Signal Processing 238 (2026) 110113

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Tensor-based higher-order multivariate singular spectrum analysis and applications to multichannel biomedical signal analysis☆

Thanh Trung Le [a,*], Karim Abed-Meraim [b,d], Nguyen Linh Trung [a], Philippe Ravier [b], Olivier Buttelli [b], Ales Holobar [c]

[a] VNU University of Engineering and Technology, Vietnam National University, Hanoi, Viet Nam
[b] University of Orleans, PRISME Laboratory, France
[c] University of Maribor, FERI Faculty, Slovenia
[d] Institut Universitaire de France, France

ARTICLE INFO

Keywords:
Singular spectrum analysis
Multivariate time series
Tensor decomposition
Tensor singular value decomposition
Spectral clustering
Biomedical signal analysis
EMG
ECG
EEG

ABSTRACT

Singular spectrum analysis (SSA) is a nonparametric spectral estimation method that decomposes time series signals into interpretable components. With the rise of big time series, the demand for effective and scalable SSA techniques has become increasingly urgent. In this paper, we propose a novel multiway extension of SSA, called higher-order multivariate SSA (HO-MSSA), specifically designed for multivariate and multichannel time series signal analysis via tensor decomposition. HO-MSSA utilizes time-delay embedding and tensor singular value decomposition to transform multichannel time series signals into trajectory tensors, which are then decomposed into elementary components in the Fourier domain, rather than the time domain as in traditional SSA methods. These components are grouped into disjoint subsets using spectral clustering, enabling the reconstruction of the underlying source signals. Experimental results demonstrate that HO-MSSA outperforms state-of-the-art SSA methods in various biomedical applications, including electromyography (EMG), electrocardiography (ECG), and electroencephalogram (EEG) signals.

1. Introduction

Biomedical signal decomposition involves deconstructing complex signals, such as electromyography (EMG), electrocardiography (ECG), or electroencephalogram (EEG) signals, into simpler, more interpretable components [2]. Accordingly, it helps in identifying underlying patterns and enhances our understanding of intricate biological processes. Singular spectrum analysis (SSA) emerges as a powerful technique for time-series analysis and biomedical signal decomposition [3–5]. In particular, SSA decomposes a signal into several interpretable components, such as slowly time-varying trends, harmonic components, and noises. Notably, there are no assumptions about parametric models or stationarity-type conditions for the signals, making SSA a model-free method with broad applicability. We refer the readers to [3–5] for good references on the SSA literature.

In the era of big data, there has been an increasing interest in the analysis of multivariate and high-dimensional datasets [6]. Several multivariate SSA (MSSA) techniques have been introduced for analyzing multivariate time series, following the approach of Golyandina et al. in [3] for univariate time series and in [7] for multivariate time series. Noteworthy among these techniques are methods like vertical MSSA and horizontal MSSA in [8], 2D-SSA in [9], 2D-SSA in [10], MA-SSA in [11] and principal component trajectories (PCT) in [12] (refer to [13] for an extensive overview of MSSA). However, the aforementioned MSSA techniques typically rely on the singular value decomposition (SVD) on the trajectory matrix to extract elementary components. Consequently, they may not fully exploit features of multivariate time series and might overlook crucial aspects such as spatial correlation, multilinear relationships, and higher-order statistical information.

In parallel, tensor decomposition (TD) has emerged as a powerful data processing tool with applications across various fields [14–17]. Tensors, which are multidimensional/multiway arrays, provide natural representations for multivariate and high-dimensional data [14]. Tensor decomposition allows for the factorization of tensors into basic components (e.g., vectors, matrices, or "simpler" tensors). Accordingly,

☆ A short part of this paper was presented at the 32nd EURASIP European Signal Processing Conference (EUSIPCO 2024), held in Lyon, France [1]. Thanh Trung Le was funded by the Postdoctoral Scholarship Programme of Vingroup Innovation Foundation (VINIF), Viet Nam, code VINIF.2024.STS.40.

* Corresponding author.
E-mail address: thanhletrung@vnu.edu.vn (T.T. Le).

https://doi.org/10.1016/j.sigpro.2025.110113
Received 17 October 2024; Received in revised form 29 January 2025; Accepted 16 May 2025
Available online 2 June 2025

TD has been successfully applied to data and signal processing, particularly in multichannel biomedical signal analysis, including EMG [18,19], ECG [20,21], and EEG signals [22,23]. Our study aims to develop a multilinear, multiway MSSA approach through the lens of TD. In the literature, the first tensor-based SSA method, referred to as TSSA, was introduced in [24]. This method utilizes the CP/PARAFAC decomposition [25] of the trajectory tensor. Specifically, TSSA employs Hankelization and segmentation techniques to embed time series into a third-order trajectory tensor. For the grouping step, TSSA applies empirical mode decomposition to identify the underlying signal components in the desired groups. Since then, several attempts have been made to develop tensor-based MSSA methods, including [26–33]. Most of these methods leverage one of two classical TDs: CP/PARAFAC [25] and Tucker/HOSVD/MLSVD [34]. However, each of these methods has limitations when applied to MSSA. For methods utilizing CP/PARAFAC (e.g., [24,26,27]), the elementary components are rank-1 elements of the CP/PARAFAC decomposition, which can be challenging to identify in practice. Specifically, estimating the number of elementary components for MSSA (i.e., the CP rank) is an NP-hard problem [35]. Additionally, determining the best rank-r approximation of a tensor with order greater than 3 may be ill-posed [36]. Methods utilizing Tucker decomposition (e.g., [28,29]) face issues of non-uniqueness in the sense that the core tensor of Tucker decomposition can be modified without changing the representation, provided that inverse modifications are applied to its loading factors [14]. As a result, the elementary components derived from Tucker decomposition are generally not unique. Moreover, truncating the Tucker decomposition for a fixed rank is also not optimal [37]. The work [33] applied MLSVD, which is a special case of Tucker decomposition, to generalize MSSA. However, the MLSVD decomposition may not be unique if the n-mode singular values of the underlying tensor exhibit a high degree of multiplicity. Some tensor-based SSA methods, such as in [30–32], employ circular Hankelization for embedding of time-series signals. In particular, the columns of trajectory matrices are formed from full-length time series and their circular shifts. Accordingly, these trajectory matrices and tensors can become (very) huge, which limits their practical utility for embedding big time series. In addition, some of them are tailored to specific datasets or applications, such as EEG analysis [29], fault diagnosis [26,27,31], and hyperspectral image analysis [30], limiting their applicability. Therefore, the need for a generalized tensor-based MSSA approach is evident.

In this paper, we introduce a novel multiway extension of SSA designed to handle multichannel time series, which we refer to as higher-order multivariate SSA (HO-MSSA).[1] Specifically, we employ a modified time-delay embedding (TDE) technique that can adapt to various window lengths and step sizes, enabling the transformation of any multichannel time series into trajectory tensors. The decomposition of these trajectory tensors into elementary components is performed by using tensor singular value decomposition (t-SVD) technique [40], a multiway extension of SVD for higher-order tensors. Subsequently, a spectral clustering method is employed to group these elementary components. To extract the underlying time-series signals, we introduce a new block diagonal averaging technique applied to frontal slices of the reconstructed tensor components.

Compared to existing tensor-based SSA methods, HO-MSSA offers several notable advantages. First, the embedding technique we use, TDE in (8), can be seen as a generalized version of the Hankelization and segmentation techniques commonly used in current tensor-based SSA methods (e.g., [24,26–30,33]). Accordingly, it preserves their desirable properties in classical SSA and previous tensor-based SSA methods, while also facilitating low-rank approximations for many types of signals (see Section 3.1 for further details). Next, TDE provides a more compact representation than circular Hankelization method used in [30–32]. As mentioned above, the circular Hankelization can result in a large trajectory tensor, but by adjusting the window size W and step size δ, TDE generates a more moderate trajectory tensor that can be efficiently and effectively factorized using state-of-the-art tensor decomposition methods. Furthermore, TDE often produces a lower-rank representation of the signal compared to the circular Hankelization, see Fig. 6 for an example.

Second, in the context of SSA, t-SVD likely offers a superior decomposition compared to the classical CP and Tucker decompositions. The effectiveness of SVD in traditional SSA is due to three key factors: (i) its Eckart–Young–Mirsky optimality, which ensures that this decomposition provides the best low-rank approximation of trajectory matrices; (ii) the singular values in SVD represent the spectrum of trajectory matrices, serving as a characteristic measure of the contribution of underlying components; and (iii) the orthogonality of the loading factors (i.e., the left and right eigenvector matrices) helps in distinguishing the underlying components [3]. In multichannel and multivariate SSA, CP and Tucker decompositions no longer exhibit these properties, whereas t-SVD still retains them. Third, unlike other SSA methods, the proposed method leverages the appealing properties of tubes in the f-diagonal tensor in the t-SVD of trajectory tensors. This enables the use of modern clustering techniques from machine learning to effectively perform the grouping step. Last but not least, the effectiveness of HO-MSSA is demonstrated over other tensor-based SSA methods through several applications in decomposing and analyzing multichannel biomedical signals, including EMG, ECG, and EEG. Last but not least, the effectiveness of HO-MSSA is demonstrated over other tensor-based SSA methods through several applications in decomposing and analyzing multichannel biomedical signals, including EMG, ECG, and EEG.

2. Preliminaries

2.1. Notations

In this paper, we use the following conventions. Lowercase letters represent scalars (e.g., x), while boldface capital letters indicate vectors (e.g., $\mathbf{x}$). Matrices and tensors are denoted using boldface capital letters (e.g., $\mathbf{X}$) and bold calligraphic letters (e.g., $\boldsymbol{\mathcal{X}}$), respectively. The $(i_1, i_2, \dots, i_n)$th element of $\boldsymbol{\mathcal{X}}$ is denoted as $\boldsymbol{\mathcal{X}}(i_1, i_2, \dots, i_n)$. The transpose operation is represented as $(\cdot)^\top$, and the Frobenius norm as $\|\cdot\|_F$. The functions "$\mathrm{fft}(\cdot)$" and "$\mathrm{ifft}(\cdot)$" denote fast Fourier transform and its inverse operator. We denote by $\widehat{\boldsymbol{\mathcal{X}}} = \mathrm{fft}(\boldsymbol{\mathcal{X}}, [], 3)$ a third-order tensor obtained by taking the Fourier transform along the 3rd dimension of $\boldsymbol{\mathcal{X}}$. Symbols "$\bullet$", "$\boxplus$", and "$*$" represent the t-product, tensor concatenation and circular convolution, respectively. In particular, the t-product of two tensors $\boldsymbol{\mathcal{A}}$ of size $I_1 \times I_2 \times I_3$ and $\boldsymbol{\mathcal{B}}$ of size $I_2 \times L \times I_3$ results in a tensor $\boldsymbol{\mathcal{C}}$ of size $I_1 \times L \times I_3$

$$\boldsymbol{\mathcal{C}} = \boldsymbol{\mathcal{A}} \bullet \boldsymbol{\mathcal{B}} \Leftrightarrow \boldsymbol{\mathcal{C}}(i, j, :) = \sum_{l=1}^{L} \boldsymbol{\mathcal{A}}(i, l, :) * \boldsymbol{\mathcal{B}}(l, j, :). \tag{1}$$

The tensor concatenation of two tensors $\boldsymbol{\mathcal{A}}$ of size $I_1 \times I_2 \times J$ and $\boldsymbol{\mathcal{B}}$ of size $I_1 \times I_2 \times K$ results in a tensor $\boldsymbol{\mathcal{C}}$ of size $I_1 \times I_2 \times (J + K)$

$$\boldsymbol{\mathcal{C}} = \boldsymbol{\mathcal{A}} \boxplus \boldsymbol{\mathcal{B}} \Leftrightarrow \boldsymbol{\mathcal{C}}(:, :, l) = \begin{cases} \boldsymbol{\mathcal{A}}(:, :, l) & \text{if } l \le J \\ \boldsymbol{\mathcal{B}}(:, :, k - J) & \text{if } k > J. \end{cases} \tag{2}$$

[1] In the literature, a variant of SSA was discussed in [38,39], also referred to as higher-order SSA. This variant integrates basic SSA with higher-order statistics of univariate time series. Our proposed method deviates from this approach by employing tensor analysis on multiple trajectory matrices derived from multivariate time series. In our context, the term "higher-order" in HO-MSSA refers to both the increased dimensionality and the order of the trajectory tensor formed by stacking all trajectory matrices.

2.2. Singular spectrum analysis

Singular spectrum analysis (SSA) is performed in four main steps, namely embedding, decomposition, grouping, and diagonal averaging [3].

Step 1. (*Embedding*): Embedding, also known as Hankelization, transforms a time series vector $\mathbf{x} \in \mathbb{R}^{N\times 1}$ into the following Hankel matrix ($J = N - W$)

$$\mathcal{H}(\mathbf{x}) = \begin{bmatrix} \mathbf{x}(1) & \mathbf{x}(2) & \dots & \mathbf{x}(J+1) \\ \mathbf{x}(2) & \mathbf{x}(3) & \dots & \mathbf{x}(J+2) \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{x}(W) & \mathbf{x}(W+1) & \dots & \mathbf{x}(N) \end{bmatrix}, \tag{3}$$

where W is a chosen window length. Here, $\mathcal{H}(\mathbf{x})$ is a trajectory matrix and its columns are called W-lagged vectors of $\mathbf{x}$. In SSA, the window length W should be sufficiently large so that each W-lagged vector incorporates an essential part of the behavior of the time-series signal $\mathbf{x}$.[2]

Step 2. (*Decomposition*): At this step, we perform the singular value decomposition (SVD) of the trajectory matrix

$$\mathcal{H}(\mathbf{x}) = \mathbf{U}\mathbf{S}\mathbf{V}^\top = \sum_{k=1}^{K} \lambda_k \mathbf{u}_k \mathbf{v}_k^\top, \tag{4}$$

where $K = \operatorname{rank}(\mathcal{H}(\mathbf{x}))$; $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_K]$ and $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_K]$ are left and right singular orthogonal vector matrices; and $\lambda_k = \mathbf{S}(k,k)$ is the kth singular value of $\mathcal{H}(\mathbf{x})$. The collection $(\lambda_k, \mathbf{u}_k, \mathbf{v}_k)$ is kth eigentriple of the SVD of $\mathcal{H}(\mathbf{x})$. Rows and columns of $\mathcal{H}(\mathbf{x})$ are subseries of the original time series $\mathbf{x}$. Therefore, the left and right singular vectors also have temporal structures and hence can also be regarded as time series.

Step 3. (*Grouping*): The purpose of this step is to separate additive components of time series, achieved by partitioning the set of indices $\{1, 2, \dots, K\}$ into R (with $R \le K$) disjoint subsets $I_1, I_2, \dots, I_R$ and forming

$$\mathcal{H}(\mathbf{x}) = \sum_{r=1}^{R} \mathbf{X}_r, \text{ where } \mathbf{X}_r = \sum_{i\in I_r} \lambda_i \mathbf{u}_i \mathbf{v}_i^\top. \tag{5}$$

One of the most widely-used technique for extracting components $\{\mathbf{X}_r\}_{r=1}^R$ is to use the matrix of W-correlations [3]. The eigentriples belonging to the same group can correspond to highly correlated components of the time series.

Step 4. (*Diagonal Averaging*): If components of the series are distinctly separated and the indices divided accordingly, then all the matrices in (5) are Hankel matrices, facilitating the direct extraction of corresponding time series. However, in practice, such perfect separation may not be satisfied and we transform an arbitrary matrix into a Hankel matrix and subsequently into a signal, as follows

$$\hat{\mathbf{x}}_r(n) = \begin{cases} \dfrac{1}{n}\displaystyle\sum_{j=1}^{n} \mathbf{X}_r(j, n-j+1) & \text{if } 1 \le n < W \\ \dfrac{1}{W}\displaystyle\sum_{j=1}^{W} \mathbf{X}_r(j, n-j+1) & \text{if } W \le n < J \\ \dfrac{1}{N-n+1}\displaystyle\sum_{j=n-J+1}^{W} \mathbf{X}_r(j, n-j+1) & \text{otherwise.} \end{cases} \tag{6}$$

This step is called diagonal averaging or de-hankelization.

[2] In MSSA, the trajectory matrix may take various forms, such as stacked Hankel, circular Hankel, Hankel-block-Hankel, or quasi-Hankel matrices, depending on the employed techniques. For further details on these forms, we refer readers to [41] for a good reference.

2.3. Tensors and tensor singular value decomposition

Tensor is a multidimensional array and it provides a natural representation for multivariate and high-dimensional data [14]. Tensor singular value decomposition (t-SVD) is a multiway extension of SVD for factorizing higher-order tensors [40]. Under the t-SVD format, a tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{n_1\times n_2\times n_3}$ is decomposed into three tensors $\boldsymbol{\mathcal{U}}, \boldsymbol{\mathcal{S}}$, and $\boldsymbol{\mathcal{V}}$ as follows:

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{U}} \bullet \boldsymbol{\mathcal{S}} \bullet \boldsymbol{\mathcal{V}}^\top, \tag{7}$$

where $\boldsymbol{\mathcal{U}} \in \mathbb{R}^{n_1\times n_1\times n_3}$ and $\boldsymbol{\mathcal{V}} \in \mathbb{R}^{n_2\times n_2\times n_3}$ are orthogonal tensors (i.e., $\boldsymbol{\mathcal{U}} \bullet \boldsymbol{\mathcal{U}}^\top = \boldsymbol{\mathcal{U}}^\top \bullet \boldsymbol{\mathcal{U}} = \boldsymbol{\mathcal{I}}$); $\boldsymbol{\mathcal{S}} \in \mathbb{R}^{n_1\times n_2\times n_3}$ is an f-diagonal tensor whose frontal slices are diagonal. Note that when $n_3 = 1$, the t-product becomes the matrix product and (7) boils down to the classical SVD.

The t-SVD algebraic framework is quite different from the classical multilinear algebra in other types of tensor decomposition. Leveraging the t-product and Fourier transform, it extends various linear and multilinear operations from matrices to tensors, including transpose, orthogonality, and inverse [40].

3. Proposed method

In this section, we propose a novel multiway (tensor) extension of MSSA for handling multichannel and multivariate time-series signals, called HO-MSSA (Higher-Order MSSA). Similarly to standard SSA, HO-MSSA contains four main steps: time delay embedding, tensor SVD, grouping, and reconstruction, as illustrated in Fig. 1.

3.1. Time delay embedding

In this step, for each channel/measurement $\mathbf{x}$, we employ the following time delay embedding (TDE)

$$\mathrm{TDE}_{W,\delta}(\mathbf{x}) = \begin{bmatrix} \mathbf{x}(1) & \mathbf{x}(\delta+1) & \dots & \mathbf{x}(I\delta+1) \\ \mathbf{x}(2) & \mathbf{x}(\delta+2) & \dots & \mathbf{x}(I\delta+2) \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{x}(W) & \mathbf{x}(\delta+W) & \dots & \mathbf{x}(I\delta+W) \end{bmatrix}, \tag{8}$$

with a window length $W \ge 2$, a step size $\delta \ge 1$ and $I = \lfloor (N-W)/\delta \rfloor$. Here, (8) can be viewed as a generalized version of Hankelization (3) and segmentation [42] techniques for embedding time-series signals. When $\delta = 1$, (8) corresponds to the Hankel matrix $\mathcal{H}(\mathbf{x})$ in (3), while it becomes the segmentation if $\delta = W$. We also refer to (8) as the trajectory or TDE matrix.

Interestingly, (8) can facilitate low-rank approximations to several types of time-series signals. A notable class consists of signals that can be expressed as polynomial and exponential components [43], see Proposition 1.

Proposition 1. *Consider a polynomial* $p_r(t) = \sum_{q=0}^{Q_r} c_q t^q$ *of degree* Q_r *with* $c_p \in \mathbb{R}$ *and a signal* $\mathbf{x}$ *of form*

$$\mathbf{x}(t) = \sum_{r=1}^{R} p_r(t) \sin(\omega_r t + \psi_r) z_r^t, \tag{9}$$

with $z_r, \omega_r, \psi_r \in \mathbb{R}$. *Then,* $\operatorname{rank}(\mathrm{TDE}_{W,\delta}(\mathbf{x})) \le \min(W, 2R + \sum_{r=1}^{R} Q_r)\ \forall W, \delta$.

Example 1. If $\mathbf{x}(t) = z^t$, then $\operatorname{rank}(\mathrm{TDE}_{W,\delta}(\mathbf{x})) = 1\ \forall W, \delta$.

Proof. As indicated in [42], if $\mathbf{x}$ takes the form (9), the rank of the corresponding classical Hankel matrix is given by $\operatorname{rank}(\mathrm{TDE}_{W,1}(\mathbf{x})) = \min(W, 2R + \sum_{r=1}^{R} Q_r)$. Since $\mathrm{TDE}_{W,\delta}(\mathbf{x})$ in (8) is formed by selecting a subset columns from $\mathrm{TDE}_{W,1}(\mathbf{x})$, we always have $\operatorname{rank}(\mathrm{TDE}_{W,\delta}(\mathbf{x})) \le \operatorname{rank}(\mathrm{TDE}_{W,1}(\mathbf{x}))\ \forall W, \delta$. Therefore, its rank is upper bounded by $\min(W, 2R + \sum_{r=1}^{R} Q_r)$. □

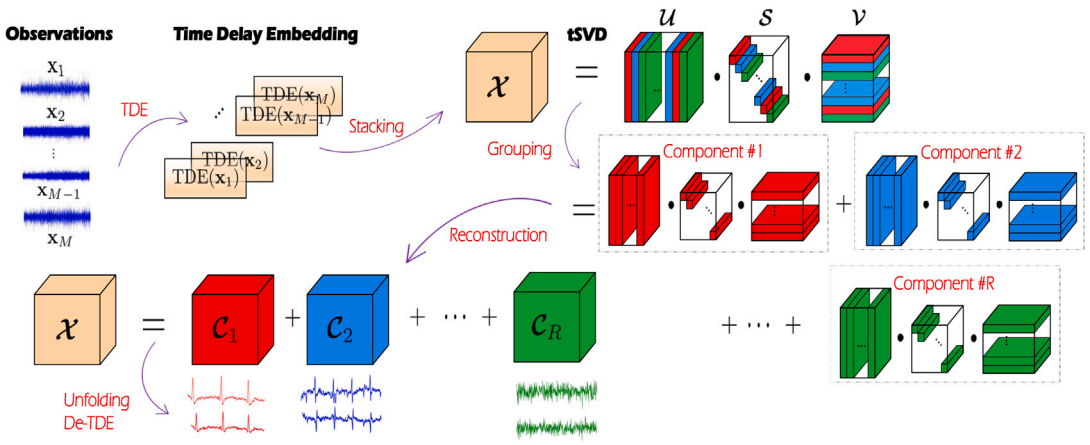


Fig. 1. Main steps of HO-MSSA for multichannel biomedical signal analysis. First, each observation is transformed into a matrix using the time delay embedding (TDE) technique. Next, we construct a trajectory tensor $\mathcal{X}$ by stacking the TDE matrices of M observations along the third dimension. We then decompose $\mathcal{X}$ using t-SVD into elementary tubal-rank-1 components. These components are grouped into disjoint clusters using spectral clustering, which are subsequently used to reconstruct the underlying source signals.

In addition, the TDE (8) is connected to the short time Fourier transform (STFT), which is a useful time–frequency representation tool. Specifically, their relation is given by

$$\mathbf{X}_{\text{STFT}}(\mathbf{x}) = \mathbf{F}_{N\times W}\,\text{diag}(\boldsymbol{\omega})\text{TDE}_{W,\delta}(\mathbf{x}), \tag{10}$$

where $\mathbf{F}_{N\times W}$ contains the first W columns of the discrete Fourier transform matrix and $\boldsymbol{\omega}$ denotes the window function in STFT. Since $\mathbf{F}_{N\times W}$ is a full column rank, $\text{rank}(\mathbf{X}_{\text{STFT}}(\mathbf{x})) = \text{rank}(\text{TDE}_{W,\delta}(\mathbf{x}))$. In [44,45], the authors presented various signals with low-rank STFT matrices. Therefore, TDE also admits the low rank representation to such signals.

It is also worth noting that many signals obtained from our human body show periodic, quasiperiodic, or cyclostationary behavior, reflecting the cyclical patterns inherent in physiological processes [5]. Therefore, TDE can offer a valuable technique to obtain a low-rank representation to biomedical signals. Refer to Fig. 10 for an illustration of an ECG signal and its low-rank TDE matrix.

3.2. Tensor SVD decomposition

We construct the trajectory tensor $\mathcal{X} \in \mathbb{R}^{W\times(I+1)\times M}$ by stacking along the third way TDE matrices of M observations $\{\mathbf{x}_m\}_{m=1}^{M}$ as follows

$$\mathcal{X} = \text{TDE}_{W,\delta}(\mathbf{x}_1) \boxplus \text{TDE}_{W,\delta}(\mathbf{x}_2) \cdots \boxplus \text{TDE}_{W,\delta}(\mathbf{x}_M). \tag{11}$$

Taking the t-SVD decomposition of $\mathcal{X}$ results in:

$$\begin{aligned}\mathcal{X} &\overset{\text{t-SVD}}{=} \mathcal{U} \bullet \mathcal{S} \bullet \mathcal{V}^{\top} \\ &= \sum_{k=1}^{K} \mathcal{U}(:,k,:) \bullet \mathcal{S}(k,k,:) \bullet \mathcal{V}(k,:,:)^{\top} = \sum_{k=1}^{K} \mathcal{X}_k,\end{aligned} \tag{12}$$

where $K \le \min(W, I+1)$ is the tubal rank of $\mathcal{X}$; $\mathcal{U} \in \mathbb{R}^{W\times K\times M}$ and $\mathcal{V} \in \mathbb{R}^{K\times(I+1)\times M}$ are orthogonal tensors; and $\mathcal{S} \in \mathbb{R}^{K\times K\times M}$ is an f-diagonal tensor. Here, the tubal rank-1 tensors $\{\mathcal{X}_k\}_{k=1}^{K}$ are regarded as elementary components in this application, see Fig. 2 for an illustration. In (12), the value of K is identified as the number of non-zero tubes of $\mathcal{S}$, i.e., $K = \sum_k \mathbf{1}[\mathcal{S}(k,k,:) \neq \mathbf{0}]$ where $\mathbf{1}$ is an indicator function. In this step, we can apply Algorithm 1 (other t-SVD algorithms can be founded in [40,46,47]) to perform the tensor decomposition (12) effectively.

The t-SVD decomposition (12) offers two properties that enable HO-MSSA to effectively decompose multichannel time series signals into identifiable components. First, t-SVD provides the optimal low tubal-rank representation for trajectory tensors. Among all tensors $\mathcal{Y}$ of tubal rank $d \le K$, the tensor $\mathcal{Y} = \sum_{k=1}^{d} \mathcal{X}_k$ offers the best approximation to $\mathcal{X}$ by minimizing $\|\mathcal{X} - \mathcal{Y}\|_F^2$ [40]. This allows for effective separation of main components from additive noises. Second, the orthogonality property aids in distinguishing components that are (approximately) orthogonal. Many time series components (e.g., trends and harmonic components with different frequencies) tend to become asymptotically orthogonal as the length of the time series increases.

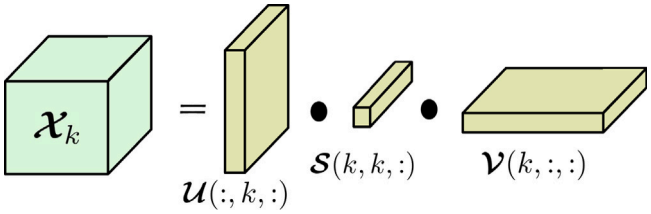
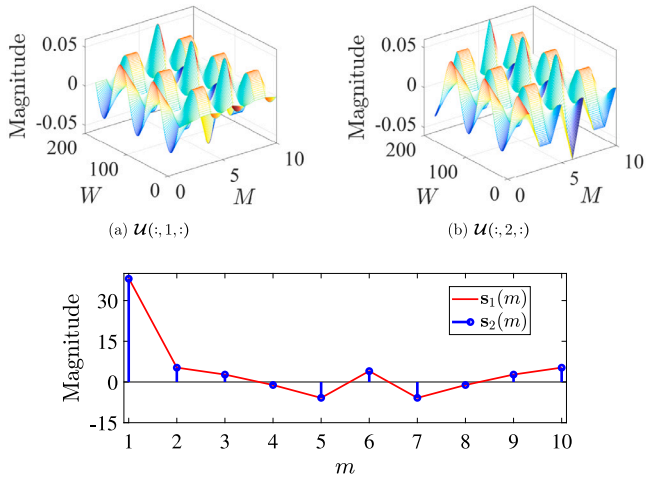


Fig. 2. Elementary tensor of tubal rank 1.

Algorithm 1: t-SVD

Input: $\mathcal{X}$
Output: $\mathcal{U}$, $\mathcal{S}$, $\mathcal{V}$
Main Procedure:
  $\widehat{\mathcal{X}} = \text{fft}(\mathcal{X}, [], 3)$
  for $i = 1, \ldots, \lceil \frac{M+1}{2} \rceil$ **do**
    $[\mathbf{U}_i, \mathbf{S}_i, \mathbf{V}_i] = \text{SVD}(\widehat{\mathcal{X}}(:,:,i), K)$
    $\widehat{\mathcal{U}}(:,:,i) = \mathbf{U}_i$, $\widehat{\mathcal{S}}(:,:,i) = \mathbf{S}_i$, $\widehat{\mathcal{V}}(:,:,i) = \mathbf{V}_i$
  end
  for $j = \lceil \frac{M+1}{2} \rceil + 1, \ldots, M$ **do**
    $\widehat{\mathcal{U}}(:,:,j) = \text{conj}(\mathbf{U}_{M-j+2})$
    $\widehat{\mathcal{S}}(:,:,j) = \mathbf{S}_{M-j+2}$
    $\widehat{\mathcal{V}}(:,:,j) = \text{conj}(\mathbf{V}_{M-j+2})$
  end
  $\mathcal{U} = \text{ifft}(\widehat{\mathcal{U}}, [], 3)$, $\mathcal{S} = \text{ifft}(\widehat{\mathcal{S}}, [], 3)$, $\mathcal{V} = \text{ifft}(\widehat{\mathcal{V}}, [], 3)$
End

We note that t-SVD factorizes the trajectory tensor $\mathcal{X}$ in the Fourier domain rather than the time domain, as in basic SSA and other SSA variants. Indeed, the connection between TDE and STFT in (10) reveals that the Fourier transform $\widehat{\text{TDE}}_{W,\delta}(\mathbf{x})$ of the trajectory matrix $\text{TDE}_{W,\delta}(\mathbf{x})$ can be expressed as $\widehat{\text{TDE}}_{W,\delta}(\mathbf{x}) = \widehat{\mathbf{U}}\mathbf{S}\mathbf{V}^{\top}$ where $\widehat{\mathbf{U}} = \mathbf{F}_{N\times W}\text{diag}(\boldsymbol{\omega})\mathbf{U}$, and $\mathbf{U}, \mathbf{S}$, and $\mathbf{V}$ are three matrix factors of the SVD of $\text{TDE}_{W,\delta}(\mathbf{x})$. It suggests that the decomposition of trajectory matrices can be performed in the Fourier domain, as described in Algorithm 1. Also when $M = 1$ (univariate time series), t-SVD simplifies to the classical SVD, thus making HO-MSSA equivalent to SSA.



(c) Two tubes $\mathbf{s}_1$ and $\mathbf{s}_2$ are almost the same.

Fig. 3. Performing t-SVD on the trajectory tensor built from 10 sinusoidal signals of the same frequency (by differing just by phase and amplitudes). The tubal rank is 2.

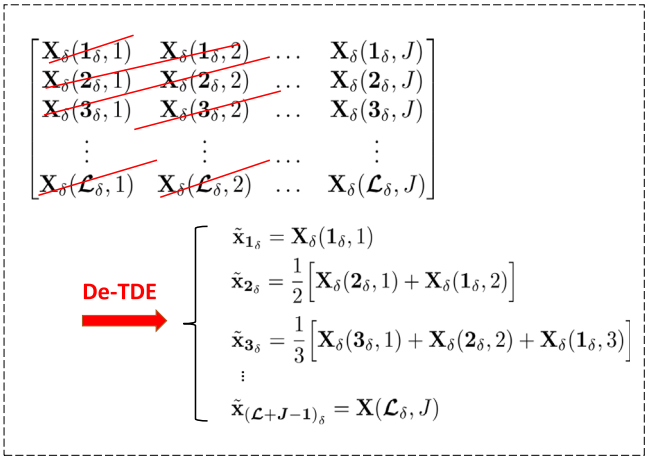


Fig. 4. Block diagonal averaging or De-TDE.

3.3. Grouping

The aim of this step is to divide the set of elementary tensors $\{\mathcal{X}_k\}_{k=1}^K$ (i.e., $\{\mathcal{U}(:,k,:), \mathcal{S}(k,k,:), \mathcal{V}(k,:,:)\}_{k=1}^K$) into R disjoint clusters $\{I_r\}_{r=1}^R$ and then form:

$$\mathcal{X} = \sum_{r=1}^{R} \mathcal{C}_r \quad \text{where} \quad \mathcal{C}_r = \sum_{i \in I_r} \mathcal{X}_i. \tag{13}$$

Let $\mathbf{s}_k = \text{vec}(\mathcal{S}(k,k,:)) \in \mathbb{R}^{M\times 1}$ $\forall k$. We exploit the following observations: (i) the set of first elements $\{\mathbf{s}_k(1)\}_{k=1}^K$ sorts in decreasing order, i.e., $\mathbf{s}_1(1) \geq \mathbf{s}_2(1) \geq \cdots \geq \mathbf{s}_K(1)$, which plays a similar role as singular values of the trajectory matrix in basic SSA; (ii) the remaining elements in $\mathbf{s}_k(2:\text{end})$ are symmetric in the sense that $\mathbf{s}_k(m) = \mathbf{s}_k(M-m+2)$, $\forall m > 1$, $k = 1,2,\ldots,K$; (iii) the value of $\mathbf{s}_k(m)$ with $m \neq 1$ can be negative, unlike the singular values of trajectory matrices; (iv) $\mathbf{s}_k$ and $\mathbf{s}_l$ tend to be "close" if they belong to the same component (see Fig. 3 for an example).

In the case of univariate time series analysis, t-SVD simplifies to SVD, with $\{\mathbf{s}_k\}_{k=1}^K$ representing the singular values of the trajectory matrix. The relevance and physical significance of the extracted components are theoretically related to the concept of separability [41]. Specifically, two components are regarded as strongly separable when they are (approximately) orthogonal and the singular values of their trajectory matrices are disjoint. The orthogonality is inherently achieved through the property of t-SVD, while the disjointness of the tensor tubes is further enhanced. Considering properties (i)-(iv) mentioned above, the likelihood of $\mathbf{s}_k$ coinciding with others from different components is low, thus facilitating the separability. Accordingly, the set of $\{\mathbf{s}_k\}_{k=1}^K$ derived from t-SVD of the trajectory tensor can effectively serve as features to perform the grouping (13). Therefore, in this step, we apply the spectral clustering method [48] to categorize K vectors $\{\mathbf{s}_k\}_{k=1}^K$ into R clusters. Specifically, this method utilizes the spectrum of the similarity or (normalized) Laplacian matrix of $\{\mathbf{s}_k\}_{k=1}^K$ to reduce dimensionality before clustering in lower dimensions via spectral embedding. In situations where the number of clusters R is unknown, this method can determine it by evaluating the eigengap of the (normalized) Laplacian matrix. The computational complexity of this step is $\mathcal{O}(K^2 \max(K,M))$. Noting that, other clustering methods in machine learning can also perform this step.

3.4. Reconstruction

After extracting $\mathcal{C}_r$ for $r = 1,2,\ldots,R$, we reconstruct the corresponding time-series signals in a manner such that their TDE matrices closely approximate the frontal slices of $\mathcal{C}_r$. Below, we present an extended version of (6), called block diagonal averaging or de-TDE, to recover the underlying signal $\mathbf{x}$ from its TDE matrix, accommodating various window lengths W and step sizes in the range $1 \leq \delta \leq W$. Refer to Fig. 4 for an illustration.

Let $\mathcal{L} = \lfloor W/\delta \rfloor$, $\boldsymbol{j}_\delta = [\delta(j-1)+1 : \delta j]$ and $\mathbf{x}_{\boldsymbol{j}_\delta} = \mathbf{x}(\boldsymbol{j}_\delta) \in \mathbb{R}^{\delta\times 1}$. We divide the TDE matrix of $\mathbf{x}$ in (8) into two parts

$$\text{TDE}_{W,\delta}(\mathbf{x}) = \begin{bmatrix} \mathbf{X}_\delta \\ \hline \mathbf{B} \end{bmatrix} = \left[\begin{array}{cccc} \mathbf{x}_{\mathbf{1}_\delta} & \mathbf{x}_{\mathbf{2}_\delta} & \cdots & \mathbf{x}_{\boldsymbol{J}_\delta} \\ \mathbf{x}_{\mathbf{2}_\delta} & \mathbf{x}_{\mathbf{3}_\delta} & \cdots & \mathbf{x}_{(\boldsymbol{J}+\mathbf{1})_\delta} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{x}_{\boldsymbol{\mathcal{L}}_\delta} & \mathbf{x}_{(\boldsymbol{\mathcal{L}}+\mathbf{1})_\delta} & \cdots & \mathbf{x}_{(\boldsymbol{\mathcal{L}}+\boldsymbol{J}-\mathbf{1})_\delta} \\ \hline \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_J \end{array}\right], \tag{14}$$

where $\mathbf{B}$ has a small number of rows (i.e., $W - \mathcal{L}\delta < \delta$ rows). It is a null matrix if $W = \mathcal{L}\delta$. Accordingly, in this step, we employ $\mathbf{X}_\delta$ and $\mathbf{b}_J$ to reconstruct the time-series $\mathbf{x}$ because the contribution of $\mathbf{B}$ to this recovery is negligible. In particular, exploiting the block Hankel structure of $\mathbf{X}_\delta$ in (14), we apply the following block diagonal averaging for the recovery of $\mathbf{x}$

$$\hat{\mathbf{x}} = \left[\hat{\mathbf{x}}_{\mathbf{1}_\delta}^\top \quad \hat{\mathbf{x}}_{\mathbf{2}_\delta}^\top \quad \cdots \quad \hat{\mathbf{x}}_{(\boldsymbol{\mathcal{L}}+\boldsymbol{J}-\mathbf{1})_\delta}^\top \quad \mathbf{b}_J^\top\right]^\top, \tag{15}$$

where

$$\hat{\mathbf{x}}_{\boldsymbol{n}_\delta} = \begin{cases} \dfrac{1}{n}\displaystyle\sum_{j=1}^{n} \mathbf{X}_\delta(\boldsymbol{j}_\delta, n-j+1) & \text{if } 1 \leq n < \mathcal{L}, \\ \dfrac{1}{\mathcal{L}}\displaystyle\sum_{j=1}^{\mathcal{L}} \mathbf{X}_\delta(\boldsymbol{j}_\delta, n-j+1) & \text{if } \mathcal{L} \leq n < J, \\ \dfrac{1}{\mathcal{L}+J-n}\displaystyle\sum_{j=n-J+1}^{\mathcal{L}} \mathbf{X}_\delta(\boldsymbol{j}_\delta, n-j+1) & \text{otherwise.} \end{cases} \tag{16}$$

Here, $\mathbf{X}_\delta$ represents any frontal slice of the tensor $\mathcal{C}_r$.

4. Numerical results

In this section, we present several experiments using both synthetic and real data to demonstrate the effectiveness of HO-MSSA over the classical SSA, vertical MSSA [8], MA-SSA [11], PARAFAC-based MSSA [27] and MLSVD-based MSSA [33] in analyzing multichannel biomedical signals.

4.1. Experiments with synthetic data

We begin by evaluating the performance of HO-MSSA in the context of multichannel electromyography (EMG) denoising, using synthetic

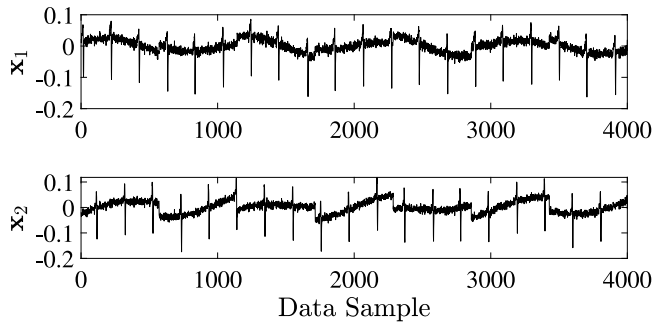


Fig. 5. Synthetic EMG signals corrupted by a sinusoidal signal, a sawtooth signal, and Gaussian noises.

data. Suppose that the mth observation $\mathbf{x}_m$ is generated under the following model

$$\mathbf{x}_m = \mathbf{y}_m + \alpha_m \mathbf{c}_1 + \beta_m \mathbf{c}_2 + \sigma_m \mathbf{n}_m, \tag{17}$$

where $\mathbf{c}_1, \mathbf{c}_2, \mathbf{y}_m$, and $\mathbf{n}_m$ represent a sinusoidal signal, a sawtooth signal, an EMG signal, and a Gaussian noise (zero mean and unit variance), respectively. Our objective is to detect the non-EMG components $\mathbf{c}_1$ and $\mathbf{c}_2$, allowing us to recover the underlying EMG signals $\{\mathbf{y}_m\}_{m=1}^M$.

Following previous works in [49,50], we simulate pure EMG signals as follows

$$\mathbf{y}_m[t] = \sum_{r=1}^{R} \mathbf{y}_{mr}[t] = \sum_{r=1}^{R} \sum_{\ell=0}^{L-1} \mathbf{A}_{mr}[\ell] \mathbf{s}_r[t-\ell]. \tag{18}$$

Here, $\mathbf{A}_{mr}$ represents the rth motor unit (MU) at the mth sensor, with a duration set to $L = 35$ in this experiment. The source signal $\mathbf{s}_r[t] = \sum_j \delta[t - \psi_{rj}]$ denotes the spike train of the rth MU, where spikes occur at times ψ_{rj} and $\delta[\cdot]$ is the delta function.[3] For further details on the mathematical modeling of EMG signals, we refer the readers to [49] for a good reference. In our setting, two pure EMG signals are comprised of two sources with an excitation level set at 5% of the maximum voluntary contraction and the sampling frequency of $f = 2048$ Hz. Each measurement contains a sequence of 4096 data samples, corresponding to a duration of 2 s. The number of spikes from each source is 20. In Matlab, we simulate two non-EMG components $\mathbf{c}_1$ and $\mathbf{c}_2$ as $\sin(2 \times \pi \times 4/N \times [1 : N])$ and $\text{sawtooth}(2 \times \pi \times 7/N \times [1 : N])$ where $N = 4096$. The coefficients $[\alpha_1, \alpha_2]$, $[\beta_1, \beta_2]$, and $[\sigma_1, \sigma_2]$ are fixed at $[1, 1], [1, -0.5], [1, 1]$, respectively. See Fig. 5 for an illustration of two observations $\mathbf{x}_1$ and $\mathbf{x}_2$.

We first illustrate the effect of the TDE in (8), comparing it to classical Hankelization in (3) and circular Hankelization as presented in [31]. We consider two case studies for the window W, corresponding to $W = 5L$ (short) and $W = N/2$ (long). The step size δ is set to L, $2L$, and $5L$. Fig. 6(a) and (b) depict the spectrum of trajectory matrices for the EMG and sawtooth components of $\mathbf{x}_1$ with these embedding techniques. Note that classical Hankelization corresponds to our $\text{TDE}_{W,\delta}$ with $\delta = 1$. We can see that the decay rate of the spectrum of trajectory matrices is significantly faster than that of the classical and circular Hankel transforms. It suggests that our $\text{TDE}_{W,\delta}$ can enhance the low rank approximation to signals more effectively than classical and circular Hankelization techniques. Particularly when $\delta = L$, the signals reconstructed using the de-TDE (15) closely match the original observations. However, with a larger step size (i.e., $\delta = 5L$), the reconstructed signals deviate significantly from the original observations, as shown in Fig. 6(c) and (d).

[3] The motor unit (MU) is the smallest functional unit of the neuromuscular command which structures it at the muscle level. This command controls MU by modulating its firing frequency (spike train).

In the following experiment, we investigate the effect of the embedding dimension I on the separability of components. Here, we fix the window size W to $N/2$ and vary the step size δ from 1 to $5L$ (where $L = 35$ is the length of the motor unit filter). As the embedding dimension $I = \lfloor (N-W)/\delta \rfloor$, its value ranges from 11 to 2048. We follow the previous works [51,52] to evaluate the mean squared separation error, which is based on the reconstructed error (RE) of the estimated components:

$$\text{RE} = \frac{1}{N} \sum_{t=1}^{N} \big(\mathbf{c}(t) - \hat{\mathbf{c}}(t)\big)^2, \tag{19}$$

where $\mathbf{c}$ and $\hat{\mathbf{c}}$ denote the true and estimated components, respectively. Fig. 7 illustrates the reconstructed error for four underlying components, including two non-EMG components $\mathbf{c}_1, \mathbf{c}_2$, and two EMG components $\mathbf{y}_1$, $\mathbf{y}_2$, as function of the embedding dimension I. As the embedding dimension I increases (i.e., as the step size δ decreases), the reconstruction error and hence separability of all four components improves. Note that selecting the optimal parameters, such as the embedding dimension, window length, and step size, depends on the specific data types and applications. Consequently, their effect on component separability may vary across different tasks and datasets.

Next, we use a time window $W = 5L$ and a step size $\delta = L$ to construct the trajectory tensor $\mathcal{X}$ of size $175 \times 110 \times 2$, and proceed to factorize $\mathcal{X}$ via the t-SVD decomposition using Algorithm 1. This decomposition results in two orthogonal tensors $\mathcal{U} \in \mathbb{R}^{175\times110\times2}$, $\mathcal{V} \in \mathbb{R}^{110\times110\times2}$, and an f-diagonal tensor $\mathcal{S} \in \mathbb{R}^{110\times110\times2}$. We obtain a set of 110 elementary tensors $\{\mathcal{X}_k\}_{k=1}^{110}$, where $\mathcal{X}_k = \mathcal{U}(:,k,:) \bullet \mathcal{S}(k,k,:) \bullet \mathcal{V}(k,:,:) \in \mathbb{R}^{175\times110\times2}$. We then apply spectral clustering to group these components into four main classes: EMG signal, sinusoid signal, sawtooth signal, and noise. This clustering method is implemented by using normalized graph Laplacian with a Gaussian similarity function defined as $\mathbf{L}(\mathbf{s}_i, \mathbf{s}_j) = \exp(-\|\mathbf{s}_i - \mathbf{s}_j\|_2^2/2)$. In Matlab, we use the command "`spectralcluster(L,4, 'Distance', 'precomputed', 'LaplacianNormalization', 'symmetric')`", where "$\mathbf{L} = \exp(-\text{dist}.\{2)$" and "`dist = squareform(pdist(S))`". Subsequently, time-series signals are extracted from each measurement through block diagonal averaging applied to the frontal slices of the corresponding reconstructed tensor. Fig. 8 illustrates the recovered signals, which closely match the original ones. These results highlight the potential of HO-MSSA for effectively decomposing multichannel time-series signals.

4.2. Experiments with real data

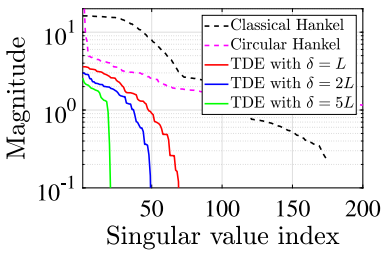
We next consider two real-life applications to further demonstrate the effectiveness of HO-MSSA compared to other (M)SSA methods: fetal heartbeat extraction from maternal abdominal ECG recordings and EOG artifact removal from EEG signals.

4.2.1. Fetal heartbeat extraction from maternal abdominal ECG recordings

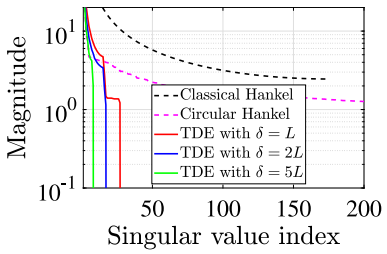
Fetal heartbeat extraction from maternal abdominal electrocardiogram (ECG) recordings is a challenging blind source separation problem due to (i) the low amplitude of fetal ECG signals, (ii) the potential for noise from various sources, and (iii) other factors [53]. We explore the application of HO-MSSA to this task using an open ECG dataset.[4] This dataset contains five abdominal and three thoracic recordings acquired from various regions of the mother's body, with a sampling frequency of 250 Hz. We focus on five abdominal recordings, each consisting of 800 data samples only. Refer to Fig. 9 for an illustration.

In our experiment, we set the window length W to 400 (i.e., half of the signal length), with a step size $\delta = 1$, resulting in a trajectory tensor $\mathcal{X}$ of size $400 \times 401 \times 5$. Fig. 10 shows the TDE matrix (a frontal slice of $\mathcal{X}$) and its corresponding singular values for an ECG recording. We can see that the spectrum exhibits fast decay, with focusing on the first few

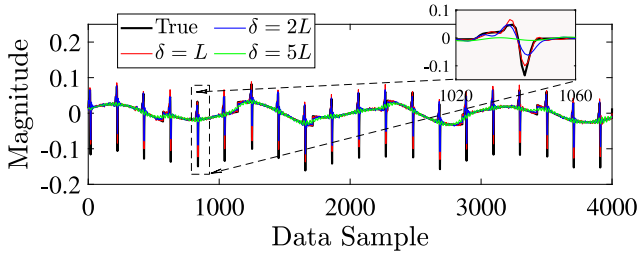
[4] https://ftp.esat.kuleuven.be/pub/SISTA/data/biomedical.



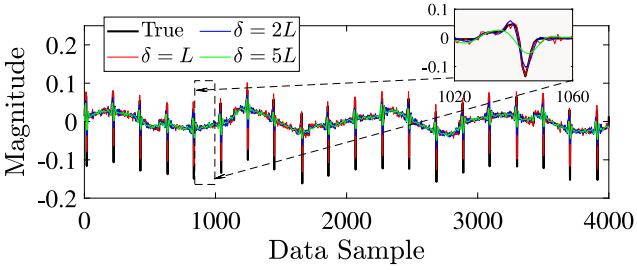
(a) Spectrum of trajectory matrices of EMG components with the window $W = 5L$.



(b) Spectrum of trajectory matrices of sawtooth components with the window $W = 5L$.



(c) Recovered signals with the window $W = 5L$.



(d) Recovered signals with the window $W = N/2$.

Fig. 6. Transformations of time delay embedding for the signal $\mathbf{x}_1$ using different window lengths W and step sizes δ.

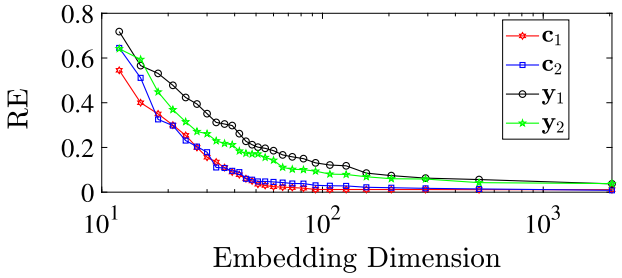


Fig. 7. Effect of the embedding dimension on the separability of components.

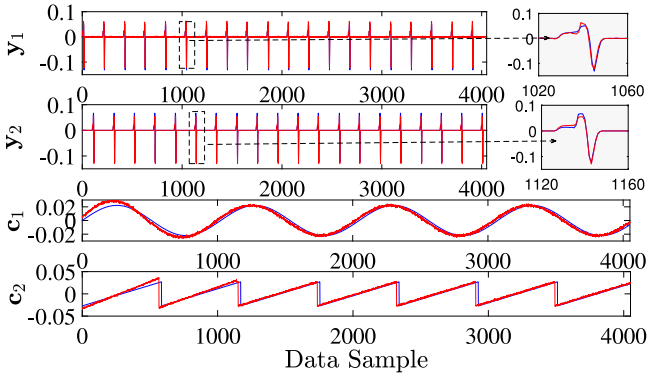


Fig. 8. Extracted components: (——) ground truth; (——) recovered component.

singular values. This property of the TDE matrix facilitates a low-rank approximation of the ECG signal and hence the trajectory tensor $\mathcal{X}$.

As the groundtruth is unavailable, we evaluate the results by comparing them with those obtained using TenSOFO, a successful tensor-based method for ECG source separation in [50,54]. Additionally, we can assess the experimental results by confirming the fact that the fetal heart rate is consistently higher than that of the mother.

All the SSA methods compared in this study rely on the window length, which is also set to $W = 400$. For the PARAFAC-based MSSA method, the authors recommend setting the CP rank to the number of source signals, which is 2 for this experiment. For the MLSVD-based MSSA method, the rank is determined by maximizing the spectral gap of the mode-1 unfolding matrix of the trajectory tensor, with an estimated value of 17. The algorithmic parameters of TenSOFO are kept as default. We apply the same HO-MSSA framework as described in Section 4.1 for this task. The only difference is that we estimate the tubal rank of $\mathcal{X}$ instead of using the full tubal rank t-SVD representation. Specifically, the tubal rank of $\mathcal{X}$ is determined to be $K = 92$, representing to 92 elementary components. We use the "tubalrank.m" function from the tensor–tensor product toolbox to estimate the value of K.[5] The experimental results are shown in Fig. 11. It is evident that SSA, vertical MSSA, and PARAFAC-based MSSA methods are not effective for this task, as they struggle to provide accurate results. Specifically, both classical SSA and vertical MSSA methods produce two components that seem to correspond to the mother's cardiac cycle, but their waveforms do not accurately reflect the heartbeat. The

[5] https://github.com/canyilu/tensor-tensor-product-toolbox.

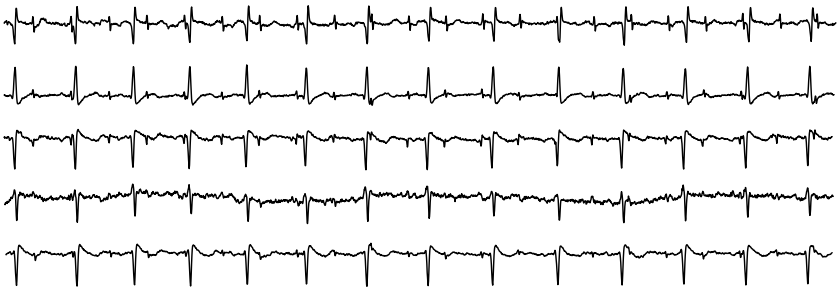
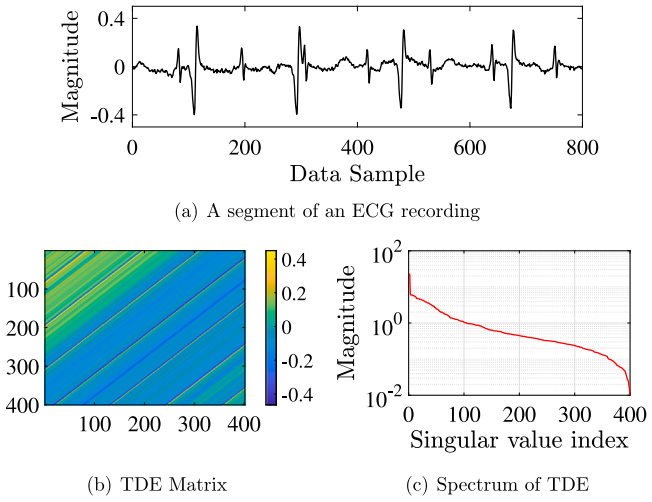


Fig. 9. Five abdominal ECG recordings.



(a) A segment of an ECG recording

(b) TDE Matrix

(c) Spectrum of TDE

Fig. 10. Time delay embedding (TDE) promotes low-rank approximation to ECG signals.

results from the PARAFAC-based MSSA make it difficult to realize any heartbeats. The MLSVD-based MSSA is limited to estimating only the mother's heartbeats. By contrast, HO-MSSA produces two distinct beats that likely correspond to the fetal and maternal heartbeats. Additionally, these results closely match those obtained using TenSOFO [50]. It suggests that HO-MSSA is capable of extracting fetal heartbeats from ECG recordings of the mother.

4.2.2. EOG artifact removal from EEG signals

A common issue in electroencephalographic (EEG) applications is that artifacts can significantly distort EEG signals, such as ocular artifacts from eye movements and blinks [55]. Therefore, it is crucial to remove such artifacts before further analysis. We demonstrate the second application of HO-MSSA for removing electrooculographic (EOG) artifacts from EEG signals, with an open EEG dataset.[6] Particularly, EEG signals were recorded from 27 participants, using the 10–20 system with 19 channels. The signals were resampled with a sampling frequency of 200 Hz, and filtered by a bandpass filter between 0.5 and 40 Hz, along with a notch filter at 50 Hz. In this task, we focus on six EEG recording channels (i.e., Fp1, Fp2, F3, F4, F7, F8) and analyze data from the first 15 s, which corresponds to 3000 data samples. Refer to Fig. 12 for an illustration of the EEG recordings from Participant #1.

To detect and remove EOG artifacts, we apply the same HO-MSSA framework of as in previous tasks. For each participant, we construct a trajectory tensor $\mathcal{X}$ of size $1500 \times 1501 \times 6$, perform the t-SVD decomposition of $\mathcal{X}$ to extract elementary components, and then apply spectral clustering to group them into two main classes: EEG and artifacts (EOG and noises). The number of elementary components is also determined by the tubal rank K of $\mathcal{X}$ and its value changes depending on the participants. Table 1 reports the estimated tubal rank K of the EEG tensors for 27 patients. Given that the EEG signals contain various artifacts and noises, the estimated tubal rank K tends to be moderate. Specifically, the averaged value of K over 27 patients is 532.95 ± 8.23 (mean $\pm$ standard derivation). Despite the moderate tubal rank, the variance ratio $\mathrm{VR}(k)$ defined as[7]

$$\mathrm{VR}(k) = \frac{\sum_{i=1}^{k} \|\mathbf{s}_i\|_2^2}{\sum_{j=1}^{I} \|\mathbf{s}_j\|_2^2} \tag{20}$$

reaches nearly 100% when k exceeds 30% of the tensor dimension $I = 1500$, as shown in Fig. 13. This suggests that the majority of the signal's variance is captured by a relatively small number of components compared to the data dimension, indicating that 30% strongest tensor components are dominant.

Table 1
Estimated tubal ranks of EEG tensors.

Patient	Rank	Patient	Rank	Patient	Rank
#1	540	#10	541	#19	528
#2	551	#11	538	#20	529
#3	529	#12	529	#21	546
#4	527	#13	544	#22	547
#5	548	#14	540	#23	541
#6	536	#15	523	#24	526
#7	523	#16	524	#25	532
#8	527	#17	529	#26	525
#9	527	#18	531	#27	534

To evaluate the algorithm's performance, we use the Mean Absolute Error (MAE) metric in the frequency domain, as presented in previous work [56]:

$$\mathrm{MAE}(B_f) = \sum_{f \in B_f} \left| P_{\mathrm{con}}(f) - P_{\mathrm{rec}}(f) \right| \Big/ \left| B_f \right|, \tag{21}$$

where $P_{\mathrm{con}}(f)$ and $P_{\mathrm{rec}}(f)$ are the power spectrum density of the contaminated (noisy) signal $\mathbf{x}$ and the recovered (denoised) signal $\mathbf{y}$, and B_f indicates the frequency band of interest. We aim for the MAE metric to be as small as possible. As our EEG signals are filtered by a bandpass filter of 0.5–40 Hz, the MAE metric is specifically determined in five frequency bands, including delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (12–30 Hz) and low-gamma (30–40 Hz). As suggested in [56,57], together with the MAE metric, we also use the following metric to assess the correlation between $\mathbf{x}$ and $\mathbf{y}$ in the frequency domain

$$\rho(f) = \frac{\frac{1}{2}\left| \sum_{f-\delta_f}^{f+\delta_f} \hat{\mathbf{x}}^*(k)\hat{\mathbf{y}}(k) + \hat{\mathbf{x}}(k)\hat{\mathbf{y}}^*(k) \right|}{\sqrt{\left| \sum_{f-\delta_f}^{f+\delta_f} \hat{\mathbf{x}}(k)\hat{\mathbf{x}}^*(k) . \sum_{f-\delta_f}^{f+\delta_f} \hat{\mathbf{y}}(k)\hat{\mathbf{y}}^*(k) \right|}}, \tag{22}$$

[6] https://data.mendeley.com/datasets/wb6yvr725d/3.

[7] Note that for any tensor $\mathcal{X}$ with tubal rank K, it holds that $\|\mathcal{X}\|_F^2 = \|\mathcal{S}\|_F^2 = \sum_{k=1}^{K} \|\mathbf{s}_k\|_2^2$ where $\mathbf{s}_k = \mathcal{S}(k,k,:)$.

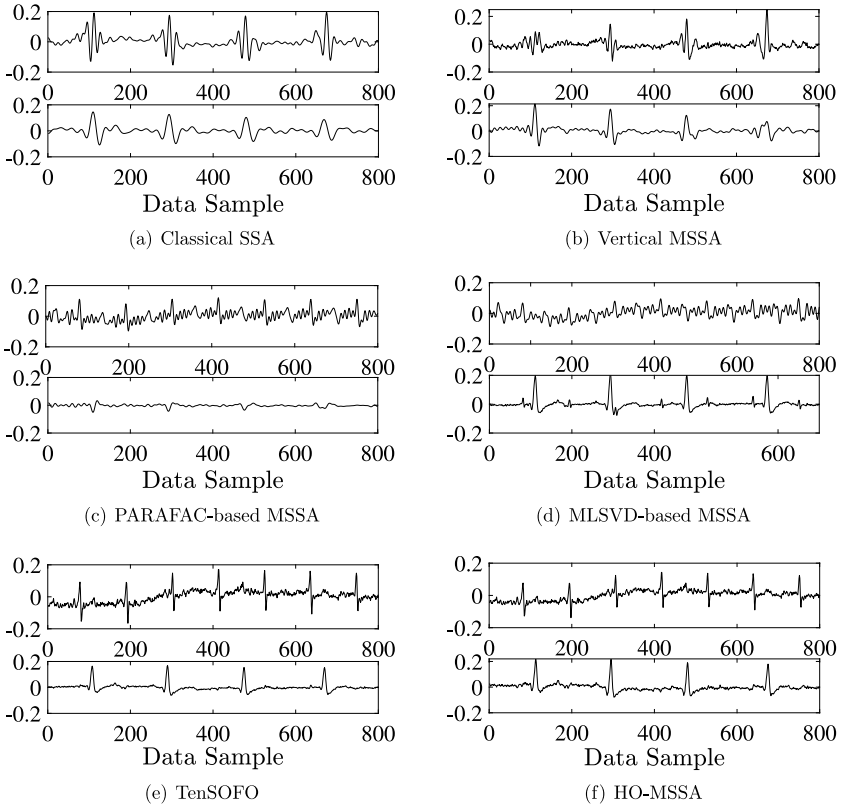


Fig. 11. ECG separation results: fetal heartbeats (above) and maternal heartbeats (below).

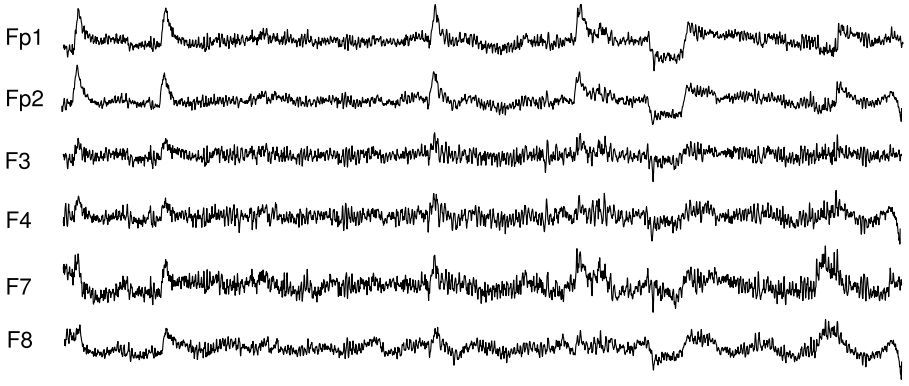


Fig. 12. EEG signals mixed with EOG artifacts: 06 recording channels.

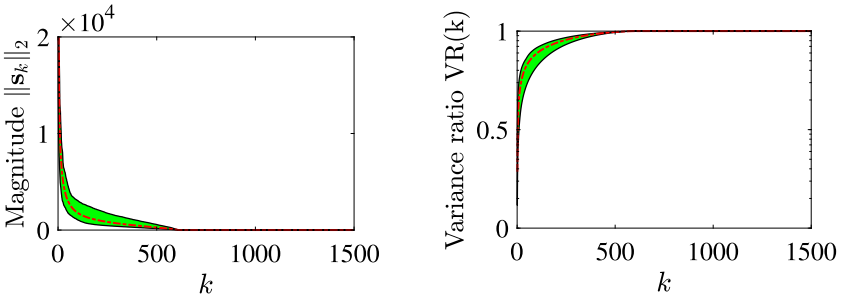
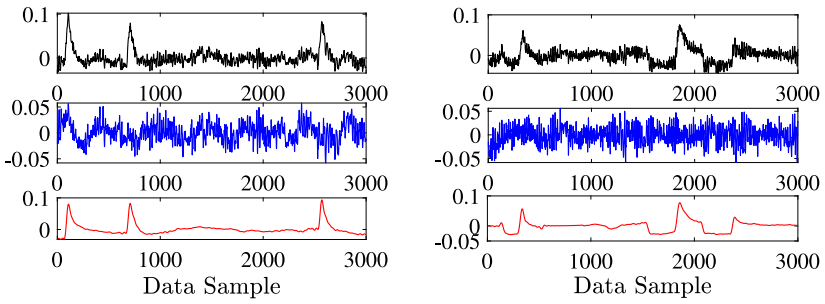
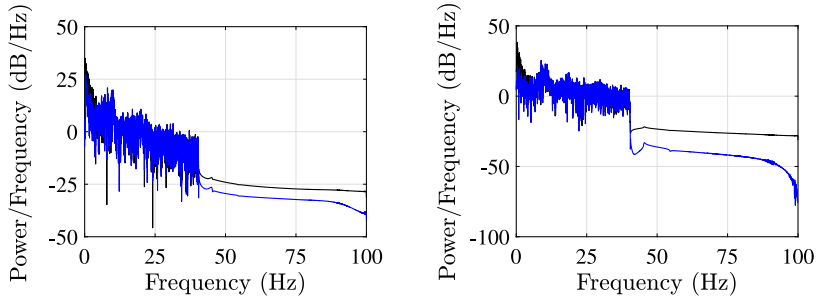


Fig. 13. The norm of $\{\mathbf{s}_k\}_{k=1}^{1500}$ of the f-diagonal tensor $\mathcal{S}$ from the t-SVD of EEG trajectory tensors, along with the corresponding variance ratio VR(k). The red dashed lines represent the average values, while the green shaded area indicates their domain.



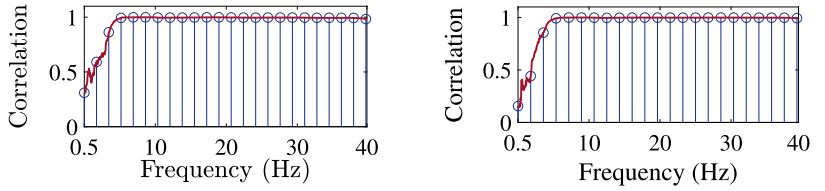
(a) Pat. #1: Signals

(b) Pat. #2: Signals



(c) Pat. #1: Power spectral density

(d) Pat. #2: Power spectral density



(e) Pat. #1: Correlation between the contaminated and denoised EEG signals in the frequency domain from 0.5 to 40 Hz.

(f) Pat. #2: Correlation between the contaminated and denoised EEG signals in the frequency domain from 0.5 to 40 Hz.

Fig. 14. Results of EOG artifact removal from EEG signals using HO-MSSA: (——) contaminated EEG signals; (——) denoised EEG signals; (——) estimated EOG artifacts. Sub-figures on the left and right are for Participant #1 and #2, respectively. Note that our EEG signals are filtered using a bandpass filter with a frequency range of 0.5 to 40 Hz.

Table 2
EOG artifact removal from EEG signals: Performance of (M)SSA algorithms in terms of the MAE metric.

Method \ Band	delta (0.5–4 Hz)	theta (4–8 Hz)	alpha (8–12 Hz)	beta (12–30 Hz)	low-gamma (30–40 Hz)
SSA	80.37 ± 25.86	12.73 ± 4.342	5.802 ± 0.949	2.269 ± 1.908	0.624 ± 0.121
MA-SSA	97.49 ± 34.41	8.522 ± 3.215	4.891 ± 1.556	1.074 ± 0.593	0.357 ± 0.098
Vertical MSSA	126.9 ± 40.92	3.287 ± 1.026	3.714 ± 1.207	1.231 ± 0.954	0.093 ± 0.072
PARAFAC-based	89.27 ± 20.45	7.233 ± 3.956	4.722 ± 1.045	1.334 ± 0.809	0.528 ± 0.138
Tucker-based	152.5 ± 50.97	2.463 ± 0.841	1.503 ± 0.942	0.959 ± 0.072	0.088 ± 0.076
HO-MSSA	102.1 ± 37.78	1.301 ± 0.374	0.705 ± 0.462	0.133 ± 0.085	0.051 ± 0.029

(mean ± standard deviation).

Here, $(\cdot)^*$ denotes the complex conjugate operator, $2\delta_f$ defines the frequency window, $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ represent the Fourier coefficients of $\mathbf{x}$ and $\mathbf{y}$, respectively. The value of $\rho(f)$ ranges from 0 to 1, where 0 indicates no correlation and 1 indicates perfect correlation. In this task, we set the value of $2\delta_f$ to 3 Hz which covers approximately 87 Fourier transform coefficients. This 3 Hz across window is then moved across the entire spectrum of 40 Hz.

Performance of the SSA methods are shown graphically in Fig. 14 and statistically in Table 2. For Participants #1 and #2, HO-MSSA effectively separates artifacts from their EEG recordings, as illustrated by red lines in Fig. 14(a) and (b). The power spectral density of the reconstructed/denoised EEG signals (blue lines) closely matches to the reference signals, particularly in the theta, alpha, beta and low-gamma bands, as shown in Fig. 14(c) and (d). This is further validated by the evaluation metric $\rho(f)$, where the correlation value is close to 1 for frequencies above 4 Hz, as seen in Fig. 14(e) and (f).

The average results for HO-MSSA, and other SSA methods (using the same experimental setups as in previous tasks) across 27 participants with respect to the MAE metric are reported in Table 2. We can observe that, in the theta, alpha, beta, and low-gamma bands, our method achieves the lowest MAE values, outperforming other (M)SSA methods. Performance of all (M)SSA methods in the delta band (0.5–4 Hz) is poor. However, this band is not considered in our analysis and the results do not completely reflect the performance of the

algorithms. It is due to that delta waves are typically associated with sleep stages [58], while the EEG signals in this study were recorded from awake participants.

5. Conclusions

In this paper, we proposed a novel extension of singular spectrum analysis called higher-order multivariate SSA (HO-MSSA) for multichannel time-series analysis. Our method employs a variant of time-delay embedding to transform signals into trajectory tensors, enabling effective decomposition using tensor SVD. Through spectral clustering and block diagonal averaging technique, we extracted interpretable time-series signals. The demonstrated success in separating fetal ECG signals from maternal ECG signals and EOG artifact removal from EEG signals highlights the potential of our method for multivariate and high-dimensional biomedical data analysis.

CRediT authorship contribution statement

Thanh Trung Le: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Karim Abed-Meraim:** Writing – review & editing, Validation, Methodology. **Nguyen Linh Trung:** Writing – review & editing, Validation. **Philippe Ravier:** Writing – review & editing, Validation. **Olivier Buttelli:** Writing – review & editing, Validation. **Ales Holobar:** Writing – review & editing, Validation.

Data availability

Data will be made available on request.

References

[1] T.T. Le, K. Abed-Meraim, N.L. Trung, Higher-order singular spectrum analysis for multichannel biomedical signal analysis, in: Proc. Eur. Signal Proc. Conf., 2024, pp. 1337–1341.

[2] R.M. Rangayyan, S. Krishnan, Biomedical Signal Analysis, John Wiley & Sons, 2024.

[3] N. Golyandina, V. Nekrutkin, A.A. Zhigljavsky, Analysis of Time Series Structure: SSA and Related Techniques, CRC Press, 2001.

[4] J.B. Elsner, A.A. Tsonis, Singular Spectrum Analysis: A New Tool in Time Series Analysis, Springer Science & Business Media, 2013.

[5] S. Sanei, H. Hassani, Singular Spectrum Analysis of Biomedical Signals, CRC Press, 2015.

[6] I. Koch, Analysis of Multivariate and High-Dimensional Data, Cambridge University Press, 2013.

[7] N. Golyandina, D. Stepanov, SSA-based approaches to analysis and forecast of multidimensional time series, in: Proc. St. Petersburg Works. Simulation, Vol. 293, 2005, p. 298.

[8] H. Hassani, R. Mahmoudvand, Multivariate singular spectrum analysis: A general view and new vector forecasting approach, Int. J. Energy Stat. 1 (01) (2013) 55–83.

[9] N.E. Golyandina, K. Usevich, 2D-extension of Singular Spectrum Analysis: algorithm and elements of theory, in: Matrix Methods: Theory, Algorithms and Applications, 2010, pp. 449–473.

[10] J. Zabalza, et al., Novel two-dimensional singular spectrum analysis for effective feature extraction and data classification in hyperspectral imaging, IEEE Trans. Geosci. Remote Sens. 53 (8) (2015) 4418–4433.

[11] S. Jain, R. Panda, R.K. Tripathy, Multivariate sliding-mode singular spectrum analysis for the decomposition of multisensor time series, IEEE Sens. Lett. 4 (6) (2020) 1–4.

[12] D. Dylewsky, E. Kaiser, S.L. Brunton, J.N. Kutz, Principal component trajectories for modeling spectrally continuous dynamics as forced linear systems, Phys. Rev. E 105 (1) (2022) 015312.

[13] N. Golyandina, A. Korobeynikov, A. Shlemov, K. Usevich, Multivariate and 2D extensions of singular spectrum analysis with the Rssa package, J. Stat. Softw. 67 (2) (2015) 1–78.

[14] T.G. Kolda, B.W. Bader, Tensor decompositions and applications, SIAM Rev. 51 (3) (2009) 455–500.

[15] A. Cichocki, et al., Tensor decompositions for signal processing applications: From two-way to multiway component analysis, IEEE Signal Process. Mag. 32 (2) (2015) 145–163.

[16] N.D. Sidiropoulos, L.D. Lathauwer, X. Fu, K. Huang, E.E. Papalexakis, C. Faloutsos, Tensor decomposition for signal processing and machine learning, IEEE Trans. Signal Process. 65 (13) (2017) 3551–3582.

[17] T.T. Le, K. Abed-Meraim, N.L. Trung, A. Hafiane, A contemporary and comprehensive survey on streaming tensor decomposition, IEEE Trans. Knowl. Data Eng. 35 (11) (2023) 10897–10921.

[18] S. Padhy, A tensor-based approach using multilinear SVD for hand gesture recognition from sEMG signals, IEEE Sens. J. 21 (5) (2020) 6634–6642.

[19] T.T. Le, K. Abed-Meraim, P. Ravier, O. Buttelli, A. Holobar, Tensorial convolutive blind source separation, in: Proc. IEEE Int. Conf. Acoust. Speech Signal Process., 2024, pp. 5575–5579.

[20] S. Padhy, S. Dandapat, Third-order tensor based analysis of multilead ECG for classification of myocardial infarction, Biomed. Signal Process. Control. 31 (2017) 71–78.

[21] J.H.d.M. Goulart, P.M.R. de Oliveira, R.C. Farias, V. Zarzoso, P. Comon, Alternating group lasso for block-term tensor decomposition and application to ECG source separation, IEEE Trans. Signal Process. 68 (2020) 2682–2696, http://dx.doi.org/10.1109/TSP.2020.2985591.

[22] F. Cong, Q.-H. Lin, L.-D. Kuang, X.-F. Gong, P. Astikainen, T. Ristaniemi, Tensor decomposition of EEG signals: A brief review, J. Neurosci. Methods 248 (2015) 59–69.

[23] L.T. Thanh, N.T.A. Dao, N.V. Dung, et al., Multi-channel EEG epileptic spike detection by a new method of tensor decomposition, J. Neural Eng. 17 (1) (2020) 016023.

[24] S. Kouchaki, S. Sanei, E.L. Arbon, D.-J. Dijk, Tensor based singular spectrum analysis for automatic scoring of sleep EEG, IEEE Trans. Neural Syst. Rehabil. Eng. 23 (1) (2014) 1–9.

[25] R.A. Harshman, Foundations of the PARAFAC procedure: Models and conditions for an explanatory multimodal factor analysis, UCLA Work. Pap. Phon. 16 (1–84) (1970).

[26] C. Yi, Y. Lv, M. Ge, H. Xiao, X. Yu, Tensor singular spectrum decomposition algorithm based on permutation entropy for rolling bearing fault diagnosis, Entropy 19 (4) (2017) 139.

[27] D. Yang, et al., Improved tensor-based singular spectrum analysis based on single channel blind source separation algorithm and its application to fault diagnosis, Appl. Sci. 7 (4) (2017) 418.

[28] W.-p. Huang, B.P. Kwan, W. Ding, B. Min, R.C. Cheung, L. Qi, H. Yan, High performance hardware architecture for singular spectrum analysis of hankel tensors, Microprocess. Microsyst. 64 (2019) 120–127.

[29] M.J. Antony, B.P. Sankaralingam, S. Khan, A. Almjally, N.A. Almujally, R.K. Mahendran, Brain–computer interface: The HOL–SSA decomposition and two-phase classification on the HGD EEG data, Diagnostics 13 (17) (2023) 2852.

[30] H. Fu, G. Sun, A. Zhang, B. Shao, J. Ren, X. Jia, Tensor singular spectral analysis for 3D feature extraction in hyperspectral images, IEEE Trans. Geosci. Remote Sens. 61 (2023) 1–14.

[31] J. Huang, L. Cui, Tensor singular spectrum decomposition: Multisensor denoising algorithm and application, IEEE Trans. Instrum. Meas. 72 (2023) 1–15.

[32] J. Huang, F. Zhang, B. Safaei, Z. Qin, F. Chu, The flexible tensor singular value decomposition and its applications in multisensor signal fusion processing, Mech. Syst. Signal Process. 220 (2024) 111662.

[33] S. Joshi, P. Dreesen, P. Bonizzi, J. Karel, R. Peeters, M. Boussé, Novel tensor-based singular spectrum decomposition, in: Proc. Eur. Signal Proc. Conf., 2024, pp. 1327–1331.

[34] L.R. Tucker, Some mathematical notes on three-mode factor analysis, Psychometrika 31 (3) (1966) 279–311.

[35] C.J. Hillar, L.-H. Lim, Most tensor problems are NP-hard, J. ACM 60 (6) (2013) 45.

[36] V. De Silva, L.-H. Lim, Tensor rank and the ill-posedness of the best low-rank approximation problem, SIAM J. Matrix Anal. Appl. 30 (3) (2008) 1084–1127.

[37] L. De Lathauwer, B. De Moor, J. Vandewalle, On the best rank-1 and rank-$(r_1,r_2,\ldots,r_n)$ approximation of higher-order tensors, SIAM J. Matrix Anal. Appl. 21 (2000) 1324–1342.

[38] Y. Jian, X.-C. Xiao, Higher-order singular-spectrum analysis of nonlinear time series, Acta Phys. Sin. 47 (6) (1998) 897–905.

[39] S. Luo, et al., An intelligent fault diagnosis model for rotating machinery based on multi-scale higher order singular spectrum analysis and GA-VPMCD, Measurement 87 (2016) 38–50.

[40] M. Kilmer, C. Martin, Factorization strategies for third-order tensors, Linear Algebra Appl. 435 (3) (2011) 641–658.

[41] N. Golyandina, Particularities and commonalities of singular spectrum analysis as a method of time series analysis and signal processing, WIREs Comp. Stat. 12 (4) (2020) e1487.

[42] M. Boussé, O. Debals, L. De Lathauwer, A tensor-based method for large-scale blind source separation using segmentation, IEEE Trans. Signal Process. 65 (2) (2017) 346–358.

[43] R. Badeau, G. Richard, B. David, Performance of ESPRIT for estimating mixtures of complex exponentials modulated by polynomials, IEEE Trans. Signal Process. 56 (2) (2008) 492–504.

[44] K. Usevich, V. Emiya, D. Brie, C. Chaux, Characterization of finite signals with low-rank STFT, in: Proc. IEEE Stat. Signal Process. Works., 2018, pp. 393–397.

[45] V. Emiya, R. Hamon, C. Chaux, Being low-rank in the time-frequency plane, in: Proc. IEEE Int. Conf. Acoust. Speech Signal Process., 2018, pp. 4659–4663.
[46] Z. Zhang, S. Aeron, Exact tensor completion using T-SVD, IEEE Trans. Signal Process. 65 (6) (2017) 1511–1526.
[47] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, S. Yan, Tensor robust principal component analysis with a new tensor nuclear norm, IEEE Trans. Pattern Anal. Mach. Intell. 42 (4) (2020) 925–938.
[48] U. Von Luxburg, A tutorial on spectral clustering, Stat. Comput. 17 (2007) 395–416.
[49] D. Farina, A. Holobar, Characterization of human motor units from surface EMG decomposition, Proc. IEEE 104 (2) (2016) 353–373.
[50] T.T. Le, K. Abed-Meraim, P. Ravier, O. Buttelli, A. Holobar, Tensor decomposition meets blind source separation, Signal Process. (2024) 109483.
[51] M. Atikur Rahman Khan, A note on window length selection in singular spectrum analysis, Aust. N. Z. J. Stat. 55 (2) (2013) 87–108.
[52] N. Golyandina, On the choice of parameters in singular spectrum analysis and related subspace-based methods, Stat. Interface (2010) 259–279.
[53] J.D.K. Abel, S. Dhanalakshmi, R. Kumar, A comprehensive survey on signal processing and machine learning techniques for non-invasive fetal ECG extraction, Multimedia Tools Appl. 82 (1) (2023) 1373–1400.
[54] L.T. Thanh, K. Abed-Meraim, P. Ravier, O. Buttelli, A. Holobar, Joint INDSCAL decomposition meets blind source separation, in: Proc. IEEE Int. Conf. Acoust. Speech Signal Process., 2024, pp. 5570–5574.
[55] M. Fatourechi, A. Bashashati, R.K. Ward, G.E. Birch, EMG and EOG artifacts in brain computer interface systems: A survey, Clin. Neurophysiol. 118 (3) (2007) 480–494.
[56] H. Peng, B. Hu, Q. Shi, M. Ratcliffe, Q. Zhao, Y. Qi, G. Gao, Removal of ocular artifacts in EEG—An improved approach combining DWT and ANC for portable applications, IEEE J. Biomed. Heal. Inf. 17 (3) (2013) 600–607.
[57] V. Krishnaveni, S. Jayaraman, L. Anitha, K. Ramadoss, Removal of ocular artifacts from EEG using adaptive thresholding of wavelet coefficients, J. Neural Eng. 3 (4) (2006) 338.
[58] F. Amzica, M. Steriade, Electrophysiological correlates of sleep delta waves, Electroencephalogr. Clin. Neurophysiol. 107 (2) (1998) 69–83.