Generic uniqueness of a structured matrix factorization and applications in blind source separation

Ignat Domanov and Lieven De Lathauwer, *Fellow, IEEE*

***Abstract*—Algebraic geometry, although little explored in signal processing, provides tools that are very convenient for investigating generic properties in a wide range of applications. Generic properties are properties that hold "almost everywhere". We present a set of conditions that are sufficient for demonstrating the generic uniqueness of a certain structured matrix factorization. This set of conditions may be used as a checklist for generic uniqueness in different settings. We discuss two particular applications in detail. We provide a relaxed generic uniqueness condition for joint matrix diagonalization that is relevant for independent component analysis in the underdetermined case. We present generic uniqueness conditions for a recently proposed class of deterministic blind source separation methods that rely on mild source models. For the interested reader we provide some intuition on how the results are connected to their algebraic geometric roots.**

***Index Terms*—structured matrix factorization, structured rank decomposition, blind source separation, direction of arrival, uniqueness, algebraic geometry**

I. Introduction

A. Blind source separation and uniqueness

The matrix factorization $\mathbf{X} = \mathbf{M}\mathbf{S}^T$ is well known in the blind source separation (BSS) context: the rows of $\mathbf{S}^T$ and $\mathbf{X}$ represent unknown source signals and their observed linear mixtures, respectively. The task of the BSS problem is to estimate the source matrix $\mathbf{S}$ and the mixing matrix $\mathbf{M}$ from $\mathbf{X}$. If no prior information is available on the matrices $\mathbf{M}$ or $\mathbf{S}$, then they cannot be uniquely identified from $\mathbf{X}$. Indeed, for any nonsingular matrix $\mathbf{T}$,

$$\mathbf{X} = \mathbf{M}\mathbf{S}^T = (\mathbf{M}\mathbf{T})(\mathbf{S}\mathbf{T}^{-T})^T = \overline{\mathbf{M}}\overline{\mathbf{S}}^T. \qquad (1)$$

Applications may involve particular constraints on $\mathbf{M}$ and/or $\mathbf{S}$, so that in the resulting class of structured matrices the solution of (1) becomes unique. Commonly used constraints include sparsity [1], constant modulus [2] and Vandermonde structure [3].

Sufficient conditions for uniqueness can be deterministic or generic. Deterministic conditions concern particular matrices $\mathbf{M}$ and $\mathbf{S}$. Generic conditions concern the situation that can be expected in general; a generic property is a property that holds everywhere except for a set of measure 0. (A formal definition will be given in Subsection I-C below.)

To illustrate the meaning of deterministic and generic uniqueness let us consider decomposition (1) in which $\mathbf{X} \in \mathbb{C}^{K\times N}$, $\mathbf{M} \in \mathbb{C}^{K\times R}$ and the columns of $\mathbf{S} \in \mathbb{C}^{N\times R}$ are obtained by sampling the exponential signals $z_1^{t-1}, \dots, z_R^{t-1}$ at $t = 1, \dots, N$. Then $(\mathbf{S})_{nr} = (z_r^{n-1})$, i.e. $\mathbf{S}$ is a Vandermonde matrix. A deterministic condition under which decomposition (1) is unique (up to trivial indeterminacies) is [3]: (i) the Vandermonde matrix $\mathbf{S}$ has strictly more rows than columns and its generators z_j are distinct and (ii) the matrix $\mathbf{M}$ has full column rank. (In this paper we say that an $K \times R$ matrix *has full column rank* if its column rank is R, which implies $K \geq R$.) This deterministic condition can easily be verified for any particular $\mathbf{M}$ and $\mathbf{S}$. A generic variant is: (i) the Vandermonde matrix $\mathbf{S}$ has $N > R$ and (ii) the (unstructured) matrix $\mathbf{M}$ has $K \geq R$. Indeed, under these dimensionality conditions the deterministic conditions are satisfied everywhere, except in a set of measure 0 (which contains the particular cases of coinciding generators z_r and the cases in which the columns of $\mathbf{M}$ are not linearly independent despite the fact that $\mathbf{M}$ is square or even tall). Note that generic properties do not allow one to make statements about specific matrices; they only show the general picture.

As mentioned before, BSS has many variants, which differ in the types of constraints that are imposed. Different constraints usually mean different deterministic uniqueness conditions, and the derivation of these is work that is difficult to automate. In this paper we focus on generic uniqueness conditions. We propose a framework with which generic uniqueness can be investigated in a broad range of cases. Indeed, it will become clear that if we limit ourselves to generic uniqueness, the derivation of conditions can to some extent be automated. We discuss two concrete applications which may serve as examples.

Our approach builds on results in algebraic geometry. Algebraic geometry has so far been used in system theory in [4], [5] and it also has direct applications in tensor-based BSS via the generic uniqueness of tensor decompositions [6]–[8].

This work was supported by Research Council KU Leuven: C1 project c16/15/059-nD, CoE PFV/10/002 (OPTEC) and PDM postdoc grant, by F.W.O.: project G.0830.14N, G.0881.14N, by the Belgian Federal Science Policy Office: IUAP P7 (DYSCO II, Dynamical systems, control and optimization, 2012-2017), by 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.

The authors are with Group Science, Engineering and Technology, KU Leuven-Kulak, E. Sabbelaan 53, 8500 Kortrijk, Belgium. Lieven De Lathauwer is also with Dept. of Electrical Engineering ESAT/STADIUS KU Leuven, Kasteelpark Arenberg 10, bus 2446, B-3001 Leuven-Heverlee, Belgium (e-mail: Ignat.Domanov@kuleuven-kulak.be; Lieven.DeLathauwer@kuleuven-kulak.be).

Our paper makes a contribution in further connecting algebraic geometry with applications in signal processing.

B. Notation

Throughout the paper $\mathbb{F}$ denotes the field of real or complex numbers; bold lowercase letters denote vectors, while bold uppercase letters represent matrices; a column of a matrix $\mathbf{A}$ and an entry of a vector $\mathbf{b}$ are denoted by $\mathbf{a}_j$ and b_j, respectively; the superscripts $\cdot^*$, $\cdot^T$ and $\cdot^H$ are used for the conjugate, transpose, and Hermitian transpose, respectively; "$\otimes$" denotes the Kronecker product.

C. Statement of the problem and organization of the paper

A structured matrix factorization. In this paper we consider the following structured factorization of a $K \times N$ matrix $\mathbf{Y}$,

$$\mathbf{Y} = \mathbf{A}(\mathbf{z})\mathbf{B}(\mathbf{z})^T, \quad \mathbf{z} \in \Omega \tag{2}$$

where Ω is a subset of $\mathbb{F}^n$ and $\mathbf{A}(\mathbf{z})$ and $\mathbf{B}(\mathbf{z})$ are known matrix-valued functions defined on Ω.

W.l.o.g. we can assume that the parameter vector $\mathbf{z} = [z_1 \ \dots \ z_n]^T$ is ordered such that $\mathbf{B}(\mathbf{z})$ depends on the last $s \leq n$ entries, while $\mathbf{A}(\mathbf{z})$ depends on $m \geq 0$ entries that are not necessarily the first or the last. That is,

$$\mathbf{A}(\mathbf{z}) = \mathbf{A}(z_{i_1}, \dots, z_{i_m}), \quad \mathbf{B}(\mathbf{z}) = \mathbf{B}(z_{n-s+1}, \dots, z_n)$$

for some $1 \leq i_1 < i_2 < \cdots < i_m \leq n$. In general, the entries used to parameterize $\mathbf{A}$ and $\mathbf{B}$ are allowed to overlap so that $m + s \geq n$. The case where $\mathbf{A}$ and $\mathbf{B}$ depend on separated parameter sets corresponds to $m + s = n$; in this case $\mathbf{A}$ depends strictly on the first m of the entries of $\mathbf{z}$.

Our study is limited to $K \times R$ matrices $\mathbf{A}(\mathbf{z})$ that generically have full column rank. We do not make any other assumptions on the form of $\mathbf{A}(\mathbf{z})$. In particular, we do not impose restrictions on how the entries depend on $\mathbf{z}$. We are however more explicit about the form of the $N \times R$ matrix $\mathbf{B}(\mathbf{z})$. We assume that each of its columns $\mathbf{b}_r(\mathbf{z})$ is generated by l parameters that are independent of the parameters used to generate the other columns, i.e., $\mathbf{B}(\mathbf{z}) = [\mathbf{b}_1(\boldsymbol{\zeta}_1) \ \dots \ \mathbf{b}_R(\boldsymbol{\zeta}_R)]$ with $\boldsymbol{\zeta}_1, \dots, \boldsymbol{\zeta}_R \in \mathbb{F}^l$. Note that the independence implies that $s = Rl$ and that $[\boldsymbol{\zeta}_1^T \ \dots \ \boldsymbol{\zeta}_R^T]^T$ and $[z_{n-s+1} \ \dots \ z_n]^T$ are the same up to index permutation.

For the sake of exposition, let us first consider a class of matrices $\mathbf{B}(\mathbf{z})$ that is smaller than the class that we will be able to handle in our derivation of generic uniqueness conditions. Namely, let us first consider matrices $\mathbf{B}^{rat}(\mathbf{z})$, of which the n-th row is obtained by evaluating a known rational function $\frac{p_n(\cdot)}{q_n(\cdot)}$ at some points $\boldsymbol{\zeta}_1, \dots, \boldsymbol{\zeta}_R \in \mathbb{F}^l$, $1 \leq n \leq N$:

$$\mathbf{B}^{rat}(\mathbf{z}) = \begin{bmatrix} \frac{p_1(\boldsymbol{\zeta}_1)}{q_1(\boldsymbol{\zeta}_1)} & \cdots & \frac{p_1(\boldsymbol{\zeta}_R)}{q_1(\boldsymbol{\zeta}_R)} \\ \vdots & \vdots & \vdots \\ \frac{p_N(\boldsymbol{\zeta}_1)}{q_N(\boldsymbol{\zeta}_1)} & \cdots & \frac{p_N(\boldsymbol{\zeta}_R)}{q_N(\boldsymbol{\zeta}_R)} \end{bmatrix},$$

where

$$p_1, \dots, p_N, \ q_1, \dots, q_N \ \text{ are polynomials in } l \text{ variables.}$$

Note that we model a column of $\mathbf{B}^{rat}$ through the values taken by N functions $\frac{p_1(\cdot)}{q_1(\cdot)}, \dots, \frac{p_N(\cdot)}{q_N(\cdot)}$ at one particular point $\boldsymbol{\zeta}_r$. On the other hand, a row of $\mathbf{B}^{rat}$ is modeled as values taken by one particular function $\frac{p_n(\cdot)}{q_n(\cdot)}$ at R points $\boldsymbol{\zeta}_1, \dots, \boldsymbol{\zeta}_R$.

The structure that we consider in our study for the $N \times R$ matrix $\mathbf{B}(\mathbf{z})$ is more general than the rational structure of $\mathbf{B}^{rat}(\mathbf{z})$ in the sense that we additionally allow (possibly nonlinear) transformations of $\boldsymbol{\zeta}_1, \dots, \boldsymbol{\zeta}_R$. Formally, we assume that the columns of $\mathbf{B}(\mathbf{z})$ are sampled values of known vector functions of the form

$$\mathbf{b}(\boldsymbol{\zeta}) = \left[\frac{p_1(\mathbf{f}(\boldsymbol{\zeta}))}{q_1(\mathbf{f}(\boldsymbol{\zeta}))} \ \dots \ \frac{p_N(\mathbf{f}(\boldsymbol{\zeta}))}{q_N(\mathbf{f}(\boldsymbol{\zeta}))}\right]^T, \quad \boldsymbol{\zeta} \in \mathbb{F}^l, \tag{3}$$

at points $\boldsymbol{\zeta}_1, \dots, \boldsymbol{\zeta}_R \in \mathbb{F}^l$, such that

$$\mathbf{B}(\mathbf{z}) = [\mathbf{b}(\boldsymbol{\zeta}_1) \dots \mathbf{b}(\boldsymbol{\zeta}_R)] = \begin{bmatrix} \frac{p_1(\mathbf{f}(\boldsymbol{\zeta}_1))}{q_1(\mathbf{f}(\boldsymbol{\zeta}_1))} & \cdots & \frac{p_1(\mathbf{f}(\boldsymbol{\zeta}_R))}{q_1(\mathbf{f}(\boldsymbol{\zeta}_R))} \\ \vdots & \vdots & \vdots \\ \frac{p_N(\mathbf{f}(\boldsymbol{\zeta}_1))}{q_N(\mathbf{f}(\boldsymbol{\zeta}_1))} & \cdots & \frac{p_N(\mathbf{f}(\boldsymbol{\zeta}_R))}{q_N(\mathbf{f}(\boldsymbol{\zeta}_R))} \end{bmatrix},$$

where

$$\mathbf{f}(\boldsymbol{\zeta}) = (f_1(\boldsymbol{\zeta}), \dots, f_l(\boldsymbol{\zeta})) \in \mathbb{F}^l,$$
$$f_1, \dots, f_l \text{ are scalar functions of } l \text{ variables.}$$

The functions $f_1, \dots, f_l$ are subject to an analyticity assumption that will be specified in Theorem 1 further. Although our general result in Theorem 1 will be formulated in terms of functions $f_1, \dots, f_l$ in l variables, in the applications in Sections III–IV we will only need entry-wise transformations:

$$\mathbf{f}(\boldsymbol{\zeta}) = \mathbf{f}(\zeta_1, \dots, \zeta_l) = (f_1(\zeta_1), \dots, f_l(\zeta_l)) \tag{4}$$

with $f_1, \dots, f_l$ analytic functions in one variable.

As an example of how the model for $\mathbf{B}(\mathbf{z})$ can be used, consider R vectors that are obtained by sampling the exponential signals $e^{i\zeta_1(t-1)}, \dots, e^{i\zeta_R(t-1)}$ (with $\zeta_1, \dots, \zeta_R \in \mathbb{R}$) at $t = 1, \dots, N$. In this case $\mathbf{B}(\mathbf{z})$ is an $N \times R$ Vandermonde matrix with unit norm generators; its rth column is $\mathbf{b}(\zeta_r) = [1 \ e^{i\zeta_r} \ \dots \ e^{i\zeta_r(N-1)}]^T$. We have $e^{i\zeta_r(n-1)} = \frac{p_n(f(\zeta_r))}{q_n(f(\zeta_r))}$, where $f(\zeta) = e^{i\zeta}$, $p_n(x) = x^{n-1}$, and $q_n(x) = 1$ for $\zeta \in \mathbb{R}$ and $x \in \mathbb{C}$.

Generic uniqueness of the decomposition. We interpret factorization (2) as a decomposition into a sum of structured rank-1 matrices

$$\mathbf{Y} = \mathbf{A}(\mathbf{z})\mathbf{B}(\mathbf{z})^T = \sum_{r=1}^{R} \mathbf{a}_r(\mathbf{z})\mathbf{b}(\boldsymbol{\zeta}_r)^T, \ \mathbf{z} \in \Omega, \tag{5}$$

where $\mathbf{a}_r(\mathbf{z})$ denotes the rth column of $\mathbf{A}(\mathbf{z})$. It is clear that in (5) the rank-1 terms can be arbitrarily permuted. We say that *decomposition* (5) *is unique* when it is only subject to this trivial indeterminacy. We say that decomposition (5) is *generically unique* if it is unique for a generic choice of $\mathbf{z} \in \Omega$, that is

$$\mu_n\{\mathbf{z} \in \Omega : \ \text{decomposition (5) is not unique}\} = 0, \tag{6}$$

where μ_n is a measure that is absolutely continuous (a.c.) with respect to the Lebesgue measure on $\mathbb{F}^n$.

In this paper we present conditions on the polynomials $p_1, \dots, p_N$, $q_1, \dots, q_N$, the function $\mathbf{f}$ and the set Ω which guarantee that decomposition (5) is generically unique. As a technical assumption, since in the case where $\mu_n(\Omega) = 0$

condition (6) cannot be used to infer generic uniqueness from a subset of Ω, we assume that $\mu_n(\Omega) > 0$.

Organization and results. In Section II we state the main result of this paper in general terms, namely, Theorem 1 presents conditions that guarantee that the structured decomposition (5) is generically unique. The proof of Theorem 1 is given in Appendix A. Besides the technical derivation, Appendix A provides some intuition behind the high-level reasoning and makes the connection with the trisecant lemma in algebraic geometry, for readers who are interested. In Sections III–IV we use Theorem 1 to obtain new uniqueness results in the context of two different applications. This is done by first expressing the specific BSS problem as a decomposition of the form (5), for which the list of conditions in Theorem 1 is checked. Section III concerns an application in independent component analysis. More precisely, it concerns joint matrix diagonalization in the underdetermined case (more sources than observations) and presents a new, relaxed bound on the number of sources under which the solution of this basic subproblem is generically unique. This bound is a simple expression in the number of matrices and their dimension. Section IV presents generic uniqueness results for a recently introduced class of deterministic blind source separation algorithms that may be seen as a variant of sparse component analysis which makes use of a non-discrete dictionary of basis functions. Appendix B contains the short proof of a technical lemma in Section IV. The paper is concluded in Section V.

II. Main result

The following theorem is our main result on generic uniqueness of decomposition (5). It states that, generically, the R structured rank-1 terms of the $K \times N$ matrix $\mathbf{Y}$ can be uniquely recovered if $K \geq R$ and $R \leq \widehat{N} - \widehat{l}$. Here, $\widehat{N} \leq N$ is a lower bound on the dimension of the linear vector space $\operatorname{span}\{\mathbf{r}(\mathbf{x}) : \ q_1(\mathbf{x}) \cdots q_N(\mathbf{x}) \neq 0, \ \mathbf{x} \in \mathbb{F}^l\}$ generated by vectors of the form

$$\mathbf{r}(\mathbf{x}) = \begin{bmatrix} \frac{p_1(\mathbf{x})}{q_1(\mathbf{x})} & \cdots & \frac{p_N(\mathbf{x})}{q_N(\mathbf{x})} \end{bmatrix}^T. \tag{7}$$

(Note that the definition of $\mathbf{r}(\mathbf{x})$ does not involve a nonlinear transformation $\mathbf{f}$, even when such a nonlinear transformation is used for modelling $\mathbf{b}(\boldsymbol{\zeta})$.) On the other hand, the value $\widehat{l} \leq l$ is an upper bound on the number of "free parameters" actually needed to parameterize a generic vector of the form (7). (Indeed, although $\mathbf{r}(\mathbf{x})$ is generated by l independent parameters, it may be possible to do it with less in particular cases. For instance, let $N = 3$, $q_1(\mathbf{x}) = q_2(\mathbf{x}) = q_3(\mathbf{x}) = 1$ and $p_1(\mathbf{x}) = x_1 + x_3$, $p_2(\mathbf{x}) = x_2 - x_3$, $p_3(\mathbf{x}) = x_1 + x_2$, so that $\mathbf{r}(\mathbf{x}) = \mathbf{W}\mathbf{x}$ with $\mathbf{W} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & 1 & 0 \end{bmatrix}$. Since $\operatorname{rank}(\mathbf{W}) = 2$, $\mathbf{r}(\mathbf{x})$ can be parameterized by $2 < 3$ independent parameters.)

In the theorem and throughout the paper we use $\mathbf{J}(\mathbf{r}, \mathbf{x}) \in \mathbb{F}^{N \times l}$ and $\mathbf{J}(\mathbf{f}, \boldsymbol{\zeta}) \in \mathbb{F}^{l \times l}$ to denote the Jacobian matrices of $\mathbf{r}$ and $\mathbf{f}$, respectively,

$$(\mathbf{J}(\mathbf{r}, \mathbf{x}))_{ij} = \frac{\partial \frac{p_i}{q_i}}{\partial x_j}, \quad (\mathbf{J}(\mathbf{f}, \boldsymbol{\zeta}))_{ij} = \frac{\partial f_i}{\partial \zeta_j}.$$

Further,

$$\operatorname{Range}(\mathbf{r}) = \{\mathbf{r}(\mathbf{x}) : \ q_1(\mathbf{x}) \cdots q_N(\mathbf{x}) \neq 0, \ \mathbf{x} \in \mathbb{C}^l\} \subset \mathbb{C}^N$$

denotes the set of all values of $\mathbf{r}(\mathbf{x})$ for $\mathbf{x} \in \mathbb{C}^l$. We say that the set $\operatorname{Range}(\mathbf{r})$ is invariant under scaling if

$$\operatorname{Range}(\mathbf{r}) \supseteq \lambda \cdot \operatorname{Range}(\mathbf{r}) \text{ for all } \lambda \in \mathbb{C}.$$

Theorem 1. *Let Ω be a subset of $\mathbb{F}^n$ and $\mu_n(\Omega) > 0$. Assume that*

1) *the matrix $\mathbf{A}(\mathbf{z})$ has full column rank for a generic choice of $\mathbf{z} \in \Omega$, that is,*
$$\mu_n\{\mathbf{z} \in \Omega : \ \operatorname{rank} \mathbf{A}(\mathbf{z}) < R\} = 0;$$
2) *the coordinate functions $f_1, \ldots, f_l$ of $\mathbf{f}$ can be represented as*
$$f_1(\boldsymbol{\zeta}) = \frac{f_{1,num}(\boldsymbol{\zeta})}{f_{1,den}(\boldsymbol{\zeta})}, \ldots, f_l(\boldsymbol{\zeta}) = \frac{f_{l,num}(\boldsymbol{\zeta})}{f_{l,den}(\boldsymbol{\zeta})},$$
where the functions
$$f_{1,num}(\boldsymbol{\zeta}), f_{1,den}(\boldsymbol{\zeta}), \ldots, f_{l,num}(\boldsymbol{\zeta}), f_{l,den}(\boldsymbol{\zeta})$$
are analytic on $\mathbb{C}^l$;
3) *there exists $\boldsymbol{\zeta}^0 \in \mathbb{C}^l$ such that $\det \mathbf{J}(\mathbf{f}, \boldsymbol{\zeta}^0) \neq 0$;*
4) *the dimension of the subspace spanned by the vectors of form (7) is at least $\widehat{N}$,*
$$\dim \operatorname{span}\{\mathbf{r}(\mathbf{x}) : \ q_1(\mathbf{x}) \cdots q_N(\mathbf{x}) \neq 0, \ \mathbf{x} \in \mathbb{C}^l\} \geq \widehat{N};$$
5) *$\operatorname{rank} \mathbf{J}(\mathbf{r}, \mathbf{x}) \leq \widehat{l}$ for a generic choice of $\mathbf{x} \in \mathbb{C}^l$;*
6) *$R \leq \widehat{N} - \widehat{l}$ or $R \leq \widehat{N} - \widehat{l} - 1$, depending on whether the set $\operatorname{Range}(\mathbf{r})$ is invariant under scaling or not.*

Then decomposition (5) is generically unique.

Proof. See Appendix A. □

Assumptions 1–6 can be used as a checklist for demonstrating the generic uniqueness of decompositions that can be put in the form (2). We will discuss two application examples in Sections III–IV. We comment on the following aspects of assumptions 2–6.

• In this paper we will use Theorem 1 in the case where $\mathbf{f}(\boldsymbol{\zeta})$ is of the form (4). For such $\mathbf{f}$ the matrix $\mathbf{J}(\mathbf{f}, \boldsymbol{\zeta})$ is diagonal, yielding that $\det \mathbf{J}(\mathbf{f}, \boldsymbol{\zeta}) = f_1'(\zeta_1) \cdots f_l'(\zeta_l)$. Moreover, in this paper $f_1, \ldots, f_l$ are non-constant, so $\det \mathbf{J}(\mathbf{f}, \boldsymbol{\zeta})$ is not identically zero. Thus, *assumption* 3 in Theorem 1 will hold automatically.

• For the reader who wishes to apply Theorem 1 in cases where $\mathbf{f}$ is not of the form (4), we recall the definition of an analytic (or holomorphic) function of several variables used in *assumption 2*. A function $f : \mathbb{C}^l \to \mathbb{C}$ of l complex variables is analytic [9, page 4] if it is analytic in each variable separately, that is, if for each $j = 1, \ldots, l$ and accordingly fixed $\zeta_1, \cdots \zeta_{j-1}, \zeta_{j+1}, \ldots, \zeta_l$ the function

$$z \mapsto f(\zeta_1, \cdots \zeta_{j-1}, z, \zeta_{j+1}, \ldots, \zeta_l)$$

is analytic on $\mathbb{C}$ in the classical one-variable sense. Examples of analytic functions of several variables can be obtained by taking compositions of multivariate polynomials and analytic functions in one variables, e.g. $f(\zeta_1, \zeta_2) = \sin(\cos(\zeta_1 \zeta_2)) + \zeta_1$.

• To check *assumption 4* in Theorem 1 it is sufficient to present (or prove the existence of) $\widehat{N}$ linearly independent vectors $\{\mathbf{r}(\mathbf{x}_i)\}_{i=1}^{\widehat{N}}$. It is clear that larger $\widehat{N}$ yield a better bound on R in *assumption* 6. In all cases considered in this paper $\widehat{N} = N$. The situation $\widehat{N} < N$ may appear when the $N \times 1$ vector-function $\mathbf{b}(\boldsymbol{\zeta})$ models a periodic, (locally) odd or even function, etc.

• The goal of *assumption 5* is to check whether generic signals of the form (7) can be re-parameterized with fewer (i.e. $\widehat{l} < l$) parameters. In this case, the Jacobian $\mathbf{J}(\mathbf{r}, \mathbf{x})$ has indeed rank strictly less than l. It is clear that *assumption* 5 in Theorem 1 holds trivially for $\widehat{l} = l$ and that smaller $\widehat{l}$ yield a better bound on R in *assumption* 6. In this paper we set either $\widehat{l} = l$ (namely in the proof of Theorem 5) or, in the case where it is clear that $\mathbf{J}(\mathbf{r}, \mathbf{x})$ does not have full column rank (namely in the proof of Theorems 2 and 6), $\widehat{l} = l - 1$.

• Although the Theorem holds both for $\mathbb{F} = \mathbb{C}$ and $\mathbb{F} = \mathbb{R}$, we formulated *assumptions* 3, 4 and 5 in Theorem 1 for $\boldsymbol{\zeta}^0 \in \mathbb{C}^l$ and $\mathbf{x} \in \mathbb{C}^l$. In these assumptions $\mathbb{C}^l$ can also be replaced by $\mathbb{R}^l$. We presented the complex variants, even for the case $\mathbb{F} = \mathbb{R}$, since they may be easier to verify than their real counterparts, as $\boldsymbol{\zeta}^0$ and $\mathbf{x}$ are allowed to take values in a larger set. On the other hand, the analyticity on $\mathbb{C}^l$ in *assumption* 2 is a stronger assumption than analyticity on $\mathbb{R}^l$ and is needed in the form it is given.

III. An application in independent component analysis

We consider data described by the model $\mathbf{x} = \mathbf{M}\mathbf{s}$, where $\mathbf{x}$ is the I-dimensional vector of observations, $\mathbf{s}$ is the R-dimensional unknown source vector and $\mathbf{M}$ is the I-by-R unknown mixing matrix. We assume that the sources are mutually uncorrelated but individually correlated in time. It is known that the spatial covariance matrices of the observations satisfy [10]

$$\begin{aligned} \mathbf{C}_1 &= \mathrm{E}(\mathbf{x}_t \mathbf{x}_{t+\tau_1}^H) = \mathbf{M}\mathbf{D}_1\mathbf{M}^H = \sum_{r=1}^{R} d_{1r}\mathbf{m}_r\mathbf{m}_r^H, \\ &\vdots \\ \mathbf{C}_P &= \mathrm{E}(\mathbf{x}_t \mathbf{x}_{t+\tau_P}^H) = \mathbf{M}\mathbf{D}_P\mathbf{M}^H = \sum_{r=1}^{R} d_{Pr}\mathbf{m}_r\mathbf{m}_r^H, \end{aligned} \tag{8}$$

in which $\mathbf{D}_p = \mathrm{E}(\mathbf{s}_t \mathbf{s}_{t+\tau_p}^H)$ is the R-by-R diagonal matrix with the elements of the vector $(d_{p1}, \ldots, d_{pR})$ on the main diagonal. The estimation of $\mathbf{M}$ from the set $\{\mathbf{C}_p\}$ is known as Second-Order Blind Identification (SOBI) [10] or as Second-Order Blind Identification of Underdetermined Mixtures (SOBIUM) [11] depending on whether the matrix $\mathbf{M}$ has full column rank or not. Variants of this problem are discussed in, e.g., [12], [13], [14], [15, Chapter 7]. It is clear that if the matrices $\mathbf{M}$ and $\mathbf{D}_1, \ldots, \mathbf{D}_P$ satisfy (8), then the matrices $\overline{\mathbf{M}} = \mathbf{M}\boldsymbol{\Lambda}\mathbf{P}$ and $\overline{\mathbf{D}}_1 = \mathbf{P}^T\mathbf{D}_1\mathbf{P}, \ldots, \overline{\mathbf{D}}_P = \mathbf{P}^T\mathbf{D}_P\mathbf{P}$ also satisfy (8) for any permutation matrix $\mathbf{P}$ and diagonal unitary matrix $\boldsymbol{\Lambda}$. We say that (8) has a unique solution when it is only subject to this trivial indeterminacy.

Generic uniqueness of solutions of (8) has been studied 1) in [16] and [8, Subsection 1.4.2] in the case where the superscript "H" in (8) is replaced by the superscript "T" (for quantities $\mathbf{x}$, $\mathbf{M}$ are $\mathbf{s}$ that can be either real valued or complex valued); 2) in [11], [17] (where $\mathbf{x}$, $\mathbf{M}$ are $\mathbf{s}$ are complex valued). In [8], [11], [17] the matrix equations in (8) were interpreted as a so-called canonical polyadic decomposition of a (partially symmetric) tensor. In the following theorems we interpret the equations in (8) as matrix factorization problem (2). The new interpretation only relies on elementary linear algebra; it does not make use of advanced results on tensor decompositions while it does lead to more relaxed bounds on R than in [11], [17] for $I \geq 5$. We consider the variants $\tau_p \neq 0$, $1 \leq p \leq P$, and $\tau_1 = 0$ in Theorems 2 and 3, respectively.

Theorem 2. *Assume that* $\tau_1 \neq 0$ *and*

$$R \leq \min(2P, (I-1)^2). \tag{9}$$

Then (8) *has a unique solution for generic matrices* $\mathbf{M}$ *and* $\mathbf{D}_1, \ldots, \mathbf{D}_P$*, i.e.,*

$$\mu_k\{(\mathrm{vec}(\mathbf{D}), \mathrm{vec}(\mathbf{M})) : \textit{ solution of } (8) \textit{ is not unique}\} = 0, \tag{10}$$

where $\mathbf{D}$ *denotes the* $P \times R$ *matrix with entries* d_{pr}*,* $k = IR + PR$*, and* μ_k *is a measure that is a.c. with respect to the Lebesgue measure on* $\mathbb{C}^k$ *.*

Proof. (i) First we rewrite the equations in (8) as matrix decomposition (5)[1]. In step (ii) we will apply Theorem 1 to (5).

Since $\mathbf{C}_p^H = \sum_{r=1}^{R} d_{pr}^* \mathbf{m}_r \mathbf{m}_r^H$, the pth equation in (8) is equivalent to the following pair of equations

$$\begin{aligned} \mathrm{Re}\,\mathbf{C}_p &= \frac{\mathbf{C}_p + \mathbf{C}_p^H}{2} = \sum_{r=1}^{R} \mathrm{Re}\, d_{pr}\mathbf{m}_r\mathbf{m}_r^H, \\ \mathrm{Im}\,\mathbf{C}_p &= \frac{\mathbf{C}_p - \mathbf{C}_p^H}{2i} = \sum_{r=1}^{R} \mathrm{Im}\, d_{pr}\mathbf{m}_r\mathbf{m}_r^H. \end{aligned}$$

Since $\mathrm{vec}(\mathbf{m}\mathbf{m}^H) = \mathbf{m}^* \otimes \mathbf{m}$, we further obtain that

$$\begin{aligned} &\mathrm{vec}(\mathrm{Re}\,\mathbf{C}_p)^T = \\ &\qquad [\mathrm{Re}\, d_{p1} \ldots \mathrm{Re}\, d_{pR}][\mathbf{m}_1^* \otimes \mathbf{m}_1 \ldots \mathbf{m}_R^* \otimes \mathbf{m}_R]^T, \\ &\mathrm{vec}(\mathrm{Im}\,\mathbf{C}_p)^T = \\ &\qquad [\mathrm{Im}\, d_{p1} \ldots \mathrm{Im}\, d_{pR}][\mathbf{m}_1^* \otimes \mathbf{m}_1 \ldots \mathbf{m}_R^* \otimes \mathbf{m}_R]^T. \end{aligned}$$

Hence, the P equations in (8) can be rewritten as $\mathbf{Y} = \mathbf{A}\mathbf{B}^T$, where

$$\begin{aligned} &\mathbf{Y} = \\ &[\mathrm{vec}(\mathrm{Re}\,\mathbf{C}_1) \ldots \mathrm{vec}(\mathrm{Re}\,\mathbf{C}_P)\ \ \mathrm{vec}(\mathrm{Im}\,\mathbf{C}_1) \ldots \mathrm{vec}(\mathrm{Im}\,\mathbf{C}_P)]^T, \\ &\mathbf{A} = \begin{bmatrix} \frac{\mathbf{D}+\mathbf{D}^*}{2} \\ \frac{\mathbf{D}-\mathbf{D}^*}{2i} \end{bmatrix} \in \mathbb{R}^{K \times R},\ K = 2P, \text{ and} \\ &\mathbf{B} = [\mathbf{m}_1^* \otimes \mathbf{m}_1\ \ \ldots\ \ \mathbf{m}_R^* \otimes \mathbf{m}_R] \in \mathbb{R}^{N \times R},\ N = I^2. \end{aligned}$$

Now we choose l, $\boldsymbol{\zeta}$, p_n, q_n, and $\mathbf{f}$ such that the columns of $\mathbf{B}$ are of the form (3). Note that the trivial parameterization

[1]Our derivation of a matrix version of (8) is similar to the derivation in [17, Subsection 5.2].

$\mathbf{b}(\boldsymbol{\zeta}) = \boldsymbol{\zeta}^* \otimes \boldsymbol{\zeta}$ with $\boldsymbol{\zeta} \in \mathbb{C}^I$ is not of the form (3) because of the conjugation. However, since for $\mathbf{m} = \operatorname{Re}\mathbf{m} + i \operatorname{Im}\mathbf{m}$,

$$\mathbf{m}^* \otimes \mathbf{m} = (\operatorname{Re}\mathbf{m} - i \operatorname{Im}\mathbf{m}) \otimes (\operatorname{Re}\mathbf{m} + i \operatorname{Im}\mathbf{m}),$$

the parameterization

$$\begin{aligned}\mathbf{b}(\boldsymbol{\zeta}) =& ([\zeta_1 \ \dots \ \zeta_I]^T - i[\zeta_{i+1} \ \dots \ \zeta_{2I}]^T \otimes \\ & ([\zeta_1 \ \dots \ \zeta_I]^T + i[\zeta_{i+1} \ \dots \ \zeta_{2I}]^T), \quad \boldsymbol{\zeta} \in \mathbb{R}^l\end{aligned}$$

with $l = 2I$, is of the form (3). As a matter of fact, each component of $\mathbf{b}(\boldsymbol{\zeta})$ is a polynomial p_n in $\zeta_1, \dots, \zeta_l$, $1 \le n \le N$, so we can set $\mathbf{f}(\boldsymbol{\zeta}) = \boldsymbol{\zeta}$, and $q_1(\boldsymbol{\zeta}) = \cdots = q_N(\boldsymbol{\zeta}) = 1$.

It is clear that the matrix $\mathbf{A}$ can be parameterized independently of $\mathbf{B}$ by $m = 2PR$ real parameters, namely, by the entries of the $P \times R$ matrices $\frac{\mathbf{D}+\mathbf{D}^*}{2}$ and $\frac{\mathbf{D}-\mathbf{D}^*}{2i}$. Thus, the equations in (8) can be rewritten as decomposition (5) with $\mathbf{z} \in \Omega = \mathbb{R}^n$, where $n = m + s = 2PR + lR = 2PR + 2IR$. Moreover, one can easily verify that (8) has a unique solution if and only if decomposition (5) is unique. In turn, since, obviously, (10) is equivalent to

$$\begin{aligned}\mu_n \{(\operatorname{vec}((\mathbf{D} + \mathbf{D}^*)/2), \operatorname{vec}((\mathbf{D} - \mathbf{D}^*)/2i), \\ \operatorname{Re}\mathbf{m}_1, \operatorname{Im}\mathbf{m}_1, \dots, \operatorname{Re}\mathbf{m}_R, \operatorname{Im}\mathbf{m}_R) : \\ \text{solution of (8) is not unique}\} = 0,\end{aligned}$$

it follows that (10) can be rewritten as (6).

(ii) To prove (6) we check assumptions 1–6 in Theorem 1. Assumption 1: it is clear that if $\mathbf{D}$ is generic, then, by the assumption $2P \ge R$, the matrix $\mathbf{A}$ has full column rank. Assumptions 2–3 are trivial since $\mathbf{f}$ is the identity mapping. Assumption 4: since the rank-1 matrices of the form $\mathbf{m}\mathbf{m}^H$ span the whole space of $I \times I$ matrices and $\mathbf{b}(\operatorname{Re}\mathbf{m}, \operatorname{Im}\mathbf{m}) = \operatorname{vec}(\mathbf{m}\mathbf{m}^H)$ it follows that assumption 4 holds for $\widehat{N} = I^2$. Assumption 5: an elementary computation shows that for a generic $\boldsymbol{\zeta}$, $\mathbf{J}(\mathbf{r}, \mathbf{x})[x_{I+1} \ \dots \ x_{2I} \ -x_1 \ \dots \ -x_I] = \mathbf{0}$, implying that $\operatorname{rank}(\mathbf{J}(\mathbf{r}, \mathbf{x})) \le l - 1$, so we set $\widehat{l} = l - 1$. Assumption 6: since $\widehat{N} - \widehat{l} = I^2 - 2I + 1$, assumption 6 holds by (9) since $\lambda \mathbf{r}(\boldsymbol{\zeta}) = \lambda \mathbf{b}(\boldsymbol{\zeta}) = \mathbf{b}(\sqrt{\lambda}\boldsymbol{\zeta}) = \mathbf{r}(\sqrt{\lambda}\boldsymbol{\zeta})$. □

Now we consider the case $\tau_1 = 0$. The only difference with the case $\tau_1 \neq 0$ is that the diagonal matrix $\mathbf{D}_1 = \mathrm{E}(\mathbf{s}_t \mathbf{s}_{t+\tau_1}^H)$ is real, yielding that (8) can be parameterized by R real and $IR + (P-1)R$ complex parameters, or equivalently, by $n = R + 2IR + 2(P-1)R$ real parameters.

Theorem 3. *Assume that $\tau_1 = 0$ and $R \le \min(2P - 1, (I - 1)^2)$. Then* (8) *has a unique solution for generic real matrix $\mathbf{D}_1$ and generic complex matrices $\mathbf{M}$ and $\mathbf{D}_2, \dots, \mathbf{D}_P$, i.e.,*

$$\begin{aligned}\mu_n \Big\{ d_{11}, \dots, d_{1R}, \Big(\operatorname{vec}((\overline{\mathbf{D}} + \overline{\mathbf{D}}^*)/2), \operatorname{vec}((\overline{\mathbf{D}} - \overline{\mathbf{D}}^*)/2i), \\ \operatorname{Re}\mathbf{m}_1, \operatorname{Im}\mathbf{m}_1, \dots, \operatorname{Re}\mathbf{m}_R, \operatorname{Im}\mathbf{m}_R) : \\ \textit{solution of } (8) \textit{ is not unique} \Big\} = 0,\end{aligned}$$

where $\overline{\mathbf{D}} \in \mathbb{C}^{(P-1)\times R}$ denotes a matrix with entries d_{pr} ($p > 1$), $n = (2I + 2P - 1)R$, and μ_n is a measure that is a.c. with respect to the Lebesgue measure on $\mathbb{R}^n$.

Proof. The proof is essentially the same as that of Theorem 2. □

TABLE I
UPPER BOUNDS ON THE NUMBER OF SOURCES IN SOBI

	I	3	4	5	6	7	8	9
Theorem 2	$\mathbb{F} = \mathbb{C}$	4	9	16	25	36	49	64
[11, Eq. (15)]	$\mathbb{F} = \mathbb{C}$	4	9	14	21	30	40	51
[8, Proposition 1.11]	$\mathbb{F} = \mathbb{R}^*$	3	6	10	15	21	28	36

*or $\mathbb{F} = \mathbb{C}$ if the superscript "H" in (8) is replaced by the superscript "T"

Assuming that $R \le P$, we check up to which value of R condition (9) in Theorem 2 and conditions $R(R-1) \le I^2(I-1)^2/2$ in [11] and $R \le (I^2 - I)/2$ in [8] hold. The results are shown in Table I. Note that under the condition in [11] the mixing matrix $\mathbf{M}$ can be found from an eigenvalue decomposition in the exact case. Hence, it is not surprising that this condition is more restrictive. The condition in [8] is more restrictive since, if $\mathbf{D}_p$ is complex, the unsymmetric matrix $\mathbf{M}\mathbf{D}_p\mathbf{M}^H$ has more distinct entries than the complex symmetric matrix $\mathbf{M}\mathbf{D}_p\mathbf{M}^T$.

IV. An application in deterministic signal separation using mild source models

A. Context and contribution

We have recently proposed tensor-based algorithms for the deterministic blind separation of signals that can be modeled as exponential polynomials (i.e., sums and/or products of exponentials, sinusoids and/or polynomials) [18] or as rational functions [19]. These signal models are meant to be little restrictive; on the other hand, they enable a unique source separation under certain conditions. The approach is somewhat related to sparse modelling [1]. In sparse modelling, matrix $\mathbf{M}$ in (1) is known but has typically more columns than rows while most of the entries of $\mathbf{S}$ are zero. That is, the nonzero entries of $\mathbf{S}$ make sparse combinations of the columns of $\mathbf{M}$ (called the "dictionary") to model $\mathbf{X}$. The uniqueness of the model depends on the degree of linear independence of the columns of $\mathbf{M}$ and the degree of sparsity of the rows of $\mathbf{S}$ [1]. In [18], [19] on the other hand, the basis vectors are estimated as well, by optimization over continuous variables. By way of example, in the case of sparse modelling of a sine wave, the columns of $\mathbf{M}$ could be chosen as sampled versions of $\sin((\omega_0 + k\Delta\omega)t)$ for a number of values k (say $k = -K, \dots, -1, 0, 1, \dots, K$ so that $R = 2K+1$), and ω_0 and $\Delta\omega$ are fixed. On the other hand, in [18] one optimizes over a continuous variable ω to determine the best representation $\sin(\omega t)$; in this way the accuracy is not bounded by $\Delta\omega$.

In [18], [19] deterministic uniqueness conditions are given for exponential polynomial and rational source models. Here, we propose generic uniqueness conditions for the case that the mixing matrix has full column rank.

We actually consider a more general family of models, namely we assume that the source signals $s_1(t), \dots, s_R(t)$ can be modeled as the composition of a known multivariate rational function and functions of the type t, $\cos(\omega t + \phi)$, $\sin(\omega t + \phi)$, and a^t. We assume that the discrete-time signals

are obtained by sampling at the points $t = 1, \dots, N$. The observed data are a mixture of the sources:

$$\mathbf{X} = \mathbf{M} \begin{bmatrix} s_1(1) & \dots & s_1(N) \\ \vdots & \vdots & \vdots \\ s_R(1) & \dots & s_R(N) \end{bmatrix} = \mathbf{M}\mathbf{S}^T. \quad (11)$$

B. An example

To simplify the presentation we will consider the concrete case where the source signals can be modelled as

$$s_r(t) = \frac{a_r^t}{t} + \frac{b_r + t}{c_r + t}\cos(\alpha_r t + \phi_r) + \cos(\beta_r t), \quad t \in \mathbb{R} \quad (12)$$

for a priori unknown parameters a_r, b_r, c_r, α_r, ϕ_r and β_r. That is, $s_r(t)$ is the composition of the known rational function

$$R(x_1, \dots, x_6) = \frac{x_1}{x_2} + \frac{x_3 + x_2}{x_4 + x_2}x_5 + x_6$$

and the functions $x_1(t) = a_r^t$, $x_2(t) = t$, $x_3(t) = b_r$, $x_4(t) = c_r$, $x_5(t) = \cos(\alpha_r t + \phi_r)$, and $x_6 = \cos\beta_r t$. The general case can be studied similarly.

In the remaining part of this subsection we show that if (i) $R \leq N - 6$, (ii) the parameters a_r, b_r, c_r, α_r, ϕ_r, and β_r are generic, and (iii) the mixing matrix $\mathbf{M}$ has full column rank, then the mixing matrix and the sources $s_1(t), \dots, s_R(t)$ can be uniquely recovered.

We rewrite (11) as matrix decomposition (5). We set $\mathbf{Y} = \mathbf{X}$ and $\mathbf{A}(\mathbf{z}) = \mathbf{M}$. It is clear that the signals in (12) can be parameterized as

$$s(t) = \frac{\zeta_1^t}{t} + \frac{\zeta_2 + t}{\zeta_3 + t}\cos(\zeta_4 t + \zeta_5) + \cos(\zeta_6 t), \quad t \in \mathbb{R}, \quad (13)$$

where $\boldsymbol{\zeta} = [\zeta_1 \ \dots \ \zeta_6]^T = [a \ b \ c \ \alpha \ \phi \ \beta]^T$, so we set $\mathbf{b}(\boldsymbol{\zeta}) = [s(1) \ \dots \ s(N)]^T$. First, we bring $\mathbf{b}(\boldsymbol{\zeta})$ into the form (3). Then we will check assumptions 1–6 in Theorem 1.

The following identities are well-known:

$$\cos\zeta = \frac{1 - \tan^2\frac{\zeta}{2}}{1 + \tan^2\frac{\zeta}{2}}, \quad \sin\zeta = \frac{2\tan\frac{\zeta}{2}}{1 + \tan^2\frac{\zeta}{2}}. \quad (14)$$

We will need the following generalization of (14).

Lemma 4. *There exist a polynomial P_n and rational functions Q_n and R_n such that*

$$\cos\zeta n = P_n(\cos\zeta) = Q_n\left(\tan\frac{\zeta}{2}\right), \quad (15)$$

$$\sin\zeta n = R_n\left(\tan\frac{\zeta}{2}\right). \quad (16)$$

Proof. See Appendix B. □

From (13) and Lemma 4 it follows that

$$s(n) = \frac{\zeta_1^n}{n} + \frac{\zeta_2 + n}{\zeta_3 + n}\cos(\zeta_4 n + \zeta_5) + \cos(\zeta_6 n) =$$
$$\frac{\zeta_1^n}{n} + \frac{\zeta_2 + n}{\zeta_3 + n}(\cos\zeta_4 n\cos\zeta_5 - \sin\zeta_4 n\sin\zeta_5) + \cos(\zeta_6 n) =$$
$$\frac{\zeta_1^n}{n} + \frac{\zeta_2 + n}{\zeta_3 + n}\left(Q_n\left(\tan\frac{\zeta_4}{2}\right)\frac{1 - \tan^2\frac{\zeta_5}{2}}{1 + \tan^2\frac{\zeta_5}{2}} - \right.$$
$$\left. R_n\left(\tan\frac{\zeta_4}{2}\right)\frac{2\tan\frac{\zeta_5}{2}}{1 + \tan^2\frac{\zeta_5}{2}}\right) + P_n(\cos\zeta_6) = \frac{p_n(\mathbf{f}(\boldsymbol{\zeta}))}{q_n(\mathbf{f}(\boldsymbol{\zeta}))},$$

where

$$\frac{p_n(\mathbf{x})}{q_n(\mathbf{x})} = \frac{x_1^n}{n} + \frac{x_2 + n}{x_3 + n}\left(Q_n(x_4)\frac{1 - x_5^2}{1 + x_5^2} - \right.$$
$$\left. R_n(x_4)\frac{2x_5}{1 + x_5}\right) + P_n(x_6),$$
$$\mathbf{f}(\zeta_1, \zeta_2, \zeta_3, \zeta_4, \zeta_5, \zeta_6) = [\zeta_1 \ \zeta_2 \ \zeta_3 \ \tan\frac{\zeta_4}{2} \ \tan\frac{\zeta_5}{2} \ \cos\zeta_6]^T.$$

Thus, $\mathbf{b}(\boldsymbol{\zeta}) = [s(1) \ \dots \ s(N)]^T$ is of the form (3) and $l = 6$. Now we check assumptions 1–6 in Theorem 1: 1) holds by our assumption (iii); 2) and 3) are trivial; 4) holds for $\widehat{N} = N$ since the vectors $\mathbf{b}(\zeta_1, \zeta_2, \dots, \zeta_6) - \mathbf{b}(0, \zeta_2, \dots, \zeta_6) = [\frac{\zeta_1}{1} \ \dots \ \frac{\zeta_1^N}{N}]^T$ span the entire space $\mathbb{F}^N$; 5) holds for $\widehat{l} = l = 6$; 6) holds by assumption (i).

C. Separation of exponential polynomials and separation of rational functions

The cases where the sources in (11) can be expressed as sampled exponential polynomials

$$s(n) = \sum_{f=1}^{F}(p_{0f} + p_{1f}n + \dots + p_{d_f f}n^{d_f})a_f^n = \sum_{f=1}^{F} P_f(n)a_f^n, \quad n = 1, \dots, N \quad (17)$$

and sampled rational functions

$$s(n) = \frac{a_0 + a_1 n + \dots + a_p n^p}{b_0 + b_1 n + \dots + b_q n^q}, \quad n = 1, \dots, N \quad (18)$$

were studied in [18] and [19], respectively.

The following two theorems complement results on generic uniqueness from [18] and [19]. In contrast to papers [18] and [19] we do not exploit specific properties of Hankel or Löwner matrices in our derivation. We only use the source models (17)–(18) for verifying the assumptions in Theorem 1.

Theorem 5. *Assume that the mixing matrix $\mathbf{M}$ has full column rank and that*

$$R \leq N - (d_1 + \dots d_F + 2F), \quad (19)$$

then $\mathbf{M}$ and R generic sources of form (17) *can be uniquely recovered from the observed data $\mathbf{X} = \mathbf{M}\mathbf{S}^T$.*

Proof. We set

$$\boldsymbol{\zeta} = [a_1 \ p_{01} \ \dots \ p_{d_1 1} \ \dots \ a_F \ p_{0F} \ \dots \ p_{d_F F}]^T \in \mathbb{F}^l,$$
$$l = (2 + d_1) + \dots + (2 + d_F) = d_1 + \dots + d_F + 2F$$

and check the assumptions in Theorem 1 for $\mathbf{Y} = \mathbf{X}$, $\mathbf{A}(\mathbf{z}) = \mathbf{M}$ and $\mathbf{b}(\boldsymbol{\zeta}) = [s(1) \ \dots \ s(N)]^T$: 1)–3) are trivial; 4) since the vectors $\mathbf{b}(\zeta, 1, 0, \dots, 0) = [\zeta \ \dots \ \zeta^N]^T$ span the entire space $\mathbb{F}^N$, we set $\widehat{N} = N$; 5) we set $\widehat{l} = l$; 6) holds by (19) since $\mathrm{Range}(\mathbf{r})$ is invariant under scaling. □

Theorem 6. *Assume that the mixing matrix $\mathbf{M}$ has full column rank, $q \geq 1$, and that*

$$R \leq N - (p + q + 1), \quad (20)$$

then $\mathbf{M}$ *and* R *generic sources of form* (18) *can be uniquely recovered from the observed data* $\mathbf{X} = \mathbf{M}\mathbf{S}^T$.

Proof. We set

$$\boldsymbol{\zeta} = [a_0 \ \dots \ a_p \ b_0 \ \dots \ b_q]^T \in \mathbb{F}^l, \quad l = p + q + 2$$

and check the assumptions in Theorem 1 for $\mathbf{Y} = \mathbf{X}$, $\mathbf{A}(\mathbf{z}) = \mathbf{M}$ and $\mathbf{r}(\boldsymbol{\zeta}) = \mathbf{b}(\boldsymbol{\zeta}) = [s(1) \ \dots \ s(N)]^T$: 1)–3) are trivial; 4) since an $N \times N$ matrix with $(k+1)$th column (for $k = 0, \dots, N-1$) given by

$$\mathbf{b}(\underbrace{1, 0, \dots, 0}_{p+1}, k, 1, 0, \dots, 0) = [(k+1)^{-1} \ \dots \ (k+N)^{-1}]^T,$$

is nonsingular [20, p. 38], we set $\widehat{N} = N$; 5) an elementary computation shows that for a generic $\mathbf{x}$, $\mathbf{J}(\mathbf{r}, \mathbf{x})\mathbf{x} = \mathbf{0}$, implying that $\operatorname{rank}(\mathbf{J}(\mathbf{r}, \mathbf{x})) \leq l - 1$, so we set $\widehat{l} = l - 1$; 6) holds by (20) since $\operatorname{Range}(\mathbf{r})$ is invariant under scaling. □

We assume that the matrix $\mathbf{M}$ is generic and compare the bounds in Theorem 5 and Theorem 6 with the generic bounds in [18] and [19], respectively. Since $\mathbf{M}$ is generic, it has full column rank if and only if $R \leq K$. Thus, we compare the bound $R \leq \min(N - (d_1 + \dots d_F + 2F), K)$ with the bound $R(d_1 + \dots d_F + F) \leq \lfloor \frac{N+1}{2} \rfloor$, $2 \leq K$ in [18], and the bound $R \leq \min(N - (p + q + 1), K)$ with the bound $R \leq \frac{1}{\max(p,q)} \lfloor \frac{N+1}{2} \rfloor$, $2 \leq K$ in [19]. On one hand, the bounds in [18] and [19] can be used in the undetermined case ($2 \leq K$), while our bounds work only in the overdetermined case ($R \leq K$). On the other hand, roughly speaking, our bounds are of the form $R \leq N - c$ while the bounds in [18] and [19] are of the form $R \leq N/c$, where c is the number of parameters that describe a generic signal. In this sense our new uniqueness conditions are significantly more relaxed.

V. Conclusion

Borrowing insights from algebraic geometry, we have presented a theorem that can be used for investigating generic uniqueness in BSS problems that can be formulated as a particular structured matrix factorization. We have used this tool for deriving generic uniqueness conditions in (i) SOBIUM-type independent component analysis and (ii) a class of deterministic BSS approaches that rely on parametric source models. In a companion paper we will use the tool to obtain generic results for structured tensor and coupled matrix/tensor factorizations.

Appendix A
Proof of Theorem 1

In this appendix we consider the decomposition

$$\mathbf{Y} = \mathbf{A}\mathbf{B}^T = \sum_{r=1}^{R} \mathbf{a}_r \mathbf{b}_r^T, \quad \mathbf{b}_r \in S, \tag{21}$$

where the matrix $\mathbf{A}$ has full column rank and S denotes a known subset of $\mathbb{F}^N$.

In Theorem 7 below, we present two conditions that guarantee the uniqueness of decomposition (21). These conditions will be checked in the proof of Theorem 1 for generic points in $S = \{\mathbf{b}(\boldsymbol{\zeta}) : \ q_1(\mathbf{f}(\boldsymbol{\zeta})) \cdots q_N(\mathbf{f}(\boldsymbol{\zeta})) \neq 0\}$, where $\mathbf{b}(\boldsymbol{\zeta})$ is defined in (3). The latter proof is given in Subsection A-C. The step from the deterministic formulation in Subsection A-A to the generic result in Subsection A-C is taken in Subsection A-B.

A. A deterministic uniqueness result

Theorem 7. *Assume that*

1) *the matrix* $\mathbf{A}$ *has full column rank;*
2) *the columns* $\mathbf{b}_1, \dots, \mathbf{b}_R$ *of the matrix* $\mathbf{B}$ *satisfy the following condition:*

$$\begin{aligned} &\textit{if at least two of the values } \ \lambda_1, \dots, \lambda_R \in \mathbb{F} \\ &\textit{are nonzero, then } \ \lambda_1 \mathbf{b}_1 + \dots + \lambda_R \mathbf{b}_R \notin S. \end{aligned} \tag{22}$$

Then decomposition (21) *is unique.*

Proof. We need to show that if there exist $\overline{\mathbf{A}}$ and $\overline{\mathbf{B}}$ such that

$$\mathbf{Y} = \overline{\mathbf{A}}\overline{\mathbf{B}}^T = \sum_{r=1}^{R} \bar{\mathbf{a}}_r \bar{\mathbf{b}}_r^T, \quad \bar{\mathbf{b}}_r \in S \tag{23}$$

then decompositions (21) and (23) coincide up to permutation of the rank-1 terms.

First we show that assumption 2 implies that $\mathbf{B}$ has full column rank. Assume that there exist $\lambda_1, \dots, \lambda_R$ for which $\lambda_1 \mathbf{b}_1 + \dots + \lambda_R \mathbf{b}_R = \mathbf{0}$, such that at least one of these values being nonzero would imply that $\mathbf{B}$ does not have full column rank.

Then for any $\mu \notin \{0, -\lambda_1\}$, $\frac{\lambda_1+\mu}{\mu}\mathbf{b}_1 + \frac{\lambda_2}{\mu}\mathbf{b}_2 + \dots + \frac{\lambda_R}{\mu}\mathbf{b}_R = \mathbf{b}_1 \in S$. Hence, by assumption 2, at most one of the values $\lambda_1 + \mu, \lambda_2, \dots, \lambda_R$ is nonzero. Since $\mu \neq -\lambda_1$, we have that $\lambda_2 = \dots \lambda_R = 0$. Since $\lambda_1 \mathbf{b}_1 = \lambda_1 \mathbf{b}_1 + \dots + \lambda_R \mathbf{b}_R = \mathbf{0}$, it follows that $\lambda_1 = 0$ or $\mathbf{b}_1 = \mathbf{0}$. One can easily verify that $\mathbf{b}_1 = \mathbf{0}$ is in contradiction to assumption 2. Hence $\lambda_1 = 0$. Thus the matrix $\mathbf{B}$ has full column rank.

Since the matrices $\mathbf{A}$ and $\mathbf{B}$ have full column rank, it follows from the identity

$$\mathbf{Y} = \mathbf{A}\mathbf{B}^T = \overline{\mathbf{A}}\overline{\mathbf{B}}^T \tag{24}$$

that the matrices $\overline{\mathbf{A}}$ and $\overline{\mathbf{B}}$ also have full column rank. Hence,

$$\overline{\mathbf{A}}^\dagger \mathbf{A}\mathbf{B}^T = \overline{\mathbf{B}}^T \tag{25}$$

where $\overline{\mathbf{A}}^\dagger = \left(\overline{\mathbf{A}}^H \overline{\mathbf{A}}\right)^{-1} \overline{\mathbf{A}}^H$ denotes the left inverse of $\overline{\mathbf{A}}$. By assumption 2, each row of the matrix $\overline{\mathbf{A}}^\dagger \mathbf{A}$ contains at most one nonzero entry. Since the matrices $\mathbf{B}$ and $\overline{\mathbf{B}}$ have full column rank, the square matrix $\overline{\mathbf{A}}^\dagger \mathbf{A}$ is nonsingular. Thus, each row and each column of $\overline{\mathbf{A}}^\dagger \mathbf{A}$ contains exactly one nonzero entry. Hence there exist an $R \times R$ nonsingular diagonal matrix $\boldsymbol{\Lambda}$ and an $R \times R$ permutation matrix $\mathbf{P}$ such that $\overline{\mathbf{A}}^\dagger \mathbf{A} = \boldsymbol{\Lambda}\mathbf{P}$. From (25) it follows that

$$\boldsymbol{\Lambda}\mathbf{P}\mathbf{B}^T = \overline{\mathbf{A}}^\dagger \mathbf{A}\mathbf{B} = \overline{\mathbf{B}}^T. \tag{26}$$

Substituting (26) into (24) and taking into account that the matrix $\mathbf{B}$ has full column rank we obtain

$$\mathbf{A}\mathbf{P}^T \boldsymbol{\Lambda}^{-1} = \overline{\mathbf{A}}. \tag{27}$$

Equations (26)–(27) imply that decompositions (21) and (23) coincide up to permutation of the rank-1 terms. □

Theorem 7 has already been proved for the particular cases where decomposition (21) represents the CPD of a third-order tensor [21, Section IV], the CPD of a partially symmetric of order higher than three [22, Theorem 4.1], the CPD of an unstructured tensor of order higher than three [23, Theorem 4.2], and the decomposition in multilinear rank-$(L, L, 1)$ terms [18, Theorem 2.4].

B. A generic variant of assumption 2 in Theorem 7

Condition (22) means that the subspace $\operatorname{span}\{\mathbf{b}_1, \ldots, \mathbf{b}_R\}$ has dimension R and may intersect the set S only at "trivial" points $\lambda_r \mathbf{b}_r$, that is

$$\text{the vectors } \mathbf{b}_1, \ldots, \mathbf{b}_R \text{ are linearly independent and} \quad (28)$$

$$\operatorname{span}\{\mathbf{b}_1, \ldots, \mathbf{b}_R\} \cap S \subseteq \{\lambda \mathbf{b}_r : \ \lambda \in \mathbb{F}, \ 1 \leq r \leq R\}. \quad (29)$$

Property (29) is the key to proving uniqueness of (21). We can easily find $\operatorname{span}\{\mathbf{b}_1, \ldots, \mathbf{b}_R\}$ from the matrix $\mathbf{Y}$ if it can be assumed that the matrix $\mathbf{A}$ has full column rank. On the other hand, property (29) means that the only points in $\operatorname{span}\{\mathbf{b}_1, \ldots, \mathbf{b}_R\}$ that have the hypothesized structure (encoded in the definition of the set S), are the vectors $\mathbf{b}_r$, $1 \leq r \leq R$ (up to trivial indeterminacies). However, conditions (22) and (29) are most often hard to check for particular points $\mathbf{b}_1, \ldots, \mathbf{b}_R$. The checking may become easier if we focus on the generic case, and this is where algebraic geometry comes in. More precisely, if $S = V$ is an algebraic variety, then the classical trisecant lemma states that if R is sufficiently small, then (29) holds for "generic" $\mathbf{b}_1, \ldots, \mathbf{b}_R \in S$. A set $V \subseteq \mathbb{C}^N$ is an algebraic variety if it is the set of solutions of a system of polynomial equations. It is clear that algebraic varieties form an interesting class of subsets of $\mathbb{C}^N$; however, is not easy to verify whether a given subset of $\mathbb{C}^N$ is a variety or not. On the other hand, it is known that a set obtained by evaluating a known rational vector-function (such as $\mathbf{r}(\mathbf{x})$ in (7)) can be extended to a variety by taking the closure, i.e., by including its boundary. This is indeed what we will do in the proof of Lemma 9 below. First we give a formal statement of the trisecant lemma.

Lemma 8. *([24, Corollary 4.6.15], [25, Theorem 1.4]) Let $V \subset \mathbb{C}^N$ be an irreducible algebraic variety and $R \leq \dim \operatorname{span}\{V\} - \dim V$ or $R \leq \dim \operatorname{span}\{V\} - \dim V - 1$ depending on whether V is invariant under scaling or not. Let G_V denote a set of points $(\mathbf{v}_1, \ldots, \mathbf{v}_R)$ such that*

$$\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_R\} \cap V \not\subset \{\lambda \mathbf{v}_r : \ \lambda \in \mathbb{C}, \ 1 \leq r \leq R\}.$$

Then the Zariski closure of G_V is a proper subvariety of $V \times \cdots \times V$ (R times), that is, there exists a polynomial $h(\mathbf{v}_1, \ldots, \mathbf{v}_R)$ in RN variables whose zero set does not contain $V \times \cdots \times V$ but does contain G_V.

It is the last sentence in the trisecant lemma that makes it a powerful tool for proving generic properties. Let us explain in more detail how this works. We can use G_V to denote a set that poses problems in terms of uniqueness, in the sense that $\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_R\}$ does *not* intersect V only in the points that correspond to the pure sources. The trisecant lemma states now that G_V belongs to the zero set of a polynomial h that is not identically zero and hence nonzero almost everywhere, i.e. the problematic cases occur in a measure-zero situation. In order to make the connection with Theorem 1 we will need the following notations:

$$\operatorname{Range}(\mathbf{b}) := \{\mathbf{b}(\boldsymbol{\zeta}) : \ q_1(\mathbf{f}(\boldsymbol{\zeta})) \cdots q_N(\mathbf{f}(\boldsymbol{\zeta})) \neq 0, \ \boldsymbol{\zeta} \in \mathbb{F}^l\},$$
$$\operatorname{Range}(\mathbf{r}) := \{\mathbf{r}(\mathbf{x}) : \ q_1(\mathbf{x}) \cdots q_N(\mathbf{x}) \neq 0, \ \mathbf{x} \in \mathbb{F}^l\}.$$

Lemma 9. *Let assumptions 2–6 in Theorem 1 hold. Then assumption 2 in Theorem 7 holds for $S = \operatorname{Range}(\mathbf{b})$ and $\mathbf{b}_1 = \mathbf{b}(\boldsymbol{\zeta}_1), \ldots, \mathbf{b}_R = \mathbf{b}(\boldsymbol{\zeta}_R) \in S$, where the vectors $\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_R \in \mathbb{F}^l$ are generic.*

Proof. Since (28)–(29) is equivalent to (22) it is sufficient to show that $\mu_{Rl}(W_{\mathbf{b}}) = \mu_{Rl}(G_{\mathbf{b}})$=0, where

$$W_{\mathbf{b}} = \{[\boldsymbol{\zeta}_1^T \ \ldots \ \boldsymbol{\zeta}_R^T]^T : \mathbf{b}_1 = \mathbf{b}(\boldsymbol{\zeta}_1), \ldots, \mathbf{b}_R = \mathbf{b}(\boldsymbol{\zeta}_R) \text{ are linearly dependent}\},$$
$$G_{\mathbf{b}} = \{[\boldsymbol{\zeta}_1^T \ \ldots \ \boldsymbol{\zeta}_R^T]^T : \text{ (29) does not hold for } \mathbf{b}_1 = \mathbf{b}(\boldsymbol{\zeta}_1), \ldots, \mathbf{b}_R = \mathbf{b}(\boldsymbol{\zeta}_R)\}.$$

It is a well-known fact that the zero set of a nonzero analytic function on $\mathbb{C}^{Rl}$ has measure zero both on $\mathbb{C}^{Rl}$ and $\mathbb{R}^{Rl}$. Thus, to prove $\mu_{Rl}(W_{\mathbf{b}}) = \mu_{Rl}(G_{\mathbf{b}})$=0, we will show that there exist analytic functions w and g of Rl complex variables such that

$$w \text{ is not identically zero but vanishes on } W_{\mathbf{b}}, \quad (30)$$
$$g \text{ is not identically zero but vanishes on } G_{\mathbf{b}}. \quad (31)$$

We consider the following three cases: 1) $\mathbb{F} = \mathbb{C}$ and $\mathbf{f}(\boldsymbol{\zeta}) = \boldsymbol{\zeta}$; 2) $\mathbb{F} = \mathbb{C}$ and $\mathbf{f}(\boldsymbol{\zeta})$ is arbitrary; 3) $\mathbb{F} = \mathbb{R}$.

1) *Case $\mathbb{F} = \mathbb{C}$ and $\mathbf{f}(\boldsymbol{\zeta}) = \boldsymbol{\zeta}$.* In this case $\mathbf{b}(\boldsymbol{\zeta}) = \mathbf{r}(\boldsymbol{\zeta})$, thus, the sets $W_{\mathbf{b}}$ and $G_{\mathbf{b}}$ take the following form:

$$W_{\mathbf{b}} = W_{\mathbf{r}} = \{[\boldsymbol{\zeta}_1^T \ \ldots \ \boldsymbol{\zeta}_R^T]^T : \mathbf{b}_1 = \mathbf{r}(\boldsymbol{\zeta}_1), \ldots, \mathbf{b}_R = \mathbf{r}(\boldsymbol{\zeta}_R) \text{ are linearly dependent}\},$$
$$G_{\mathbf{b}} = G_{\mathbf{r}} = \{[\boldsymbol{\zeta}_1^T \ \ldots \ \boldsymbol{\zeta}_R^T]^T : \text{ (29) does not hold for } S = \operatorname{Range}(\mathbf{r}) \text{ and } \mathbf{b}_1 = \mathbf{r}(\boldsymbol{\zeta}_1), \ldots, \mathbf{b}_R = \mathbf{r}(\boldsymbol{\zeta}_R)\}.$$

Here we prove that there exist polynomials d_{num} and h_{num} in Rl variables such that (30)–(31) hold for $w = d_{num}$ and $g = h_{num}$.

First we focus on $G_{\mathbf{r}}$. Let V denote the Zariski closure of $\operatorname{Range}(\mathbf{r}) \subset \mathbb{C}^N$. Since $\operatorname{Range}(\mathbf{r})$ is the image of the open (hence irreducible) subset $\{\boldsymbol{\zeta} : \ q_1(\boldsymbol{\zeta}) \cdots q_N(\boldsymbol{\zeta}) \neq 0, \ \boldsymbol{\zeta} \in \mathbb{C}^l\}$ under the rational map

$$\mathbf{r} : \ \boldsymbol{\zeta} \mapsto \left[\frac{p_1(\boldsymbol{\zeta})}{q_1(\boldsymbol{\zeta})}, \ldots, \frac{p_N(\boldsymbol{\zeta})}{q_N(\boldsymbol{\zeta})}\right]^T,$$

it follows that $\operatorname{Range}(\mathbf{r})$ is also an irreducible set. Hence $V \subset \mathbb{C}^N$ is an irreducible variety and the dimension of V is equal to $\operatorname{rank} \mathbf{J}(\mathbf{r}, \boldsymbol{\zeta})$ at a generic point $\boldsymbol{\zeta} \in \mathbb{C}^l$ [26, p. 186]. Hence, by assumption 5 in Theorem 1,

$$\dim V \leq \widehat{l}. \quad (32)$$

Since, by definition, $\operatorname{Range}(\mathbf{r})$ consists of all vectors of the form (7), from assumption 4 in Theorem 1 it follows that

$$\dim \operatorname{span} \operatorname{Range}(\mathbf{r}) =$$
$$\dim \operatorname{span}\{\mathbf{r}(\boldsymbol{\zeta}) : \ q_1(\boldsymbol{\zeta}) \cdots q_N(\boldsymbol{\zeta}) \neq 0, \ \boldsymbol{\zeta} \in \mathbb{C}^l\} \geq \widehat{N}.$$

Since $V \supseteq \mathrm{Range}(\mathbf{r})$, it follows that

$$\dim \mathrm{span}\, V \geq \dim \mathrm{span}\, \mathrm{Range}(\mathbf{r}) \geq \widehat{N}. \tag{33}$$

Thus, by assumption 6 in Theorem 1 and (32)–(33),

$$R \leq \widehat{N} - \widehat{l} \leq \dim \mathrm{span}\, V - \dim V.$$

Thus, we have shown that V satisfies the assumptions in Lemma 8. Let now the set G_V and the polynomial $h(\mathbf{v}_1, \ldots, \mathbf{v}_R)$ be as in Lemma 8. Since V is the Zariski closure of $\mathrm{Range}(\mathbf{r})$, it follows that $V \times \cdots \times V$ is the Zariski closure of $\mathrm{Range}(\mathbf{r}) \times \cdots \times \mathrm{Range}(\mathbf{r})$. Since, by Lemma 8, the zero set of $h(\mathbf{v}_1, \ldots, \mathbf{v}_R)$ does not contain $V \times \cdots \times V$, it follows that the zero set of $h(\mathbf{v}_1, \ldots, \mathbf{v}_R)$ does not contain $\mathrm{Range}(\mathbf{r}) \times \cdots \times \mathrm{Range}(\mathbf{r})$. Hence, there exist $\boldsymbol{\zeta}_1^0, \ldots, \boldsymbol{\zeta}_R^0 \in \mathbb{C}^l$ such that $h(\mathbf{r}(\boldsymbol{\zeta}_1^0), \ldots, \mathbf{r}(\boldsymbol{\zeta}_R^0)) \neq 0$. On the other hand, since $\mathrm{Range}(\mathbf{r})$ is a subset of V, from the definitions of $G_{\mathbf{r}}$ and G_V it follows that

$$(\mathbf{r}(\boldsymbol{\zeta}_1), \ldots, \mathbf{r}(\boldsymbol{\zeta}_R)) \in G_V \quad \text{for all} \quad [\boldsymbol{\zeta}_1^T \ \ldots \ \boldsymbol{\zeta}_R^T]^T \in G_{\mathbf{r}}.$$

Hence, by Lemma 8,

$$h(\mathbf{r}(\boldsymbol{\zeta}_1), \ldots, \mathbf{r}(\boldsymbol{\zeta}_R)) = 0 \quad \text{for all} \quad [\boldsymbol{\zeta}_1^T \ \ldots \ \boldsymbol{\zeta}_R^T]^T \in G_{\mathbf{r}}. \tag{34}$$

Since the function $h(\mathbf{r}(\boldsymbol{\zeta}_1), \ldots, \mathbf{r}(\boldsymbol{\zeta}_R))$ is a composition of the polynomial h in RN variables and RN rational functions $\frac{p_1(\boldsymbol{\zeta}_1)}{q_1(\boldsymbol{\zeta}_1)}, \ldots, \frac{p_N(\boldsymbol{\zeta}_R)}{q_N(\boldsymbol{\zeta}_R)}$, it follows that $h(\mathbf{r}(\boldsymbol{\zeta}_1), \ldots, \mathbf{r}(\boldsymbol{\zeta}_R))$ can be written as a ratio of two polynomials in the entries of $\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_R$,

$$h(\mathbf{r}(\boldsymbol{\zeta}_1), \ldots, \mathbf{r}(\boldsymbol{\zeta}_R)) = \frac{h_{num}(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_R)}{h_{den}(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_R)}. \tag{35}$$

By (34)–(35), h_{num} vanishes on $G_{\mathbf{r}}$ and is not identically zero. That is, (31) holds for $g = h_{num}$.

Now we focus on $W_{\mathbf{r}}$. By assumption 6 in Theorem 1, $\widehat{N} \geq R + \widehat{l}$, so assumption 4 implies that there exist $\boldsymbol{\zeta}_1^0, \ldots, \boldsymbol{\zeta}_R^0 \in \mathbb{C}^l$ such that the vectors $\mathbf{r}(\boldsymbol{\zeta}_1^0), \ldots, \mathbf{r}(\boldsymbol{\zeta}_R^0)$ are linearly independent. Hence there exists an $R \times R$ submatrix $\mathbf{R}_{sub}(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_R)$ of $[\mathbf{r}(\boldsymbol{\zeta}_1) \ \ldots \ \mathbf{r}(\boldsymbol{\zeta}_R)]$ whose determinant $d(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_R)$ is not zero at the point $(\boldsymbol{\zeta}_1^0, \ldots, \boldsymbol{\zeta}_R^0)$. On the other hand, $d(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_R)$ vanishes on $W_{\mathbf{r}}$ by definition. Since a determinant is a multivariate polynomial, and since the entries of $\mathbf{R}_{sub}(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_R)$ are rational functions of $\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_R$, $d(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_R)$ can be written as a ratio of two polynomials d_{num} and d_{den} in the entries of $\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_R$. It is clear that $d_{num}(\boldsymbol{\zeta}_1^0, \ldots, \boldsymbol{\zeta}_R^0) \neq 0$ and that d_{num} vanishes on $W_{\mathbf{r}}$. That is, (30) holds for $w = d_{num}$.

2) *Case $\mathbb{F} = \mathbb{C}$ and $\mathbf{f}(\boldsymbol{\zeta})$ is arbitrary.* We restrict ourselves to the case $G_{\mathbf{b}}$. Namely, we use the polynomial h_{num} and the function $\mathbf{f}$ to construct an analytic function $g = u_{num}$ in Rl variables that satisfies (31). The function w that satisfies (30) can be constructed in the same way as g but from the polynomial d_{num} and the function $\mathbf{f}$.

First we prove the existence and analyticity of g. From the definitions of $G_{\mathbf{b}}$ and $G_{\mathbf{r}}$ it follows that if $(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_R) \in G_{\mathbf{b}}$, then $(\mathbf{f}(\boldsymbol{\zeta}_1), \ldots, \mathbf{f}(\boldsymbol{\zeta}_R)) \in G_{\mathbf{r}}$. Hence, by case 1 above and assumption 2 in Theorem 1, the set $G_{\mathbf{b}}$ is contained in the zero set of the function

$$u(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_R) = h_{num}\left(\mathbf{f}(\boldsymbol{\zeta}_1), \ldots, \mathbf{f}(\boldsymbol{\zeta}_R)\right) = h_{num}\left(\frac{f_{1,num}(\boldsymbol{\zeta}_1)}{f_{1,den}(\boldsymbol{\zeta}_1)}, \ldots, \frac{f_{l,num}(\boldsymbol{\zeta}_R)}{f_{l,den}(\boldsymbol{\zeta}_R)}\right). \tag{36}$$

Since h_{num} is a polynomial, the function u can be represented as a ratio $u = u_{num}/u_{den}$, where the functions u_{num} and u_{den} are defined on the whole space $\mathbb{C}^{Rl}$. Since both u_{num} and u_{den} consist of the composition of some polynomials and $2Rl$ functions $f_{1,num}(\boldsymbol{\zeta}_1), f_{1,den}(\boldsymbol{\zeta}_1), \ldots, f_{l,num}(\boldsymbol{\zeta}_R), f_{l,den}(\boldsymbol{\zeta}_R)$ which are analytic on $\mathbb{C}^{RL}$, it follows that u_{num} and u_{den} are analytic on $\mathbb{C}^{RL}$ [27, p. 6]. We set $g = u_{num}$. It is clear that g vanishes on $G_{\mathbf{b}}$.

Now we prove that g is not identically zero. Since $u = g/u_{den}$, it is sufficient to show that u is not zero at some point. Let $\boldsymbol{\zeta}^0$ be a point as in assumption 3 in Theorem 1. Then, by the inverse function theorem, there exists a neighborhood $\mathcal{N}(\boldsymbol{\zeta}^0, \varepsilon) \subset \mathbb{F}^l$ of the point $\boldsymbol{\zeta}^0$ such that for any $\mathbf{p} \in \mathcal{N}(\boldsymbol{\zeta}^0, \varepsilon)$ the equation $\mathbf{f}(\boldsymbol{\zeta}) = \mathbf{p}$ has the solution $\boldsymbol{\zeta} = \mathbf{f}^{-1}(\mathbf{p})$. Hence the equation $(\mathbf{f}(\boldsymbol{\zeta}_1), \ldots, \mathbf{f}(\boldsymbol{\zeta}_R)) = (\mathbf{p}_1, \ldots, \mathbf{p}_R)$ has the solution $(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_R) = (\mathbf{f}^{-1}(\mathbf{p}_1), \ldots, \mathbf{f}^{-1}(\mathbf{p}_R))$ for all $(\mathbf{p}_1, \ldots, \mathbf{p}_R) \in \mathcal{N}(\boldsymbol{\zeta}^0, \varepsilon) \times \cdots \times \mathcal{N}(\boldsymbol{\zeta}^0, \varepsilon)$. Since $\mu_{Rl}(\mathcal{N}(\boldsymbol{\zeta}^0, \varepsilon) \times \cdots \times \mathcal{N}(\boldsymbol{\zeta}^0, \varepsilon)) = \mu_l(\mathcal{N}(\boldsymbol{\zeta}^0, \varepsilon))^R > 0$ [28, Theorem B, p.144] and, by step 1), h_{num} is not identically zero, there exists a point $(\mathbf{p}_1^0, \ldots, \mathbf{p}_R^0) \in \mathcal{N}(\boldsymbol{\zeta}^0, \varepsilon) \times \cdots \times \mathcal{N}(\boldsymbol{\zeta}^0, \varepsilon)$ such that $h_{num}(\mathbf{p}_1^0, \ldots, \mathbf{p}_R^0) \neq 0$. Hence, by (36), $u(\mathbf{f}^{-1}(\mathbf{p}_1^0), \ldots, \mathbf{f}^{-1}(\mathbf{p}_R^0)) = h_{num}(\mathbf{p}_1^0, \ldots, \mathbf{p}_R^0) \neq 0$. That is, (31) holds for $g = u_{num}$.

3) *Case $\mathbb{F} = \mathbb{R}$.* To distinguish between the complex and the real case we denote $G_{\mathbf{b}}$ and $W_{\mathbf{b}}$ in case 3 by $G_{\mathbf{b},\mathbb{R}}$ and $W_{\mathbf{b},\mathbb{R}}$, respectively. Similarly, the sets $G_{\mathbf{b}}$ and $W_{\mathbf{b}}$ considered in case 2, i.e. for $\mathbb{F} = \mathbb{C}$, are denoted by $G_{\mathbf{b},\mathbb{C}}$ and $W_{\mathbf{b},\mathbb{C}}$, respectively. Let $g_{\mathbb{C}} = u_{num}$ and $w_{\mathbb{C}} = w$ denote the analytic functions constructed in case 2. Then $g_{\mathbb{C}}$ and $w_{\mathbb{C}}$ are not identically zero and $g_{\mathbb{C}}$ vanishes on $G_{\mathbf{b},\mathbb{C}}$ and $w_{\mathbb{C}}$ vanishes on $W_{\mathbf{b},\mathbb{C}}$.

Since $W_{\mathbf{b},\mathbb{R}}$ is a subset of $W_{\mathbf{b},\mathbb{C}}$, it follows that $w_{\mathbb{C}}$ vanishes on $W_{\mathbf{b},\mathbb{R}}$. Thus, (30) holds for $w = w_{\mathbb{C}}$.

It has not been proven that set $G_{\mathbf{b},\mathbb{R}}$ is a subset of $G_{\mathbf{b},\mathbb{C}}$ but in any case $G_{\mathbf{b},\mathbb{R}} = (G_{\mathbf{b},\mathbb{R}} \cap G_{\mathbf{b},\mathbb{C}}) \cup (G_{\mathbf{b},\mathbb{R}} \setminus G_{\mathbf{b},\mathbb{C}})$. We show that (31) holds for $g = g_{\mathbb{C}} \cdot w_{\mathbb{C}}$. Indeed, by case 2, $g_{\mathbb{C}}$ and hence $g = g_{\mathbb{C}} w_{\mathbb{C}}$ vanish on $G_{\mathbf{b},\mathbb{C}} \supseteq G_{\mathbf{b},\mathbb{R}} \cap G_{\mathbf{b},\mathbb{C}}$. On the other hand, if $[\boldsymbol{\zeta}_1^T \ \ldots \ \boldsymbol{\zeta}_R^T]^T \in G_{\mathbf{b},\mathbb{R}} \setminus G_{\mathbf{b},\mathbb{C}}$, then there exist $\lambda_1, \ldots, \lambda_R \in \mathbb{R}$, $\lambda \in \mathbb{C} \setminus \mathbb{R}$ and $r \in \{1, \ldots, R\}$ such that $\lambda_1 \mathbf{b}(\boldsymbol{\zeta}_1) + \cdots + \lambda_R \mathbf{b}(\boldsymbol{\zeta}_R) = \lambda \mathbf{b}(\boldsymbol{\zeta}_r)$, yielding that $[\boldsymbol{\zeta}_1^T \ \ldots \ \boldsymbol{\zeta}_R^T]^T \in W_{\mathbf{b},\mathbb{C}}$. Thus, $G_{\mathbf{b},\mathbb{R}} \setminus G_{\mathbf{b},\mathbb{C}} \subseteq W_{\mathbf{b},\mathbb{C}}$ and $w_{\mathbb{C}}$ vanishes on $G_{\mathbf{b},\mathbb{R}} \setminus G_{\mathbf{b},\mathbb{C}}$ as well. That is, (31) holds for $g = g_{\mathbb{C}} w_{\mathbb{C}}$. □

C. Proof of Theorem 1

We show that for a generic $\mathbf{z} \in \Omega$ that satisfies the conditions in Theorem 1, conditions 1–2 in Theorem 7 are also satisfied for

$$\mathbf{A} = \mathbf{A}(\mathbf{z}), \quad S = \{\mathbf{b}(\boldsymbol{\zeta}) : \ q_1(\mathbf{f}(\boldsymbol{\zeta})) \cdots q_N(\mathbf{f}(\boldsymbol{\zeta})) \neq 0\}, \text{ and}$$
$$\mathbf{b}_1 = \mathbf{b}(\boldsymbol{\zeta}_1), \ldots, \mathbf{b}_R = \mathbf{b}(\boldsymbol{\zeta}_R),$$

where, by our notational convention from Subsection I-C, the vectors $\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_R$ are such that $[\boldsymbol{\zeta}_1^T \ \ldots \ \boldsymbol{\zeta}_R^T]^T$ coincides with the last $s = Rl$ entries of $\mathbf{z} \in \Omega$. The generic uniqueness that we want to prove in Theorem 1, then follows from Theorem 7. We have the following.

1) Condition 1 of Theorem 7 holds for generic $\mathbf{z} \in \Omega$ by assumption 1 in Theorem 1.

2) By Lemma 9, condition 2 of Theorem 7 holds for generic $\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_R \in \mathbb{F}^l$, or equivalently, for generic $[\boldsymbol{\zeta}_1^T \ \ldots \ \boldsymbol{\zeta}_R^T]^T \in \mathbb{F}^s$. Hence, condition 2 of Theorem 7 holds for generic $\mathbf{z} \in \Omega$. (Indeed, if $\widetilde{\Omega}$ denotes a set of points $\mathbf{z} \in \Omega$ such that condition 2 of Theorem 7 does not hold and π_s denotes the projection onto the last s coordinates of $\mathbb{F}^n$, then, by Lemma 9, $\mu_s\{\pi_s(\widetilde{\Omega})\} = 0$, which implies [28, Theorem B, p.144] that $\mu_n\{\widetilde{\Omega}\} = 0$.)

APPENDIX B
PROOF OF LEMMA 4

We use the fact that $\cos n\zeta$ and $\frac{\sin n\zeta}{\sin \zeta}$ are polynomials in $\cos \zeta$ [29, p. 642]:

$$\cos n\zeta = \sum_{k=0}^{\lfloor n/2 \rfloor} C_n^{2k} (\cos^2 \zeta - 1)^k \cos^{n-2k} \zeta = P_n(\cos \zeta),$$

$$\frac{\sin n\zeta}{\sin \zeta} = \sum_{k=0}^{\lfloor (n-1)/2 \rfloor} C_n^{2k+1} (\cos^2 \zeta - 1)^k \cos^{n-2k-1} \zeta,$$

where $\lfloor x \rfloor$ denotes the integer part of x. Substituting (14) into these equations we obtain that there exist rational functions R_n and Q_n such that (15)–(16) hold.

ACKNOWLEDGMENT

The authors wish to thank Giorgio Ottaviani and Ed Dewey for their assistance in algebraic geometry.

REFERENCES

[1] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Rev.*, vol. 51, no. 1, pp. 34–81, 2009.
[2] A.-J. van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Trans. Signal Process.*, vol. 44, no. 5, pp. 1136–1155, May 1996.
[3] R. Roy and T. Kailath, "ESPRIT—Estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, 1989.
[4] R. Hermann and C. F. Martin, "Applications of algebraic geometry to systems theory–Part I," *IEEE Trans. Autom. Control*, vol. 22, no. 1, pp. 19–25, 1977.
[5] C. Martin and R. Hermann, "Applications of algebraic geometry to systems theory, part ii: Feedback and pole placement for linear hamiltonian systems," *Proc. IEEE*, vol. 65, no. 6, pp. 841–848, June 1977.
[6] A. Cichocki, D. Mandic, C. Caiafa, A.-H. Phan, G. Zhou, Q. Zhao, and L. De Lathauwer, "Tensor decompositions for signal processing applications. From two-way to multiway component analysis," *IEEE Signal Process. Mag.*, vol. 32, pp. 145–163, March 2015.
[7] P. Comon, "Tensors : A brief introduction," *IEEE Signal Process. Mag.*, vol. 31, no. 3, pp. 44–53, 2014.
[8] I. Domanov and L. De Lathauwer, "Generic uniqueness conditions for the canonical polyadic decomposition and INDSCAL," *SIAM J. Matrix Anal. Appl.*, vol. 36, no. 4, pp. 1567–1589, 2015.
[9] S. Krantz, *Function Theory of Several Complex Variables.Reprint of the 1992 edition*. AMS Chelsea Publishing, Providence, RI, 2001.
[10] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 434–444, Feb 1997.
[11] L. De Lathauwer and J. Castaing, "Blind identification of underdetermined mixtures by simultaneous matrix diagonalization," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1096–1105, March 2008.
[12] D.-T. Pham and J.-F. Cardoso, "Blind separation of instantaneous mixtures of nonstationary sources," *IEEE Trans. Signal Process.*, vol. 49, no. 9, pp. 1837–1848, Sep 2001.
[13] A. Yeredor, "Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation," *IEEE Trans. Signal Process.*, vol. 50, no. 7, pp. 1545–1553, Jul 2002.
[14] ——, "Blind source separation via the second characteristic function," *Signal Processing*, vol. 80, no. 5, pp. 897–902, 2000.
[15] P. Comon and C. Jutten, Eds., *Handbook of Blind Source Separation, Independent Component Analysis and Applications*. Academic Press, Oxford, UK, 2010.
[16] A. Stegeman, J. Ten Berge, and L. De Lathauwer, "Sufficient conditions for uniqueness in CANDECOMP/PARAFAC and INDSCAL with random component matrices," *Psychometrika*, vol. 71, no. 2, pp. 219–229, June 2006.
[17] M. Sørensen and L. De Lathauwer, "New Uniqueness Conditions for the Canonical Polyadic Decomposition of Third-Order Tensors," *SIAM J. Matrix Anal. Appl.*, vol. 36, no. 4, pp. 1381–1403, 2015.
[18] L. De Lathauwer, "Blind Separation of Exponential Polynomials and the Decomposition of a Tensor in Rank-$(L_r, L_r, 1)$ Terms," *SIAM J. Matrix Anal. Appl.*, vol. 32, no. 4, pp. 1451–1474, 2011.
[19] O. Debals, M. Van Barel, and L. De Lathauwer, "Löwner-based blind signal separation of rational functions with applications," *IEEE Trans. Signal Process. (accepted)*.
[20] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, Cambridge, 2013.
[21] T. Jiang and N. D. Sidiropoulos, "Kruskal's permutation lemma and the identification of Candecomp/Parafac and bilinear models with constant modulus constraints," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2625–2636, September 2004.
[22] A. Stegeman, "On uniqueness of the canonical tensor decomposition with some form of symmetry," *SIAM J. Matrix Anal. Appl.*, vol. 32, no. 2, pp. 561–583, 2011.
[23] ——, "On uniqueness of the nth order tensor decomposition into rank-1 terms with linear independence in one mode," *SIAM J. Matrix Anal. Appl.*, vol. 31, no. 5, pp. 2498–2516, 2010.
[24] H. Flenner, L. O'Carroll, and W. Vogel, *Joins and intersections*, ser. Springer Monogr. Math. Berlin: Springer, 1999.
[25] L. Chiantini and C. Ciliberto, "Weakly defective varieties," *Trans. Amer. Math. Soc.*, vol. 354, pp. 151–178, 2002.
[26] M. Beltrametti, E. Carletti, D. Gallarati, and G. Monti Bragadin, *Lectures on Curves, Surfaces and Projective Varieties. A Classical View of Algebraic Geometry*. EMS Textbooks in Mathematics. European Mathematical Society (EMS), Zürich, 2009.
[27] R. Gunning and H. Rossi, *Analytic Functions of Several Complex Variables*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1965.
[28] P. Halmos, *Measure Theory*. Graduate Texts in Mathematics 18, Springer, 1950.
[29] A. P. Prudnikov, Y. A. Brychkov, O. I. Marichev, and N. Queen, *Integrals and series. Vol. 1. Elementary functions*. CRC, 1998.



Ignat Domanov Ignat Domanov received the Masters degree from Donetsk State University, Ukraine, the Ph.D degree in Physics and Mathematics from Institute of Applied Mathematics and Mechanics, Ukraine, and the Ph.D. degree in Engineering from KU Leuven, Belgium, in 1998, 2004, and 2013, respectively. Since 2013 he has been a Postdoctoral Fellow with the KU Leuven, Belgium. His research interests include applied linear algebra, tensor decompositions and tensor-based signal processing.



Lieven De Lathauwer Lieven De Lathauwer received the Masters degree in electromechanical engineering and the Ph.D. degree in applied sciences from KU Leuven, Belgium, in 1992 and 1997, respectively. He is currently Professor with KU Leuven, Belgium. Dr. De Lathauwer is an Associate Editor of the SIAM Journal on Matrix Analysis and Applications and has been an Associate Editor for IEEE Transactions on Signal Processing. His research concerns the development of tensor tools for engineering applications.