Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT®

Signal Processing 86 (2006) 846–858

SIGNAL PROCESSING

www.elsevier.com/locate/sigpro

Common pole estimation in multi-channel exponential data modeling

Jean-Michel Papy[a,*], Lieven De Lathauwer[a,b], Sabine Van Huffel[a]

[a]*Katholieke Universiteit Leuven, ESAT, division SCD-SISTA Kasteelpark Arenberg 10, 3001 Leuven-Heverlee, Belgium*
[b]*Lab. ETIS, UMR 8051 (ENSEA, UCP, CNRS), 6, avenue du Ponceau, BP 44, F 95014 Cergy-Pontoise Cedex, France*

Received 4 October 2004; received in revised form 1 June 2005; accepted 7 July 2005
Available online 8 August 2005

Abstract

In this paper we develop a technique for detecting and retrieving the common harmonics in a multi-channel setup. The different (complex) signals are arranged in a set of Hankel matrices. First, we compute the singular value decomposition (SVD) of these matrices in order to denoise and normalize the relevant subspaces. Then a second SVD is applied to detect and select the common subspace. The poles of the harmonics are computed from the common subspace using the total least squares (TLS) technique. This method is very flexible, robust and outperforms existing subspace-based methods. Moreover, it can be applied to an arbitrary number of channels.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Multi-channel; Harmonic retrieval; Singular value decomposition; Common poles; Total least squares; Subspace

1. Introduction

Many real-world signals are naturally modeled as a sum of exponentially damped sinusoids. This is, for instance, the case in nuclear magnetic resonance (NMR) spectroscopy where the quantification of the complex time-domain signals [1–3] is of great importance for brain tumor detection, in material health monitoring where the model parameters help to characterize the damage occurring in a structure [4], or in audio processing where such a model is used for low bit rate audio compression [5].

In general, this type of signals exhibit a rank deficient correlation matrix, and therefore lend themselves to subspace-based processing. One of the most efficient techniques consist in first computing the *singular value decomposition*

*Corresponding author. Tel.: +32 16 32 10 67; fax: +32 16 32 19 70.
E-mail address: jean-michel.papy@esat.kuleuven.ac.be (J.-M. Papy).

0165-1684/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2005.07.015

(SVD) of the Hankel data matrix; then a system of equations is built via a similarity transformation of the dominant subspace to a Vandermonde vector basis by means of the shift-invariance property; finally, the total least squares (TLS) solution of this system is computed and yields the frequencies and damping factors. This method is called HTLS [1], a special case of the ESPRIT algorithm [6] and a TLS variant of Kung et al.'s algorithm [7]. In the subspace-based processing framework, a multi-channel set-up make it possible to extract the common information in a more accurate way if the signals are contaminated by noise. In this paper we address the case in which we have several channels, each of which can be modeled as a sum of discrete-time complex exponentials. In [8], a multi-channel harmonic retrieval algorithm, called HTLSstack, has been studied. With the HTLSstack technique it is possible to estimate all the discrete-time complex exponential signals present in a multi-channel signal. However, it is often of great interest to quantify and extract only the common information present in the multi-channel signal in the case where the channels are contaminated by different undesired harmonics. This is possible by determining the subspace common to all the data matrices. In the two-channel case several methods already exist. For instance in [9], an algorithm based on the QR decomposition for determining the common subspace of two matrices via canonical angles is provided. It is possible, from this subspace, to retrieve the parameters of interest by computing the TLS solution. Alternatively, in [10] a method, called HTLS-SEP, based on SVD and TLS, is used to estimate the complex exponentials common to two signals. However, these two methods only apply in the two-channel case.

The method which is studied in this paper applies to an arbitrary number of channels. It consists in using unitary matrices for determining a set of linearly independent vectors spanning the common space if it exists. Then the parameters of the common complex exponentials are computed via the TLS solution. This is achieved without setting a too severe constraint on the determination of the model order of each channel. In Section 2, we summarize in three steps the HTLS technique which serves as basis for the new algorithm. Then, the multi-channel case is addressed in Section 3; after the description of the problem in Section 3.1, the algorithm is outlined in Section 3.2, then the underlying theory is given in Section 3.3 and discussed in Section 3.3.1. The computational complexity is discussed in Section 3.3.2. Finally, the algorithm is applied to simulated signals in Section 4.

2. Hankel total least squares method

2.1. Signal model

Assume a single-channel noiseless time-domain signal containing N complex samples x_n, $n = 0, 1, \ldots, N-1$. Let this time series be modeled as a finite sum of K different exponentially damped complex sinusoids:

$$x_n = \sum_{k=1}^{K} a_k \exp\{j\varphi_k\} \exp\{(-\alpha_k + j\omega_k)t_n\}, \tag{1}$$

where $j = \sqrt[2]{-1}$, $t_n = n\Delta t$ is the time lapse between the time origin and the sample x_n and Δt is the sampling time interval. The parameters of this model are the amplitudes a_k, the phases φ_k, the damping factors α_k and the pulsations ω_k. The frequencies v_k can be obtained from the pulsations by means of the equality: $\omega_k = 2\pi v_k$. One can rewrite Eq. (1) in a more compact form:

$$x_n = \sum_{k=1}^{K} c_k z_k^n, \tag{2}$$

where $c_k = a_k \exp\{j\varphi_k\}$ denotes the kth complex amplitude including the phase, and $z_k = \exp\{(-\alpha_k + j\omega_k)\Delta t\}$ is called the kth pole of the signal with $z_k \neq z_{k'}$ for $k \neq k'$.

We will now explain how this model can be handled in a subspace framework.

2.2. Subspace representation

First, given x_n and K, the data samples are arranged into an $L \times M$ Hankel matrix as

follows:

$$\boldsymbol{H}(x_n) = \begin{pmatrix} x_0 & x_1 & x_2 & \cdots & x_{M-1} \\ x_1 & x_2 & \ddots & \cdots & \vdots \\ x_2 & \ddots & \ddots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & x_{N-2} \\ x_{L-1} & \cdots & \cdots & x_{N-2} & x_{N-1} \end{pmatrix}, \quad (3)$$

with $\{L, M\}$ chosen such that $N = L + M - 1$ and $L > K$. This Hankel matrix can be straightforwardly decomposed in a product of three matrices:

$$\begin{aligned} \boldsymbol{H}(x_n) &= \begin{pmatrix} 1 & \cdots & 1 \\ z_1^1 & \cdots & z_K^1 \\ z_1^2 & \cdots & z_K^2 \\ \vdots & \vdots & \vdots \\ z_1^{L-1} & \cdots & z_K^{L-1} \end{pmatrix} \begin{pmatrix} c_1 & & \\ & & 0 \\ & \ddots & \\ 0 & & \\ & & c_K \end{pmatrix} \\ &\quad \times \begin{pmatrix} 1 & z_1^1 & z_1^2 & \cdots & z_1^{M-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & z_K^1 & z_K^2 & \cdots & z_K^{M-1} \end{pmatrix} \\ &= \boldsymbol{S}\boldsymbol{C}\boldsymbol{T}^{\mathsf{T}}. \end{aligned} \quad (4)$$

Eq. (4) is called a *Vandermonde decomposition* (VDMD) in which the poles z_k are also called the *generators* and the superscript T denotes the transpose. The rank deficiency of $\boldsymbol{H}$ is also reflected by the SVD defined as follows:

$$\boldsymbol{H} = (\widehat{\boldsymbol{U}} \;\; \boldsymbol{U}_0) \begin{pmatrix} \widehat{\boldsymbol{\Sigma}} & 0 \\ 0 & \boldsymbol{\Sigma}_0 \end{pmatrix} \begin{pmatrix} \widehat{\boldsymbol{V}}^{\mathsf{H}} \\ \boldsymbol{V}_0^{\mathsf{H}} \end{pmatrix}, \quad (5)$$

where $\widehat{\boldsymbol{U}} \in \mathbb{C}^{L\times K}$, $\boldsymbol{U}_0 \in \mathbb{C}^{L\times(L-K)}$, $\widehat{\boldsymbol{\Sigma}} \in \mathbb{R}^{K\times K}$ contains the K non-zero singular values in decreasing order of magnitude, $\boldsymbol{\Sigma}_0 \in \mathbb{R}^{(L-K)\times(M-K)}$, $\widehat{\boldsymbol{V}} \in \mathbb{C}^{M\times K}$, $\boldsymbol{V}_0 \in \mathbb{C}^{K\times(M-K)}$. The superscript H denotes the Hermitian transpose. In the absence of noise $\boldsymbol{\Sigma}_0$ is a null matrix and the SVD of $\boldsymbol{H}$ reduces to the product $\widehat{\boldsymbol{U}}\widehat{\boldsymbol{\Sigma}}\widehat{\boldsymbol{V}}^{\mathsf{H}}$. The columns of $\boldsymbol{S}$ (resp. $\boldsymbol{T}$) span the same subspace as the columns of $\widehat{\boldsymbol{U}}$ (resp. $\widehat{\boldsymbol{V}}^*$). If the signal is corrupted by noise, $\boldsymbol{\Sigma}_0$ is full rank. In this case, the truncated SVD of $\boldsymbol{H}$,

$$\widehat{\boldsymbol{H}}_{L\times M} = \widehat{\boldsymbol{U}}_{L\times K}\, \widehat{\boldsymbol{\Sigma}}_{K\times K}\, \widehat{\boldsymbol{V}}^{\mathsf{H}}_{K\times M}, \quad (6)$$

is the best rank-K approximation of $\boldsymbol{H}$.

2.3. Harmonic retrieval via total least squares solution

The matrices $\boldsymbol{S}$ and $\boldsymbol{T}$ possess a shift-invariance property that can be expressed as

$$\begin{aligned} \boldsymbol{S}_{\downarrow}\boldsymbol{Z} &= \boldsymbol{S}^{\uparrow}, \\ \boldsymbol{T}_{\downarrow}\boldsymbol{Z} &= \boldsymbol{T}^{\uparrow}, \end{aligned} \quad (7)$$

where the up (down) arrow placed behind a matrix stands for deleting the top (bottom) row of the considered matrix and $\boldsymbol{Z} = \mathrm{diag}(z_1, z_2, \ldots, z_K) \in \mathbb{C}^{K\times K}$. In the noise-free case, $\widehat{\boldsymbol{U}}$ equals $\boldsymbol{S}$ up to a multiplication by a square non-singular matrix $\boldsymbol{Q} \in \mathbb{C}^{K\times K}$:

$$\widehat{\boldsymbol{U}} = \boldsymbol{S}\boldsymbol{Q}. \quad (8)$$

The matrices $\widehat{\boldsymbol{U}}^{\uparrow}$ (resp. $\widehat{\boldsymbol{U}}_{\downarrow}$) are related to the Vandermonde matrices $\boldsymbol{S}^{\uparrow}$ (resp. $\boldsymbol{S}_{\downarrow}$) in the following way:

$$\begin{aligned} \widehat{\boldsymbol{U}}^{\uparrow} &= \boldsymbol{S}^{\uparrow}\boldsymbol{Q}, \\ \widehat{\boldsymbol{U}}_{\downarrow} &= \boldsymbol{S}_{\downarrow}\boldsymbol{Q}. \end{aligned} \quad (9)$$

Combining Eqs. (7) and (9) results in the shift-invariance property of $\widehat{\boldsymbol{U}}$:

$$\widehat{\boldsymbol{U}}^{\uparrow} = \widehat{\boldsymbol{U}}_{\downarrow}\boldsymbol{Q}^{-1}\boldsymbol{Z}\boldsymbol{Q} = \widehat{\boldsymbol{U}}_{\downarrow}\widetilde{\boldsymbol{Z}}. \quad (10)$$

Note that the same reasoning applies to the matrix $\widehat{\boldsymbol{V}}^*$. In the case of white noise, the equality does not hold anymore. The TLS solution of the overdetermined set of linear equations $\widehat{\boldsymbol{U}}^{\uparrow} \approx \widehat{\boldsymbol{U}}_{\downarrow}\widetilde{\boldsymbol{Z}}$, whose aim is to minimize the correction applied to the data in order to make the set compatible, is given by

$$\widehat{\widetilde{\boldsymbol{Z}}} = -\boldsymbol{W}_{12}\boldsymbol{W}_{22}^{-1}, \quad (11)$$

where

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{W}_{11} & \boldsymbol{W}_{12} \\ \boldsymbol{W}_{21} & \boldsymbol{W}_{22} \end{bmatrix} \quad (12)$$

is obtained from the thin SVD of the matrix $[\widehat{\boldsymbol{U}}_{\downarrow}\ \widehat{\boldsymbol{U}}^{\uparrow}]$:

$$[\widehat{\boldsymbol{U}}_{\downarrow}\ \widehat{\boldsymbol{U}}^{\uparrow}] \overset{\text{SVD}}{=} \boldsymbol{Y}\boldsymbol{\Gamma}\boldsymbol{W}^{\mathsf{H}}, \tag{13}$$

where $\boldsymbol{Y} \in \mathbb{C}^{(L-1)\times 2K}$ is the (semi-unitary) left singular vector matrix, $\boldsymbol{W} \in \mathbb{C}^{2K\times 2K}$ is the (unitary) right singular vector matrix, and $\boldsymbol{\Gamma} \in \mathbb{R}^{2K\times 2K}$ is the (diagonal) singular value matrix. Once $\widetilde{\boldsymbol{Z}}$ is estimated, its eigenvalues λ_k give an estimate of the signal poles:

$$\lambda_k = \hat{z}_k = \exp\{(-\hat{\alpha}_k + 2\pi j\hat{\nu}_k)\Delta t\}. \tag{14}$$

Finally, the complex amplitudes can be estimated by solving the set of equations

$$\begin{pmatrix} 1 & \cdots & 1 \\ \hat{z}_1^1 & \cdots & \hat{z}_K^1 \\ \hat{z}_1^2 & \cdots & \hat{z}_K^2 \\ \vdots & \vdots & \vdots \\ \hat{z}_1^{N-1} & \cdots & \hat{z}_K^{N-1} \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_K \end{pmatrix} = \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{N-1} \end{pmatrix} \tag{15}$$

in the least squares (LS) sense.

3. Determining the set of signal poles common to all channels

3.1. Problem statement

Let us first assume exact data. The influence of the noise will be discussed later on. Given a G-channel discrete-time signal $x_n^{(g)}$ where $g = 1,\ldots,G$ denotes the channel number and $n = 0,\ldots,N-1$ is the time index. Let each $x_n^{(g)}$ be modeled as a finite sum of K_g weighted discrete-time complex exponentials (1), with K_g given, and let these signals have κ common poles $\{z_k\}$, $(k = 1,\ldots,\kappa)$. For each channel g we can write

$$x_n^{(g)} = \sum_{k=1}^{K_g} c_{g,k}\, z_{g,k}^n = \sum_{k=1}^{\kappa} c_k\, z_k^n + \sum_{k=1}^{K_g-\kappa} \tilde{c}_{g,k}\, \tilde{z}_{g,k}^n \quad n = 0,\ldots,N-1, \tag{16}$$

with $\{c_{g,k}\} = \{c_k, \tilde{c}_{g,k}\}$, $\{z_{g,k}\} = \{z_k, \tilde{z}_{g,k}\}$ and $\tilde{z}_{g,k} \neq \tilde{z}_{g',k'}$ for all $k \neq k'$ or $g \neq g'$.

We give a summary of the algorithm in the next section. The explanation is given in Section 3.3.

3.2. Outline of the algorithm

Algorithm: Common pole estimation

Given G multi-channel signals $x_n^{(g)}$, $g = 1,\ldots,G$ of N samples $n = 0,\ldots,N-1$ with corresponding model order K_g, L and M such that $N = L + M - 1$:

Step 1: Map each signal into an Hankel matrix: $x_n^{(g)} \xrightarrow{\mathscr{H}_L^M(\cdot)} \mathscr{H}_L^M(x_n^{(g)}) = \boldsymbol{H}_g$, where $\mathscr{H}_L^M(\cdot)$ denotes the Hankel mapping operator.

Step 2: Compute the SVD of each $\boldsymbol{H}_g$ separately: $\boldsymbol{H}_g = \boldsymbol{U}_g\boldsymbol{\Sigma}_g\boldsymbol{V}_g^{\mathsf{H}}$.

Step 3: Retain the K_g leftmost singular vectors of the left singular vectors matrix $\boldsymbol{U}_g$: $\boldsymbol{U}_{g\ (L\times M)} \longrightarrow \tilde{\boldsymbol{U}}_{g\ (L\times K_g)}$. If K_g is not given, it can, for example, be determined by using a rank determinator, a scree plot or an information criterion.

Step 4: Stack the vectors obtained in the previous step in a $(L \times \sum_{g=1}^{G} K_g)$-matrix: $\boldsymbol{U}_{\text{tot}} = [\tilde{\boldsymbol{U}}_1 \,|\, \tilde{\boldsymbol{U}}_2 \,|\, \cdots \,|\, \tilde{\boldsymbol{U}}_G]$.

Step 5: Compute the SVD of $\boldsymbol{U}_{\text{tot}}$: $\boldsymbol{U}_{\text{tot}} = \tilde{\boldsymbol{V}}\tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{W}}^{\mathsf{H}}$.

Step 6: Display the entries of $\tilde{\boldsymbol{\Sigma}}$: if there exist κ common poles, the κ first singular values are equal to $\sqrt{G}$.

Step 7: Form a matrix $\widehat{\boldsymbol{U}}$ with the κ leftmost column vectors of $\tilde{\boldsymbol{V}}$ and use (11)–(14) to retrieve the parameters of interest.

This algorithm does not present any difficulty of implementation. However, the main issue is to understand the key-points 4–6. An explanation is given in the following section. Point 7 follows naturally from point 6 by using the well-known TLS method described in Section 2.3.

3.3. Determining a set of vectors spanning the common subspace

In the noise-free case we have, in analogy with (4), (5),

$$\boldsymbol{H}_g \overset{\text{VDMD}}{=} \boldsymbol{S}_g\boldsymbol{C}_g\boldsymbol{T}_g^{\mathsf{T}} \overset{\text{SVD}}{=} \widehat{\boldsymbol{U}}_g\widehat{\boldsymbol{\Sigma}}_g\widehat{\boldsymbol{V}}_g^{\mathsf{H}} \quad g = 1,\ldots,G, \tag{17}$$

where $\boldsymbol{H}_g \in \mathbb{C}^{L\times M_g}$, $\boldsymbol{S}_g \in \mathbb{C}^{L\times K_g}$, $\boldsymbol{C}_g \in \mathbb{C}^{K_g\times K_g}$, $\boldsymbol{T}_g \in \mathbb{C}^{M_g\times K_g}$, $\widehat{\boldsymbol{U}}_g \in \mathbb{C}^{L\times K_g}$, $\widehat{\boldsymbol{\Sigma}}_g \in \mathbb{R}^{K_g\times K_g}$, $\widehat{\boldsymbol{V}}_g \in \mathbb{C}^{M_g\times K_g}$. The Vandermonde matrices $\boldsymbol{S}_g$ ($g = 1,\ldots,G$) have κ common vectors:

$$\boldsymbol{S}_g = \begin{pmatrix} 1 & \cdots & 1 & 1 & \cdots & 1 \\ z_1 & \cdots & z_\kappa & \tilde{z}_{g,1} & \cdots & \tilde{z}_{g,K_g-\kappa} \\ z_1^2 & \cdots & z_\kappa^2 & \tilde{z}_{g,1}^2 & \cdots & \tilde{z}_{g,K_g-\kappa}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ z_1^{L-1} & \cdots & z_\kappa^{L-1} & \tilde{z}_{g,1}^{L-1} & \cdots & \tilde{z}_{g,K_g-\kappa}^{L-1} \end{pmatrix} = [\boldsymbol{S}|\widetilde{\boldsymbol{S}}_g], \tag{18}$$

where the generators of $\boldsymbol{S}$ are the common signal poles, and the generators of $\widetilde{\boldsymbol{S}}_g$ are signal poles that are not present in all channels.

We recall that, generically, none of the column vectors of a given Vandermonde matrix $\boldsymbol{S}_g$ can be expressed as a linear combination of the other ones when the generators are different. In other words, the column vectors of any $\boldsymbol{S}_g$ can be seen as an nonorthogonal basis of a complex vector subspace of dimension K_g belonging to a complex vector space of dimension L. It follows that the intersection of the G subspaces, spanned by the column vectors of each matrix $\boldsymbol{S}_g$, contains the column vectors of $\boldsymbol{S}$, but not more than that. Thus, the column vectors of $\boldsymbol{S}$ form an nonorthogonal basis for the common subspace C. We formally write

$$\bigcap_{g=1}^{G} \mathrm{Span}(s_{g,1},\ldots,s_{g,K_g}) = C \tag{19}$$

with

$$\mathrm{Dim}(C) = \kappa, \tag{20}$$

where $s_{g,k}$ is the kth column vector of $\boldsymbol{S}_g$. Obviously, the column vectors of the semi-unitary matrix $\widehat{\boldsymbol{U}}_g$ and the column vectors of the Vandermonde matrix $\boldsymbol{S}_g$ span exactly the same subspace. Therefore, property (19) also applies to the semi-unitary matrices $\widehat{\boldsymbol{U}}_g$:

$$\bigcap_{g=1}^{G} \mathrm{Span}(u_{g,1},\ldots,u_{g,K_g}) = C, \tag{21}$$

where $u_{g,k}$ is the kth column vector of $\widehat{\boldsymbol{U}}_g$. In what follows, we explain how to find the κ orthogonal vectors spanning the subspace C with a minimal number of operations and κ unknown. The process will rely on the proof of the existence of several unitary and semi-unitary matrices, that will not be explicitly written.

From Eqs. (20) and (21) it follows that, by means of simple rotations of the K_g vectors $\{u_{g,k}\}$, it is possible to find a basis of κ vectors, common to the G subspaces. Therefore, if we stack the G matrices $\tilde{\boldsymbol{U}}_g$ ($g = 1,\ldots,G$) in a matrix $\boldsymbol{U}_{\mathrm{tot}} \in \mathbb{C}^{L\times\sum_{g=1}^{G}K_g}$:

$$\boldsymbol{U}_{\mathrm{tot}} = [\widehat{\boldsymbol{U}}_1|\widehat{\boldsymbol{U}}_2|\cdots|\widehat{\boldsymbol{U}}_G], \tag{22}$$

we can always find a set of G unitary matrices $\boldsymbol{Q}_g \in \mathbb{R}^{K_g\times K_g}$ ($g = 1,\ldots,G$) such that

$$\left[\, \widehat{\boldsymbol{U}}_1\boldsymbol{Q}_1 \,\middle|\, \cdots \,\middle|\, \widehat{\boldsymbol{U}}_G\boldsymbol{Q}_G \,\right] = \left[\, \widehat{\boldsymbol{U}}_1 \,\middle|\cdots\middle|\, \widehat{\boldsymbol{U}}_G \,\right]\boldsymbol{Q} = \left[\, \boldsymbol{U} \,\middle|\, \boldsymbol{U}_1^{\perp} \,\middle|\cdots\middle|\, \boldsymbol{U} \,\middle|\, \boldsymbol{U}_G^{\perp} \,\right] \tag{23}$$

where $\boldsymbol{Q} = \mathrm{diag}\{\boldsymbol{Q}_1,\ldots,\boldsymbol{Q}_G\}$, $\boldsymbol{U} \in \mathbb{C}^{L\times\kappa}$ contains the set of column vectors spanning the common subspace, and $\boldsymbol{U}_g^{\perp} \in \mathbb{C}^{L\times(K_g-\kappa)}$ span the orthogonal complement of $\boldsymbol{U}$ in the column space of $\widehat{\boldsymbol{U}}_g$. By means of a permutation matrix $\boldsymbol{P} \in \mathbb{R}^{\sum_{g=1}^{G}K_g\times\sum_{g=1}^{G}K_g}$, the matrices $\boldsymbol{U}$ can be moved to the left:

$$\left[\, \boldsymbol{U} \,\middle|\, \boldsymbol{U}_1^{\perp} \,\middle|\cdots\middle|\, \boldsymbol{U} \,\middle|\, \boldsymbol{U}_G^{\perp} \,\right] = \left[\, \underbrace{\boldsymbol{U} \,\middle|\cdots\middle|\, \boldsymbol{U}}_{G\text{ times}} \,\middle|\, \underbrace{\boldsymbol{U}_1^{\perp} \,\middle|\cdots\middle|\, \boldsymbol{U}_G^{\perp}}_{\boldsymbol{U}^{\perp}} \,\right]\boldsymbol{P} \tag{24}$$

Note that $\boldsymbol{U}^{\perp} \in \mathbb{C}^{L\times\sum_{g=1}^{G}(K_g-\kappa)}$ is a full rank matrix with semi-unitary blocks that are mutually not necessarily orthogonal. If the SVD of $\boldsymbol{U}^{\perp}$ is given by $\boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{W}^{\mathsf{H}}$, then we obtain

$$\left[\, \boldsymbol{U} \,\middle|\cdots\middle|\, \boldsymbol{U} \,\middle|\, \boldsymbol{U}^{\perp} \,\right] = \boldsymbol{U}\cdot\sqrt{G}\,\boldsymbol{I}\cdot\Big[\underbrace{\boldsymbol{I}/\sqrt{G} \,\middle|\cdots\middle|\, \boldsymbol{I}/\sqrt{G}}_{\text{rows normalisation}} \,\Big|\, \mathbf{0}_1\Big] + \boldsymbol{V}\cdot\Sigma\cdot\Big[\, \underbrace{\mathbf{0}_2 \,\middle|\cdots\middle|\, \mathbf{0}_2}_{\mathbf{0}_3} \,\Big|\, \boldsymbol{W}^{\mathsf{H}} \,\Big] \tag{25}$$

where $\boldsymbol{I}$ is a $(\kappa \times \kappa)$-identity matrix, $\mathbf{0}_1 \in \mathbb{C}^{\kappa \times \sum_{g=1}^{G}(K_g-\kappa)}$, $\mathbf{0}_2 \in \mathbb{C}^{\sum_{g=1}^{G}(K_g-\kappa)\times\kappa}$, $\mathbf{0}_3 \in \mathbb{C}^{\sum_{g=1}^{G}(K_g-\kappa)\times G.\kappa}$, $\boldsymbol{V} \in \mathbb{C}^{L\times\sum_{g=1}^{G}(K_g-\kappa)}$ is a semi-unitary matrix, $\boldsymbol{\Sigma} \in \mathbb{R}^{\sum_{g=1}^{G}(K_g-\kappa)\times\sum_{g=1}^{G}(K_g-\kappa)}$ is a diagonal matrix, and $\boldsymbol{W}^{\mathsf{H}} \in \mathbb{C}^{\sum_{g=1}^{G}(K_g-\kappa)\times\sum_{g=1}^{G}(K_g-\kappa)}$ is a unitary matrix. Combining (23)–(25) finally yields

$$\underbrace{\left[\,\boldsymbol{U}\,\middle|\,\boldsymbol{V}\,\right]}_{\widetilde{\boldsymbol{U}}}\underbrace{\begin{bmatrix}\sqrt{G}\boldsymbol{I} & 0\\ 0 & \boldsymbol{\Sigma}\end{bmatrix}}_{\widetilde{\boldsymbol{\Sigma}}}\underbrace{\begin{bmatrix}\boldsymbol{I}/\sqrt{G} & \cdots & \boldsymbol{I}/\sqrt{G} & 0_1\\ & 0_3 & & \boldsymbol{W}^{\mathsf{H}}\end{bmatrix}}_{\widetilde{\boldsymbol{W}}^{\mathsf{H}}}\boldsymbol{P}\boldsymbol{Q}^{\mathsf{T}} \tag{26}$$

where $\widetilde{\boldsymbol{U}}$ and $\widetilde{\boldsymbol{W}}^{\mathsf{H}}$ are semi-unitary, $\widetilde{\boldsymbol{\Sigma}}$ is diagonal and the diagonal core matrix $\boldsymbol{\Sigma} = \mathrm{diag}\{\sigma_1,\ldots,\sigma_{\sum(K_g-\kappa)}\}$ is such that

$$\sqrt{G} > \sigma_1 \geqslant \cdots \geqslant \sigma_{\sum(K_g-\kappa)}. \tag{27}$$

Thus, we can draw a set of preliminary conclusions. First, it is clear that if there are κ poles common to the G channels, there exists a column subspace common to the G matrices $\boldsymbol{H}_g$ and therefore the $L \times \sum_g^G K_g$ matrix $\boldsymbol{U}_{\mathrm{tot}}$ is of rank $\min(L, \sum_g^G K_g - (G-1)\kappa)$. Second, it is also clear that

$$\widetilde{\boldsymbol{U}}_{\mathrm{tot}} = \boldsymbol{U}\cdot\sqrt{G}\,\boldsymbol{I}\cdot\left[\,\boldsymbol{I}/\sqrt{G}\;\cdots\;\boldsymbol{I}/\sqrt{G}\,\middle|\;0_1\,\right]\boldsymbol{P}\boldsymbol{Q}^{\mathsf{T}} \tag{28}$$

minimizes the function

$$\|\boldsymbol{U}_{\mathrm{tot}} - \widetilde{\boldsymbol{U}}_{\mathrm{tot}}\|^2 \tag{29}$$

with

$$\|\widetilde{\boldsymbol{U}}_{\mathrm{tot}}\|^2 = \kappa.G, \tag{30}$$

and hence a set of column vectors spanning the common subspace, and from which the common signal poles can be obtained, can be found in the best rank-κ approximation of $\boldsymbol{U}_{\mathrm{tot}}$. Finally, the κ dominant singular values of $\boldsymbol{U}_{\mathrm{tot}}$ are all equal to $\sqrt{G}$.

3.3.1. Discussion

In practice the data might be corrupted by noise, and the orders K_g might not be known. For the estimation of K_g one can refer to [11,12]. First, it is a fact that the level of noise must allow the possibility of properly separating the noise and the signal subspace. With respect to this constraint for a given matrix $\boldsymbol{\Sigma}_g$, it is possible to define $\varepsilon \in \mathbb{R}$ such that

$$\begin{aligned}&\sigma_i \leqslant \varepsilon, \quad i = K_g+1,\ldots,M,\\ &\sigma_i > \varepsilon, \quad i = 1,\ldots,K_g.\end{aligned} \tag{31}$$

When the model orders are slightly overestimated, a good estimation of the common subspace is usually still possible. Indeed, if for some of the matrices $\boldsymbol{H}_g$ an additional left singular vector u_{g,K_g+1} is taken into account, the lack of structure in these additional vectors ensures that the principle of derivation above still holds. Of course, when more vectors are added, the probability of creating an artificial intersection increases.

On the other hand, the influence of the noise tends to decrease the singular values that are theoretically equal to $\sqrt{G}$. Consequently, one has to fix a threshold $\delta \in \mathbb{R}$ which can be defined such that

$$\tilde{\sigma}_1,\ldots,\tilde{\sigma}_\kappa \geqslant \sqrt{G} - \delta, \tag{32}$$

$\tilde{\sigma}_1,\ldots,\tilde{\sigma}_\kappa$ being the set of singular values corresponding to the common subspace.

The matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ must be interpreted with care. The columns of $\boldsymbol{U}$ span the subspace in which lie the *common* Vandermonde vectors. The columns of $[\boldsymbol{U}\,\boldsymbol{V}]$ span the subspace in which lie *all* the Vandermonde vectors, the column span of $\boldsymbol{V}$ contains only a component of the non-common poles.

3.3.2. Complexity

The computational complexity of this algorithm is determined by the computation of the SVD. Using the standard Golub–Reinsch SVD [9] algorithm, the cost for an SVD of an $L \times M$ matrix, including the computation of the three matrices $\boldsymbol{U}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{V}$, is $4L^2M + 8LM^2 + 9M^3$ floating point operations (flops) [9, p. 254]. In Step 2 of the algorithm we typically take an

approximately square Hankel matrix [1], i.e $L = M$ if N is odd and $L = M + 1$ if N is even. Let us take $L = M$ for simplicity. In this case each of the SVDs requires $O(L^3)$ flops; the total cost is $O(GL^3)$. In Step 4 of the algorithm, we compute the SVD of an $L \times \Sigma_{g=1}^{G} K_g$ matrix with $L \geqslant \Sigma_{g=1}^{G} K_g$. Therefore the computational complexity is $O(L^2.\Sigma_{g=1}^{G} K_g)$. Step 7 of the algorithm involves the SVD of an $(L-1) \times 2K_c$ matrix, as well as the eigenvalue decomposition of a $K_c \times K_c$ matrix. Since $K_c \ll (L-1) \sim L$ the complexity of Step 7 is $O(L^2 K_c)$.

As discussed earlier in the paper there is no competing approach in the multi-channel case ($G > 2$). We can, however, compare the computational complexity of our algorithm with the one of HTLS-SEP in the two-channel case. The latter algorithm requires the SVD of two $L \times M$ matrices, the SVD of an $L \times 2M$ matrix and finally the SVD of an $L \times M$ matrix in order to estimate the common subspace. The method used for finding the signal poles is similar to our algorithm. Note that HTLS-SEP also requires the Hankel matrices to be approximately square ($L = M$). We see that HTLS-SEP requires one more SVD than our technique, namely the SVD of an $L \times 2M$ matrix. Moreover, more matrix manipulations and multiplications are necessary. In conclusion, although both algorithms have roughly the same complexity ($O(L^3)$), HTLS-SEP is always slower than our new algorithm.

4. Application to simulated signals

In this section we illustrate the performance of the algorithm and we assess the reliability of the common subspace estimate. Complex circular symmetric white Gaussian noise (WGN) is added to the signal in each channel. We conduct Monte Carlo experiments consisting of $N_{\text{runs}} = 500$ runs. The performance is mainly evaluated in terms of the *relative root mean squared error* (RRMSE), defined by

$$\text{RRMSE}(\gamma) = \left[\frac{1}{N_{\text{runs}}} \sum_{i=1}^{N_{\text{runs}}} |\gamma - \hat{\gamma}_i|^2\right]^{1/2} \frac{100}{\gamma} [\%], \tag{33}$$

where $\hat{\gamma}_i$ is the estimate of parameter γ, obtained in the ith run. We call a weighted complex discrete-time exponential $c_k z_k^n$ a peak, simply because in the frequency domain $c_k z_k^n$ has a bell shape. The number of peaks in the frequency domain is usually equal to the number of signal poles. As performance measure, we will also use the root mean squared distance between two parameters $\hat{\gamma}$ and $\hat{\eta}$ defined as follows:

$$\text{d}(\hat{\gamma}, \hat{\eta}) = \left[\frac{1}{N_{\text{runs}}} \sum_{i=1}^{N_{\text{runs}}} |\hat{\gamma}_i - \hat{\eta}_i|^2\right]^{1/2}, \tag{34}$$

where $\hat{\gamma}_i$ (resp. $\hat{\eta}_i$) is the estimate of parameter γ (resp. η), obtained in the ith run. If $\hat{\gamma} = \gamma$ is not an estimate but a fixed value, then $\text{d}(\gamma, \hat{\eta})$ measures the deviation of the estimate $\hat{\eta}$ from the fixed value γ. Therefore, $\text{d}(\eta, \hat{\eta})$ equals the *root mean squared error* (RMSE) of the parameter η.

4.1. A five-channel example

In this example we consider five signals, having two poles in common. Each signal is of order three, which means that there is one perturbing discrete-time complex exponential per channel. Fig. 1 shows the signals in the frequency domain. Their relevant parameters are given in Table 1. All the phases φ are drawn from a uniform distribution $[0, 2\pi]$, and the amplitudes are perturbed by a WGN with standard deviation 0.1. Note that the third channel contains two closely spaced peaks. Applying steps 1–6 of our algorithm for different levels of noise yields the results shown in Fig. 2. The simulation parameters were set as follows: $N = 128$, $L = 65$, $M = 64$, $\text{G} = 5$, $\{K_g\}_{g=1}^{5} = \{3, 3, 3, 3, 3\}$ and $\kappa = 2$. The subspace of interest spanned by all the Vandermonde vectors has dimension $\sum K_g - (G-1).\kappa = 7$. The $\kappa = 2$ first singular values of $\boldsymbol{U}_{\text{tot}}$ correspond to the common subspace and the following $\sum K_g - G\kappa = 5$ singular values correspond to a subspace in which the non-common Vandermonde vectors are projected. Finally, the remaining $(G-1)\kappa = 8$ singular values of $\boldsymbol{U}_{\text{tot}}$ correspond to the noise subspace. The singular values of the matrix $\boldsymbol{U}_{\text{tot}}$ exhibit a quite robust behavior. First, we notice that, over a broad range of noise variance, there is a clear gap between the smallest singular value related to the

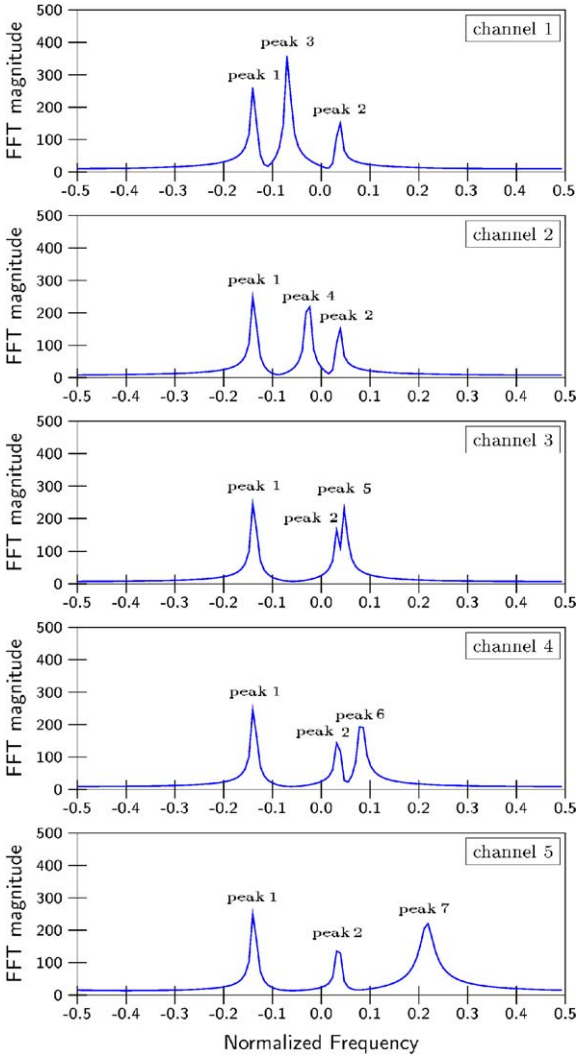


Fig. 1. Frequency domain signals in the five-channel case.

Table 1
Set of parameters for the five-channel example

		ν (Hz)	α (s^{-1})	a (a.u.)
Channel 1	Peak 1	−1379	208	6.1
	Peak 2	−353	117	2.8
	Peak 3	−685	256	9.9
Channel 2	Peak 1	−1379	208	6.1
	Peak 2	−353	117	2.8
	Peak 4	−271	197	6.0
Channel 3	Peak 1	−1379	208	6.1
	Peak 2	−353	117	2.8
	Peak 5	478	208	5.0
Channel 4	Peak 1	−1379	208	6.1
	Peak 2	−353	117	2.8
	Peak 6	815	319	7.3
Channel 5	Peak 1	−1379	208	6.1
	Peak 2	−353	117	2.8
	Peak 7	2162	808	17

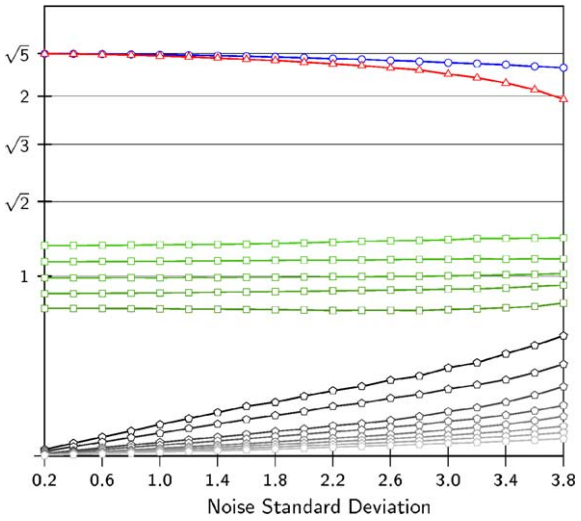


Fig. 2. The $\sum_g^G K_g = 15$ singular values of $\boldsymbol{U}_{\text{tot}}$ in the five-channel case. The two top singular values (○, △) are related to the common subspace. The five following singular values (□) are related to the subspace in which lie a projection of the undesired (not common) poles. The last eight singular (⬠) are related to the noise subspace.

common subspace and the largest one related to the orthogonal subspace. The two largest singular values deviate very slowly from their initial value ($\sqrt{5}$). Moreover, the five next singular values keep an almost constant value for any level of noise. Finally, it is interesting to see that only the last singular values increase proportionally to the noise standard deviation. The best performance can be expected when the number of common poles and the number of perturbing harmonics are small. A noise standard deviation equal to 3.4 corresponds on average to a peak SNR of −5.2 dB for the smallest common peak and a peak SNR of 1.57 dB for common peak 1. At this noise level, the smallest singular value related to the

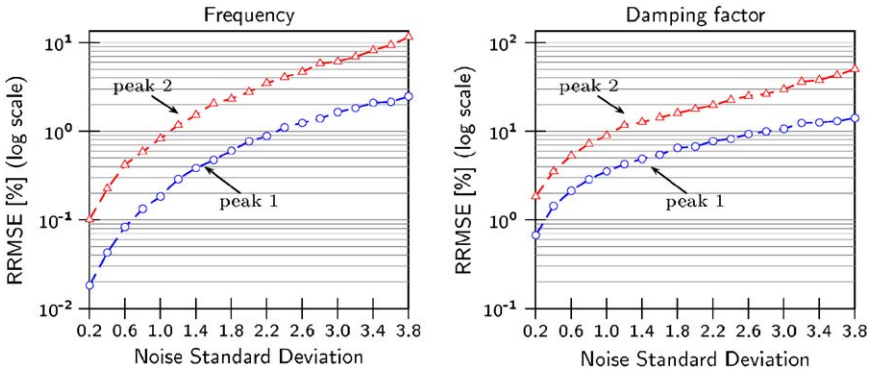


Fig. 3. Estimation of the common frequencies and damping factors in the five-channel case.

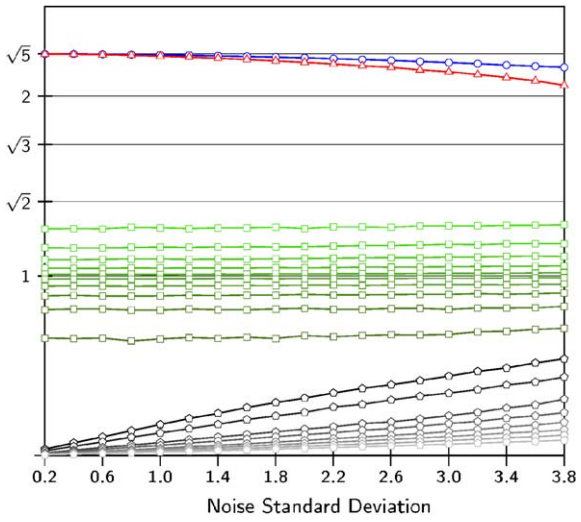


Fig. 4. The $\sum_{g=1}^{G} K_g = 20$ singular values of $\boldsymbol{U}_{\text{tot}}$ in the five-channel case with an overestimation of one pole for each channel. The two top singular values ($\circ, \triangle$) are related to the common subspace. The five following singular values ($\square$) are related to the subspace in which a projection of the undesired (not common) poles is lying. The last eight singular values ($\pentagon$) are related to the noise space.

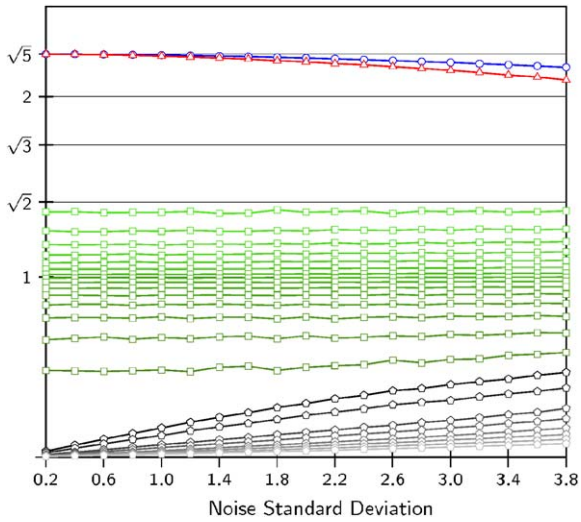


Fig. 5. The $\sum_{g=1}^{G} K_g = 25$ singular values of $\boldsymbol{U}_{\text{tot}}$ in the five-channel case with an overestimation of two poles for each channel. The two top singular values ($\circ, \triangle$) are related to the common subspace. The five following singular values ($\square$) are related to the subspace in which a projection of the undesired (not common) poles is lying. The last eight singular values ($\pentagon$) are related to the noise space.

commonsubspace crosses the value $\sqrt{4}$. At this point we must have the prior knowledge that there are no poles common to four channels. Fig. 3 shows the estimates of the frequencies and damping factors of the common poles. We conclude that, over a wide SNR range, (1) it is possible to separate the signal subspace, spanned by all Vandermonde vectors from the noise subspace, (2) it is possible to extract the common subspace from the signal subspace, and estimate the common poles. Our simulation consisted of 500 noise realizations per noise level, but actually the estimates are already very smooth for a few runs (i.e. 50–100), which means that this method does not suffer from outliers.

In most real-world situations it is very difficult to determine the exact number of signal poles for each channel. Therefore we test the robustness of the algorithm against an overestimation of the model orders. In Figs. 4 and 5 we plot the singular values of the matrix $\boldsymbol{U}_{\text{tot}}$ in the case where the model order of each channel has been overestimated with, respectively, one and two poles. Except for the overestimation of K_g, the simulation parameters remained the same. We see that there is still a gap between the singular values related to the common subspace and the other singular values. We can observe two main effects. First, as explained in Section 3.3.1, by overestimating model orders, we include some vectors that are orthogonal to the common subspace, such that a smaller part of the data is projected away. This increases the probability of obtaining a higher-dimensional "common" subspace. This is reflected by an increase of the biggest singular values related to the non-common subspace. On the other hand, the overestimation helps to catch more accurately the true common subspace. Indeed, we can see in Figs. 4 and 5 that the singular values related to the common subspace are slightly closer to $\sqrt{5}$ than in the case without overestimation. This result is confirmed by Figs. 6 and 7, which show the estimates of the common frequencies and damping factors.

4.2. A two-channel example

We also compare the new technique with the HTLS-SEP algorithm for two-channel data [10].

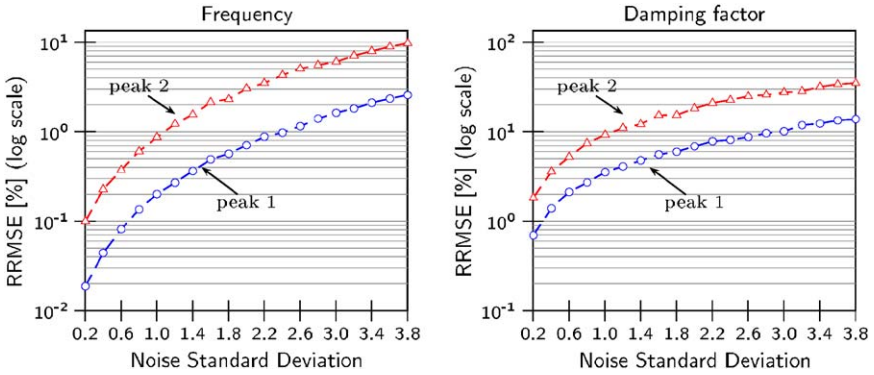


Fig. 6. Estimation of the common frequencies and damping factors in the five-channel case. The model order of each channel has been increased by 1.

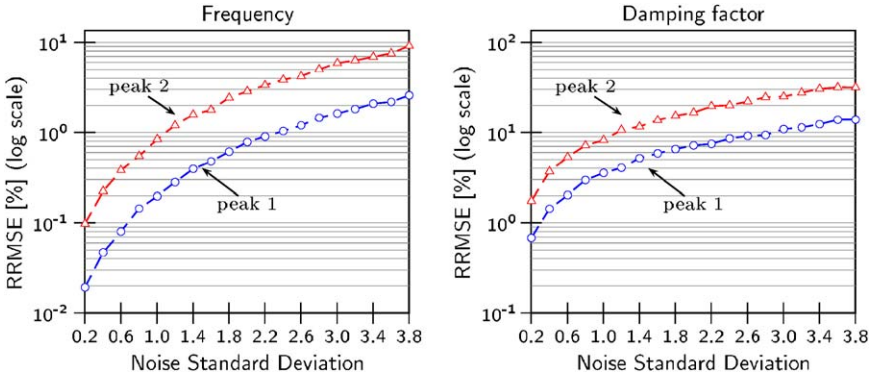


Fig. 7. Estimation of the common frequencies and damping factors in the five-channel case. The model order of each channel has been increased by 2.

The two signals consist of two complex exponentials, one exponential being common to both signals. The parameters of the three peaks are given in Table 2. The sampling frequency is $v_s = 10^4$ Hz and the number of samples for each signal is $N = 128$. The frequency domain representation of this set of two signals is given in Fig. 8. The parameters of the data are: $L = 65$, $M_g = 64$, $G = 2$, $K_1 = 2$, $K_2 = 2$ and $\kappa = 1$. Thus, as shown in Fig. 9, the first singular value σ_1 slowly deviates from its theoretical value, which is $\sqrt{2}$ according to (26), until the noise standard deviation reaches 2. Within this range, σ_1 stays far from σ_2, the largest singular value related to a subspace that is not common. It is clearly possible to detect that there is only one pole. The estimates of its frequency and damping factor are shown in Fig. 10. The results of our new method are consistently better than the results of HTLS-SEP. First, the frequency estimates computed by HTLS-SEP become inconsistent when the noise standard deviation reaches 1.1, while for the new method, this happens when the noise standard deviation reaches 1.5. Second, in the noise range $[0.1 - 1.1]$, where both methods are consistent, the estimates of the frequency and the damping factor are about 30.39% more accurate for the new algorithm. The main reason is that we work with normalized subspaces, which decreases the influence of the amplitudes of the peaks. This confer to the algorithm a very high robustness.

Table 2
Set of parameters for a three-peak example

	v (Hz)	α (s^{-1})	a (a.u)	φ (deg)
Peak 1	−1379	208	2.5	15
Peak 2	−685	256	9.9	15
Peak 3	−271	197	6.0	15

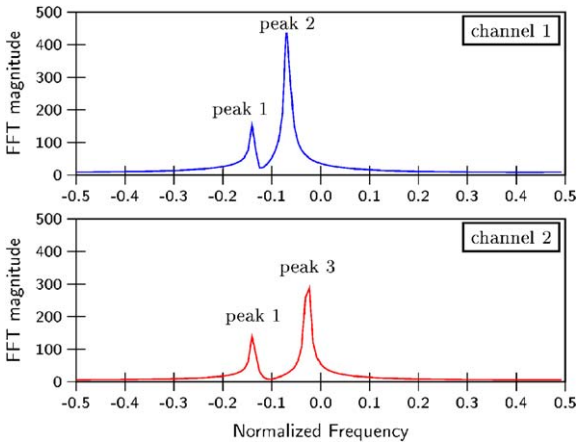


Fig. 8. Frequency domain representation of the two signals, in the two-channel example. "Peak 1" is the common peak.

5. Conclusion

In this paper, we have studied a novel subspace-based method for common pole estimation in the

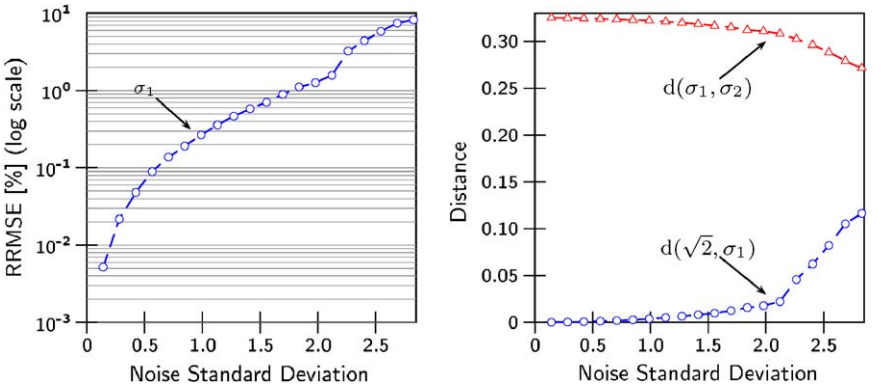


Fig. 9. The left plot shows the RRMSE of the singular value related to the common signal pole in the two-channel case. The right plot shows the deviation of σ_1 from $\sqrt{2}$ (○) and the average distance from σ_1 to σ_2 (△).

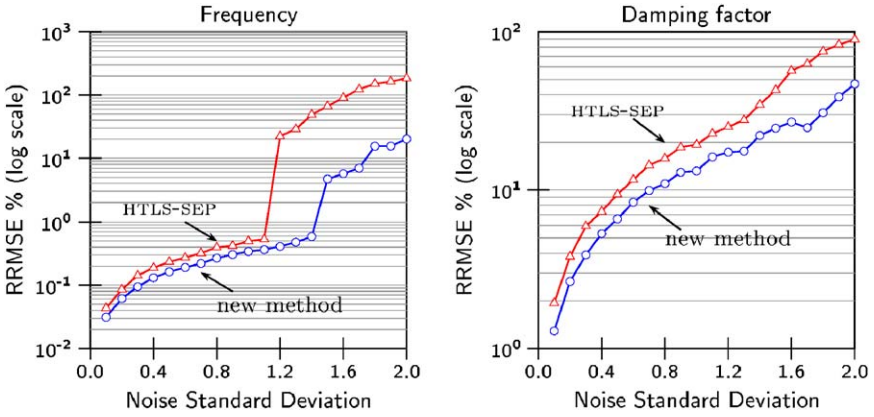


Fig. 10. Estimates of the frequency and damping factor of the common pole in the two-channel case.

multi-channel case. We have proven that the use of semi-unitary matrices allows one to determine the number of common poles and get an accurate estimate of a subspace from which this these poles can be computed. This method is easy to implement, very robust and can be applied to an arbitrary number of channels. Moreover, we have shown that a slight overestimation of the individual model orders does not affect the accuracy of the common pole estimation. In the particular case in which only two channels are involved, this technique turns out to be more efficient than the HTLS-SEP technique developed for the two-channel case.

Acknowledgements

The research reported in this paper is supported by several institutions: (1) the Flemish Government: (a) Research Council K.U. Leuven: GOA-MEFISTO-666, GOA-Ambiorics, IDO/99/003 and /02/009, (b) F.W.O. project G.0240.99, G.0078.01, G.0407.02, G.0269.02, G.0270.02 (c) F.W.O. Research Communities ICCoS and ANMMM, (d) Tournesol project T2004.13, (2) the Belgian Federal Science Policy Office: IUAP P5/22, (3) EU: PDT-COIL, BIOPATTERN and eTUMOUR. Lieven De Lathauwer holds a permanent research position with the French Centre National de la Recherche Scientifique (C.N.R.S.). He also holds a honorary research position with the K.U. Leuven, Belgium. Sabine Van Huffel is Full Professor with the K.U. Leuven.

References

[1] S. Van Huffel, Enhanced resolution based on minimum variance estimation and exponential data modeling, Signal Processing 33 (3) (1993) 333–355.
[2] S. Van Huffel, H. Chen, C. Decanniere, P. Van Hecke, Algorithm for time-domain NMR data fitting based on total least squares, J. Magnetic Resonance, Series A 110 (1994) 228–237.
[3] H. Chen, S. Van Huffel, J. Vandewalle, Bandpass prefiltering for exponential data fitting with known frequency region of interest, Signal Processing 48 (1996) 135–154.
[4] L. Rippert, Optical fiber for damage monitoring in carbon fiber reinforced plastic composite materials, Ph.D. Thesis, Katholieke universiteit Leuven, March 2003.
[5] R. Boyer, K. Abed-Meraim, Audio modeling based on delayed sinusoids, IEEE Trans. Speech Audio Process. 12 (1) (January 2004).
[6] R. Roy, T. Kailath, Esprit-estimation of signal parameters via rotational invariance techniques, IEEE Trans. Acoust. Speech Signal Process. 37 (1989) 984–995.
[7] S.-Y. Kung, K.S. Arun, D.V. Bhaskar Rao, State-space and singular value decomposition-based approximation methods for the harmonic retrieval problem, J. Opt. Soc. Amer. 73 (12) (1983) 1799–1811.

[8] L. Vanhamme, S. Van Huffel, Multichannel quantification of biomedical magnetic resonance spectroscopy signals, in: F.T. Luk, (Ed.), Proceedings of SPIE, Advanced Signal Processing Algorithms, Architectures, and Implementations VIII, vol. 3461, San Diego, California, 19–24 July 1998, pp. 237–248.

[9] G.H. Golub, C.F. Van Loan, Matrix Computations, third ed., The Johns Hopkins University Press, Baltimore, MD, 1996.

[10] G. Morren, P. Lemmerling, S. Van Huffel, G. Naulaers, H. Devlieger, P. Casaer, Detection of autoregulation in the brain of premature infants using a novel subspace-based technique, in: Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC2001), Istanbul, Turkey, October 2001.

[11] D. Kundu, A. Mitra, Estimating the number of signals of the damped exponential models, Comput. Statist. Data Anal. 36 (2) (April 2001) 245–256.

[12] A.-K. Seghouane, D.L. Lieven, A bootstrap model selection criterion based on kullback's symmetric divergence, in: Proceedings of the IEEE Workshop on Statistical Signal Processing (SSP'03), St. Louis, Missouri, September 2003.