KU LEUVEN

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Citation/Reference	Debals, O., De Lathauwer, L. (2015), **Löwner-based Blind Signal Separation of Rational Functions with Applications** IEEE Transactions on Signal Processing, vol. 64, no. 8, p. 1909-1918.
Archived version	Author manuscript: the content is identical to the content of the published paper, but without the final typesetting by the publisher
Published version	http://dx.doi.org/10.1109/TSP.2015.2500179
Journal homepage	http://www.signalprocessingsociety.org/publications/periodicals/tsp/
Author contact	Otto.Debals@esat.kuleuven.be *+ 32 (0)16 3 20364*
IR	https://lirias.kuleuven.be/handle/123456789/505541

(article begins on next page)

Löwner-based Blind Signal Separation of Rational Functions with Applications

Otto Debals, *Student Member, IEEE*, Marc Van Barel, *Member, IEEE*, and Lieven De Lathauwer, *Fellow, IEEE*

***Abstract*—A new blind signal separation (BSS) technique is proposed, enabling a deterministic separation of signals into rational functions. Rational functions can take on a wide range of forms, such as the well-known pole-like shape. The approach is a possible alternative for the well-known independent component analysis when the theoretical sources are not independent, such as for frequency spectra, or when only a small number of samples is available. The technique uses a low-rank decomposition on the tensorized version of the observed data matrix. The deterministic tensorization with Löwner matrices is comprehensively analyzed in this paper. Uniqueness properties are investigated, and a connection with the separation into exponential polynomials is made. Finally, the technique is illustrated for fetal electrocardiogram extraction and with an application in the domain of fluorescence spectroscopy, enabling the identification of chemical analytes using only a single excitation-emission matrix.**

***Index Terms*—Blind signal separation (BSS), block term decomposition, independent component analysis, Löwner matrix, rational functions, tensors.**

I. Introduction

THIS paper deals with the separation of linear mixtures of different source signals, known as blind signal separation (BSS). The general solution to this problem is not unique and various approaches have been proposed, ranging from applying independence assumptions to non-negativity and sparsity constraints [1]. The former has received the name independent component analysis (ICA) in which one assumes the sources to be statistically independent [2], [3]. ICA has already been applied in numerous applications in for example image processing, finance, telecommunications and biomedical sciences [3]–[6]. It is a stochastic technique, tensorizing the observed matrix data using higher-order statistics [7]. Many applications do not involve stochastic and independent signals however, and one can also tensorize the data deterministically. In a source-related deterministic setting, as in this paper, the separation of the observed signals is based on a specific source model.

A first possible source model could be the class of polynomials, as they have a simple form and well-understood properties. However, polynomial signal models often require a large number of coefficients. Second and more essential, polynomial models generally do not allow a unique separation, as will be illustrated in this paper. With the family of exponential polynomials (sums and/or products of exponentials, sinusoids and/or polynomials), a source model is proposed in [8] which is applicable for blind signal separation by using a deterministic Hankel tensorization.

This paper proposes the class of rational functions, contributing to a general framework of deterministic blind signal separation. They encompass the class of polynomials, and are able to model complicated structures with a fairly low degree in both the numerator and denominator; much like autoregressive–moving-average models are more powerful in the field of system identification than pure moving-average models [9]. Rational functions are mainly known because of their pole-like behavior (suitable when modeling frequency spectra, time-of-flight data, distribution functions, etc.) but can take on an extremely wide range of shapes in the complex domain; also smooth curves and signals with both low- and high-varying regions can be modeled. Furthermore, by considering uniqueness properties, we show in this paper that rational functions are suitable in a BSS context. It is also illustrated that the sampling points need not be equidistant for the proposed technique, contrary to the Hankel technique from [8].

Another basis for BSS is sparse separation in which one uses a fixed signal dictionary with the weights optimized according to some sparse objectives [10]–[12]. The proposed technique in this paper goes a step further as there is no need for an initial signal dictionary: the dictionary itself is estimated too.

The assumption of rationality is implemented using the theory and properties of the Löwner matrix. This kind of matrix is well-known in the domain of system identification regarding rational interpolation [13]–[15], but is not acknowledged in other application domains; its definition is given in Section II. The observed data matrix is tensorized using Löwner matrices, and the obtained tensor is analyzed using a block term tensor decomposition [16]–[19]. Block component analysis (BCA) describes the use of block term decompositions to identify underlying components [20].

It is shown in this paper that the solution with the assump-

Manuscript received March 09, 2015; revised August 31, 2015; accepted October 25, 2015. Date of publication November 12, 2015. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Anthony So. This work was supported by a Ph.D. grant of the Agency for Innovation by Science and Technology (IWT), by the Research Council KU Leuven: C1 project c16/15/059-nD, CoE PFV/10/002 (OPTEC), by the F.W.O.: project G.0830.14N and G.0881.14N, by the Belgian Federal Science Policy Office: IUAP P7/19 (DYSCO II, Dynamical systems, control and optimization, 2012-2017), and by the 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.

O. Debals and L. De Lathauwer are with the Group of Science, Engineering and Technology, KU Leuven Kulak, B-8500 Kortrijk, Belgium and with the Department of Electrical Engineering (ESAT), B-3001 Leuven, Belgium. They are also with iMinds Medical IT, Leuven, Belgium (e-mail: otto.debals@esat.kuleuven.be; lieven.delathauwer@kuleuven-kulak.be).

M. Van Barel is with the Department of Computer Science, KU Leuven, Celestijnenlaan 200A, B-3001 Leuven, Belgium (e-mail: marc.vanbarel@cs.kuleuven.be).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org

Digital Object Identifier 10.1109/TSP.2015.2500179

tion of rational sources is unique under mild conditions. This is important as it explains that the assumption is powerful and natural, while in the case of non-negative constraints, for example, additional sparsity constraints need to be imposed to recover a unique solution [1]. Furthermore, techniques have been developed for ICA to recover more source signals than there are observed signals, i.e., for underdetermined mixtures [21], [22]. Because of the strong uniqueness results, this is readily extended in the separation method described. Simulations are presented in the final section of this paper.

The technique was briefly described in [23]. We present the method with two illustrations using real-life datasets. The first illustration is antepartum fetal heart rate monitoring with the separation of mother and fetal electrocardiogram signals from multilead cutaneous potential recordings. An excellent separation is obtained, even for short sequences with coinciding heart beats. The second illustration covers the detection of chemical components in mixtures, using emission-excitation data from fluorescence spectroscopy. With the technique, only a single sample is sufficient to determine the concentrations and frequency spectra of the different chemical components.

We start by fixing the notation and discussing some basic definitions and decompositions in the field of multilinear algebra. In Section II the problem statement is examined and the Löwner-based technique is introduced. Section III contains a more advanced analysis. Uniqueness properties are considered in Section IV. A connection to the method from [8] is investigated in Section V, and numerical experiments are performed in Section VI.

A. Notation and Basic Operations

Vectors (denoted by a bold, lowercase letter, e.g., $\mathbf{a}$) and matrices (denoted by a bold, uppercase letter, e.g., $\mathbf{A}$) can be generalized to higher orders, obtaining tensors. A general Nth order tensor of size $I_1 \times I_2 \times \cdots \times I_N$ is denoted by a calligraphic letter as $\mathcal{A} \in \mathbb{K}^{I_1 \times I_2 \times \cdots \times I_N}$ (with $\mathbb{K}$ we mean $\mathbb{R}$ or $\mathbb{C}$). $\mathcal{A}$ is a multidimensional array with numerical values $a_{i_1 i_2 \cdots i_N} = \mathcal{A}(i_1, i_2, \ldots, i_N)$. The mode-$n$ vectors of $\mathcal{A}$ are constructed by fixing all but one index, e.g., $\mathbf{a} = \mathcal{A}(i_1, \ldots, i_{n-1}, :, i_{n+1}, \ldots, i_N)$.

A number of products can be defined in the domain of tensors. The mode-n tensor–matrix product between a tensor $\mathcal{A} \in \mathbb{K}^{I_1 \times I_2 \times \cdots \times I_N}$ and a matrix $\mathbf{B} \in \mathbb{K}^{J \times I_n}$ is defined as

$$(\mathcal{A} \cdot_n \mathbf{B})_{i_1 \cdots i_{n-1} j i_{n+1} \cdots i_N} = \sum\nolimits_{i_n=1}^{I_n} a_{i_1 i_2 \cdots i_N} b_{j i_n}.$$

The outer product of two tensors $\mathcal{A} \in \mathbb{K}^{I_1 \times I_2 \times \cdots \times I_N}$ and $\mathcal{B} \in \mathbb{K}^{J_1 \times J_2 \times \cdots \times J_M}$ is given as

$$(\mathcal{A} \otimes \mathcal{B})_{i_1 i_2 \cdots i_N j_1 j_2 \cdots j_M} = a_{i_1 i_2 \cdots i_N} b_{j_1 j_2 \cdots j_M}.$$

The matrices $\mathbf{A}^{\mathrm{T}}$ and $\mathbf{A}^{\dagger}$ denote the transpose and Moore-Penrose pseudo-inverse of $\mathbf{A}$, respectively. The Frobenius norm of a tensor is defined as the root of the sum of the squares of the elements: $\|\mathcal{A}\| = (\sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} |a_{i_1 \cdots i_N}|^2)^{1/2}$.

B. Rank Definitions and Basic Tensor Decompositions

The column (row) rank of a matrix $\mathbf{A}$ is the maximum number of linearly independent columns (rows) of $\mathbf{A}$. Note that the column rank is equal to the row rank for a matrix. We make a distinction between the rank and the multilinear rank.

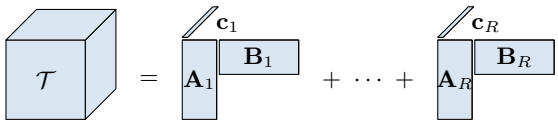


Fig. 1. Decomposition of a tensor $\mathcal{T}$ in rank-$(L_r, L_r, 1)$ terms.

First, a tensor $\mathcal{T}$ has rank 1 if it can be written as the outer product of some nonzero vectors: $\mathcal{T} = \mathbf{a}^{(1)} \otimes \mathbf{a}^{(2)} \otimes \ldots \otimes \mathbf{a}^{(N)}$. If one writes a tensor $\mathcal{T}$ as a linear combination of R rank-1 tensors, one obtains a Polyadic Decomposition (PD):

$$\begin{aligned}\mathcal{T} &= \sum\nolimits_{r=1}^{R} \mathbf{a}_r^{(1)} \otimes \mathbf{a}_r^{(2)} \otimes \cdots \otimes \mathbf{a}_r^{(N)} \\ &\triangleq \left[\!\!\left[\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \ldots, \mathbf{A}^{(N)} \right]\!\!\right].\end{aligned}$$

If this R is minimal, the rank of $\mathcal{T}$ is defined as R, denoted by $r(\mathcal{T})$. The decomposition then becomes canonical (CPD).

Second, the mode-n rank of a tensor $\mathcal{T}$ is the dimension of the subspace spanned by its mode-n vectors. It is equal to the rank of the matrix constructed by stacking all the mode-n vectors one after the other. If the mode-1 rank, mode-2 rank and mode-3 rank of a third-order tensor are equal to L, M and N, respectively, it is said to have trilinear rank (L, M, N). This becomes the multilinear rank when generalized to arbitrary order, obtaining the N-tuple $(R_1, R_2, \ldots, R_N)$. Connected to this multilinear rank is the Tucker Decomposition:

$$\mathcal{T} = \mathcal{G} \cdot_1 \mathbf{A}^{(1)} \cdot_2 \mathbf{A}^{(2)} \cdots_N \mathbf{A}^{(N)} \triangleq \left[\!\!\left[\mathcal{G}; \mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \ldots, \mathbf{A}^{(N)} \right]\!\!\right]$$

with $\mathcal{G} \in \mathbb{K}^{R_1 \times R_2 \times \ldots \times R_N}$ being a (typically small) core tensor. Related are the multilinear singular value decomposition and the low multilinear rank approximation; for details we refer to [24]–[27].

In this paper we make use of the block term decomposition (BTD), which starts from the idea of linearly combining tensors of low multilinear rank [16], [17]. For third-order tensors, one obtains the BTD in R rank-(L_r, M_r, N_r) terms. A special instance is the decomposition of a tensor $\mathcal{T} \in \mathbb{K}^{I_1 \times I_2 \times I_3}$ in rank-$(L_r, L_r, 1)$ terms (with $M_r = L_r$ and $N_r = 1$, for $1 \leq r \leq R$):

$$\mathcal{T} = \sum\nolimits_{r=1}^{R} \mathbf{E}_r \otimes \mathbf{c}_r, \tag{1}$$

with the matrix $\mathbf{E}_r \in \mathbb{K}^{I_1 \times I_2}$ having rank L_r and vector $\mathbf{c}_r \in \mathbb{K}^{I_3}$ being nonzero. Each matrix $\mathbf{E}_r$ can be factorized to give

$$\mathcal{T} = \sum\nolimits_{r=1}^{R} (\mathbf{A}_r \mathbf{B}_r^{\mathrm{T}}) \otimes \mathbf{c}_r, \tag{2}$$

with $\mathbf{A}_r \in \mathbb{K}^{I_1 \times L_r}$, $\mathbf{B}_r \in \mathbb{K}^{I_2 \times L_r}$. It is visualized in Fig. 1.

II. Löwner-Based Blind Signal Separation

In this section, the technique for BSS with Löwner matrices is discussed and the use of the previously defined BTD in rank-$(L_r, L_r, 1)$ terms is explained. Section II-A discusses the model setup and Section II-B introduces Löwner matrices. Section II-C explains the tensorization technique. Finally, two different ways to recover the original sources are discussed.

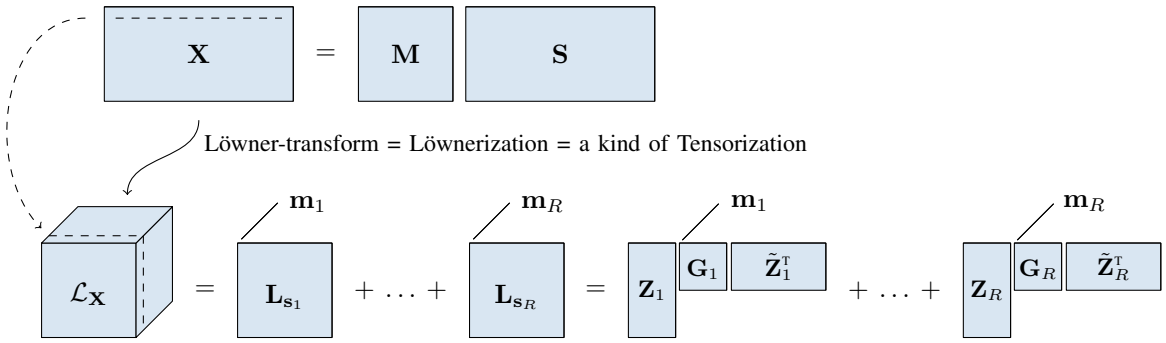


Fig. 2. The observed data matrix is tensorized to stacked Löwner matrices, which are decomposed with a Block Term Decomposition in rank-$(L_r, L_r, 1)$ terms. The mixing vectors $\mathbf{m}_1, \ldots, \mathbf{m}_R$ appear as factor vectors in the third mode, and the Löwner matrices of the sources appear in the first and second mode. The factor matrices have an interesting structure, explained in Section III.

A. The Blind Signal Separation problem

Assume we have R source signals being linearly mixed into K observed signals. For each signal N samples are available. We consider the following data model in BSS:

$$\mathbf{X} = \mathbf{MS} + \mathbf{N},$$

with $\mathbf{X} \in \mathbb{K}^{K\times N}$ containing the observed data, $\mathbf{S} \in \mathbb{K}^{R\times N}$ the R unknown source signals, $\mathbf{M} \in \mathbb{K}^{K\times R}$ the unknown mixing matrix and $\mathbf{N} \in \mathbb{K}^{K\times N}$ representing additive noise. The general goal in BSS is to recover the unknown sources in $\mathbf{S}$ and the unknown mixing vectors in $\mathbf{M}$, given only the observed data $\mathbf{X}$. We investigate the behavior related to added Gaussian noise in the experiments of Section VI but omit $\mathbf{N}$ in the next analyses for convenience.

Broadly speaking, we will map each observed signal (each row in $\mathbf{X}$) to a Löwner matrix. By stacking these Löwner matrices, one obtains a tensor which is of low multilinear rank because of the working hypothesis, i.e., the source signals can be modeled by rational functions of low degrees. This hypothesis is satisfied in many applications. By decomposing the tensorized data, one can immediately identify the mixing vectors. The reconstruction of the sources follows. Fig. 2 gives a comprehensive overview of the technique.

B. Löwner Matrices

We first define the Löwner matrix for a function sampled in a point set T consisting of N distinct points:

Definition 1. *Given a function $f(t)$ sampled on points $T = \{t_1, t_2, \ldots, t_N\}$. We partition the point set T into two distinct point sets $X = \{x_1, x_2, \ldots, x_I\}$ and $Y = \{y_1, y_2, \ldots, y_J\}$ with $I+J = N$, and define the elements of the Löwner matrix $\mathbf{L} \in \mathbb{K}^{I\times J}$ as*

$$L_{i,j} = \frac{f(x_i) - f(y_j)}{x_i - y_j} \qquad \forall i, j.$$

We thus obtain the following matrix:

$$\mathbf{L} = \begin{bmatrix} \frac{f(x_1)-f(y_1)}{x_1-y_1} & \frac{f(x_1)-f(y_2)}{x_1-y_2} & \cdots & \frac{f(x_1)-f(y_J)}{x_1-y_J} \\ \frac{f(x_2)-f(y_1)}{x_2-y_1} & \frac{f(x_2)-f(y_2)}{x_2-y_2} & \cdots & \frac{f(x_2)-f(y_J)}{x_2-y_J} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{f(x_I)-f(y_1)}{x_I-y_1} & \frac{f(x_I)-f(y_2)}{x_I-y_2} & \cdots & \frac{f(x_I)-f(y_J)}{x_I-y_J} \end{bmatrix}.$$

In the literature, a parameter α is often used with $I = \alpha$ and $J = N - \alpha$. Matrix $\mathbf{L}$ is square when N is even and $\alpha = N/2$. Unless denoted otherwise, we assume $I = \alpha = \lceil N/2 \rceil$.

A Löwner matrix can also be constructed for point sets with coinciding sample points, for which we refer to other literature [28], [29]. Notice that the Löwner matrix corresponding to a constant function becomes the zero matrix.

The Löwner matrix has interesting properties in connection to rational functions. Let the degree of an irreducible rational function be defined as the maximum of the degrees of the polynomial in its numerator and the polynomial in its denominator. The following has been proven in [13], [30], [31]:

Theorem 1. *Given a Löwner matrix $\mathbf{L}$ of size $I \times J$ constructed from a function $f(t)$ sampled in a point set $T = \{t_1, \ldots, t_N\}$ with $N = I + J$. If $f(t)$ is a rational function of degree δ and if $I, J \geq \delta$, then $\mathbf{L}$ has rank δ:*

$$\operatorname{rank}(\mathbf{L}) = \delta = \deg(f).$$

This theorem is easy to verify for simple rational functions of low degree. For example, $f(t) = \frac{c}{t-p}$ gives $L_{i,j} = -c \cdot \frac{1}{x_i - p} \cdot \frac{1}{y_j - p}$. The corresponding matrix $\mathbf{L}$ is of rank 1, as it can be written as an outer product of two vectors. The next section gives an insight into the structure of the Löwner matrix for a more general rational function.

Note that Theorem 1 is valid for any point set partitioning: two straightforward partitionings are the interleaved partitioning with $X = \{t_1, t_3, \ldots\}$ and $Y = \{t_2, t_4, \ldots\}$ and the block partitioning with $X = \{t_1, \ldots, t_I\}$ and $Y = \{t_{I+1}, \ldots, t_N\}$.

C. Tensorization and Block Term Decomposition

Consider again the BSS model with $\mathbf{X} = \mathbf{MS}$. Let us map each row of $\mathbf{X}$ to a Löwner matrix of size $I \times J$, and stack these matrices in the third dimension obtaining a tensor $\mathcal{L}_{\mathbf{X}}$ of size $I \times J \times K$. We call this transformation the *Löwnerization* of matrix $\mathbf{X}$. As the BSS transformation is linear, we can write:

$$\mathcal{L}_{\mathbf{X}} = \sum\nolimits_{r=1}^{R} \mathbf{L}_{\mathbf{s}_r} \otimes \mathbf{m}_r, \tag{3}$$

with $\mathbf{L}_{\mathbf{s}_r}$ the Löwner matrix and $\mathbf{m}_r$ the mixing vector of source r. If source $\mathbf{s}_r$, i.e., each rth row of $\mathbf{S}$, can be modeled as a rational function of some (low) degree δ_r, the matrix $\mathbf{L}_{\mathbf{s}_r}$ will have a (low) rank δ_r provided there are enough samples

for $I, J \geq \max_r \delta_r$. The matrix $\mathbf{L}_{\mathbf{s}_r}$ admits a factorization with some $\mathbf{A}_r \in \mathbb{K}^{I\times\delta_r}$ and $\mathbf{B}_r \in \mathbb{K}^{J\times\delta_r}$:

$$\mathcal{L}_{\mathbf{X}} = \sum\nolimits_{r=1}^{R} (\mathbf{A}_r\mathbf{B}_r^{\mathrm{T}}) \otimes \mathbf{m}_r, \tag{4}$$

which is precisely the decomposition in rank-$(L_r, L_r, 1)$ terms from (2) with $\delta_r = L_r$. In Section III we look into the factorization of $\mathbf{L}_{\mathbf{s}_r} = \mathbf{A}_r\mathbf{B}_r^{\mathrm{T}} = \mathbf{Z}_r\mathbf{G}_r\tilde{\mathbf{Z}}_r^{\mathrm{T}}$. In Section IV the uniqueness properties are investigated regarding the use of a BTD in rank-$(L_r, L_r, 1)$ terms.

D. Recovery of the Mixing Matrix and the Sources

The columns of the mixing matrix appear as the factor vectors in the third mode of the decomposition. The source signals can be reconstructed in two main ways. First, the estimated matrix $\hat{\mathbf{M}}$ can be inverted to calculate $\hat{\mathbf{S}} = \hat{\mathbf{M}}^{\dagger}\mathbf{X}$. This is the most straightforward method, and the default method for ICA. However, it is not applicable for underdetermined mixtures (with fewer observed signals than sources) or when the mixing matrix does not have full column rank. The columns of $\hat{\mathbf{S}}$ are then determined up to arbitrary vectors in the null space of $\hat{\mathbf{M}}$, as in underdetermined ICA [21], [22].

A more general but more cumbersome technique is to recover the sources from the separated Löwner matrices in (3) and (4). Each recovered matrix $\mathbf{E}_r = \mathbf{A}_r\mathbf{B}_r^{\mathrm{T}}$ contains information about the corresponding source in a finite difference format. A linear system can be constructed to recover the source signal from their corresponding Löwner matrices:

$$\mathbf{s}_r = \arg\min_{\mathbf{s}_r} \frac{1}{2}\left\|\mathbf{F}\mathbf{s}_r - \mathrm{vec}(\hat{\mathbf{L}}_{\mathbf{s}_r})\right\|^2 \quad \text{for } 1 \leq r \leq R,$$

with the matrix $\mathbf{F}$ being the reshaped matrix version of the following tensor $\mathcal{F}$ with size $I \times J \times N$:

$$\forall i,j,n: \ f_{i,j,n} = \begin{cases} \frac{1}{x_i - y_j} & \text{if } n = \phi(i), \\ \frac{-1}{x_i - y_j} & \text{if } n = \theta(j), \\ 0 & \text{elsewhere,} \end{cases}$$

with $\phi : i \to \{n \in \{1,\ldots,N\} : t_n = x_i\}$ and $\theta : j \to \{n \in \{1,\ldots,N\} : t_n = y_j\}$. $\mathbf{F}$ is constructed by vectorizing the third-order slices of $\mathcal{F}$ and stacking them as columns. With the point set $T = \{x_1, y_1, x_2, y_2\}$ for example, we have $\phi(1) = 1$, $\phi(2) = 3$, $\theta(1) = 2$ and $\theta(2) = 4$, and one obtains the following linear system:

$$\begin{bmatrix} (\hat{\mathbf{L}}_r)_{1,1} \\ (\hat{\mathbf{L}}_r)_{2,1} \\ (\hat{\mathbf{L}}_r)_{1,2} \\ (\hat{\mathbf{L}}_r)_{2,2} \end{bmatrix} = \begin{bmatrix} \frac{1}{x_1-y_1} & \frac{-1}{x_1-y_1} & 0 & 0 \\ 0 & \frac{-1}{x_2-y_1} & \frac{1}{x_2-y_1} & 0 \\ \frac{1}{x_1-y_2} & 0 & 0 & \frac{-1}{x_1-y_2} \\ 0 & 0 & \frac{1}{x_2-y_2} & \frac{-1}{x_2-y_2} \end{bmatrix} \begin{bmatrix} s_r(x_1) \\ s_r(y_1) \\ s_r(x_2) \\ s_r(y_2) \end{bmatrix}.$$

Let now μ_r be the DC component of the rth source. Observe that the vector $[\mu_r, \ldots, \mu_r]^{\mathrm{T}} \in \mathbb{K}^N$ is an element of the null space of the matrix $\mathbf{F}$. The vector $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_R]^{\mathrm{T}} \in \mathbb{K}^R$ can be found solving an additional linear system:

$$\boldsymbol{\mu} = \arg\min_{\boldsymbol{\mu}} \frac{1}{2}\left\|\mathbf{X} - \hat{\mathbf{M}}\left(\hat{\mathbf{S}} + \boldsymbol{\mu}\mathbf{e}^{\mathrm{T}}\right)\right\|^2 = \frac{1}{N}\left(\hat{\mathbf{M}}^{\dagger}\mathbf{X} - \hat{\mathbf{S}}\right)\mathbf{e}$$

with $\mathbf{e} = [1, \ldots, 1]^{\mathrm{T}} \in \mathbb{R}^N$. The vector $\boldsymbol{\mu}$ is determined up to a vector in the null space of $\hat{\mathbf{M}}$, generating fewer indeterminacies than the first method.

III. Factorization of Löwner Matrices

There is well-known theory about the factorization of Hankel matrices with the Vandermonde decomposition [8], [32], [33]. Each Hankel matrix $\mathbf{H} \in \mathbb{C}^{I\times J}$ can be written as $\mathbf{V}\mathbf{G}\tilde{\mathbf{V}}^{\mathrm{T}}$ with Vandermonde matrices $\mathbf{V} \in \mathbb{C}^{I\times H}$ and $\tilde{\mathbf{V}} \in \mathbb{C}^{J\times H}$, and a block-diagonal matrix $\mathbf{G} \in \mathbb{C}^{H\times H}$. The rank H of $\mathbf{H}$ is the degree of the underlying exponential polynomial. For Löwner matrices, a general factorization has been developed in [28], [34]. We present the factorization in a way that facilitates the use in signal processing for modeling signals by rational functions of low degree. The factorization relies on partial fractions and Cauchy matrices. A Cauchy matrix $\mathbf{C}_{\mathbf{u},\mathbf{v}} \in \mathbb{K}^{I\times J}$ based on two vectors $\mathbf{u} \in \mathbb{K}^I$, $\mathbf{v} \in \mathbb{K}^J$ with $\forall i,j : u_i \neq v_j$ consists of elements $c_{i,j} = 1/(u_i - v_j)$.

We assume a rational source $s(t)$ with the following partial fraction decomposition:

$$s(t) = a(t) + \sum_{f=1}^{F}\sum_{d=1}^{D_f} \frac{c_{f,d}}{(t-p_f)^d}, \tag{5}$$

meaning that $s(t)$ has F complex poles p_f which can have a multiplicity D_f higher than one. Equation (5) is general in the sense that it covers all rational functions. The first term $a(t)$ in (5) denotes a polynomial of degree W. We define $L = W + \sum_{f=1}^{F} D_f$. As a working assumption, W is zero for most sources; Section IV considers the uniqueness conditions.

In each of the three following subsections we illustrate the factorization of the Löwner matrices corresponding to (5) in a constructive way, first discussing the case of non-coinciding poles and afterwards discussing coinciding poles, i.e., poles with a multiplicity higher than one. In a third subsection polynomials are discussed. Subsection III-D concludes the section by applying the results to blind signal separation.

A. Case of Rational Functions With Non-Coinciding Poles

Consider a rational function with $W = 0$ and with F poles p_f with multiplicities $D_f = 1$ for $1 \leq f \leq F$ (thus $L = F$), collected in a vector $\mathbf{p}$ and with corresponding coefficients c_f:

$$s(t) = a + \sum_{f=1}^{F} \frac{c_f}{t - p_f}.$$

If sampled in a point set T with N distinct points and partitions X and Y, the corresponding Löwner matrix is given by:

$$\begin{aligned} L_{i,j} &= \left(\sum_{f=1}^{F}\frac{c_f}{x_i - p_f} - \sum_{f=1}^{F}\frac{c_f}{y_j - p_f}\right)/(x_i - y_j), \\ &= \sum\nolimits_{f=1}^{F} -c_f \cdot \frac{1}{x_i - p_f}\cdot\frac{1}{y_j - p_f} \qquad \forall i,j. \end{aligned}$$

Hence, the matrix $\mathbf{L}$ can be factorized as

$$\mathbf{L} = \mathbf{Z}\cdot\mathrm{diag}(-c_1, \ldots, -c_F)\cdot\tilde{\mathbf{Z}}^{\mathrm{T}},$$

with

$$\mathbf{Z} = \begin{bmatrix} \frac{1}{x_1-p_1} & \cdots & \frac{1}{x_1-p_F} \\ \vdots & \ddots & \vdots \\ \frac{1}{x_I-p_1} & \cdots & \frac{1}{x_I-p_F} \end{bmatrix}, \ \tilde{\mathbf{Z}} = \begin{bmatrix} \frac{1}{y_1-p_1} & \cdots & \frac{1}{y_1-p_F} \\ \vdots & \ddots & \vdots \\ \frac{1}{y_J-p_1} & \cdots & \frac{1}{y_J-p_F} \end{bmatrix}.$$

One can see that $\mathbf{Z} = \mathbf{C}_{\mathbf{x},\mathbf{p}}$ and $\tilde{\mathbf{Z}} = \mathbf{C}_{\mathbf{y},\mathbf{p}}$.

The assumption of non-coinciding poles with its factorizaton of the Löwner matrix suffices in many cases. We present the general case with coinciding poles in the following subsection for the sake of completeness.

B. General Case of Rational Functions With Coinciding Poles

We first study the Löwner matrix of the term corresponding to $\sum_{d=1}^{D_f} \frac{c_{f,d}}{(t-p_f)^d}$ in (5), i.e., for a pole p_f with multiplicity $D_f \geq 1$. The matrix is given by

$$\begin{aligned} L_{i,j} &= \left(\sum_{d=1}^{D_f} \frac{c_{f,d}}{(x_i - p_f)^d} - \sum_{d=1}^{D_f} \frac{c_{f,d}}{(y_j - p_f)^d} \right) / (x_i - y_j), \\ &= \sum_{d=1}^{D_f} c_{f,d} \frac{(y_j - p_f)^d - (x_i - p_f)^d}{(x_i - y_j)(x_i - p_f)^d (y_j - p_f)^d}, \qquad \forall i, j. \end{aligned}$$

Because $(x_i - y_j) = ((x_i - p_f) - (y_j - p_f))$, one can derive

$$L_{i,j} = \sum_{d=1}^{D_f} \left(-c_{f,d} \sum_{e=1}^{d} \frac{1}{(x_i - p_f)^e (y_j - p_f)^{d-e+1}} \right).$$

The result yields the factorization $\mathbf{L} = \mathbf{Z}_{f,D_f} \mathbf{G}_{f,D_f} \tilde{\mathbf{Z}}_{f,D_f}^{\mathrm{T}}$ with $\mathbf{Z}_{f,D_f}$ and $\tilde{\mathbf{Z}}_{f,D_f}$ being Vandermonde matrices:

$$\mathbf{Z}_{f,D_f} = \begin{bmatrix} \frac{1}{x_1 - p_f} & \cdots & \frac{1}{(x_1 - p_f)^{D_f}} \\ \vdots & \ddots & \vdots \\ \frac{1}{x_I - p_f} & \cdots & \frac{1}{(x_I - p_f)^{D_f}} \end{bmatrix},$$

$$\tilde{\mathbf{Z}}_{f,D_f} = \begin{bmatrix} \frac{1}{y_1 - p_f} & \cdots & \frac{1}{(y_1 - p_f)^{D_f}} \\ \vdots & \ddots & \vdots \\ \frac{1}{y_J - p_f} & \cdots & \frac{1}{(y_J - p_f)^{D_f}} \end{bmatrix}.$$

These matrices are variants of the confluent Cauchy matrices from [34]. We also have

$$\mathbf{G}_{f,D_f} = \begin{bmatrix} -c_{f,1} & -c_{f,2} & -c_{f,3} & \cdots & -c_{f,D_f} \\ -c_{f,2} & -c_{f,3} & -c_{f,4} & \cdots & 0 \\ -c_{f,3} & -c_{f,4} & -c_{f,5} & \cdots & 0 \\ \vdots & \vdots & \vdots & \iddots & \vdots \\ -c_{f,D_f} & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

C. Case of Polynomials

Let $s(t)$ be a polynomial of degree W, i.e., $s(t) = \sum_{w=1}^{W} a_w t^w = a_W t^W + \ldots + a_1 t + a_0$, with $a_w \in \mathbb{K}$, $1 \leq w \leq W$. The Löwner matrix corresponding to a point set T with partitions X and Y is given by:

$$\begin{aligned} L_{i,j} &= \left(\sum_{w=1}^{W} a_w x_i^w - \sum_{w=1}^{W} a_w y_j^w \right) / (x_i - y_j), \\ &= \sum_{w=1}^{W} a_w \sum_{v=0}^{w-1} x_i^v y_j^{w-v-1}, \qquad \forall i, j. \end{aligned}$$

The matrix admits to a factorization $\mathbf{L} = \mathbf{Z}_W \mathbf{G}_W \tilde{\mathbf{Z}}_W^{\mathrm{T}}$ with $\mathbf{Z}_W \in \mathbb{K}^{I \times W}$, $\tilde{\mathbf{Z}}_W \in \mathbb{K}^{J \times W}$:

$$\begin{aligned} \mathbf{Z}_W &= \begin{bmatrix} 1 & x_1 & \cdots & x_1^{W-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_I & \cdots & x_I^{W-1} \end{bmatrix} \\ \tilde{\mathbf{Z}}_W &= \begin{bmatrix} 1 & y_1 & \cdots & y_1^{W-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & y_J & \cdots & y_J^{W-1} \end{bmatrix} \end{aligned} \tag{6}$$

and with

$$\mathbf{G}_W = \begin{bmatrix} a_1 & a_2 & \cdots & a_W \\ a_2 & a_3 & \cdots & 0 \\ \vdots & \vdots & \iddots & \vdots \\ a_W & 0 & \ldots & 0 \end{bmatrix} \in \mathbb{K}^{W \times W}.$$

D. Summary and Implication For Blind Source Separation

Regarding the complete rational signal $s(t)$ from (5), its associated Löwner matrix $\mathbf{L}$ admits to a general decomposition

$$\mathbf{L} = \mathbf{Z} \mathbf{G} \tilde{\mathbf{Z}}^{\mathrm{T}} \tag{7}$$

in which $\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_W & \mathbf{Z}_{1,D_1} & \mathbf{Z}_{2,D_2} & \cdots & \mathbf{Z}_{F,D_F} \end{bmatrix} \in \mathbb{K}^{I \times L}$ and $\tilde{\mathbf{Z}} = \begin{bmatrix} \tilde{\mathbf{Z}}_W & \tilde{\mathbf{Z}}_{1,D_1} & \tilde{\mathbf{Z}}_{2,D_2} & \cdots & \tilde{\mathbf{Z}}_{F,D_F} \end{bmatrix} \in \mathbb{K}^{J \times L}$ with $L = W + \sum_{f=1}^{F} D_f$. The matrix $\mathbf{G} \in \mathbb{K}^{L \times L}$ is a block-diagonal matrix with Hankel and upper antitriangular matrices $\mathbf{G}_W$ and $\mathbf{G}_{f,D_f}$ for $1 \leq f \leq F$ on its diagonal.

Instead of a single signal, suppose we have R different sources $s_r(t)$. From this point on, we use the subscript r to denote the specific source. Each Löwner matrix $\mathbf{L}_r$, constructed with the same point set partitions admits to a decomposition of the form (7).

Let us assume that $I, J \geq \max(L_1, \ldots, L_R)$. First, the matrices $\mathbf{Z}_r$ and $\tilde{\mathbf{Z}}_r$ clearly have full rank for distinct points. Second, $\mathbf{G}_r$ is nonsingular since its diagonal blocks are nonsingular. Each Löwner matrix $\mathbf{L}_r$ is then of rank L_r and the rth term in (3) is rank-$(L_r, L_r, 1)$ for $1 \leq r \leq R$. Hence, eq. (3) is a decomposition of $\mathcal{L}_{\mathbf{X}}$ in rank-$(L_r, L_r, 1)$ terms. Furthermore, the matrix multiplication $\mathbf{A}_r \mathbf{B}_r^{\mathrm{T}}$ in (4) can be written as $\mathbf{Z}_r \mathbf{G}_r \tilde{\mathbf{Z}}_r^{\mathrm{T}}$ with the previously explained structure.

The next section will explain whether (and when) this decomposition in rank-$(L_r, L_r, 1)$ terms is unique.

IV. Uniqueness

The analysis in this section is the Löwner counterpart of the Hankel case [8]. We first recall [16, Theorem 4.1] regarding the uniqueness of a BTD in rank-$(L_r, L_r, 1)$ terms:

Theorem 2. *Consider a decomposition of $\mathcal{T} \in \mathbb{K}^{I \times J \times K}$ in rank-$(L_r, L_r, 1)$ terms as in* (1) *and* (2)*, with $I, J \geq \sum_{r=1}^{R} L_r$. If $\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_R \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_2 & \cdots & \mathbf{B}_R \end{bmatrix}$ have full column rank and $\mathbf{C} = \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_R \end{bmatrix}$ does not have proportional columns, then the decomposition is essentially unique.*

We call the decomposition 'essentially unique' when one can only permute the rth and r'th terms in (1) when $L_r = L_{r'}$ and when one can only scale $\mathbf{E}_r$ provided that $\mathbf{c}_r$ is

counterscaled. In [8] the theorem is generalized into the following theorem for the block term decomposition, including a necessary condition:

Theorem 3. *Consider a decomposition of $\mathcal{T} \in \mathbb{K}^{I\times J\times K}$ in rank-$(L_r, L_r, 1)$ terms as in* (1) *and* (2)*. Define $\mathbf{W}(\mathbf{w}) = \sum_{r=1}^{R} w_r \mathbf{E}_r$. Assume the following conditions to be satisfied:*

(C1) For every $\mathbf{w}$ that has at least two nonzero entries, we have that $\operatorname{rank}(\mathbf{E}(\mathbf{w})) > \max_{r|w_r\neq 0}(L_r)$.

(C2) The columns of $\mathbf{C}$ are linearly independent.

The decomposition is then essentially unique. On the other hand, if condition (C1) is not satisfied, then the decomposition is not essentially unique.

We now apply Theorem 2 to BSS for rational sources with coinciding poles and polynomial terms:

Theorem 4. *Consider a matrix $\mathbf{M} \in \mathbb{K}^{K\times R}$ that does not have proportional columns, and a matrix $\mathbf{S} \in \mathbb{K}^{R\times N}$ in which every row has a structure as in* (5)*. Assume that $\lfloor \frac{N+1}{2} \rfloor \geq \sum_{r=1}^{R} L_r$. If all the poles p_{r,f_r} are distinct for $1 \leq f_r \leq F_r$, $1 \leq r \leq R$ in* (5) *and if at most one source contains a polynomial term (at most one $W_r \neq 0$), then the decomposition $\mathbf{X} = \mathbf{MS}$ is essentially unique.*

Proof. The constraint $\lfloor \frac{N+1}{2} \rfloor \geq \sum_{r=1}^{R} L_r$ allows us to map the rows of $\mathbf{X}$ to Löwner matrices with sizes $I \times J$ and with $I, J \geq \sum_{r=1}^{R} L_r$. With distinct poles in (5), the matrices $\mathbf{Z} = \begin{bmatrix}\mathbf{Z}_{r=1} & \mathbf{Z}_{r=2} & \cdots & \mathbf{Z}_{r=R}\end{bmatrix} \cdot \operatorname{diag}(\mathbf{G}_{r=1}, \ldots, \mathbf{G}_{r=R})$ and $\tilde{\mathbf{Z}} = \begin{bmatrix}\tilde{\mathbf{Z}}_{r=1} & \tilde{\mathbf{Z}}_{r=2} & \cdots & \tilde{\mathbf{Z}}_{r=R}\end{bmatrix}$ have full column rank. This is assuming that at most one $W_r \neq 0$, i.e., at most one $\mathbf{Z}_{W_r}$ and $\tilde{\mathbf{Z}}_{W_r}$ from (6) is included in $\mathbf{Z}$ and $\tilde{\mathbf{Z}}$, respectively. Indeed, $\mathbf{Z}_{W_r}$ and $\mathbf{Z}_{W_{r'}}$ from (6) have mutually linear dependent columns for $r \neq r'$ (likewise for $\tilde{\mathbf{Z}}_{W_r}$). The uniqueness result then follows from Theorem 2. □

The proof shows that it is not possible to separate polynomials, as pointed out in [8] too. A polynomial from a single source can be identified though.

It is also important to remark that if the sources have distinct poles, uniqueness of the factorization $\mathbf{X} = \mathbf{MS}$ is guaranteed when enough samples are available. Even so, only $2 \times \sum_{r=1}^{R} L_r$ samples are needed, with L_r mostly small.

Furthermore, uniqueness results can be obtained if the sources share common poles too [8]. Part VI-D gives an example with the second fetal electrocardiogram experiment. A sufficient property can be found in [35, condition (5.18)].

Other general and useful results are presented in [36], while the procedure from [37] can be used to deduce generic uniqueness conditions.

V. Connection With Hankel-Based Tensorization and Vandermonde Decomposition

In Subsection V-A we give a connection between Löwner and Hankel matrices, with the latter being used in blind signal separation of exponential polynomials [8]. A discussion about the choice of sampling points is given in a second subsection, illustrating that an equidistant sampling is not needed for Löwner-based BSS.

A. Transformation Between Löwner and Hankel

A strong connection between Löwner and Hankel matrices was given in [28] and afterwards generalized in [29] and used in [34]. We repeat the theorem in a customized way:

Theorem 5. *For fixed point sets X and Y with distinct points, the mapping $\mathcal{F} : \mathbf{H} \to \mathbf{L} = \mathbf{W_x H W_y}^{\mathrm{T}}$ is an isomorphism (i.e. there is a one-to-one relationship) between the class of all $I \times J$ Hankel matrices and the class of $I \times J$ Löwner matrices corresponding to X and Y, with*

$$(\mathbf{W_x})_{k,l} = \frac{1}{l!}\frac{\mathrm{d}^l a_k(z)}{\mathrm{d}z^l}\bigg|_{z=0}, \quad a_k(z) = \prod_{i\neq k}(z - x_i);$$

$$(\mathbf{W_y})_{k,l} = \frac{1}{l!}\frac{\mathrm{d}^l b_k(z)}{\mathrm{d}z^l}\bigg|_{z=0}, \quad b_k(z) = \prod_{j\neq k}(z - y_j).$$

The matrices $\mathbf{W_x}$ and $\mathbf{W_y}$ depend only on the point sets X and Y and not on the actual signals. As the matrices $\mathbf{W_x}$ and $\mathbf{W_y}$ are of full column rank in the generic case, an important consequence of the isomorphism is that $\operatorname{rank}(\mathbf{L}) = \operatorname{rank}(\mathbf{H})$. The relationship enables, for example, the transposition of uniqueness results between the different techniques.

The condition number of $\mathbf{W_X}$ and $\mathbf{W_Y}$ depends heavily on the point set however. For equidistant samples on the real axis, the matrices are highly ill-conditioned, so that the explicit use of the transformation can pose numerical difficulties.

B. Choice of Sampling Points

The results obtained in the previous sections (such as Theorem 4 or the factorizations in Section III) are independent of the sample points used. Any sampling in the complex domain can be used.

For the Hankel case, one needs measurements sampled in an equidistant way; otherwise the Vandermonde decomposition is not applicable. This is a fundamental difference with respect to the Löwner case for which any collection of sampling points can be used: Chebyshev nodes, logarithmic distribution, … The arbitrary choice of sampling points enables the use of compressed sensing techniques. One can, for example, randomly sample the observed signals at a number of time instances, useful in the case of high-cost measurement settings.

VI. Experiments and Applications

In the first Subsection VI-A an example of a separation is given with one source containing a polynomial term. We also investigate the behavior for a low number of samples and with respect to the presence of noise. Subsection VI-B presents the separation for an underdetermined mixture with more source signals than observed signals. A third Subsection VI-C explains the illustration of fluorescence spectroscopy. We conclude with a description of the use of Löwner-based tensorization for fetal electrocardiogram extraction in Subsection VI-D.

To present the recovered sources and to determine the relative error, we use an optimal scaling and permutation step (the default indeterminacies in BSS) with respect to the theoretical sources. The relative error is then defined as the relative difference in Frobenius norm, e.g., if $\hat{\mathbf{S}}$ are the recovered sources

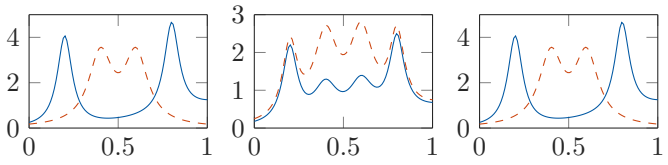


Fig. 3. Results for the first experiment of section VI-A. Left: the two original sources. Middle: the observed signals. Right: the perfectly recovered sources.

after this step, we have a relative error $\epsilon_{\mathbf{S}} = \|\mathbf{S} - \hat{\mathbf{S}}\|/\|\mathbf{S}\|$. Second, the signal-to-noise ratio (SNR) is defined as the power of the signal to the power of the noise, with the noise being Gaussian additive noise (unless stated otherwise). To calculate a BTD in rank-$(L_r, L_r, 1)$ terms, various approaches similar to CPD algorithms exist such as alternating least squares, unconstrained nonlinear optimization, or non-linear least squares [17], [38], [39]. We employ the latter by using Tensorlab [40]. A generalized eigenvalue decomposition is used for the initialization [16]. In all experiments only a few iterations are needed to reach convergence. For information about complexity we refer to [38], [39]. By default, the sampling points are chosen equidistantly on the real axis in $[0, 1]$. To construct the Löwner matrix, we use square matrices and partition the point set T in two interleaved partitions X and Y. An extensive analysis did not give a clear answer on which partitioning method is preferred; both methods described in section II-B give a similar performance. The sources are by default recovered by using the inverted mixing matrix, except for the underdetermined case in subsection VI-B.

A. General Experiment

We start with the separation of the two following sources:

$$s_1(t) = \left((t-0.2)^2 + 0.05^2\right)^{-1} + \left((t-0.8)^2 + 0.05^2\right)^{-1} + t^2,$$
$$s_2(t) = \left((t-0.4)^2 + 0.08^2\right)^{-1} + \left((t-0.6)^2 + 0.08^2\right)^{-1}.$$

The first source of degree 6 has two conjugated pole pairs $0.2 \pm 0.05j$ and $0.8 \pm 0.05j$, and also includes a polynomial term of degree 2. The second source has two conjugated pole pairs $0.4 \pm 0.08j$ and $0.6 \pm 0.08j$ and has degree 4. The two sources are divided by a factor 100 and 50, respectively, to obtain suitable magnitudes. Fig. 3 shows the signals. A mixture with $\mathbf{M} = [0.5,\ 0.3;\ 0.5,\ 0.7]$ is used, and $L_1 = 6$ and $L_2 = 4$. In a first case, we take 100 equidistant samples in $[0, 1]$. The recovered sources for the noiseless case are presented in Fig. 3 and the results for the noisy case are shown in Fig. 4 on the left. A second case, presented in function of SNR in Fig. 4 on the right, only uses 20 samples.

Fig. 5 shows the results for an experiment in which the number of source signals is varied, given ten observed signals. The ith source is given by $s_i(t) = \left((t - r_i)^2 + q^2\right)^{-1}$, with r_i equidistantly spaced in $[0, 1]$. Two cases are considered: $q = 0.01$ (mildly overlapping) and $q = 0.1$ (highly overlapping). An SNR of 25 is used. $\mathbf{M}$ is taken column-wise orthonormal.

B. Underdetermined System

We examine the identification of more sources than there are observed signals available. Consider the following three

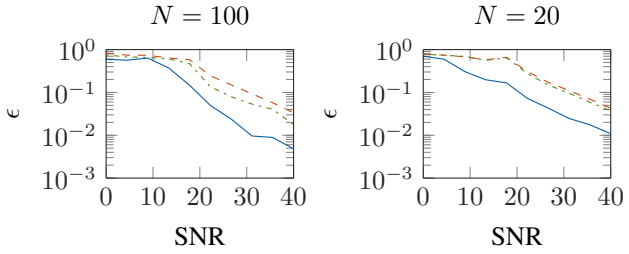


Fig. 4. Recovery results for 100 and 20 samples when Gaussian noise is added, in function of SNR. The median relative errors across 100 experiments are shown, for the mixing matrix (——), and for the recovered sources for the two different recovery methods of II-D: by using the inverted mixing matrix (– – –) and by using the Löwner matrices (·-·-·).

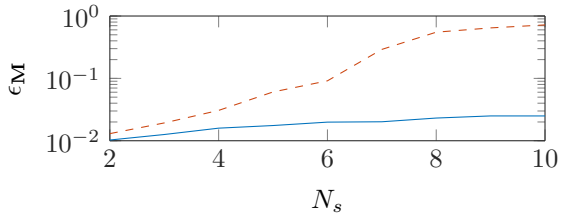


Fig. 5. Recovery results for a varying number of mildly (——) and highly (– – –) overlapping source signals. Median relative errors across 100 experiments are shown.

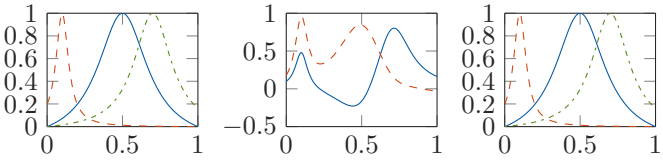


Fig. 6. Results for the underdetermined experiment. Left: the three original sources. Middle: the two observed mixed signals. Right: the three recovered sources after optimal scaling and permutation. A perfect recovery has been obtained of more sources than observed signals.

source signals:

$$s_1(t) = \left((t-0.1)^2 + 0.05^2\right)^{-1} \quad \text{(pole pair } 0.1 \pm 0.05j\text{)}$$
$$s_2(t) = \left((t-0.5)^2 + 0.20^2\right)^{-1} \quad \text{(pole pair } 0.5 \pm 0.20j\text{)}$$
$$s_3(t) = \left((t-0.7)^2 + 0.15^2\right)^{-1} \quad \text{(pole pair } 0.7 \pm 0.15j\text{)}$$

sampled in $t \in [0, 1]$ with $N = 100$ equidistant points. The signals are mixed into two observed signals using the mixing matrix $\mathbf{M} = [-0.5,\ 0.5,\ 1;\ 0.9,\ 0.9,\ -0.2]$. We use $L_1 = L_2 = L_3 = 2$. A perfect reconstruction of the three sources is obtained, as Fig. 6 illustrates.

In Fig. 7 we include results for varying SNR and different values of (L_1, L_2, L_3). Note that several choices of the degrees lead to good results. This shows that the choice of the (multilinear) rank(s) is not very critical [8], [20]. A trial-and-error method can be used to deduct L_r, knowing that the multilinear rank of $\mathcal{L}_{\mathbf{X}}$ is bounded by $(\sum_{r=1}^{R} L_r, \sum_{r=1}^{R} L_r, R)$.

C. Fluorescence Spectroscopy

Source separation, also known as curve resolution, is a valuable technique used in fluorescence spectroscopy. It enables the estimation of relative concentrations and pure analyte

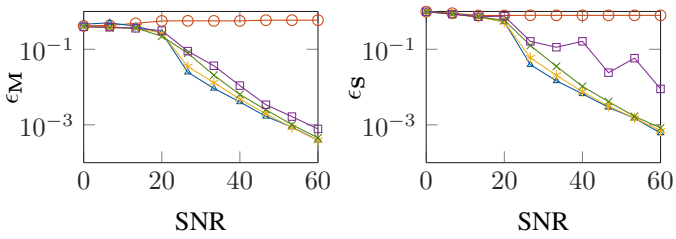


Fig. 7. Results for the underdetermined mixture, in function of SNR and for varying ranks with (L_1, L_2, L_3) being (2,1,1) (—○—); (2,2,2) (—△—); (3,2,2) (—✳—); (4,3,2) (—×—) and (4,4,4) (—□—). The sources are determined up to a constant, cf. subsection II-D. Median relative errors across 100 experiments are shown.

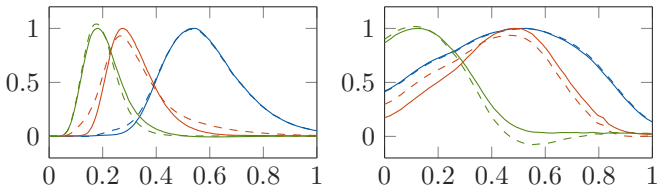


Fig. 8. Results for the fluorescence experiment. On the left the emission spectra are shown (being approximated with rational functions of degree 2) and on the right the excitation spectra are shown. The real signals are given by solid lines (——) and the reconstructed signals by dashed lines (- - -).

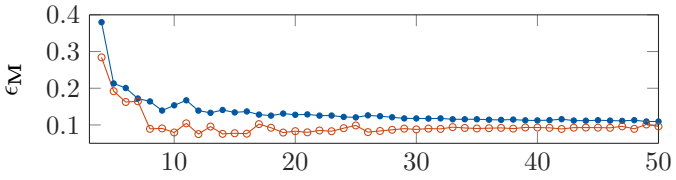


Fig. 9. Results for the fluorescence experiment of the relative error on the mixing matrix (—●—) and the recovered sources (—○—) in function of the number of excitations used.

spectra from fluorescence measurements of chemical analytes in mixtures. Consider a sufficiently diluted chemical solution containing different amounts of R chemical components. By exciting the mixture at K different excitation wavelengths and measuring the spectrum of the emitted light at N different emission wavelenghts, one obtains an intensity matrix called $\mathbf{X} \in \mathbb{R}^{K \times N}$. Through the Beer-Lambert law [41], the spectra of the mixture are linearly dependent on the spectra of the underlying chemical components and on their concentrations, and one can show that $\mathbf{X} = \mathbf{A}\mathbf{\Sigma}\mathbf{B} = \mathbf{M}\mathbf{S}$ with $\mathbf{A} \in \mathbb{R}^{K \times R}$ containing the excitation spectra of the R underlying chemical components in the columns, $\mathbf{B} \in \mathbb{R}^{R \times N}$ containing the emission spectra of the components in the rows, and $\mathbf{\Sigma} \in \mathbb{R}^{R \times R}$ a diagonal matrix with the concentrations. When interpreted in terms of source signals in $\mathbf{S}$ with a mixture matrix $\mathbf{M}$, the matrix $\mathbf{\Sigma}$ can be included both in $\mathbf{S}$ or $\mathbf{M}$. We allocate the emission spectra (rather than the excitation spectra) to the source signals. As the different spectra of the underlying components are unknown, this is a standard BSS problem.

Typical techniques using multilinear algebra resort to the use of multiple mixtures: the different excitation-emission matrices (EEM) are stacked, and a solution is obtained with a CPD [42]–[44]. Our technique only requires a single EEM, reducing the measurement cost and enabling analysis when only a single EEM is available. This is done with the understanding that the emission spectra can be well approximated by rational functions. It is also a possible alternative to time-dependent spectroscopy techniques.

The dataset[1] used in this paper contains the components phenylalanine, tyrosine and tryptophan [42], [45], [46]. The measurements are done with excitation wavelengths in 260-300 nm and emission wavelengths in 250-450 nm, both with steps of 1 nm. We mix the analytes with concentrations of 0.5, 0.2 and 0.3, respectively. In Fig. 8 the pure emission and excitation spectra of the components are visualized; the emission spectra have been Löwnerized and approximated by rational functions of a significantly low degree. The technique is used to separate the observations in three components with $L_r = 2$ for $r = 1, 2, 3$. The results are given in Fig. 8, and one can see that the components are separated very well given that only a single mixture has been used. A relative error on the emission (excitation) spectra of 0.1094 (0.096) has been obtained. For comparison, ICA yields a relative error of 0.640 (0.305), 0.818 (0.491) and 0.606 (0.682) for FastICA, RobustICA and JADE, respectively.

A final remarkable thing to mention is that in theory, we do not need fluorescence measurements across 50 different excitation wavelengths. The more observations in practice however, the more accurate the separation will be. Fig. 9 illustrates the findings if less than 50 excitation wavelengths are used. The extracted rows from the observed data matrix are selected as equidistant as possible.

D. Fetal Electrocardiogram Extraction

In this application, the proposed technique is used for the extraction of antepartum fetal electrocardiogram (fetal ECG or FECG) from multilead cutaneous potential recordings. While examining ECG recordings measured on the pregnant woman's skin (cutaneous), one tries to eliminate the dominant heartbeat of the mother. Seeing the problem as a blind signal separation problem, one can resort to the use of ICA [47]. ICA falls short however when only few samples or heartbeats are available. Second, for coinciding beats, the basic independence assumption of ICA is not valid.

It seems that ECG beats (with their easily recognizable QRS complexes) can be well modeled by rational functions [48]–[50]. The Löwner technique needs no preprocessing, as opposed to the technique in, e.g., [51]. We carry out two experiments with real-life datasets to illustrate the technique.

The first dataset consists of 8 measurement signals (of which 5 abdominal and 3 thoracic signals), available at DaISy[2] [47], [52]. For the sake of simplicity, only the 5 abdominal signals and only the first 500 samples are used, with the observations shown in Fig. 10. Each signal has been scaled to unit norm. For recovery, a separation into two source signals is not enough and at least three source signals are needed; this is also the

[1] Available from http://www.models.kvl.dk/Amino_Acid_fluo

[2] Available from http://homes.esat.kuleuven.be/~smc/daisy/daisydata.html.

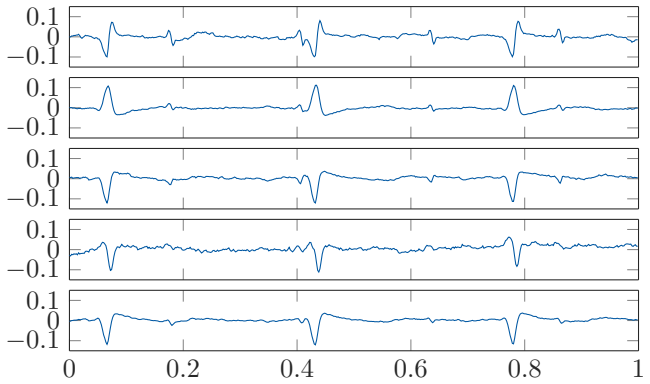


Fig. 10. Visualization of the 5 abdominal ECG recordings used in the first FECG experiment.

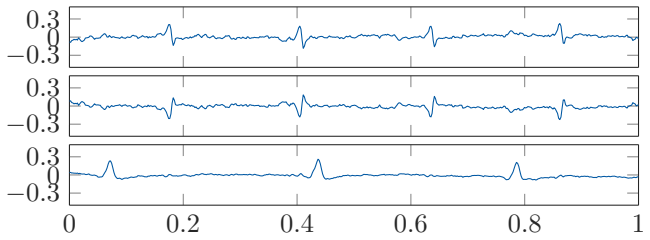


Fig. 11. The separation of the ECG recordings into three recovered source signals for the first FECG experiment. One clearly notices the separation of the fetal heart beats (above) and the heart beats of the mother (below). Typically, the fetal heart rate is significantly higher than the mother's heart rate.

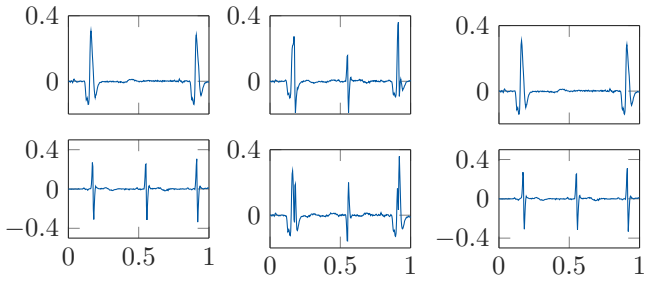


Fig. 12. Illustrations for the second FECG experiment. Left we have a limited amount of heart beats of mother (above) and fetus (below) where some beats coincide. The mixed signals are shown in the middle. To the right, the recovered sources are shown. An excellent recovery is obtained with a relative error on the sources of only 0.013.

case when applying ICA [47]. For the BTD, a rank of 20 for each source signal has been used. The three recovered sources are visualized in Fig. 11 with a clear separation of the two different ECG sources.

The second dataset contains a limited number of heartbeats with the beats of the mother and fetus coinciding. A mixing matrix $\mathbf{M} = [1\ 1; 1\ -0.8]$ is used to mix the signals. Fig. 12 visualizes the signals and the recovered sources. When using the proposed technique with again a rank 20 for each source signal, an excellent recovery is obtained with a relative source error of 0.013. To compare, ICA recovers the signals up to a relative error of 0.126, 0.125 and 0.29 for FastICA, RobustICA and JADE, respectively.

VII. Conclusion

A novel technique for blind signal separation has been proposed for signals that can be modeled as rational functions. The tensor-based technique makes use of a deterministic tensorization with Löwner matrices and the obtained tensor is decomposed with a block term decomposition. In the paper, the factorization of the Löwner matrices has been analyzed, together with the uniqueness conditions. The proposed method can be applied to any collection of sampling points and not only for equidistant points, as has been discussed while relating the method to another separation technique using a source model of exponential polynomials. Two synthetic experiments (including an underdetermined mixture) and two real-life illustrations with fluorescence spectroscopy and fetal electrocardiogram extraction have been used to verify the proposed technique. The method has been compared against ICA, demonstrating the power of the deterministic Löwner technique when the source signals are not independent or when only a limited number of samples are available.

References

[1] A. Cichocki, R. Zdunek, A. Phan, and S. Amari, *Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation*. Wiley, 2009.

[2] C. Jutten and J. Herault, "Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture," *Signal Processing*, vol. 24, no. 1, pp. 1 – 10, 1991.

[3] P. Comon and C. Jutten, *Handbook of blind source separation: Independent component analysis and applications*. Academic Press, 2010.

[4] V. D. Calhoun, T. Adali, G. D. Pearlson, and K. A. Kiehl, "Neuronal chronometry of target detection: fusion of hemodynamic and event-related potential data," *Neuroimage*, vol. 30, no. 2, pp. 544–553, 2006.

[5] A. Cichocki and S. Amari, *Adaptive blind signal and image processing: learning algorithms and applications*. John Wiley, 2002, vol. 1.

[6] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural networks*, vol. 13, no. 4, pp. 411–430, 2000.

[7] J. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications," *Proceedings of the IEEE*, vol. 79, no. 3, pp. 278–305, Mar 1991.

[8] L. De Lathauwer, "Blind separation of exponential polynomials and the decomposition of a tensor in rank-$(L_r, L_r, 1)$ terms," *SIAM Journal on Matrix Analysis and Applications*, vol. 32, no. 4, pp. 1451–1474, 2011.

[9] J. D. Hamilton, *Time series analysis*. Princeton university press Princeton, 1994, vol. 2.

[10] T. Yokota, R. Zdunek, A. Cichocki, and Y. Yamashita, "Smooth nonnegative matrix and tensor factorizations for robust multi-way data analysis," *Signal Processing*, no. 0, pp. –, 2015.

[11] Y. Li, A. Cichocki, and S.-i. Amari, "Analysis of sparse representation and blind source separation," *Neural computation*, vol. 16, no. 6, pp. 1193–1234, 2004.

[12] M. Zibulevsky and B. Pearlmutter, "Blind source separation by sparse decomposition in a signal dictionary," *Neural computation*, vol. 13, no. 4, pp. 863–882, 2001.

[13] A. C. Antoulas and B. D. O. Anderson, "On the scalar rational interpolation problem," *IMA Journal of Mathematical Control and Information*, vol. 3, no. 2-3, pp. 61–88, 1986.

[14] A. C. Antoulas and D. C. Sorensen, "Approximation of large-scale dynamical systems: An overview," *International Journal of Applied Mathematics and Computer Science*, vol. 11, pp. 1093–1121, 2001.

[15] T. Antoulas, "A tutorial introduction to the Loewner framework for model reduction," 2014, 9th Elgersburg Workshop Mathematische Systemtheorie.

[16] L. De Lathauwer, "Decompositions of a higher-order tensor in block terms — Part II: Definitions and uniqueness," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1033–1066, 2008.

[17] L. De Lathauwer and D. Nion, "Decompositions of a higher-order tensor in block terms — Part III: Alternating least squares algorithms," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1067–1083, 2008.

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

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

[20] L. De Lathauwer, "Block component analysis, a new concept for blind source separation," in *Latent Variable Analysis and Signal Separation*, ser. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2012, vol. 7191, pp. 1–8.

[21] A. Ferréol, L. Albera, and P. Chevalier, "Fourth-order blind identification of underdetermined mixtures of sources (FOBIUM)," *Signal Processing, IEEE Transactions on*, vol. 53, no. 5, pp. 1640–1653, 2005.

[22] L. De Lathauwer, J. Castaing, and J.-F. Cardoso, "Fourth-order cumulant-based blind identification of underdetermined mixtures," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2965–2973, 2007.

[23] O. Debals, M. Van Barel, and L. De Lathauwer, "Blind signal separation of rational functions using Löwner-based tensorization," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2015, pp. 4145–4149.

[24] L. De Lathauwer, "Signal processing based on multilinear algebra," Ph.D. dissertation, Katholieke Universiteit Leuven, Belgium, 1997.

[25] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1253–1278, 2000.

[26] ——, "On the best rank-1 and rank-$(R_1, R_2, \ldots, R_n)$ approximation of higher-order tensors," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1324–1342, 2000.

[27] N. Vervliet, O. Debals, L. Sorber, and L. De Lathauwer, "Breaking the curse of dimensionality using decompositions of incomplete tensors: Tensor-based scientific computing in big data analysis," *Signal Processing Magazine, IEEE*, vol. 31, no. 5, pp. 71–79, September 2014.

[28] M. Fiedler, "Hankel and Löwner matrices," *Linear Algebra and its Applications*, vol. 58, pp. 75–95, 1984.

[29] Z. Vavřín, "A unified approach to Löwner and Hankel matrices," *Linear Algebra and its Applications*, vol. 143, pp. 171–222, 1991.

[30] V. Belevitch, "Interpolation matrices," *Philips Res. Rep*, vol. 25, pp. 337–369, 1970.

[31] K. Löwner, "Über monotone matrixfunktionen," *Mathematische Zeitschrift*, vol. 38, no. 1, pp. 177–216, 1934.

[32] D. Vandevoorde, "A fast exponential decomposition algorithm and its applications to structured matrices," Ph.D. dissertation, Rensselaer Polytechnic Institute, Troy, NY, 1998.

[33] D. L. Boley, F. T. Luk, and D. Vandevoorde, "A fast method to diagonalize a hankel matrix," *Linear algebra and its applications*, vol. 284, no. 1, pp. 41–52, 1998.

[34] Z. Vavřín, "Confluent Cauchy and Cauchy-Vandermonde matrices," *Linear algebra and its applications*, vol. 258, pp. 271–293, 1997.

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

[36] 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 Journal on Matrix Analysis and Applications*, vol. 36, no. 2, pp. 496 – 522, 2015.

[37] I. Domanov and L. De Lathauwer, "Generic uniqueness of a structured matrix factorization and applications in blind source separation," ESAT-STADIUS, KU Leuven (Leuven, Belgium), Tech. Rep. 14-153, 2014.

[38] L. Sorber, M. Van Barel, and L. De Lathauwer, "Optimization-based algorithms for tensor decompositions: Canonical polyadic decomposition, decomposition in rank-$(L_r, L_r, 1)$ terms, and a new generalization," *SIAM Journal on Optimization*, vol. 23, no. 2, pp. 695–720, 2013.

[39] ——, "Structured data fusion," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 4, pp. 586–600, June 2015.

[40] ——. "Tensorlab v2.0." Available online. URL: http://www.tensorlab.net.

[41] S. E. Leurgans and R. T. Ross, "Multilinear models: applications in spectroscopy," *Statistical Science*, pp. 289–310, 1992.

[42] R. Bro, "Parafac. tutorial and applications," *Chemometrics and intelligent laboratory systems*, vol. 38, no. 2, pp. 149–171, 1997.

[43] C. M. Andersen and R. Bro, "Practical aspects of parafac modeling of fluorescence excitation-emission data," *Journal of Chemometrics*, vol. 17, no. 4, pp. 200–215, 2003.

[44] A. K. Smilde, R. Bro, P. Geladi, and J. Wiley, *Multi-way analysis with applications in the chemical sciences*. Wiley Chichester, UK, 2004.

[45] R. Bro, "Multi-way analysis in the food industry. models, algorithms, and applications." Ph.D. dissertation, University of Amsterdam (NL) & Royal Veterinary and Agricultural University (DK), 1998.

[46] H. A. Kiers, "A three–step algorithm for candecomp/parafac analysis of large data sets with multicollinearity," *Journal of Chemometrics*, vol. 12, no. 3, pp. 155–171, 1998.

[47] L. De Lathauwer, B. De Moor, and J. Vandewalle, "Fetal electrocardiogram extraction by blind source subspace separation," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 5, pp. 567–572, 2000.

[48] S. Fridli, L. Lócsi, and F. Schipp, "Rational function systems in ECG processing," in *Computer Aided Systems Theory–EUROCAST 2011*. Springer, 2012, pp. 88–95.

[49] L. Lócsi, "Approximating poles of complex rational functions," *Acta Univ. Sapientiae Math*, vol. 1, pp. 169–182, 2009.

[50] ——, "Rational function systems with applications in signal processing," Ph.D. dissertation, Eötvös Loránd University, Budapest, Hungaryf, 2014.

[51] E. Karvounis, M. Tsipouras, D. Fotiadis, and K. Naka, "An automated methodology for fetal heart rate extraction from the abdominal electrocardiogram," *IEEE Transactions on Information Technology in Biomedicine*, vol. 11, no. 6, pp. 628–638, Nov 2007.

[52] D. Callaerts, "Signal separation methods based on singular value decomposition and their application to the real-time extraction of the fetal electrocardiogram from cutaneous recordings," Ph.D. dissertation, KU Leuven, Dec 1989.



Otto Debals obtained the M.Sc. degree in mathematical engineering from KU Leuven, Belgium, in 2013.

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



Marc Van Barel received the Ph.D. degree in computer science from the Faculty of Engineering, KU Leuven, Belgium, in 1989.

He is currently a Full Professor at KU Leuven and head of Numerical Approximation and Linear Algebra Group (NALAG), one of the three research groups of the Division Numerical Analysis and Applied Mathematics within the Department of Computer Science. His research concerns the development of the theory and algorithms for matrix and tensor computations and multivariate polynomial and rational approximation.

Professor Van Barel is co-editor of six special issues of journals devoted to numerical linear algebra. He is also Associate Editor of the SIAM *Journal in Matrix Analysis and Applications* and editorial board member of *Linear and Multilinear Algebra*, of *Calcolo* and of *Numerical Algorithms*.



Lieven De Lathauwer received the Ph.D. degree from the Faculty of Engineering, KU Leuven, Belgium, in 1997.

From 2000 to 2007 he was a Research Associate with the Centre National de la Recherche Scientifique, France. He is currently a Professor with KU Leuven. He is affiliated with the Group Science, Engineering and Technology of Kulak, with the Stadius Center for Dynamical Systems, Signal Processing and Data Analytics of the Electrical Engineering Department (ESAT) and with iMinds Medical IT. His research concerns the development of tensor tools for engineering applications.

Professor De Lathauwer is an Associate Editor of the SIAM *Journal on Matrix Analysis and Applications* and served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING.