, $\mathbf{W}_{J\times QMR}$ has collinear columns implying $k(\mathbf{W}_{J\times QMR}) = 1$. Consequently, the allocation matrices for all users have to be chosen such that $\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}} \tilde{\mathbf{\Phi}}_{QMR}$ is full column-rank with the purpose of still ensuring the uniqueness condition by Theorem 3.8.

If we set the same allocation matrices for all users, i.e. $\mathbf{C}^{(\mathbf{H},1)} = \ldots = \mathbf{C}^{(\mathbf{H},Q)}$ and $\mathbf{C}^{(\mathbf{S},1)} = \ldots = \mathbf{C}^{(\mathbf{S},Q)}$, then $\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}} \tilde{\mathbf{\Phi}}_{QMR}$ has collinear columns implying $k\left(\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}} \tilde{\mathbf{\Phi}}_{QMR}\right) = 1$. In order to still guarantee the uniqueness in accordance with Theorem 3.8, the code tensor are chosen such that $\mathbf{W}_{J\times QMR}$ is full column-rank.

Rewriting the multiuser TST system as a constrained CP model whose the loading matrices are given in (3.97), we can exploit the uniqueness results analogously to the case with one user. Both theorems 3.5 and 3.6 are directly applied to deduce the following theorems.

Theorem 3.9. *Suppose the perfect knowledge of the code tensor $\mathcal{W}^{(q)}$ and the allocation matrices $\mathbf{C}^{(\mathbf{H},q)}$ and $\mathbf{C}^{(\mathbf{S},q)}$ for all users. Assuming that the code tensor is chosen such that $\mathbf{W}_{J\times QMR}$ is full column-rank, if the following condition is satisfied*

$$\operatorname{rank}(\mathbf{H}_{K\times QM}) \operatorname{rank}(\mathbf{S}_{N\times QR}) + k\left(\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}} \tilde{\mathbf{\Phi}}_{QMR}\right) \geq MR + 2, \tag{3.108}$$

we can guarantee the uniqueness of $\mathbf{S}_{N\times QR}$ and $\mathbf{H}_{K\times QM}$ assuming the knowledge of one symbol per user.

Theorem 3.10. *Suppose the perfect knowledge of the code tensor $\mathcal{W}^{(q)}$ and the allocation matrices $\mathbf{C}^{(\mathbf{H},q)}$ and $\mathbf{C}^{(\mathbf{S},q)}$ for all users. Assuming that the code tensor is chosen such that $\mathbf{W}_{J\times QMR}$ is full column-rank, if the following condition is satisfied*

$$k(\mathbf{H}_{K\times QM} \otimes \mathbf{S}_{N\times QR}) + k\left(\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}} \tilde{\mathbf{\Phi}}_{QMR}\right) \geq MR + 1, \tag{3.109}$$

we can guarantee the uniqueness of $\mathbf{S}_{N\times QR}$ and $\mathbf{H}_{K\times QM}$ assuming the knowledge of one symbol per user.

Both theorems 3.9 and 3.10 allow to ensure the uniqueness of all transmitted symbols and channel coefficients even when $\mathbf{S}_{N\times QR}$ and $\mathbf{H}_{K\times QM}$ are not full column-rank. Hence, it is not required to impose $N \geq QR$ and $K \geq QM$ as happens for Theorem 3.8.

Analogously to the single user case, the condition (3.109) of Theorem 3.10 is less constrained than the condition (3.108) of Theorem 3.9. When the matrices $\mathbf{S}_{N\times QR}$ and $\mathbf{H}_{K\times QM}$ are full column-rank, Theorem 3.10 leads to Theorem 3.8.

It is important to emphasize that we have considered the same number of transmit antennas M for each user. However, we can appropriately set the antenna allocation matrix $\mathbf{C}^{(\mathbf{H},q)}$ for each user q in order to obtain different numbers of received antennas per user (lower than M).

Remark that we can also extend the STF system for multiple users in the transmission analogously to the generalization proposed for the TST system, as shown in Appendix B.

3.6 Resume of uniqueness results

A global synthesis of the uniqueness conditions for both systems are presented in Table 3.1. As discussed previously, the uniqueness conditions for the TST system are satisfied by restrictions on the code tensor and the allocation matrices. Two options to satisfy the uniqueness condition for the TST system are proposed in Table 3.1: by imposing a stronger condition over the allocations matrices ($\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}$ or $\tilde{\mathbf{\Phi}}^{\mathrm{T}}_{QMR}\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)$ must be full row-rank implying $P \geq MR$ and $P \geq QMR$, respectively) and over the code tensor ($\mathbf{W}_{J\times MR}$ or $\mathbf{W}_{J\times QMR}$ must be full column-rank implying $J \geq MR$ and $J \geq QMR$, respectively). Remark that it is still possible to satisfy the uniqueness condition by employing other structures for the code and allocation matrices.

Table 3.1: Summary of the uniqueness results for the TST and STF systems.

TST system		**STF system**
multiple users	**one user**	
$P \geq QMR$	$P \geq MR$	$N \geq R$, $K \geq FM$
$\tilde{\mathbf{\Phi}}^{\mathrm{T}}_{QMR}\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)$ full row-rank	$\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}$ full row-rank	$P \geq MR$
$w^{(q)}_{m,r} \neq 0, \forall q,m,r$	$w_{m,r} \neq 0, \forall m,r$	$\mathbf{C}^{(\mathcal{H})^{\mathrm{T}}}_{f..} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}_{f..}$ full row-rank, $\forall f$
$N \geq QR$, $K \geq QM$	$N \geq R$, $K \geq M$	$w_{m,r} \neq 0, \forall m,r$
$J \geq QMR$	$J \geq MR$	
$\sum_{p=1}^{P} c^{(\mathbf{S},q)}_{p,r} c^{(\mathbf{H},q)}_{p,m} \geq 1, \forall q,m,r$	$\sum_{p=1}^{P} c^{(\mathbf{S})}_{p,r} c^{(\mathbf{H})}_{p,m} \geq 1, \forall m,r$	
$\mathbf{W}_{J\times QMR}$ full column-rank	$\mathbf{W}_{J\times MR}$ full column-rank	

Contrary to the TST system, the uniqueness condition for the STF system is only satisfied by restriction on the allocation tensors with $\mathbf{C}^{(\mathcal{H})^{\mathrm{T}}}_{f..} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}_{f..}$ full row-rank for all $f \in \{1,\ldots,F\}$ (implying $P \geq MR$). The structure of the STF system does not allow the same flexibility than the one of the TST system by means of the third-order code tensor $\mathcal{W} \in \mathbb{C}^{M\times R\times J}$.

It is important to remark that the decomposition of the received signal tensor for the TST system as a constrained CP model allows to obtain a sufficient uniqueness condition more flexible by employing the results for CP model. The same reasoning is not possible for the STF system because the channel tensor $\mathcal{H} \in \mathbb{C}^{F\times K\times M}$ and the allocation tensors $\mathbf{C}^{(\mathcal{H})} \in \mathbb{C}^{F\times P\times M}$ $\mathbf{C}^{(\mathbf{S})} \in \mathbb{C}^{F\times P\times R}$ rely on the frequency diversity F.

According to Table 3.1, the proposed conditions can still ensure the model uniqueness when less receive antennas K than transmit antennas M are employed for the TST system. Differently, we have to use F times more receive antennas for the STF system.

Chapter 4

Semi-blind receivers

Several algorithms were proposed in the tensor decomposition context with the purpose of computing the matrix factors from which the original tensor was constructed or in another way, of recovering a CP model that best approximates this tensor. Brief overview of the main algorithms fitting the CP model and a comparison between several algorithms are presented in [104, 105]. In general terms, the convergence speed and the solution found for the parameter estimates rely on conditions over loading matrices and its dimensions. Most of these algorithms are developed to compute the CP decomposition. However, these algorithms may be exploited to recover other decompositions of interest and also to derive blind receivers for MIMO systems [50, 94, 106, 61, 58, 107, 62].

The proposed semi-blind receivers are based on the well-known Alternating Least Squares (ALS) and Levenberg-Marquardt (LM) algorithms, which iteratively estimate the symbol and channel matrices in the presence of additive noise. We also propose semi-blind receivers based on the structure of the Kronecker product, which allow to jointly estimate both matrices in only one iteration. The identifiability conditions for each method are established in order to ensure the symbol recovery and channel detection. The computational complexity for all algorithms is computed taking into account the most onerous operations per iteration.

In the simulation part, we present a performance analysis of the TST and STF systems by employing the zero-forcing (ZF) receivers assuming a perfect knowledge of the channel coefficients. The influence of the uniqueness and identifiability conditions proposed in this thesis on the performance of the ALS algorithm is analyzed. The performance of all proposed receivers are compared in terms of symbol recovery and convergence speed. Finally, we provide a comparison between the proposed TST coding with well-known tensor approaches, such as: Khatri-Rao Space-Time (KRST) [53] and Space-time Multiplexing (STM) [55] codes.

All receivers are developed by assuming the perfect knowledge of the coding and allocation structures at both transceiver and receiver. The received signal tensor $\mathcal{X}$ is corrupted by an additive white Gaussian noise $\mathcal{V}$, thereby $\tilde{\mathcal{X}} = \mathcal{X} + \mathcal{V}$ represents the noisy version of $\mathcal{X}$.

4.1 Alternating Least Squares (ALS)

The Alternating Least Squares algorithm, or simply denoted by ALS, was originally introduced in the tensor decomposition context by Harshman [51] and Carrol and Chang [52]. The ALS approach estimates each matrix factor by fixing the other matrix factors, which allows to convert a nonlinear optimization problem into several independent linear LS problems. The ALS solution is derived from the minimization of a cost function with respect to each matrix factor independently. The cost function is given by the squared errors between the received signal tensor model and the noisy received signals.

Observe that this approach does not solve the global problem since the estimation of one matrix is given by fixing the other matrix. For this, the ALS convergence to the global minimum can not be guaranteed. However, the cost function is strictly monotonic decreasing. The overall LS problem is reduced to alternating LS sub-problems, which permits that the ALS technique is simple and easy to be implemented.

The ALS algorithm is extensively used for CP decompositions, nevertheless it can be easily adapted for other decompositions of interest taking into account appropriately the matrix unfoldings. From matrix unfoldings of the received signal tensor deduced in Chapter 3, the ALS method is employed to estimate two matrices in alternating way: symbol $\mathbf{S} \in \mathbb{C}^{N\times R}$ and channel $\mathbf{H} \in \mathbb{C}^{K\times M}$ for the TST system or $\mathbf{H}_{K\times FM} \in \mathbb{C}^{K\times FM}$ for the STF system. In this section, we derive semi-blind ALS and non-blind ZF receivers for the TST and STF systems. The conditions to ensure the LS identifiability of channel and symbol estimates are analyzed.

4.1.1 TST system

Let us consider the TST system modeled by the PT-(2,4) and the fourth-order constrained CP models to develop two TST receivers called, respectively, ALS-PT and ALS-CP. For each receiver, we investigate the necessary conditions to ensure the LS identifiability of the estimates.

TST system modeled with the PARATUCK-(2,4) model

Rewriting the unfolding matrices of the received signal tensor $\mathcal{X}$ given in (3.11) as

$$\mathbf{X}_{K\times PJN} = \mathbf{H}\,\mathbf{G}_{M\times PJR}\,\left(\mathbf{I}_{PJ}\otimes\mathbf{S}^{\mathrm{T}}\right), \tag{4.1}$$

$$\mathbf{X}_{N\times PJK} = \mathbf{S}\,\mathbf{G}_{R\times PJM}\,\left(\mathbf{I}_{PJ}\otimes\mathbf{H}^{\mathrm{T}}\right), \tag{4.2}$$

the problem of channel and symbol estimation according to the ALS approach can be formulated as a set of two independent linear LS problems as follows

$$\begin{cases} \hat{\mathbf{H}} = \underset{\mathbf{H}}{\operatorname{argmin}} \left\| \tilde{\mathbf{X}}_{K\times PJN} - \mathbf{H}\,\mathbf{G}_{M\times PJR}\left(\mathbf{I}_{PJ}\otimes\mathbf{S}^{\mathrm{T}}\right)\right\|_F^2 \\ \hat{\mathbf{S}} = \underset{\mathbf{S}}{\operatorname{argmin}} \left\| \tilde{\mathbf{X}}_{N\times PJK} - \mathbf{S}\,\mathbf{G}_{R\times PJM}\left(\mathbf{I}_{PJ}\otimes\mathbf{H}^{\mathrm{T}}\right)\right\|_F^2 \end{cases}. \tag{4.3}$$

Both minimizations in (4.3) lead to channel $\hat{\mathbf{H}}$ and symbol $\hat{\mathbf{S}}$ estimates given by

$$\hat{\mathbf{H}} = \tilde{\mathbf{X}}_{K\times PJN}\left(\left(\mathbf{I}_{PJ}\otimes\mathbf{S}\right)\mathbf{G}_{PJR\times M}\right)^{\mathrm{T}\dagger}, \tag{4.4}$$

$$\hat{\mathbf{S}} = \tilde{\mathbf{X}}_{N\times PJK}\left(\left(\mathbf{I}_{PJ}\otimes\mathbf{H}\right)\mathbf{G}_{PJM\times R}\right)^{\mathrm{T}\dagger} \tag{4.5}$$

respectively. The unique existence of the right-inverse of $\mathbf{G}_{M\times PJR}\left(\mathbf{I}_{PJ}\otimes\mathbf{S}^{\mathrm{T}}\right)$ and $\mathbf{G}_{R\times PJM}$ $\left(\mathbf{I}_{PJ}\otimes\mathbf{H}^{\mathrm{T}}\right)$ is guaranteed by assuming that $(\mathbf{I}_{PJ}\otimes\mathbf{S})\,\mathbf{G}_{PJR\times M}$ and $(\mathbf{I}_{PJ}\otimes\mathbf{H})\,\mathbf{G}_{PJM\times R}$ are full column-rank, respectively.

We apply the ALS algorithm to alternately estimate the channel and symbol matrices using (4.4) and (4.5) at each iteration. Remark that the estimation of one matrix at the i-th iteration is conditioned to the knowledge of previously estimated value of the other matrix at the $(i-1)$-th iteration.

Analysis of identifiability conditions:

In order to guarantee the uniqueness of the ALS solution, we have to ensure that the left-inverses in (4.4) and (4.5) exist and are unique. Assuming that $\mathbf{S}$ and $\mathbf{H}$ are full column-rank, identifiability of their LS estimates requires that $\mathbf{G}_{PJR\times M}$ and $\mathbf{G}_{PJM\times R}$ be also full column-rank. From this, we enunciate the following theorems.

Theorem 4.1. (Identifiability Condition). [66] *Supposing that* $\mathbf{S}$ *and* $\mathbf{H}$ *are full column-rank, a necessary condition for identifiability of* $\mathbf{H}$ *and* $\mathbf{S}$ *by, respectively,* (4.4) *and* (4.5) *is given by*

$$PJ \geq \max\left(\left\lceil \frac{R}{M}, \frac{M}{R} \right\rceil\right), \tag{4.6}$$

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

Proof: Let us rewrite both equations (4.1) and (4.2) as $\mathbf{X}_{K\times PJN} = \mathbf{H}\,\mathbf{Z}_1^{\mathrm{T}}$ and $\mathbf{X}_{N\times PJK} = \mathbf{S}\,\mathbf{Z}_2^{\mathrm{T}}$, where $\mathbf{Z}_1 \triangleq (\mathbf{I}_{PJ}\otimes\mathbf{S})\,\mathbf{G}_{PJR\times M}$ and $\mathbf{Z}_2 \triangleq (\mathbf{I}_{PJ}\otimes\mathbf{H})\,\mathbf{G}_{PJM\times R}$. Uniqueness of the LS solution for $\mathbf{H}$ and $\mathbf{S}$ requires respectively that $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are full column-rank. Assuming that $\mathbf{S}$ and $\mathbf{H}$ are full column-rank implies that $\mathbf{I}_{PJ}\otimes\mathbf{S}$ and $\mathbf{I}_{PJ}\otimes\mathbf{H}$ are full column-rank as well. Consequently, $\mathrm{rank}(\mathbf{Z}_1) = \mathrm{rank}(\mathbf{G}_{PJR\times M})$ and $\mathrm{rank}(\mathbf{Z}_2) = \mathrm{rank}(\mathbf{G}_{PJM\times R})$, which means that $\mathbf{G}_{PJR\times M}$ and $\mathbf{G}_{PJM\times R}$ must be full column-rank to ensure the identifiability of $\mathbf{H}$ and $\mathbf{S}$, implying $PJR \geq M$ and $PJM \geq R$, or equivalently (4.6). ∎

This condition (4.6) defines a constraint that the design parameters (P, J, M, R) must satisfy. It is interesting to notice that the supplementary diversity introduced by the time-spreading mode (j) of the code tensor allows us to get a more relaxed condition on the number P of time blocks than the one obtained in [59].

If we assume that $\mathbf{S}$ and $\mathbf{H}$ are full column-rank, then the identifiability conditions can be ensured just setting appropriately the allocation matrices and the tensor code. The choice of the code tensor will be explained later. Considering the Vandermonde structure to construct each j-th matrix-slice $\mathbf{W}_{..j}$ of the code tensor $\mathcal{W}$, its elements can be given by

$$w_{m,r,j} = e^{i2\pi\, j\frac{mr}{MR}}. \tag{4.7}$$

An important reason behind this choice for the code tensor is that this structure guarantees the existence of a minimum value of the spreading length J ensuring the identifiability in the LS sense of the channel $\mathbf{H}$ and symbol $\mathbf{S}$ matrices, as shown in the next theorem.

Theorem 4.2. (Identifiability Condition). [66] *Suppose the perfect knowledge of the code tensor* $\mathcal{W}$ *and the allocation matrices* $\mathbf{C}^{(\mathbf{H})}$ *and* $\mathbf{C}^{(\mathbf{S})}$. *Assuming that* $\mathbf{S}$ *and* $\mathbf{H}$ *are full column-rank, with the Vandermonde structure* (4.7) *for the code tensor, and* ($\mathbf{C}^{(\mathbf{S})}_{p\cdot} = \mathbf{1}^{\mathrm{T}}_R$, $\mathbf{C}^{(\mathbf{H})}_{p\cdot} = \mathbf{1}^{\mathrm{T}}_M$) *for a given* $p \in \{1, ..., P\}$, *Table 4.1 gives the minimum value of the spreading length ensuring the LS identifiability of* $\mathbf{H}$ *and* $\mathbf{S}$ *by* (4.4) *and* (4.5), *for* M *and* $R \in \{1, ..., 8\}$.

Table 4.1: Minimum value of J for LS identifiability of $\mathbf{S}$ and $\mathbf{H}$

M \ R	1	2	3	4	5	6	7	8
1	1	2	3	4	5	6	7	8
2	2	1	2	3	3	4	5	5
3	3	2	1	2	2	3	4	4
4	4	3	2	1	2	2	3	3
5	5	3	2	2	1	2	2	2
6	6	4	3	2	2	1	2	2
7	7	5	4	3	2	2	1	2
8	8	5	4	3	2	2	2	1

Proof: For $\mathbf{C}^{(\mathbf{S})}_{p\cdot} = \mathbf{1}^{\mathrm{T}}_R$ and $\mathbf{C}^{(\mathbf{H})}_{p\cdot} = \mathbf{1}^{\mathrm{T}}_M$, (3.4) gives $\mathbf{G}_{\cdot\cdot p,j} = \mathbf{W}_{\cdot\cdot j}$ and then from (3.12) we can extract the following two sub-matrices associated with the block p from $\mathbf{G}_{PJM\times R}$ and $\mathbf{G}_{PJR\times M}$ respectively,

$$\mathbf{W}_{JM\times R} \triangleq \begin{bmatrix} \mathbf{W}_{\cdot\cdot 1} \\ \vdots \\ \mathbf{W}_{\cdot\cdot J} \end{bmatrix} \in \mathbb{C}^{JM\times R}, \quad \mathbf{W}_{JR\times M} \triangleq \begin{bmatrix} \mathbf{W}^{\mathrm{T}}_{\cdot\cdot 1} \\ \vdots \\ \mathbf{W}^{\mathrm{T}}_{\cdot\cdot J} \end{bmatrix} \in \mathbb{C}^{JR\times M}. \tag{4.8}$$

Remark that both matrices $\mathbf{W}_{JM\times R}$ and $\mathbf{W}_{JR\times M}$ represent two matrix unfolding of $\mathcal{W}$. First, we have to notice the symmetric role played by M and R, with the block $\mathbf{W}_{\cdot\cdot j}$ given by

$$\mathbf{W}_{\cdot\cdot j} = \begin{bmatrix} u^{j} & u^{2j} & \cdots & u^{Rj} \\ u^{2j} & u^{4j} & \cdots & u^{2Rj} \\ \vdots & \vdots & \ddots & \vdots \\ u^{Mj} & u^{2Mj} & \cdots & u^{RMj} \end{bmatrix}, \tag{4.9}$$

where $u = e^{\frac{i2\pi}{MR}}$.

When $M = R$, $\mathbf{W}_{\cdot\cdot 1}$ is non-singular, which implies that $\mathbf{G}_{PJM\times R}$ and $\mathbf{G}_{PJR\times M}$ are full column-rank, and therefore the LS estimate of $\mathbf{S}$ and $\mathbf{H}$ is unique with $J_{\min} = 1$, which corresponds to Theorem 3 in [59]. In the case where $M > R$, the block $\mathbf{W}_{\cdot\cdot 1}$ and therefore $\mathbf{G}_{PJM\times R}$ are full column-rank, whereas $\mathbf{W}^{\mathrm{T}}_{\cdot\cdot 1}$ is full row-rank equal to R. The number $J_{\min}$ of blocks $\mathbf{W}^{\mathrm{T}}_{\cdot\cdot j}$ to be considered in $\mathbf{W}_{JR\times M}$ to guaranty its full column rank property is given in Table 4.1 for $M, R \in \{1, ..., 8\}$. Note that setting $J \geq J_{\min}$ it is added $M - R$ independent columns to $\mathbf{W}^{\mathrm{T}}_{\cdot\cdot 1}$, which leads to $\mathbf{W}_{JR\times M}$ full column-rank. When $R > M$, the same reasoning can be applied to determine the minimum number $J_{\min}$ of blocks $\mathbf{W}_{\cdot\cdot j}$ to be considered in $\mathbf{W}_{JM\times R}$ for guarantying its full column rank property, which explains the symmetric form of Table 4.1. ∎

Remark 4.1.

- *The constraints* $\mathbf{C}_{p\cdot}^{(\mathbf{S})} = \mathbf{1}_R^{\mathrm{T}}$ *and* $\mathbf{C}_{p\cdot}^{(\mathbf{H})} = \mathbf{1}_M^{\mathrm{T}}$ *mean that all data streams (*R*) are sent using all transmit antennas (*M*), during the* p*-th time block.*
- *Introducing the time spreading mode in the Vandermonde code tensor allows us to derive a minimum value of the spreading length* J *that ensures the identifiability of* $\mathbf{S}$ *and* $\mathbf{H}$ *in the case* $M \neq R$*, which is not possible in [59] with* $J = 1$*.*
- *Note that Theorem 4.2 gives a sufficient condition over* $\mathcal{W}$*,* $\mathbf{C}^{(\mathbf{H})}$ *and* $\mathbf{C}^{(\mathbf{S})}$*, while Theorem 4.1 establishes a necessary condition over* (P, J, M, R) *for LS identifiability of* $\mathbf{S}$ *and* $\mathbf{H}$*. Both theorems assume that* $\mathbf{S}$ *and* $\mathbf{H}$ *are full column-rank.*
- *According to the assumption for the channel coefficients discussed in Chapter 3,* $\mathbf{H}$ *is almost surely full column-rank if* $K \geq M$*.*

Observe that when Theorem 4.2 is satisfied, the inequality (4.6) of Theorem 4.1 is implicitly achieved.

TST system modeled by a constrained CP-4 model

Let us recall the received signal tensor $\mathcal{X} \in \mathbb{C}^{K\times N\times P\times J}$ modeled by the PT-(2,4) model given in (3.5)

$$x_{k,n,p,j} = \sum_{m=1}^{M}\sum_{r=1}^{R} w_{m,r,j}\, h_{k,m}\, s_{n,r}\, c_{p,m}^{(\mathbf{H})}\, c_{p,r}^{(\mathbf{S})}, \tag{4.10}$$

we can rewrite the received signal tensor by the following constrained CP-4 model

$$x_{k,n,p,j} = \sum_{r=1}^{MR} \overline{a}_{k,r}\, \overline{b}_{n,r}\, \overline{c}_{p,r}\, \overline{d}_{j,r}, \tag{4.11}$$

whose matrix factors $\overline{\mathbf{A}} \in \mathbb{C}^{K\times MR}$, $\overline{\mathbf{B}} \in \mathbb{C}^{N\times MR}$, $\overline{\mathbf{C}} \in \mathbb{C}^{P\times MR}$ and $\overline{\mathbf{D}} \in \mathbb{C}^{J\times MR}$ are given by

$$\begin{aligned}
\overline{\mathbf{A}} &= \mathbf{H} \otimes \mathbf{1}_R^{\mathrm{T}} = \mathbf{H}\boldsymbol{\Omega}^{(1)},\\
\overline{\mathbf{B}} &= \mathbf{1}_M^{\mathrm{T}} \otimes \mathbf{S} = \mathbf{S}\boldsymbol{\Omega}^{(2)},\\
\overline{\mathbf{C}} &= \left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}},\\
\overline{\mathbf{D}} &= \mathbf{W}_{J\times MR},
\end{aligned} \tag{4.12}$$

where

$$\begin{aligned}
\boldsymbol{\Omega}^{(1)} &\triangleq \mathbf{I}_M \otimes \mathbf{1}_R^{\mathrm{T}} \in \mathbb{R}^{M\times MR},\\
\boldsymbol{\Omega}^{(2)} &\triangleq \mathbf{1}_M^{\mathrm{T}} \otimes \mathbf{I}_R \in \mathbb{R}^{R\times MR}.
\end{aligned} \tag{4.13}$$

Thus, we can deduce the following matrices from the unfolding matrix for the CP model

given in (2.21) as

$$\mathbf{X}_{K\times PJN} = \mathbf{H}\boldsymbol{\Omega}^{(1)} \left(\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}} \right)^{\mathrm{T}} \diamond \mathbf{W}_{J\times MR} \diamond \mathbf{S}\boldsymbol{\Omega}^{(2)} \right)^{\mathrm{T}}, \tag{4.14}$$

$$\mathbf{X}_{N\times PJK} = \mathbf{S}\boldsymbol{\Omega}^{(2)} \left(\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}} \right)^{\mathrm{T}} \diamond \mathbf{W}_{J\times MR} \diamond \mathbf{H}\boldsymbol{\Omega}^{(1)} \right)^{\mathrm{T}}, \tag{4.15}$$

$$\mathbf{X}_{P\times JKN} = \left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}} \right)^{\mathrm{T}} \left(\mathbf{W}_{J\times MR} \diamond \mathbf{H}\boldsymbol{\Omega}^{(1)} \diamond \mathbf{S}\boldsymbol{\Omega}^{(2)} \right)^{\mathrm{T}}, \tag{4.16}$$

$$\mathbf{X}_{J\times KNP} = \mathbf{W}_{J\times MR} \left(\mathbf{H}\boldsymbol{\Omega}^{(1)} \diamond \mathbf{S}\boldsymbol{\Omega}^{(2)} \diamond \left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}} \right)^{\mathrm{T}} \right)^{\mathrm{T}}. \tag{4.17}$$

Applying property (A.6) to (4.16) and (4.17), and using (4.13), we obtain

$$\begin{aligned} \mathbf{X}_{P\times JKN} &= \left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}} \right)^{\mathrm{T}} \left(\mathbf{W}_{J\times MR} \diamond (\mathbf{H} \otimes \mathbf{S}) \left(\boldsymbol{\Omega}^{(1)} \diamond \boldsymbol{\Omega}^{(2)} \right) \right)^{\mathrm{T}} \\ &= \left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}} \right)^{\mathrm{T}} \left(\mathbf{W}_{J\times MR} \diamond (\mathbf{H} \otimes \mathbf{S}) \right)^{\mathrm{T}}, \end{aligned} \tag{4.18}$$

$$\begin{aligned} \mathbf{X}_{J\times KNP} &= \mathbf{W}_{J\times MR} \left((\mathbf{H} \otimes \mathbf{S}) \left(\boldsymbol{\Omega}^{(1)} \diamond \boldsymbol{\Omega}^{(2)} \right) \diamond \left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}} \right)^{\mathrm{T}} \right)^{\mathrm{T}} \\ &= \mathbf{W}_{J\times MR} \left((\mathbf{H} \otimes \mathbf{S}) \diamond \left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}} \right)^{\mathrm{T}} \right)^{\mathrm{T}}. \end{aligned} \tag{4.19}$$

where $\boldsymbol{\Omega}^{(1)} \diamond \boldsymbol{\Omega}^{(2)} = \mathbf{I}_{MR}$.

The minimization of LS cost functions for channel and symbol estimation can be also reformulated from the matrix unfoldings $\mathbf{X}_{K\times PJN}$ and $\mathbf{X}_{N\times PJK}$, respectively, in (4.14) and (4.15) by

$$\begin{cases} \hat{\mathbf{H}} = \underset{\mathbf{H}}{\operatorname{argmin}} \left\| \tilde{\mathbf{X}}_{K\times PJN} - \mathbf{H}\boldsymbol{\Omega}^{(1)} \left(\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}} \right)^{\mathrm{T}} \diamond \mathbf{W}_{J\times MR} \diamond \mathbf{S}\boldsymbol{\Omega}^{(2)} \right)^{\mathrm{T}} \right\|_F^2 \\ \hat{\mathbf{S}} = \underset{\mathbf{S}}{\operatorname{argmin}} \left\| \tilde{\mathbf{X}}_{N\times PJK} - \mathbf{S}\boldsymbol{\Omega}^{(2)} \left(\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}} \right)^{\mathrm{T}} \diamond \mathbf{W}_{J\times MR} \diamond \mathbf{H}\boldsymbol{\Omega}^{(1)} \right)^{\mathrm{T}} \right\|_F^2 \end{cases}. \tag{4.20}$$

Resulting in both estimations

$$\hat{\mathbf{H}} = \tilde{\mathbf{X}}_{K\times PJN} \left(\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}} \right)^{\mathrm{T}} \diamond \mathbf{W}_{J\times MR} \diamond \mathbf{S}\boldsymbol{\Omega}^{(2)} \right)^{\mathrm{T}^{\dagger}} \boldsymbol{\Omega}^{(1)^{\dagger}}, \tag{4.21}$$

$$\hat{\mathbf{S}} = \tilde{\mathbf{X}}_{N\times PJK} \left(\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}} \right)^{\mathrm{T}} \diamond \mathbf{W}_{J\times MR} \diamond \mathbf{H}\boldsymbol{\Omega}^{(1)} \right)^{\mathrm{T}^{\dagger}} \boldsymbol{\Omega}^{(2)^{\dagger}}. \tag{4.22}$$

From (4.13) and the right-inverse definition, we obtain that

$$\begin{aligned} \boldsymbol{\Omega}^{(1)^{\dagger}} &= {}^{1}/_{R} \left(\mathbf{I}_M \otimes \mathbf{1}_R \right) = {}^{1}/_{R}\, \boldsymbol{\Omega}^{(1)^{\mathrm{T}}} \in \mathbb{C}^{MR\times M}, \\ \boldsymbol{\Omega}^{(2)^{\dagger}} &= {}^{1}/_{M} \left(\mathbf{1}_M \otimes \mathbf{I}_R \right) = {}^{1}/_{M}\, \boldsymbol{\Omega}^{(2)^{\mathrm{T}}} \in \mathbb{C}^{MR\times R}. \end{aligned} \tag{4.23}$$

Analogously to (4.4) and (4.5), assuming that the right-inverses exist and are unique, we can iteratively estimate the channel and symbol matrices by (4.21) and (4.22), respectively.

Analysis of identifiability conditions:

Analogously to the analysis for the TST modeled by the PT-(2,4) model, we can study the conditions to guarantee the identifiability of channel and symbol matrices from (4.21) and (4.22).

Let us define the following matrices

$$\mathbf{P}_1 \triangleq \left(\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}} \diamond \mathbf{W}_{J\times MR} \diamond \mathbf{S}\boldsymbol{\Omega}^{(2)}\right)^{\mathrm{T}} \in \mathbb{C}^{MR\times PJN}, \tag{4.24}$$

$$\mathbf{P}_2 \triangleq \left(\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}} \diamond \mathbf{W}_{J\times MR} \diamond \mathbf{H}\boldsymbol{\Omega}^{(1)}\right)^{\mathrm{T}} \in \mathbb{C}^{MR\times PJK}, \tag{4.25}$$

if we assume that $\mathbf{P}_1$ and $\mathbf{P}_2$ are full row-rank, implying $PJN \geq MR$ and $PJK \geq MR$, then we can rewrite (4.21) and (4.22) using (4.23) as

$$\hat{\mathbf{H}} = {}^{1}\!/_{R}\ \tilde{\mathbf{X}}_{K\times PJN}\, \mathbf{P}_1^{\dagger}\, \boldsymbol{\Omega}^{(1)^{\mathrm{T}}}, \tag{4.26}$$

$$\hat{\mathbf{S}} = {}^{1}\!/_{M}\, \tilde{\mathbf{X}}_{N\times PJK}\, \mathbf{P}_2^{\dagger}\, \boldsymbol{\Omega}^{(2)^{\mathrm{T}}}, \tag{4.27}$$

where $\mathbf{P}_1^{\dagger}$ and $\mathbf{P}_2^{\dagger}$ denote, respectively, the right-inverse of $\mathbf{P}_1$ and $\mathbf{P}_2$, implying $\mathbf{P}_1\mathbf{P}_1^{\dagger} = \mathbf{I}_{MR}$ and $\mathbf{P}_2\mathbf{P}_2^{\dagger} = \mathbf{I}_{MR}$.

Therefore, a necessary condition for LS identifiability of both matrices $\mathbf{H}$ and $\mathbf{S}$ by (4.26) and (4.27) depends on the unique existence of the pseudo-inverse of $\mathbf{P}_1$ and $\mathbf{P}_2$ respectively, which imposes a condition over the design parameters (N, K, P, J, M, R) given as follows.

Theorem 4.3. (Identifiability Condition). *A necessary condition for identifiability of* $\mathbf{H}$ *and* $\mathbf{S}$ *by, respectively,* (4.26) *and* (4.27) *is given by*

$$PJ \geq \max\left(\left\lceil \frac{MR}{N}, \frac{MR}{K} \right\rceil\right). \tag{4.28}$$

Remark that the condition given by Theorem 4.3 does not require the assumption of $\mathbf{S}$ and $\mathbf{H}$ being full column-rank, which leads to $N \geq R$ and $K \geq M$, as it is required in Theorem 4.1. From Lemma 2.2, if $\mathbf{S}$, $\mathbf{H}$, $\mathbf{C}^{(\mathbf{S})}$, $\mathbf{C}^{(\mathbf{H})}$ and $\mathbf{W}_{J\times MR}$ satisfy the next double condition:

$$k\left(\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}}\right) + k(\mathbf{W}_{J\times MR}) + k\big(\mathbf{S}\boldsymbol{\Omega}^{(2)}\big) \geq MR + 2, \tag{4.29}$$

$$k\left(\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}}\right) + k(\mathbf{W}_{J\times MR}) + k\big(\mathbf{H}\boldsymbol{\Omega}^{(1)}\big) \geq MR + 2, \tag{4.30}$$

then the right-inverse of $\mathbf{P}_1$ and $\mathbf{P}_2$ exists and we can write $\mathbf{H}$ and $\mathbf{S}$ as (4.26) and (4.27) respectively. From two conditions (4.29) and (4.30), we can deduce the next theorem.

Theorem 4.4. (Identifiability Condition). *Suppose the perfect knowledge of the code tensor* $\mathcal{W}$ *and the allocation matrices* $\mathbf{C}^{(\mathbf{H})}$ *and* $\mathbf{C}^{(\mathbf{S})}$. *If the code tensor* $\mathcal{W}$ *is chosen such that* $\mathbf{W}_{J\times MR}$ *is full column-rank implying* $J \geq MR$, *and* $\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}$ *has no zero-rows, then the LS identifiability of* $\mathbf{H}$ *and* $\mathbf{S}$ *by, respectively,* (4.26) *and* (4.27) *is ensured.*

Proof: The identifiability of channel and symbol matrices in the LS sense requires that $\mathbf{P}_1$ and $\mathbf{P}_2$ given by (4.24) and (4.25) are full row-rank which can be achieved by the conditions (4.29) and (4.30) respectively. By system feature, both matrices $\mathbf{H}$ and $\mathbf{S}$ have no zero-columns and using $\mathbf{\Omega}^{(1)} = \mathbf{I}_M \otimes \mathbf{1}_R^{\mathrm{T}}$ and $\mathbf{\Omega}^{(2)} = \mathbf{1}_M^{\mathrm{T}} \otimes \mathbf{I}_R$, we obtain $k\big(\mathbf{H}\mathbf{\Omega}^{(1)}\big) = 1$ and $k\big(\mathbf{S}\mathbf{\Omega}^{(2)}\big) = 1$ because the product by $\mathbf{\Omega}^{(1)}$ and $\mathbf{\Omega}^{(2)}$ results in column repetitions of $\mathbf{H}$ and $\mathbf{S}$ respectively, i.e.

$$\begin{aligned} \mathbf{H}\mathbf{\Omega}^{(1)} &= \mathbf{H} \otimes \mathbf{1}_R^{\mathrm{T}} \in \mathbb{C}^{K \times MR}, \\ \mathbf{S}\mathbf{\Omega}^{(2)} &= \mathbf{1}_M^{\mathrm{T}} \otimes \mathbf{S} \in \mathbb{C}^{N \times MR}. \end{aligned} \tag{4.31}$$

If $\mathbf{W}_{J \times MR}$ is full column-rank and $\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}$ has no zero-rows, then $k(\mathbf{W}_{J \times MR}) = MR$ and $k\Big(\Big(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\Big)^{\mathrm{T}}\Big) \geq 1$, respectively. Therefore, we can simultaneously satisfy both conditions (4.29) and (4.30), and finally ensure the identifiability of $\mathbf{H}$ and $\mathbf{S}$ by (4.26) and (4.27). ∎

Remark 4.2.

- *Both Theorems 4.2 and 4.4 guarantee the LS identifiability of channel and symbol matrices by choosing appropriately* $\mathcal{W}$, $\mathbf{C}^{(\mathbf{S})}$ *and* $\mathbf{C}^{(\mathbf{H})}$. *However, Theorem 4.2 imposes that* $\mathbf{S}$ *and* $\mathbf{H}$ *are full column-rank unlike Theorem 4.4.*
- *Analogously to Theorems 4.1 and 4.2, while Theorem 4.3 provides a necessary condition over* (N, K, P, J, M, R), *Theorem 4.4 affords a sufficient condition over* $\mathcal{W}$, $\mathbf{C}^{(\mathbf{S})}$ *and* $\mathbf{C}^{(\mathbf{H})}$ *for the LS identifiability of* $\mathbf{H}$ *and* $\mathbf{S}$ *by, respectively,* (4.26) *and* (4.27).
- *Remember that* $k\Big(\Big(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\Big)^{\mathrm{T}}\Big) \geq 1$ *is equivalent to satisfy* (3.69). *Thus, as discussed previously, it has a practical interpretation to construct the allocation structures.*

Note that the TST system modeled by a constraint CP-4 permits to derive another identifiability condition which does not require $K \geq M$ and $N \geq R$, since we analyze the existence of pseudo-inverse matrices by means of a joint condition under $\mathbf{C}^{(\mathbf{S})}$, $\mathbf{C}^{(\mathbf{H})}$, $\mathbf{W}_{J \times MR}$, $\mathbf{S}$ and $\mathbf{H}$. Analogously to Theorems 4.1 and 4.2, satisfying Theorem 4.4 leads to the inequality (4.28) of Theorem 4.3.

Defining the elements of code tensor by

$$w_{m,r,j} = e^{i2\pi\, j \frac{r+(m-1)R}{MR}}. \tag{4.32}$$

for $J \geq MR$, the matrix $\mathbf{W}_{J \times MR}$ is full column-rank thanks to the Vandermonde structure, implying $k(\mathbf{W}_{J \times MR}) = MR$.

Generalization to multiuser case

Consider the received signal tensor for the TST system with multiple users given by (3.80). We obtain the following matrices by transposing two unfolding matrices given in (3.82) and (3.81), respectively

$$\mathbf{X}_{K \times PJN} = \mathbf{H}_{K \times QM}\, \mathbf{G}_{QM \times QPJR}\, \left(\mathbf{\Omega}^{\mathrm{T}} \,|\otimes|\, \mathbf{S}_{QR \times N}\right), \tag{4.33}$$

$$\mathbf{X}_{N \times PJK} = \mathbf{S}_{N \times QR}\, \mathbf{G}_{QR \times QPJM}\, \left(\mathbf{\Omega}^{\mathrm{T}} \,|\otimes|\, \mathbf{H}_{QM \times K}\right), \tag{4.34}$$

with $\mathbf{\Omega} \triangleq \mathbf{1}_Q^{\mathrm{T}} \otimes \mathbf{I}_{PJ} \in \mathbb{C}^{PJ \times QPJ}$.

According to the ALS approach, the problem of channel and symbol estimation can be formulated as a set of the following linear problems

$$\begin{cases} \hat{\mathbf{H}}_{K\times QM} = \underset{\mathbf{H}_{K\times M}}{\operatorname{argmin}} \left\| \tilde{\mathbf{X}}_{K\times PJN} - \mathbf{H}_{K\times QM}\, \mathbf{G}_{QM\times QPJR}\, \left(\mathbf{\Omega}^{\mathrm{T}} \,|\otimes|\, \mathbf{S}_{QR\times N}\right) \right\|_F^2 \\ \hat{\mathbf{S}}_{N\times QR} = \underset{\mathbf{S}_{N\times QR}}{\operatorname{argmin}} \left\| \tilde{\mathbf{X}}_{N\times PJK} - \mathbf{S}_{N\times QR}\, \mathbf{G}_{QR\times QPJM}\, \left(\mathbf{\Omega}^{\mathrm{T}} \,|\otimes|\, \mathbf{H}_{QM\times K}\right) \right\|_F^2 \end{cases}. \tag{4.35}$$

Both minimizations (4.35) result in

$$\hat{\mathbf{H}}_{K\times QM} = \tilde{\mathbf{X}}_{K\times PJN}((\mathbf{\Omega} \,|\otimes|\, \mathbf{S}_{N\times QR})\ \mathbf{G}_{QPJR\times QM})^{\mathrm{T}^\dagger}, \tag{4.36}$$

$$\hat{\mathbf{S}}_{N\times QR} = \tilde{\mathbf{X}}_{N\times PJK}((\mathbf{\Omega} \,|\otimes|\, \mathbf{H}_{K\times QM})\ \mathbf{G}_{QPJM\times QR})^{\mathrm{T}^\dagger}. \tag{4.37}$$

Assuming that $(\mathbf{\Omega} \,|\otimes|\, \mathbf{S}_{N\times QR})\ \mathbf{G}_{QPJR\times QM}$ and $(\mathbf{\Omega} \,|\otimes|\, \mathbf{H}_{K\times QM})\ \mathbf{G}_{QPJM\times QR}$ are full column-rank, we ensure the right-inverse uniqueness of both transposed matrices. Consequently, we can estimate the channel and symbol matrices associated with each user by (4.36) and (4.37), respectively.

Analysis of identifiability conditions:

Supposing that $\mathbf{S}_{N\times QR}$ and $\mathbf{H}_{K\times QM}$ are full column-rank, identifiability of their LS estimates requires that $\mathbf{G}^{(q)}_{PJR\times M}$ and $\mathbf{G}^{(q)}_{PJM\times R}$ be also full column-rank for all $q \in \{1, \cdots, Q\}$. From this, we can enunciate the following theorems.

Theorem 4.5. (Identifiability Condition). *Supposing that $\mathbf{S}_{N\times QR}$ and $\mathbf{H}_{K\times QM}$ are full column-rank, a necessary condition for identifiability of $\mathbf{H}_{K\times QM}$ and $\mathbf{S}_{N\times QR}$ by, respectively,* (4.36) *and* (4.37) *is given by*

$$PJ \geq \max\left(\left\lceil \frac{R}{M}, \frac{M}{R} \right\rceil \right). \tag{4.38}$$

Proof: Rewriting both equations (4.33) and (4.34) as $\mathbf{X}_{K\times PJN} = \mathbf{H}_{K\times QM}\, \mathbf{Z}_1^{\mathrm{T}}$ and $\mathbf{X}_{N\times PJK} = \mathbf{S}_{N\times QR}\, \mathbf{Z}_2^{\mathrm{T}}$, where $\mathbf{Z}_1 \triangleq (\mathbf{\Omega} \,|\otimes|\, \mathbf{S}_{N\times QR})\ \mathbf{G}_{QPJR\times QM}$ and $\mathbf{Z}_2 \triangleq (\mathbf{\Omega} \,|\otimes|\, \mathbf{H}_{K\times QM})\ \mathbf{G}_{QPJM\times QR}$, thus the uniqueness of the LS solution for $\mathbf{H}_{K\times QM}$ and $\mathbf{S}_{N\times QR}$ requires that $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are full column-rank, respectively.

Both matrices $\mathbf{\Omega} \,|\otimes|\, \mathbf{S}_{N\times QR}$ and $\mathbf{\Omega} \,|\otimes|\, \mathbf{H}_{K\times QM}$ are full column-rank when $\mathbf{S}_{N\times QR}$ and $\mathbf{H}_{K\times QM}$ are full column-rank, respectively. Consequently, it gets $\operatorname{rank}(\mathbf{Z}_1) = \operatorname{rank}(\mathbf{G}_{PJR\times M})$ and $\operatorname{rank}(\mathbf{Z}_2) = \operatorname{rank}(\mathbf{G}_{PJM\times R})$, which means that $\mathbf{G}_{QPJR\times QM}$ and $\mathbf{G}_{QPJM\times QR}$ must be full column-rank to ensure the identifiability of $\mathbf{H}_{K\times QM}$ and $\mathbf{S}_{N\times QR}$, implying $PJR \geq M$ and $PJM \geq R$, or equivalently (4.38). ∎

Remark 4.3.

- *The condition* (4.38) *is equal to* (4.6) *obtained considering the transmission by only one user, i.e.* $Q = 1$.

- *As the channel coefficients are drawn from a continuous distribution,* $\mathbf{H}_{K\times QM}$ *is almost surely full column-rank if* $K \geq QM$*.*

Both matrices $\mathbf{G}_{QPJR\times QM}$ and $\mathbf{G}_{QPJM\times QR}$, defined in in (3.84), are full column-rank when $\mathbf{G}^{(q)}_{PJR\times M}$ and $\mathbf{G}^{(q)}_{PJM\times R}$ are full column-rank for all $q \in \{1, \cdots, Q\}$. Supposing the perfect knowledge of the code tensor and the allocation matrices of all Q users, the channel and symbol estimation by, respectively, (4.36) and (4.37) requires that $\mathbf{G}_{QPJR\times QM}$ and $\mathbf{G}_{QPJM\times QR}$ are full column-rank. Hence, the identifiability condition on the allocation matrices and tensor code for each user is directly derived from Theorem 4.1 deduced for $Q = 1$.

4.1.2 STF system

We regard the STF system given by (3.27) for the STF receiver. Transposing the unfolding matrices of the received signal tensor $\mathcal{X} \in \mathbb{C}^{F\times K\times N\times P}$ given in (3.32) gives

$$\mathbf{X}_{K\times PFN} = \mathbf{H}_{K\times FM}\,\mathbf{T}_{FM\times PFR}\left(\mathbf{I}_{PF}\otimes\mathbf{S}^{\mathrm{T}}\right), \tag{4.39}$$

$$\mathbf{X}_{N\times PFK} = \mathbf{S}\,\mathbf{T}_{R\times PFM}\left(\mathbf{I}_{P}\otimes \mathrm{bdiag}\big(\mathbf{H}^{\mathrm{T}}_{1..},\ldots,\mathbf{H}^{\mathrm{T}}_{F..}\big)\right). \tag{4.40}$$

The problem of channel and symbol estimation can be formulated in accordance with the ALS approach as two independent LS problems given by

$$\begin{cases} \hat{\mathbf{H}}_{K\times FM} = \underset{\mathbf{H}_{K\times FM}}{\operatorname{argmin}} \left\| \tilde{\mathbf{X}}_{K\times PFN} - \mathbf{H}_{K\times FM}\,\mathbf{T}_{FM\times PFR}\left(\mathbf{I}_{PF}\otimes\mathbf{S}^{\mathrm{T}}\right)\right\|_F^2 \\ \hat{\mathbf{S}} = \underset{\mathbf{S}}{\operatorname{argmin}} \left\| \tilde{\mathbf{X}}_{N\times PFK} - \mathbf{S}\,\mathbf{T}_{R\times PFM}\left(\mathbf{I}_{P}\otimes \mathrm{bdiag}\big(\mathbf{H}^{\mathrm{T}}_{1..},\ldots,\mathbf{H}^{\mathrm{T}}_{F..}\big)\right)\right\|_F^2 \end{cases}. \tag{4.41}$$

The minimizations in (4.41) result in

$$\hat{\mathbf{H}}_{K\times FM} = \tilde{\mathbf{X}}_{K\times PFN}\left(\left(\mathbf{I}_{PF}\otimes\mathbf{S}\right)\mathbf{T}_{PFR\times FM}\right)^{\mathrm{T}^{\dagger}}, \tag{4.42}$$

$$\hat{\mathbf{S}} = \tilde{\mathbf{X}}_{N\times PFK}\left(\left(\mathbf{I}_{P}\otimes \mathrm{bdiag}(\mathbf{H}_{1..},\ldots,\mathbf{H}_{F..})\right)\mathbf{T}_{PFM\times R}\right)^{\mathrm{T}^{\dagger}}, \tag{4.43}$$

where $\left(\mathbf{I}_{PF}\otimes\mathbf{S}\right)\mathbf{T}_{PFR\times FM}$ and $\left(\mathbf{I}_{P}\otimes \mathrm{bdiag}(\mathbf{H}_{1..},\ldots,\mathbf{H}_{F..})\right)\mathbf{T}_{PFM\times R}$ are assumed to be full column-rank in order to ensure the right-inverse uniqueness of both transposed matrices.

Analogously to the ALS algorithm for the TST system, we can estimate alternately $\mathbf{H}_{K\times FM} = \left[\mathbf{H}_{1..}\ \cdots\ \mathbf{H}_{F..}\right]$ and $\mathbf{S}$ from (4.42) and (4.43) at each iteration respectively.

Analysis of identifiability conditions:

The identifiability of channel and symbol by (4.42) and (4.43) respectively depends on the unique existence of the pseudo-inverses. Observe that a block diagonal matrix is full rank if and only if each matrix on the diagonal is full rank as well. Assuming that $\mathbf{H}_{f..}$ for all $f \in \{1, ..., F\}$ and $\mathbf{S}$ are full column-rank, the identifiability of LS estimates of symbol and channel requires that $\mathbf{T}_{PFM\times R}$ and $\mathbf{T}_{PFR\times FM}$ be full column-rank respectively. The design parameters have to respect the following theorem.

Theorem 4.6. (Identifiability Condition). *Assuming that* $\mathbf{S}$ *and* $\mathbf{H}_{f\cdot\cdot}$ *for all* $f \in \{1,...,F\}$ *are full column-rank, a necessary condition for identifiability of* $\mathcal{H}$ *and* $\mathbf{S}$ *by, respectively,* (4.42) *and* (4.43) *is given by*

$$P \geq \max\left(\left\lceil \frac{R}{FM}, \frac{M}{R} \right\rceil\right). \tag{4.44}$$

Proof: Let us rewrite both equations (4.39) and (4.40) as $\mathbf{X}_{K\times PFN} = \mathbf{H}_{K\times FM}\mathbf{Z}_1^{\mathrm{T}}$ and $\mathbf{X}_{N\times PFK} = \mathbf{S}\mathbf{Z}_2^{\mathrm{T}}$, where $\mathbf{Z}_1 \triangleq (\mathbf{I}_{PF} \otimes \mathbf{S})\,\mathbf{T}_{PFR\times FM}$ and $\mathbf{Z}_2 \triangleq (\mathbf{I}_P \otimes \mathrm{bdiag}(\mathbf{H}_{1\cdot\cdot},\ldots,\mathbf{H}_{F\cdot\cdot}))\,\mathbf{T}_{PFM\times R}$. Uniqueness of the LS solution for $\mathcal{H}$ and $\mathbf{S}$ requires that $\mathbf{Z}_1$ and $\mathbf{Z}_2$ be full column-rank respectively. The assumption of $\mathbf{S}$ and $\mathbf{H}_{f\cdot\cdot}$ for all $f \in \{1,...,F\}$ being full column-rank implies that $\mathbf{I}_{PF} \otimes \mathbf{S}$ and $\mathbf{I}_P \otimes \mathrm{bdiag}(\mathbf{H}_{1\cdot\cdot},\ldots,\mathbf{H}_{F\cdot\cdot})$ are full column-rank as well. Consequently, $\mathrm{rank}(\mathbf{Z}_1) = \mathrm{rank}(\mathbf{T}_{PFR\times FM})$ and $\mathrm{rank}(\mathbf{Z}_2) = \mathrm{rank}(\mathbf{T}_{PFM\times R})$, which means that $\mathbf{T}_{PFR\times FM}$ and $\mathbf{T}_{PFM\times R}$ must be full column-rank to ensure the identifiability of $\mathcal{H}$ and $\mathbf{S}$, implying $PFR \geq FM$ and $PFM \geq R$, or equivalently (4.44). ∎

Remark 4.4.

- *As expected, eliminating the time* ($J = 1$) *and frequency* ($F = 1$) *redundancies of the TST and STF systems respectively, we have the same condition parameters on* (P, M, R) *for the system proposed in [59].*
- *Observe that when the channel coefficients are drawn from a continuous distribution,* $\mathbf{H}_{f\cdot\cdot}$ *for all* $f \in \{1,...,F\}$ *is almost surely full column-rank if* $K \geq M$.

By convenience, let us recall $\mathbf{T}_{f\cdot\cdot p}$, $\mathbf{T}_{PFM\times R}$ and $\mathbf{T}_{PFR\times FM}$ defined in (3.30) and (3.33),

$$\mathbf{T}_{f\cdot\cdot p} = \mathrm{diag}\left(\mathbf{c}_{f,p\cdot}^{(\mathcal{H})}\right)\,\mathbf{W}\,\mathrm{diag}\left(\mathbf{c}_{f,p\cdot}^{(\mathbf{S})}\right) \in \mathbb{C}^{M\times R} \tag{4.45}$$

and

$$\mathbf{T}_{PFM\times R} = \begin{bmatrix} \mathbf{T}_{1\cdot\cdot 1} \\ \vdots \\ \mathbf{T}_{F\cdot\cdot 1} \\ \vdots \\ \mathbf{T}_{1\cdot\cdot P} \\ \vdots \\ \mathbf{T}_{F\cdot\cdot P} \end{bmatrix} \in \mathbb{C}^{PFM\times R}, \quad \mathbf{T}_{PFR\times FM} = \begin{bmatrix} \mathbf{T}_{1\cdot\cdot 1}^{\mathrm{T}} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{T}_{F\cdot\cdot 1}^{\mathrm{T}} \\ & \vdots & \\ \mathbf{T}_{1\cdot\cdot P}^{\mathrm{T}} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{T}_{F\cdot\cdot P}^{\mathrm{T}} \end{bmatrix} \in \mathbb{C}^{PFR\times FM}. \tag{4.46}$$

Setting $\mathbf{c}_{f,p\cdot}^{(\mathbf{S})} = \mathbf{1}_R$ and $\mathbf{c}_{f,p\cdot}^{(\mathcal{H})} = \mathbf{1}_M$ for a given $f \in \{1,...,F\}$ and $p \in \{1,...,P\}$, (4.45) gives $\mathbf{T}_{f\cdot\cdot p} = \mathbf{W}$. Furthermore, if we suppose that $\mathbf{W}$ is full column-rank implying $M \geq R$, then the unfolded matrix $\mathbf{T}_{PFM\times R}$ is also full column-rank. Consequently, assuming $K \geq M$ and from (4.43), the LS estimate of $\mathbf{S}$ is unique.

In the same way, setting $\mathbf{C}_{\cdot p\cdot}^{(\mathbf{S})} = \mathbf{1}_{F\times R}$ and $\mathbf{C}_{\cdot p\cdot}^{(\mathcal{H})} = \mathbf{1}_{F\times M}$ for a given $p \in \{1,...,P\}$, we can extract from $\mathbf{T}_{PFR\times FM}$ the following sub-matrix associated with the p-th block

$$\mathbf{P}_{FR\times FM} \triangleq \begin{bmatrix} \mathbf{W}^{\mathrm{T}} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{W}^{\mathrm{T}} \end{bmatrix} \in \mathbb{C}^{FR\times FM}. \tag{4.47}$$

Assume that $\mathbf{W}$ is full row-rank implying $M \leq R$, $\mathbf{P}_{FR\times FM}$ is full column-rank which leads to $\mathbf{T}_{PFR\times FM}$ full column-rank as well. Thus, assuming $\mathbf{S}$ full column-rank and from (4.42), the LS estimate of $\mathcal{H}$ is unique.

From the last reasonings, we can directly enunciate a sufficient condition for ensuring the LS identifiability of both channel tensor and symbol matrix by (4.42) and (4.43) when $M = R$ as follows.

Theorem 4.7. (Identifiability Condition when $M = R$). *Suppose perfect knowledge of the code matrix* $\mathbf{W}$ *and the allocation structures* $\mathcal{C}^{(\mathcal{H})}$ *and* $\mathcal{C}^{(\mathbf{S})}$. *Setting* $\mathbf{C}^{(\mathbf{S})}_{:p:} = \mathbf{1}_{F\times R}$ *and* $\mathbf{C}^{(\mathcal{H})}_{:p:} = \mathbf{1}_{F\times M}$ *for a given* $p \in \{1, ..., P\}$, *and assuming that* $\mathbf{W}$, $\mathbf{S}$ *and* $\mathbf{H}_{f::}$ *for all* $f \in \{1, ..., F\}$ *are full column-rank, thus we can ensure the LS identifiability of* $\mathcal{H}$ *and* $\mathbf{S}$ *by, respectively,* (4.42) *and* (4.43) *when* $M = R$.

The spreading code for the STF system is a $(M \times R)$-matrix instead of a $(M \times R \times J)$-tensor as for the TST system. The time diversity J allows to guarantee the LS estimate for both symbol and channel when $M \neq R$. For the TST coding, it is possible to set a minimum value of J according to Table 4.1 which can ensure that both matrices $\mathbf{G}_{PJM\times R}$ and $\mathbf{G}_{PJR\times M}$ will be full column-rank.

As discussed previously, when $\mathbf{S}$ and $\mathbf{H}_{f::}$ for all $f \in \{1, ..., F\}$ are full column-rank the LS identifiability of channel and symbol by (4.42) and (4.43) requires that $\mathbf{T}_{PFR\times FM}$ and $\mathbf{T}_{PFM\times R}$ are full column-rank respectively. Differently to TST coding, the allocation structures and a minimum value of P are essential to guarantee that these matrices are full column-rank. The next two theorems are proposed to ensure the LS identifiability in two cases: $M < R$ and $M > R$.

Theorem 4.8. (Identifiability Condition when $M < R$). *Suppose perfect knowledge of the code matrix* $\mathbf{W}$ *and the allocation structures* $\mathcal{C}^{(\mathcal{H})}$ *and* $\mathcal{C}^{(\mathbf{S})}$. *Set* $\mathbf{C}^{(\mathbf{S})}_{:p^*:} = \mathbf{1}_{F\times R}$ *and* $\mathbf{C}^{(\mathcal{H})}_{:p^*:} = \mathbf{1}_{F\times M}$ *for a given* $p^* \in \{1, ..., P\}$, *and assume that* $\mathbf{W}^{\mathrm{T}}$, $\mathbf{S}$ *and* $\mathbf{H}_{f::}$ *for all* $f \in \{1, ..., F\}$ *are full column-rank. If the slice matrices associated with a given* $f^* \in \{1, ..., F\}$ *are fixed such that:* $\mathbf{C}^{(\mathcal{H})}_{f^*::} = \mathbf{1}_{P\times M}$ *and* $\mathbf{C}^{(\mathbf{S})}_{f^*::}$ *is full column-rank implying* $P \geq R$, *then we can ensure the LS identifiability of* $\mathcal{H}$ *and* $\mathbf{S}$ *by, respectively,* (4.42) *and* (4.43) *when* $M < R$.

Proof: As proved above, setting $\mathbf{C}^{(\mathbf{S})}_{:p^*:} = \mathbf{1}_{F\times R}$ and $\mathbf{C}^{(\mathcal{H})}_{:p^*:} = \mathbf{1}_{F\times M}$ for a given $p^* \in \{1, ..., P\}$, and assuming that $\mathbf{W}^{\mathrm{T}}$ and $\mathbf{S}$ are full column-rank, thus the LS estimate of $\mathcal{H}$ is unique.

From (4.45), (4.46) and the definition of Khatri-Rao product, let us rewrite $\mathbf{T}_{PFM\times R}$ considering only the rows associated with $f^* \in \{1, ..., F\}$ as follows

$$\begin{aligned}\begin{bmatrix} \mathbf{T}_{f^*::1} \\ \vdots \\ \mathbf{T}_{f^*::P} \end{bmatrix} &= \begin{bmatrix} \mathrm{diag}\left(\mathbf{c}^{(\mathcal{H})}_{f^*,1:}\right) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathrm{diag}\left(\mathbf{c}^{(\mathcal{H})}_{f^*,P:}\right) \end{bmatrix} \begin{bmatrix} \mathbf{W}\,\mathrm{diag}\left(\mathbf{c}^{(\mathbf{S})}_{f^*,1:}\right) \\ \vdots \\ \mathbf{W}\,\mathrm{diag}\left(\mathbf{c}^{(\mathbf{S})}_{f^*,P:}\right) \end{bmatrix} \\ &= \mathrm{diag}\left(\mathrm{vec}\left(\mathbf{C}^{(\mathcal{H})^{\mathrm{T}}}_{f^*::}\right)\right)\left(\mathbf{C}^{(\mathbf{S})}_{f^*::} \diamond \mathbf{W}\right) \in \mathbb{C}^{PM\times R}. \end{aligned} \tag{4.48}$$

Rewriting (4.48) with $\mathbf{C}^{(\mathcal{H})}_{f^{*}\cdot\cdot} = \mathbf{1}_{P\times M}$, we obtain

$$\begin{bmatrix} \mathbf{T}_{f^{*}\cdot 1} \\ \vdots \\ \mathbf{T}_{f^{*}\cdot P} \end{bmatrix} = \mathbf{C}^{(\mathbf{S})}_{f^{*}\cdot\cdot} \diamond \mathbf{W}. \tag{4.49}$$

According to the Lemma 2.2, if $k\Big(\mathbf{C}^{(\mathbf{S})}_{f^{*}\cdot\cdot}\Big) + k(\mathbf{W}) \geq R+1$, then $\mathbf{C}^{(\mathbf{S})}_{f^{*}\cdot\cdot} \diamond \mathbf{W}$ is full column-rank for a given $f^{*} \in \{1,...,F\}$. Supposing that $\mathbf{W}$ is full row-rank, $\mathbf{W}$ has M independent columns and rows. In accordance with the k-rank definition, it implies $1 \leq k(\mathbf{W}) \leq M$, hence the maximum value of the k-rank of $\mathbf{W}$ will be equal to $M < R$ which is not enough to obtain $\mathbf{C}^{(\mathbf{S})}_{f^{*}\cdot\cdot} \diamond \mathbf{W}$ full column-rank.

However, if $\mathbf{C}^{(\mathbf{S})}_{f^{*}\cdot\cdot}$ is full column-rank implying $P \geq R$ then $\mathbf{C}^{(\mathbf{S})}_{f^{*}\cdot\cdot} \diamond \mathbf{W}$ is full column-rank, which leads to $\mathbf{T}_{PFM\times R}$ full column-rank as well. Thus, from (4.43) and assuming $\mathbf{H}_{f\cdot\cdot}$ for all $f \in \{1,...,F\}$ is full column-rank, the LS estimate of $\mathbf{S}$ is unique. Finally, satisfying the conditions enunciated in this theorem, the LS estimate of $\mathbf{S}$ and $\mathcal{H}$ is simultaneously ensured unique for $M < R$. ∎

Remark 4.5. *When $\mathbf{C}^{(\mathcal{H})}_{f^{*}\cdot\cdot} = \mathbf{1}_{P\times M}$ for a given f^{*} and $k(\mathbf{W}) = M < R$, $\mathbf{T}_{PFM\times R}$ is full column-rank if $k\Big(\mathbf{C}^{(\mathbf{S})}_{f^{*}\cdot\cdot}\Big) \geq R-M+1$. Thus, the theorem condition over $\mathbf{C}^{(\mathbf{S})}_{f^{*}\cdot\cdot}$ is relaxed because it is not required to be full column-rank, i.e. $k\Big(\mathbf{C}^{(\mathbf{S})}_{f^{*}\cdot\cdot}\Big) = R$.*

Theorem 4.9. (Identifiability Condition when $M > R$). *Suppose perfect knowledge of the code matrix $\mathbf{W}$ and of the allocation structures $\mathcal{C}^{(\mathcal{H})}$ and $\mathcal{C}^{(\mathbf{S})}$. Assume that $\mathbf{W}$, $\mathbf{S}$ and $\mathbf{H}_{f\cdot\cdot}$ for all $f \in \{1,...,F\}$ are full column-rank, and set $\mathcal{C}^{(\mathbf{S})} = \mathbf{1}_{F\times P\times R}$ and $\mathbf{c}^{(\mathcal{H})}_{f^{*},p^{*}\cdot} = \mathbf{1}_{M}$ for a given $p^{*} \in \{1,...,P\}$ and $f^{*} \in \{1,...,F\}$. If the matrix-slice $\mathbf{C}^{(\mathcal{H})}_{f\cdot\cdot}$ is full column-rank for all $f \in \{1,...,F\}$ implying $P \geq M$, then we can ensure the LS identifiability of $\mathcal{H}$ and $\mathbf{S}$ by, respectively,* (4.42) *and* (4.43) *when $M > R$.*

Proof: Analogously to (4.48), we can write a sub-matrix associated with $f \in \{1,...,F\}$ as

$$\begin{aligned} \begin{bmatrix} \mathbf{T}^{\mathrm{T}}_{f\cdot\cdot 1} \\ \vdots \\ \mathbf{T}^{\mathrm{T}}_{f\cdot\cdot P} \end{bmatrix} &= \begin{bmatrix} \operatorname{diag}\Big(\mathbf{c}^{(\mathbf{S})}_{f,1\cdot}\Big) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \operatorname{diag}\Big(\mathbf{c}^{(\mathbf{S})}_{f,P\cdot}\Big) \end{bmatrix} \begin{bmatrix} \mathbf{W}^{\mathrm{T}} \operatorname{diag}\Big(\mathbf{c}^{(\mathcal{H})}_{f,1\cdot}\Big) \\ \vdots \\ \mathbf{W}^{\mathrm{T}} \operatorname{diag}\Big(\mathbf{c}^{(\mathcal{H})}_{f,P\cdot}\Big) \end{bmatrix} \\ &= \operatorname{diag}\Big(\operatorname{vec}\Big(\mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}_{f\cdot\cdot}\Big)\Big) \Big(\mathbf{C}^{(\mathcal{H})}_{f\cdot\cdot} \diamond \mathbf{W}^{\mathrm{T}}\Big) \in \mathbb{C}^{PR\times M}. \end{aligned} \tag{4.50}$$

We can rewrite $\mathbf{T}_{PFR\times FM}$ given in (4.46) permuting in a convenient way its rows, applying

(4.50) and also setting $\mathcal{C}^{(\mathbf{S})} = \mathbf{1}_{F\times P\times R}$ as follows

$$\mathbf{T}_{FPR\times FM} = \begin{bmatrix} \begin{bmatrix} \mathbf{T}_{1\cdot\cdot 1}^{\mathrm{T}} \\ \ddots \\ \mathbf{T}_{1\cdot\cdot P}^{\mathrm{T}} \end{bmatrix} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \begin{bmatrix} \mathbf{T}_{F\cdot\cdot 1}^{\mathrm{T}} \\ \ddots \\ \mathbf{T}_{F\cdot\cdot P}^{\mathrm{T}} \end{bmatrix} \end{bmatrix} \in \mathbb{C}^{FPR\times FM} \tag{4.51}$$

$$= \begin{bmatrix} \mathbf{C}_{1\cdot\cdot}^{(\mathcal{H})} \diamond \mathbf{W}^{\mathrm{T}} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{C}_{F\cdot\cdot}^{(\mathcal{H})} \diamond \mathbf{W}^{\mathrm{T}} \end{bmatrix}. \tag{4.52}$$

Observe (4.52) that each f-th matrix on diagonal, $\mathbf{C}_{f\cdot\cdot}^{(\mathcal{H})} \diamond \mathbf{W}^{\mathrm{T}}$, have to be full column-rank in order to $\mathbf{T}_{PFR\times FM}$ be full column-rank as well. From Lemma 2.2, if $k\Big(\mathbf{C}_{f\cdot\cdot}^{(\mathcal{H})}\Big) + k\big(\mathbf{W}^{\mathrm{T}}\big) \geq M+1$ for all $f \in \{1,...,F\}$, then $\mathbf{C}_{f\cdot\cdot}^{(\mathcal{H})} \diamond \mathbf{W}^{\mathrm{T}}$ is full column-rank for all $f \in \{1,...,F\}$ and consequently, $\mathbf{T}_{PFR\times FM}$ is full column-rank.

Analogously to the previous theorem, considering that $\mathbf{W}$ is full column-rank and by the k-rank definition, we have $1 \leq k\left(\mathbf{W}^{\mathrm{T}}\right) \leq R$. Choosing $\mathbf{C}_{f\cdot\cdot}^{(\mathcal{H})}$ full column-rank for all f, we can guarantee that $\mathbf{T}_{PFR\times FM}$ is full column-rank. Consequently, from (4.42) and assuming that $\mathbf{S}$ is full column-rank, the LS estimate of $\mathcal{H}$ is unique.

As proved previously for Theorem 4.7, additionally setting $\mathbf{c}_{f^*,p^*\cdot}^{(\mathcal{H})} = \mathbf{1}_M$ for a given $p^* \in \{1,...,P\}$ and $f^* \in \{1,...,F\}$, and assuming that $\mathbf{W}$ and $\mathbf{H}_{f\cdot\cdot}$ for all $f \in \{1,...,F\}$ are full column-rank, we obtain that the LS estimate of $\mathbf{S}$ is unique. Taking into account both conditions, the LS estimate of $\mathbf{S}$ and $\mathcal{H}$ is simultaneously unique for $M > R$. ∎

Remark 4.6. *Analogously to the last remark for Theorem 4.8, the condition that $\mathbf{C}_{f\cdot\cdot}^{(\mathcal{H})}$ is full column-rank for all $f \in \{1,...,F\}$ becomes more flexible when $k\big(\mathbf{W}^{\mathrm{T}}\big) = R < M$. It is enough to satisfy $k\Big(\mathbf{C}_{f\cdot\cdot}^{(\mathcal{H})}\Big) \geq M - R + 1$ for all f.*

4.1.3 ALS receivers

Semi-blind receivers using a two-step ALS algorithm are derived to jointly estimate channel and symbols for the TST and STF systems. Table 4.2 summarizes the ALS algorithm for both systems. Notice that the estimation of symbol matrix at it-th iteration is conditioned to the knowledge of channel matrix (or matrix unfolding for the STF system) estimated at the $(it-1)$-th iteration.

4.1.4 Zero-Forcing (ZF) receivers

Considering the perfect knowledge of channel for both TST and STF systems, we can deduce the Zero-Forcing (ZF) receivers for estimating the symbol matrix directly from (4.5), (4.27) and

Table 4.2: ALS algorithm for semi-blind joint symbol and channel estimation.

TST system: ALS-TST	**STF system: ALS-STF**
1. Initialization ($it = 0$): randomly initialize $\hat{\mathbf{H}}_{(0)}$ and $\hat{\mathcal{H}}_{(0)}$	
2. $it = it + 1$	
3. Compute an LS estimate of $\mathbf{S}$ from $\hat{\mathbf{H}}_{(it-1)}$ and $\hat{\mathcal{H}}_{(it-1)}$, using:	
(4.5) or (4.27)	(4.43)
4. Compute an LS estimate of $\mathbf{H}$ and $\mathcal{H}$ from $\hat{\mathbf{S}}_{(it)}$, using:	
(4.4) or (4.26)	(4.42)
5. Repeat steps (2)-(4) until convergence.	
6. Eliminate the scaling ambiguity with $\alpha = s_{1,1}/\hat{s}_{1,1}$:	
$\hat{\mathbf{S}} = \alpha\,\hat{\mathbf{S}}$	
$\hat{\mathbf{H}} = 1/\alpha\,\hat{\mathbf{H}}$	$\hat{\mathbf{H}}_{K\times FM} = 1/\alpha\,\hat{\mathbf{H}}_{K\times FM}$
7. Project the estimated symbols onto the symbol alphabet.	

(4.43).

- ZF-TST receiver: For TST system modeled by the PT-(2,4) and a constrained CP-4 model respectively

$$\hat{\mathbf{S}} = \tilde{\mathbf{X}}_{N\times PJK}\left(\left(\mathbf{I}_{PJ}\otimes\mathbf{H}\right)\mathbf{G}_{PJM\times R}\right)^{\dagger^{\mathrm{T}}}, \tag{4.53}$$

$$\hat{\mathbf{S}} = {}^{1}\!/_{M}\,\tilde{\mathbf{X}}_{N\times PJK}\left(\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}}\diamond\mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}}\diamond\mathbf{W}_{J\times MR}\diamond\mathbf{H}\boldsymbol{\Omega}^{(1)}\right)^{\mathrm{T}^{\dagger}}\boldsymbol{\Omega}^{(2)^{\mathrm{T}}}, \tag{4.54}$$

  where $\mathbf{G}_{PJM\times R}$ is given by (3.4) and (3.12), $\boldsymbol{\Omega}^{(1)} = \mathbf{I}_M\otimes\mathbf{1}_R^{\mathrm{T}}$ and $\boldsymbol{\Omega}^{(2)} = \mathbf{1}_M^{\mathrm{T}}\otimes\mathbf{I}_R$.
- ZF-STF receiver: For STF system.

$$\hat{\mathbf{S}} = \tilde{\mathbf{X}}_{N\times PFK}\left(\left(\mathbf{I}_P\otimes\mathrm{bdiag}(\mathbf{H}_{1\cdot\cdot},\ldots,\mathbf{H}_{F\cdot\cdot})\right)\mathbf{T}_{PFM\times R}\right)^{\dagger^{\mathrm{T}}}, \tag{4.55}$$

  where $\mathbf{T}_{PFM\times R}$ is given by (4.45) and (4.46).

4.2 Levenberg-Marquardt (LM)

The Levenberg-Marquardt method [108], denoted by LM, is a well-known alternative to the Gauss-Newton (GN) method of solving nonlinear LS problems. The main idea of this method is to combine the advantages of two minimization methods, the Steepest Descent (SD) and the GN methods, by switching between both techniques through a damping parameter. The LM method was employed to estimate the parameters of CP model in [105] and of other tensor models in [106, 61, 62].

The minimization of the sum of the squared errors between the received signal model and the noisy received signals is a nonlinear LS problem. Hence, the LM approach is an estimation method which, differently to the ALS approach, provides an update for all parameters by successive approximations.

In this section, we introduce two receivers based on the LM algorithm for both TST and STF systems. Firstly, we present the Gradient and Jacobian expressions for both systems and in the sequence, derive the proposed receivers.

4.2.1 Gradient and Jacobian expressions

Let us assume that $\mathbf{x}(\mathbf{p}) \in \mathbb{C}^{L\times 1}$ is a vector of the model output of received signals, which depends on a vector of parameters to estimate $\mathbf{p} \in \mathbb{C}^{Q\times 1}$ with Q and L denoting the number of model parameters and received signal samples respectively. Thus, the estimation error vector is denoted by $\mathbf{r}(\mathbf{p}) = \mathbf{x}(\mathbf{p}) - \tilde{\mathbf{x}}$, where $\tilde{\mathbf{x}}$ represents the noisy version of the received signal vector. The cost function to be minimized can be formulated as follows

$$\begin{aligned}\phi(\mathbf{p}) &= \frac{1}{2}\left\|\mathbf{x}(\mathbf{p}) - \tilde{\mathbf{x}}\right\|_2^2 \\ &= \frac{1}{2}\left\|\mathbf{r}(\mathbf{p})\right\|_2^2 = \frac{1}{2}\,\mathbf{r}^{\mathrm{H}}(\mathbf{p})\;\mathbf{r}(\mathbf{p})\,.\end{aligned} \tag{4.56}$$

We may define the complex gradient vector $\mathbf{g}(\mathbf{p}) = \nabla\phi \in \mathbb{C}^{Q\times 1}$ of $\phi\,(\mathbf{p})$ with respect to $\mathbf{p}$ by

$$\nabla_q\phi \triangleq \frac{\partial\phi}{\partial a_q} + i\frac{\partial\phi}{\partial b_q} \quad\Longrightarrow\quad \mathbf{g}(\mathbf{p}) = \mathbf{J}^{\mathrm{H}}(\mathbf{p})\,\mathbf{r}(\mathbf{p})\,, \tag{4.57}$$

where $p_q = a_q + ib_q$ is the q-th element of $\mathbf{p}$ and $\mathbf{J}(\mathbf{p}) \in \mathbb{C}^{L\times Q}$ is the Jacobian matrix containing the first partial derivatives of $\mathbf{r}(\mathbf{p})$,

$$\left[\mathbf{J}(\mathbf{p})\right]_{l,q} \triangleq \frac{\partial r_l(\mathbf{p})}{\partial p_q} \quad\Longleftrightarrow\quad \mathbf{J}(\mathbf{p}) = \frac{\partial\mathbf{r}(\mathbf{p})}{\partial\mathbf{p}} = \frac{\partial\mathbf{x}(\mathbf{p})}{\partial\mathbf{p}}. \tag{4.58}$$

TST system

Let us consider the TST system given by (3.5). From (3.11), we define the following vectors by

$$\mathbf{x}_{NPJK}(\mathbf{p}) \triangleq \mathrm{vec}(\mathbf{X}_{PJK\times N}(\mathbf{p})) = \left(\mathbf{I}_N \otimes \left(\mathbf{I}_{PJ} \otimes \mathbf{H}\right)\mathbf{G}_{PJM\times R}\right)\,\mathbf{p}_{\mathbf{S}} \in \mathbb{C}^{NPJK\times 1}, \tag{4.59}$$

$$\mathbf{x}_{KPJN}(\mathbf{p}) \triangleq \mathrm{vec}(\mathbf{X}_{PJN\times K}(\mathbf{p})) = \left(\mathbf{I}_K \otimes \left(\mathbf{I}_{PJ} \otimes \mathbf{S}\right)\mathbf{G}_{PJR\times M}\right)\,\mathbf{p}_{\mathbf{H}} \in \mathbb{C}^{KPJN\times 1}, \tag{4.60}$$

with

$$\begin{aligned}\mathbf{p}_{\mathbf{S}} &\triangleq \mathrm{vec}\left(\mathbf{S}^{\mathrm{T}}\right) \in \mathbb{C}^{NR\times 1}, \\ \mathbf{p}_{\mathbf{H}} &\triangleq \mathrm{vec}\left(\mathbf{H}^{\mathrm{T}}\right) \in \mathbb{C}^{KM\times 1}\end{aligned} \tag{4.61}$$

and the vector of all parameters to estimate is represented by

$$\mathbf{p} \triangleq \begin{bmatrix}\mathbf{p}_{\mathbf{S}} \\ \mathbf{p}_{\mathbf{H}}\end{bmatrix} \in \mathbb{C}^{(NR+KM)\times 1}. \tag{4.62}$$

Applying the property of the Kronecker product given in (A.4), we can rewrite (4.59) and (4.60), respectively, as

$$\begin{aligned}\mathbf{x}_{NPJK}(\mathbf{p}) &= \left(\mathbf{I}_N \otimes \left(\mathbf{I}_{PJ} \otimes \mathbf{H}\right)\right)\left(\mathbf{I}_N \otimes \mathbf{G}_{PJM\times R}\right)\,\mathbf{p}_{\mathbf{S}} \\ &= \left(\mathbf{I}_{NPJ} \otimes \mathbf{H}\right)\left(\mathbf{I}_N \otimes \mathbf{G}_{PJM\times R}\right)\,\mathbf{p}_{\mathbf{S}},\end{aligned} \tag{4.63}$$

$$\begin{aligned}\mathbf{x}_{KPJN}(\mathbf{p}) &= \left(\mathbf{I}_K \otimes \left(\mathbf{I}_{PJ} \otimes \mathbf{S}\right)\right)\left(\mathbf{I}_K \otimes \mathbf{G}_{PJR\times M}\right)\,\mathbf{p}_{\mathbf{H}} \\ &= \left(\mathbf{I}_{KPJ} \otimes \mathbf{S}\right)\left(\mathbf{I}_K \otimes \mathbf{G}_{PJR\times M}\right)\,\mathbf{p}_{\mathbf{H}}.\end{aligned} \tag{4.64}$$

Note that the expressions in (4.63) and (4.64) represent two different vectorial forms of the received signal tensor model which its elements can be related by

$$\begin{aligned} x_{k,n,p,j} &= [\mathbf{x}_{NPJK}(\mathbf{p})]_{(n-1)PJK+(p-1)JK+(j-1)K+k} \\ &= [\mathbf{x}_{KPJN}(\mathbf{p})]_{(k-1)PJN+(p-1)JN+(j-1)N+n}\,. \end{aligned} \tag{4.65}$$

Then, we can relate both vectors through a permutation matrix as follows

$$\mathbf{x}_{NPJK}(\mathbf{p}) = \mathbf{\Pi}\,\mathbf{x}_{KPJN}(\mathbf{p})\,, \tag{4.66}$$

where $\mathbf{\Pi} \in \mathbb{R}^{NPJK\times KPJN}$ is defined as

$$\mathbf{\Pi} \triangleq \sum_{n=1}^{N}\sum_{k=1}^{K} \mathbf{e}_n^{(N)}\mathbf{e}_k^{(K)^{\mathrm{T}}} \otimes \mathbf{I}_{PJ} \otimes \mathbf{e}_k^{(K)}\mathbf{e}_n^{(N)^{\mathrm{T}}} \tag{4.67}$$

and $\mathbf{e}_i^{(I)}$ denotes a column vector of length I with 1 in the i-th position and 0 in every other position.

Using (4.63) and (4.64) in (4.56), we have two equivalent expressions for the cost function

$$\phi(\mathbf{p}) = \frac{1}{2}\left\|\left(\mathbf{I}_{NPJ}\otimes\mathbf{H}\right)\left(\mathbf{I}_N\otimes\mathbf{G}_{PJM\times R}\right)\,\mathbf{p_S} - \tilde{\mathbf{x}}_{NPJK}\right\|^2 \tag{4.68}$$

$$= \frac{1}{2}\left\|\left(\mathbf{I}_{KPJ}\otimes\mathbf{S}\right)\left(\mathbf{I}_K\otimes\mathbf{G}_{PJR\times M}\right)\,\mathbf{p_H} - \tilde{\mathbf{x}}_{KPJN}\right\|^2. \tag{4.69}$$

We calculate the Jacobian matrix with respect to $\mathbf{p_S}$ and $\mathbf{p_H}$ using the definition in (4.58) by

$$\mathbf{J_S} \triangleq \frac{\partial\,\mathbf{r}(\mathbf{p})}{\partial\,\mathbf{p_S}} = \left(\mathbf{I}_{NPJ}\otimes\mathbf{H}\right)\left(\mathbf{I}_N\otimes\mathbf{G}_{PJM\times R}\right) \in \mathbb{C}^{NPJK\times NR}, \tag{4.70}$$

$$\mathbf{J_H} \triangleq \frac{\partial\,\mathbf{r}(\mathbf{p})}{\partial\,\mathbf{p_H}} = \mathbf{\Pi}\,\left(\mathbf{I}_{KPJ}\otimes\mathbf{S}\right)\left(\mathbf{I}_K\otimes\mathbf{G}_{PJR\times M}\right) \in \mathbb{C}^{NPJK\times KM}, \tag{4.71}$$

with $\mathbf{\Pi} \in \mathbb{R}^{NPJK\times KPJN}$ defined in (4.67).

Note that the permutation matrix is employed with the purpose of imposing a specific order of indices and in this case, equal to n, j, p, k instead of k, p, j, n. Finally, we can construct the overall Jacobian matrix $\mathbf{J}(\mathbf{p}) \in \mathbb{C}^{NPJK\times(NR+KM)}$ denoted by

$$\mathbf{J}(\mathbf{p}) \triangleq \begin{bmatrix}\mathbf{J_S} & \mathbf{J_H}\end{bmatrix}. \tag{4.72}$$

From the definition (4.57), we can calculate the gradient of $\phi(\mathbf{p})$ with respect to $\mathbf{p_S}$ and $\mathbf{p_H}$,

$$\begin{aligned} \mathbf{g_S} \triangleq \frac{\partial\,\phi(\mathbf{p})}{\partial\,\mathbf{p_S}} &= \mathbf{J_S^H}\,\left(\mathbf{J_S}\,\mathbf{p_S} - \tilde{\mathbf{x}}_{NPJK}\right) \\ &= \left(\mathbf{J_S^H}\,\mathbf{J_S}\right)\mathbf{p_S} - \mathbf{J_S^H}\,\tilde{\mathbf{x}}_{NPJK} \in \mathbb{C}^{NR\times 1}, \end{aligned} \tag{4.73}$$

$$\begin{aligned} \mathbf{g_H} \triangleq \frac{\partial\,\phi(\mathbf{p})}{\partial\,\mathbf{p_H}} &= \mathbf{J_H^H}\,\left(\mathbf{J_H}\,\mathbf{p_H} - \tilde{\mathbf{x}}_{NPJK}\right) \\ &= \left(\mathbf{J_H^H}\,\mathbf{J_H}\right)\mathbf{p_H} - \mathbf{J_H^H}\,\tilde{\mathbf{x}}_{NPJK} \in \mathbb{C}^{KM\times 1}. \end{aligned} \tag{4.74}$$

And analogously to (4.72), we can define the overall gradient vector $\mathbf{g}(\mathbf{p}) \in \mathbb{C}^{(NR+KM)\times 1}$ by

$$\mathbf{g}(\mathbf{p}) \triangleq \begin{bmatrix} \mathbf{g_S} \\ \mathbf{g_H} \end{bmatrix}. \tag{4.75}$$

Generalization to multiuser case

From (3.81) and (3.82), we define the following vectors by

$$\mathbf{x}_{NPJK}(\mathbf{p}) \triangleq \text{vec}(\mathbf{X}_{PJK\times N}(\mathbf{p})) = (\mathbf{I}_N \otimes (\mathbf{\Omega} \,|\otimes|\, \mathbf{H}_{K\times QM})\, \mathbf{G}_{QPJM\times QR})\ \mathbf{p_S} \in \mathbb{C}^{NPJK\times 1}, \tag{4.76}$$

$$\mathbf{x}_{KPJN}(\mathbf{p}) \triangleq \text{vec}(\mathbf{X}_{PJN\times K}(\mathbf{p})) = (\mathbf{I}_K \otimes (\mathbf{\Omega} \,|\otimes|\, \mathbf{S}_{N\times QR})\, \mathbf{G}_{QPJR\times QM})\ \mathbf{p_H} \in \mathbb{C}^{KPJN\times 1}, \tag{4.77}$$

with

$$\begin{aligned} \mathbf{p_S} &\triangleq \text{vec}(\mathbf{S}_{QR\times N}) \in \mathbb{C}^{NQR\times 1}, \\ \mathbf{p_H} &\triangleq \text{vec}(\mathbf{H}_{QM\times K}) \in \mathbb{C}^{KQM\times 1} \end{aligned} \tag{4.78}$$

and the vector of all parameters to estimate is defined in (4.62) with $\mathbf{p} \in \mathbb{C}^{(NQR+KQM)\times 1}$.

From the permutation matrix and the cost function given, respectively, in (4.67) and (4.56), we write two equivalent expressions for the cost function using (4.76) and (4.77)

$$\phi(\mathbf{p}) = \frac{1}{2} \left\| (\mathbf{I}_N \otimes (\mathbf{\Omega} \,|\otimes|\, \mathbf{H}_{K\times QM})\, \mathbf{G}_{QPJM\times QR})\ \mathbf{p_S} - \tilde{\mathbf{x}}_{NPJK} \right\|^2 \tag{4.79}$$

$$= \frac{1}{2} \left\| (\mathbf{I}_K \otimes (\mathbf{\Omega} \,|\otimes|\, \mathbf{S}_{N\times QR})\, \mathbf{G}_{QPJR\times QM})\ \mathbf{p_H} - \tilde{\mathbf{x}}_{KPJN} \right\|^2, \tag{4.80}$$

with $\mathbf{\Pi} \in \mathbb{R}^{NPJK\times KPJN}$.

According to the definition given by (4.57) and (4.58), we can write the Jacobian matrix and the gradient vector with respect to $\mathbf{p_S}$ and $\mathbf{p_H}$ as

$$\mathbf{J_S} = \mathbf{I}_N \otimes (\mathbf{\Omega} \,|\otimes|\, \mathbf{H}_{K\times QM})\, \mathbf{G}_{QPJM\times QR} \in \mathbb{C}^{NPJK\times NQR}, \tag{4.81}$$

$$\mathbf{J_H} = \mathbf{\Pi}\ (\mathbf{I}_K \otimes (\mathbf{\Omega} \,|\otimes|\, \mathbf{S}_{N\times QR})\, \mathbf{G}_{QPJR\times QM}) \in \mathbb{C}^{NPJK\times KQM}, \tag{4.82}$$

$$\mathbf{g_S} = \left(\mathbf{J_S^H}\, \mathbf{J_S}\right) \mathbf{p_S} - \mathbf{J_S^H}\, \tilde{\mathbf{x}}_{NPJK} \in \mathbb{C}^{NQR\times 1}, \tag{4.83}$$

$$\mathbf{g_H} = \left(\mathbf{J_H^H}\, \mathbf{J_H}\right) \mathbf{p_H} - \mathbf{J_H^H}\, \tilde{\mathbf{x}}_{NPJK} \in \mathbb{C}^{KQM\times 1}. \tag{4.84}$$

The overall Jacobian matrix $\mathbf{J}(\mathbf{p}) \in \mathbb{C}^{NPJK\times(NQR+KQM)}$ and gradient vector $\mathbf{g}(\mathbf{p}) \in \mathbb{C}^{(NQR+KQM)\times 1}$ are defined respectively in (4.72) and (4.75).

STF system

Now, we consider the STF system [62] given by (3.27) to develop analogously the equivalent expressions for the STF context. From the vectorization of both unfolded matrices given in (3.32), we obtain the following vectors

$$\mathbf{x}_{NPFK}(\mathbf{p}) \triangleq \text{vec}\left(\mathbf{X}_{PFK\times N}(\mathbf{p})\right) = (\mathbf{I}_N \otimes \mathbf{\Omega}_1 \mathbf{T}_{PFM\times R})\ \mathbf{p_S} \in \mathbb{C}^{NPFK\times 1}, \tag{4.85}$$

$$\mathbf{x}_{KPFN}(\mathbf{p}) \triangleq \text{vec}\left(\mathbf{X}_{PFN\times K}(\mathbf{p})\right) = (\mathbf{I}_K \otimes \mathbf{\Omega}_2 \mathbf{T}_{PFR\times FM})\ \mathbf{p_H} \in \mathbb{C}^{KPFN\times 1}, \tag{4.86}$$

where $\mathbf{\Omega}_1 \triangleq \mathbf{I}_P \otimes \text{bdiag}(\mathbf{H}_{1\cdot\cdot}, \ldots, \mathbf{H}_{F\cdot\cdot})$, $\mathbf{\Omega}_2 \triangleq \mathbf{I}_{PF} \otimes \mathbf{S}$, $\mathbf{p_S} \in \mathbb{C}^{NR\times 1}$ and $\mathbf{p_H} \triangleq \text{vec}(\mathbf{H}_{FM\times K}) \in \mathbb{C}^{KFM\times 1}$, and the overall estimate vector $\mathbf{p} \in \mathbb{C}^{(NR+KFM)\times 1}$ given by (4.62).

Analogously to (4.65)-(4.67), we can relate each element of two vectors in (4.85) and (4.86) by

$$\begin{aligned} x_{f,p,n,k} &= [\mathbf{x}_{NPFK}(\mathbf{p})]_{(n-1)PFK+(p-1)FK+(f-1)K+k} \\ &= [\mathbf{x}_{KPFN}(\mathbf{p})]_{(k-1)PFN+(p-1)FN+(f-1)N+n} \end{aligned} \tag{4.87}$$

and also by

$$\mathbf{x}_{NPFK}(\mathbf{p}) = \mathbf{\Pi}\, \mathbf{x}_{KPFN}(\mathbf{p}) \tag{4.88}$$

with $\mathbf{\Pi} \in \mathbb{R}^{NPFK\times KPFN}$ given as

$$\mathbf{\Pi} \triangleq \sum_{n=1}^{N}\sum_{k=1}^{K} \mathbf{e}_n^{(N)}\mathbf{e}_k^{(K)^{\mathrm{T}}} \otimes \mathbf{I}_{PF} \otimes \mathbf{e}_k^{(K)}\mathbf{e}_n^{(N)^{\mathrm{T}}}. \tag{4.89}$$

We can write two equivalent expressions for the cost function defined in (4.56) using (4.85) and (4.86) as

$$\phi(\mathbf{p}) = \frac{1}{2}\left\|\left(\mathbf{I}_N \otimes \mathbf{\Omega}_1\mathbf{T}_{PFM\times R}\right)\, \mathbf{p_S} - \tilde{\mathbf{x}}_{NPFK}\right\|^2 \tag{4.90}$$

$$= \frac{1}{2}\left\|\left(\mathbf{I}_K \otimes \mathbf{\Omega}_2\mathbf{T}_{PFR\times FM}\right)\, \mathbf{p_H} - \tilde{\mathbf{x}}_{KPFN}\right\|^2. \tag{4.91}$$

According to the definition given by (4.57) and (4.58), we can write the Jacobian matrix and the gradient vector with respect to $\mathbf{p_S}$ and $\mathbf{p_H}$ as

$$\mathbf{J_S} = \mathbf{I}_N \otimes \mathbf{\Omega}_1\mathbf{T}_{PFM\times R} \in \mathbb{C}^{NPFK\times NR}, \tag{4.92}$$

$$\mathbf{J_H} = \mathbf{\Pi}\, \left(\mathbf{I}_K \otimes \mathbf{\Omega}_2\mathbf{T}_{PFR\times FM}\right) \in \mathbb{C}^{NPFK\times KFM}, \tag{4.93}$$

$$\mathbf{g_S} = \left(\mathbf{J_S}^{\mathrm{H}}\,\mathbf{J_S}\right)\mathbf{p_S} - \mathbf{J_S}^{\mathrm{H}}\,\tilde{\mathbf{x}}_{NPFK} \in \mathbb{C}^{NR\times 1}, \tag{4.94}$$

$$\mathbf{g_H} = \left(\mathbf{J_H}^{\mathrm{H}}\,\mathbf{J_H}\right)\mathbf{p_H} - \mathbf{J_H}^{\mathrm{H}}\,\tilde{\mathbf{x}}_{NPFK} \in \mathbb{C}^{KFM\times 1}. \tag{4.95}$$

with $\mathbf{\Pi} \in \mathbb{R}^{NPFK\times KPFN}$ given in (4.89). Remark that the index order is fixed as n, p, f, k.

In analogy to the TST case, the overall Jacobian matrix $\mathbf{J}(\mathbf{p}) \in \mathbb{C}^{NPFK\times(NR+KFM)}$ and gradient vector $\mathbf{g}(\mathbf{p}) \in \mathbb{C}^{(NR+KFM)\times 1}$ are respectively represented by (4.72) and (4.75) as well.

4.2.2 LM and ALM receivers

As previously mentioned, the LM method combines the SD and GN methods. According to the SD, the sum of the squared errors is reduced by updating the parameters in the negative direction of the gradient descent. For the GN method, the sum of the squared errors is reduced by considering a linear approximation of the estimation error at each iteration, which leads to a locally quadratic approximation of the cost function.

Let us assume that $\mathbf{p}_{(it)}$ is an estimation of $\mathbf{p}$ calculated at the it-th iteration, and $\mathbf{p}_{(it+1)}$ is obtained by

$$\mathbf{p}_{(it+1)} = \mathbf{p}_{(it)} + \Delta\mathbf{p}, \tag{4.96}$$

where $\Delta\mathbf{p}$ enforces the descending condition, i.e. $\phi\big(\mathbf{p}_{(it)}+\Delta\mathbf{p}\big)\leq\phi\big(\mathbf{p}_{(it)}\big)$. For small $\|\Delta\mathbf{p}\|$, we can write $\mathbf{r}\big(\mathbf{p}_{(it+1)}\big)$ in Taylor series truncated at the order term 1 as

$$\mathbf{r}\big(\mathbf{p}_{(it+1)}\big)\simeq\mathbf{r}\big(\mathbf{p}_{(it)}\big)+\mathbf{J}\big(\mathbf{p}_{(it)}\big)\,\Delta\mathbf{p}, \tag{4.97}$$

where $\mathbf{r}\big(\mathbf{p}_{(it)}\big)=\mathbf{x}\big(\mathbf{p}_{(it)}\big)-\tilde{\mathbf{x}}$ represents the estimation error at the it-th iteration.

From (4.56) and using this approximation of $\mathbf{r}\big(\mathbf{p}_{(it+1)}\big)$ in (4.97), we can write the cost function by

$$\begin{aligned}\phi\big(\mathbf{p}_{(it)}+\Delta\mathbf{p}\big)&=\frac{1}{2}\,\mathbf{r}^{\mathrm{H}}\big(\mathbf{p}_{(it+1)}\big)\,\mathbf{r}\big(\mathbf{p}_{(it+1)}\big) &(4.98)\\ &\simeq\phi\big(\mathbf{p}_{(it)}\big)+\frac{1}{2}\,\Delta\mathbf{p}^{\mathrm{H}}\,\mathbf{J}^{\mathrm{H}}\big(\mathbf{p}_{(it)}\big)\,\mathbf{r}\big(\mathbf{p}_{(it)}\big)+\\ &\quad+\frac{1}{2}\,\mathbf{r}^{\mathrm{H}}\big(\mathbf{p}_{(it)}\big)\,\mathbf{J}\big(\mathbf{p}_{(it)}\big)\,\Delta\mathbf{p}+\frac{1}{2}\,\big\|\mathbf{J}\big(\mathbf{p}_{(it)}\big)\,\Delta\mathbf{p}\big\|_F^2\\ &\triangleq\tilde{\phi}\big(\mathbf{p}_{(it)}+\Delta\mathbf{p}\big)\,. &(4.99)\end{aligned}$$

The GN method is based on the linear approximation (4.97) in the neighborhood of $\mathbf{p}_{(it)}$, then $\Delta\mathbf{p}_{\mathrm{GN}}$ is given by

$$\Delta\mathbf{p}_{\mathrm{GN}}=\min_{\Delta\mathbf{p}}\tilde{\phi}\big(\mathbf{p}_{(it)}+\Delta\mathbf{p}\big)\,, \tag{4.100}$$

the gradient and the Hessian matrix of $\tilde{\phi}\left(\mathbf{p}_{(it)}+\Delta\mathbf{p}\right)$ are

$$\frac{\partial\tilde{\phi}\big(\mathbf{p}_{(it)}+\Delta\mathbf{p}\big)}{\partial\Delta\mathbf{p}^*}=\mathbf{J}^{\mathrm{H}}\big(\mathbf{p}_{(it)}\big)\,\mathbf{r}\left(\mathbf{p}_{(it)}\right)+\mathbf{J}^{\mathrm{H}}\big(\mathbf{p}_{(it)}\big)\,\mathbf{J}\big(\mathbf{p}_{(it)}\big)\,\Delta\mathbf{p}, \tag{4.101}$$

$$\frac{\partial^2\tilde{\phi}\big(\mathbf{p}_{(it)}+\Delta\mathbf{p}\big)}{\partial\Delta\mathbf{p}\,\partial\Delta\mathbf{p}^{\mathrm{H}}}=\mathbf{J}^{\mathrm{H}}\big(\mathbf{p}_{(it)}\big)\,\mathbf{J}\big(\mathbf{p}_{(it)}\big)\,, \tag{4.102}$$

respectively.

Note that if $\mathbf{J}\big(\mathbf{p}_{(it)}\big)$ is full column-rank, implying $L\geq Q$, then the Hessian matrix is positive-definitive and it means that the function $\tilde{\phi}\big(\mathbf{p}_{(it)}+\Delta\mathbf{p}\big)$ has only one minimum. This minimum can be calculated by solving

$$\begin{aligned}\frac{\partial\tilde{\phi}\big(\mathbf{p}_{(it)}+\Delta\mathbf{p}\big)}{\partial\Delta\mathbf{p}^*}=\mathbf{0}\quad\Rightarrow\quad&\mathbf{J}^{\mathrm{H}}\big(\mathbf{p}_{(it)}\big)\,\mathbf{J}\big(\mathbf{p}_{(it)}\big)\,\Delta\mathbf{p}_{\mathrm{GN}}=-\mathbf{J}^{\mathrm{H}}\big(\mathbf{p}_{(it)}\big)\,\mathbf{r}\big(\mathbf{p}_{(it)}\big)\\ &=-\mathbf{g}\big(\mathbf{p}_{(it)}\big)\,, &(4.103)\end{aligned}$$

where $\mathbf{g}\left(\mathbf{p}_{(it)}\right)$ is the gradient of ϕ in $\mathbf{p}_{(it)}$, which is given by (4.57).

We emphasize that $\mathbf{J}^{\mathrm{H}}\big(\mathbf{p}_{(it)}\big)\,\mathbf{J}\big(\mathbf{p}_{(it)}\big)$ must be positive-definite at all iterations to guarantee $\phi\big(\mathbf{p}_{(it+1)}\big)\leq\phi\big(\mathbf{p}_{(it)}\big)$ for any iteration it. Nevertheless, this matrix can have null-eigenvalues in practice. According to the Levenberg-Marquardt method, $\Delta\mathbf{p}_{\mathrm{LM}}$ can be calculated by

$$\big(\mathbf{J}^{\mathrm{H}}\big(\mathbf{p}_{(it)}\big)\,\mathbf{J}\big(\mathbf{p}_{(it)}\big)+\lambda_{(it)}\,\mathbf{I}_Q\big)\,\Delta\mathbf{p}_{\mathrm{LM}}=-\mathbf{g}\big(\mathbf{p}_{(it)}\big) \tag{4.104}$$

$$\Longrightarrow\quad\Delta\mathbf{p}_{\mathrm{LM}}=-\left(\mathbf{J}^{\mathrm{H}}\big(\mathbf{p}_{(it)}\big)\,\mathbf{J}\big(\mathbf{p}_{(it)}\big)+\lambda_{(it)}\,\mathbf{I}_Q\right)^{-1}\mathbf{g}\big(\mathbf{p}_{(it)}\big)\,, \tag{4.105}$$

where $\lambda_{(it)} \geq 0$ is the damping parameter at the it-th iteration.

Observe that for large values of $\lambda_{(it)}$, we have

$$\Delta \mathbf{p}_{\mathrm{LM}} \simeq -\frac{1}{\lambda^{(it)}}\, \mathbf{g}\big(\mathbf{p}_{(it)}\big) = -\frac{1}{\lambda_{(it)}}\, \frac{\partial \phi\big(\mathbf{p}_{(it)}\big)}{\partial \mathbf{p}_{(it)}}, \tag{4.106}$$

i.e. a short step in the steepest descent direction which is interesting when $\mathbf{p}_{(it)}$ is far from the solution in order to obtain fast initial progress.. If $\lambda_{(it)}$ is very small, then $\Delta \mathbf{p}_{\mathrm{LM}} \simeq \Delta \mathbf{p}_{\mathrm{GN}}$, which is a good step in the final stages of the iteration.

Therefore, in the same way that the damping parameter $\lambda_{(it)}$ can avoid the matrix singularity through increasing of the value of eigenvalues of $\mathbf{J}^{\mathrm{H}}\big(\mathbf{p}_{(it)}\big)\, \mathbf{J}\big(\mathbf{p}_{(it)}\big)$, it can control the system behavior through its influence on both direction and size of the step. A form of choosing the initial damping parameter $\lambda_{(0)}$ is from the maximum value of the diagonal of $\mathbf{J}^{\mathrm{H}}\big(\mathbf{p}_{(0)}\big)\, \mathbf{J}\big(\mathbf{p}_{(0)}\big)$, i.e.

$$\lambda_{(0)} = \tau \max_{l} \left[\mathbf{J}^{\mathrm{H}}\big(\mathbf{p}_{(0)}\big)\, \mathbf{J}\big(\mathbf{p}_{(0)}\big)\right]_{l,l}, \tag{4.107}$$

where τ is empirically chosen by the user, the choice of $\lambda_{(0)}$ can affect the convergence of algorithm.

The damping parameter must be adapted at each iteration according to an appropriate rule, because an inappropriate choice of this parameter may result in a divergence or a slowly convergence. A well-known method to control the factor $\lambda_{(it)}$ is by the following gain ratio [109]

$$\rho_{(it)} = \frac{\phi\big(\mathbf{p}_{(it)}\big) - \phi\big(\mathbf{p}_{(it)} + \Delta \mathbf{p}_{\mathrm{LM}}\big)}{\phi\big(\mathbf{p}_{(it)}\big) - \tilde{\phi}\big(\mathbf{p}_{(it)} + \Delta \mathbf{p}_{\mathrm{LM}}\big)}, \tag{4.108}$$

where the numerator denotes the effective variation of the cost function and the denominator denotes the predicted variation by the approximation of the cost function given in (4.99) from the linear model (4.97). Thus, $\lambda_{(it)}$ can be updated at each iteration it according to an usual procedure [109]:

$$\begin{cases} \text{(a) If } \rho_{(it)} \geq 0 : \mathbf{p}_{(it)} \text{ is accepted,} \\ \quad \lambda_{(it)} = \lambda_{(it-1)} \max\left(\frac{1}{3}, 1 - \big(2\rho_{(it)} - 1\big)^3\right) \text{ and } \nu = 2 \\ \text{(b) Otherwise: } \mathbf{p}_{(it)} \text{ is rejected,} \\ \quad \lambda_{(it)} = \nu\, \lambda_{(it-1)} \text{ and } \nu \leftarrow 2\,\nu \end{cases} . \tag{4.109}$$

Rewriting the denominator of (4.108) using (4.57), (4.99) and (4.104) gives

$$\begin{aligned}
\phi\big(\mathbf{p}_{(it)}\big) - \tilde{\phi}\big(\mathbf{p}_{(it)} + \Delta\mathbf{p}_{\mathrm{LM}}\big) &= -\frac{1}{2}\,\Delta\mathbf{p}_{\mathrm{LM}}^{\mathrm{H}}\Big(\mathbf{J}^{\mathrm{H}}\mathbf{r} + \frac{1}{2}\mathbf{J}^{\mathrm{H}}\mathbf{J}\,\Delta\mathbf{p}_{\mathrm{LM}}\Big) \\
&\quad - \frac{1}{2}\left(\mathbf{r}^{\mathrm{H}}\mathbf{J} + \frac{1}{2}\Delta\mathbf{p}_{\mathrm{LM}}^{\mathrm{H}}\,\mathbf{J}^{\mathrm{H}}\mathbf{J}\right)\Delta\mathbf{p}_{\mathrm{LM}} \\
&= -\frac{1}{2}\,\Delta\mathbf{p}_{\mathrm{LM}}^{\mathrm{H}}\Big(\mathbf{J}^{\mathrm{H}}\mathbf{r} + \frac{1}{2}\left(\mathbf{J}^{\mathrm{H}}\mathbf{J} + \lambda_{(it)}\mathbf{I}_Q - \lambda_{(it)}\mathbf{I}_Q\right)\Delta\mathbf{p}_{\mathrm{LM}}\Big) \\
&\quad - \frac{1}{2}\left(\mathbf{r}^{\mathrm{H}}\mathbf{J} + \frac{1}{2}\Delta\mathbf{p}_{\mathrm{LM}}^{\mathrm{H}}\left(\mathbf{J}^{\mathrm{H}}\mathbf{J} + \lambda_{(it)}\mathbf{I}_Q - \lambda_{(it)}\mathbf{I}_Q\right)\right)\Delta\mathbf{p}_{\mathrm{LM}} \\
&= -\frac{1}{2}\,\Delta\mathbf{p}_{\mathrm{LM}}^{\mathrm{H}}\left(\frac{1}{2}\mathbf{g} - \frac{1}{2}\lambda_{(it)}\Delta\mathbf{p}_{\mathrm{LM}}\right) \\
&\quad - \frac{1}{2}\left(\frac{1}{2}\mathbf{g}^{\mathrm{H}} - \frac{1}{2}\lambda_{(it)}\Delta\mathbf{p}_{\mathrm{LM}}^{\mathrm{H}}\right)\Delta\mathbf{p}_{\mathrm{LM}} \\
&= \frac{1}{2}\,\lambda_{(it)}\Delta\mathbf{p}_{\mathrm{LM}}^{\mathrm{H}}\Delta\mathbf{p}_{\mathrm{LM}} - \frac{1}{4}\left(\Delta\mathbf{p}_{\mathrm{LM}}^{\mathrm{H}}\mathbf{g} + \mathbf{g}^{\mathrm{H}}\Delta\mathbf{p}_{\mathrm{LM}}\right), \qquad (4.110)
\end{aligned}$$

for simplicity, we assumed that $\mathbf{J} \triangleq \mathbf{J}\big(\mathbf{p}_{(it)}\big)$, $\mathbf{r} \triangleq \mathbf{r}\big(\mathbf{p}_{(it)}\big)$ and $\mathbf{g} \triangleq \mathbf{g}\big(\mathbf{p}_{(it)}\big) = \mathbf{J}^{\mathrm{H}}\mathbf{r}$. Note that $-\Delta\mathbf{p}_{\mathrm{LM}}^{\mathrm{H}}\mathbf{g} = \Delta\mathbf{p}_{\mathrm{LM}}^{\mathrm{H}}\left(\mathbf{J}^{\mathrm{H}}\mathbf{J} + \lambda_{(it)}\mathbf{I}_Q\right)\Delta\mathbf{p}_{\mathrm{LM}} > 0$ and $-\mathbf{g}^{\mathrm{H}}\Delta\mathbf{p}_{\mathrm{LM}} > 0$, which imply $\phi\big(\mathbf{p}_{(it)}\big) - \tilde{\phi}\big(\mathbf{p}_{(it)} + \Delta\mathbf{p}_{\mathrm{LM}}\big)$ always positive, independent of $\lambda_{(it)}$.

A large value of $\rho_{(it)}$ indicates that $\tilde{\phi}\big(\mathbf{p}_{(it)} + \Delta\mathbf{p}_{\mathrm{LM}}\big)$ is close to $\phi\big(\mathbf{p}_{(it)}\big)$ and also that $\mathbf{p}_{(it)}$ is close to a local minimum point. Hence, it is possible to reduce $\lambda_{(it)}$, which means that the LM method gets closer to the GN method. On the other hand, a small value of $\rho_{(it)}$ (even negative) indicates that $\tilde{\phi}\big(\mathbf{p}_{(it)} + \Delta\mathbf{p}_{\mathrm{LM}}\big)$ is not close to $\phi\big(\mathbf{p}_{(it)}\big)$ and also that $\mathbf{p}_{(it)}$ is far from the solution. It is thus necessary to increase $\lambda_{(it)}$, or equivalently to get closer to the SD method.

The LM algorithm is summarized in Table 4.3 for both the TST and STF systems. Observe that the signal ($\mathbf{S}$) and channel ($\mathbf{H}$ or $\mathcal{H}$) are simultaneously estimated by $\mathbf{p}_{(it)} = \mathbf{p}_{(it-1)} + \Delta\mathbf{p}_{\mathrm{LM}}$ and using (4.62). Both vectors $\tilde{\mathbf{x}}_{NPJK}$ and $\tilde{\mathbf{x}}_{NPFK}$ represent vectorizations of the noisy received signal tensors $\tilde{\mathcal{X}} \in \mathbb{C}^{K\times N\times P\times J}$ and $\tilde{\mathcal{X}} \in \mathbb{C}^{F\times P\times N\times K}$, respectively.

In order to estimate separately the symbol matrix and the channel matrix (or channel tensor for the STF system), and also to simplify the complexity of LM algorithm, we deduce the Alternating LM (ALM) algorithm which considers an approximation of the Hessian matrix of $\tilde{\phi}\left(\mathbf{p}_{(it)} + \Delta\mathbf{p}\right)$ at each iteration it (4.102), i.e. $\mathbf{J}^{\mathrm{H}}\big(\mathbf{p}_{(it)}\big)\,\mathbf{J}\big(\mathbf{p}_{(it)}\big)$.

From the overall Jacobian matrix and gradient vector defined in (4.72) and (4.75) respectively, we can write the Hessian matrix corresponding to parameter estimates at the it-th iteration as

$$\mathbf{J}^{\mathrm{H}}\big(\mathbf{p}_{(it)}\big)\,\mathbf{J}\big(\mathbf{p}_{(it)}\big) = \begin{bmatrix} \mathbf{J}_{\mathbf{S}}^{\mathrm{H}}\Big(\mathbf{p}_{\mathbf{H}_{(it)}}\Big)\,\mathbf{J}_{\mathbf{S}}\Big(\mathbf{p}_{\mathbf{H}_{(it)}}\Big) & \mathbf{J}_{\mathbf{S}}^{\mathrm{H}}\Big(\mathbf{p}_{\mathbf{H}_{(it)}}\Big)\,\mathbf{J}_{\mathbf{H}}\Big(\mathbf{p}_{\mathbf{S}_{(it)}}\Big) \\ \mathbf{J}_{\mathbf{H}}^{\mathrm{H}}\Big(\mathbf{p}_{\mathbf{S}_{(it)}}\Big)\,\mathbf{J}_{\mathbf{S}}\Big(\mathbf{p}_{\mathbf{H}_{(it)}}\Big) & \mathbf{J}_{\mathbf{H}}^{\mathrm{H}}\Big(\mathbf{p}_{\mathbf{S}_{(it)}}\Big)\,\mathbf{J}_{\mathbf{H}}\Big(\mathbf{p}_{\mathbf{S}_{(it)}}\Big) \end{bmatrix}. \qquad (4.111)$$

Let us consider an approximation of the partitioned structure of the Jacobian matrix given by

$$\mathbf{J}^{\mathrm{H}}\big(\mathbf{p}_{(it)}\big)\,\mathbf{J}\big(\mathbf{p}_{(it)}\big) \simeq \begin{bmatrix} \mathbf{J}_{\mathbf{S}}^{\mathrm{H}}\Big(\mathbf{p}_{\mathbf{H}_{(it)}}\Big)\,\mathbf{J}_{\mathbf{S}}\Big(\mathbf{p}_{\mathbf{H}_{(it)}}\Big) & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{\mathbf{H}}^{\mathrm{H}}\Big(\mathbf{p}_{\mathbf{S}_{(it)}}\Big)\,\mathbf{J}_{\mathbf{H}}\Big(\mathbf{p}_{\mathbf{S}_{(it)}}\Big) \end{bmatrix}. \qquad (4.112)$$

Table 4.3: LM algorithm for semi-blind joint symbol and channel estimation.

TST system: LM-TST	**STF system: LM-STF**
1. Initialization ($it = 0$): randomly initialize $\hat{\mathbf{S}}_{(0)}$ and $\hat{\mathbf{H}}_{(0)}$ or $\hat{\mathcal{H}}_{(0)}$ with $\nu = 2$, $\mathbf{p}_{(0)} = \begin{bmatrix} \mathbf{p}_{\mathbf{S}_{(0)}} \\ \mathbf{p}_{\mathbf{H}_{(0)}} \end{bmatrix}$ and $\lambda_{(0)} = \max_l \left[\mathbf{J}^{\mathrm{H}}\big(\mathbf{p}_{(0)}\big)\, \mathbf{J}\big(\mathbf{p}_{(0)}\big) \right]_{l,l}$.	
2. $it = it + 1$.	
3. Compute $\Delta\mathbf{p}_{\mathrm{LM}_{(it)}}$ using (4.105).	
4. Compute $\hat{\mathbf{S}}_{(it)}$ and $\hat{\mathbf{H}}_{(it)}$ or $\hat{\mathcal{H}}_{(it)}$ from $\mathbf{p}_{(it)} = \mathbf{p}_{(it-1)} + \Delta\mathbf{p}_{\mathrm{LM}_{(it)}}$ where $\Delta\mathbf{p}_{\mathrm{LM}_{(it)}} = \begin{bmatrix} \Delta\mathbf{p}_{\mathrm{LM},\mathbf{S}_{(it)}} \\ \Delta\mathbf{p}_{\mathrm{LM},\mathbf{H}_{(it)}} \end{bmatrix}$ and $\mathbf{p}_{(it)} = \begin{bmatrix} \mathbf{p}_{\mathbf{S}_{(it)}} \\ \mathbf{p}_{\mathbf{H}_{(it)}} \end{bmatrix}$.	
5. Compute $\mathbf{J}\big(\mathbf{p}_{(it)}\big) = \left[\mathbf{J}_{\mathbf{S}}\Big(\mathbf{p}_{\mathbf{H}_{(it)}}\Big) \;\; \mathbf{J}_{\mathbf{H}}\Big(\mathbf{p}_{\mathbf{S}_{(it)}}\Big) \right]$ and $\mathbf{g}\big(\mathbf{p}_{(it)}\big) = \begin{bmatrix} \mathbf{g}_{\mathbf{S}}\big(\mathbf{p}_{(it)}\big) \\ \mathbf{g}_{\mathbf{H}}\big(\mathbf{p}_{(it)}\big) \end{bmatrix}$ using:	
(4.70), (4.71), (4.73), (4.74)	(4.92), (4.93), (4.94), (4.95)
6. Compute $\phi\big(\mathbf{p}_{(it)}\big)$ using:	
(4.68) or (4.69)	(4.90) or (4.91)
7. Compute $\rho_{(it)}$ using (4.108) and (4.110).	
8. Update $\lambda_{(it)}$ according to (4.109).	
9. Repeat steps (2) to (8) until convergence.	
10. Eliminate the scaling ambiguity with $\alpha = s_{1,1}/\hat{s}_{1,1}$:	
$\hat{\mathbf{S}} = \alpha\, \hat{\mathbf{S}}$,	
$\hat{\mathbf{H}} = 1/\alpha\, \hat{\mathbf{H}}$	$\hat{\mathbf{H}}_{K\times FM} = 1/\alpha\, \hat{\mathbf{H}}_{K\times FM}$.
11. Project the estimated symbols onto the symbol alphabet.	

Note that it is equivalent to consider $\mathbf{J}_{\mathbf{S}}^{\mathrm{H}}\Big(\mathbf{p}_{\mathbf{H}_{(it)}}\Big)\, \mathbf{J}_{\mathbf{H}}\Big(\mathbf{p}_{\mathbf{S}_{(it)}}\Big) = \mathbf{0}_{NR\times KM}$ for the TST system and $\mathbf{J}_{\mathbf{S}}^{\mathrm{H}}\Big(\mathbf{p}_{\mathbf{H}_{(it)}}\Big)\, \mathbf{J}_{\mathbf{H}}\Big(\mathbf{p}_{\mathbf{S}_{(it)}}\Big) = \mathbf{0}_{NR\times KFM}$ for the STF system.

Applying (4.112) to (4.104) gives

$$\left(\mathbf{J}_{\mathbf{S}}^{\mathrm{H}}\Big(\mathbf{p}_{\mathbf{H}_{(it)}}\Big)\, \mathbf{J}_{\mathbf{S}}\Big(\mathbf{p}_{\mathbf{H}_{(it)}}\Big) + \lambda_{\mathbf{S}_{(it)}}\, \mathbf{I}\right) \Delta\mathbf{p}_{\mathrm{LM},\mathbf{S}} = -\mathbf{g}_{\mathbf{S}}\big(\mathbf{p}_{(it)}\big) \tag{4.113}$$

$$\Longrightarrow \quad \Delta\mathbf{p}_{\mathrm{LM},\mathbf{S}} = -\left(\mathbf{J}_{\mathbf{S}}^{\mathrm{H}}\Big(\mathbf{p}_{\mathbf{H}_{(it)}}\Big)\, \mathbf{J}_{\mathbf{S}}\Big(\mathbf{p}_{\mathbf{H}_{(it)}}\Big) + \lambda_{\mathbf{S}_{(it)}}\, \mathbf{I}\right)^{-1} \mathbf{g}_{\mathbf{S}}\big(\mathbf{p}_{(it)}\big)\,, \tag{4.114}$$

$$\left(\mathbf{J}_{\mathbf{H}}^{\mathrm{H}}\Big(\mathbf{p}_{\mathbf{S}_{(it)}}\Big)\, \mathbf{J}_{\mathbf{H}}\Big(\mathbf{p}_{\mathbf{S}_{(it)}}\Big) + \lambda_{\mathbf{H}_{(it)}}\, \mathbf{I}\right) \Delta\mathbf{p}_{\mathrm{LM},\mathbf{H}} = -\mathbf{g}_{\mathbf{H}}\big(\mathbf{p}_{(it)}\big) \tag{4.115}$$

$$\Longrightarrow \quad \Delta\mathbf{p}_{\mathrm{LM},\mathbf{H}} = -\left(\mathbf{J}_{\mathbf{H}}^{\mathrm{H}}\Big(\mathbf{p}_{\mathbf{S}_{(it)}}\Big)\, \mathbf{J}_{\mathbf{H}}\Big(\mathbf{p}_{\mathbf{S}_{(it)}}\Big) + \lambda_{\mathbf{H}_{(it)}}\, \mathbf{I}\right)^{-1} \mathbf{g}_{\mathbf{H}}\big(\mathbf{p}_{(it)}\big)\,. \tag{4.116}$$

From (4.114) and (4.116), we can separately estimate $\Delta\mathbf{p}_{\mathrm{LM},\mathbf{S}}$ and $\Delta\mathbf{p}_{\mathrm{LM},\mathbf{H}}$, and also we can employ different damping parameters $\lambda_{(it)}$ for each estimate parameter $\mathbf{S}$ and $\mathbf{H}$ (or $\mathcal{H}$). In this sense, it is possible to control independently each factor $\lambda_{\mathbf{S},(it)}$ and $\lambda_{\mathbf{H},(it)}$, and this flexibility may allow to reach faster convergence speed.

In order to derive a gain ratio $\rho_{(it)}$ for each parameter, we rewrite (4.110) as a function of

$\Delta\mathbf{p}_{\mathrm{LM},\mathbf{S}}$, $\Delta\mathbf{p}_{\mathrm{LM},\mathbf{H}}$, $\mathbf{g}_{\mathbf{S}}$ and $\mathbf{g}_{\mathbf{H}}$ using

$$\Delta\mathbf{p}_{\mathrm{LM}} \triangleq \begin{bmatrix} \Delta\mathbf{p}_{\mathrm{LM},\mathbf{S}} \\ \Delta\mathbf{p}_{\mathrm{LM},\mathbf{H}} \end{bmatrix}, \tag{4.117}$$

$$\mathbf{g} \triangleq \begin{bmatrix} \mathbf{g}_{\mathbf{S}} \\ \mathbf{g}_{\mathbf{H}} \end{bmatrix}. \tag{4.118}$$

From (4.110), we obtain

$$\begin{aligned} \phi\big(\mathbf{p}_{(it)}\big) - \tilde{\phi}\big(\mathbf{p}_{(it)} + \Delta\mathbf{p}_{\mathrm{LM}}\big) &= \frac{1}{2}\,\lambda_{(it)}\Delta\mathbf{p}_{\mathrm{LM}}^{\mathrm{H}}\Delta\mathbf{p}_{\mathrm{LM}} - \frac{1}{4}\left(\Delta\mathbf{p}_{\mathrm{LM}}^{\mathrm{H}}\mathbf{g} + \mathbf{g}^{\mathrm{H}}\Delta\mathbf{p}_{\mathrm{LM}}\right) \\ &= \frac{1}{2}\,\lambda_{(it)}\left(\Delta\mathbf{p}_{\mathrm{LM},\mathbf{S}}^{\mathrm{H}}\Delta\mathbf{p}_{\mathrm{LM},\mathbf{S}} + \Delta\mathbf{p}_{\mathrm{LM},\mathbf{H}}^{\mathrm{H}}\Delta\mathbf{p}_{\mathrm{LM},\mathbf{H}}\right) - \\ &\quad \frac{1}{4}\left(\Delta\mathbf{p}_{\mathrm{LM},\mathbf{S}}^{\mathrm{H}}\mathbf{g}_{\mathbf{S}} + \mathbf{g}_{\mathbf{S}}^{\mathrm{H}}\Delta\mathbf{p}_{\mathrm{LM},\mathbf{S}} + \Delta\mathbf{p}_{\mathrm{LM},\mathbf{H}}^{\mathrm{H}}\mathbf{g}_{\mathbf{H}} + \mathbf{g}_{\mathbf{H}}^{\mathrm{H}}\Delta\mathbf{p}_{\mathrm{LM},\mathbf{H}}\right). \end{aligned} \tag{4.119}$$

Considering the influence of each parameter separately and employing different damping parameters, we have

$$\rho_{\mathbf{S}_{(it)}} = \frac{\phi\big(\mathbf{p}_{(it)}\big) - \phi\big(\mathbf{p}_{(it)} + \Delta\mathbf{p}_{\mathrm{LM}}\big)}{\frac{1}{2}\,\lambda_{\mathbf{S}_{(it)}}\left(\Delta\mathbf{p}_{\mathrm{LM},\mathbf{S}}^{\mathrm{H}}\Delta\mathbf{p}_{\mathrm{LM},\mathbf{S}}\right) - \frac{1}{4}\left(\Delta\mathbf{p}_{\mathrm{LM},\mathbf{S}}^{\mathrm{H}}\mathbf{g}_{\mathbf{S}} + \mathbf{g}_{\mathbf{S}}^{\mathrm{H}}\Delta\mathbf{p}_{\mathrm{LM},\mathbf{S}}\right)} \tag{4.120}$$

and

$$\rho_{\mathbf{H}_{(it)}} = \frac{\phi\big(\mathbf{p}_{(it)}\big) - \phi\big(\mathbf{p}_{(it)} + \Delta\mathbf{p}_{\mathrm{LM}}\big)}{\frac{1}{2}\,\lambda_{\mathbf{H}_{(it)}}\left(\Delta\mathbf{p}_{\mathrm{LM},\mathbf{H}}^{\mathrm{H}}\Delta\mathbf{p}_{\mathrm{LM},\mathbf{H}}\right) - \frac{1}{4}\left(\Delta\mathbf{p}_{\mathrm{LM},\mathbf{H}}^{\mathrm{H}}\mathbf{g}_{\mathbf{H}} + \mathbf{g}_{\mathbf{H}}^{\mathrm{H}}\Delta\mathbf{p}_{\mathrm{LM},\mathbf{H}}\right)}. \tag{4.121}$$

Let us now describe the proposed ALM algorithm in Table 4.4 for the TST and STF systems employing the expressions deduced above. Observe that the initial damping parameter for each estimate parameter ($\lambda_{\mathbf{S}_{(0)}}$ and $\lambda_{\mathbf{H}_{(0)}}$) is calculated from (4.107) taking into account the Jacobian matrix with respect to $\mathbf{p}_{\mathbf{S}_{(0)}}$ and $\mathbf{p}_{\mathbf{H}_{(0)}}$ respectively.

4.3 Kronecker based Least Squares (KLS)

Let us introduce a direct (non-iterative) procedure, denoted by the Kronecker Least Squares, which is based on the structure of the Kronecker product between two matrices. From an appropriate unfolded matrix and using the Kronecker structure, it is possible to jointly estimate both symbol and channel matrices.

We derive the semi-blind KLS approach for the TST and STF systems from the unfolded matrix $\mathbf{X}_{KN\times JP}$ and $\mathbf{X}_{KN\times FP}$ given by (3.16) and (3.83), and (3.38) respectively. Observe that this approach can be extend to other systems depending on a convenient choice of the unfolded matrix.

4.3.1 TST system

Let us rewrite (3.16) assuming that $\mathbf{G}_{RM\times JP}$ is full row-rank as

$$\mathbf{X}_{NK\times JP}\,\mathbf{G}_{RM\times JP}^{\dagger} = \mathbf{S}\otimes\mathbf{H} \triangleq \mathbf{Y} \in \mathbb{C}^{NK\times RM}, \tag{4.122}$$

Table 4.4: ALM algorithm for semi-blind joint symbol and channel estimation.

TST system: ALM-TST	**STF system: ALM-STF**
1. Initialization ($it = 0$): randomly initialize $\hat{\mathbf{S}}_{(0)}$ and $\hat{\mathbf{H}}_{(0)}$ or $\hat{\mathcal{H}}_{(0)}$, $\nu = 2$, $\lambda_{\mathbf{S}_{(0)}} = \max_l \left[\mathbf{J}_{\mathbf{S}}^{\mathrm{H}}(\mathbf{p}_{\mathbf{H}_{(0)}})\,\mathbf{J}_{\mathbf{S}}(\mathbf{p}_{\mathbf{H}_{(0)}})\right]_{l,l}$, $\lambda_{\mathbf{H}_{(0)}} = \max_l \left[\mathbf{J}_{\mathbf{H}}^{\mathrm{H}}(\mathbf{p}_{\mathbf{S}_{(0)}})\,\mathbf{J}_{\mathbf{H}}(\mathbf{p}_{\mathbf{S}_{(0)}})\right]_{l,l}$.	
2. $it = it + 1$.	
3. Compute $\mathbf{J}_{\mathbf{S}}\left(\mathbf{p}_{\mathbf{H}_{(it-1)}}\right)$ and $\mathbf{g}_{\mathbf{S}}\left(\mathbf{p}_{(it-1)}\right)$ using:	
(4.70), (4.73)	(4.92), (4.94)
4. Compute $\Delta\mathbf{p}_{\mathrm{LM},\mathbf{S}_{(it)}}$ using (4.114).	
5. Compute $\hat{\mathbf{S}}_{(it)}$ from $\mathbf{p}_{\mathbf{S}_{(it)}} = \mathbf{p}_{\mathbf{S}_{(it-1)}} + \Delta\mathbf{p}_{\mathrm{LM},\mathbf{S}_{(it)}}$.	
6. Update $\mathbf{p}_{(it)} = \begin{bmatrix} \mathbf{p}_{\mathbf{S}_{(it)}} \\ \mathbf{p}_{\mathbf{H}_{(it-1)}} \end{bmatrix}$.	
7. Compute $\phi\left(\mathbf{p}_{(it)}\right)$ using:	
(4.68)	(4.90)
8. Compute $\rho_{\mathbf{S}_{(it)}}$ using (4.120).	
9. Update $\lambda_{\mathbf{S}_{(it)}}$ according to (4.109).	
10. Compute $\mathbf{J}_{\mathbf{H}}\left(\mathbf{p}_{\mathbf{S}_{(it)}}\right)$ and $\mathbf{g}_{\mathbf{H}}\left(\mathbf{p}_{(it)}\right)$ using:	
(4.71), (4.74)	(4.93), (4.95)
11. Compute $\Delta\mathbf{p}_{\mathrm{LM},\mathbf{H}_{(it)}}$ using (4.116).	
12. Compute $\hat{\mathbf{H}}_{(it)}$ from $\mathbf{p}_{\mathbf{H}_{(it)}} = \mathbf{p}_{\mathbf{H}_{(it-1)}} + \Delta\mathbf{p}_{\mathrm{LM},\mathbf{H}_{(it)}}$.	
13. Update $\mathbf{p}_{(it)} = \begin{bmatrix} \mathbf{p}_{\mathbf{S}_{(it)}} \\ \mathbf{p}_{\mathbf{H}_{(it)}} \end{bmatrix}$.	
14. Compute $\phi\left(\mathbf{p}_{(it)}\right)$ using:	
(4.69)	(4.91)
15. Compute $\rho_{\mathbf{H}_{(it)}}$ using (4.121).	
16. Update $\lambda_{\mathbf{H}_{(it)}}$ according to (4.109).	
17. Repeat steps (2) to (16) until convergence.	
18. Eliminate the scaling ambiguity with $\alpha = s_{1,1}/\hat{s}_{1,1}$:	
$\hat{\mathbf{S}} = \alpha\,\hat{\mathbf{S}}$,	
$\hat{\mathbf{H}} = 1/\alpha\,\hat{\mathbf{H}}$	$\hat{\mathbf{H}}_{K\times FM} = 1/\alpha\,\hat{\mathbf{H}}_{K\times FM}$
19. Project the estimated symbols onto the symbol alphabet.	

with

$$\mathbf{G}_{RM\times JP} = \left(\mathbf{W}_{J\times RM} \diamond \left(\mathbf{C}^{(\mathbf{S})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{H})^{\mathrm{T}}}\right)^{\mathrm{T}}\right)^{\mathrm{T}}, \tag{4.123}$$

$$\mathbf{W}_{J\times RM} \triangleq \left[\mathrm{vec}(\mathbf{W}_{\cdot\cdot 1}) \;\; \cdots \;\; \mathrm{vec}(\mathbf{W}_{\cdot\cdot J})\right]^{\mathrm{T}}. \tag{4.124}$$

Note that $\mathbf{G}_{RM\times JP}$ and $\mathbf{W}_{J\times RM}$ denote a matrix unfolding of $\mathcal{G} \in \mathbb{C}^{M\times R\times P\times J}$ and of $\mathcal{W} \in \mathbb{C}^{M\times R\times J}$, which is slightly modified from (3.18) and (3.19) by convenience. Note that it is possible to simultaneously estimate $\mathbf{S}$ and $\mathbf{H}$ from the Kronecker product between both matrices using (4.122).

We can use some information a priori about symbol matrix $\mathbf{S}$ and the special structure of the Kronecker product to separate the mixture given by channel coefficients and transmitted symbols, which leads to semi-blind estimation of both matrices.

Suppose that the (n, r)-th element of the symbol matrix is known at the receiver and given by $s_{n,r} = 1$ for simplicity. From the structure of the Kronecker product, we can directly estimate $\mathbf{H}$. In the same way, assuming that the n-th row of the symbol matrix is given by $\mathbf{S}_{n\cdot} = \mathbf{1}_R^{\mathrm{T}}$, we obtain R estimations of $\mathbf{H}$ and from the mean square of all estimations, we can better estimate $\mathbf{H}$.

From the structure of the Kronecker product

$$\mathbf{Y} \triangleq \begin{bmatrix} \mathbf{Y}^{(1,1)} & \cdots & \mathbf{Y}^{(1,R)} \\ \vdots & \ddots & \vdots \\ \mathbf{Y}^{(N,1)} & \cdots & \mathbf{Y}^{(N,R)} \end{bmatrix} \in \mathbb{C}^{NK \times RM}, \tag{4.125}$$

we can write $\mathbf{Y}^{(n,r)} \triangleq (\mathbf{S} \otimes \mathbf{H})^{(n,r)} = s_{n,r}\,\mathbf{H}$ by fixing n and r, for all $n \in \{1, \cdots, N\}$ and $r \in \{1, \cdots, R\}$. Applying the least squares method and considering the channel estimate and the received signal tensor, we can write

$$\hat{s}_{n,r} = \frac{\mathrm{vec}(\mathbf{H})^{\mathrm{H}}\,\mathrm{vec}\big(\mathbf{Y}^{(n,r)}\big)}{\|\mathrm{vec}(\mathbf{H})\|^2}, \tag{4.126}$$

where $\mathbf{Y}^{(n,r)} \in \mathbb{C}^{K \times M}$ is constructed from $\mathbf{Y} = \mathbf{X}_{NK \times JP}\,\mathbf{G}^{\dagger}_{RM \times JP} \in \mathbb{C}^{NK \times RM}$. The expression (4.126) allows to directly compute an estimate of each symbol $s_{n,r}$. Table 4.5 illustrates the KLS algorithm based on this idea by assuming one symbol known at the receiver.

Analysis of identifiability conditions:

The estimation of the Kronecker product between symbol and channel matrices by (4.122) requires that the unfolded matrix $\mathbf{G}_{RM \times JP}$ is full row-rank, implying $RM \leq JP$. From (4.123) and Lemma 2.2, it is enough to satisfy the following inequality

$$k(\mathbf{W}_{J \times RM}) + k\left(\left(\mathbf{C}^{(\mathbf{S})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{H})^{\mathrm{T}}}\right)^{\mathrm{T}}\right) \geq RM + 1. \tag{4.127}$$

Remark that the condition (4.127) is also implicit in Theorem 3.4 concerning the uniqueness condition. Contrarily to the ALS algorithm, satisfying the identifiability condition (4.127) leads to ensure the uniqueness condition as well. Another advantage is that the symbol and channel matrices can be estimated by just one iteration.

Generalization to multiuser case

Assuming that $\mathbf{G}_{QRM \times JP}$ is full row-rank, we can write the following matrix by transposing the unfolded matrix of the received signal tensor given in (3.83) as

$$\mathbf{X}_{NK \times JP}\,\mathbf{G}^{\dagger}_{QRM \times JP} = \mathbf{S}_{N \times QR}\,|\otimes|\,\mathbf{H}_{K \times QM} \triangleq \mathbf{Y} \in \mathbb{C}^{NK \times QRM}, \tag{4.128}$$

with

$$\mathbf{G}_{QRM\times JP} = \left(\mathbf{W}_{J\times QRM} \diamond \left[\left(\mathbf{C}^{(\mathbf{S},1)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{H},1)^{\mathrm{T}}}\right)^{\mathrm{T}} \ \cdots \ \left(\mathbf{C}^{(\mathbf{S},Q)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{H},Q)^{\mathrm{T}}}\right)^{\mathrm{T}}\right]\right)^{\mathrm{T}}$$
$$= \left(\mathbf{W}_{J\times QRM} \diamond \left(\left(\mathbf{C}^{(\mathbf{S})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{H})^{\mathrm{T}}}\right)^{\mathrm{T}} \tilde{\mathbf{\Phi}}_{QRM}\right)\right)^{\mathrm{T}}, \tag{4.129}$$

$$\mathbf{W}_{J\times QRM} \triangleq \left[\mathbf{W}^{(1)}_{J\times RM} \ \cdots \ \mathbf{W}^{(Q)}_{J\times RM}\right]. \tag{4.130}$$

Observe that $\mathbf{G}_{QRM\times JP}$ and $\mathbf{W}_{J\times QRM}$ represent slightly modified version of (3.88) and (3.89), respectively.

Considering $\mathbf{Y} \triangleq \left[\mathbf{Y}^{(1)} \ \cdots \ \mathbf{Y}^{(Q)}\right]$ a partitioned matrix and applying the definition of a partition-wise Kronecker product (2.3), we can rewrite (4.128) as

$$\left[\mathbf{Y}^{(1)} \ \cdots \ \mathbf{Y}^{(Q)}\right] = \mathbf{S}_{N\times QR} \,|\otimes|\, \mathbf{H}_{K\times QM}$$
$$= \left[\mathbf{S}^{(1)} \otimes \mathbf{H}^{(1)} \ \cdots \ \mathbf{S}^{(Q)} \otimes \mathbf{H}^{(Q)}\right], \tag{4.131}$$

where $\mathbf{Y}^{(q)} \triangleq \mathbf{S}^{(q)} \otimes \mathbf{H}^{(q)} \in \mathbb{C}^{NK\times RM}$.

From the Kronecker product associated with each user q, i.e.

$$\mathbf{Y}^{(q)} \triangleq \begin{bmatrix} \mathbf{Y}^{(q,1,1)} & \cdots & \mathbf{Y}^{(q,1,R)} \\ \vdots & \ddots & \vdots \\ \mathbf{Y}^{(q,N,1)} & \cdots & \mathbf{Y}^{(q,N,R)} \end{bmatrix} = \mathbf{S}^{(q)} \otimes \mathbf{H}^{(q)} \in \mathbb{C}^{NK\times RM}, \tag{4.132}$$

we obtain $\mathbf{Y}^{(q,n,r)} \triangleq \left(\mathbf{S}^{(q)} \otimes \mathbf{H}^{(q)}\right)^{(n,r)} = s^{(q)}_{n,r}\,\mathbf{H}^{(q)}$ by fixing q, n and r. According to the least squares approach analogously to (4.126), the transmitted symbols associated with each user q, $s^{(q)}_{n,r}$, can be estimated by

$$\hat{s}^{(q)}_{n,r} = \frac{\mathrm{vec}\left(\mathbf{H}^{(q)}\right)^{\mathrm{H}} \mathrm{vec}\left(\mathbf{Y}^{(q,n,r)}\right)}{\left\|\mathrm{vec}(\mathbf{H}^{(q)})\right\|^2}, \tag{4.133}$$

taking into account the channel estimation for each user and with $\mathbf{Y}^{(q,n,r)} \in \mathbb{C}^{K\times M}$ constructed from $\mathbf{Y} \triangleq \left[\mathbf{Y}^{(1)} \ \cdots \ \mathbf{Y}^{(Q)}\right] = \mathbf{X}_{NK\times JP}\,\mathbf{G}^{\dagger}_{QRM\times JP} \in \mathbb{C}^{NK\times QRM}$.

Analysis of identifiability conditions:

The estimation of channel and symbol matrices from (4.128) requires that $\mathbf{G}_{QRM\times JP}$ is full row-rank, implying $QRM \leq JP$. Applying Lemma 2.2 into (4.129), the property of full rank can be ensured from the following inequality

$$k(\mathbf{W}_{J\times QRM}) + k\left(\left(\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}\right)^{\mathrm{T}} \tilde{\mathbf{\Phi}}_{QRM}\right) \geq QRM + 1. \tag{4.134}$$

Remark that the condition on the allocation and code structures given in Theorem 3.8, concerning the uniqueness result, leads to (4.134).

If $\mathbf{C}^{(\mathbf{S},1)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{H},1)^{\mathrm{T}}}$ has no zero-rows for all $q \in \{1, \cdots, Q\}$ and the code tensor for each user $\mathcal{W}^{(q)}$ is set such that $\mathbf{W}_{J\times QRM}$ is full column-rank, then the condition (4.134) is satisfied and $\mathbf{G}_{QRM\times JP}$ is full row-rank.

4.3.2 STF system

Analogously to the TST coding, if we assume the unfolded matrix $\mathbf{T}_{FMR\times FP}$ is full row-rank, we can rewrite (3.38) as

$$\mathbf{X}_{KN\times FP}\,\mathbf{T}^{\dagger}_{FMR\times FP} = \mathbf{H}_{K\times FM}\otimes\mathbf{S} \triangleq \mathbf{Y}\in\mathbb{C}^{KN\times FMR}. \tag{4.135}$$

Based on the same idea proposed previously, we can estimate the symbol matrix $\mathbf{S}$ and the channel tensor $\mathcal{H}$ from (4.135) thanks to the structure of the Kronecker product between both matrices $\mathbf{H}_{K\times FM}$ and $\mathbf{S}$, and a prior information at the receiver.

Defining $\mathbf{Y} \triangleq \left[\mathbf{Y}^{(1)}_{KN\times MR}\ \cdots\ \mathbf{Y}^{(F)}_{KN\times MR}\right]$ as a partitioned matrix, observe that we can rewrite (4.135) as

$$\begin{aligned}\left[\mathbf{Y}^{(1)}_{KN\times MR}\ \ \cdots\ \ \mathbf{Y}^{(F)}_{KN\times MR}\right] &= \mathbf{H}_{K\times FM}\otimes\mathbf{S}\\ &= \left[\mathbf{H}_{1..}\otimes\mathbf{S}\ \ \cdots\ \ \mathbf{H}_{F..}\otimes\mathbf{S}\right],\end{aligned} \tag{4.136}$$

with $\mathbf{Y}^{(f)}_{KN\times MR} \triangleq \mathbf{H}_{f..}\otimes\mathbf{S}\in\mathbb{C}^{KN\times MR}$.

From the Kronecker product associated with each subcarrier f, a permutation matrix of rows $\mathbf{\Pi}^{(1)}\in\mathbb{R}^{NK\times KN}$ and a permutation matrix of columns $\mathbf{\Pi}^{(2)}\in\mathbb{R}^{MR\times RM}$, we can write

$$\begin{aligned}\mathbf{Y}^{(f)}_{NK\times RM} \triangleq \begin{bmatrix}\mathbf{Y}^{(f,1,1)} & \cdots & \mathbf{Y}^{(f,1,R)}\\ \vdots & \ddots & \vdots\\ \mathbf{Y}^{(f,N,1)} & \cdots & \mathbf{Y}^{(f,N,R)}\end{bmatrix} &= \mathbf{\Pi}^{(1)}\,\mathbf{Y}^{(f)}_{KN\times MR}\,\mathbf{\Pi}^{(2)}\\ &= \mathbf{S}\otimes\mathbf{H}_{f..}\in\mathbb{C}^{NK\times RM}\end{aligned} \tag{4.137}$$

where $\mathbf{Y}^{(f,n,r)} \triangleq s_{n,r}\,\mathbf{H}_{f..}\in\mathbb{C}^{K\times M}$ and, $\mathbf{\Pi}^{(1)}\in\mathbb{R}^{NK\times KN}$ and $\mathbf{\Pi}^{(2)}\in\mathbb{R}^{MR\times RM}$ are defined as

$$\begin{aligned}\mathbf{\Pi}^{(1)} &\triangleq \sum_{n=1}^{N}\sum_{k=1}^{K}\mathbf{e}^{(N)}_{n}\mathbf{e}^{(K)\mathrm{T}}_{k}\otimes\mathbf{e}^{(K)}_{k}\mathbf{e}^{(N)\mathrm{T}}_{n}\\ \mathbf{\Pi}^{(2)} &\triangleq \sum_{m=1}^{M}\sum_{r=1}^{R}\mathbf{e}^{(M)}_{m}\mathbf{e}^{(R)\mathrm{T}}_{r}\otimes\mathbf{e}^{(R)}_{r}\mathbf{e}^{(M)\mathrm{T}}_{m}.\end{aligned} \tag{4.138}$$

Thus, we can estimate the channel matrix associated with each f, i.e. $\mathbf{H}_{f..}$ for $f=1,...,F$, by assuming the knowledge of at least one symbol. As consequence, we can better estimate the channel tensor considering the knowledge of more than one symbol, as for example one row of the symbol matrix. From the channel estimate and the received signal tensor, each transmitted symbol $s_{n,r}$ can be estimated by

$$\hat{s}_{n,r} = \frac{\mathrm{vec}(\mathbf{H}_{K\times FM})^{\mathrm{H}}\,\mathrm{vec}\big(\mathbf{Y}^{(n,r)}\big)}{\left\|\mathrm{vec}(\mathbf{H}_{K\times FM})\right\|^{2}}, \tag{4.139}$$

where $\mathbf{Y}^{(n,r)} \triangleq \left[\mathbf{Y}^{(1,n,r)}\ \cdots\ \mathbf{Y}^{(F,n,r)}\right]\in\mathbb{C}^{K\times FM}$ is obtained from $\mathbf{Y} \triangleq \mathbf{X}_{KN\times FP}\,\mathbf{T}^{\dagger}_{FMR\times FP}\in\mathbb{C}^{KN\times FMR}$. Table 4.5 presents the KLS algorithm for the STF system taking into account the knowledge of one symbol.

Table 4.5: KLS algorithm for semi-blind joint symbol and channel estimation.

TST system: KLS-TST	**STF system: KLS-STF**
1. Compute an LS estimate of $\mathbf{S} \otimes \mathbf{H}$ and $\mathbf{H}_{K\times FM} \otimes \mathbf{S}$ from:	
$\hat{\mathbf{Y}} = \tilde{\mathbf{X}}_{NK\times JP}\, \mathbf{G}^{\dagger}_{RM\times JP}$	$\hat{\mathbf{Y}} = \tilde{\mathbf{X}}_{KN\times FP}\, \mathbf{T}^{\dagger}_{FMR\times FP}$
2. Estimate $\mathbf{H}$ or $\mathcal{H}$ using the knowledge of one symbol $s_{1,1} = 1$:	
$\hat{\mathbf{H}} = \hat{\mathbf{Y}}^{(1,1)}$	$\hat{\mathbf{H}}_{f\cdot\cdot} = \hat{\mathbf{Y}}^{(f,1,1)}$ for $f = 1, ..., F$
3. Compute $\mathbf{S}$ from $\hat{\mathbf{H}}$ or $\hat{\mathbf{H}}_{K\times FM}$:	
$\hat{s}_{n,r} = \frac{\text{vec}\left(\hat{\mathbf{H}}\right)^{\text{H}} \text{vec}\left(\hat{\mathbf{Y}}^{(n,r)}\right)}{\left\|\text{vec}\left(\hat{\mathbf{H}}\right)\right\|^2}$	$\hat{s}_{n,r} = \frac{\text{vec}\left(\hat{\mathbf{H}}_{K\times FM}\right)^{\text{H}} \text{vec}\left(\hat{\mathbf{Y}}^{(n,r)}\right)}{\left\|\text{vec}\left(\hat{\mathbf{H}}_{K\times FM}\right)\right\|^2}$ with $\hat{\mathbf{H}}_{K\times FM} = \left[\hat{\mathbf{H}}_{1\cdot\cdot} \ \cdots \ \hat{\mathbf{H}}_{F\cdot\cdot}\right]$, $\hat{\mathbf{Y}}^{(n,r)} = \left[\hat{\mathbf{Y}}^{(1,n,r)} \ \ldots \ \hat{\mathbf{Y}}^{(F,n,r)}\right]$
4. Project the estimated symbols onto the symbol alphabet.	

Analysis of identifiability conditions:

Observe that the unique estimation of $\mathbf{H}_{K\times FM} \otimes \mathbf{S}$ by (4.135) requires that $\mathbf{T}_{FMR\times FP}$ is full row-rank, implying $MR \leq P$. Hence, it follows the same reasoning made in Theorem 3.7 for the uniqueness condition, leading to the same condition obtained for the code matrix and allocation tensors.

From (3.39) and (3.40), if the code matrix $\mathbf{W} \in \mathbb{C}^{M\times R}$ is composed just by nonzero elements and the allocation tensors are chosen such way that $\mathbf{C}^{(\mathcal{H})^{\text{T}}}_{f\cdot\cdot} \diamond \mathbf{C}^{(\mathbf{S})^{\text{T}}}_{f\cdot\cdot}$ is full row-rank for all $f \in \{1, ..., F\}$, then the unfolded matrix $\mathbf{T}_{FMR\times FP}$ is full row-rank.

4.4 Complexity analysis of algorithms

In Table 4.6, we compare all algorithms in terms of its computational complexity for the TST and STF systems, taking into account more onerous operations at each iteration such as matrix inversions and complex multiplications.

Table 4.6: Computational complexity per iteration.

	Computational complexity	
Algorithms	**TST system**	**STF system**
ALS	$\mathcal{O}((JP)^2) + \mathcal{O}(R^3) + \mathcal{O}(M^3)$	$\mathcal{O}((FP)^2) + \mathcal{O}(R^3) + \mathcal{O}((FM)^3)$
LM	$\mathcal{O}((JP)^2) + \mathcal{O}((NR + KM)^3)$	$\mathcal{O}((FP)^2) + \mathcal{O}((NR + FKM)^3)$
ALM	$\mathcal{O}((JP)^2) + \mathcal{O}((NR)^3) + \mathcal{O}((KM)^3)$	$\mathcal{O}((FP)^2) + \mathcal{O}((NR)^3) + \mathcal{O}((FKM)^3)$
KLS[‡]	$\mathcal{O}((RM)^2) + \mathcal{O}((RM)^3)$	$\mathcal{O}((FRM)^2) + \mathcal{O}((FRM)^3)$

[‡]Remember that the KLS algorithm is a non-iterative procedure.

For the ALS algorithm, we compute two pseudo-inverses for each estimation of symbols and channels at each iteration, which are implicit matrix inversions and complex multiplications corresponding approximately to $\mathcal{O}((JP)^2) + \mathcal{O}(R^3) + \mathcal{O}(M^3)$ and $\mathcal{O}((FP)^2) + \mathcal{O}(R^3)$

+ $\mathcal{O}((FM)^3)$ for the TST and STF systems, respectively. The LM algorithm involves a matrix inversion to estimate each new parameter variation corresponding approximately to $\mathcal{O}((NR+KM)^3)$ and $\mathcal{O}((NR+FKM)^3)$ for the TST and STF systems, respectively. As symbols and channels are alternately estimated in the ALM algorithm, there are two matrix inversions which correspond respectively to $\mathcal{O}((NR)^3) + \mathcal{O}((KM)^3)$ and $\mathcal{O}((NR)^3) + \mathcal{O}((FKM)^3)$ for the TST and STF systems. The LM and ALM algorithms approximately involve $\mathcal{O}((JP)^2)$ and $\mathcal{O}((FP)^2)$ complex multiplications to compute the Jacobian matrix for the TST and STF systems, respectively.

The KLS algorithm involves only one pseudo-inverse, which are implicit matrix inversion and complex multiplications corresponding approximately to $\mathcal{O}((RM)^2) + \mathcal{O}((RM)^3)$ and $\mathcal{O}((FRM)^2) + \mathcal{O}((FRM)^3)$ for the TST and STF systems, respectively. Differently from the ALS, LM and ALM algorithms, the KLS algorithm is a non-iterative method and consequently, involves one iteration. In general, the KLS is the least complex. According to Table 4.6, the ALS is the least complex iterative algorithm and LM algorithm is the most complex. The simplification done to achieve the ALM algorithm allows us to reduce the complexity of the LM algorithm as desired. In order to illustrate the values in Table 4.6, we give some examples of design parameters for four values of P, K, $J = F$, and R in Figures 4.1, 4.2, 4.3 and 4.4, respectively.

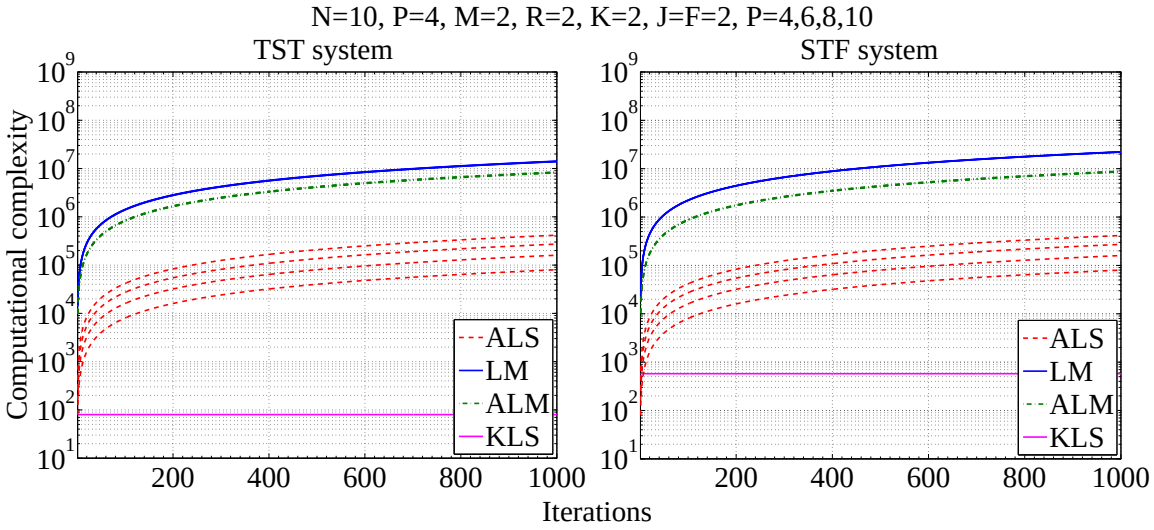


Figure 4.1: Computational complexity: Influence of the number P of data blocks.

The difference in the complexity between the LM and ALM are not significant when compared to the ALS algorithm for both systems. The number of data blocks P has more influence on the variation of complexity of the ALS algorithm than the LM and ALM. In the same way that the variation of the number of receive antennas K affects more the complexity of the LM and ALM algorithms than the ALS for both systems.

The variations of J do not significantly change the complexity of the LM and ALM for the TST system as happens with variations of F for the STF system. From Figure 4.4, we can observe that the increase of the number of data streams directly leads to an increase in the

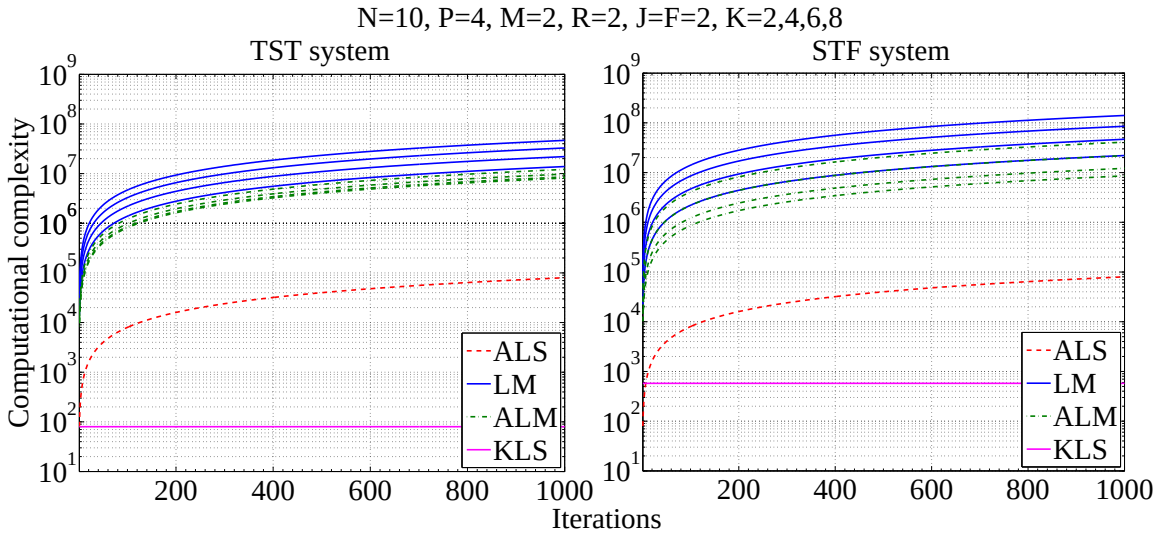


Figure 4.2: Computational complexity: Influence of the number K of receive antennas.

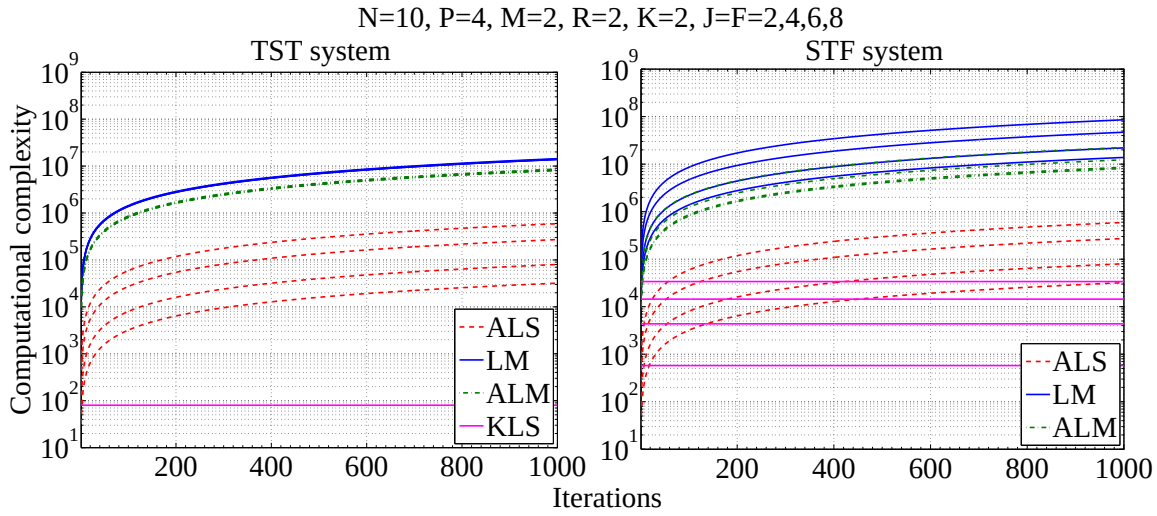


Figure 4.3: Computational complexity: Influence of the spreading length J and of the number F of subcarriers.

complexity of all algorithms. From the variations of P, K, J, R, observe that the KLS algorithm is always the least complex for the TST system. However, for the STF system, the increase of the number of data streams and of subcarriers may lead to greater complexity for the KLS than the ALS algorithm, depending on the number of iterations required for the ALS convergence.

4.5 General discussion

Table 4.7 presents a resume of the identifiability conditions of the ALS algorithm for both systems developed in Section 4.1. By construction, the LM and ALM algorithms do not require to satisfy identifiability conditions and the KLS method requires that two matrix unfolding re-

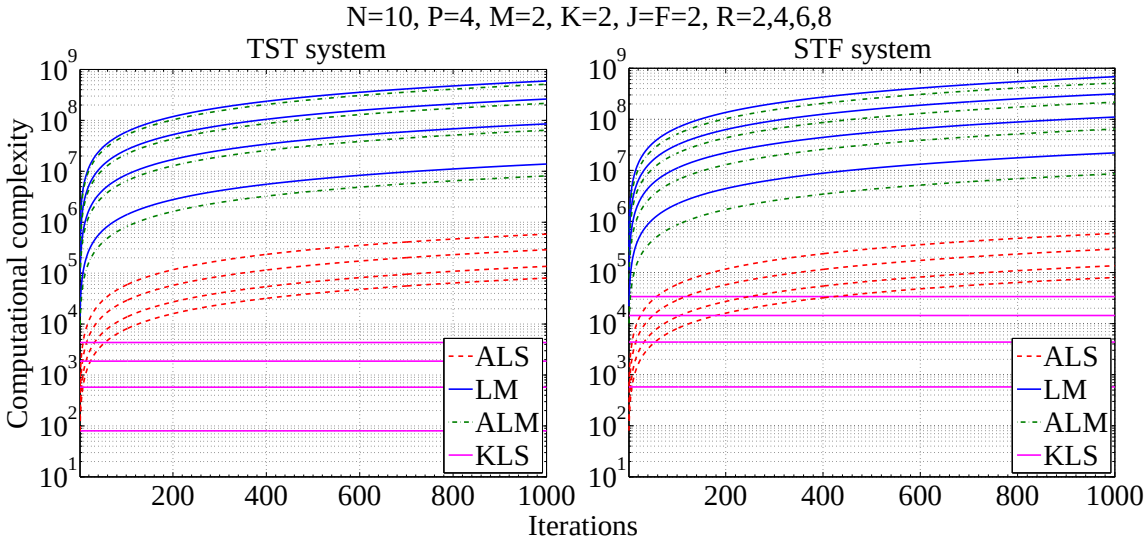


Figure 4.4: Computational complexity: Influence of the number R of data streams.

lated to the code and allocation structures are full rank, which lead to the uniqueness conditions in Table 3.1.

Table 4.7: Summary of the identifiability conditions with $N \geq R$ and $K \geq M$ (or, $N \geq QR$ and $K \geq QM$).

ALS	**TST system**		**STF system**
	multiple users	**one user**	
for $M = R$	$w_{m,r,j}^{(q)} = e^{i2\pi j \frac{mr}{MR}}$ $\mathbf{c}_{p^*.}^{(\mathbf{S},q)} = \mathbf{1}_R^{\mathrm{T}}$, $\mathbf{c}_{p^*.}^{(\mathbf{H},q)} = \mathbf{1}_M^{\mathrm{T}}$	$w_{m,r,j} = e^{i2\pi j \frac{mr}{MR}}$ $\mathbf{c}_{p^*.}^{(\mathbf{S})} = \mathbf{1}_R^{\mathrm{T}}$, $\mathbf{c}_{p^*.}^{(\mathbf{H})} = \mathbf{1}_M^{\mathrm{T}}$	$w_{m,r} = e^{i2\pi \frac{mr}{MR}}$ $\mathbf{C}_{\cdot p^*.}^{(\mathbf{S})} = \mathbf{1}_{F\times R}$, $\mathbf{C}_{\cdot p^*.}^{(\mathcal{H})} = \mathbf{1}_{F\times M}$
for $M < R$	$J \geq J_{\min}$ (see Tab. 4.1) $w_{m,r,j}^{(q)} = e^{i2\pi j \frac{mr}{MR}}$ $\mathbf{c}_{p^*.}^{(\mathbf{S},q)} = \mathbf{1}_R^{\mathrm{T}}$, $\mathbf{c}_{p^*.}^{(\mathbf{H},q)} = \mathbf{1}_M^{\mathrm{T}}$	$J \geq J_{\min}$ (see Tab. 4.1) $w_{m,r,j} = e^{i2\pi j \frac{mr}{MR}}$ $\mathbf{c}_{p^*.}^{(\mathbf{S})} = \mathbf{1}_R^{\mathrm{T}}$, $\mathbf{c}_{p^*.}^{(\mathbf{H})} = \mathbf{1}_M^{\mathrm{T}}$	$P \geq R$ $w_{m,r} = e^{i2\pi \frac{mr}{MR}}$ $\mathbf{C}_{\cdot p^*.}^{(\mathbf{S})} = \mathbf{1}_{F\times R}$, $\mathbf{C}_{\cdot p^*.}^{(\mathcal{H})} = \mathbf{1}_{F\times M}$ $\mathbf{C}_{f^*..}^{(\mathcal{H})} = \mathbf{1}_{P\times M}$ $\exists f^* \,\vert\, \mathbf{C}_{f^*..}^{(\mathbf{S})}$ full column-rank
for $M > R$			$P \geq M$ $w_{m,r} = e^{i2\pi \frac{mr}{MR}}$ $\mathcal{C}^{(\mathbf{S})} = \mathbf{1}_{F\times P\times R}$, $\mathbf{c}_{f^*,p^*.}^{(\mathcal{H})} = \mathbf{1}_M^{\mathrm{T}}$ $\mathbf{C}_{f..}^{(\mathcal{H})}$ full column-rank, $\forall f$

The time spreading mode in the code tensor permits to derive a minimum value of the spreading code length J that ensures the LS identifiability of channel and symbol matrices for the TST system when $M \neq R$. As the spreading code for the STF system is a matrix, the structure of the allocation tensors has to guarantee the identifiability of channel and symbol estimates when $M \neq R$, which leads to a minimum value of P ($P \geq R$ or $P \geq M$) and a stronger restriction on allocation tensors as shown in Table 4.7.

The identifiability condition presented in this table for the TST system is based on the received signal tensor modeled with the PT-(2,4) model. Rewriting as a constrained CP model, it allows to provide another condition by relaxing the restriction on the allocation matrices and

imposing $\mathbf{W}_{J\times MR}$ full column-rank ($J \geq MR$).

From the analysis of identifiability conditions, we can conclude that the TST system is more flexible to design the allocation structures than the STF system. For the TST system, it is also possible to guarantee the solution uniqueness by imposing a stronger restriction on the code tensor instead of the allocation structures. Remark that the identifiability conditions for the TST system with one user can be directly derived from the conditions for the multiuser TST system by $Q = 1$.

The transmission rate (in bits per channel use) is given by

$$R_b = \tau \log_2(\mu), \tag{4.140}$$

where μ denotes the cardinality of the information symbol constellation and the ratio τ is equal to $\tau = {}^{R}/_{P}$ and $\tau = {}^{R}/_{PF}$ for the TST and STF systems, respectively. Thereby, the transmission rate for the STF system is limited by $\tau = {}^{1}/_{MF}$ since $P \geq MR$ in accordance with the uniqueness results discussed in Chapter 3. Therefore, the transmission rate for the STF system depends on the number of transmit antennas M, which does not happen for the TST system due to the flexibility of uniqueness condition by an appropriate choice of the code tensor with $J \geq MR$ (see Table 3.1).

In terms of spectral efficiency, the STF system in the FDMA context divides the available bandwidth B into F disjoint frequency bands, i.e. ${}^{B}/_{F}$ Hz for each subcarrier and the data streams are repeated over these bands. Considering the symbol period T, the total bandwidth used for the system is $B_{\text{STF}} = {}^{F}/_{T}$ Hz. For the TST system in the CDMA context, all symbols are transmitted using the same frequency band. For the available bandwidth B, the reciprocal of B defines the duration of a pulse, i.e. the chip interval ${}^{T}/_{J}$, which gets $B_{\text{TST}} = {}^{J}/_{T}$ Hz. Thus, each system can theoretically achieve the same spectral efficiency for the same total bandwidth B and the same symbol period T. Nevertheless, the FDMA-MIMO system can achieve higher spectral efficiency than CDMA in practice.

4.6 Simulation results

In this section, we present some simulation results considering the proposed receivers based on four algorithms: ALS, LM, ALM, and KLS, which jointly estimate the symbol matrix $\mathbf{S} \in \mathbb{C}^{N\times R}$ and the channel matrix $\mathbf{H} \in \mathbb{C}^{K\times M}$ or tensor $\mathcal{H} \in \mathbb{C}^{F\times K\times M}$ in the presence of the additive noise assumed to be zero-mean complex-valued white Gaussian.

The total number of Monte Carlo runs (L) is fixed equal to 2000, which corresponds to 2000 random wireless channels, each one with different symbol sequences randomly drawn from a PSK constellation, different random noise sequences and different random allocation structures subject to the uniqueness and identifiability results. Assuming flat Rayleigh fading propagation channels, a different random initialization $\hat{\mathbf{H}}_{l=0}$ is also used for each run. The code matrix $\mathbf{W} \in \mathbb{C}^{M\times R}$ and the code tensor $\mathcal{W} \in \mathbb{C}^{M\times R\times J}$ are constructed based on the Vandermonde structure.

The performance associated to the proposed receivers is evaluated by means of Monte Carlo simulations in terms of bit-error rate (BER) and normalized mean square error (NMSE) on

channel estimation defined as

$$\mathrm{NMSE}_{\mathrm{dB}}^{\mathbf{H}} = 10\log_{10}\left(\frac{1}{L}\sum_{l=1}^{L}\frac{\left\|\mathbf{H}_l - \hat{\mathbf{H}}_{l(\infty)}\right\|_F^2}{\|\mathbf{H}_l\|_F^2}\right), \tag{4.141}$$

where $\hat{\mathbf{H}}_{l(\infty)}$ is the channel matrix (or an unfolded matrix $\mathbf{H}_{K\times FM}$ for the STF system) estimated at convergence of the l-th run. The BER is calculated by averaging the results obtained for all data streams and all Monte Carlo runs. The signal-to-noise ratio (SNR) is determined by

$$\mathrm{SNR}_{\mathrm{dB}} = 10\log_{10}\frac{\|\mathcal{X}\|_F^2}{\|\mathcal{V}\|_F^2}, \tag{4.142}$$

where $\mathcal{X}$ and $\mathcal{V}$ are, respectively, the received signal tensor without noise and the additive noise tensor. The SNR is set by adjusting an adequate noise variance.

The convergence of algorithms is decided when the errors between the noisy received signal tensor and its values reconstructed using the channel and symbol matrices estimated at two successive iterations are such as

$$\frac{\left|e_{(it)} - e_{(it-1)}\right|}{e_{(it-1)}} \leq 10^{-6}, \tag{4.143}$$

with $e_{(it)} = \left\|\tilde{\mathcal{X}} - \hat{\mathcal{X}}_{(it)}\right\|_F$, $\tilde{\mathcal{X}}$ and $\hat{\mathcal{X}}_{(it)}$ denote, respectively, the noisy received signal tensor and its estimate at it-th iteration. In another way, the convergence is considered when the error between iterations $it-1$ and it does not significantly change.

The convergence speed of algorithms is evaluated in terms of the NMSE of received signal estimation, constructed by symbol and channel estimates at each iteration it, defined as

$$\mathrm{NMSE}_{\mathrm{dB}}^{\mathcal{X}} = 10\log_{10}\left(\frac{1}{L}\sum_{l=1}^{L}\frac{\left\|\tilde{\mathcal{X}}_l - \hat{\mathcal{X}}_{l(it)}\right\|_F^2}{\left\|\tilde{\mathcal{X}}_l\right\|_F^2}\right), \tag{4.144}$$

where $\tilde{\mathcal{X}}_l$ and $\hat{\mathcal{X}}_{l(it)}$ are, respectively, the noisy received signal tensor and its estimate at the it-th iteration of the l-th run.

According to the uniqueness analysis investigated in Chapter 3, we assume the knowledge of one symbol in order to eliminate the scaling ambiguity of the channel and symbol estimates inherent to the models related to the TST and STF systems. The KLS procedure is an non-iterative algorithm and does not present the ambiguity problem since it explores the structure of the Kronecker product from a priori knowledge of some transmitted symbols.

In all simulations, the code ($\mathbf{W} \in \mathbb{C}^{M\times R}$ or $\mathcal{W} \in \mathbb{C}^{M\times R\times J}$) and the allocation structures ($\mathbf{C}^{(\mathbf{S})} \in \mathbb{C}^{P\times R}$ and $\mathbf{C}^{(\mathbf{H})} \in \mathbb{C}^{P\times M}$, or $\mathcal{C}^{(\mathbf{S})} \in \mathbb{C}^{F\times P\times R}$ and $\mathcal{C}^{(\mathcal{H})} \in \mathbb{C}^{F\times P\times M}$) are assumed to be known at both transceiver and receiver. By default, we consider the system configuration below:

For the TST system:

$$\begin{cases} \bullet \text{ Vandermonde structure for } \mathbf{W}_{\cdot\cdot j} : w_{m,r,j} = e^{i2\pi j\frac{mr}{MR}} \\ \bullet \ {\mathbf{C}^{(\mathbf{H})}}^{\mathrm{T}} \diamond {\mathbf{C}^{(\mathbf{S})}}^{\mathrm{T}} \text{ full row-rank} \\ \bullet \ \mathbf{c}_{1\cdot}^{(\mathbf{S})} = \mathbf{1}_R^{\mathrm{T}}, \mathbf{c}_{1\cdot}^{(\mathbf{H})} = \mathbf{1}_M^{\mathrm{T}} \end{cases}. \tag{4.145}$$

For the STF system, when $M = R$:

$$\begin{cases} \bullet \text{ Vandermonde structure for } \mathbf{W}_{M\times R} : w_{m,r} = e^{i2\pi\frac{mr}{MR}} \\ \bullet \ \mathbf{C}_{f..}^{(\mathcal{H})^{\mathrm{T}}} \diamond \mathbf{C}_{f..}^{(\mathbf{S})^{\mathrm{T}}} \text{ full row-rank for all } f \\ \bullet \ \mathbf{C}_{.1.}^{(\mathbf{S})} = \mathbf{1}_{F\times R}, \mathbf{C}_{.1.}^{(\mathcal{H})} = \mathbf{1}_{F\times M} \end{cases} \tag{4.146}$$

and when $M < R$:

$$\begin{cases} \bullet \text{ Vandermonde structure for } \mathbf{W}_{M\times R} : w_{m,r} = e^{i2\pi\frac{mr}{MR}}, \\ \bullet \ \mathbf{C}_{f..}^{(\mathcal{H})^{\mathrm{T}}} \diamond \mathbf{C}_{f..}^{(\mathbf{S})^{\mathrm{T}}} \text{ full row-rank for all } f \\ \bullet \ \mathbf{C}_{.1.}^{(\mathbf{S})} = \mathbf{1}_{F\times R}, \mathbf{C}_{1..}^{(\mathbf{S})} \text{ full column-rank} \\ \bullet \ \mathbf{C}_{.1.}^{(\mathcal{H})} = \mathbf{1}_{F\times M}, \mathbf{C}_{1..}^{(\mathcal{H})} = \mathbf{1}_{P\times M} \end{cases} . \tag{4.147}$$

Firstly, we study the influence of several system parameters for the TST and STF systems as: the spreading code length (J), the number of phases of the PSK modulation for representing the data symbols, the number of subcarriers (F), of blocks (P), of data streams (R), and of receive antennas (K). In the sequence both systems are compared regarding the symbol recovery. For this performance analysis, the ZF receivers, defined in (4.53) and (4.55), are employed assuming a perfect knowledge of the channel coefficients. Let us recall the receivers for the TST and STF systems denoted, respectively, by ZF-TST and ZF-STF.

$$\hat{\mathbf{S}} = \tilde{\mathbf{X}}_{N\times PJK} \left(\left(\mathbf{I}_{PJ} \otimes \mathbf{H} \right) \mathbf{G}_{PJM\times R} \right)^{\dagger^{\mathrm{T}}}, \tag{4.148}$$

$$\hat{\mathbf{S}} = \tilde{\mathbf{X}}_{N\times PFK} \left(\left(\mathbf{I}_{P} \otimes \mathrm{bdiag}(\mathbf{H}_{1..}, \ldots, \mathbf{H}_{F..}) \right) \mathbf{T}_{PFM\times R} \right)^{\dagger^{\mathrm{T}}}. \tag{4.149}$$

Since the uniqueness and identifiability conditions are only sufficient conditions, we study the influence of these restrictions on the performance of the ALS algorithm in terms of BER and channel NMSE. The influence of a priori information on the performance of the KLS algorithm is evaluated by employing the knowledge of one symbol and one row of the symbol matrix. We establish a comparison between the performance of all proposed receivers in terms of symbol recovery and convergence speed. The TST coding is compared with well-known tensor approaches, such as KRST [53] and STM codes [55], using the ALS algorithm. The BER performance of the TST system for multiple users in the transmission is computed.

4.6.1 Performance analysis of the TST system

Let us consider the ZF receiver to analyze the influence of the system parameters on the performance of the TST system. Figure 4.5 shows the BER versus SNR for five values of the spreading code length ($J \in \{1, 2, 4, 6, 10\}$). Note that the BER is canceled for a $SNR = \{18; 12; 12; 10\}$ dB for $J = \{2; 4; 6; 10\}$, respectively. From this figure, we can conclude that an increase of J induces a significant performance improvement in terms of symbol recovery. It is important to remark that the case $J = 1$ corresponds to the ZF receiver proposed in [59], the improvement obtained is due to the extra time spreading introduced by the TST coding. Evidently, an increase of J provides a better performance at the cost of a computation complexity increase.

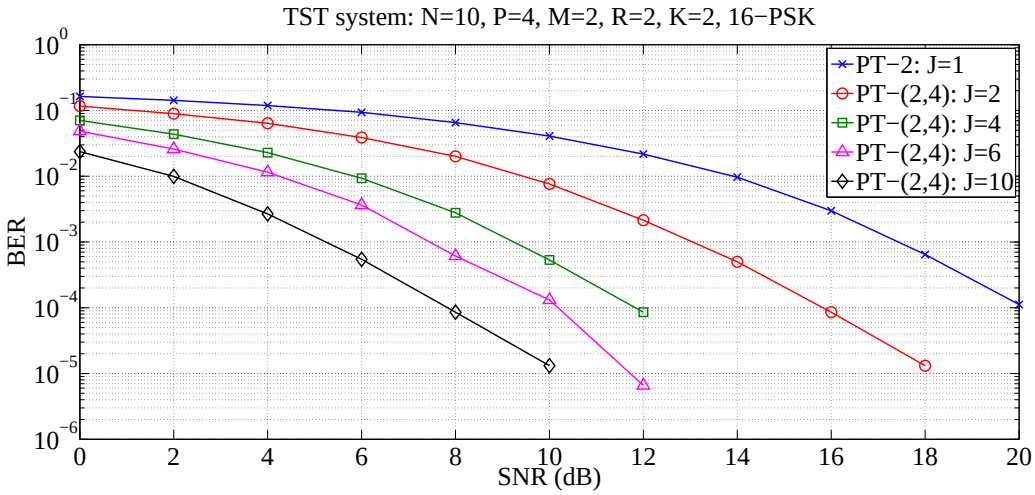


Figure 4.5: ZF-TST receiver: Influence of the spreading code length.

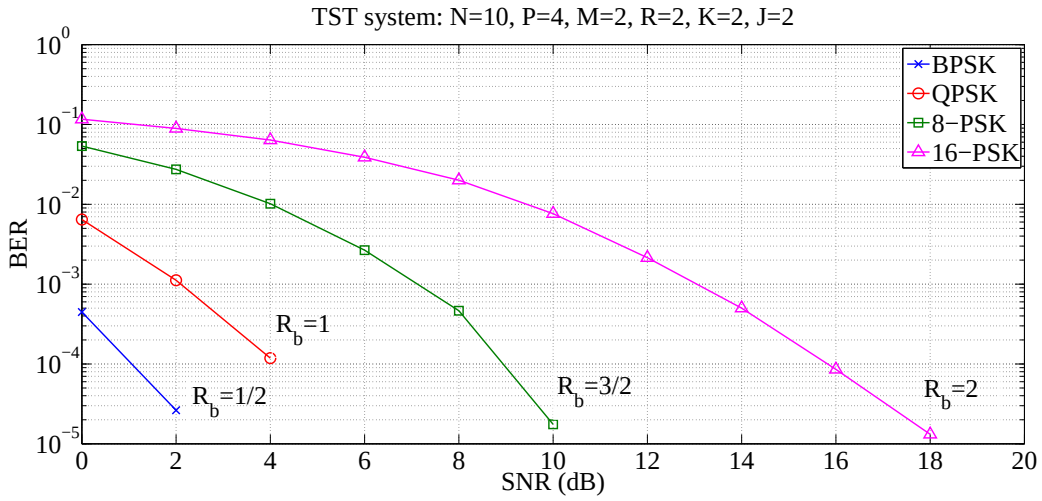


Figure 4.6: ZF-TST receiver: Influence of the number of phases of PSK modulation.

Figure 4.6 shows the BER performance for four PSK constellations: 2- (BPSK), 4- (QAM or QPSK), 8- and 16-PSK. We verify that the symbol recovery is improved for low numbers of phase. The greater the number of phases, the greater will be the number of bits to represent each symbol. Higher-order PSK modulation leads to higher transmission rate but also higher bit-error rates, since the total energy per symbol is divided per bit and hence the energy per bit is reduced. In practice, the use of different PSK constellations depends on the transmission rate required and the difficult to implement it.

In Figure 4.7 we have plotted the BER versus SNR for two different values of the number of blocks ($P \in \{4, 10\}$) and of the spreading code length ($J \in \{1, 2\}$). Note that the BER performance can be significantly improved by increasing either the number of data blocks or the spreading code length, since both actions induce an increase of time diversity. However, the increase of the number of data blocks leads to a decrease of the transmission rate given by (4.140) (proportional to the ratio $^{R}/_{P}$). We obtain the same BER performance for $J = 2$ and $P = 4$, and $J = 1$ and $P = 10$, but the first case provides a transmission rate twice higher

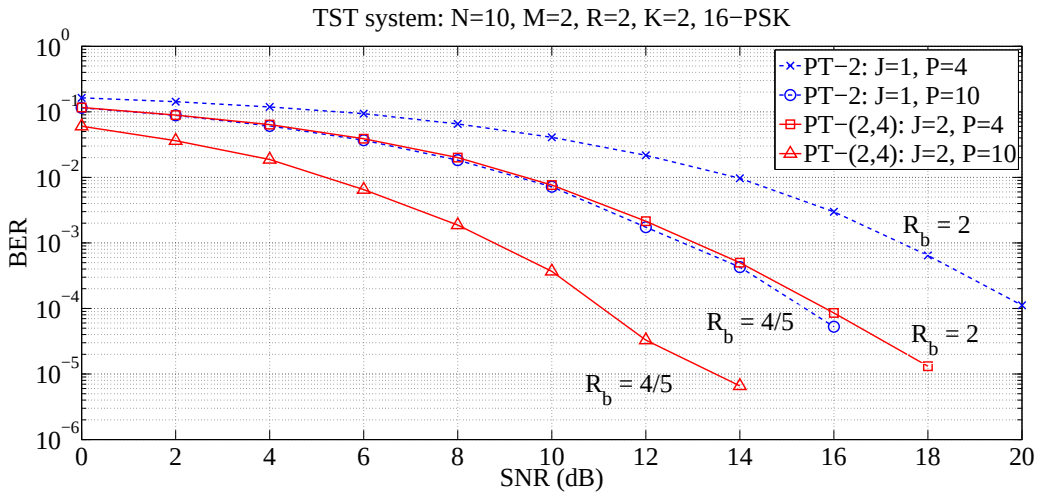


Figure 4.7: ZF-TST receiver: Influence of the number of data blocks.

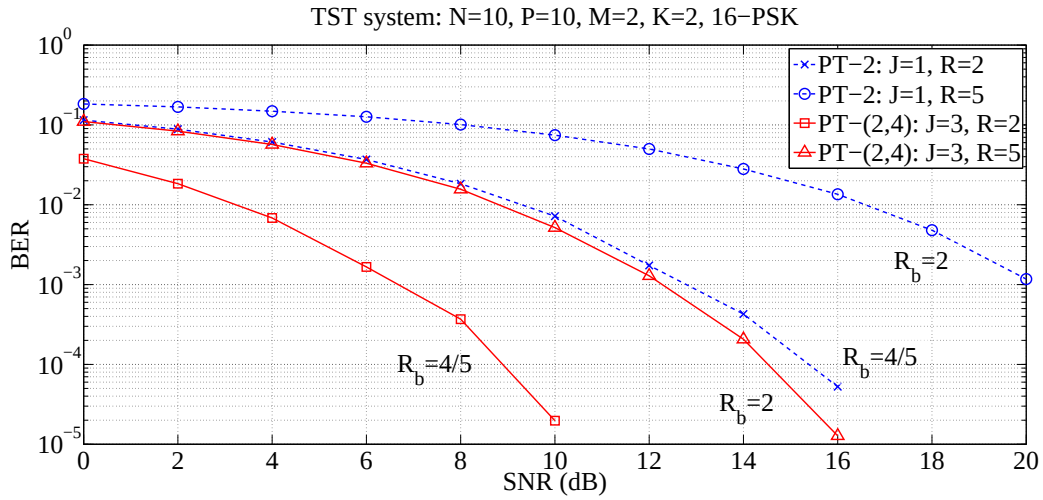


Figure 4.8: ZF-TST receiver: Influence of the number of data streams.

(exactly, 2,5 times).

The transmission rate decreases from 2 to 4/5 bit per channel use when R is decreased from 5 to 2. On the other hand, as expected and shown in Figure 4.8, the BER performance is improved when R is reduced from 5 to 2, due to the fact that fewer symbols have to be estimated using the same number of received signals. It illustrates the trade-off between error performance and transmission rate that can be achieved with the proposed TST coding. It is interesting to note that the error performance for $J = 1$ and $R = 2$ is close to the one obtained with $J = 3$ and $R = 5$. Hence, the TST coding by adjusting J allows to obtain almost the same performance but providing a higher transmission rate.

Figure 4.9 shows the BER versus SNR for two different values of the number of receive antennas ($K \in \{2, 4\}$) and two spreading code lengths ($J \in \{1, 2\}$). As expected, the use of more receive antennas leads to a better BER performances. It is interesting to notice that the TST coding provides the same performance employing half of the number of receive antennas ($K = 2$) than the receiver proposed in [59] with $K = 4$ antennas thanks to extra time diversity.

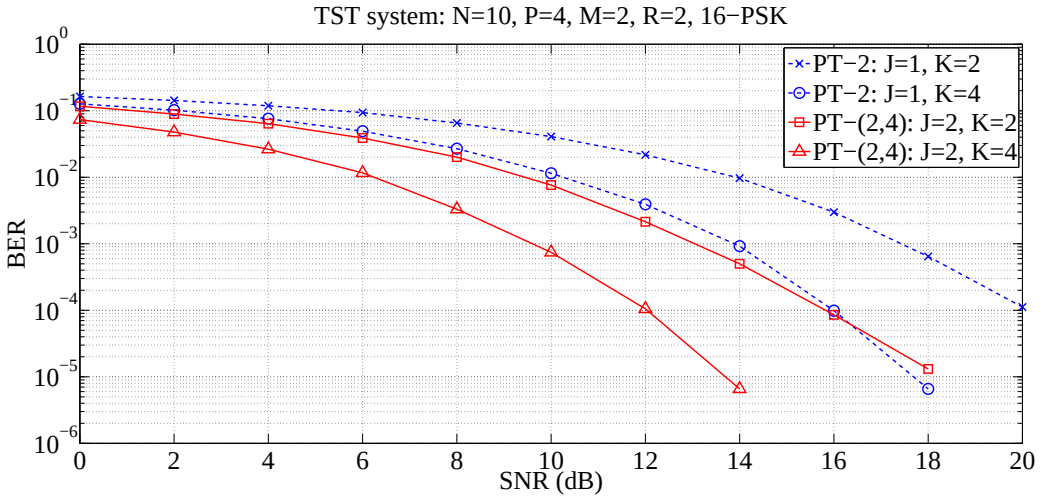


Figure 4.9: ZF-TST receiver: Influence of the number of receive antennas.

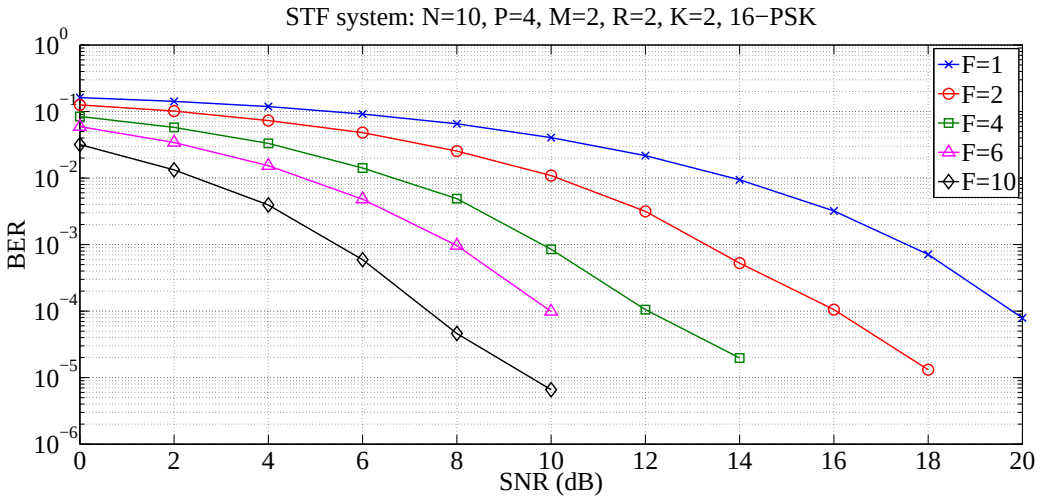


Figure 4.10: ZF-STF receiver: Influence of the number of subcarriers.

4.6.2 Performance analysis of the STF system

The ZF receiver is employed to analyze the performance of the STF system analogously to the TST system. In Figure 4.10 we have plotted the BER versus SNR for five values of the number of subcarriers ($F \in \{1, 2, 4, 6, 10\}$). It confirms the performance improvement by increasing of the frequency diversity as well as the time diversity for the TST system. The frequency diversity is eliminated when $F = 1$ and it also leads to the receiver of [59]. Observe that the number of subcarriers must be supported by the total available bandwidth, since the system bandwidth has to be divided into F disjoint frequency bands.

Figure 4.11 shows the BER versus SNR for four PSK modulations: BPSK, QPSK, 8-PSK, and 16-PSK. Analogously to the TST system, low values of phase provide better BER performances.

According to Figures 4.12 and 4.13, we obtain almost the same performance for ($F = 1$, $P = 10$) and ($F = 2$, $P = 4$), and ($F = 1$, $R = 2$) and ($F = 3$, $R = 5$), respectively. A way to reduce the error rate without changing the transmission rate $R_b = 2$ is by the use of more

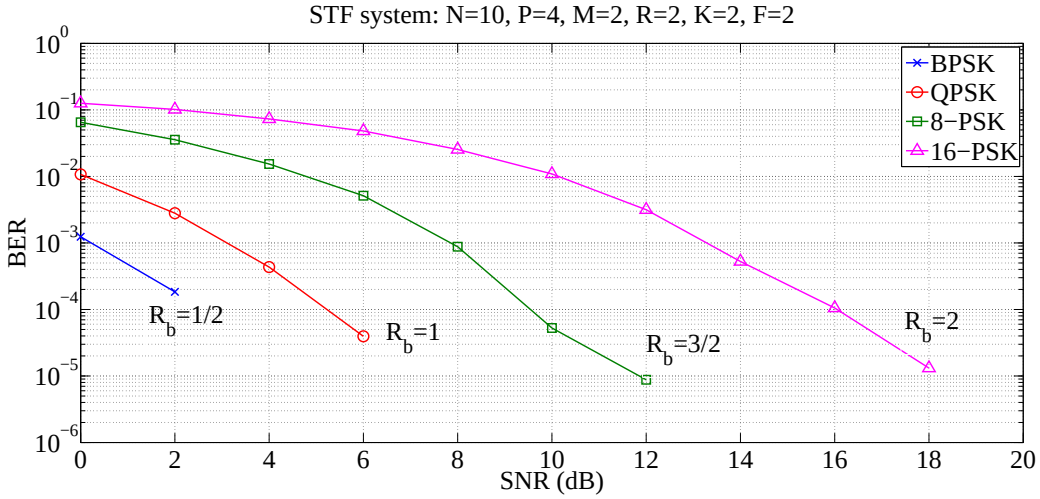


Figure 4.11: ZF-STF receiver: Influence of the number of phases of PSK modulation.

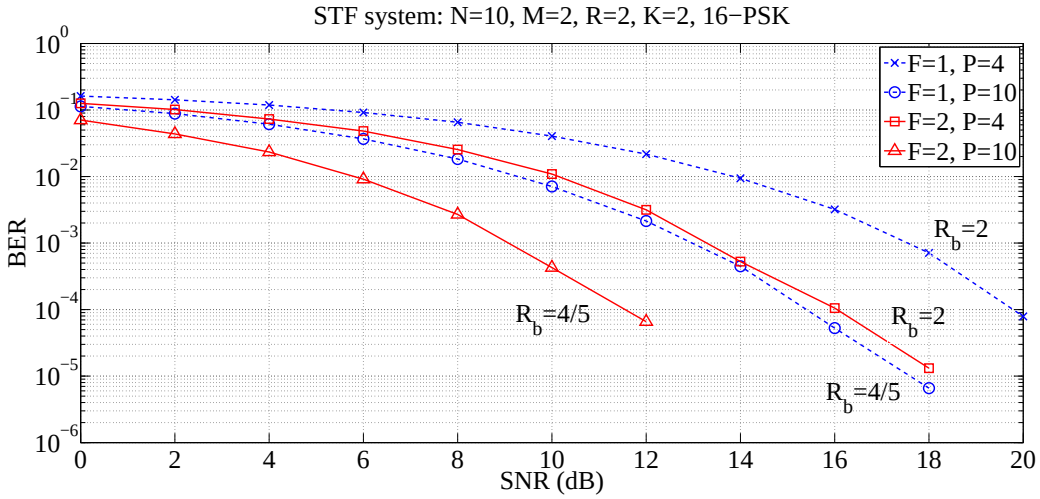


Figure 4.12: ZF-STF receiver: Influence of the number of data blocks.

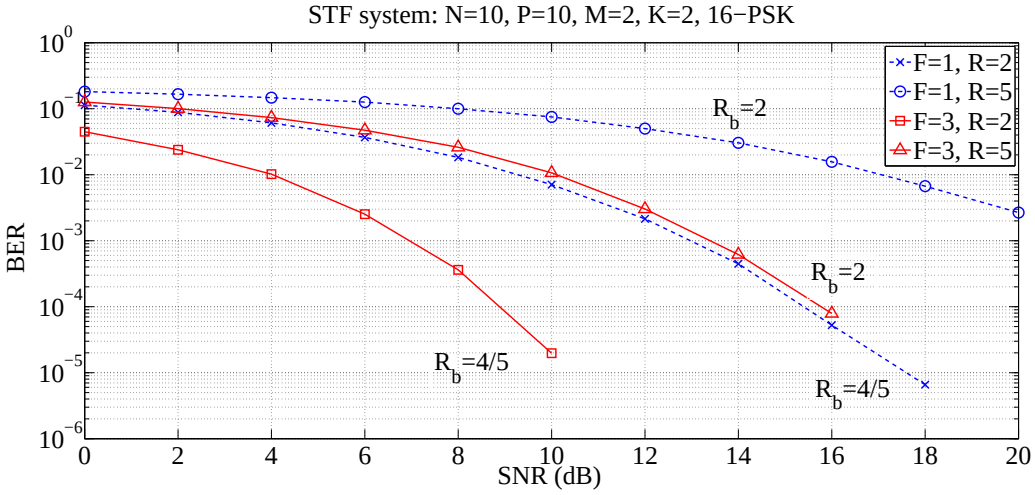


Figure 4.13: ZF-STF receiver: Influence of the number of data streams.

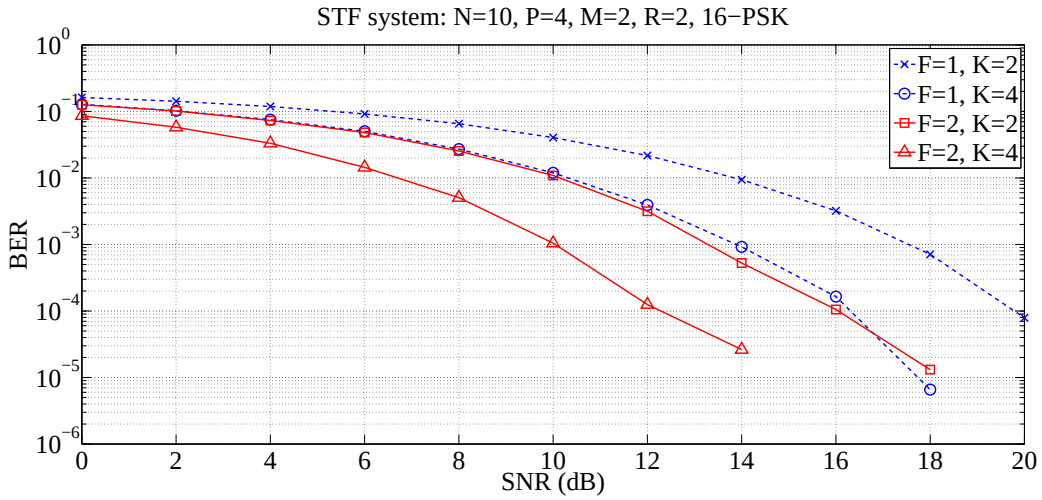


Figure 4.14: ZF-STF receiver: Influence of the number of receive antennas.

subcarriers (with $F = 2$ and $F = 3$). The transmission rate can be increased by appropriately setting R and/or P, however the BER performance is impaired.

Figure 4.14 shows the BER versus SNR for two values of the number of subcarriers ($F \in \{1, 2\}$) and two values of the number of receive antennas ($K \in \{2, 4\}$). The use of two subcarriers allows to obtain the same BER performance employing half of the receive antennas.

4.6.3 Comparison between TST and STF systems

In order to analyze the influence of J and F on diversity gain of the TST and STF systems, we fix all parameters including $J = F$ and employ the random allocation structures such that $\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}$ and $\mathbf{C}_{f..}^{(\mathcal{H})^{\mathrm{T}}} \diamond \mathbf{C}_{f..}^{(\mathbf{S})^{\mathrm{T}}}$ for all $f \in \{1, ..., F\}$ are full row-rank, which lead to the same diversity gain for both systems as viewed in Chapter 3.

The BER curves obtained with the ZF receiver for the TST and STF systems are plotted in Figure 4.15 for five values of $J = F \in \{1, 2, 4, 6, 10\}$. For $J = F = 1$, verify the same

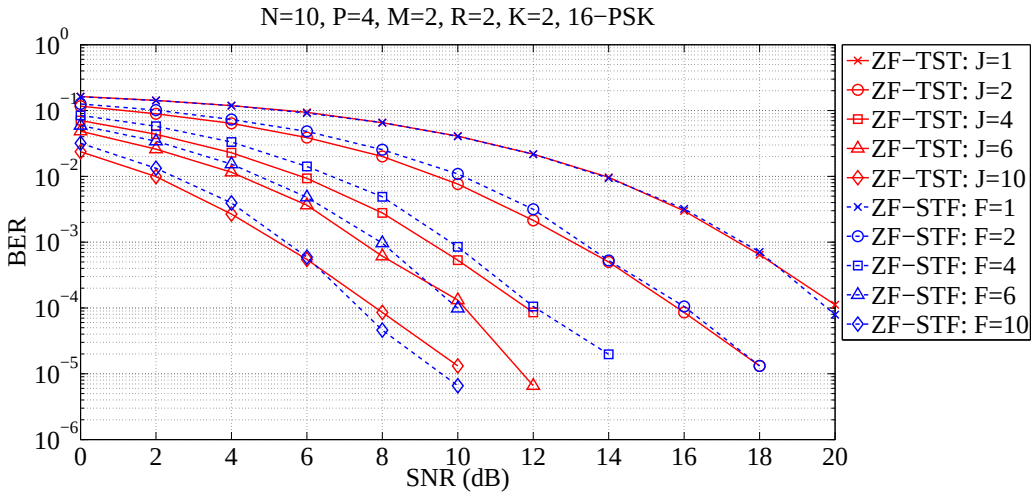


Figure 4.15: TST × STF systems: Influence of the diversity gain by J and F.

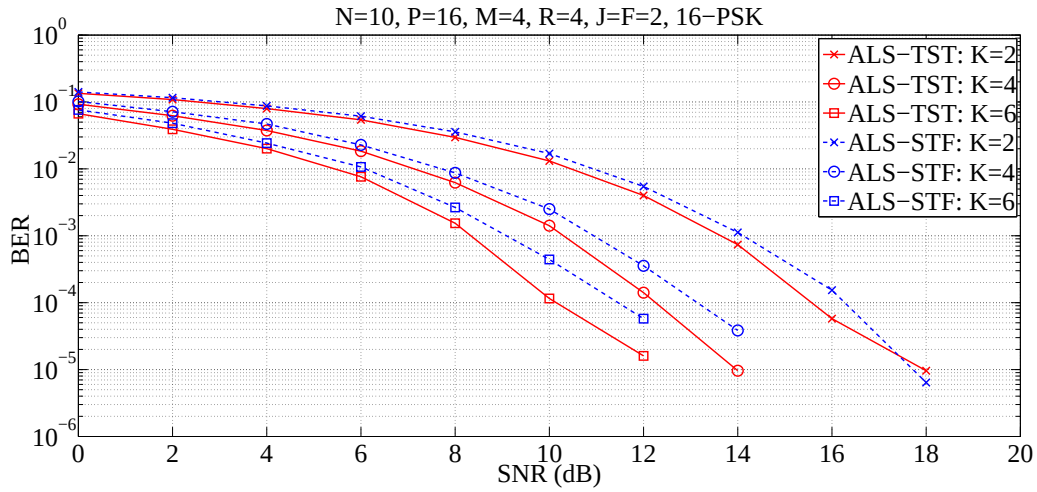


Figure 4.16: Influence of the number of receive antennas: BER versus SNR.

performance for both systems as expected. When $J = 1$ and $F = 1$, the extra time and frequency diversities are respectively eliminated resulting an equivalent system. Assuming a perfect knowledge of the channel coefficients, we have the same number of system parameters to estimate, i.e. NR. When J and F are increased from 2 to 10, we note that both systems tend to provide almost the same curves of BER, which allows to confirm that the time and frequency diversities can lead to an equivalent diversity gain.

In Figures 4.16 and 4.17 we have respectively plotted the BER and the channel NMSE versus SNR for three values of the number of receive antennas ($K \in \{2, 4, 6\}$). The purpose is to study the uniqueness conditions discussed in Chapter 3 and the LS-identifiability of the symbol and channel estimates deduced in this chapter for the ALS algorithm. According to Tables 3.1 and 4.7, the uniqueness condition for the STF system imposes $K \geq FM$ and the identifiability conditions for both systems are derived by assuming $K \geq M$.

From Figures 4.16 and 4.17, observe that even when $K < M$ the ALS algorithm for both systems can still estimate uniquely the symbols and channel coefficients. It is important to at-

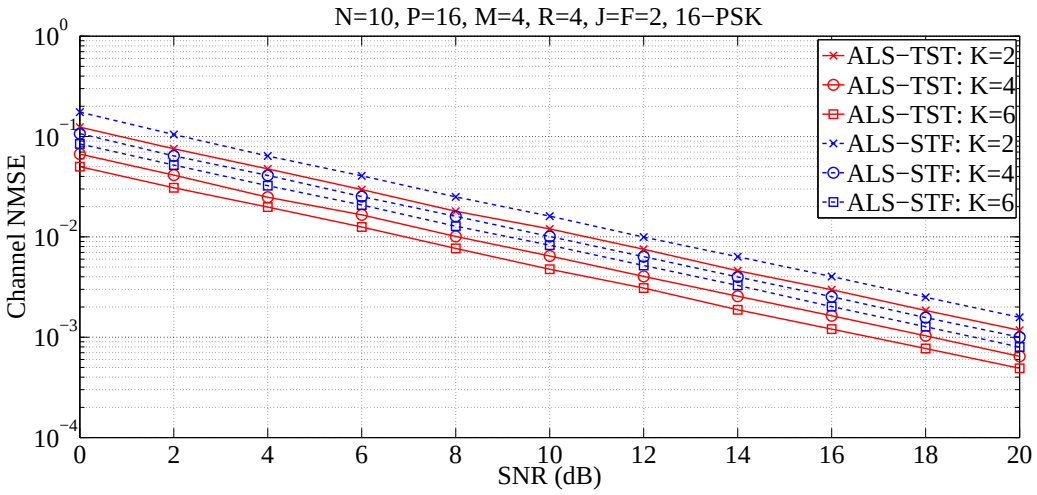


Figure 4.17: Influence of the number of receive antennas: Channel NMSE versus SNR.

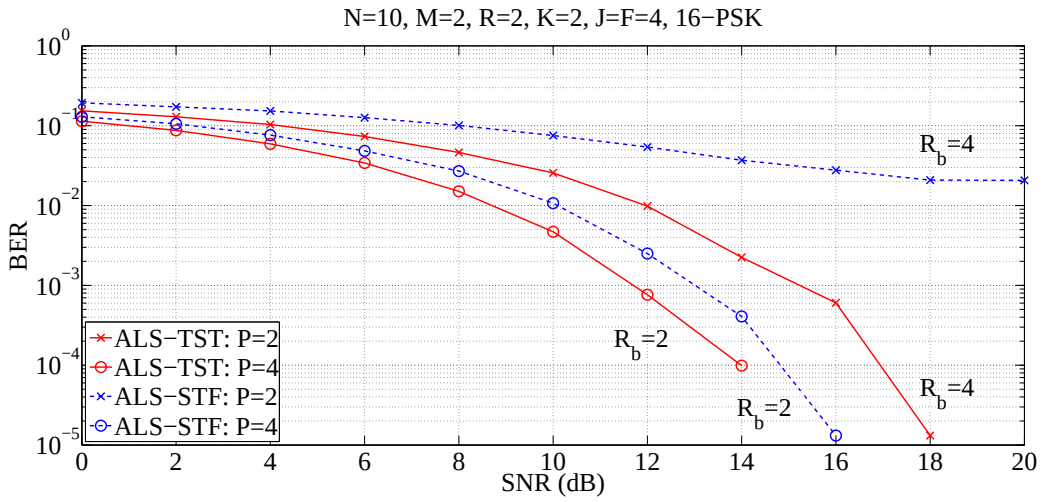


Figure 4.18: Influence of the number of data blocks: BER versus SNR.

tend that the proposed conditions of uniqueness and identifiability are only sufficient conditions. In this sense, even if the conditions are not satisfied we may perfectly estimate the symbol and channel. Note that the STF system has more channel coefficients to estimate and, therefore, the ALS algorithm for the TST system provides a better performance in terms of symbol and channel estimation.

Figures 4.18 and 4.19 show, respectively, the BER and channel NMSE versus SNR obtained for $P \in \{2, 4\}$. The ALS algorithms are employed to analyze the influence of P on the uniqueness conditions for both systems established in Chapter 3 (see Table 3.1). Both allocation structures are randomly chosen such that $\mathbf{C}^{(\mathbf{H})^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S})^{\mathrm{T}}}$ and $\mathbf{C}_{f..}^{(\mathcal{H})^{\mathrm{T}}} \diamond \mathbf{C}_{f..}^{(\mathbf{S})^{\mathrm{T}}}$ for all $f \in \{1, ..., F\}$ are full row-rank implying $P \geq MR$.

According to these figures, the channel and symbol of the STF system can not be properly estimated when $P = 2 < MR$, despite the convergence of the ALS algorithm. As the uniqueness condition for the TST system depends on both allocation matrices and code tensor, the symbol and channel estimations can be achieved even when $P < MR$ by appropriately adjust-

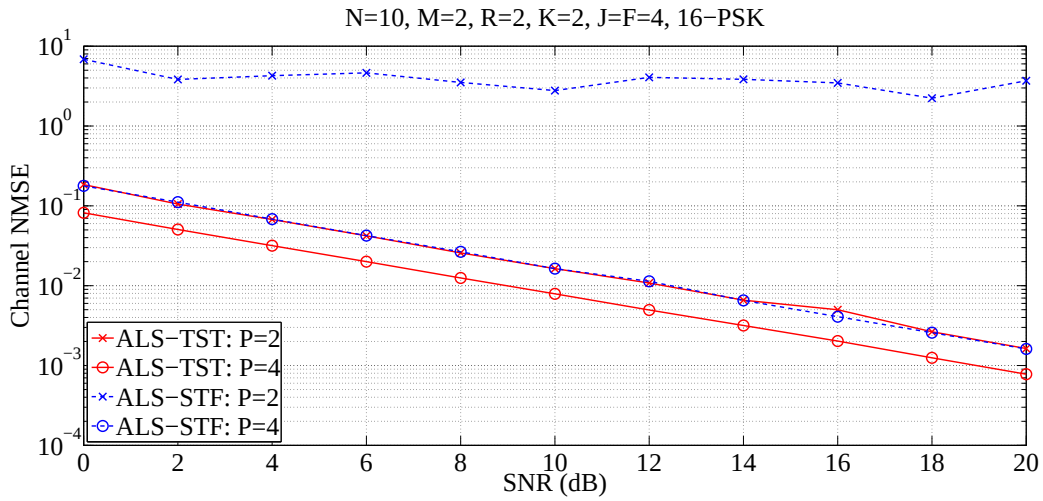


Figure 4.19: Influence of the number of data blocks: Channel NMSE versus SNR.

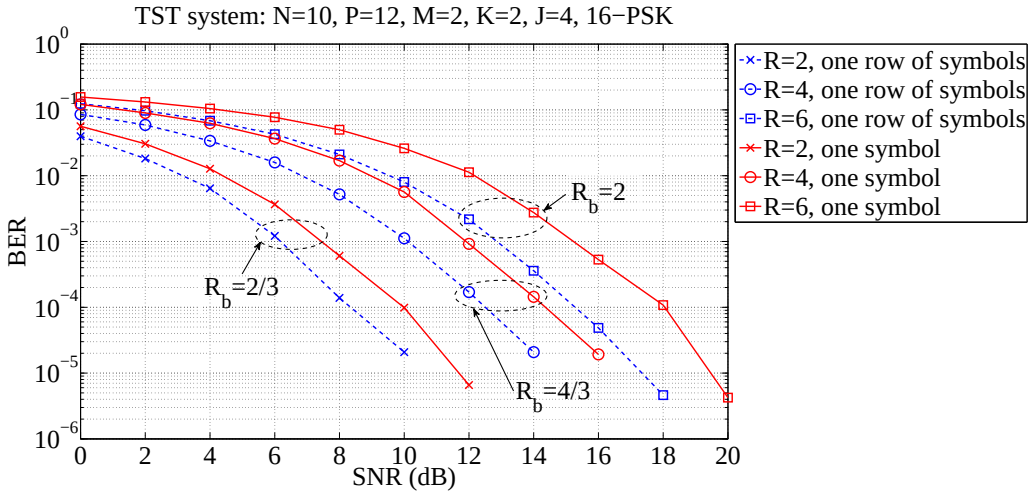


Figure 4.20: KLS-TST Receiver: Influence of a priori information in the symbol recovery.

ing the code tensor ($J \geq MR$). Furthermore, observe that the TST system provides better BER performance since the STF system has four times more channel parameters ($F = 4$) to estimate. Hence, the TST system affords a better trade-off between transmission rate and error performance than the STF system.

4.6.4 KLS algorithm: Influence of a priori information

In order to evaluate the influence of a priori information considered in the KLS algorithm, we consider three different values of $R \in \{2, 4, 6\}$ and assume the knowledge of only one symbol and one row of symbols to estimate the channel coefficients. Figures 4.20 and 4.21, and 4.22 and 4.23 show the BER and channel NMSE versus SNR for the TST and STF systems, respectively.

As expected, the knowledge of more than one symbol improves the channel estimation and consequently, the symbol recovery. The KLS algorithm for the TST system estimates better the channel coefficients than the one for the STF system (mainly for $R > 2$). Because the STF

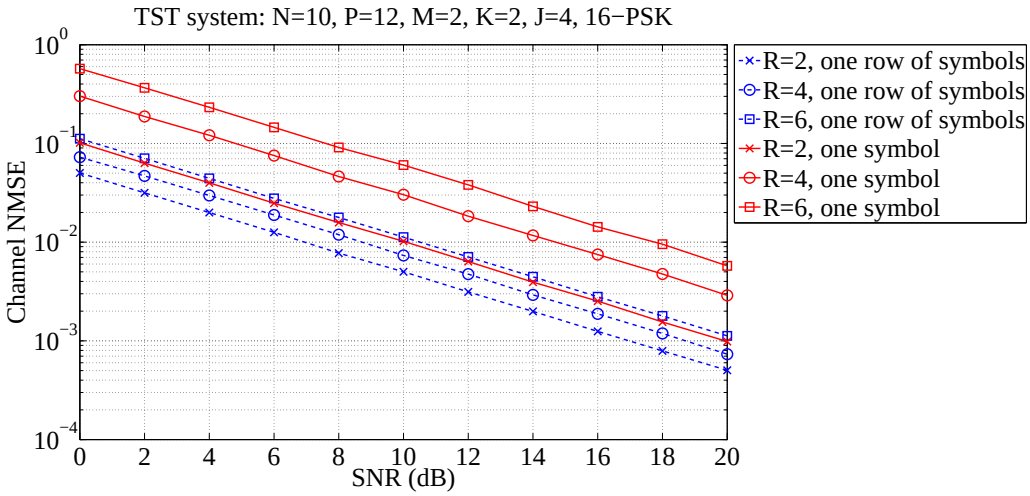


Figure 4.21: KLS-TST Receiver: Influence of a priori information in the channel estimation.

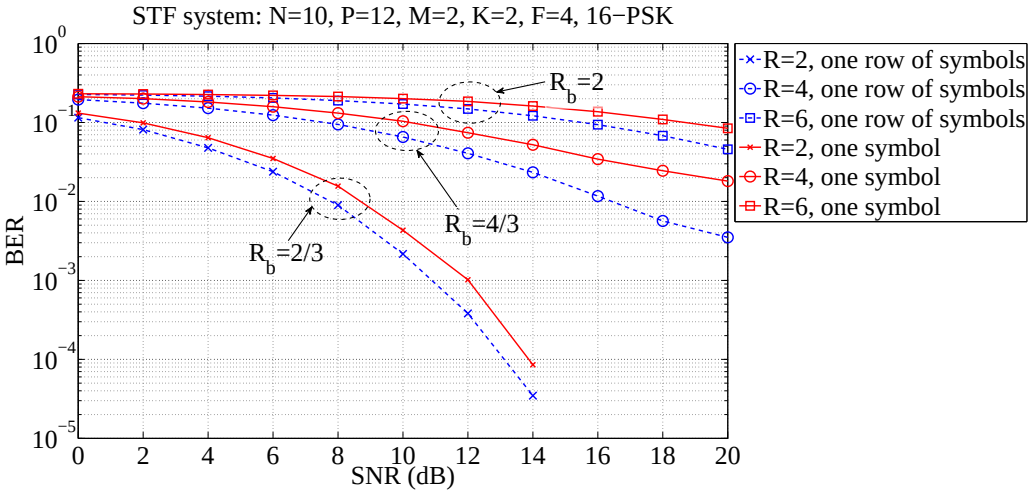


Figure 4.22: KLS-STF Receiver: Influence of a priori information in the symbol recovery.

system has F times more channel coefficients to estimate by using the same priori information about the transmitted symbols. Consequently, the poor channel estimation for the STF system leads to a poor symbol recovery.

4.6.5 Comparison between different algorithms

In the sequel, the ALS, LM, ALM and KLS algorithms are compared in terms of the symbol recovery and convergence speed for both TST and STF systems. The KLS algorithm is employed assuming the knowledge of one symbol, i.e. the first transmitted symbol being $s_{1,1} = 1$, and of one row of symbols, i.e. the first row of symbols being $\mathbf{S}_{1\cdot} = \mathbf{1}_R^{\mathrm{T}}$, at both transceiver and receiver.

In Figures 4.24, 4.25, 4.26 and 4.27 we have plotted the BER versus SNR for two different values of the spreading code length $J \in \{2, 6\}$, of the number of subcarriers $F \in \{2, 8\}$, and of the number of data streams $R \in \{2, 4\}$. From these figures, we observe that the ALS, LM and

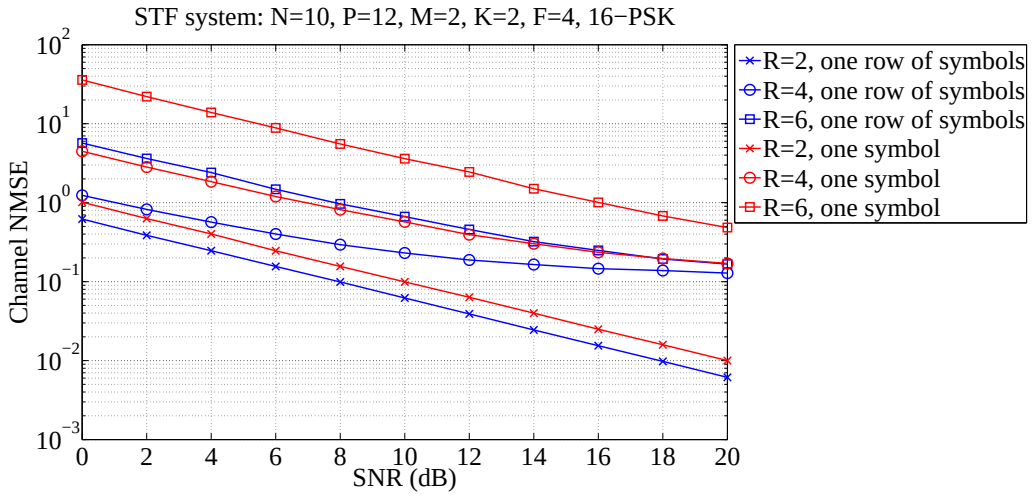


Figure 4.23: KLS-STF Receiver: Influence of a priori information in the channel estimation.

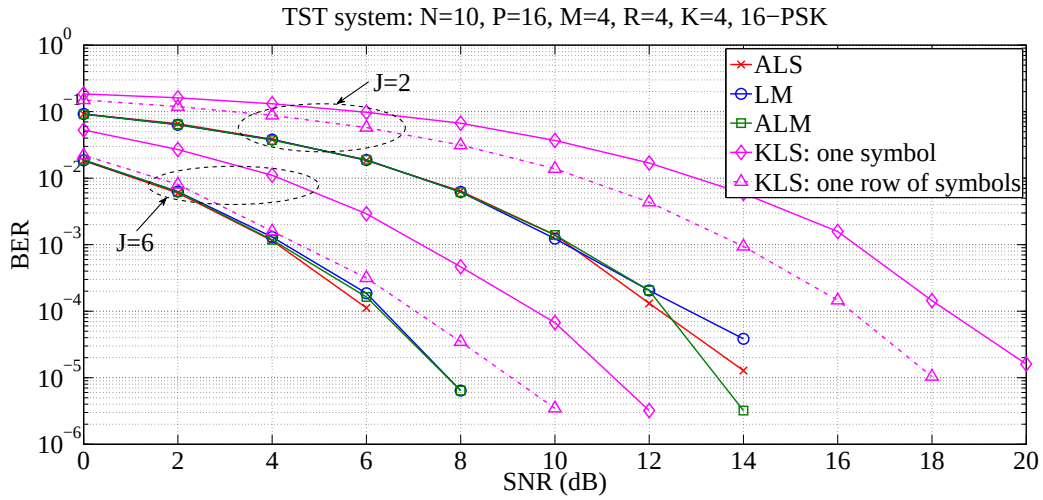


Figure 4.24: TST system: Performance of all receivers for two values of J.

ALM algorithms present close performances at the convergence given by (4.143) for different values of J, F and R.

According to Figure 4.24, the KLS-TST receiver employing the knowledge of one row of symbols can considerably improve the symbol recovery by increasing of the number of spreading code length from 2 to 6. From Figure 4.25, we can observe that the BER performance of the KLS-STF receiver can not be improved by increasing F as happens for the KLS-TST with J. For a BER equal to 10^{-3}, the gap between KLS-STF (assuming the knowledge of one row of symbols) and other receivers (ALS-STF, LM-STF and ALM-STF) is around 4 dB for both values $F = 2$ and $F = 8$. It can be explained by the larger number of channel coefficients to estimate for the STF system as mentioned above.

From Figures 4.26 and 4.27, we can note that the decrease of the number of data streams from 4 to 2 reduces the difference between the BER curves obtained with the KLS algorithms and other algorithms (ALS, LM and ALM) for both systems. For a BER equal to 10^{-3}, the gap between KLS-TST (assuming the knowledge of one row of symbols) and other receivers is

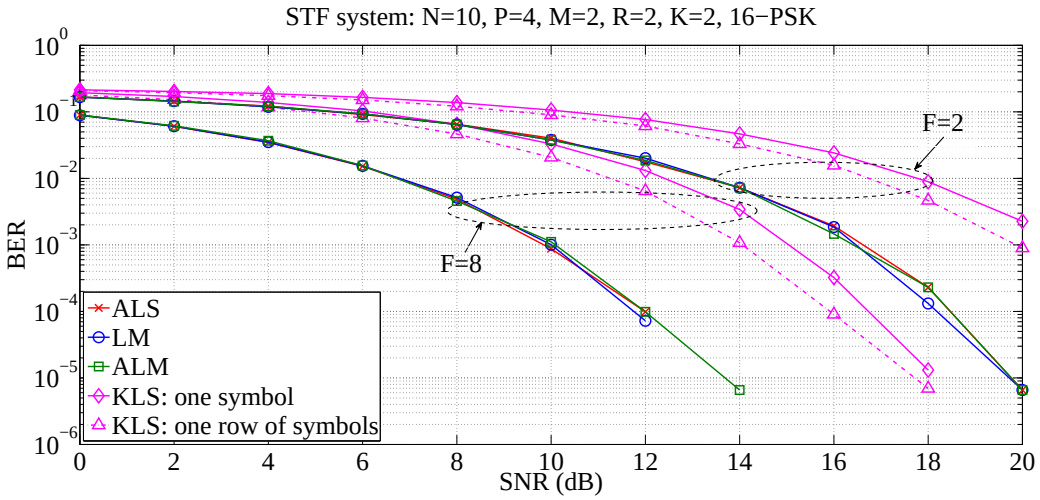


Figure 4.25: STF system: Performance of all receivers for two values of F.

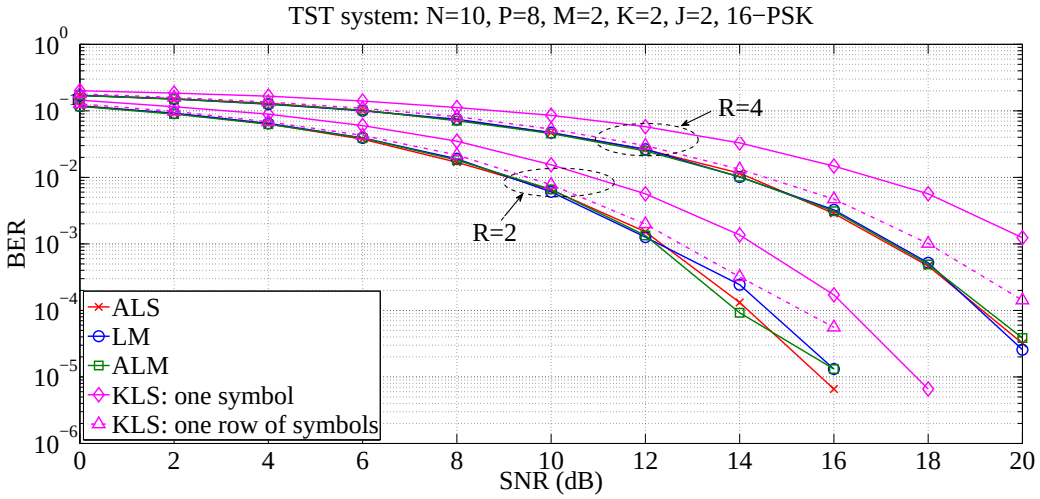


Figure 4.26: TST system: Performance of all receivers for two values of R.

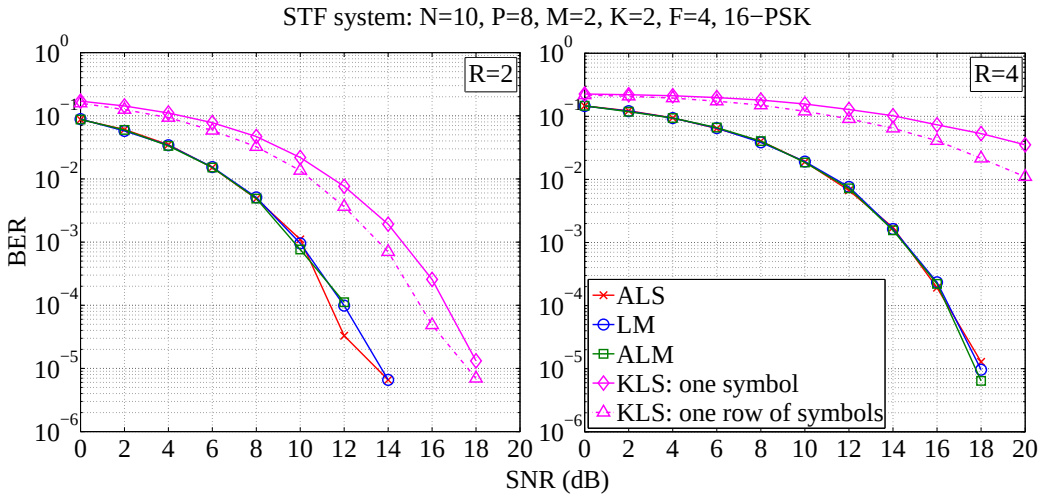


Figure 4.27: STF system: Performance of all receivers for two values of R.

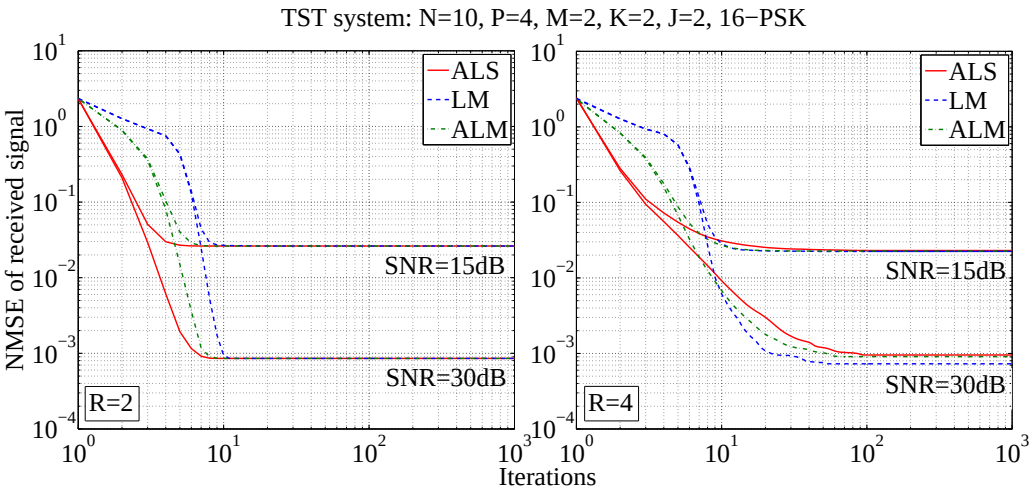


Figure 4.28: TST system: Convergence speed of all algorithms.

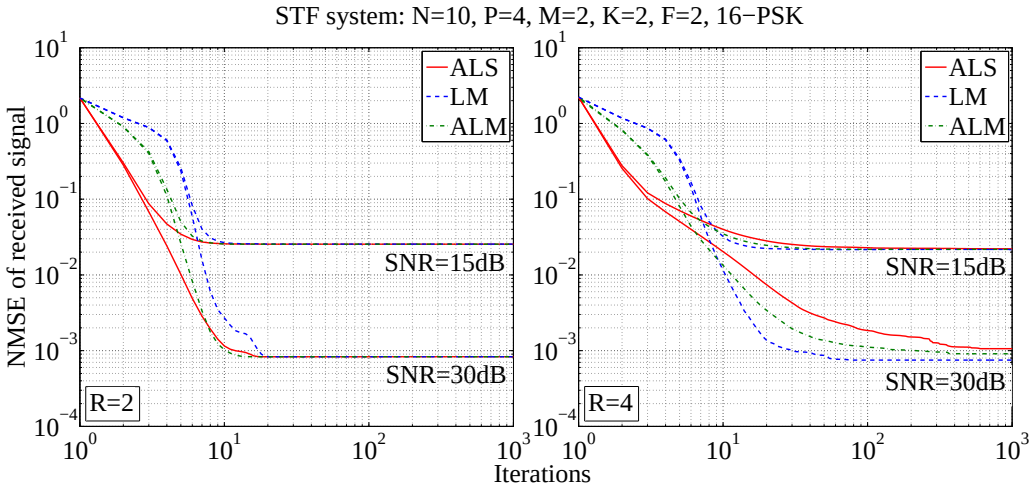


Figure 4.29: STF system: Convergence speed of all algorithms.

around 2 dB and 3 dB for $R = 2$ and $R = 4$, respectively. For a BER equal to 10^{-2}, the gap between KLS-STF and other receivers is approximately 4 dB and more than 9 dB for $R = 2$ and $R = 4$, respectively.

In order to compare the convergence speed of the ALS, LM and ALM algorithms, we have fixed the same total number of iterations for all algorithms and have plotted the NMSE of received signal at each iteration. Figures 4.28 and 4.29 show the NMSE of received signal versus iteration for two values of $R \in \{2, 4\}$ and of SNR (SNR=15dB and SNR=30dB). Observe that all algorithms tend to converge to the same value of the NMSE of received signal estimation constructed by the symbol and channel estimates.

Note also that the number of data streams drastically affects the behavior of the convergence speed of all algorithms for both systems. For smaller values of R ($R = 2$), according to Figures 4.28 and 4.29, the LM and the ALS algorithms respectively have the slowest and the fastest convergence. However, there is an inversion of the convergence behavior when the number of data streams is increased to $R = 4$. The LM becomes faster than the ALS algorithm mainly for

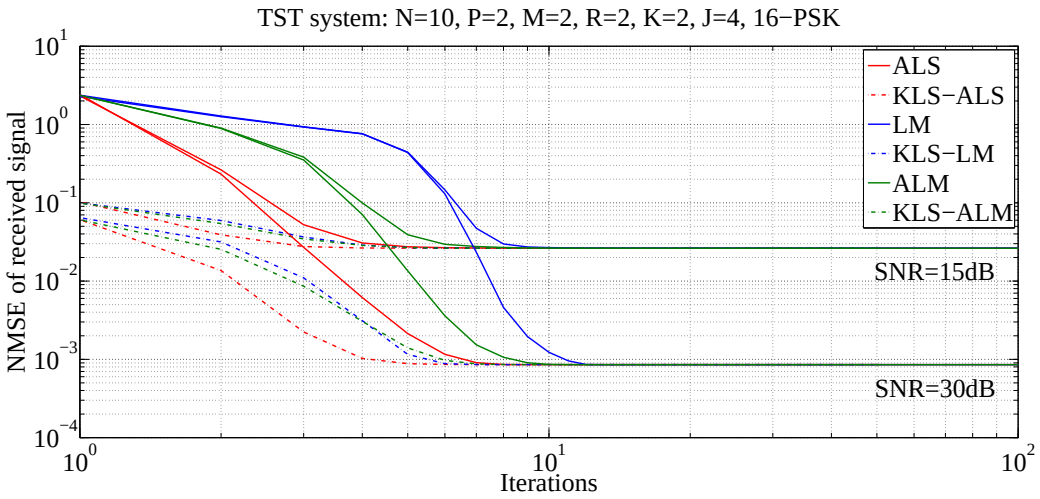


Figure 4.30: TST system: Convergence of all algorithms with initialization based on the KLS method.

higher values of SNR. For SNR=30dB, the LM converges in around 50 and 60 iterations and the ALS needs more than 100 and 600 iterations for the TST and STF systems, respectively. An interesting point is that the ALM algorithm provides convergence curves between both ALS and LM algorithms.

A disadvantage of the ALS algorithm is that it involves the pseudo-inverse calculus and the identifiability of the estimates depends on the unique existence of the pseudo-inverse. The LM is the most complex algorithm and can converge slower than the ALS and ALM for low values of the number of data streams R. Therefore, the ALM is a good option providing a good trade-off between computational complexity and convergence compared to the ALS and LM algorithms.

All previous simulations were obtained from a random initialization of the symbol and channel matrices (or channel tensor). An interesting option to improve the convergence speed of the ALS, LM and ALM algorithms is to employ the KLS method as an initial stage with the purpose of providing a better initialization of the estimates than random initialization.

In Figures 4.30 and 4.31 we have plotted the NMSE of received signal versus iterations in order to evaluate the influence of the initialization based on the KLS method. These initializations take into account the knowledge of only one symbol, remember that this information is already used to eliminate the scalar ambiguities. Thus, it is not required additional information known a priori. According to both figures, we can verify the improvement obtained by the initialization based on the KLS method for both systems.

The employment of the KLS method as initial step always allows to obtain faster convergence and can provide a convergence to a global optimum. It is very important because the ALS, LM and ALM algorithms are strongly dependent on the initialization, can converge very slowly to a global minimum or even a local minimum.

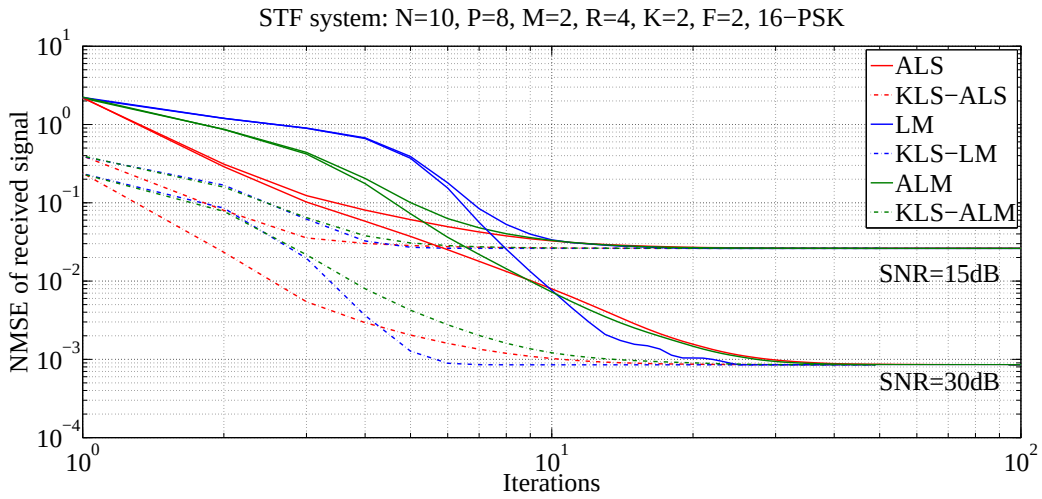


Figure 4.31: STF system: Convergence of all algorithms with initialization based on the KLS method.

4.6.6 Comparison with KRST and STM

In order to evaluate the performance of our proposed ST coding, we compare with two other tensor approaches: Khatri-Rao Space-Time (KRST) and Space-Time Multiplexing (STM) codes proposed, respectively, in [53] and [55]. These approaches rely on tensor decompositions for modeling of the received signals and achieve variable rate-diversity trade-offs for any transmit/receive antenna configuration or signal constellation.

The KRST coding [53] combines a linear spatial precoding with a data spreading only over the time dimension. In [55], a third order tensor coding performs jointly spatial multiplexing and ST coding, leading to an additional spreading over all transmit antennas. Differently to the KRST, the transceivers associated with the TST, STF and STM techniques transmit a linear combination of R data streams composed of N symbols each. It introduces some flexibility at the transceivers by choosing a number of data streams different from the number of transmit antennas. Furthermore, the TST and STF systems provide different degrees of space and time spreading and multiplexing, which depend on the choice of the allocation structures.

In both works [53] and [55], a joint semi-blind channel estimation and symbol detection is afforded thanks to the ALS method. In this sense, the ALS receivers for the TST and STF systems are considered in the next simulations. In order to have an adequate comparison with the proposed system under the same conditions, we fix the same design parameters for all systems and employ the Vandermonde structure for the codes. In [53], the transmitted symbols are precoded by a constellation rotation (CR) matrix. According to the design rule in [53] and by simplicity, this matrix is set equal to the identity matrix for achieving full diversity gain.

Figures 4.32 and 4.33 show the BER versus SNR for QPSK and 16-PSK, respectively. According to these figures, the ALS-TST based receiver outperforms the ALS-STF, ALS-STM and ALS-KRST based receivers, due to its higher diversity gain. The ALS-KRST receiver provides the worst performance of BER, mainly for QPSK constellation. Remark that both TST and STF systems allow to improve even more the symbol recovery by increasing J and F. The KRST and STM codes do not provide this flexibility introduced by extra diversities and the

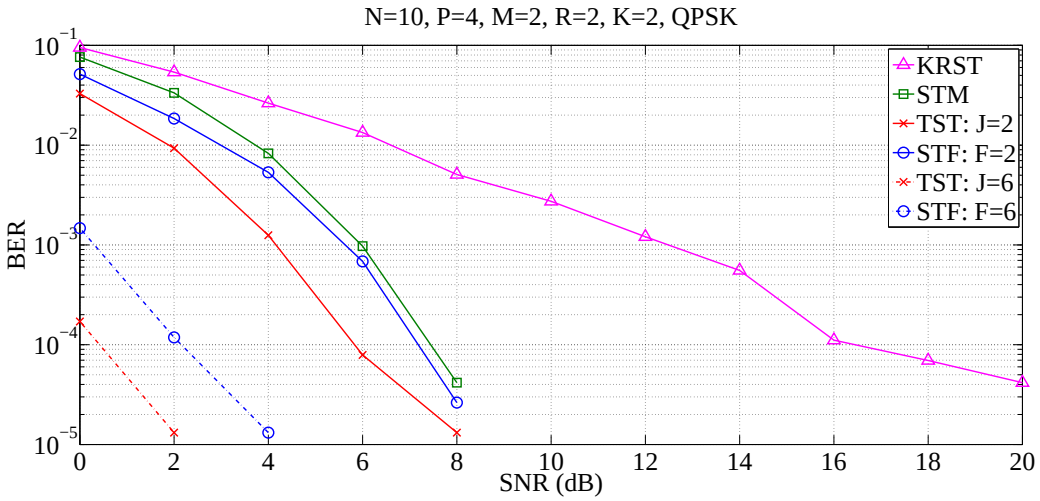


Figure 4.32: Comparison with KRST and STM for QPSK.

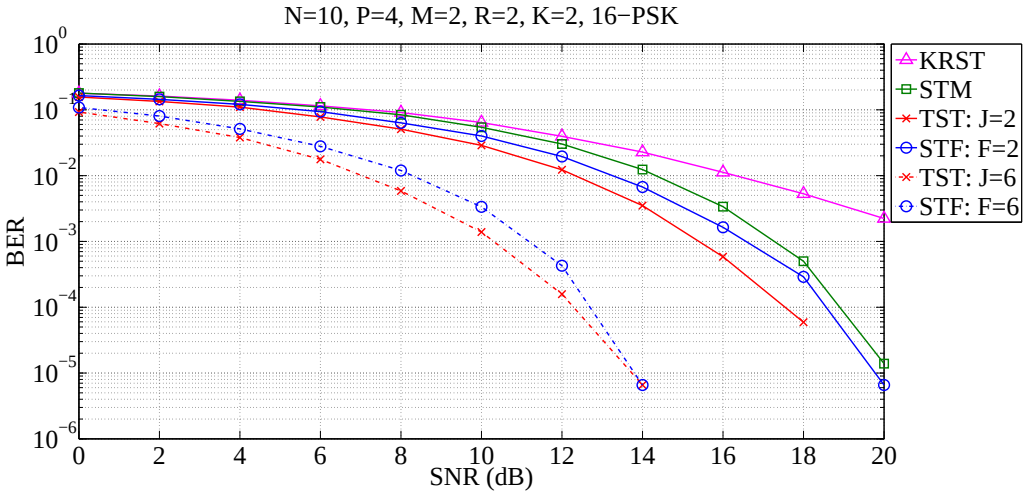


Figure 4.33: Comparison with KRST and STM for 16-PSK.

maximum achievable diversity gain of both codes is $K\min(P,M)$. For $P > M$ and $M = R$, the TST and STF systems can achieve maximum diversity gains JP and FP times more than the ones for the STM and KRST.

Observe that the transmission rate of the KRST system is limited by the number of transmit antennas M because the number of data streams R is forced to be equal to M. Another point that deserves attention is that the uniqueness of symbol and channel estimates for both KRST and STM systems are ensured from the knowledge of one row of symbols. Hence, for $N = 10$, 10% of the transmitted symbols are known at the transceiver and receiver, which leads to 10% of reduction of the transmission rate. We have previously shown that both TST and STF systems require only one symbol to eliminate the scaling ambiguities. In Figures 4.32 and 4.33 for $R = 2$, 5% of the transmitted symbols for the TST and STF are known at the transceiver and receiver, and the transmission rate is reduced 5%. However, this reduction can even be smaller for higher values of R.

4.6.7 Generalization of TST systems to multiuser case

Next, we evaluate the performance of the multiuser TST system with the proposed semi-blind receivers and taking into account the knowledge of only one symbol per user. Let us consider two design configurations as follows:

$$\textit{Case 1:}\begin{cases}\bullet\ \text{Vandermonde structure for } \mathbf{W}_{J\times QRM} : w^{(q)}_{m,r,j} = e^{i2\pi j\frac{(q-1)RM+(r-1)M+m}{QRM}}\\ \bullet\ \mathbf{C}^{(\mathbf{H},q)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S},q)^{\mathrm{T}}} \text{ has no zero-rows for all } q\\ \bullet\ \mathbf{c}^{(\mathbf{S},q)}_{1\cdot} = \mathbf{1}^{\mathrm{T}}_R,\ \mathbf{c}^{(\mathbf{H},q)}_{1\cdot} = \mathbf{1}^{\mathrm{T}}_M,\ \text{for all } q\end{cases} \tag{4.150}$$

and

$$\textit{Case 2:}\begin{cases}\bullet\ \text{Vandermonde structure for } \mathbf{W}^{(q)}_{J\times RM} : w^{(q)}_{m,r,j} = e^{i2\pi j\frac{mr}{RM}}\\ \bullet\ \left[\left(\mathbf{C}^{(\mathbf{H},1)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S},1)^{\mathrm{T}}}\right)^{\mathrm{T}} \cdots \left(\mathbf{C}^{(\mathbf{H},Q)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S},Q)^{\mathrm{T}}}\right)^{\mathrm{T}}\right] \text{ full column-rank}\\ \bullet\ \mathbf{c}^{(\mathbf{S},q)}_{1\cdot} = \mathbf{1}^{\mathrm{T}}_R,\ \mathbf{c}^{(\mathbf{H},q)}_{1\cdot} = \mathbf{1}^{\mathrm{T}}_M,\ \text{for all } q\end{cases}. \tag{4.151}$$

Both configurations represent two different forms to satisfy the uniqueness condition investigated in Chapter 3, being resumed in Table 3.1. For Case 1, the code tensor for each user is chosen such that $\mathbf{W}_{J\times QRM}$ is full column-rank, implying $J \geq QRM$. The allocation matrices for each user are randomly chosen such that the k-rank of $\left[\left(\mathbf{C}^{(\mathbf{H},1)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S},1)^{\mathrm{T}}}\right)^{\mathrm{T}} \cdots \left(\mathbf{C}^{(\mathbf{H},Q)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S},Q)^{\mathrm{T}}}\right)^{\mathrm{T}}\right]$ is at least equal to 1.

For Case 2, the same code tensor is fixed for all users, i.e. $\mathcal{W}^{(1)} = \cdots = \mathcal{W}^{(Q)}$, which leads to $k\ (\mathbf{W}_{J\times QRM}) = 1$. The allocation matrices are randomly chosen such that $\left[\left(\mathbf{C}^{(\mathbf{H},1)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S},1)^{\mathrm{T}}}\right)^{\mathrm{T}} \cdots \left(\mathbf{C}^{(\mathbf{H},Q)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S},Q)^{\mathrm{T}}}\right)^{\mathrm{T}}\right]$ is full column-rank, implying $P \geq QRM$.

We show in Figures 4.34 and 4.35 the BER performance for all proposed receivers considering two and four users in the transmission. In Figure 4.34 the BER is computed by averaging the results obtained for all users. We present this average BER for all users in the red curves and the BER results for each user in the blue curves in Figure 4.35.

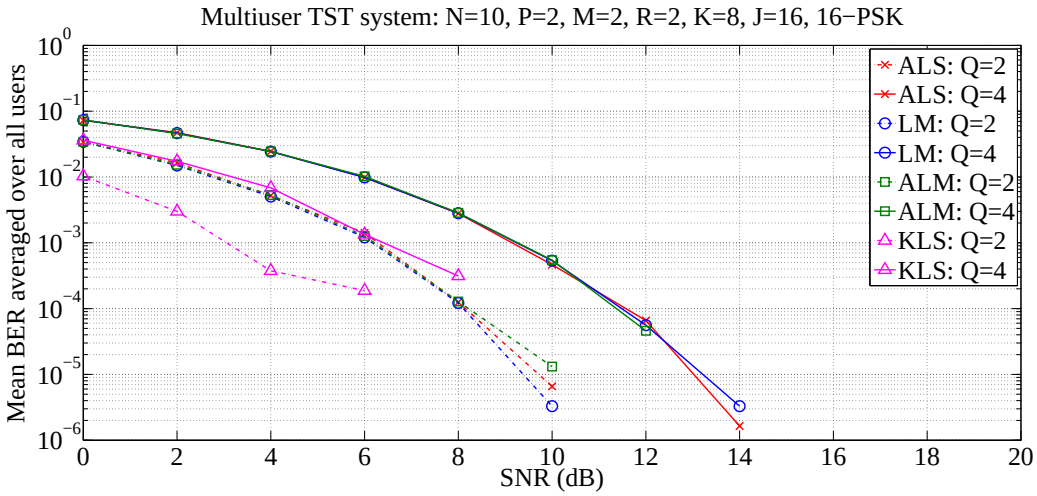


Figure 4.34: Multiuser TST system: Performance of all receivers for *Case 1*. For better visualization, the BER curves for each user are separately plotted for each algorithm in Fig. 4.35.

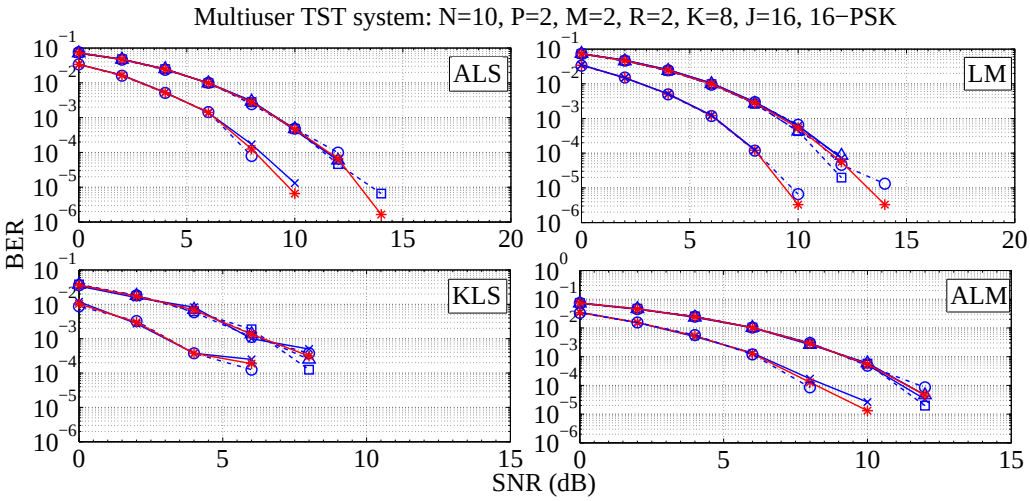


Figure 4.35: Multiuser TST system: Performance of all receivers for *Case 1*, (–): averaged over all users and (–): each user. The red curves are also plotted in Fig. 4.34 with the purpose of comparing all algorithms.

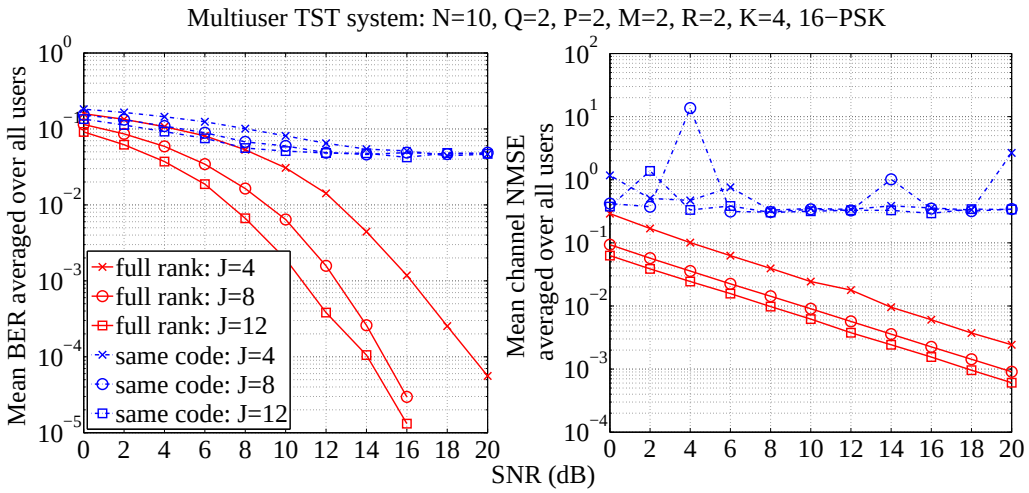


Figure 4.36: ALS-TST receiver: Condition on the code tensor for *Case 1*.

From these figures, we can verify that the transmitted symbols for each user are correctly and equally recovered through the knowledge of one symbol per user. As shown for the TST and STF systems with one user, the performance of ALS, ALM and LM are closed. Interestingly, the KLS algorithm provides the best symbol estimation, which can be justified by the large value of the spreading code length ($J = 16$). In the sense that the channel estimate is computed taking into account more versions of the received signal for a fixed number of parameters to be estimated.

We show two plots in order to analyze the proposed uniqueness conditions for multiuser TST systems. In Figure 4.36 we plot the BER and channel NMSE versus SNR for three values of the spreading code length ($J \in \{4, 8, 12\}$), and in Figure 4.37 we have the same plot for three values of the number of data blocks ($P \in \{4, 8, 12\}$) instead.

In Figure 4.36 the allocation matrices for all users are chosen according to *Case 1*. However, the tensor code is set such that $\mathbf{W}_{J\times QRM}$ is full rank for the red curves and $\mathcal{W}^{(1)} = \ldots = \mathcal{W}^{(Q)}$ for the blue curves, i.e. we employ the same code tensor for each user. When all users employ the same code tensor, it leads to $k(\mathbf{W}_{J\times QRM}) = 1$. Hence the proposed uniqueness conditions are not satisfied if $J < QRM$ or $\mathcal{W}^{(1)} = \ldots = \mathcal{W}^{(Q)}$. However, we can observe from Figure 4.36 that the channel coefficients and all transmitted symbols are correctly estimated even when $\mathbf{W}_{J\times QRM}$ is full rank with $J < QRM$.

We employ in Figure 4.37 the same code tensor for all users in accordance with *Case 2*. For the red curves, the allocation matrices for all users are set according to *Case 2* and contrarily for the blue curves, we do not impose that $\left[\left(\mathbf{C}^{(\mathbf{H},1)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S},1)^{\mathrm{T}}}\right)^{\mathrm{T}} \ \ldots \ \left(\mathbf{C}^{(\mathbf{H},Q)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S},Q)^{\mathrm{T}}}\right)^{\mathrm{T}}\right]$ is full rank. Figure 4.37 shows that we can correctly estimate the channel coefficients and all transmitted symbols even when $\left[\left(\mathbf{C}^{(\mathbf{H},1)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S},1)^{\mathrm{T}}}\right)^{\mathrm{T}} \ \ldots \ \left(\mathbf{C}^{(\mathbf{H},Q)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S},Q)^{\mathrm{T}}}\right)^{\mathrm{T}}\right]$ is full rank with $P < QRM$. However, random allocation structures can not ensure the uniqueness specially when $P \leq QRM$.

From both figures 4.36 and 4.37, we can observe that when $\mathbf{W}_{J\times QRM}$ or $\left[\left(\mathbf{C}^{(\mathbf{H},1)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S},1)^{\mathrm{T}}}\right)^{\mathrm{T}} \ \ldots \ \left(\mathbf{C}^{(\mathbf{H},Q)^{\mathrm{T}}} \diamond \mathbf{C}^{(\mathbf{S},Q)^{\mathrm{T}}}\right)^{\mathrm{T}}\right]$ is full rank, the channel and symbols are appropriately recovered even with $J < QRM$ or $P < QRM$. We also note that when the code and allocation structures

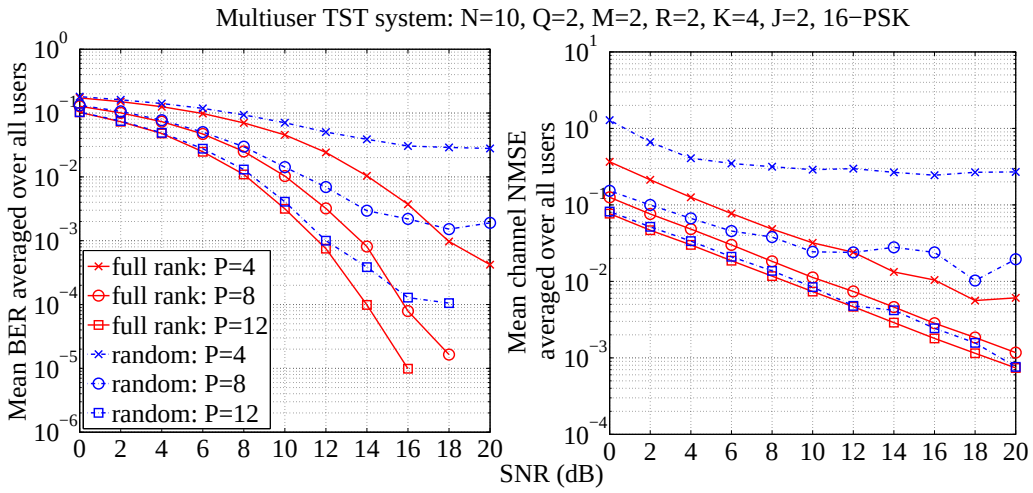


Figure 4.37: ALS-TST receiver: Condition on the allocation matrices for *Case 2*.

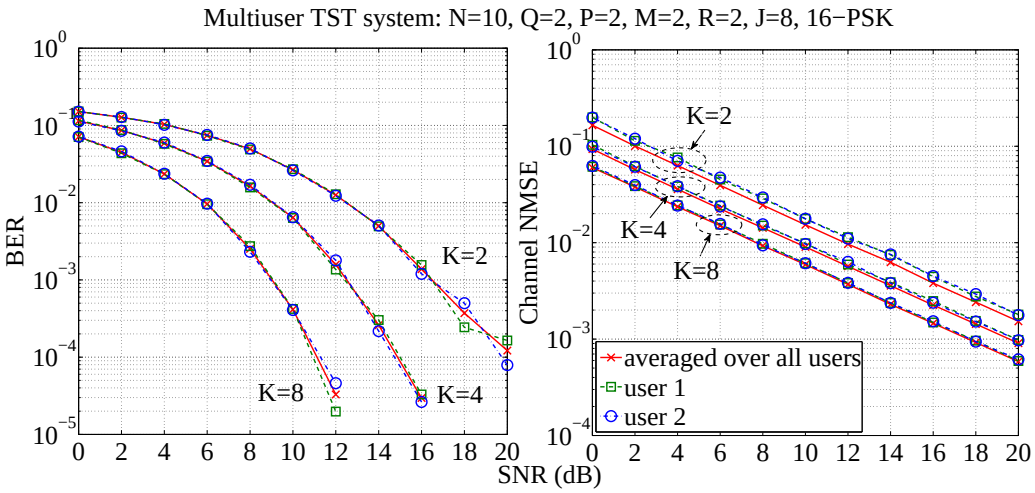


Figure 4.38: ALS-TST receiver: Influence of the number of receive antennas for *Case 1*.

satisfy the uniqueness conditions investigated in Chapter 3, the channel and symbol estimations present better performances.

According to the identifiability condition for the ALS algorithm, we have indicated to set $K \geq QM$ which means that the number of receive antennas has to be at least equal to the total number of transmit antennas for all users. Hence the number of available receive antennas can restrict the total number of users. Nevertheless, we can verify from Figure 4.38 that this condition on the number of receive antennas is not necessary. The symbol and channel estimations can be still achieved when $K < QM$.

Chapter 5

Conclusion and Perspectives

A new tensor decomposition, PARATUCK-(N_1, N), is introduced in this thesis, which generalizes the well-known PARATUCK-2 model. The uniqueness condition have been derived for our proposed model. The generalization of two lemmas [67, 68] concerning the Khatri-Rao product has been derived and employed to deduce the uniqueness and identifiability results.

We have proposed a new tensor space-time coding for MIMO wireless communication systems. The associated transceiver is characterized by a third-order code tensor and two allocation matrices that allow space-time spreading-multiplexing of the transmitted symbols. The proposed transmission system can be viewed as an extension of the ST transmission system of [107] that relies on the PARATUCK-2 tensor model for the received signals. This extension is derived from the introduction of an extra time diversity.

A performance analysis of the TST system is deduced with the purpose of evaluating the diversity of information transmitted, which allows us to express a maximum diversity gain in terms of some system parameters and taking into account the structures of antenna and data stream allocations per block. This performance analysis has been extended for a space-time-frequency (STF) system [62] and the maximum diversity gain has also been achieved. The uniqueness conditions of a generalized PARATUCK-2 model is also established for the STF system. A comparison between the TST and STF systems has been presented in an unified way in terms of diversity gain, identifiability and uniqueness conditions.

We have observed from the diversity gain analysis that systems with different allocation structures can provide different performances and that the diversity gain for both systems depends on the code and mainly on the allocation structures. We can obtain different performances for the symbol estimation employing the same diversity gain provided by the extra time and frequency diversities of the TST and STF systems due to the difference of the number of system parameters to estimate.

Semi-blind receivers have been proposed based on the ALS, LM and ALM algorithms for the TST and STF systems. The identifiability conditions for the ALS algorithm have been derived for both systems. The ALS, ALM and LM algorithms for both systems provide the same BER performance at the convergence. Thus, the difference between these algorithms is basically concerning the complexity and convergence speed of the algorithms. We have shown that the behavior of theses algorithms is affected by the variation of the number of data streams.

Despite a higher computational complexity, the advantage of both LM and ALM algorithms is that identifiability conditions are not required as happens with the ALS. We have observe that the ALM algorithm provides curves of convergence and complexity between the curves for the ALS and LM, leading to a good trade-off.

The uniqueness and identifiability conditions established in this work have been analyzed in terms of the design parameters of both systems, taking into account the number of data blocks and of receive antennas. The flexibility of the uniqueness condition for the TST system allows to estimate the symbols even when $P < MR$ by adjusting $J \geq MR$, differently to the STF system. Furthermore, the transmission rate for the STF system is consequently limited by the number of transmit antennas. Thus, the advantage of the TST system is that the symbol recovery can be ensured without affecting the transmission rate. Another relevant conclusion corroborated by our simulations is that both systems can perfectly estimate the transmitted symbols even when there are more transmit antennas than the receive antennas, i.e. $K < M$.

According to our performance analysis, TST coding increases the maximum diversity gain. The introduction of one extra time diversity J via the third mode of the code tensor induces a significant performance improvement in terms of symbol recovery and channel detection comparatively to existing tensor-based solutions such as: KRST [53], STM [55] and ST-PARATUCK-2 [59], as illustrated by means of simulation results.

A direct non-iterative receiver, herein referred to as KLS, is proposed for the TST and STF systems based on the structure of the Kronecker product and assuming a priori knowledge of some transmitted symbols. As expected, the knowledge of more than one symbol improves the channel estimation and consequently, the symbol recovery. The disadvantage of these algorithms based on the Kronecker structure is that the channel estimation depends on a priori information known at the receiver and a poor channel estimation leads to a poor symbol recovery.

The KLS receiver proposed for the TST system is interesting because can provide a BER performance close to the performances obtained with the receivers based on the ALS, LM and ALM algorithms or even a superior performance for high values of the spreading code length J. An increase of the extra time diversity allows to improve the BER performance and for the STF system, an increase of the extra frequency diversity leads to an increase of the channel coefficients to estimate which impairs the channel and symbol estimation. The identifiability conditions for the KLS algorithm are equivalents to the uniqueness conditions proposed in this thesis. One of the main advantages of the KLS algorithm is that it provides low computational complexity.

The performance of all algorithms based on the ALS and LM methods depends on the algorithm initialization, the convergence speed can be strongly affected and further these algorithms can not converge to a global optimum. The KLS method showed interesting results as an alternative procedure to initialize algorithms in order to accelerate the convergence speed ensuring an initialization more closed to the optimal solution. Moreover, this method can be exploited for different tensor approaches.

We have proposed an uplink processing based on the TST coding with allocation resources and derived semi-blind receivers from the ALS, LM, ALM and KLS methods. The main advantage is that it is possible to perfectly recovery the transmitted symbols for all users and channel

coefficients by the knowledge of only one symbol per user without permutation ambiguity. The uniqueness conditions for multiuser TST systems have been established analogously to the case with one user. Even if the proposed conditions are only sufficient, we have verified that these conditions always ensure the uniqueness of the estimates.

Some perspectives of this thesis can be highlighted as follows:

- In this thesis, the wireless communication channel is modeled by a random attenuation of the transmitted signal, followed by additive noise, which relies on an instantaneous MIMO channel. An interesting generalization would consist in considering a more complex situation with multipath propagation and convolutive channel in order to render the channel model more realistic analogously to [110, 61]. It would lead to a different tensor decomposition implying the study of the uniqueness conditions for this new model.

- From some simulations and the performance analysis developed in Chapter 3, we observe that the allocation structures can provide different performances in terms of channel estimation and symbol recovery. It instigates to derive an optimal structure for the allocation of transmit antennas and of data streams at each time block in order to achieve a performance optimization. We also could evaluate the multiplexing gain for the TST coding and derive a tradeoff between both multiplexing and diversity gains, in the sense that diversity and multiplexing gains tend to provide low error rates and high transmission rates, respectively.

- A natural extension of this work is to consider the problem of symbol recovery without the knowledge of allocation and/or code structures at the receiver in military applications. This extension would involve the study of new uniqueness conditions for ensuring the identifiability of channel and symbol, properly eliminating the ambiguities.

- Use of the PT-(N_1, N) decomposition for modeling other practical applications, which would allow to exploit the uniqueness results derived in this thesis.

- The development of new receivers based on the KLS method is an interesting topic for future research. This procedure can be employed as an initial stage of algorithms or can be combine with other methods for estimating system parameters, which would allow to accelerate the convergence speed and/or to ensure the convergence to an optimal solution. The KLS receiver can be also improved by introducing of an orthogonal code tensor to replace the precoding with the allocation structures.

- An interesting generalization would consist in deducing new coding structures based on the TST and STF systems with the purpose of exploiting the angular diversity from the employ of directional antennas. Directional antennas have been used in advanced systems to optimize and maximize transmission/ reception in some directions.

Appendix A

Basic properties

The Kronecker, column-wise Kronecker (called Khatri-Rao), partition-wise Kronecker and Hadamard products are denoted by $\otimes$, $\diamond$, $|\otimes|$ and $\odot$, respectively. We have the following definitions and properties:

$$\mathbf{A} \diamond \mathbf{C} \triangleq \left[\begin{array}{ccc} \mathbf{A}_{\cdot 1} \otimes \mathbf{C}_{\cdot 1} & \ldots & \mathbf{A}_{\cdot R} \otimes \mathbf{C}_{\cdot R} \end{array} \right] = \left[\begin{array}{c} \mathbf{C} D_1(\mathbf{A}) \\ \vdots \\ \mathbf{C} D_I(\mathbf{A}) \end{array} \right] \in \mathbb{C}^{IS \times R}, \tag{A.1}$$

$$\tilde{\mathbf{A}} \,|\otimes|\, \tilde{\mathbf{B}} \triangleq \left[\begin{array}{ccc} \mathbf{A}^{(1)} \otimes \mathbf{B}^{(1)} & \ldots & \mathbf{A}^{(Q)} \otimes \mathbf{B}^{(Q)} \end{array} \right] \in \mathbb{C}^{IJ \times QRS}, \tag{A.2}$$

$$\mathrm{vec}(\mathbf{B}\mathbf{C}\mathbf{A}^{\mathrm{T}}) = (\mathbf{A} \otimes \mathbf{B})\,\mathrm{vec}(\mathbf{C}) \in \mathbb{C}^{\mathrm{IJ} \times 1}, \tag{A.3}$$

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{E} \otimes \mathbf{F}) = (\mathbf{A}\mathbf{E} \otimes \mathbf{B}\mathbf{F}) \in \mathbb{C}^{IJ \times KL}, \tag{A.4}$$

$$(\mathbf{A} \otimes \mathbf{B})(\tilde{\mathbf{E}} \,|\otimes|\, \tilde{\mathbf{F}}) = (\mathbf{A}\tilde{\mathbf{E}} \,|\otimes|\, \mathbf{B}\tilde{\mathbf{F}}) \in \mathbb{C}^{IJ \times QKL}, \tag{A.5}$$

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{H} \diamond \mathbf{F}) = (\mathbf{A}\mathbf{H} \diamond \mathbf{B}\mathbf{F}) \in \mathbb{C}^{IJ \times L}, \tag{A.6}$$

$$(\mathbf{A} \odot \mathbf{G})_{i,r} \triangleq a_{i,r}\, g_{i,r}, \tag{A.7}$$

for $\mathbf{A}, \mathbf{G} \in \mathcal{C}^{I \times R}$, $\mathbf{B} \in \mathcal{C}^{J \times S}$ and $\mathbf{C} \in \mathcal{C}^{S \times R}$, $\mathbf{E} \in \mathcal{C}^{R \times K}$, $\mathbf{F} \in \mathcal{C}^{S \times L}$, $\mathbf{H} \in \mathcal{C}^{R \times L}$, $\tilde{\mathbf{A}} = \left[\begin{array}{ccc} \mathbf{A}^{(1)} & \cdots & \mathbf{A}^{(Q)} \end{array} \right] \in \mathcal{C}^{I \times QR}$ and $\tilde{\mathbf{B}} = \left[\begin{array}{ccc} \mathbf{B}^{(1)} & \cdots & \mathbf{B}^{(Q)} \end{array} \right] \in \mathcal{C}^{J \times QS}$, $\tilde{\mathbf{E}} = \left[\begin{array}{ccc} \mathbf{E}^{(1)} & \cdots & \mathbf{E}^{(Q)} \end{array} \right] \in \mathcal{C}^{R \times QK}$ and $\tilde{\mathbf{F}} = \left[\begin{array}{ccc} \mathbf{F}^{(1)} & \cdots & \mathbf{F}^{(Q)} \end{array} \right] \in \mathcal{C}^{S \times QL}$.

Appendix B

Generalization of STF system to multiuser case

Analogously to the TST case, we can generalize the STF MIMO wireless communication system for Q users in the transmission. Each user transmits R input data streams using M different antennas. We consider a precoding matrix $\mathbf{W}^{(q)}$ and two allocation tensors for each user q: the stream allocation tensor $\mathcal{C}^{(\mathbf{S},q)} \in \mathbb{R}^{F\times P\times R}$ and the antenna allocation tensor $\mathcal{C}^{(\mathcal{H},q)} \in \mathbb{R}^{F\times P\times M}$.

We can write the signal associated with the q-th user and the f-th subcarrier transmitted from the m-th antenna using the n-th symbol period of the p-th block by

$$u^{(q)}_{f,m,n,p} = \sum_{r=1}^{R} w^{(q)}_{m,r}\, s^{(q)}_{n,r}\, c^{(\mathcal{H},q)}_{f,p,m}\, c^{(\mathbf{S},q)}_{f,p,r} = \sum_{r=1}^{R} t^{(q)}_{f,m,r,p}\, s^{(q)}_{n,r} \tag{B.1}$$

with

$$t^{(q)}_{f,m,r,p} \triangleq w^{(q)}_{m,r}\, c^{(\mathcal{H},q)}_{f,p,m}\, c^{((\mathbf{S},q)}_{f,p,r}. \tag{B.2}$$

In the noiseless case of scattering-rich multipath fading channel, the received signal associated with the q-th user can be written using (B.1) as

$$x^{(q)}_{f,k,n,p} = \sum_{m=1}^{M} h^{(q)}_{f,k,m}\, u^{(q)}_{f,m,n,p} = \sum_{m=1}^{M}\sum_{r=1}^{R} t^{(q)}_{f,m,r,p}\, h^{(q)}_{f,k,m}\, s^{(q)}_{n,r}. \tag{B.3}$$

From (3.28), we can write the overall received signal tensor $\mathcal{X}^{(f)} \in \mathbb{C}^{K\times N\times P}$ associated with the f-th subcarrier by summing all received signals as

$$\mathcal{X}^{(f)} = \sum_{q=1}^{Q} \mathcal{X}^{(f,q)} = \sum_{q=1}^{Q} \mathcal{T}^{(f,q)} \times_1 \mathbf{H}^{(f,q)} \times_2 \mathbf{S}^{(q)} \tag{B.4}$$

where $\mathcal{X}^{(f,q)} \triangleq \mathcal{X}^{(q)}_{f\cdots} \in \mathbb{C}^{K\times N\times P}$, $\mathcal{T}^{(f,q)} \triangleq \mathcal{T}^{(q)}_{f\cdots} \in \mathbb{C}^{M\times R\times P}$ and $\mathbf{H}^{(f,q)} \triangleq \mathbf{H}^{(q)}_{f\cdot\cdot} \in \mathbb{C}^{K\times M}$.

From (3.32) and (3.38), we can express three matrix unfoldings of $\mathcal{X}$ as follows

$$\mathbf{X}_{PFK\times N} \triangleq \sum_{q=1}^{Q} \mathbf{X}^{(q)}_{PFK\times N} = \left(\mathbf{\Omega}_P \,|\otimes|\, \tilde{\mathbf{H}}_{FK\times QFM}\right) \mathbf{T}_{QPFM\times QR}\, \mathbf{S}_{QR\times N} \in \mathbb{C}^{PFK\times N}, \tag{B.5}$$

$$\mathbf{X}_{PFN\times K} \triangleq \sum_{q=1}^{Q} \mathbf{X}^{(q)}_{PFN\times K} = (\mathbf{\Omega}_{PF} \,|\otimes|\, \mathbf{S}_{N\times QR})\, \mathbf{T}_{QPFR\times QFM}\, \mathbf{H}_{QFM\times K} \in \mathbb{C}^{PFN\times K}, \tag{B.6}$$

$$\mathbf{X}_{KN\times FP} \triangleq \sum_{q=1}^{Q} \mathbf{X}^{(q)}_{KN\times FP} = (\mathbf{H}_{K\times QFM} \,|\otimes|\, \mathbf{S}_{N\times QR})\, \mathbf{T}_{QFMR\times FP} \in \mathbb{C}^{KN\times FP}, \tag{B.7}$$

with $\mathbf{S}_{N\times QR} \triangleq \begin{bmatrix}\mathbf{S}^{(1)} & \cdots & \mathbf{S}^{(Q)}\end{bmatrix} \in \mathbb{C}^{N\times QR}$, $\mathbf{H}_{K\times QFM} \triangleq \begin{bmatrix}\mathbf{H}^{(1)}_{K\times FM} & \cdots & \mathbf{H}^{(Q)}_{K\times FM}\end{bmatrix} \in \mathbb{C}^{K\times QFM}$, $\mathbf{\Omega}_Z \triangleq \mathbf{1}_Q^{\mathrm{T}} \otimes \mathbf{I}_Z \in \mathbb{C}^{Z\times QZ}$,

$$\tilde{\mathbf{H}}_{FK\times QFM} = \begin{bmatrix}\mathrm{bdiag}\Big(\mathbf{H}^{(1)}_{1\cdot\cdot}, \ldots, \mathbf{H}^{(1)}_{F\cdot\cdot}\Big) & \cdots & \mathrm{bdiag}\Big(\mathbf{H}^{(Q)}_{1\cdot\cdot}, \ldots, \mathbf{H}^{(Q)}_{F\cdot\cdot}\Big)\end{bmatrix} \in \mathcal{C}^{FK\times QFM}, \tag{B.8}$$

$$\mathbf{T}_{QPFM\times QR} \triangleq \underset{\in \mathbb{C}^{QPFM\times QR}}{\begin{bmatrix}\mathbf{T}^{(1)}_{PFM\times R} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{T}^{(Q)}_{PFM\times R}\end{bmatrix}}, \quad \mathbf{T}_{QPFR\times QFM} \triangleq \underset{\in \mathbb{C}^{QPFR\times QFM}}{\begin{bmatrix}\mathbf{T}^{(1)}_{PFR\times FM} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{T}^{(Q)}_{PFR\times FM}\end{bmatrix}}, \tag{B.9}$$

$$\mathbf{T}_{QFMR\times FP} \triangleq \begin{bmatrix}{\mathbf{T}^{(1)}_{FMR\times FP}}^{\mathrm{T}} & \cdots & {\mathbf{T}^{(Q)}_{FMR\times FP}}^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}} \in \mathbb{C}^{QFMR\times FP}. \tag{B.10}$$

Each matrix $\mathbf{T}^{(q)}_{FMR\times FP}$ can be written using (3.39) and (3.40) as

$$\mathbf{T}^{(q)}_{FMR\times FP} \triangleq \underset{\in \mathbb{C}^{FMR\times FP}}{\begin{bmatrix}\Big[\mathrm{vec}\Big({\mathbf{T}^{(q)}_{1\cdot\cdot 1}}^{\mathrm{T}}\Big) \cdots \mathrm{vec}\Big({\mathbf{T}^{(q)}_{1\cdot\cdot P}}^{\mathrm{T}}\Big)\Big] & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \Big[\mathrm{vec}\Big({\mathbf{T}^{(q)}_{F\cdot\cdot 1}}^{\mathrm{T}}\Big) \cdots \mathrm{vec}\Big({\mathbf{T}^{(q)}_{F\cdot\cdot P}}^{\mathrm{T}}\Big)\Big]\end{bmatrix}} \tag{B.11}$$

with

$$\Big[\mathrm{vec}\Big({\mathbf{T}^{(q)}_{f\cdot\cdot 1}}^{\mathrm{T}}\Big) \cdots \mathrm{vec}\Big({\mathbf{T}^{(q)}_{f\cdot\cdot P}}^{\mathrm{T}}\Big)\Big]^{\mathrm{T}} = \mathrm{vec}\Big(\mathbf{W}^{(q)\mathrm{T}}\Big)^{\mathrm{T}} \diamond \Big(\mathbf{C}^{(\mathcal{H},q)\mathrm{T}}_{f\cdot\cdot} \diamond \mathbf{C}^{(\mathbf{S},q)\mathrm{T}}_{f\cdot\cdot}\Big)^{\mathrm{T}} \in \mathbb{C}^{P\times MR}. \tag{B.12}$$

Applying (B.11) and (B.12) to (B.10), we can rewrite $\mathbf{T}_{QFMR\times FP}$ as

$$\mathbf{T}_{QFMR\times FP} = \left[\begin{array}{ccc}\Big(\mathrm{vec}\Big(\mathbf{W}^{(1)\mathrm{T}}\Big)^{\mathrm{T}} \diamond \Big(\mathbf{C}^{(\mathcal{H},1)\mathrm{T}}_{1\cdot\cdot} \diamond \mathbf{C}^{(\mathbf{S},1)\mathrm{T}}_{1\cdot\cdot}\Big)^{\mathrm{T}}\Big)^{\mathrm{T}} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \Big(\mathrm{vec}\Big(\mathbf{W}^{(1)\mathrm{T}}\Big)^{\mathrm{T}} \diamond \Big(\mathbf{C}^{(\mathcal{H},1)\mathrm{T}}_{F\cdot\cdot} \diamond \mathbf{C}^{(\mathbf{S},1)\mathrm{T}}_{F\cdot\cdot}\Big)^{\mathrm{T}}\Big)^{\mathrm{T}} \\ \hdashline & \vdots & \\ \hdashline \Big(\mathrm{vec}\Big(\mathbf{W}^{(Q)\mathrm{T}}\Big)^{\mathrm{T}} \diamond \Big(\mathbf{C}^{(\mathcal{H},Q)\mathrm{T}}_{1\cdot\cdot} \diamond \mathbf{C}^{(\mathbf{S},Q)\mathrm{T}}_{1\cdot\cdot}\Big)^{\mathrm{T}}\Big)^{\mathrm{T}} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \Big(\mathrm{vec}\Big(\mathbf{W}^{(Q)\mathrm{T}}\Big)^{\mathrm{T}} \diamond \Big(\mathbf{C}^{(\mathcal{H},Q)\mathrm{T}}_{F\cdot\cdot} \diamond \mathbf{C}^{(\mathbf{S},Q)\mathrm{T}}_{F\cdot\cdot}\Big)^{\mathrm{T}}\Big)^{\mathrm{T}}\end{array}\right]. \tag{B.13}$$

Uniqueness analysis

The uniqueness conditions for the STF system with multiple users can be studied in accordance with Theorem 3.8 for the TST system. Let us consider the unfolded matrix $\mathbf{X}_{KN\times FP}$ defined in (B.7) to analyze the uniqueness of this model.

From the selection matrix $\tilde{\mathbf{\Phi}}_{QFMR} \in \mathbb{C}^{QFMQR\times QFMR}$ defined in (3.90), we can relate the Kronecker and partition-wise Kronecker product by

$$\mathbf{H}_{K\times QFM} \,|\otimes|\, \mathbf{S}_{N\times QR} = (\mathbf{H}_{K\times QFM} \otimes \mathbf{S}_{N\times QR})\, \tilde{\mathbf{\Phi}}_{QFMR}, \tag{B.14}$$

where $\tilde{\mathbf{\Phi}}_{QFMR}$ selects $QFMR$ columns of $\mathbf{H}_{K\times QFM} \otimes \mathbf{S}_{N\times QR} \in \mathbb{C}^{KN\times QFMQR}$.

Applying (B.14) to (B.7), it gets

$$\mathbf{X}_{KN\times FP} = (\mathbf{H}_{K\times QFM} \,|\otimes|\, \mathbf{S}_{N\times QR})\, \mathbf{T}_{QFMR\times FP} = (\mathbf{H}_{K\times QFM} \otimes \mathbf{S}_{N\times QR})\, \tilde{\mathbf{\Phi}}_{QFMR} \mathbf{T}_{QFMR\times FP}. \tag{B.15}$$

Considering $\hat{\mathbf{S}}_{N\times QR} = \begin{bmatrix} \hat{\mathbf{S}}^{(1)} & \cdots & \hat{\mathbf{S}}^{(Q)} \end{bmatrix}$ and $\hat{\mathbf{H}}_{K\times QFM} = \begin{bmatrix} \hat{\mathbf{H}}^{(1)}_{K\times FM} & \cdots & \hat{\mathbf{H}}^{(Q)}_{K\times FM} \end{bmatrix}$ as alternative solutions that satisfy (B.7), we can write $\hat{\mathbf{S}}_{N\times QR} = \mathbf{S}_{N\times QR}\, \mathbf{U}$ and $\hat{\mathbf{H}}_{K\times QFM} = \mathbf{H}_{K\times QFM}\, \mathbf{V}$, with

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}^{(1,1)} & \cdots & \mathbf{U}^{(1,Q)} \\ \vdots & \ddots & \vdots \\ \mathbf{U}^{(Q,1)} & \cdots & \mathbf{U}^{(Q,Q)} \end{bmatrix} \in \mathbb{C}^{QR\times QR}, \qquad \mathbf{V} = \begin{bmatrix} \mathbf{V}^{(1,1)} & \cdots & \mathbf{V}^{(1,Q)} \\ \vdots & \ddots & \vdots \\ \mathbf{V}^{(Q,1)} & \cdots & \mathbf{V}^{(Q,Q)} \end{bmatrix} \in \mathbb{C}^{QFM\times QFM} \tag{B.16}$$

non-singular. From the Kronecker property (A.5), $\mathbf{X}_{KN\times FP}$ can be rewritten using as

$$(\mathbf{H}_{K\times QFM} \otimes \mathbf{S}_{N\times QR})\, (\mathbf{V} \,|\otimes|\, \mathbf{U})\, \mathbf{T}_{QFMR\times FP} = (\mathbf{H}_{K\times QFM} \otimes \mathbf{S}_{N\times QR})\, \tilde{\mathbf{\Phi}}_{QFMR}\, \mathbf{T}_{QFMR\times FP}. \tag{B.17}$$

Theorem B.1. *Suppose that $\mathbf{S}_{N\times QR}$ and $\mathbf{H}_{K\times QFM}$ are full column-rank, and the perfect knowledge of the code matrix $\mathbf{W}^{(q)}$ and the allocation tensors $\mathcal{C}^{(\mathbf{S},q)}$ and $\mathcal{C}^{(\mathcal{H},q)}$ for all users. If we choose $\mathbf{W}^{(q)}$ and $\mathcal{C}^{(\mathbf{S},q)}$ and $\mathcal{C}^{(\mathcal{H},q)}$ such that $w^{(q)}_{m,r} \neq 0$ for all $m \in \{1,...,M\}$, $r \in \{1,...,R\}$ and $q \in \{1,...,Q\}$, and $\left[\left(\mathbf{C}^{(\mathcal{H},1)^{\mathrm{T}}}_{f..} \diamond \mathbf{C}^{(\mathbf{S},1)^{\mathrm{T}}}_{f..} \right)^{\mathrm{T}} \cdots \left(\mathbf{C}^{(\mathcal{H},Q)^{\mathrm{T}}}_{f..} \diamond \mathbf{C}^{(\mathbf{S},Q)^{\mathrm{T}}}_{f..} \right)^{\mathrm{T}} \right]$ full column-rank for all $f \in \{1,...,F\}$ implying $P \geq QMR$, then we can uniquely estimate $\mathbf{S}_{N\times QR}$ and $\mathbf{H}_{K\times QFM}$ up to a scalar factor α, i.e.*

$$\mathbf{S}^{(q)} = \alpha\, \hat{\mathbf{S}}^{(q)}, \qquad \mathbf{H}^{(q)}_{K\times FM} = \frac{1}{\alpha}\, \hat{\mathbf{H}}^{(q)}_{K\times FM}. \tag{B.18}$$

Proof: If $\mathbf{S}_{N\times QR}$, $\mathbf{H}_{K\times QFM}$ and $\mathbf{T}^{\mathrm{T}}_{QFMR\times FP}$ are full column-rank, then (B.17) can be rewritten as

$$\begin{aligned} \mathbf{V} \,|\otimes|\, \mathbf{U} &= \tilde{\mathbf{\Phi}}_{QFMR} \qquad \Longrightarrow \\ \begin{bmatrix} \mathbf{V}^{(1,q)} \\ \vdots \\ \mathbf{V}^{(Q,q)} \end{bmatrix} \otimes \begin{bmatrix} \mathbf{U}^{(1,q)} \\ \vdots \\ \mathbf{U}^{(Q,q)} \end{bmatrix} &= \mathbf{\Phi}_q \in \mathbb{C}^{QFMQR\times FMR} \\ &= \begin{bmatrix} \mathbf{E}_{(q-1)FMQ+q} & \mathbf{E}_{(q-1)FMQ+Q+q} & \cdots & \mathbf{E}_{(q-1)FMQ+(M-1)Q+q} \end{bmatrix}, \end{aligned} \tag{B.19}$$

and from the definition in (3.92), (B.19) leads to

$$\begin{aligned} & v_{l,l}^{(q,q)}\,\mathbf{U}^{(q,q)} = \mathbf{I}_R, \quad \forall l = 1,...,FM, \forall q = 1,...,Q, \\ \Longrightarrow \quad & \mathbf{V}^{(q,q)} \otimes \mathbf{U}^{(q,q)} = \mathbf{I}_{FMR}, \quad \forall q = 1,...,Q, \end{aligned} \tag{B.20}$$

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}^{(1,1)} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{V}^{(Q,Q)} \end{bmatrix}, \qquad \mathbf{U} = \begin{bmatrix} \mathbf{U}^{(1,1)} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{U}^{(Q,Q)} \end{bmatrix}. \tag{B.21}$$

Thus, we obtain

$$\mathbf{S}^{(q)} = \hat{\mathbf{S}}^{(q)}\,\mathbf{U}^{(q,q)}, \qquad \mathbf{H}^{(q)} = \hat{\mathbf{H}}^{(q)}\,\mathbf{V}^{(q,q)}. \tag{B.22}$$

The only solution for $\mathbf{V}^{(q,q)} \otimes \mathbf{U}^{(q,q)} = \mathbf{I}_{FMR}$, for $q = 1,...,Q$, happens when both matrices $\mathbf{U}^{(q,q)}$ and $\mathbf{V}^{(q,q)}$ are identity matrices up to scalar factors that compensate each other, i.e. $\mathbf{U}^{(q,q)} = \alpha_q\,\mathbf{I}_R$ and $\mathbf{V}^{(q,q)} = 1/\alpha_q\,\mathbf{I}_{FM}$, which leads to (B.18).

Observe that $\mathbf{T}_{QFMR\times FP}$ given in (B.13) is full row-rank if

$$\begin{bmatrix} \mathrm{vec}\Big(\mathbf{W}^{(1)^{\mathrm{T}}}\Big)^{\mathrm{T}} \diamond \Big(\mathbf{C}_{f,:}^{(\mathcal{H},1)^{\mathrm{T}}} \diamond \mathbf{C}_{f,:}^{(\mathbf{S},1)^{\mathrm{T}}}\Big)^{\mathrm{T}} & \cdots & \mathrm{vec}\Big(\mathbf{W}^{(Q)^{\mathrm{T}}}\Big)^{\mathrm{T}} \diamond \Big(\mathbf{C}_{f,:}^{(\mathcal{H},Q)^{\mathrm{T}}} \diamond \mathbf{C}_{f,:}^{(\mathbf{S},Q)^{\mathrm{T}}}\Big)^{\mathrm{T}} \end{bmatrix} \in \mathbb{C}^{P\times QMR} \tag{B.23}$$

is full column-rank for all $f \in \{1,\ldots,F\}$, implying $P \geq QMR$.

We can rewrite (B.23)

$$\begin{aligned} & \Big[\mathrm{vec}(\mathbf{W}_{R\times QM})^{\mathrm{T}} \diamond \Big[\Big(\mathbf{C}_{f,:}^{(\mathcal{H},1)^{\mathrm{T}}} \diamond \mathbf{C}_{f,:}^{(\mathbf{S},1)^{\mathrm{T}}}\Big)^{\mathrm{T}} \cdots \Big(\mathbf{C}_{f,:}^{(\mathcal{H},Q)^{\mathrm{T}}} \diamond \mathbf{C}_{f,:}^{(\mathbf{S},Q)^{\mathrm{T}}}\Big)^{\mathrm{T}}\Big]\Big] = \\ & \qquad \Big[\mathrm{vec}(\mathbf{W}_{R\times QM})^{\mathrm{T}} \diamond \Big(\mathbf{C}_{f,:}^{(\mathcal{H})^{\mathrm{T}}} \diamond \mathbf{C}_{f,:}^{(\mathbf{S})^{\mathrm{T}}}\Big)^{\mathrm{T}}\ \tilde{\mathbf{\Phi}}_{QMR}\Big] \end{aligned} \tag{B.24}$$

with

$$\mathbf{W}_{R\times QM} \triangleq \Big[\mathbf{W}^{(1)^{\mathrm{T}}} \cdots \mathbf{W}^{(Q)^{\mathrm{T}}}\Big] \in \mathbb{C}^{R\times QM}, \tag{B.25}$$

$$\mathbf{C}_{f,:}^{(\mathcal{H})} \triangleq \Big[\mathbf{C}_{f,:}^{(\mathcal{H},1)} \cdots \mathbf{C}_{f,:}^{(\mathcal{H},Q)}\Big] \in \mathbb{C}^{P\times QM}, \tag{B.26}$$

$$\mathbf{C}_{f,:}^{(\mathbf{S})} \triangleq \Big[\mathbf{C}_{f,:}^{(\mathbf{S},1)} \cdots \mathbf{C}_{f,:}^{(\mathbf{S},Q)}\Big] \in \mathbb{C}^{P\times QR}, \tag{B.27}$$

$\mathbf{C}_{f,:}^{(\mathcal{H})}$ and $\mathbf{C}_{f,:}^{(\mathbf{S})}$ represent the global allocation matrices associated with the f-th subcarrier which concatenate the antenna and stream allocation matrices for all users, $\tilde{\mathbf{\Phi}}_{QMR} \in \mathbb{C}^{QMQR\times QMR}$ denotes a selection matrix which selects QMR columns of $\Big(\mathbf{C}_{f,:}^{(\mathcal{H})^{\mathrm{T}}} \diamond \mathbf{C}_{f,:}^{(\mathbf{S})^{\mathrm{T}}}\Big)^{\mathrm{T}}$.

Applying Lemma 2.2 to (B.24), if the elements of the code matrix are nonzero and $\Big[\Big(\mathbf{C}_{f,:}^{(\mathcal{H},1)^{\mathrm{T}}} \diamond \mathbf{C}_{f,:}^{(\mathbf{S},1)^{\mathrm{T}}}\Big)^{\mathrm{T}} \cdots \Big(\mathbf{C}_{f,:}^{(\mathcal{H},Q)^{\mathrm{T}}} \diamond \mathbf{C}_{f,:}^{(\mathbf{S},Q)^{\mathrm{T}}}\Big)^{\mathrm{T}}\Big]$ is full column-rank for all f, then $\mathbf{T}_{QFMR\times FP}$ will be full row-rank as well. It concludes the proof. ∎

Bibliography

[1] C. Oestges and B. Clerckx. *MIMO wireless communications: From real-world propagation to space-time code design*. Academic Press, 2007.

[2] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[3] J. G. Proakis. *Digital communications*. McGraw-Hill, 2000.

[4] A. Goldsmith. *Wireless communications*. Cambridge University Press, 2005.

[5] H. Jafarkhani. *Space-time coding*. Cambridge University Press, 2005.

[6] V. Tarokh, N. Seshadri, and A. R. Calderbank. Space-time codes for high data rate wireless communication: performance criterion and code construction. *IEEE Transactions on Information Theory*, 44(2):744–765, 1998.

[7] Lizhong Zheng and D. N. C. Tse. Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels. *IEEE Transactions on Information Theory*, 49(5):1073–1096, 2003.

[8] A.J. Paulraj, D.A. Gore, R.U. Nabar, and H. Bolcskei. An overview of MIMO communications: A key to gigabit wireless. *Proc. of the IEEE*, 92:198–218, 2004.

[9] I. E. Telatar. Capacity of multi-antenna gaussian channels. AT&T-Bell Labs Internal Tech. Memo., June 1995.

[10] G. J. Foschini. Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. *Bell Labs Technical Journal*, 1(2):41–59, 1996.

[11] G.J. Foschini and M.J. Gans. On limits of wireless communications when using multiple antennas. *Wireless Personal Communications*, 6(3):311–335, 1998.

[12] D. N. C. Tse, P. Viswanath, and L. Zheng. Diversity-multiplexing tradeoff in multiple-access channels. *IEEE Transactions on Information Theory*, 50(9):1859–1873, 2004.

[13] S. M. Alamouti. A simple transmit diversity technique for wireless communications. *IEEE Journal on Selected Areas in Communications*, 16(10):1451–1458, 1998.

[14] V. Tarokh, H. Jafarkhani, and A. R. Calderbank. Space-time block codes from orthogonal designs. *IEEE Transactions on Information Theory*, 45(5):1456–1467, 1999.

[15] H. Jafarkhani. A quasi-orthogonal space-time block code. *IEEE Transactions on Communications*, 49(1):1–4, 2001.

[16] M. O. Damen, A. Tewfik, and J. C. Belfiore. A construction of a space-time code based on number theory. *IEEE Transactions on Information Theory*, 48(3):753–760, 2002.

[17] B. Hassibi and B. M. Hochwald. High-rate codes that are linear in space and time. *IEEE Transactions on Information Theory*, 48(7):1804–1824, 2002.

[18] H. El Gamal and M. O. Damen. Universal space-time coding. *IEEE Transactions on Information Theory*, 49(5):1097–1119, 2003.

[19] M. O. Damen, H. El Gamal, and N. C. Beaulieu. Linear Threaded Algebraic Space-Time Constellations. *IEEE Transactions on Information Theory*, 49(10):2372–2388, 2003.

[20] Xue-Bin Liang. Orthogonal designs with maximal rates. *IEEE Transactions on Information Theory*, 49(10):2468–2503, 2003.

[21] N. Sharma and C. B. Papadias. Improved quasi-orthogonal codes through constellation rotation. *IEEE Transactions on Communications*, 51(3):332–335, 2003.

[22] Chau Yuen, Yong Liang Guan, and Tjeng Thiang Tjhung. Quasi-orthogonal STBC with minimum decoding complexity. *IEEE Transactions on Wireless Communications*, 4(5):2089–2094, 2005.

[23] E. Biglieri, Yi Hong, and E. Viterbo. On fast-decodable space–time block codes. *IEEE Transactions on Information Theory*, 55(2):524–530, 2009.

[24] Anzhong Wong and Jian-Kang Zhang. Novel rotated quasi-orthogonal space-time block codes with the fixed nearest neighbor number. *IEEE Signal Processing Letters*, 17(11):965–968, 2010.

[25] D. Agrawal, V. Tarokh, A. Naguib, and N. Seshadri. Space-time OFDM for high data-rate wireless communication over wideband channels. In *Proc. IEEE Vehicular Technology Conf. (VTC 1998)*, pages 2232–2236, Ottawa, Canada, 1998.

[26] B. Lu and X. Wang. Space-time code design in OFDM systems. In *Proc. IEEE Global Telecommunications Conf. (GLOBECOM)*, volume 2, pages 1000–1004, San Francisco, CA, USA, 2000.

[27] R. Blum, Y. Li, J. Winters, and Q. Yan. Improved space-time coding for MIMO-OFDM wireless communications. *IEEE Transactions on Communications*, 49(11):1873–1878, 2001.

[28] H. Bölcskei and A. Paulraj. Space-frequency coded broadband OFDM systems. In *Proc. IEEE Wireless Communications Networking Conf. (WCNC)*, volume 1, pages 1–6, Chicago, IL, USA, 2000.

[29] H. Bölcskei and A. J. Paulraj. Space-frequency codes for broadband fading channels. In *Proc. IEEE International Symposium on Information Theory (ISIT)*, page 219, Washington, DC, USA, 2001.

[30] W. Su, Z. Safar M. Olfat, and K. J. R. Liu. Obtaining full-diversity space-frequency codes from space-time codes via mapping. *IEEE Transactions on Signal Processing (Special Issue on MIMO Wireless Communications)*, 51(11):2905–2916, 2003.

[31] Y. Gong and K. B. Letaief. Space-frequency-time coded OFDM for broadband wireless communications. In *Proc. IEEE Global Telecommunications Conf. (GLOBECOM)*, volume 1, pages 519–523, San Antonio, TX, USA, 2001.

[32] Z. Liu, Y. Xin, and G. Giannakis. Space-time-frequency coded OFDM over frequency selective fading channels. *IEEE Transactions on Signal Processing*, 50(10):2465–2476, 2002.

[33] A. F. Molisch, M. Z. Win, and J. H. Winters. Space-time-frequency (STF) coding for MIMO-OFDM systems. *IEEE Communications Letters*, 6(9):370–372, 2002.

[34] G. R. Stuber, J. R. Barry, S. W. Mclaughlin, Y. Li, M. A. Ingram, and T. G. Pratt. Broadband MIMO-OFDM wireless communications. In *Proc. of IEEE*, volume 92, pages 271–294, 2004.

[35] L. Shao and S. Roy. Rate-one space-frequency block codes with maximum diversity for MIMO-OFDM. *IEEE Transactions on Wireless Communications*, 4(4):1674–1686, 2005.

[36] W. Su, Z. Safar, and K. J. R. Liu. Full-rate full-diversity space-frequency codes with optimum coding advantage. *IEEE Transactions on Information Theory*, 51(4):229–249, 2005.

[37] W. Su, Z. Safar, and K. J. R. Liu. Towards maximum achievable diversity in space, time and frequency: Performance analysis and code design. *IEEE Transactions on Wireless Communications*, 4(4):1847–1857, 2005.

[38] S. Senthuran, A. Anpalagan, O. Das, and S. Khan. On the performance of a constellation-rotated time-spreaded space-frequency coding scheme for 4×1 MISO OFDM transceivers. *Springer Wireless Personal Communications*, 60:251–262, 2011.

[39] B. Gupta and D. S. Saini. Space-Time/Space-Frequency/Space-Time-Frequency Block Coded MIMO-OFDM System with Equalizers in Quasi Static Mobile Radio Channels Using Higher Tap Order. *Springer Wireless Personal Communications*, 69(4):1947–1968, 2013.

[40] B. Hochwald, T. L. Marzetta, and C. B. Papadias. A transmitter diversity scheme for wideband CDMA systems based on space-time spreading. *IEEE Journal on Selected Areas in Communications*, 19(1):48–60, 2001.

[41] A. G. Dabak and R. Negi. Space time block coded transmit antenna diversity for WCDMA. U. S. Patent, 2002.

[42] H. Huang, H. Viswanathan, and G. J. Foschini. Multiple antennas in cellular CDMA systems: transmission, detection, and spectral effciency. *IEEE Transactions on Wireless Communications*, 1(3):383–392, 2002.

[43] X. Yu and D. Xu. Performance Analysis of Space-Time Coded CDMA System Over Nakagami Fading Channels with Perfect and Imperfect CSI. *Springer Wireless Personal Communications*, 62(3):633–653, 2012.

[44] N. Prasad, M. K. Varanasi, L. Venturino, and X. Wang. An Analysis of the MIMO-SDMA Channel With Space-Time Orthogonal and Quasi-Orthogonal User Transmissions and Efficient Successive Cancellation Decoders. *IEEE Transactions on Information Theory*, 54(12):5427–5446, 2008.

[45] D. Godard. An analytical constant modulus algorithm. *IEEE Transactions on Communications*, 28(11):1867–1875, 1980.

[46] A.-J. van der Veen and A. Paulraj. An analytical constant modulus algorithm. *IEEE Transactions on Signal Processing*, 44(5):1136–1155, 1996.

[47] S. Mayrargue. A blind spatio-temporal equalizer for a radio-mobile channel using the constant modulus algorithm (CMA). In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, volume 6, pages IV:317–320,, Adelaide, South Australia, 1994.

[48] E. Moulines and P. Duhamel and J.-F. Cardoso and S. Mayrargue. Subspace methods for the blind identification of multichannel FIR filters. *IEEE Transactions on Signal Processing*, 43(2):516–525, 1995.

[49] A.-J. van der Veen and S. Talwar and A. Paulraj. Blind estimation of multiple digital signals transmitted over FIR channels. *IEEE Signal Processing Letters*, 2(5):99–102, 1995.

[50] N. D. Sidiropoulos, G. B. Giannakis, and R. Bro. Blind PARAFAC receivers for DS-CDMA systems. *IEEE Transactions on Signal Processing*, 48(3):810–823, 2000.

[51] R. A. Harshman. Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-modal factor analysis. *UCLA Working Papers in Phonetics*, 16:1–84, 1970.

[52] J. D. Carroll and J. J. Chang. Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. *Psychometrika*, (35):283–319, 1970.

[53] N. D. Sidiropoulos and R. S. Budampati. Khatri-Rao space-time codes. *IEEE Transactions on Signal Processing*, 50(10):2396–2407, 2002.

[54] A. de Baynast, L. de Lathauwer, and B. Aazhang. Blind PARAFAC receivers for multiple access-multiple antenna systems. In *Proc. IEEE 58th Fall Vehicular Technology Conf. (VTC 2003)*, volume 2, pages 1128–1132, 2003.

[55] A. L. F. de Almeida, G. Favier, and J. C. M. Mota. Space-time multiplexing codes: A tensor modeling approach. In *Proc. IEEE 7th Workshop Signal Processing Advances in Wireless Communications SPAWC '06*, Cannes, France, 2006.

[56] A. L. F. de Almeida, G. Favier, and J. C. M. Mota. Multiuser MIMO system using block space-time spreading and tensor modeling. *Signal Processing*, 88(10):2388–2402, 2008.

[57] A. L. F. de Almeida, G. Favier, and J. C. M. Mota. Constrained tensor modeling approach to blind multiple-antenna CDMA schemes. *IEEE Transactions on Signal Processing*, 56(6):2417–2428, 2008.

[58] A. L. F. de Almeida, G. Favier, and J. C. M. Mota. A constrained factor decomposition with application to MIMO antenna systems. *IEEE Transactions on Signal Processing*, 56(6):2429–2442, 2008.

[59] A. L. F. de Almeida, G. Favier, and J. C. M. Mota. Space-time spreading-multiplexing for MIMO wireless communication systems using the PARATUCK-2 tensor model. *Signal Processing*, 89(11):2103–2116, 2009.

[60] L.L. de Lathauwer and A. de Baynast. Blind deconvolution of DS-CDMA signals by means of decomposition in rank-(1,L,L) terms. *IEEE Transactions on Signal Processing*, 56(4):1562–1571, 2008.

[61] D. Nion and L. de Lathauwer. A block component model-based blind DS-CDMA receiver. *IEEE Transactions on Signal Processing*, 56(11):5567–5579, 2008.

[62] A. L. F. de Almeida, G. Favier, and L. R. Ximenes. Space-Time-Frequency (STF) MIMO Communication Systems With Blind Receiver Based on a Generalized PARATUCK2 Model. *IEEE Transactions on Signal Processing*, 61(8):1895–1909, 2013.

[63] L. de Lathauwer. Decompositions of a higher-order tensor in block terms-Part II: Definitions and uniqueness. *SIAM Journal on Matrix Analysis and Applications*, 30(3):1033–1066, 2008.

[64] R. A. Harshman and M. E. Lundy. Uniqueness proof for a family of models sharing features of Tucker's three-mode factor analysis and PARAFAC/CANDECOMP. *Psychometrika*, 61(1):133–154, 1996.

[65] G. Favier, M. N. da Costa, A. L. F. de Almeida, and J. M. T. Romano. Tensor coding for CDMA-MIMO wireless communication systems. In *Proc. of European Signal Processing Conference (EUSIPCO)*, Barcelone, Spain, 2011.

[66] G. Favier, M. N. da Costa, A. L. F. de Almeida, and J. M. T. Romano. Tensor Space-time (TST) coding for MIMO wireless communication systems. *Signal Processing*, 92(4):1079–1092, 2012.

[67] N. D. Sidiropoulos and R. Bro. On the uniqueness of multilinear decomposition of N-way arrays. *Journal of Chemometrics*, 14:229–239, 2000.

[68] N. D. Sidiropoulos and X. Liu. Identifiability results for blind beamforming in incoherent multipath with small delay spread. *IEEE Transactions on Signal Processing*, 49(1):228–236, 2001.

[69] A. Stegeman. On uniqueness of the n-th order tensor decomposition into rank-1 terms with linear independence in one mode. *SIAM Journal on Matrix Analysis and Applications*, 31(5):2498–2516, 2010.

[70] X. Guo, S. Miron, D. Brie, and A. Stegeman. Uni-mode and partial uniquennes conditions for CANDECOMP/PARAFAC of three-way arrays with linearly dependent loadings. *IEEE Transactions on Signal Processing*, 33(1):111–129, 2012.

[71] N. D. Sidiropoulos, R. Bro, and G. B. Giannakis. Parallel factor analysis in sensor array processing. *IEEE Transactions on Signal Processing*, 48(8):2377–2388, 2000.

[72] J. W. Brewer. Kronecker products and matrix calculus in system theory. *IEEE Transactions on Circuits and Systems*, 25(9):772–781, 1978.

[73] S. Liu. Matrix results on the Khatri-Rao and Tracy-Singh products. *Linear Algebra and its Applications*, 289:267–277, 1999.

[74] C.G. Khatri and C.R. Rao. Solutions to some functional equations and their applications to characterization of probability distributions. *Sankhya*, 30:167–180, 1968.

[75] C. R. Rao and S. K. Mitra. *Generalized inverse of matrices and its applications*. John Wiley & Sons, 1972.

[76] J. B. Kruskal. Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra and its Applications*, 18:111–117, 1977.

[77] R. A. Harshman and M. E. Lundy. *Research methods in multimode data analysis*, chapter The PARAFAC model for three-way factor analysis and multidimensional scaling, pages 122–215. New York: Praeger, 1984.

[78] A. Stegeman and N. D. Sidiropoulos. On Kruskal's uniqueness condition for the CANDECOMP/PARAFAC decomposition. *Linear Algebra and its Applications*, 420:540–552, 2007.

[79] L. de Lathauwer. Decompositions of a higher-order tensor in block terms-Part I: Lemmas for partitioned matrices. *SIAM Journal on Matrix Analysis and Applications*, 30(3):1022–1032, 2008.

[80] F. L. Hitchcock. The expression of a tensor or a polyadic as a sum of products. *Journal Mathematical Physics*, 6(1):164–189, 1927.

[81] P. Comon and B. Mourrain. Decomposition of quantics in sums of powers of linear forms. *Signal Processing*, 53:93–108, 1996. Special Issue on Higher Order Statistics.

[82] L. De Lathauwer, B. De Moor, and J. Vandewalle. Computation of the canonical decomposition by means of a simultaneous generalized Schur decomposition. *SIAM Journal on Matrix Analysis and Applications*, 26(2):295–327, 2004.

[83] P. Comon, J. M. F. Ten Berge, L. De Lathauwer, and J. Castaing. Generic and typical ranks of multi-way arrays. *Linear Algebra and its Applications*, 430(11):2997–3007, 2009.

[84] R. A. Harshman. Determination and proof of minimum uniqueness conditions for PARAFAC1. *UCLA Working Papers in Phonetics*, 22:111–117, 1972.

[85] J. B. Kruskal. Rank, decomposition, and uniqueness for 3-way and N-way arrays. In R. Coppi & S. Bolasco, editor, *Multiway data analysis*, pages 7–18, Netherlands: North-Holland, Amsterdan, 1989.

[86] J. M. F. Ten Berge. Kruskal's polynomial for $2 \times 2 \times 2$ arrays and a generalization to $2 \times n \times n$ arrays. *Psychometrika*, 56:631–636, 1991.

[87] X. Liu and N. D. Sidiropoulos. Cramér-Rao Lower Bounds for Low-Rank Decomposition of Multidimensional Arrays. *IEEE Transactions on Signal Processing*, 49(9):2074–2086, 2001.

[88] J. M. F. Ten Berge and N. D. Sidiropoulos. On uniqueness in CANDECOMP/PARAFAC. *Psychometrika*, 67(3):399–409, 2002.

[89] T. Jiang and N. D. Sidiropoulos. Kruskal's Permutation Lemma and the Identification of CANDECOMP/PARAFAC and Bilinear Models with Constant Modulus Constraints. *IEEE Transactions on Signal Processing*, 52(9):2625–2636, 2004.

[90] L. de Lathauwer. A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization. *SIAM Journal on Matrix Analysis and Applications*, 28(3):642–666, 2006.

[91] A. Stegeman, J. M. F. Ten Berge, and L. de Lathauwer. Sufficient conditions for uniqueness in CANDECOMP/PARAFAC and INDSCAL with random component matrices. *Psychometrika*, 71(2):219–229, 2006.

[92] A. Stegeman. On uniqueness conditions for CANDECOMP/PARAFAC and INDSCAL with full column-rank in one mode. *Linear Algebra and its Applications*, 431:211–227, 2009.

[93] J. A. Rhodes. A concise proof of Kruskal's theorem on tensor decomposition. *Linear Algebra and its Applications*, 432:1818–1824, 2010.

[94] N. D. Sidiropoulos and G. Z. Dimic. Blind multiuser detection in W-CDMA systems with large delay spread. *IEEE Signal Processing Letters*, 8(3):87–89, 2001.

[95] J. M. F. Ten Berge. Partial uniqueness in CANDECOMP/PARAFAC. *Journal of Chemometrics*, 18(1):12–16, 2004.

[96] R. Bro, R. A. Harshman, N. D. Sidiropoulos, and M. E. Lundy. Modeling multi-way data with linearly dependent loadings. *Journal of Chemometrics*, 23:324–340, 2009.

[97] A. Stegeman and A. L. F. de Almeida. Uniqueness conditions for constrained three-way factor decompositions with linearly dependent loadings. *SIAM Journal on Matrix Analysis and Applications*, 31:1469–1499, 2009.

[98] D. Brie, S. Miron, F. Caland, and C. Mustin. An uniqueness condition for the 4-way CANDECOMP/PARAFAC model with collinear loadings in three modes. In *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4108–4111, Prague, Czech Republic, 2011.

[99] L. Tucker. Some mathematical notes of three-mode factor analysis. *Psychometrika*, 31(3):279–311, 1966.

[100] Rasmus Bro. *Multi-way analysis in the food industry. Models, algorithms and applications.* PhD thesis, University of Amsterdam, Netherlands, 1998.

[101] A. Kibangou and G. Favier. Blind joint identification and equalization of Wiener-Hammerstein communication channels using PARATUCK-2 tensor decomposition. In *Proc. of European Signal Processing Conference (EUSIPCO)*, Poznan, Poland, September 2007.

[102] E. A. Lee and D. G. Messerschmitt. *Digital communications*. Kluwer Academic, 1993.

[103] G.L. Turin. The characteristic function of Hermitian quadratic forms in complex normal random variables. *Biometrika*, 47(1-2):199–201, 1960.

[104] N. M. Faber, R. Bro, and P. K. Hopke. Recent developments in CANDECOMP/PARAFAC algorithms: a critical review. *Chemometrics and Intelligent Laboratory Systems*, 65(1):119–137, 2003.

[105] G. Tomasi and R. Bro. A comparison of algorithms for fitting the PARAFAC model. *Computational Statistics & Data Analysis*, 50(7):1700–1734, 2006.

[106] D. Nion and L. de Lathauwer. Levenberg-Marquardt computation of the block factor model for blind multi-user access in wireless communications. In *Proc. of European Signal Processing Conference (EUSIPCO)*, Florence, Italy, 2006.

[107] A. L. F. de Almeida, G. Favier, and J. C. M. Mota. Constrained Tucker-3 model for blind beamforming. *Signal Processing*, 89(6):1240–1244, 2009.

[108] K. Levenberg. A Method for the Solution of Certain Non-Linear Problems in Least Squares. *The Quarterly of Applied Mathematics*, 2:164–168, 1944.

[109] K. Madsen, H.B. Nielsen, and O. Tingleff. Methods for non-linear least squares problems. Department of Mathematical Modelling, Technical University of Denmark, 2004. second ed.

[110] A. L. F. de Almeida, G. Favier, and J. C. M. Mota. PARAFAC-based unified tensor modeling for wireless communication systems with application to blind multiuser equalization. *Signal Processing*, 87(2):337–351, 2007.