Higher-Order Singular Spectrum Analysis For Multichannel Biomedical Signal Analysis

Thanh Trung Le[*,†], Karim Abed-Meraim[†,‡], Nguyen Linh Trung[*]

[*] VNU University of Engineering and Technology, Hanoi, Vietnam
[†]University of Orléans, Orléans, France
[‡] Institut Universitaire de France (IUF), Paris, France

***Abstract*—Singular spectrum analysis (SSA) is a nonparametric spectral estimation method that decomposes a time series signal into interpretable components. In this paper, we introduce a new multilinear generalization of SSA, called higher-order multivariate SSA (HO-MSSA), specifically designed to handle multivariate and multichannel time-series signals. The proposed method employs time-delay embedding and tensor SVD to transform multichannel time-series signals into trajectory tensors, subsequently decomposing them into distinct components. Similar to the classical SSA, these components hold significant interpretability. The effectiveness of HO-MSSA is illustrated through its application to the analysis of multichannel biomedical signals.**

***Index Terms*—Singular spectrum analysis, multivariate time series, tensor decomposition, spectral clustering.**

I. Introduction

Singular spectrum analysis (SSA) stands out as a powerful technique for time series analysis [1]. SSA decomposes a signal into several interpretable components, such as slowly time-varying trends, harmonic components, and noise. Notably, there are no assumptions about parametric models or stationarity-type conditions for the time series, making SSA a model-free method with broad applicability. We refer the readers to [1], [2] for good references on the SSA literature.

Recently, there has been an increasing interest in the analysis of multivariate and high-dimensional data, particularly with the rise in big data [3]. In the literature, several multivariate SSA (MSSA) techniques have emerged for analyzing multivariate time series, following the approach of Golyandina *et al.* in [1] for univariate time series and in [4] for multivariate time series. Noteworthy among these techniques are methods like vertical MSSA and horizontal MSSA in [5], 2D-SSA in [6], 2D-SSA in [7], and principal component trajectories (PCT) [8], among others (see [9] 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.[1] 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 statistics.

In parallel, tensors, which are multiway arrays, provide natural representations for multivariate and high-dimensional data [10]. Tensor decomposition allows for the factorization of tensors into basic components (e.g., vectors, matrices, or "simpler" tensors), making it a powerful data processing tool with applications in various areas [10]–[13]. Our study aims to develop a tensor-based MSSA approach. Previous studies have made efforts to develop tensor-based (M)SSA methods [14]–[19]. However, they are often tailored to specific datasets or applications, such as sleep EEG [14], [15], fault diagnosis [16]–[18], and hyperspectral image analysis [19], limiting their applicability. Therefore, the need for a generalized tensor-based (M)SSA approach is evident.

In this paper, we introduce a multilinear extension of SSA designed to handle multichannel time series, which we refer to as higher-order multivariate SSA (HO-MSSA).[2] Specifically, we employ a modified time-delay embedding 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 SVD (tSVD) technique [22], 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. Last but not least, we validate the efficacy of HO-MSSA with the application of extracting fetal ECG from maternal ECG signals.

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 $(i_1, i_2, \ldots, i_n)$-th element of $\boldsymbol{\mathcal{X}}$ is denoted as $\boldsymbol{\mathcal{X}}(i_1, i_2, \ldots, i_n)$. The transpose operation is represented as $(\cdot)^\top$, and the Frobenius norm as $\|\cdot\|_F$. Symbols $\bullet$, $\boxplus$, and $*$ represents the t-product, tensor concatenation and circular convolution, respectively. The functions "$\mathrm{fft}(\cdot)$" and "$\mathrm{ifft}(\cdot)$" denote fast Fourier transform and its inverse operator.

[1]The trajectory matrix in MSSA may take various forms, including Hankel, stacked Hankel, Hankel-block-Hankel, or quasi-Hankel matrices, depending on the techniques employed.

[2]In the literature, a variant of SSA is discussed in [20], [21], also referred to as higher-order SSA. This variant integrates basic SSA with higher-order statistics of univariate time series. Notably, 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.

II. Preliminaries

A. Singular Spectrum Analysis (SSA)

SSA is performed in four main steps, namely embedding, decomposition, grouping, and diagonal averaging [1].

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

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

where W is the window length. Here, $\mathbf{X}$ 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 $\mathbf{x}$.

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

$$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^\top = \textstyle\sum_{k=1}^{K} \lambda_k \mathbf{u}_k \mathbf{v}_k^\top, \quad (2)$$

where $K = \mathrm{rank}(\mathbf{X})$; $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_K]$ and $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_K]$ are left and right singular vector matrices, and they are orthogonal matrices; and $\lambda_k = \mathbf{S}(k,k)$ is the k-th singular value of $\mathbf{X}$. The collection $(\lambda_k, \mathbf{u}_k, \mathbf{v}_k)$ is k-th eigentriple of the SVD of $\mathbf{X}$. Rows and columns of $\mathbf{X}$ 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

$$\mathbf{X} = \textstyle\sum_{r=1}^{R} \mathbf{X}_r, \text{ where } \mathbf{X}_r = \sum_{i\in I_r} \lambda_i \mathbf{u}_i \mathbf{v}_i^\top. \quad (3)$$

One of the most widely-used technique for extracting components $\{\mathbf{X}_r\}_{r=1}^R$ is to use the matrix of W-correlations [1]. A necessary condition for the (approximate) separability of two time series is the (approximate) zero W-correlation of the reconstructed components. 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 (3) are Hankel matrices, facilitating the direct extraction of corresponding time series. However, in practice, such perfect separation may not be satisfied and we require a method to transform an arbitrary matrix into a Hankel matrix and subsequently into a signal, as follows

$$\tilde{\mathbf{x}}_r(n) = \begin{cases} \dfrac{1}{n}\displaystyle\sum_{j=1}^{n} \mathbf{X}_r(j, n-j+1) & \text{if } 1 \le n < W \\ \dfrac{1}{W}\displaystyle\sum_{j=1}^{W} \mathbf{X}_r(j, n-j+1) & \text{if } W \le n < J \\ \dfrac{1}{N-n+1}\displaystyle\sum_{j=n-J+1}^{W} \mathbf{X}_r(j, n-j+1) & \text{otherwise.} \end{cases} \quad (4)$$

This step is called diagonal averaging or de-hankelization.

B. Tensor Singular Value Decomposition

Tensor SVD (t-SVD) is a multiway extension of SVD for factorizing higher-order tensors [22]. 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:

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{U}} \bullet \boldsymbol{\mathcal{S}} \bullet \boldsymbol{\mathcal{V}}^\top, \quad (5)$$

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. The t-product "$\bullet$" is defined as follows

$$\boldsymbol{\mathcal{C}} = \boldsymbol{\mathcal{A}} \bullet \boldsymbol{\mathcal{B}} \Leftrightarrow \boldsymbol{\mathcal{C}}(i,j,:) = \textstyle\sum_{l=1} \boldsymbol{\mathcal{A}}(i,l,:) * \boldsymbol{\mathcal{B}}(l,j,:). \quad (6)$$

Note that when $n_3 = 1$, the t-product becomes the matrix product and (5) 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 [22].

III. Higher-Order Multivariate Singular Spectrum Analysis

In this section, we introduce HO-MSSA (Higher-Order MSSA), a multilinear extension of SSA for handling multi-channel and multivariate time-series signals. Similar to basis SSA, HO-MSSA contains four main steps, including time delay embedding, tensor SVD, grouping and reconstruction, as illustrated in Fig. 1.

Step 1. (***Time Delay Embedding***): In this step, we employ the following time delay embedding (TDE) transformation

$$\mathrm{TDE}_{W,\delta}(\mathbf{x}) = \begin{bmatrix} x(1) & x(\delta+1) & \dots & x(I\delta+1) \\ x(2) & x(\delta+2) & \dots & x(I\delta+2) \\ \vdots & \vdots & \dots & \vdots \\ x(W) & x(\delta+W) & \dots & x(I\delta+W) \end{bmatrix}, \quad (7)$$

with a window length $W \ge 2$ and a step size $\delta \ge 1$. Here, TDE can be considered as a generalized version of Hankelization (1) and segmentation [23] techniques for embedding time-series signals. When $\delta = 1$, (7) corresponds to Hankelization (1), while it becomes the segmentation if $\delta = W$.

Interestingly, Boussé *et al.* in [23] presented a common class of time-series signals in which (1) can promote low-rank approximation to signals. Specifically, if $\mathbf{x}$ is a sum of products of polynomial and exponential signals, then the corresponding Hankel matrix $\mathrm{TDE}_{W,1}(\mathbf{x})$ is low rank. This property can be summarized in 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* $x(t) = \sum_{r=1}^{R} p_r(t) z_r^t \sin(\omega_r t)$, *where* $z_r, \omega_r \in \mathbb{R}$. *Then,* $\mathrm{rank}(\mathrm{TDE}_{W,1}(\mathbf{x})) = \min(W, 2R + \sum_{r=1}^{R} Q_r)$.

As $\mathrm{TDE}_{W,\delta}(\cdot)$ is a submatrix of $\mathrm{TDE}_{W,1}(\cdot)$ with $\delta > 1$, $\mathrm{rank}(\mathrm{TDE}_{W,\delta}(\cdot)) \le \mathrm{rank}(\mathrm{TDE}_{W,1}(\cdot))\ \forall W, \delta$ and hence its rank is upper bounded by $2R + \sum_{r=1}^{R} Q_r$. In other words, the TDE in (7) further enhances low rank representation for polynomial and exponential signals.

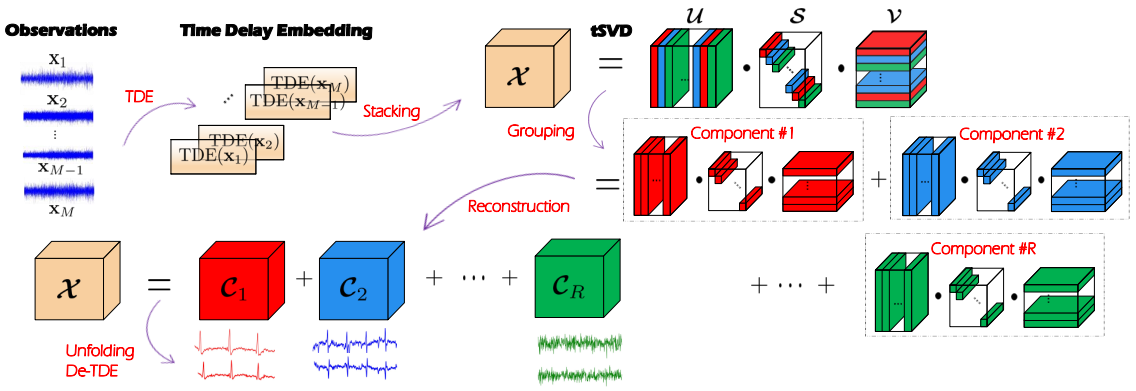


Fig. 1: Main steps of HO-MSSA for multichannel biomedical signal analysis.

Moreover, TDE has a connection to the short time Fourier transform (STFT) as follows

$$\widehat{\mathbf{X}} = \mathbf{F}_{N\times W}\mathrm{diag}(\boldsymbol{\omega})\mathrm{TDE}_{W,\delta}(\mathbf{x}), \tag{8}$$

where $\mathbf{F}_{N\times W}$ contains the first W columns of the Fourier transform matrix and $\boldsymbol{\omega}$ denotes the window in STFT. Since $\mathbf{F}_{N\times W}$ is full column rank, $\mathrm{rank}(\mathrm{TDE}_{W,\delta}(\mathbf{x})) = \mathrm{rank}(\widehat{\mathbf{X}})$. In [24], Usevich *et al.* presented several finite signals with low-rank STFT. Therefore, apart from polynomial and exponential signals, TDE also promotes low-rank approximation to several other types of data, such as convoluted 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 [2]. Therefore, TDE can offer a valuable method to obtain a low-rank representation of biomedical signals. Refer to Fig. 5 for an illustration of an ECG signal and its low-rank TDE matrix.

Step 2. (***Tensor SVD Decomposition***): In this step, we construct the trajectory tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{W\times(I+1)\times M}$ by stacking TDE matrices of M observations $\{\mathbf{x}_m\}_{m=1}^M$ as follows

$$\boldsymbol{\mathcal{X}} = \mathrm{TDE}_{W,\delta}(\mathbf{x}_1) \boxplus \mathrm{TDE}_{W,\delta}(\mathbf{x}_2) \cdots \boxplus \mathrm{TDE}_{W,\delta}(\mathbf{x}_M). \tag{9}$$

Subsequently, taking the t-SVD of $\boldsymbol{\mathcal{X}}$ results in:

$$\begin{aligned}\boldsymbol{\mathcal{X}} &\overset{\text{t-SVD}}{=} \boldsymbol{\mathcal{U}} \bullet \boldsymbol{\mathcal{S}} \bullet \boldsymbol{\mathcal{V}}^\top \\ &= \textstyle\sum_{k=1}^K \boldsymbol{\mathcal{U}}(:,k,:) \bullet \boldsymbol{\mathcal{S}}(k,k,:) \bullet \boldsymbol{\mathcal{V}}(k,:,:)^\top = \sum_{k=1}^K \boldsymbol{\mathcal{X}}_k,\end{aligned} \tag{10}$$

where $K \le \min(W, I+1)$ is the tubal rank of $\boldsymbol{\mathcal{X}}$; $\boldsymbol{\mathcal{U}} \in \mathbb{R}^{W\times K\times M}$ and $\boldsymbol{\mathcal{V}} \in \mathbb{R}^{K\times(I+1)\times M}$ are orthogonal tensors; and $\boldsymbol{\mathcal{S}} \in \mathbb{R}^{K\times K\times M}$ is an f-diagonal tensor; and tubal rank-1 tensors $\{\boldsymbol{\mathcal{X}}_k\}_{k=1}^K$ are regarded as elementary tensors in this application, see Fig. 2 for an illustration. Expression (10) is regarded as the truncated version of t-SVD. The value of K is identified as the number of non-zero tubes of $\boldsymbol{\mathcal{S}}$, i.e., $K = \sum_k \mathbf{1}[\boldsymbol{\mathcal{S}}(k,k,:) \neq \mathbf{0}]$ where $\mathbf{1}$ is an indicator function. Here, we can apply Algorithm 1 to perform the tensor decomposition (10) effectively.

Similar to SVD in basis SSA, t-SVD also offers the best low-tubal-rank representation to trajectory tensors. In particular, among all tensors $\boldsymbol{\mathcal{Y}}$ of tubal rank $d \le K$, the tensor $\boldsymbol{\mathcal{Y}} = \sum_{k=1}^d \boldsymbol{\mathcal{X}}_k$ provides the best approximation to $\boldsymbol{\mathcal{X}}$ in the sense that $\|\boldsymbol{\mathcal{X}} - \boldsymbol{\mathcal{Y}}\|_F^2$ is minimum [22, Theorem 4.3]. Note that, t-SVD factorizes the trajectory tensor $\boldsymbol{\mathcal{X}}$ in the Fourier domain rather than the time domain, as in basis SSA and other variants. Indeed, expression (8) reveals that the Fourier transform $\widehat{\mathbf{X}}$ of the original trajectory matrix $\mathbf{X}$ can be expressed as $\widehat{\mathbf{X}} = \widehat{\mathbf{U}}\mathbf{S}\mathbf{V}^\top$ where $\widehat{\mathbf{U}} = \mathbf{F}_{N\times W}\mathrm{diag}(\boldsymbol{\omega})\mathbf{U}$, and $\mathbf{U}, \mathbf{S}$, and $\mathbf{V}$ are three factors of the SVD of $\mathbf{X}$. It suggests that the decomposition of trajectory matrices can be performed in the Fourier domain, as outlined in Algorithm 1. Specifically, when $M = 1$ (univariate time series), t-SVD simplifies to the classical SVD, thus making HO-MSSA equivalent to SSA.

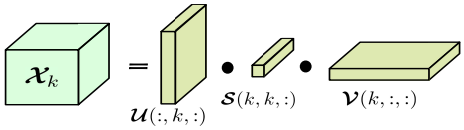


Fig. 2: Elementary tensor of tubal rank 1.

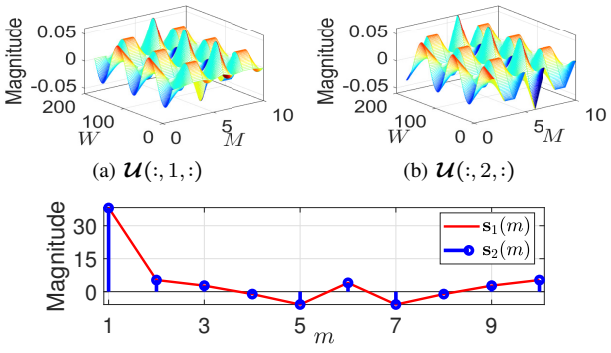
Algorithm 1: t-SVD [25]

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

Step 3. (***Grouping***): The aim of this step is to divide the set of elementary tensors $\{\boldsymbol{\mathcal{X}}_k\}_{k=1}^K$ (i.e., $\{\boldsymbol{\mathcal{U}}(:,k,:), \boldsymbol{\mathcal{S}}(k,k,:), \boldsymbol{\mathcal{V}}(k,:,:)\}_{k=1}^K$) into R disjoint clusters $\{I_r\}_{r=1}^R$ and then form:

$$\boldsymbol{\mathcal{X}} = \textstyle\sum_{r=1}^R \boldsymbol{\mathcal{C}}_r \quad \text{where} \quad \boldsymbol{\mathcal{C}}_r = \sum_{i\in I_r} \boldsymbol{\mathcal{X}}_i. \tag{11}$$

Let $\mathbf{s}_k = \mathrm{vec}(\boldsymbol{\mathcal{S}}(k,k,:)) \in \mathbb{R}^{M\times 1}$ $\forall k$. We exploit the following observations: (i) the set of first elements $\{\mathbf{s}_k(1)\}_{k=1}^K$ sorts in decreasing order, i.e., $\mathbf{s}_1(1) \ge \mathbf{s}_2(1) \ge \cdots \ge \mathbf{s}_K(1)$, which plays a similar role as singular values of the trajectory matrix in basis SSA; (ii) the remaining elements in $\mathbf{s}_k(2:\mathrm{end})$ are



(a) $\boldsymbol{\mathcal{U}}(:,1,:)$ (b) $\boldsymbol{\mathcal{U}}(:,2,:)$

(c) Two tubes $\boldsymbol{\mathcal{S}}(1,1,:)$ and $\boldsymbol{\mathcal{S}}(2,2,:)$ 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 corresponding tubal-rank is 2.

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).

Consequently, when $M > 2$, the set of $\{\mathbf{s}_k\}_{k=1}^K$ can effectively serve as features to perform the grouping (11).[3] Here, we apply the spectral clustering method [26] 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. Noting that, other clustering methods in machine learning can also perform this task.

Step 4. (***Reconstruction***): After extracting $\boldsymbol{\mathcal{C}}_r$ 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 $\boldsymbol{\mathcal{C}}_r$. Below, we present an extended version of (4), called block diagonal averaging, to recover the underlying signal $\mathbf{x}$ from its TDE matrix, accommodating various window lengths W and step sizes in the range $1 \le \delta \le W$. Refer to Fig. 4 for an illustration.

Let $L = \lfloor W/\delta \rfloor$, $\boldsymbol{j}_\delta = [\delta(j-1)+1 : \delta j]$ 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 (7) into two parts

$$\mathrm{TDE}_{W,\delta}(\mathbf{x}) = \left[\frac{\mathbf{X}_\delta}{\mathbf{B}}\right] = \left[\begin{array}{cccc} \mathbf{x}_{\mathbf{1}_\delta} & \mathbf{x}_{\mathbf{2}_\delta} & \cdots & \mathbf{x}_{\boldsymbol{J}_\delta} \\ \mathbf{x}_{\mathbf{2}_\delta} & \mathbf{x}_{\mathbf{3}_\delta} & \cdots & \mathbf{x}_{(\boldsymbol{J}+\mathbf{1})_\delta} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{x}_{\boldsymbol{L}_\delta} & \mathbf{x}_{(\boldsymbol{L}+\mathbf{1})_\delta} & \cdots & \mathbf{x}_{(\boldsymbol{L}+\boldsymbol{J}-\mathbf{1})_\delta} \\ \hline \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_J \end{array}\right], \tag{12}$$

where $\mathbf{B}$ has a small number of rows (i.e., $W - L\delta < \delta$ rows). It is empty if $W = L\delta$. 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,

[3]When $M = 1$, HO-MSSA simplifies to basis SSA and $\{\mathbf{s}_k\}_{k=1}^K$ represents the singular values of the trajectory matrix. This spectrum also proves valuable for (eigentriple) grouping, particularly when processing harmonic signals.

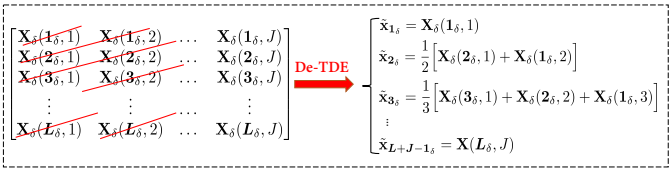
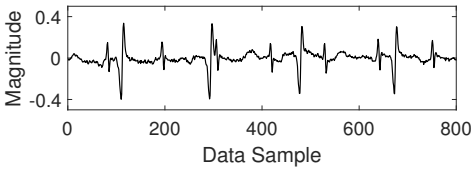
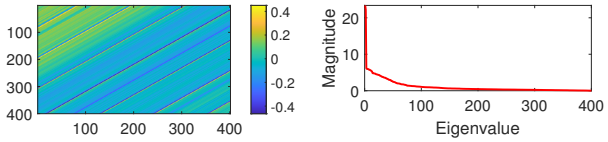


Fig. 4: Block Diagonal Averaging



(a) ECG Signal



(b) TDE Matrix (c) Spectrum of TDE

Fig. 5: TDE promotes low-rank approximation to ECG signals.

exploiting the block Hankel structure of $\mathbf{X}_\delta$ in (12), we apply the following block diagonal averaging for the recovery of $\mathbf{x}$

$$\tilde{\mathbf{x}} = [\tilde{\mathbf{x}}_{\mathbf{1}_\delta}^\top \;\; \tilde{\mathbf{x}}_{\mathbf{2}_\delta}^\top \;\; \cdots \;\; \tilde{\mathbf{x}}_{(\boldsymbol{L}+\boldsymbol{J}-\mathbf{1})_\delta}^\top \;\; \mathbf{b}_J^\top]^\top, \text{ where} \tag{13}$$

$$\tilde{\mathbf{x}}_{\boldsymbol{n}_\delta} = \begin{cases} \dfrac{1}{n}\displaystyle\sum_{j=1}^{n} \mathbf{X}_\delta(\boldsymbol{j}_\delta, n-j+1) & \text{if } 1 \le n < L \\ \dfrac{1}{L}\displaystyle\sum_{j=1}^{L} \mathbf{X}_\delta(\boldsymbol{j}_\delta, n-j+1) & \text{if } L \le n < J \\ \dfrac{1}{L+J-n}\displaystyle\sum_{j=n-J+1}^{L} \mathbf{X}_\delta(\boldsymbol{j}_\delta, n-j+1) & \text{otherwise.} \end{cases} \tag{14}$$

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

IV. HO-MSSA Based Fetal ECG Extraction From Maternal ECG

In this section, we demonstrate the effectiveness of HO-MSSA in extracting fetal electrocardiogram (fetal ECG) from maternal ECG recordings. The dataset used for this task contains five abdominal and three thoracic recordings acquired from various regions of the mother's body, each sampled at a rate of 250 Hz.[4] Here, we focus on five abdominal recordings, each consisting of 800 data samples only. Refer to Fig. 5(a) for an illustration of an abdominal recording.

We set the window length W to 400, with a step-size δ of 1, resulting in a trajectory tensor $\boldsymbol{\mathcal{X}}$ of size $400 \times 401 \times 5$. Figs. 5(b) and (c) show the TDE matrix and its corresponding singular values for an ECG recording. We can see that the spectrum exhibits rapid decay, with focusing on the first few singular values. This property of the TDE matrix facilitates a low-rank approximation of the ECG signal and hence the trajectory tensor $\boldsymbol{\mathcal{X}}$.

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

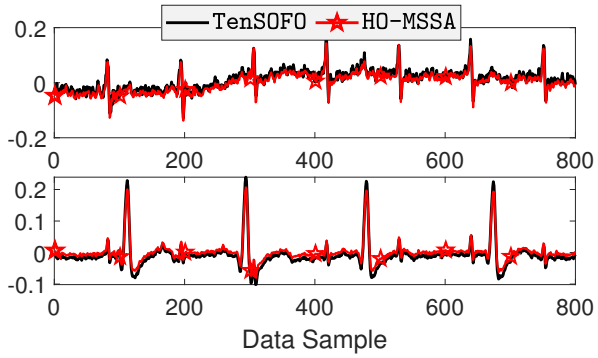


Fig. 6: ECG separation results for the first ECG recording: fetal heartbeats (above) and maternal heartbeats (below).

We proceed by factorizing $\boldsymbol{\mathcal{X}}$ using t-SVD. The tubal rank of $\boldsymbol{\mathcal{X}}$ is determined to be $K = 92$,[5] representing to 92 elementary tensor components. Subsequently, we utilize spectral clustering to categorize these components into three main groups: fetal ECG, maternal ECG, and noises. This clustering method is implemented using normalized graph Laplacian with a Gaussian similarity function defined as $A(\mathbf{s}_i, \mathbf{s}_j) = \exp(-\|\mathbf{s}_i - \mathbf{s}_j\|_2^2/(2\sigma^2))$, where σ controls the width of the neighborhoods, set to 1 for this experiment.[6] Afterwards, time-series signals are extracted from each ECG recording through block diagonal averaging on the frontal slices of the corrresponding reconstructed tensor. We also compare the performance of HO-MSSA with the recently proposed tensor-based ECG extraction method called TenSOFO in [27]. The experimental results are illustrated in Fig. 6. Similar to TenSOFO, HO-MSSA can separate fetal heartbeats from maternal heartbeats, with the fetal heart rate consistently higher than that of the mother. Note that, the primary objective of this section is to offer an illustrative example of HO-MSSA for biomedical signal analysis. A more comprehensive performance analysis and comparison will be provided in our forthcoming journal.

V. Conclusions

In this study, we proposed a novel extension of singular spectrum analysis (SSA) 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 t-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 shows HO-MSSA's potential for multivariate and high-dimensional data analysis that will be investigated thoroughly in future work.

References

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

[5]We use the function "tubalrank.m" from t-product toolbox to estimate K. Access it at: https://github.com/canyilu/tensor-tensor-product-toolbox.

[6]In Matlab, we can simply use the following command "spectralcluster ($\mathbf{A}$,3,'Distance','precomputed','LaplacianNormalization','symmetric')" where "$\mathbf{A}$ = exp(−dist.^2)" and "dist = squareform(pdist($\mathbf{S}$))".

[2] S. Sanei and H. Hassani, *Singular Spectrum Analysis of Biomedical Signals*. CRC press, 2015.
[3] I. Koch, *Analysis of Multivariate and High-Dimensional Data*. Cambridge University Press, 2013.
[4] N. Golyandina and D. Stepanov, "SSA-based approaches to analysis and forecast of multidimensional time series," in *Proc. St. Petersburg Works. Simulation*, vol. 293, 2005, p. 298.
[5] H. Hassani and R. Mahmoudvand, "Multivariate singular spectrum analysis: A general view and new vector forecasting approach," *Int. J. Energy Stat.*, vol. 1, no. 01, pp. 55–83, 2013.
[6] N. E. Golyandina and K. Usevich, "2d-extension of singular spectrum analysis: algorithm and elements of theory," in *Matrix Methods: Theory, Algorithms and Applications*, 2010, pp. 449–473.
[7] J. Zabalza, J. Ren, J. Zheng, J. Han, H. Zhao, S. Li, and S. Marshall, "Novel two-dimensional singular spectrum analysis for effective feature extraction and data classification in hyperspectral imaging," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 8, pp. 4418–4433, 2015.
[8] D. Dylewsky, E. Kaiser, S. L. Brunton, and J. N. Kutz, "Principal component trajectories for modeling spectrally continuous dynamics as forced linear systems," *Phys. Rev. E*, vol. 105, no. 1, p. 015312, 2022.
[9] N. Golyandina, A. Korobeynikov, A. Shlemov, and K. Usevich, "Multivariate and 2D extensions of singular spectrum analysis with the Rssa package," *J. Stat. Softw.*, vol. 67, no. 2, pp. 1–78, 2015.
[10] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.
[11] A. Cichocki *et al.*, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 145–163, 2015.
[12] N. D. Sidiropoulos *et al.*, "Tensor decomposition for signal processing and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3551–3582, 2017.
[13] L. T. Thanh, K. Abed-Meraim, N. L. Trung, and A. Hafiane, "A contemporary and comprehensive survey on streaming tensor decomposition," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 11, pp. 10 897–10 921, 2023.
[14] S. Kouchaki and S. Sanei, "Tensor based singular spectrum analysis for nonstationary source separation," in *Proc. IEEE Int. Works. Mach. Learn. Signal Process.*, 2013, pp. 1–5.
[15] S. Kouchaki, S. Sanei, E. L. Arbon, and D.-J. Dijk, "Tensor based singular spectrum analysis for automatic scoring of sleep EEG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, no. 1, pp. 1–9, 2014.
[16] C. Yi, Y. Lv, M. Ge, H. Xiao, and X. Yu, "Tensor singular spectrum decomposition algorithm based on permutation entropy for rolling bearing fault diagnosis," *Entropy*, vol. 19, no. 4, p. 139, 2017.
[17] 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.*, vol. 7, no. 4, p. 418, 2017.
[18] J. Huang and L. Cui, "Tensor singular spectrum decomposition: Multisensor denoising algorithm and application," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–15, 2023.
[19] H. Fu, G. Sun, A. Zhang, B. Shao, J. Ren, and X. Jia, "Tensor singular spectral analysis for 3d feature extraction in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, pp. 1–14, 2023.
[20] Y. Jian *et al.*, "Higher-order singular-spectrum analysis of nonlinear time series," *Acta Phys. Sin.*, vol. 47, no. 6, pp. 897–905, 1998.
[21] 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*, vol. 87, pp. 38–50, 2016.
[22] M. Kilmer and C. Martin, "Factorization strategies for third-order tensors," *Linear Algebra Appl.*, vol. 435, no. 3, pp. 641–658, 2011.
[23] M. Boussé, O. Debals, and L. De Lathauwer, "A tensor-based method for large-scale blind source separation using segmentation," *IEEE Trans. Signal Process.*, vol. 65, no. 2, pp. 346–358, 2017.
[24] K. Usevich, V. Emiya, D. Brie, and C. Chaux, "Characterization of finite signals with low-rank STFT," in *Proc. IEEE Stat. Signal Process. Works.*, 2018, pp. 393–397.
[25] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan, "Tensor robust principal component analysis with a new tensor nuclear norm," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 4, pp. 925–938, 2020.
[26] U. Von Luxburg, "A tutorial on spectral clustering," *Stat. Comput.*, vol. 17, pp. 395–416, 2007.
[27] T. T. Le, K. Abed-Meraim, P. Ravier, O. Buttelli, and A. Holobar, "Tensor decomposition meets blind source separation," *Signal Process.*, p. 109483, 2024.