Blind Source Separation of Single Channel Mixture Using Tensorization and Tensor Diagonalization

Anh-Huy Phan[1], Petr Tichavský[2(✉)], and Andrzej Cichocki[1,3]

[1] Lab for Advanced Brain Signal Processing, Brain Science Institute - RIKEN, Wako, Japan
[2] Institute of Information Theory and Automation, Prague, Czech Republic
tichavsk@utia.cas.cz
[3] Skolkovo Institute of Science and Technology (SKOLTECH), Moscow, Russia

Abstract. This paper deals with estimation of structured signals such as damped sinusoids, exponentials, polynomials, and their products from single channel data. It is shown that building tensors from this kind of data results in tensors with hidden block structure which can be recovered through the tensor diagonalization. The tensor diagonalization means multiplying tensors by several matrices along its modes so that the outcome is approximately diagonal or block-diagonal of 3-rd order tensors. The proposed method can be applied to estimation of parameters of multiple damped sinusoids, and their products with polynomial.

Keywords: Blind source separation · Block tensor decomposition · Tensor diagonalization · Three-way folding · Three-way toeplitzation · Damped sinusoids

1 Introduction

Separation of hidden components or sources from mixtures with one or a few sensors appears in many real problems in signal processing. So far, there is no straightforward method for separation of arbitrary signals. However, the problem can be tackled in some cases when strong assumptions have been imposed onto the components or the data, e.g., non-negativity together with non-overlapping (i.e., orthogonality) as in nonnegative matrix factorisation [1,2], and statistical independence as in independent component analysis [3]. In some cases, the signal components of interest have some specific structures. For example, separation of the damped sinusoids appears in wide range of applications, e.g., electrical, mechanical, electromechanical, geophysical, chemical. The traditional methods for estimating damped sinusoid parameters are based on the linear self-prediction (auto regression) using signal samples, where the Prony least squares

The work of P. Tichavsky was supported by The Czech Science Foundation through Projects No. 14-13713S and 17-00902S.

P. Tichavský et al. (Eds.): LVA/ICA 2017, LNCS 10169, pp. 36–46, 2017.
DOI: 10.1007/978-3-319-53547-0_4

autoregressive model fitting, and the Pade approximation procedure are two well-known methods [4]. The other algorithmic families are based the signal subspace method, which computes self-linear prediction coefficients as eigenvectors of autocorrelation matrix of the measurement, e.g., Pisarenko, MUSIC, ESPRIT, or the singular value decomposition (SVD) of the Hankel-type matrix, e.g., the Kumaresan-Tufts and Matrix Pencil methods [5,6]. Extension of the methods to tensor decomposition can be found in [7–10].

Alternatively, another type of methods has been developed for this BSS problem, based on exploiting the structure of the components rather based on their statistical properties. More specifically, when the extracted components can be expressed in some low rank formats, e.g., low rank matrix or tensor, low multilinear rank tensor, by tensorization of the mixture, one can apply appropriate tensor decomposition methods to retrieve hidden factor matrices, which are then used to reconstruct the original sources [11,12]. De Lathauwer proposed to use the Hankelization to convert signals to be low-rank matrices, and resort the BSS problem to the canonical tensor decomposition or the $(L, L, 1)$-block term decomposition [11,12]. An advantage of this approach is that it allows to separate signals with relatively short length samples, even with dozens of samples.

In the same direction with the latter method, the method proposed in this paper first tensorizes the mixtures, and converts the BSS problem of R sources into block tensor decompositions (BTD) of R block terms, each term corresponds to a source. Tensorization for single channel data can be the simple re-shaping (folding), Toeplitzation or Hankelization. The blocks structure can be revealed through the tensor diagonalization technique, which is explained in an accompanying papers [13]. The proposed method not only separate sinusoids as, e.g., [14], but also allows to estimate a wider class of signals, e.g., products of polynomials and damped sinusoids. In addition, the proposed algorithm works even when the number of samples is small, or when the number of sources becomes significantly large.

In the simulation section, we test the proposed method and compare its performance with state-of-the-art methods of estimating parameters of damped sinusoids. We also show one example to which the damping estimation methods are not applicable.

2 General Framework for BSS of Single Channel Mixture

Consider a single mixture $y(t)$ which is composed of R component signals $x_r(t)$, $r = 1, \ldots, R$, and corrupted by additive Gaussian noise $e(t)$

$$y(t) = a_1 x_1(t) + a_2 x_2(t) + \ldots + a_R x_R(t) + e(t). \tag{1}$$

The aim is to extract the unknown sources (components) $x_r(t)$ from the observed signal $y(t)$. For this problem, we first apply some linear tensorization to the source vectors $\boldsymbol{x}_r = [x_r(1), \ldots, x_r(t), \ldots]$, in order to obtain 3-rd order tensors $\mathcal{X}_r$, which are assumed to be low rank or low multilinear rank tensor, that is,

$$\mathcal{X}_r = \mathcal{G}_r \times_1 \mathbf{U}_r \times_2 \mathbf{V}_r \times_3 \mathbf{W}_r,$$

where $\mathcal{G}_r$ are core tensors of size $R_1 \times R_2 \times R_3$. Such tensorizations of interest can be reshaping, which is also known as segmentation [15,16], Hankelization [11], and Toeplization. From the mixing model in (1), and due to linearity of the tensorization, we have the following relation between tensorization $\mathcal{Y}$ of the mixture and those of the hidden components

$$\begin{aligned} \mathcal{Y} &= a_1 \mathcal{X}_1 + a_2 \mathcal{X}_2 + \ldots + a_R \mathcal{X}_R + \mathcal{E} \\ &= \sum_{r=1}^{R} (a_r \mathcal{G}_r) \times_1 \mathbf{U}_r \times_2 \mathbf{V}_r \times_3 \mathbf{W}_r + \mathcal{E} \end{aligned} \tag{2}$$

where $\mathcal{E}$ is tensorization of the noise $e(t)$.

Now, by decomposition of $\mathcal{Y}$ into R blocks, i.e., BTD with R block terms, we can find approximations of $\mathcal{X}_r$ up to scaling. Finally, reconstruction of signals $\hat{x}_r(t)$ from $\mathcal{X}_r$ can be done straightforwardly because the tensorization is linear. Instead of applying the block term decomposition, we address the above tensor model as a tensor diagonalization, where each block of the core tensor corresponds to a hidden source.

For the multi-channel BSS, $y_m(t) = \sum_r a_{kr} x_r(t)$, channel mixtures $\boldsymbol{y}_m = [y_m(1), \ldots, y_m(t), \ldots]$, are tensorized separately, then all together they construct a tensor of order-4 which admits the BTD-$(R_1, R_2, R_3, 1)$

$$\mathcal{Y} = \sum_{r=1}^{R} \mathcal{G}_r \times_1 \mathbf{U}_r \times_2 \mathbf{V}_r \times_3 \mathbf{W}_r \times_4 \boldsymbol{a}_r + \mathcal{E} \tag{3}$$

where $\boldsymbol{a}_r$ are column vectors of the mixing matrix.

We are particular interested in the sinusoid signals and its modulated variant, e.g., the exponentially decaying signals

$$x(t) = \exp(-\gamma t) \sin(\omega\, t + \phi)\,, \tag{4}$$

$$x(t) = t^n \sin(\omega\, t + \phi)\,, \qquad x(t) = t^n \exp(-\gamma t)\,, \tag{5}$$

for $t = 1, 1, \ldots, L$, $\omega \neq 0$, $n = 1, 2, \ldots$.

In the next sections, we present tensorizations to yield low-rank tensors from the above signals, and confirm their efficiency through examples for BSS of single mixture.

3 Tensorization of Sinusoid Signals

The tensorizations presented in this section are for the sinusoids but can also be applied to the other signals in (4)–(5) to yield tensors of multilinear rank-$(2, 2, 2)$, $(2(n+1), 2(n+1), 2(n+1))$ or $(n+1, n+1, n+1)$.

3.1 Two-Way and Three-Way Foldings

The simplest tensorization is reshaping (folding), which rearranges a vector to a matrix or tensor. This type of tensorization preserves the number of original data entries and their sequential ordering. It can be shown that reshaping of a sinusoid results a rank-2 matrix, or a multilinear rank-$(2, 2, 2)$ tensor.

Lemma 1 (Two-way folding). *A matrix of size $I \times J$ which is reshaped from a sinusoid signal $x(t)$ of length $L = IJ$ is of rank-2 and can be decomposed as*

$$\mathbf{Y} = \begin{bmatrix} y(1) & y(I+1) & \cdots & y(K-I+1) \\ y(2) & y(I+2) & \cdots & y(K-I+2) \\ \vdots & & \ddots & \vdots \\ y(I) & y(2I) & \cdots & y(K) \end{bmatrix} = \mathbf{U}_{\omega,I}\,\mathbf{S}\,\mathbf{U}_{\omega I,J}^T \tag{6}$$

where $\mathbf{S}$ is invariant to the folding size I, and depends only on the phase ϕ and takes the form

$$\mathbf{S} = \begin{bmatrix} \sin(\phi) & \cos(\phi) \\ \cos(\phi) & -\sin(\phi) \end{bmatrix}, \qquad \mathbf{U}_{\omega,I} = \begin{bmatrix} 1 & 0 \\ \vdots & \vdots \\ \cos(k\omega) & \sin(k\omega) \\ \vdots & \vdots \\ \cos((I-1)\omega) & \sin((I-1)\omega) \end{bmatrix}. \tag{7}$$

Lemma 2 (Three-way folding). *An order-3 tensor of size $I \times J \times K$, where $I, J, K > 2$, reshaped from a sinusoid signal of length $L = IJK$, can take a form of a multilinear rank-$(2, 2, 2)$ or rank-3 tensor*

$$\mathcal{Y} = \mathcal{H} \times_1 \mathbf{U}_{\omega,I} \times_2 \mathbf{U}_{\omega I,J} \times_3 \mathbf{U}_{\omega IJ,K} \tag{8}$$

where $\mathcal{H} = \mathcal{G} \times_3 \mathbf{S}$ is a small-scale tensor of size $2 \times 2 \times 2$, and

$$\mathcal{G}(:,:,1) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad \mathcal{G}(:,:,2) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{9}$$

Proof of Lemma 1 is obvious, while Lemma 2 can be deduced from Lemma 1 because the three-way folding can be performed over two foldings [17]. Note that in the complex field, the above tensors are of rank-2. An advantage of this tensorization is that it does not increase the number of samples; therefore, the method does not need extra space. However, when the signals are of short duration, the reshaped tensors are relatively small, and decomposition of these tensors does not give good approximation of the original sources. For such a case, tensorizations, which can increase the number of entries, e.g., the Toeplitzation and Hankelization, are recommended.

3.2 Toeplitzation

Definition 1 (Three-way Toeplitzation). *Tensorization of the signal $x(t)$ of length L to an order-3 tensor $\mathcal{X}$ of size $I \times J \times K$, where $I + J + K = L + 2$,*

$$\mathcal{X}(i, j, k) = x(I + J + k - i - j)\,. \tag{10}$$

For this tensorization, each horizontal slice $\mathcal{X}(i, :, :)$ is a Toeplitz matrix composed of vectors $[x(I+1-i), \ldots, x(I+J-i)]$ and $[x(I+J-i), \ldots, y(I+J+K-1-i)]$.

Lemma 3. *Order-3 Toeplitz tensors tensorized from a sinusoid signal is of multilinear rank-$(2, 2, 2)$, and can be represented as*

$$\mathcal{X} = \frac{1}{\sin^3(\omega)} \mathcal{G} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \mathbf{U}_3 \tag{11}$$

where $\mathcal{G}$ is a tensor of size $2 \times 2 \times 2$

$$\mathcal{G}(:, :, 1) = \begin{bmatrix} \sin(\omega(I_1+2)+2\phi) & -\sin(\omega(I_1+1)+2\phi) \\ -\sin(\omega(I_1+1)+2\phi) & \sin(\omega I_1+2\phi) \end{bmatrix},$$
$$\mathcal{G}(:, :, 2) = \begin{bmatrix} -\sin(\omega(I_1+1)+2\phi) & \sin(\omega I_1+2\phi) \\ \sin(\omega I_1+2\phi) & -\sin(\omega(I_1-1)+2\phi) \end{bmatrix},$$

and the three factor matrices are given by

$$\mathbf{U}_1 = \begin{bmatrix} x(1) & x(2) \\ \vdots & \vdots \\ x(I) & x(I+1) \end{bmatrix}, \quad \mathbf{U}_2 = \begin{bmatrix} x(I) & x(I+1) \\ \vdots & \vdots \\ x(I+J-1) & x(I+J) \end{bmatrix}, \quad \mathbf{U}_3 = \begin{bmatrix} x(I+J-1) & x(I+J-2) \\ \vdots & \vdots \\ x(L) & x(L-1) \end{bmatrix}.$$

Proof. The proof can be seen from the fact that

$$\mathcal{G} \times_1 \begin{bmatrix} x(i) \; x(i+1) \end{bmatrix} = \sin(\omega) \begin{bmatrix} -x(I-i+3) & x(I-i+2) \\ x(I-i+2) & -x(I-i+1) \end{bmatrix} \tag{12}$$

and

$$[x(I+j-1), x(I+j)] \begin{bmatrix} -x(I-i+3) & x(I-i+2) \\ x(I-i+2) & x(I-i+1) \end{bmatrix} \begin{bmatrix} x(I+J+k-2) \\ x(I+J+k-3) \end{bmatrix}$$
$$= \sin^2(\omega)\, x(I+J+k-i-j).$$

□

In Appendix, we show that the three-way folding and Toeplitzation of $x(t) = t$ also yield multilinear rank-$(2, 2, 2)$ tensors or tensors of rank-3. A more general result is that the tensors of $x(t) = t^n$ have multilinear rank of $(n+1, n+1, n+1)$. The closed-form expressions of the low multilinear representations are given in Lemmas 5–8. In addition, we note that the above three-way tensorizations to sinusoids yield tensors $\mathcal{X}_{\sin}$ of multilinear rank-$(2, 2, 2)$, and tensorization of exponentially decaying signal $\exp(-\omega t)$ yields a rank-1 tensor $\mathcal{X}_{\exp}$. Following Lemma 4, the Hadamard product $\mathcal{X}_{\sin} \circledast \mathcal{X}_{\exp}$ remains a tensor of multilinear rank-$(2, 2, 2)$, implying that the damped sinusoids can be represented by multilinear rank-$(2, 2, 2)$ tensors. Similarly, the signal $t\exp(-\gamma t)$ in (5) has multilinear rank-$(2, 2, 2)$, and $t\sin(\omega\, t + \phi)$ in (5) has multilinear rank-(4, 4, 4).

Lemma 4 (Hadamard product of two tensors of low multilinear ranks). *Given two tensors of the same size, $\boldsymbol{\mathcal{X}}$ of multilinear rank-(R, S, T) and $\boldsymbol{\mathcal{Y}}$ of multilinear rank-(R', S', T'), the Hadamard product of them can be represented by a tensor which has multilinear rank at most (RR', SS', TT').*

Proof. We represent $\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{G}} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C}$, and $\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{H}} \times_1 \mathbf{U} \times_2 \mathbf{V} \times_3 \mathbf{W}$, where $\boldsymbol{\mathcal{G}}$ is a tensor of size $R \times S \times T$, and $\boldsymbol{\mathcal{H}}$ is of size $R' \times S' \times T'$. Then the Hadamard product $\boldsymbol{\mathcal{X}} \circledast \boldsymbol{\mathcal{Y}}$ yields a tensor, for which entries are defined as

$$\begin{aligned}\boldsymbol{\mathcal{X}}(i,j,k)\boldsymbol{\mathcal{Y}}(i,j,k) &= \left(\sum_{r,s,t} g_{r,s,t}\, a_{ir} b_{js} c_{kt}\right)\left(\sum_{r',s',t'} h_{r',s',t'}\, u_{ir'} v_{js'} w_{kt'}\right) \\ &= \sum_{r,r'}\sum_{s,s'}\sum_{t,t'} (g_{r,s,t} h_{r',s',t'})\,(a_{ir}u_{ir'})(b_{js}v_{js'})(c_{kt}w_{kt'}) \\ &= \sum_{\bar{r}}^{RR'}\sum_{\bar{s}}^{SS'}\sum_{\bar{t}}^{TT'} z_{\bar{r},\bar{s},\bar{t}}\, d_{i,\bar{r}}\, e_{j,\bar{s}}\, f_{j,\bar{t}},\end{aligned}$$

where $\boldsymbol{\mathcal{Z}} = \boldsymbol{\mathcal{G}} \otimes \boldsymbol{\mathcal{H}}$ is of size $RR' \times SS' \times TT'$, and $\mathbf{D}$, $\mathbf{E}$ and $\mathbf{F}$ are row-wise Khatri-Rao products of the two corresponding factor matrices of $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}$, i.e., $\boldsymbol{d}_i = \boldsymbol{a}_i \otimes \boldsymbol{u}_i$, $\boldsymbol{e}_j = \boldsymbol{b}_j \otimes \boldsymbol{v}_j$, and $\boldsymbol{f}_j = \boldsymbol{c}_k \otimes \boldsymbol{w}_k$. In summary, we obtain

$$\boldsymbol{\mathcal{X}} \circledast \boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{Z}} \times_1 \mathbf{D} \times_2 \mathbf{E} \times_3 \mathbf{F}\,. \tag{13}$$

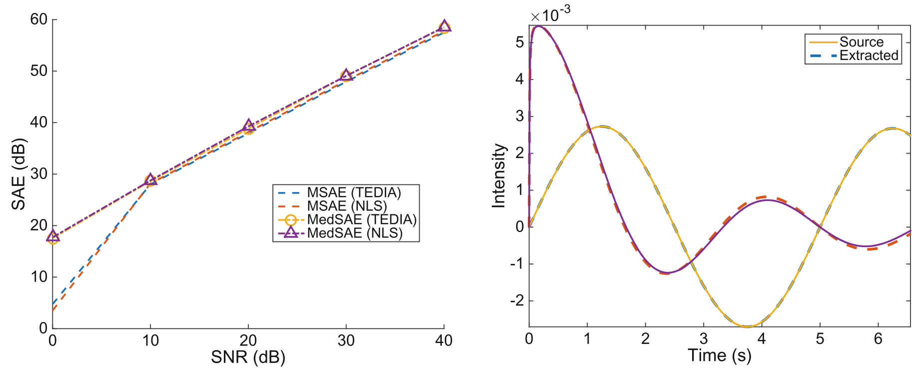
□

4 Simulations

Example 1. In this first example, we considered a signal of length $L = 414$, which were composed of two source signals $x_1(t)$ and $x_2(t)$ and corrupted by additive Gaussian noise $e(t)$

$$\begin{aligned}&y(t) = a_1 x_1(t) + a_2 x_2(t) + e(t), \\ &x_1(t) = \exp(\frac{-2t}{L}) \sin(\frac{2\pi f_1}{f_s} t), \qquad x_2(t) = t\, \exp(\frac{-4t}{L}),\end{aligned}$$

where $f_1 = 5\,\text{Hz}$, and $f_s = 135\,\text{Hz}$, and the mixing coefficient $a_r = \frac{1}{\|\boldsymbol{x}_r\|_2}$ $(r = 1, 2)$. We performed the three-way Toeplitzation for the mixture $y(t)$ to give an order-3 tensor of size $192 \times 32 \times 192$. Since the signal $x_2(t)$ can be represented by a multilinear rank-$(2, 2, 2)$ tensor (see Lemma 6), we can extract the two source signals $x_1(t)$ and $x_2(t)$ through block tensor decompositions. We verified the separation for various noise levels, SNR $= 0, 10, \ldots, 40$ dB, and assessed the performance over 100 independent runs for each noise level. In Fig. 1(a), we compared the mean and median Squared Angular Errors (SAE) achieved using TEDIA and the non-linear least squares (NLS) algorithm which utilises the



(a) Example 1 (b)

Fig. 1. (a) Mean and median squared angular errors (MSAE and MedSAE) of the considered algorithms at various SNRs in Example 1. (b) Sources and estimated signals in Example 2.

Gauss-Newton algorithm with dogleg trust region for the block tensor decomposition [18][1]. The performances were relatively stable and linearly decreased with the signal-noise-ratios in logarithmic scale.

Example 2. This example aims at showing that the proposed method can separate slowly time-varying signals. We considered two mixture signals of length $L = 262144$, $y_r(t) = a_{r1}x_1(t)+a_{r2}x_2(t)$, $r = 1, 2$, composed of two source signals

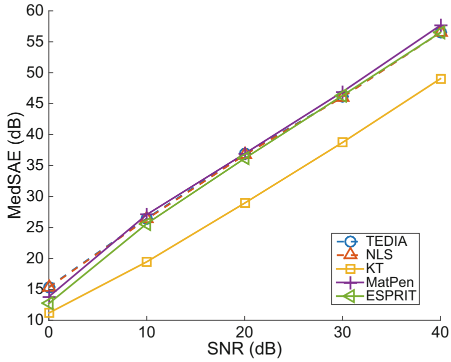
$$x_1(t) = \exp(\frac{-5t}{L}) \sin(\frac{2\pi f_1}{f_s} t), \qquad x_2(t) = \frac{\sin(2\pi f_2/f_s t)}{t + 200}$$

by a mixing matrix $\mathbf{A} = \begin{bmatrix} -1.1896 & 0.1072 \\ 0.0946 & -1.3018 \end{bmatrix}$, where $f_1 = 0.2\,\text{Hz}$, $f_1 = 0.3\,\text{Hz}$, and $f_s = 40000\,\text{Hz}$. We reshaped the mixtures to give an order-4 tensor of size $64 \times 64 \times 64 \times 2$. For this tensorization, the signal $x_1(t)$ yields a tensor of multilinear rank-$(2, 2, 2)$, while the signal $x_2(t)$ is well approximated by a tensor of multilinear rank-$(5, 5, 5)$, because it is elementwise product of a sinusoid and the rational function $1/(t + 200)$, which can be well approximated by a second-order polynomial. Hence we can apply the BTD with two blocks of rank-$(2, 2, 2)$ and rank-$(5, 5, 5)$ to retrieve the source signals. The obtained squared angular errors for the two sources are respective 118.7 and 54.1 dB.

Example 3 (BSS from a short length mixture). In this example, we used a signal of the length $L = 414$ composed of $R = 3$ component signals $x_r(t)$, $r = 1, 2, 3$, and corrupted by additive Gaussian noise $e(t)$

$$y(t) = a_1x_1(t) + a_2x_2(t) + a_3x_3(t) + e(t)$$

[1] The NLS algorithm is available in the Tensorlab toolbox at www.tensorlab.net.



(a) Example 3

Fig. 2. Performance comparison of the considered algorithms at various SNRs in Example 3.

where

$$x_r(t) = \exp(\frac{-5t}{rL}) \sin(\frac{2\pi f_r}{f_s} t + \frac{(r-1)\pi}{R})$$

where the frequencies $f_r = 5, 7, 9\,\text{Hz}$, and the sampling frequency $f_s = 135\,\text{Hz}$. The mixing coefficients a_r were simply set to $a_r = \frac{1}{\|\boldsymbol{x}_r\|_2}$ so that contributions of x_r to $y(t)$ are equivalent.

Tensorization using the three-way folding did not work for such short length signal. Instead, we applied the three-way Toeplitzation, and constructed a tensor of size $192 \times 32 \times 192$. It means that we increased the number of tensor entries to 1,179,648.

Similar to the previous example, after the tensorization, we compressed tensor of the measurement to one of size $6 \times 6 \times 6$ using the HOOI algorithm, and then applied the tensor diagonalization. In Fig. 2(a), we compare performance of the separation via the squared angular error SAE when SNR $= 0,\ 10,\ \ldots 40\,\text{dB}$. The results were assessed over 100 independent runs for each noise level. An important observation is that performances achieved using TEDIA and the NLS algorithm for BTD are almost at the same level. Even when the signal-noise-ratio SNR $= 0\,\text{dB}$, we were able to retrieve the sources with sufficiently good performance.

Performance of the parametric methods including the Kumaresan-Tufts KT algorithm [5], the Matrix Pencil [6], and ESPRIT methods [19] is also compared in Fig. 2(a). The results indicate that our non-parametric method achieved higher performance than the KT algorithm, and was comparable to the ESPRIT algorithm, and slightly worse than the Matrix Pencil algorithm, which performed best.

5 Conclusions

We have presented a method for single mixture blind source separation of low rank signals through the block term decomposition and tensor diagonalization, using three-way folding and Toeplitzation. In particular, we have also shown that the tensorizations of the sinusoid signals and its variants have low multi-linear ranks. For separation of damped sinusoid signals, our method achieved performance which is comparable to the parametric algorithms. The proposed method can also separate other kind of low-rank signals as illustrated in Example 1. In general, the method is also able to separate signals whose multilinear ranks are different.

A Appendix: Low-Rank Representation of the Sequence $x(t) = t^n$

Lemma 5 (Three-way folding of $x(t) = t$). *An order-3 tensor of size $I \times J \times K$, reshaped (folded) from the sequence $1, 2, \ldots, IJK$, where $I, J, K > 2$, has multilinear rank-$(2,2,2)$ and rank-3, and can be represented as*

$$\mathcal{Y} = \mathcal{G} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \mathbf{U}_3 \tag{14}$$

where $\mathcal{G}$ is a tensor of size $2 \times 2 \times 2$

$$\mathcal{G}(:,1,:) = \begin{bmatrix} -2 & -1 \\ 1 & 0 \end{bmatrix}, \qquad \mathcal{G}(:,2,:) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

and the three factor matrices are given by

$$\mathbf{U}_1 = \begin{bmatrix} 0 & 1 \\ \vdots & \vdots \\ i & i+1 \\ \vdots & \vdots \\ (I-1) & I \end{bmatrix}, \ \mathbf{U}_2 = \begin{bmatrix} 0 & 1 \\ \vdots & \vdots \\ jI & jI+1 \\ \vdots & \vdots \\ (J-1)I & (J-1)I+1 \end{bmatrix}, \ \mathbf{U}_3 = \begin{bmatrix} 0 & 1 \\ \vdots & \vdots \\ kIJ & kIJ+1 \\ \vdots & \vdots \\ (K-1)IJ & (K-1)IJ+1 \end{bmatrix}.$$

Lemma 6 (Toeplitzation of $x(t) = t$). *An order-3 Toeplitz tensor of size $I \times J \times K$, tensorized from the sequence $1, 2, \ldots, L$, where $L = I + J + K - 2$, has multilinear rank-$(2,2,2)$ and rank-3, and can be represented as*

$$\mathcal{Y} = \mathcal{G} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \mathbf{U}_3 \tag{15}$$

where $\mathcal{G}$ is a tensor of size $2 \times 2 \times 2$

$$\mathcal{G}(:,:,1) = \begin{bmatrix} I+4 & -(I+3) \\ -(I+3) & I+2 \end{bmatrix}, \qquad \mathcal{G}(:,:,2) = \begin{bmatrix} -(I+3) & I+2 \\ I+2 & -(I+1) \end{bmatrix},$$

and the three factor matrices are given by

$$\mathbf{U}_1 = \begin{bmatrix} 1 & 2 \\ \vdots & \vdots \\ i & i+1 \\ \vdots & \vdots \\ I & I+1 \end{bmatrix}, \quad \mathbf{U}_2 = \begin{bmatrix} I & I+1 \\ \vdots & \vdots \\ j & j+1 \\ \vdots & \vdots \\ I+J-1 & I+J \end{bmatrix}, \quad \mathbf{U}_3 = \begin{bmatrix} I+J-1 & I+J-2 \\ \vdots & \vdots \\ k & k-1 \\ \vdots & \vdots \\ L & L-1 \end{bmatrix}.$$

Lemma 7 (Three-way folding of $x(t) = t^n$). *An order-3 tensor of size $I \times J \times K$, reshaped (folded) from the sequence $x(t) = t^n$, where $n = 1, 2, \ldots$ and $I, J, K > 2$, has multilinear rank-$(n+1, n+1, n+1)$.*

Proof. By exploiting the closed-form expression of $x(t) = t$ in Lemma 5, and the property of the Hadamard product stated in Lemma 4 or in (13), we can prove that the tensor reshaped from $x(t) = t^n$ can be fully explained by three factor matrices which have $n + 1$ columns, and are defined as $\mathbf{U}_1 = \mathbf{F}(I, n)$, $\mathbf{U}_2 = \mathbf{F}(IJ, n)$ and $\mathbf{U}_3 = \mathbf{F}(IJK, n)$, where

$$\mathbf{F}(I, n) = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 1 \\ \vdots & & \vdots & & \\ i^n & \cdots & i^k (i+1)^{n-k} & \cdots & (i+1)^n \\ \vdots & & \vdots & & \vdots \\ (I-1) & \cdots & (I-1)^k I^{n-k} & \cdots & I^n \end{bmatrix}.$$

□

Lemma 8 (Toeplitzation of $x(t) = t^n$). *An order-3 Toeplitz tensor of size $I \times J \times K$ of the sequence $x(t) = t^n$, has multilinear rank-$(n+1, n+1, n+1)$.*

Proof. Skipped for lack of space. □

References

1. Cichocki, A., Zdunek, R., Phan, A.-H., Amari, S.: Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation. Wiley, Chichester (2009)
2. Cichocki, A., Mandic, D., Caiafa, C., Phan, A.-H., Zhou, G., Zhao, Q., Lathauwer, L.D.: Multiway component analysis: tensor decompositions for signal processing applications. IEEE Sig. Process. Mag. **32**(2), 145–163 (2015)
3. Comon, P., Jutten, C. (eds.): Handbook of Blind Source Separation: Independent Component Analysis and Applications. Academic Press, Cambridge (2010)
4. Proakis, J., Manolakis, D.: Digital Signal Processing: Principles, Algorithms, Applications. Macmillan, London (1992)
5. Kumaresan, R., Tufts, D.: Estimating the parameters of exponentially damped sinusoids and pole-zero modelling in noise. IEEE Trans. Acoust. Speech Sig. Proces. **30**(7), 837–840 (1982)

6. Hua, Y., Sarkar, T.: Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoid in noise. IEEE Trans. Acoust. Speech Sig. Process. **38**(5), 814–824 (1990)
7. Papy, J.M., De Lathauwer, L., Van Huffel, S.: Exponential data fitting using multilinear algebra: the single-channel and the multichannel case. Numer. Linear Algebra Appl. **12**(9), 809–826 (2005)
8. Nion, D., Sidiropoulos, N.D.: Tensor algebra and multidimensional harmonic retrieval in signal processing for MIMO radar. IEEE Trans. Sig. Process. **58**, 5693–5705 (2010)
9. Wen, F., So, H.C.: Robust multi-dimensional harmonic retrieval using iteratively reweighted HOSVD. IEEE Sig. Process. Lett. **22**, 2464–2468 (2015)
10. Sørensen, M., De Lathauwer, L.: Multidimensional harmonic retrieval via coupled canonical polyadic decomposition - Part 1: model and identifiability. IEEE Trans. Sig. Process. **65**(2), 517–527 (2017)
11. De Lathauwer, L.: Blind separation of exponential polynomials and the decomposition of a tensor in rank-(L_r,L_r,1) terms. SIAM J. Matrix Anal. Appl. **32**(4), 1451–1474 (2011)
12. De Lathauwer, L.: Block component analysis, a new concept for blind source separation. In: Proceedings of the 10th International Conference on LVA/ICA, Tel Aviv, 12–15 March, pp. 1–8 (2012)
13. Tichavský, P., Phan, A.-H., Cichocki, A.: Tensor diagonalization-a new tool for PARAFAC and block-term decomposition (2014). arXiv preprint arXiv:1402.1673
14. Kouchaki, S., Sanei, S.: Tensor based singular spectrum analysis for nonstationary source separation. In: Proceedings of the 2013 IEEE International Workshop on Machine Learning for Signal Processing, Southampton, UK (2013)
15. Debals, O., Lathauwer, L.: Stochastic and deterministic tensorization for blind signal separation. In: Vincent, E., Yeredor, A., Koldovský, Z., Tichavský, P. (eds.) LVA/ICA 2015. LNCS, vol. 9237, pp. 3–13. Springer, Heidelberg (2015). doi:10.1007/978-3-319-22482-4_1
16. Boussé, M., Debals, O., De Lathauwer, L.: A tensor-based method for large-scale blind source separation using segmentation. IEEE Trans. Sig. Process. **65**(2), 346–358 (2017)
17. Cichocki, A., Namgil, L., Oseledets, I., Phan, A.-H., Zhao, Q., Mandic, D.P.: Tensor networks for dimensionality reduction and large-scale optimization. Found. Trends® Mach. Learn. **9**(4–5) (2017)
18. Sorber, L., Van Barel, M., De Lathauwer, L.: Structured data fusion. IEEE J. Sel. Top. Sign. Proces. **9**(4), 586–600 (2015)
19. Hayes, M.: Statistical Digital Signal Processing and Modeling. Wiley, Hoboken (1996)