



HAL
open science

Joint Independent Subspace Analysis by Coupled Block Decomposition: Non-Identifiable Cases

Dana Lahat, Christian Jutten

► **To cite this version:**

Dana Lahat, Christian Jutten. Joint Independent Subspace Analysis by Coupled Block Decomposition: Non-Identifiable Cases. ICASSP 2018 - IEEE International Conference on Acoustics, Speech and Signal Processing, Apr 2018, Calgary, Canada. pp.2186-2190, 10.1109/ICASSP.2018.8461358 . hal-01884799

HAL Id: hal-01884799

<https://hal.science/hal-01884799>

Submitted on 1 Oct 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

JOINT INDEPENDENT SUBSPACE ANALYSIS BY COUPLED BLOCK DECOMPOSITION: NON-IDENTIFIABLE CASES

Dana Lahat and Christian Jutten

Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, 38000 Grenoble, France

ABSTRACT

This paper deals with the identifiability of joint independent subspace analysis of real-valued Gaussian stationary data with uncorrelated samples. This model is not identifiable when each mixture is considered individually. Algebraically, this model amounts to coupled block decomposition of several matrices. In previous work, we showed that if all the cross-correlations in this model were square matrices, the model was generally identifiable. In this paper, we show that this does not necessarily hold when the cross-correlation matrices are rectangular. In this current contribution, we first show that, in certain cases, the balance of degrees of freedom (d.o.f.) between model and observations does not allow identifiability; this situation never occurs in the square case. Second, we explain why for certain block sizes, even if the balance of d.o.f. seems adequate, the model is never identifiable.

Index Terms— Blind source separation, coupled decompositions, low-rank approximation, uniqueness, identifiability

1. INTRODUCTION

This theoretical paper deals with certain aspects of the identifiability of joint independent subspace analysis (JISA) [1, 2, 3]. JISA is a recently-proposed model that extends independent subspace analysis (ISA) (e.g., [4, 5]) by exploiting statistical links among latent multivariate random processes in several ISA problems. Another way to look at JISA is as a variant of independent vector analysis (IVA) [6] in which the random processes within each mixture are possibly multivariate. Both ISA and IVA are themselves extensions of independent component analysis (ICA) [7]. As such, JISA is a very general framework that is able to exploit any of the types of diversity as in single-mixture ICA, such as complex-valued data, sample non-stationarity and/or dependence, and higher-order statistics (HOS), to name a few [1, 8, 3]. Naturally, each type of diversity will further enhance identifiability (e.g., [9, 10]). Accordingly, the algebraic formulations corresponding to these models are more elaborate (e.g., [1, Section VI],[8]). In addition, JISA inherits the enhanced identifiability, interpretability, and versatility of IVA with respect to (w.r.t.) ICA (e.g., [11, 12, 13, 14, 15, 16]). These properties make JISA a potentially useful tool in various applications.

In this paper, however, our focus is on a different aspect, of understanding the added value of the link among datasets w.r.t. an ensemble of individual unrelated ISA problems. For this aim, we focus on the minimal JISA model that provides such insights. Namely, JISA of real-valued Gaussian stationary data with uncorrelated samples. As shown, e.g., in [17], this model amounts to coupled block decomposition (CBD) of an ensemble of matrices. This model does

not exploit any of the additional single-set types of diversity that we mentioned earlier. One can readily verify that in this case, each individual ISA mixture is not identifiable (see also Section 2.2). Therefore, this minimal model allows us to isolate the contribution of the multiset diversity [18].

IVA identifiability under these minimal model assumptions was discussed in [11, 12, 13]. These results, along with, e.g., [19], provide evidence that coupled decompositions can achieve uniqueness and identifiability that exceed those of their underlying individual building blocks. Recently, coupled decompositions have been attracting increasing attention as a means to analyse more elaborate linked datasets and achieve data fusion in various applications; see, e.g., [18, 20, 21] and references therein. Therefore, understanding the properties of coupled decomposition, their advantages and disadvantages, is a matter of ongoing research.

In previous work, we provided numerical [1, 17] and theoretical [22] evidence that this minimal JISA model can be identifiable. More specifically, we have proven that in the special case where the non-zero cross-correlation matrices are square and nonsingular (which implies that all blocks in the CBD are square), and the mixing matrices have full column rank, JISA is generally identifiable, except for very special cases in which the covariance profiles of two independent sources belong to the same equivalence class. In this paper, we do not impose this constraint. Instead, we allow the non-zero cross-correlation matrices to be rectangular, as long as they are full-rank. In this current contribution, we show that this relaxation results in a new phenomenon: for certain block dimensions, the CBD is never unique, which implies that the corresponding JISA model is never identifiable. Furthermore, this phenomenon occurs even if the balance of degrees of freedom (d.o.f.) between model unknowns and constraints imposed by the observations seems adequate. The goal of this paper is to characterize these cases, and to explain them.

1.1. Notations and Conventions

Scalars, column vectors, and matrices, are denoted a , \mathbf{a} , and \mathbf{A} , respectively. The direct product of two matrices, \mathbf{A} and \mathbf{B} , is denoted $\mathbf{A} \oplus \mathbf{B} \triangleq \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix}$. \mathbf{I}_α , $\mathbf{1}_\beta$, and $\mathbf{0}_{\gamma \times \delta}$ denote an $\alpha \times \alpha$ identity matrix, an all-ones $\beta \times 1$ vector, and a $\gamma \times \delta$ all-zeros matrix, respectively. \cdot^\top denotes transpose. $\mathbf{A}^{-[k]}$ and $\mathbf{A}^{-[k]\top}$ stand for $(\mathbf{A}^{[k]})^{-1}$ and $(\mathbf{A}^{-[k]})^\top$, respectively. The Khatri-Rao product for partitioned matrices [23] is defined as

$$\mathbf{A} \boxtimes \mathbf{B} \triangleq \begin{bmatrix} \mathbf{A}_{11} \otimes \mathbf{B}_{11} & \mathbf{A}_{12} \otimes \mathbf{B}_{12} & \cdots \\ \mathbf{A}_{21} \otimes \mathbf{B}_{21} & \mathbf{A}_{22} \otimes \mathbf{B}_{22} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \quad (1)$$

where \mathbf{A}_{ij} and \mathbf{B}_{ij} denote the (i, j) th $m_i \times n_j$ and $p_i \times q_j$ blocks of partitioned matrices \mathbf{A} and \mathbf{B} , respectively, and where $\mathbf{A}_{ij} \otimes \mathbf{B}_{ij}$ and $\mathbf{A} \boxtimes \mathbf{B}$ are $m_i p_i \times n_j q_j$ and $(\sum m_i p_i) \times (\sum n_j q_j)$, respectively.

This work is supported by the project CHESSE, 2012-ERC-AdG-320684. GIPSA-Lab is a partner of the LabEx PERSYVAL-Lab (ANR-11-LABX-0025).

2. BACKGROUND AND PROBLEM FORMULATION

2.1. JISA Model

Consider an ensemble of $K \geq 2$ datasets, modeled as

$$\mathbf{x}^{[k]} = \mathbf{A}^{[k]} \mathbf{s}^{[k]}, \quad k = 1, \dots, K \quad (2a)$$

$$= \sum_{i=1}^R \mathbf{A}_i^{[k]} \mathbf{s}_i^{[k]} \quad (2b)$$

where $\mathbf{x}^{[k]} \in \mathbb{R}^{I^{[k]} \times 1}$ is a random vector representing the observations at dataset k , and $R \geq 2$ is the number of statistically independent contributions in each dataset. The multivariate Gaussian random vectors $\mathbf{s}^{[k]} = [\mathbf{s}_1^{[k]\top}, \dots, \mathbf{s}_R^{[k]\top}]^\top \in \mathbb{R}^{I^{[k]} \times 1}$ and $\mathbf{s}_i^{[k]} \in \mathbb{R}^{m_i^{[k]} \times 1}$ in dataset k have non-degenerate probability density function (pdfs) that cannot be written as a product of several non-trivial pdfs. These random vectors generate multivariate Gaussian stationary random processes with uncorrelated samples. In addition, the block (partitioned) matrix $\mathbf{A}^{[k]} = [\mathbf{A}_1^{[k]} | \dots | \mathbf{A}_R^{[k]}] \in \mathbb{R}^{I^{[k]} \times (\sum_{i=1}^R m_i^{[k]})}$ is nonsingular (in fact, the model may be identifiable even if this assumption is relaxed; however, this assumption simplifies our discussion), where $\mathbf{A}_i^{[k]} \in \mathbb{R}^{I^{[k]} \times m_i^{[k]}}$. The dimensions of these latent variables satisfy $m_i^{[k]} \geq 1$, where $\mathbf{m}^{[k]} = [m_1^{[k]}, \dots, m_R^{[k]}]^\top$, $\mathbf{m}_i = [m_i^{[1]}, \dots, m_i^{[K]}]^\top$, and $I^{[k]} = \sum_{i=1}^R m_i^{[k]} = \mathbf{m}^{[k]\top} \mathbf{1}_R$.

In the JISA model that we consider in this paper, the cross-correlation between $\mathbf{s}_i^{[k]}$ and $\mathbf{s}_j^{[l]}$, where $1 \leq i, j \leq R$, $1 \leq k, l \leq K$, satisfies

$$\mathbf{S}_{ij}^{[k,l]} \triangleq E\{\mathbf{s}_i^{[k]} \mathbf{s}_j^{[l]\top}\} = \begin{cases} \mathbf{S}_{ii}^{[k,l]} & i = j \\ \mathbf{0} & i \neq j \end{cases} \quad (3)$$

where $\mathbf{S}_{ii}^{[k,l]} \in \mathbb{R}^{m_i^{[k]} \times m_i^{[l]}}$ are full-rank $\forall i, k, l$. The $m_i^{[k]} \times m_j^{[l]}$ cross-correlation matrices $\mathbf{S}_{ij}^{[k,l]}$ can be collected in the (k, l) th block of the $\mathbf{m}_i^\top \mathbf{1}_K \times \mathbf{m}_j^\top \mathbf{1}_K$ matrix

$$\mathbf{S}_{ij} \triangleq \begin{bmatrix} \mathbf{S}_{ij}^{[1,1]} & \dots & \mathbf{S}_{ij}^{[K,1]} \\ \vdots & & \vdots \\ \mathbf{S}_{ij}^{[K,1]} & \dots & \mathbf{S}_{ij}^{[K,K]} \end{bmatrix} = \begin{cases} \mathbf{S}_{ii} & i = j \\ \mathbf{0} & i \neq j \end{cases} \quad (4)$$

as well as in the (i, j) th block of the $\mathbf{m}^{[k]\top} \mathbf{1}_R \times \mathbf{m}^{[l]\top} \mathbf{1}_R$ matrix

$$\mathbf{S}^{[k,l]} \triangleq E\{\mathbf{s}^{[k]} \mathbf{s}^{[l]\top}\} = \mathbf{S}_{11}^{[k,l]} \oplus \dots \oplus \mathbf{S}_{RR}^{[k,l]}. \quad (5)$$

Throughout this paper, we assume that \mathbf{S}_{ii} and $\mathbf{S}_{ii}^{[k,k]}$ are positive-definite $\forall i, k$. The cross-correlation between observations in datasets k and l satisfies

$$\mathbf{X}^{[k,l]} \triangleq E\{\mathbf{x}^{[k]} \mathbf{x}^{[l]\top}\} = \mathbf{A}^{[k]} \mathbf{S}^{[k,l]} \mathbf{A}^{[l]\top} \quad (6a)$$

$$= \sum_{i=1}^R \mathbf{A}_i^{[k]} \mathbf{S}_{ii}^{[k,l]} \mathbf{A}_i^{[l]\top} \quad \forall k, l \quad (6b)$$

where the second equality in (6a) is due to (2a), and the transition to (6b) is due to (5). Since $\mathbf{A}^{[k]}$ are nonsingular $\forall k$, (6) can be written as

$$\mathbf{A}^{-[k]} \mathbf{X}^{[k,l]} \mathbf{A}^{-[l]\top} = \mathbf{S}^{[k,l]} = \bigoplus_{i=1}^R \mathbf{S}_{ii}^{[k,l]} \quad \forall k, l \quad (7)$$

where the last equality is due to (5). We refer to the decomposition in (6b) and (7), of all the matrices in $\{\mathbf{X}^{[k,l]}\}_{k,l=1}^K$ at once, as CBD.

It follows from (2b) that one cannot distinguish between the pairs $(\mathbf{A}_i^{[k]}, \mathbf{s}_i^{[k]})$ and $(\mathbf{A}_i^{[k]} \mathbf{Z}_{ii}^{-[k]}, \mathbf{Z}_{ii}^{[k]} \mathbf{s}_i^{[k]})$, where $\mathbf{Z}_{ii}^{[k]}$ is an arbitrary nonsingular $m_i^{[k]} \times m_i^{[k]}$ matrix. However, given $\mathbf{x}_i^{[k]} \triangleq \mathbf{A}_i^{[k]} \mathbf{s}_i^{[k]} \forall i, k$, we can rewrite (2) as $\mathbf{x}^{[k]} = \sum_{i=1}^R \mathbf{x}_i^{[k]} \forall k$, which does not suffer from this ambiguity. Therefore, given a sequence of samples drawn from $\{\mathbf{x}^{[k]}\}_{k=1}^K$, and given $\{m_i^{[k]}\}_{i=1, k=1}^R, K$, we define the problem associated with this model as obtaining estimates of $\mathbf{x}_1, \dots, \mathbf{x}_R$ that are as statistically independent (uncorrelated) as possible, where $\mathbf{x}_i \triangleq [\mathbf{x}_i^{[1]\top}, \dots, \mathbf{x}_i^{[K]\top}]^\top \forall i$. Accordingly, we propose the following *definition for JISA identifiability*: if all (un-ordered) sets of maximally independent estimates of $\mathbf{x}_1, \dots, \mathbf{x}_R$ that satisfy the JISA model assumptions are identical, we say that the JISA model is identifiable.

Similarly, each of the R summands in (6b) satisfies

$$\mathbf{A}_i^{[k]} \mathbf{S}_{ii}^{[k,l]} \mathbf{A}_i^{[l]\top} = (\mathbf{A}_i^{[k]} \mathbf{Z}_{ii}^{-[k]}) (\mathbf{Z}_{ii}^{[k]} \mathbf{S}_{ii}^{[k,l]} \mathbf{Z}_{ii}^{[l]\top}) (\mathbf{Z}_{ii}^{-[l]\top} \mathbf{A}_i^{[l]\top}). \quad (8)$$

In analogy to our definition of JISA identifiability, we now state the algebraic counterpart for *CBD uniqueness*: if any choice of $\{\mathbf{A}_i^{[k]}\}_{i=1, k=1}^R, K$ and $\{\mathbf{S}_{ii}^{[k,l]}\}_{k,l=1}^K$ that satisfy (6) for fixed $\{\mathbf{X}^{[k,l]}\}_{k,l=1}^K$ yields the same R summands $\forall k, l$, we say that the decomposition in (6) and (7) is unique.

If the JISA model satisfies all the above model assumptions, then the cross-correlations $\{\mathbf{X}^{[k,l]}\}_{k,l=1}^K$ are sufficient statistics for estimating the model's parameters [1]. In this case, the uniqueness of the CBD problem in (6) and (7) amounts to JISA identifiability.

2.2. Degrees of Freedom

In our JISA model of interest, the model unknowns consists of $\{\mathbf{A}^{[k]}\}_{k=1}^K$ and $\{\mathbf{S}_{ii}\}_{i=1}^R$. These matrices provide

$$\begin{aligned} N_{\text{model}} &= \sum_{k=1}^K \left(\underbrace{\left(\sum_{i=1}^R m_i^{[k]} \right)^2}_{\text{due to } \mathbf{A}^{[k]}} - \underbrace{\sum_{i=1}^R m_i^{2[k]}}_{\text{due to } \mathbf{Z}_{ii}^{[k]}} \right) \\ &\quad + \frac{1}{2} \sum_{i=1}^R \left(\underbrace{\left(\sum_{k=1}^K m_i^{[k]} \right) \left(\sum_{k=1}^K m_i^{[k]} + 1 \right)}_{\text{due to } \mathbf{S}_{ii}} \right) \quad (9) \end{aligned}$$

free scalar parameters. In (9), we take into account the model's unavoidable inherent ambiguities (8). The data (i.e., observations, measurements, etc.) are represented by $\{\mathbf{X}^{[k,l]}\}_{k,l=1}^K$, which can be regarded as the blocks of the covariance matrix of $[\mathbf{x}^{[1]\top}, \dots, \mathbf{x}^{[K]\top}]^\top \in \mathbb{R}^{(\sum_{i=1}^R \sum_{k=1}^K m_i^{[k]}) \times 1}$. Hence, the model has to fit

$$N_{\text{data}} = \frac{1}{2} \left(\sum_{i=1}^R \sum_{k=1}^K m_i^{[k]} \right) \left(\left(\sum_{i=1}^R \sum_{k=1}^K m_i^{[k]} \right) + 1 \right) \quad (10)$$

scalar constraints. One can readily verify that when $K = 1$ and $R \geq 2$, $N_{\text{data}} - N_{\text{model}} < 0$. When $N_{\text{data}} - N_{\text{model}} < 0$, the number of constraints provided by the observations is smaller than the number of unknowns in the model, and thus the model is not identifiable. Nonnegative values of $N_{\text{data}} - N_{\text{model}}$, on the other hand, do not guarantee identifiability.

2.3. Problem Formulation

In previous work [1, 17, 22], we considered the case of square nonsingular $\mathbf{S}_{ii}^{[k,l]} \forall i, k, l$, which implies $m_i^{[k]} = m_i \forall k$. In this case, the balance of d.o.f. is always nonnegative. Given these assumptions, we defined in [22] a subspace associated with signal i as *reducible* if and only if (iff) there exist K nonsingular matrices (transformations) $\mathbf{T}_{ii}^{[k]}$ such that $\mathbf{T}_{ii}^{[k]} \mathbf{S}_{ii}^{[k,l]} \mathbf{T}_{ii}^{[l]\top} = \begin{bmatrix} \mathbf{S}_{i_1 i_1}^{[k,l]} & \mathbf{0} \\ \mathbf{0}_{\alpha^{[k]} \times \beta^{[l]}} & \mathbf{S}_{i_2 i_2}^{[k,l]} \end{bmatrix} \forall k, l$, where $\alpha^{[k]}$ and $\beta^{[k]} \triangleq m_i - \alpha^{[k]}$ are positive integers $\forall k$. In this case, we say that \mathbf{S}_{ii} is reducible. Otherwise, \mathbf{S}_{ii} is said to be irreducible. Assuming irreducibility $\forall i$, we showed in [22] that JISA was generally identifiable, that is, except for very special cases in which, for some $i \neq j$, $m_i = m_j$ and the blocks of \mathbf{S}_{ii} and \mathbf{S}_{jj} satisfy a certain equivalence relation. In other words, $m_i \neq m_j$ guarantees identifiability.

In this paper, we consider a more relaxed scenario, of rectangular full-rank $\mathbf{S}_{ii}^{[k,l]}$. We show that in this case, different dimensions for (i, j) do not guarantee identifiability. Namely, for certain values of $\{m_i^{[k]}\}_{i=1, k=1}^R, K$, the d.o.f. balance is negative. Then, in certain other cases, the balance is nonnegative but the model is never identifiable, regardless of the value of $\mathbf{S}_{ii} \forall i$. The main part of this paper, in Section 3, is dedicated to explaining this latter phenomenon. Identifiability of the remaining cases can be obtained using arguments similar to those in [22], and is beyond the scope of this work.

2.4. Fisher Information Matrix

In [1], we have shown¹ that asymptotically, that is, when the number of samples drawn from the random processes goes to infinity, for every pair (i, j) with $i \neq j$, the estimation error of the parameters in the model that we have just defined is proportional to the inverse of the symmetric positive semi-definite $2\mathbf{m}_i^\top \mathbf{m}_j \times 2\mathbf{m}_i^\top \mathbf{m}_j$ matrix

$$\mathbf{H} = \begin{bmatrix} \mathbf{S}_{jj} \boxplus \mathbf{S}_{ii}^{-1} & \mathbf{I} \\ \mathbf{I} & \mathbf{S}_{jj}^{-1} \boxplus \mathbf{S}_{ii} \end{bmatrix} \quad (11)$$

(the transition from [1, Equation (32)] to (11) is trivial) where $\mathbf{S}_{jj} \boxplus \mathbf{S}_{ii}^{-1}$ is an $\mathbf{m}_j^\top \mathbf{m}_j \times \mathbf{m}_i^\top \mathbf{m}_i$ matrix whose (k, l) th block has size $m_i^{[k]} m_j^{[k]} \times m_i^{[l]} m_j^{[l]}$. Hence, \mathbf{H} is related to the Fisher information matrix (for further details, see [1]). Matrix \mathbf{H} is generally different for each pair of (i, j) . Matrix \mathbf{H} is well-defined, because it was derived based on the assumption that \mathbf{S}_{ii} and \mathbf{S}_{jj} were positive-definite covariance matrices. If \mathbf{H} is singular for at least one pair of $i \neq j$, we say that the model is not identifiable. Consequently, the question of identifiability boils down to characterizing the singular points of \mathbf{H} for all pairs of (i, j) . In [22], we have shown that \mathbf{H} is singular iff the following system of coupled matrix equations,

$$\tilde{\mathbf{S}}_{ii}^{[k,l]} \mathbf{L}^{[l]} = \mathbf{L}^{[k]} \tilde{\mathbf{S}}_{jj}^{[k,l]} \quad k, l = 1, \dots, K \quad (12)$$

has non-zero solutions, that is, iff there exist $\{\mathbf{L}^{[k]}\}_{k=1}^K$ not all zero that satisfy (12), where $\mathbf{L}^{[k]}$ are $m_i^{[k]} \times m_j^{[k]}$ matrices, and

$$\tilde{\mathbf{S}}_{ii}^{[k,l]} = (\mathbf{S}_{ii}^{[k,k]})^{-\frac{1}{2}} \mathbf{S}_{ii}^{[k,l]} (\mathbf{S}_{ii}^{[l,l]})^{-\frac{1}{2}\top} \quad \forall k, l \quad (13)$$

are normalized versions of \mathbf{S}_{ii} satisfying $\tilde{\mathbf{S}}_{ii}^{[k,k]} = \mathbf{I}_{m_i^{[k]}} \forall k$, as well as all our other model assumptions. We shall use this formulation in our analysis in Section 3.

¹The generalization of all the results in [1] from $m_i^{[k]} = m_i \forall k$ to $m_i^{[k]} \neq m_i^{[l]}$ for $l \neq k$ is trivial and straightforward.

3. MAIN RESULT: NON-IDENTIFIABLE CASES

In this section, we present our claims using a case study. Let $\mathbf{m}_i = [1, 1, \alpha]^\top$, $\mathbf{m}_j = [1, 1, \beta]^\top$. This example corresponds to $K = 3$ datasets, with block dimensions given by $m_i^{[1]} = m_i^{[2]} = m_j^{[1]} = m_j^{[2]} = 1$, $m_i^{[3]} = \alpha$, $m_j^{[3]} = \beta$, for some pair (i, j) , $1 \leq i, j \leq R$, $i \neq j$, where the total number of latent low-rank terms in each dataset is $R \geq 2$. In this example, $R = 2$, $i = 1$, $j = 2$. Substituting these parameters in (9) and (10), the balance of d.o.f. between model unknowns and observational constraints is

$$N_{\text{data}} - N_{\text{model}} = 2\alpha + 2\beta - \alpha\beta. \quad (14)$$

Specific values of (14) for $1 \leq \alpha, \beta \leq 7$ are given in Table 1. For this model, depending on α and β , we identify three types of identifiability, as we now explain. The negative values in Table 1 correspond to non-identifiable cases. The non-highlighted nonnegative entries in Table 1 correspond to cases that are generally identifiable. The highlighted entries in Table 1 correspond to cases that are never identifiable although their d.o.f. balance is nonnegative. The aim of this paper is to explain this type of non-identifiability.

Table 1. $N_{\text{data}} - N_{\text{model}} = 2\alpha + 2\beta - \alpha\beta$. The highlighted cells correspond to cases that are never identifiable although their d.o.f. balance is nonnegative.

$\alpha \backslash \beta$	1	2	3	4	5	6	7
1	3	4	5	6	7	8	9
2	4	4	4	4	4	4	4
3	5	4	3	2	1	0	-1
4	6	4	2	0	-2	-4	-6
5	7	4	1	-2	-5	-8	-11
6	8	4	0	-4	-8	-12	-16
7	9	4	-1	-6	-11	-16	-21

Before going into our theoretical analysis, we point out that each of these claims for (non)identifiability can be validated numerically by generating arbitrary \mathbf{S}_{ii} and \mathbf{S}_{jj} that satisfy the model assumptions of Section 2 and looking at the eigenvalues of \mathbf{H} .

In a first step (Section 3.1), we show that given these values of α and β , one can always write $\tilde{\mathbf{S}}_{ii}$ and $\tilde{\mathbf{S}}_{jj}$ as a direct sum of a smaller square matrix and an identity matrix. In a second step (Section 3.2), we show that this structure is always associated with a singular \mathbf{H} . We discuss and explain these results in Section 3.3.

3.1. First Step

The main idea is that the arbitrary matrices $\mathbf{Z}_{ii}^{[k]}$, defined in (8), act as transformations (see also [24]). Let

$$\mathbf{Z}_{ii}^{[k]} \triangleq \mathbf{Q}_{ii}^{-[k]} (\mathbf{S}_{ii}^{[k,k]})^{-\frac{1}{2}} \quad \forall i, k \quad (15)$$

where $\mathbf{Q}_{ii}^{[k]}$ is a nonsingular $m_i^{[k]} \times m_i^{[k]}$ matrix $\forall i, k$. Then

$$\mathbf{Z}_{ii}^{[k]} \mathbf{S}_{ii}^{[k,l]} \mathbf{Z}_{ii}^{[l]\top} = \mathbf{Q}_{ii}^{-[k]} (\mathbf{S}_{ii}^{[k,k]})^{-\frac{1}{2}} \mathbf{S}_{ii}^{[k,l]} (\mathbf{S}_{ii}^{[l,l]})^{-\frac{1}{2}\top} \mathbf{Q}_{ii}^{-[l]\top} \quad (16a)$$

$$= \mathbf{Q}_{ii}^{-[k]} \tilde{\mathbf{S}}_{ii}^{[k,l]} \mathbf{Q}_{ii}^{-[l]\top} \quad \forall i, k, l \quad (16b)$$

where the last transition is due to (13). For the sake of simplicity, we exemplify using $\alpha = 3 = \beta$. However, the same arguments apply

to all pairs of (α, β) corresponding to highlighted entries in Table 1. In this example, $\tilde{\mathbf{S}}_{ii}^{[3,1]}$ and $\tilde{\mathbf{S}}_{ii}^{[3,2]}$ are 3×1 vectors; $\tilde{\mathbf{S}}_{ii}^{[2,1]}$, $\mathbf{Q}_{ii}^{[1]}$, and $\mathbf{Q}_{ii}^{[2]}$ are nonzero scalars, and $\mathbf{Q}_{ii}^{[3]}$ is 3×3 . Hence, the ensemble of transformation $\{\mathbf{Q}_{ii}^{[k]}\}_{k=1}^K$ on $\{\tilde{\mathbf{S}}_{ii}^{[k,l]}\}_{k,l}^K$ on the right-hand side (RHS) of (16b) can be written as

$$\left(\bigoplus_{k=1}^K \mathbf{Q}_{ii}^{-[k]}\right) \tilde{\mathbf{S}}_{ii} \left(\bigoplus_{k=1}^K \mathbf{Q}_{ii}^{-[k]}\right)^\top = \begin{bmatrix} q_{ii}^{-[1]} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & q_{ii}^{-[2]} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_{ii}^{-[3]} \end{bmatrix} \begin{bmatrix} 1 & \tilde{\mathbf{s}}_{ii}^{[2,1]} & \tilde{\mathbf{S}}_{ii}^{[3,1]\top} \\ \tilde{\mathbf{s}}_{ii}^{[2,1]} & 1 & \tilde{\mathbf{S}}_{ii}^{[3,2]\top} \\ \tilde{\mathbf{S}}_{ii}^{[3,1]} & \tilde{\mathbf{S}}_{ii}^{[3,2]} & \mathbf{I}_3 \end{bmatrix} \begin{bmatrix} q_{ii}^{-[1]} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & q_{ii}^{-[2]} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_{ii}^{-[3]} \end{bmatrix}^\top \quad (17)$$

The key point is that the operation of $\mathbf{Q}_{ii}^{-[3]}$ on the lower-left 3×2 submatrix of $\tilde{\mathbf{S}}_{ii}$ (highlighted in (17)) can be interpreted as a QR decomposition: $\begin{bmatrix} \tilde{\mathbf{s}}_{ii}^{[3,1]} & \tilde{\mathbf{s}}_{ii}^{[3,2]} \end{bmatrix} = \mathbf{Q}_{ii}^{[3]} \mathbf{R}$, where \mathbf{R} has an upper triangular structure $\begin{bmatrix} * & * \\ 0 & * \\ 0 & 0 \end{bmatrix}$, “*” stands for any unspecified non-zero scalar value, and $\mathbf{Q}_{ii}^{[3]}$ is unitary. The nonnegative scalars $q_{ii}^{[1]}$ and $q_{ii}^{[2]}$ do not have any effect on the structure. We further note that $\mathbf{Q}_{ii}^{-[3]} \tilde{\mathbf{S}}_{ii}^{[3]} \mathbf{Q}_{ii}^{-[3]\top} = \mathbf{I}_3$ for any unitary $\mathbf{Q}_{ii}^{[3]}$. If we choose $q_{ii}^{[1]} = 1 = q_{ii}^{[2]}$, it follows that one can always transform \mathbf{S}_{ii} to a structure

$$\tilde{\mathbf{S}}_{ii} = \begin{bmatrix} 1 & * & * & 0 & 0 \\ * & 1 & * & * & 0 \\ * & * & 1 & 0 & 0 \\ 0 & * & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (18)$$

via (15). The same procedure can be applied to \mathbf{S}_{jj} . Summarizing our derivations so far, we have shown that given $\mathbf{m}_i = (1, 1, 3) = \mathbf{m}_j$, both \mathbf{S}_{ii} and \mathbf{S}_{jj} can always be assumed to have the structure in (18).

Similar arguments can now applied to the other highlighted entries in Table 1. We now note that each of these highlighted cells corresponds to $2 = m_i^{[1]} + m_i^{[2]} < m_i^{[3]} = \alpha$ and $2 = m_j^{[1]} + m_j^{[2]} < m_j^{[3]} = \beta$. Hence, the QR decomposition of the lower-left $m_i^{[3]} \times 2$ submatrix of $\tilde{\mathbf{S}}_{ii}$ yields an upper-triangular matrix \mathbf{R} with $m_i^{[3]} - 2$ rows of zeros at the bottom. Consequently, $\tilde{\mathbf{S}}_{ii}$ can always be written as a direct sum of an order- $(\sum_{k=1}^2 m_i^{[k]} + 2)$ matrix and $\mathbf{I}_{m_i^{[3]}-2}$. The same holds for j . This concludes the first step of our case study.

3.2. Second Step

We now explain why a structure that is a direct sum of a square matrix and the identity, for both i and j , implies that the corresponding \mathbf{H} is always singular. In our example, one can readily verify that for any $\tilde{\mathbf{S}}_{ii}$ and $\tilde{\mathbf{S}}_{jj}$ with the structure in (18), the system of equations

$$\text{in (12) holds for any } \mathbf{L}^{[1]} = \mathbf{0} = \mathbf{L}^{[2]} \text{ and } \mathbf{L}^{[3]} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \xi \end{bmatrix} \neq \mathbf{0},$$

$\xi \neq 0$. As explained in Section 2.4, a non-zero solution to (12) corresponds to a singular \mathbf{H} , and thus the overall JISA model is not identifiable. This concludes our proof that our minimal JISA model with $\mathbf{m}_i = \mathbf{m}_j = (1, 1, 3)$, $j \neq i$, is never identifiable. As for the other highlighted values in Table 1, one can readily verify that given the

direct-sum structure explained in Section 3.1 for both $\tilde{\mathbf{S}}_{ii}$ and $\tilde{\mathbf{S}}_{jj}$, there always exists a non-zero solution to (12), with $\mathbf{L}^{[1]} = \mathbf{0} = \mathbf{L}^{[2]}$ and $\mathbf{L}^{[3]} = \mathbf{0}_{2 \times 3} \oplus \mathbf{\Xi}$, where $\mathbf{\Xi}$ is an arbitrary $(m_i^{[3]} - 2) \times (m_j^{[3]} - 2)$ matrix.

3.3. Discussion

A question that may arise is whether the fact that the effective form of $\tilde{\mathbf{S}}_{ii}$ and $\tilde{\mathbf{S}}_{jj}$ contains zeros changes the balance of d.o.f., and if so, can this explain the non-identifiability. To answer this question, we point out that the zeros in (18) can affect only N_{model} , by decreasing the number of model unknowns. Hence, the value of $N_{\text{data}} - N_{\text{model}}$ can only increase. Therefore, this cannot serve as an alternative explanation to non-identifiability.

Until now, we ignored the question of reducibility. By observing the transformed structure in (18), we see that $\tilde{\mathbf{S}}_{ii}$ does not satisfy the definition of reducibility in Section 2.3. However, (18) reveals that the last entry of $\mathbf{s}_i^{[3]}$ is uncorrelated with all other entries of $\mathbf{s}_i^{[k]} \forall k$. Hence, $\mathbf{s}_i^{[3]}$ consists of two uncorrelated terms, and is thus reducible. These results imply that the definition of (ir)reducibility in Section 2.3 is inadequate for the more general case of JISA with rectangular blocks. In other words, the new phenomenon is due to a new and more elaborate type of reducibility that was not considered in earlier work.

Clearly, this type of reducibility and non-identifiability is not limited to models with structure given by $\mathbf{m}_i = [1, 1, \alpha]^\top$, $\mathbf{m}_j = [1, 1, \beta]^\top$. One can readily verify that this phenomenon occurs whenever a model has at least one pair of $i \neq j$ for which $\sum_{k=1}^{K-1} m_i^{[k]} < m_i^{[K]}$, and similarly for j . For example, if $\mathbf{m}_i = [1, 2, 4]^\top$ and $\mathbf{m}_j = [3, 5, 9]^\top$, where $N_{\text{data}} - N_{\text{model}} = 21$, the model is not identifiable. Hence, in the general case of rectangular cross-correlations, having different dimensions for the different random vectors can result in completely non-identifiable models. This is in striking contrast to the square case, i.e., $m_i^{[k]} = m_i$ and $m_j^{[k]} = m_j \forall k$, in which identifiability is guaranteed if $m_i \neq m_j$ for any $i \neq j$.

This type of reducibility and non-identifiability is not a concern when $m_i^{[k]} = m_i \forall i, k$, because in this case, each $\mathbf{S}_{ii}^{[k,l]}$ is square (and nonsingular, by our model assumptions) such that its QR decomposition yields a nonsingular triangular matrix.

A possible remedy is to use additional types of diversity. If the samples are piecewise nonstationary, such that there exist at least two distinct matrices $\mathbf{S}_{ii}^{[k,l](t)}$ for fixed i, k, l , where $t = 1, \dots, T$ are indices of disjoint $T \geq 2$ sampling intervals, then this type of non-identifiability cannot occur, in general. The reason is that once the normalization $\tilde{\mathbf{S}}_{ii}^{[k,k](t)} = \mathbf{I}$ has been applied for some fixed $t = t'$, the other $\tilde{\mathbf{S}}_{ii}^{[k,k](t')}$, for $t'' \neq t'$, cannot, in general, be diagonalized. Hence, this result can serve as a further motivation for exploiting all types of diversity in the data.

To conclude, in this paper, we presented new results on the identifiability of JISA and CBD, when dependent random processes have different dimensions. We have shown that while in certain cases different dimensions guarantee identifiability, in other cases they can completely destroy it. We explained this by showing that this setup may be associated with more elaborate types of (ir)reducibility. These results indicate that while block decompositions carry more flexibility in representing data, they should be treated with care. These results also further emphasize that block decompositions cannot be handled as straightforward extensions of rank-1 methods.

4. REFERENCES

- [1] D. Lahat and Ch. Jutten, "Joint independent subspace analysis using second-order statistics," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4891–4904, Sep. 2016.
- [2] D. Lahat and Ch. Jutten, "Joint blind source separation of multidimensional components: Model and algorithm," in *Proc. EUSIPCO*, Lisbon, Portugal, Sep. 2014, pp. 1417–1421.
- [3] R. F. Silva, S.M. Plis, T. Adalı, and V. D. Calhoun, "Multidataset independent subspace analysis extends independent vector analysis," in *Proc. ICIP*, Paris, France, Oct. 2014, pp. 2864–2868.
- [4] P. Comon, "Supervised classification, a probabilistic approach," in *Proc. ESANN*, Brussels, Belgium, Apr. 1995, pp. 111–128.
- [5] J.-F. Cardoso, "Multidimensional independent component analysis," in *Proc. ICASSP*, Seattle, WA, May 1998, vol. 4, pp. 1941–1944.
- [6] T. Kim, T. Eltoft, and T.-W. Lee, "Independent vector analysis: An extension of ICA to multivariate components," in *Independent Component Analysis and Blind Signal Separation*, 2006, vol. 3889 of *LNCS*, pp. 165–172, Springer Berlin Heidelberg.
- [7] P. Comon, "Independent component analysis, a new concept?," *Signal Process.*, vol. 36, no. 3, pp. 287–314, Apr. 1994.
- [8] D. Lahat and C. Jutten, "Joint analysis of multiple datasets by cross-cumulant tensor (block) diagonalization," in *Proc. SAM*, Rio de Janeiro, Brazil, Jul. 2016.
- [9] T. Adalı, M. Anderson, and G.-S. Fu, "Diversity in independent component and vector analyses: Identifiability, algorithms, and applications in medical imaging," *IEEE Signal Process. Mag.*, pp. 18–33, May 2014.
- [10] N. D. Sidiropoulos and R. Bro, "On communication diversity for blind identifiability and the uniqueness of low-rank decomposition of N -way arrays," in *Proc. ICASSP*, Istanbul, Turkey, Jun. 2000, vol. 5, pp. 2449–2452.
- [11] J. Vía, M. Anderson, X.-L. Li, and T. Adalı, "Joint blind source separation from second-order statistics: Necessary and sufficient identifiability conditions," in *Proc. ICASSP*, Prague, Czech Republic, May 2011, pp. 2520–2523.
- [12] M. Anderson, G.-S. Fu, R. Phlypo, and T. Adalı, "Independent vector analysis: Identification conditions and performance bounds," *IEEE Trans. Signal Process.*, vol. 62, no. 17, pp. 4399–4410, Sep. 2014.
- [13] D. Lahat and Ch. Jutten, "An alternative proof for the identifiability of independent vector analysis using second order statistics," in *Proc. ICASSP*, Shanghai, China, Mar. 2016.
- [14] D. Kitamura, N. Ono, H. Sawada, H. Kameoka, and H. Saruwatari, "Determined blind source separation unifying independent vector analysis and nonnegative matrix factorization," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 24, no. 9, pp. 1626–1641, Sep. 2016.
- [15] S. Bhinge, Y. Levin-Schwartz, and T. Adalı, "Data-driven fusion of multi-camera video sequences: Application to abandoned object detection," in *Proc. ICASSP*, New Orleans, LA, USA, Mar. 2017, pp. 1697–1701.
- [16] F. Nesta, S. Mosayyebpour, Z. Koldovský, and K. Paleček, "Audio/video supervised independent vector analysis through multimodal pilot dependent components," in *Proc. EUSIPCO*, Kos Island, Greece, Aug. 2017.
- [17] D. Lahat and Ch. Jutten, "Joint independent subspace analysis: A quasi-Newton algorithm," in *Latent Variable Analysis and Signal Separation*, 2015, vol. 9237 of *LNCS*, pp. 111–118, Springer International Publishing Switzerland.
- [18] D. Lahat, T. Adalı, and Ch. Jutten, "Multimodal data fusion: An overview of methods, challenges and prospects," *Proc. IEEE*, vol. 103, no. 9, pp. 1449–1477, Sep. 2015.
- [19] M. Sørensen and L. De Lathauwer, "Coupled canonical polyadic decompositions and (coupled) decompositions in multilinear rank- $(L_{r,n}, L_{r,n}, 1)$ terms—part I: Uniqueness," *SIAM J. Matrix Anal. Appl.*, vol. 36, no. 2, pp. 496–522, Apr. 2015.
- [20] E. Acar, M. A. Rasmussen, F. Savorani, T. Næs, and R. Bro, "Understanding data fusion within the framework of coupled matrix and tensor factorizations," *Chemom. Intell. Lab. Syst.*, vol. 129, pp. 53–63, 2013.
- [21] I. Van Mechelen and A. K. Smilde, "A generic linked-mode decomposition model for data fusion," *Chemom. Intell. Lab. Syst.*, vol. 104, no. 1, pp. 83–94, Nov. 2010.
- [22] D. Lahat and Ch. Jutten, "A generalization to Schur's lemma with an application to joint independent subspace analysis," Tech. Rep. hal-01247899, GIPSA-Lab, Grenoble, France, Dec. 2015.
- [23] R. A. Horn and R. Mathias, "Block-matrix generalizations of Schur's basic theorems on Hadamard products," *Linear Algebra and its Applications*, vol. 172, pp. 337–346, Jul. 1992.
- [24] L. De Lathauwer, "Decompositions of a higher-order tensor in block terms. Part II: Definitions and uniqueness," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, pp. 1033–1066, 2008.