Analytical Multi-Modulus Algorithms Based on Coupled Canonical Polyadic Decompositions

Otto Debals, *Student Member, IEEE*, Muhammad Sohail, and Lieven De Lathauwer, *Fellow, IEEE*

***Abstract*—We present new techniques for multiple-input-multiple-output blind signal separation and blind system identification of multi-modulus source signals. Multi-modulus signals, such as 16-QAM, are very common in telecommunications. We algebraically transform the problem into a set of coupled tensor decompositions for which uniqueness results and algebraic solutions exist. An exact solution is guaranteed to be obtained by a matrix eigenvalue decomposition in the noiseless case. The proposed technique is deterministic and does not rely on statistics. Furthermore, we explain that certain source signals can still be recovered in the case of a rank-deficient mixing matrix. As a side result, we generalize a rank-1 detection procedure from a previously proposed tensor decomposition method. In the multi-modulus context, the generalization allows a reduction of the required number of samples for separation.**

***Index Terms*—Blind signal separation, constant modulus, multi-modulus, tensor, coupled tensor decompositios**

I. Introduction

Given a set of signals $\mathbf{X}$, the multiple-input-multiple-output (MIMO) blind signal separation (BSS) problem consists of the identification of the mixing matrix $\mathbf{M}$ and/or the original sources $\mathbf{S}$. In the instantaneous linear MIMO BSS model, one considers the following model with $\mathbf{N}$ the additive noise matrix:

$$\mathbf{X} = \mathbf{MS} + \mathbf{N}. \tag{1}$$

In blind system identification (BSI), an unknown system transfer function is considered instead of an instantaneous mixing matrix. In case of a linear time-invariant system, the observed signals are obtained through a convolution. Hence, the problem is also known as blind deconvolution.

For both BSS and BSI, solutions can be obtained under different working hypotheses. One hypothesis is mutual statistical independence of the source signals, leading to independent component analysis (ICA) [1], [2]. A deterministic hypothesis for multiple-invariance sensor array processing can be found in [3]; other deterministic assumptions are discussed in [4]. In telecommunications, blind separation and identification avoid the use of training sequences. A lot of work has been done in exploiting prior knowledge on the moduli of the samples of telecommunication signals. An example is the constant modulus algorithm (CMA), discussed in [5] and [6]. It was originally presented in a context of convolutive single-input-single-output (SISO). In CMA, all source samples are assumed to have the same constant modulus (CM), as is the case for real-valued binary phase shift keying (BPSK) signals, or complex-valued continuous phase/frequency shift keying (CPSK/CFSK) and complex 4-QAM (quadrature amplitude modulation) signals. Several important improvements followed in [7] and [8]. In [9], an algorithm for CMA was developed which was later coined as the multi-modulus algorithm [10], [11]. The algorithm applies a cost function with multiple linear moduli to fix the rotational ambiguity for 4-QAM signals. Despite the terminology, we will show that the method is not suitable for the problem discussed in this paper. Various other cost functions have been analyzed for CMA, grouped in different families [12], [13], [14].

Instead of optimizing a certain CM cost function, the analytical constant modulus algorithm (ACMA) transforms the MIMO BSS problem analytically into a simultaneous matrix diagonalization [15]. This was one of the first appearances of the constant-modulus problem in a MIMO context. ACMA was later extended to the convolutive case [16], [17]. In [17], a connection was made to tensor algebra, as it was shown that ACMA boils down to the computation of a tensor decomposition of a partially Hermitian third-order tensor.

Many signals, such as higher-order QAM, do not satisfy the constant modulus property. One has to resort to more general techniques allowing signal samples with a broader range of possible moduli. Signals of which the sample moduli are not restricted to a single constant modulus but rather to a finite set of moduli are called multi-modulus signals. The samples then lie on a number of concentric circles. An example is a rectangular 16-QAM signal with samples drawn from $\pm\{1,3\} \pm \{1,3\}j$, having squared moduli of 2, 10 or 18. Circular 16-QAM involves four moduli. Both constellations are illustrated in Fig. 1.

It was shown in [18] that the behavior of CMA is poor for the multi-modulus MIMO BSS and BSI problem, especially when the source signals are correlated and nonuniform. Because the MMA algorithm from Oh [9] uses the dispersion of real and imaginary parts separately, it was presented as more suitable for higher-order QAM constellations [19], [20], [21]. Note that the definitions of the dispersion moduli used

This research is funded by (1) a Ph.D. grant of the Agency for Innovation by Science and Technology (IWT), (2) Research Council KU Leuven: C1 project c16/15/059-nD, CoE PFV/10/002 (OPTEC), (3) F.W.O.: project G.0830.14N and G.0881.14N, (4) the Belgian Federal Science Policy Office: IUAP P7/19 (DYSCO II, Dynamical systems, control and optimization, 2012-2017), (5) EU: The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Advanced Grant: BIOTENSORS (no. 339804). This paper reflects only the authors' views and the Union is not liable for any use that may be made of the contained information.

Otto Debals and Lieven De Lathauwer are with both the Group of Science, Engineering and Technology, KU Leuven Kulak, E. Sabbelaan 53, B-8500 Kortrijk, Belgium and with the Department of Electrical Engineering (ESAT), KU Leuven, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium. (e-mail: otto.debals@esat.kuleuven.be, lieven.delathauwer@kuleuven-kulak.be). Muhammad Sohail is a master student at KU Leuven. (e-mail: muhammad.sohail@student.kuleuven.be).

Manuscript received September, 2016.

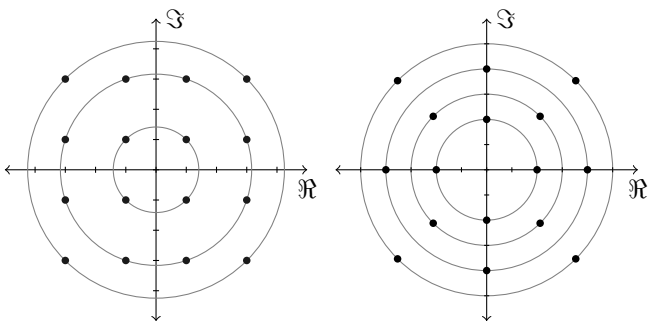


Fig. 1: Constellation diagram for rectangular (left) and circular (right) 16-QAM, in which three and four squared moduli are present, respectively. CMA is only applicable when samples lie on a single circle. For rectangular 16-QAM, the squared moduli are 2, 10 and 18. For circular 16-QAM, the squared moduli are 4, 9, 16 and 25.

in the latter methods differ intrinsically from the definition of the moduli of multi-modulus signals. Appendix A contains a further clarification. Others applied Givens rotations on the MMA cost function [22] or introduced customizations to the default MMA algorithm [23].

This paper proposes a new method for MIMO BSS and BSI of multi-modulus signals as a sound generalization of the ACMA algorithm. Unique to our MIMO approach is, first, that a more suitable cost function is used than the one in MMA. It is related to finite alphabet methods [24], [25], [26] and has been used in a convolutive SISO context in [18], [27].

Second, the problem is algebraically transformed into a set of coupled tensor decompositions. Coupled tensor decompositions, and more specifically coupled canonical polyadic decompositions (cCPDs), form an important and emerging concept in signal processing and data analytics. They enable the integration of data from multiple datasets, often called data fusion [28], [29], [30]. Using cCPDs, one-dimensional harmonic retrieval and shift-invariance techniques can be generalized to multi-dimensional harmonic retrieval and multiple-invariance techniques, respectively. Furthermore, one can deal with non-uniform linear arrays, incomplete arrays and convolutive mixtures instead of uniform linear arrays, dense arrays and instantaneous mixtures, respectively [31], [32].

Because of the reformulation into the tensor framework, algorithms and uniqueness conditions can be readily obtained [33], [34]. Algebraic methods using matrix eigenvalue decompositions are available such that an exact solution can be obtained in the noiseless case. In the presence of noise, refinement optimization techniques can be used. Furthermore, we present explicit lower bounds for the number of samples. Note that the method does not make use of statistics, and is therefore not dependent on the corresponding estimation error.

As a side result of this paper, we generalize a rank-1 detection procedure from a previously proposed algebraic CPD algorithm [35]. While the original method finds Kronecker-structured vectors in a space spanned by a number of Kronecker-structured vectors, the generalization allows a search space spanned by both Kronecker-structured and arbitrary vectors. The generalized method is applied in the context of the multi-modulus technique to obtain a reduction of the number of samples required for separation.

The paper is organized as follows. Multilinear algebra is introduced in Section II. The BSS method is proposed in Section III. An advanced version is discussed in Section IV making use of a generalized rank-1 detection procedure. The methods are extended to BSI in Section V. Simulations are presented in Section VI and a discussion and conclusion follows in Section VII.

II. Multilinear algebra and notation

Tensors, denoted by calligraphic letters, e.g., $\mathcal{A}$, are higher-order generalizations of vectors (denoted by boldface lowercase letters, e.g., $\mathbf{a}$) and matrices (denoted by boldface uppercase letters, e.g., $\mathbf{A}$). Scalars are written as italic lowercase letters, e.g., a. The entry with row index i and column index j of a matrix $\mathbf{A} \in \mathbb{C}^{I\times J}$ is denoted by a_{ij}. Likewise, the $(i_1, i_2, \ldots, i_N)$th entry of an Nth-order tensor $\mathcal{A} \in \mathbb{C}^{I_1\times I_2\times\ldots\times I_N}$ is denoted by $a_{i_1 i_2 \ldots i_N}$. The jth column of a matrix $\mathbf{A} \in \mathbb{C}^{I\times J}$ is denoted by $\mathbf{a}_j$. The superscripts $\cdot^{\mathrm{T}}$, $\cdot^{\star}$, $\cdot^{\mathrm{H}}$, $\cdot^{-1}$ and $\cdot^{\dagger}$ represent the transpose, complex conjugate, complex conjugated transpose, inverse and Moore–Penrose pseudoinverse, respectively.

The Kronecker and column-wise Khatri–Rao products are denoted by $\otimes$ and $\odot$, respectively. The operation $\otimes$ stands for the outer product. The Kronecker and outer product are related, as for vectors $\mathbf{a} \in \mathbb{C}^I$ and $\mathbf{b} \in \mathbb{C}^J$ it holds that $\mathbf{a} \otimes \mathbf{b} = \mathrm{vec}(\mathbf{b} \otimes \mathbf{a})$, with vec the column-wise vectorization of a matrix or tensor. Concatenations along the first and second mode are denoted as $\left[\cdot\,;\,\cdot\right]$ and $\left[\cdot \quad \cdot\right]$, respectively.

A Kronecker-structured vector $\mathbf{v} \in \mathbb{C}^{IJ}$ is defined as a vector which can be written as a Kronecker product of two non-zero vectors $\mathbf{a} \in \mathbb{C}^I$ and $\mathbf{b} \in \mathbb{C}^J$ such that $\mathbf{v} = \mathbf{a} \otimes \mathbf{b}$. It can be seen that such a vector is a reshaped rank-1 matrix.

If an Nth-order tensor $\mathcal{A}$ can be written as an outer product of non-zero vectors, i.e., $\mathcal{A} = \mathbf{a}^{(1)} \otimes \mathbf{a}^{(2)} \otimes \cdots \otimes \mathbf{a}^{(N)}$, it has rank 1. If a tensor can be written as a sum of R rank-1 terms, the decomposition is called a polyadic decomposition (PD):

$$\mathcal{A} = \sum_{r=1}^{R} \mathbf{a}_r^{(1)} \otimes \mathbf{a}_r^{(2)} \otimes \ldots \otimes \mathbf{a}_r^{(N)} \triangleq \left[\!\!\left[\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \ldots, \mathbf{A}^{(N)} \right]\!\!\right].$$

The factor vectors $\mathbf{a}_r^{(n)}$ are the columns of the factor matrices $\mathbf{A}^{(n)}$. If R is minimal, the decomposition is a canonical polyadic decomposition (CPD) and R is defined as the rank of the tensor. One advantage of the CPD is that it is unique under very mild conditions [36], [37]. There exists ample literature on tensors and tensor decompositions; we refer the interested reader to [38], [39], [40].

III. Tensor-based multi-modulus BSS

We recall the blind signal separation problem:

$$\mathbf{X} = \mathbf{M}\mathbf{S} + \mathbf{N}.$$

Let us assume R sources, K observed signals and N samples for each signal; hence, $\mathbf{X}, \mathbf{N} \in \mathbb{C}^{K\times N}$, $\mathbf{M} \in \mathbb{C}^{K\times R}$ and $\mathbf{S} \in \mathbb{C}^{R\times N}$. We assume that $K = R$, as a preprocessing step based on principal component analysis (PCA) can be carried

out when $K > R$. The underdetermined case is not discussed. Note that the distinction between K and R is still made to improve readability.

By defining the separation matrix $\mathbf{W} \in \mathbb{C}^{K\times R}$ as $\mathbf{W}^{\mathrm{T}} = \mathbf{M}^{\dagger}$ and by omitting $\mathbf{N}$ for convenience, one can write $\mathbf{S} = \mathbf{W}^{\mathrm{T}}\mathbf{X}$. The transpose is meant to simplify notation further on. Considering a source sample $s_{rn} = s_r(n)$, one can write

$$\forall r, n: \quad s_{rn} = \mathbf{w}_r^{\mathrm{T}}\mathbf{x}_n, \tag{2}$$

with $\mathbf{w}_r$ the rth column of $\mathbf{W}$ and $\mathbf{x}_n$ the nth column of $\mathbf{X}$.

Constant modulus signals have the property that $\forall r, n: |s_{rn}|^2 = s_{rn} \cdot s_{rn}^{\star} = c$, with c the *a priori* known squared constant modulus. For multi-modulus signals, we have a more general constraint:

$$\forall r, n: |s_{rn}|^2 = s_{rn} \cdot s_{rn}^{\star} \in \{c_1, \ldots, c_P\},$$

in which P is the number of possible moduli, or, equivalently, the number of concentric circles on which the source samples can lie; e.g., $P = 3$ in the case of rectangular 16-QAM. Hence, one has:

$$\forall r, n: \prod_{p=1}^{P} \left(|s_{rn}|^2 - c_p \right) = 0. \tag{3}$$

From these constraints, a minimization problem can be constructed with objective function J:

$$J = \sum_{r=1}^{R}\sum_{n=1}^{N} \left[\prod_{p=1}^{P} \left(|s_{rn}|^2 - c_p \right)^2 \right]. \tag{4}$$

Standard optimization techniques can be used to minimize J. Throughout the next subsections, however, we will solve the problem in an algebraic way. The constraint from Eq. (3) yields a linear system with structured solutions in subsection III-A. Whereas in [27] an approximate suboptimal quadratic eigenvalue method [41] is used, we will show that by applying a classical subspace-based technique (subsections III-B and III-C), the problem can be solved by means of coupled CPDs (subsection III-D).

A. Translation into a constrained set of linear equations

We derive the method for $P = 2$ while the generalization for $P > 2$ is straightforward. Let us consider a single source $\mathbf{s}_r$, drop the subscript r, and substitute (2) in (3):

$$\forall n: \quad \left(\mathbf{w}^{\mathrm{T}}\mathbf{x}_n\mathbf{x}_n^{\mathrm{H}}\mathbf{w}^{\star} - c_1\right)\left(\mathbf{w}^{\mathrm{T}}\mathbf{x}_n\mathbf{x}_n^{\mathrm{H}}\mathbf{w}^{\star} - c_2\right) = 0, \tag{5}$$

then we obtain the following:

$$\forall n: \left(\mathbf{w}^{\mathrm{T}}\mathbf{x}_n\mathbf{x}_n^{\mathrm{H}}\mathbf{w}^{\star}\right)^2 - (c_1 + c_2)\mathbf{w}^{\mathrm{T}}\mathbf{x}_n\mathbf{x}_n^{\mathrm{H}}\mathbf{w}^{\star} + c_1c_2 = 0.$$

As one can show that for arbitrary vectors $\mathbf{d}, \mathbf{e} \in \mathbb{C}^I$ and $\mathbf{f}, \mathbf{g} \in \mathbb{C}^J$ the equality $(\mathbf{d}^{\mathrm{T}}\mathbf{e})(\mathbf{f}^{\mathrm{T}}\mathbf{g}) = (\mathbf{e}\otimes\mathbf{f})^{\mathrm{T}}(\mathbf{d}\otimes\mathbf{g})$ holds, we can transform the preceding equation into:

$$\begin{aligned}\forall n: \quad & (\mathbf{x}_n \otimes \mathbf{x}_n^{\star} \otimes \mathbf{x}_n \otimes \mathbf{x}_n^{\star})^{\mathrm{T}}(\mathbf{w}\otimes\mathbf{w}^{\star}\otimes\mathbf{w}\otimes\mathbf{w}^{\star}) - \\ & (c_1 + c_2)(\mathbf{x}_n \otimes \mathbf{x}_n^{\star})^{\mathrm{T}}(\mathbf{w}\otimes\mathbf{w}^{\star}) = -c_1c_2.\end{aligned}$$

Stacking for n from 1 to N yields:

$$\begin{aligned}&\begin{bmatrix} (\mathbf{x}_1 \otimes \mathbf{x}_1^{\star} \otimes \mathbf{x}_1 \otimes \mathbf{x}_1^{\star})^{\mathrm{T}} \\ \vdots \\ (\mathbf{x}_N \otimes \mathbf{x}_N^{\star} \otimes \mathbf{x}_N \otimes \mathbf{x}_N^{\star})^{\mathrm{T}} \end{bmatrix} (\mathbf{w}\otimes\mathbf{w}^{\star}\otimes\mathbf{w}\otimes\mathbf{w}^{\star}) - \\ &(c_1 + c_2)\begin{bmatrix} (\mathbf{x}_1 \otimes \mathbf{x}_1^{\star})^{\mathrm{T}} \\ \vdots \\ (\mathbf{x}_N \otimes \mathbf{x}_N^{\star})^{\mathrm{T}} \end{bmatrix} (\mathbf{w}\otimes\mathbf{w}^{\star}) = -c_1c_2 \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}.\end{aligned}$$

This set of equations can be written as follows:

$$\mathbf{R}\mathbf{u} + \mathbf{P}\mathbf{v} = \alpha\mathbf{1}, \tag{6}$$

with $\mathbf{1}$ the vector with all ones, the scalar $\alpha = -c_1c_2$ and

$$\mathbf{R} = \begin{bmatrix} (\mathbf{x}_1 \otimes \mathbf{x}_1^{\star} \otimes \mathbf{x}_1 \otimes \mathbf{x}_1^{\star})^{\mathrm{T}} \\ \vdots \\ (\mathbf{x}_N \otimes \mathbf{x}_N^{\star} \otimes \mathbf{x}_N \otimes \mathbf{x}_N^{\star})^{\mathrm{T}} \end{bmatrix} = (\mathbf{X}\odot\mathbf{X}^{\star}\odot\mathbf{X}\odot\mathbf{X}^{\star})^{\mathrm{T}},$$

$$\mathbf{P} = -(c_1 + c_2)\begin{bmatrix} (\mathbf{x}_1 \otimes \mathbf{x}_1^{\star})^{\mathrm{T}} \\ \vdots \\ (\mathbf{x}_N \otimes \mathbf{x}_N^{\star})^{\mathrm{T}} \end{bmatrix} = -(c_1 + c_2)(\mathbf{X}\odot\mathbf{X}^{\star})^{\mathrm{T}},$$

$$\mathbf{u} = \mathbf{w}\otimes\mathbf{w}^{\star}\otimes\mathbf{w}\otimes\mathbf{w}^{\star}, \quad \mathbf{v} = \mathbf{w}\otimes\mathbf{w}^{\star}. \tag{7}$$

Equivalently, one can write

$$\mathbf{T}\mathbf{z} = \alpha\mathbf{1},$$

with $\mathbf{T} = \begin{bmatrix}\mathbf{R} & \mathbf{P}\end{bmatrix}$ and $\mathbf{z} = \begin{bmatrix}\mathbf{u}; \mathbf{v}\end{bmatrix}$. Each vector $\mathbf{w}$ extracting a multi-modulus source satisfies constraint (5) and gives a solution to (6). Vice versa, if we find a solution vector of system (6) with $\mathbf{u}$ and $\mathbf{v}$ constrained as in (7), $\mathbf{w}$ is a valid separation vector of the MIMO BSS problem. In total, we need to find R different vectors $\mathbf{w}_r$, yielding the separation matrix $\mathbf{W}$.

B. Omitting identical columns from the linear system

The Kronecker products introduce identical columns in $\mathbf{R}$, which should be removed. We define a reduced matrix $\mathbf{R}^{\oslash}$ including only the distinct columns of $\mathbf{R}$ multiplied by the number of occurrences, cf. [42]. Equivalently, $\mathbf{u}^{\oslash}$ is introduced containing only the unique elements of $\mathbf{u}$. Then holds:

$$\mathbf{R}\mathbf{u} + \mathbf{P}\mathbf{v} = \alpha\mathbf{1} \quad \Leftrightarrow \quad \mathbf{R}^{\oslash}\mathbf{u}^{\oslash} + \mathbf{P}\mathbf{v} = \alpha\mathbf{1}. \tag{8}$$

Considering K observed signals, the matrix $\mathbf{R}^{\oslash}$ has size $N \times \frac{1}{4}K^2(K+1)^2$. Note that there is no need to operate on $\mathbf{P}$ and $\mathbf{v}$, unless both $\mathbf{M}$ and $\mathbf{S}$ are real. Alternatively, we can consider the system $\mathbf{T}^{\oslash}\mathbf{z}^{\oslash} = \alpha\mathbf{1}$ with the matrix $\mathbf{T}^{\oslash} = \begin{bmatrix}\mathbf{R}^{\oslash} & \mathbf{P}\end{bmatrix}$ and the vector $\mathbf{z}^{\oslash} = \begin{bmatrix}\mathbf{u}^{\oslash}; \mathbf{v}\end{bmatrix}$.

C. Solving the system

The solutions of the system are found in the following procedure through the computation of a null space. To obtain a homogeneous linear system of equations, the right-hand side of (8) with all elements equal to α is rotated to the first coordinate axis. The first row of the system can then be dropped.

Let $\mathbf{Q}$ be any unitary matrix such that $\mathbf{Q1} = \left[N^{\frac{1}{2}}; \mathbf{0}\right]$. $\mathbf{Q}$ can correspond to a Householder or discrete Fourier transformation [43]. Now consider the left multiplication of the system with $\mathbf{Q}$ such that

$$\begin{aligned} &\mathbf{QR}^{\oslash}\mathbf{u}^{\oslash} + \mathbf{QPv} = \alpha\mathbf{Q1} \qquad (9)\\ \Leftrightarrow\ &\begin{bmatrix}\tilde{\mathbf{r}}_1^{\oslash}\\ \tilde{\mathbf{R}}^{\oslash}\end{bmatrix}\mathbf{u}^{\oslash} + \begin{bmatrix}\tilde{\mathbf{p}}_1\\ \tilde{\mathbf{P}}\end{bmatrix}\mathbf{v} = \begin{bmatrix}\alpha N^{\frac{1}{2}}\\ \mathbf{0}\end{bmatrix}. \end{aligned}$$

The vectors $\tilde{\mathbf{r}}_1^{\oslash}$ and $\tilde{\mathbf{p}}_1$ denote the first row of $\mathbf{QR}^{\oslash}$ and $\mathbf{QP}$, respectively, while $\tilde{\mathbf{R}}^{\oslash}$ and $\tilde{\mathbf{P}}$ consist of the other $N-1$ rows. By omitting the first equation, the following system is obtained, with $\tilde{\mathbf{R}}^{\oslash} \in \mathbb{C}^{(N-1)\times\frac{1}{4}K^2(K+1)^2}$ and $\tilde{\mathbf{P}} \in \mathbb{C}^{(N-1)\times K^2}$:

$$\begin{aligned} \tilde{\mathbf{R}}^{\oslash}\mathbf{u}^{\oslash} + \tilde{\mathbf{P}}\mathbf{v} = \mathbf{0} \quad &\Leftrightarrow \quad \begin{bmatrix}\tilde{\mathbf{R}}^{\oslash} & \tilde{\mathbf{P}}\end{bmatrix}\begin{bmatrix}\mathbf{u}^{\oslash}\\ \mathbf{v}\end{bmatrix} = \mathbf{0}, \qquad (10)\\ &\Leftrightarrow \quad \tilde{\mathbf{T}}^{\oslash}\mathbf{z}^{\oslash} = \mathbf{0}, \qquad (11)\end{aligned}$$

with the matrix $\tilde{\mathbf{T}}^{\oslash} = \begin{bmatrix}\tilde{\mathbf{R}}^{\oslash} & \tilde{\mathbf{P}}\end{bmatrix} \in \mathbb{C}^{(N-1)\times\frac{1}{4}K^2(K+1)^2+K^2}$. Ignoring the structure in $\mathbf{z}^{\oslash}$, we now focus on the null space or kernel of the matrix $\tilde{\mathbf{T}}^{\oslash}$.

In the noiseless case, the dimension of the kernel is at least R. Indeed, each separation vector $\mathbf{w}_r$ yields a solution $\mathbf{z}_r^{\oslash} = \left[\mathbf{u}_r^{\oslash}; \mathbf{v}_r\right]$ via (11), and these R vectors are linearly independent if the R separation vectors $\mathbf{w}_r$ are independent. This is equivalent with $\mathbf{W}$ having full column rank which is a valid working assumption if $K \geq R$.

Let us first consider the case in which the matrix $\tilde{\mathbf{T}}^{\oslash}$ has more rows than columns minus the kernel dimension (i.e., $N - 1 \geq \frac{1}{4}K^2(K+1)^2 + K^2 - R$; we refer to Section IV for an alternative procedure if this is not satisfied). The dimension of the kernel is at most R unless there are significant (phase) relations between the source signals. This can occur if the (phases of the) samples are not sufficiently random, cf. the derivations and discussions in [6], [15], [44], [45]. Furthermore, we give the following theorem, which we prove in Appendix B:

Theorem 1. *If $N - 1 \geq \frac{1}{4}K^2(K+1)^2 + K^2 - R$, $\tilde{\mathbf{T}}^{\oslash}$ has full column rank for generic $\mathbf{S}$ in* (1)*.*

Summarizing, one can assume that the dimension of the kernel of $\tilde{\mathbf{T}}^{\oslash}$ is exactly R, being the number of separation vectors $\mathbf{w}_r$. Consider a basis $\{\mathbf{f}_1^{\oslash}, \ldots, \mathbf{f}_R^{\oslash}\}$ of the kernel of $\tilde{\mathbf{T}}^{\oslash}$. If the observed signals are perturbed by noise, one can use the R smallest singular vectors of $\tilde{\mathbf{T}}^{\oslash}$ for $\{\mathbf{f}_1^{\oslash}, \ldots, \mathbf{f}_R^{\oslash}\}$.

D. Recovery of the separation matrix

The basis vectors can be partitioned as follows, and also expanded again by reintroducing the repeated elements:

$$\mathbf{f}_r^{\oslash} = \begin{bmatrix}\mathbf{f}_r^{\mathbf{u}^{\oslash}}\\ \mathbf{f}_r^{\mathbf{v}}\end{bmatrix} \in \mathbb{C}^{\frac{1}{4}K^2(K+1)^2+K^2} \ \Leftrightarrow\ \mathbf{f}_r = \begin{bmatrix}\mathbf{f}_r^{\mathbf{u}}\\ \mathbf{f}_r^{\mathbf{v}}\end{bmatrix} \in \mathbb{C}^{K^4+K^2}.$$

As $\mathbf{W}$ is assumed to have full column rank, $\{\mathbf{z}_1, \ldots, \mathbf{z}_R\}$ is a linearly independent set. Hence, it is a basis itself, and each basis vector $\mathbf{f}_r$ is a linear combination of the vectors from $\{\mathbf{z}_1, \ldots, \mathbf{z}_R\}$:

$$\begin{aligned} \forall r: \quad &\mathbf{f}_r = \lambda_{r,1}\mathbf{z}_r + \ldots + \lambda_{r,R}\mathbf{z}_r, \qquad (12)\\ \Leftrightarrow \forall r: \quad &\begin{bmatrix}\mathbf{f}_r^{\mathbf{u}}\\ \mathbf{f}_r^{\mathbf{v}}\end{bmatrix} = \lambda_{r,1}\begin{bmatrix}\mathbf{u}_1\\ \mathbf{v}_1\end{bmatrix} + \ldots + \lambda_{r,R}\begin{bmatrix}\mathbf{u}_R\\ \mathbf{v}_R\end{bmatrix}. \end{aligned}$$

The coefficients can be collected in a matrix $\mathbf{\Lambda} \in \mathbb{C}^{R\times R}$. Note that this matrix has full rank. Equivalently, we can write:

$$\begin{aligned} \forall r: \quad \mathbf{f}_r^{\mathbf{u}} &= \lambda_{r,1}\left(\mathbf{w}_1 \otimes \mathbf{w}_1^{\star} \otimes \mathbf{w}_1 \otimes \mathbf{w}_1^{\star}\right) + \ldots\\ &\quad + \lambda_{r,R}\left(\mathbf{w}_R \otimes \mathbf{w}_R^{\star} \otimes \mathbf{w}_R \otimes \mathbf{w}_R^{\star}\right),\\ \forall r: \quad \mathbf{f}_r^{\mathbf{v}} &= \lambda_{r,1}\left(\mathbf{w}_1 \otimes \mathbf{w}_1^{\star}\right) + \ldots + \lambda_{r,R}\left(\mathbf{w}_R \otimes \mathbf{w}_R^{\star}\right). \end{aligned}$$

These equations can now be expressed in a tensor format. Let us reshape the vectors $\mathbf{f}_r^{\mathbf{u}} \in \mathbb{C}^{K^4}$ to fourth-order tensors, and stack them along a fifth mode for $r = 1, \ldots, R$ to obtain the tensor $\mathcal{F}^{\mathbf{u}} \in \mathbb{C}^{K\times K\times K\times K\times R}$. Likewise, we construct a third-order tensor $\mathcal{F}^{\mathbf{v}} \in \mathbb{C}^{K\times K\times R}$ by stacking the matricized versions of $\mathbf{f}_r^{\mathbf{v}}$ along a third mode. Both tensors $\mathcal{F}^{\mathbf{u}}$ and $\mathcal{F}^{\mathbf{v}}$ have rank R, and we obtain two CPDs (note that the column-wise matricization of $\mathbf{w} \otimes \mathbf{w}^{\star}$ is equal to $\mathbf{w}^{\star} \otimes \mathbf{w}$):

$$\mathcal{F}^{\mathbf{u}} = \sum_{r=1}^{R} \mathbf{w}_r^{\star} \otimes \mathbf{w}_r \otimes \mathbf{w}_r^{\star} \otimes \mathbf{w}_r \otimes \boldsymbol{\lambda}_r = \left[\!\left[\mathbf{W}^{\star}, \mathbf{W}, \mathbf{W}^{\star}, \mathbf{W}, \mathbf{\Lambda}\right]\!\right],$$

$$\mathcal{F}^{\mathbf{v}} = \sum_{r=1}^{R} \mathbf{w}_r^{\star} \otimes \mathbf{w}_r \otimes \boldsymbol{\lambda}_r = \left[\!\left[\mathbf{W}^{\star}, \mathbf{W}, \mathbf{\Lambda}\right]\!\right].$$

with $\boldsymbol{\lambda}_r$ the rth column of $\mathbf{\Lambda}$. Summarizing, the equations show that the solutions of the original system of (6) can be partitioned, and, when reshaped into tensors, can be decomposed using a CPD to recover the separation vectors. This is an interesting result, in multiple ways, as we can now exploit knowledge on tensors and multilinear algebra.

E. Interpretation in a tensor framework

Let us interpret the previous results in a tensor setting. First, a CPD is *essentially unique* under mild conditions. This means that the factor matrices can be recovered up to scaling and permutation. Note that these are the basic ambiguities in BSS. We can apply the uniqueness results already obtained by Harshman: as it is assumed that the matrix $\mathbf{W}$ has full rank and because $\mathbf{\Lambda}$ does not contain proportional columns (as it has full rank), the decomposition of $\mathcal{F}^{\mathbf{v}}$ is unique [46]. Furthermore, as both $\mathbf{W}$ and $\mathbf{\Lambda}$ have full rank, it is shown in [47] that the CPD of $\mathcal{F}^{\mathbf{v}}$ can be computed by means of the eigenvalue decomposition (EVD). An algebraic technique using the generalized EVD (GEVD) can be found in [48] and references therein.

Hence, the CPD of $\mathcal{F}^{\mathbf{v}}$ is essentially unique and decomposing $\mathcal{F}^{\mathbf{v}}$ is enough to recover $\mathbf{W}$. There is no need to decompose or even construct the larger tensor $\mathcal{F}^{\mathbf{u}}$. Besides deterministic conditions on $\mathbf{W}$ and/or $\mathbf{\Lambda}$, there also exist generic uniqueness conditions, for which we refer the interested reader to [49], [37].

Second, while it might be sufficient to use $\mathcal{F}^{\mathbf{v}}$, a higher accuracy might be obtained by resorting to the decomposition of the larger tensor $\mathcal{F}^{\mathbf{u}}$.

Third, the CPDs are partially symmetric. For example, $\mathbf{W}$ is a factor matrix of $\mathcal{F}^{\mathbf{u}}$ in the second and fourth mode. The factor matrices $\mathbf{W}$ and $\mathbf{W}^\star$ are also related to each other. This can be exploited by customized tensor decomposition algorithms.

Fourth, the tensors are coupled through the matrices $\mathbf{W}$ and $\mathbf{\Lambda}$. This structure can be taken into account by using, e.g., the structured data fusion framework implemented in Tensorlab [28], [50], among others. An overview of data fusion models is given in [30]. By considering the coupling, the uniqueness conditions can be relaxed [33]. The accuracy might increase too.

Furthermore, it is possible to reduce the coupled decompositions of $\mathcal{F}^{\mathbf{u}}$ and $\mathcal{F}^{\mathbf{v}}$ to a single tensor decomposition through relaxation. For example, consider the unfolding of the fifth-order tensor $\mathcal{F}^{\mathbf{u}}$ into a third-order tensor $\tilde{\mathcal{F}}^{\mathbf{u}} \in \mathbb{C}^{K^3 \times K \times R}$. The tensor $\tilde{\mathcal{F}}^{\mathbf{u}}$ then admits the CPD $\tilde{\mathcal{F}}^{\mathbf{u}} = [\![\mathbf{Z}, \mathbf{W}^\star, \mathbf{\Lambda}]\!]$ with $\mathbf{Z} = \mathbf{W} \otimes \mathbf{W}^\star \otimes \mathbf{W} \in \mathbb{C}^{K^3 \times R}$. Now consider the concatenation of the tensors $\tilde{\mathcal{F}}^{\mathbf{u}}$ and $\mathcal{F}^{\mathbf{v}}$ along the first mode, defining the concatenated tensor $\mathcal{G} \in \mathbb{C}^{(K^3+K)\times K \times R}$. This tensor $\mathcal{G}$ admits the CPD $\mathcal{G} = [\![\mathbf{B}, \mathbf{W}^\star, \mathbf{\Lambda}]\!]$ with $\mathbf{B} = \left[(\mathbf{W} \otimes \mathbf{W}^\star \otimes \mathbf{W})^{\mathrm{T}} \ \mathbf{W}^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{C}^{(K^3+K)\times R}$. Alternatively, $\mathcal{G}$ can be reshaped to a tensor $\mathcal{H} \in \mathbb{C}^{(K^2+1)\times K^2 \times R}$ which admits the CPD $\mathcal{H} = [\![\mathbf{C}, \mathbf{D}, \mathbf{\Lambda}]\!]$, with the matrices $\mathbf{D} = \mathbf{W} \otimes \mathbf{W}^\star \in \mathbb{C}^{K^2 \times R}$ and $\mathbf{C} = \left[\mathbf{D}; \mathbf{1}\right] \in \mathbb{C}^{(K^2+1)\times R}$ where $\mathbf{1}$ is the vector of size $1 \times R$ with all ones. Hence, the coupled tensor decomposition of $\mathcal{F}^{\mathbf{u}}$ and $\mathcal{F}^{\mathbf{v}}$ has been reduced to a single CPD of $\mathcal{G}$ or $\mathcal{H}$. It is also possible to express the CPD of $\mathcal{F}^{\mathbf{u}}$ as the coupled decomposition of third-order tensors. For details on coupled CPDs we refer the reader to [33], [34].

Finally, note that the structured data fusion framework from Tensorlab allows the user to add regularization to the tensor decompositions or additionally impose structure on $\mathbf{W}$, such as sparseness or non-negativity.

F. Note concerning rank deficiency of the mixing matrix

So far, a full-rank mixing matrix $\mathbf{M}$ has been considered, together with its corresponding full-rank separation matrix $\mathbf{W}$. It has been discussed under which circumstances the dimension of the null space of $\tilde{\mathbf{T}}^{\oslash}$ corresponds to the number of different separation vectors that can be found. In this subsection, it is shown that even if $\mathbf{M}$ is rank-deficient, some source signals can be recovered.

Let us illustrate this with an example that involves four multi-modulus source signals and four observed signals. Let us assume that $\mathbf{m}_4 = \alpha\mathbf{m}_2 + \beta\mathbf{m}_3$ with $\alpha \neq 0, \beta \neq 0$. The vectors $\mathbf{m}_1, \mathbf{m}_2$ and $\mathbf{m}_3$ are linearly independent, and $\mathbf{M} \in \mathbb{C}^{4\times 4}$ has rank 3. Then one can write:

$$\mathbf{X} = \begin{bmatrix}\mathbf{m}_1 & \mathbf{m}_2 & \mathbf{m}_3 & \mathbf{m}_4\end{bmatrix} \begin{bmatrix}\mathbf{s}_1 \\ \mathbf{s}_2 \\ \mathbf{s}_3 \\ \mathbf{s}_4\end{bmatrix} = \underbrace{\begin{bmatrix}\mathbf{m}_1 & \mathbf{m}_2 & \mathbf{m}_3\end{bmatrix}}_{\tilde{\mathbf{M}}} \underbrace{\begin{bmatrix}\mathbf{s}_1 \\ \mathbf{s}_2 + \alpha\mathbf{s}_4 \\ \mathbf{s}_3 + \beta\mathbf{s}_4\end{bmatrix}}_{\tilde{\mathbf{S}}}.$$

Note that, in general, a linear combination of multi-modulus signals is not a multi-modulus signal. Hence, we are unable to recover $\mathbf{s}_2 + \alpha\mathbf{s}_4$ and $\mathbf{s}_3 + \beta\mathbf{s}_4$ using the proposed method. The vector $\mathbf{s}_1$ can be recovered, as there exists a column vector $\mathbf{w} \in \mathbb{C}^4$ such that $\mathbf{w}^{\mathrm{T}}\mathbf{X} = \mathbf{w}^{\mathrm{T}}\tilde{\mathbf{M}}\tilde{\mathbf{S}} = \mathbf{s}_1$.

Generally, let us assume that $\mathbf{M} \in \mathbb{C}^{K\times R}$ has rank $U < R$. $\mathbf{X} = \mathbf{M}\mathbf{S}$ can then be expressed as $\mathbf{X} = \tilde{\mathbf{M}}\tilde{\mathbf{S}}$ with full-rank matrices $\tilde{\mathbf{M}} \in \mathbb{C}^{K\times U}$ and $\tilde{\mathbf{S}} \in \mathbb{C}^{U\times N}$. Let us assume that $\mathbf{M}$ has Z column vectors that are, each, linearly independent of the $R-1$ other vectors. Only the source signals corresponding to these mixing vectors can be recovered, as explained in the example. The null space of $\tilde{\mathbf{T}}^{\oslash}$ has dimension Z, and the coupled tensor decompositions can be used to recover the Z corresponding separation vectors. Note that strictly $Z < U$; if Z was equal to the rank U of $\mathbf{M}$, each of the other $R - Z$ column vectors of $\mathbf{M}$ would be linearly dependent on one or more of the Z column vectors, which is a contradiction.

Note that the source signals corresponding to the other mixing vectors are either removed from $\tilde{\mathbf{S}}$ or are replaced with a linear combination of multi-modulus signals (which is not multi-modulus, in general), as illustrated in the previous example. As the multi-modulus property is lost, they can not be recovered.

IV. Multi-modulus BSS using a rank-1 detection procedure

We ignored the structure of the vectors $\mathbf{z}_r$ in Section III-C so that (11) could be seen as a linear least-squares problem. Let us now exploit the rank-1 structure of $\mathbf{z}_r$ to extract the relevant vectors $\mathbf{w}_r$. The technique of detecting Kronecker-structured vectors (or reshaped rank-1 matrices) is explained in general in the first subsection. In the second subsection, we explain how the rank-1 detection technique can be used in conjunction with the previously proposed multi-modulus BSS procedure. The method allows us to reduce the number of samples needed.

A. Rank-1 detection procedure

Consider a matrix $\mathbf{E} \in \mathbb{C}^{IJ\times R_{\text{tot}}}$ that can be written as

$$\mathbf{E} = (\mathbf{A} \odot \mathbf{B})\,\mathbf{G}^{\mathrm{T}}, \tag{13}$$

with $\mathbf{A} \in \mathbb{C}^{I\times R}$, $\mathbf{B} \in \mathbb{C}^{J\times R}$, and $\mathbf{G} \in \mathbb{C}^{R_{\text{tot}}\times R}$. We assume that $\mathbf{A} \odot \mathbf{B}$ and $\mathbf{G}$ have full column rank [35]. Equation (13) is equivalent to a matrix representation of a polyadic decomposition. Given $\mathbf{E}$, consider now the generic problem of recovering $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{G}$ (up to (counter)scaling and permutation). In [35], this problem was solved by searching for the matrix $\mathbf{W} = (\mathbf{G}^{\mathrm{T}})^{\dagger} \in \mathbb{C}^{R_{\text{tot}}\times R}$ such that

$$\mathbf{E}\mathbf{W} = \mathbf{A} \odot \mathbf{B}.$$

The column space of $\mathbf{E}$ can be represented by a basis of only Kronecker-structured vectors. The goal is then to detect these vectorized rank-1 matrices; hence, the 'rank-1 detection' terminology. Note that an additional mild necessary condition on $\mathbf{A}$ and $\mathbf{B}$ was specified in [35].

Now consider a more general $\mathbf{E} \in \mathbb{C}^{IJ\times R_{\text{tot}}}$ of the form

$$\mathbf{E} = (\mathbf{A} \odot \mathbf{B})\,\mathbf{G}^{\mathrm{T}} + \mathbf{C}\mathbf{H}^{\mathrm{T}} = \begin{bmatrix}\mathbf{A} \odot \mathbf{B} & \mathbf{C}\end{bmatrix}\begin{bmatrix}\mathbf{G} & \mathbf{H}\end{bmatrix}^{\mathrm{T}}, \tag{14}$$

with $\mathbf{C} \in \mathbb{C}^{IJ\times R_{\text{add}}}$ and $\mathbf{H} \in \mathbb{C}^{R_{\text{tot}}\times R_{\text{add}}}$. The subscript of R_{add} stands for 'additional'. We assume both $\begin{bmatrix}\mathbf{A} \odot \mathbf{B} & \mathbf{C}\end{bmatrix}$

and $\begin{bmatrix}\mathbf{G} & \mathbf{H}\end{bmatrix}$ have full column rank. The column space of $\mathbf{E}$ cannot be represented by a basis of only Kronecker-structured vectors anymore — also other vectors are needed — and the problem is not equivalent to finding a CPD anymore.

It is however still possible to detect the different Kronecker-structured vectors in the column space of $\mathbf{E}$ constructed from (14) up to machine precision. We will use Algorithm 1 given in [35, Algorithm 2.1, steps 3-10]. The algorithm was discussed in [35] for only $R_{\text{add}} = 0$. Perhaps surprisingly, the algorithm is guaranteed to work in the exact case as well for $R_{\text{add}} > 0$, provided R_{tot} is sufficiently small with respect to I, J (see further).

For a detailed explanation of the algorithm, we refer to [35]. Here, we briefly comment on some steps. Consider the mapping $\Phi : (\mathbf{X}, \mathbf{Y}) \in \mathbb{C}^{I\times J} \times \mathbb{C}^{I\times J} \to \Phi(\mathbf{X}, \mathbf{Y}) \in \mathbb{C}^{I\times I\times J\times J}$ with its resulting tensor $\mathcal{P}$ defined by

$$(\mathcal{P})_{ijkl} = (\Phi(\mathbf{X}, \mathbf{Y}))_{ijkl} = x_{ij}y_{il} + y_{ik}x_{jl} - x_{il}y_{jk} - y_{il}x_{jk}.$$

Let us review two properties of the mapping Φ. First, it is clear that the mapping Φ is bilinear in $\mathbf{X}$ and $\mathbf{Y}$, i.e., if $\mathbf{X} = \sum_{m=1}^{M} \alpha_m \mathbf{X}_m$ and $\mathbf{Y} = \sum_{n=1}^{N} \beta_n \mathbf{Y}_n$ then

$$\mathcal{P} = \Phi(\mathbf{X}, \mathbf{Y}) = \sum_{m=1}^{M}\sum_{n=1}^{N} \alpha_m\beta_n \Phi(\mathbf{X}_m, \mathbf{Y}_n).$$

Second, consider the case where $\mathbf{Y} = \mathbf{X}$. If and only if $\mathbf{X}$ has rank 1, $\mathcal{P} = \Phi(\mathbf{X}, \mathbf{X})$ is an all-zero tensor. Hence, the mapping is able to distinguish between matrices with rank 1 and matrices with rank strictly greater than 1.

By applying Φ on reshaped versions of the columns of $\mathbf{E}$ as detailed in Steps 1 and 2, one obtains a matrix $\mathbf{P} \in \mathbb{C}^{I^2J^2\times R_{\text{tot}}^2}$ in Step 3. Under the necessary assumption that the tensors $\mathcal{P}_{u\tilde{u}}$ are linearly independent for $1 \le u < \tilde{u} \le R_{\text{tot}}$, a CPD can be obtained from the kernel of $\mathbf{P}$ as follows. If the tensors are not linearly independent, the method does not work.

Rather than the whole kernel of $\mathbf{P}$, we consider a specific subspace. Let us define the 'symmetric kernel' of $\mathbf{P}$ as the intersection of the kernel and the space spanned by vectorized symmetric matrices. If there are R linearly independent Kronecker-structured vectors in the column space of $\mathbf{E}$, then the dimension of the symmetric kernel of $\mathbf{P}$ is equal to R. $\mathbf{M}$ in Step 4 contains a basis of this symmetric kernel. Through the multilinearity of Φ, it is shown in [35] that a reshaped tensor version of $\mathbf{M}$ admits the CPD $\mathcal{M} = [\![\mathbf{W}, \mathbf{W}, \Theta]\!]$.

Note that the kernel of $\mathbf{P}^{\oslash} \in \mathbb{C}^{I^2J^2\times\frac{1}{2}R_{\text{tot}}(R_{\text{tot}}+1)}$ is equal to the kernel of $\mathbf{P}^{\oslash\text{H}}\mathbf{P}^{\oslash} \in \mathbb{C}^{\frac{1}{2}R_{\text{tot}}(R_{\text{tot}}+1)\times\frac{1}{2}R_{\text{tot}}(R_{\text{tot}}+1)}$; calculating the kernel of the matrix $\mathbf{P}^{\oslash\text{H}}\mathbf{P}^{\oslash}$ might be computationally more efficient as the latter matrix is smaller. $\mathbf{P}^{\oslash\text{H}}\mathbf{P}^{\oslash}$ can also be constructed in an efficient way. Additionally, the number of rows from $\mathbf{P}^{\oslash}$ can be reduced to $\frac{1}{4}I(I+1)J(J+1)$ by removing redundant rows. Step 3 in Algorithm 1 can be computationally demanding for large R_{tot}, as the number of elements in $\mathbf{P}^{\oslash\text{T}}\mathbf{P}^{\oslash}$ is of the order of magnitude $\mathcal{O}(R_{\text{tot}}^4)$.

A critical assumption in Algorithm 1 is that the tensors $\mathcal{P}_{u\tilde{u}}$ are linearly independent for $u < \tilde{u}$. A generic bound was given in [35, p. 656] for $R_{\text{add}} = 0$:

$$\frac{R(R-1)}{2} \le \frac{I(I-1)J(J-1)}{4}. \tag{15}$$

We now conjecture the following bound for $R_{\text{add}} \ge 0$:

$$\frac{R_{\text{tot}}(R_{\text{tot}}+1)}{2} - R \le \frac{I(I-1)J(J-1)}{4}. \tag{16}$$

We have verified this bound empirically for various combinations of R, R_{add}, I and J. One can see that the bound in (15) is the specific case for $R_{\text{add}} = 0$ or, equivalently, $R_{\text{tot}} = R$.

Algorithm 1: Rank-1 detection procedure

Input : Matrix $\mathbf{E}$

Output: Matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{W}$ such that $\mathbf{EW} = \mathbf{A} \odot \mathbf{B}$

1) Construct a tensor $\mathcal{E} \in \mathbb{C}^{I\times J\times R_{\text{tot}}}$ as follows:
$$\forall i, j, u: \quad (\mathcal{E})_{iju} = (\mathbf{E})_{(i-1)J+j,u}.$$
2) Compute $\mathcal{P}_{u\tilde{u}} \in \mathbb{C}^{I\times I\times J\times J}, 1 \le u, \tilde{u} \le R_{\text{tot}}$ as follows:
$$\forall i, j, k, l: \quad (\mathcal{P}_{u\tilde{u}})_{ijkl} = e_{iku}e_{jl\tilde{u}} + e_{ik\tilde{u}}e_{jlu} - e_{ilu}e_{jk\tilde{u}} - e_{il\tilde{u}}e_{jku}.$$
3) Reshape each tensor $\mathcal{P}_{u\tilde{u}}$ to a vector of length I^2J^2 and stack the different vectors in the columns of a matrix $\mathbf{P} \in \mathbb{C}^{I^2J^2\times R_{\text{tot}}^2}$. Compute the symmetric kernel of $\mathbf{P}$ by
   - Removing identical columns to obtain $\mathbf{P}^{\oslash}$;
   - Computing the kernel of $\mathbf{P}^{\oslash}$ (or equivalently $\mathbf{P}^{\oslash\text{H}}\mathbf{P}^{\oslash}$);
   - Expanding the R computed kernel vectors again as discussed in subsection III-D to obtain a matrix $\mathbf{M} \in \mathbb{C}^{R_{\text{tot}}^2\times R}$.
4) Reshape the matrix $\mathbf{M}$ to the tensor $\mathcal{M} \in \mathbb{C}^{R_{\text{tot}}\times R_{\text{tot}}\times R}$.
5) Compute a rank-R CPD $\mathcal{M} = [\![\mathbf{W}, \mathbf{W}, \mathbf{\Theta}]\!]$.
6) Compute $\mathbf{Z} = \mathbf{EW}$. Reshape each vector $\mathbf{z}_r$ to a matrix of size $J \times I$ and compute a best rank-1 approximation to obtain $\mathbf{b}_r$ and $\mathbf{a}_r$.

B. Application in the multi-modulus setting

Recall the problem in subsection III-D of finding R structured vectors $\mathbf{z}_r = [\mathbf{w}_r \otimes \mathbf{w}_r^\star \otimes \mathbf{w}_r \otimes \mathbf{w}_r^\star;\ \mathbf{w}_r \otimes \mathbf{w}_r^\star]$ in the kernel of $\mathbf{P}$. One can see that $\mathbf{z}_r$ is a Kronecker-structured vector, as we can write $\mathbf{z}_r = [\mathbf{w}_r \otimes \mathbf{w}_r^\star;\ 1] \otimes (\mathbf{w}_r \otimes \mathbf{w}_r^\star) = [\mathbf{v}_r;\ 1] \otimes \mathbf{v}_r$.

In subsection III-C a tall $\tilde{\mathbf{T}}^{\oslash}$ is assumed, i.e., $N - 1 \ge \frac{1}{4}K^2(K+1)^2 + K^2 - R$. This limits the dimension of the kernel of $\tilde{\mathbf{T}}^{\oslash}$ to R such that the kernel is exactly spanned by the vectors $\mathbf{z}_r$. Using $\mathbf{z}_r = [\mathbf{v}_r;\ 1] \otimes \mathbf{v}_r$, the system in (12) can be expressed in the form (13). This can be solved using the algebraic Algorithm 1 as discussed in [35]. As $\mathbf{a}_r = [\mathbf{w}_r \otimes \mathbf{w}_r^\star;\ 1]$ and $\mathbf{b}_r = \mathbf{w}_r \otimes \mathbf{w}_r^\star$, the different $\mathbf{w}_r$ can be recovered up to the standard scaling and permutation ambiguities by using a best rank-1 approximation on a reshaped version of either $\mathbf{a}_r$ or $\mathbf{b}_r$, or in a coupled way.

However, if $N - 1 < \frac{1}{4}K^2(K+1)^2 + K^2 - R$, the dimension of the kernel is strictly larger than R, simply because $\tilde{\mathbf{T}}^{\oslash}$ is a wide matrix. The kernel is not only spanned by the vectors $\mathbf{z}_r$ but also by other arbitrary vectors. Let us define the dimension as $R_{\text{tot}} > R$. Generically, $R_{\text{tot}} = R + R_{\text{add}}$ with $R_{\text{add}} = \left(\frac{1}{4}K^2(K+1)^2 + K^2 - R\right) - (N-1)$. Rather than (13), (14) now describes the structure of the kernel of $\tilde{\mathbf{T}}^{\oslash}$. While the system in (12) and thus the approach in Section III is not valid in this setting, Algorithm 1 can still be used

as explained in Section IV-A to recover the R Kronecker-structured vectors $\mathbf{z}_r = [\mathbf{v}_r; 1] \otimes \mathbf{v}_r$ from the kernel of $\tilde{\mathbf{T}}^{\oslash}$ with dimension $R_{\text{tot}} > R$.

The general bound in (16) does not apply directly because the matrices $\mathbf{A}$ and $\mathbf{B}$ in (14) have a particular structure, as explained previously. We conjecture the following bound in the multi-modulus context, which we have verified for different values of R and N:

$$\frac{R_{\text{tot}}(R_{\text{tot}}+1)}{2} - R \leq \frac{K^2(K^2-1)}{4}\left[\frac{K^2(K^2-1)}{2}+1\right] - K\binom{K}{4},$$

with $R_{\text{tot}} = R + R_{\text{add}} = \frac{1}{4}K^2(K+1)^2 + K^2 - N + 1$ and with $\binom{K}{4} = \frac{K!}{4!(K-4)!}$ if $R \geq 4$ and zero otherwise. The required number of samples for 2, 3 and 4 source signals, each with two possible source moduli, is reduced to 8, 11 and 20 samples instead of 12, 43 and 113 samples, respectively.

The mapping was applied by considering the Kronecker structure $\mathbf{z}_r = [\mathbf{v}_r; 1] \otimes \mathbf{v}_r$. To obtain a more powerful version of the rank-1 detection technique, one can take into account the full Kronecker structure in $\mathbf{z}_r = [\mathbf{w}_r \otimes \mathbf{w}_r^\star \otimes \mathbf{w}_r \otimes \mathbf{w}_r^\star;\ \mathbf{w}_r \otimes \mathbf{w}_r^\star]$, working in analogy with [33], [34]. The mapping Φ is applied four times in total: three times on the first part of $\mathbf{z}_r$, focusing on $(\mathbf{w}_r \otimes \mathbf{w}_r^\star \otimes \mathbf{w}_r) \otimes \mathbf{w}_r^\star$, $(\mathbf{w}_r \otimes \mathbf{w}_r^\star) \otimes (\mathbf{w}_r \otimes \mathbf{w}_r^\star)$ and $\mathbf{w}_r \otimes (\mathbf{w}_r^\star \otimes \mathbf{w}_r \otimes \mathbf{w}_r^\star)$, and once on the second part of $\mathbf{z}_r$ consisting of $\mathbf{w}_r \otimes \mathbf{w}_r^\star$. Each mapping will provide a different matrix $\mathbf{P}$ in Step 3. By vertically stacking these matrices in a large matrix, the null space of interest can be estimated. A more relaxed bound can be obtained on the required number of samples. A detailed derivation is considered to be out of scope of this paper.

Summarizing, we have shown that the rank-1 detection procedure can extract the separation vectors from the kernel of $\tilde{\mathbf{T}}^{\oslash}$, even if the kernel is spanned by not only Kronecker-structured vectors, enabling a reduction of the number of samples required for the separation of the multi-modulus source signals.

V. Blind Deconvolution of Multi-Modulus Signals

In this section, we consider the equalization of MIMO systems of multi-modulus signals. The problem is more general than the instantaneous problem from Section III. The data model is first discussed in Subsection V-A. Subsections V-B and V-C discuss two different methods to solve the BSI problem.

A. Data model

We consider a following data model with system order L_s:

$$\forall k, n: \ x_k(n) = \sum_{r=1}^{R}\sum_{l=0}^{L_s} h_{kr}^{(l)} s_r(n-l).$$

where the noise term is omitted. The model can be written as

$$\forall n: \ \mathbf{x}_n = \sum_{l=0}^{L_s} \mathbf{H}^{(l)} \mathbf{s}_{n-l},$$

in which the matrices $\mathbf{H}^{(l)} \in \mathbb{C}^{K\times R}$ contain the unknown system coefficients for $l = 0, \ldots, L_s$; hence, the Z-transform $\mathcal{H}$ of the system can be written as follows:

$$\mathbf{H}[z] = \mathbf{H}^{(0)} + \mathbf{H}^{(1)} z^{-1} + \ldots + \mathbf{H}^{(L_s)} z^{-L_s}.$$

For $L_s = 0$, one obtains the instantaneous case from Section III. Alternatively, consider $\mathbf{H} = \begin{bmatrix}\mathbf{H}^{(0)} & \cdots & \mathbf{H}^{(L_s)}\end{bmatrix} \in \mathbb{C}^{K\times(L_s+1)R}$ and let us stack the vectors $\mathbf{s}_n, \mathbf{s}_{n-1}, \ldots, \mathbf{s}_{n-L_s}$ in one large vector $\mathbf{s}_n^{\square} \in \mathbb{C}^{(L_s+1)R}$ for $n = L_s+1, \ldots, N$. By concatenating the vectors $\mathbf{s}_n^{\square}$, a block-Toeplitz matrix $\mathbf{S}^{\square} \in \mathbb{C}^{(L_s+1)R\times(N-L_s)}$ is obtained such that

$$\mathbf{X} = \mathbf{H}\mathbf{S}^{\square}. \tag{17}$$

The superscript $\cdot^{\square}$ stands for the block-Toeplitz encapsulation.

We assume that it is possible to equalize the channel by means of a finite impulse response (FIR) filter $\mathbf{F}(z)$ with filter order L_x:

$$\mathbf{s}_n = \sum_{l=0}^{L_x} \mathbf{F}^{(l)} \mathbf{x}_{n-l},$$

with $\mathbf{F}^{(l)} \in \mathbb{C}^{R\times K}$ for $l = 0, \ldots, L_x$. This assumption is valid when $K > R$ and under some additional conditions on the coefficients [51], [52]. This bound can be further relaxed to $K \geq R$ when considering paraunitary systems as these can be equalized by an FIR filter of the same length; hence, $L_s = L_x = L$ [17], [53]. We consider general non-paraunitary systems in this section unless stated otherwise. Again, by collecting the vectors $\mathbf{x}_n, \mathbf{x}_{n-1}, \ldots, \mathbf{x}_{n-L_x}$ in a block-Toeplitz matrix $\mathbf{X}^{\square} \in \mathbb{C}^{(L_x+1)K\times(N-L_x)}$, it is possible to write $\mathbf{S} = \mathbf{F}\mathbf{X}^{\square}$ with $\mathbf{F} = \begin{bmatrix}\mathbf{F}^{(0)} & \cdots & \mathbf{F}^{(L_x)}\end{bmatrix} \in \mathbb{C}^{R\times(L_x+1)K}$.

B. Deconvolution using an increased number of source signals

Consider the case where $K \geq R(L_s+1)$; the matrix $\mathbf{H}$ has then more rows than columns and one can write $\mathbf{S}^{\square} = \mathbf{W}\mathbf{X}$ with $\mathbf{W} = \mathbf{H}^{\dagger} \in \mathbb{C}^{(L_s+1)R\times K}$. As a delayed multi-modulus signal is still multi-modulus, the techniques from Section III and IV can be used, assuming there are $(L_s+1)R$ different source signals instead of only R signals. Note that the block-Toeplitz structure in $\mathbf{S}^{\square}$ is not taken into account.

Otherwise, if $R < K < (L_s+1)R$, a smoothing technique can be applied in a preprocessing step. By delaying the observed signals, one can artificially construct a system such that a tall mixing matrix $\tilde{\mathbf{H}}$ is obtained. The techniques from Section III and IV can again be used to recover the original source signals. For further details on the smoothing technique we refer to e.g. [52], [51].

C. Deconvolution using the original number of source signals by exploiting the block-Toeplitz structure

The number of samples required for a unique decomposition depends directly on the number of source signals R, as discussed in Sections III and IV. By exploiting the block-Toeplitz structure of $\mathbf{S}^{\square}$, it is possible to convert the deconvolution problem into an instantaneous BSS problem involving only R source signals instead of $(L_s+1)R$, as required in the previous section.

Consider the problem from (17). Ignoring the multi-modulus constraint, this is a block-Toeplitz factorization of the matrix $\mathbf{X}$. This is just one factorization; let us consider an alternative factorization of $\mathbf{X}$:

$$\mathbf{X} = \mathbf{H}\mathbf{S}^{\square} = \mathbf{H}\mathbf{Z}^{-1}\mathbf{Z}\mathbf{S}^{\square} = \mathbf{U}\mathbf{V}^{\square},$$

with $\mathbf{U} = \mathbf{H}\mathbf{Z}^{-1}$, $\mathbf{V}^{\square} = \mathbf{Z}\mathbf{S}^{\square}$ and $\mathbf{Z} \in \mathbb{C}^{(L_s+1)R\times(L_s+1)R}$ an arbitrary invertible matrix. Consider the following matrix $\mathbf{\Sigma}$:

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{s}_{L_s+2} & \cdots & \cdots & \mathbf{s}_{N-L_s-1} & \cdots & \mathbf{s}_N \\ \vdots & \ddots & & & \ddots & \vdots \\ \mathbf{s}_1 & \cdots & \mathbf{s}_{L_s+2} & \cdots & \cdots & \mathbf{s}_{N-L_s-1} \end{bmatrix}^{\mathrm{T}}.$$

We can now proceed with a result that can for instance be found in [51], [54], [55]: if both $\mathbf{H}$ and $\mathbf{\Sigma}$ have full column rank, then $\mathbf{X} = \mathbf{U}\mathbf{V}^{\square}$ is a block-Toeplitz factorization of $\mathbf{X}$ if and only if $\mathbf{Z}$ is of the following form:

$$\mathbf{Z} = \mathbf{I}_{L_s+1} \otimes \mathbf{G},$$

in which $\mathbf{G} \in \mathbb{C}^{R\times R}$ is an invertible matrix. Let $\mathbf{V}$ be the data matrix corresponding to the block-Toeplitz matrix $\mathbf{V}^{\square}$, in the same way as $\mathbf{S}^{\square}$ is constructed from $\mathbf{S}$. Then as $\mathbf{V}^{\square} = (\mathbf{I}_{L_s+1} \otimes \mathbf{G})\,\mathbf{S}^{\square}$, we have that $\mathbf{V} = \mathbf{G}\mathbf{S}$. Hence, $\mathbf{V}^{\square}$ can be found from a block-Toeplitz factorization of $\mathbf{X}$ (which is a linear least-squares problem) and $\mathbf{V}$ can be derived from $\mathbf{V}^{\square}$. After $\mathbf{V}$ has been determined, what remains to do is to find an invertible matrix $\mathbf{G}$ such that $\mathbf{S}$ is multi-modulus. This is exactly the instantaneous BSS problem which has been discussed in Sections III and IV.

Stated otherwise: if $\mathbf{H}$ and $\mathbf{\Sigma}$ have full column rank, then the block-Toeplitz matrix factorization $\mathbf{X} = \mathbf{U}\mathbf{V}^{\square}$ reduces the deconvolution to an instantaneous BSS problem that involves only R source signals. For the solution of the linear least-squares block-Toeplitz problem to be unique, it is necessary that $\mathbf{\Sigma}$ has full column rank, which we can interpret as a working assumption. Second, the matrix $\mathbf{H} \in \mathbb{C}^{K\times(L_s+1)R}$ should have full column rank which requires that $K \geq (L_s+1)R$. Otherwise, if $R < K < (L_s+1)R$, it is again possible to apply the smoothing technique from Section V-B.

VI. Simulations

We investigate the behavior of the proposed multi-modulus method for a varying number of samples and signal-to-noise ratio (SNR), as well as the behavior of Algorithm 1. The multi-modulus technique is compared to three other methods: robustICA [56], ACMA [15] and a recently proposed MMA technique using Givens rotations [22], [57]. ICA has already been used before in a constant or multi-modulus context [58], and its use is valid for the following simulations as the source signals are mutually stochastically independent.

To calculate a CPD, various approaches exist such as the popular alternating least-squares (ALS) algorithm. We use a non-linear least-squares (NLS) approach, implemented in Tensorlab [50]. A generalized eigenvalue decomposition is used for the initialization, together with five random initializations. Each time, the final solution with the minimal cost function value is retained. The observed signals are first prewhitened.

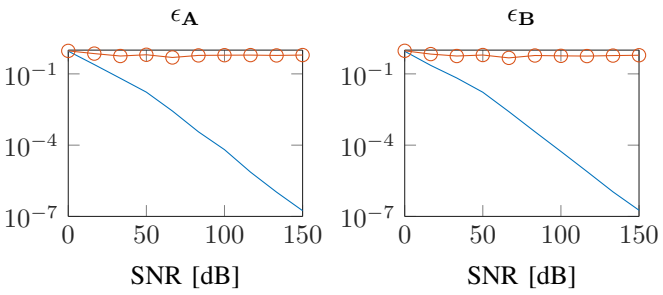


Fig. 2: Relative errors on $\mathbf{A}$ and $\mathbf{B}$ for the rank-1 detection simulation from Section VI-A using Algorithm 1 (——) and using a CPD with rank $R = 2$ (—○—) in function of the SNR.

For the coupled decompositions, the coupling constraints are incorporated using the dedicated `ccpd_nls` algorithm from Tensorlab, with the following cost function:

$$\min_{\mathbf{A},\mathbf{B},\mathbf{C}} \frac{\omega_1}{2} \left|\left|\mathcal{F}^{\mathbf{u}} - [\![\mathbf{B},\mathbf{A},\mathbf{B},\mathbf{A},\mathbf{C}]\!]\right|\right|^2 + \frac{\omega_2}{2} \left|\left|\mathcal{F}^{\mathbf{v}} - [\![\mathbf{B},\mathbf{A},\mathbf{C}]\!]\right|\right|^2,$$

in which the weights ω_1 and ω_2 are the inverse squared Frobenius norms of $\mathcal{F}^{\mathbf{u}}$ and $\mathcal{F}^{\mathbf{v}}$, respectively. The complex conjugation symmetry is not exploited; hence, the matrices $\mathbf{W}$ and $\mathbf{W}^{\star}$ are considered as different factor matrices $\mathbf{A}$ and $\mathbf{B}$, respectively. The matrix $\mathbf{\Lambda}$ is estimated as $\mathbf{C}$. For the initialization, we use a combination of the GEVD initializations of the lower-order and higher-order tensor, together with five additional random initializations.

To determine the relative error, we correct for scaling and permutation (the default indeterminacies in BSS) with respect to the theoretical sources. The relative error is then defined as the relative difference in Frobenius norm, e.g., if $\hat{\mathbf{S}}$ are the recovered sources after this step, we have a relative error $\epsilon_{\mathbf{S}} = \|\mathbf{S} - \hat{\mathbf{S}}\|/\|\mathbf{S}\|$. Second, the SNR is defined as the ratio of the power of the signal to the power of the Gaussian additive noise. For each experiment, the medians of the relative errors are given across 100 simulations. In each run, new realizations of source signals, noise signals and mixing matrix / channel coefficients are generated.

A. Simulation for rank-1 detection algorithm

A first simulation considers the rank-1 detection method in Algorithm 1. The matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{G}$ and $\mathbf{H}$ from (14) are randomly generated using a Gaussian distribution with $I = J = 10$, $U = 40$ and $R = 2$. Hence, while the dimension of the column space of $\mathbf{E}$ from (14) is 40, the dimension of the intersection of the column space and the space of Kronecker-structured vectors is only 2. Algorithm 1 is compared with a method which extracts the two Kronecker-structured vectors by applying a CPD with rank $R = 2$ on a reshaped version of $\mathbf{E}$. Fig. 2 illustrates the estimation error on $\mathbf{A}$ and $\mathbf{B}$ in function of the SNR on $\mathbf{E}$. It shows that the CPD approach fails as the influence of $\mathbf{C}\mathbf{H}^{\mathrm{T}}$ is not negligible, while Algorithm 1 delivers good results.

B. Simulations for the instantaneous case

A first experiment regarding multi-modulus BSS considers two ($R = 2$) multi-modulus continuous-phase shift keying

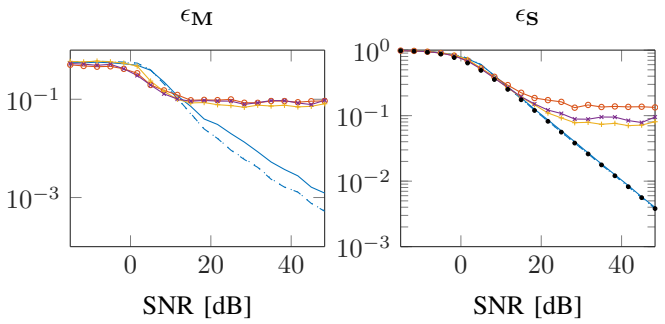


Fig. 3: Results for the first experiment in Section VI-B with CPSK source signals. Both the relative error on the mixing matrix (left) and on the source matrix (right) are shown for varying SNRs. The results obtained by the proposed methods are shown in blue. We discuss variants using only the lower-order tensor (——), only the higher-order tensor (- - -) and both in a coupled way (·····). The coupled and higher-order methods show a similar performance. They are compared with ICA (—○—), ACMA (—+—) and MMA using Givens rotations (—×—). The black dots (•) represent the case in which the exact mixing matrix is used to recover the source signals, a.k.a. the zero-forcing solution.

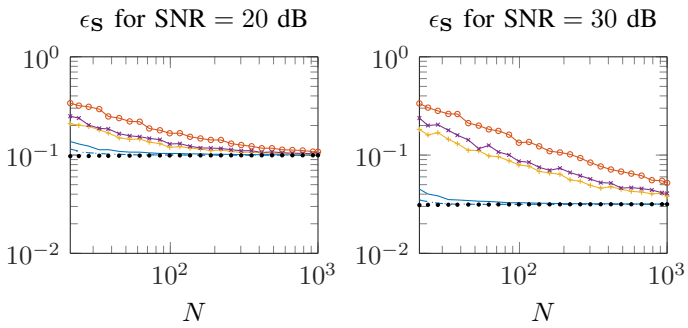


Fig. 4: Results for the second experiment in Section VI-B. The relative error on the source matrix is shown for SNRs of 20 dB (left) and 30 dB (right) for an increasing number of samples N. The labels are the same as in Fig. 3.

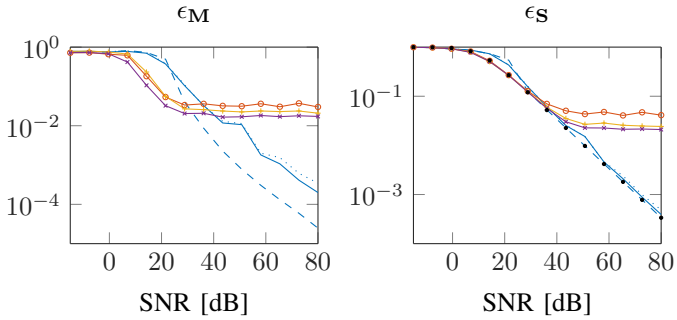


Fig. 5: Results for the third experiment in Section VI-B with 16-QAM source signals. Both the relative error on the mixing matrix (left) and on the source matrix (right) are shown for varying SNR. The labels are the same as in Fig. 3.

(CPSK) source signals of 100 samples each ($N = 100$), mixed into two ($K = 2$) observed signals. Each source sample has a uniform random phase drawn from $[0, 2\pi[$ and a modulus of 1 ($c_1 = 1$) or 2 ($c_2 = 4$) with equal probability. A unitary mixing matrix is used. Fig. 3 visualizes the relative errors on the mixture matrix and the source signals for varying SNRs. The proposed technique shows optimal asymptotic performance for increasing SNR, in contrast to the other methods. The proposed technique performs slightly worse at low SNR in terms of mixing matrix error; however, more important is the source recovery error, where the difference is negligible. The accuracy of ICA is limited because of the statistics estimation error. CMA is not suited in a multi-modulus context, and MMA optimizes a non-suitable cost function.

In a second experiment, we take a similar setting but vary the number of samples for an SNR of 20 dB and 30 dB, as shown in Fig. 4. The relative error is only weakly dependent on the number of samples and the solutions are about as good as the zero-forcing solutions. The experiment shows that not many samples are needed to reach asymptotic performance.

In the third experiment, we consider four rectangular 16-QAM source signals with $N = 2000$ samples. The samples are drawn from $\pm\{1, 3\} \pm \{1, 3\}i$ with equal probability, having a squared modulus of $c_1 = 2$, $c_2 = 10$ or $c_3 = 18$; hence, $P = 3$. We use a generalization of the technique for two moduli explained in Section III, allowing now three different moduli. Four instantaneously mixed signals are observed. For this experiment, a mixing matrix with condition number 10 is used to illustrate the performance in ill-conditioned situations. For $P = 3$, a set of coupled decompositions of a third-order tensor, a fifth-order tensor and a seventh-order tensor is obtained. Fig. 5 visualizes the results when decomposing the third-order tensor, the seventh-order tensor and both tensors coupled. In terms of the source error, the solutions are about as good as the zero-forcing solutions for increasing SNR and are only slightly worse for low SNRs. Fig. 6 visualizes the source signals, the observed signals and the recovered signals at 40 dB.

C. Simulations for the convolutive case

For the convolutive case, we consider a single experiment with two ($R = 3$) multi-modulus CPSK source signals of 500 samples each. The samples have a squared modulus of $c_1 = 1$ or $c_2 = 4$ and a random phase. A random system is used of order $L_s = 2$. Six signals ($K = 6$) are observed.

As it is not straightforward to compare transfer functions, we only report the error on the recovered deconvolved source signals in Fig. 7. Both the results from the method of subsection V-B with $(L_s + 1)R$ and the block-Toeplitz factorization method of subsection V-C with R source signals are given. For the former method, the technique works well for middle to high SNR, and performs slightly worse compared to the other algorithms for low SNR. The block-Toeplitz factorization procedure performs optimally.

VII. Conclusion and Discussion

We have proposed a new technique for multiple-input-multiple-output blind signal separation and blind system identification of multi-modulus signals such as 16-QAM signals. The method includes an analytical transformation to a set of CPDs by using subspace methods. We have shown that

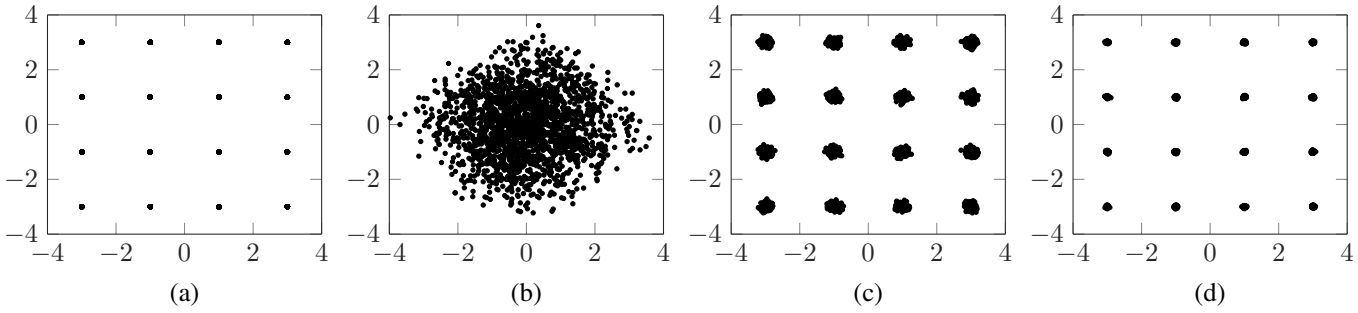


Fig. 6: Visualization in the complex plane of the different signals in the third experiment from Section VI-B at 40 dB. In (a), one of the 16-QAM source signals is shown. One of the observed mixed signals is shown in (b). In (c) and (d), an estimate of a source signal is shown (after compensating for scaling and permutation) by applying the ICA technique and the proposed technique, respectively.

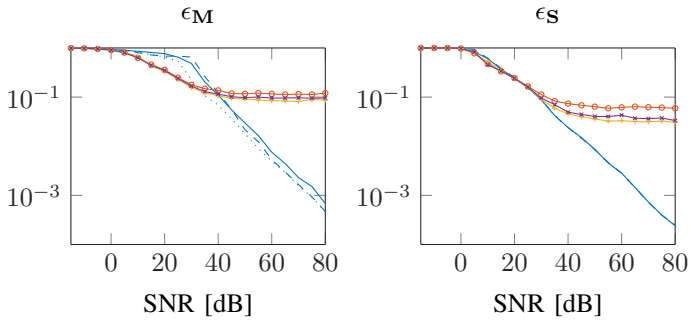


Fig. 7: Results for the experiment from Section VI-C with the blind deconvolution of two multi-modulus signals. On the left, the procedure of Section V-B is applied with $(L_s+1)R$ source signals. On the right, the procedure of Section V-C based on a block-Toeplitz factorization is applied with R source signals. The latter method clearly outperforms the former. The labels are the same as in Fig. 3.

a number of source signals can still be recovered in the case of a rank-deficient mixing matrix. As a side-result, an existing rank-1 detection procedure has been generalized to find Kronecker-structured vectors in a space spanned by both Kronecker-structured vectors and arbitrary vectors, and we include a conjectured bound on the number of vectors that can be handled. The generalized technique has been applied in the multi-modulus context to allow for a reduction in number of samples required for separation. The proposed methods can be interpreted as generalizations of the analytical constant modulus algorithm (ACMA). Some advantages of the tensor framework have been discussed, and the algorithm has been tested in different situations and compared to methods from the literature. An algebraic or exact solution can be obtained in the noiseless case for the proposed method, and the method reaches optimal asymptotic performance for increasing SNR.

In this paper, each source signal was assumed to be multi-modulus. If only Z source signals with $Z < R$ are multi-modulus, the dimension of the null space of $\tilde{\mathbf{T}}^{\oslash}$ in (11) is Z in general. The Z corresponding separation vectors can then be found using the techniques proposed in this paper, despite the presence of non-multi-modulus source signals.

In [59], a method was suggested for a BSS problem in which the source signal samples are assumed to have either constant modulus or zero modulus. The technique proposed in this paper can be extended to allow such signals. Eq. (5) then changes to

$$\forall n: \ \left(\mathbf{w}^{\mathrm{T}}\mathbf{x}_n\mathbf{x}_n^{\mathrm{H}}\mathbf{w}^{\star}-c_1\right)\left(\mathbf{w}^{\mathrm{T}}\mathbf{x}_n\mathbf{x}_n^{\mathrm{H}}\mathbf{w}^{\star}-c_2\right)\mathbf{w}^{\mathrm{T}}\mathbf{x}_n\mathbf{x}_n^{\mathrm{H}}\mathbf{w}^{\star}=0.$$

One can work in analogy with Section III to obtain an algebraic method.

Appendix A
Comparison with the MMA cost function

In [9], a cost function with multiple dispersion constants D_R and D_I (also defined as the in-phase and quadrature moduli, respectively) was designed for separating QAM signals:

$$J=\sum_{r=1}^{R}\sum_{n=1}^{N}\left[\left(\Re\{s_{rn}\}^2-D_R^2\right)^2+\left(\Im\{s_{rn}\}^2-D_I^2\right)^2\right]. \tag{18}$$

It has been coined the multi-modulus algorithm (MMA) cost function in [10], [11]. Typically, one uses $D_R = D_I$, as a compromise value which acts to project the constellation points onto the same circle. The cost function has been reused in many subsequent publications on the separation or deconvolution of multi-modulus signals such as [19], [20], [21].

Let us recall the cost function from Eq. (4) for $P=2$:

$$\tilde{J}=\sum_{r=1}^{R}\sum_{n=1}^{N}\left[\left(|s_{rn}|^2-c_1\right)\left(|s_{rn}|^2-c_2\right)\right]. \tag{19}$$

This cost function is only implicitly used in this paper. Given that $|s_{rn}|^2=\Re\{s_{rn}\}^2+\Im\{s_{rn}\}^2$, one can rework Eq. (19) to the following:

$$\tilde{J}=\sum_{r=1}^{R}\sum_{n=1}^{N}\Big[\left(\Re\{s_{rn}\}^2-F_R^2\right)^2+\left(\Im\{s_{rn}\}^2-F_I^2\right)^2 +\frac{1}{2}\left(4\Re\{s_{rn}\}^2\Im\{s_{rn}\}^2-c_1^2-c_2^2\right)\Big]$$

with $F_R=F_I=\sqrt{\frac{c_1+c_2}{2}}$. The constant terms c_1^2 and c_2^2 are not relevant for the minimization; hence, they can be omitted. We have now shown that both cost functions are not equivalent, but rather differ in one term:

$$\tilde{J}\equiv J+2\sum_{r=1}^{R}\sum_{n=1}^{N}\Re\{s_{rn}\}^2\Im\{s_{rn}\}^2.$$

The results in the paper have shown that the cost function J in (18) is not able to find the optimal solution, while the proposed cost function $\tilde{J}$ in (19) allows optimal asymptotic performance.

APPENDIX B
PROOF OF THEOREM 1

We first give the following lemma from [37], [60], [61], [62] which is used in the proof:

Lemma 1. *Given an analytic function $f : \mathbb{C}^N \to \mathbb{C}$. If there exists an element $\mathbf{x} \in \mathbb{C}^N$ such that $f(\mathbf{x}) \neq 0$, then the set $\{\mathbf{x}|f(\mathbf{x}) = 0\}$ is of Lebesgue measure zero.*

Let us also define the following matrices related to (6):

$$\mathbf{Z}_{\mathbf{X}} = \begin{bmatrix} \mathbf{X} \odot \mathbf{X}^\star \odot \mathbf{X} \odot \mathbf{X}^\star \\ \mathbf{X} \odot \mathbf{X}^\star \end{bmatrix}, \qquad \mathbf{Z}_{\mathbf{S}} = \begin{bmatrix} \mathbf{S} \odot \mathbf{S}^\star \odot \mathbf{S} \odot \mathbf{S}^\star \\ \mathbf{S} \odot \mathbf{S}^\star \end{bmatrix},$$
$$\mathbf{Z}_{\mathbf{X}}^{\boxslash} = \begin{bmatrix} (\mathbf{X} \odot \mathbf{X}^\star \odot \mathbf{X} \odot \mathbf{X}^\star)^{\boxslash} \\ \mathbf{X} \odot \mathbf{X}^\star \end{bmatrix}, \quad \mathbf{Z}_{\mathbf{S}}^{\boxslash} = \begin{bmatrix} (\mathbf{S} \odot \mathbf{S}^\star \odot \mathbf{S} \odot \mathbf{S}^\star)^{\boxslash} \\ \mathbf{S} \odot \mathbf{S}^\star \end{bmatrix}.$$

It can be seen that $\mathbf{Z}_{\mathbf{X}}^{\boxslash} = [\mathbf{R}^{\oslash}\ \mathbf{P}]^{\mathrm{T}}$ and $\mathbf{Z}_{\mathbf{X}} = [\mathbf{R}\ \mathbf{P}]^{\mathrm{T}}$. The operator $^{\boxslash}$ removes repeated rows, much like the operator $^{\oslash}$ from Section III removes repeated columns.

Proof of Theorem 1. The inequality states that $\tilde{\mathbf{T}}^{\oslash}$ should not have more columns than rows; hence, $\tilde{\mathbf{T}}^{\oslash}$ is assumed to be a tall matrix. Note that $\mathbf{Q}$ in (9) is a unitary matrix. To prove that $\tilde{\mathbf{T}}^{\oslash} = \mathbf{Q}\mathbf{Z}_{\mathbf{X}}^{\boxslash\,\mathrm{T}}$ has full column rank, it is thus sufficient to prove that $\mathbf{Z}_{\mathbf{X}}^{\boxslash\,\mathrm{T}}$ has full column rank or, equivalently, that $\mathbf{Z}_{\mathbf{X}}^{\boxslash}$ has full row rank.

From $\mathbf{X} = \mathbf{M}\mathbf{S}$, it can be seen that $\mathbf{Z}_{\mathbf{X}} = \mathbf{B}\mathbf{Z}_{\mathbf{S}}$ with $\mathbf{B} \in \mathbb{C}^{K^6 \times R^6}$ a block-diagonal matrix containing the blocks $\mathbf{M}\otimes\mathbf{M}^\star\otimes\mathbf{M}\otimes\mathbf{M}^\star$ and $\mathbf{M}\otimes\mathbf{M}^\star$ on the diagonal. Furthermore, it can be seen that $\mathbf{Z}_{\mathbf{X}}^{\boxslash} = \mathbf{B}^{\otimes}\mathbf{Z}_{\mathbf{S}}^{\boxslash}$. $\mathbf{B}^{\otimes}$ is obtained by removing the non-unique rows and columns from $\mathbf{B}$. As we assume that $\mathbf{M}$ has full rank, $\mathbf{B}^{\otimes}$ has full rank as well. Hence, it remains to show that $\mathbf{Z}_{\mathbf{S}}^{\boxslash}$ has full row rank for generic $\mathbf{S}$.

The latter problem is equivalent to showing that $\mathbf{Z}_{\mathbf{S}}^{\boxslash}$ has full row rank for at least one choice of $\mathbf{S}$. This follows from Lemma 1, as $\mathbf{S}$ has full row rank if and only if $\mathbf{S}\mathbf{S}^{\mathrm{T}}$ has a non-zero determinant. Hence, the analytic function mentioned in the lemma is $f(\mathbf{A}) = \det(\mathbf{A}\mathbf{A}^{\mathrm{T}})$.

We look for an example in the set of Vandermonde matrices. Let us assume that $\mathbf{S}$ is a transposed Vandermonde matrix:

$$\mathbf{S} = \begin{bmatrix} 1 & d_1 & \cdots & d_1^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & d_R & \cdots & d_R^{N-1} \end{bmatrix}.$$

Then, $\mathbf{S} \odot \mathbf{S}^\star$ is again a transposed Vandermonde matrix [60]:

$$\mathbf{S} \odot \mathbf{S}^\star = \begin{bmatrix} 1 & d_1 d_1^\star & \cdots & d_1^{N-1} d_1^{\star N-1} \\ 1 & d_1 d_2^\star & \cdots & d_1^{N-1} d_1^{\star N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & d_R d_R^\star & \cdots & d_R^{N-1} d_R^{\star N-1} \end{bmatrix},$$

with R^2 generators $d_i d_j^\star$ for $1 \leq i, j \leq R$. Similarly, $\mathbf{S} \odot \mathbf{S}^\star \odot \mathbf{S} \odot \mathbf{S}^\star$ is a transposed Vandermonde matrix with generators $d_i d_j^\star d_k d_l^\star$ for $1 \leq i, j, k, l \leq R$. Combined, $\mathbf{Z}_{\mathbf{S}}$ and $\mathbf{Z}_{\mathbf{S}}^{\boxslash}$ are Vandermonde matrices with $R^4 + R^2$ and $\frac{1}{4}R^2(R+1)^2 + R^2$ generators, respectively:

$$\begin{cases} d_i d_j^\star d_k d_l^\star, & 1 \leq i \leq k \leq R \ \text{ and } \ 1 \leq j \leq l \leq R, \\ d_i d_j^\star, & 1 \leq i, j \leq R. \end{cases}$$

We can now choose different values d_r such that the $\frac{1}{4}R^2(R+1)^2 + R^2$ generators are distinct. Then, $\mathbf{Z}_{\mathbf{S}}^{\boxslash}$ has full row rank. □

ACKNOWLEDGMENT

The authors wish to thank Ignat Domanov for discussions on the rank-1 detection problem and for providing useful tools to solve Theorem 1, and Frederik Van Eeghem and Mikael Sørensen for discussions on the block-Toeplitz factorization.

REFERENCES

[1] P. Comon and C. Jutten, *Handbook of blind source separation: Independent component analysis and applications*. Academic Press, 2010.
[2] K. Abed-Meraim, W. Qiu, and Y. Hua, "Blind system identification," *Proceedings of the IEEE*, vol. 85, no. 8, pp. 1310–1322, 1997.
[3] N. Sidiropoulos, R. Bro, and G. Giannakis, "Parallel factor analysis in sensor array processing," *IEEE Transactions on Signal Processing*, vol. 48, no. 8, pp. 2377–2388, 2000.
[4] O. Debals and L. De Lathauwer, "Stochastic and deterministic tensorization for blind signal separation," in *Latent Variable Analysis and Signal Separation*, ser. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2015, vol. 9237, pp. 3–13.
[5] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Transactions on Communications*, vol. 28, no. 11, pp. 1867–1875, 1980.
[6] J. Treichler and B. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 31, no. 2, pp. 459–472, 1983.
[7] R. Gooch and J. Lundell, "The CM array: An adaptive beamformer for constant modulus signals," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 11, 1986, pp. 2523–2526.
[8] J. Benesty and P. Duhamel, "Fast constant modulus adaptive algorithm," *IEE Proceedings F (Radar and Signal Processing)*, vol. 138, no. 4, pp. 379–387, 1991.
[9] K. N. Oh and Y. Chin, "New blind equalization techniques based on constant modulus algorithm," in *IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 2, Nov 1995, pp. 865–869.
[10] J. Yang, J.-J. Werner, and G. Dumont, "The multimodulus blind equalization algorithm," in *IEEE International Conference on Digital Signal Processing (DSP)*, vol. 1, 1997, pp. 127–130.
[11] J. Yang, "Multimodulus algorithms for blind equalization," Ph.D. dissertation, University of British Columbia, 1997.
[12] X.-L. Li and X.-D. Zhang, "A family of generalized constant modulus algorithms for blind equalization," *IEEE Transactions on Communications*, vol. 54, no. 11, pp. 1913–1917, 2006.
[13] A. Goupil and J. Palicot, "New algorithms for blind equalization: The constant norm algorithm family," *IEEE Transactions on Signal Processing*, vol. 55, no. 4, pp. 1436–1444, 2007.
[14] S. Abrar, "A family of reduced-constellation algorithms for blind equalization of square-QAM signals," in *Proceedings of the International Conference on Microelectronics (ICM)*. IEEE, 2005, pp. 296–300.
[15] A.-J. Van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Transactions on Signal Processing*, vol. 44, no. 5, pp. 1136–1155, 1996.
[16] A.-J. Van der Veen and A. Trindade, "Combining blind equalization with constant modulus properties," in *Proceedings of the Asilomar Conference on Signals, Systems and Computers*, vol. 2. IEEE, 2000, pp. 1568–1572.
[17] L. De Lathauwer, "Algebraic techniques for the blind deconvolution of constant modulus signals," in *Proceedings of the 12th European Signal Processing Conference (EUSIPCO)*, 2004, pp. 225–228.
[18] W. Sethares, G. Rey, C. R. Johnson Jr *et al.*, "Approaches to blind equalization of signals with multiple modulus," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 1989, pp. 972–975.

[19] S. Daumont and D. Le Guennec, "An analytical multimodulus algorithm for blind demodulation in a time-varying MIMO channel context," *International Journal Of Digital Multimedia Broadcasting*, vol. 2010, pp. 1–11, 2010.

[20] L. He, M. Amin, C. Reed Jr, and R. Malkemes, "A hybrid adaptive blind equalization algorithm for QAM signals in wireless communications," *IEEE Transactions on Signal Processing*, vol. 52, no. 7, pp. 2058–2069, 2004.

[21] K.-C. Hung, D. Lin, and C.-N. Ke, "Variable-step-size multimodulus blind decision-feedback equalization for high-order QAM based on boundary MSE estimation," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 4, 2004, pp. 881–884.

[22] S. A. W. Shah, K. Abed-Meraim, and T. Y. Al-Naffouri, "Multi-modulus algorithms using hyperbolic and Givens rotations for blind deconvolution of MIMO systems," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2015, pp. 2155–2159.

[23] S. Abrar, A. Roy, and J. Axford, "Sliced multi-modulus blind equalization algorithm," *ETRI Journal*, vol. 27, no. 3, pp. 257–266, 2005.

[24] T.-H. Li and K. Mbarek, "A blind equalizer for nonstationary discrete-valued signals," *IEEE Transactions on Signal Processing*, vol. 45, no. 1, pp. 247–254, 1997.

[25] S. Zhou and G. Giannakis, "Finite-alphabet based channel estimation for OFDM and related multicarrier systems," *IEEE Transactions on Communications*, vol. 49, no. 8, pp. 1402–1414, Aug 2001.

[26] L. Rota and P. Comon, "Blind equalizers based on polynomial criteria," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 4, 2004, pp. 441–444.

[27] S. P. Neugebauer, "A deterministic blind equalization method for multi-modulus signals," in *Proceedings of the Fourth IEEE Workshop on Sensor Array and Multichannel Signal Processing*, 2006, pp. 551–555.

[28] L. Sorber, M. Van Barel, and L. De Lathauwer, "Structured data fusion," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 4, pp. 586–600, June 2015.

[29] D. Lahat, T. Adali, and C. Jutten, "Multimodal data fusion: An overview of methods, challenges, and prospects," *Proceedings of the IEEE*, vol. 103, no. 9, pp. 1449–1477, Sept 2015.

[30] E. Acar, R. Bro, and A. Smilde, "Data fusion in metabolomics using coupled matrix and tensor factorizations," *Proceedings of the IEEE*, vol. 103, no. 9, pp. 1602–1620, Sept 2015.

[31] M. Sørensen and L. De Lathauwer, "Multiple invariance ESPRIT for nonuniform linear arrays: A coupled canonical polyadic decomposition approach," *IEEE Transactions on Signal Processing*, vol. 64, no. 14, pp. 3693–3704, July 2016.

[32] ——, "Shift invariance, incomplete arrays and coupled CPD: A case study," in *Proceedings of the Ninth IEEE Sensor Array and Multichannel Signal Processing Workshop*, 2016.

[33] ——, "Coupled canonical polyadic decompositions and (coupled) decompositions in multilinear rank-$(L_{r,n}, L_{r,n}, 1)$ terms — Part I: Uniqueness," *SIAM Journal on Matrix Analysis and Applications*, vol. 36, no. 2, pp. 496–522, 2015.

[34] M. Sørensen, I. Domanov, and L. De Lathauwer, "Coupled canonical polyadic decompositions and (coupled) decompositions in multilinear rank-$(L_{r,n}, L_{r,n}, 1)$ terms — Part II: Algorithms," *SIAM Journal on Matrix Analysis and Applications*, vol. 36, no. 3, pp. 1015–1045, 2015.

[35] L. De Lathauwer, "A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization," *SIAM Journal on Matrix Analysis and Applications*, vol. 28, no. 3, pp. 642–666, 2006.

[36] J. B. Kruskal, "Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Linear Algebra and its Applications*, vol. 18, no. 2, pp. 95 – 138, 1977.

[37] I. Domanov and L. De Lathauwer, "On the uniqueness of the canonical polyadic decomposition of third-order tensors — Part II: Uniqueness of the overall decomposition," *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 3, pp. 876–903, 2013.

[38] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.

[39] A. Cichocki, D. Mandic, A. H. Phan, C. Caiafa, G. Zhou, Q. Zhao, and L. De Lathauwer, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Processing Magazine*, vol. 32, no. 2, pp. 145–163, 2015.

[40] N. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," June 2016, Technical Report 16–34, ESAT-STADIUS, KU Leuven, Belgium.

[41] F. Tisseur and K. Meerbergen, "The quadratic eigenvalue problem," *SIAM review*, vol. 43, no. 2, pp. 235–286, 2001.

[42] V. Zarzoso and P. Comon, "Blind and semi-blind equalization based on the constant power criterion," *IEEE Transactions on Signal Processing*, vol. 53, no. 11, pp. 4363–4375, 2005.

[43] G. H. Golub and C. F. Van Loan, *Matrix computations*, 4th ed. JHU Press, 2013.

[44] Y. Wang, Y. Pati, Y. Cho, A. Paulraj, and T. Kailath, "A matrix factorization approach to signal copy of constant modulus signals arriving at an antenna array," in *Proc. 28th Conference on Information Sciences and Systems*, 1994, pp. 178–183.

[45] S. Talwar, M. Viberg, and A. Paulraj, "Blind estimation of multiple co-channel digital signals arriving at an antenna array," in *Proceedings of the Asilomar Conference on Signals, Systems and Computers*. IEEE, 1993, pp. 349–353.

[46] R. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an explanatory multimodal factor analysis," *UCLA Working Papers in Phonetics*, vol. 16, pp. 1–84, 1970.

[47] R. A. Harshman, "Determination and proof of minimum uniqueness conditions for PARAFAC1," *UCLA Working Papers in phonetics*, vol. 22, no. 111-117, p. 3, 1972.

[48] S. E. Leurgans, R. T. Ross, and R. B. Abel, "A decomposition for three-way arrays," *SIAM Journal on Matrix Analysis and Applications*, vol. 14, no. 4, pp. 1064–1083, 1993.

[49] I. Domanov and L. De Lathauwer, "On the uniqueness of the canonical polyadic decomposition of third-order tensors — Part I: Basic results and uniqueness of one factor matrix," *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 3, pp. 855–875, 2013.

[50] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer. (2016, March) Tensorlab 3.0. [Online]. Available: http://www.tensorlab.net

[51] H. Liu and G. Xu, "Smart antennas in wireless systems: uplink multiuser blind channel and sequence detection," *IEEE Transactions on Communications*, vol. 45, no. 2, pp. 187–199, 1997.

[52] A.-J. Van der Veen, S. Talwar, and A. Paulraj, "A subspace approach to blind space-time signal processing for wireless communication systems," *IEEE Transactions on Signal Processing*, vol. 45, no. 1, pp. 173–190, 1997.

[53] P. P. Vaidyanathan, "Multirate digital filters, filter banks, polyphase networks, and applications: A tutorial," *Proceedings of the IEEE*, vol. 78, no. 1, pp. 56–93, 1990.

[54] H. Liu and G. Xu, "Closed-form blind symbol estimation in digital communications," *IEEE Transactions on Signal Processing*, vol. 43, no. 11, pp. 2714–2723, 1995.

[55] M. Sørensen and L. De Lathauwer, "Convolutive low-rank factorizations via coupled low-rank and Toeplitz structured matrix/tensor decompositions," March 2016, Technical Report 16–37, ESAT-STADIUS, KU Leuven, Belgium.

[56] V. Zarzoso and P. Comon, "Robust independent component analysis by iterative maximization of the kurtosis contrast with algebraic optimal step size," *IEEE Transactions on Neural Networks*, vol. 21, no. 2, pp. 248–261, 2010.

[57] S. A. W. Shah, K. Abed-Meraim, and T. Y. Al-Naffouri, "Multi-modulus algorithms using hyperbolic and Givens rotations for MIMO deconvolution," *arXiv preprint arXiv:1506.06650*, 2015.

[58] M. Novey and T. Adali, "Complex fixed-point ICA algorithm for separation of QAM sources using Gaussian mixture model," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 2, 2007, pp. II–445.

[59] A.-J. Van der Veen and J. Tol, "Separation of zero/constant modulus signals," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 5, 1997, pp. 3445–3448.

[60] M. Sørensen, L. Lathauwer, P. Comon, S. Icart, and L. Deneire, "Canonical polyadic decomposition with a columnwise orthonormal factor matrix," *SIAM Journal on Matrix Analysis and Applications*, vol. 33, no. 4, pp. 1190–1213, 2012.

[61] T. Jiang, N. D. Sidiropoulos, and J. M. ten Berge, "Almost-sure identifiability of multidimensional harmonic retrieval," *IEEE Transactions on Signal Processing*, vol. 49, no. 9, pp. 1849–1859, 2001.

[62] R. C. Gunning and H. Rossi, *Analytic functions of several complex variables*. American Mathematical Society, 2009, vol. 368.



Otto Debals obtained the M.Sc. degree in Mathematical Engineering from KU Leuven, Belgium, in 2013. He is a Ph.D. candidate affiliated with the Group Science, Engineering and Technology of Kulak, KU Leuven, with the STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics of the Electrical Engineering Department (ESAT), KU Leuven and with iMinds Medical IT. His research concerns the tensorization of matrix data, with further interests in tensor decompositions, optimization, blind signal separation and blind system identification.



Muhammad Sohail Muhammad Sohail obtained his M.Sc. degree in Electrical Engineering from KU Leuven, Belgium in 2015. Currently he is pursuing an advanced masters in Artificial Intelligence from KU Leuven. He has interests in signal processing, machine learning, and applications of tensor techniques.



Lieven De Lathauwer received the Ph.D. degree from the Faculty of Engineering, KU Leuven, Belgium, in 1997. From 2000 to 2007 he was Research Associate with the Centre National de la Recherche Scientifique, France. He is currently Professor with KU Leuven. He is affiliated with the Group Science, Engineering and Technology of Kulak, with the Stadius Center for Dynamical Systems, Signal Processing and Data Analytics of the Electrical Engineering Department (ESAT) and with iMinds Medical IT. He is Associate Editor of the 'SIAM Journal on Matrix Analysis and Applications' and has served as Associate Editor for the 'IEEE Transactions on Signal Processing'. His research concerns the development of tensor tools for engineering applications.