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

High-Order Statistical Methods for Blind Channel Identification and Source Detection with Applications to Wireless Communications

Autor:

CARLOS ESTÊVÃO ROLIM FERNANDES

Orientadores:

JOÃO CESAR M. MOTA
GÉRARD FAVIER

SOPHIA ANTIPOLIS, FRANÇA
MAIO, 2008

CARLOS ESTÊVÃO ROLIM FERNANDES

HIGH-ORDER STATISTICAL METHODS FOR BLIND CHANNEL IDENTIFICATION AND SOURCE DETECTION WITH APPLICATIONS TO WIRELESS COMMUNICATIONS

Tese realizada em regime de cotutela e submetida simultaneamente à Coordenação do Programa de Pós-Graduação em Engenharia de Teleinformática da Universidade Federal do Ceará (UFC) e à *Ecole Doctorale Sciences et Technologies de l'Information et de la Communication* da *Université de Nice Sophia Antipolis* (UNSA) como requisito parcial para obtenção do grau de Doutor em Engenharia de Teleinformática pela UFC e do grau de Doutor em Ciências pela UNSA.

Áreas de concentração:
Sinais e sistemas (UFC)
Automática e Processamento de Sinais e de Imagens (UNSA)
Orientadores:
João Cesar M. MOTA
Gérard FAVIER

Composição da banca examinadora:

João Cesar M. MOTA	*Professor titular da Universidade Federal do Ceará*	Orientador
Gérard FAVIER	*Directeur de Recherche au CNRS*	Orientador
Pierre COMON	*Directeur de Recherche au CNRS*	Presidente
Eric MOREAU	*Professeur de l'Université de Toulon*	Relator
Luc DENEIRE	*Maître de Conférences à l'Université de Nice Sophia Antipolis*	Examinador
Charles C. CAVALCANTE	*Professor visitante da Universidade Federal do Ceará*	Examinador
Guilherme de A. BARRETO	*Pesquisador do DETI - Universidade Federal do Ceará*	Examinador
Francisco R. P. CAVALCANTI	*Professor adjunto da Universidade Federal do Ceará*	Examinador
João Marcos T. ROMANO	*Professor titular da Universidade Estadual de Campinas*	Relator

SOPHIA ANTIPOLIS, FRANÇA
30 DE MAIO DE 2008

E ainda que tivesse o dom da profecia, e conhecesse todos os mistérios e toda a ciência, e ainda que tivesse toda a fé, a ponto de transportar montanhas, se não tivesse amor, nada seria.

(Coríntios, 13:2)

Esta tese é a expressão de meu amor incondicional por:

Hélida, minha metade

Vinícius, fruto de nosso amor

Fernando e Cynthia, meus pais, meu exemplo

ii

Acknowledgements

son amitié et pour les discussions de tous ordres, mon ami Sofiane Boudaoud ainsi que sa petite famille, Madina et Ziad, avec qui on garde tellement de bonnes choses en commun. Merci pour tous les encouragements, les rigolades, les prises de tête... merci pour ton exemple. Enfin, une mention très spéciale à André F. de Almeida (et Julliana), amigo e colega desde o tempo do mestrado, sempre disponível para uma discussão calorosa. Pesquisador de alto nível e batalhador de sempre, obrigado pela mão amiga de todas as horas.

Enfin, mes sincères remerciements à Sébastien Icart, Ronald Phlyppo, Laure Amate, Pietro Bonizzi, Vincenzo Angelino, Laurent Galluccio, Paolo Piro, Lionel Nicolas, Marie Oger, Marie-Andrée Agostini, Mathias Hesse, Mikael Sørensen, Aline Cabasson, Frédéric Payan, Xavier Luciani, David Simoncin, Cédric DeCésare, mais aussi les "*beleza boys*" de la MIT French Division, Vincent Garcia et Sylvain Boltz. Um abraço forte ao meu amigo Léo Hidd e sua sempre meiga Lili, aos gaúchos Elton e Khati e ainda aos outros "*brazucas*" que em algum momento tiveram sua passagem por Nice: Raul e Larissa, Leosam e tantos outros. Merci à vous tous, avec qui j'ai vécu des moments inoubliables. Vos ferez toujours partie des plus beaux souvenirs de ces années sur la Côte d'Azur.

Deixo ainda um abraço carinhoso de agradecimento a Neco, Emanuelle e ao pequeno Diego. Je garde toujours l'espoir de voir nos enfants jouer ensemble un de ces jours (mas, com camisas tricolores ou rubro-negras?).

Flora e Olivier, vocês são anjos na nossa vida. Aquelas pessoas que Deus manda de encomenda. Chegaram, se misturaram, se encrustaram e são parte da gente. Jóias sem preço. E as jóias são assim, eternas! ... mais, heureusement que vous êtes partis avant nous. Ça rend notre départ moins difficile.

Quanto a minha famìlia, agradeço primeiramente a meus pais, Fernando e Cynthia, meus grandes encorajadores e cúmplices de cada passo, de cada sucesso. Um dia eu lhes disse e hoje repito publicamente: vocês são meu guia, meu fio condutor, meu espelho de Deus na Terra. Amo vocês e essa vitoria é de vocês. Muito obrigado também às minhas irmãs Juliana e Mariana e meus amados sobrinhos, Ana Beatriz, Carlos Arthur e Lívia. Saibam que vocês foram sempre o farol desta embarcação. Vocês são o nosso Porto Seguro.

Não poderia terminar sem agradecer e abraçar meu companheiro, meu "amigo de fé", meu irmão de tantas jornadas, das aventuras de menino no quintal aos desafios de brincar de fazer ciência no estrangeiro. Alexandre, obrigado pela presença, pelo companheirismo e pela amizade, de agora e de sempre. Obrigado por cuidar de minha duplamente irmã. Aquela que brincava de bonecas com a minha mulher, e que hoje cuida do meu filho como se fosse o dela. Aline, você é como minha irmã caçula e ao mesmo tempo a cunhada mais velha. E é por isso mesmo que eu amo vocês.

Por fim, dedico este trabalho aos personagens principais deste sonho. Na alegria e na tristeza, na saúde e na doença, um apoio incondicional. Uma mulher de carne, osso e coragem. Hélida, posso ver seu nome ao lado meu na capa desta tese. E onde mais meu nome estiver. Você é a luz que complementa meu espírito. Você é a luz que me deu Vinícius, meu sangue e minha vida. E que continua a me dar a Felicidade todo dia, a cada olhar, a cada respirar.

Que o nome de Deus-Pai Todo-Poderoso seja louvado através deste humilde trabalho.

Carlos Estêvão Rolim Fernandes
Nice, June 2008.

Abstract

Modern telecommunication systems offer services demanding very high transmission rates. Channel identification appears as a major concern in this context. Looking forward better tradeoffs between the quality of information recovery and suitable bit-rates, the use of *blind* techniques is of great interest. Making use of the special properties of the 4th-order output cumulants, this thesis introduces new statistical signal processing tools with applications in radio-mobile communication systems. Exploiting the highly symmetrical structure of the output cumulants, we address the blind channel identification problem by introducing a multilinear model for the 4th-order output cumulant tensor, based on the Parallel Factor (Parafac) analysis. The components of the new tensor model have a Hankel structure, in the SISO case. For (memoryless) MIMO channels, redundant tensor factors are exploited in the estimation of the channel coefficients.

In this context, we develop blind identification algorithms based on a single-step least squares (SS-LS) minimization problem. The proposed methods fully exploit the multilinear structure of the cumulant tensor as well as its symmetries and redundancies, thus enabling us to avoid any kind of pre-processing. Indeed, the SS-LS approach induces a solution based on a sole optimization procedure, without intermediate stages, contrary to the vast majority of methods found in the literature. Using only the 4th-order cumulants, and exploiting the Virtual Array concept, we treat the source localization problem in multiuser sensor array processing. Exploiting a particular arrangement of the cumulant tensor, an original contribution consists in providing additional virtual sensors by improving the array resolution by means of an enhanced array structure that commonly arises when using 6th-order statistics. We also consider the problem of estimating the physical parameters of a multipath MIMO communication channel. Using a fully blind approach, we first treat the multipath channel as a convolutive MIMO model and propose a new technique to estimate its coefficients. This non-parametric technique generalizes the methods formerly proposed for the SISO and (memoryless) MIMO cases. Using a tensor formalism to represent the multipath MIMO channel, we estimate the physical multipath parameters by means of a combined ALS-MUSIC technique based on subspace algorithms. Finally, we turn our attention to the problem of determining the order of FIR channels in the context of MISO systems. We introduce a complete combined procedure for the detection and estimation of frequency-selective MISO communication channels. The new algorithm successively detects the signal sources, determines the order of their individual transmission channels and estimates the associated channel coefficients using a deflationary approach.

Keywords: *Blind channel identification, channel order determination, multipath MIMO channel estimation, Parafac decomposition, source localization, tensors, wireless communication systems.*

Résumé

L es systèmes de télécommunications modernes exigent des débits de transmission très élevés. Dans ce cadre, le problème d'identification de canaux est un enjeu majeur. L'utilisation de techniques aveugles est d'un grand intérêt pour avoir le meilleur compromis entre un taux binaire adéquat et la qualité de l'information récupérée. En utilisant les propriétés des cumulants d'ordre 4 des signaux de sortie du canal, cette thèse introduit de nouvelles méthodes de traitement du signal tensoriel avec des applications pour les systèmes de communication radio-mobiles. En utilisant la structure symétrique des cumulants de sortie, nous traitons le problème de l'identification aveugle de canaux en introduisant un modèle multilinéaire pour le tenseur des cumulants d'ordre 4, basé sur une décomposition de type Parafac. Dans le cas SISO, les composantes du modèle tensoriel ont une structure de Hankel. Dans le cas de canaux MIMO instantanés, la redondance des facteurs tensoriels est exploitée pour l'estimation des coefficients du canal.

Dans ce contexte, nous développons des algorithmes d'identification aveugle basés sur une minimisation de type moindres carrés à pas unique (SS-LS). Les méthodes proposées exploitent la structure multilinéaire du tenseur de cumulants aussi bien que les relations de symétrie et de redondance, ce qui permet d'éviter toute sorte de traitement au préalable. En effet, l'approche SS-LS induit une solution basée sur une seule et unique procédure d'optimisation, sans les étapes intermédiaires requises par la majorité des méthodes existant dans la littérature. En exploitant seulement les cumulants d'ordre 4 et le concept de réseau virtuel, nous abordons aussi le problème de la localisation de sources dans le cadre d'un réseau d'antennes multiutilisateur. Une contribution originale consiste à augmenter le nombre de capteurs virtuels en exploitant un arrangement particulier du tenseur de cumulants, de manière à améliorer la résolution du réseau, dont la structure équivaut à celle qui est typiquement issue de l'utilisation des statistiques d'ordre 6. Nous traitons par ailleurs le problème de l'estimation des paramètres physiques d'un canal de communication de type MIMO à trajets multiples. Dans un premier temps, nous considérons le canal à trajets multiples comme un modèle MIMO convolutif et proposons une nouvelle technique d'estimation des coefficients. Cette technique non-paramétrique généralise les méthodes proposées dans les chapitres précédents pour les cas SISO et MIMO instantané. En représentant le canal multi-trajet à l'aide d'un formalisme tensoriel, les paramètres physiques sont obtenus en utilisant une technique combinée de type ALS-MUSIC, basée sur un algorithme de sous-espaces. Enfin, nous considérons le problème de la détermination d'ordre de canaux de type RIF, dans le contexte des systèmes MISO. Nous introduisons une procédure complète qui combine la détection des signaux avec l'estimation des canaux de communication MISO sélectifs en fréquence. Ce nouvel algorithme, basé sur une technique de déflation, est capable de détecter successivement les sources, de déterminer l'ordre de chaque canal de transmission et d'estimer les coefficients associés.

Mots-clés: *canaux MIMO à trajets multiples, décomposition Parafac, détermination dordre, estimation de canaux, identification aveugle de canaux, localisation de sources, systèmes de communication sans-fils, tenseurs*

Resumo

O s sistemas de telecomunicações atuais oferecem servios que demandam taxas de transmissão muito elevadas. O problema da identificação de canal aparece nesse contexto com um problema da maior importância. O uso de técnicas cegas tem sido de grande interesse na busca por um melhor compromisso entre uma taxas binária adequada e a qualidade da informação recuperada. Apoiando-se em propriedades especiais dos cumulantes de 4ª ordem dos sinais à saída do canal, esta tese introduz novas ferramentas de processamento de sinais com aplicações em sistemas de comunicação rádio-móveis. Explorando a estrutura simétrica dos cumulantes de saída, o problema da identificação cega de canais é abordado a partir de um modelo multilinear do tensor de cumulantes 4ª ordem, baseado em uma decomposição em fatores paralelos (Parafac). No caso SISO, os componentes do novo modelo tensorial apresentam uma estrutura Hankel. No caso de canais MIMO sem memória, a redundância dos fatores tensoriais é explorada na estimação dos coeficientes dos canal.

Neste contexto, novos algoritmos de identificação cega de canais são desenvolvidos nesta tese com base em um problema de otimização de mínimos quadrados de passo único (SS-LS). Os métodos propostos exploram plenamente a estrutura multilinear do tensor de cumulantes bem como suas simetrias e redundâncias, evitando assim qualquer forma de pré-processamento. Com efeito, a abordagem SS-LS induz uma solução baseada em um único procedimento de minimização, sem etapas intermediárias, contrariamente ao que ocorre na maior parte dos métodos existentes na literatura. Utilizando apenas os cumulantes de ordem 4 e explorando o conceito de Arranjo Virtual, trata-se também o problema da localização de fontes, num contexto multiusuário. Uma contribuição original consiste em aumentar o número de sensores virtuais com base em uma decomposição particular do tensor de cumulantes, melhorando assim a resolução do arranjo, cuja estrutura é tipicamente obtida quando se usa estatísticas de ordem 6. Considera-se ainda a estimação dos parâmetros físicos de um canal de comunicação MIMO com muti-percursos. Através de uma abordagem completamente cega, o canal multi-percurso é primeiramente tratado como um modelo convolutivo e uma nova técnica é proposta para estimar seus coeficientes. Esta técnica não-paramétrica generaliza os métodos previamente propostos para os casos SISO e MIMO (sem memória). Fazendo uso de um formalismo tensorial para representar o canal de multi-percursos MIMO, seus parâmetros físicos podem ser obtidos através de uma técnica combinada de tipo ALS-MUSIC, baseada em um algoritmo de subespaço. Por fim, será considerado o problema da determinação de ordem de canais FIR, particularmente no caso de sistemas MISO. Um procedimento completo é introduzido para a detecção e estimação de canais de comunicação MISO seletivos em freqüência. O novo algoritmo, baseado em uma abordagem de deflação, detecta sucessivamente cada fonte de sinal, determina a ordem de seu canal de transmissão individual e estima os coeficientes associados.

Palavras-chave: *canais de multi-percursos MIMO, decomposição Parafac, determinação de ordem, estimação de canais, identificação cega de canais, localização de fontes, sistemas de de communicação sem-fio, tensores*

Contents

List of Figures

List of Tables

Mathematical notation and acronyms

List of mathematical notations and operators

$(\cdot)^*$ Conjugate operation.

$(\cdot)^{\mathsf{T}}$ Transpose operation.

$(\cdot)^{\mathsf{H}}$ Conjugate transpose (*Hermitian*).

$\mathbb{E}\{\cdot\}$ Expectation operator.

$\mathbb{R}$ Set of the Real numbers

$\mathbb{C}$ Set of the Complex numbers

$|\cdot|$ Absolute value of the scalar argument (corresponds to the magnitude of a complex number). Applied element-wise when the argument is a matrix.

$\|\cdot\|$ Euclidean norm of the vector argument, i.e. $\|\mathbf{a}\| = (\mathbf{a}^{\mathsf{H}}\mathbf{a})^{1/2}$, for $\mathbf{a} \in \mathbb{C}^{m \times 1}$, and $\|\mathbf{a}\| = (\mathbf{a}\mathbf{a}^{\mathsf{H}})^{1/2}$, for $\mathbf{a} \in \mathbb{C}^{1 \times m}$.

$Tr(\cdot)$ Trace of the matrix argument.

$\|\cdot\|_F$ Frobenius norm of the matrix argument, defined as $\|\mathbf{A}\|_F = Tr(\mathbf{A}\mathbf{A}^{\mathsf{H}})^{1/2}$, for $\mathbf{A} \in \mathbb{C}^{m \times n}$.

$\mathbf{I}_n$ The $n \times n$ identity matrix.

$(\cdot)^{\#}$ The *Moore-Penrose* pseudoinverse, defined for $\mathbf{A} \in \mathbb{C}^{m \times n}$ in such a way that
$\mathbf{A}^{\#}\mathbf{A} = \mathbf{I}_n$, if $m \geq n$, and $\mathbf{A}\mathbf{A}^{\#} = \mathbf{I}_m$, otherwise.

$Diag\,(\cdot)$ Diagonal matrix built from the entries of the vector argument.

$D_i\,(\cdot)$ Diagonal matrix built from ith row of the matrix argument.

$diag\,(\cdot)$ Column-vector built from the diagonal elements of the matrix argument.

$D_i\,(\cdot)$ Diagonal matrix built from the ith row of the matrix argument.

$rank\,(\cdot)$ Rank of the matrix argument.

r_X Rank of matrix $\mathbf{X}$, i.e. $r_X = rank\,(\mathbf{X})$.

k_X k-rank of matrix $\mathbf{X}$.

$off\,(\cdot)$ Off-diagonal operator, i.e. the root of the sum of the squared magnitudes of
the non-diagonal elements of the matrix argument.

$\mathbf{A}_{\cdot j}$ The jth column of matrix $\mathbf{A}$, i.e. if $\mathbf{A} \in \mathbb{C}^{m \times n}$ then $\mathbf{A} = [\mathbf{A}_{\cdot 1} \ldots \mathbf{A}_{\cdot n}]$ with
$\mathbf{A}_{\cdot j} \in \mathbb{C}^{m \times 1}$.

$\mathbf{A}_{i\cdot}$ The ith row of matrix $\mathbf{A}$, i.e. if $\mathbf{A} \in \mathbb{C}^{m \times n}$ then $\mathbf{A} = [\mathbf{A}_{1\cdot}^{\mathsf{T}} \ldots \mathbf{A}_{m\cdot}^{\mathsf{T}}]^{\mathsf{T}}$ with
$\mathbf{A}_{i\cdot} \in \mathbb{C}^{1 \times n}$.

$vec\,(\cdot)$ *Vectorization* operator: stacks the columns of its matrix argument into a
single column vector, i.e. for $\mathbf{A} \in \mathbb{C}^{m \times n}$, $vec\,(\mathbf{A}) = [\mathbf{A}_{\cdot 1}^{\mathsf{T}} \ldots \mathbf{A}_{\cdot n}^{\mathsf{T}}]^{\mathsf{T}} \in \mathbb{C}^{mn \times 1}$.

$unvec\,(\cdot, \cdot)$ *Unvectorization* operator: builds a matrix from the first (vector) argument
with the number of columns being equal to the second (scalar) argument.

$\circ$ Outer product.

$\otimes$ Kronecker product, defined for $\mathbf{A} \in \mathbb{C}^{m \times n}$ and $\mathbf{B} \in \mathbb{C}^{p \times q}$ as:

$$\mathbf{A} \otimes \mathbf{B} \triangleq \begin{pmatrix} a_{11}\mathbf{B} & \ldots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \ldots & a_{mn}\mathbf{B} \end{pmatrix} \in \mathbb{C}^{mp \times nq}.$$

◇   Khatri-Rao product (column-wise Kronecker product), defined for $\mathbf{A} \in \mathbb{C}^{m \times n}$ and $\mathbf{B} \in \mathbb{C}^{p \times n}$ as:

$$\mathbf{A} \diamond \mathbf{B} \triangleq [\mathbf{A}_{\cdot 1} \otimes \mathbf{B}_{\cdot 1} \ldots \mathbf{A}_{\cdot n} \otimes \mathbf{B}_{\cdot n}] \in \mathbb{C}^{mp \times n}.$$

$\mathrm{Re}(\cdot)$ Real part of complex number

$\mathrm{Im}(\cdot)$ Imaginary part of complex number

$\jmath$ Imaginary unit, i.e. $\jmath = (-1)^{1/2}$.

∂ Partial derivative

$f_{\mathrm{Z}}(z_1, \ldots, z_k)$ joint p.d.f. of the k random variables in the set Z.

$\Omega_{\mathrm{Z}}(\mathbf{u})$ Moment generating function of the random variables in the set Z (characteristic function).

$M_{m,\mathrm{Z}}$ mth-order moment of the random variables in the set Z.

$\Upsilon_{\mathrm{Z}}(\mathbf{u})$ Cumulant generating function of the random variables in the set Z (second characteristic function).

$c_{m,\mathrm{Z}}$ mth-order cumulant of the random variables in the set Z.

$cum\,[\cdot]$ joint cumulant operator of the random variables in the argument (the number of variables gives the order of the cumulant), i.e. $cum\,[z_1, \ldots, z_k] = c_{k,\mathrm{Z}}$.

$\gamma_{2,v}$ Variance of the random process $v(t)$, i.e. $\gamma_{2,v} = cum\,[v(t), v^*(t)]$.

$\gamma_{3,v}$ Skewness of the random process $v(t)$, i.e. $\gamma_{3,v} = cum\,[v^*(t), v(t), v^*(t)]$.

$\gamma_{4,v}$ Kurtosis of the random process $v(t)$, i.e. $\gamma_{4,v} = cum\,[v^*(t), v(t), v^*(t), v(t)]$.

$\mathtt{det}(\cdot)$ Determinant of the matrix argument.

$\mathbf{e}_i^{(I)}$ The ith basis vector of $\mathbb{R}^I$, i.e.

$$\mathbf{e}_i^{(I)} = \left[\underbrace{0, \ldots 0}_{(i-1)\times}, 1, \underbrace{0, \ldots 0}_{(I-i)\times} \right]^{\mathsf{T}}$$

List of acronyms

ALS	Alternating least squares
AR	Autoregressive
ARMA	Autoregressive Moving-average
BER	Bit error rate
BIOME	Blind identification of overcomplete mixtures
DF	Direction finding
DOA	Direction-of-arrival
DOD	Direction-of-departure
ESPRIT	Estimation of signal parameters via rotational invariance techniques
EVD	Eigenvalue decomposition
FIR	Finite impulse response
FOBIUM	Fourth-order blind identification of underdetermined mixtures
FOOBI	Fourth-order only blind identification
FOSI	Fourth-order system identification
HOS	High-order statistics
HO	High-order
ICA	Independent component analysis
ICAR	Independent component analysis using the redundancies of the Quadricovariance matrix
i.i.d	Independent and identically distributed
ISI	Intersymbol interference
JADE	Joint approximate diagonalization of eigenmatrices
JD	Joint diagonalization
LS	Least squares
MA	Moving-average
MISO	Multiple-input single-output
MIMO	Multiple-input multiple-output
MMSE	Minimum mean squared error
MUSIC	Multiple signal classification
NB	Narrowband
NMSE	Normalized mean squared error
Parafac	Parallel Factor analysis

PBCI	Parafac-based blind channel identification
PBMCI	Parafac-based blind MIMO channel identification
p.d.f.	probability density function
QPSK	Quadrature phase shift keying
QALS	Quadrilinear ALS
RMSE	Root mean squared error
rv	random variable
SER	Symbol error rate
SISO	Single-input single-output
SINR	Signal-to-interference-plus-noise ratio
SMF	Spatial matched filter
SNR	Signal-to-noise ratio
SOBI	Second-order blind identification
SOS	Second-order statistics
SS-LS	Single-step least squares
ST	Space-time
SVD	Singular value decomposition
TALS	Trilinear ALS
TLS	Total least squares
ULA	Uniform linear antenna
VA	Virtual array
VS	Virtual sensor

Introduction

Parametric channel modelling and estimation are of primary importance in digital telecommunication systems. The knowledge of the channel model can be used to design equalizers to deconvolve the received signals. Channel identification and equalization consist in the retrieval of unknown information about the transmission channel and source signals, respectively. In order to reach a desired quality of service, broadband wireless communication systems classically perform channel identification and/or equalization using pilot symbols, i.e. training sequences composed of *a priori* known signals. This supervised approach introduces an overhead to the transmission system that may not be suitable for certain radiocommunication systems since it reduces the effective transmission rate. On the other hand, unsupervised (or "*blind*") approaches take only the output signals into account with possibly some *a priori* hypothesis on the input signals.

Most of the known channel identification algorithms assume the channel order (memory) is known. This is not always necessarily true and any mismatch may have very costly consequences. Actually, the order of the radio mobile channel is closely related to the delay spread profile produced by the multipath propagation scenario. Long delay spread profiles characterize highly frequency-selective channels and introduce intersymbol interference (ISI) in the sampling process. Typical effects of under- or over-estimating the channel order include bit error rate (BER) floors, signal-to-noise ratio (SNR) penalties and numerical instabilities.

High-order statistics (HOS) have been an important research topic in diverse fields including data communication, speech and image processing and geophysical data processing. When dealing with stationary complex input signals, the second-order statistics (SOS) may be unable to keep the phase information of a *nonminimum* phase system and the use of HOS is generally mandatory for blindly identifying finite impulse response (FIR) channels, unless additional information about the input signal is known, such as the non-circularity property, for instance. The high-order spectra have the ability to preserve both magnitude and (nonminimum-) phase information. Moreover, it is well-known that all the cumulant spectra of order greater than 2 vanish for Gaussian signals, which makes HOS-based identification methods insensitive to an additive Gaussian noise [1, 2].

A major problem treated in this thesis concerns the blind identification of channel parameters, in the context of radiocommunication systems. Several relationships exist connecting

high-order cumulants of a linear process to the parameters of its generating model. A vast amount of papers can be found on this subject and numerous solutions have been proposed for the identification of linear *autoregressive* (AR), *moving-average* (MA) and ARMA models, exploiting only the cumulants of output signals. In particular, Brillinger and Rosenblatt [3] established the exact expressions for computing cumulants in terms of the coefficients of an FIR system. The well-known $C(q, k)$ solution proposed by Giannakis [4] requires very few statistics but is quite sensitive to cumulant estimation errors. Other approaches include techniques that use additional cumulant information yielding improved solutions, such as the methods by Mendel and Giannakis [5], Friedlander [6] and Comon [7], whose method is optimal in the total least squares (TLS) sense. See also [8, 9, 10, 11] among others.

Since the introduction of the independent component analysis (ICA) concept in the seminal paper by Comon [12], research efforts have been spent for generalizing simultaneous diagonalization criteria and establishing links with canonical tensor decompositions (c.f. [13, 14] and references therein). For instance, in [15], De Lathauwer et. al reformulated the canonical decomposition of high-order tensors as a simultaneous generalized Schur decomposition. The Parallel Factor (Parafac) analysis of a P-dimensional tensor with rank F consists in the decomposition of the tensor into a sum of F rank-one tensors, each one being written as an outer product of P vectors [16]. In fact, output cumulants are multi-index objects having a symmetric tensor representation [17] and the blind identification of linear mixtures is closely related to the (simultaneous) diagonalization of cumulant tensors [18, 19]. In Chapter 1, we present a survey of the main HOS concepts and properties; some algebraic tools and algorithms are also reviewed and a synthetic presentation of the Parafac tensor decomposition is included along with the proposition of an extended version of the alternating least squares (ALS) algorithm for the estimation of the Parafac components of tensors of any order.

The key-point in the use of the Parafac decomposition is its uniqueness property, which can be ensured under simple conditions that are stated in the Kruskal Theorem [20]. Furthermore, canonical tensor decompositions do not impose any kind of orthogonality constraints and the factorization of tensors composed of high-order output cumulants has the advantage of avoiding the so-called *prewhitening* operation by fully exploiting the multidimensional nature of the cumulant tensor. Moreover, the tensor rank is not bounded by the tensor dimensions as it is the case for matrices, which conceptually allows for the blind identification of underdetermined mixtures. A formal relationship between Parafac decomposition and simultaneous matrix diagonalization has been established in [21] showing that the components of the tensor decomposition can be obtained from a simultaneous matrix diagonalization by congruence transformation, leading to weaker uniqueness conditions and yielding algorithms that identify a greater number of user channels with a given number of receive antennas, but still proceeding with two computation stages to recover the channel coefficients [22], i.e. one needs to compute a unitary factor (spatial pre-whitening) before extracting the channel coefficients from an estimated matrix product.

Our main focus in Chapter 2 is in exploiting the redundancies in the factors of the 4th-order cumulant tensor decomposition by solving a single-step least-squares (SS-LS) problem, under very mild assumptions. To this end, we treat the 4th-order cumulants as a Parafac tensor with components having a particular Hankel structure. Introducing this new cumulant tensor modeling enables us to develop an iterative blind identification algorithm for the case of FIR single-input single-output (SISO) communication channels. In this context, the proposed method estimates the channel coefficients by solving a sole minimization problem, contrary to previously known techniques, thus avoiding classical pre-processing operations. On the other hand, we also treat the case of instantaneous (memoryless) multiple-input multiple-output (MIMO) mixtures by extending our 4th-order cumulant Parafac tensor model and then introducing another important contribution of this thesis, corresponding to a SS-LS algorithm for the blind identification of the MIMO channel coefficients. In the FIR-SISO case, the SS-LS Parafac-based algorithm represents a new tensor-based scheme for the blind estimation of the channel coefficients. Although the 4th-order cumulant symmetries have being exploited with a tensor formalism for a long time [23, 24], the SS-LS approach also consists, to our knowledge, in the first contribution proposing to improve the LS solution of the Parafac decomposition, in both SISO and MIMO contexts, using the redundancies of the 4th-order cumulant tensor.

Mobile radiocommunication systems are often characterized by multipath propagation, which introduces ISI, thus causing serious limitations in capacity and performance. In multiuser/multiantenna systems, this scenario can be represented by a convolutive MIMO channel model, characterized by the multipath physical parameters (delays, attenuations, and angles of departure and arrival). In this context, equalization algorithms generally make use of the channel coefficients and the multipath parameters can be of interest for source localization purposes, among other applications. In this thesis, we have been interested in estimating both the MIMO channel coefficients and the physical parameters describing the multipath propagation scenario, using a fully-blind two-stage approach. First, we extend the 4th-order output cumulant tensor model for the convolutive MIMO channel case and, using a SS-LS algorithm, we perform a non-parametric estimation of the channel coefficients. This cumulant tensor model along with the new blind identification technique can be viewed as a generalization of the models and methods proposed in Chapter 2 for the SISO and the instantaneous MIMO cases, hence consisting in a major contribution of chapter 3. In a second stage, we introduce a tensor notation to represent the structure of a convolutive multiuser radio channel, which allows us to identify the spatial and temporal signatures of the channel by using a 3rd-order Parafac decomposition. Using an ALS-based algorithm followed by a MUSIC-like search for the multipath parameters, we end up with a new combined ALS-MUSIC technique that allows for the recovery of the physical structure of the MIMO channel, as well as its coefficients without the ambiguities due to the Parafac decomposition.

Actually, MUSIC-like algorithms play an important role in determining the location of signal

sources in sensor array processing. In this context, direction finding (DF) techniques have been of particular interest for source separation and interference suppression (beamforming) applications. We also treat this problem in chapter 3, where we come up with a new high-resolution DF algorithm based on the 4th-order cumulants only. Exploiting the virtual antenna array concept [25, 26], we show that we can get some additional virtual sensors using the Khatri-Rao structure of an unbalanced arrangement of the cumulant tensor, thus providing some additional free dimensions to the antenna array. Without resorting to 6th-order statistics, our new source localization algorithm uses the SS-LS approach to estimate the extended virtual array, allowing for resolution gains comparable to the 6-MUSIC algorithm [27] with cumulant estimation burden equivalent to the 4-MUSIC algorithm [28, 29].

Finally, we turn our attention to the problem of determining the order of radiocommunication channels in the context of multiple-input single-output (MISO) systems, using only the 4th-order cumulants of the output data sequence. Channel order estimation is a classic model selection problem strongly related to determining the number of signals embedded in noisy observations in narrow-band array processing. This is often referred to as the signal (or source) detection problem [30, 31, 32]. A classical solution for the channel order selection problem relies on the Sphericity Test [33], which is a well-known algorithm for estimating the number of parameters in a model using the eigenvalues of a correlation matrix in order to determine test statistics. This algorithm finds several applications in the context of passive arrays, such as in [34, 35], where an important modification of the Sphericity Test has been proposed. This approach is based on the sample correlation matrix, which is estimated from a finite number of output samples and hence subject to statistical variations.

In chapter 4, we address the problem of channel order determination as a series of hypothesis tests based on scalar statistics. Using the multivariate estimator of the 4th-order output cumulants, we exploit the insensitiveness of a Chi-square test statistic with respect to the non-linearity of a stochastic process. This property enables us to observe the amount of signal energy in the representation space of the 4th-order cumulants and thereby deduce the order of a FIR-SISO communication channel. Our approach leads to a new channel order detection method and we provide a performance analysis along with a criterion to establish decision thresholds, according to a desired level of statistical tolerance. Afterwards, we come up with another major contribution of the chapter, which consists in introducing the concept of MISO channel nested detectors based on a deflation-type procedure using the 4th-order output cumulants. The nested detector devices run combined algorithms that select the order and estimate the coefficients associated to the different emitters composing the MISO channel. By treating successively shorter and shorter channels, we can also determine the number of sources.

Chapter contents and contributions

This thesis is fundamentally based on the following axis:

1. Blind channel identification exploiting the symmetry properties of the 4th-order output cumulants in a single LS minimization problem;

2. Source localization in multiuser narrowband array processing;

3. Structured channel parameter estimation in a multipath propagation scenario;

4. Channel order estimation and signal detection in the context of MISO channels.

The thesis is divided in four chapters, organized as follows:

Chapter 1: We present a survey on high-order statistics tools in Signal Processing. Some definitions and properties of HOS are introduced and reported to the context of telecommunication systems, including important relationships between higher-order cumulants and the parameters of a linear system model. Useful linear algebraic tools and (simultaneous) matrix decompositions are discussed. Finally, we present a brief introduction to multilinear tensor decomposition tools using a generalized formulation for tensors of any order and extending the ALS algorithm to this general case. Uniqueness conditions are presented and the particular cases of 3rd- and 4th-order tensors are also discussed.

Chapter 2: New blind channel identification algorithms are proposed exploiting 4th-order cumulant redundancies in order to perform the cumulant tensor decomposition by solving a single least squares minimization problem. We analyze the cumulant tensors in the convolutive SISO as well as in the instantaneous MIMO cases and propose particular cumulant tensor models for treating each case. Then, we propose Parafac-based SS-LS algorithms to estimate the channel coefficients. The algorithms, based on 4th-order cumulants only, are also able to treat certain underdetermined mixtures. Known algorithms based on the joint-diagonalization technique are also described and performance comparisons are provided by means of computer simulations to assess the applicability of the proposed algorithms in both SISO and MIMO cases.

Chapter 3: In this chapter, we are first interested in the problem of blind multiuser localization in the context of multiple antenna array processing, under the far-field assumption, using only the array output signals. Exploiting the Virtual Array concept, we propose a high-resolution DF algorithm exploiting an unbalanced structure of the cumulant tensor, based on the estimation of an array matrix formed from a double Khatri-Rao product, using the SS-LS technique. Then, we also treat the problem of estimating the physical parameters of a multipath MIMO communication channel characterized by specular reflections due to remote scatterers. A two-stage approach is proposed: before extracting the physical parameters of the multipath channel

structure using a proposed ALS-MUSIC approach, we extend the 4th-order cumulant tensor model to the case of a convolutive MIMO channel and estimate its coefficients using a SS-LS algorithm.

Chapter 4: This chapter treats the problem of estimating the channel order and detecting the number of sources in a MISO channel. First, we propose a sequence of hypothesis tests for selecting order of a FIR-SISO communication channel. Relying on some properties of the 4th-order cumulant, we introduce a test variable that is sensitive to the non-linearity of a stochastic process. Exploiting this property enables us to detect the channel order. We discuss the choice of a decision criterion and propose a new algorithm for order determination. In the context of MISO channels, we introduce the concept of nested detectors that successively test for the presence of shorter and shorter channels, determining their respective orders and estimating their associated coefficients using HOS-based blind identification techniques.

The main original contributions of this thesis are listed below:

- Proposition of a generalized version of the ALS algorithm for the estimation of the Parafac components of a Pth-order tensor (section 1.3.2);

- Definition of a tensor model for 4th-order output cumulants, in the FIR-SISO case, admitting a Parafac decomposition with components having a particular Hankel structure (section 2.2);

- Development of a Parafac-based blind channel identification (PBCI) algorithm using a SS-LS approach (section 2.3);

- Proposition of a Parafac-based blind (memoryless) MIMO channel identification (PBMCI) algorithm using the SS-LS approach to exploit the redundancies in the Parafac components of the 4th-order output cumulant tensor (sections 2.4 and 2.5.2);

- Definition of a 3rd-order tensor model of the 4th-order output cumulants and proposition of a Parafac SS-LS algorithm for blind channel identification in the memoryless MIMO case (sections 2.4 and 2.5.2);

- Derivation of a 3rd-order virtual array based on an unbalanced unfolding of the 4th-order cumulant tensor structure; description of a method for estimating the VA using the SS-LS approach (section 3.2);

- Unification of the 4th-order cumulant tensor models by means of a generalized tensor formulation including the convolutive MIMO case, in which the Parafac components have a block-Hankel structure (section 3.3.1);

- Development of a non-parametric Parafac-based SS-LS algorithm for the blind identification of convolutive MIMO channels (section 3.4.1);

- Introduction of a Parafac tensor representation for the structured multipath MIMO channel model based on the parameters characterizing the signal propagation (section 3.3.2);

- Realization of a combined ALS-MUSIC method for the estimation of the structured multipath MIMO channel parameters (section 3.4.2);

- Determination of a Chi-square test statistic based on the energy of the 4th-order cumulants (section 4.1); proposition of a blind method for determining the order of a SISO channel (section 4.1.2);

- Development of a combined blind procedure for signal detection, order determination and channel identification in the context of MISO channels (section 4.2);

Chapter 1

Mathematical Tools

The telecommunications history dates back to the 19th century, when Samuel F. B. Morse started telegraphic transmissions and Alexander Graham Bell invented the telephone. Since then, the information technologies experimented dramatic developments and the technological challenges changed a lot: ubiquitous access, powerful computation and high transfer rates. The new world scenario created a very harmful environment to data transmissions, especially in the wireless and mobile communication contexts. Dense urban agglomerations, hot spot user areas, and high-speed transportation means are some of the factors causing the physical phenomena responsible for signal deterioration. Actually, these phenomena lead to well-known troubles of imperfect information recovery, known by the generic name of *interference*.

In multipath propagation environments, distorted frequency-response channels may cause the energy of the electromagnetic pulses to spread in time thus corrupting adjacent pulses and introducing inter-symbol interference (ISI). Wireless communication systems are known to face several problems related to multipath propagation including ISI as a very severe performance and capacity limiting factor. In order to suppress the effects of interference and assure the information recovery at the receiver side, knowledge of the transmission channel is necessary. Several mathematical models have been developed in order to tentatively predict the behavior of real systems with the purpose of designing filtering structures (*equalizers*) that compensates or reduces the ISI.

In this context, second- and high-order statistics (HOS) appear as powerful signal processing tools, playing a very important role in several applications that involve system information recovery. However, second-order statistics (SOS) contain no phase information and, as a consequence, nonminimum phase signals cannot be correctly identified by those techniques. On the other hand, HOS of Gaussian signals are either zero (odd-ordered moments) or contain redundant information. This is a remarkable information since measurement noise is often Gaussian and many real-life signals have non-zero HOS. Several important papers on HOS have been written since the Sixties, but it is from the Seventies that the subject starts to experience its greatest growing of interest with applications involving different contexts such as economics,

speech, seismic data processing, plasma physics, control, optics and obviously, communications. More recently, the introduction of the multiuser and multiple output communication systems served as an application background to the new developments in the area. Section 1 is a brief tutorial on this subject, presenting the important definitions, properties and some relationships between high-order moments and cumulants.

Anyway, processing HOS data often implies implementing algebraic methods and generally requires the use of numerical algorithms. Most of the engineering problems and physical applications make use of numerical methods, especially those associated with stability and perturbation analysis. In particular, the eigenvalue problem for square matrices is of crucial importance in several domains of Signal Processing. As we will see in section 2, several matrix factorization techniques play an important role in these scenarios and will be divided in two main strategies [36]: *diagonalization* and *triangularization*. The former is classically solved by Jacobi-like algorithms, still powerful and popular due to the high inherent parallelism. The latter one is applicable to any (square) matrix and its classical implementations are based on QR-type algorithms. In some domains, such as blind sources separation and system identification, simultaneous matrix factorization tools are strongly desirable. Section 2 also discusses simultaneous decomposition techniques that are suitable for processing long data records sharing common structure properties but differing in the individual information contents.

On the other hand, multi-linear algebraic tools have been developed and applications in multiuser systems using HOS are now a current research topic. In particular, the *trilinear* Parallel Factor analysis (Parafac) has become very popular in the fields of Psychometrics and Chemometrics [37, 38] but it also has been widely used in Signal Processing applications (c.f. [39, 40, 41, 42, 43, 44] among others). The major importance of using Parafac is due to its uniqueness property, ensured under very mild conditions that have been stated by Kruskal [20]. In section 3, we present a synthetic review and some fundamental aspects of the Parafac decomposition, using a general formulation for the case of a Pth-order tensor. We also briefly discuss the estimation of Parafac components by describing the associated alternating least squares (ALS) algorithm.

1 High-order statistics

Moments and cumulants are descriptive constants of a probability distribution. In this section, we present a global overview on the main aspects of HOS including definitions, properties and important relationships.

Moments and cumulants

Let us consider a set of k real random variables (r.v.) $Z = \{z_1, \ldots, z_k\}$ with known joint probability density function (p.d.f.), $f_Z(z_1, \ldots, z_k)$, and a sequence of k integer numbers

$(n_1, n_2, \cdots, n_k)$ such that $n_1 + n_2 + \cdots + n_k = m$. The mth-order moment of Z is defined as

$$M_{m,\mathrm{Z}} = \mathbb{E}\left\{ z_1^{n_1} z_2^{n_2} \ldots z_k^{n_k} \right\}, \tag{1.1}$$

where $\mathbb{E}\{\cdot\}$ stands for the *expectation* operator. The moments $M_{m,\mathrm{Z}}$ defined in (1.1) can be obtained as the coefficients of the Taylor series expansion around the origin of the characteristic function $\Omega_{\mathrm{Z}}(\mathbf{u})$, also known as the *moments generating function*, given as follows [45]:

$$
\begin{aligned}
\Omega_{\mathrm{Z}}(\mathbf{u}) &\triangleq \int_{-\infty}^{\infty} f_{\mathrm{Z}}(z_1, \ldots, z_k) \exp\left(j\mathbf{u}^\mathsf{T}\mathrm{Z} \right) dz_1 \ldots dz_k \\
&= \mathbb{E}\left\{ \exp\left(j\mathbf{u}^\mathsf{T}\mathrm{Z} \right) \right\},
\end{aligned}
\tag{1.2}
$$

where $\mathbf{u} = (u_1 \ldots u_k)^\mathsf{T}$. For complex random variables, we consider the joint distribution of their real and imaginary parts, and the expressions given here become more complicated [46].

The second characteristic function, known as the *cumulants generating* function, $\Upsilon_{\mathrm{Z}}(\mathbf{u})$, is then defined as the *natural logarithm* of the moments generating function, i.e. $\Upsilon_{\mathrm{Z}}(\mathbf{u}) = \ln[\Omega_{\mathrm{Z}}(\mathbf{u})]$. Thus, the mth-order cumulants of Z are obtained as the coefficients of its Taylor series expansion around the origin and they can be computed as the partial derivatives of $\Upsilon_{\mathrm{Z}}(\mathbf{u})$:

$$c_{m,\mathrm{Z}} = \frac{\partial^m \Upsilon_{\mathrm{Z}}(\mathbf{u})}{\partial u_1^{n_1} \ldots \partial u_k^{n_k}}. \tag{1.3}$$

Let us assume, without loss of generality, that $n_1 = \cdots = n_k = 1$ and thus $m = k$. Now, denote by $\mathbb{P}_i^{(\eta)}$ a *partition* of length η of the set $\mathcal{I} = \{1, \ldots, m\}$. The partition $\mathbb{P}_i^{(\eta)}$ is an unordered collection of η nonintersecting nonempty sets P_j such that $\bigcup_{j=1}^{\eta} \mathrm{P}_j = \mathcal{I}$. Let $\mathcal{P}_\eta$ be the set containing all partitions of $\mathcal{I}$ with length η, so that $\mathrm{P}_j \subset \mathbb{P}_i^{(\eta)} \subset \mathcal{P}_\eta$, $1 \leq \eta \leq m$. Consider the set $\mathcal{P}$ comprising all possible groups of partitions $\mathcal{P}_\eta$ of the set $\mathcal{I}$, i.e. $\mathcal{P} = \{\mathcal{P}_1, \ldots, \mathcal{P}_m\}$. For instance, when $m = 3$ we have $\mathcal{I} = \{1, 2, 3\}$ and thus:

$$\mathcal{P}_1 = \left\{ \underbrace{\{1,2,3\}}_{\mathbb{P}_1^{(1)}} \right\}, \quad \mathcal{P}_2 = \left\{ \underbrace{\left[\{1\},\{2,3\}\right]}_{\mathbb{P}_1^{(2)}}, \underbrace{\left[\{2\},\{1,3\}\right]}_{\mathbb{P}_2^{(2)}}, \underbrace{\left[\{3\},\{1,2\}\right]}_{\mathbb{P}_3^{(2)}} \right\}, \quad \mathcal{P}_3 = \left\{ \underbrace{\left[\{1\},\{2\},\{3\}\right]}_{\mathbb{P}_1^{(3)}} \right\}$$

so that

$$\mathcal{P} = \left\{ \mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3 \right\} = \left\{ \left[\{1,2,3\}\right], \left[\{1\},\{2,3\}\right], \left[\{2\},\{1,3\}\right], \left[\{3\},\{1,2\}\right], \left[\{1\},\{2\},\{3\}\right] \right\}.$$

Using the above definitions, we can state the following formulas explicitly relating moments to cumulants and vice-versa, respectively [2]:

$$c_{m,\mathrm{Z}} = \sum_{\eta=1}^{m} (-1)^{\eta-1}(\eta-1)! \sum_{\mathbb{P}_i^{(\eta)} \subset \mathcal{P}_\eta} \prod_{j=1}^{\eta} M_{m,\mathrm{Z}_{\mathrm{P}_j}} \qquad \mathrm{P}_j \subset \mathbb{P}_i^{(\eta)} \tag{1.4}$$

$$M_{m,\mathrm{Z}} = \sum_{\eta=1}^{m} \sum_{\mathbb{P}_i^{(\eta)} \subset \mathcal{P}_\eta} \prod_{j=1}^{\eta} c_{m,\mathrm{Z}_{\mathrm{P}_j}} \qquad \mathrm{P}_j \subset \mathbb{P}_i^{(\eta)} \tag{1.5}$$

where Z_{P_j} is a subset of $Z = \{z_1, \ldots, z_m\}$ formed with the elements having indices in P_j.

From definition (1.1), we observe that if the variables $z_1, \ldots, z_m$ are independent with zero mean (i.e. centered around the origin) then *all the odd-order moments are identically zero*. From (1.4), we note that this property can be extended to cumulants because, if m is odd then every partition $\mathbb{P}_i^{(\eta)}$ will always contain at least one subset P_j with an odd number of elements. That is the reason why in so many situations we are constrained to use at least fourth-order statistics. For centered independent processes, the expression for the 4th-order cumulant reduces to

$$c_{4,Z} = \mathbb{E}\{z_1 z_2 z_3 z_4\} - \mathbb{E}\{z_1 z_2\}\mathbb{E}\{z_3 z_4\} - \mathbb{E}\{z_1 z_3\}\mathbb{E}\{z_2 z_4\} - \mathbb{E}\{z_1 z_4\}\mathbb{E}\{z_2 z_3\}, \qquad (1.6)$$

with $Z = \{z_1, z_2, z_3, z_4\}$. For notational convenience, we define the following convention, to be used throughout the rest of this thesis:

$$c_{m,Z_{\mathbb{P}}} \triangleq cum\,[z_{j_1}, \ldots, z_{j_m}], \qquad (1.7)$$

where $\mathbb{P} = \{j_1, \ldots, j_m\}$ and the operator $cum\,[\cdot]$ stand for the mth-order joint cumulant of the random variables $z_{j_1}, \ldots, z_{j_m}$. Thus, considering a zero-mean random process $\{v(t)\}$, we have

$$M_{m,\Upsilon} = \mathbb{E}\left\{\prod_{i=0}^{m-1} v(t + t_i)\right\} \qquad (1.8)$$

$$C_{m,\Upsilon} = cum\,\left[v(t + t_0), v(t + t_1), \ldots, v(t + t_{m-1})\right], \qquad (1.9)$$

where $\Upsilon = \{v(t + t_0), v(t + t_1), \ldots, v(t + t_{m-1})\}$ and hence the involved r.v. are time-shifted samples of the process $\nu(t)$. Notice that, for stationary processes, the mth-order statistics depend only on the $m - 1$ time-lags $\tau_1 = t_1 - t_0$, ..., $\tau_{m-1} = t_{m-1} - t_0$. This allows us to introduce the following notations:

$$M_{m,v}(\tau_1, \ldots, \tau_{m-1}) = \mathbb{E}\left\{v(t)\prod_{i=1}^{m-1} v(t + \tau_i)\right\} \qquad (1.10)$$

$$C_{m,v}(\tau_1, \ldots, \tau_{m-1}) = cum\,[v(t), v(t + \tau_1), \ldots, v(t + \tau_{m-1})] \qquad (1.11)$$

where $\{v(t)\}$ is a zero-mean stationary random process. Stationarity will be further discussed later in this section.

Polyspectra

The *polyspectrum* of a stationary process $v(t)$ is defined as the $(m - 1)$-dimensional discrete Fourier Transform of the mth-order cumulant, i.e.

$$S_{m,v}(\omega_1, \omega_2, \ldots, \omega_{m-1}) \triangleq \sum_{\tau_1=-\infty}^{\infty} \cdots \sum_{\tau_{m-1}=-\infty}^{\infty} C_{m,v}(\tau_1, \ldots, \tau_{m-1}) \times \exp\left(-j\sum_{i=1}^{m-1}\omega_i\tau_i\right) \quad (1.12)$$

where $C_{m,v}(\tau_1, \ldots, \tau_{m-1})$ is assumed to be absolutely summable. For $m = 2$, equation (1.12) coincides with the classic *power spectrum* $S_{2,v}(\omega)$. When $m = 3$, we have the so-called *bispectrum* $S_{3,v}(\omega_1, \omega_2)$, whereas for $m = 4$ we get $S_{4,v}(\omega_1, \omega_2, \omega_3)$, named *trispectrum*.

Next, we present a survey of the most important properties of cumulants and polyspectra. Further details and proofs to these properties can be easily found in the literature (c.f. [1, 2]) and will be omitted here.

1.1 Properties and comments

Due to some special properties, the use of cumulants yields advantages that may not necessarily be exploited when dealing with moments. Consider a set of k r.v. $Z = \{z_1, \ldots, z_k\}$ and the following properties hold for cumulants:

P1 *Linearity*:

$$cum\left[\lambda_1 z_1, \ldots, \lambda_k z_k\right] = cum\left[z_1, \ldots, z_k\right] \prod_{i=1}^{k} \lambda_i, \tag{1.13}$$

where $\lambda_i, i = 1, \ldots, k$ are constants

P2 *Additivity*: If x_0, y_0 are mutually independent random variables, then

$$cum\left[x_0 + y_0, z_1, \ldots, z_k\right] = cum\left[x_0, z_1, \ldots, z_k\right] + cum\left[y_0, z_1, \ldots, z_k\right]. \tag{1.14}$$

P3 *Symmetry*: Let the set $\{n_1, \ldots, n_k\}$ be any permutation of $\{1, \ldots, k\}$. Then, it holds

$$cum\left[z_1, \ldots, z_k\right] = cum\left[z_{n_1}, \ldots, z_{n_k}\right] \tag{1.15}$$

P4 *Independence*: If any (nonempty) subset of Z is independent of the remaining r.v. in Z, then

$$cum\left[z_1, \ldots, z_k\right] = 0. \tag{1.16}$$

Comments and discussion

1. *Stationarity*: A stochastic process $\{v(t)\}$ is said to be *strictly* stationary if the joint distribution of any set of random variables $\{v(t + t_0), \ldots, v(t + t_{n-1})\}$ is independent of the time t, for all n. In other words, all the moments of $v(t)$ are time-invariant and depend only on the differences between the time-lags $\tau_i = t_i - t_0$, $i \in [1, n-1]$. In practice, it is usual to deal with the weaker concept of *wide-sense stationarity* (WSS), also known as second-order stationarity, which only ensures the mean and the correlation function to be independent of the observation instants. Recalling property P4, it follows that for a stationary centered i.i.d. process the mth-order cumulant is a multidimensional impulse function, i.e.

$$cum\left[v(t), v(t + \tau_1), \ldots, v(t + \tau_{m-1})\right] = cum\left[v^m(t)\right] \delta(\tau_1) \ldots \delta(\tau_{m-1}), \tag{1.17}$$

where $\delta(\cdot)$ stands for the *Dirac's impulse* function. We thus denote $\gamma_{m,v} = cum\left[\nu^m(t)\right]$ and, assuming $\upsilon(t)$ is zero-mean, we get

$$\gamma_{2,v} = cum\left[v^2(t)\right] = C_{2,v}(0) = \mathbb{E}\left\{v^2(t)\right\}, \tag{1.18}$$

$$\gamma_{3,v} = cum\left[v^3(t)\right] = C_{3,v}(0,0) = \mathbb{E}\left\{v^3(t)\right\}, \tag{1.19}$$

$$\gamma_{4,v} = cum\left[v^4(t)\right] = C_{4,v}(0,0,0) = \mathbb{E}\left\{v^4(t)\right\} - 3\mathbb{E}\left\{v^2(t)\right\}^2, \tag{1.20}$$

where $\gamma_{2,v}$, $\gamma_{3,v}$ and $\gamma_{4,v}$ stand respectively for the *variance*, the *skewness* and the *kurtosis* of the random process $\upsilon(t)$.

The *skewness* measures the lack of symmetry in a given distribution. It equals zero for symmetric (centered) distributions. A normalized version of the skewness is reported in the literature as follows:

$$\bar{\gamma}_{3,v} = \frac{C_{3,v}(0,0)}{\left(C_{2,v}(0)\right)^{3/2}}.$$

The *kurtosis* is a measure of *flatness* and, indirectly, of gaussianity. It equals zero for Gaussian processes and has positive or negative value depending on whether the probability density of the process is peaked (over-Gaussian) or flat (under-Gaussian) with respect to a normal distribution, respectively. A normalized version of the kurtosis is defined as

$$\bar{\gamma}_{4,v} = \frac{C_{4,v}(0,0,0)}{\left(C_{2,v}(0)\right)^{2}}.$$

2. *Symmetries*: In addition to properties P1 to P4, cumulants and polyspectra present several symmetry properties. Using property P3, we notice from (1.11), that $C_{m,v}(\tau_1, \ldots, \tau_{m-1}) = C_{m,v}(\tau_{n_1}, \ldots, \tau_{n_{m-1}})$, where $\{n_1, \ldots, n_k\}$ can be any permutation of the set $\{1, \ldots, m-1\}$. Hence, there are $(m-1)!$ different ways to order the time-lags τ_i yielding the same cumulant value. In addition, notation (1.11) was defined with respect to the time-shift t_0, using the convention $\tau_i = t_i - t_0$, $i \in [1, n-1]$. However, any other choice of t_i, $i \in [1, n-1]$, should lead to the same result, giving us m additional ways to get the same cumulant. In conclusion, mth-order cumulants define a representation space $\mathbb{C}^{\tau_1 \times \cdots \times \tau_{m-1}}$ in which the function $C_{m,v}(\tau_1, \ldots, \tau_{m-1})$ has $m(m-1)! = m!$ regions of symmetry, each one containing all the mth-order cumulant information and providing no additional information with respect to the other regions. For instance, when $m = 3$ each cumulant appears $3! = 6$ times in the space $\mathbb{R}^{\tau_1 \times \tau_2}$, as illustrated in fig. 1.1. Each of these 6 regions of symmetry in the plane $\tau_1 \times \tau_2$ contains all the non-redundant 3rd-order information. For $m = 4$, the space $\mathbb{R}^{\tau_1 \times \tau_2 \times \tau_3}$ is divided into $4! = 24$ redundant regions of symmetry.

3. *Cumulants of complex processes:* The above formulas were stated for the case of real-valued random processes. In the complex case, the random variables may be conjugated or not. Starting from equation (1.11), where no conjugations were made, we can state several definitions until reaching the one where all the variables are conjugated. Throughout the
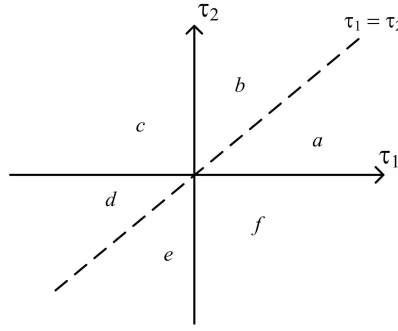
Figure 1.1: Regions of symmetry for the 3rd-order cumulants in the plane $\tau_1 \times \tau_2$.

rest of this work, except when otherwise stated, we will employ a particular definition[1] in which the random variables are alternately conjugated, so that (1.11) becomes:

$$C_{m,v}(\tau_1, \ldots, \tau_{m-1}) = \begin{cases} cum\left[v^*(t), v(t+\tau_1), \ldots, v^*(t+\tau_{m-2}), v(t+\tau_{m-1})\right], & \text{if } m \text{ is even;} \\ cum\left[v^*(t), v(t+\tau_1), \ldots, v(t+\tau_{m-2}), v^*(t+\tau_{m-1})\right], & \text{if } m \text{ is odd,} \end{cases}$$
(1.21)

and, according to the above, $\gamma_{m,v}$ is denoted as follows:

$$\gamma_{m,v} = C_{m,v}(0, 0, \ldots, 0).$$
(1.22)

4. *Second-order moments and cumulants*: using (1.21) and the relationship (1.4), second-order moments and cumulants ($m = 2$) can be written as

$$M_{2,v}(\tau) = \mathbb{E}\left\{v^*(t)v(t+\tau)\right\} = cum\left[v^*(t), v(t+\tau)\right] = C_{2,v}(\tau),$$
(1.23)

which is the *autocorrelation* function of $v(t)$. We also define

$$M_{2,v}^{(d)}(\tau) = \mathbb{E}\left\{v(t)v(t+\tau)\right\} = cum\left[v(t), v(t+\tau)\right] = C_{2,v}^{(d)}(\tau),$$
(1.24)

so that $M_{2,v}^{(d)*}(\tau) = \mathbb{E}\left\{v^*(t)v^*(t+\tau)\right\} = \left(\mathbb{E}\left\{v(t)v(t+\tau)\right\}\right)^* = C_{2,v}^{(d)*}(\tau)$.

5. *Fourth-order cumulants*: For centered processes, when $m = 4$, relation (1.4) reduces to (1.6). Recalling $v(t)$ is assumed zero-mean stationary, and using the notation (1.21), we get the following:

$$\begin{aligned} C_{4,v}(\tau_1, \tau_2, \tau_3) &= cum\left[v^*(t), v(t+\tau_1), v^*(t+\tau_2), v(t+\tau_3)\right] \\ &= \mathbb{E}\left\{v^*(t)v(t+\tau_1)v^*(t+\tau_2)v(t+\tau_3)\right\} - \\ &\quad \mathbb{E}\left\{v^*(t)v(t+\tau_1)\right\}\mathbb{E}\left\{v^*(t+\tau_2)v(t+\tau_3)\right\} - \\ &\quad \mathbb{E}\left\{v^*(t)v^*(t+\tau_2)\right\}\mathbb{E}\left\{v(t+\tau_1)v(t+\tau_3)\right\} - \\ &\quad \mathbb{E}\left\{v^*(t)v(t+\tau_3)\right\}\mathbb{E}\left\{v(t+\tau_1)v^*(t+\tau_2)\right\} \end{aligned}$$
(1.25)

---

[1] The use of this definition is motivated by the fact that for certain signals of interest fourth-order cumulants are zero when an odd number of conjugated terms is taken into account.

and consequently

$$
\begin{aligned}
C_{4,v}(\tau_1,\tau_2,\tau_3) \; = \; & M_{4,v}(\tau_1,\tau_2,\tau_3) - M_{2,v}(\tau_1)M_{2,v}(\tau_3-\tau_2) - \\
& M_{2,v}^{(d)*}(\tau_2)M_{2,v}^{(d)}(\tau_3-\tau_1) - M_{2,v}(\tau_3)M_{2,v}(\tau_1-\tau_2).
\end{aligned}
\tag{1.26}
$$

Replacing $\tau_1 = \tau_2 = \tau_3 = 0$ in (1.26) we can re-write (1.20) as

$$
\begin{aligned}
\gamma_{4,v} \; = \; & C_{4,v}(0,0,0) = M_{4,v}(0,0,0) - 2M_{2,v}^2(0) - \left| M_{2,v}^{(d)}(0) \right|^2 \\
= \; & \mathbb{E}\left\{ |v(t)|^4 \right\} - 2\mathbb{E}\left\{ |v(t)|^2 \right\}^2 - \left| \mathbb{E}\left\{ v^2(t) \right\} \right|^2
\end{aligned}
\tag{1.27}
$$

6. *Gaussianity*: Cumulants of *any* order greater than two of Gaussian processes are *zero* [1, 2]. Hence, cumulants can be viewed as a measure of the distance of a process from gaussianity.

7. *Circularity*: Let us define the complex-valued random vector $\mathbf{z} = [z_1 \ldots z_k]^{\mathsf{T}}$. The vector $\mathbf{z}$ is said to be *circular* if and only if

$$
\Omega_{\mathbf{z}}(e^{j\theta}\mathbf{u}) = \Omega_{\mathbf{z}}(e^{\mathbf{u}}), \qquad \forall\,\theta, \quad j = \sqrt{-1},
\tag{1.28}
$$

which means that the moments of the variables $\mathbf{z}$ and $e^{j\theta}\mathbf{z}$ are equal. The circularity of a complex random variable can therefore be viewed as the invariance of its probability density to a rotation of an angle θ. In particular, for a scalar complex circular random variable, it follows that moments and cumulants with a different number of conjugated and non-conjugated terms are zero, e.g. $\mathbb{E}\{z\} = 0$, $\mathbb{E}\{z^2\} = 0$, $\mathbb{E}\{z^2 z^*\} = 0$ and so on.

1.2 Estimation of moments and cumulants from real data

Practical applications of high-order statistics require the use of methods for estimating their values from the available data. In this context, *ergodicity*[2] is key assumption, allowing us to estimate moments from finite data sequences. Therefore, disposing of N data samples of a centered random variable $v(t)$, the simplest estimator of the mth-order moments of $v(t)$ is given as follows:

$$
\hat{M}_{m,v}(\tau_1,\ldots,\tau_{m-1}) =
\begin{cases}
\frac{1}{n_2-n_1+1} \sum\limits_{n=n_1}^{n_2} v^*(t)v(t+\tau_1)\cdots v^*(t+\tau_{m-2})v(t+\tau_{m-1}), & \text{if } m \text{ is even;} \\[2ex]
\frac{1}{n_2-n_1+1} \sum\limits_{n=n_1}^{n_2} v^*(t)v(t+\tau_1)\cdots v(t+\tau_{m-2})v^*(t+\tau_{m-1}), & \text{if } m \text{ is odd,}
\end{cases}
\tag{1.29}
$$

where $n_1 = \max(0, -\tau_1, \ldots, -\tau_{m-1})$ and $n_2 = \min(N, N-\tau_1, \ldots, N-\tau_{m-1})$. Let us also define

$$
\hat{M}_{m,v}^{(d)}(\tau_1,\ldots,\tau_{m-1}) = \frac{1}{n_2-n_1+1} \sum_{n=n_1}^{n_2} v(t)v(t+\tau_1)\cdots v(t+\tau_{m-1}), \quad \forall\, m \in \mathbb{N}^*.
\tag{1.30}
$$

---

[2] Statistical expectations of ergodic processes coincide with their time averages

Furthermore, in order to evaluate the performance of an estimator $\hat{\vartheta}$ of a given deterministic quantity ϑ, the mean, variance and bias of the estimator are useful metrics, and we define them as follows, respectively:

$$\mu_{\hat{\vartheta}} = \mathbb{E}\left\{\hat{\vartheta}\right\} \tag{1.31}$$

$$\sigma_{\hat{\vartheta}} = \mathbb{E}\left\{\left(\hat{\vartheta} - \mu_{\hat{\vartheta}}\right)^2\right\} \tag{1.32}$$

$$bias\left(\hat{\vartheta}\right) = \mathbb{E}\left\{\hat{\vartheta}\right\} - \vartheta \tag{1.33}$$

Taking $\vartheta = M_{m,v}(\tau_1, \ldots, \tau_{m-1})$, we notice that the mth-order moment estimator defined in (1.29) is *non-biased*, since $\mathbb{E}\left\{\hat{M}_{m,v}(\tau_1, \ldots, \tau_{m-1})\right\} = M_{m,v}(\tau_1, \ldots, \tau_{m-1})$. In addition, (1.29) is also said to be a consistent estimator because

$$\lim_{N \to \infty} \sigma_{\hat{M}_{m,v}(\tau_1, \ldots, \tau_{m-1})} = 0. \tag{1.34}$$

In this thesis, we are particularly interested in the case where $m = 4$. In order to estimate 4th-order cumulants, we first get the 2nd- and 4th-order moments estimates using (1.29) and (1.30), which yields

$$\begin{cases} \hat{M}_{4,v}(\tau_1, \tau_2, \tau_3) = \frac{1}{n_2 - n_1 + 1} \sum\limits_{n=n_1}^{n_2} v^*(n)v(n + \tau_1)v^*(n + \tau_2)v(n + \tau_3) \\ \hat{M}_{2v}(\tau) = \frac{1}{n_2 - n_1 + 1} \sum\limits_{n=n_1}^{n_2} v^*(n)v(n + \tau) \\ \hat{M}_{2v,d}(\tau) = \frac{1}{n_2 - n_1 + 1} \sum\limits_{n=n_1}^{n_2} v(n)v(n + \tau), \end{cases} \tag{1.35}$$

and then $\hat{C}_{4,v}(\tau_1, \tau_2, \tau_3)$ is obtained from the relationship (1.26). It is possible to show that the cumulant estimator is biased, but its bias tends to zero as N tends to infinity (c.f. [47] and references therein). In addition, $\hat{C}_{4,v}(\tau_1, \tau_2, \tau_3)$ is said to be a consistent estimator, since its variance goes to zero as N goes to infinity.

For polyspectra estimations, cumulant estimates must be obtained first and then converted to the frequency-domain, using (1.12). Further information on cumulant and polyspectra estimation can be found in [17, 48, 47].

2 Linear algebraic tools

In this section, we consider the computation of eigenvectors that simultaneously satisfy a number of given matrices. This so-called *generalized eigenproblem* consists in the search for solutions revealing the *common* eigenstructure of a set of symmetric matrices. This problem is of great importance in the domain of blind source separation but we also find applications in the field of system identification and equalization.

2.1 The symmetric-definite eigenvalue problem

Eigenanalysis is a basic algebraic tool in any domain of Signal Processing. The search for the nontrivial solutions of the linear system $\mathbf{Ax} = \lambda\mathbf{x}$ gave rise to enormous developments since the Jacobi's works in the middle of the 19th century [49]. In particular, computing the eigenvalues of a matrix $\mathbf{A}$ consists in finding the roots of the characteristic equation:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0. \tag{1.36}$$

However, in most of the cases, explicitly solving the characteristic equation is an *ill-conditioned* problem that should be avoided. Moreover, if there exist solutions to (1.36), those are necessarily iterative because there is no closed-form expression for the roots of a general polynomial of degree $n > 4$. As a result, numerical solutions and stability issues have been strongly addressed in the literature aiming an accurate and efficient computation of eigenvalues and eigenvectors.

A straightforward solution to this problem consists in a Schur-type decomposition of the square matrix $\mathbf{A}$ to the form $\mathbf{A} = \mathbf{Q}\mathbf{L}^\mathsf{T}\mathbf{Q}^\mathsf{H}$, where $\mathbf{Q}$ is unitary[3] and $\mathbf{L}$ is lower triangular. Several zero-introducing methods are reported in the literature for implementing this *triangularization* strategy, including Householder, Givens and Gram-Schmidt [50, 49]. A special case occurs when $\mathbf{A}$ is normal, i.e. it commutes with its conjugate-transpose ($\mathbf{AA}^\mathsf{H} = \mathbf{A}^\mathsf{H}\mathbf{A}$). A normal matrix is unitarily diagonalizable i.e. there exists a unitary matrix $\mathbf{Q}$ such that $\mathbf{Q}^\mathsf{H}\mathbf{AQ} = \mathbf{\Delta}$, where $\mathbf{\Delta}$ is a diagonal matrix with the eigenvalues of $\mathbf{A}$ composing its main diagonal. This *diagonalization* approach is easily implemented by means of successive applications of unitary similarity transformations.

Next, we briefly discuss these two classes of techniques that cope with the vast majority of the cases concerned in signal processing applications. Notice, however, that the choice of an adequate algorithm depends on a number of characteristics of the concerned matrix. In particular, some features to be considered involve main properties (symmetric, unitary, Hermitian, etc), structure (sparse, Toeplitz, etc) and type of elements (real or complex) among others.

Triangularization strategies

A valuable approach for computing the eigenvalues of a square matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ consists in reducing $\mathbf{A}$ to a triangular form by means of unitary similarity transformations. This solution is based on the *Schur's Unitary Triangularization Theorem*, which states that there exists a unitary $n \times n$ matrix $\mathbf{Q}$ such that $\mathbf{Q}^\mathsf{H}\mathbf{AQ} = \mathbf{L}^\mathsf{T}$ is upper triangular. Since the determinant of a triangular matrix equals the product of the diagonal entries, the eigenvalues of $\mathbf{A}$ are given by the diagonal elements of $\mathbf{L}$, i.e. $\lambda_i = [\mathbf{L}]_{ii}$. Furthermore, provided that $\mathbf{A}$ is full rank, the columns of $\mathbf{Q}$ form an orthonormal basis for the column space of $\mathbf{A}$. Therefore, the Schur decomposition can be viewed as a way to compute eigenvalues and eigenvectors of $\mathbf{A}$, as it is shown in Appendix A.1.

---

[3] A matrix $\mathbf{Q} \in \mathbb{C}^{n \times n}$ is unitary iff $\mathbf{Q}\mathbf{Q}^\mathsf{H} = \mathbf{Q}^\mathsf{H}\mathbf{Q} = \mathbf{I}$. As a result, if $\mathbf{Q}$ is unitary then $\mathbf{Q}^\mathsf{H} = \mathbf{Q}^{-1}$.

Diagonalization strategies

Let $\boldsymbol{\Delta}$ be unitarily similar to $\mathbf{A}$, i.e. $\boldsymbol{\Delta} = \mathbf{Q}^{\mathsf{H}}\mathbf{A}\mathbf{Q}$, with $\mathbf{Q}$ unitary. If $\mathbf{A}$ is normal then also is $\boldsymbol{\Delta}$, i.e. if $\mathbf{A}\mathbf{A}^{\mathsf{H}} = \mathbf{A}^{\mathsf{H}}\mathbf{A}$ then $\boldsymbol{\Delta}\boldsymbol{\Delta}^{\mathsf{H}} = \boldsymbol{\Delta}^{\mathsf{H}}\boldsymbol{\Delta}$. On the other hand, Schur's Triangularization Theorem says that a unitary $\mathbf{Q}$ exists for which $\boldsymbol{\Delta}$ is upper triangular. Notice, however, that a matrix that is both normal and upper triangular can only be diagonal and hence, the Schur factorization actually *diagonalizes* normal matrices. Therefore, diagonalization techniques can be used to determine the eigenstructure of a normal matrix. In particular, Hermitian matrices (symmetric in the real case) often receive special attention since they have real eigenvalues.

Numerical solutions to this problem are known to be iterative and the idea is based on the repeated application of unitary similarity transformations so that matrix $\mathbf{A}$ is systematically changed toward a diagonal form until reaching a tolerance level with respect to a certain criterion. Such a criterion must be carefully defined in order to stop iteration at a point where $\mathbf{A}$ is close enough to being diagonal. A classical measure of how much a matrix differs from being diagonal, based on its non-diagonal values, is given in [50], as follows:

$$off(\mathbf{A}) \triangleq \left(\sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|^2 \right)^{1/2} \tag{1.37}$$

$$= \left(\|\mathbf{A}\|_F^2 - \|diag(\mathbf{A})\|^2 \right)^{1/2}, \tag{1.38}$$

where

$$\|\mathbf{A}\|_F \triangleq \left(\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2} \tag{1.39}$$

is the *Frobenius* norm of $\mathbf{A}$, $diag(\mathbf{A})$ is the vector consisting of the diagonal elements of $\mathbf{A}$ and $\|\cdot\|$ stands for the *Euclidean* norm. Note that $\|\mathbf{A}\|_F^2 = Tr(\mathbf{A}\mathbf{A}^{\mathsf{H}})$ and, since the trace of a matrix is invariant under a similarity transformation, the following holds for any unitary $n \times n$ matrix $\mathbf{Q}$:

$$\|\mathbf{A}\|_F^2 = Tr(\mathbf{A}\mathbf{A}^{\mathsf{H}}) = Tr(\mathbf{Q}^{\mathsf{H}}\mathbf{A}\mathbf{A}^{\mathsf{H}}\mathbf{Q}) = Tr(\mathbf{Q}^{\mathsf{H}}\mathbf{A}\mathbf{Q}\mathbf{Q}^{\mathsf{H}}\mathbf{A}^{\mathsf{H}}\mathbf{Q}) = \|\mathbf{Q}^{\mathsf{H}}\mathbf{A}\mathbf{Q}\|_F^2. \tag{1.40}$$

Therefore, the Frobenius norm is also preserved under a similarity transformation. As a result, we have the following relationship

$$\|\mathbf{A}\|_F^2 = off\left(\mathbf{Q}^{\mathsf{H}}\mathbf{A}\mathbf{Q}\right)^2 + \left\|diag\left(\mathbf{Q}^{\mathsf{H}}\mathbf{A}\mathbf{Q}\right)\right\|^2. \tag{1.41}$$

Once the sum in the right-hand side of the above equation is constant it is straightforward to conclude that minimizing the norm of the off-diagonal terms implies maximizing the norm of the diagonal elements and vice-versa.

Diagonalization methods appear among the earliest solutions for the eigenvalue problem and have been formerly addressed by Jacobi, whose seminal ideas became classic and very attractive

for application in parallel computing. In Appendix A.2, we describe the classical Jacobi method for computing similarity transformations that make $\mathbf{A}$ closer and closer to being diagonal, in an iterative procedure.

In the next section, we show that the simultaneous diagonalization of two symmetric matrices is equivalent to a generalized eigenvalue problem. Under certain conditions, this problem can be viewed as a generalization of the *Lower-Diagonal-Upper* factorization ($\mathbf{L}\boldsymbol{\Delta}\mathbf{M}^\mathsf{T}$) and it indicates a link between the EVD techniques and the optimization methods [51, 52]. We will also show how this problem can be extended to a set of K symmetric matrices.

2.2   The generalized eigenvalue problem

Given a Hermitian matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ and a Hermitian positive-definite matrix $\mathbf{B} \in \mathbb{C}^{n\times n}$, we want to find a factor $\mathbf{Q}$ such that $\mathbf{Q}^\mathsf{H}\mathbf{A}\mathbf{Q}$ and $\mathbf{Q}^\mathsf{H}\mathbf{B}\mathbf{Q}$ are diagonal $n \times n$ matrices. We start by looking for an intermediate transformation $\mathbf{W}_1$, referred to as *whitening* transformation, such that

$$\mathbf{W}_1^\mathsf{H}\mathbf{B}\mathbf{W}_1 = \mathbf{I}. \tag{1.42}$$

In the above equation, $\mathbf{W}_1$ can be computed by several means including classical singular value decomposition (SVD) or eigenvalue decomposition (EVD). Let, for instance, the EVD of $\mathbf{B}$ be given by $\boldsymbol{\Gamma}^\mathsf{H}\mathbf{B}\boldsymbol{\Gamma} = \boldsymbol{\Lambda}_B$, then

$$\mathbf{W}_1 = \boldsymbol{\Gamma}\boldsymbol{\Lambda}_B^{-1/2}, \tag{1.43}$$

where $\boldsymbol{\Lambda}_B^{-1/2}$ is real since $\mathbf{B}$ is assumed Hermitian and positive-definite. However, we search a factor $\mathbf{Q}$ that is supposed to diagonalize both $\mathbf{A}$ and $\mathbf{B}$, simultaneously. Hence, the whitening transformation $\mathbf{W}_1$ must also be applied to $\mathbf{A}$, which yields $\mathbf{A}_1 = \mathbf{W}_1^\mathsf{H}\mathbf{A}\mathbf{W}_1$. We can now diagonalize $\mathbf{A}_1$ by computing a second transformation $\mathbf{W}_2$, as follows,

$$\mathbf{W}_2^\mathsf{H}\mathbf{A}_1\mathbf{W}_2 = \mathbf{W}_2^\mathsf{H}\mathbf{W}_1^\mathsf{H}\mathbf{A}\mathbf{W}_1\mathbf{W}_2 = \boldsymbol{\Delta}, \tag{1.44}$$

where $\mathbf{W}_2^\mathsf{H}\mathbf{A}_1\mathbf{W}_2 = \boldsymbol{\Delta}$ is the EVD of $\mathbf{A}_1$ with $\boldsymbol{\Delta}$ being a diagonal matrix and $\mathbf{W}_2$ a unitary transformation (because $\mathbf{A}_1$ is Hermitian). We conclude that $\mathbf{Q} = \mathbf{W}_1\mathbf{W}_2$ *jointly* diagonalizes both $\mathbf{A}$ and $\mathbf{B}$, so that

$$\begin{cases} \mathbf{Q}^\mathsf{H}\mathbf{A}\mathbf{Q} = \boldsymbol{\Delta} \\ \mathbf{Q}^\mathsf{H}\mathbf{B}\mathbf{Q} = \mathbf{I}, \end{cases} \tag{1.45}$$

where we have used the fact that $\mathbf{W}_2$ is unitary to obtain the latter equation. Furthermore, $\mathbf{Q}\mathbf{Q}^\mathsf{H} = \mathbf{W}_1\mathbf{W}_2\mathbf{W}_2^\mathsf{H}\mathbf{W}_1^\mathsf{H} = \mathbf{W}_1\mathbf{W}_1^\mathsf{H}$ and, from (1.43), we get $\mathbf{W}_1\mathbf{W}_1^\mathsf{H} = \boldsymbol{\Gamma}\boldsymbol{\Lambda}_B^{-1}\boldsymbol{\Gamma}^\mathsf{H}$, hence $\mathbf{Q}\mathbf{Q}^\mathsf{H} = \mathbf{B}^{-1}$. Thus, we can handle (1.45) to get $(\mathbf{B}\mathbf{Q})\mathbf{Q}^\mathsf{H}\mathbf{A}\mathbf{Q} = (\mathbf{B}\mathbf{Q})\boldsymbol{\Delta}$ and it follows that

$$\mathbf{A}\mathbf{Q} = \mathbf{B}\mathbf{Q}\boldsymbol{\Delta}. \tag{1.46}$$

The above equation states a generalized eigenvalue problem [50]. The columns of $\mathbf{Q}$ form a basis of eigenvectors associated with the generalized eigenvalues, disposed at the diagonal entries of $\mathbf{\Delta}$. Moreover, since $\mathbf{B}$ is nonsingular[4], we have

$$\mathbf{B}^{-1}\mathbf{A}\mathbf{Q} = \mathbf{Q}\mathbf{\Delta}. \tag{1.47}$$

Therefore, equation (1.47) shows that the diagonalizing factor $\mathbf{Q}$ exists and can be exactly obtained from the EVD of $\mathbf{B}^{-1}\mathbf{A}$ or, equivalently, from two consecutive matrix decompositions as indicated by (1.43) and (1.44), avoiding matrix inversion. This problem is referred to as the *symmetric-definite generalized eigenproblem* and a number of algorithms are available in the literature to compute $\mathbf{Q}$ satisfying (1.45).

Simultaneous diagonalization

We have just considered the simultaneous diagonalization of two symmetric matrices ($K = 2$). We now consider the general case where $K > 2$. In other words, we search the nonsingular factor $\mathbf{Q}$ that simultaneously diagonalizes a set $\mathbb{A}$ of symmetric matrices $\mathbf{A}^{(k)} \in \mathbb{C}^{n \times n}$, $k = 1, \ldots, K$. For $K > 2$, it does not necessarily exist a common set of eigenvectors and $\mathbf{Q}$ is said to reveal the *average eigenstructure* of the set $\mathbb{A}$ [53]. Existing numerical methods can be used to compute an orthonormal basis $\mathbf{Q} = [\mathbf{Q}_{\cdot 1} \cdots \mathbf{Q}_{\cdot n}]$ that approximately diagonalizes the matrices $\mathbf{A}^{(k)}$, as best as possible, following a Jacobi-like approach. The idea behind these methods is to optimize a cost function aiming to minimize the off-diagonal elements of $\mathbf{A}^{(k)}$, so that all matrices in the set $\mathbb{A}$ become systematically and simultaneously closer to being diagonal. A very straightforward way to define the cost function describing this criterion is as follows:

$$\mathsf{J}(\mathbf{Q}, \mathbb{A}) \triangleq \sum_{k=1}^{K} \mathit{off}\left(\mathbf{Q}^{\mathsf{H}}\mathbf{A}^{(k)}\mathbf{Q}\right) \tag{1.48}$$

where the operator $\mathit{off}(\cdot)$ is given in (1.38) and the matrix $\mathbf{Q}$ that minimizes (1.48) is referred to as *joint diagonalizer* of the set $\mathbb{A}$ [53].

From the previous section, recall that minimizing the norm of the off-diagonal terms of a matrix implies maximizing the norm of its diagonal elements, as suggested by (1.41). As a result, computing a joint diagonalizer $\mathbf{Q}$ to the set $\mathbb{A}$ by minimizing (1.48) is equivalent to *maximizing* the following criterion

$$\mathsf{J}(\mathbf{Q}, \mathbb{A}) \triangleq \sum_{k=1}^{K} \left\| \mathit{diag}\left(\mathbf{Q}^{\mathsf{H}}\mathbf{A}^{(k)}\mathbf{Q}\right) \right\|^{2}. \tag{1.49}$$

---

[4] If $\mathbf{B}$ is singular, the method discussed here can be applied by using a rank-reduction technique in such a way that the first $r_{\mathbf{B}} = \mathit{rank}(\mathbf{B}) < n$ columns of $\mathbf{\Gamma}$ (eigenvectors of $\mathbf{B}$) are used to form $\mathbf{\Gamma}_{r_{\mathbf{B}}}$ and then (1.43) becomes $\mathbf{\Gamma}_{r_{\mathbf{B}}} \mathit{diag}(\lambda_1, \ldots, \lambda_{r_{\mathbf{B}}})^{-1/2} = \mathbf{W}_1 \in \mathbb{C}^{n \times r_{\mathbf{B}}}$. Thus, $\mathbf{W}_2 \in \mathbb{C}^{r_{\mathbf{B}} \times r_{\mathbf{B}}}$ and hence $\mathbf{W}_1\mathbf{W}_2 = \mathbf{Q} \in \mathbb{C}^{n \times r_{\mathbf{B}}}$. In this case, (1.47) becomes $\mathbf{B}^{\#}\mathbf{A}\mathbf{Q} = \mathbf{Q}\mathbf{\Delta}$, where $\mathbf{B}^{\#}$ is the pseudo-inverse of $\mathbf{B}$ defined as $\mathbf{B}^{\#} = \mathbf{\Gamma}_{r_{\mathbf{B}}} \mathit{diag}(\lambda_1, \ldots, \lambda_{r_{\mathbf{B}}})^{-1} \mathbf{\Gamma}_{r_{\mathbf{B}}}^{\mathsf{H}}$ so that $\mathbf{B}\mathbf{B}^{\#} = \mathbf{B}^{\#}\mathbf{B} = \mathbf{\Gamma}_{r_{\mathbf{B}}}\mathbf{\Gamma}_{r_{\mathbf{B}}}^{\mathsf{H}}$. This procedure is referred to as *reduced-rank simultaneous diagonalization*.

For $K = 1$, the above problem corresponds to a simple matrix diagonalization and we can easily compute $\mathbf{Q}$ using a classical Jacobi algorithm. The same holds for $K = 2$. For $K > 2$, however, orthogonal similarity transformations might not be able to produce K perfectly diagonal matrices. An extended version of the Jacobi diagonalization algorithm was proposed in [51, 52] for performing an *approximative joint diagonalization* of the matrices in $\mathbb{A}$ by maximizing (1.49) through the successive application of plane rotations (*Givens* rotations). The extended Jacobi algorithm for approximative simultaneous diagonalization is described in Appendix A.2.

Least Squares equivalence

The criterion (1.49) for the simultaneous diagonalization of the set $\mathbb{A}$ is shown to be equivalent to the following:

$$\mathsf{J}\left(\mathbf{Q}, \mathbb{A}\right) = \sum_{k=1}^{K} \left\| \mathbf{A}^{(k)} - \mathbf{Q}\boldsymbol{\Delta}_k \mathbf{Q}^{\mathsf{H}} \right\|_F^2. \tag{1.50}$$

Although the original proposition of Cardoso and Souloumiac [51] was merely intuitive, the above result, first demonstrated by Wax [54], proves that joint diagonalization actually coincides with a least squares problem and it is therefore optimal in that sense.

3   Multilinear algebraic tools

3.1   Parafac tensor decomposition

The Parallel Factor (Parafac) analysis of a Pth order tensor with rank F consists in the decomposition of the tensor into a sum of F rank-one tensors, each one being written as an outer product of P vectors [16]. Let us consider the Pth-order tensor $\mathcal{T}^{(P)}$ of dimensions $I_1 \times \ldots \times I_P$ having the following F-component decomposition:

$$t_{i_1 \ldots i_P} = \sum_{f=1}^{F} a_{i_1 f}^{(1)} \ldots a_{i_P f}^{(P)} \tag{1.51}$$

where $i_p \in [1, I_p]$, with $p \in [1, P]$. The sum expressed in (1.51) is the scalar representation of the Parafac decomposition of tensor $\mathcal{T}^{(P)}$. The rank of a tensor is defined as the minimum number F of factors needed to decompose it in the form (1.51). The tensor $\mathcal{T}^{(P)}$ can be written as the sum of F outer products[5] involving P vectors, as follows:

$$\begin{aligned} \mathcal{T}^{(P)} &= \sum_{i_1=1}^{I_1} \cdots \sum_{i_P=1}^{I_P} t_{i_1 \ldots i_P} \mathbf{e}_{i_1}^{(I_1)} \circ \cdots \circ \mathbf{e}_{i_P}^{(I_P)} \\ &= \sum_{f=1}^{F} \mathbf{A}_{\cdot f}^{(1)} \circ \cdots \circ \mathbf{A}_{\cdot f}^{(P)}, \end{aligned} \tag{1.52}$$

---

[5]The outer product of two arrays $\mathcal{A}^{(P)} \in \mathbb{C}^{I_1 \times \ldots \times I_P}$ and $\mathcal{B}^{(Q)} \in \mathbb{C}^{J_1 \times \ldots \times J_Q}$ consists of a tensor of order $P + Q$ in which the element in position $i_1, i_2, \ldots, i_P, j_1, j_2, \ldots, j_Q$ equals the product $a_{i_1 i_2 \ldots i_P} b_{j_1 j_2 \ldots j_Q}$.

where the notation $\mathbf{e}_{i_p}^{(I_p)}$, $i_p \in [1, I_p]$, $p \in [1, P]$, stands for the i_pth canonical basis vector of $\mathbb{R}^{I_p}$, i.e. $\mathbf{e}_{i_p}^{(I_p)} = [0, \dots, 0, 1, 0, \dots, 0]^{\mathsf{T}} \in \mathbb{R}^{I_p \times 1}$, with the nonzero element '1' placed at the i_pth entry. In addition, the P matrices $\mathbf{A}^{(p)} \in \mathbb{C}^{I_p \times F}$, $p \in [1, P]$, formed of the elements $a_{i_p f}^{(p)}$, $i_p \in [1, I_p]$, $f \in [1, F]$, contain all the tensor information and will be referred to as (canonical) Parafac components. The fth column of matrix $\mathbf{A}^{(p)}$ is defined as follows:

$$\mathbf{A}_{\cdot f}^{(p)} \triangleq \sum_{i_p=1}^{I_p} a_{i_p f}^{(p)} \mathbf{e}_{i_p}^{(I_p)}, \quad f \in [1, F]. \tag{1.53}$$

We define a d-dimensional *slice* of tensor $\mathcal{T}^{(P)}$ as the set of elements obtained by freezing $P - d$ of its P indexes and making the d other ones to vary in their respective ranges. As a result, *one-dimensional* (1D) tensor slices can be viewed as vectors and *two-dimensional* (2D) tensor slices are matrices.

Establishing conditions to ensure uniqueness of the Parafac decomposition is of major importance. Uniqueness represents a great advantage of Parafac over matrix decompositions, since Parafac does not produce rotational ambiguities. In addition, there are generally no orthogonality constraints such as in SVD, even in the symmetric (Hermitian) case, where such constraints also apply to EVD [6]. In particular, the decomposition of a tensor $\mathcal{T}^{(P)}$ with components $\mathbf{A}^{(1)}, \dots, \mathbf{A}^{(P)}$ is said to be *essentially unique* if any other set of matrices $\{\bar{\mathbf{A}}^{(1)}, \dots, \bar{\mathbf{A}}^{(P)}\}$ satisfying (1.52) is such that

$$\bar{\mathbf{A}}^{(p)} = \mathbf{A}^{(p)} \mathbf{\Lambda}_p \mathbf{\Pi}, \quad \forall \, p \in [1, P], \tag{1.54}$$

where $\mathbf{\Lambda}_p$, $p \in [1, P]$, are diagonal scaling matrices satisfying

$$\prod_{p=1}^{P} \mathbf{\Lambda}_p = \mathbf{I}_F \tag{1.55}$$

and $\mathbf{\Pi}$ is an $F \times F$ permutation matrix [55]. In other words, essential uniqueness means uniqueness up to column scaling and permutation. A sufficient uniqueness condition has been stated by Kruskal in [20] for the case of a 3rd-order tensor. For a generic Pth-order tensor, Sidiropoulos and Bro extended the Kruskal Uniqueness Theorem as follows [38]:

Theorem 1.1 *The Parafac decomposition of a Pth-order tensor with rank $F > 1$, is essentially unique if*

$$\sum_{p=1}^{P} k_{A^{(p)}} \geq 2F + (P - 1), \tag{1.56}$$

where $k_{A^{(p)}}$ stands for the k-rank of the Parafac component $\mathbf{A}^{(p)}$, $p \in [1, P]$.

---

[6] It is well known that the SVD of a (complex) matrix yields a factorization of the type $\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}^{\mathsf{H}}$, with $\mathbf{D}$ diagonal and $\mathbf{U}$ and $\mathbf{V}$ unitary matrices. When dealing with Hermitian matrices, this orthogonality constraint also applies to EVD.

The k-rank of an $n \times m$ matrix $\mathbf{X}$ equals the largest integer k_X for which *any* set of k_X columns of $\mathbf{X}$ is independent. From this definition, we notice that $k_X \leq r_X \leq \min(n, m)$, where $r_X = rank(\mathbf{X})$. Several authors have addressed the Parafac uniqueness problem and different proofs have been given to the above theorem [20, 38, 56]. In addition, the Kruskal condition is shown to be necessary for $F = 2$ and $F = 3$ and, if $k_{A^{(p)}} = r_{A^{(p)}}$, $\forall p \in [1, P]$, then (1.56) is also necessary for $F = 4$ [38]. In the general case, some necessary uniqueness conditions include $k_{A^{(p)}} \neq 0$, $\forall p \in [1, P]$, which means that the Parafac components $\mathbf{A}^{(p)}$ should not have any all-zero column [57]. Implications of this fact will be further discussed later in this section.

In section 3.2 below, we discuss the important issue of estimating Parafac components. Specifically, we will introduce the basic principles for implementing the ALS algorithm under a general framework. To this end, we need to express tensor $\mathcal{T}^{(P)}$ using matrix representations. In which follows, we derive a generalized formulation for the matrix representations of a Pth-order (Parafac) tensor allowing us to extend trilinear estimation algorithms to the order P. Matrix representations are obtained by *unfolding* the tensor so that all the tensor elements are placed in a 2D array. We define the Pth unfolded tensor representation of $\mathcal{T}^{(P)}$ as a $(I_1 \cdots I_{P-1}) \times I_P$ matrix $\mathbf{T}_{[P]}$, such that $[\mathbf{T}_{[P]}]_{r,i_P} = t_{i_1 \ldots i_P}$ where the row number is given by $r = (i_1 - 1)I_2 \cdots I_{P-1} + (i_2 - 1)I_3 \cdots I_{P-1} + \ldots + (i_{P-2} - 1)I_{P-1} + i_{P-1}$.

Using the canonical basis vector notation, this is equivalent to write:

$$
\begin{aligned}
\mathbf{T}_{[P]} &\triangleq \sum_{i_1=1}^{I_1} \cdots \sum_{i_P=1}^{I_P} t_{i_1 \ldots i_P} \mathbf{e}_{(i_1-1)I_2 \cdots I_{P-1}+(i_2-1)I_3 \cdots I_{P-1}+\ldots+(i_{P-2}-1)I_{P-1}+i_{P-1}}^{(I_1 \cdots I_{P-1})} \mathbf{e}_{i_P}^{(I_P)\mathsf{T}} \\
&= \sum_{i_1=1}^{I_1} \cdots \sum_{i_P=1}^{I_P} t_{i_1 \ldots i_P} \left(\mathbf{e}_{i_1}^{(I_1)} \diamond \mathbf{e}_{i_2}^{(I_2)} \diamond \cdots \diamond \mathbf{e}_{i_{P-1}}^{(I_{P-1})} \right) \mathbf{e}_{i_P}^{(I_P)\mathsf{T}}
\end{aligned}
\tag{1.57}
$$

where we have used the fact that $\mathbf{e}_i^{(I)} \diamond \mathbf{e}_j^{(J)} = \mathbf{e}_{(i-1)J+j}^{(IJ)}$. Replacing (1.51) in the above equation and using definition (1.53), we easily get:

$$
\begin{aligned}
\mathbf{T}_{[P]} &= \sum_{f=1}^{F} \left(\mathbf{A}_{\cdot f}^{(1)} \diamond \cdots \diamond \mathbf{A}_{\cdot f}^{(P-1)} \right) \mathbf{A}_{\cdot f}^{(P)\mathsf{T}} \\
&= \left(\mathbf{A}^{(1)} \diamond \cdots \diamond \mathbf{A}^{(P-1)} \right) \mathbf{A}^{(P)\mathsf{T}} \in \mathbb{C}^{(I_1 \cdots I_{P-1}) \times I_P}.
\end{aligned}
\tag{1.58}
$$

The index P in $\mathbf{T}_{[P]}$ is clearly associated with the Parafac component $\mathbf{A}^{(P)}$, which is right-multiplied by the Khatri-Rao product in the above equation. This notation suggests that $\mathbf{A}^{(P)}$ can be estimated from $\mathbf{T}_{[P]}$, provided that initial estimates of $\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(P-1)}$ are given. Extending the above reasoning, we can define $P - 1$ other unfolded tensor representations, denoted $\mathbf{T}_{[p]}$, $p \in [1, P-1]$, which can be used to estimate the remaining components $\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(P-1)}$.

For instance, when $p = 1$ we have:

$$\mathbf{T}_{[1]} = \left(\mathbf{A}^{(2)} \diamond \cdots \diamond \mathbf{A}^{(P)}\right) \mathbf{A}^{(1)^{\mathsf{T}}} \in \mathbb{C}^{(I_2 \cdots I_P) \times I_1}. \tag{1.59}$$

The general formulation is given as follows:

$$\mathbf{T}_{[p]} = \left(\mathbf{A}^{(p+1)} \diamond \cdots \diamond \mathbf{A}^{(P)} \diamond \mathbf{A}^{(1)} \diamond \cdots \diamond \mathbf{A}^{(p-1)}\right) \mathbf{A}^{(p)^{\mathsf{T}}}, \quad p \in [2, P-1]. \tag{1.60}$$

Note that $\mathbf{T}_{[p]} \in \mathbb{C}^{(I_{p+1} \cdots I_P I_1 \cdots I_{p-1}) \times I_p}$. In practice, in order to form the column $[\mathbf{T}_{[p]}]_{.,i_p}$, we arrange the elements $t_{i_1 \ldots i_P}$ by fixing the index i_p while varying the first $p-1$ and the last $P-p$ indices in consecutive nested loops with i_P being the most inner one (fastest) and i_1 the most outer one (slowest).

Finally, we remark that for each $p \in [1, P]$, there exist $(P-1)!$ equivalent (but different) ways to define a matrix denoted $\mathbf{T}_{[p]}$ (by permuting the $P-1$ indices of the Parafac components in the multiple Khatri-Rao product of (1.60)). Actually, (1.60) is only one of these equivalent formulations. For the sake of a uniform notation, throughout the rest of this thesis, we convention to denote by $\mathbf{T}_{[p]}$ the matrix representation of tensor $\mathcal{T}^{(P)}$ obtained from the unfolding procedure above described, taking the order of the indices into account. As a consequence, equations (1.58) to (1.60) hold. Any other notation will be disregarded.

Among many algorithms proposing a solution to estimate the factors of the Parafac decomposition, the ALS algorithm is probably the most famous one. In the next section, we address this subject under a general framework by describing an ALS algorithm that estimates the Parafac components of a pth-order tensor. We will also briefly discuss the particular cases of $P = 4$ and $P = 3$, which yield quadrilinear and trilinear ALS algorithms, respectively. In chapter 2, we will exploit the redundancies in the factors of the 4th-order cumulant tensor decomposition in the minimization problem in order to develop new channel identification algorithms.

3.2 The Alternating Least Squares (ALS) algorithm

The main idea behind the ALS algorithm is to divide the parameters to be estimated into several sets in order to facilitate the use of simpler estimation algorithms. Then, each set is estimated by iteratively minimizing, in the least squares sense, a single cost function conditioned to the previous estimates of the other parameters. The algorithm iterates until no improvements are observed (c.f. [58] and references therein). In fact, the ALS algorithm is shown to be monotonically convergent, i.e. it can only improve or keep the same fit of the model. Its main drawbacks include slow convergence and possible convergence to local minima due to inadequate initializations. But these problems are more likely to occur in difficult cases[7].

---

[7] Situations with strongly correlated Parafac components or with too many components.

Using the unfolded tensor representations defined in the previous section, with the general formulation given in (1.60), we can define the following LS criteria:

$$\psi_p(\mathbf{A}^{(p)}) \triangleq \left\| \mathbf{T}_{[p]} - \left(\hat{\mathbf{A}}_r^{(p+1)} \diamond \cdots \hat{\mathbf{A}}_r^{(P)} \diamond \hat{\mathbf{A}}_{r-1}^{(1)} \diamond \cdots \hat{\mathbf{A}}_{r-1}^{(p-1)} \right) \mathbf{A}^{(p)\mathsf{T}} \right\|_F^2, \tag{1.61}$$

where r stands for the iteration number and, for notational convenience, we have omitted the dependence of ψ_p on $\hat{\mathbf{A}}_r^{(p+1)}, \ldots, \hat{\mathbf{A}}_r^{(P)}, \hat{\mathbf{A}}_{r-1}^{(1)}, \ldots, \hat{\mathbf{A}}_{r-1}^{(p-1)}$. Then, each Parafac component $\mathbf{A}_r^{(p)}$, $p \in [1, P]$, can be estimated by minimizing the function $\psi_p(\mathbf{A}^{(p)})$, assuming that previous estimates of the Parafac components $\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(p-1)}$ are available from the preceding iteration, while estimates of $\mathbf{A}^{(p+1)}, \ldots, \mathbf{A}^{(P)}$ have been previously obtained during the current iteration. The optimal LS solution to this problem is given by:

$$\begin{aligned} \hat{\mathbf{A}}_r^{(p)\mathsf{T}} &= \arg\min_{\mathbf{A}^{(p)}} \{\psi_p(\mathbf{A}^{(p)})\} \\ &= \left(\hat{\mathbf{A}}_r^{(p+1)} \diamond \cdots \hat{\mathbf{A}}_r^{(P)} \diamond \hat{\mathbf{A}}_{r-1}^{(1)} \diamond \cdots \hat{\mathbf{A}}_{r-1}^{(p-1)} \right)^{\#} \mathbf{T}_{[p]}, \quad p \in [1, P]. \end{aligned} \tag{1.62}$$

The Generalized Parafac-ALS algorithm is summarized below. We start with $p = P$ by assuming the initial guesses $\hat{\mathbf{A}}_0^{(1)}, \ldots, \hat{\mathbf{A}}_0^{(P-1)}$ are known. Random initialization can be a good choice, but it does not always yield a good first estimate [55].

Algorithm 1.1 (Generalized Parafac-ALS algorithm)

Determine a threshold $\varepsilon > 0$, initialize the Parafac components $\hat{\mathbf{A}}_0^{(1)}, \ldots, \hat{\mathbf{A}}_0^{(P-1)}$ and execute the steps below, starting with $r = 1$:

1. *Using (1.62) with $p = P$, compute $\hat{\mathbf{A}}_r^{(P)}$, using the estimates $\hat{\mathbf{A}}_{r-1}^{(1)}, \ldots, \hat{\mathbf{A}}_{r-1}^{(P-1)}$ from the preceding iteration, so that:*

$$\hat{\mathbf{A}}_r^{(P)\mathsf{T}} = \left(\hat{\mathbf{A}}_{r-1}^{(1)} \diamond \cdots \hat{\mathbf{A}}_{r-1}^{(P-1)} \right)^{\#} \mathbf{T}_{[P]};$$

2. *For $p = P - 1, \ldots, 2$, compute $\hat{\mathbf{A}}_r^{(p)}$ using the estimates $\hat{\mathbf{A}}_r^{(p+1)}, \ldots, \hat{\mathbf{A}}_r^{(P)}$ previously computed during the current iteration and $\hat{\mathbf{A}}_{r-1}^{(1)}, \ldots, \hat{\mathbf{A}}_{r-1}^{(p-1)}$ from the preceding iteration, as follows:*

$$\hat{\mathbf{A}}_r^{(p)\mathsf{T}} = \left(\hat{\mathbf{A}}_r^{(p+1)} \diamond \cdots \hat{\mathbf{A}}_r^{(P)} \diamond \hat{\mathbf{A}}_{r-1}^{(1)} \diamond \cdots \hat{\mathbf{A}}_{r-1}^{(p-1)} \right)^{\#} \mathbf{T}_{[p]};$$

3. *For $p = 1$, take all the estimates previously computed in the current iteration into account and get:*

$$\hat{\mathbf{A}}_r^{(1)\mathsf{T}} = \left(\hat{\mathbf{A}}_r^{(2)} \diamond \cdots \hat{\mathbf{A}}_r^{(P)} \right)^{\#} \mathbf{T}_{[1]};$$

4. *Update $r \leftarrow r + 1$ and repeat steps 1 to 4 until criterion (1.63) is satisfied.*

Classically, the estimation process is terminated when we can no longer observe significant variations on either the parameters or the fit of the model. According to [58], convergence of fit does not necessarily imply convergence of parameters, although this is usually the case in practical situations. As suggested therein, in order to avoid costly fit calculations after each iteration, the relative change in the value of the estimated parameters can also be indicative of convergence. That is why we adopt, throughout the rest of this thesis, except when otherwise stated, the following stop criterion:

$$\left| e(r) - e(r-1) \right| \le \varepsilon, \tag{1.63}$$

where

$$e(r) = \sum_{p=1}^{P} \frac{\left\| \hat{\mathbf{A}}_r^{(p)} - \hat{\mathbf{A}}_{r-1}^{(p)} \right\|_F^2}{\left\| \hat{\mathbf{A}}_r^{(p)} \right\|_F^2}, \tag{1.64}$$

and ε is an arbitrary small positive constant.

Lastly, we note that the full-column rank property of the multiple Khatri-Rao product $\mathbf{A}^{(p+1)} \diamond \cdots \mathbf{A}^{(P)} \diamond \mathbf{A}^{(1)} \diamond \cdots \mathbf{A}^{(p-1)}$, for all $p \in [1, P]$, is a necessary condition for the uniqueness of the Parafac decomposition [56]. It is also possible to show that satisfying the Kruskal condition implies satisfying this necessary full-column rank condition. In fact, it has been shown in the context of 3rd-order tensors that, if the Kruskal condition is satisfied, then the terms $(\mathbf{A}^{(1)} \diamond \mathbf{A}^{(2)})$, $(\mathbf{A}^{(1)} \diamond \mathbf{A}^{(3)})$ and $(\mathbf{A}^{(2)} \diamond \mathbf{A}^{(3)})$ are full-column rank, provided that the k-ranks of the Parafac components are nonzero [59]. This result is easily extendable to any order $P > 3$ [60]. Consequently, as r increases, the multiple Khatri-Rao product in (1.62) is ensured to converge to a full-column rank matrix for any $p \in [1, P]$, if the Kruskal condition is satisfied.

3.3 Particular cases

Fourth-order tensor

Let us consider a 4th-order tensor $\mathcal{T}^{(4)}$ of dimensions $I \times J \times K \times L$ with scalar representation given from (1.51) as follows:

$$t_{ijkl} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} d_{lf}, \tag{1.65}$$

where, for convenience of notation, we used a_{if}, b_{jf}, c_{kf}, and d_{lf}, with $i \in [1, I]$, $j \in [1, J]$, $k \in [1, K]$ and $l \in [1, L]$, to denote the elements of the Parafac components $\mathbf{A} \in \mathbb{C}^{I \times F}$, $\mathbf{B} \in \mathbb{C}^{J \times F}$, $\mathbf{C} \in \mathbb{C}^{K \times F}$ and $\mathbf{D} \in \mathbb{C}^{L \times F}$. Taking this notation into account and rewriting equations (1.58) to (1.60) with $P = 4$ (and $P = 3$), we easily obtain the unfolded tensor representations shown in Table 1.1, which summarizes the formulæ for the unfolded forms of the 4th- (and 3rd-) order Parafac tensor.

Table 1.1: Unfolded representations of 4th- and 3rd-order Parafac tensors

Unfolded tensor representations	$P = 4$	(dim.)	$P = 3$	(dim.)
$\mathbf{T}_{[4]}$	$(\mathbf{A} \diamond \mathbf{B} \diamond \mathbf{C})\mathbf{D}^{\mathsf{T}}$	$IJK \times L$		
$\mathbf{T}_{[3]}$	$(\mathbf{D} \diamond \mathbf{A} \diamond \mathbf{B})\mathbf{C}^{\mathsf{T}}$	$LIJ \times K$	$(\mathbf{A} \diamond \mathbf{B})\mathbf{C}^{\mathsf{T}}$	$IJ \times K$
$\mathbf{T}_{[2]}$	$(\mathbf{C} \diamond \mathbf{D} \diamond \mathbf{A})\mathbf{B}^{\mathsf{T}}$	$KLI \times J$	$(\mathbf{C} \diamond \mathbf{A})\mathbf{B}^{\mathsf{T}}$	$KI \times J$
$\mathbf{T}_{[1]}$	$(\mathbf{B} \diamond \mathbf{C} \diamond \mathbf{D})\mathbf{A}^{\mathsf{T}}$	$JKL \times I$	$(\mathbf{B} \diamond \mathbf{C})\mathbf{A}^{\mathsf{T}}$	$JK \times I$

Uniqueness, up to column scaling and permutation, is ensured under the condition stated in Theorem 1.1. Thus, we conclude from (1.56) that, if

$$k_A + k_B + k_C + k_D \geq 2F + 3, \tag{1.66}$$

then, any set $\{\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}}, \bar{\mathbf{D}}\}$ satisfying the equations in Table 1.1 is of the form

$$\bar{\mathbf{A}} = \mathbf{A}\boldsymbol{\Lambda}_1\boldsymbol{\Pi}, \quad \bar{\mathbf{B}} = \mathbf{B}\boldsymbol{\Lambda}_2\boldsymbol{\Pi}, \quad \bar{\mathbf{C}} = \mathbf{C}\boldsymbol{\Lambda}_3\boldsymbol{\Pi}, \quad \text{and} \quad \bar{\mathbf{D}} = \mathbf{D}\boldsymbol{\Lambda}_4\boldsymbol{\Pi}, \tag{1.67}$$

where $\boldsymbol{\Pi}$ is an $F \times F$ permutation matrix and $\boldsymbol{\Lambda}_p$, $p \in [1,4]$, are diagonal scaling matrices satisfying (1.55), i.e. $\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2\boldsymbol{\Lambda}_3\boldsymbol{\Lambda}_4 = \mathbf{I}_F$.

Quadrilinear Parafac-ALS (QALS) algorithm

Algorithm 1.2 (QALS algorithm)

Determine a threshold $\varepsilon > 0$, initialize $\hat{\mathbf{A}}_0$, $\hat{\mathbf{B}}_0$ and $\hat{\mathbf{C}}_0$ and compute the Parafac components as follows, starting with $r = 1$:

1. $\hat{\mathbf{D}}_r^{\mathsf{T}} = \left(\hat{\mathbf{A}}_{r-1} \diamond \hat{\mathbf{B}}_{r-1} \diamond \hat{\mathbf{C}}_{r-1}\right)^{\#} \mathbf{T}_{[4]}$;

2. $\hat{\mathbf{C}}_r^{\mathsf{T}} = \left(\hat{\mathbf{D}}_r \diamond \hat{\mathbf{A}}_{r-1} \diamond \hat{\mathbf{B}}_{r-1}\right)^{\#} \mathbf{T}_{[3]}$;

3. $\hat{\mathbf{B}}_r^{\mathsf{T}} = \left(\hat{\mathbf{C}}_r \diamond \hat{\mathbf{D}}_r \diamond \hat{\mathbf{A}}_{r-1}\right)^{\#} \mathbf{T}_{[2]}$;

4. $\hat{\mathbf{A}}_r^{\mathsf{T}} = \left(\hat{\mathbf{B}}_r \diamond \hat{\mathbf{C}}_r \diamond \hat{\mathbf{D}}_r\right)^{\#} \mathbf{T}_{[1]}$;

5. Update $r \leftarrow r + 1$ and repeat the previous steps until (1.63) is satisfied.

Using the expressions in Table 1.1, we can easily obtain from (1.61) the LS criteria to be minimized in order to estimate the four desired Parafac components. The Quadrilinear Parafac-ALS algorithm, presented above, follows from Algorithm 1.1 straightforwardly. Each Parafac component is updated with the three other ones fixed to their most up-to-date estimated values.

Third-order tensor

Originally proposed in the context of 3rd-order tensors, the Parafac decomposition of a tensor $\mathcal{T}^{(3)}$ of dimensions $I \times J \times K$ with scalar representation given as

$$t_{ijkl} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} \tag{1.68}$$

yields components $\mathbf{A} \in \mathbb{C}^{I \times F}$, $\mathbf{B} \in \mathbb{C}^{J \times F}$ and $\mathbf{C} \in \mathbb{C}^{K \times F}$, which are unique up to column scaling and permutation, if (1.56) is satisfied, i.e.

$$k_A + k_B + k_C \geq 2(F + 1). \tag{1.69}$$

In this case, the unfolded representations of $\mathcal{T}^{(3)}$, derived from equations (1.58) to (1.60) with $P = 3$, are shown in Table 1.1. Hence, if (1.69) holds, then any set $\{\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}}\}$ satisfying the equations in Table 1.1 is of the form

$$\bar{\mathbf{A}} = \mathbf{A}\boldsymbol{\Lambda}_1\boldsymbol{\Pi}, \quad \bar{\mathbf{B}} = \mathbf{B}\boldsymbol{\Lambda}_2\boldsymbol{\Pi}, \quad \text{and} \quad \bar{\mathbf{C}} = \mathbf{C}\boldsymbol{\Lambda}_3\boldsymbol{\Pi}, \tag{1.70}$$

where $\boldsymbol{\Pi}$ is an $F \times F$ permutation matrix and $\boldsymbol{\Lambda}_p$, $p \in [1,3]$, are diagonal scaling matrices satisfying $\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2\boldsymbol{\Lambda}_3 = \mathbf{I}_F$.

Trilinear Parafac-ALS (TALS) algorithm

The idea behind the ALS algorithm is now straightforward. For the particular case of a third-order tensor, the LS cost functions follow from (1.61) with $P = 3$ and the Trilinear Parafac-ALS (TALS) algorithm can be summarized as follows:

Algorithm 1.3 (TALS algorithm)

Determine a threshold $\varepsilon > 0$, initialize $\hat{\mathbf{A}}_0$ and $\hat{\mathbf{B}}_0$ and compute the Parafac components as follows, starting with $r = 1$:

1. $\hat{\mathbf{C}}_r^T = \left(\hat{\mathbf{A}}_{r-1} \diamond \hat{\mathbf{B}}_{r-1}\right)^{\#} \mathbf{T}_{[3]}$;

2. $\hat{\mathbf{B}}_r^T = \left(\hat{\mathbf{C}}_r \diamond \hat{\mathbf{A}}_{r-1}\right)^{\#} \mathbf{T}_{[2]}$;

3. $\hat{\mathbf{A}}_r^T = \left(\hat{\mathbf{B}}_r \diamond \hat{\mathbf{C}}_r\right)^{\#} \mathbf{T}_{[1]}$;

4. Update $r \leftarrow r + 1$ and repeat the previous steps until (1.63) is satisfied.

4   Summary

In the first part of this chapter, we presented a review of the main statistical tools used in this thesis for applications in telecommunication systems. We reviewed classical definitions such as moments, cumulants and polyspectra and several statistical properties have been discussed, including stationarity and symmetry relationships. Since blind signal processing methods do not generally assume any a priori knowledge about output statistics, we also addressed the issue of estimating cumulants from real data. Cumulants (of order higher than two) present special properties that moments do not. In fact, those properties enable us to handle nonminimum-phase channels using only the output (complex) signals. In addition, high-order output cumulants are blind with respect to additive Gaussian noise. Computational burden may be a drawback, due to the amounts of data required for satisfactory estimates.

Afterwards, we presented a description of some important algebraic tools used throughout the rest of this thesis. We first treated the triangularization strategies and then, more particularly, we discussed the diagonalization methods for Hermitian matrices using the classical Jacobi algorithm, based on the application of successive plane rotations. Motivated by applications in blind source separation and blind system identification using HOS, we also studied techniques for the approximative simultaneous factorization of several matrices sharing some properties but differing in individual information contents. In this context, we described an extended version of the Jacobi algorithm, which is shown to be optimal in the least squares sense.

Finally, we turned our attention to multilinear algebraic tools of great interest in statistical signal processing. In particular, we presented an overview of the Parafac tensor decomposition, which has been recently used for modeling communication systems. We described the decomposition of a generic Pth-order tensor with rank F as the sum of F rank-one tensors and discussed the uniqueness issue by introducing the Kruskal Theorem and defining the notion of k-rank. Concerning the estimation of the Parafac components, we revisited the well-known Alternating Least Square algorithm by proposing a generalized procedure for estimating the components of a tensor of any order.

Chapter 2

Blind Channel Identification using Tensor Decomposition

Symmetry properties of fourth-order cumulants yield enormous redundancies in the components of the Parallel Factor (Parafac) decomposition of the cumulant tensors. In this chapter, we develop new blind channel identification algorithms that exploit those redundancies, performing the cumulant tensor decomposition by solving a single-step (SS) least squares (LS) problem. We start with a single-input single-output (SISO) finite impulse response (FIR) channel and then we extend the principle to a multiple-input multiple-output (MIMO) instantaneous mixture. Our solution is based on the 4th-order output cumulants only and it is shown to hold for certain underdetermined mixtures, i.e. systems with more sources than sensors. In the MIMO case, a simplified approach using a reduced-order tensor is also discussed. Computer simulations are provided to assess the performance of the proposed algorithms in both SISO and MIMO cases, comparing them to other existing solutions. Initialization and convergence issues are also addressed.

As we have seen in the preceding chapter, several algorithms propose solutions to fit a Pth-order Parafac model. The well-known Alternating Least Squares (ALS) algorithm iteratively minimizes, in an alternate way, P least squares (LS) cost functions. Our main focus in this chapter is in exploiting the redundancies of the 4th-order cumulants in the Parafac decomposition of the cumulant tensors. This allows us to propose new iterative single-step least squares (SS-LS) Parafac-based Blind Channel Identification (PBCI) algorithms that have the advantage of being based on the solution of a sole optimization problem, contrary to the methods described in the literature. For that reason, SS-LS PBCI consists in a new scheme for the estimation of FIR systems [61, 62]. In addition, using the same underlying idea, under mild assumptions, we also propose algorithms to treat the case of instantaneous MIMO channels[63, 64]. These techniques are, to our knowledge, the first to consider the cumulant redundancies in order to improve the LS solution of the Parafac decomposition of the cumulant tensor.

In the sequel, we will be interested first in recovering the impulse response of a complex

FIR-SISO channel from the Parafac decomposition of a 3rd-order tensor composed of 4th-order output cumulants. Using the SS-LS approach, the permutation and scaling ambiguities intrinsic to the Parafac decomposition are solved and the uniqueness issue is addressed [61, 62, 64]. After that, we consider the problem of blind MIMO channel (mixture) identification in the context of a multiuser system characterized by instantaneous complex channels. A quadrilinear ALS solution is described based on the decomposition of a 4th-order tensor composed of 4th-order spatial cumulants. A trilinear approach is also discussed using a third-order tensor of 4th-order cumulants. Then, we finally present a simplified SS-LS Parafac-based Blind MIMO Channel Identification (PBMCI) algorithm. Although our main goal is not focused on underdetermined mixtures, we make use of some tensor properties to show that under certain conditions our algorithm is able to identify channels with more sources than sensors. Computer simulations illustrate the performance gains that our method provides with respect to other existing solutions. We also assess the algorithms performances by recovering the input signals using a minimum mean squared error (MMSE) equalizer built from the estimated channel. In the MIMO case, a *semi-blind* MMSE equalizer will be implemented, using a few pilot symbols.

This chapter is organized as follows: in section 2, we introduce the signal model and define the output cumulants as a tensor; in section 3, we describe a joint-diagonalization based algorithm and propose a Parafac-based algorithm to estimate the SISO channel parameters based on a SS-LS minimization procedure; we also discuss come connections between the (simultaneous) matrix diagonalization approach and our cumulant tensor decomposition; section 4 introduces a multiuser and multiantenna channel model and, in section 5, we propose blind channel estimation algorithms coping with the MIMO case; methods using joint-diagonalization techniques are also described; section 6 presents some computer simulation results to illustrate the proposed methods; conclusions are drawn in section 7, along with some perspectives.

1   Brief history of the HOS-based blind identification methods

Blind identification methods aim to determine an unknown model from the system output only. Known applications range from data communications, beamforming and echo cancelation to image restoration, speech recognition and seismic applications, among others. In which concerns statistical techniques, second-order statistics (SOS) based methods usually require models with multiple outputs, such as oversampled or multiantenna systems.

On the other hand, it is well-known that cumulants of order higher than two can be viewed as tensors with a highly symmetrical structure [17]. Among the earliest works exploiting the cumulant symmetries with a tensor formalism, Cardoso introduced the concept of eigenvalue structure of 4th-order cumulant tensors [23, 24]. He used the uniqueness property of the cumulant tensors as an advantage over singular value decomposition (SVD), but prewhitening was

needed. Later on, an extended Jacobi technique for approximate simultaneous diagonalization was proposed by Cardoso and Soloumiac in [51]. This latter paper introduced the JADE algorithm that uses second and 4th-order statistics to estimate an instantaneous multiple-input multiple-output (MIMO) channel in the context of blind beamforming. The joint diagonalization technique became very popular and has been used by Belouchrani et al. to propose the second-order blind identification (SOBI) algorithm [65], which uses a set of correlation matrices to identify stationary sources with different spectral contents, also in the context of memoryless MIMO channels. On the other hand, the fourth-order system identification (FOSI) algorithm [66] treats single-input single output (SISO) FIR channels and also involves an a priori transformation over the cumulant matrices, which is often a source of increased complexity and estimation errors.

During the last decade, several joint-diagonalization criteria have been introduced [13, 14]. Important modifications of the technique proposed in [24] were provided by De Lathauwer et al. in [67], resorting to joint diagonalization techniques. A link between the Parafac decomposition and the simultaneous matrix diagonalization approach has been discussed in [21], leading to algorithms subject to weaker uniqueness conditions. These ideas gave rise to the FOOBI family of algorithms [22], which are theoretically able to identify a greater number of user channels for a given number of receive antennas. The FOOBI algorithms exploit the Quadricovariance matrix making use of its column-wise Kronecker structure. Also using the joint diagonalization approach, the ICAR algorithm proposed in [68] is based on the redundancies of the Quadricovariance structure to estimate the mixture matrix, but only in the overdetermined case, i.e. the case of systems with more sensors than sources. The principle behind the ICAR method [68] has also been used in [69] and [70], resorting to 6th- and higher-order statistics, respectively, in order to include the case of underdetermined mixtures. The ICAR and the FOOBI algorithms will be further discussed in section 5.1. These techniques, while avoiding prewhitening, still break the problem into two optimization stages, which remain necessary to extract the MIMO channel coefficients from an initial estimate based on an eigenvalue decomposition (EVD).

Using the Parafac decomposition and exploiting the symmetry properties of the 4th-order cumulants, we eliminate the need for prewhitening in the SISO case and, in both SISO and MIMO cases, we found solutions to the blind identification problem by searching the minima of a single LS cost function. In addition, the Parafac-based approach allows us to treat the underdetermined case, although the uniqueness condition imposes an upper bound on the number of identifiable sources. Actually, the blind identification of underdetermined mixtures has received a special attention from the signal processing community under different tensorial approaches that include, among others, the decomposition of *quantics* in sums of powers of linear forms [42]; the use of congruent transformation [70] exploiting the virtual array concept [25, 26]; and the use of high-order derivatives of the multivariate characteristic function [71]. Besides, a frequency domain framework for MIMO system identification using Parafac was introduced in [72] using

HOS-based tensors. More recently, that approach was further developed and also included the underdetermined case [73, 74, 60].

2 Single-user signal model and 4th-order output cumulants

Let us represent a FIR-SISO communication channel in which the output signal $y(n)$, after sampling at the symbol rate, is written as follows:

$$
\begin{aligned}
y(n) &= x(n) + \upsilon(n), \\
x(n) &= \sum_{\ell=0}^{L} h(\ell)s(n - \ell),
\end{aligned} \tag{2.1}
$$

with $h(0) = 1$ (which is equivalent to a simple unit-norm constraint). Moreover, the following assumptions hold:

A1 : The non-measurable, complex-valued, discrete input sequence $s(n)$ is stationary, ergodic, independent and identically distributed (iid) with symmetric distribution, zero-mean and non-zero kurtosis $\gamma_{4,s}$.

A2 : The additive Gaussian noise sequence $\upsilon(n)$, with zero-mean, unknown variance σ_υ^2 and unknown autocorrelation function, is assumed to be independent of the input signal $s(n)$.

A3 : The complex coefficients $h(\ell)$ represent the equivalent discrete channel impulse response, including the pulse shaping filter, the transmission channel and the receive filter.

A4 : The FIR filter representing the channel is assumed to be causal with memory $L \neq 0$ and no temporal sparsity, i.e. $h(\ell) \neq 0, \, \forall\, \ell \in [0, L]$, and $h(\ell) = 0$ otherwise.

From definition (1.25), the 4th-order cumulants of the output signal $y(n)$ are given as follows:

$$
c_{4,y}(\tau_1, \tau_2, \tau_3) \triangleq cum\left[y^*(n), y(n + \tau_1), y^*(n + \tau_2), y(n + \tau_3)\right]. \tag{2.2}
$$

Using the channel model (2.1), taking assumptions A1 and A2 into account and making use of the multilinearity property of cumulants, it can be shown that (c.f. Appendix B):

$$
c_{4,y}(\tau_1, \tau_2, \tau_3) = \gamma_{4,s} \sum_{\ell=0}^{L} h^*(\ell)h(\ell + \tau_1)h^*(\ell + \tau_2)h(\ell + \tau_3), \tag{2.3}
$$

where $\gamma_{4,s} = c_{4,s}(0, 0, 0)$, according to definition (1.20). Based on (2.3) and on assumption A4, we note that:

$$
c_{4,y}(\tau_1, \tau_2, \tau_3) = 0, \, \forall\, |\tau_1|, |\tau_2|, |\tau_3| > L. \tag{2.4}
$$

Hence, by taking the time-lags τ_1, τ_2 and τ_3 in the interval $[-L, L]$, we consider all possible nonzero values of $c_{4,y}(\tau_1, \tau_2, \tau_3)$. This choice yields a maximum redundancy information model.

Let us define the 3rd-order tensor $\mathcal{C}^{(3,y)} \in \mathbb{C}^{(2L+1) \times (2L+1) \times (2L+1)}$, with scalar representation given by $c_{ijk} = c_{4,y}(i - L - 1, j - L - 1, k - L - 1)$, $i, j, k \in [1, 2L + 1]$. Using (2.3), in analogy with (1.51), it is not difficult to show that tensor $\mathcal{C}^{(3,y)}$, depicted in fig. 2.1a, can be written as a sum of $L + 1$ outer products, each one involving 3 vectors, so that

$$\mathcal{C}^{(3,y)} = \gamma_{4,s} \sum_{\ell=0}^{L} \mathbf{H}_{\cdot \ell+1} \circ \mathbf{H}_{\cdot \ell+1}^* \circ \left(h^*(\ell) \mathbf{H}_{\cdot \ell+1} \right), \tag{2.5}$$

with

$$\mathbf{H}_{\cdot \ell+1} = \sum_{p=1}^{2L+1} h(\ell + p - L - 1) \mathbf{e}_p^{(2L+1)}, \tag{2.6}$$

where $\mathbf{e}_i^{(I)}$ denotes the ith canonical basis vector of $\mathbb{R}^I$. The above equations can be easily obtained from (1.52) and (1.53) with $P = 3$, taking the scalar representation of $\mathcal{C}^{(3,y)}$ into account. Clearly, equation (2.5) represents the Parafac decomposition of the tensor $\mathcal{C}^{(3,y)}$, with components $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ given as follows:

$$\mathbf{A} = \mathbf{H}, \quad \mathbf{B} = \mathbf{H}^* \quad \text{and} \quad \mathbf{C} = \gamma_{4,s} \mathbf{H} Diag\left(\mathbf{h}^* \right), \tag{2.7}$$

where $Diag\left(\cdot \right)$ denotes a diagonal matrix built from the entries of the vector argument and the channel coefficient vector $\mathbf{h}$ is defined as:

$$\mathbf{h} = \left[h(0) \ \ldots \ h(L) \right]^\mathsf{T} \in \mathbb{C}^{(L+1)}. \tag{2.8}$$

The channel coefficient matrix $\mathbf{H} \in \mathbb{C}^{(2L+1) \times (L+1)}$ can be explicitly expressed as follows:

$$\mathbf{H} \triangleq \mathcal{H}(\mathbf{h}) = \left[\mathbf{H}_{\cdot 1} \ \mathbf{H}_{\cdot 2} \ldots \mathbf{H}_{\cdot L+1} \right] = \begin{pmatrix} 0 & 0 & \cdots & h(0) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & h(0) & \cdots & h(L-1) \\ h(0) & h(1) & \cdots & h(L) \\ \vdots & \vdots & \ddots & \vdots \\ h(L-1) & h(L) & \cdots & 0 \\ h(L) & 0 & \cdots & 0 \end{pmatrix} \tag{2.9}$$

where $\mathcal{H}(\,\cdot\,)$ is an operator that builds a Hankel matrix from its vector argument, as shown above.

3 Blind SISO channel identification algorithms

Due to the symmetric structure of the cumulant tensor $\mathcal{C}^{(3,y)}$, the Parafac decomposition given in (2.5) can be obtained by means of a simultaneous diagonalization of the cumulant tensor (2D) slices, subject to a prior orthonormalization. The FOSI algorithm proposed in [66] adopts this latter approach without using a tensor formulation. In this section, we present the FOSI algorithm, highlighting that connection with the simultaneous diagonalization approach. After that, we also present the SS-LS blind channel identification algorithm and discuss the uniqueness issue.
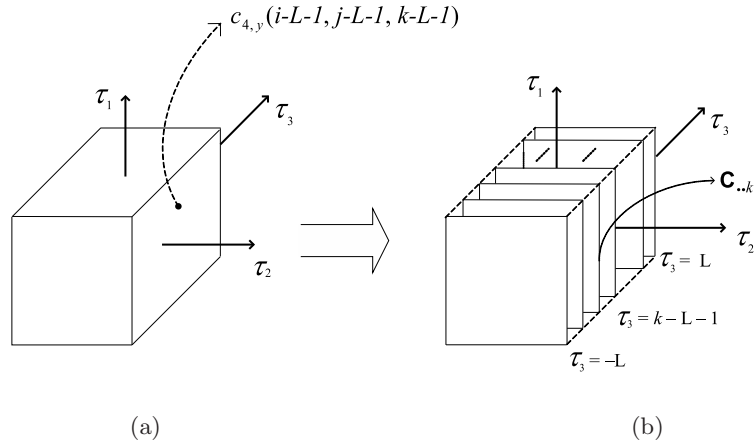
Figure 2.1: (a) 3rd-order tensor $\mathcal{C}^{(3,y)}$ of 4th-order output cumulants; (b) frontal slices of tensor $\mathcal{C}^{(3,y)}$.

3.1 A Joint-diagonalization based approach

Joint diagonalization (JD) has been a reference tool in signal processing, finding applications in several fields including blind source separation, blind identification of quadratic models [75] and source localization [76]. The concept introduced in [51], in the context of blind beamforming, gave rise to the *Joint Approximate Diagonalization of Eigenmatrices* (JADE) algorithm, which exploits the assumption of statistical independence of the sources and utilizes extended Jacobi techniques (c.f. Appendix A) over a set of 4th-order cumulant matrices [52]. This basic principle has been shown to be very useful in applications using second and higher-order statistics [65, 66].

Let us define the frontal slices $\mathbf{C}_{..k} \in \mathbb{C}^{(2L+1)\times(2L+1)}$, $k \in [1, 2L + 1]$, of the cumulant tensor $\mathcal{C}^{(3,y)}$ as illustrated in fig. 2.1(b). Taking the scalar representation c_{ijk} into account, we have:

$$
\begin{aligned}
\mathbf{C}_{..k} &= \sum_{i=1}^{2L+1} \sum_{j=1}^{2L+1} c_{ijk} \mathbf{e}_i^{(2L+1)} \mathbf{e}_j^{(2L+1)\mathsf{T}} \\
&= \gamma_{4,s} \sum_{\ell=0}^{L} h^*(\ell) h(\ell + k - L - 1) \mathbf{H}_{.\ell} \mathbf{H}_{.\ell}^{\mathsf{H}} \\
&= \gamma_{4,s} \mathbf{H} D_k(\mathbf{\Sigma}) \mathbf{H}^{\mathsf{H}}, \qquad k \in [1, 2L + 1],
\end{aligned}
\tag{2.10}
$$

where $D_i(\cdot)$ denotes a diagonal matrix built from the ith row of the matrix argument and we have used the following definition:

$$
\mathbf{\Sigma} = \mathbf{H} Diag(\mathbf{h}^*) \quad \in \mathbb{C}^{(2L+1)\times(L+1)}.
\tag{2.11}
$$

Noticing from (2.11) that $D_k(\mathbf{\Sigma}) = D_k(\mathbf{H}) D_{L+1}(\mathbf{H}^*)$, let us define the set of *modified* cumulant matrices $\bar{\mathbf{C}}_k \triangleq \mathbf{W} \mathbf{C}_{..k} \mathbf{W}^\mathsf{H} \in \mathbb{C}^{(L+1)\times(L+1)}$, so that from (2.10) we have:

$$\bar{\mathbf{C}}_k = \gamma_{4,s} \, \mathbf{W} \mathbf{H} \left(D_k(\mathbf{H}) D_{L+1}(\mathbf{H}^*) \right) \mathbf{H}^\mathsf{H} \mathbf{W}^\mathsf{H}. \tag{2.12}$$

Left- and right-multiplying the term $D_k(\mathbf{H})$ by $D_{L+1}(\mathbf{H})$ and $D_{L+1}(\mathbf{H})^{-1}$, respectively, we end up with

$$\bar{\mathbf{C}}_k = \mathbf{Q} \, \bar{\mathbf{\Delta}}_k \, \mathbf{Q}^\mathsf{H}, \tag{2.13}$$

where $\bar{\mathbf{\Delta}}_k = \gamma_{4,s} D_k(\mathbf{H}) D_{L+1}(\mathbf{H})^{-1}$ and the transformation $\mathbf{W} \in \mathbb{C}^{(L+1)\times(2L+1)}$ is such that

$$\mathbf{Q} = \mathbf{W} \, \mathbf{H} \, D_{L+1}(\mathbf{H}) \tag{2.14}$$

is a $(L+1) \times (L+1)$ unitary matrix. Matrix $\mathbf{W}$ can be computed from the above equation by noting that $\mathbf{Q} \mathbf{Q}^\mathsf{H} = \mathbf{I}$, which leads to:

$$\mathbf{W} \left(\gamma_{4,s} \mathbf{H} \, |D_{L+1}(\mathbf{H})|^2 \, \mathbf{H}^\mathsf{H} \right) \mathbf{W}^\mathsf{H} = \mathbf{I}, \tag{2.15}$$

where $|\cdot|$ denotes the element-wise absolute value. Using $k = L+1$ in (2.10), we note that $\mathbf{C}_{..L+1} = \gamma_{4,s} \mathbf{H} \, |D_{L+1}(\mathbf{H})|^2 \, \mathbf{H}^\mathsf{H}$ and hence $\mathbf{W}$ can be viewed as a whitening transformation that decorrelates the elements of the cumulant matrix $\mathbf{C}_{..L+1}$. A solution to this problem is given by the inverse square-root[1] of $\mathbf{C}_{..L+1}$, as follows:

$$\mathbf{W} = \mathbf{C}_{..L+1}^{-1/2} = \mathbf{\Lambda}^{-1/2} \, \mathbf{U}^\mathsf{H}, \tag{2.16}$$

where the matrices $\mathbf{U} \in \mathbb{C}^{(2L+1)\times(L+1)}$ and $\mathbf{\Lambda} \in \mathbb{C}^{(L+1)\times(L+1)}$ represent the *economy-size* EVD of $\mathbf{C}_{..L+1}$. Although the source Kurtosis is assumed unknown, it must be ensured to be positive ($\gamma_{4,s} > 0$) due to the above square-root calculation. If this is not the case, then matrix $-\mathbf{C}_{..L+1}$ should be used instead.

The computation of matrix $\mathbf{W}$ is a very common operation in HOS-based methods [65, 77, 78], often referred to as *prewhitening*. However, it usually requires resorting to SOS and, even if this is not the case, this additional stage is time-consuming and often responsible for increased estimation errors [79, 80].

Noticing from (2.13) that $\mathbf{Q}$ is the matrix that simultaneously diagonalizes the set $\bar{\mathbf{C}}_k$, $k = 1, \ldots, 2L+1$, it can be computed by maximizing the following cost function:

$$\mathsf{J}(\mathbf{Q}, \bar{\mathsf{c}}) = \sum_{\substack{k=1 \\ k \neq L+1}}^{2L+1} \left\| diag\left(\mathbf{Q}^\mathsf{H} \bar{\mathbf{C}}_k \mathbf{Q} \right) \right\|^2, \tag{2.17}$$

---

[1] The square-root of a given Hermitian matrix $\mathbf{X}$ is so that $(\mathbf{X}^{1/2})(\mathbf{X}^{1/2})^\mathsf{H} = \mathbf{X}$. This operation presents an orthonormal ambiguity since any matrix $\bar{\mathbf{X}}^{1/2} = \mathbf{X}^{1/2}\mathbf{U}$ is also a square-root of $\mathbf{X}$ if $\mathbf{U}$ is unitary.

where $\bar{\mathsf{C}} = \{\bar{\mathbf{C}}_1, \ldots, \bar{\mathbf{C}}_L, \bar{\mathbf{C}}_{L+2}, \ldots, \bar{\mathbf{C}}_{2L+1}\}$. This solution is shown to be optimal in the LS sense [54] and can be obtained using the joint-diagonalization algorithm (A.2) described in Appendix A based on the extended Jacobi technique. Finally, the channel coefficient matrix is obtained from (2.14) and (2.16), as follows:

$$\bar{\mathbf{H}} = \mathbf{W}^{\#}\mathbf{Q} = \mathbf{U}\,\boldsymbol{\Lambda}^{1/2}\,\mathbf{Q}, \tag{2.18}$$

up to trivial indeterminacies (column scaling and permutations).

In order to recover the channel coefficient vector from the estimated channel matrix $\bar{\mathbf{H}}$, we need to get rid of the permutation ambiguity. Considering the structure of $\mathbf{H}$, given in (2.9), we note from (2.10) and (2.11) that $\mathbf{C}_{..1} = \mathbf{H}\,\boldsymbol{\Delta}_1\,\mathbf{H}^{\mathsf{H}}$, where $\boldsymbol{\Delta}_1 = \delta_1 Diag\,([0, \ldots 0, 1])$ is an $(L+1) \times (L+1)$ diagonal matrix, with $\delta_1 = \gamma_{4,s}\,h(0)\,h^*(L)$. Thus, $\mathbf{C}_{..1}$ is a rank-1 matrix and can be written as $\mathbf{C}_{..1} = \delta_1\,\mathbf{H}_{\cdot L+1}\,\mathbf{H}^{\mathsf{H}}_{\cdot L+1}$. Using the column-permuted and scaled channel matrix $\bar{\mathbf{H}}$, computed in (2.18), we denote by $\ell_{(1)}$ the column of $\bar{\mathbf{H}}$ that is a weighted version of the $(L+1)$th column of $\mathbf{H}$, so that

$$\bar{\mathbf{H}}_{\cdot\ell_{(1)}} = \xi_{\ell_{(1)}}\mathbf{H}_{\cdot L+1}, \tag{2.19}$$

where $\xi_{\ell_{(1)}}$ is a nonzero complex scalar factor. This allows us to write $\mathbf{C}_{..1} = \bar{\delta}_1\,\bar{\mathbf{H}}_{\cdot\ell_{(1)}}\,\bar{\mathbf{H}}^{\mathsf{H}}_{\cdot\ell_{(1)}}$, with $\bar{\delta}_1 = \gamma_{4,s}\,h(0)\,h^*(L)|\xi_{\ell_{(1)}}|^{-2}$, and we can conclude that $\ell_{(1)}$ determines the position of the only nonzero diagonal entry of matrix $\bar{\boldsymbol{\Delta}}_1$, which can be obtained from the set of equations (2.13). Analogously, using (2.10) and (2.11) with $k = 2L + 1$, we can conclude that $\mathbf{C}_{..2L+1}$ is also a rank-1 matrix and can be written from the 1st column of $\mathbf{H}$. Denoting by $\ell_{(2L+1)}$ the column of $\bar{\mathbf{H}}$ that is a weighted version of the 1st column of $\mathbf{H}$, we get

$$\bar{\mathbf{H}}_{\cdot\ell_{(2L+1)}} = \xi_{\ell_{(2L+1)}}\mathbf{H}_{\cdot 1}, \tag{2.20}$$

so that $\mathbf{C}_{..2L+1} = \bar{\delta}_{2L+1}\,\bar{\mathbf{H}}_{\cdot\ell_{(2L+1)}}\,\bar{\mathbf{H}}^{\mathsf{H}}_{\cdot\ell_{(2L+1)}}$, with $\bar{\delta}_{2L+1} = \gamma_{4,s}\,h(L)\,h^*(0)|\xi_{\ell_{(2L+1)}}|^{-2}$, where $\xi_{\ell_{(2L+1)}}$ is a nonzero complex scalar factor. As a result, $\ell_{(2L+1)}$ determines the position of the sole nonzero diagonal element of $\bar{\boldsymbol{\Delta}}_{2L+1}$, which can be obtained from (2.13).

Considering the Hankel structure of $\mathbf{H}$ given in (2.9), equations (2.19) and (2.20) suggest that the channel coefficient vector $\mathbf{h}$ can be recovered from the $L + 1$ top elements of $\bar{\mathbf{H}}_{\cdot\ell_{(1)}}$ and the $L + 1$ bottom elements of $\bar{\mathbf{H}}_{\cdot\ell_{(2L+1)}}$. These two solutions can be computed by means of the FOSI algorithm, summarized below in Algorithm 2.1. A third channel estimate can be obtained by averaging these solutions. Finally, we note that in [66] the 4th-order cumulants are computed with exclusively non-negative time-lags, i.e. $0 \leq \tau_1, \tau_2, \tau_3 \leq L$ or, equivalently, $L + 1 \leq i, j, k \leq 2L + 1$. Therefore, matrix dimensions are different from those shown here.

Algorithm 2.1 (Fourth-Order System Identification)

1. Compute the EVD of $\hat{\mathbf{C}}_{..L+1}$ and denote by $\mathbf{U}$ the $(2L+1) \times 1$ eigenvectors associated with the $L+1$ largest eigenvalues, arranged in the diagonal matrix $\mathbf{\Lambda}$; Deduce the sign of $\gamma_{4,s}$ from the diagonal elements of $\mathbf{\Lambda}$; If the $sign(\gamma_{4,s}) < 0$, then use $-\hat{\mathbf{C}}_{..L+1}$ instead.

2. Estimate the orthonormalizing transformation $\hat{\mathbf{W}} = \mathbf{\Lambda}^{-1/2}\mathbf{U}^H$ and its pseudo-inverse $\hat{\mathbf{W}}^{\#} = \mathbf{U}\mathbf{\Lambda}^{1/2}$.

3. Compute the set of modified cumulant matrices $\bar{\mathbf{C}}_k \triangleq \mathbf{W}\hat{\mathbf{C}}_{..k}\mathbf{W}^H$, $k = 1, \ldots, L, L+2, \ldots, 2L+1$;

4. Using Algorithm A.2 compute a unitary matrix $\mathbf{Q}$ that simultaneously diagonalizes $\bar{\mathbf{C}} = \{\bar{\mathbf{C}}_1, \ldots, \bar{\mathbf{C}}_L, \bar{\mathbf{C}}_{L+2}, \ldots, \bar{\mathbf{C}}_{2L+1}\}$. Denote by $\hat{\mathbf{\Delta}}_k$ the corresponding diagonal matrices.

5. Take the diagonal elements of $\hat{\mathbf{\Delta}}_1$ and $\hat{\mathbf{\Delta}}_{2L+1}$ and denote by $\ell_{(1)}$ and $\ell_{(2L+1)}$ the column number of their largest absolute values, respectively;

6. From (2.18), compute the channel matrix estimate as $\hat{\mathbf{H}} = \mathbf{W}^{\#}\mathbf{Q}$.

7. Determine two different channel estimates by taking the $L+1$ top elements of $\hat{\mathbf{H}}_{.\ell_{(1)}}$ and the $L+1$ bottom elements of $\hat{\mathbf{H}}_{.\ell_{(2L+1)}}$. Normalize the two resulting vectors with respect to their corresponding first entries.

3.2 The Single-Step Least-Squares approach

Using the Parafac components expressed in (2.7), the unfolded tensor representations of $\mathcal{C}^{(3,y)}$, with dimensions $(2L+1)^2 \times (2L+1)$, can be written from the general formulæ (1.58), (1.59) and (1.60), as follows:

$$\mathbf{C}_{[1]} = \gamma_{4,s}\Big(\mathbf{H} \diamond \mathbf{H}^*\Big)\mathbf{\Sigma}^{\mathsf{T}} \tag{2.21}$$

$$\mathbf{C}_{[2]} = \gamma_{4,s}\Big(\mathbf{H}^* \diamond \mathbf{\Sigma}\Big)\mathbf{H}^{\mathsf{T}} \tag{2.22}$$

$$\mathbf{C}_{[3]} = \gamma_{4,s}\Big(\mathbf{\Sigma} \diamond \mathbf{H}\Big)\mathbf{H}^{\mathsf{H}}, \tag{2.23}$$

where $\mathbf{\Sigma}$ is defined in (2.11).

Equation (2.5) shows that the rank of tensor $\mathcal{C}^{(3,y)}$ equals $L+1$. Assumption A4 ensures that $L \geq 1$. Due to its Hankel structure, $\mathbf{H}$ is full-rank and then $k_{\mathbf{A}} = k_{\mathbf{B}} = r_{\mathbf{H}} = L+1$. Taking assumption A4 into account, we deduce from (2.11) that $k_{\mathbf{C}} = r_{\mathbf{\Sigma}} = L+1$. From the Kruskal uniqueness condition (1.69), we conclude that $k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} = 3L+3 \geq 2F+2 = 2L+4$, which

Table 2.1: Parafac formulæ for the 3rd-order tensor $\mathcal{C}^{(3,y)}$

Slicing direction	2D slices	Unfolded representations
Horizontal	$\mathbf{C}_{i\cdot\cdot} = \gamma_{4,s}\mathbf{H}^* D_i\left(\mathbf{H}\right)\mathbf{\Sigma}^\mathsf{T}$	$\mathbf{C}_{[1]} = \gamma_{4,s}\left(\mathbf{H}\diamond\mathbf{H}^*\right)\mathbf{\Sigma}^\mathsf{T}$
Vertical	$\mathbf{C}_{\cdot j\cdot} = \gamma_{4,s}\mathbf{\Sigma} D_j\left(\mathbf{H}^*\right)\mathbf{H}^\mathsf{T}$	$\mathbf{C}_{[2]} = \gamma_{4,s}\left(\mathbf{H}^*\diamond\mathbf{\Sigma}\right)\mathbf{H}^\mathsf{T}$
Frontal	$\mathbf{C}_{\cdot\cdot k} = \gamma_{4,s}\mathbf{H} D_k\left(\mathbf{\Sigma}\right)\mathbf{H}^\mathsf{H}$	$\mathbf{C}_{[3]} = \gamma_{4,s}\left(\mathbf{\Sigma}\diamond\mathbf{H}\right)\mathbf{H}^\mathsf{H}$

is always true. Thus, any set $\{\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}}\}$ satisfying the Parafac decomposition of the cumulant tensor $\mathcal{C}^{(3,y)}$ has the form (1.70), with components $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ given in (2.7).

Table 2.1 summarizes the Parafac decomposition of tensor $\mathcal{C}^{(3,y)}$, including the tensor unfolded representations and the 2D slices obtained from the three possible slicing directions. Considering the unfolded matrices in the right column, traditional ALS algorithms can be used to estimate the three Parafac components of $\mathcal{C}^{(3,y)}$, leading to the matrices $\mathbf{H}$ and $\mathbf{\Sigma}$ and then to the channel parameters. However, we can improve the efficiency of the estimation procedure by coupling both estimation steps, i.e. taking the relationships between the channel coefficient vector $\mathbf{h}$ and the matrices $\mathbf{H}$ and $\mathbf{\Sigma}$ into account, thus eliminating column scaling and permutation ambiguities [61, 62, 64].

A new SS-LS PBCI algorithm

Next, we present a very useful property of the Khatri-Rao product and then we propose a single-step least squares algorithm to estimate the channel coefficient vector $\mathbf{h}$ by means of the previously described tensor decomposition.

Property 1

Let $\mathbf{Z} \in \mathbb{C}^{m\times n}$ be written as $\mathbf{Z} = \mathbf{X}Diag\left(\mathbf{v}\right)\mathbf{Y}$, where $\mathbf{X} \in \mathbb{C}^{m\times q}$, $\mathbf{Y} \in \mathbb{C}^{q\times n}$ and $\mathbf{v} \in \mathbb{C}^{q\times 1}$. Then it holds:

$$vec(\mathbf{Z}) = (\mathbf{Y}^\mathsf{T}\diamond\mathbf{X})\mathbf{v} \in \mathbb{C}^{mn\times 1}. \tag{2.24}$$

∎

In the above equation, $vec\left(\cdot\right)$ denotes the *vectorization* operator. Replacing (2.11) in (2.21), matrix $\mathbf{C}_{[1]}$ can be written as follows:

$$\mathbf{C}_{[1]} = \gamma_{4,s}(\mathbf{H}\diamond\mathbf{H}^*)Diag\left(\mathbf{h}^*\right)\mathbf{H}^\mathsf{T}.$$

Applying property (2.24) to the above equation, we get:

$$vec(\mathbf{C}_{[1]}) = \gamma_{4,s}(\mathbf{H}\diamond\mathbf{H}\diamond\mathbf{H}^*)\mathbf{h}^*. \tag{2.25}$$

Thus, the channel coefficient vector $\mathbf{h}$ can be obtained as the argument that minimizes the following LS cost function

$$\psi(\mathbf{h}^*, \hat{\mathbf{h}}^{(r-1)}) \triangleq \left\| vec(\mathbf{C}_{[1]}) - \gamma_{4,s}\hat{\mathbf{G}}^{(r-1)}\mathbf{h}^* \right\|_F^2 \qquad (2.26)$$

by means of an iterative procedure, where $\hat{\mathbf{G}}$ is given as:

$$\hat{\mathbf{G}}^{(r-1)} = \hat{\mathbf{H}}^{(r-1)} \diamond \hat{\mathbf{H}}^{(r-1)} \diamond \hat{\mathbf{H}}^{(r-1)*}. \qquad (2.27)$$

Considering the structure of the channel matrix $\mathbf{H}$ given in (2.9), we denote $\hat{\mathbf{H}}^{(r-1)} = \mathcal{H}(\hat{\mathbf{h}}^{(r-1)})$. At each iteration $r \geq 1$, we have:

$$\hat{\mathbf{h}}^{(r)} = \arg\min\psi(\mathbf{h}^*, \hat{\mathbf{h}}^{(r-1)}), \qquad (2.28)$$

from which we get:

$$\hat{\mathbf{h}}^{(r)*} = \gamma_{4,s}^{-1}\hat{\mathbf{G}}^{(r-1)\#}vec\left(\mathbf{C}_{[1]}\right). \qquad (2.29)$$

The algorithm is initialized with a Hankel matrix $\hat{\mathbf{H}}^{(0)}$ of which the first column is $[\mathbf{0}_{(L)}^\mathsf{T}\ \hat{\mathbf{h}}^{(0)\mathsf{T}}]^\mathsf{T}$ and the last row is $[\hat{h}^{(0)}(L)\ \mathbf{0}_{(L)}^\mathsf{T}]$, where $\hat{\mathbf{h}}^{(0)} = [1\ \mathbf{v}^\mathsf{T}]^\mathsf{T}$, $\mathbf{v} \in \mathbb{C}^{(L)}$ is a Gaussian random vector and $\mathbf{0}_{(L)}$ is an all-zero vector of dimension L. The algorithm is stopped when significant variations of the estimated parameters are no longer observed, i.e. when $|e(r) - e(r-1)| \leq \varepsilon$, where $e(r) = \|\hat{\mathbf{h}}^{(r)} - \hat{\mathbf{h}}^{(r-1)}\|^2/\|\hat{\mathbf{h}}^{(r)}\|^2$ and ε is an arbitrary small positive constant. Taking the model constraint $h(0) = 1$ into account, we normalize, at each iteration r, the preceding estimate $\hat{\mathbf{h}}^{(r-1)}$ with respect to its first entry $\hat{h}^{(r-1)}(0)$, before using it to update $\hat{\mathbf{H}}^{(r-1)}$ from (2.9). Then, $\hat{\mathbf{G}}^{(r-1)}$ is computed from (2.27) using $\hat{\mathbf{H}}^{(r-1)}$. The normalization step eliminates the scaling ambiguity and renders the solution independent from the source kurtosis. Forcing the Hankel structure of $\mathbf{H}$ with the operator $\mathcal{H}(\cdot)$ allows us to avoid column permutation in the estimated Parafac components. The Single-Step LS Parafac-based Blind Channel Identification (SS-PBCI) algorithm can be summarized as follows:

Algorithm 2.2 (Single-Step LS PBCI algorithm)

Determine a threshold $\varepsilon > 0$ and initialize $\hat{\mathbf{h}}^{(0)}$ as described above.

For $r \geq 1$, execute the steps below:

1. *Use (2.9) to build* $\hat{\mathbf{H}}^{(r-1)} = \mathcal{H}\left(\frac{1}{\hat{h}^{(r-1)}(0)}\hat{\mathbf{h}}^{(r-1)}\right)$;

2. *Using (2.27), compute* $\hat{\mathbf{G}}^{(r-1)}$;

3. *Compute the channel vector estimate as* $\hat{\mathbf{h}}^{(r)*} = \hat{\mathbf{G}}^{(r-1)\#}vec\left(\mathbf{C}_{[1]}\right)$;

4. *Reiterate until* $|e(r) - e(r-1)|^2 \leq \varepsilon$, $e(r) = \|\hat{\mathbf{h}}^{(r)} - \hat{\mathbf{h}}^{(r-1)}\|/\|\hat{\mathbf{h}}^{(r)}\|$.

According to the above discussion, the identifiability of the channel coefficient vector $\hat{\mathbf{h}}^{(r)}$ depends on the uniqueness of the LS solution (2.29), which in turn depends on the full-rank property of the double Khatri-Rao product defined in (2.27). It is possible to show [60, 59] that satisfying the Kruskal condition implies the full-column rank property, which is actually a necessary condition for the uniqueness of the Parafac decomposition [56]. As a result, matrix $\hat{\mathbf{G}}^{(r-1)}$ can be said to be full-column rank, which ensures the uniqueness of the proposed solution.

The SS-LS strategy ensures the Hankel structure of $\mathbf{H}$ at each iteration, taking advantage of its full-rank property to make the tensor decomposition essentially unique and the channel parameters estimation free from ambiguities. Furthermore, one sole LS minimization is needed, contrary to the classical trilinear ALS algorithm. For that reason, our method should also be expected to increase convergence speed.

4   Multiuser channel model and 4th-order spatial cumulants

Let us consider an instantaneous MIMO channel with Q signal sources and M receive antennas. The signals received at the front-end of the antenna array at the time-instant n are modeled as a complex vector $\mathbf{y}(n) \in \mathbb{C}^M$, which is written as:

$$\mathbf{y}(n) = \mathbf{H}\mathbf{s}(n) + \boldsymbol{v}(n), \tag{2.30}$$

where the elements of the complex instantaneous mixing matrix $\mathbf{H} \in \mathbb{C}^{M \times Q}$ are the MIMO channel coefficients h_{mq}, i.e. $[\mathbf{H}]_{mq} = h_{mq}$. The following assumptions hold:

B1 : The source signals $s_q(n)$ are stationary, ergodic and mutually independent with symmetric distribution, zero-mean and non-zero kurtosis $\gamma_{4,s_q} = c_{4,s_q}(0,0,0)$.

B2 : The vector $\boldsymbol{v}(n) \in \mathbb{C}^{M \times 1}$ is the additive Gaussian noise at the output of the antenna array. It is independent from the input signals and has unknown spatial correlation.

B3 : The transmission channel is characterized by a Rayleigh flat fading propagation environment, i.e. the channel coefficients $h_{m,q}$ are complex constants with real and imaginary parts driven from a continuous Gaussian distribution.

By blind channel (or mixture) identification, we understand the problem of estimating the channel model coefficients with no a priori knowledge on the array manifold, i.e. estimate the column vectors $\mathbf{H}_{\cdot q}$, $q \in [1,Q]$, in an arbitrary order, up to a nonzero complex gain, using the 4th-order output statistics only. Actually, it is well-known that solutions to the blind channel identification problem only exist up to a column scaling and permutation indeterminacy.

Assumption B3 allows us to say that $\mathbf{H}$ is full-rank with probability one. Moreover, since any combination of the columns of $\mathbf{H}$ can be viewed as another random matrix driven from a

continuous distribution, $\mathbf{H}$ is also said to be full k-rank, i.e. $k_{\mathbf{H}} = r_{\mathbf{H}} = \min(M, Q)$ [41]. In addition, although the source modulation schemes are generally known in a telecommunication context, we do not make any constraints on the sign of the source Kurtoses γ_{4,s_q}, contrary to other known methods [70, 22].

We define the 4th-order spatial cumulants of the array outputs as follows:

$$C_{4,y}(i, j, k, l) \triangleq cum\left[y_i^*(n), y_j(n), y_k^*(n), y_l(n)\right]. \tag{2.31}$$

Under the above mentioned assumptions, it is straightforward to show that:

$$C_{4,y}(i, j, k, l) = \sum_{q=1}^{Q} \gamma_{4,s_q} h_{iq}^* h_{jq} h_{kq}^* h_{lq}, \tag{2.32}$$

Notice that the spatial cumulants defined in (2.31) only exist for $1 \le i, j, k, l \le M$. Let us define the 4th-order tensor $\mathcal{C}^{(4,y)} \in \mathbb{C}^{M \times M \times M \times M}$ with scalar representation given by $C_{4,y}(i, j, k, l)$, i.e. the element in position (i, j, k, l) can be written as in (2.32). Recalling the general formulation (1.52) with $P = 4$, we can write $\mathcal{C}^{(4,y)}$ as a sum of Q rank-1 tensors that can be written as outer products involving four vectors, as follows:

$$\mathcal{C}^{(4,y)} = \sum_{q=1}^{Q} \mathbf{H}_{\cdot q}^* \circ \mathbf{H}_{\cdot q} \circ \mathbf{H}_{\cdot q}^* \circ (\gamma_{4,s_q} \mathbf{H}_{\cdot q}), \tag{2.33}$$

with $\mathbf{H}_{\cdot q} = \sum_{m=1}^{M} h_{m,q} \mathbf{e}_m^{(M)}$, $q \in [1, Q]$. Equation (2.33) is the Parafac decomposition of tensor $\mathcal{C}^{(4,y)}$ with the four Parafac components depending on $\mathbf{H}$ and being given by:

$$\mathbf{A} = \mathbf{H}^*, \quad \mathbf{B} = \mathbf{H}, \quad \mathbf{C} = \mathbf{H}^* \quad \text{and} \quad \mathbf{D} = \mathbf{H}\mathbf{\Gamma}_{4,s}, \tag{2.34}$$

where $\mathbf{\Gamma}_{4,s} = Diag\left(\gamma_{4,s_1}, \ldots, \gamma_{4,s_Q}\right)$.

Uniqueness

Notice from (2.33) that the rank of the 4th-order tensor $\mathcal{C}^{(4,y)}$ is Q. In addition, since $\mathbf{H}$ is assumed to be full k-rank, we have $k_{A^{(1)}} = k_{A^{(2)}} = k_{A^{(3)}} = k_{A^{(4)}} = r_H = \min(M, Q)$. From (1.66), we conclude that the Kruskal uniqueness condition reduces to

$$4r_{\mathbf{H}} \ge 2Q + 3. \tag{2.35}$$

We will consider the two following cases:

- The MIMO channel is an overdetermined system, i.e. $M \ge Q$. In this case $r_H = Q$ and (2.35) states that the Parafac decomposition of $\mathcal{C}^{(4,y)}$ is essentially unique if $Q \ge 3/2$, i.e. $Q > 1$. There are no further constraints on the number of sensors.

- The MIMO channel is an underdetermined system, i.e. $M < Q$. In this case $r_H = M$ and hence equation (2.35) becomes

$$Q \le \frac{4M - 3}{2}. \tag{2.36}$$

Although equation (2.36) is not a necessary condition, it establishes an upper bound on the number of sources we are guaranteed to identify using tensor $\mathcal{C}^{(4,y)}$. Under that condition, $\mathcal{C}^{(4,y)}$ can be expressed as the sum of Q rank-1 tensors, up to trivial permutation and scaling ambiguities. In other words, the Parafac components of $\mathcal{C}^{(4,y)}$ are written as in (1.67) with $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ given by (2.34).

Reduced-order cumulant tensor

It is possible to reduce the 4th-order tensor defined in (2.33) to a 3rd-order one by combining the 3D slices of tensor $\mathcal{C}^{(4,y)}$. We can thus reduce the complexity of the above described tensor decomposition. Let us freeze, without loss of generality, the index k of the cumulant tensor $\mathcal{C}^{(4,y)}$ and define the 3D tensors $\mathcal{C}_k^{(3,y)} \in \mathbb{C}^{M \times M \times M}$. Replacing the scalar representation (2.32) in the general formulation (1.52), with $P = 3$, we end up with

$$\mathcal{C}_k^{(3,y)} = \sum_{q=1}^{Q} \mathbf{H}_{\cdot q}^* \circ \mathbf{H}_{\cdot q} \circ \left(\gamma_{4,s_q} h_{kq}^* \mathbf{H}_{\cdot q} \right) \tag{2.37}$$

Summing the above tensors for all $k \in [1, M]$ we get

$$\mathcal{C}^{(3,y)} = \sum_{k=1}^{M} \mathcal{C}_k^{(3,y)} = \sum_{q=1}^{Q} \mathbf{H}_{\cdot q}^* \circ \mathbf{H}_{\cdot q} \circ \left(\gamma_{4,s_q} \mathbf{H}_{\cdot q} \sum_{k=1}^{M} h_{kq}^* \right). \tag{2.38}$$

The 3rd-order tensor $\mathcal{C}^{(3,y)} \in \mathbb{C}^{M \times M \times M}$ has a straightforward Parafac decomposition with the following components:

$$\mathbf{A} = \mathbf{H}^*, \mathbf{B} = \mathbf{H} \quad \text{and} \quad \mathbf{C} = \mathbf{H} \mathbf{\Delta} \mathbf{\Gamma}_{4,s}, \tag{2.39}$$

where $\mathbf{\Delta}$ is a diagonal matrix given by:

$$\mathbf{\Delta} = \sum_{k=1}^{M} D_k \left(\mathbf{H}^* \right). \tag{2.40}$$

Note that $k_{\mathbf{A}} = k_{\mathbf{B}} = k_{\mathbf{C}} = r_{\mathbf{H}} = \min(M, Q)$ and the Kruskal uniqueness condition (1.69) becomes $3r_{\mathbf{H}} \geq 2Q + 2$. This yields an upper bound on the number of identifiable sources, which is given by

$$\begin{cases} Q \geq 2, & \text{for } M \geq Q \quad \text{and} \\ Q \leq (3M - 2)/2 & \text{for } M < Q. \end{cases} \tag{2.41}$$

Under the above condition, the Parafac components of $\mathcal{C}^{(3,y)}$ are written as in (1.70) with $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ given in (2.39).

5   Blind MIMO channel identification algorithms

In this section, we propose two algorithms to estimate the instantaneous MIMO mixing matrix, up to column scaling and permutations. This is achieved by means of a SS-LS minimization

procedure, thanks to the symmetry properties of the 4th-order cumulant. The algorithms proposed in the sequel utilize only one of the unfolded representations of the cumulant tensors by exploiting the relationships (2.34), or (2.39) in the case of the 3rd-order tensor. After that, we present the procedures for estimating the Parafac components of the cumulant tensors separately by means of the classical trilinear and quadrilinear ALS-type algorithms, described in section 1.3.1.

Other (non-ALS) algorithms have been reported in the literature to solve the canonical tensor decomposition problem, notably by means of simultaneous diagonalization of matrices. In fact, it has been shown in [21] that the canonical tensor components can be derived from a simultaneous matrix diagonalization and, most importantly, this leads to weaker uniqueness conditions. Exploiting the symmetric structure of the quadricovariance matrix, the joint diagonalization approach has been used in [68] and [22], giving rise o the ICAR and FOOBI algorithms, respectively. While the FOOBI algorithms induce weaker uniqueness conditions, allowing for the identification of more sources than sensors, the ICAR approach only treats the overdetermined case, by exploiting the redundancies in the 4th-order cumulant. In order to include the case of underdetermined mixtures, the underlying principle behind the ICAR method has also been applied to 6th- [69] and higher-order statistics [70].

Although avoiding prewhitening, both ICAR and FOOBI algorithms come up with solutions that require going through two different optimization stages in order to extract MIMO parameters from an initial EVD-based estimate. The ICAR and FOOBI algorithms are briefly described in the sequel. After that, the SS-LS approach is discussed and an algorithm that minimizes one single LS cost function is proposed, under very mild assumptions.

5.1 Joint-diagonalization based algorithms

Making use of the multilinearity property of the cumulants, several methods have been recently proposed utilizing the JD technique to exploit the Hermitian structure of a certain representation of the cumulant tensor [69, 70, 81, 68, 22]. In fact, from the scalar representation of $\mathcal{C}^{(4,y)}$, we can form the matrix $\mathbf{Q}^{(4,y)} \in \mathbb{C}^{M^2 \times M^2}$, so that

$$\left[\mathbf{Q}^{(4,y)}\right]_{(j-1)M+i,\,(k-1)M+l} = C_{4,y}(i,j,k,l), \tag{2.42}$$

from which, using (2.32), we easily get:

$$\mathbf{Q}^{(4,y)} = \left(\mathbf{H} \diamond \mathbf{H}^*\right)\mathbf{\Gamma}_{4,s}\left(\mathbf{H} \diamond \mathbf{H}^*\right)^{\mathsf{H}}. \tag{2.43}$$

Since the source Kurtoses are assumed to be nonzero[2] and the coefficients of $\mathbf{H} \in \mathbb{C}^{M \times Q}$ are driven from a complex continuous Gaussian distribution, the rank of $\mathbf{Q}^{(4,y)}$ is ensured to be

---

[2] Additionally, all the source Kurtoses are assumed to have equal sign. In the case of under-Gaussian sources ($\gamma_{4,s_q} < 0, \ \forall q \in [1, Q]$), one should replace $\mathbf{Q}^{(4,y)}$ by $-\mathbf{Q}^{(4,y)}$.

equal to the number of signal sources Q [68]. Often referred to as the Quadricovariance matrix [82, 83], $\mathbf{Q}^{(4,y)}$ admits the following (economy-size) EVD:

$$\mathbf{Q}^{(4,y)} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}}, \tag{2.44}$$

where the columns of $\mathbf{U} \in \mathbb{C}^{M^2 \times Q}$ are the eigenvectors of $\mathbf{Q}^{(4,y)}$ associated with the Q largest real-valued eigenvalues, arranged in the diagonal matrix $\mathbf{\Lambda} \in \mathbb{R}^{Q \times Q}$. From (2.44), we get:

$$\mathbf{Q}^{(4,y)^{1/2}} = \mathbf{U}\mathbf{\Lambda}^{1/2}, \tag{2.45}$$

and, since the square root of a Hermitian matrix is unique up to a unitary factor $\mathbf{W}$, we deduce from (2.43) that

$$\mathbf{U}\mathbf{\Lambda}^{1/2} = \left(\mathbf{H} \diamond \mathbf{H}^*\right)\mathbf{\Gamma}_{4,s}^{1/2}\mathbf{W}^{\mathsf{H}}. \tag{2.46}$$

The relationship (2.46) is the core equation for some recently proposed methods proposing solutions to recover the channel coefficients by retrieving the unitary matrix $\mathbf{W}$. This is the case of the so-called *Independent Component Analysis using the Redundancies in the Quadricovariance* (ICAR) algorithm, which uses a JD approach in order to estimate $\mathbf{W}$, exploiting symmetry relationships of the 4th-order cumulants [68]. Also using the JD technique, the *Fourth-Order-Only Blind Identification* (FOOBI) algorithm exploits the rank-1 Kronecker structure intrinsic to the columns of $\mathbf{Q}^{(4,y)^{1/2}}\mathbf{W}$ [22]. These methods are further discussed in the sequel.

The ICAR algorithm (*Albera et al.* [68])

Using the ICAR concept, after the EVD of the Quadricovariance matrix, the channel coefficients are estimated by means of two additional stages: computation of $\mathbf{W}$ using the JD technique and subsequent estimation of $\mathbf{H}$ from $(\mathbf{H} \diamond \mathbf{H}^*)$. This latter Khatri-Rao product can be written as:

$$\left(\mathbf{H} \diamond \mathbf{H}^*\right) = \begin{pmatrix} \mathbf{H}^* D_1\left(\mathbf{H}\right) \\ \vdots \\ \mathbf{H}^* D_M\left(\mathbf{H}\right) \end{pmatrix}, \tag{2.47}$$

from which we readily deduce, using (2.45) and (2.46), that $\mathbf{Q}^{(4,y)^{1/2}} = \left[\mathbf{\Omega}_1^{\mathsf{T}} \ \cdots \ \mathbf{\Omega}_M^{\mathsf{T}}\right]^{\mathsf{T}}$, where each block $\mathbf{\Omega}_m \in \mathbb{C}^{M \times Q}$, $m \in [1, M]$, is given as follows:

$$\mathbf{\Omega}_m = \mathbf{H}^* D_m\left(\mathbf{H}\right)\mathbf{\Gamma}_{4,s}^{1/2}\mathbf{W}^{\mathsf{H}}, \quad m \in [1, M], \tag{2.48}$$

and shown to be full-rank. Let $\mathbf{\Omega}_m^{\#}$ denote the pseudoinverse of the above defined matrix, i.e. $\mathbf{\Omega}_m^{\#} = (\mathbf{\Omega}_m^{\mathsf{H}}\mathbf{\Omega}_m)^{-1}\mathbf{\Omega}_m^{\mathsf{H}}$, and define the following set of $M(M-1)$ matrices:

$$\mathbf{\Theta}_{m_1,m_2} = \mathbf{\Omega}_{m_1}^{\#}\mathbf{\Omega}_{m_2}, \quad 1 \leq m_1 \neq m_2 \leq M, \tag{2.49}$$

so that, after some straightforward manipulations, we end up with

$$\mathbf{\Theta}_{m_1,m_2} = \mathbf{W} D_{m_1}\left(\mathbf{H}\right) D_{m_2}\left(\mathbf{H}\right)\mathbf{W}^{\mathsf{H}}, \quad 1 \leq m_1 \neq m_2 \leq M. \tag{2.50}$$

Due to (2.49), the ICAR approach is limited to treat only the overdetermined case $(Q \leq M)$. The above equation shows that the unitary factor $\mathbf{W}$ is a *joint-diagonalizer* of the $M(M-1)$ matrices $\mathbf{\Theta}_{m_1,m_2}$ defined in (2.49). Thus, $\mathbf{W}$ can be approximately computed from the JD of those matrices, using the extended Jacobi technique, by means of the Algorithm A.2, described in the Appendix A. However, we know from [65] that the solution $\mathbf{W}_o$ of the JD problem is unique only up to column permutations and scaling. Hence, by multiplying the output of the JD algorithm by any of the $M \times Q$ blocks $\mathbf{\Omega}_m$ defined in (2.48) yields

$$\mathbf{\Omega}_m \mathbf{W}_o = \mathbf{H}^* D_m (\mathbf{H}) \mathbf{\Gamma}_{4,s}^{1/2} \mathbf{\Lambda} \mathbf{\Pi} \tag{2.51}$$

where matrix $\mathbf{\Lambda}$ is nonsingular and diagonal and $\mathbf{\Pi}$ is a permutation matrix. Otherwise, by right-multiplying $\mathbf{Q}^{(4,y)^{1/2}}$ by $\mathbf{W}_o$, and using (2.45) and (2.46), we get an estimate of $(\mathbf{H} \diamond \mathbf{H}^*)$, up to the trivial indeterminacies, i.e.

$$\mathbf{Q}^{(4,y)^{1/2}} \mathbf{W}_o = \left(\mathbf{H} \diamond \mathbf{H}^* \right) \mathbf{\Gamma}_{4,s}^{1/2} \mathbf{\Lambda} \mathbf{\Pi}. \tag{2.52}$$

To obtain $\mathbf{H}$ from the above equation, note that $unvec\left(\mathbf{H}_{.q} \otimes \mathbf{H}_{.q}^*, M\right) = \mathbf{H}_{.q}^* \mathbf{H}_{.q}^\mathsf{T}$, which is a rank-1 matrix, with $\otimes$ denoting the Kronecker product and $unvec(\mathbf{x}, n)$ being the *unvectorization* operator, which builds from the vector $\mathbf{x}$ a matrix with n columns. As a result, by mapping the qth column of $\mathbf{Q}^{(4,y)^{1/2}} \mathbf{W}_o$ into a $M \times M$ matrix $\mathbf{B}_q$, the column $\mathbf{H}_{.q}$ can be obtained, up to a scaling factor, as the eigenvector associated with the largest eigenvalue of $\mathbf{B}_q^*$. This solution is referred to as the ICAR3 algorithm, and is summarized below.

Algorithm 2.3 (ICAR algorithm)

1. *Compute* $\mathbf{U} \in \mathbb{C}^{M^2 \times Q}$ *and* $\mathbf{\Lambda} \in \mathbb{R}^{Q \times Q}$ *from the EVD of* $\hat{\mathbf{Q}}^{(4,y)}$ *as in (2.44); Deduce the sign of the source Kurtoses from the diagonal elements of* $\mathbf{\Lambda}$; *If it is negative, use* $-\hat{\mathbf{Q}}^{(4,y)}$ *instead.*

2. *Take the square-root of* $\hat{\mathbf{Q}}^{(4,y)}$ *as in (2.45), i.e.* $\hat{\mathbf{Q}}^{(4,y)^{1/2}} = \mathbf{U}\mathbf{\Lambda}^{1/2}$;

3. *Deduce the M matrices* $\mathbf{\Omega}_m \in \mathbb{C}^{M \times Q}$, *by taking for each* $m \in [1, M]$ *the rows* $(m-1)M + 1$ *to* mM *of* $\hat{\mathbf{Q}}^{(4,y)^{1/2}}$;

4. *Using (2.49), form the set of* $M(M-1)$ *matrices* $\mathbf{\Theta}_{m_1,m_2}$, *for all* $1 \leq m_1 \neq m_2 \leq M$;

5. *Using Algorithm A.2 compute the matrix* $\mathbf{W}_o$ *that simultaneously diagonalizes* $\mathbf{\Theta}_{m_1,m_2}$, *for all* $1 \leq m_1 \neq m_2 \leq M$;

6. *Compute* $\mathbf{Q}^{(4,y)^{1/2}} \mathbf{W}_o$ *and denote its* qth *column by* $\mathbf{b}_q \in \mathbb{C}^{M^2 \times 1}$; *Deduce* $\mathbf{B}_q = unvec(\mathbf{b}_q, M) \in \mathbb{C}^{M \times M}$ *for all* $q \in [1, Q]$;

7. *Estimate the* qth *column* $\hat{\mathbf{H}}_{.q}$ *of the mixture matrix as the eigenvector of* $\mathbf{B}_q^*$ *associated with its largest eigenvalue, for all* $q \in [1, Q]$.

Other solutions to extract $\hat{\mathbf{H}}$ from the matrix $\mathbf{Q}^{(4,y)1/2}\mathbf{W}_o$ are reported in [68] but will not be considered here. Finally, we note that extending the ICAR concept to 6th- or higher-order statistics allows for treating some underdetermined cases, which is the case of the BIRTH and the BIOME algorithms, respectively [69, 70]. Since we limit our analyses to the methods using 4th-order statistics, these two latter algorithms will not be further discussed. In addition, the FOBIUM algorithm [81] can be viewed as an extension to high orders of the classical SOS-based SOBI algorithm [65]. FOBIUM includes a 4th-order based prewhitening step and it is unable to deal with sources that have similar trispectra, analogously to the SOBI algorithm, which is theoretically insensitive to the presence of sources with the same spectral densities. For that reason, these algorithm will not be considered in this thesis.

The FOOBI algorithm (*De Lathauwer et al.* [22])

The FOOBI algorithm exploits the Khatri-Rao structure obtained when we right multiply $\mathbf{Q}^{(4,y)1/2}$ by $\mathbf{W}$. The basic principle behind this method relies on a rank-1 detecting device, which takes the form of a 4th-order tensor $\Phi(\mathbf{X}, \mathbf{Y}) \in \mathbb{C}^{M \times M \times M \times M}$ with scalar representation given by

$$\left[\Phi(\mathbf{X}, \mathbf{Y})\right]_{ijkl} = \mathbf{X}_{ij}\mathbf{Y}^*_{kl} + \mathbf{Y}_{ij}\mathbf{X}^*_{kl} - \mathbf{X}_{ik}\mathbf{Y}^*_{jl} - \mathbf{Y}_{ik}\mathbf{X}^*_{jl}, \tag{2.53}$$

where $\mathbf{X}$ and $\mathbf{Y}$ are $M \times M$ matrices. It has been shown in [22] that $\Phi(\mathbf{X}, \mathbf{X})$ is an all-zero tensor if and only if the rank of $\mathbf{X}$ is at most equal to one (see also [24, 21]). Let us denote by $\mathbf{b}_q$ the qth $M^2 \times 1$ column of $\mathbf{Q}^{(4,y)1/2} = \mathbf{U}\mathbf{\Lambda}^{1/2}$ and define the Q Hermitian matrices $\mathbf{B}_q = unvec(\mathbf{b}_q, M) \in \mathbb{C}^{M \times M}$, $q \in [1, Q]$. Moreover, define the 4th-order tensors $\mathcal{F}_{q_1,q_2} = \Phi(\mathbf{B}_{q_1}, \mathbf{B}_{q_2})$, $1 \le q_1 \le q_2 \le Q$. Assuming that the tensors $\mathcal{F}_{q_1,q_2}$, $1 \le q_1 < q_2 \le Q$, are linearly independent, it is possible to show that there exist Q real-valued linearly independent symmetric matrices $\mathbf{V}_q \in \mathbb{R}^{Q \times Q}$, satisfying

$$\sum_{q_1=1}^{Q}\sum_{q_2=1}^{Q}\left[\mathbf{V}_q\right]_{q_1,q_2}\mathcal{F}_{q_1,q_2} = \mathbf{0}_{M \times M \times M \times M}, \tag{2.54}$$

and being simultaneously diagonalized by $\mathbf{W}$, i.e. $\mathbf{V}_q = \mathbf{W}\mathbf{D}_q\mathbf{W}^\mathsf{T}$, $q \in [1, Q]$, with $\mathbf{D}_q$ real-valued and diagonal. Therefore, $\mathbf{W}$ can be obtained from the JD of the linearly independent symmetric matrices $\mathbf{V}_q$, by means of the extended Jacobi Algorithm A.2.

Matrices $\mathbf{V}_q$ can be computed from the set of equations (2.54), which can be rewritten as:

$$\mathbf{F}\mathbf{v}_q = \mathbf{0}_{M^4}, \qquad \text{where,} \tag{2.55}$$

$$\mathbf{v}_q = \left[[\mathbf{V}_q]_{1,1}, \ldots, [\mathbf{V}_q]_{Q,Q}, 2[\mathbf{V}_q]_{1,2}, 2[\mathbf{V}_q]_{1,3}, \ldots, 2[\mathbf{V}_q]_{Q-1,Q}\right]^\mathsf{T}, \qquad \text{and} \tag{2.56}$$

$$\mathbf{F} = \left[\mathbf{f}_{1,1} \ldots \mathbf{f}_{Q,Q} \mathbf{f}_{1,2} \mathbf{f}_{1,3} \ldots \mathbf{f}_{Q-1,Q}\right] \in \mathbb{C}^{M^4 \times Q(Q-1)/2}, \tag{2.57}$$

where the vectors $\mathbf{f}_{q_1,q_2} \in \mathbb{C}^{M^4 \times 1}$ are the *vectorized* versions of the tensors $\mathcal{F}_{q_1,q_2}$, given as follows:

$$\left[\mathbf{f}_{q_1,q_2}\right]_{(i-1)M^3+(j-1)M^2+(k-1)M+l} = \left[\mathcal{F}_{q_1,q_2}\right]_{ijkl}. \tag{2.58}$$

The LS solution of the linear system (2.55) is given by the Q right singular vectors $\mathbf{v}_q \in \mathbb{R}^{Q(Q-1)/2 \times 1}$ associated with its smallest singular values of $\mathbf{F}$. These singular vectors are then mapped into upper triangular matrices $\bar{\mathbf{V}}_q \in \mathbb{R}^{Q \times Q}$ in the order suggested by (2.56), i.e.

$$
\bar{\mathbf{V}}_q = \begin{pmatrix}
[\mathbf{v}_q]_1 & [\mathbf{v}_q]_{Q+1} & [\mathbf{v}_q]_{Q+2} & \cdots & [\mathbf{v}_q]_{2Q-2} & [\mathbf{v}_q]_{2Q-1} \\
0 & [\mathbf{v}_q]_2 & [\mathbf{v}_q]_{2Q} & \cdots & [\mathbf{v}_q]_{3Q-4} & [\mathbf{v}_q]_{3Q-3} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & [\mathbf{v}_q]_{Q-1} & [\mathbf{v}_q]_{Q(Q-3)/2} \\
0 & 0 & 0 & \cdots & 0 & [\mathbf{v}_q]_Q
\end{pmatrix}
\tag{2.59}
$$

and we finally get

$$
\mathbf{V}_q = \frac{\left(\bar{\mathbf{V}}_q + \bar{\mathbf{V}}_q^{\mathsf{T}} \right)}{2}.
\tag{2.60}
$$

Algorithm 2.4 (FOOBI algorithm)

1. *Compute* $\mathbf{U} \in \mathbb{C}^{M^2 \times Q}$ *and* $\mathbf{\Lambda} \in \mathbb{R}^{Q \times Q}$ *from the EVD of* $\hat{\mathbf{Q}}^{(4,y)}$ *as in (2.44); Deduce the sign of the source Kurtoses from the diagonal elements of* $\mathbf{\Lambda}$; *If it is negative, use* $-\hat{\mathbf{Q}}^{(4,y)}$ *instead.*

2. *Take the square-root of* $\hat{\mathbf{Q}}^{(4,y)}$ *as in (2.45), i.e.* $\hat{\mathbf{Q}}^{(4,y)^{1/2}} = \mathbf{U}\mathbf{\Lambda}^{1/2}$; *Denote by* $\mathbf{b}_q$ *the qth column of* $\hat{\mathbf{Q}}^{(4,y)^{1/2}}$ *and form Q Hermitian matrices* $\mathbf{B}_q = unvec\,(\mathbf{b}_q, M) \in \mathbb{C}^{M \times M}$, $q \in [1, Q]$;

3. *Using (2.53), form the 4th-order tensors* $\mathcal{F}_{q_1,q_2} = \Phi(\mathbf{B}_{q_1}, \mathbf{B}_{q_2})$, *for all* $1 \le q_1 \le q_2 \le Q$ *and build vectors* $\mathbf{f}_{q_1,q_2} \in \mathbb{C}^{M^4 \times 1}$ *as follows:*

$$
\left[\mathbf{f}_{q_1,q_2} \right]_{(i-1)M^3 + (j-1)M^2 + (k-1)M + l} = \left[\mathcal{F}_{q_1,q_2} \right]_{ijkl}.
$$

4. *Form the matrix* $\mathbf{F} \in \mathbb{C}^{M^4 \times Q(Q-1)/2}$ *by concatenating the vectors* $\mathbf{f}_{q_1,q_2}$ *for all* $1 \le q_1 \le q_2 \le Q$; *Compute the Q right singular vectors* $\mathbf{v}_q \in \mathbb{R}^{Q(Q-1)/2 \times 1}$ *associated with the smallest singular values of* $\mathbf{F}$;

5. *From each* $\mathbf{v}_q$, $q \in [1, Q]$, *form a triangular matrix* $\bar{\mathbf{V}}_q \in \mathbb{R}^{Q \times Q}$ *as indicated in (2.59); Compute Q matrices* $\mathbf{V}_q = (\bar{\mathbf{V}}_q + \bar{\mathbf{V}}_q^T)/2$;

6. *Using Algorithm A.2 compute the orthogonal matrix* $\mathbf{W}$ *that simultaneously diagonalizes the set* $\mathbf{V}_q$, $q = 1, \dots, Q$; *Compute* $\mathbf{Q}^{(4,y)^{1/2}}\mathbf{W}$ *and denote its qth column by* $\mathbf{z}_q \in \mathbb{C}^{M^2 \times 1}$;

7. *Form the Q matrices* $\mathbf{Z}_q = unvec\,(\mathbf{z}_q, M) \in \mathbb{C}^{M \times M}$, $q \in [1, Q]$; *Estimate the qth column* $\hat{\mathbf{H}}_{\cdot q}$ *of the mixture matrix as the left singular vector of* $\mathbf{Z}_q^*$ *associated with the largest singular value, for all* $q \in [1, Q]$.

The matter of estimating $\mathbf{H}$ is addressed by exploiting the rank-1 structure that characterizes the columns of $\mathbf{Q}^{(4,y)1/2}\mathbf{W}$. Denote by $\mathbf{z}_q \in \mathbb{C}^{M^2 \times 1}$ the qth column of $\mathbf{Q}^{(4,y)1/2}\mathbf{W}$ and build Q matrices $\mathbf{Z}_q = unvec(\mathbf{z}_q, M) \in \mathbb{C}^{M \times M}$, $q \in [1, Q]$. Since $\mathbf{Z}_q$ are rank-1 matrices for all $q \in [1, Q]$, the qth column of the mixture matrix $\hat{\mathbf{H}}_{.q}$ can be estimated from the left singular vector of $\mathbf{Z}_q^*$ associated with the largest singular value, for all $q \in [1, Q]$. The above described method is referred to as the FOOBI-1 algorithm. We note that a second FOOBI-like algorithm has been proposed in [22] based on a simultaneous off-diagonalization. Since both FOOBI solutions have demonstrated very similar performance, we will only consider FOOBI-1 throughout the rest of this chapter.

A major interest in the FOOBI algorithm relies on the condition ensuring the uniqueness of its solution. In fact, assuming that the tensors $\mathcal{F}_{q_1,q_2}$, $1 \le q_1 < q_2 \le Q$ are linearly independent implies an upper bound on the number of identifiable sources. However, linear independence of the tensors $\mathcal{F}_{q_1,q_2}$ is claimed to be guaranteed under the following condition:

$$Q(Q-1) \le M^2(M-1)^2/2, \qquad (2.61)$$

which allows this algorithm for identifying, with a given number of sensors, more sources than the classical ALS approaches can deal with.

5.2   Single-step least squares PBMCI algorithms

The main idea behind the algorithms proposed in the sequel is to exploit the fact that all the Parafac components of the cumulant tensors depend on the channel matrix $\mathbf{H}$. The parameter estimation algorithms make use of only one among the unfolded representations of the tensors $\mathcal{C}^{(4,y)}$ and $\mathcal{C}^{(3,y)}$. According to the unfolding procedure introduced in section 1.3.1, using equation (1.58) with $P = 4$, and taking (2.34) into account, we get the following for $\mathcal{C}^{(4,y)}$:

$$\mathbf{C}_{[4]}^{(4,y)} = \left(\mathbf{H}^* \diamond \mathbf{H} \diamond \mathbf{H}^*\right)\mathbf{\Gamma}_{4,s}\mathbf{H}^\mathsf{T} \quad \in \mathbb{C}^{M^3 \times M}. \qquad (2.62)$$

The solution to the above is said to be unique if any matrix $\bar{\mathbf{H}}$ satisfying (2.62) is such that $\bar{\mathbf{H}} = \mathbf{H}\mathbf{\Lambda}\mathbf{\Pi}$, where $\mathbf{\Pi}$ is a permutation matrix and $\mathbf{\Lambda}$ a diagonal matrix. A sufficient uniqueness condition for this decomposition has been given in (2.36). For $\mathcal{C}^{(3,y)}$, using (2.39), we get the following from (1.58) with $P = 3$:

$$\mathbf{C}_{[3]}^{(3,y)} = \left(\mathbf{H}^* \diamond \mathbf{H}\right)\mathbf{\Gamma}_{4,s}\mathbf{\Delta}\mathbf{H}^\mathsf{T} \quad \in \mathbb{C}^{M^2 \times M}. \qquad (2.63)$$

In practice, we compose $\mathbf{C}_{[4]}^{(4,y)}$ by filling in its lth column $[\mathbf{C}_{[4]}^{(4,y)}]_{.l}$, $l \in [1, M]$, with the elements $C_{4,y}(i,j,k,l)$ by varying the indices $i, j, k \in [1, M]$ in nested loops with k being the innermost one (fastest) and i the outermost one (slowest). For the 3rd-order tensor, we proceed likewise, except for the index k, which is kept fixed for each value in the interval $[1, M]$. Matrix $\mathbf{C}_{[3]}^{(3,y)}$ is then obtained by summing the M resulting matrices.

4D SS-LS PBMCI algorithm

Equation (2.62) enables us to estimate the MIMO channel matrix by iteratively minimizing a single LS cost function, which is written as follows:

$$\psi(\hat{\mathbf{H}}_{r-1}, \mathbf{H}) \triangleq \|\mathbf{C}_{[4]}^{(4,y)} - (\hat{\mathbf{H}}_{r-1}^* \diamond \hat{\mathbf{H}}_{r-1} \diamond \hat{\mathbf{H}}_{r-1}^*)\mathbf{\Gamma}_{4,s}\mathbf{H}^\mathsf{T}\|_F^2. \tag{2.64}$$

where r denotes the iteration number. The iterative minimization of $\psi(\hat{\mathbf{H}}_{r-1}, \mathbf{H})$ yields the following LS solution:

$$\begin{aligned}
\hat{\mathbf{H}}_r^\mathsf{T} &\triangleq \arg\min_{\mathbf{H}} \psi(\hat{\mathbf{H}}_{r-1}, \mathbf{H}) \\
&= \mathbf{\Gamma}_{4,s}^{-1}(\hat{\mathbf{H}}_{r-1}^* \diamond \hat{\mathbf{H}}_{r-1} \diamond \hat{\mathbf{H}}_{r-1}^*)^\# \mathbf{C}_{[4]}^{(4,y)}
\end{aligned} \tag{2.65}$$

where $\hat{\mathbf{H}}_0$ is initialized as a complex $M \times Q$ Gaussian random matrix. In order to improve estimation at iteration $r \geq 1$, before computing $\hat{\mathbf{H}}_r$, we normalize each column of the previous estimate by its respective norm i.e. $[\hat{\mathbf{H}}_{r-1}]_{\cdot q} \leftarrow [\hat{\mathbf{H}}_{r-1}]_{\cdot q}/\|[\hat{\mathbf{H}}_{r-1}]_{\cdot q}\|$. This normalization step also renders the solution (2.65) independent of the source Kurtosis matrix $\mathbf{\Gamma}_{4,s}$. The algorithm is stopped when $|e(r) - e(r-1)|^2 \leq \varepsilon$, where $e(r) = \|\hat{\mathbf{H}}_r - \hat{\mathbf{H}}_{r-1}\|/\|\hat{\mathbf{H}}_r\|$ and ε is an arbitrary small positive constant.

Our developments in this section only considered the unfolded matrix $\mathbf{C}_{[4]}^{(4,y)}$, without loss of generality. Using any other unfolded representation $\mathbf{C}_{[p]}^{(4,y)}$, $p = 1, 2, 3$, should lead to similar results. The above described method will be referred to as the Single-Step LS Parafac-based Blind MIMO Channel Identification (SS-LS PBMCI) algorithm [63, 64]. The SS-LS approach can also be formulated from tensor $\mathcal{C}^{(3,y)}$ defined in (2.37). This is discussed in the sequel.

3D SS-LS PBMCI algorithm

The SS approach can also be formulated using tensor $\mathcal{C}^{(3,y)}$ defined in (2.37). Equation (2.63) yields the following LS cost function:

$$\psi(\hat{\mathbf{H}}_{r-1}, \mathbf{H}) \triangleq \|\mathbf{C}_{[3]}^{(3,y)} - (\hat{\mathbf{H}}_{r-1}^* \diamond \hat{\mathbf{H}}_{r-1})\mathbf{\Gamma}_{4,s}\hat{\mathbf{\Delta}}_{r-1}\mathbf{H}^\mathsf{T}\|_F^2. \tag{2.66}$$

Iteratively minimizing (2.66) leads to:

$$\hat{\mathbf{H}}_r^\mathsf{T} = \mathbf{\Gamma}_{4,s}^{-1}\hat{\mathbf{\Delta}}_{r-1}^{-1}(\hat{\mathbf{H}}_{r-1}^* \diamond \hat{\mathbf{H}}_{r-1})^\# \mathbf{C}_{[3]}^{(3,y)}. \tag{2.67}$$

Here again, $\hat{\mathbf{H}}_0$ is initialized as a complex $M \times Q$ Gaussian random matrix and $\hat{\mathbf{H}}_{r-1}$ is normalized before computing the next estimate $\hat{\mathbf{H}}_r$. Due to this normalization step, the algorithm is independent of the diagonal matrices $\mathbf{\Gamma}_{4,s}$ and $\hat{\mathbf{\Delta}}_{r-1}$, which do not need to be computed. We will refer to this method as the 3D SS-LS PBMCI algorithm.

5.3   Classical ALS-type PBMCI algorithms

Classical ALS-type algorithms can also be used to solve the blind MIMO channel identification problem. In particular, the QALS and TALS algorithms described in section 1.3.3 provide solutions to the Parafac decomposition of tensors $\mathcal{C}^{(4,y)}$ and $\mathcal{C}^{(3,y)}$, respectively. Methods utilizing these algorithms are discussed in the sequel.

Quadrilinear ALS-PBMCI algorithm

The Parafac components of $\mathcal{C}^{(4,y)}$ can be estimated from its unfolded tensor representations. While $\mathbf{C}_{[4]}^{(4,y)}$ has been given in (2.62), matrices $\mathbf{C}_{[p]}^{(4,y)}$, $p = 1, 2, 3$, can be straightforwardly deduced from Table 1.1 using (2.34). The unfolded representations of tensor $\mathcal{C}^{(4,y)}$ are explicitly given in Table 2.2. A solution to this set of equations can be iteratively obtained by means of the QALS algorithm described in section 1.3.3 (Algorithm 1.2). Denoting by $r = \infty$ the iteration at which convergence is reached, and taking the column scaling and permutation into account, we can write:

$$\begin{aligned}
\hat{\mathbf{A}}_\infty &= \hat{\mathbf{H}}_1^* \mathbf{\Lambda}_1 \mathbf{\Pi}, \\
\hat{\mathbf{B}}_\infty &= \hat{\mathbf{H}}_2 \mathbf{\Lambda}_2 \mathbf{\Pi}, \\
\hat{\mathbf{C}}_\infty &= \hat{\mathbf{H}}_3^* \mathbf{\Lambda}_3 \mathbf{\Pi}, \\
\hat{\mathbf{D}}_\infty &= \hat{\mathbf{H}}_4 \mathbf{\Gamma}_{4,s} \mathbf{\Lambda}_4 \mathbf{\Pi}.
\end{aligned} \tag{2.68}$$

The channel estimates $\hat{\mathbf{H}}_1$ and $\hat{\mathbf{H}}_3$ can be obtained, up to the trivial ambiguities, by simple conjugation of $\hat{\mathbf{A}}_\infty$ and $\hat{\mathbf{C}}_\infty$, respectively. Another solution is obtained by averaging $\hat{\mathbf{A}}_\infty^*$, $\hat{\mathbf{B}}_\infty$, $\hat{\mathbf{C}}_\infty^*$ and $\hat{\mathbf{D}}_\infty$. The above procedure will be referred to as the Quadrilinear ALS Parafac-based Blind MIMO Channel Identification (QALS-PBMCI) algorithm.

The QALS algorithm does not exploit the interdependencies between the Parafac components. In spite of that, we can initialize it with a complex $M \times Q$ Gaussian random matrix $\hat{\mathbf{A}}_0$ and then deduce $\hat{\mathbf{B}}_0$ and $\hat{\mathbf{C}}_0$ using (2.34). After that, the Algorithm 1.2 starts by computing $\hat{\mathbf{D}}_0$. The above described procedure will be referred to as the Quadrilinear ALS Parafac-based Blind MIMO Channel Identification (QALS-PBMCI) algorithm.

Trilinear ALS-PBMCI algorithm

A similar ALS approach can be implemented for decomposing the 3rd-order tensor $\mathcal{C}^{(3,y)}$ defined in (2.37), by means of the Trilinear Parafac-ALS (TALS) algorithm, described in section 1.3.3. In this case, the unfolded tensor representations can be deduced from Table 1.1 using (2.39) and they are explicitly given in Table 2.2. The solution is iteratively obtained from the Algorithm 1.3, which can be initialized with a complex $M \times Q$ Gaussian random matrix $\hat{\mathbf{A}}_0$, with $\hat{\mathbf{B}}_0$ being deduced from (2.39) and (2.40). After convergence of the algorithm, the Parafac components of

Table 2.2: Parafac formulæ for the 3rd- and 4th-order spatial cumulant tensors

Unfolded tensor representations	$P = 4$	$P = 3$
$\mathbf{C}_{[4]}$	$\left(\mathbf{H}^* \diamond \mathbf{H} \diamond \mathbf{H}^*\right)(\mathbf{H}\boldsymbol{\Gamma}_{4,s})^{\mathsf{T}}$	
$\mathbf{C}_{[3]}$	$\left((\mathbf{H}\boldsymbol{\Gamma}_{4,s}) \diamond \mathbf{H}^* \diamond \mathbf{H}\right)\mathbf{H}^{\mathsf{H}}$	$\left(\mathbf{H}^* \diamond \mathbf{H}\right)(\mathbf{H}\boldsymbol{\Delta}\boldsymbol{\Gamma}_{4,s})^{\mathsf{T}}$
$\mathbf{C}_{[2]}$	$\left(\mathbf{H}^* \diamond (\mathbf{H}\boldsymbol{\Gamma}_{4,s}) \diamond \mathbf{H}^*\right)\mathbf{H}^{\mathsf{T}}$	$\left((\mathbf{H}\boldsymbol{\Delta}\boldsymbol{\Gamma}_{4,s}) \diamond \mathbf{H}^*\right)\mathbf{H}^{\mathsf{T}}$
$\mathbf{C}_{[1]}$	$\left(\mathbf{H} \diamond \mathbf{H}^* \diamond (\mathbf{H}\boldsymbol{\Gamma}_{4,s})\right)\mathbf{H}^{\mathsf{H}}$	$\left(\mathbf{H} \diamond (\mathbf{H}\boldsymbol{\Delta}\boldsymbol{\Gamma}_{4,s})\right)\mathbf{H}^{\mathsf{H}}$

$\mathcal{C}^{(3,y)}$ have the following form:

$$
\begin{aligned}
\hat{\mathbf{A}}^{(\infty)} &= \hat{\mathbf{H}}_1^* \boldsymbol{\Lambda}_1 \boldsymbol{\Pi}, \\
\hat{\mathbf{B}}^{(\infty)} &= \hat{\mathbf{H}}_2 \boldsymbol{\Lambda}_2 \boldsymbol{\Pi}, \\
\hat{\mathbf{C}}^{(\infty)} &= \hat{\mathbf{H}}_3 \hat{\boldsymbol{\Delta}}_3 \boldsymbol{\Gamma}_{4,s} \boldsymbol{\Lambda}_3 \boldsymbol{\Pi},
\end{aligned}
\tag{2.69}
$$

from which we can deduce three channel matrix estimates, up to column scaling and permutations. Averaging these three estimates yields a fourth solution. This method will be called the Trilinear ALS Parafac-based Blind MIMO Channel Identification (TALS-PBMCI) algorithm.

Column scaling and permutation indeterminacies, although not explicitly solved by any of the algorithms described in this section, do not represent a concern in the context of blind mixture identification, still allowing for the recovery of the source signals in the overdetermined case. Finally, in order to evaluate the capacity of the algorithms in terms of the bounds on the number of identifiable sources, we show in Table 2.3 the theoretical maximum number of users that each algorithm is capable to identify using a given number of receive antennas (varying from $M = 2$ to $M = 7$). The bounds for the 4D SS-LS and the QALS algorithms are derived from (2.36), while those for the 3D SS-LS and the TALS algorithms come from (2.41). The identifiability condition of the FOOBI algorithm has been given in (2.61). The ICAR algorithm is omitted since it is constrained to the overdetermined case ($Q < M$). In spite of the bounds stated in Table 2.3, the tradeoff between capacity and estimation performance of these algorithms remains an open issue and it will be addressed in the computer simulations section.

Table 2.3: Identifiability conditions of MIMO channel identification algorithms

	$M =$	2	3	4	5	6	7
3D SS-LS and TALS	$Q \leq$	2	3	5	6	8	9
4D SS-LS and QALS	$Q \leq$	2	4	6	8	10	12
FOOBI	$Q \leq$	2	4	9	14	21	30

6   Computer simulations

In this section, we present some computer simulation results in order to assess the performance of the blind identification algorithms proposed in this chapter. We will first consider the case of a SISO-FIR communication channel. We compare the performance of the proposed SS-LS PBCI method with the results obtained using the well-known *Fourth-Order System Identification* (FOSI) algorithm [66], which is based on a joint diagonalization technique. We also compare our method with an algebraic solution that is optimal in the total least squares (TLS) sense, proposed in [7].

After that, we consider in section 6.2 a *quasi-static* MIMO scenario in which the propagation channel is characterized by a Rayleigh flat fading, so that the channel coefficients are drawn from a continuous complex Gaussian distribution and assumed to be time-invariant within the duration of a time-slot consisting of N symbols. The QALS-PBMCI and TALS-PBMCI algorithms are compared with the SS-LS approach. We also present a comparative study to illustrate how the 4D SS-LS PBMCI algorithm performs with respect to some methods reported in the literature. Finally, although our main interest is on mixture identification, we also provide results concerning the recovery of the transmitted symbols using the channel estimates obtained from the proposed methods in both SISO and MIMO cases.

6.1   SISO channel identification

In the SISO case, the parametric channel estimation performance will be evaluated by means of the normalized mean squared error (NMSE) of the estimator, computed as follows:

$$\text{NMSE} = \frac{1}{R} \sum_{r=1}^{R} \frac{\|\hat{\mathbf{h}}_{\langle r \rangle}^{(\infty)} - \mathbf{h}\|^2}{\|\mathbf{h}\|^2}, \tag{2.70}$$

where R is the number of Monte Carlo simulations and $\hat{\mathbf{h}}_{\langle r \rangle}^{(\infty)}$ is the channel estimate obtained after convergence for the rth simulation, assuming perfect knowledge of the channel memory L. Except otherwise stated, 4th-order cumulants are estimated using $N = 1000$ output data samples (length of one time-slot). For each Monte Carlo simulation, a different complex channel coefficient vector has been randomly generated in such a way that minimum-phase, nonminimum-phase as well as maximum-phase channels are allowed to occur. Furthermore, we allow for possibly sparse channels, i.e. channel having some coefficients with very small magnitude, which may lead to numerical instabilities. The results illustrated in the following curves represent the average of $R = 200$ Monte Carlo runs. The input signal is QPSK modulated. In which follows, the results concerning the FOSI algorithm have been obtained by averaging the two solutions proposed in [66], as suggested by the authors.

In fig. 2.2, the NMSE is plotted against signal-to-noise ratio (SNR) for the SS-LS PBCI, FOSI and TLS algorithms. The curves on the left-hand side show that the SS-LS approach performs
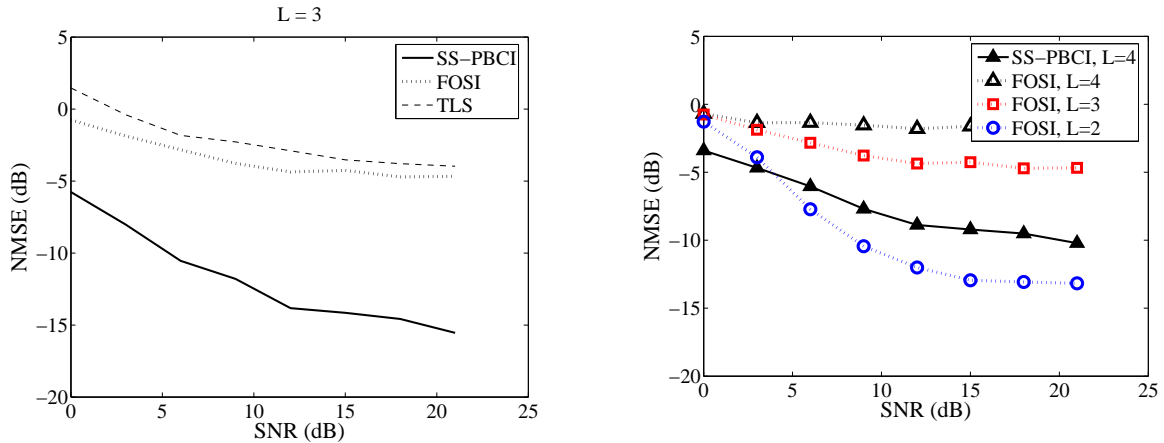
Figure 2.2: NMSE performance with QPSK modulation.

better than both, the FOSI algorithm and the TLS solution, for channels with memory $L = 3$. The relative behavior of the algorithms shown in that figure has also been verified for $L = 2$ and $L = 4$. On the right-hand side of fig. 2.2, we compare the results of SS-LS PBCI for $L = 4$ with those of the FOSI algorithm for $L = 2$, 3 and 4. Note that the estimation errors obtained with SS-LS PBCI for $L = 4$ are smaller than those of FOSI for $L = 4$ and $L = 3$. Furthermore, for low SNR values, the performance provided by the SS-LS PBCI algorithm for channels with $L = 4$ can be considered equivalent (or better) than those obtained with FOSI for channels with $L = 2$. We can therefore conclude that SS-LS PBCI is able to deal with more complicated channel scenarii (larger delay spread) while providing better performance than the other two algorithms, especially in highly noisy situations.
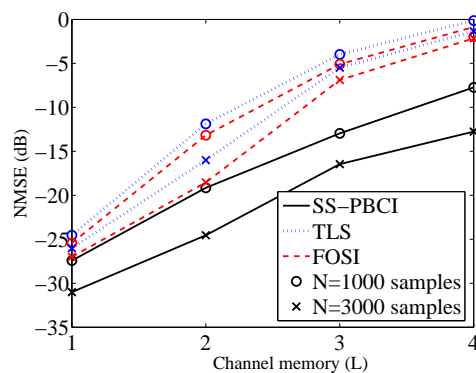
Figure 2.3: NMSE $\times$ channel memory with SNR = 21dB.

In order to evaluate the effect of the output data sequence length used to estimate the 4th-order cumulants over the identification performance of the algorithms, we plot in fig. 2.3 the NMSE against the channel memory for $N = 1000$ and $N = 3000$ output symbols, with
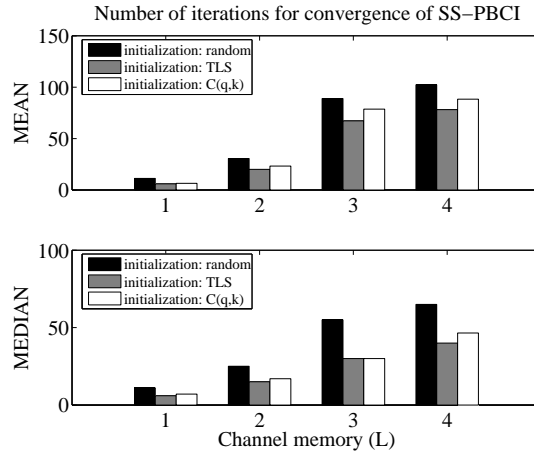
Figure 2.4: Convergence analysis for SS-PBCI with three different initializations (SNR = 21dB).

SNR = 21dB. In a general manner, using SS-LS PBCI with $N = 1000$ yields better results than using TLS or FOSI algorithms with $N = 3000$.

It is interesting to note that the number of iterations required for convergence of the SS-LS PBCI algorithm can be reduced by initializing it with an algebraic solution such as the TLS solution. In fig. 2.4 we show the mean and median number of iterations needed for convergence of SS-LS PBCI with SNR = 21dB using three different initializations: 1) a Gaussian random vector; 2) the TLS solution and 3) the $C(q, k)$ solution [4]. Using both TLS and $C(q, k)$ solutions to initialize SS-LS PBCI decreases the number of iterations in comparison with a random initialization. Finally, it is worth to mention that the NMSE performance after convergence remains unchanged, i.e. initialization only affects convergence speed.

Recovery of the input signal

Several equalization approaches exist to recover the input data sequence using the estimated channel. The optimal solution in the minimum mean squared error (MMSE) sense is given by the Wiener solution [36]. The coefficient vector $\mathbf{w}_o \in \mathbb{C}^{(K+1) \times 1}$ of the optimal equalizer is given by:

$$\mathbf{w}_o = \left(\mathbf{T}^{\mathsf{H}} \mathbf{T} + \sigma_v^2 \mathbf{I}_{(L+1)} \right)^{-1} \mathbf{T}^{\mathsf{H}} \mathbf{s}_d, \tag{2.71}$$

where $\mathbf{T}$ is a $(K + L + 1) \times (K + 1)$ Toeplitz matrix built from the channel coefficients, of which the first row is given by $\mathbf{T}_{1\cdot} = [h(0)\ \mathbf{0}_{K\times 1}^{\mathsf{T}}]$ and the first column is $\mathbf{T}_{\cdot 1} = [\mathbf{h}^{\mathsf{T}}\ \mathbf{0}_{K\times 1}^{\mathsf{T}}]^{\mathsf{T}}$. The vector $\mathbf{s}_d \in \mathbb{R}^{(K+L+1) \times 1}$ is given by $[0, \ldots, 0, 1, 0, \ldots, 0]^{\mathsf{T}}$, where the only nonzero element corresponds to the dth entry and d represents the equalization delay, which is usually taken as $d = (K + L + 1)/2$ if $K + L$ is odd or $d = (K + L + 2)/2$ if $K + L$ is even [36]. The input signal is recovered as follows:

$$\hat{s}(n) = \sum_{k=0}^{K} w_o(k) y(n - k). \tag{2.72}$$
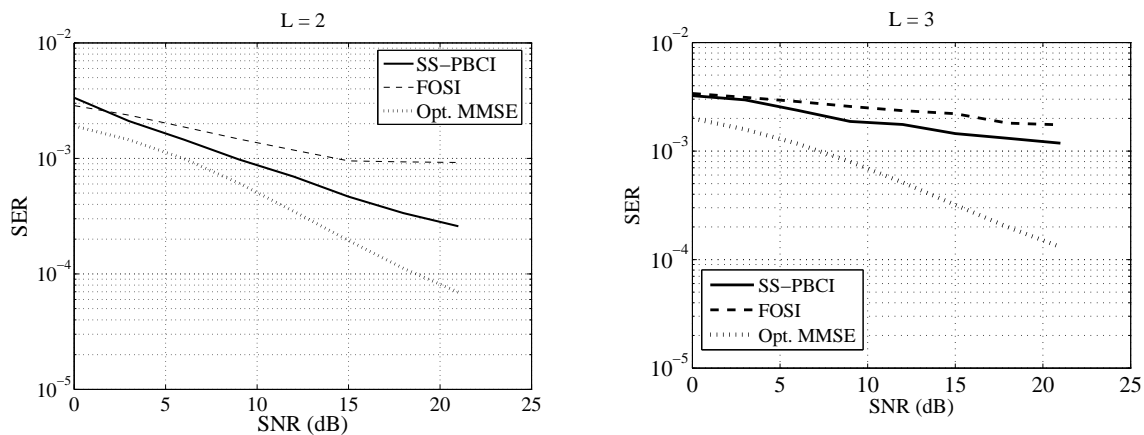
Figure 2.5: Symbol error rate (SER) performance in the SISO case with QPSK modulation for $L = 2$ (left) and $L = 3$ (right).

In fig. 2.5, we present the performance of SS-LS PBCI and FOSI algorithms in terms of the symbol error rate (SER) for channels with $L = 2$ (left) and $L = 3$ (right) with a QPSK modulated input signal. The dotted lines concern the results obtained with the optimal MMSE equalizer assuming perfect knowledge of the channel coefficients. For a target SER of 10^{-3}, with $L = 2$ (left), SS-LS PBCI provides a gain of about 5dB in SNR with respect to FOSI. For $L = 3$ (right), despite the expected performance loss of both algorithms, this gain is around 7dB in SNR for a target SER of 2×10^{-3}.

6.2 MIMO channel identification

In this section, we assume a *quasi-static* Rayleigh flat-fading transmission scenario where the MIMO channel coefficients are drawn from a continuous complex Gaussian distribution and are assumed to be time-invariant within the duration of a time-slot with length equal to N symbol periods. At each time-slot, a new channel is randomly selected. In which follows, the length of the time-slot has been set to $N = 1000$ symbol periods and the output data samples received in this interval are used to estimate the spatial cumulants. Our results are averaged over 300 time-slots.

Overdetermined mixtures

In order to evaluate the performance of the Parafac-based blind MIMO channel identification algorithms in the overdetermined case ($M > Q$), we utilize the identification performance index given in [84, 85], which is based on the matrix $\mathbf{\Phi}^{\langle r \rangle} = \mathbf{H}^{\#}\hat{\mathbf{H}}_{\langle r \rangle}$, where $\hat{\mathbf{H}}_{\langle r \rangle}$ is an estimate of the channel matrix, up to column scaling and permutations, obtained after convergence for the simulation $r \in [1, R]$. Therefore, in the ideal case $\hat{\mathbf{H}}_{\langle r \rangle} = \mathbf{H\Lambda\Pi}$, and hence $\mathbf{\Phi}^{\langle r \rangle}$ should take the form of a scaled permutation matrix. The identification performance index is computed as
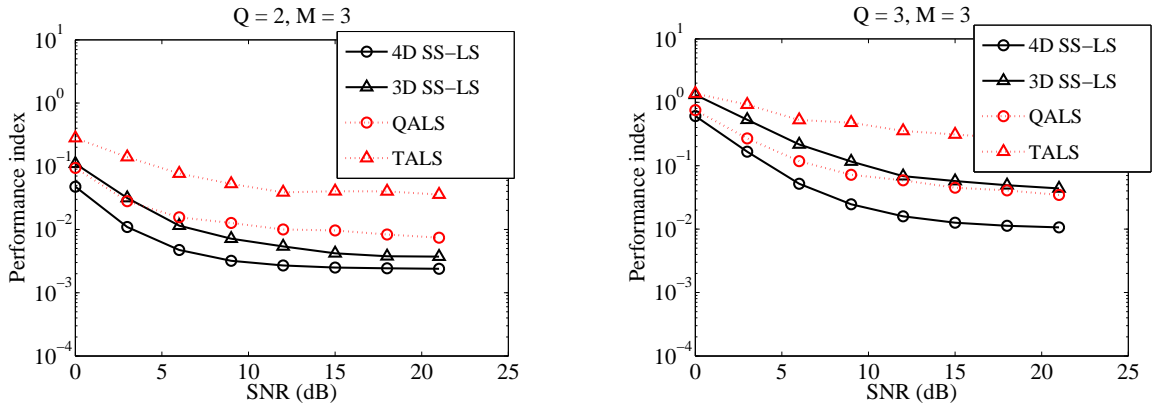
Figure 2.6: Average identification performance index × SNR.

follows:

$$\xi(\boldsymbol{\Phi}^{\langle r \rangle}) \triangleq \frac{1}{2} \left[\left(\sum_i \left(\sum_j \frac{|\phi_{i,j}^{\langle p \rangle}|^2}{\max_\ell |\phi_{i,\ell}^{\langle r \rangle}|^2} \right) - 1 \right) + \left(\sum_j \left(\sum_i \frac{|\phi_{i,j}^{\langle r \rangle}|^2}{\max_\ell |\phi_{\ell,j}^{\langle r \rangle}|^2} \right) - 1 \right) \right], \qquad (2.73)$$

where $\phi_{i,j}^{\langle r \rangle}$ are the entries of $\boldsymbol{\Phi}^{\langle r \rangle}$. The identification performance index $\xi(\cdot)$ equals zero if its matrix argument has the exact structure of a scaled permutation matrix, and small values indicate proximity to the desired solution. In our case, $\xi(\boldsymbol{\Phi}^{\langle r \rangle})$ tends towards zero when the channel estimate approximates the actual MIMO channel matrix, up to column scaling and permutation. Actually, the identification performance index defined in (2.73) is generally viewed as a measure of the quality of source restoration, irrespective of the trivial indeterminacies, indicating the global level of interference rejection at the output of a spatial filter built from the estimated channel coefficients. In the following figures, we plot the value in dB for the average performance index, i.e. $(1/R) \sum_{r=1}^{R} \xi(\boldsymbol{\Phi}^{\langle r \rangle})$, where R is the number of time-slots (Monte Carlo simulations).

In a first simulation experiment, we evaluate the PBMCI approach by comparing the proposed algorithms 4D SS-LS and 3D SS-LS with their ALS-based counterparts (QALS and TALS respectively). We obtained the plots in fig. 2.6 using $M = 3$ antenna elements. These pictures show the average identification performance index computed using (2.73) in function of the SNR for $Q = 2$ (left) and $Q = 3$ (right) QPSK modulated sources. Notice that the methods based on 4th-order tensors (4D SS-LS and QALS) performed generally better than their 3rd-order versions (3D SS-LS and TALS). As expected, increasing the number of sources leads to worse performance, but 4D SS-LS seems to be less affected than the other methods.

In fig. 2.7, we show the mean number of iterations needed for convergence of the four algorithms when $Q = 2$ (left) and $Q = 3$ sources (right) with SNR = $21dB$. Although 4D SS-LS takes generally more iterations to converge than QALS, the former one is a more attractive option due to its smaller computational complexity, since it involves only one LS minimization instead of four. Note that increasing the number of users for a given number of antennas con-
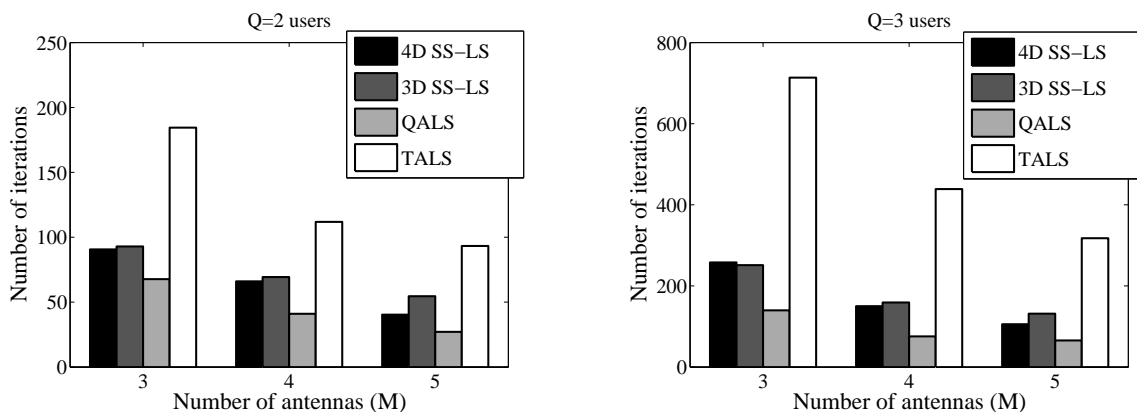
Figure 2.7: Mean number of iterations for convergence with SNR = 21dB.
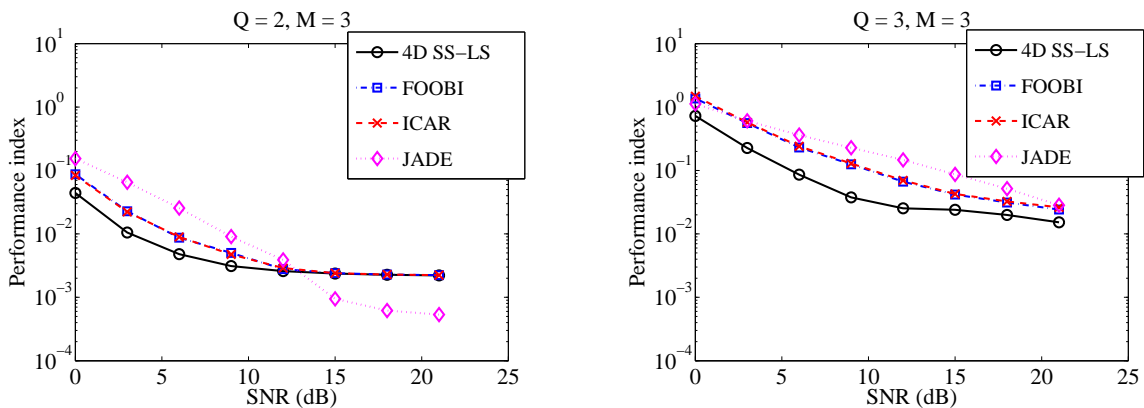
Figure 2.8: Comparison with other algorithms.

siderably increases the number of iterations needed for the tensor-based algorithms to converge. As expected, the methods based on the 4th-order tensor converge faster than those based on the 3rd-order one. Finally, we observe that the algorithms take more iterations to converge when the number of antennas decreases, due to the loss of spatial diversity.

In the sequel, we present some results comparing the SS-LS approach with some algorithms reported in the literature. In particular, we confront the identification performance obtained with the 4D SS-LS PBMCI algorithm against the results provided by the classical JADE [51] algorithm, the FOOBI [22] and the ICAR [68] methods. From fig. 2.8, we note that the 4D SS-LS PBMCI algorithm presents satisfactory results vis-a-vis of the other three methods. By increasing the noise level, JADE's performance degrades and becomes worse than the other methods for SNR values below 12dB. For $Q = 3$ sources and $M = 3$ antennas, fig. 2.8 (right) shows that our approach performs better than the other tested algorithms. In conclusion, the SS-LS approach seems to be a very interesting solution, especially when the number of users increases and the noise level becomes important.
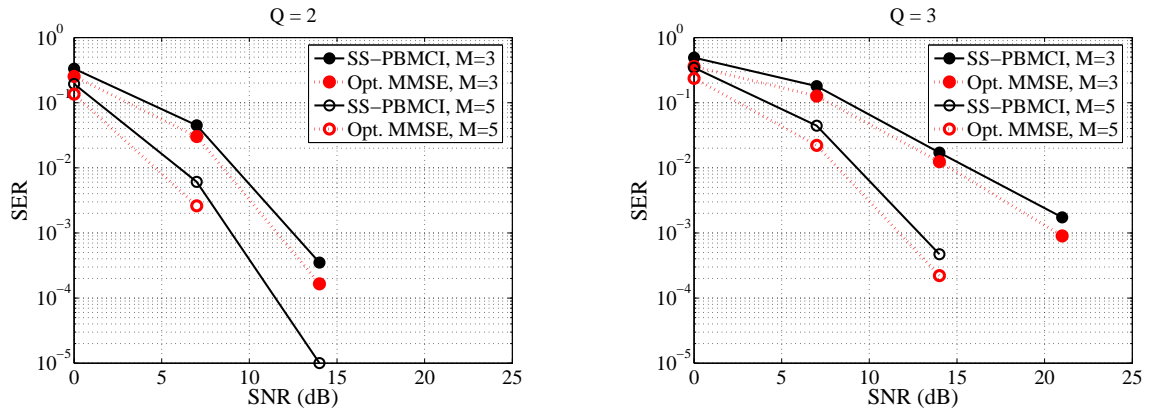
Figure 2.9: SER vs. SNR in the MIMO case.

Finally, concerning the recovery of the source signals, we illustrate in fig. 2.9, the performance of the 4D SS-LS PBMCI algorithm in terms of the average SER per user, for $Q = 2$ users (left) and $Q = 3$ users (right). The source symbols were recovered using a *semiblind* MMSE filter $\mathbf{W} \in \mathbb{C}^{Q \times M}$ built from the estimated MIMO channel matrices, as follows: $\mathbf{W} = \hat{\mathbf{H}}^{\mathsf{H}}(\hat{\mathbf{H}}\hat{\mathbf{H}}^{\mathsf{H}} + \hat{\mathbf{R}}_v)^{-1}$, where $\hat{\mathbf{R}}_v$ is the estimated noise covariance matrix. In order to get rid of the problem of scaling (phase) and permutation ambiguities, P pilot symbols have been used, with $P = 10$ in the case of $Q = 2$ users and $P = 15$ for $Q = 3$ users. The results are compared with those obtained with the optimal MMSE receiver using perfect knowledge of the channel coefficients. Note that, for $Q = 2$ users (left) as well as for $Q = 3$ users (right), the performance of 4D SS-LS is quite close to the optimal MMSE reference. With $Q = 2$ users, the average SER performance has the same global behavior as with $Q = 3$ users, except for a vertical shift in the curves, indicating an expected performance loss due to the increase in the number of users.

Underdetermined mixtures

In order to evaluate the identification algorithms in the underdetermined case, we will consider a uniform linear antenna (ULA) array with M identical sensors, equally spaced of half a wavelength, receiving signal from Q narrow-band sources, assumed to be in the far-field of the antenna array, with azimuth angles given by θ_q, $q \in [1, Q]$, and no elevation angle. The signals are transmitted in the baseband with unit-variance using a QPSK modulation. The mixing matrix coefficients are given by:

$$\left[\mathbf{H}\right]_{m,q} = e^{j\pi(m-1)\cos\theta_q}, \qquad m \in [1, M], \quad q \in [1, Q], \tag{2.74}$$

where $j = \sqrt{-1}$. The estimation accuracy is measured in terms of the NMSE, computed as:

$$\mathrm{NMSE} = \frac{1}{R} \sum_{r=1}^{R} \frac{\|\hat{\mathbf{H}}_{\langle r \rangle}^{(\infty)} - \mathbf{H}\|_F^2}{\|\mathbf{H}\|_F^2}, \tag{2.75}$$
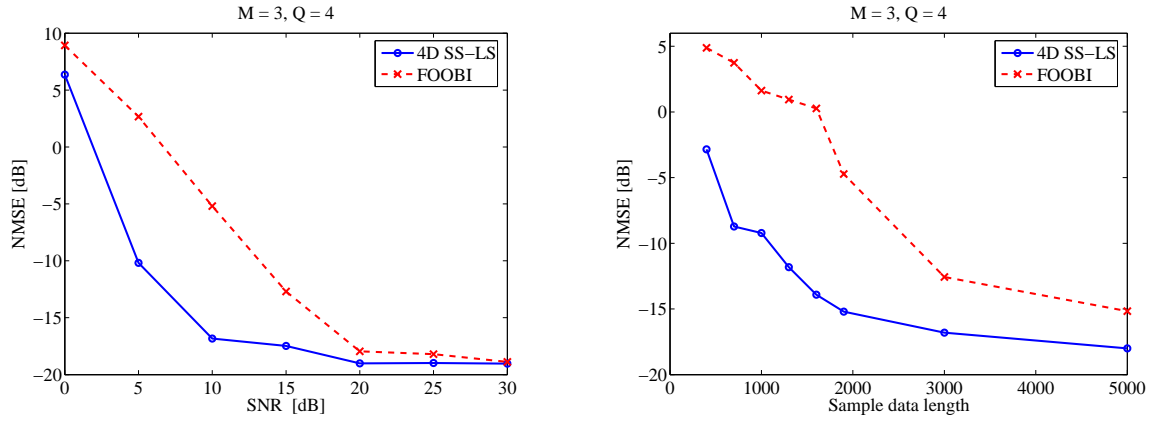
Figure 2.10: NMSE vs. SNR with $N = 1000$ symbols (left) and NMSE vs. Sample data length with SNR=5dB (right).
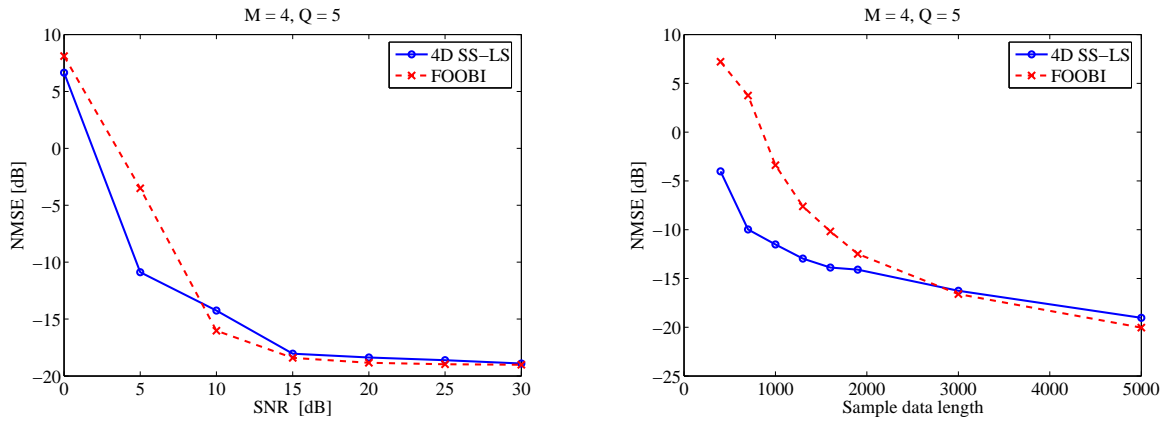
Figure 2.11: NMSE vs. SNR with $N = 1000$ symbols (left) and NMSE vs. Sample data length with SNR=5dB (right).

where R is the number of Monte Carlo simulations and $\hat{\mathbf{H}}_{\langle r \rangle}^{(\infty)}$ is the optimally ordered and scaled channel estimate obtained after convergence for the rth simulation. Our simulations include 300 Monte Carlo runs of each experiment. Firstly, we have used $M = 3$ sensors and $Q = 4$ sources with angles of arrival given by $\theta_1 = 55°$, $\theta_2 = -5°$, $\theta_3 = -50°$ and $\theta_4 = 25°$, respectively. The curves in fig. 2.10 have been obtained with $N = 1000$ symbols for a SNR ranging from 0 to 30dB (left) and with a SNR of 5dB with N varying from 400 to 5000 symbols (right). These results show that 4D SS-LS algorithm gives better precision than FOOBI, specially for low SNR levels. Note from Table 2.3, that in this particular scenario ($M = 3$, $Q = 4$) both methods have the same bounds in terms of the maximum number of identifiable sources.

For the next experiment, we used $M = 4$ sensors and included an extra source with angle of arrival given by $\theta_5 = -25°$. In fig. 2.11, we observe that FOOBI's performance with respect to the 4D SS-LS algorithm is better, and their NMSE practically coincide for high SNR levels (left) as well as for large sample sizes (right). Note from Table 2.3 that in this case ($M = 4$, $Q = 5$), FOOBI has weaker uniqueness conditions than the SS-LS approach and hence more freedom to
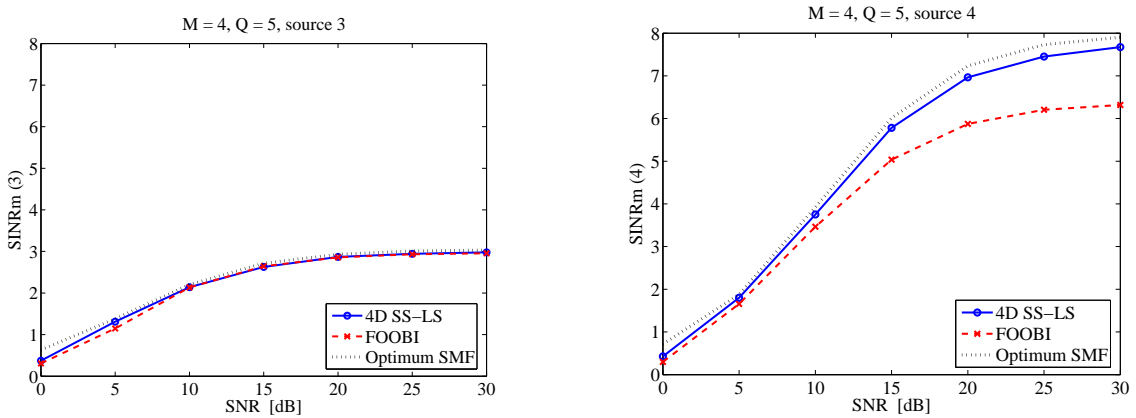
Figure 2.12: Maximum SINR vs. SNR with $N = 1000$ symbols: Best (left) and worst (right) cases.

treat additional sources.

The performance of the identification algorithms can also be assessed in terms of the quality of the extraction of the Q independent components. To this end, we apply a spatial matched filter (SMF) built from the estimated channel coefficients as $\mathbf{W} = \hat{\mathbf{R}}_y^{-1} \hat{\mathbf{H}}$. A performance evaluation criterion can be defined as the maximum signal-to-interference-plus-noise ratio (SINR) per source [25]. The SINR of each source $q \in [1, Q]$ at the ith separator output can be computed as follows:

$$\mathrm{SINR}_q(\mathbf{w}_i) = \sigma_{s_q}^2 \frac{|\mathbf{w}_i^{\mathsf{H}} \hat{\mathbf{H}}_{\cdot q}|^2}{\mathbf{w}_i^{\mathsf{H}} \mathbf{R}_{vq} \mathbf{w}_i}, \tag{2.76}$$

where $\mathbf{w}_i = [\mathbf{W}]_{\cdot i}$, $\sigma_{s_q}^2$ is the variance of source q and $\mathbf{R}_{vq}$ is the total noise plus interference matrix for source q, corresponding to $\mathbf{R}_y = \mathbb{E}\left\{\mathbf{y}(n)\mathbf{y}^{\mathsf{H}}(n)\right\}$ in the absence of source q. The maximum value of $\mathrm{SINR}_q(\mathbf{w}_i)$ for all $i \in [1, Q]$ is indicative of the quality of restitution of source q. In fig. 2.12, we plot the results in function of the SNR level, for the sources $q = 3$ and $q = 4$, which gave respectively the best and worst fit with respect to the optimal SMF, which built from the exact channel coefficients (dotted lines). We conclude that some sources can be better recovered than others in spite of the global performance of the separator.

7 Summary

In this chapter, a new blind SISO channel identification algorithm has been presented based on the Parafac decomposition of a 3rd-order tensor formed of 4th-order output cumulants. The proposed PBCI algorithm relies on a SS-LS minimization problem. The Parafac decomposition fully exploits the three-dimensional nature of the cumulant tensor and has the advantage of avoiding any kind of pre-processing. Uniqueness and convergence issues have been addressed. Computer simulations show that our approach provides better estimation performance than both the TLS solution and the FOSI algorithm, which is based on a simultaneous matrix diagonalization. Furthermore, the convergence of the PBCI algorithm can be accelerated when it is initialized with the TLS solution.

We have also addressed the problem of blind MIMO channel (mixture) identification in the context of a multiuser system characterized by instantaneous complex channels. We have presented a simplified SS-LS MIMO channel identification algorithm based on the Parafac decomposition of a 4th-order tensor composed of 4th-order spatial output cumulants. Quadrilinear and trilinear ALS solutions have been described and compared with the SS-LS method. We have established uniqueness conditions bounding the number of identifiable sources and showing that, under certain conditions, our algorithm can identify underdetermined mixtures. Computer simulations have been presented assessing the performance of the proposed algorithms and comparing it with other MIMO channel identification algorithms, showing that the SS-LS approach is of great interest in several practical situations.

Chapter 3

Parafac-based methods for Array Processing and Multipath Parameter Estimation

High -resolution subspace-based direction finding (DF) methods, such as the well-known MUSIC [86, 87] and ESPRIT [88] algorithms, have become very popular in narrowband (NB) array processing. Exploiting the orthogonality between the signal and noise subspaces, these methods based on the second-order statistics (SOS) provide asymptotically infinite resolution and are very interesting solutions for localizing multiple sources when the spatial correlation of the additive noise is known [89, 90, 91, 92, 93]. However, the performance of SOS-based methods can be seriously deteriorated when dealing with several sources with low signal-to-noise ratio (SNR) and small angular separation using finite data sample sequences [90, 92, 93] or in presence of spatial noise with unknown correlation function [94]. In addition, they can only treat overdetermined mixtures (more sensors than sources).

Source localization is a crucial aspect in sensor array processing. Determining the location of signal emitters allows for the implementation of source separation techniques as well as beamforming for interference suppression. During the last two decades, the use of high-order statistics (HOS) has been widely considered for the estimation of the direction-of-arrival (DOA) in the context of multiuser NB array processing. Several solutions to the source localization and DF problems have been proposed for non-Gaussian signals based on the 4th-order cumulants of the array output data [82, 95, 96]. Extensions of the MUSIC algorithm to the 4th- and higher (even) orders gave rise to the 4-MUSIC [28, 97, 29] and, more recently, the 2κ-MUSIC ($\kappa \geq 2$) methods [27]. In addition to noise robustness, these methods offer better resolution and allow for an increased number of sources to be localized, including certain underdetermined cases. Although characterized by a higher variance [98], the HOS-based MUSIC-like algorithms increase the number of virtual sensors and the effective aperture of the receive antenna array at the

cost of an increased complexity due to the estimation of the high-order statistical information [99, 25, 26]. Other known 4th-order DF approaches include [82, 100, 101], among others.

In the first part of this chapter, we treat the problem of blind multiuser localization in the context of multiple antenna array processing. Assuming that the sources are located at the far-field of the antenna array, our goal is to estimate signal DOAs using only the array output signals. Specifically, we propose a new high-resolution DF algorithm that artificially adds sensors to a virtual antenna array without resorting to statistics of order higher than fourth. In fact, using the 4th-order cumulants only, the proposed method estimates the array matrix and, exploiting the structure of the cumulant tensor, creates an enhanced virtual array that yields an augmented observation space, thus providing additional degrees of freedom to the antenna array and allowing for improved resolution. Based on the single-step least-squares (SS-LS) Parallel Factor (Parafac) decomposition technique introduced in Chapter 2, the new source localization algorithm exploits an array having a double Kronecker structure, which commonly only arises when using 6th-order statistics. However, since we do not need to estimate cumulants of order higher than fourth, our approach keeps the variance of the cumulant estimators at a moderate level, even for quite short output data sequences. Uniqueness and identifiability conditions will be discussed in order to assess the capacity of the proposed technique in terms of the maximum number of resolvable sources. Computer simulations are provided to illustrate the performance of the proposed method compared with the classical MUSIC approaches.

On the other hand, in mobile radiocommunication contexts, signals are often transmitted through multiple propagation paths, characterized by specular reflections and scattering due to physical objects placed in the environment. In such a multipath propagation scenario, the wavefronts may reach the receive array front-end with different delays, spreading the energy of the signals over the time and corrupting temporally adjacent pulses. The so-called delay spread profile induces an altered channel impulse response thus yielding intersymbol interference (ISI), which accounts for important capacity and performance limitations in wireless communication systems. In order to ensure a correct information recovery, we may need to reduce or suppress the effects of ISI, which generally requires some knowledge about the transmission channel. Classically estimated by using known sequences embedded in the transmitted signals, the use of channel coefficients allow for the application of several linear and nonlinear methods aiming the recovery of the input symbols [102]. However, such a supervised (*trained*) approach may be very costly in the context of time-varying channels, even when variations are slow.

Due to the (possibly) nonminimum phase property of the radio channel, the identification problem has been often addressed using a HOS formulation, which gave rise to several non-supervised (*blind*) approaches [5, 103, 104]. Other blind HOS-based approaches include well-known adaptive techniques that are intended to recover the transmitted symbols without the previous channel estimation stage [105, 106, 107, 108]. Since the pioneer paper by Tong et. al [109], blind methods have also been proposed for the single user case based only on SOS, ex-

ploiting subspace properties of multiple output channels corresponding to the different sensors of an antenna array and/or to an oversampled output signal (fractionally-spaced equalization) [110, 111, 112, 113, 114]. See also [115] and references therein for the particular case of single-input multiple-output (SIMO) channels. Still relying on the oversampling of the channel output, extensions of the subspace method to multiuser configurations have been developed in [116, 117] and SOS-based approaches using linear prediction have been proposed in [118, 119].

Starting from section 3, we will focus on the problem of blind multipath channel parameter estimation. Our approach includes two successive stages. We first introduce a 4th-order output cumulant tensor model for the convolutive MIMO channel case, which is an important contribution of this chapter as long as it generalizes the results obtained in Chapter 2 for the SISO and the memoryless MIMO cases. After that, using the physical multipath parameters to model the channel coefficients, we introduce a 3rd-order Parafac tensor representation for the convolutive multiuser radio channel. Using the proposed generalized cumulant tensor framework, we extend the SS-LS algorithm to the non-parametric estimation of the convolutive MIMO channel coefficients. Uniqueness conditions show that the proposed algorithm copes with a flexible range of possible channel configurations, each configuration corresponding to a given number of transmit and receive antennas and a fixed channel memory. Then, an ALS-based algorithm is used to estimate the spatial and temporal channel signatures using the estimated channel coefficients. Finally, the multipath parameters are extracted by means of subspace-based algorithms, which enables us to recover the channel coefficients without the trivial indeterminacies associated with the Parafac decomposition. This new parametric estimation procedure is also an original contribution of this chapter and will be referred to as ALS-MUSIC algorithm.

The two-stage procedure introduced in this chapter for estimating the convolutive MIMO channel and its multipath parameters is, to our knowledge, the first fully blind technique proposing solution to the channel identification problem in a multiuser radio-mobile context, making use of the 4th-order cumulant symmetries in a single LS minimization problem and exploiting the multipath structure of the channel using a tensor formalism in both stages, without resorting to oversampling.

The remaining of this chapter is organized as follows: in section 1, we formulate the array output signal model along with the basic definitions of signal and noise subspaces; we also discuss the Virtual Array concept and present a survey of classic MUSIC-like algorithms, including the general formulation for the case of statistics of any even order; in section 2 we derive a new high-resolution DF algorithm exploiting the double Kronecker structure of the unfolded 4th-order cumulant tensor; the SS-LS cumulant tensor decomposition approach is revisited and the uniqueness issue is addressed; in section 3, we turn our attention to the problem of estimating the multipath parameters of a MIMO channel; we start by introducing a general formulation for the 4th-order output cumulant tensor model in the convolutive MIMO case and then we treat
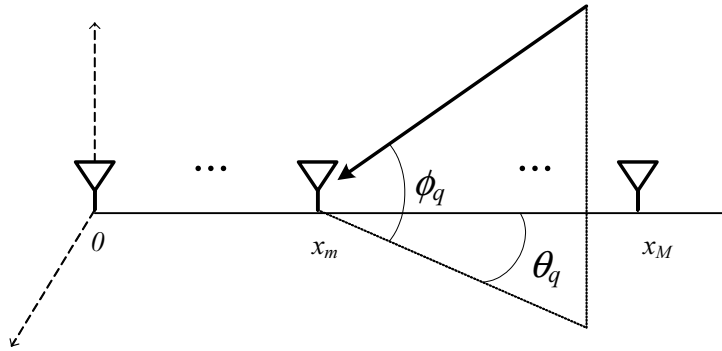
Figure 3.1: Linear antenna array and incoming signal.

the multiuser radio channel as a Parafac tensor formed from its physical multipath parameters; section 4 proposes combined algorithms for the blind identification of the convolutive multipath MIMO channel in two stages: using first a non-parametric 4th-order cumulant-based SS-LS approach and then a parametric ALS-based algorithm that uses the previous estimation of the channel coefficients; after that, the multipath parameters are extracted by means of MUSIC-like subspace-based algorithms that allow for the complete reconstruction of the MIMO channel impulse responses without scaling ambiguities; finally, in section 5, we provide computer simulations illustrating the methods discussed throughout the chapter and assessing their performance under different channel configurations. Conclusions are drawn in section 6, along with some future work perspectives.

1 The source localization problem in NB array processing

Let us consider a linear array of M identical NB sensors receiving the contributions of Q zero-mean stationary sources, assumed to be placed at the far-field of the array. Denoting by $\mathbf{y}(n) \in \mathbb{C}^{M \times 1}$ the vector of complex signals measured at the output of the antenna array, we have:

$$
\begin{aligned}
\mathbf{y}(n) &= \sum_{q=1}^{Q} s_q(n)\mathbf{a}(\theta_q, \phi_q) + \boldsymbol{v}(n) \\
&= \mathbf{A}\mathbf{s}(n) + \boldsymbol{v}(n)
\end{aligned}
\tag{3.1}
$$

where the vector $\mathbf{s}(n) \in \mathbb{C}^{Q \times 1}$ is formed of the complex amplitudes of the source signals $s_q(n)$, which are stationary, ergodic and mutually independent with symmetric distribution, zero-mean and non-zero kurtosis $\gamma_{4,s_q} = c_{4,s_q}(0,0,0)$, $q \in [1, Q]$, with azimuth and elevation angles given by θ_q and ϕ_q, respectively, as illustrated in fig. 3.1. Moreover, the signals $s_q(n)$ are assumed to be independent of the additive Gaussian noise, which is arranged in the vector $\boldsymbol{v}(n) \in \mathbb{C}^{M \times 1}$, stationary with zero-mean and unknown spatial correlation. Matrix $\mathbf{A} \in \mathbb{C}^{M \times Q}$ represents the spatial responses of the linear antenna array, concatenating the source steering vectors

$\mathbf{a}(\theta_q, \phi_q) \in \mathbb{C}^{M \times 1}$, carrying the DOA information (θ_q, ϕ_q) associated with each source $q \in [1, Q]$. The array matrix $\mathbf{A}$ can therefore be written as

$$\mathbf{A} = \left[\mathbf{a}(\theta_1, \phi_1) \dots \mathbf{a}(\theta_Q, \phi_Q) \right] \in \mathbb{C}^{M \times Q}, \tag{3.2}$$

where the mth element of vector $\mathbf{a}(\theta_q, \phi_q)$ corresponds to the response of the array element m with respect to the source q. Assuming a planewave propagation with no coupling between sensors [120, 121], we can write:

$$a_m(\theta_q, \phi_q) = \exp \left\{ \frac{\jmath \, 2 \, \pi \, x_m \cos \theta_q \cos \phi_q}{\lambda} \right\}, \tag{3.3}$$

where $\jmath = \sqrt{-1}$ and x_m is the distance of each array element $m \in [1, M]$ with respect to a given reference sensor, assumed by convention to be the first antenna, i.e. $x_1 = 0$. The signal wavelength λ is given by $\lambda = c/f_c$, where f_c is the carrier frequency and the constant c is the propagation speed of the light. Due to (3.3), $\mathbf{A}$ has a particular unit-modulus property and, since $x_1 = 0$, the directional matrix gets a all-one first row, i.e. $\mathbf{A}_1. = [1, 1, \dots, 1]$. In the case of Uniform Linear Antenna (ULA) arrays, the sensors are equally spaced from each other along the array axis and distanced of Δx with respect to adjacent sensors, so that (3.3) becomes as follows:

$$a_m(\theta_q, \phi_q) = \exp \left\{ \frac{\jmath \, 2 \, \pi \, (m-1) \Delta x \cos \theta_q \cos \phi_q}{\lambda} \right\}. \tag{3.4}$$

In this case, the spatial response array matrix $\mathbf{A}$ has the following Vandermonde structure:

$$\mathbf{A} = \begin{pmatrix} 1 & \dots & 1 \\ a_1(\theta_1, \phi_1) & \dots & a_1(\theta_K, \phi_K) \\ a_1^2(\theta_1, \phi_1) & \dots & a_1^2(\theta_K, \phi_K) \\ \vdots & \ddots & \vdots \\ a_1^{M-1}(\theta_1, \phi_1) & \dots & a_1^{M-1}(\theta_K, \phi_K) \end{pmatrix}, \tag{3.5}$$

where the second row is the generating vector, from which the whole matrix can be deduced.

1.1 Array output statistics

Considering the above mentioned assumptions, let us define the covariance matrix $\mathbf{C}^{(2,y)} \in \mathbb{C}^{M \times M}$, so that $[\mathbf{C}^{(2,y)}]_{i,j} = C_{2,y}(i, j)$, $i, j \in [1, M]$, where the 2nd-order spatial cumulant of the array output is defined as $C_{2,y}(i, j) \triangleq cum \left[y_i(n), y_j^*(n) \right]$. From (3.1), we have:

$$\begin{aligned} \mathbf{C}^{(2,y)} &= \mathbb{E} \left\{ \mathbf{y}(n) \mathbf{y}^\mathsf{H}(n) \right\} \\ &= \mathbf{A} \boldsymbol{\Gamma}_{2,s} \mathbf{A}^\mathsf{H} + \mathbf{C}^{(2,v)} \end{aligned} \tag{3.6}$$

where $\boldsymbol{\Gamma}_{2,s} = \mathbb{E} \left\{ \mathbf{s}(n) \mathbf{s}^\mathsf{H}(n) \right\}$ and $\mathbf{C}^{(2,v)} = \mathbb{E} \left\{ \boldsymbol{v}(n) \boldsymbol{v}^\mathsf{H}(n) \right\}$. From the assumption of mutually independence of the sources, it follows that $\boldsymbol{\Gamma}_{2,s}$ is a diagonal matrix with diagonal entries given by the source variances, $\gamma_{2,s_q} = \mathbb{E} \left\{ |s_q(n)|^2 \right\}$, $q \in [1, Q]$.

Moreover, by defining the 4th-order tensor $\mathcal{C}^{(4,y)} \in \mathbb{C}^{M \times M \times M \times M}$ with scalar representation given by $C_{4,y}(i,j,k,l) \triangleq cum\left[y_i^*(n), y_j(n), y_k^*(n), y_l(n)\right]$, we can build the Quadricovariance matrix $\mathbf{C}^{(4,y)} \in \mathbb{C}^{M^2 \times M^2}$, as $[\mathbf{C}^{(4,y)}]_{(j-1)M+i,\,(k-1)M+l} = C_{4,y}(i,j,k,l)$, yielding the structure given below [82, 83]:

$$\mathbf{C}^{(4,y)} = \left(\mathbf{A} \diamond \mathbf{A}^*\right) \mathbf{\Gamma}_{4,s} \left(\mathbf{A} \diamond \mathbf{A}^*\right)^{\mathsf{H}}, \tag{3.7}$$

where $\mathbf{\Gamma}_{4,s} = Diag\left(\gamma_{4,s_1}, \ldots, \gamma_{4,s_Q}\right)$ and $\diamond$ denotes the Khatri-Rao product.

Comparing (3.7) with (3.6), we note strong similarities in the structures of $\mathbf{C}^{(4,y)}$ and (the noiseless part of) $\mathbf{C}^{(2,y)}$. While both are diagonal quadratic forms, the latter one is built from the source steering vectors, and $\mathbf{C}^{(4,y)}$ involves a column-wise Kronecker product of those vectors. This structural analogy is the basic idea allowing for extending some array processing methods based on SOS to the 4th-order [82, 47]. In addition, since the above analysis only evokes the linearity and the additivity properties of cumulants, it can be extended to statistics of any (even) order. In fact, complex-valued 2κth-order output cumulants, defined as $C_{2\kappa,y}(i_1, \ldots, i_{2\kappa}) \triangleq cum\left[y_{i_1}(n), \ldots, y_{i_\kappa}(n), y_{i_{\kappa+1}}^*(n), \ldots, y_{i_{2\kappa}}^*(n)\right]$, $\kappa \geq 2$, can always be represented by a Hermitian matrix $\mathbf{C}_\ell^{(2\kappa,y)} \in \mathbb{C}^{M^\kappa \times M^\kappa}$, which admits the following decomposition:

$$\mathbf{C}_\ell^{(2\kappa,y)} = \left(\mathbf{A}^{\diamond \ell} \diamond \mathbf{A}^{*\diamond \kappa - \ell}\right) \mathbf{\Gamma}_{2\kappa,s} \left(\mathbf{A}^{\diamond \ell} \diamond \mathbf{A}^{*\diamond \kappa - \ell}\right)^{\mathsf{H}}, \quad \ell \in [1, \kappa], \tag{3.8}$$

where $\mathbf{\Gamma}_{2\kappa,s} = Diag\left(\gamma_{2\kappa,s_1}, \ldots, \gamma_{2\kappa,s_Q}\right)$ and $\gamma_{2\kappa,s_q}$ is the 2κth-order cumulant of the input signal $s_q(n)$. The notation $\mathbf{X}^{\diamond n}$ stands for a multiple Khatri-Rao product involving a matrix $\mathbf{X}$ so that $\mathbf{X}^{\diamond n} = \mathbf{X} \diamond \mathbf{X} \diamond \ldots \diamond \mathbf{X}$, where the Khatri-Rao product symbol $\diamond$ appears $n-1$ times. Throughout the rest of this chapter, we omit the index ℓ, choosing by convention $\ell = \kappa/2$ when κ is even and $\ell = (\kappa+1)/2$ for odd values of κ.

In practical applications, the channel statistics are not known at the array output and must be estimated from the received data sequences, based on the *ergodicity* assumption. Cumulant estimation is an important issue and has been briefly discussed in section 1.1.2 of Chapter 1, in the case of 4th-order cumulants. Exact expressions exist for computing the variance of cumulant estimators of order 2κ, generally involving very complicated calculations using cumulants of order up to 4κ [48, 122, 17]. In this context, an important result shows that, for orders higher than 3, as the sample data length increases, the cumulant estimators tend to be Gaussian random variables [3]. However, the convergence towards Gaussianity may be very slow, as κ increases. Consequently, when dealing with HOS, the use of short sample data sequences may lead to significant errors with respect to the asymptotic results [98].

Signal and noise subspaces

Assuming that the sources are spaced far enough apart from each other, we can consider that the source steering vectors $\mathbf{a}(\theta_q, \phi_q)$, $q \in [1, Q]$, are mutually independent. Under this assumption, the space spanned by these vectors is a subspace of the *observation space*, with M dimensions,

which the received signal vector $\mathbf{y}(n)$ belongs to. Therefore, the column space of matrix $\mathbf{A}$, with Q dimensions, will be referred to as the *signal subspace*, while its orthogonal complement, with $M - Q$ dimensions, $M > Q$, will be named the *noise subspace*.

From (3.6) we conclude that the rank of $\mathbf{C}^{(2,y)}$ is ensured to be equal to the number of sources Q if $Q \leq M$. In this case, we notice that the signal subspace contains all the necessary information for representing the transmitted signals. By projecting the observations on the signal subspace, we can reduce the amount of noise without loss of useful information ($M > Q$). In addition, the number of sources is given by the dimension of the signal subspace.

On the other hand, the rank of $\mathbf{C}^{(4,y)}$ equals Q whenever $Q \leq M^2$. In fact, (3.7) suggests that the observation space of the 4th-order cumulants, with M^2 dimensions, is spanned by the Kronecker products of the received signal vectors $\mathbf{y}(n) \otimes \mathbf{y}^*(n)$. In this case, the signal subspace is defined as the space spanned by the Kronecker products of the steering vectors $\mathbf{a}(\theta_q, \phi_q) \otimes \mathbf{a}^*(\theta_q, \phi_q)$, with Q dimensions. Its orthogonal complement, with $M^2 - Q$ dimensions, will be referred to as the noise subspace.

1.2   The Virtual Array concept

By replacing the received signal vectors by their Kronecker product, we actually increase the dimension of the observation space, thus allowing for a greater number of separable sources [123]. To illustrate this principle, let us consider the Kronecker product $\mathbf{a}(\theta_q, \phi_q) \otimes \mathbf{a}^*(\theta_q, \phi_q)$ and take the element in position $(m_1 - 1)M + m_2$, with $m_1, m_2 \in [1, M]$. Using a ULA array, we can write this element from (3.4), as follows:

$$\left[\mathbf{a}(\theta_q, \phi_q) \otimes \mathbf{a}^*(\theta_q, \phi_q)\right]_{(m_1-1)M+m_2} = a_{m_1}(\theta_q, \phi_q)\, a_{m_2}^*(\theta_q, \phi_q)$$

$$= \exp\left\{\frac{\jmath\, 2\,\pi\, (m_1 - m_2)\Delta x \cos\theta_q \cos\phi_q}{\lambda}\right\}. \quad (3.9)$$

The above equation clearly shows that the Kronecker product of the steering vectors results in an augmented ULA array with *virtual sensors* (VS) placed at the array axis and distanced of $(m_1 - m_2)\Delta x$ with respect to the reference sensor, for all $m_1, m_2 \in [1, M]$. Notice that the M elements for which $m_1 = m_2$ are located at the array origin and are said to be virtual sensors of multiplicity M. In the case of ULA arrays with space diversity only, the 2nd-order *virtual array* (VA) defined in (3.9) is shown to have $2M - 1$ different VS, meaning it can deal with up to $2M - 2$ independent sources [25]. In the general case, using an optimal array geometry, it is possible to get up to $M^2 - M + 1$ different VS [25].

The theory of Virtual Arrays has been introduced independently in [99] and [123] using 4th-order statistics. The concept has been further developed in [25] and [26], for the case of 4th-and higher-order cumulants, respectively. However, some basic results had already been given in [124]. The steering vectors of a κth-order VA are given as follows:

$$\mathbf{a}_\kappa(\theta_q, \phi_q) = \left[\mathbf{a}(\theta_q, \phi_q)^{\otimes \ell} \otimes \mathbf{a}^*(\theta_q, \phi_q)^{\otimes \kappa - \ell}\right], \quad (3.10)$$

where $\mathbf{x}^{\otimes n}$ denotes a multiple Kronecker product involving a vector $\mathbf{x}$ so that $\mathbf{x}^{\otimes n} = \mathbf{x} \otimes \mathbf{x} \otimes \ldots \otimes \mathbf{x}$, where the Kronecker product symbol $\otimes$ appears $n-1$ times. The index ℓ has been omitted from the left-hand side of (3.10) since we have chosen to take, by convention, $\ell = \kappa/2$ for κ even and $\ell = (\kappa + 1)/2$ when κ is odd.

One major interest in using high-order (HO) VAs is in exploiting the Kronecker structure that naturally arises in the HOS representations. Despite the increased variance of the HOS estimators, the HO VAs are known to provide gains in terms of resolution, which can be measured by means of the spatial correlation between two sources. Let us consider the sources q_1 and q_2, with DOAs given by the angles $(\theta_{q_1}, \phi_{q_1})$ and $(\theta_{q_2}, \phi_{q_2})$, respectively. The spatial correlation coefficient of sources q_1 and q_2 is given by the normalized inner product of the steering vectors $\mathbf{a}_\kappa(\theta_{q_1}, \phi_{q_1})$ and $\mathbf{a}_\kappa(\theta_{q_2}, \phi_{q_2})$, i.e.

$$\varsigma_\kappa(q_1, q_2) \triangleq \frac{\mathbf{a}_\kappa(\theta_{q_1}, \phi_{q_1})^\mathsf{H} \mathbf{a}_\kappa(\theta_{q_2}, \phi_{q_2})}{\left[\mathbf{a}_\kappa(\theta_{q_1}, \phi_{q_1})^\mathsf{H} \mathbf{a}_\kappa(\theta_{q_1}, \phi_{q_1})\right]^{1/2} \left[\mathbf{a}_\kappa(\theta_{q_2}, \phi_{q_2})^\mathsf{H} \mathbf{a}_\kappa(\theta_{q_2}, \phi_{q_2})\right]^{1/2}}. \tag{3.11}$$

Notice that for any complex $n \times 1$ vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{d}$ it holds: $(\mathbf{a} \otimes \mathbf{b}^*)^\mathsf{H}(\mathbf{c} \otimes \mathbf{d}^*) = (\mathbf{a}^\mathsf{H}\mathbf{c})(\mathbf{d}^\mathsf{H}\mathbf{b})$; we can thus write $(\mathbf{a}^{\otimes \ell} \otimes \mathbf{a}^{*\otimes \kappa - \ell})^\mathsf{H}(\mathbf{b}^{\otimes \ell} \otimes \mathbf{b}^{*\otimes \kappa - \ell}) = (\mathbf{a}^\mathsf{H}\mathbf{b})^\ell (\mathbf{b}^\mathsf{H}\mathbf{a})^{\kappa - \ell}$, for all $\mathbf{a}, \mathbf{b} \in \mathbb{C}^{n \times 1}$. Replacing this latter relationship into (3.11) and taking its modulus, we obtain:

$$\left|\varsigma_\kappa(q_1, q_2)\right| = \left(\frac{|\mathbf{a}(\theta_{q_1}, \phi_{q_1})^\mathsf{H} \mathbf{a}(\theta_{q_2}, \phi_{q_2})|}{|\mathbf{a}(\theta_{q_1}, \phi_{q_1})| \, |\mathbf{a}(\theta_{q_2}, \phi_{q_2})|}\right)^\kappa, \tag{3.12}$$

from which it is straightforward to deduce that $|\varsigma_\kappa(q_1, q_2)| = |\varsigma_1(q_1, q_2)|^\kappa$. Note that for $\kappa = 1$ the virtual steering vector defined in (3.10) coincides with the actual array response vector $\mathbf{a}(\theta_q, \phi_q)$. Thus, for an array with space diversity only, the spatial correlation coefficient of the κth-order Virtual Array associated with a given direction (θ_0, ϕ_0) only depends on κ and on the normalized amplitude response $|\varsigma_1(q, 0)|$ of the actual antenna array of M sensors, for each pair (θ_q, ϕ_q). Since $0 \leq |\varsigma_1(q, 0)| \leq 1$, we conclude that for the direction (θ_0, ϕ_0), the spatial correlation of the HO VA decreases with κ, thus improving its angular resolution. This fact is illustrated in fig. 3.2 for a ULA array with 3 sensors spaced of $\lambda/2$, where we plot the antenna response of the VAs with $\kappa = 1$, $\kappa = 2$ and $\kappa = 3$, to a source at $\theta_0 = 5°$ (no elevation angle). The antenna pattern is obtained from the inner products of the associated steering vectors. Considering the beamwidth (in degrees) at the point of 3dB attenuation of the main lobe for the DOA of 5°, we observe gains of about 13.5° and 19.8° using the HO VAs with $\kappa = 2$ and $\kappa = 3$, respectively, with respect to the considered array of M sensors ($\kappa = 1$).

1.3 MUSIC-like DF algorithms

In its basic form, the Multiple Signal Classification (MUSIC) technique has been introduced to provide asymptotically unbiased estimates of the parameters of multiple wavefronts arriving at an antenna array [86, 87]. Exploiting the orthogonality between the signal and noise subspaces,
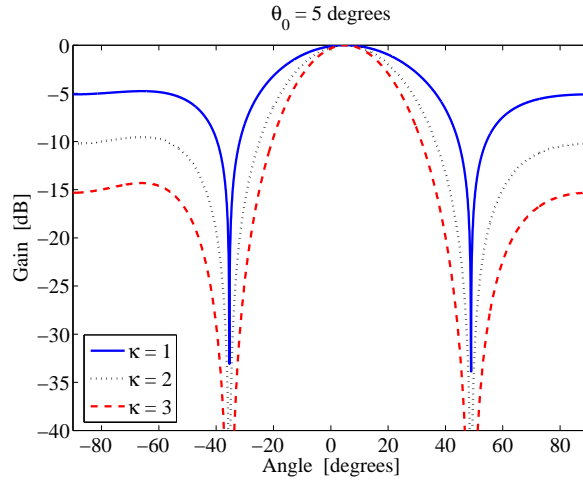
Figure 3.2: Antenna pattern of different VAs for a ULA with 3 sensors spaced of $\lambda/2$.

the MUSIC algorithm aims to determine, among others, the number of signal sources, their location (DOAs) and the cross-correlations among the directional waveforms.

2-MUSIC

The SOS-based MUSIC algorithm (so-called 2-MUSIC) algorithm is of particular interest: 1) in the noiseless case, in which (3.6) leads to $\mathbf{C}^{(2,y)} = \mathbf{A}\boldsymbol{\Gamma}_{2,s}\mathbf{A}^{\mathsf{H}}$; 2) in the case of a spatially white additive noise, when $\mathbf{C}^{(2,v)}$ is proportional to the identity matrix, i.e. $\mathbf{C}^{(2,v)} = \gamma_{2,v}\mathbf{I}_M$, with $\gamma_{2,v}$ being the noise variance. In this latter case, the smallest eigenvalues of $\mathbf{C}^{(2,y)}$ equal the noise variance and the corresponding eigenvectors span the noise subspace of the noiseless observation space. The eigenvectors associated with the largest eigenvalues span the signal subspace and the EVD of the estimated covariance matrix is shown to yield the maximum likelihood estimator of the number of sources [87]. Implementations of the 2-MUSIC algorithm generally assume that the additive noise is white and $M > Q$.

Taking the EVD of the Covariance matrix $\mathbf{C}^{(2,y)}$, we get:

$$\mathbf{C}^{(2,y)} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^{\mathsf{H}} \tag{3.13}$$

where $\boldsymbol{\Sigma}$ is a real-valued $M \times M$ diagonal matrix and $\mathbf{U} = [\mathbf{U}_s\ \mathbf{U}_n] \in \mathbb{C}^{M \times M}$ is a unitary matrix, with $\mathbf{U}_s \in \mathbb{C}^{M \times Q}$ and $\mathbf{U}_n \in \mathbb{C}^{M \times M-Q}$ corresponding to the signal and noise subspaces, respectively. Taking (3.6) into account and noticing that $\mathbf{U}_n^{\mathsf{T}}\mathbf{A} = \mathbf{0}_{(M-Q)\times Q}$, we can define the following localization function:

$$P_2(\theta, \phi) = \frac{1}{\left\|\mathbf{w}(\theta,\phi)^{\mathsf{H}}\mathbf{U}_n\right\|^2}, \tag{3.14}$$

where $\mathbf{U}_n$ is chosen as the columns of $\mathbf{U}$ associated with the $M - Q$ smallest eigenvalues and the orthogonal projector $\mathbf{w}(\theta, \phi) \in \mathbb{C}^{M \times 1}$ has the form of the steering vector $\mathbf{a}(\theta, \phi)$ defined in

(3.3). The function $P_2(\theta, \phi)$ clearly measures the orthogonality between the signal and noise subspaces for the source q; the desired value of the orthogonal projector is the one maximizing (3.14) [82]. We can therefore conclude that the local maxima of $P_2(\theta, \phi)$ are associated with the DOAs of the Q source signals.

In practice, the algorithm must previously estimate the covariance matrix of the output observations $\hat{\mathbf{C}}^{(2,y)}$ (the equations given in section 1.2 of Chapter 1 can be used to this end) and the localization function is built from the estimated noise subspace $\hat{\mathbf{U}}_n$. This approach provides asymptotically infinite resolution because only the true array steering vectors associated to each source strictly belong to the column space of $\mathbf{A}$. The main drawback is the limitation to treat only overdetermined mixtures, since it can only localize $M - 1$ sources.

HO-MUSIC

As we have seen in section 1.1, the Quadricovariance matrix $\mathbf{C}^{(4,y)}$ has structural properties very similar to those of the Covariance matrix $\mathbf{C}^{(2,y)}$, but involving the Kronecker products of the source steering vectors. According to the VA concept discussed in section 1.2, the Kronecker structure of $\mathbf{C}^{(4,y)}$ yields an increased number of virtual antenna elements, thus allowing for the localization of more sources than sensors, the amount of which varies in function of the array geometry. This is the main principle behind the extension of the MUSIC algorithm to the 4th- (and higher-) orders [82, 47].

Let us consider the EVD of $\mathbf{C}^{(4,y)}$, as follows:

$$\mathbf{C}^{(4,y)} = \begin{bmatrix} \mathbf{U}_s \ \mathbf{U}_n \end{bmatrix} \boldsymbol{\Sigma} \begin{bmatrix} \mathbf{U}_s \ \mathbf{U}_n \end{bmatrix}^{\mathsf{H}} \tag{3.15}$$

where $\boldsymbol{\Sigma}$ is a real-valued $M^2 \times M^2$ diagonal matrix and the columns of $\mathbf{U}_s \in \mathbb{C}^{M^2 \times Q}$ and $\mathbf{U}_n \in \mathbb{C}^{M^2 \times (M^2 - Q)}$ correspond to the 4th-order signal and noise subspaces, obtained from the eigenvectors associated with the Q largest and the $M^2 - Q$ smallest eigenvalues of $\mathbf{C}^{(4,y)}$, respectively. Exploiting the orthogonality between the observation space and the noise subspace, we define the following localization function:

$$P_4(\theta, \phi) = \frac{1}{\left\| \mathbf{w}_2(\theta, \phi)^{\mathsf{H}} \mathbf{U}_n \right\|^2}, \tag{3.16}$$

where $\mathbf{w}_2(\theta, \phi) = \mathbf{a}(\theta, \phi) \otimes \mathbf{a}^*(\theta, \phi)$. The source DOA parameters (θ_q, ϕ_q) can be deduced from the parameters of the orthogonal projectors $\mathbf{w}_2(\theta, \phi) \in \mathbb{C}^{M^2 \times 1}$ that maximize the 4th-order localization function $P_4(\theta, \phi)$, which in practice is built from the estimated noise subspace $\hat{\mathbf{U}}_n$, obtained from the estimated cumulant matrix $\hat{\mathbf{C}}^{(4,y)}$.

The 4-MUSIC algorithm has been introduced in [28, 97, 29] in the context of overdetermined mixtures. The concept has been discussed in [82] as well, including the case of more sources than sensors. The principle underlying the 4-MUSIC algorithm makes use of two basic properties of 4th-order cumulants: the linearity and the additivity under linear independence. Since these

properties hold for cumulants of any (even) order, the algorithmic formalism is easily extendable to higher-order statistics. In fact, using the 2κth-order cumulant matrix $\mathbf{C}^{(2\kappa,y)}$ defined in (3.8), it is straightforward to build a general localization function of the form:

$$P_{2\kappa}(\theta, \phi) = \frac{1}{\left\| \mathbf{w}_\kappa(\theta, \phi)^{\mathsf{H}} \, \mathbf{U}_n \right\|^2}, \tag{3.17}$$

where the orthogonal projector $\mathbf{w}_\kappa(\theta, \phi) \in \mathbb{C}^{M^\kappa \times 1}$ takes the form of (3.10) and $\mathbf{U}_n$ is the $M^\kappa \times (M^\kappa - Q)$ matrix that concatenates the eigenvectors of $\mathbf{C}^{(2\kappa,y)}$ associated with its $M^\kappa - Q$ smallest eigenvalues. Source DOAs can be found by searching for the local maxima of $P_{2\kappa}(\theta, \phi)$. See [27] for a survey on the 2κ-MUSIC algorithms.

2   DF algorithms based on cumulant tensor decomposition

As we have seen in the preceding sections, HO-MUSIC-like algorithms have the ability to make use of the virtual antenna array that naturally arises from the HOS structure at the array output. However, due to the high variance of the HO cumulant estimators, the use of finite sample data yields considerable deviations with respect to asymptotic results. On the other hand, exploiting additional sensors allows for improving resolution and capacity in terms of the number of resolvable sources, at the cost of an increased complexity due to the estimation of higher-order cumulants.

In this section, we propose a high-resolution DF algorithm that creates a 3rd-order virtual array, only exploiting the Kronecker structure of the 4th-order cumulant tensor. Our solution is based on the single-step least-squares (SS-LS) Parafac decomposition technique introduced in Chapter 2, which exploits the symmetry properties of 4th-order output cumulants to perform the Parafac decomposition of a cumulant tensor [63, 64]. This approach involves a channel estimation stage prior to source localization, but it allows for an improved resolution due to an enhanced VA, artificially constructed from the estimated channel, without resorting to 6th-order statistics. While keeping the cumulant estimation variance at a lower level compared with the 2κ-MUSIC algorithms, $\kappa > 2$, the proposed technique is robust to an additive Gaussian noise with unknown spatial correlation, contrary to the 2-MUSIC method. In addition, for ULA arrays, the SS-LS approach is shown to resolve as many sources as the 4-MUSIC algorithm.

Let us rewrite the scalar representation of the 4th-order tensor $\mathcal{C}^{(4,y)}$, defined in section 1.1, as follows:

$$C_{4,y}(i, j, k, l) = \sum_{q=1}^{Q} \gamma_{4,s_q} \, a_i^*(\theta_q, \phi_q) \, a_j(\theta_q, \phi_q) \, a_k^*(\theta_q, \phi_q) \, a_l(\theta_q, \phi_q) \tag{3.18}$$

for $1 \leq i, j, k, l \leq M$ and $q \in [1, Q]$, where the nonzero source Kurtoses γ_{4,s_q} are assumed unknown. It follows from (3.18) that $\mathcal{C}^{(4,y)}$ is a 4th-order Parafac tensor with rank Q. Its canonical components can be straightforwardly deduced and are all written in terms of the

array matrix $\mathbf{A}$ and the diagonal Kurtosis matrix $\mathbf{\Gamma}_{4,s}$ (see section 4 of Chapter 2)[1]. Let us now define the unfolded tensor representation $\mathbf{C}_{[1]} \in \mathbb{C}^{M^3 \times M}$, as follows:

$$\left[\mathbf{C}_{[1]}\right]_{(j-1)M^2+(k-1)M+l,\, i} = C_{4,y}(i,j,k,l), \tag{3.19}$$

which can be easily shown to be written as follows (see the unfolding procedure introduced in section 3.1 of Chapter 1):

$$\mathbf{C}_{[1]} = (\mathbf{A} \diamond \mathbf{A}^* \diamond \mathbf{A})\, \mathbf{\Gamma}_{4,s} \mathbf{A}^{\mathsf{H}} \tag{3.20}$$

$$= \mathbf{A}^{(3)}\, \mathbf{\Gamma}_{4,s}\, \mathbf{A}^{\mathsf{H}} \tag{3.21}$$

where $\mathbf{A}^{(3)}$ is the $M^3 \times Q$ 3rd-order VA matrix, defined as $\mathbf{A}^{(3)} = \mathbf{A} \diamond \mathbf{A}^* \diamond \mathbf{A}$, with $\mathbf{A}$ being defined in (3.5).

2.1 The SS-LS Parafac algorithm

Using the unfolded tensor representation $\mathbf{C}_{[1]}$, the SS-LS Parafac algorithm introduced in Chapter 2 can be used to estimate the array matrix $\mathbf{A}$, as well as the VA matrix $\mathbf{A}^{(3)}$. The algorithm iteratively minimizes a single LS cost function, given by:

$$\psi(\hat{\mathbf{A}}_{r-1}, \mathbf{A}) \triangleq \left\| \mathbf{C}_{[1]} - \left(\hat{\mathbf{A}}_{r-1} \diamond \hat{\mathbf{A}}_{r-1}^* \diamond \hat{\mathbf{A}}_{r-1} \right) \mathbf{\Gamma}_{4,s}\, \mathbf{A}^{\mathsf{H}} \right\|_F^2, \tag{3.22}$$

in which r is the iteration number and $\|\cdot\|_F$ denotes the Frobenius norm. The iterative minimization of $\psi(\hat{\mathbf{A}}_{r-1}, \mathbf{A})$ yields the following LS solution:

$$\begin{aligned} \hat{\mathbf{A}}_r^{\mathsf{H}} &\triangleq \arg\min_{\mathbf{A}} \psi(\hat{\mathbf{A}}_{r-1}, \mathbf{A}) \\ &= \mathbf{\Gamma}_{4,s}^{-1}\, \hat{\mathbf{A}}_{r-1}^{(3)\#}\, \mathbf{C}_{[1]}, \end{aligned} \tag{3.23}$$

with

$$\hat{\mathbf{A}}_{r-1}^{(3)} = \hat{\mathbf{A}}_{r-1} \diamond \hat{\mathbf{A}}_{r-1}^* \diamond \hat{\mathbf{A}}_{r-1}. \tag{3.24}$$

Note that we only have to initialize $\hat{\mathbf{A}}_0$. In fact, at each iteration $r \geq 1$, we deduce $\hat{\mathbf{A}}_{r-1}^{(3)}$ from (3.24) and then, we compute $\hat{\mathbf{A}}_r$ from (3.23).

Iterative LS algorithms are known to be very sensitive to the initialization of the parameters [58]. Exploiting the unit-modulus property of the array steering matrix, the following modification of the algorithmic procedure is expected to improve convergence. After initializing $\hat{\mathbf{A}}_0$ with an $M \times Q$ matrix drawn from a (complex) Gaussian distribution, perform the following:

At each iteration $r \geq 1$, before computing $\hat{\mathbf{A}}_r$, divide each entry of the preceding estimate by its own magnitude, i.e.

$$[\hat{\mathbf{A}}_{r-1}]_{mq} \leftarrow \frac{[\hat{\mathbf{A}}_{r-1}]_{mq}}{\left|[\hat{\mathbf{A}}_{r-1}]_{mq}\right|}\ ;$$

---

[1] This result can also be obtained from the general formulations (1.52) and (1.53), introduced in Chapter 1, with $P = 4$.

Normalize each column by its first-row element:

$$[\hat{\mathbf{A}}_{r-1}]_{\cdot q} \leftarrow \frac{[\hat{\mathbf{A}}_{r-1}]_{\cdot q}}{[\hat{\mathbf{A}}_{r-1}]_{1\,q}} \; ;$$

Deduce $\hat{\mathbf{A}}_{r-1}^{(3)}$ from (3.24) and compute the array matrix estimate at iteration r as follows:

$$\hat{\mathbf{A}}_r \leftarrow \left[\hat{\mathbf{A}}_{r-1}^{(3)\#} \, \mathbf{C}_{[1]} \right]^{\mathsf{H}} . \tag{3.25}$$

Notice that, due to the normalization step, the computation of $\hat{\mathbf{A}}_r$ becomes independent of the source Kurtosis matrix $\mathbf{\Gamma}_{4,s}$. The algorithm is stopped when $|e(r) - e(r-1)|^2 \leq \varepsilon$, where $e(r) = \|\hat{\mathbf{A}}_r - \hat{\mathbf{A}}_{r-1}\|_F / \|\hat{\mathbf{A}}_r\|_F$ and ε is an arbitrary small positive constant.

2.2   Uniqueness and identifiability

Due to the Vandermonde structure of the array matrix, given in (3.5), and assuming the signal sources are not closely located, matrix $\mathbf{A}$ can be shown to be full k-rank [125], so that $k_A = r_A = \min(M, Q)$. In this case, the Kruskal Theorem (1.56) yields $Q \leq (4M-3)/2$, for $M < Q$, leading to the following general sufficient conditions for the uniqueness of the Parafac decomposition of tensor $\mathcal{C}^{(4,y)}$:

$$2 \leq Q \leq 2M - 2. \tag{3.26}$$

Although (3.26) is not a necessary condition, it establishes an upper bound on the number of guaranteed resolvable sources. This bound limits the number of sources that we can treat using the 3rd-order VA matrix $\mathbf{A}^{(3)}$, regardless of the number of virtual sensors.

In the case of a ULA array with M sensors, the number of different virtual sensors associated with the κth-order VA is shown to be equal to $\kappa(M-1)+1$ [26]. In this context, the 3rd-order VA matrix $\mathbf{A}^{(3)}$ admits a maximum capacity of $3M - 3$ sources. However, since the SS-LS approach is bounded by the uniqueness condition (3.26), we can never charge the VA with more than $2M - 2$ sources, so that its noise subspace has at least M free dimensions (i.e. linearly independent basis vectors). Moreover, when using an M-element ULA array, the capacity of the 4-MUSIC algorithm is associated with the number of VS sensors of a 2nd-order VA, which coincides with the upper bound of the SS-LS approach. However, if 4-MUSIC operates with maximal capacity, the noise subspace of the 2nd-order VA has only one free dimension.

DOA estimation

The source DOAs can be recovered from the VA matrix $\mathbf{A}^{(3)}$ by using a 6th-order MUSIC-like localization function $P_6(\theta, \phi)$, such as defined in (3.17) with $\kappa = 3$, i.e.

$$P_6(\theta, \phi) = \frac{1}{\left\| \mathbf{w}_3(\theta, \phi)^{\mathsf{H}} \, \mathbf{U}_n \right\|^2}, \tag{3.27}$$

where $\mathbf{w}_3(\theta,\phi) = \mathbf{a}(\theta,\phi) \otimes \mathbf{a}^*(\theta,\phi) \otimes \mathbf{a}(\theta,\phi)$, with $\mathbf{a}(\theta,\phi)$ defined in (3.3), and $\mathbf{U}_n$ is a $M^3 \times (M^3 - Q)$ matrix representing the noise subspace and formed of the left singular vectors of $\mathbf{A}^{(3)}$ associated with its $M^3 - Q$ smallest singular values.

The source DOA parameters (θ_q, ϕ_q) can be recovered from the parameters of the orthogonal projectors $\mathbf{w}_3(\theta,\phi) \in \mathbb{C}^{M^3 \times 1}$ leading to the local maxima of the 6th-order localization function $P_6(\theta,\phi)$, defined in (3.27).

3   Multipath MIMO channel modeling

Let us consider a multiple-input multiple-output (MIMO) wireless communication system with Q transmit and M receive sensors disposed both in ULA arrays, with sensors spaced of half wavelength. The transmit antenna array is assumed to be placed at the far-field of the receive array and the transmission is subject to specular multipath propagation due to remote scatterers, as illustrated in fig. 3.3, so that the channel between each transmit and receive antenna can be represented by a superposition of K plane waves, associated with different scatterers, located far apart each other. The location of each scatterer determines the angles (ϑ_k, φ_k) and (θ_k, ϕ_k) defining the directions of departure (DOD) and arrival of the kth path with respect to the transmit and receive arrays, respectively. For notational simplicity, we assume that the antenna arrays and the scatterers are approximately coplanar so that the elevation angles φ_k and ϕ_k approach zero and will therefore be omitted in the sequel. The transmitted signals are assumed to be narrowband with respect to the array aperture so that the signals over the kth path are subject to the a single attenuation factor β_k, and achieve the same relative propagation delay, equal to τ_k. In order to capture all the incoming delayed signals, we utilize a known pulse shape filter $g(\ell)$, with finite temporal support equal to $L+1$ symbol periods, so that $g(\ell) \neq 0$, for $\ell \notin [0, L]$. We also assume that the multipath delay spread profile is finite with a known maximum path delay τ_{max} that is larger than the inverse of the coherence bandwidth so that the channel can be viewed as a frequency-selective MIMO model. Finally, the channel is stationary over the length of the observation interval.

The received signal at the output of the array element m, sampled at the symbol rate, can be written as follows:

$$y_m(n) = \sum_{q=1}^{Q}\sum_{k=1}^{K} \beta_k\, a_q'(\vartheta_k)\, a_m(\theta_k) \sum_{\ell=0}^{L} g(\ell - \tau_k)\, s_q(n-\ell) + v_m(n) \qquad (3.28)$$

$$= \sum_{q=1}^{Q}\sum_{\ell=0}^{L} h_{mq}(\ell)\, s_q(n-\ell) + v_m(n), \quad m \in [1, M], \qquad (3.29)$$

where the channel coefficient $h_{mq}(\ell)$ has been defined as follows:

$$h_{mq}(\ell) \triangleq \sum_{k=1}^{K} \beta_k\, a_q'(\vartheta_k)\, a_m(\theta_k)\, g(\ell - \tau_k), \qquad \forall\, \ell \in [0, L], \qquad (3.30)$$
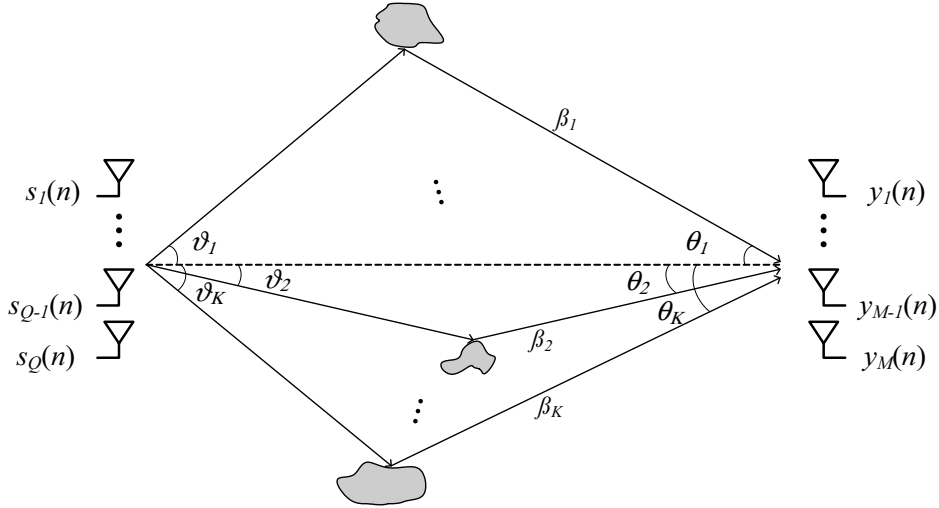
Figure 3.3: Multipath propagation scenario.

and $h_{mq}(\ell) = 0$ elsewhere, with β_k being the complex fading gain associated with the kth path and $a'_q(\vartheta_k)$ and $a_m(\theta_k)$ denoting the spatial responses of the transmit and receive array elements q and m, respectively, with respect to the kth path, with DOD equal to ϑ_k and DOA given by the angle θ_k. The non-measurable complex-valued discrete input signals $s_q(n)$ are stationary, ergodic and mutually independent with symmetric distribution, zero-mean and non-zero kurtosis $\gamma_{4,s_q} = c_{4,s_q}(0,0,0)$. The additive noise $v_m(n)$ at the receive array output is assumed Gaussian and independent from the input signals, with zero-mean and unknown spatial correlation.

Let us write the channel coefficients $h_{mq}(\ell)$ in a vector form, so that $\mathbf{h}^{(m,q)}(\ell) = [h_{mq}(\ell), \ldots, h_{mq}(\ell + L)]^{\mathsf{T}}$. By stacking the row-vectors $\mathbf{h}^{(m,q)}(\ell)^{\mathsf{T}}$ for all $m \in [1, M]$, we can build the following matrices:

$$\mathbf{H}^{(q)}(\ell) = \left[\mathbf{h}^{(1,q)}(\ell) \ \ldots \ \mathbf{h}^{(M,q)}(\ell)\right]^{\mathsf{T}} \in \mathbb{C}^{M \times (L+1)} \tag{3.31}$$

$$= \begin{pmatrix} h_{1q}(\ell) & h_{1q}(\ell+1) & \ldots & h_{1q}(\ell+L) \\ \vdots & \vdots & \ddots & \vdots \\ h_{Mq}(\ell) & h_{Mq}(\ell+1) & \ldots & h_{Mq}(\ell+L) \end{pmatrix}$$

Note that for $\ell = 0$, matrix $\mathbf{H}^{(q)}(0)$ contains the impulse responses of the channels linking the transmit antenna q with each receive antenna $m \in [1, M]$. Hence, by concatenating $\mathbf{H}^{(q)}(0)$ for all $q \in [1, Q]$, we can define a channel matrix $\mathbf{H} \in \mathbb{C}^{M \times Q(L+1)}$, that can be written as

$$\mathbf{H} \triangleq \left[\mathbf{H}^{(1)}(0) \ \ldots \ \mathbf{H}^{(Q)}(0)\right], \tag{3.32}$$

and contains all the channel impulse responses characterizing the $M \times Q$ MIMO system. Finally, by stacking temporally shifted versions of $\mathbf{H}^{(q)}(0)$, we can define a matrix $\mathbf{H}^{(q)} \in \mathbb{C}^{M(2L+1) \times (L+1)}$

for each transmit antenna $q \in [1, Q]$, as follows:

$$\mathbf{H}^{(q)} \triangleq \begin{pmatrix} \mathbf{H}^{(q)}(-L) \\ \vdots \\ \mathbf{H}^{(q)}(0) \\ \vdots \\ \mathbf{H}^{(q)}(L) \end{pmatrix}, \tag{3.33}$$

where each block-row of size $M \times (L+1)$ can be written as $\mathbf{H}^{(q)}(\ell+\Delta\ell)$, with $\ell = 0$, $\Delta\ell \in [-L, L]$. Recalling that $h_{mq}(\ell) = 0$, $\forall \ell \notin [0, L]$, we notice from (3.33) and (3.31) that matrix $\mathbf{H}^{(q)}$ has a *block*-Hankel structure, since

$$\mathbf{H}^{(q)}(\Delta\ell) = \begin{pmatrix} 0 & \cdots & 0 & h_{1q}(0) & \cdots & h_{1q}(L+\Delta\ell) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & h_{Mq}(0) & \cdots & h_{Mq}(L+\Delta\ell) \end{pmatrix}, \qquad \text{for } \Delta\ell \leq 0 \quad \text{and} \tag{3.34}$$

$$\mathbf{H}^{(q)}(\Delta\ell) = \begin{pmatrix} h_{1q}(\Delta\ell) & \cdots & h_{1q}(L) & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ h_{Mq}(\Delta\ell) & \cdots & h_{Mq}(L) & 0 & \cdots & 0 \end{pmatrix}, \qquad \text{for } \Delta\ell \geq 0. \tag{3.35}$$

By concatenating $\mathbf{H}^{(q)}$, for all $q \in [1, Q]$, we define the matrix $\mathcal{H} \in \mathbb{C}^{M(2L+1) \times Q(L+1)}$, as follows:

$$\mathcal{H} \triangleq \left[\mathbf{H}^{(1)} \ldots \mathbf{H}^{(Q)} \right]^{\mathsf{T}}, \tag{3.36}$$

which, due to (3.34) and (3.35), has the following structure:

$$\mathcal{H} = \begin{pmatrix} \mathbf{H}^{(1)}(-L) & \mathbf{H}^{(2)}(-L) & \cdots & \mathbf{H}^{(Q)}(-L) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}^{(1)}(0) & \mathbf{H}^{(2)}(0) & \cdots & \mathbf{H}^{(Q)}(0) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}^{(1)}(L) & \mathbf{H}^{(2)}(L) & \cdots & \mathbf{H}^{(Q)}(L) \end{pmatrix} = \begin{pmatrix} \square & & \square & \cdots & & \square \\ \vdots & & \vdots & \ddots & & \vdots \\ \square & & \square & \cdots & & \square \\ \square & & \square & \cdots & & \square \\ \square & & \square & \cdots & & \square \\ \vdots & & \vdots & \ddots & & \vdots \\ \square & & \square & \cdots & \square \end{pmatrix}.$$

Notice that the $(L+1)$th block-row of size $M \times Q(L+1)$ of matrix $\mathcal{H}$ corresponds to matrix $\mathbf{H}$ defined in (3.32). Thus, it is possible to deduce $\mathcal{H}$ from the block-columns $\mathbf{H}^{(q)}(0)$ of matrix $\mathbf{H}$. To this end, we first build $\mathbf{H}^{(q)}$ by stacking the time-shifted versions $\mathbf{H}^{(q)}(\Delta\ell)$, for all $\Delta\ell \in [-L, L]$, as suggested by (3.33). Then, we obtain $\mathcal{H}$ by concatenating the resulting matrices $\mathbf{H}^{(q)}$ for all $q \in [1, Q]$, as indicated in (3.36).

3.1 A Space-Time 4th-order cumulant tensor model

The *space-time* (ST) 4th-order output cumulants are defined as follows[2]:

$$c_{m_1m_2m_3m_4}^{(4,y)}(l_1, l_2, l_3) \triangleq cum\left[y_{m_1}^*(n), y_{m_2}(n + l_1), y_{m_3}^*(n + l_2), y_{m_4}(n + l_3)\right], \tag{3.37}$$

and, due to the linearity and additivity properties of cumulants, considering the assumption of Gaussian noise, we can write:

$$c_{m_1m_2m_3m_4}^{(4,y)}(l_1, l_2, l_3) = \sum_{q=1}^{Q}\gamma_{4,s_q}\sum_{\ell=0}^{L} h_{m_1q}^*(\ell)\, h_{m_2q}(\ell + l_1)\, h_{m_3q}^*(\ell + l_2)\, h_{m_4q}(\ell + l_3), \tag{3.38}$$

where $\gamma_{4,s_q} = cum\left[s_q^*(n), s_q(n), s_q^*(n), s_q(n)\right]$, $m_1, m_2, m_3, m_4 \in [1, M]$ and $|l_1|, |l_2|, |l_3| \le L$.

Let us denote $i_{p+1} = (u_p - 1)M + m_{p+1}$, with $u_p = l_p + L + 1$, for all $p \in [1, 3]$, and define $C_{i_1i_2i_3i_4}^{(4,y)} \triangleq c_{m_1m_2m_3m_4}^{(4,y)}(l_1, l_2, l_3)$, where $i_1 = m_1$. Then, using $f = (q-1)(L+1) + \ell + 1$, equation (3.38) yields:

$$C_{i_1i_2i_3i_4}^{(4,y)} = \sum_{f=1}^{F} a_{i_1f}\, b_{i_2f}\, c_{i_3f}\, d_{i_4f}, \tag{3.39}$$

where $F = Q(L + 1)$ and we have performed the following substitutions:

$$\begin{cases} a_{i_1f} = \gamma_{4,s_q}\, h_{m_1q}^*(\ell), & m_1 = i_1 \in [1, M], \\ b_{i_2f} = h_{m_2q}(\ell + l_1), & m_2 \in [1, M], \quad l_1 \in [-L, L], \quad i_2 \in [1, M(2L+1)], \\ c_{i_3f} = h_{m_3q}^*(\ell + l_2), & m_3 \in [1, M], \quad l_2 \in [-L, L], \quad i_3 \in [1, M(2L+1)], \\ d_{i_4f} = h_{m_4q}(\ell + l_3), & m_4 \in [1, M], \quad l_3 \in [-L, L], \quad i_4 \in [1, M(2L+1)]. \end{cases} \tag{3.40}$$

From (3.39), we note that $C_{i_1i_2i_3i_4}^{(4,y)}$ can be viewed as the scalar representation of a 4th-order tensor $\mathcal{C}^{(4,y)}$ with rank equal to $Q(L+1)$, which admits a Parafac decomposition with components given by the matrices $\mathbf{A} \in \mathbb{C}^{M \times Q(L+1)}$ and $\mathbf{B}, \mathbf{C}, \mathbf{D} \in \mathbb{C}^{M(2L+1) \times Q(L+1)}$, of which the elements are given in (3.40). Tensor $\mathcal{C}^{(4,y)}$ has one dimension equal to M and three other equal dimensions of size $M(2L + 1)$. Using the canonical basis vector notation, matrix $\mathbf{A}$ can be written as follows:

$$\begin{aligned} \mathbf{A} &= \sum_{f=1}^{F}\sum_{i_1=1}^{M} a_{i_1f}\, \mathbf{e}_{i_1}^{(M)}\, \mathbf{e}_f^{(F)\mathsf{T}} \\ \mathbf{A} &= \sum_{q=1}^{Q}\sum_{\ell=0}^{L}\sum_{m_1=1}^{M} \gamma_{4,s_q}\, h_{m_1q}^*(\ell)\, \mathbf{e}_{m_1}^{(M)}\left(\mathbf{e}_q^{(Q)} \diamond \mathbf{e}_{\ell+1}^{(L+1)}\right)^{\mathsf{T}}, \end{aligned} \tag{3.41}$$

where $\mathbf{e}_i^{(I)}$ denotes the ith canonical basis vector of $\mathbb{R}^I$ and we have used the fact that $\mathbf{e}_i^{(I)} \diamond \mathbf{e}_j^{(J)} = \mathbf{e}_{(i-1)J+j}^{(IJ)}$. From the above equation, we conclude that

$$\mathbf{A} = \mathbf{H}^* \boldsymbol{\Gamma}_{4,s} \in \mathbb{C}^{M \times Q(L+1)}, \tag{3.42}$$

---

[2] Notice that, due to the assumption of stationarity of the input signals, all the output statistics only depend on the differences between the considered time-shifts. As a result, only three temporal indices are needed in definition (3.37).

where $\mathbf{H}$ has been defined in (3.32) and $\mathbf{\Gamma}_{4,s} = Diag\left(\mathbf{\Gamma}_{4,s_1}, \ldots, \mathbf{\Gamma}_{4,s_Q}\right)$ is a $Q(L+1) \times Q(L+1)$ diagonal matrix, with $\mathbf{\Gamma}_{4,s_q} = \gamma_{4,s_q} \mathbf{I}_{L+1}$. Similarly, matrix $\mathbf{B}$ can be represented as follows:

$$\mathbf{B} = \sum_{f=1}^{F} \sum_{i_2=1}^{M(2L+1)} b_{i_2 f}\, \mathbf{e}_{i_2}^{(M)}\, \mathbf{e}_f^{(F)\mathsf{T}}$$

$$\mathbf{B} = \sum_{q=1}^{Q} \sum_{\ell=0}^{L} \sum_{l_1=-L}^{L} \sum_{m_2=1}^{M} h_{m_2 q}(\ell + l_1) \left(\mathbf{e}_{l_1+L+1}^{(2L+1)} \diamond \mathbf{e}_{m_2}^{(M)}\right) \left(\mathbf{e}_q^{(Q)} \diamond \mathbf{e}_{\ell+1}^{(L+1)}\right)^{\mathsf{T}}, \qquad (3.43)$$

and we note that:

$$\mathbf{B} = \mathcal{H} \in \mathbb{C}^{M(2L+1) \times Q(L+1)}, \qquad (3.44)$$

where $\mathcal{H}$ is the block-Hankel matrix defined in (3.36). From (3.40), it is now straightforward to deduce that

$$\mathbf{C} = \mathcal{H}^* \in \mathbb{C}^{M(2L+1) \times Q(L+1)}, \qquad (3.45)$$

$$\mathbf{D} = \mathcal{H} \in \mathbb{C}^{M(2L+1) \times Q(L+1)}. \qquad (3.46)$$

Comments on the space-time cumulant tensor formulation

It is interesting to compare the above described cumulant tensor $\mathcal{C}^{(4,y)}$, with the formulations introduced in Chapter 2 for the convolutive SISO and instantaneous MIMO cases (sections 2 and 4, respectively). Notice that when $M = Q = 1$, with $L \geq 1$, the signal model (3.29) corresponds to the output of a purely temporal SISO channel represented by a single FIR filter. In this case, using (3.36), $\mathcal{H}$ reduces to the Hankel channel matrix given in (2.9), while from (3.32), we note that $\mathbf{H}$ becomes the channel coefficient vector $\mathbf{h}^{\mathsf{T}}$, defined in (2.8). In the convolutive SISO case, the 4th-order cumulant definition given in (3.37) is equivalent to (2.2) and tensor $\mathcal{C}^{(4,y)}$ becomes the 3rd-order tensor defined in (2.5).

On the other hand, considering the memoryless case ($L = 0$) with $Q > 1$ and $M > 1$, the signal model (3.29) can be viewed as the output of an instantaneous MIMO channel. Indeed, using (3.36) and (3.32), we note that the expression for $\mathcal{H}$ coincides with $\mathbf{H}$ in this case, and both are equivalent to the channel coefficient matrix used in (2.30). Under these conditions, the 4th-order cumulants defined in (3.37) take the form of (2.31) and tensor $\mathcal{C}^{(4,y)}$ is equivalent to the purely spatial cumulant tensor defined in (2.33), with four identical dimensions of size M.

In conclusion, the space-time cumulant tensor $\mathcal{C}^{(4,y)}$ can be viewed as a generalized cumulant tensor model that includes the convolutive SISO ($Q = 1$, $M = 1$, $L \geq 1$) and the instantaneous MIMO ($Q > 1$, $M > 1$, $L = 0$) as particular cases.

Uniqueness conditions

A sufficient uniqueness condition for the Parafac decomposition of the cumulant tensor $\mathcal{C}^{(4,y)}$ can be derived from the Kruskal Theorem, introduced in section 3.1 of Chapter 1. Due

to its block-Hankel structure, $\mathcal{H}$ is ensured to be full k-rank, and thus $k_{\mathcal{H}} = r_{\mathcal{H}} = \min(M(2L+1), Q(L+1))$. Assuming that the specular reflectors are located far apart each other, and remote to both the transmit and receive arrays, we can consider that the incoming signals are spatially distinguishable. Further, assuming that there is at least one path delay τ_k that is not a multiple of the symbol period[3], we can ensure that $h_{mq}(\ell) \neq 0$, $\forall \ell \in [0, L]$. Under these conditions, matrix $\mathbf{H}$ is also guaranteed to be full k-rank, and hence $k_{\mathbf{H}} = r_{\mathbf{H}} = \min(M, Q(L+1))$. Under these assumptions, the Kruskal uniqueness theorem yields:

$$3k_{\mathcal{H}} + k_{\mathbf{H}} \geq 2F + 3, \tag{3.47}$$

where $F = Q(L+1)$, and it follows that:

$$3 \min\left(M(2L+1), Q(L+1)\right) + \min\left(M, Q(L+1)\right) \geq 2Q(L+1) + 3. \tag{3.48}$$

Assuming $L \geq 1$, the following cases can be considered:

1. $M \geq Q(L+1)$, which implies $M(2L+1) \geq Q(L+1)$.

   In this case, (3.48) becomes $2Q(L+1) \geq 3$, which is satisfied for all $Q \geq 1$ and $L \geq 1$, so that uniqueness is ensured for all

$$1 \leq Q \leq \frac{M}{L+1}, \qquad L \geq 1; \tag{3.49}$$

2. $M < Q(L+1)$ and $M(2L+1) \geq Q(L+1)$.

   In this case, (3.48) yields $Q(L+1) + M \geq 3$, which is always satisfied with $Q \geq 1$, $M \geq 1$ and $L \geq 1$, i.e. the uniqueness condition is guaranteed when:

$$\frac{M}{L+1} < Q \leq \frac{M(2L+1)}{L+1}, \quad M \geq 1, \quad L \geq 1; \tag{3.50}$$

3. $M(2L+1) < Q(L+1)$, which implies $M < Q(L+1)$.

   In this case, (3.48) gives $3M(2L+1) + M \geq 2Q(L+1) + 3$, which is satisfied when

$$\frac{M(2L+1)}{L+1} < Q \leq \frac{2M(3L+2) - 3}{2(L+1)}, \tag{3.51}$$

Putting together equations (3.49) to (3.51), it follows that the uniqueness of the Parafac decomposition of $\mathcal{C}^{(4,y)}$ is guaranteed under the following general sufficient condition:

$$Q \leq \frac{2M(3L+2) - 3}{2(L+1)} \quad M \geq 1, \quad L \geq 1. \tag{3.52}$$

When $L = 0$, we have a memoryless MIMO channel and the cases **1** and **2** correspond to $Q \leq M$ and $Q = M$, respectively. In such cases, (3.48) is satisfied provided that $2 \leq Q \leq M$. Case **3**

---

[3] This assumption is due to the frequency-selective nature of the channel. In the case of a flat-fading channel, this assumption is not necessary and we should use $L = 0$.

corresponds to a strictly underdetermined mixture ($Q > M$), and yields the following uniqueness condition:

$$2 \leq Q \leq \frac{4M - 3}{2}, \quad L = 0, \tag{3.53}$$

which coincides, as it should be expected, with the uniqueness condition stated in (2.36) for instantaneous MIMO mixtures in the underdetermined case;

The uniqueness conditions for the Parafac decomposition of the generalized cumulant tensor $\mathcal{C}^{(4,y)}$ are summarized in Table 3.1 for some values of M and L. Note that the first row ($L = 0$) corresponds to the instantaneous MIMO case, while the first column ($M = 1$) corresponds to the purely temporal case, without spatial diversity. In this latter case, we notice that when $L > 1$, uniqueness is ensured for $Q \leq 2$, which enables us to identify convolutive multiple-input single-output (MISO) channels with up to 2 signal sources.

Table 3.1: Uniqueness conditions for the Parafac decomposition of $\mathcal{C}^{(4,y)}$.

M	1	2	3	4	5	6	7	8
$L = 0$ max Q	-	2	4	6	8	10	12	14
$L = 1$ max Q	1	4	6	9	11	14	16	19
$L = 2$ max Q	2	4	7	10	12	15	18	20
$L = 3$ max Q	2	5	7	10	13	16	18	21
$L = 4$ max Q	2	5	8	10	13	16	19	22

3.2 Parafac modeling of the multipath MIMO channel

The introduction of the propagation channel structure in the signal model (3.28) allows us to model the multipath transmission as a specular channel with multiple planar wavefronts, each one being characterized by an attenuation, a propagation delay and a spatial signature, associated with the angles of departure and arrival. Using such a parametric model, the blind identification problem reduces the estimation of these multipath parameters. This allows us to exploit some prior information about the structure of the wireless channel, which is often available in radiocommunication contexts, such as the knowledge of pulse shape filter and the transmit and receive array manifolds.

Considering the parametric multipath channel model (3.30), let us define:

$$\begin{cases} [\mathbf{A}_{\mathrm{R}}]_{m,k} = a_m(\theta_k), & m \in [1, M] \\ [\mathbf{A}_{\mathrm{T}}]_{q,k} = a'_q(\vartheta_k), & q \in [1, Q] \\ [\mathbf{G}]_{\ell+1, k} = g(\ell - \tau_k), & \ell \in [0, L]. \\ [\mathbf{b}]_k = \beta_k & k \in [1, K], \end{cases} \tag{3.54}$$

and

$$\mathbf{F} = \mathbf{G} Diag\,(\mathbf{b})\ \in \mathbb{C}^{(L+1)\times K}, \tag{3.55}$$

where $\mathbf{A}_\mathrm{R} \in \mathbb{C}^{M\times K}$, $\mathbf{A}_\mathrm{T} \in \mathbb{C}^{Q\times K}$, $\mathbf{G} \in \mathbb{C}^{(L+1)\times K}$ and $\mathbf{b} \in \mathbb{C}^{K\times 1}$. Notice that the channel coefficients $h_{mq}(\ell)$ defined in equation (3.30) can be viewed as the scalar representation of a 3rd-order tensor $\mathcal{H} \in \mathbb{C}^{M\times Q\times(L+1)}$ that admits a Parafac decomposition with rank K and components given by $\mathbf{A}_\mathrm{R}$, $\mathbf{A}_\mathrm{T}$ and $\mathbf{F}$. This model assumes that the channel is stationary over the interval of one time-slot and $\mathcal{H}$ can be viewed as a particularization of the tensor channel model proposed in [126], where a block-fading channel has been considered yielding a Parafac tensor properly formed by combining the signals received during multiple time-slots. In that case, training sequences have been used to separate the signals from the channel information.

Uniqueness conditions

The spatial signature matrices $\mathbf{A}_\mathrm{R}$ and $\mathbf{A}_\mathrm{T}$ have a Vandermonde structure, as shown in (3.5), and can be written as follows:

$$\mathbf{A}_\mathrm{R} = \begin{pmatrix} 1 & \dots & 1 \\ \mathrm{a}_1 & \dots & \mathrm{a}_K \\ \mathrm{a}_1^2 & \dots & \mathrm{a}_K^2 \\ \vdots & \ddots & \vdots \\ \mathrm{a}_1^{M-1} & \dots & \mathrm{a}_K^{M-1} \end{pmatrix};\quad \mathbf{A}_\mathrm{T} = \begin{pmatrix} 1 & \dots & 1 \\ \mathrm{a'}_1 & \dots & \mathrm{a'}_K \\ \mathrm{a'}_1^2 & \dots & \mathrm{a'}_K^2 \\ \vdots & \ddots & \vdots \\ \mathrm{a'}_1^{Q-1} & \dots & \mathrm{a'}_K^{Q-1} \end{pmatrix}, \tag{3.56}$$

where we have defined $\mathrm{a}_k = e^{J\pi\cos\theta_k}$ and $\mathrm{a'}_k = e^{J\pi\cos\vartheta_k}$ from (3.4), considering transmit and receive ULA arrays with sensors spaced of half wavelength and no elevation angle. In this context, assuming that the incoming signals are spatially distinguishable is equivalent to have generating vectors of $\mathbf{A}_\mathrm{R}$ and $\mathbf{A}_\mathrm{T}$ with distinct nonzero elements, i.e. $\mathrm{a}_{k_1} \neq \mathrm{a}_{k_2} \neq 0$ and $\mathrm{a'}_{k_1} \neq \mathrm{a'}_{k_2} \neq 0$, for all $k_1 \neq k_2 \in [1, K]$. Under this condition, it has been shown in [125] that a Vandermonde matrix is full k-rank and, therefore: $k_{\mathbf{A}_\mathrm{R}} = r_{\mathbf{A}_\mathrm{R}} = \min(M, K)$ and $k_{\mathbf{A}_\mathrm{T}} = r_{\mathbf{A}_\mathrm{T}} = \min(Q, K)$.

In addition, the pulse shape filter $g(\ell)$ is known and, due to the frequency-selective nature of the channel, we have assumed that there is at least one path delay τ_k that is not a multiple of the symbol period. This allows us to ensure that $g(\ell - \tau_k) \neq 0$, $\forall\,\ell \in [0, L], k \in [1, K]$. Moreover, the path delays are distinct $\tau_{k_1} \neq \tau_{k_2}$, $\forall\,k_1 \neq k_2$, and a Rayleigh fading is assumed so that the gains β_k are modeled as a complex random variable with independent real and imaginary parts driven from a continuous Gaussian distribution. Under these conditions, matrix $\mathbf{F} = \mathbf{G} Diag\,(\mathbf{b})$ is also ensured to be full k-rank, and hence $k_\mathbf{F} = r_\mathbf{F} = \min(L+1, K)$.

Using the Kruskal Theorem, we derive a sufficient condition for the uniqueness of the Parafac decomposition of the 3rd-order tensor $\mathcal{H}$, as follows:

$$k_{\mathbf{A}_\mathrm{R}} + k_{\mathbf{A}_\mathrm{T}} + k_\mathbf{F} \geq 2K + 2 \tag{3.57}$$

and hence

$$\min(M, K) + \min(Q, K) + \min(L + 1, K) \geq 2K + 2. \tag{3.58}$$

When $M > K$, $Q > K$ and $L + 1 > K$, the above condition yields $K \geq 2$, meaning that in this case at least two delayed signals must be collected at the receive filters. In the SISO case ($M = Q = 1$), the uniqueness condition is not satisfied, as it should be expected, since the channel lacks of information in the spatial domain.

4 Blind identification of multipath MIMO channels

By characterizing the transmission channel in terms of its multipath parameters (attenuations, propagation delays and spatial signatures), the identification of the MIMO channel becomes equivalent to the estimation of these propagation parameters. In this section, we propose a two-stage approach for estimating the multipath MIMO channel. Firstly, in section 4.1, we use a non-parametric model to blindly identify the convolutive channel coefficients. This stage is based on a SS-LS algorithm and can be viewed as a generalization of the blind identification methods introduced in Chapter 2. After that, in section 4.2, we propose to recover the multipath channel parameters by means of an ALS-based algorithm exploiting the specular structure of the channel model.

4.1 A non-parametric Parafac-based SS-LS algorithm

Let us denote $I = M(2L + 1)$. The 4th-order tensor $\mathcal{C}^{(4,y)} \in \mathbb{C}^{M \times I \times I \times I}$, with scalar representation given by (3.39), can be unfolded in the form of matrix $\mathbf{C}_{[1]} \in \mathbb{C}^{I^3 \times M}$, so that

$$\left[\mathbf{C}_{[1]}\right]_{(i_2 - 1)I^2 + (i_3 - 1)I + i_4, \, i_1} = C^{(4,y)}_{i_1 i_2 i_3 i_4}, \tag{3.59}$$

for all $i_1 \in [1, M]$ and $i_2, i_3, i_4 \in [1, I]$. From the above definition, and using (3.39), it is easy to note that $\mathbf{C}_{[1]} = \left(\mathbf{B} \diamond \mathbf{C} \diamond \mathbf{D}\right) \mathbf{A}^{\mathsf{T}}$, and hence, using equations (3.42) to (3.46), we end up with:

$$\mathbf{C}_{[1]} = \left(\mathcal{H} \diamond \mathcal{H}^* \diamond \mathcal{H}\right) \mathbf{\Gamma}_{4,s} \, \mathbf{H}^{\mathsf{H}}. \tag{3.60}$$

From (3.60), we can define the following iterative LS cost function:

$$\psi(\widehat{\mathcal{H}}_{r-1}, \mathbf{A}) \triangleq \left\| \mathbf{C}_{[1]} - \left(\widehat{\mathcal{H}}_{r-1} \diamond \widehat{\mathcal{H}}^*_{r-1} \diamond \widehat{\mathcal{H}}_{r-1}\right) \mathbf{A}^{\mathsf{T}} \right\|^2_F, \tag{3.61}$$

where r is the iteration number and $\mathbf{A} = \mathbf{H}^* \mathbf{\Gamma}_{4,s}$, according to (3.42). Minimizing (3.61) yields:

$$\begin{aligned}
\hat{\mathbf{A}}^{\mathsf{T}}_r &= \arg \min_{\mathbf{H}} \psi(\widehat{\mathcal{H}}_{r-1}, \mathbf{H}) \\
&= \left(\widehat{\mathcal{H}}_{r-1} \diamond \widehat{\mathcal{H}}^*_{r-1} \diamond \widehat{\mathcal{H}}_{r-1}\right)^{\#} \mathbf{C}_{[1]}.
\end{aligned} \tag{3.62}$$

The Kruskal Theorem ensures uniqueness up to column scaling and permutation ambiguities. Therefore, under the conditions stated in section 3.1, the Parafac decomposition of tensor $\mathcal{C}^{(4,y)}$

is subject to these trivial indeterminacies, so that any matrices $\bar{\mathcal{H}}$ and $\bar{\mathbf{H}}$ satisfying (3.60) are such that $\bar{\mathcal{H}} = \mathcal{H}\boldsymbol{\Lambda}_1\boldsymbol{\Pi}$ and $\bar{\mathbf{H}} = \mathbf{H}\boldsymbol{\Lambda}_2\boldsymbol{\Pi}$, where $\boldsymbol{\Lambda}_1$, $\boldsymbol{\Lambda}_2$ are $Q(L+1) \times Q(L+1)$ diagonal matrices and $\boldsymbol{\Pi}$ is a permutation matrix. Exploiting the block-Hankel structure of $\mathcal{H}$, given in (3.36), we can avoid intra-block permutations, thus reducing the permutation ambiguity to block-column permutations. In practice, before computing the iteration $r \geq 1$ using (3.62), matrix $\widehat{\mathcal{H}}_{r-1}$ is built from $\widehat{\mathbf{H}}_{r-1}$, as follows:

1. Deduce $\widehat{\mathbf{H}}^{(q)}_{r-1}(0)$, $q \in [1, Q]$, from the columns of $\hat{\mathbf{A}}^*_{r-1}$, as follows:

$$\widehat{\mathbf{H}}^{(q)}_{r-1}(0) = \left[[\hat{\mathbf{A}}^*_{r-1}]_{\cdot (q-1)(L+1)+1} \quad \cdots \quad [\hat{\mathbf{A}}^*_{r-1}]_{\cdot q(L+1)} \right];$$

2. For each $q \in [1, Q]$, build the matrices $\widehat{\mathbf{H}}^{(q)}_{r-1}(\Delta\ell)$, for all $\Delta\ell \in [-L, L]$, by shifting the columns of $\widehat{\mathbf{H}}^{(q)}_{r-1}(0)$, as indicated in (3.34) and (3.35);

3. From (3.33), build $\widehat{\mathbf{H}}^{(q)}_{r-1}$, $q \in [1, Q]$, by stacking $\widehat{\mathbf{H}}^{(q)}_{r-1}(\Delta\ell)$, for all $\Delta\ell \in [-L, L]$;

4. Obtain $\widehat{\mathcal{H}}_{r-1}$ by concatenating $\widehat{\mathbf{H}}^{(q)}_{r-1}$ for all $q \in [1, Q]$, as indicated in (3.36).

The algorithm is initialized with a $M \times Q(L+1)$ Gaussian random matrix $\widehat{\mathbf{H}}_0$. The iterations are stopped when $|e(r) - e(r-1)|^2 \leq \varepsilon$, where $e(r) = \|\widehat{\mathbf{H}}_r - \widehat{\mathbf{H}}_{r-1}\|^2_F / \|\widehat{\mathbf{H}}_r\|^2_F$ and ε is an arbitrary small positive constant.

4.2 Parametric estimation of multipath MIMO channels

In this section, we propose an ALS-based algorithm to jointly estimate the multipath propagation delays and the angles of departure and arrival. This approach is based on the Parafac modeling of the multipath MIMO channel presented in section 3.2, and assumes a prior estimation of the channel coefficients by means of a blind technique such as the SS-LS algorithm introduced in the previous section.

Let us consider the 3rd-order tensor $\mathcal{H} \in \mathbb{C}^{Q \times M \times (L+1)}$, with scalar representation given by (3.30) and Parafac components given by the matrices $\mathbf{A}_{\mathrm{T}}$, $\mathbf{A}_{\mathrm{R}}$ and $\mathbf{F}$, defined in (3.55) and (3.56). Taking $q \in [1, Q]$, $m \in [1, M]$ and $\ell \in [0, L]$, we define three unfolded representations of $\mathcal{H}$, as follows:

$$\begin{cases} [\mathbf{H}_{[1]}]_{\ell M+m, q} &= h_{mq}(\ell), \\ [\mathbf{H}_{[2]}]_{(m-1)Q+q, \ell+1} &= h_{mq}(\ell), \\ [\mathbf{H}_{[3]}]_{(q-1)(L+1)+\ell+1, m} &= h_{mq}(\ell), \end{cases} \tag{3.63}$$

where $\mathbf{H}_{[1]} \in \mathbb{C}^{M(L+1) \times Q}$, $\mathbf{H}_{[2]} \in \mathbb{C}^{QM \times (L+1)}$ and $\mathbf{H}_{[3]} \in \mathbb{C}^{Q(L+1) \times M}$. From (3.30), using the

canonical basis vector notation and definitions (3.55) and (3.56), we can write:

$$
\begin{aligned}
\mathbf{H}_{[1]} &= \sum_{q=1}^{Q} \sum_{m=1}^{M} \sum_{\ell=0}^{L} h_{mq}(\ell) \left(\mathbf{e}_{\ell+1}^{(L+1)} \diamond \mathbf{e}_m^{(M)} \right) \mathbf{e}_q^{(Q)\mathsf{T}} \\
&= \sum_{k=1}^{K} \left[\left(\sum_{\ell=0}^{L} \beta_k g(\ell - \tau_k) \mathbf{e}_{\ell+1}^{(L+1)} \right) \diamond \left(\sum_{m=1}^{M} a_m(\theta_k) \mathbf{e}_m^{(M)} \right) \right] \left(\sum_{q=1}^{Q} a_q'(\vartheta_k) \mathbf{e}_q^{(Q)} \right)^{\mathsf{T}},
\end{aligned}
\tag{3.64}
$$

and hence

$$
\mathbf{H}_{[1]} = \left(\mathbf{F} \diamond \mathbf{A}_{\mathrm{R}} \right) \mathbf{A}_{\mathrm{T}}^{\mathsf{T}}.
\tag{3.65}
$$

After similar manipulations with $\mathbf{H}_{[2]}$ and $\mathbf{H}_{[3]}$, we get:

$$
\mathbf{H}_{[2]} = \left(\mathbf{A}_{\mathrm{R}} \diamond \mathbf{A}_{\mathrm{T}} \right) \mathbf{F}^{\mathsf{T}},
\tag{3.66}
$$

$$
\mathbf{H}_{[3]} = \left(\mathbf{A}_{\mathrm{T}} \diamond \mathbf{F} \right) \mathbf{A}_{\mathrm{R}}^{\mathsf{T}}.
\tag{3.67}
$$

Equations (3.65) to (3.67) allow us to write the following iterative cost functions:

$$
\psi_1(\mathbf{A}_{\mathrm{T}}, \hat{\mathbf{F}}^{(r-1)}, \hat{\mathbf{A}}_{\mathrm{R}}^{(r-1)}) = \left\| \mathbf{H}_{[1]} - \left(\hat{\mathbf{F}}^{(r-1)} \diamond \hat{\mathbf{A}}_{\mathrm{R}}^{(r-1)} \right) \mathbf{A}_{\mathrm{T}}^{\mathsf{T}} \right\|_F^2,
\tag{3.68}
$$

$$
\psi_2(\mathbf{F}, \hat{\mathbf{A}}_{\mathrm{R}}^{(r-1)}, \hat{\mathbf{A}}_{\mathrm{T}}^{(r)}) = \left\| \mathbf{H}_{[2]} - \left(\hat{\mathbf{A}}_{\mathrm{R}}^{(r-1)} \diamond \hat{\mathbf{A}}_{\mathrm{T}}^{(r)} \right) \mathbf{F}^{\mathsf{T}} \right\|_F^2,
\tag{3.69}
$$

$$
\psi_3(\mathbf{A}_{\mathrm{R}}, \hat{\mathbf{A}}_{\mathrm{T}}^{(r)}, \hat{\mathbf{F}}^{(r)}) = \left\| \mathbf{H}_{[3]} - \left(\hat{\mathbf{A}}_{\mathrm{T}}^{(r)} \diamond \hat{\mathbf{F}}^{(r)} \right) \mathbf{A}_{\mathrm{R}}^{\mathsf{T}} \right\|_F^2,
\tag{3.70}
$$

where r is the iteration number. The LS solution of these equations is given by

$$
\hat{\mathbf{A}}_{\mathrm{T}}^{(r)\mathsf{T}} = \left(\hat{\mathbf{F}}^{(r-1)} \diamond \mathbf{A}_{\mathrm{R}}^{(r-1)} \right)^{\#} \mathbf{H}_{[1]},
\tag{3.71}
$$

$$
\hat{\mathbf{F}}^{(r)\mathsf{T}} = \left(\hat{\mathbf{A}}_{\mathrm{R}}^{(r-1)} \diamond \hat{\mathbf{A}}_{\mathrm{T}}^{(r)} \right)^{\#} \mathbf{H}_{[2]},
\tag{3.72}
$$

$$
\hat{\mathbf{A}}_{\mathrm{R}}^{(r)\mathsf{T}} = \left(\hat{\mathbf{A}}_{\mathrm{T}}^{(r)} \diamond \hat{\mathbf{F}}^{(r)} \right)^{\#} \mathbf{H}_{[3]},
\tag{3.73}
$$

where $\hat{\mathbf{F}}^{(0)}$ and $\mathbf{A}_{\mathrm{R}}^{(0)}$ can be initialized as Gaussian random matrices or using some previous knowledge about the multipath parameters, if available (e.g. DOAs for $\mathbf{A}_{\mathrm{R}}^{(0)}$). The algorithm is stopped when $|e(r) - e(r-1)|^2 \leq \varepsilon$, where ε is an arbitrary small positive constant and the estimation error $e(r)$ for the iteration $r \geq 1$ is given by $e(r) = e_1(r) + e_2(r) + e_3(r)$, where

$$
e_1(r) = \frac{\left\| \hat{\mathbf{A}}_{\mathrm{T}}^{(r)} - \hat{\mathbf{A}}_{\mathrm{T}}^{(r-1)} \right\|_F^2}{\left\| \hat{\mathbf{A}}_{\mathrm{T}}^{(r)} \right\|_F^2}, \quad e_2(r) = \frac{\left\| \hat{\mathbf{A}}_{\mathrm{R}}^{(r)} - \hat{\mathbf{A}}_{\mathrm{R}}^{(r-1)} \right\|_F^2}{\left\| \hat{\mathbf{A}}_{\mathrm{R}}^{(r)} \right\|_F^2}, \quad e_3(r) = \frac{\left\| \hat{\mathbf{F}}^{(r)} - \hat{\mathbf{F}}^{(r-1)} \right\|_F^2}{\left\| \hat{\mathbf{F}}^{(r-1)} \right\|_F^2}.
\tag{3.74}
$$

Subspace-based algorithms for multipath parameter extraction

We now describe specific techniques for extracting the multipath channel parameters from the estimates $\hat{\mathbf{A}}_{\mathrm{T}} = \hat{\mathbf{A}}_{\mathrm{T}}^{(r)}$, $\hat{\mathbf{A}}_{\mathrm{R}} = \hat{\mathbf{A}}_{\mathrm{R}}^{(r)}$ and $\hat{\mathbf{F}} = \hat{\mathbf{F}}^{(r)}$, obtained after the convergence of the algorithm. Exploiting the known transmit and receive array manifolds as well as the knowledge of the pulse shape filter, we build orthogonal projectors based on the signal subspace structure. In

the sequel, we describe three subspace techniques for the estimation of DOAs, DODs and path delays, based on MUSIC-like localization functions

Let us start with the case $M > K$ and denote by $\hat{\mathbf{U}}_n \in \mathbb{C}^{M \times (M-K)}$ the matrix formed from the left singular vectors of $\hat{\mathbf{A}}_R$ associated with its $M - K$ null singular values. Define $w_m(\theta) = e^{\jmath \pi (m-1) \cos \theta}$ and build a column-vector $\mathbf{w}(\theta) \in \mathbb{C}^{M \times 1}$. For each given value of θ, the orthogonal projector $\mathbf{w}(\theta)$ reproduces the column-wise structure of (3.56). The DOAs θ_k associated with each path $k \in [1, K]$, can be obtained as the arguments yielding the K local minima of the localization function $P_2(\theta)$, defined in (3.14). For the case $M \geq K$, a generalized localization function $P_{2\kappa}(\theta)$ has been defined in (3.17), where the orthogonal projector $\mathbf{w}_\kappa(\theta)$ takes the form of a multiple Kronecker product, as indicated in (3.10) and $\mathbf{U}_n \in \mathbb{C}^{M^\kappa \times (M^\kappa - K)}$ is obtained from the left singular vectors associated with the $M^\kappa - K$ null singular values of the κth-order virtual array $\hat{\mathbf{A}}_R^{(\kappa)}$. Using ULA arrays, this technique allows us for estimating up to $\kappa(M - 1)$ multipath DOAs [26]. Finally, note that the DODs ϑ_k, $k \in [1, K]$ can also be estimated using the above described technique, with Q replacing M and $\hat{\mathbf{A}}_T$ instead of $\hat{\mathbf{A}}_R$.

Similarly, path delays can also be extracted from the estimated matrix $\hat{\mathbf{F}}$ using the orthogonality between signal and noise subspaces. Exploiting the knowledge of the pulse shape filter $g(\ell)$, we can construct orthogonal projectors $\mathbf{w}(\tau) \in \mathbb{C}^{(L+1) \times 1}$, $L+1 > K$, so that $w_\ell(\tau) = g(\tau - \ell + 1)$, $\ell \in [1, L+1]$, for all $\tau \in [0, \tau_{max}]$, where τ_{max} is a known upper bound of the path delays. Taking the SVD of $\hat{\mathbf{F}}$, we build $\hat{\mathbf{U}}_n \in \mathbb{C}^{M \times (M-K)}$ with the left singular vectors associated with the $L + 1 - K$ null singular values and find the path delays τ_k, $k \in [1, K]$, as the arguments yielding the $L + 1$ local minima of $P_2(\tau)$. A straightforward extension of the technique for the case $L + 1 \leq K$ is possible by utilizing an augmented matrix $\hat{\mathbf{F}}^{(\kappa)} = \hat{\mathbf{F}}^{\diamond \kappa}$, and a corresponding projector $\mathbf{w}_\kappa(\tau) = \mathbf{w}(\tau)^{\otimes \kappa}$. For instance, with $\kappa = 2$, matrix $\hat{\mathbf{F}}^{(2)} = \hat{\mathbf{F}} \diamond \hat{\mathbf{F}}$ has $(L + 1)^2$ rows, of which $L(L + 1)/2$ are repeated. Hence, $\hat{\mathbf{F}}^{(2)}$ can be easily shown to have rank equal to $r_{\hat{\mathbf{F}}^{(2)}} = \min(I, K)$, where $I = (L+1)^2 - L(L+1)/2 = (L+1)(L+2)/2$, which is the number of distinct rows, meaning that we can estimate the delays of up to $K \leq I - 1$ paths.

We also notice that, under the Kruskal condition (3.58), the Parafac decomposition of the 3rd-order tensor $\mathcal{H}$ is ensured to be unique up to column scaling and permutations, which means that any matrices $\bar{\mathbf{A}}_T$, $\bar{\mathbf{A}}_R$ and $\bar{\mathbf{F}}$, satisfying equations (3.65) to (3.67), are of the following form:

$$\begin{cases} \bar{\mathbf{A}}_T &= \mathbf{A}_T \mathbf{\Lambda}_T \mathbf{\Pi}, \\ \bar{\mathbf{A}}_R &= \mathbf{A}_R \mathbf{\Lambda}_R \mathbf{\Pi}, \\ \bar{\mathbf{F}} &= \mathbf{F} \mathbf{\Lambda}_F \mathbf{\Pi}, \end{cases} \tag{3.75}$$

where $\mathbf{\Pi}$ is a permutation matrix and $\mathbf{\Lambda}_T, \mathbf{\Lambda}_R, \mathbf{\Lambda}_F$ are complex diagonal scaling matrices satisfying $\mathbf{\Lambda}_T \mathbf{\Lambda}_R \mathbf{\Lambda}_F = \mathbf{I}_K$. Due to their Vandermonde structure shown in (3.56), the spatial array response matrices, $\mathbf{A}_T$ and $\mathbf{A}_R$, have an all-one first row. This property allows us to get rid of the diagonal scaling ambiguities. By taking the first row of $\bar{\mathbf{A}}_T$, we build the diagonal matrix

$\bar{\mathbf{\Lambda}}_{\mathrm{T}} = D_1\left(\bar{\mathbf{A}}_{\mathrm{T}}\right)$ and, using (3.75), we have $\bar{\mathbf{\Lambda}}_{\mathrm{T}} = D_1\left(\mathbf{A}_{\mathrm{T}}\mathbf{\Lambda}_{\mathrm{T}}\mathbf{\Pi}\right) = \mathbf{\Pi}^{\mathsf{T}} D_1\left(\mathbf{A}_{\mathrm{T}}\right)\mathbf{\Lambda}_{\mathrm{T}}\mathbf{\Pi}$, where $D_i\left(\cdot\right)$ denotes the diagonal matrix built from the ith row of the matrix argument. Since $D_1\left(\mathbf{A}_{\mathrm{T}}\right) = \mathbf{I}_K$, we get $\bar{\mathbf{\Lambda}}_{\mathrm{T}} = \mathbf{\Pi}^{\mathsf{T}}\mathbf{\Lambda}_{\mathrm{T}}\mathbf{\Pi}$. Analogous manipulations yield $\bar{\mathbf{\Lambda}}_{\mathrm{R}} = D_1\left(\bar{\mathbf{A}}_{\mathrm{R}}\right) = \mathbf{\Pi}^{\mathsf{T}}\mathbf{\Lambda}_{\mathrm{R}}\mathbf{\Pi}$. It follows that $\bar{\mathbf{\Lambda}}_{\mathrm{T}}\bar{\mathbf{\Lambda}}_{\mathrm{R}} = \mathbf{\Pi}^{\mathsf{T}}\mathbf{\Lambda}_{\mathrm{F}}^{-1}\mathbf{\Pi} = \bar{\mathbf{\Lambda}}_{\mathrm{F}}^{-1}$, which is also a diagonal matrix. Note that the above procedure leaves the column permutation unchecked.

Let us now denote by $\breve{\mathbf{A}}_{\mathrm{R}}$, $\breve{\mathbf{A}}_{\mathrm{T}}$ and $\breve{\mathbf{G}}$ the parameter matrices reconstructed from the estimated DOAs $(\hat{\theta}_k)$, DODs $(\hat{\vartheta}_k)$ and path delays $(\hat{\tau}_k)$, according to (3.54), for all $k \in [1, K]$. Recalling that $\mathbf{F} = \mathbf{G}\,Diag\left(\mathbf{b}\right)$, we are able to estimate the attenuation vector $\widehat{\mathbf{b}}$, as follows:

$$\widehat{\mathbf{b}} = diag\left(\breve{\mathbf{G}}^{\#}\hat{\mathbf{F}}\hat{\mathbf{\Lambda}}_{\mathrm{F}}^{-1}\right), \qquad L + 1 \geq K, \tag{3.76}$$

where the matrix $\hat{\mathbf{F}}$, obtained after the convergence of the ALS-based algorithm, is an estimate of $\bar{\mathbf{F}}$ and $\hat{\mathbf{\Lambda}}_{\mathrm{F}}^{-1}$ is computed as $\hat{\mathbf{\Lambda}}_{\mathrm{F}}^{-1} = \hat{\mathbf{\Lambda}}_{\mathrm{T}}\hat{\mathbf{\Lambda}}_{\mathrm{R}}$, where $\hat{\mathbf{\Lambda}}_{\mathrm{T}} = D_1(\hat{\mathbf{A}}_{\mathrm{T}})$ and $\hat{\mathbf{\Lambda}}_{\mathrm{R}} = D_1(\hat{\mathbf{A}}_{\mathrm{R}})$ represent the estimates of $\bar{\mathbf{\Lambda}}_{\mathrm{T}}$ and $\bar{\mathbf{\Lambda}}_{\mathrm{R}}$, respectively. The operator $diag\left(\cdot\right)$ forms a column-vector from the diagonal elements of the matrix argument.

At last, we remark that the remaining column permutation, although not resolvable, is not relevant in the present context. However, in order to completely characterize the multipath channel, we need to indicate the correspondences linking the parameters associated with a given path $k \in [1, K]$, i.e. given the estimates $\hat{\theta}_{k_1}$, $\hat{\vartheta}_{k_2}$ and $\hat{\tau}_{k_3}$, which values of k_1, k_2 and k_3, are associated with the kth path. In other words, we need to find k_1, k_2 and k_3, for each $k \in [1, K]$, so that $[\breve{\mathbf{A}}_{\mathrm{R}}]_{.k_1}$, $[\breve{\mathbf{A}}_{\mathrm{T}}]_{.k_2}$ and $[\breve{\mathbf{G}}]_{.k_3}$ are scaled versions of $[\hat{\mathbf{A}}_{\mathrm{R}}]_{.k}$, $[\hat{\mathbf{A}}_{\mathrm{T}}]_{.k}$ and $[\hat{\mathbf{F}}]_{.k}$, respectively. In practice, we solve this problem using the normalized inner product, so that

$$k_1 = \arg \min_{u \in [1, K]} \left| 1 - \frac{[\breve{\mathbf{A}}_{\mathrm{R}}]_{.u}^{\mathsf{H}}[\hat{\mathbf{A}}_{\mathrm{R}}]_{.k}}{\|[\breve{\mathbf{A}}_{\mathrm{R}}]_{.u}\| \, \|[\hat{\mathbf{A}}_{\mathrm{R}}]_{.k}\|} \right|, \qquad k_2 = \arg \min_{u \in [1, K]} \left| 1 - \frac{[\breve{\mathbf{A}}_{\mathrm{T}}]_{.u}^{\mathsf{H}}[\hat{\mathbf{A}}_{\mathrm{T}}]_{.k}}{\|[\breve{\mathbf{A}}_{\mathrm{T}}]_{.u}\| \, \|[\hat{\mathbf{A}}_{\mathrm{T}}]_{.k}\|} \right|,$$

and

$$k_3 = \arg \min_{u \in [1, K]} \left| 1 - \frac{[\breve{\mathbf{G}}]_{.u}^{\mathsf{T}}[\hat{\mathbf{F}}]_{.k}}{\|[\breve{\mathbf{G}}]_{.u}\| \, \|[\hat{\mathbf{F}}]_{.k}\|} \right|, \quad \forall \, k \in [1, K].$$

The attenuation vector $\widehat{\mathbf{b}}$ can only be estimated from (3.76) after reordering the columns of $\breve{\mathbf{G}}$, as explained above, in order to keep the same permutation as $\hat{\mathbf{F}}$.

5   Simulation results

In this section, we present computer simulation results aiming to illustrate the use and assess the performance of the techniques discussed throughout this chapter. First, in the context of a flat fading channel and using a ULA receive array, we simulate a radio propagation scenario for the application of direction finding algorithms. The SS-LS approach, proposed in section 2, will be used to estimate the signal DOAs from a 3rd-order virtual array. Performance comparisons will be provided using the 2-, 4- and 6-MUSIC algorithms [86, 127, 27].

After that, we will simulate a multipath radio propagation channel with multiple transmit and receive antennas, both using ULA arrays. In this context, we will be first interested in estimating the coefficients of the convolutive filters representing the connections between each transmit and receive antenna. To this end, the generalized 4th-order cumulant tensor described in section 3.1 will be decomposed by means of the SS-LS algorithm proposed in section 4.1. Then, starting from the estimated channel model, we recover the spatial and temporal signatures of the MIMO channel using the Parafac-based algorithm proposed in section 4.2. Finally, we extract the multipath channel parameters (DOAs, DODs, path delays and attenuations) using the subspace-based methods also described in section 4.2.

Direction finding algorithms for array processing

In which follows, we evaluate the performance of the method proposed in section 2 in terms of the quality of DOA estimation. We will use the root mean-squared error (RMSE) performance criterion, defined for each source q as follows [27]:

$$\text{RMSE}(q) \triangleq \sqrt{\frac{1}{R} \sum_{r=1}^{R} \left| \hat{\theta}_q^{\langle r \rangle} - \theta_q \right|^2}, \qquad q \in [1, Q], \tag{3.77}$$

where R is the number of Monte Carlo simulations and $\hat{\theta}_q^{\langle r \rangle}$ is the estimation of θ_q for the simulation r. The DOA estimates $\hat{\theta}_q^{\langle r \rangle}$, $q \in [1, Q]$, are deduced from the angle arguments of the orthogonal projectors $\mathbf{w}_\kappa(\theta)$ leading to the local maxima of the corresponding localization function $P_{2\kappa}(\theta)$. Local maxima can be obtained by searching the critical points, i.e. where the first derivative is zero, with a negative second derivative. In this context, an estimate $\hat{\theta}_q^{\langle r \rangle}$, $q \in [1, Q]$, is said to be aberrant if $1/P_{2\kappa}(\hat{\theta}_q^{\langle r \rangle})$ is greater than a certain threshold. For the simulations performed in this section, we adopted the value of 0.1 for this threshold, as suggested by [26, 27]. Aberrant estimates can also happen when the algorithm cannot resolve all the sources. In this case, the number of local maxima of the localization function is smaller than Q. In the following results, the probability of having aberrant estimates has been omitted, since only negligible values have been attained.

We first simulated the case of a ULA array with $M = 3$ narrowband sensors spaced of $\lambda/2$, receiving $Q = 4$ sources with azimuth angles given by $\theta_1 = -55°$, $\theta_2 = -25°$, $\theta_3 = 5°$, $\theta_4 = 50°$, and no elevation angle. The array output signals are corrupted by additive Gaussian and spatially white noise. The curves in fig. 3.4 show, for several values of SNR, the RMSE for the worst (left) and the best (right) estimated sources. In order to evaluate the impact of cumulant estimation errors on the tested algorithms, we show in fig. 3.5 the maximal (left) and minimal (right) RMSE as a function of the sample data length, for a fixed SNR value of 15dB. In this case, the SS-LS and the 4-MUSIC algorithms operate with their maximal capacity in terms of the number of sources. By exploiting the larger noise subspace of the 3rd-order virtual array, the SS-LS approach provides better results than the 4-MUSIC algorithm, using the same
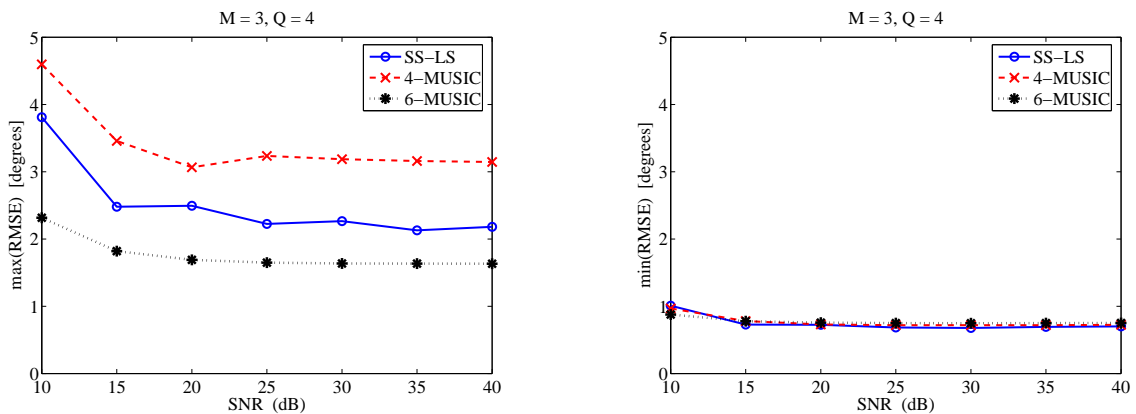
Figure 3.4: Maximal (left) and minimal (right) RMSE as a function of the SNR.
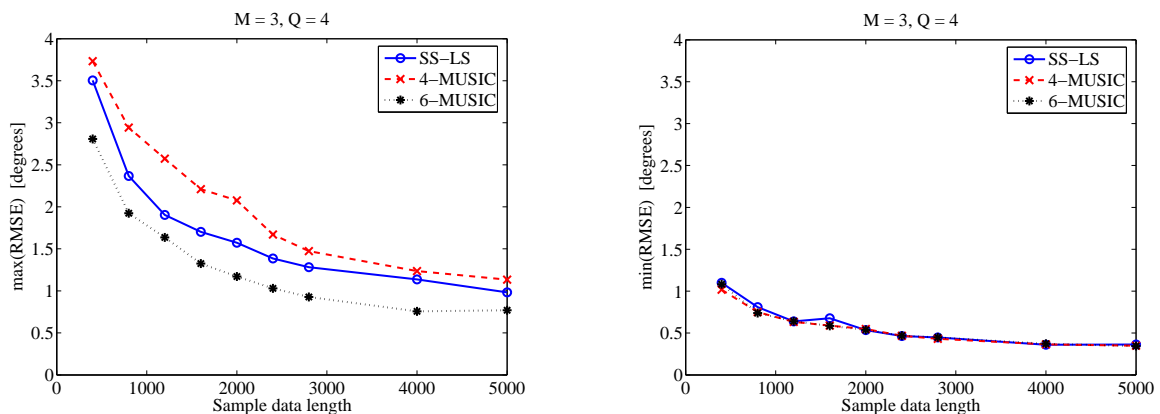

Figure 3.5: Maximal (left) and minimal (right) RMSE as a function of the sample data length.

output statistics. In this scenario, the 6-MUSIC algorithm is not at its identifiability bound and, in the worst case (curves at left), it gives better results than the other techniques, at the cost of having to estimate 6th-order cumulants.

By adding a fourth sensor ($M = 4$) to the antenna array (with $\lambda/2$ spacing), we set up another simulation scenario with $Q = 5$ sources. In this case, the additional source arrives from the direction $\theta_5 = 20°$, with no elevation angle. In fig. 3.6, we show the maximal (left) and minimal (right) RMSE as a function of the SNR, for $N = 1000$. These curves demonstrate that the three algorithms achieve better performance, with very similar results when the VAs do not operate with maximal capacity. In fig. 3.7, the results for the worst (left) and the best (right) estimated sources are given for several values of the sample data length, with a fixed SNR of 15dB. In this case, the 6-MUSIC algorithm does not yield any noticeable advantage.

We have also tested the algorithms in presence of Gaussian noise with unknown spatial correlation. In this case, we used a $\lambda/2$-spaced 3-element ULA array receiving $Q = 2$ sources
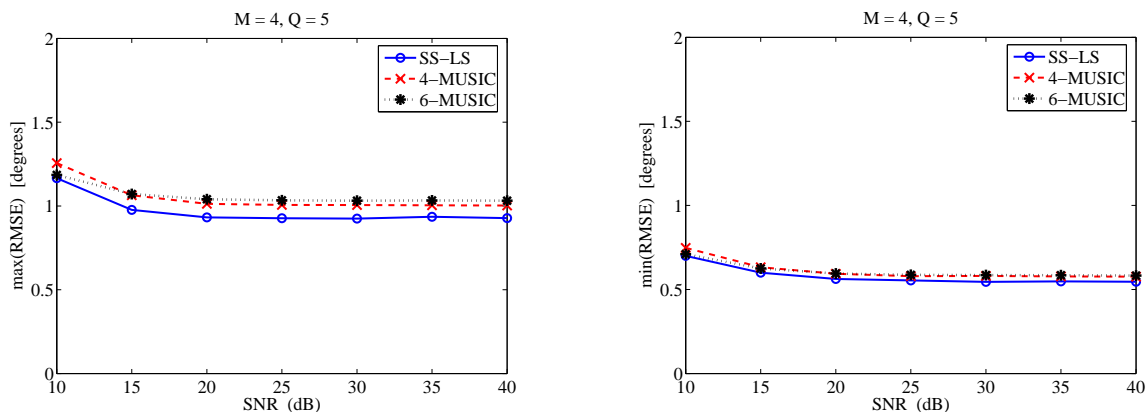
Figure 3.6: Maximal (left) and minimal (right) RMSE as a function of the SNR.
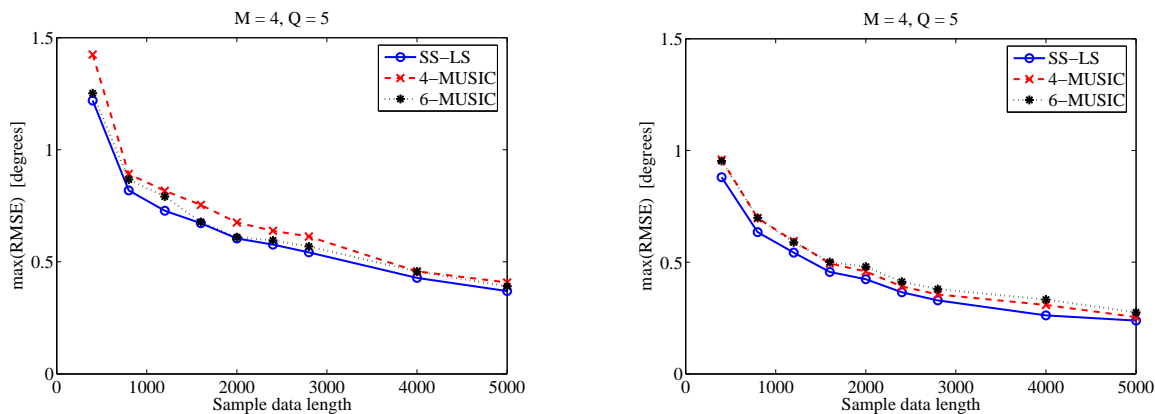
Figure 3.7: Maximal (left) and minimal (right) RMSE as a function of the sample data length.

with DOAs given by $\theta_1 = 5°$ and $\theta_2 = 50°$, respectively. Since this is an overdetermined case, we used the SS-LS approach to estimate the user DOAs from both, the 3rd-order virtual array $\hat{\mathbf{A}}^{(3)}$ ($\kappa = 3$) and the estimated array matrix $\hat{\mathbf{A}}$ ($\kappa = 1$). The additive Gaussian noise has been modeled so that its spatial correlation matrix is given by $[\mathbf{R}_v]_{ij} = \sigma_v^2 \rho^{|i-j|}$, $i, j \in [1, M]$, where σ_v^2 is the noise variance per antenna and ρ is the spatial correlation coefficient of the noise. In fig. 3.8, we compare our results with the 2- and 4-MUSIC algorithms using $N = 1000$ output symbols, with a SNR of 5dB, for different values of the noise spatial correlation. Note that, for $\kappa = 1$ as well as for $\kappa = 3$, the SS-LS approach performed very closely to the 4-MUSIC algorithm, showing good robustness with respect to spatially colored noise, as it should be expected. The 2-MUSIC algorithm, on the other hand, degrades as ρ increases, since the SOS are not able to handle an additive noise with unknown spatial correlation.

Finally, we illustrate the resolution gains in terms of beamwidth provided by the 3rd-order VA used in the SS-LS approach with respect to the 2nd-order VA used by the 4-MUSIC algorithm.
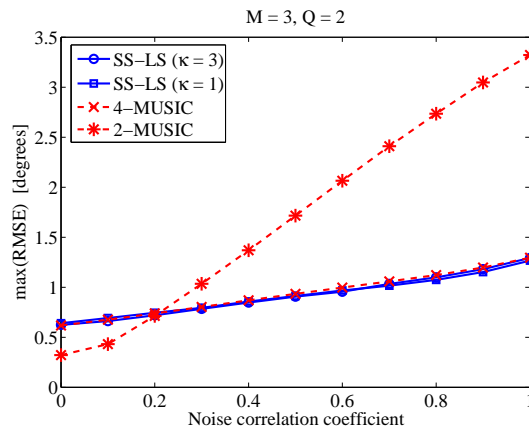
Figure 3.8: Maximal RMSE vs. noise spatial correlation ($N = 1000$ and SNR=5dB).
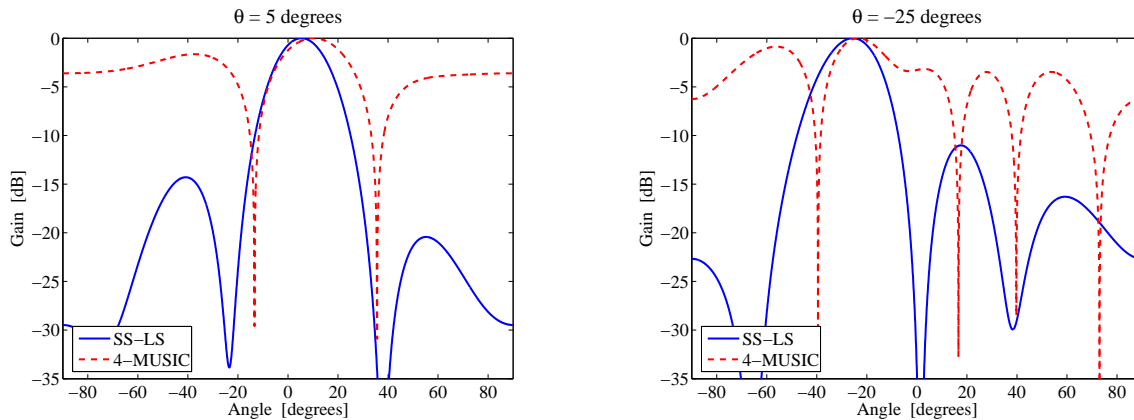
Figure 3.9: Antenna pattern obtained from the 3rd- and 2nd-order VAs for a ULA with 4 sensors.

In fig. 3.9 we show, for a ULA with 4 sensors spaced of $\lambda/2$, the array response of the 3rd-and 2nd-order VAs for a source arriving from direction $\theta_0 = 5°$ (left) and $\theta_0 = -25°$ (right). These curves have been obtained from the magnitude of the inner product of the VA steering vectors, computed as shown in (3.11). In both cases, for a 3dB attenuation, the beamwidth of the estimated 3rd-order VA (SS-LS) in the given direction is narrower than the one obtained with 4-MUSIC. In practice, this latter one is computed from the eigenvector of $\hat{\mathbf{C}}^{(4,y)}$ associated with its largest eigenvalue. We used $N = 1000$ output symbols with additive white Gaussian noise at a SNR level of 10dB. The gain in terms of the beamwidth for an attenuation of 3dB is about $8°$ for the curves at left and $3°$ for the figure at right.

Multipath MIMO channel estimation

Let us consider a Q-sensor transmit and an M-sensor receive ULA arrays located far apart each other, transmitting signals over a multipath specular radio channel such as defined in (3.28). The multipath channel is characterized by K remote scatterers each one determining a set of physical parameters $(\theta_k, \vartheta_k, \tau_k, \beta_k)$, according to the model (3.28). In which follows, we will be first interested in the non-parametric representation of the multipath channel and in estimating the channel impulse response coefficients $h_{mq}(\ell)$ of the convolutive MIMO model, up to a scaling factor and a permutation on the order of the transmit antennas. After that, exploiting the physical model given in (3.30), we make use of the ALS-based algorithm described in section 4.2 to estimate the spatial and temporal channel signatures and extract the multipath parameters by means of subspace-based techniques.

Non-parametric estimation of convolutive MIMO channel

In order to assess the quality of the non-parametric MIMO channel estimation, we will use the normalized mean squared error (NMSE) performance criterion[4], defined as:

$$\text{NMSE} = \frac{1}{R} \sum_{r=1}^{R} \sum_{q=1}^{Q} \text{NMSE}^{\langle r \rangle}(q), \tag{3.78}$$

where R is the number of Monte Carlo simulations and $\text{NMSE}^{\langle r \rangle}(q) = \min_{q_1 \in [1,Q]} \left(\epsilon_{\langle r \rangle}^{(q_1)}(q) \right)$ and

$$\epsilon_{\langle r \rangle}^{(q_1)}(q) = \frac{\|\widehat{\mathbf{H}}_{\langle r \rangle}^{(q_1)}(0) - \mathbf{H}^{(q)}(0)\|_F^2}{\|\mathbf{H}^{(q)}(0)\|_F^2}, \tag{3.79}$$

with $\widehat{\mathbf{H}}_{\langle r \rangle}^{(q_1)}(0)$ being the q_1th $M \times (L+1)$ block of the estimated matrix $\hat{\mathbf{A}}_{\langle r \rangle}$, obtained from (3.62) after convergence of the simulation r, assuming that $\hat{\mathbf{H}}_{\langle r \rangle}^{(q_1)}(0)$ has been optimally scaled with respect to $\mathbf{H}^{(q)}(0)$, defined in (3.31).

The following simulation results have been obtained with entirely synthetic 4th-order output cumulant data. In order to reproduce the effects of the additive Gaussian noise corrupting the output signals and to emulate the errors due to cumulant estimation from finite-length output data sequences, we have modeled the 4th-order output cumulant as follows:

$$\widehat{\mathbf{C}}_{[1]} = \mathbf{C}_{[1]} + \sigma_e^2 \mathbf{E} + \sigma_v^2 \mathbf{\Upsilon} \tag{3.80}$$

where $\mathbf{C}_{[1]} \in \mathbb{C}^{M^3(2L+1)^3 \times M}$ is the true output cumulant matrix computed from (3.60), while $\mathbf{E}$ and $\mathbf{\Upsilon}$ are complex random matrices of the same size as $\mathbf{C}_{[1]}$ and elements driven from standard Gaussian distributions. In this context, σ_e^2 represents the variance of the 4th-order cumulant

---

[4] Note that the NMSE is used only for performance evaluation purposes, since its computation involves the knowledge of the true channel coefficients.

estimators and σ_v^2 amounts to the additive noise power and is computed in such a way to ensure a given SNR level, given as follows:

$$\text{SNR} = \frac{\|\mathbf{C}_{[1]}\|_F}{\|\sigma_1^2 \mathbf{E}_1\|_F}. \tag{3.81}$$

In our simulations we have used the fixed value $\sigma_e^2 = 0.01$ whereas the SNR has been taken in the range of 5 to 35dB. The curves shown in the sequel have been obtained from the average of $R = 100$ Monte Carlo simulations.
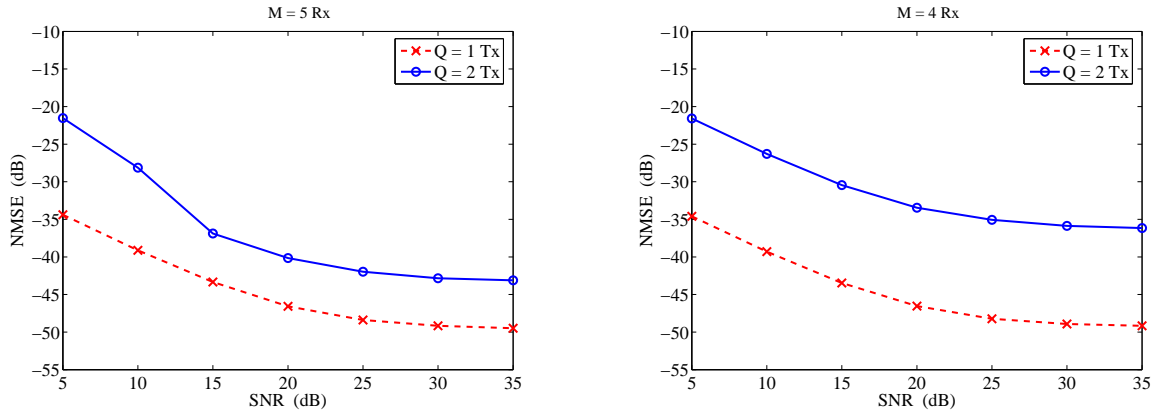
Figure 3.10: NMSE vs. SNR for channel configuration A with $L = 1$.
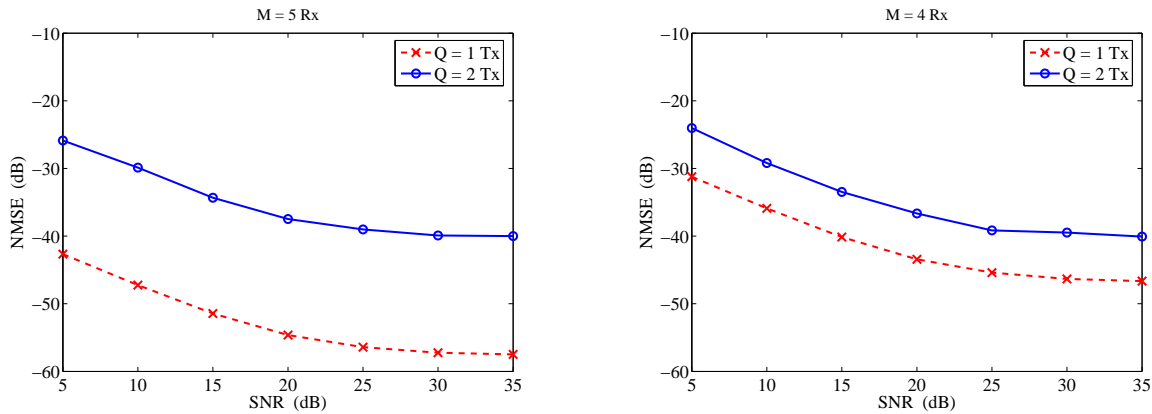
Figure 3.11: NMSE vs. SNR for channel configuration B with $L = 2$.

Two sets of channel configuration parameters have been considered and are described in Table 3.2. In fig. 3.10, we show the NMSE performance of the SS-LS algorithm for $M = 5$ (left) and $M = 4$ (right) receive antennas, under the channel configuration A ($K = 2$ multipath), using a pulse shape filter of order $L = 1$ (note that $\tau_{max} < T_s$ in this case). By increasing the number of transmit antennas from $Q = 1$ to $Q = 2$, the channels become more complex and
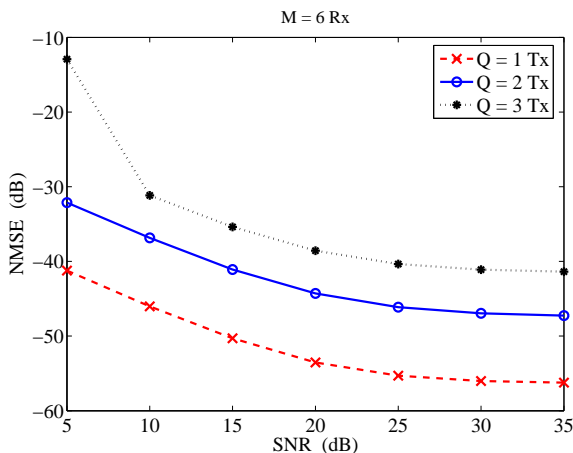
Figure 3.12: NMSE vs. SNR for channel configuration A with $L = 1$.

the curves clearly show an identification performance loss due to the co-channel interference. In fig. 3.11, we show similar results obtained under the channel configuration B, in which multipath propagation is characterized by $K = 3$ rays. In this case, $T_s < \tau_{max} < 2T_s$ and we use $L = 2$. Again, the curves show worse results for $Q = 2$, as it should be expected, and a performance degradation can also be observed with respect to the case of $L = 1$. In both figures, 3.10 and 3.11, comparing the left and right graphs, we notice an improvement due to addition of a receive antenna for a given value of Q. Finally, in fig. 3.12, we included the case of $Q = 3$ transmit antennas under channel configuration A, with $L = 1$ and $M = 6$ receive antennas. The curves show that the SS-LS algorithm correctly identified the MIMO channel coefficients in this case, in spite of the performance loss.

Table 3.2: Channel configuration parameters.

	Configuration A	Configuration B
Number of paths	2	3
DOAs	$\theta_1 = 40°, \theta_2 = -30°$	$\theta_1 = 50°, \theta_2 = -5°, \theta_3 = -45°$
DODs	$\vartheta_1 = 50°, \vartheta_2 = -5°$	$\vartheta_1 = 45°, \vartheta_2 = -35°, \vartheta_3 = -10°$
Path delays[†]	$\tau_1 = 0.3T_s, \tau_2 = 0.85T_s$	$\tau_1 = 0.35T_s, \tau_2 = 0.8T_s, \tau_3 = 1.4T_s$

[†] T_s stands for the symbol period.

Subspace-based algorithms for multipath parameter extraction

In the sequel, we will be interested in estimating the physical multipath parameters of the MIMO channel using the combined ALS-MUSIC algorithm proposed in section 4.2. In this section, in
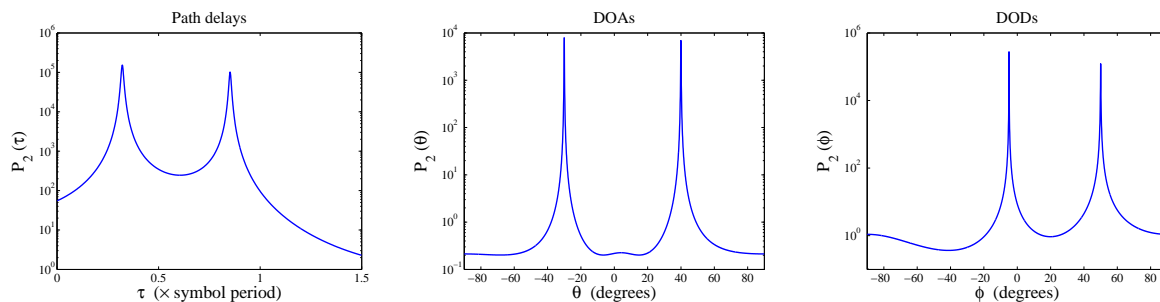
Figure 3.13: Physical multipath parameter extraction using ALS-MUSIC method for channel configuration A with $Q = 3$ Tx, $M = 5$ Rx and $L = 2$.
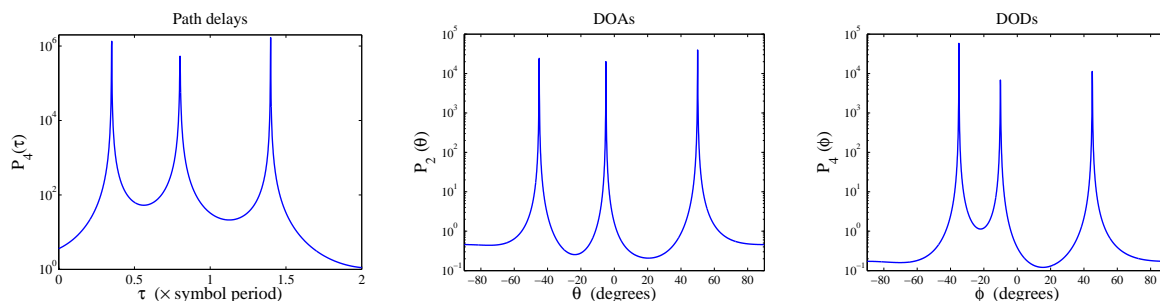
Figure 3.14: Physical multipath parameter extraction using ALS-MUSIC method for channel configuration B with $Q = 3$ Tx, $M = 5$ Rx and $L = 2$.

order to estimate the multipath channel parameters, we have first obtained an estimate of the convolutive MIMO coefficients using the non-parametric SS-LS algorithm based on the synthetic 4th-order output cumulant model given (3.80). The following results have been obtained with $\sigma_e^2 = 0.01$ and a SNR of 20B. After that, we decompose the estimated channel coefficient tensor using the ALS algorithm and extract the physical multipath parameters from the resulting Parafac components using a MUSIC-like algorithm. In fig. 3.13 we show the localization function $P_2(\cdot)$ as a function of τ, θ and ϕ, respectively, for one realization of the experiment using channel configuration A ($K = 2$) with $Q = 3$, $M = 5$ and $L = 2$. In fig. 3.14 we show the results for channel configuration B ($K = 3$) with the same parameters ($Q = 3$, $M = 5$, $L = 2$). Notice that, in this latter case, $\mathbf{A}_T$ and $\mathbf{G}$ are square matrices since $K = Q = L + 1$. For that reason, we need to use the 4th-order localization functions $P_4(\cdot)$, as a function of τ and ϕ, in order to estimate the DODs and path delays. For the DOAs, we keep $P_2(\theta)$, since $M > K$.

6 Summary

The goals of this chapter have been two fold: firstly, we have been interested in the blind source localization problem in the context of multiuser narrowband array processing, under the assumption of sources located at the far-field of the antenna array. The DOA estimation problem has been treated using the 4th-order cumulants only. A high-resolution DF algorithm has been proposed, exploiting the structure of the cumulant tensor. This method is based on the estimation of an array matrix formed from a double (column-wise) Kronecker product, thus creating an enhanced virtual array that commonly only arises when using 6th-order statistics. This yields an augmented observation space, which allows for more free dimensions at the antenna array and provides a resolution improvement, without resorting to statistics of order higher than fourth. Consequently, the proposed method works well even for relatively short output data sequences and it is robust with respect to an additive Gaussian noise with an unknown spatial correlation. Making use of the symmetry properties of 4th-order output cumulants, the estimation of the enhanced virtual array utilizes the SS-LS technique to perform the Parafac decomposition of the cumulant tensor. In the case of ULA arrays, this yields as many resolvable sources as the 4-MUSIC algorithm but with better DOA estimation performance, as confirmed by our simulation results.

In the second part of this chapter, we turned our attention to the estimation of the physical parameters describing a multipath MIMO communication channel characterized by specular reflections due to remote scatterers. This problem has been treated with a two-stage approach. We start by modeling the physical channel between each transmit-receive antenna pair as an FIR filter with coefficients that are given as a function of the multipath MIMO channel parameters. At that stage, we have proposed a 4th-order cumulant tensor formulation that generalizes the notation introduced in Chapter 2 for the convolutive SISO and the memoryless MIMO cases. This allows us to solve the blind channel estimation problem from a non-parametric perspective, using the SS-LS approach and exploiting the block-Hankel structure of the channel matrix, without oversampling the channel output. After that, we consider the structured multipath MIMO channel coefficients to form a tensor that admits a Parafac decomposition with rank equal to the number of multipaths. An ALS-based algorithm has been used to estimate the channel spatial and temporal signatures and the multipath parameters have been extracted by means of subspace-based algorithms, enabling us to recover the MIMO channel coefficients without ambiguities. Illustrative simulations have been shown demonstrating the applicability of the proposed methods in practical situations.

Chapter 4

Order selection, signal detection and blind identification of MISO channels

onvolutive propagation channels are typical in radiocommunication systems. In this context, the channel memory is known to be closely related to the delay spread profile produced by the multipath propagation scenario and it indicates the length of the channel impulse response. Long delay spread profiles characterize highly frequency-selective channels and introduce intersymbol interference (ISI) in the sampling process. The order of the radio mobile channel relates the length of the channel impulse response with the symbol period by indicating the number of past symbols being convolved with the current one. In practice, the channel order is given by the number of symbol periods fitting the (truncated) channel delay spread profile. Most channel parameter estimation algorithms require the knowledge of the channel order or, at least an upper bound of it [128]. Channel order mismatches may have very costly consequences, including bit error rate (BER) floors, signal-to-noise ratio (SNR) penalties and numerical instabilities [129, 130, 131, 31].

In this chapter, we address the problem of determining the order of finite impulse response (FIR) channels in the context of a multiple-input single-output (MISO) communication system, using only the 4th-order cumulants of the output data sequence. Channel order estimation is a classic model selection problem strongly related to determining the number of signals embedded in noisy observations in narrow-band array processing. This has been often referred to as the signal (or source) detection problem [30, 31, 32]. Classical procedures for model order determination are based on multiple hypothesis testing and make use of the eigenvalues of a sample correlation matrix. This is the case of the well-known sphericity test [132, 33], which estimates the number of model parameters by testing on adjacent groups of eigenvalues, in the case of a Gaussian process [133]. This approach has been widely used for solving the signal detection problem in the context of passive arrays [34, 35]. Other traditional techniques are based on information theoretical criteria that minimize the information lost due to the model approximation [134, 135]. The Akaike's Information Criterion (AIC) [136] as well as

the Rissanen's Minimum Description Length (MDL) [137], also known as the Schwarz Bayesian Criterion (SBC) [138], can both be used to test the equality of the smallest eigenvalues of the sample correlation matrix. Since AIC is not consistent [138], it tends to overestimate the model order, even for high SNR values [139]. In general, MDL performs better than AIC, but it tends to an underestimation at low and medium SNR [140]. Similar results have also been demonstrated in the context of source detection in passive arrays [141, 142].

Opposite to previous works, the order detection method proposed in this chapter is based on a scalar Chi-square test statistic that is sensitive to the non-linearity of a stochastic process, since it is derived from the multivariate estimator of the 4th-order output cumulants. Exploiting this property enables us to detect the order of a single-input single-output (SISO) communication channel by measuring the channel energy in the space of representation of the 4th-order cumulants. Making use of existing results on the asymptotic variance of the test variable, we present in section 1.2 a performance analysis of the proposed detector. In addition, we establish a criterion for fixing decision thresholds based on a given level of statistical tolerance. Some implementations of the proposed method are provided in the context of radiocommunication channels demonstrating very good fit between the empirical results and the theoretical curves.

Nevertheless, in the case of MISO communication channels, the propagation scenario can be viewed as a highly underdetermined convolutive mixture (more sources than sensors). Overdetermined mixtures have been exhaustively studied in the literature, including instantaneous [143, 144, 145, 146, 51, 12] as well as convolutive mixtures (c.f. [111, 112, 114, 116, 118, 115] and references therein). The case of underdetermined mixtures, on the other hand, has only recently been treated [67, 42, 70, 71, 73], and systems with one single output sensor have received considerably less attention. In this work, we introduce a deflation-type approach to solve the MISO channel order detection and identification problem. The proposed method is based on the use of a combined procedure that jointly selects the order, detects the number of sources and identifies the coefficients of the MISO channels. The so-called HOS-based nested detector searches for the longest channel, determines its order and estimates the associated coefficients. Then, based on the deflation principle, it successively tests and detects the presence of shorter and shorter channels. Adopting the Chi-square channel order selection method proposed in the first part of this chapter, our detector makes use of blind identification techniques for properly estimating the MISO channel coefficients using the 4th-order cumulants. A preliminary version of this deflation approach has been introduced in [147], where only the case of 2×1 MISO systems has been considered and the order detection test statistics were based on the eigenvalues of 4th-order output cumulant matrices.

This chapter is organized as follows: in section 1, we derive a Chi-square test statistic based on 4th-order cumulant information; the principles underlying our SISO channel order detector are introduced and an asymptotic performance analysis is carried out; then, in section 2, we

introduce the concept of nested detectors for combined order selection, signal detection and blind identification in the context of convolutive MISO communication channels; this idea exploits the residual 4th-order information remaining after subtraction of the reconstructed cumulants of previously estimated sources; in section 3, computer simulations results are provided to illustrate the good performance obtained in terms of channel order detection and identification; we finally draw our conclusions in section 4 along with some perspectives for future works.

1 HOS-based Chi-square test for SISO channel order selection

In order to introduce the main ideas behind our channel order detector, we first consider the case of a single input. Let us consider the baseband representation of a radiocommunication channel in which the output signal $y(n)$, after sampling at the symbol rate, is written as follows:

$$
\begin{aligned}
x(n) &= \sum_{\ell=0}^{L} h(\ell) s(n-\ell), \\
y(n) &= x(n) + v(n),
\end{aligned}
\tag{4.1}
$$

where the complex coefficients $h(\ell)$ represent the equivalent discrete impulse response of the channel, including pulse shaping and receiving filters. The model (4.1) introduces memory to the received signal corresponding to L times the symbol period. Throughout the rest of this work, we denote the channel order by $L+1$. The following assumptions hold:

A1 : The discrete input sequence $s(n)$ is complex-valued, non-measurable, ergodic, stationary, independent and identically distributed (iid) with symmetric distribution, zero-mean, unit variance and non-zero kurtosis $\gamma_{4,s}$.

A2 : The additive noise sequence $v(n)$ is normally distributed with zero-mean and unknown autocorrelation function. It is assumed to be independent from $s(n)$.

A3 : The FIR filter representing the channel is assumed to be causal, i.e. $h(\ell) = 0$, $\forall\, \ell \notin [0, L]$. In addition, $h(\ell) \neq 0$ for $\ell = L$ and $\ell = 0$.

A4 : The channel order is bounded by a known value K, i.e. $L + 1 < K$.

The 4th-order output cumulants are defined as follows:

$$
c_{4,y}(i,j,k) \triangleq cum\left[y^*(n), y(n+i), y^*(n+j), y(n+k) \right].
\tag{4.2}
$$

Using the channel model (4.1), taking assumptions A1 and A2 into account and making use of the multilinearity property of cumulants, we get [3]:

$$
c_{4,y}(i,j,k) = \gamma_{4,s} \sum_{\ell=0}^{L} h^*(\ell) h(\ell+i) h^*(\ell+j) h(\ell+k)
\tag{4.3}
$$

where $\gamma_{4,s} = c_{4,s}(0,0,0)$. Let us define the 4th-order output cumulant vectors $\mathbf{c}_k \in \mathbb{C}^{P \times 1}$, for $k \in [1, K]$, with elements given by $[\mathbf{c}_k]_p = C_k^{(p)}$, $p \in [1, P]$, where

$$C_k^{(p)} \triangleq c_{4,y}(i_p - 1, j_p - 1, k - 1), \quad (i_p, j_p) \in \mathcal{J}, \tag{4.4}$$

and each pair (i_p, j_p) is formed of strictly positive natural numbers, with $i_p, j_p \leq K$, belonging to the index set $\mathcal{J} = \{(i_1, j_1); \ldots; (i_P, j_P)\}$, with cardinality P. Notice that the decision about which cumulants are to be included in $\mathbf{c}_k$ is done by carefully choosing the elements of the index set $\mathcal{J}$. This choice may vary in function of the assumptions considered in each particular application. Further discussion on this subject is postponed to section 2.1, in the context of the blind identification of MISO channels. Due to (4.3) and assumption A3, we have:

$$C_k^{(p)} = 0, \ \forall \ |i_p|, |j_p|, |k| > L + 1. \tag{4.5}$$

Hence, the 4th-order output cumulants are zero whenever either i_p, j_p, or k are larger than the channel order $L + 1$. Furthermore, from the above definitions, we also note that $\mathbf{c}_k$ can include some purely real-valued components. More precisely, we have:

$$C_k^{(p)} = C_k^{(p)*} \quad \begin{cases} \text{if} \quad k = 1 \quad \text{and} \quad j_p = i_p \\ \qquad \text{or} \\ \text{if} \quad i_p = 1 \quad \text{and} \quad j_p = k. \end{cases} \tag{4.6}$$

In most of the real-life situations, the true value of the output cumulants is not available and have to be estimated from the output signal samples $y(n)$, $n = 0, \ldots, N - 1$. As discussed in section 1.1.2 of Chapter 1, due to the ergodicity assumption, the 2nd- and 4th-order moments can be estimated by replacing the expectations by the corresponding time averages, as in (1.35). The cumulant estimator $\hat{C}_k^{(p)}$ thus obtained take the form of (1.26). It is possible to show that this estimator is biased, but its bias tends towards zero as N goes to infinity [47]. Cumulant estimators are also consistent, since their variance is shown to be asymptotically zero.

We can hence define the complex-valued estimator $\hat{\mathbf{c}}_k$, which is written as:

$$\hat{\mathbf{c}}_k = \left[\hat{C}_k^{(1)} \ \cdots \ \hat{C}_k^{(P)} \right]^{\mathsf{T}}. \tag{4.7}$$

Let $\mathbf{V}_k$ and $\mathbf{W}_k$ be the $P \times P$ positive-definite *circular* and *non-circular* covariance matrices of the estimator $\hat{\mathbf{c}}_k$, defined respectively as follows:

$$\mathbf{V}_k \triangleq \mathbb{E}\left\{ \left(\hat{\mathbf{c}}_k - \mathbf{c}_k \right) \left(\hat{\mathbf{c}}_k - \mathbf{c}_k \right)^{\mathsf{H}} \right\}; \tag{4.8}$$

$$\mathbf{W}_k \triangleq \mathbb{E}\left\{ \left(\hat{\mathbf{c}}_k - \mathbf{c}_k \right) \left(\hat{\mathbf{c}}_k - \mathbf{c}_k \right)^{\mathsf{T}} \right\}. \tag{4.9}$$

Note that both $\mathbf{V}_k$ and $\mathbf{W}_k$ are complex-valued matrices, with $\mathbf{V}_k$ being Hermitian, while $\mathbf{W}_k$ is a symmetric matrix. Let us now define the following real-valued 4th-order cumulant vector:

$$\mathbf{z}_k = \begin{bmatrix} \mathrm{Re}(\mathbf{c}_k)^\mathsf{T} & \mathrm{Im}(\mathbf{c}_k)^\mathsf{T} \end{bmatrix}^\mathsf{T} \quad \in \mathbb{R}^{2P \times 1}, \tag{4.10}$$

where the operators $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ return the real and imaginary parts of the vector arguments, respectively. Consider the estimator $\hat{\mathbf{z}}_k$ with covariance matrix $\mathbf{\Sigma}_k \triangleq \mathbb{E}\left\{ (\hat{\mathbf{z}}_k - \mathbf{z}_k)(\hat{\mathbf{z}}_k - \mathbf{z}_k)^\mathsf{T} \right\} \in \mathbb{R}^{2P \times 2P}$, which can be readily deduced from (4.8) and (4.9), as follows:

$$\mathbf{\Sigma}_k = \frac{1}{2} \begin{pmatrix} \mathrm{Re}(\mathbf{V}_k + \mathbf{W}_k) & \mathrm{Im}(\mathbf{V}_k + \mathbf{W}_k)^\mathsf{T} \\ \mathrm{Im}(\mathbf{V}_k + \mathbf{W}_k) & \mathrm{Re}(\mathbf{V}_k - \mathbf{W}_k) \end{pmatrix}. \tag{4.11}$$

We can now define the following scalar multivariate function:

$$\xi_k = \left(\mathbf{z}_k - \hat{\mathbf{z}}_k \right)^\mathsf{T} \mathbf{\Sigma}_k^{-1} \left(\mathbf{z}_k - \hat{\mathbf{z}}_k \right) \tag{4.12}$$

which depends, though omitted here, on the channel coefficient vector $\mathbf{h} = [h(0) \ \ldots \ h(L)]^\mathsf{T}$.

It appears that Porat and Friedlander have been the first to use the above function in the context of channel parameter estimation [28, 97]. Actually, they showed that if the estimated parameter vector $\hat{\mathbf{h}}$ yields a global minimum of ξ_k, then $\hat{\mathbf{h}}$ is an asymptotically minimum variance estimate [97]. However, this approach requires the calculation of the covariance matrix $\mathbf{\Sigma}_k$, which depends on $\mathbf{h}$ and involves knowledge of exact output cumulants of order up to eight. Due to the enormous computational complexity involved in calculating $\mathbf{\Sigma}_k$, a simpler solution has been proposed in [148] based on an estimated covariance matrix. Expressions for computing $\mathbf{V}_k$ and $\mathbf{W}_k$ in the case of input signals with discrete probability distributions and spatially uncorrelated Gaussian noise have also been given [29]. Other exact expressions for the computation of the covariance matrices are available in the literature for symmetrically distributed sources and also in the general case (c.f. [149] for the former and [150] for the latter). These expressions are very important for the theoretical analysis of the estimator when the model parameters are assumed known, but they have also found application in the context of Gaussianity tests [151]. On the other hand, they can also be useful for algorithmic purposes when only output measurements are available. In this case, ergodicity can be exploited and time averages can be used to estimate moments and cumulants. These estimates take the place of the actual cumulant values in the covariance expressions (this approach is often used, e.g. in [29]).

Finally, note that $\mathbf{z}_k$ may have some zero elements corresponding to the entries $[\mathbf{z}_k]_{P+p}$ for which p and k are such that the conditions stated in (4.6) are satisfied. Each zero element in $\mathbf{z}_k$ induces a zero row and a zero column in matrix $\mathbf{\Sigma}_k$. To avoid singularity of the covariance matrix, we need to eliminate the element in position $P + p$ of the vector $\mathbf{z}_k$ for each triplet (i_p, j_p, k) satisfying (4.6). This yields a reduced vector $\mathbf{z}_k \in \mathbb{R}^{2P - m_k}$, where m_k is the number of purely real-valued elements in $\mathbf{c}_k$. Consequently, in order to obtain the covariance matrix of $\mathbf{z}_k$,

we have to suppress from $\mathbf{\Sigma}_k$ the rows and the columns corresponding to each triplet (i_p, j_p, k) satisfying (4.6), which yields a $(2P - m_k) \times (2P - m_k)$ matrix. This is an important step in order to ensure the non-singularity of $\mathbf{\Sigma}_k$. Throughout the rest of this chapter, we denote by $\mathbf{z}_k$ and $\mathbf{\Sigma}_k$ the reduced versions of these variables, thus assuming that the covariance matrix is nonsingular and the real-valued cumulant vector has $2P - m_k$ elements.

1.1   A Chi-square statistic for channel order detection

Given a sample output data sequence $y(n)$, $n = 0, \dots, N - 1$, it can be shown that, as N goes to infinity, the estimator $\hat{\mathbf{c}}_k$ approaches a complex multivariate random variable that follows an approximately Gaussian distribution with mean equal to $\mathbf{c}_k$ and covariance matrix $\mathbf{V}_k$, i.e. $\hat{\mathbf{c}}_k \sim \mathcal{N}(\mathbf{c}_k, \mathbf{V}_k)$ [152]. As a consequence, we have:

$$\hat{\mathbf{z}}_k \sim \mathcal{N}(\mathbf{z}_k, \mathbf{\Sigma}_k), \quad \text{as } N \to \infty. \tag{4.13}$$

Hence, $\hat{\mathbf{z}}_k$ can be viewed as a realization of an asymptotically Gaussian random vector, which can be standardized as follows:

$$\boldsymbol{\omega}_k = \mathbf{\Sigma}_k^{-1/2}(\mathbf{z}_k - \hat{\mathbf{z}}_k) \tag{4.14}$$

so that $\boldsymbol{\omega}_k \in \mathbb{R}^{2P - m_k}$ is asymptotically normal with zero mean and unit variance, i.e. $\boldsymbol{\omega}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. It is now easy to see that the scalar random variable ξ_k defined in (4.12), can be written as:

$$\xi_k = \boldsymbol{\omega}_k^\mathsf{T} \boldsymbol{\omega}_k. \tag{4.15}$$

The above results enable us to conclude that ξ_k asymptotically follows a Chi-square distribution with $d_k = 2P - m_k$ degrees of freedom [133], i.e.

$$\xi_k \sim \mathcal{X}_{(d_k)}^2. \tag{4.16}$$

Therefore, its probability density function (pdf) can be written as follows:

$$f_\xi(\xi_k) = \begin{cases} \frac{1}{2^{d_k/2}\Gamma(d_k/2)} \xi_k^{(d_k/2)-1} e^{-\xi_k/2}, & \text{for } \xi_k > 0 \\ 0 & \text{for } \xi_k \leq 0, \end{cases} \tag{4.17}$$

where $\Gamma(\cdot)$ denotes the well-known Gamma function, defined as $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} \mathrm{d}t$. From the above equation, it is not difficult to deduce that:

$$\mu_{\xi_k} = \mathbb{E}\{\xi_k\} = d_k \tag{4.18}$$

and

$$\sigma_{\xi_k}^2 = \mathbb{E}\{(\xi_k - \mu_{\xi_k})^2\} = 2d_k. \tag{4.19}$$

1.2   Order detection algorithm

In the sequel, we formulate the problem of determining the channel order as a series of successive hypothesis tests on adjacent groups of variables ρ_k, $k \in [1, K]$, $K > L + 1$, aiming to determine whether $k > L + 1$ or not. Contrary to the approach based on exhaustive search for the minimum variance, our proposition exploits the fact that $\mathbf{z}_k = \mathbf{0}$ if $k > L + 1$. Thus, replacing the covariance matrix $\boldsymbol{\Sigma}_k$ by its estimate, equation (4.14) becomes:

$$\bar{\boldsymbol{\omega}}_k = -\hat{\boldsymbol{\Sigma}}_k^{-1/2} \hat{\mathbf{z}}_k, \quad \text{for } k > L + 1, \tag{4.20}$$

and we can define

$$\rho_k = \bar{\boldsymbol{\omega}}_k^\mathsf{T} \bar{\boldsymbol{\omega}}_k = \hat{\mathbf{z}}_k^\mathsf{T} \, \hat{\boldsymbol{\Sigma}}_k^{-1} \, \hat{\mathbf{z}}_k, \quad k > L + 1. \tag{4.21}$$

The above defined variable can be viewed as a measure of the energy in the space of representation of the 4th-order cumulants. It means that, since $\hat{\mathbf{z}}_k$ is a consistent estimator with asymptotically zero bias, ρ_k should be able to detect the presence of signal sources with nonzero 4th-order cumulants whenever $k \leq L + 1$.

From definition (4.21), we conclude that, as the output data sequence length goes to infinity, ρ_k tends to be $\mathcal{X}^2_{(d_k)}$, if $k > L + 1$. In this case, the pdf of ρ_k is given by (4.17). However, for $k \leq L + 1$, the true cumulant vector $\mathbf{z}_k$ is unknown and we cannot center the random vector $\boldsymbol{\omega}_k$, hence ρ_k assumes a *non-central* Chi-square distribution, denoted $_{nC}\chi^2_{(d_k)}(\lambda_k)$, with $d_k = 2P - m_k$ degrees of freedom and parameter $\lambda_k = \mathbf{z}_k^\mathsf{T}\boldsymbol{\Sigma}_k^{-1}\mathbf{z}_k$, which is related to the mean of the test variable as follows: $\mathbb{E}\{\rho_k\} = \lambda_k + d_k$. From (4.21), we have: $\rho_k \sim \,_{nC}\chi^2_{(d_k)}(\lambda_k)$, with $\lambda_k = 0$ for $k > L + 1$ and $\lambda_k \neq 0$ for $k \leq L + 1$.

Remark on the number of degrees of freedom d_k

Situations may arise where the estimated covariance matrix $\hat{\boldsymbol{\Sigma}}_k$ is ill-conditioned due, for instance, to negligible values of the cross-correlation between the real and imaginary parts of some of the 4th-order cumulants composing the test statistic ρ_k. Such situations are difficult to predict and the literature lacks of guidelines on how to proceed in order to skip them. In spite of that, numerical instabilities can be avoided by controlling the condition number of $\hat{\boldsymbol{\Sigma}}_k$, i.e. the ratio between its greatest and smallest eigenvalues. In practice, when the condition number is high, we eliminate the smallest eigenvalues and the associated eigenvectors, until the condition number falls below a certain threshold. Thus, we can write the economy-size eigenvalue decomposition (EVD) of $\hat{\boldsymbol{\Sigma}}_k \in \mathbb{R}^{(2P-m_k)\times(2P-m_k)}$, as follows:

$$\hat{\boldsymbol{\Sigma}}_k = \mathbf{U}_k \mathbf{D}_k \mathbf{U}_k^\mathsf{T}, \tag{4.22}$$

where $\mathbf{D}_k = Diag\,(\delta_1, \ldots, \delta_{2P-m_k-\mu_k})$ and $\mathbf{U}_k \in \mathbb{R}^{(2P-m_k)\times(2P-m_k-\mu_k)}$ is the matrix concatenating the eigenvectors of $\hat{\mathbf{\Sigma}}_k$ associated with its $2P - m_k - \mu_k$ largest eigenvalues, with μ_k corresponding to the number of small eigenvalues eliminated in order to attain a moderate condition number. From (4.22), we have $\hat{\mathbf{\Sigma}}_k^{-1/2} = \mathbf{D}_k^{-1/2}\mathbf{U}_k^\mathsf{T} \in \mathbb{R}^{(2P-m_k-\mu_k)\times(2P-m_k)}$, and thus (4.20) becomes:

$$\bar{\boldsymbol{\omega}}_k = -\mathbf{D}_k^{-1/2}\mathbf{U}_k^\mathsf{T}\hat{\mathbf{z}}_k \quad \in \mathbb{R}^{(2P-m_k-\mu_k)\times 1}. \tag{4.23}$$

In this case, since the dimension of the real-valued vector $\bar{\boldsymbol{\omega}}_k$ has been modified, the number of degrees of freedom of the test statistic $\rho_k = \bar{\boldsymbol{\omega}}_k^\mathsf{T}\bar{\boldsymbol{\omega}}_k$ is reduced to $d_k = 2P - m_k - \mu_k$.

Hypothesis test and performance analysis

Let us build our channel order test by defining the null hypotheses $H_0(k)$ and corresponding alternative hypotheses $H_1(k)$ as follows:

$H_0(k):$ Channel order is strictly smaller than k $(k > L + 1)$ $\quad \Rightarrow \quad \rho_k \sim \mathcal{X}^2_{(d_k)}$

$H_1(k):$ Channel order is k $(k = L + 1)$ $\qquad\qquad\qquad \Rightarrow \quad \rho_k \sim_{nC} \chi^2_{(d_k)}(\lambda_k)$

Under $H_0(k)$, we have $\mathbb{E}\left\{\rho_k\right\} = 2P - m_k$ and hence we should expect that $\rho_k < \eta_k$, where η_k is a decision threshold associated with the number of degrees of freedom of the test statistic ρ_k. Under $H_1(k)$, $\mathbb{E}\left\{\rho_k\right\} = \lambda_k + 2P - m_k$, and we should get $\rho_k \geq \eta_k$. The test is successively performed for $k = K, K - 1, \ldots, 1$. Our goal is to find the largest value of k so that the null hypothesis $H_0(k)$ is rejected, i.e. $\rho_k \geq \eta_k$, which implies $\hat{L} = k - 1$. The non-rejection of $H_0(k)$ for a given k $(\rho_k < \eta_k)$ induces a new test on ρ_{k-1}. If the null hypothesis is rejected only when $k = 1$, then $\hat{L} = 0$ and the channel is said to be memoryless. The rejection of $H_0(k)$ for all $k \in [1, K]$ indicates that no signal is being transmitted (only additive Gaussian noise is observed at the antenna output).

Let us denote by p_k and q_k the probabilities of the event $\rho_k \geq \eta_k$, under hypotheses $H_0(k)$ and $H_1(k)$, respectively, i.e.

$$p_k \quad \triangleq \quad \mathrm{P}[\rho_k \geq \eta_k | H_0(k)], \quad k \in [1, K], \tag{4.24}$$

$$q_k \quad \triangleq \quad \mathrm{P}[\rho_k \geq \eta_k | H_1(k)], \quad k \in [1, K]. \tag{4.25}$$

In addition, we denote by $P(k)$ the probability of getting $\hat{L} = k - 1$, $k \in [1, K]$, which is defined as the joint probability of the events $\rho_K < \eta_K$, $\rho_{K-1} < \eta_{K-1}$, $\ldots$, $\rho_{k+1} < \eta_{k+1}$ and $\rho_k \geq \eta_k$, i.e.

$$P(k) = \begin{cases} \mathrm{P}[\rho_K \geq \eta_K]; & \text{for } k = K, \\ \mathrm{P}[\rho_K < \eta_K; \rho_{K-1} < \eta_{K-1}; \ldots; \rho_{k+1} < \eta_{k+1}; \rho_k \geq \eta_k]; & \text{for } 1 \leq k < K. \end{cases} \tag{4.26}$$

The test on the variable ρ_k may have one of the following results, for each $k \in [1, K]$:

   i. $\rho_k < \eta_k \Rightarrow \hat{L} < k - 1$. The test continues. This happens with probability $1 - p_k$. The next step is to test ρ_{k-1}. In this case, if $k \leq L + 1$, we underestimate the channel order (detection missed) with unknown probability $1 - q_k$.

   ii. $\rho_k \geq \eta_k \Rightarrow \hat{L} = k - 1$. The test stops. This happens with probability q_k if indeed $k = L + 1$. In this case, the channel order is correctly detected. Otherwise, if $k > L + 1$, we overestimate the channel order with probability p_k.

Theorem 4.1 *The total probability of false alarm (overestimation) of the channel order is given as follows:*

$$P_F = \sum_{q=1}^{K-L-1} \sum_{i=1}^{N_q} (-1)^{q+1} \prod_{k \in \mathrm{I}_i^{(q)}} p_k, \tag{4.27}$$

with $\mathrm{I}_i^{(q)}$ denoting a q-subset of the set $\{K, K - 1, \ldots, L + 2\}$, i.e. a subset with q disjoint elements. The family of all q-subsets is given by $\mathcal{I}^{(q)} = \{\mathrm{I}_1^{(q)}, \mathrm{I}_2^{(q)}, \ldots, \mathrm{I}_{N_q}^{(q)}\}$, where N_q is the number of combinations of size q from a set with $K - L - 1$ elements, which equals the binomial coefficient, i.e. $N_q = \binom{K-L-1}{q} = \frac{(K-L-1)!}{(K-L-1-q)!q!}$.

Proof: Overestimation happens when we get $\hat{L} = k - 1$ for any $k > L + 1$. Hence, the total probability of overestimation of the channel order is given as follows:

$$P_F = \sum_{k=L+2}^{K} \mathrm{P}[\hat{L} = k - 1] = \sum_{q=1}^{K-L-1} P(K + 1 - q). \tag{4.28}$$

Assuming that the events $\rho_k > \eta_k$ are mutually independent for all $k \in [1, K]$, equation (4.26) yields:

$$P(k) = \begin{cases} p_K, & k = K, \\ (1 - p_K) \ldots (1 - p_{k+1}) p_k, & k = K - 1, \ldots, L + 2 \end{cases} \tag{4.29}$$

Replacing (4.29) into (4.28), equation (4.27) follows straightforwardly.

$\square$

Theorem 4.2 *The probability of detection of the channel order is given as follows:*

$$P_D = (1 - p_K)(1 - p_{K-1}) \ldots (1 - p_{L+2}) q_{L+1} \tag{4.30}$$

Proof: In order to correctly detect the channel order, we need: $\rho_K < \eta_K$, $\rho_{K-1} < \eta_{K-1}$, ... and $\rho_{L+2} < \eta_{L+2}$ under $H_0(k)$ and $\rho_{L+1} \geq \eta_{L+1}$ under $H_1(k)$. The joint probability of these events is given by (4.26), so that (4.30) follows immediately from (4.29) with $k = L + 1$.

$\square$

Theorem 4.3 *The probability of missing the channel order detection (underestimation) is given as follows:*

$$P_M = (1 - p_K)(1 - p_{K-1}) \dots (1 - p_{L+2})(1 - q_{L+1}) \tag{4.31}$$

Proof: If $k < L + 1$, the channel order is always underestimated. This means that $P_M = \mathrm{P}[\rho_K < \eta_K; \rho_{K-1} < \eta_{K-1}; \dots; \rho_{L+2} < \eta_{L+2}; \rho_{L+1} < \eta_{L+1}]$, which yields (4.31).

$\square$

The total probability of error in the order detection is given by $P_E = P_F + P_M$.

Decision thresholds

According to the Neyman-Pearson criterion [135, 153], a decision rule can be established in order to maximize the probability of detection P_D while not allowing the probability of false alarm P_F to exceed a certain value. However, in our particular case, P_F and P_D can not be computed explicitly as in (4.27) and (4.30), respectively, because we do not know the channel order L. We can nevertheless limit the tolerance of each ρ_k-test, by establishing a bound α for the acceptable level of the probability p_k, defined in (4.24), so that decision thresholds η_k can be established in order to ensure that $\mathrm{P}[\rho_k \geq \eta_k] \leq \alpha$, $k \in [K, \hat{L} + 2]$.

Considering definition (4.21) and under the null-hypothesis ($k > L + 1$), the test variable ρ_k is asymptotically Chi-square distributed with $d_k = 2P - m_k$ degrees of freedom, since $\mathbf{z}_k$ is theoretically zero for all $k > L + 1$. Otherwise, ρ_k follows a non-central Chi-square distribution, also with d_k degrees of freedom, but with an unknown parameter $\lambda_k = \mathbf{z}_k^\mathsf{T} \boldsymbol{\Sigma}_k^{-1} \mathbf{z}_k$, the value of which depending on the true real-valued cumulant vector $\mathbf{z}_k$. In other words, the distribution of the test variable ρ_k under the null hypothesis is given by (4.17), for a given choice of the index set $\mathcal{J}$. Thus, equation (4.24) becomes:

$$\begin{aligned} p_k &= \int_{\eta_k}^{+\infty} f_\rho(\rho_k) \mathrm{d}\rho_k, \quad k \in [1, K] \\ &= \frac{1}{2^{d_k/2} \Gamma(d_k/2)} \int_{\eta_k}^{+\infty} \rho_k^{(d_k/2)-1} e^{-\rho_k/2} \mathrm{d}\rho_k \end{aligned} \tag{4.32}$$

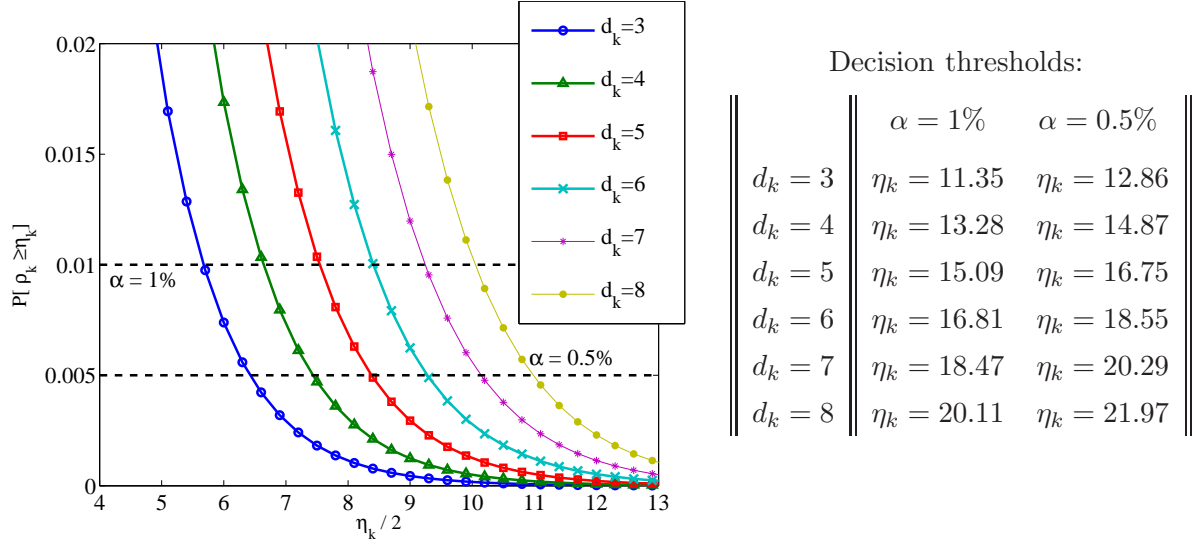| Decision thresholds:		
	$\alpha = 1\%$	$\alpha = 0.5\%$
$d_k = 3$	$\eta_k = 11.35$	$\eta_k = 12.86$
$d_k = 4$	$\eta_k = 13.28$	$\eta_k = 14.87$
$d_k = 5$	$\eta_k = 15.09$	$\eta_k = 16.75$
$d_k = 6$	$\eta_k = 16.81$	$\eta_k = 18.55$
$d_k = 7$	$\eta_k = 18.47$	$\eta_k = 20.29$
$d_k = 8$	$\eta_k = 20.11$	$\eta_k = 21.97$

Figure 4.1: pdf curves for ρ_k and decision thresholds: variation in function of the number of degrees of freedom.

Substituting $\rho_k = 2u$, we have:

$$
\begin{aligned}
p_k &= \frac{1}{\Gamma(d_k/2)} \int_{\eta_k/2}^{+\infty} u^{(d_k/2)-1} e^{-u} \mathrm{d}u, \\
&= \frac{\bar{\Gamma}(d_k/2, \eta_k/2)}{\Gamma(d_k/2)},
\end{aligned}
\tag{4.33}
$$

where $\bar{\Gamma}(d_k/2, \eta_k/2)$ is the *upper incomplete* Gamma function, defined as $\bar{\Gamma}(z, n) = \int_n^\infty t^{z-1} e^{-t} \mathrm{d}t$ [154]. Equation (4.33) enables us to plot the pdf curves of the marginal events $\rho_k \geq \eta_k$ for different values of the parameter d_k, as illustrated in fig. 4.1.

Decision thresholds η_k can now be established in such a way to limit the probability of events $\rho_k \geq \eta_k$, $k > L + 1$, to a low level α, while allowing $\mathrm{P}[\rho_{L+1} \geq \eta_{L+1}]$ to be as high as possible, depending on the value of $\mathbf{z}_k$. With such an aim, we choose η_k so that

$$
\bar{\Gamma}(d_k/2, \eta_k/2) \leq \alpha \Gamma(d_k/2).
\tag{4.34}
$$

Using the curves of fig. 4.1 corresponding to the exact number of degrees of freedom of the test variables ρ_k, we can define decision rules by placing the threshold levels at the points where each curve crosses the acceptable level of tolerance α. These threshold levels are indicated in the table in fig. 4.1 for the considered values of d_k, with $\alpha = 1\%$ and $\alpha = 0.5\%$.

Notice from (4.28) that the global level of false alarm probability still depends on α and $K - L$ so that a bad choice of K may reduce the power of the test (increased P_F should be expected for $K \gg L$). Possible solutions to this drawback include the use the Benjamini-Hochberg procedure for controlling the global level of false alarm (c.f. [32, 31, 155] and references therein). This approach will not be discussed here.

The proposed channel order detection algorithm can be summarized as follows:

Algorithm 4.1 (HOS-based channel order detector algorithm)

Define $\mathcal{J}$ and α and initialize the algorithm with $k = K$:

1. *Estimate the 4th-order output cumulants corresponding to the indices in $\mathcal{J}$ and compose $\hat{\mathbf{c}}_k \in \mathbb{C}^{P \times 1}$ as in (4.7);*

2. *Determine m_k as the number of real-valued cumulants in $\hat{\mathbf{c}}_k$; compose the real-valued vector $\hat{\mathbf{z}}_k \in \mathbb{R}^{(2P - m_k) \times 1}$*

3. *Estimate the 8th-, 6th- and 2nd-order output cumulants corresponding to the indices in $\mathcal{J}$; Compute the approximate circular and non-circular covariance matrices defined in (4.8) and (4.9), respectively [149];*

4. *Deduce $\hat{\boldsymbol{\Sigma}}_k$ from (4.11), take its EVD and test its condition number; Keep eliminating the smallest eigenvalue until the condition number becomes smaller than a certain threshold; define μ_k as the number of eliminated eigenvalues;*

5. *Determine the number of degrees of freedom $d_k = 2P - m_k - \mu_k$ and compute $\bar{\boldsymbol{\omega}}_k$ from (4.23);*

6. *Compute $\rho_k = \bar{\boldsymbol{\omega}}_k^{\mathsf{T}} \bar{\boldsymbol{\omega}}_k$ and derive η_k from (4.34) using d_k;*

7. *Test the hypotheses:*

   - *if $\rho_k < \eta_k$, take $k \leftarrow k - 1$ and repeat the procedure from step 1.*
   - *if $\rho_k \geq \eta_k$, then $\hat{L} = k - 1$. Stop the algorithm.*

Simulation examples

Let us consider a channel with $L + 1 = 3$ constant coefficients chosen at random from a complex Gaussian distribution. Using N noiseless output data samples, we compute the 4th-order cumulant estimates to form the vector $\hat{\mathbf{c}}_k$, defined in (4.7), for $(i_p, j_p) \in \mathcal{J}$ and $k = 1, \ldots, K$. In this example, we take $\mathcal{J} = \{(1,1); (1,2); \ldots; (1,P)\}$ and $P = K = 4$. After that, we compose the vector $\hat{\mathbf{z}}_k$ by taking the real and imaginary parts of $\hat{\mathbf{c}}_k$, as in (4.10), and then we eliminate the zero components corresponding to the imaginary part of the purely real-valued cumulants. Any triplet (i_p, j_p, k) satisfying the conditions given in (4.6) yields a purely real-valued cumulant, and we denote by m_k the number of such components in the vector $\hat{\mathbf{c}}_k$. For $k = 4$ and $k = 3$, due to our choice of $\mathcal{J}$, we have $m_4 = 0$ and $m_3 = 1$ (since $\hat{C}_k^{(3)}$ is purely real-valued for $k = 3$) and hence $d_4 = 8$ and $d_3 = 7$.
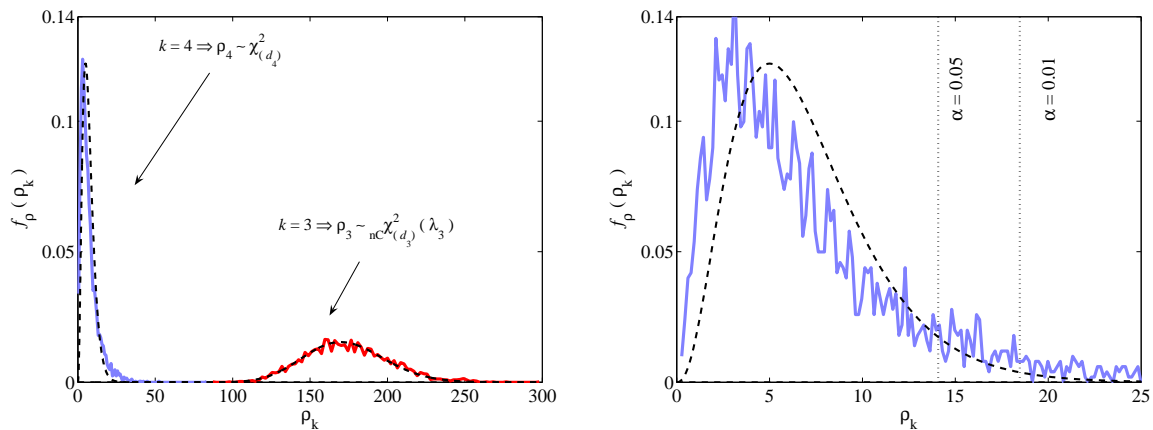
Figure 4.2: Histogram curves of ρ_3 and ρ_4 (left) in a noiseless static SISO channel scenario with $L = 2$, $N = 10000$ and 3000 Monte Carlo repetitions; dotted lines are the theoretical pdf curves; detail on the area associated with the null-hypothesis (right) showing the decision thresholds for $\alpha = 1\%$ and $\alpha = 5\%$.

Using the expressions given in [149] (for the case of symmetrically distributed sources) we can estimate the circular and non-circular covariance matrices $\hat{\mathbf{V}}_k$ and $\hat{\mathbf{W}}_k$, defined in (4.8) and (4.9), respectively, by replacing the true cumulant values by their estimates, obtained using time averages of the output signal. Thus, $\hat{\mathbf{\Sigma}}_k$ follows from (4.11) with subsequent elimination of the m_k rows and columns corresponding to the imaginary part of the purely real-valued cumulants. Then, the test variable ρ_k can be computed from (4.21). Decision thresholds can be deduced from the criterion (4.34), in function of d_k, as indicated in fig. 4.1.

In order to validate in practice the theoretical behavior of the test variable, we repeated the above described experiment by performing 3000 Monte Carlo simulations, varying the input data sequence $s(n)$ from one repetition to another, but keeping the same channel parameters all over the simulations (*static* channel scenario). In this case, we used $N = 10000$ and fig. 4.2 shows the histogram curves for the test variables ρ_4 and ρ_3. The figure at right consists in a zoom on the area of the ρ_4 pdf curve at left. Decision thresholds for $\alpha = 1\%$ and $\alpha = 5\%$ are depicted. The dotted lines correspond to the theoretical pdf curves of the associated variables. The curves for $k < 3$ have been omitted from fig. 4.2 because in this case $\lambda_k \gg \eta_k$ (none underestimation cases occurred).

In a second example, we performed another 200 Monte Carlo simulations varying the input data sequence as well as the channel parameters from one repetition to another (*quasi-static* channel scenario). The performance of the proposed order detection test can be assessed in terms of the detection and error rates, which are summarized in fig. 4.3 for $\alpha = 1\%$ and $\alpha = 0.5\%$ with $N = 500, 1000, 3000, 5000$ and 10000. Decision thresholds for $\alpha = 1\%$ and $\alpha = 0.5\%$ have
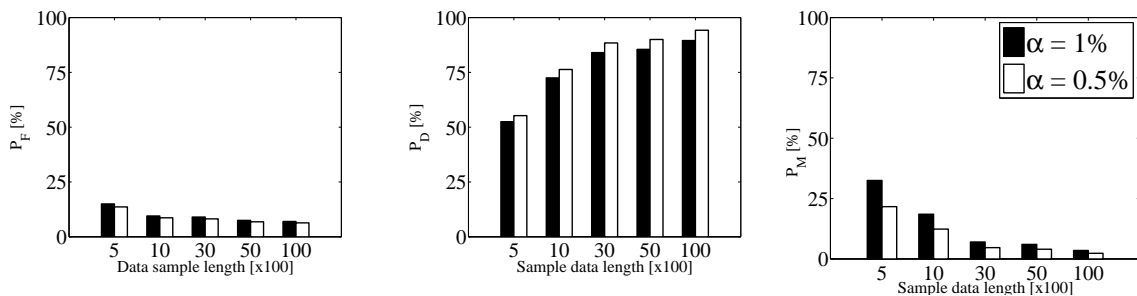
Figure 4.3: Performance of the channel order detection algorithm in a noiseless static channel SISO scenario with $L = 2$ over 200 Monte Carlo simulations: rates of false alarm, good detection and miss-detection.

been set to $\eta_4 = \eta_3 = 18.45$ and $\eta_4 = \eta_3 = 20.29$ respectively.

2 Order determination and identification of MISO channels

Let us now consider other signal sources sharing the same carrier frequency at the neighborhood of the receive antenna. In this section, we assume that an unknown number Q of co-channel users are located far apart from each other, hence utilizing physically different channels. In this context, the radiocommunication channel can be modeled as a MISO system and we will be interested in estimating the number of sources and identifying their respective channel parameters. The signal measured at time instant n at the output of the receive antenna, is written as:

$$x_q(n) = \sum_{\ell=0}^{L_q} h_q(\ell)s_q(n-\ell), \quad h_q(0) = 1,$$
$$y(n) = \sum_{q=1}^{Q} x_q(n) + \upsilon(n), \tag{4.35}$$

where the order of each channel q is $L_q + 1$. Assumptions A1 to A4 from section 1 still hold for each source $s_q(n)$ and their respective channel coefficients $h_q(\ell)$. Here, we further assume that input signals $s_q(n)$ are mutually (spatially and temporally) independent with non-zero kurtoses. There exists a known bound K for the orders of all the Q channels, i.e. $K > L_q + 1, \forall q \in [1, Q]$. We also assume that channel orders are not equal and, without loss of generality, we consider that $L_1 > \ldots > L_Q$, so that $L_1 = \max(L_q)$. The case of same length channels will be discussed in section 2.3.

Taking the above assumptions into account and using the multilinearity property of cumulants, the 4th-order cumulant of the output signal $y(n)$, defined in (4.2), can be expressed as

the sum of the marginal cumulant contributions of each source as:

$$c_{4,y}(\tau_1, \tau_2, \tau_3) = \sum_{q=1}^{Q} C_q(\tau_1, \tau_2, \tau_3) \tag{4.36}$$

in which $C_q(\tau_1, \tau_2, \tau_3)$ depends on the unknown user channel parameters $h_q(\ell)$ and can be written as follows [7, 3]:

$$C_q(\tau_1, \tau_2, \tau_3) = \gamma_{4,s_q} \sum_{\ell=0}^{L_q} h_q^*(\ell)h_q(\ell + \tau_1)h_q^*(\ell + \tau_2)h_q(\ell + \tau_3), \quad q \in [1, Q], \tag{4.37}$$

where $\gamma_{4,s_q} = c_{4,s_q}(0, 0, 0)$. Since we assume that $h_q(\ell) = 0, \forall \ell \notin [0, L_q]$, we have:

$$C_q(\tau_1, \tau_2, \tau_3) = 0, \quad \forall |\tau_1|, |\tau_2|, |\tau_3| > L_q. \tag{4.38}$$

Using (4.36), we can define the 4th-order output cumulant vector, as follows:

$$\mathbf{c}_k = \sum_{q=1}^{Q} \mathbf{c}_{k,q} \quad \in \mathbb{C}^{P \times 1}, \quad \text{for each } k \in [1, K], \tag{4.39}$$

so that $[\mathbf{c}_{k,q}]_p = C_q(i_p - 1, j_p - 1, k - 1), q \in [1, Q], p \in [1, P]$, with $(i_p, j_p) \in \mathcal{J}$, where the index set $\mathcal{J} = \{(i_1, j_1); \ldots; (i_P, j_P)\}$ is formed of strictly positive natural numbers $i_p, j_p \leq K$. Note from (4.38) that $\mathbf{c}_{k,q} = \mathbf{0}_P, \forall k > L_q + 1$. Hence, since we have assumed $L_1 > \ldots > L_Q$, we have:

$$\mathbf{c}_k = \begin{cases} \mathbf{0}_P, & \text{if } k > L_1 + 1 \\ \mathbf{c}_{k,1}, & \text{if } L_2 + 1 < k \leq L_1 + 1 \\ \mathbf{c}_{k,1} + \mathbf{c}_{k,2}, & \text{if } L_3 + 1 < k \leq L_2 + 1 \\ \quad \vdots & \quad \vdots \\ \mathbf{c}_{k,1} + \ldots + \mathbf{c}_{k,Q}, & \text{if } k \leq L_Q + 1. \end{cases} \tag{4.40}$$

In the sequel, we describe a deflation-type approach based on the use of nested channel order detectors, in which the test statistics are formed from the residual 4th-order cumulants, after subtracting the contribution of the previously estimated channels. We will also briefly discuss the impacts of allowing the user channels to have the same order. We remark that, in the context of 2×1 MISO systems, an earlier description of this deflation-based technique has been introduced in [147], using a different order detection technique, with test statistics obtained from the eigenvalue decomposition of a cumulant matrix.

2.1   Nested MISO order-detectors and blind identification

Equation (4.40) shows that, for an appropriate range of k, the output cumulant vector contains information on the the longest channel exclusively. More precisely, if $k = L_1 + 1$, $\mathbf{c}_k$ can be

written only from the coefficients $h_1(\ell)$. As a result, if we can determine the order $L_1 + 1$ of the longest channel, then $\hat{\mathbf{c}}_{L+1,1} = \hat{\mathbf{c}}_{L+1}$ and we can obtain the channel coefficients associated with the corresponding source. After that, using the estimated channel coefficients, the marginal cumulants associated with that particular source can be reconstructed approximately for all $k \in [1, L_1 + 1]$. To this end, kurtosis estimation is also needed. Then, by subtracting the contribution $\hat{\mathbf{c}}_{k,1}$ from $\hat{\mathbf{c}}_k$, we get an identical situation with $Q - 1$ sources and the same reasoning can be successively applied until the Q channels are identified. The algorithm is stopped when no residual information remains in the estimated output cumulant vector.

The procedure suggested above summarizes the idea behind the proposed nested MISO detectors. The following three main steps are repeated for each user $q \in [1, Q]$:

1. Channel order detection: determine $L_q + 1$;

2. Blind channel identification: estimate channel coefficients $\hat{h}_q(\ell)$, $\ell \in [0, L_q]$;

3. Estimation of marginal cumulants: reconstruct $\hat{\mathbf{c}}_{k,q}$ for all $k \in [1, L_q + 1]$ using the estimated channel coefficients.

Before proceeding to user $q + 1$, the marginal contribution of user q is subtracted from the estimated output cumulant vector $\hat{\mathbf{c}}_k$. Using the channel-order detection method proposed in section 1, we can solve the problem of determining the channel order $L_q + 1$ by formulating a multiple hypothesis test based on the variables ρ_k that are computed from the estimated MISO channel output cumulant vectors, $\hat{\mathbf{c}}_k$. Details about these test variables (step 1) are postponed to section 2.2. In the sequel, we discuss the matter of estimating the channel coefficients $\hat{h}_q(\ell)$ (step 2) using singular value decomposition (SVD) techniques. Afterwards, we will show that the reconstruction of the marginal cumulant vectors $\hat{\mathbf{c}}_{k,q}$ for all $k \in [1, L_q + 1]$ (step 3) is straightforward.

MISO channel parameter estimation

A. Using a rank-1 approximation

We know from (4.37) that:

$$C_q(i_p - 1, j_p - 1, L_q) = \gamma_{4,s_q} h_q^*(0)\, h_q(i_p - 1)\, h_q^*(j_p - 1)\, h_q(L_q), \quad q \in [1, Q]. \qquad (4.41)$$

and hence

$$\mathbf{c}_{L_q+1,q} = \gamma_{4,s_q} h_q^*(0) h_q(L_q) \mathbf{g}^{(q)}, \qquad (4.42)$$

where $[\mathbf{g}^{(q)}]_p = h_q(i_p - 1)h_q^*(j_p - 1)$, $p \in [1, P]$. In order to recover the channel parameters, we need to impose some minimal conditions on the index set $\mathcal{J}$ of the cumulants utilized by

the algorithm. Simple conditions ensuring correct parameter estimation can be as follows:

$$\begin{cases} i_p = 1, & \forall\, p \in [1, P] \\ j_p = p, & \forall\, p \in [1, P], \ P = K. \end{cases} \tag{4.43}$$

Satisfying the above conditions, we get $[\mathbf{g}^{(q)}]_p = h_q(0)h_q^*(p-1)$ and (4.42) becomes $\mathbf{c}_{L_q+1,q} = \gamma_{4,s_q}|h_q(0)|^2 h_q(L_q)\mathbf{h}^{(q)}$, with $\mathbf{h}^{(q)} = [h_q(0), \ldots, h_q(L_q), 0, \ldots, 0]^{\mathsf{T}} \in \mathbb{C}^{K \times 1}$. Thus we can construct the matrix $\mathbf{C}_q \in \mathbb{C}^{K \times K}$, as follows:

$$\mathbf{C}_q = \mathbf{c}_{L_q+1,q}\mathbf{c}_{L_q+1,q}^{\mathsf{H}} = \gamma_{4,s_q}^2|h_q(0)|^4|h_q(L_q)|^2 \mathbf{h}^{(q)}\mathbf{h}^{(q)\mathsf{H}}, \tag{4.44}$$

which is clearly a rank-1 matrix. Assuming we know the channel-order $L_q + 1$, we can estimate the vector $\mathbf{h}^{(q)} \in \mathbb{C}^{K \times 1}$, up to a complex scaling factor, by computing the eigenvector associated with the largest eigenvalue of $\mathbf{C}_q$. The constraint $h(0) = 1$ imposed in (4.35) helps us to avoid the trivial solution and solve the intrinsic scaling ambiguity. In practice, we only need to use the first $L_q + 1$ elements of $\hat{\mathbf{c}}_{L_q+1,q}$ to compose $\hat{\mathbf{C}}_q$ in order to avoid the zero-padding at the estimated channel tail.

By allowing for an increased set of 4th-order cumulants, we can improve the quality of the channel parameter estimates. It can be particularly interesting to expand the index set $\mathcal{J}$ to the following:

$$\mathcal{J} = \left\{ \ (1,1)\ldots(1,K) \quad \ldots \quad (K,1)\ldots(K,K) \ \right\} \tag{4.45}$$

so that $P = K^2$. Using the above index set, we get the following from (4.42):

$$\mathbf{C}_q = unvec\left(\mathbf{c}_{L_q+1,q}, K\right) = \gamma_{4,s_q}h_q^*(0)h_q(L_q)\mathbf{G}^{(q)}, \tag{4.46}$$

with $\mathbf{G}^{(q)} = unvec\left(\mathbf{g}^{(q)}, K\right) = \mathbf{h}^{(q)}\mathbf{h}^{(q)\mathsf{H}}$, where the notation $unvec(\mathbf{x}, n)$ stands for the *unvectorization* operator, which builds a matrix $\mathbf{X} \in \mathbb{C}^{m \times n}$ from the vector argument $\mathbf{x} \in \mathbb{C}^{mn \times 1}$. As a result, we can still use the rank-1 approximation solution by taking the singular vector associated with the largest singular value of $\mathbf{C}_q$.

B. *Using an optimal solution in the total least squares sense*

Noticing that $[\mathbf{c}_{L_q+1,q}]_{i_p j_p} = C_q(i_p - 1, j_p - 1, L_q)$, another way to recover the channel coefficients consists in solving the following linear system of equations, obtained from (4.41):

$$C_q(v-1, u-1, L_q)\,h_q(w-1) - C_q(w-1, u-1, L_q)\,h_q(v-1) = 0, \tag{4.47}$$

$$\text{with} \quad 1 \leq v < w \leq L_q + 1, \quad \text{and} \quad 1 \leq u \leq L_q + 1. \tag{4.48}$$

System (4.47) yields a set of up to $L_q(L_q + 1)^2/2$ equations with $L_q + 1$ unknowns, which can be rewritten in a matrix form as follows:

$$\mathbf{T}_q\mathbf{h}_q = \mathbf{0}, \tag{4.49}$$

where $\mathbf{0}$ is an all-zero vector, $\mathbf{h}_q = [h_q(0), \ldots, h_q(L_q)]^\mathsf{T}$ and $\mathbf{T}_q$ is built from the output cumulants in such a way that equation (4.49) is equivalent to the linear system (4.47). Matrix $\mathbf{T}_q$ has $L_q + 1$ columns and up to $L_q(L_q + 1)^2/2$ rows so that each row has only two nonzero elements given by $[\mathbf{T}_q]_{rv} = -C(w-1, u-1, L_q)$ and $[\mathbf{T}_q]_{rw} = C(v-1, u-1, L_q)$, where $r = (L_q + 1)(v + w - 3) + u$ is the row number satisfying the general rule (4.48). Since we need to have at least $L_q + 1$ equations, we must impose some minimal conditions on the index set $\mathcal{J}$, always satisfying (4.48), such as the following:

$$
\begin{cases}
i_p = p, & \forall\, p \in [1, P], \ P = K, \\
j_p = 1, & \forall\, p \in [1, P].
\end{cases}
\tag{4.50}
$$

A solution to (4.49) is obtained by computing the right singular vector of $\mathbf{T}_q$ associated with its smallest singular value. The expanded index set (4.45) can also be used to improve this solution, which has been originally proposed in [7] and is shown to be optimal in the total least squares (TLS) sense.

Reconstructing marginal cumulants from estimated channel coefficients

Rewriting (4.40) for $k = L_q + 1$, $q \in [1, Q]$, we have:

$$
\mathbf{c}_k =
\begin{cases}
\mathbf{c}_{L_1+1,1} & \text{for } k = L_1 + 1 \\
\mathbf{c}_{L_2+1,1} + \mathbf{c}_{L_2+1,2} & \text{for } k = L_2 + 1 \\
\quad \vdots & \quad \vdots \\
\mathbf{c}_{L_Q+1,1} + \ldots + \mathbf{c}_{L_Q+1,Q} & \text{for } k = L_Q + 1.
\end{cases}
\tag{4.51}
$$

so that we can estimate the marginal cumulant contribution of source q as follows:

$$
\hat{\mathbf{c}}_{L_q+1,q} = \hat{\mathbf{c}}_{L_q+1} - \sum_{i=1}^{q-1} \bar{\mathbf{c}}_{L_q+1,i}.
\tag{4.52}
$$

where $\bar{\mathbf{c}}_{L_q+1,i}$ are the reconstructed cumulant vectors obtained from (4.37) and (4.39) using the previously estimated coefficient vectors $\hat{\mathbf{h}}_i$, $i \in [1, q-1]$. To achieve this step, we need to estimate the kurtosis of source q. Using the index set defined in (4.43) for instance, this can be done as follows:

$$
\hat{\gamma}_{4,s_q} = \frac{1}{P} \sum_{p=1}^{P} \frac{[\hat{\mathbf{c}}_{\hat{L}_q+1}]_p}{\hat{h}_q^*(j_p - 1)\hat{h}_q(k-1)},
\tag{4.53}
$$

where $[\hat{\mathbf{c}}_{\hat{L}_q+1}]_p = \hat{c}_{4,y}(i_p - 1, j_p - 1, \hat{L}_q)$.

2.2 Test statistics for MISO order-detection

To start with the MISO channel detection procedure proposed above, we must first determine the order of each user channel, since all the subsequent steps depend on this parameter. Here, we will treat this problem as an FIR channel order selection problem and show that we can use the method proposed in section 1 of this chapter by computing the test variables from the estimated cumulant vectors obtained at the output of the MISO channel.

Recalling definition (4.21), we can compute the test variable ρ_k from the real cumulant vector $\hat{\mathbf{z}}_k$. The structure of vector $\hat{\mathbf{z}}_k$ is given in (4.10), consisting of the real and imaginary parts of the vector $\hat{\mathbf{c}}_k$, which is formed of the estimated 4th-order output cumulants of the MISO channel, such as defined in (4.39). Thus, from (4.40), we know that $\hat{\mathbf{z}}_{k,1} = [\text{Re}(\hat{\mathbf{c}}_{k,1})^\mathsf{T} \, \text{Im}(\hat{\mathbf{c}}_{k,1})^\mathsf{T}]^\mathsf{T}$ and $\hat{\mathbf{z}}_k = \hat{\mathbf{z}}_{k,1}$ for $k = L_1 + 1$. In addition, since $\mathbf{z}_k = \mathbf{0}_{d_k}$ for $k > L_1 + 1$, we can use (4.21) to construct our test variable from the vector $\hat{\mathbf{z}}_k$, so that ρ_k follows a Chi-square distribution with $d_k = 2P - m_k$ degrees of freedom, for $k > L_1 + 1$. As before, m_k stands for the number of purely real-valued elements in $\hat{\mathbf{c}}_k$, and is given by the number of entries satisfying the conditions in (4.6). On the other hand, as long as $\mathbf{z}_k$ is nonzero for $k = L_1 + 1$, the variable ρ_{L_1+1} has a *non-central* Chi-square distribution with parameter $\lambda_{L_1+1} = \mathbf{z}_{L_1+1,1}^\mathsf{T} \mathbf{\Sigma}_{L_1+1}^{-1} \mathbf{z}_{L_1+1,1}$, which is related to the mean of the test variable as follows: $\mathbb{E}\{\rho_{L_1+1}\} = \lambda_{L_1+1} + d_{L_1+1}$. We can hence denote:

$$\rho_k = \hat{\mathbf{z}}_k^\mathsf{T} \hat{\mathbf{\Sigma}}_k^{-1} \hat{\mathbf{z}}_k \begin{cases} \sim \chi^2_{(d_k)}, & k > L_1 + 1, \ N \to \infty \\ \sim {}_{nC}\chi^2_{(d_k)}(\lambda_k), & k \leq L_1 + 1, \ N \to \infty. \end{cases} \tag{4.54}$$

Using (4.11), we can deduce $\hat{\mathbf{\Sigma}}_k$ for $k \geq L_1 + 1$ from the estimated circular and non-circular covariance matrices, given in (4.8) and (4.9), respectively.

Equation (4.54) shows that the test statistic ρ_k enables us to correctly detect the order of the longest channel associated with the MISO mixture. After determining the order $L_1 + 1$, we can use one of the techniques described in section 2.1 to estimate the channel coefficients $\hat{h}_1(\ell)$. Then, by replacing the estimated coefficients in (4.37) and (4.39), we approximately reconstruct the vectors $\bar{\mathbf{c}}_{k,1}$ for all $k \in [1, L_1 + 1]$. Subtracting $\bar{\mathbf{c}}_{k,1}$ from $\hat{\mathbf{c}}_k$ allows for the computation of new test statistics leading to the detection of the next longest channel. This deflation principle has been discussed in section 2.1 and allows us to successively extract the sources from the MISO mixture. The algorithm continues until no residual information is detected in the remaining cumulant vector. The nested MISO channel detection procedure is summarized in Table 4.1.

2.3 MISO channels with identical lengths

Throughout section 2, we have assumed that the channel orders are strictly different for all the Q users. In the case where at least two users q_1 and q_2 have channels with the same order, $L_{q_1} = L_{q_2} = L$, the proposed MISO channel detection method does not fully apply because we cannot separate the marginal cumulants of these sources. Note that, in this case, the residual

Table 4.1: Nested MISO channel detection procedure

1. Estimate the 4th-order cumulants from the MISO channel output and compose the real cumulant vectors $\hat{\mathbf{z}}_k$, $k \in [1, K]$;

2. Initialize $q = 1$ and $k = K$;

3. Compute the test variables ρ_k, defined in (4.21) using $\hat{\mathbf{z}}_k$ and $\hat{\mathbf{\Sigma}}_k$; determine decision thresholds using the procedure described in section 1

4. Run the order-detection algorithm by testing the hypotheses:

   if $\rho_k < \eta_k$ then take $k \leftarrow k - 1$ and start over from step 3.

   if $\rho_k \geq \eta_k$ then $\hat{L}_q = k - 1$. Go to step 5.

   if $\hat{L}_q = 0$, terminate the procedure;

5. Compute $\hat{h}_q(\ell)$, $\ell \in [0, L_q]$, from $\hat{\mathbf{c}}_{L_q+1,q}$ using one of the blind channel identification techniques described in section 2.1;

6. From (4.37) and (4.39), compute the entries of $\bar{\mathbf{c}}_{k,q}$, for all $k \in [1, L_q + 1]$, using $\hat{h}_q(\ell)$ and $\hat{\gamma}_{4,s_q}$, the latter one obtained from (4.53);

7. Update the output cumulant vector as $\hat{\mathbf{c}}_k \leftarrow \hat{\mathbf{c}}_k - \bar{\mathbf{c}}_{k,q}$;

8. Update $q \leftarrow q + 1$ and repeat steps 3 to 8 until hypothesis $\rho_k < \eta_k$ is not rejected in step 4.

information in $\hat{\mathbf{c}}_k$ at $k = L + 1$ still contains a mixture of both channels, i.e. for $q \in \{q_1, q_2\}$, the residual cumulant vector at $k = L + 1$ is written as $\hat{\mathbf{c}}_{L+1} = \hat{\mathbf{c}}_{L+1,q_1} + \hat{\mathbf{c}}_{L+1,q_2}$. As long as the marginal cumulant vectors are not separable, it is not possible to estimate the individual channel coefficients and the nested detection procedure is terminated. As a conclusion, the deflation principle cannot fully apply when there are channels with the same length. However, we can still determine the channel orders greater than $L+1$ as well as their respective coefficients.

Actually, the proposed algorithm (Table 4.1) is able to run normally for all the users having channel order greater than $L + 1$. We can detect their orders, estimate their coefficients and, subtracting their contribution from the 4th-order output cumulant vector, the remaining information contained in $\hat{\mathbf{c}}_k$ can be detected by the proposed channel order detection procedure. Hence, when $q = q_1$ (or $q = q_2$), the hypothesis $H_0(k)$ ($\rho_k < \eta_k$) in step 4 should be rejected for $k = L + 1$, enabling us to determine the order of the channels with same length. In the sequel of the procedure, the blind channel estimation algorithms are obviously affected, because they assume that the estimated cumulant vector only contains information about one single channel, which is not true in this case. Thus, the estimated parameters do not correspond to the actual channel coefficients, and the recomposition of the marginal cumulants in step 6 leads to bad results. After that, the residual information $\hat{\mathbf{c}}_k - \bar{\mathbf{c}}_{k,q}$ at $k = L + 1$ should be different of zero, in spite of (4.51).

As a result, when the algorithm runs the next iteration (for $q = q_1 - 1$ or $q = q_2 - 1$), the hypothesis test should be able to detect another channel at $k = L + 1$, since the residue in $\hat{\mathbf{z}}_{L+1}$ is not zero. One can thus conclude that there are at least two channels with same order. In this case, the coefficients of the last channel estimated should be disregarded and the nested detection procedure terminated. At the present time, we are not able to identify the number of channels having the same length $L + 1$ (we can only say that there are at least two). Also, using the deflation approach herein described, we cannot estimate the coefficients of any of the channels with order equal to or lower than $L + 1$.

3 Simulation results

In this section, we will be interested in assessing the overall performance of the nested detectors described in Table 4.1. In the sequel, computer simulation results are shown to illustrate the combined procedure for MISO channel identification in terms of both, channel order detection and blind channel estimation. In order to evaluate the method, we make use of the two following criteria, which are defined for each signal source $q \in [1, Q]$:

- *i.* The empirical probability of detection: the ratio between the number of Monte Carlo simulations in which the detector results are consistent ($\hat{L}_q = L_q$) and the total number R of realizations of the experiment;

- *ii* The normalized mean squared error (NMSE) of the consistent results: computed for each channel q as follows:

$$\texttt{NMSE}(q) = \frac{1}{R_q} \sum_{r=1}^{R_q} \frac{\left\| \widehat{\mathbf{h}}_q^{\langle r \rangle} - \mathbf{h}_q \right\|^2}{\|\mathbf{h}_q\|^2}, \tag{4.55}$$

where R_q is the number of consistent results obtained for source q over R realizations and $\widehat{\mathbf{h}}_q^{\langle r \rangle}$ is the estimated channel vector associated with source q, obtained after the (consistent) realization r.

Both blind channel identification algorithms described in section 2.1 have been implemented (rank-1 approximation and TLS) using the index set defined in (4.45). However, since their results are very close, we omitted here the curves obtained from the rank-1 approximation solution.

Nested algorithms for order detection and blind estimation of MISO channels

Let us consider a frequency-selective MISO communication channel with $Q = 2$ users and one single receive antenna. A static multipath propagation scenario is assumed, which induces a delay spread of the order of 3 symbol periods on the channel of source 1 ($L_1 = 2$) and of the order of 2 symbol periods on the channel of source 2 ($L_2 = 1$). Channel coefficients have been randomly

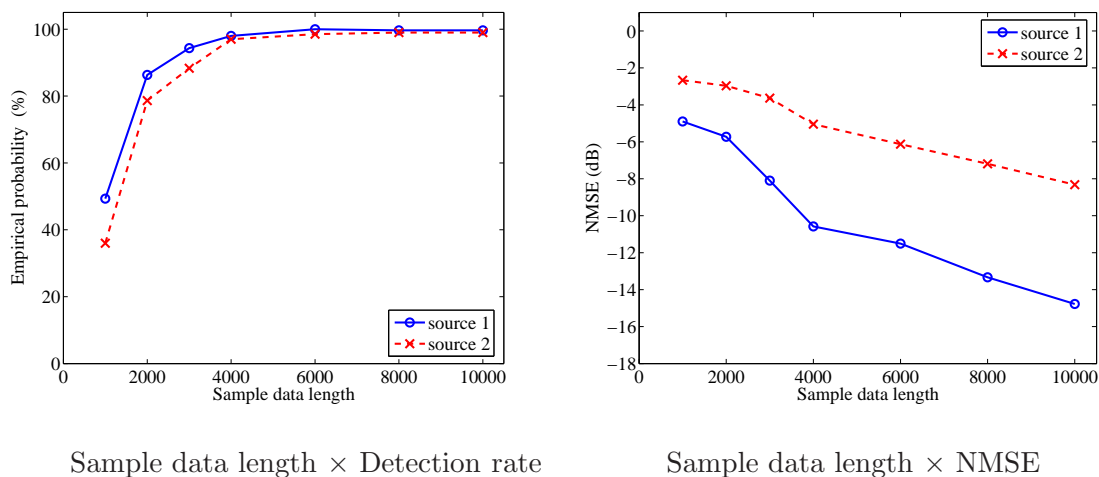Sample data length × Detection rate   Sample data length × NMSE

Figure 4.4: Blind MISO channel order detection and identification: detection success rate and NMSE as a function of the sample data length, with SNR=40dB ($L_1 = 3$, $L_2 = 1$ and $K = 4$).

generated from a continuous complex Gaussian distribution. The figures in the sequel have been obtained with the following coefficient vectors: $\mathbf{h}_1 = [1.0, \ 1.35 - 0.57\jmath, \ -0.72 + 1.49\jmath]^\mathsf{T}$; $\mathbf{h}_2 = [1.0, \ -1.14 + 0.23\jmath]^\mathsf{T}$. A known upper bound of $K = 4$ is assumed for both channel orders. We simulated the nested detection-estimation procedure presented in Table 4.1 and the results have been averaged over 300 received data blocks, with N symbols each.

In a first experiment, we fixed the SNR value in 40dB and considered values of the output sample data length (N) varying from 10^3 to 10^4. At the left-hand side of fig. 4.4, we notice a poor detection performance for both sources when the sample data length is smaller than $N = 3000$. At a glance, the results in fig. 4.4 indicate that the overall channel identification performance increases as the number of output data symbols becomes greater. This is an expected result due to the fact that the cumulant estimators are biased, but the bias is asymptotically zero. It also suggests that the statistical variance of the cumulant estimators depends on all the user channels, since it depends on the output cumulants. Increased cumulant estimation variance implies worse channel parameter estimation, specially for short data lengths.

In our second simulation, we used a sample data length of $N = 10^4$ for SNR values ranging from 5 to 40dB. As we can see in fig. 4.5, the nested detectors are quite robust with respect to additive Gaussian noise, at moderate an high SNR levels. For low SNR levels the channel estimation performance is significantly degraded. Notice that, although our order detection technique is based on the 4th-order cumulants, the computation of ρ_k involves the estimation of the covariance matrix $\hat{\boldsymbol{\Sigma}}_k$, which requires 2nd-order moments estimates. The overall detection performance should be therefore expected to suffer with increasing levels of Gaussian noise, as illustrated in fig. 4.5. In addition, we observe in both figures, 4.4 and 4.5, a severe performance loss between sources 1 and 2, suggesting that the nested MISO channel detector may suffer from
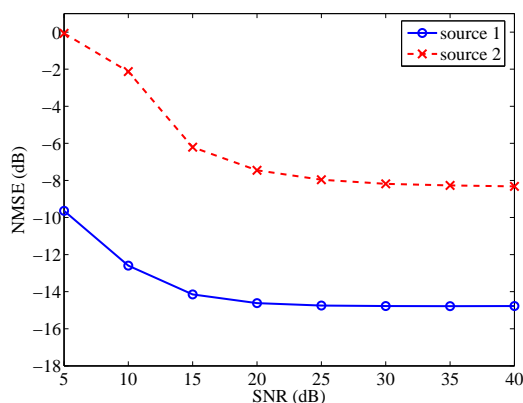
Figure 4.5: SNR $\times$ NMSE: blind MISO channel identification with $N = 10000$ symbols ($L_1 = 3$, $L_2 = 1$ and $K = 4$).

error propagation, harming the identification of shorter channels. This behavior can be explained from the fact that shorter channels are estimated from the residual 4th-order cumulants, i.e. the information remaining after the extraction of the reconstructed cumulants of the longer channels, which are in turn estimated and hence susceptible to errors.

4   Summary

This chapter treated the problem of determining the order of an FIR channel in a radiocommunications context. We have proposed a channel order detection algorithm based on HOS hypothesis testing. A Chi-square test variable has been introduced along with a discussion about the choice of decision thresholds and detection performance in terms of success rate. The proposed detection algorithm relies on some properties of the 4th-order output cumulants.

Exploiting the order selection method proposed in the first part of this chapter, we have also proposed a complete combined detection-estimation procedure in the context of a frequency-selective MISO communication system. Still using the 4th-order cumulant properties, this new algorithm successively detects the signal sources, determines the orders of their individual transmission channels and estimates the associated channel coefficients. The so-called nested detectors process the sources one after the other, extracting from the output cumulant matrix the estimated contributions of the previously identified channels and testing for the presence of shorter and shorter channels.

We have shown that the test variable defined in the first part of the chapter (in the context of a single FIR channel) can also be useful in the case of MISO mixtures. The nested detectors treat the order-detection problem separately for each user channel by applying the proposed order detection algorithm on the residual cumulant vector. Computer simulations show good

results obtained for channel order detection and identification.

The case of channels with the same order still needs further investigation. Future works on this subject include an implementation of the nested MISO detectors to the case of a multiple-input multiple-output (MIMO) communications channel. Such an approach is simpler than the one presented here and should allow us to treat the case of same length channels, including the overdetermined as well as the underdetermined mixtures.

General Conclusions and future work

Modern telecommunication systems require increased transmission rates. The correct recovery of the transmitted information is among the greatest challenges of the upcoming era of wireless and mobile communications so that strong engineering efforts are currently spent on this active research topic. This thesis has been mainly devoted to the problem of blind channel estimation in the context of radiocommunication systems. In other words, we have aimed in obtaining a complete representation of the communication channel using only the data sequences available at the channel output. Making use of 4th-order cumulants, we have been able to exploit their highly symmetrical structure using a tensor formalism. High-order cumulants also have the advantage of eliminating the effects of an additive Gaussian noise (with unknown correlation function) corrupting the output signals. Our propositions throughout the thesis have been guided by the underlying idea of avoiding multiple optimization stages or redundant computational operations.

In which follows, we summarize the most important points discussed in each chapter of this thesis, highlighting the main conclusions and contributions. After that, we present a list of possible topics that could lead to the continuation of some of our propositions.

Summary and conclusions

In **Chapter 1**, we have presented an overview of the main definitions, relationships and properties of high-order statistics (HOS). We have also briefly discussed the simultaneous matrix diagonalization issue using linear algebraic tools. Then, using a generalized formulation for tensors of any order, we have surveyed the basic notions behind the Parallel Factor (Parafac) tensor decomposition. Concerning the estimation of the tensor components, we have introduced an extended version of the alternating least squares (ALS) algorithm, generalizing it to the case of a Pth-order tensor.

In **Chapter 2**, we have addressed the problem of blind channel identification based on a multilinear decomposition of the 4th-order output cumulant tensor. A new tensor model has been proposed for the single-input single-output (SISO) case including components with a Hankel structure. Regarding the instantaneous multiple-input multiple-output (MIMO) mixtures, the

redundancies contained in the tensor components have been exploited when estimating the channel coefficients. In both SISO and MIMO cases, we proposed new blind identification algorithms based on the solution of the cumulant tensor decomposition problem by means of a single-step least squares (SS-LS) minimization. The multilinear approach fully exploits the structural properties of the cumulant tensor and allows to skip any kind of pre-processing. Indeed, the SS-LS approach has been an original contribution of this thesis, based on one sole iterative optimization, with no need for intermediate steps such as pre-whitening.

In **Chapter 3**, we have extended our blind identification solution to the case of a convolutive MIMO channel, thus unifying the proposed tensor models into a generalized space-time 4th-order cumulant tensor admitting a Parafac decomposition with a block-Hankel structure, of which the SISO and the memoryless MIMO channels are particular cases. A SS-LS algorithm has also been derived to estimate the coefficients of the convolutive MIMO channel model, which can in fact be used to represent a multiantenna communication system characterized by specular multipath propagation. Indeed, the coefficient estimation method proposed in this chapter has been introduced in the context of a two-stage technique for estimating the parameters that describe the physical structure of a multipath channel. In this broader framework, after estimating the channel coefficients, the structure of the convolutive multipath channel has been shown to be associated with the components of a Parafac tensor formed from the channel parameters. This new tensor modeling introduced in Chapter 3 has allowed us to propose a fully blind ALS-MUSIC technique to recover the path delays, attenuations as well as the angles of departure and arrival of the transmitted and received signals, respectively.

Previously, in the first part of Chapter 3, we have also been interested in determining the source locations in the context of narrowband multiuser sensor array processing, under the far-field approximation. We have described a new technique that allows for incorporating additional virtual sensors to a linear antenna array by exploiting an asymmetric arrangement of the 4th-order cumulant tensor, involving a Khatri-Rao structure generally associated with the 6th-order cumulant tensor. A high-resolution algorithm has been proposed using the SS-LS approach to estimate the direction of arrival of the incoming signals.

At last, in **Chapter 4**, we have taken the channel order detection problem into consideration. First, we derived a Chi-square test statistic from the 4th-order output cumulants and, relying on its insensitiveness with respect to a non-linear stochastic process, we have developed a new method for FIR-SISO channel order determination. Then, using the residual 4th-order information at the output of a multiple-input single-output (MISO) channel, we have introduced a combined procedure for detecting the presence of the signal sources, determining the channel orders and estimating their coefficients. This deflationary approach treating single-output channels has been another original contribution of this thesis.

Near future work and perspectives

Beyond the contributions presented in this thesis, some questions remain open issues and need further investigation. Below, we list some topics that may be viewed as the natural continuity of our work.

- The SS-LS algorithms presented in Chapter 2 belong to the class of iterative techniques, which is also the case for the joint-diagonalization based methods described in sections 2.3.1 and 2.5.1. A deeper performance comparison between these algorithms should answer to some questions about the ultimate potentialities of each technique in terms of the tradeoff between capacity and estimation performance. A formal analysis of their respective computational burden should also help to bring out their practical applicability.

- The joint-diagonalization-based methods described in sections 2.3.1 and 2.5.1 do not cope with the case of convolutive MIMO channels; extending these methods or, maybe, proposing a combination of them to treat that case is envisaged for the near future. Since we focus on telecommunication contexts, it should be interesting to evaluate the impact of this methods in terms of channel equalization as well.

- The proposed SS-LS approach has been based on the use of 4th-order output cumulants because this is the lower order allowing us to keep the advantages of HOS while eliminating the effects of an additive Gaussian noise with unknown correlation function. Notice, however, that we have only used the properties of linearity and additivity of cumulants. Moreover, symmetry relationships also apply to higher-order cumulants. As a matter of fact, it seems that the SS-LS technique for channel estimation can be generalized to the case of any (even) order cumulants. We think that such an extension deserves more attention in the near future and should yield some questions about the multiple possible choices to unfold higher-order cumulant tensors.

- In the sequel of the preceding topic, it is straightforward to imagine that, using cumulants of order higher than fourth, we should be able to create different array configurations in order to further expand the number of virtual sensors achieved using the method proposed in section 3.2.

- The proposition of the nested MISO channel detector in Chapter 4 has been based on the assumption that all the individual channels have different orders. When this condition is not satisfied, the technique does not fully apply. therefore, further investigation is due in the case of same order channels. In order to treat that case, the inclusion of additional receive antennas can be envisaged, so yielding MIMO channels, which should be simpler to treat since the information can be observed from a supplementary spatial domain.

Appendices

Appendix A

Eigenvalue decomposition algorithms

R educing a square matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ to a triangular form is an interesting way to compute its eigenvalues. Several techniques can be used to implement this strategy, including Householder, Givens and Gram-Schmidt transformations [50, 49]. In the special case of normal matrices, it can be shown that the eigenvalue decomposition reduces to a diagonalization problem, which can be implemented by means of successive applications of unitary similarity transformations. In appendix A.1 below, we describe a triangular factorization algorithm that solves the eigenvalue problem by means of a Schur decomposition; then, in appendix A.2, we present an overview of the main ideas behind diagonalization techniques for normal matrices, including the classical Jacobi algorithm and its extended version for approximative simultaneous diagonalization.

A.1   Eigenvalue computation via Schur decomposition

From the Schur Unitary Triangularization Theorem, we know that there exists a unitary $n \times n$ matrix $\mathbf{Q}$ such that $\mathbf{Q}^{\mathsf{H}} \mathbf{A} \mathbf{Q} = \mathbf{L}^{\mathsf{T}}$ is upper triangular. Using the Schur decomposition, we can compute the triangularizing factor $\mathbf{Q}$ by means of successive unitary similarity transformations.

Let us consider a initial transformation $\mathbf{Q}_0 \in \mathbb{C}^{n \times r}, 1 \leq r \leq n$, with pairwise orthonormal columns. Consider the sequence of matrices $\{\mathbf{Q}_k\}$ generated by means of a certain number of iterations as follows

$$\begin{cases} \mathbf{B}_k = \mathbf{A} \mathbf{Q}_{k-1} \\ \mathbf{Q}_k \mathbf{L}_k^{\mathsf{T}} \leftarrow \mathbf{B}_k \qquad \text{for all } k \geq 1, \end{cases} \qquad (A.1)$$

where the second iteration step is the QR factorization of the auxiliary matrix $\mathbf{B}_k$, so that $\mathbf{Q}_k$ is unitary and $\mathbf{L}_k^{\mathsf{T}}$ is upper triangular. Then, it is possible to show [50, pp. 333] that, under certain assumptions on the initial iterate $\mathbf{Q}_0$, the column space of $\mathbf{Q}_k$ "converges" to the column space of $\mathbf{V}^{(r)} = [\mathbf{V}_{\cdot 1} \ldots \mathbf{V}_{\cdot r}]$, in which $\mathbf{V}^{\mathsf{H}} \mathbf{A} \mathbf{V} = \mathbf{L}^{\mathsf{T}}$ is a Schur decomposition of $\mathbf{A}$ and $\mathbf{V}^{(r)}$ corresponds to the first r columns of $\mathbf{V}$. Making $r = n$, this implies that $\lim_{k \to \infty} \mathbf{Q}_k = \mathbf{V}$,

i.e. under reasonable conditions[1], the iterations (A.1) actually compute a Schur decomposition of $\mathbf{A}$. Therefore, defining

$$\mathbf{T}_k = \mathbf{Q}_k^{\mathsf{H}} \mathbf{A} \mathbf{Q}_k, \tag{A.2}$$

we note that as k approaches infinity $\mathbf{T}_k$ goes upper triangular, i.e. $\lim_{k \to \infty} \mathbf{T}_k = \mathbf{L}^{\mathsf{T}}$.

In order to compute the factor $\mathbf{T}_k$ from the preceding iteration $\mathbf{T}_{k-1}$, we first write $\mathbf{T}_{k-1} = \mathbf{Q}_{k-1}^{\mathsf{H}} \mathbf{A} \mathbf{Q}_{k-1}$ and then, using $\mathbf{A} \mathbf{Q}_{k-1} = \mathbf{Q}_k \mathbf{L}_k^{\mathsf{T}}$ from (A.1), we note that

$$\mathbf{T}_{k-1} = \left(\mathbf{Q}_{k-1}^{\mathsf{H}} \mathbf{Q}_k \right) \mathbf{L}_k^{\mathsf{T}} \tag{A.3}$$

is the QR factorization of $\mathbf{T}_{k-1}$ with unitary factor $\bar{\mathbf{Q}}_k = \mathbf{Q}_{k-1}^{\mathsf{H}} \mathbf{Q}_k$. However, from (A.2) we have $\mathbf{T}_k = \mathbf{Q}_k^{\mathsf{H}} \mathbf{A} (\mathbf{Q}_{k-1} \mathbf{Q}_{k-1}^{\mathsf{H}}) \mathbf{Q}_k$ and, since $\mathbf{L}_k^{\mathsf{T}} = \mathbf{Q}_k^{\mathsf{H}} \mathbf{A} \mathbf{Q}_{k-1}$, we can finally write

$$\mathbf{T}_k = \mathbf{L}_k^{\mathsf{T}} \left(\mathbf{Q}_{k-1}^{\mathsf{H}} \mathbf{Q}_k \right) = \mathbf{L}_k^{\mathsf{T}} \bar{\mathbf{Q}}_k, \tag{A.4}$$

which is exactly the product of the factors obtained from the QR decomposition of $\mathbf{T}_{k-1}$ in the reverse order. Now, the following QR iteration is straightforward

$$\begin{cases} \bar{\mathbf{Q}}_k \mathbf{L}_k^{\mathsf{T}} \leftarrow \mathbf{A}_{k-1} \\ \mathbf{A}_k = \mathbf{L}_k^{\mathsf{T}} \bar{\mathbf{Q}}_k \qquad \text{for all } k \geq 1. \end{cases} \tag{A.5}$$

Initializing the above iterative loop with $\mathbf{A}_0 = \mathbf{A}$, we notice that $\mathbf{A}_k = \bar{\mathbf{Q}}_k^{\mathsf{H}} \dots \bar{\mathbf{Q}}_1^{\mathsf{H}} \mathbf{A} \bar{\mathbf{Q}}_1 \dots \bar{\mathbf{Q}}_k$. We can thus conclude that $\mathbf{A}_k$ is unitarily similar to $\mathbf{A}$. Moreover, $\mathbf{A}_k$ is a Schur decomposition of $\mathbf{A}$ and hence upper triangular. So, its main diagonal entries are the eigenvalues of $\mathbf{A}$ and the matrix product $\bar{\mathbf{Q}}_1 \dots \bar{\mathbf{Q}}_k$ is an orthonormal basis of eigenvectors.

A.2   Jacobi algorithms for matrix diagonalization

Orthogonal transformations play a central role in the least squares solutions of overdetermined systems as well as in eigenvalue computation. Jacobi methods are very interesting for solving the symmetric eigenvalue problem because they present high parallelism features allowing for efficient implementations in modern computer architectures. Good rounding error properties, sometimes better than the QR algorithm, are also reported in the literature [53].

In this context, *Givens rotations* are a well-known tool for introducing zeros in a matrix, in a controlled manner. Givens rotations are orthogonal rank-2 transformations denoted by $\mathbf{G}_{pq}(c, s)$ and defined in function of the complex parameters c and s as follows

$$\mathbf{G}_{pq}(c, s) = \mathbf{I}_n + (c - 1)\mathbf{e}_p^{(n)} \mathbf{e}_p^{(n)\mathsf{T}} - s\,\mathbf{e}_p^{(n)} \mathbf{e}_q^{(n)\mathsf{T}} + s\,\mathbf{e}_q^{(n)} \mathbf{e}_p^{(n)\mathsf{T}} + (c - 1)\mathbf{e}_q^{(n)} \mathbf{e}_q^{(n)\mathsf{T}}, \tag{A.6}$$

---

[1] For $r = n$, the column space of the initial iterate $\mathbf{Q}_0$ must not be orthogonal to the column space of $\mathbf{V}$.

where $|c|^2 + |s|^2 = 1$ and the vectors $\mathbf{e}_p^{(n)}$ and $\mathbf{e}_q^{(n)}$ stand for the pth and qth canonical basis vectors of $\mathbb{R}^n$, respectively. Notice that it is reasonable to parameterize (A.6) using $c = \cos\theta$ and $s = \sin\theta$, thus we can denote $\mathbf{G}_{pq}(c,s) = \mathbf{G}_{pq}(\theta)$. Clearly, a pre-multiplication by $\mathbf{G}_{pq}(\theta)$ has the effect of rotating the affected space of an amount of θ radians on the 2-dimensional subspace spanned by the basis vectors $\{\mathbf{e}_p^{(n)}, \mathbf{e}_q^{(n)}\}$, referred to as the (p, q)-plan.

For real matrices, Jacobi's diagonalization approach consists in applying successive transformations on a matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ so that $\mathbf{A} \leftarrow \mathbf{Q}^\mathsf{T} \mathbf{A} \mathbf{Q}$ in such a way that $\mathbf{A}$ is closer and closer to being diagonal. The basic step in this iterative procedure consists in choosing, at each iteration, an orthogonal matrix $\mathbf{Q}$ in the form of a plane rotation as given in (A.6), where the angle θ is chosen to systematically reduce the quantity $off(\mathbf{Q}^\mathsf{T}\mathbf{A}\mathbf{Q})$, the operator $off(\cdot)$ being defined in (1.37). The algorithm is iterated for all pairs (p, q), $1 \le p < q \le n$. Denoting $\boldsymbol{\Delta} = \mathbf{Q}^\mathsf{T}\mathbf{A}\mathbf{Q}$, we note that if $\mathbf{Q} = \mathbf{G}_{pq}(\theta)$ then the elements of matrix $\boldsymbol{\Delta}$ are the same as those of $\mathbf{A}$ except for the rows p and q as well as columns p and q. Indeed,

$$\begin{pmatrix} \delta_{pp} & \delta_{pq} \\ \delta_{qp} & \delta_{qq} \end{pmatrix} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix}^\mathsf{T} \begin{pmatrix} a_{pp} & a_{pq} \\ a_{qp} & a_{qq} \end{pmatrix} \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \tag{A.7}$$

where $a_{pq} = a_{qp}$ and $\delta_{pq} = \delta_{qp}$ (recall that $\mathbf{A}$ is assumed symmetric).

Thus, for each pair $(p, q), p \ne q$, we search for the parameters c and s that make the non-diagonal elements of $\boldsymbol{\Delta}$ go to zero, i.e.

$$\delta_{pq} = (a_{pp} - a_{qq})cs + a_{pq}(c^2 - s^2) = 0. \tag{A.8}$$

The above equation yields the following quadratic form

$$t^2 + 2\kappa t - 1 = 0, \tag{A.9}$$

where we have defined $\kappa = (a_{qq} - a_{pp})/2a_{pq}$ and used the parametrization $t = \tan\theta = s/c$ in order to obtain c and s from the conventional formulas $c = 1/\sqrt{1 + t^2}$ and $s = tc$, without resorting to trigonometrical equations. Solutions for (A.9) are of the form $t = -\kappa \pm \sqrt{1 + \kappa^2}$ and it is possible to show that the smallest of them minimizes the difference between $\mathbf{A}$ and $\boldsymbol{\Delta}$ ($|\theta| \le \pi/4$) hence improving stability.

Classical Jacobi algorithm

Next, we describe an iterative procedure for computing the orthogonal matrix $\mathbf{Q}$ so that $\boldsymbol{\Delta} = \mathbf{Q}^\mathsf{T}\mathbf{A}\mathbf{Q}$ is closer and closer to diagonal. The algorithm is stopped when a certain tolerance level $\alpha > 0$ is reached, i.e. $off(\boldsymbol{\Delta}) < \alpha \, off(\mathbf{A})$.

Algorithm A.1 (Classical Jacobi algorithm)

Initialize the algorithm with $\mathbf{\Delta} \leftarrow \mathbf{A}$ and $\mathbf{Q} \leftarrow \mathbf{I}$, determine a tolerance level $\alpha > 0$ and verify if $\mathit{off}(\mathbf{\Delta}) < \alpha \, \mathit{off}(\mathbf{A})$. While this is false, repeat the procedure below:

1. *Choose the values of p and q so that $1 \leq p < q \leq n$ and $|a_{pq}| \geq |a_{ij}|$, $\forall\, i \neq p$, $j \neq q$, $i \neq j$;*

2. *If $\delta_{pq} = 0$ then $c = 1$ and $s = 0$. Go to step 4;*

3. *If $\delta_{pq} \neq 0$, determine $\kappa = (a_{qq} - a_{pp})/2a_{pq}$ and then compute the parameters $c = 1/\sqrt{1 + t^2}$ and $s = tc$ so that $\delta_{pq} = 0$, where $t = -\kappa + \sqrt{1 + \kappa^2}$ if $\kappa \geq 0$ and $t = -\kappa - \sqrt{1 + \kappa^2}$ otherwise.*

4. *Construct the Givens rotation matrix $\mathbf{G}_{pq}(c, s)$, as indicated in (A.6).*

5. $\mathbf{\Delta} \leftarrow \mathbf{G}_{pq}(c, s)^T \mathbf{\Delta} \mathbf{G}_{pq}(c, s)$

6. $\mathbf{Q} \leftarrow \mathbf{Q}\mathbf{G}_{pq}(c, s)$

It is possible to show that Jacobi's algorithm is monotonically convergent, i.e. at each new iteration, its solutions are shown to be systematically closer to being diagonal.

The extended Jacobi algorithm

In this section, we describe an extended version of the Jacobi algorithm that computes a non-singular factor $\mathbf{Q}$ that simultaneously diagonalizes a set $\mathbb{A}$ of symmetric matrices $\mathbf{A}^{(k)} \in \mathbb{C}^{n \times n}$, $k = 1, \ldots, K$. Let us define an $n \times n$ plane rotation matrix $\mathbf{G}_{pq}(c, s)$, slightly different from that of (A.6), as follows:

$$\mathbf{G}_{pq}(c, s) = \mathbf{I} + (c - 1)\mathbf{e}_p^{(n)}\mathbf{e}_p^{(n)\mathsf{T}} - s^* \, \mathbf{e}_p^{(n)}\mathbf{e}_q^{(n)\mathsf{T}} + s \, \mathbf{e}_q^{(n)}\mathbf{e}_p^{(n)\mathsf{T}} + (c^* - 1)\mathbf{e}_q^{(n)}\mathbf{e}_q^{(n)\mathsf{T}}, \qquad \text{(A.10)}$$

with $1 \leq p < q \leq n$, where $c, s \in \mathbb{C}$ are parameterized as $c = \cos\theta$ and $s = e^{-j\phi}\sin\theta$, respectively. The problem here consists in computing the complex values c and s in order to maximize the diagonal elements of $\mathbf{G}_{pq}(c, s)\mathbf{A}^{(k)}\mathbf{G}_{pq}^{\mathsf{H}}(c, s)$ for each pair (p, q), i.e.

$$(c, s) = \arg \max_{|c|^2 + |s|^2 = 1} \mathsf{J}(\mathbf{Q}, \mathbb{A}), \qquad \text{(A.11)}$$

where $\mathsf{J}(\mathbf{Q}, \mathbb{A})$ has been defined in (1.49) with $\mathbf{Q} = \mathbf{G}_{pq}(c, s)^{\mathsf{H}}$ and $\mathbb{A} = \{A_1, \ldots, A_K\}$. Denoting $\mathbf{\Delta}_k = \mathbf{Q}^{\mathsf{H}}\mathbf{A}_k\mathbf{Q}$, we notice that

$$\begin{pmatrix} \delta_{pp}^{(k)} & \delta_{pq}^{(k)} \\ \delta_{qp}^{(k)} & \delta_{qq}^{(k)} \end{pmatrix} = \begin{pmatrix} \cos\theta & -e^{j\phi}\sin\theta \\ e^{-j\phi}\sin\theta & \cos\theta \end{pmatrix}^{\mathsf{H}} \begin{pmatrix} a_{pp}^{(k)} & a_{pq}^{(k)} \\ a_{qp}^{(k)} & a_{qq}^{(k)} \end{pmatrix} \begin{pmatrix} \cos\theta & -e^{j\phi}\sin\theta \\ e^{-j\phi}\sin\theta & \cos\theta \end{pmatrix}, \quad \text{(A.12)}$$

and

$$\|diag\,(\mathbf{\Delta}_k)\|^2 = \|diag\,(\mathbf{A}_k)\|^2 - \left|a_{pp}^{(k)}\right|^2 - \left|a_{qq}^{(k)}\right|^2 + \left|\delta_{pp}^{(k)}\right|^2 + \left|\delta_{qq}^{(k)}\right|^2. \qquad \text{(A.13)}$$

Thus, recalling that $\|\boldsymbol{\Delta}_k\|_F = \|\mathbf{A}_k\|_F$, we observe that maximizing $\mathsf{J}(\mathbf{Q}, \mathbb{A})$ is equivalent to maximizing $|\delta_{pp}^{(k)}|^2 + |\delta_{qq}^{(k)}|^2$. Furthermore, using the following ordinary identity for complex numbers a and b

$$|a + b|^2 + |a - b|^2 = 2\left(|a|^2 + |b|^2\right) \tag{A.14}$$

and, taking into account the trace preservation property of orthogonal transformations, it is straightforward to conclude that

$$\mathsf{J}(\mathbf{Q}, \mathbb{A}) = \sum_{k=1}^{K} \left|\delta_{pp}^{(k)} - \delta_{qq}^{(k)}\right|^2. \tag{A.15}$$

From (A.12), we can deduce

$$
\begin{aligned}
\delta_{pp}^{(k)} &= a_{pp}^{(k)} \cos^2\theta + a_{pq}^{(k)} e^{-j\phi} \sin\theta\cos\theta + a_{qp}^{(k)} e^{j\phi} \sin\theta\cos\theta + a_{qq}^{(k)} \sin^2\theta \\
\delta_{qq}^{(k)} &= a_{pp}^{(k)} \sin^2\theta - a_{pq}^{(k)} e^{-j\phi} \sin\theta\cos\theta - a_{qp}^{(k)} e^{j\phi} \sin\theta\cos\theta + a_{qq}^{(k)} \cos^2\theta
\end{aligned}
$$

Then, after some simple manipulations, we get K equations of the form

$$(\delta_{pp}^{(k)} - \delta_{qq}^{(k)}) = (a_{pp}^{(k)} - a_{qq}^{(k)}) \cos 2\theta + (a_{qp}^{(k)} + a_{pq}^{(k)}) \cos\phi \sin 2\theta + j(a_{qp}^{(k)} - a_{pq}^{(k)}) \sin\phi \sin 2\theta$$

which may be expressed in vector form as

$$\mathbf{p} = \mathbf{F}\mathbf{v} \tag{A.16}$$

where we have used $\mathbf{F} = [\mathbf{f}_1 \ \ldots \ \mathbf{f}_K]^{\mathsf{T}}$ and the following definitions

$$\mathbf{p} = \begin{pmatrix} \delta_{pp}^{(1)} - \delta_{qq}^{(1)} \\ \vdots \\ \delta_{pp}^{(K)} - \delta_{qq}^{(K)} \end{pmatrix} \qquad \mathbf{f}_k = \begin{pmatrix} a_{pp}^{(k)} - a_{qq}^{(k)} \\ a_{qp}^{(k)} + a_{pq}^{(k)} \\ j(a_{qp}^{(k)} - a_{pq}^{(k)}) \end{pmatrix} \qquad \mathbf{v} = \begin{pmatrix} \cos 2\theta \\ \cos\phi \sin 2\theta \\ \sin\phi \sin 2\theta \end{pmatrix}. \tag{A.17}$$

We can now rewrite (A.15) as follows

$$\mathsf{J}(\mathbf{Q}, \mathbb{A}) = \mathbf{p}^{\mathsf{H}}\mathbf{p} = \mathbf{v}^{\mathsf{T}}\mathbf{F}^{\mathsf{H}}\mathbf{F}\mathbf{v} \tag{A.18}$$

and, since $\mathbf{F}^{\mathsf{H}}\mathbf{F}$ has a Hermitian structure and $\mathbf{v}$ is a real vector, it is possible to show that

$$\mathsf{J}(\mathbf{Q}, \mathbb{A}) = \mathbf{v}^{\mathsf{T}}\mathsf{Re}\left(\mathbf{F}^{\mathsf{H}}\mathbf{F}\right)\mathbf{v}, \tag{A.19}$$

where we have used the fact that if $\mathbf{A} \in \mathbb{C}^{n \times n}$ is a Hermitian matrix, then $\mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x} = \mathbf{x}^{\mathsf{T}}\mathsf{Re}(\mathbf{A})\mathbf{x}$, with $\mathbf{x} \in \mathbb{R}^n$. The value of $\mathbf{v}$ that maximizes (A.19) is given by the eigenvector $\mathbf{v} = [v_1 \ v_2 \ v_3]^{\mathsf{T}}$ associated with the largest eigenvalue of the 3×3 matrix $\mathsf{Re}(\mathbf{F}^{\mathsf{H}}\mathbf{F})$. Thus, the angles ϕ and θ are easily obtained from (A.17). However, we can avoid trigonometrical operations by computing c and s directly from the coordinates of $\mathbf{v}$ as follows

$$c = \sqrt{\frac{v_1 + 1}{2}} \qquad s = \frac{v_2 - jv_3}{\sqrt{2(v_1 + 1)}} \tag{A.20}$$

and, finally, the plane rotations are determined using (A.10).

In the sequel, we summarize the implementation of the Extended Jacobi algorithm using a *cyclic-by-row* approach, i.e. each iteration consists in the search for a plane rotation relative to a given pair (p, q), where p and q correspond to the indexes of the upper triangular part of a $n \times n$ matrix. The algorithm is said to have performed a *sweep* when it finishes all $n(n-1)/2$ possible pairs (p, q). After each sweep, a stop condition is checked. In general, several sweeps are necessary in order to reach convergence under a certain tolerance level.

Algorithm A.2 (Extended Jacobi algorithm)

Initialize the algorithm with $\mathbf{V} \leftarrow \mathbf{I}$ and $\mathbf{\Delta}_k \leftarrow \mathbf{A}_k$, $\forall\, k = 1, \ldots, K$.; determine a tolerance level $\alpha > 0$ and compute $\mathsf{J}(\mathbf{Q}, \Delta)$ from (1.48), with $\Delta = \{\mathbf{\Delta_1}, \ldots, \mathbf{\Delta_K}\}$. Verify if $\mathsf{J}(\mathbf{Q}, \Delta) < \alpha\, \mathsf{J}(\mathbf{Q}, A)$. While this is false, repeat the procedure below:

For $p = 1 : n$

    For $q = p + 1 : n - 1$

1. *If $\sum_k \delta_{pq}^{(k)} = 0$ then $c = 1$ and $s = 0$. Go to step 4;*

2. *If $\sum_k \delta_{pq}^{(k)} \neq 0$ then compose the K vectors $\mathbf{f}_k$ from (A.19) and construct the following 3×3 matrix:*

$$\mathbf{F}' = Re\left(\sum_{k=1}^{K} \mathbf{f}_k \mathbf{f}_k^H\right) \tag{A.21}$$

3. *Compute the EVD of $\mathbf{F}'$ and obtain c and s from (A.20), where v_1, v_2 and v_3 are the elements of the eigenvector associated with the largest eigenvalue.*

4. *Construct the Givens rotation matrix $\mathbf{G}_{pq}(c, s)$, as indicated in (A.10).*

5. $\mathbf{\Delta}_k \leftarrow \mathbf{G}_{pq}(c, s)^H \mathbf{\Delta}_k \mathbf{G}_{pq}(c, s)$, $k = 1, \ldots, K$.

6. $\mathbf{V} \leftarrow \mathbf{V} \mathbf{G}_{pq}(c, s)$

Appendix B

Fundamental relationships

R andom processes observed at the output of a linear system convey information about the system parameters. High-order statistics have the ability to preserve system magnitude and phase information. Recovering system coefficients from output signals is hence possible due to existing relationships between output cumulants and the system parameters. In particular, Brillinger and Rosenblatt [3] established the link between high-order cumulants and the coefficients of a linear filter. In this appendix, we derive the Brillinger and Rosenblatt relationship in a simplified approach, for the case of a *single-input single-output* (SISO) system described by a noisy MA process.

Output cumulants of linear systems

Let $y(n)$ be the complex signal observed at the output of a linear SISO system given in (2.1) with the assumptions considered in section 2 of Chapter 2. From property P4, equation (**??**), we write the mth-order cumulant of $y(n)$ as follows

$$C_{m,y}(\tau_1, \ldots, \tau_{m-1}) = C_{m,x}(\tau_1, \ldots, \tau_{m-1}) + C_{m,v}(\tau_1, \ldots, \tau_{m-1}). \qquad (B.1)$$

thus, using assumption A2,

$$C_{m,y}(\tau_1, \ldots, \tau_{m-1}) = C_{m,x}(\tau_1, \ldots, \tau_{m-1}), \quad m > 2. \qquad (B.2)$$

Using definition (1.21) and the signal model (2.1), the above equation becomes

$$C_{m,y}(\tau_1, \ldots, \tau_{m-1}) = \begin{cases} cum\left[y^*(n), y(t + \tau_1), \ldots, y^*(n + \tau_{m-2}), y(n + \tau_{m-1})\right], & \text{if } m \text{ is even;} \\ cum\left[y^*(n), y(n + \tau_1), \ldots, y(n + \tau_{m-2}), y^*(n + \tau_{m-1})\right], & \text{if } m \text{ is odd,} \end{cases}$$

and, considering only the case where m is even (for ease of notation), we get:

$$C_{m,y}(\tau_1, \ldots, \tau_{m-1}) = cum\left[\sum_{\ell_0=0}^{L} h^*(\ell_0)s^*(n - \ell_0), \sum_{\ell_1=0}^{L} h(\ell_1)s(n - \ell_1 + \tau_1), \ldots, \sum_{\ell_{m-2}=0}^{L} h^*(\ell_{m-2})s^*(n - \ell_{m-2} + \tau_{m-2}), \right.$$
$$\left. \sum_{\ell_{m-1}=0}^{L} h(\ell_{m-1})s(n - \ell_{m-1} + \tau_{m-1})\right], \qquad (B.3)$$

where we have considered the stationarity of the input signal sequence $s(n)$ (assumption A1). Notice that the MA parameters $h(0), \dots, h(L)$ are deterministic complex constants. Therefore, using the additivity property of cumulants (1.14) along with property P1 (1.13), we get

$$
C_{m,y}(\tau_1, \dots, \tau_{m-1}) = \sum_{\ell_0} \sum_{\ell_1} \cdots \sum_{\ell_{m-2}} \sum_{\ell_{m-1}} h^*(\ell_0) h(\ell_1) \cdots h^*(\ell_{m-2}) h(\ell_{m-1}) \; cum \Big[s^*(n - \ell_0),
$$
$$
s(n - \ell_1 + \tau_1), \dots, s^*(n - \ell_{m-2} + \tau_{m-2}) s(n - \ell_{m-1} + \tau_{m-1}) \Big].
$$
(B.4)

Due to the iid property of the input signal sequence (assumption A1), the term represented by the $cum [\cdot]$ operator in the above equation yields the following:

$$
cum \Big[s^*(n - \ell_0), s(n - \ell_1 + \tau_1), \dots,
$$
$$
s^*(n - \ell_{m-2} + \tau_{m-2}), s(n - \ell_{m-1} + \tau_{m-1}) \Big] = \begin{cases} \gamma_{m,s}, & \ell_0 = \ell_1 - \tau_1 = \dots = \ell_{m-1} - \tau_{m-1} \\ 0, & \text{otherwise} \end{cases}
$$
(B.5)

where $\gamma_{m,s}$ is defined in (1.22). Thus, rewriting the indices in function of ℓ_0, we get $\ell_1 = \ell_0 + \tau_1$, $\dots, \ell_{m-1} = \ell_0 + \tau_{m-1}$, and hence (B.4) becomes

$$
C_{m,y}(\tau_1, \dots, \tau_{m-1}) = \gamma_{m,s} \sum_{\ell_0} h^*(\ell_0) h(\ell_0 + \tau_1) \cdots h^*(\ell_0 + \tau_{m-2}) h(\ell_0 + \tau_{m-1}).
$$
(B.6)

Finally, replacing ℓ_0 by ℓ and considering odd and even orders m, it follows that:

$$
C_{m,y}(\tau_1, \dots, \tau_{m-1}) = \begin{cases} \gamma_{m,s} \sum\limits_{\ell=\alpha_1}^{\alpha_2} h^*(\ell) h(\ell + \tau_1) \cdots h^*(\ell + \tau_{m-2}) h(\ell + \tau_{m-1}), & \text{if } m \text{ is even;} \\ \gamma_{m,s} \sum\limits_{\ell=\alpha_1}^{\alpha_2} h^*(\ell) h(\ell + \tau_1) \cdots h(\ell + \tau_{m-2}) h^*(\ell + \tau_{m-1}), & \text{if } m \text{ is odd;} \end{cases}
$$
(B.7)

where $\alpha_1 = \max(0, -\tau_1, -\tau_2, \dots, -\tau_{m-1})$ and $\alpha_2 = \min(L, L - \tau_1, L - \tau_2, \dots, L - \tau_{m-1})$, because $h(\ell) = 0 \; \forall \ell \notin [0, L]$ (assumption A4). Equation (B.7) consists in the well-known Brillinger and Rosenblatt's relationship [3] and it represents a fundamental tool in system identification since it establishes the link between cumulants and the system coefficients.

For $m = 4$, equation (B.7) gives

$$
C_{4,y}(\tau_1, \tau_2, \tau_3) = \gamma_{4,s} \sum_{\ell=\alpha_1}^{\alpha_2} h^*(\ell) h(\ell + \tau_1) h^*(\ell + \tau_2) h(\ell + \tau_3),
$$
(B.8)

with $\alpha_1 = \max(0, -\tau_1, -\tau_2, -\tau_3)$ and $\alpha_2 = \min(L, L - \tau_1, L - \tau_2, L - \tau_3)$. Identical reasoning leads to similar results for $m = 2$. However, taking into account that (B.2) does not hold for $m = 2$ because $C_{2,v}(\tau) \neq 0$, we have $C_{2,y}(\tau) = C_{2,x}(\tau) + C_{2,v}(\tau)$, but we can still write the cumulants of the noiseless signal $x(n)$ as follows:

$$
C_{2,x}(\tau) = \gamma_{2,s} \sum_{\ell=\alpha_1}^{\alpha_2} h^*(\ell) h(\ell + \tau),
$$
(B.9)

with $\alpha_1 = \max(0, -\tau)$ and $\alpha_2 = \min(L, L - \tau)$. Obviously, in the noiseless case we have $C_{2,y}(\tau) = C_{2,x}(\tau)$.

Notice that cumulants only admit non-zero values within the range of time-lags τ_1, ..., τ_{m-1} limited by the system memory. Particularly, as we assumed a causal MA model with order $L+1$, high-order cumulants will vanish when any of the lags τ_i, or the difference between any pair of lags, go beyond L or $-L$. Fourth-order cumulants, for instance, are non-zero in a region of the $\mathbb{R}^{\tau_1 \times \tau_2 \times \tau_3}$ space limited by the following bounds: $|\tau_1|, |\tau_2|, |\tau_3| \leq L$, $|\tau_3 - \tau_2| \leq L$, $|\tau_3 - \tau_1| \leq L$ and $|\tau_1 - \tau_2| \leq L$.

Bibliography

[1] J. M. Mendel. Tutorial on higher-order statistics (spectra) in signal processing and systems theory: theoretical results and some applications. *Proceedings of the IEEE*, 79(3):278–305, mar. 1991.

[2] C. L. Nikias and J. M. Mendel. Signal processing with higher-order spectra. *IEEE Signal Processing Magazine*, 10(3):10–37, jul. 1993.

[3] D. R. Brillinger and M. Rosenblatt. Computation and interpretation of kth-order spectra. In B. Harris, editor, *Spectral Analysis of Time Series*, pages 189–232, New York, USA, 1967. Wiley.

[4] G. B. Giannakis. Cumulants: a powerful tool in signal processing. *Proceedings of the IEEE*, 75(9):1333–1334, sep. 1987.

[5] G. B. Giannakis and J. M. Mendel. Identification of non-minimum phase systems using higher-order statistics. *IEEE Trans. on Acoustics, Speech and Signal Proc.*, 37:360–377, mar. 1989.

[6] B. Friedlander and B. Porat. Adaptive IIR algorithms based on high-order statistics. *IEEE Trans. on Acoustics, Speech and Signal Proc.*, 37:485–495, 1989.

[7] P. Comon. MA identification using fourth order cumulants. *Signal Processing Elsevier*, 26(3):381–388, mar. 1992.

[8] J. K. Tugnait. Approaches to FIR system identification with noisy data using higher order statistics. *IEEE Trans. on Acoustics, Speech and Signal Proc.*, 38(7):1307–1317, jul. 1990.

[9] A. Swami and J. M. Mendel. ARMA parameter estimation using only output cumulants. *IEEE Trans. on Acoustics, Speech and Signal Proc.*, 38(7):1257–1265, jul. 1990.

[10] A. G. Stogioglou and S. McLaughlin. MA parameter estimation and cumulant enhancement. *IEEE Trans. on Signal Processing*, 44(7):1704–1718, jul. 1996.

[11] D. Dembélé and G. Favier. A new FIR identification method based on fourth-order cumulants: application to blind equalization. *The Franklin Institute Journal*, 334B(1):117–133, 1997.

[12] P. Comon. Independent component analysis, a new concept ? *Signal Processing Elsevier*, 3(36):287–314, 1994.

[13] E. Moreau. A generalization of joint-diagonalization criteria for source separation. *IEEE Trans. on Signal Processing*, 49(3):530–541, mar 2001.

[14] L. De Lathauwer, B. De Moor, and J. Vandewalle. Independent component analysis and (simultaneous) third-order tensor diagonalization. *IEEE Trans. on Signal Processing*, 49(10):2262–2271, oct 2001.

[15] L. De Lathauwer, B. De Moor, and J. Vandewalle. Computation of the canonical decomposition by means of a simultaneous generalized Schur decompositition. *SIAM J. Matrix Anal. Appl.*, 6(2):295–327, 2004.

[16] R. A. Harshman. Foundations of the PARAFAC procedure: model and conditions for an "explanatory" multi-mode factor analysis. *UCLA Working papers in phonetics*, 16(1):1–84, 1970.

[17] P. McCullagh. Tensor methods in statistics. In *Monographs on Statistics and Applied Probability*, London, 1987. Chapman and Hall.

[18] L. De Lathauwer, B. De Moor, and J. Vandewalle. Independent component analysis based on higher-order statistics only. In *Proc. of the 8th IEEE SP Workshop on Statistical Signal and Array Processing (SSAP'96)*, pages 356–359, Corfu, Greece, jun. 1996.

[19] L. De Lathauwer, B. De Moor, and J. Vandewalle. Blind source separation by simultaneous third-order tensor diagonalization. In *Proc. of the 8th European Signal Processing Conference (EUSIPCO'96)*, pages 2089–2092, Trieste, Italy, sep 1996.

[20] J. B. Kruskal. Three way arrays: rank and uniqueness of trilinear decompositions with applications to arithmetic complexity and statistics. *Linear Algebra and Its Applications*, 18:95–138, 1977.

[21] L. De Lathauwer. A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization. *SIAM J. Matrix Anal. Appl.*, 28(3):642–666, 2006.

[22] L. De Lathauwer, J. Castaing, and J.-F. Cardoso. Fourth-order cumulant-based blind identification of underdetermined mixtures. *IEEE Trans. on Signal Processing*, 55(6):2965–2973, jun. 2007.

[23] J.-F. Cardoso. Eigen-structure of the fourth-order cumulant tensor with application to the blind source separation problem. In *Proc. ICASSP*, pages 2655–2658, Albuquerque, New Mexico, USA, 1990.

[24] J.-F. Cardoso. Super-symmetric decomposition of the fourth-order cumulant tensor. Blind identification of more sources than sensors. In *Proc. ICASSP*, pages 3109–3112, Toronto, Canada, 1991.

[25] P. Chevalier and A. Ferréol. On the virtual array concept for fourth order direction finding problem. *IEEE Trans. on Signal Processing*, 47(9):2592–2595, sep. 1999.

[26] P. Chevalier, L. Albera, A. Ferréol, and P. Comon. On the virtual array concept for higher order array processing. *IEEE Trans. on Signal Processing*, 53(4):1254–1271, apr. 2005.

[27] P. Chevalier, A. Ferrol, and L. Albera. High-resolution direction finding from higher order statistics: The 2q-music algorithm. *IEEE Transactions on Signal Processing*, 54(8):2986–2997, Aug. 2006.

[28] B. Porat and B. Friedlander. Performance analysis of MA parameter estimation algorithms based on high-order moments. In *Proc. ICASSP'88*, volume 4, pages 2412–2415, New York City, USA, 1988.

[29] B. Porat and B. Friedlander. Direction finding algorithms based on high-order statistics. *IEEE Trans. on Signal Processing*, 39(9):2016–2024, sep. 1991.

[30] J. F. Böhme. Statistical array signal processing of measured sonar and seismic data. In *Proc. SPIE Advanced Signal Processing Algorithms*, San Diego, USA, jul. 1995.

[31] P.-J. Chung, N. Czink, and C. F. Mecklenbräuker. Model order selection for multipath MIMO channels using the Benjamini-Hochberg procedure. In *ITG-/IEEE-Workshop on Smart Antennas*, Ulm, Germany, mar. 2006.

[32] P.-J. Chung, J. F. Böhme, C. F. Mecklenbräuker, and A. O. Hero. Detection of the number of signals using the Benjamini-Hochberg procedure. *IEEE Transactions on Signal Processing*, 55(6):2497–2508, jun. 2007.

[33] W. S. Liggett Jr. Passive sonar: fitting models to multiple time series. In J. W. R. Griffiths and P. L. Stocklin, editors, *Proc. of NATO ASI Signal Processing*, Academic, pages 327–345, New York, USA, 1973.

[34] D. B. Williams and D. H. Johnson. Modifying the sphericity test for improved source detection with narrowband passive arrays. In *Proc. of IEEE ICASSP*, volume 12, pages 2272–2275, Dallas, USA, apr. 1987.

[35] D. B. Williams and D. H. Johnson. Using the sphericity test for source detection with narrow-band passive arrays. *IEEE Transactions on Signal Processing*, 38(11):2008–2014, nov. 1990.

[36] S. Haykin. *Adaptive Filter Theory*. Prentice Hall, 3rd edition, 1995.

[37] R. Bro. PARAFAC. tutorial and applications. *Elsevier Chemometrics and Intelligent Laboratory Systems*, 38:149–171, 1997.

[38] N. D. Sidiropoulos and R. Bro. On the uniqueness of multilinear decomposition of N-way arrays. *Journal of Chemometrics*, 14:229–239, may. 2000.

[39] L. De Lathauwer. *Signal Processing Based on Multilinear Algebra*. PhD thesis, Katholieke Universiteit Leuven, Belgium, 1997. ESAT-SISTA/TR 1997-74.

[40] N. D. Sidiropoulos, G. B. Giannakis, and R. Bro. Blind PARAFAC receivers for DS-CDMA systems. *IEEE Trans. on Signal Processing*, 48(3):810–823, mar 2000.

[41] N. D. Sidiropoulos, R. Bro, and G. B. Giannakis. Parallel factor analysis in sensor array processing. *IEEE Trans. on Signal Processing*, 48(8):2377–2388, aug. 2000.

[42] P. Comon. Blind identification and source separation in 2×3 under-determined mixtures. *IEEE Trans. on Signal Processing*, 1(52):11–22, jan. 2004.

[43] A. L. F. de Almeida, G. Favier, and J. C. M. Mota. PARAFAC models for wireless communication systems. In *4th Int. Conf. on Phys. Sig. Image Proc. (PSIP'05)*, Toulouse, France, 2005.

[44] A. L. F. de Almeida, G. Favier, and J. C. M. Mota. PARAFAC-based unified tensor modeling of wireless communication systems with application to blind multiuser equalization. *Signal Processing Elsevier*, 87(2):337–351, fev. 2007.

[45] A. Papoulis. *Probability, Random Variables and Stochastic Processes*. McGraw-Hill International, 3 edition, 1991.

[46] P. O. Amblard, M. Gaeta, and J.-L. Lacoume. Statistics for complex variables and signals, part I: variables, part II: signals. *Signal Processing*, 53:1–13 (part I), 15–25 (part II), 1996.

[47] J.-L. Lacoume, P.-O. Amblard, and P. Comon. *Statistiques d'Ordre Supérieur pour le Traitement du Signal*. Masson, 1997. in French.

[48] S. Kotz and N.L. Johnson. *Enciclopedia of Statistical Sciences*. Wiley, 1982.

[49] G. H. Golub and H. A. van der Vorst. Eigenvalue computation in the 20th century. *Journal of Computational and Applied Mathematics*, (123):35–65, 2000.

[50] G. H. Golub and C. F. van Loan. *Matrix Computations*. The Johns Hopkins University Press, Maryland, US, 3rd edition, 1996.

[51] J.-F. Cardoso and A. Souloumiac. Blind beamforming for non gaussian signals. *IEE Proceedings-F*, 140(6):362–370, dec. 1993.

[52] A. Souloumiac J.-F. Cardoso. Jacobi angles for simultaneous diagonalization. *SIAM Journal on Matrix Analysis and Applications*, 17(1):161–164, jan. 1996.

[53] A. Bunse-Gerstner, R. Byers, and V. Mehrmann. Numerical methods for simultaneous diagonalization. *SIAM Journal on Matrix Analysis and Applications*, 14(4):927–949, oct. 1993.

[54] M. Wax and J. Sheinvald. A least squares approach to joint diagonalization. *IEEE Signal Processing Letters*, 4(2):52–53, feb. 1997.

[55] R. A. Harshman and M. E. Lundy. The PARAFAC model for three-way factor analysis and multidimensional scaling. In H.G. Law, C.W. Snyder, Jr.J. Hattie, and R.P. McDonald, editors, *Research Methods for Multimode Data Analysis*, pages 122–215, New York, 1984. Praeger.

[56] X. Liu and N. D. Sidiropoulos. Cramér-Rao bounds for low-rank decomposition of multidimensional arrays. *IEEE Trans. on Signal Processing*, 49:2074–2086, 2001.

[57] A. Stegeman and N. D. Sidiropoulos. On Kruskal's uniqueness condition for the Candecomp/Parafac decomposition. *Linear Algebra and Its Applications*, 420(2-3):540–552, 2007.

[58] A. Smilde, R. Bro, and P. Geladi. *Multi-way Analysis with Applications in the Chemical Sciences*. John Wiley and Sons Ltd, 2004.

[59] A. Stegeman and J. M. F. Ten Berge. Kruskal's condition for uniqueness in candecomp/parafac when ranks and k-ranks coincide. *Computational Statistics & Data Analysis*, 50(1):210–220, jan. 2006.

[60] Y. Yu and A. P. Petropulu. PARAFAC based blind estimation of possibly underdetermined convolutive MIMO systems. *IEEE Trans. on Signal Processing*, 56(1):111–124, jan. 2008.

[61] C. E. R. Fernandes, G. Favier, and J. C. M. Mota. Parafac-based blind channel identification using 4th-order cumulants. In *Proc. of VI International Telecommunications Symposium (ITS2006)*, Fortaleza, Brazil, sep. 2006.

[62] C. E. R. Fernandes, G. Favier, and J. C. M. Mota. Tensor-based blind channel identification. In *Proc. of IEEE International Conference on Communications (ICC 2007)*, pages 2728–2732, Glasgow, Scotland, UK, jun. 2007.

[63] C. E. R. Fernandes, G. Favier, and J. C. M. Mota. Blind MIMO channel identification using cumulant tensor decomposition. In *ASILOMAR Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, nov. 2007.

[64] C. E. R. Fernandes, G. Favier, and J. C. M. Mota. Blind channel identification algorithms based on the Parafac decomposition of cumulant tensors: the single and multiuser cases. *Signal Processing, Elsevier*, 88(6):1382–1401, jun. 2008.

[65] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines. A blind source separation technique using second-order statistics. *IEEE Trans. on Signal Processing*, 45(2):434–444, feb. 1997.

[66] A. Belouchrani and B. Derras. An efficient fourth-order system identification FOSI algorithm utilizing the joint diagonalization procedure. In *Proc. of the 10-th IEEE Workshop on Statistical Signal and Array Processing*, pages 621–625, Pocono Manor, Pennsylvania, USA, aug. 2000.

[67] L. De Lathauwer, B. De Moor, J. Vandewalle, and J.-F. Cardoso. Independent component analysis of largely underdetermined mixtures. In *Proc. 4th Int. Symp. on Independent Component Analysis and Blind Signal Separation (ICA 2003)*, pages 29–34, Nara, Japan, apr. 2003.

[68] L. Albera, A. Ferréol, P. Chevalier, and P. Comon. ICAR, a tool for blind source separation using fourth order statistics only. *IEEE Transactions On Signal Processing*, 53(10):3633–3643, oct. 2005. part I.

[69] L. Albera, A. Ferréol, P. Comon, and P. Chevalier. Sixth order blind identification of under-determined mixtures - BIRTH. In *Proc. ICA'03, 4th Int. Symp. on Indep. Comp. Anal. and Blind Sig. Sep.*, pages 909–914, Nara, Japan, apr. 2003.

[70] L. Albera, A. Ferréol, P. Comon, and P. Chevalier. Blind identification of overcomplete mixtures of sources (BIOME). *Linear Algebra Applications*, 391C:3–30, nov 2004.

[71] P. Comon and M. Rajih. Blind identification of under-determined mixtures based on the characteristic function. *Signal Processing Elsevier*, 86(9):2271–2281, sep 2006.

[72] B. Chen and A. Petropulu. Frequency domain blind MIMO system identification based on second- and higher order statistics. *IEEE Trans. on Signal Processing*, 49(8):1677–1688, aug. 2001.

[73] Y. Yu and A.P. Petropulu. Robust PARAFAC based blind estimation of MIMO systems with possibly more inputs than outputs. In *Proc. ICASSP*, pages 133–136, Toulouse, France, may. 2006.

[74] T. Acar, Y. Yu, and A. P. Petropulu. Blind MIMO system estimation based on PARAFAC decomposition of higher order output tensors. *IEEE Trans. on Signal Processing*, 54(11):4156–4168, nov. 2006.

[75] K. Abed-Meraim, A. Belouchrani, and Y. Hua. Blind identification of a linear-quadratic mixture of independent components based on joint diagonalization procedure. In *Proc. of IEEE ICASSP*, Atlanta, USA, may. 1996.

[76] A. Belouchrani, M. G. Amin, and K. Abed-Meraim. Direction finding in correlated noise fields based on joint block-diagonalization of spatio-temporal correlation matrices. *IEEE Signal Processing Letters*, 4(9):266–268, sep. 1997.

[77] E. Moreau and J.-C. Pesquet. Generalized contrasts for multichannel blind deconvolution of linear systems. *IEEE Signal Processing Letters*, 4(6):182–183, jun. 1997.

[78] J.C. Pesquet M. Castella and A.P. Petropulu. Family of frequency and time-domain contrasts for blind separation of convolutive mixtures of temporally dependent dignals. *IEEE Trans. on Signal Processing*, 53(1):107–120, jan. 2005.

[79] T. D. Acar and A. P. Petropulu. Blind MIMO system identification using PARAFAC decomposition of an output HOS-based tensor. In *Asilomar Conference on Signals, Systems and Computers*, pages 1080–1084, Pacific Grove, USA, nov 2003.

[80] Y. Yu and A.P. Petropulu. Blind MIMO system estimation based on PARAFAC decomposition of HOS tensors. In *IEEE Statistical Signal Processing Workshop*, Bordeaux, France, jul. 2005.

[81] A. Ferréol, L. Albera, and P. Chevalier. Fourth order blind identification of underdetermined mixtures of sources (FOBIUM). *IEEE Trans. on Signal Processing*, 53(5):1640–1653, may. 2005.

[82] J.-F. Cardoso. Localisation et identification par la quadricovariance. *Traitement du Signal*, 7(5):397–406, 1990. (*in french*).

[83] J.-F. Cardoso. Eigen-structure of the fourth-order cumulant tensor with application to the blind source separation problem. In *In Proc. ICASSP*, pages 2655–2658, 1990.

[84] E. Moreau and O. Macchi. A one stage self-adaptive algorithm for source separation. In *Proc. ICASSP*, volume III, pages 49–52, Adelaide, Australia, apr. 1994.

[85] E. Moreau and O. Macchi. High order contrasts for self-adaptive source separation. *International Journal of Adaptive Control and Signal Processing*, 10(1):19–46, jan. 1996.

[86] R. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Trans. on Antennas and Propagation*, 34(3):276–280, mar. 1986.

[87] G. Bienvenu and L. Kopp. Optimality of high resolution array processing using the eigen-system approach. *IEEE Trans. Acoust. Speech, Signal Process.*, 31(5):1235–1247, oct. 1983.

[88] R. Roy and T. Kailath. ESPRIT – estimation of signal parameters via rotational invariance techniques. *IEEE Trans. Acoustics, Speech, Signal Processing*, 37(7):984–995, Jul. 1989.

[89] P. Germain, A. Maguer, and L. Kopp. Comparison of resolving power of array processing methods by using an analytical criterion. In *Proc. ICASSP*, Glasgow, U.K., May 1989.

[90] M. Kaveh and A. J. Barabell. The statistical performance of the MUSIC and the minimum norm algorithms in resolving plane waves in noise. *IEEE Trans. Acoust., Speech, Signal Process.*, 34(2):331–341, apr. 1986.

[91] B. Porat and B. Friedlander. Analysis of the asymptotic relative efficiency of the MUSIC algorithm. *IEEE Trans. Acoust. Speech. Signal Process.*, 36(4):532544, apr. 1988.

[92] P. Stoica and A. Nehorai. MUSIC, maximum likelihood and cramer rao bound. *IEEE Trans. Acoust., Speech, Signal Process.*, 37(12):720–741, may 1989.

[93] P. Stoica and A. Nehorai. MUSIC, maximum likelihood and cramer rao bound: Further results and comparisons. *IEEE Trans. Acoust., Speech, Signal Process.*, 38(12):2140–2150, dec. 1990.

[94] A. Paulraj and T. Kailath. Eigenstructure methods for direction of arrival estimation in the presence of unknown noise field. *IEEE Trans. Acoust., Speech, Signal Process.*, 34(1):13–20, feb. 1986.

[95] J.-F. Cardoso. Higher-order narrow band array processing. In *In Proc. Int. Workshop on Higher-Order Stat.*, pages 121–130, Chamrousse, France, 1991.

[96] E. Gonen and J.M. Mendel. Applications of cumulants to array processing. iii. blindbeam-forming for coherent signals. *IEEE Transactions on Signal Processing*, 45(9):2252–2264, sep. 1997.

[97] B. Porat and B. Friedlander. Performance analysis of parameter estimation algorithms based on high-order moments. *Journal of Adaptive Control and Signal Processing*, 3:191–229, 1989.

[98] J.-F. Cardoso and E. Moulines. Asymptotic performance analysis of direction finding algorithms based on fourth-order cumulants. *IEEE Trans. on Signal Processing*, 43(1):214–224, jan. 1995.

[99] M. C. Dogan and J. M. Mendel. Applications of cumulants to array processing - part i: Aperture extension and array calibration. *IEEE Trans. on Signal Process*, 43(5):1200–1216, may. 1995.

[100] H. H. Chiang and C. L. Nikias. The ESPRIT algorithm with high order statistics. In *in Proc. Workshop Higher Order Statistics*, pages 163–168, Vail, Jun. 1989.

[101] S. Shamsunder and G. B. Giannakis. Modeling of non-gaussian array data using cumulants: DOA estimation of more sources with less sensors. *Signal Process*, 30(3):279–297, feb. 1993.

[102] J. G. Proakis. *Digital Communications*. McGraw Hill, 4th. ed. edition, 2000.

[103] G. Giannakis, Y. Inouye, and J. Mendel. Cumulant-based identification of multichannel moving average models. *IEEE Trans. on Automat. Control*, 34:783–787, 1989.

[104] C. L. Nikias. Blind deconvolution using higher-order statistics. In Elsevier, editor, *Proc. Int. Conf. on Higher-Order Stat.*, pages 49–56, 1992.

[105] Y. Sato. A method for self-recovering equalization for multi-level amplitude modulation. *IEEE Trans. on Communic.*, 6:679–682, 1975.

[106] D. N. Godard. Self-recovering equalization and carrier tracking in two dimensional data communication system. *IEEE Trans. on Communic.*, 28:1867–1875, nov 1980.

[107] A. Benveniste, M. Goursat, and G. Ruget. Robust identification of a non-minimum phase system: Blind adjustment of a linear equalizer in data communication. *IEEE Trans. on Automatic Control*, 25(3):385–399, 1980.

[108] O. Shalvi and E. Weinstein. New criteria for blind deconvolution of non-minimum phase systems (channels). *IEEE Trans. on Inform. Theory*, 36:312–321, mar. 1990.

[109] L. Tong, G. Xu, and T. Kailath. A new approach to blind identication and equalization of multipath channels. In *Proc. 25th ASILOMAR*, pages 856–860, Pacific Grove, CA, USA, 1991.

[110] L. Tong, G. Xu, and T. Kailath. Blind identication and equalization of multipath channels. In *Proc. Int. Conf. Communications*, pages 1513–1517, jun. 1992.

[111] L. Tong, G. Xu, and T. Kailath. Blind identication based on second-order statistics: A time domain approaches. *IEEE Information Theory*, 40(2):340–349, 1994.

[112] E. Moulines, P. Duhamel, 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, feb. 1995.

[113] D. T. Slock. Blind fractionally-spaced equalization, perfect reconstruction filter banks and multichannel prediction. In *Proc. Icassp-94*, volume 4, pages 585–588, Adelaide, Australia, 1994.

[114] K. A. Meraim, J.-F. Cardoso, A. Gorokhov, P. Loubaton, and E. Moulines. On subspace methods for blind identification of single-input/multiple-output FIR filters. *IEEE Trans. on Signal Processing*, 45(1):42–55, jan. 1997.

[115] L. Tong and S. Perreau. Multichannel blind identification: From subspace to maximum likelihood methods. *Proceedings of the IEEE*, 86:10, oct. 1998.

[116] K. A. Meraim, P. Loubaton, and E. Moulines. A subspace algorithm for certain blind identification problems. *IEEE Trans. on Signal Processing*, 43(2):499–511, mar. 1997.

[117] A. Gorokhov and P. Loubaton. Subspace based techniques for second order blind separation of convolutive mixtures with temporally correlated sources. *IEEE Trans. on Circuit and Systems*, 44(9):813–820, sep. 1997.

[118] K. A. Meraim, E. Moulines, and P. Loubaton. Prediction error method for second order blind identification. *IEEE Trans. on Signal Processing*, 45(3):694–705, mar. 1997.

[119] J. K. Tugnait. On linear predictors for MIMO channels and related blind identication and equalization. *IEEE Signal Processing Letters*, 5(11):289–291, nov. 1998.

[120] M. Martone. *Multiantenna digital radio transmission*. Mobile communication series. Artech House, 2002.

[121] R. T. Compton. *Adaptive Antennas - Concepts and Performance*. Prentice Hall, Englewood Cliffs, New Jersey, USA, 1988.

[122] M. Kendall and A. Stuart. *The Advanced Theory of Statistics, Distribution Theory*. C. Griffin, 1977.

[123] P. Chevalier, A. Ferreol, and J. P. Denis. New geometrial results about fourth-order direction finding method performance. In *Proc. EUSIPCO 96*, pages 923–926, Trieste, Italy, sep 1996.

[124] J. Mendel and M. C. Dogan. Higher-order statistics applied to some array signal processing problems. In eds. M. Blanke, T. Sonderstrom, editor, *Proc. of 10th IFAC-SYSID*, pages 101–106, Copenhagen, Denmark, jul 1994.

[125] N. D. Sidiropoulos and X. Liu. Identifiability results for blind beamforming in incoherent multipath with small delay spread. *IEEE Trans. on Signal Processing*, 49(1):228–236, jan. 2001.

[126] A. L. F. de Almeida, G. Favier, and J. C. M. Mota. Estimation of frequency-selective block-fading MIMO channels using PARAFAC modeling and Alternating Least Squares. In *Proc. of ASILOMAR'06,*, Pacific Grove, CA, USA, 2006.

[127] B. Porat and B. Friedlander. Direction finding algorithms based on higher order statistics. *IEEE Trans. on Signal Process*, 39(9):2016–2024, sep. 1991.

[128] P. Loubaton, E. Moulines, and P. Regalia. *Signal Processing Advances in Wireless and Mobile Communications*, volume II - Trends in Single- and Multi-User Systems, chapter Subspace method for blind identification and deconvolution, pages 63–112. Prentice-Hall, 2000.

[129] A. P. Liavas, P. A. Regalia, and J.-P. Delmas. Blind channel approximation: Effective channel order determination. *IEEE Trans. on Signal Processing*, 47(12):3336–3344, dec. 1999.

[130] A. P. Liavas, P. A. Regalia, and J.-P. Delmas. Robustness of least-squares and subspace methods for blind channel identification/equalization with respect to effective channel undermodeling/overmodeling. *IEEE Trans. on Signal Processing*, 47(6):1636–1644, jun. 1999.

[131] J.-P. Delmas, H. Gazzah, A. P. Liavas, and P. A. Regalia. Statistical analysis of some second-order methods for blind channel identification/equalization with respect to channel undermodeling. *IEEE Trans. on Signal Processing*, 48(7):1984–1998, Jul. 2000.

[132] J. W. Mauchly. Significance tests for sphericity of a Normal n-Variate distribution. *The Annals of Mathematical Statistics*, 11(2):204–209, jun. 1940.

[133] R. J. Muirhead. *Aspects of multivariate statistical theory*, volume (XIX-673 p.). John Wiley & Sons, New York, 1982.

[134] P. Stoica and Y. Selén. Model order selection: a review of information criterion rules. *IEEE Signal Processing Magazine*, 21(4):36–47, 2004.

[135] H. L. Van Trees. *Detection, Estimation and Modulation Theory. Part I*. Wiley, New York, 1968.

[136] H. Akaike. A new look at the statistical model identication. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974.

[137] J. Rissanen. Modeling by shortest data description length. *Automatica*, 14(5):465–471, 1978.

[138] G. Schwarz. Estimating the dimensions of a model. *The Annals of Statistics*, 6(2):461–464, 1978.

[139] R. L. Kashyap. Inconsistency of the AIC rule for estimating the order of autoregressive models. *IEEE Transactions on Automatic Control*, AC-25(5):996–998, oct. 1980.

[140] M. Kaveh, H. Wang, and H. Hung. On the theoretical performance of a class of estimators of the number of narrow-band sources. *IEEE Trans. on Acoustics, Speech and Signal Processing*, 35(9):1350–1352, 1987.

[141] M. Wax and T. Kailath. Detection of signals by information theoretic criteria. *IEEE Trans on Acoust., Speech and Signal Processing*, ASSP-33:387–392, apr. 1985.

[142] M. Wax. *Detection and Estimation of Superimposed Signals*. PhD thesis, Stanford University, Stanford, California, USA, mar. 1985.

[143] P. Comon. Separation of stochastic processes. In *Proc. Workshop on Higher-Order Spectral Analysis*, pages 174–179, Vail, Colorado, jun. 1989.

[144] P. Comon. Separation of sources using high-order cumulants. In *SPIE Conf. Adv. Alg. Archi. Sig. Proc.*, pages 170–181, San Diego, CA, aug. 1989.

[145] J.-F. Cardoso. Source separation using higher order moments. In *In Proc. ICASSP*, pages 2109–2112, Glasgow, Scotland, may. 1989.

[146] P. Comon. Independent Component Analysis. In *Proc. Int. Sig. Proc. Workshop on Higher-Order Statistics*, pages 111–120, Chamrousse, France, July 1991.

[147] C. E. R. Fernandes, P. Comon, and G. Favier. Order detection and blind identification of 2×1 MISO channels. In *Proc. of IEEE ICASSP*, volume 3, pages 753–756, Honolulu, USA, apr. 2007.

[148] B. Friedlander and B. Porat. Asymptotically optimal estimation of MA and ARMA parameters of non-gaussian processes from higher-order moments. *IEEE Trans. on Automatic Control*, 35(1):27–35, 1990.

[149] L. Albera and P. Comon. Asymptotic performance of contrast-based blind source separation algorithms. In *2nd IEEE Sensor Array and Multichannel Signal Processing Workshop*, Rosslyn, USA, aug 2002.

[150] P. Comon, P. Chevalier, and V. Capdevielle. Performance of contrast-based blind source separation. In *Proc. SPAWC*, pages 345–348, Paris, 1997.

[151] P. Comon and L. Deruaz. Normality tests for coloured samples. In *In Proc. IEEE-ATHOS Work. on HOS*, pages 217–221, Begur, Spain, jun. 1995.

[152] D. R. Brillinger. *Time Series: Data Analysis and Theory*. Holden-Day, San Francisco, 1981.

[153] J. Neyman and E. S. Pearson. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London*, A 231:289–337, 1933.

[154] N. L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions*, volume 1. Wiley.

[155] P.-J. Chung, J. F. Böhme, C. F. Mecklenbräuker, and A. O. Hero. On signal detection using the Benjamini-Hochberg procedure. In *Proc. IEEE Workshop on Statistical Signal Processing*, pages 615–618, Bordeaux, France, jul. 2005.