Constrained Independent Vector Extraction of Quasi-Periodic Signals from Multiple Data Sets

Rencheng Song[a,b], Guoping Wang[c], Juan Cheng[a], Aiping Liu[d], Chang Li[a], Xun Chen[e,f]

[a]*Department of Biomedical Engineering, Hefei University of Technology, Hefei, 230009, Anhui, China*
[b]*Anhui Province Key Laboratory of Measuring Theory and Precision Instrument, Hefei University of Technology, Hefei 230009, Anhui, China*
[c]*Department of Cyberspace Security, University of Science and Technology of China, Hefei, 230026, Anhui, China*
[d]*School of Information Science and Technology, University of Science and Technology of China, Hefei, 230026, Anhui, China*
[e]*Department of Neurosurgery, the First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, 230001, Anhui, China*
[f]*Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei, 230026, Anhui, China*

Abstract

In this paper, we consider a problem to extract independent source component vector (SCV) formed by quasi-periodic signals from instantaneous mixtures in multiple data sets. We propose a method, termed as constrained independent vector extraction (CIVE), to uniquely determine the target quasi-periodic SCV. Specifically, the negentropy is taken to enforce the independence of the target SCV from the others, while the mutual information is used to determine the correlation of sources within the target SCV. A quasi-periodic constraint is further combined in the cost function to ensure the quasi-periodicity of the SCV. The demixing vectors of target SCV are solved as a constrained optimization problem by the Lagrange multiplier method. The CIVE method is designed to work under diverse probability distributions for the mixed signals. In the experiments, the CIVE method is verified with both simulated and semi-simulated data. The comparison results with other methods indicate the effectiveness, applicability and stability of the proposed method for extracting quasi-periodic SCVs.

Keywords: Joint blind source separation (JBSS), Constrained independent vector extraction(CIVE), Quasi-periodic signal

1. Introduction

Quasi-periodic signals appear commonly in communication, radar, and biomedical signal processing [1, 2]. In these fields, there is a strong demand to extract quasi-periodic components from instantaneously mixed signals in a single data set or multiple data sets. For example, in remote photoplethysmography (rPPG) applications [3, 4], it is known that the blood volume varies periodically with the heartbeat, leading to related changes in skin colors. The RGB channels obtained in each skin region of interest (ROI) are considered to be linear mixtures of pulse signal with noise signals. The purpose of rPPG is to extract the quasi-periodic pulse signal from a single skin ROI (single data set) or multiple skin ROIs (multiple data sets).

The blind source separation (BSS) method is a standard technique to demix source signals with observations from a single data set. Among these methods, the independent component analysis (ICA) [5, 6] is a very commonly-used method to solve BSS problems based on statistical independence. The negentropy [7], mutual information [8], or likelihood function [5] is usually defined in the cost functions of ICA to enforce the statistical properties of source signals. Different from the ICA which only relies on statistical assumptions, the constrained independent component analysis (CICA) [9] can further introduce prior information as a constraint into the source separation process, which transforms the original problem into a constrained optimization problem. Due to the use of constraints, CICA can only separate the target signal, which can be regarded as a blind source extraction (BSE) method.

There exist many studies using BSS or BSE to extract periodic or quasi-periodic signals from a single data set. For example, the AMUSE (algorithm for multiple unknown signals extraction) [10] and the SOBI (second-order blind identification) [11] are efficient second-order methods for separating quasi-periodic signals, which explore the time-correlation structure of target signals. Zhi-Lin Zhang [12] used CICA to get periodic signals through defining a constraint with a proper reference signal. Richard Macwan et al. [13] employed the autocorrelation as a constraint to

Email addresses: rcsong@hfut.edu.cn (Rencheng Song), wgp@mail.ustc.edu.cn (Guoping Wang), chengjuan@hfut.edu.cn (Juan Cheng), aipingl@ustc.edu.cn (Aiping Liu), changli@hfut.edu.cn (Chang Li), xunchen@ustc.edu.cn (Xun Chen)

extract a quasi-periodic signal with CICA. Thato Tsalaile and others [14] combined the sequential blind extraction algorithm [15] and the time-varying lag calculation procedure [16] to extract quasi-periodic signals. Hassan Akbari et al. developed a method based on Tucker decomposition [17, 18] and quasi-periodic nature to extract the quasi-periodic signal [19]. In general, the BSS methods have the issue of source permutations [9] if all sources need to be decoupled. The performance of BSE methods are highly dependent on the designed constraint.

The use of multiple data sets to obtain the shared quasi-periodic signals will fully utilize the correlation of the target signals in each data set, which provides more useful information for overcoming the issue of BSS or BSE. In each data set, the quasi-periodic source signal is assumed to linearly mix with other non-target source signals. The target quasi-periodic source signals are dependent on each other, and independent with other non-target source signals across all data sets. The correlated source signals compose a source component vector (SCV) [20]. The purpose of this study is to determine the target quasi-periodic SCV from the multiple data sets.

According to the problem definition, the joint blind source separation (JBSS) is a feasible scheme to solve the target problem. Independent vector analysis (IVA) [20, 21, 22, 23], the extension of ICA on multiple data sets, is a typical JBSS method to extract all independent SCVs from multiple data sets. The cost function of IVA consists of mutual information to maximize the independence between the SCVs and the dependence within the SCVs [22]. IVA is very efficient to solve problems that meet its assumptions. However, IVA may not converge if the actual problems break its assumptions. For example, the signals within the non-target SCV are independent or the non-target SCVs are dependent with each other. Besides, IVA needs to separate all SCVs, which reduces the efficiency if only the target SCV with some prior information like quasi-periodicity is desired.

Many methods [24, 25] improve the performance of IVA through incorporating prior information into the cost functions. For example, the constrained independent vector analysis (CIVA) [24, 26] combine new constraints from prior information to separate SCVs. Geometrically constrained IVA [25] added a geometrical penalty term to the cost function of IVA to extract the desired SCV. Supervised independent vector analysis (SIVA) [27] introduced additional supervising components named pilot signals as prior information. The above methods improve the performance of IVA through restricting the separation with prior information. But these methods still require to determine the entire demixing matrices.

Unlike IVA to separate all SCVs, independent vector extraction (IVE) [28, 29] can determine only one non-Gaussian independent SCV from a set of instantaneous mixtures based on the maximum likelihood principle. IVE is the extension of independent component extraction (ICE), of which objective function is the log-likelihood function of target signal. However, the ICE/IVE assumes the target signal is non-Gaussian with Gaussian background and it requires initials of the mixing vector or demixing vector. This strong assumption limits its application in solve the target problem because the initial value is usually not available and the distribution of the quasi-periodic signal may be diverse.

In this paper, we propose a novel method, named as constrained independent vector extraction (CIVE), to extract the quasi-periodic SCV from multiple data sets. The mutual information and negentropy are used together in the cost function of CIVE. Particularly, the mutual information is taken to enforce the dependence within the target SCV, while the negentropy is taken to enforce the independence of target SCV with other non-target SCVs [7]. The CIVE can only calculate the target unmixing vector instead of the full demixing matrix, thereby making the unique extraction of the target SCV possible. Finally, the autocorrelation of the target SCV is maximized in the loss as a constraint to ensure the quasi-periodicity [13]. The constrained optimization problem can be solved by the augmented Lagrange multiplier method. In experiments, we verify the efficiency, stability, and accuracy of the proposed method using both simulated and semi-simulated data. The comparison results with some other typical methods demonstrate the superior performance of CIVE to solve the target problems.

The main contribution of this paper is that we propose an effective method to determine the quasi-periodic SCV from multiple sets of instantaneous mixtures. The CIVE method can extract only the target SCV, which avoids the issue of target selection in other BSS methods, thereby improving the efficiency of the separation. In contrast, the IVA methods like CIVA still need to solve the full unmixing matrices, even if we only need one target SCV. Meanwhile, the CIVE does not require initial guesses of the demixing vectors like IVE. The CIVE method is also applicable for extracting target quasi-periodic SCVs with diverse probability distributions. Therefore, it has a wide range of applications and the comparison results demonstrate a superior performance over existing methods for solving the target problems.

The structure of following contents is as follows. In Section 2, we describe the problem and the proposed algorithm. The experiments are demonstrated in Section 3 and finally we conclude the paper in Section 4.

2. Method

2.1. The problem formulation

The formulation of the target problem is similar to that of JBSS in [23]. Suppose there are K data sets and each contains N observed signals, which are linear mixtures of a target quasi-periodic source signal and $N-1$ non-target source signals,

$$\mathbf{x}^{[k]} = \mathbf{A}^{[k]}\mathbf{s}^{[k]}, 1 \leq k \leq K,$$

where $\mathbf{x}^{[k]} = [\mathbf{x}_1^{[k]}, \cdots, \mathbf{x}_N^{[k]}]^T$ is the kth data set (with size of $N \times M$, N is the total number of sources in each single data set, and M is the total number of sampling points), the superscript T denotes the transpose, and $\mathbf{x}_i^{[k]} = [x_i^{[k]}(1), \ldots, x_i^{[k]}(M)]$ is the ith observed signal (a row vector) in the kth data set. The $\mathbf{s}^{[k]} = [\mathbf{s}_1^{[k]}, \cdots, \mathbf{s}_N^{[k]}]^T$ is the source signal matrix (with size of $N \times M$) corresponding to the kth data set, where $\mathbf{s}_i^{[k]} = [s_i^{[k]}(1), \ldots, s_i^{[k]}(M)]$ indicates the ith source signal (a row vector) in the kth data set with M sampling points. Here $\mathbf{A}^{[k]} \in \mathbb{R}^{N \times N}$, is the invertible mixing matrix of the kth data set. The source signals in the same sequence in each data set form a SCV, such as the nth SCV is $\mathbf{s}_n = [\mathbf{s}_n^{[1]}, \cdots, \mathbf{s}_n^{[K]}]^T$. For ease of description, we use the symbol $\mathbf{s}_p$ hereafter to denote the target quasi-periodic SCV. The total K source signals in a SCV are dependent on each other and they are independent with the source signals in other SCVs.

The purpose of this paper is to find K demixing vectors so as to get the estimations of $\mathbf{s}_p$. The kth demixing vector of $\mathbf{s}_p$ is denoted as $(\mathbf{w}_p^{[k]})^T$, which is the pth row of the demixing matrix $\mathbf{W}^{[k]}$ with $\mathbf{W}^{[k]} = (\mathbf{A}^{[k]})^{-1}$ [20]. The estimation of the kth target source signal is represented as $\mathbf{y}_p^{[k]} = (\mathbf{w}_p^{[k]})^T \mathbf{x}^{[k]}$, where $\mathbf{y}_p^{[k]}$ indicates the estimation of $\mathbf{s}_p^{[k]}$, and $\mathbf{y}_p = [\mathbf{y}_p^{[1]}, \cdots, \mathbf{y}_p^{[K]}]^T$ is the estimation of $\mathbf{s}_p$. We assume all source signals have unit variance with zero mean. Below we will show the algorithm to calculate the K demixing vectors.

2.2. Cost function

In order to determine the K demixing vectors of target SCV, we maximize the cost function composed of the negentropy and the mutual information as defined in Eq. (1),

$$\begin{aligned} \mathcal{I}_{\mathrm{CIVE}} &= \alpha \sum_{k=1}^{K} J(\mathbf{y}_p^{[k]}) + I(\mathbf{y}_p) \\ &= \alpha \sum_{k=1}^{K} J(\mathbf{y}_p^{[k]}) + \sum_{k=1}^{K} \mathcal{H}(\mathbf{y}_p^{[k]}) - \mathcal{H}(\mathbf{y}_p), \end{aligned} \tag{1}$$

where $\sum_{i=1}^{K} J(\mathbf{y}_p^{[i]})$ is the negentropy of all signals in the target SCV to enforce the independence of the target SCV with other SCVs, $I(\mathbf{y}_p)$ is the mutual information function to ensure the dependence of signals within the target SCV, $\mathcal{H}(\mathbf{y}_p)$ is information entropy of $\mathbf{y}_p$, and α is a weight constant to balance the two terms. Unless necessary, α is ignored in the following derivation.

Considering the difficulty of calculating $\sum_{i=1}^{K} J(\mathbf{y}_p^{[i]})$, we choose an approximation of the negentropy as $J(\mathbf{y}_p^{[k]}) \approx \rho[E\{G(y_p^{[k]})\} - E\{G(v)\}]^2$, where ρ is a constant, v is a Gaussian variable with unit variance and zero mean, and $G(\cdot)$ is a non-quadratic function [5]. As suggested by [30], we choose the function $G(\cdot)$ as

$$G(y_p^{[k]}) = \frac{\log(\cosh(a \cdot y_p^{[k]}))}{a}, a \geq 1 \tag{2}$$

where $\cosh(\cdot)$ is the hyperbolic cosine function and a is a constant with default value as 1.0.

To ensure the quasi-periodicity of target SCV, we further add an autocorrelation constraint [13] in the cost function. The autocorrelation constraint has an inequality form as

$$g(\mathbf{w}_p^{[k]}) = \xi^{[k]} - E\{(r_p^{[k]})^2\} \leq 0, 1 \leq k \leq K \tag{3}$$

where $\xi^{[k]}$ denotes the threshold for the lower bound of the autocorrelation of $\mathbf{y}_p^{[k]}$, $E\{(r_p^{[k]})^2\}$ is the mathematical expectation of the variable $(r_p^{[k]})^2$, and $r_p^{[k]}$ is the autocorrelation of $\mathbf{y}_p^{[k]}$ as

$$r_p^{[k]}(m) = \sum_{i=0}^{M-m} (y_p^{[k]}(i) y_p^{[k]}(m+i)), 1 \leq m \leq M, 1 \leq k \leq K. \tag{4}$$

Furthermore, we constrain $\mathbf{y}_p^{[k]}$ to have an unit variance

$$h(\mathbf{w}_p^{[k]}) = E\{(y_p^{[k]})^2 - 1\} = 0, 1 \leq k \leq K. \tag{5}$$

In summary, the full cost function of the target problem can be defined as

$$\begin{aligned} &\text{Maximize} : \mathcal{I}_{\mathrm{CIVE}}(\mathbf{w}_p^{[k]}) = \alpha J(\mathbf{y}_p^{[k]}) + \mathcal{H}(\mathbf{y}_p^{[k]}) - \mathcal{H}(\mathbf{y}_p) + c_k \\ &\text{Subject to} : g(\mathbf{w}_p^{[k]}) \leq 0, h(\mathbf{w}_p^{[k]}) = 0, 1 \leq k \leq K \end{aligned} \tag{6}$$

where $c_k = \alpha \sum_{i=1, i \neq k}^{K} J(\mathbf{y}_p^{[i]}) + \sum_{i=1, i \neq k}^{K} \mathcal{H}(\mathbf{y}_p^{[i]})$ contains all terms that are fixed with respect to changes in $\mathbf{w}_p^{[k]}$.

2.3. Optimization of the cost function

We use the augmented Lagrange multiplier method to solve the constrained optimization problem in Eq. (6). In detail, we rewrite the cost function as

$$\begin{aligned} &\mathcal{L}(\mathbf{w}_p^{[k]}, \mu^{[k]}, \lambda^{[k]}) = \mathcal{I}_{\mathrm{CIVE}}(\mathbf{w}_p^{[k]}) - \frac{1}{2\gamma}\{-(\mu^{[k]})^2 \\ &+[\max\{0, \mu^{[k]} + \gamma g(\mathbf{w}_p^{[k]})\}]^2\} - \lambda^{[k]} h(\mathbf{w}_p^{[k]}) + \frac{1}{2}\beta ||h(\mathbf{w}_p^{[k]})||^2, \end{aligned} \tag{7}$$

where $\mu^{[k]}$ and $\lambda^{[k]}$ are the Lagrange multipliers corresponding to $g(\mathbf{w}_p^{[k]})$ and $h(\mathbf{w}_p^{[k]})$, respectively, the γ and β are penalty parameters, and $||\cdot||$ denotes the Euclidean norm.

We take the Newton algorithm to update the demixing vectors $\mathbf{w}_p^{[k]}$. The iterative equations of $\mathbf{w}_p^{[k]}$, $\mu^{[k]}$ and $\lambda^{[k]}$ in (7) are given as follows:

$$\begin{aligned} \mathbf{w}_{p,new}^{[k]} &= \mathbf{w}_{p,old}^{[k]} - \mathbf{H}^{-1} \frac{\partial \mathcal{L}}{\partial \mathbf{w}_{p,old}^{[k]}}, \\ \mu_{new}^{[k]} &= \max\{0, \mu_{old}^{[k]} + \gamma g(\mathbf{w}_{p,old}^{[k]})\}, \\ \lambda_{new}^{[k]} &= \lambda_{old}^{[k]} + \beta h(\mathbf{w}_{p,old}^{[k]}), \end{aligned} \tag{8}$$

where $\mathbf{H} = \frac{\partial^2 \mathcal{L}}{\partial(\mathbf{w}_{p,old}^{[k]})\partial(\mathbf{w}_{p,old}^{[k]})^T}$ is the Hessian matrix. During update, variable with the subscript *old* denotes the result of last update, while variable with the subscript *new* indicates the result of the current update. More details about the augmented Lagrange multiplier method can be referred to [9, 31].

According to the update formulas in Eq. (8), we need to calculate the gradient vector and Hessian matrix of the augmented Lagrange function (7). Particularly, the gradient vector of the $\mathcal{L}(\mathbf{w}_p^{[k]}, \mu^{[k]}, \lambda^{[k]})$ with respect to $\mathbf{w}_p^{[k]}$ is

$$\begin{aligned}\frac{\partial \mathcal{L}(\mathbf{w}_p^{[k]}, \mu^{[k]}, \lambda^{[k]})}{\partial(\mathbf{w}_p^{[k]})} &= \overline{\rho} E\{\mathbf{x}^{[k]}(:,m)\tanh(y_p^{[k]})\} \\ + E\{(\phi^{[k]}(y_p^{[k]}))\mathbf{x}^{[k]}(:,m)\} &- E\{\mathbf{x}^{[k]}(:,m)\phi^{[k]}(\mathbf{y}_p(:,m))\} \\ - \mu^{[k]}\frac{\partial g(\mathbf{w}_p^{[k]})}{\partial \mathbf{w}_p^{[k]}} &- \lambda^{[k]} E\{\mathbf{x}^{[k]}(:,m) y_p^{[k]}\} \quad (9)\end{aligned}$$

where $\overline{\rho} = \rho E\{G(y_p^{[k]})\} - E\{G(v)\}$, $\phi^{[k]}(y_p^{[k]}) = -\frac{\partial \log\{p(y_p^{[k]})\}}{\partial y_p^{[k]}}$ and $\phi^{[k]}(\mathbf{y}_p(:,m)) = -\frac{\partial \log\{p(\mathbf{y}_p(:,m))\}}{\partial y_p^{[k]}}$ are score functions [22], which need to be calculated with explicit assumption of the probability distribution $p(\mathbf{y}_p)$ for the target SCV. Here $\mathbf{x}^{[k]}(:,m)$ denotes a multivariate vector formed by the variables of the observed signals in the kth data set. $\mathbf{y}_p(:,m)$ indicates the multivariate vector of the target SCV. If there is no specific declaration, we use the symbol (:,m) to indicate the multivariate vector hereafter. The assumption and calculation of $p(\mathbf{y}_p)$ are described in the next section.

The first-order derivative of $g(\mathbf{w}_p^{[k]})$ in (9) is derived as

$$\frac{\partial g(\mathbf{w}_p^{[k]})}{\partial \mathbf{w}_p^{[k]}} = -\frac{2}{M}\mathbf{x}^{[k]} \sum_{m=1}^{M} \mathbf{U}(m) r_p^{[k]}(m)(\mathbf{y}_p^{[k]})^T, \quad (10)$$

where $\mathbf{V}(m) = \begin{bmatrix} \mathbf{0}_{(M-m)\times m} & \mathbf{I}_{(M-m)\times(M-m)} \\ \mathbf{0}_{m\times m} & \mathbf{0}_{m\times(M-m)} \end{bmatrix}$ and $\mathbf{U}(m) = \mathbf{V}(m) + \mathbf{V}(m)^T$. The $\mathbf{I}_{(M-m)\times(M-m)}$ denotes a $(M-m)\times(M-m)$ identity matrix, and $\mathbf{0}_{(M-m)\times m}$ denotes a $(M-m)\times m$ zero matrix.

The Hessian matrix of $\mathcal{L}(\mathbf{w}_p^{[k]}, \mu^{[k]}, \lambda^{[k]})$ is

$$\begin{aligned}\frac{\partial^2 \mathcal{L}(\mathbf{w}_p^{[k]}, \mu^{[k]}, \lambda^{[k]})}{\partial(\mathbf{w}_p^{[k]})\partial(\mathbf{w}_p^{[k]})^T} &= \overline{\rho}\mathbf{R}_x^{[k]} E\{1 - \tanh^2(y_p^{[k]})\} \\ + E\{\frac{\partial \phi^{[k]}(y_p^{[k]})}{\partial \mathbf{w}_p^{[k]}}(\mathbf{x}^{[k]}(:,m))^T\} &- E\{\frac{\partial \phi^{[k]}(\mathbf{y}_p(:,m))}{\partial \mathbf{w}_p^{[k]}}(\mathbf{x}^{[k]}(:,m))^T\} \\ - \mu^{[k]}\frac{\partial^2 g(\mathbf{w}_p^{[k]})}{\partial(\mathbf{w}_p^{[k]})\partial(\mathbf{w}_p^{[k]})^T} &- \lambda^{[k]}, \quad (11)\end{aligned}$$

where $\mathbf{R}_x^{[k]} \triangleq E\{\mathbf{x}^{[k]}(:,m)(\mathbf{x}^{[k]}(:,m))^T\}$ is the covariance matrix of $\mathbf{x}^{[k]}(:,m)$. The second-order derivative of $g(\mathbf{w}_p^{[k]})$ in (11) is

$$\begin{aligned}&\frac{\partial^2 g(\mathbf{w}_p^{[k]})}{\partial(\mathbf{w}_p^{[k]})\partial(\mathbf{w}_p^{[k]})^T} \\ &= -\frac{2}{M}\mathbf{x}^{[k]}\{\sum_{m=1}^{M}\mathbf{U}(m)[\mathbf{y}_p^{[k]}\mathbf{U}(m)(\mathbf{y}_p^{[k]})^T + r_p^{[k]}(m)]\}(\mathbf{x}^{[k]})^T. \quad (12)\end{aligned}$$

The detailed derivation of $\frac{\partial g(\mathbf{w}_p^{[k]})}{\partial \mathbf{w}_p^{[k]}}$ and $\frac{\partial^2 g(\mathbf{w}_p^{[k]})}{\partial(\mathbf{w}_p^{[k]})\partial(\mathbf{w}_p^{[k]})^T}$ can be found in [13].

We observe that the derivations of $E\{\frac{\partial \phi^{[k]}(y_p^{[k]})}{\partial \mathbf{w}_p^{[k]}}(\mathbf{x}^{[k]}(:,m))^T\}$ and $E\{\frac{\partial \phi^{[k]}(\mathbf{y}_p(:,m))}{\partial \mathbf{w}_p^{[k]}}(\mathbf{x}^{[k]}(:,m))^T\}$ are not complete if the probability distribution $p(\mathbf{y}_p)$ of the target SCV is not provided. We will discuss it in the next section.

2.4. Probability distributions of target SCV

The target SCV is composed of quasi-periodic signals and the probability distribution $p(\mathbf{y}_p(:,m))$ of multivariate $\mathbf{y}_p(:,m)$ can be Gaussian or non-Gaussian. In this paper, we assume that $p(\mathbf{y}_p(:,m))$ is multivariate Gaussian [22] or multivariate Laplace [32]. We name the CIVE using multivariate Gaussian distribution as CIVE-G and the CIVE using Laplacian distribution as CIVE-L.

First, we finish the calculation of (9) and (11) following the multivariate Gaussian distribution,

$$\begin{aligned}&p(\mathbf{y}_p(:,m)|\mathbf{\Sigma}_p) \\ &= \frac{1}{(2\pi)^{K/2}\det(\mathbf{\Sigma}_p)^{1/2}}\exp(-\frac{1}{2}(\mathbf{y}_p(:,m))^T\mathbf{\Sigma}_p^{-1}\mathbf{y}_p(:,m)) \quad (13)\end{aligned}$$

where the covariance matrix of the target SCV $\mathbf{s}_p(:,m)$ is $\mathbf{\Sigma}_p \triangleq E\{\mathbf{s}_p(:,m)(\mathbf{s}_p(:,m))^T\}$ because the sources are assumed to have zero mean. We can get $\phi^{[k]}(\mathbf{y}_p(:,m))$ and $\phi^{[k]}(\mathbf{y}_p^{[k]}(:,m))$ in Eq. (9) as

$$\begin{aligned}\phi^{[k]}(\mathbf{y}_p(:,m)) &= -\frac{\partial \log\{p(\mathbf{y}_p(:,m))\}}{\partial y_p^{[k]}} \\ &= \mathbf{e}_k^T\mathbf{\Sigma}_p^{-1}\mathbf{y}_p(:,m) = (\mathbf{y}_p(:,m))^T\mathbf{\Sigma}_p^{-1}\mathbf{e}_k \quad (14)\end{aligned}$$

and

$$\phi^{[k]}(y_p^{[k]}) = -\frac{\partial \log\{p(y_p^{[k]})\}}{\partial y_p^{[k]}} = \frac{y_p^{[k]}}{(\sigma_p^{[k]})^2} \quad (15)$$

where $(\sigma_p^{[k]})^2$ is the variance of $y_p^{[k]}$, and $\mathbf{e}_k$ represents the kth column of the identity matrix $\mathbf{I}_{K\times K}$. The detailed derivation of (14) can be found in [22].

In practice, the approximations of $\mathbf{\Sigma}_p^{-1}$ and $(\sigma_p^{[k]})^2$ are used instead, which can be estimated using the samples of $\mathbf{y}_p$ and $\mathbf{y}_p^{[k]}$, respectively. The approximation of $\mathbf{\Sigma}_p$ is

$$\hat{\mathbf{\Sigma}}_p = \frac{1}{M}\Sigma_{m=1}^{M}\mathbf{y}_{p,old}(:,m)\mathbf{y}_{p,old}^T(:,m), \quad (16)$$

and the approximation of $(\sigma_p^{[k]})^2$ is

$$(\hat{\sigma}_p^{[k]})^2 = \frac{1}{M}\Sigma_{m=1}^{M}(y_{p,old}^{[k]}(m))^2. \quad (17)$$

Therefore, we can calculate $E\{\phi^{[k]}(\mathbf{y}_p(:,m))\mathbf{x}^{[k]}(:,m)\}$ and $E\{\phi^{[k]}(y_p^{[k]})\mathbf{x}^{[k]}(:,m)\}$ in (9) as

$$E\{\phi^{[k]}(\mathbf{y}_p(:,m))\mathbf{x}^{[k]}(:,m)\} = E\{\mathbf{x}^{[k]}(:,m)(\mathbf{y}_p(:,m))^T\}\mathbf{\Sigma}_p^{-1}\mathbf{e}_k \quad (18)$$

and

$$E\{\phi^{[k]}(y_p^{[k]})\mathbf{x}^{[k]}(:,m)\} = E\{\frac{y_p^{[k]}}{(\sigma_p^{[k]})^2}\mathbf{x}^{[k]}(:,m)\}. \quad (19)$$

Also, we can calculate $E\{\frac{\partial\phi^{[k]}(\mathbf{y}_p(:,m))}{\partial\mathbf{w}_p^{[k]}}(\mathbf{x}^{[k]}(:,m))^T\}$ and $E\{\frac{\partial\phi^{[k]}(y_p^{[k]})}{\partial\mathbf{w}_p^{[k]}}(\mathbf{x}^{[k]}(:,m))^T\}$ in (11) as

$$E\{\frac{\partial\phi^{[k]}(\mathbf{y}_p(:,m))}{\partial\mathbf{w}_p^{[k]}}(\mathbf{x}^{[k]}(:,m))^T\} = \mathbf{\Sigma}_p^{-1}\mathbf{R}_x^{[k]} \quad (20)$$

$$E\{\frac{\partial\phi^{[k]}(y_p^{[k]})}{\partial\mathbf{w}_p^{[k]}}(\mathbf{x}^{[k]}(:,m))^T\} = \frac{\mathbf{R}_x^{[k]}}{(\sigma_p^{[k]})^2}. \quad (21)$$

The derivation with $p(\mathbf{y}_p(:,m))$ following the multivariate Laplace distribution [32, 33] can be done in similar way and it is omitted here.

Although we only assume two types of probability distributions, the proposed method can also be applied to extract the target SCV under sub-Gaussian probability distributions, such as uniform probability distribution. We will prove it in the experiments.

3. Experiments

In this section, we evaluate the effectiveness of the proposed CIVE method for extracting quasi-periodic SCV. To prove the capability of the CIVE method, we mainly compare it with existing JBSS methods including IVA [21, 22], CIVA [26], and IVE [28], using simulated and semi-simulated data under different probability distributions. Finally, we also compare the CIVE with several classical single-set BSS methods, including the SOBI [11], the PiCA (periodic component analysis) [16], and the CICA [13], which are considered to be effective for extracting quasi-periodic signals. The comparison can demonstrate the advantages of the proposed method for joint extraction of quasi-periodic signals from multiple correlated signal sets. In the following experiments, the target signals are always mixed linearly with non-target signals.

3.1. Settings of comparison methods

Some implementation details and settings of all the comparison methods in the following experiments are introduced here. The IVA-G and IVA-L are implemented with the open source IVA code [1]. Particularly, IVA-G indicates the IVA based on second-order statistics and the IVA-L is the IVA based on higher-order statistics.

The IVE here indicates the OGIVEs from [28]. It requires an initial value of the true mixing vector. Suppose $\mathbf{a}_{ini}^{[k]} = \mathbf{a}^{[k]} + \mathbf{e}_{ini}$, where $\mathbf{a}^{[k]}$ is the true mixing vector, and $\mathbf{e}_{ini}$ is a perturbation vector. The error of the initial can be measured as $\varepsilon^2 = ||\mathbf{e}_{ini}||^2$. In this paper, we use the symbol IVE-1, IVE-0.1, IVE-0.001 to represent the IVE with initial error ε as 1.0, 0.1, and 0.001, respectively. The smaller the ε is, the closer the initial value to its true mixing vector $\mathbf{a}^{[k]}$. More details of the IVE and ε can be referred to [28]. We use IVE-r to represent the IVE with the initial value created from a Gaussian random number with zero mean and unit variance. The IVE is implemented with the gradient IVE code [2].

For CIVA, a single reference signal or a single reference mixing vector needs to be provided to extract the target SCV. In the experiments, we follow a similar way as IVE to define the reference mixing vector of CIVA. An initial mixing vector with error $\varepsilon = 1.0$ is used in CIVA. It should be mentioned that the reference of mixing vector for CIVA is difficult to obtain in reality. The implementation of CIVA is also done with the open source code [1].

In CIVE, there are two hyper-parameters, the lower bound $\xi^{[k]}$ of the autocorrelation of $y_p^{[k]}$, and the weight α in the cost function. For $\xi^{[k]}$, we set it as an estimated value based on observed signals. It can also be set as some constant based on experience. Since the negentropy can only be used to determine non-Gaussian component, we set $\alpha = 1$ for non-Gaussian target SCV, and $\alpha = 0$ for Gaussian target SCV.

For single-set separation methods, the settings of SOBI and PiCA are determined by the observation signals in each data set. We calculate the autocorrelations of each observation signal and get its dominant peak position. The average position t of the dominant peaks is used to determine the time lag of source signals. In SOBI, the number of time-delayed covariance matrices to be diagonalized is set to $1.5t$. In PiCA, the peak indexes of the source signal are set to those ones according to the obtained t. The CICA [13] only keeps the autocorrelation constraint and removes the constraint of a reference signal, because the reference of source signal is supposed to be unknown in the experiment. The implementation of SOBI refers to the open source code [3], while the PiCA is done with the code [4].

[1] http://mlsp.umbc.edu/resources.html
[2] https://asap.ite.tul.cz/downloads/ice/
[3] https://github.com/aludnam/MATLAB/blob/master/sobi/sobi.m
[4] https://github.com/marianux/ecg-kit/blob/master/common/PiCA.m

3.2. Metrics

In the experiments, the evaluation metrics are introduced first focusing on the accuracy of extractions.

First, we take a modification of joint inter-symbol-interference (joint ISI or $\mathrm{ISI}_{\mathrm{JNT}}$)[23, 34, 35], named as $\mathrm{ISI}_{\mathrm{JNT}}-\mathrm{v}$, to assess the accuracy of the extractions, where v indicates vector. As known, we only need a single demixing vector instead of a demixing matrix to extract the target signal for each data set. Accordingly, the definition of $\mathrm{ISI}_{\mathrm{JNT}}-\mathrm{v}$ is as below

$$\mathrm{ISI}_{\mathrm{JNT}}-\mathrm{v} \triangleq \frac{1}{N}\left(\sum_{n=1}^{N}\left(\frac{\tilde{f}_n}{\max_n\{\tilde{f}_n\}}-1\right)\right), \tag{22}$$

where N is the total number of signals in one data set, $\tilde{f}_n = \sum_{k=1}^{K}\left|f_n^{[k]}\right|$, and $f_n^{[k]}$ is the nth element of the vector $\mathbf{f}^{[k]}$ with $\mathbf{f}^{[k]} = (\mathbf{w}_p^{[k]})^T\mathbf{A}^{[k]}$. Since $\mathbf{f}^{[k]} = (\mathbf{w}_p^{[k]})^T\mathbf{A}^{[k]}$ is ideally equivalent to the product of a nonzero scalar and a row of $\mathbf{I}_{K\times K}$, it means $(1-N)/N \le \mathrm{ISI}_{\mathrm{JNT}}-\mathrm{v} \le 0$ and $(1-N)/N$ implying exact extraction.

Besides the metric of $\mathrm{ISI}_{\mathrm{JNT}}-\mathrm{v}$, we also use the correlation coefficient between the estimates and the target source signals as another assessment metric of accuracy. In detail, we use the average correlation coefficient

$$r = \frac{1}{K}\sum_{k=1}^{K} r^{[k]}, \tag{23}$$

where $r^{[k]}, 1 \le k \le K$ is the correlation coefficient of the kth target source signal. We clarify that the letter r in this section only denotes the correlation coefficient instead of the autocorrelation coefficient in the last section.

3.3. Experimental results of simulated data

To comprehensively evaluate the performance of the proposed algorithm under different probability distributions, we construct six different types of simulated data as demonstrated in Table 1. Specifically, we consider three types of target source signals, following Gaussian, Laplacian and uniform distributions, respectively, and two types of non-target source signals, which obey Gaussian and Laplacian distributions, respectively. The six types of observed signals are labelled as A1, A2, B1, B2, C1 and C2, respectively, where the capital letter indicate the type of target source signal and the digit represents the type of non-target signal. For example, A2 means the observed data set is composed by linear mixtures of target source with Gaussian distribution and non-target signals with Laplacian distribution.

Table 1: Six types of simulated data.

Non-target \ Target	Gaussian	Laplacian	Uniform
Gaussian	A1	B1	C1
Laplacian	A2	B2	C2

3.3.1. Generation of simulated data

With reference to the data generation method in [22], we generate quasi-periodic SCVs from periodic signals with varying periodicity. Fig. 1 show three examples of quasi-periodic signals with different probability distributions. The simulated data are pre-whitened for all methods.

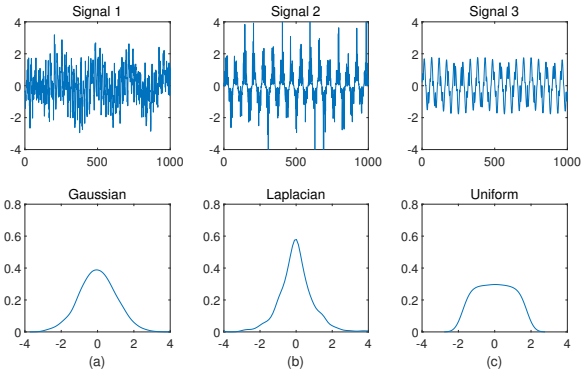

Figure 1: Quasi-periodic signals with different probability distributions. (a) shows a Gaussian quasi-periodic signal with a kurtosis of 2.98. (b) shows a Laplacian quasi-periodic signal with a kurtosis of 6.61. (c) shows a uniform quasi-periodic signal with a kurtosis of 1.85.

For target SCV following Gaussian distributions, we first generate K quasi-periodic signal $\theta(\cdot)$ by

$$\theta(m) = \sqrt{u(m)}v(m), 1 \le m \le M, \tag{24}$$

where $u(\cdot)$ is a signal following an exponential distribution, $v(\cdot)$ is a periodic signal with a random periodicity. Then, we get the Gaussian target SCV $\mathbf{s}_{p,g}$ by mixing the K quasi-periodic signals $\theta(m)$ [22].

For target SCV following Laplacian distributions, we get the quasi-periodic target SCV $\mathbf{s}_{p,l}$ with a similar way as [22],

$$\mathbf{s}_{p,l}(m) = \sqrt{u(m)}\mathbf{\Gamma}^{\frac{1}{2}}\boldsymbol{\theta}(m), 1 \le m \le M \tag{25}$$

where $\boldsymbol{\theta}(\cdot)$ is a vector formed by K quasi-periodic signals, and $\mathbf{\Gamma}$ is a randomly generated covariance matrix.

Similarly, for target SCV following uniform distributions, we get the target SCV $\mathbf{s}_{p,u}$ as

$$\mathbf{s}_{p,u}(m) = \sqrt{|v(m)|}\mathbf{\Gamma}^{\frac{1}{2}}\mathbf{z}(m), 1 \le m \le M \tag{26}$$

where $\mathbf{z}(\cdot)$ is a vector formed by K different periodic signals.

For non-target SCVs, we generate the multivariate Gaussian ones through mixing K random Gaussian signals, while multivariate Laplacian ones are constructed by

$$\mathbf{s}_{n,l}(m) = \sqrt{u(m)}\mathbf{\Gamma}^{\frac{1}{2}}\mathbf{g}(m), 1 \le m \le M \tag{27}$$

where $\mathbf{g}(\cdot)$ is a vector formed by K random Gaussian signals.

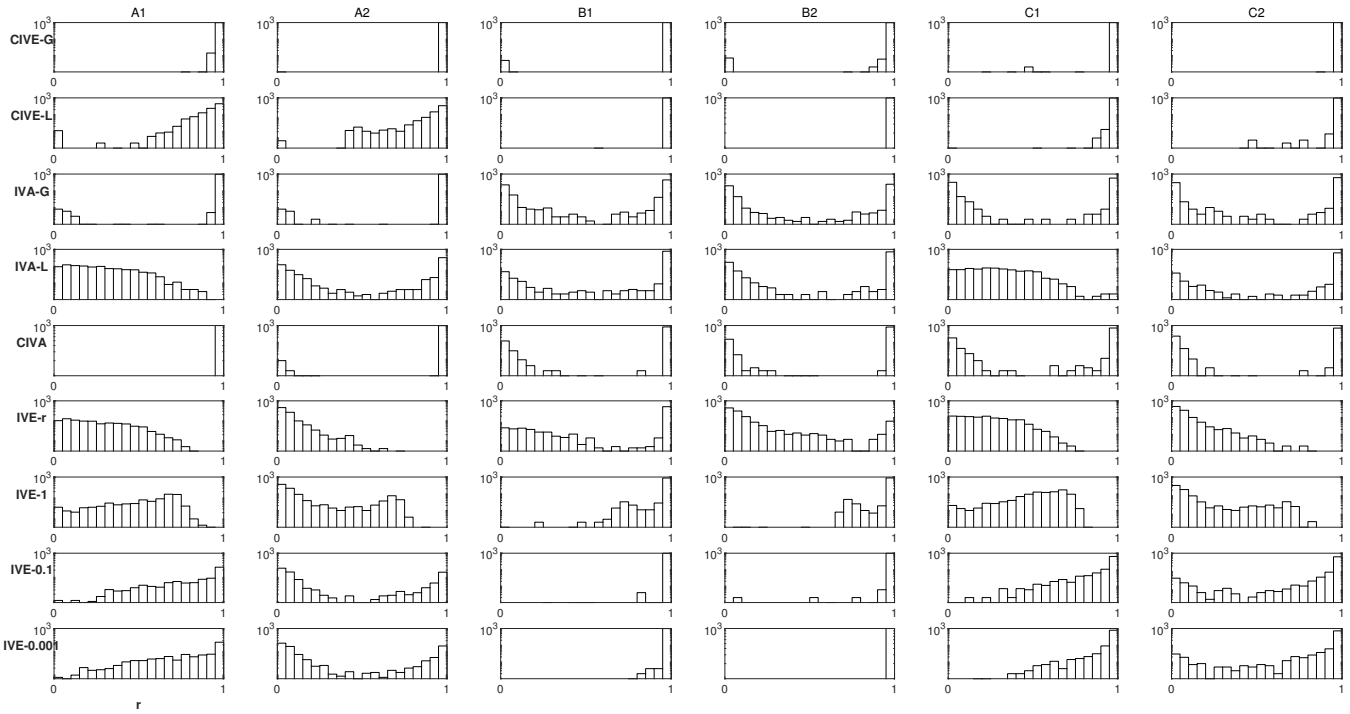

Figure 2: Histograms of r statistics for results in Table 2 using simulated data.

Table 2: The average r and average $\mathrm{ISI}_{\mathrm{JNT}}-\mathrm{v}$ for the results of 100 independent trials on simulated data with $N = 10, K = 10$ and $M = 1000$.

Metrics	$\mathrm{ISI}_{\mathrm{JNT}}-\mathrm{v}$	r	$\mathrm{ISI}_{\mathrm{JNT}}-\mathrm{v}$	r	$\mathrm{ISI}_{\mathrm{JNT}}-\mathrm{v}$	r
Data	A1		A2		B1	
CIVE-G	-0.86	0.99	**-0.89**	**1.00**	**-0.88**	0.99
CIVE-L	-0.80	0.91	-0.78	0.86	**-0.88**	**1.00**
IVA-G	-0.86	0.97	-0.88	0.98	-0.76	0.60
IVA-L	-0.28	0.27	-0.78	0.56	-0.79	0.85
CIVA	**-0.88**	**1.00**	-0.88	0.98	-0.87	0.84
IVE-r	-0.39	0.27	-0.52	0.09	-0.57	0.67
IVE-1	-0.60	0.51	-0.57	0.20	-0.81	0.96
IVE-0.1	-0.82	0.73	-0.76	0.40	**-0.88**	0.99
IVE-0.001	-0.82	0.72	-0.76	0.40	**-0.88**	**1.00**
Data	B2		C1		C2	
CIVE-G	-0.87	0.99	**-0.89**	**1.00**	**-0.89**	**1.00**
CIVE-L	**-0.89**	**1.00**	-0.87	0.99	-0.87	0.99
IVA-G	-0.54	0.49	-0.78	0.60	-0.78	0.64
IVA-L	-0.83	0.74	-0.31	0.30	-0.31	0.79
CIVA	-0.87	0.82	-0.84	0.76	-0.85	0.71
IVE-r	-0.54	0.18	-0.40	0.23	-0.40	0.09
IVE-1	-0.81	0.97	-0.60	0.53	-0.60	0.19
IVE-0.1	-0.88	0.99	-0.85	0.91	-0.85	0.84
IVE-0.001	-0.88	**1.00**	-0.86	0.95	-0.86	0.88

3.3.2. Accuracy of comparison methods

The number of total sources N, the length of source signal M and the number of data sets K are taken as $N = 10, M = 1000$, and $K = 10$, respectively. The average results of all methods after 100 independent trials are shown in Table 2. Since the number of sources to be demixed in each data set is $N = 10$, the ideal extraction indicates that the correlation coefficient r is close 1.0 and $\mathrm{ISI}_{\mathrm{JNT}}-\mathrm{v} = -0.90$. The histograms of r statistics are shown in Fig. 2. More r concentrating to 1.0 indicates better quality of signal extraction. Next, we analyze the performance of each method in detail.

1) IVA: As known, the IVA separates all SCVs. In the experiment, we select the target demixing vector of each data set as the one closest to the real one from the full demixing matrix. It should be noted that the selection of required target is not trivial in real JBSS applications since the reference is usually absent.

IVA-G is valid for separating both Gaussian and non-Gaussian source signals which have second-order correlations within the same SCV [20]. It can be seen from Table 2 that the extractions of IVA-G are accurate in A1 and A2 because the target source signals are second-order correlated. In comparison, the performance of IVA-G is degenerated in cases of B1, B2, C1 and C2 since the required second-order correlations in IVA-G may not be fully satisfied in these cases.

The IVA-L can separate SCVs with multivariate Laplacian distribution and higher-order dependency [22]. Therefore, the estimations of the target SCV by IVA-L in B1 and B2 are more accurate than that in A1, A2, C1 and C2. However, we also observe that the performance of IVA-L is not always satisfied even for the cases of B1 and B2. This is consistent with the observation described in [22]. Although the target source signals in C1 and C2 are higher-order dependent, the performance of IVA-L is degraded since the probability distribution of targets are far from Laplacian.

2) IVE: As known, IVE is an extraction method of non-Gaussian target SCV. From Table 2 and Fig. 2, we can see that the IVE achieves better performance in B1 and C1 than that in A1. For example, the metric r of IVE-0.1

in B1, C1 and A1 is 0.99, 0.92, and 0.73, respectively. The performance of IVE will improve as the signal becomes more non-Gaussian, which is consistent to its assumption [28].

From the Table 2, we also observe that the initial value of the demixing vector has a clear influence on the performance of IVE. The closer the initial value is to its reference, the more accurate the extractions are. Particularly, the IVE-r results with random initial value are totally distorted.

3) CIVA: As can be seen from Table 2, the performance of CIVA is similar to that of IVA-G. Both of the two methods achieve good performance for A1 and A2. Although the separation of CIVA for B1, B2, C1, and C2 is slightly better than IVA-G, it is still not accurate enough. This is because the objective function of CIVA is similar as that of IVA, which limits the separation performance.

4) CIVE: The CIVE-G indicates the method of CIVE with multivariate Gaussian assumption to determine the quasi-periodic SCV with second-order correlations. As can be observed in Table 2, the CIVE-G achieves amazing performance in the accuracy of extractions. In detail, the negentropy loss is ignored for cases of A1 and A2 and the CIVE-G still gets the best results. The use of mutual information and the autocorrelation constraint plays an important role for CIVE-G to extract the target signals in cases of A1 an A2 with strong second-order correlations and quasi-periodicity. The results of CIVE-G in B1, B2, C1 and C2, of which the target signals are non-Gaussian, are still satisfied. We guess the reason that the negentropy and the autocorrelation constraint still play a major role to separate the target SCV from other SCVs through enforcing the independence and quasi-periodicity.

The CIVE-L is the CIVE method with assumption of multivariate Laplacian distribution. The CIVE-L achieves the best performance in B1 and B2 when the problem assumption matches well with the method. In comparison, the extractions of CIVE-L in A1 and A2 are not as accurate as CIVE-G due to the mismatch of signal distribution. We also notice that the CIVE-L still performs well for SCV with uniform distributions in C1 and C2. Although the probability distributions do not match with the assumption, the high-order dependency of generated target signals in C1 and C2 improve the results of CIVE-L. This means that the relevant items in the loss function have important effects on the results, and the influencing factors are more complex.

3.3.3. Convergence and stability of CIVE

We evaluate the converge curves of different methods as shown in Fig. 3. Considering the convergence speeds of all methods are different, we only show the average results of the first 50 iterations. We observe that the CIVE-G has overall the best convergence among all the methods. In comparison, the CIVE-L performs better in cases of B1 and B2. The IVA-G converges quickly and accurately in A1 and A2. But it does not perform well in B1, B2, C1, and C2. The IVA-L converges slowly for all cases and it finally converges to the actual solution in B1 and B2 after around 1000 iterations. The converge curves of CIVA are similar as that of IVA-G, except for the B2 and C2, where the CIVA converges much faster than the IVA-G. We guess that the extra constraint in CIVA accelerates the convergence in these two cases. The IVE with different starting points fails to converge in the first 50 iterations. It eventually converges to a good point for B1, B2, C1, and C2 types of observation signals in the actual experiments after thousands of iterations. The above convergence results are consistent with the experimental results in Table 2.

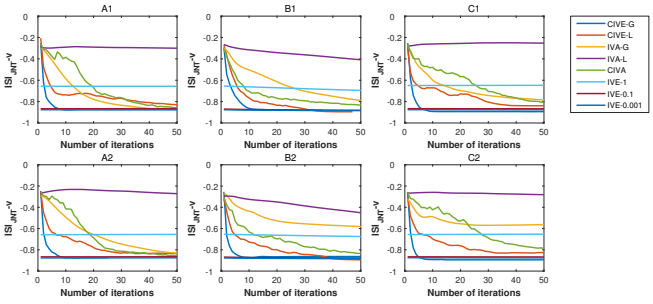

Figure 3: The converge curves of different methods using simulated data: the first 50 iterations.

In previous experiments, we take fixed values for the number of samples M, the number of SCVs N and the numbers of data sets K. In order to verify the stability of CIVE, we further run the experiment using different N, M and K. Since the results of other cases are similar, we only demonstrate the stability analysis with the case of C2. To save time, the extractions are independently implemented with 10 trials. Moreover, the results of IVE are taken from IVE-0.001 because the results of IVE-0.1, IVE-1 and IVE-r are worse.

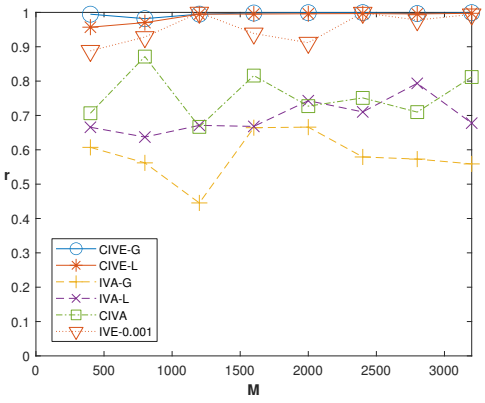

Figure 4: Mean correlation coefficients r of 10 independent trials versus the number of samples M with $N = 10$ and $K = 10$ in C2.

Fig. 4 shows the correlation coefficients r of different methods versus the number of samples M. We can see that M has little impact on the performance of CIVE-G and CIVE-L. Only the accuracy of the CIVE-L method slightly drops when M is extremely small. Fig. 5 (a) and

(b) show the performance of all methods versus different number K of data sets. Particularly, Fig. 5 (a) and (b) are obtained using $N = 4$ and $N = 20$, respectively. We can see that the CIVE-G and CIVE-L are quite stable along with different N and K. In contrast, we observe that the extractions of IVA-L in Fig. 5 (a) with $N = 4$ are more accurate than that in Fig. 5 (b) with $N = 20$. This is consistent to the conclusion in [22] that the IVA-L is easy to converge to a local minimum if the number N of SCVs is big.

3.4. Experimental results of semi-simulated data

Considering that the real quasi-periodic signals may have more complicated distributions than the above simulated data, we also evaluate the performance of the proposed method with simi-simulated data.

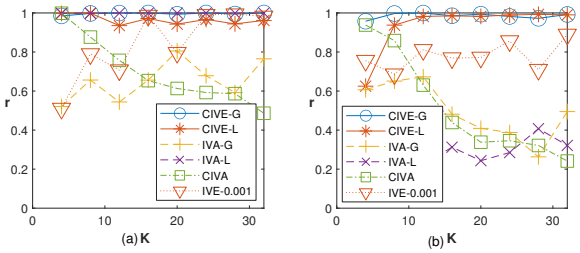

Figure 5: Mean correlation coefficients r of results from 10 independent trials versus different number K of data sets in C2 with $M = 1000$. (a)$N = 4$ and (b)$N = 20$.

3.4.1. Generation of semi-simulated data

Since the photoplethysmography (PPG) signal records the heartbeat information, it can be considered as a real quasi-periodic signal. In this paper, we randomly select PPG signals from three subjects in the public UBFC-RPPG dataset [36] to generate the semi-simulated data. The PPG signals were recorded by Contec Medical CMS50E at a sampling rate of 60 Hz. They are linearly mixed with simulated non-target Gaussian or Laplacian source signals to construct semi-simulated data sets.

Fig. 6 demonstrates some quasi-periodic SCVs composed by PPG signals from three different subjects and their corresponding distributions. We can observe that the PPG signals here can be roughly considered as sub-Gaussian signals. However, their distributions are not as symmetric as the ideal sub-Gaussian distribution.

We build total six types of semi-simulated data sets as summarized in Table 3. In detail, the capital letter (D, E, and F) indicates the subject where the PPG signal is measured, and the digit (1 for Gaussian and 2 for Laplacian) represents the type of non-target source signals.

3.4.2. Accuracy of comparison methods

We test all the comparison methods on the semi-simulated data sets with $N = 10, K = 10$ and $M = 400$. We independently run every method for 100 trials on each

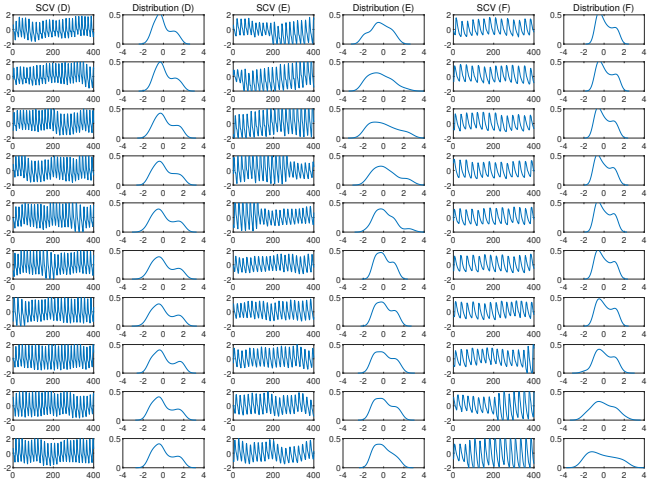

Figure 6: The quasi-periodic SCVs composed by PPG signals from three different subjects and their corresponding distributions.

Table 3: Six types of semi-simulated PPG data sets

Non-target \ Target	Subject D	Subject E	Subject F
Gaussian	D1	E1	F1
Non-Gaussian	D2	E2	F2

type of semi-simulated data. The experiment results are summarized in Table 4 and the corresponding histograms of r statistics are given in Fig. 7. We can see from Table 4 that the performance of CIVE-G is in general better than the other ones. In comparison, the IVA-L does not perform well since the PPG signals are sub-Gaussian, which is not consistent with the assumption of IVA-L. The IVA-G achieves good results for cases of D1, D2, E1 and E2, while its performance slightly drops for cases of F1 and F2. We guess that the small differences in probability distributions may have caused this result. We also observe that the CIVA is superior in comparison with IVA, which performs similarly as the CIVE for E2 and F2. Surprisingly, the results of IVE are distorted for the sub-Gaussian PPG signals even with good initials. The reason may be that the PPG signal is still too close to the Gaussian ones for IVE method.

To show more details, Fig. 8 lists the results from one of the 100 trials in E1. For ease of presentation, Fig. 8 (a) and (b) show only two estimated PPG signals within the same SCV by different methods. The correlation coefficient r for the results in Fig. 8 (a) is 1.00 for CIVE-G, and 0.99 for IVA-G and CIVA, respectively. Similarly, the r in Fig. 8 (b) is 1.00 for CIVE-G, and 0.99 for IVA-G and CIVA, respectively. The experimental results on semi-simulated data sets indicate that the proposed CIVE method overall outperforms the comparison methods, which verifies the ability of CIVE for extracting real quasi-periodic signals from multiple data sets.

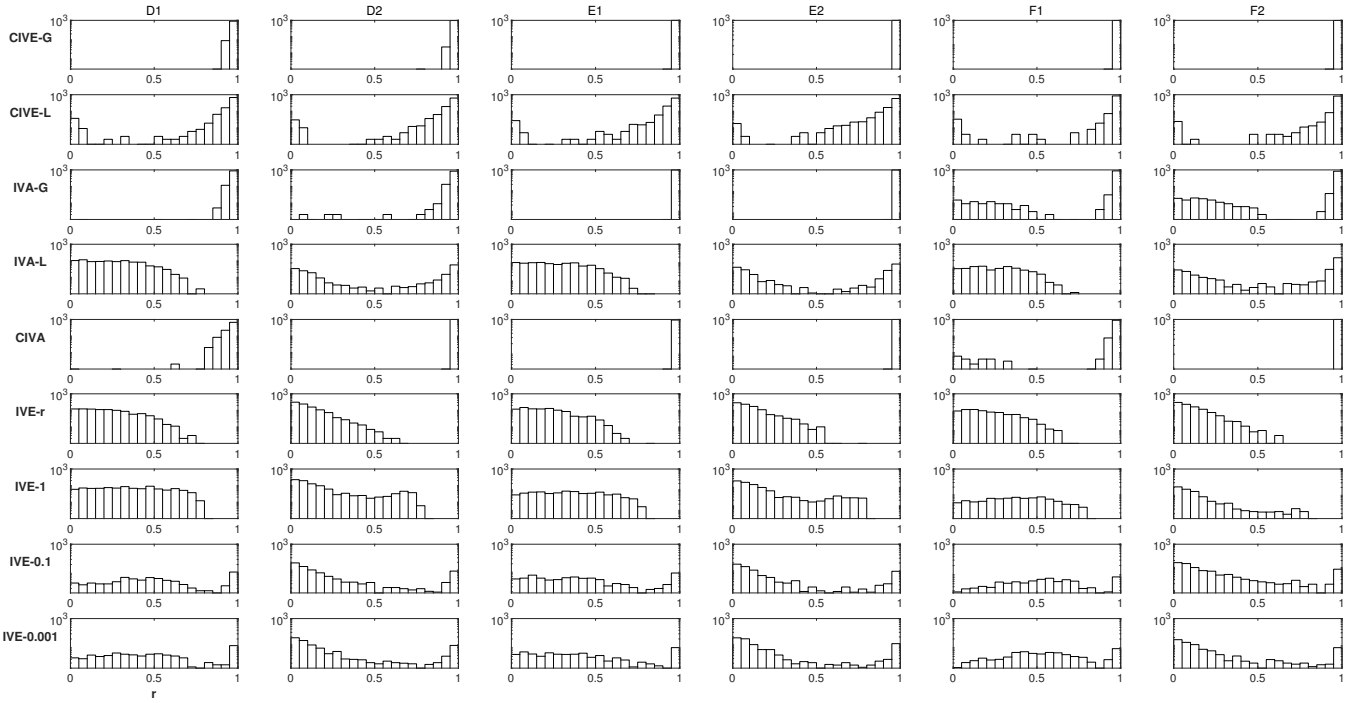

Figure 7: Histograms of r statistics for results in Table 4 using semi-simulated data.

Table 4: The average r and average $\mathrm{ISI}_{\mathrm{JNT}}-\mathrm{v}$ for the results from 100 independent trials on semi-simulated PPG data sets.

Metrics	$\mathrm{ISI}_{\mathrm{JNT}}-\mathrm{v}$	r	$\mathrm{ISI}_{\mathrm{JNT}}-\mathrm{v}$	r	$\mathrm{ISI}_{\mathrm{JNT}}-\mathrm{v}$	r
Data	D1		D2		E1	
CIVE-G	**-0.85**	**0.97**	-0.85	**0.98**	**-0.86**	**0.99**
CIVE-L	-0.82	0.91	-0.82	0.91	-0.82	0.91
IVA-G	-0.84	**0.97**	-0.84	0.96	**-0.86**	0.98
IVA-L	-0.56	0.26	-0.82	0.49	-0.56	0.27
CIVA	-0.82	0.95	**-0.86**	**0.98**	-0.84	0.98
IVE-r	-0.56	0.23	-0.74	0.12	-0.56	0.21
IVE-1	-0.57	0.36	-0.73	0.22	-0.57	0.36
IVE-0.1	-0.62	0.49	-0.74	0.34	-0.61	0.45
IVE-0.001	-0.61	0.50	-0.74	0.33	-0.61	0.45
Data	E2		F1		F2	
CIVE-G	**-0.86**	**0.99**	**-0.87**	**0.99**	**-0.87**	**0.99**
CIVE-L	-0.81	0.91	-0.85	0.93	-0.85	0.94
IVA-G	**-0.86**	0.98	-0.84	0.91	-0.83	0.89
IVA-L	-0.82	0.50	-0.56	0.27	-0.82	0.58
CIVA	**-0.86**	**0.99**	-0.85	0.96	**-0.87**	**0.99**
IVE-r	-0.73	0.13	-0.56	0.22	-0.74	0.12
IVE-1	-0.74	0.22	-0.59	0.39	-0.73	0.23
IVE-0.1	-0.75	0.34	-0.64	0.53	-0.74	0.34
IVE-0.001	-0.74	0.35	-0.63	0.52	-0.74	0.33

3.5. Comparison with classical single-set BSS methods

To demonstrate the benefit of CIVE for joint SCV extraction in comparison with classical single-set BSS methods, we also compare it with the SOBI [11], the PiCA [16] and the CICA [13], which are considered to be effective for extracting quasi-periodic signals for a single set.

3.5.1. Generation of data

Since there is only a single quasi-periodic signal in each data set in the previous experiments, it is difficult to fully compare the performance of CIVE with the above single-set separation methods. Accordingly, we generate data set here with all sources as quasi-periodic signals. Since the single-set separation method has no explicit requirement for the distribution, for simplicity, we only consider the case of A2 in Table 1. Namely, the target SCV is composed of dependent Gaussian quasi-periodic signals, while the non-target signals are independent Laplacian-distributed quasi-periodic signals. In order to make the data more realistic, we also add Gaussian white noise to all the observation signals.

In order to verify the robustness of each method against noise, the signal-to-noise ratio (SNR) of observation signals varies from 15 dB to 40 dB. Meanwhile, considering the competitive performance of CIVA in the above experiments, we also add it to the test as a reference of typical joint separation method.

3.5.2. Accuracy of comparison methods

Table 5: The average r and average $\mathrm{ISI}_{\mathrm{JNT}}-\mathrm{v}$ for the results from 100 independent trials on simulated data at different SNRs.

Metrics	$\mathrm{ISI}_{\mathrm{JNT}}-\mathrm{v}$	r	$\mathrm{ISI}_{\mathrm{JNT}}-\mathrm{v}$	r	$\mathrm{ISI}_{\mathrm{JNT}}-\mathrm{v}$	r
SNR (dB)	15		20		25	
CIVE-G	-0.71	**0.59**	**-0.76**	**0.75**	**-0.77**	**0.84**
CIVA	-0.65	0.30	-0.75	0.64	-0.76	0.73
SOBI	-0.73	0.53	-0.74	0.69	-0.73	0.79
PiCA	-0.60	0.20	-0.62	0.25	-0.64	0.37
CICA	**-0.73**	0.10	**-0.76**	0.08	-0.73	0.07
SNR (dB)	30		35		40	
CIVE-G	**-0.76**	**0.90**	**-0.72**	**0.94**	**-0.69**	**0.96**
CIVA	-0.75	0.81	-0.72	0.86	-0.68	0.91
SOBI	-0.72	0.87	-0.69	0.92	-0.66	0.95
PiCA	-0.64	0.44	-0.62	0.51	-0.61	0.56
CICA	-0.75	0.06	-0.57	0.05	-0.68	0.04

From Table 5, we observe that the CIVE-G achieves

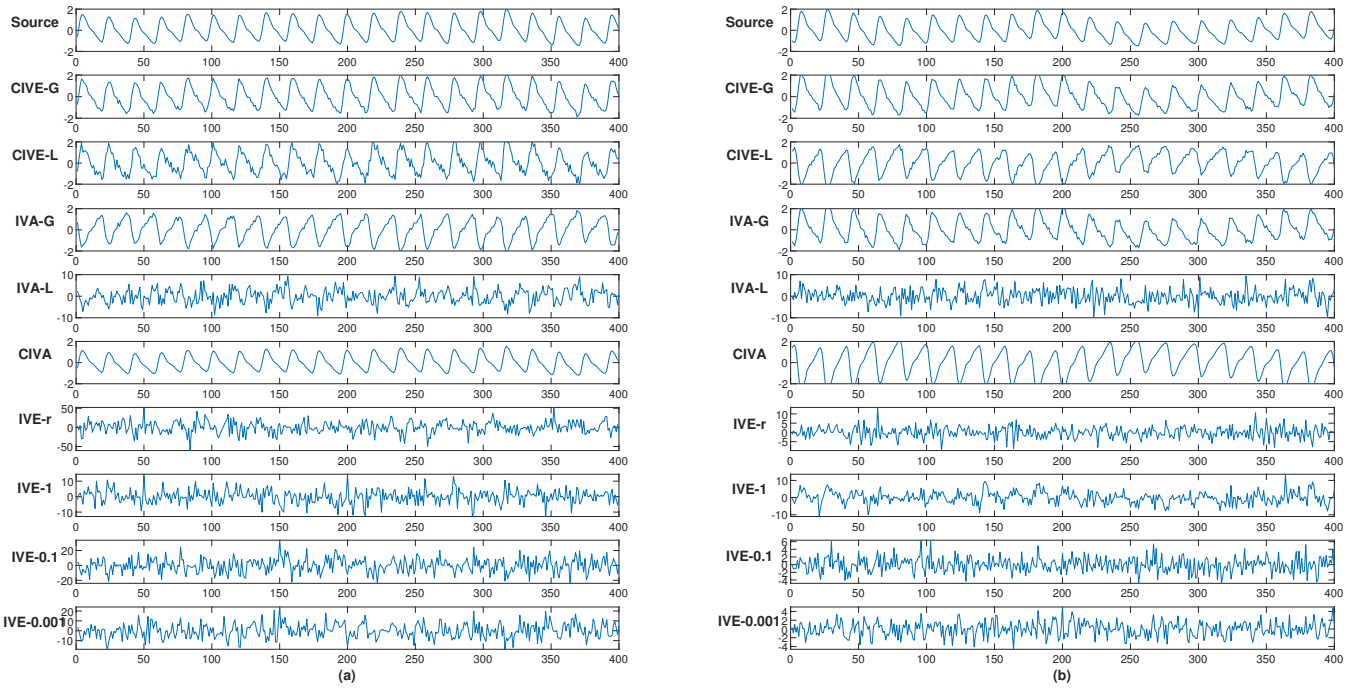

Figure 8: Extraction results for semi-simulated PPG data sets in a single trial. (a) and (b) represent the estimations of the first and second target PPG signals within the same SCV by different methods, respectively.

overall the best performance compared to other methods at different SNRs. In detail, the CIVA and the SOBI get similar performance, which outperform the other two comparison methods. Since the source signals in each set are all quasi-periodic, it causes a great interference to the extraction and selection of the target pseudo-periodic signal in single-set separation methods. As known, similar spectral densities appearing in the mixed source signals may restrict the performance of SOBI [37]. Meanwhile, the quality of the input peak indexes severely limits the performance of PiCA, especially when the SNR is low. The CICA performs the worst, of which the correlation coefficient r is 0.10 and the $\mathrm{ISI}_{\mathrm{JNT}}-\mathrm{v}$ is -0.73 when SNR is 15 dB. This indicates that the CICA can hardly distinguish the target signal from non-target ones which are also quasi-periodic. Finally, we need to declare that the results of the above single-set methods are the best ones calculated with all separated signals. In practice, the selection of the target signal is also a well-known issue for single-set separation methods if the reference is lack. In contrast, the proposed method can uniquely determine the target quasi-periodic SCV using the correlation of the target signals among each set.

In summary, the above results verify the effectiveness of the proposed CIVE methods to extract quasi-periodic SCV signals with different probability distributions and noise levels. In addition to these experimental results, we have also verified that the CIVE is efficient to extract SCVs of different periods, and it also performs well in real rPPG applications. Due to the limitation of space, we will not show all the results in this paper. Finally, although the CIVE methods outperform the other ones in terms of both accuracy and convergence speed, we should also note that the calculation of the autocorrelation constraint is still a bottleneck of the entire method, which will slow down the overall calculation time of the algorithm. It still needs to speed up the calculation of the autocorrelation constraint in our future research.

4. Conclusion

In this paper, we have introduced the constrained independent vector extraction method to determine quasi-periodic source signals that are independent from other mixed signals and dependent across multiple data sets. The cost function is composed of negentropy and mutual information together with an autocorrelation constraint to uniquely extract the target quasi-periodic SCV. The CIVE method has been tested to extract the target SCV from both simulated and semi-simulated data sets. Experimental results have verified that the proposed method outperforms the comparison methods in terms of both accuracy and convergence speed. Besides, the CIVE method does not require initial value and it works well for source signals under diverse distributions. Therefore, the proposed approach can be reliably used to extract quasi-periodic signals from multiple data sets.

Acknowledgments

This work was supported in part by the Anhui Key Project of Research and Development Plan (Grant

202104d07020005), the National Natural Science Foundation of China (Grants 61922075 and 41901350), the Fundamental Research Funds for the Central Universities (Grants PA2021KCPY0051 and JZ2021HGTB0078), and the Provincial Natural Science Foundation of Anhui (Grant 2008085QF285).

References

[1] W. A. Gardner, A. Napolitano, L. Paura, Cyclostationarity: Half a century of research, Signal Processing 86 (4) (2006) 639–697.

[2] A. Napolitano, Cyclostationarity: New trends and applications, Signal Processing 120 (2016) 385–408.

[3] H. Qi, Z. Guo, X. Chen, Z. Shen, Z. J. Wang, Video-based human heart rate measurement using joint blind source separation, Biomedical Signal Processing and Control 31 (2017) 309–320.

[4] K. Lee, J. Lee, C. Ha, M. Han, H. Ko, Video-based contactless heart-rate detection and counting via joint blind source separation with adaptive noise canceller, Applied Sciences 9 (20) (2019) 4349.

[5] A. Hyvärinen, E. Oja, Independent component analysis: algorithms and applications, Neural Networks 13 (4-5) (2000) 411–430.

[6] P. Comon, Independent component analysis, a new concept?, Signal Processing 36 (3) (1994) 287–314.

[7] M. Novey, T. Adali, Complex ICA by negentropy maximization, IEEE Transactions on Neural Networks 19 (4) (2008) 596–609.

[8] J.-F. Cardoso, Infomax and maximum likelihood for blind source separation, IEEE Signal Processing Letters 4 (4) (1997) 112–114.

[9] W. Lu, J. C. Rajapakse, Constrained independent component analysis, in: Advances in Neural Information Processing Systems, 2001, pp. 570–576.

[10] L. Tong, V. Soon, Y. Huang, R. Liu, AMUSE: a new blind identification algorithm, in: IEEE International Symposium on Circuits and Systems, IEEE, 1990, pp. 1784–1787.

[11] A. Belouchrani, K. Abed-Meraim, J. Cardoso, E. Moulines, Second-order blind separation of temporally correlated sources, in: Proc. Int. Conf. Digital Signal Processing, Citeseer, 1993, pp. 346–351.

[12] Z.-L. Zhang, Morphologically constrained ICA for extracting weak temporally correlated signals, Neurocomputing 71 (7-9) (2008) 1669–1679.

[13] R. Macwan, Y. Benezeth, A. Mansouri, Remote photoplethysmography with constrained ICA using periodicity and chrominance constraints, Biomedical Engineering Online 17 (1) (2018) 22.

[14] T. Tsalaile, R. Sameni, S. Sanei, C. Jutten, J. Chambers, Sequential blind source extraction for quasi-periodic signals with time-varying period, IEEE Transactions on Biomedical Engineering 56 (3) (2008) 646–655.

[15] X.-L. Li, X.-D. Zhang, Sequential blind extraction adopting second-order statistics, IEEE Signal Processing Letters 14 (1) (2006) 58–61.

[16] R. Sameni, C. Jutten, M. B. Shamsollahi, Multichannel electrocardiogram decomposition using periodic component analysis, IEEE Transactions on Biomedical Engineering 55 (8) (2008) 1935–1940.

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

[18] A. Cichocki, D. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, H. A. Phan, Tensor decompositions for signal processing applications: From two-way to multiway component analysis, IEEE Signal Processing Magazine 32 (2) (2015) 145–163.

[19] H. Akbari, M. B. Shamsollahi, R. Phlypo, Fetal ECG extraction using πTucker decomposition, in: 2015 International Conference on Systems, Signals and Image Processing (IWSSIP), IEEE, 2015, pp. 174–178.

[20] M. Anderson, X.-L. Li, T. Adalı, Nonorthogonal independent vector analysis using multivariate Gaussian model, in: International Conference on Latent Variable Analysis and Signal Separation, Springer, 2010, pp. 354–361.

[21] T. Kim, T. Eltoft, T.-W. Lee, Independent vector analysis: An extension of ICA to multivariate components, in: International Conference on Independent Component Analysis and Signal Separation, Springer, 2006, pp. 165–172.

[22] M. Anderson, T. Adali, X.-L. Li, Joint blind source separation with multivariate Gaussian model: Algorithms and performance analysis, IEEE Transactions on Signal Processing 60 (4) (2011) 1672–1683.

[23] M. Anderson, X.-L. Li, T. Adalı, Complex-valued independent vector analysis: Application to multivariate Gaussian model, Signal Processing 92 (8) (2012) 1821–1831.

[24] S. Bhinge, Q. Long, Y. Levin-Schwartz, Z. Boukouvalas, V. D. Calhoun, T. Adalı, Non-orthogonal constrained independent vector analysis: Application to data fusion, in: 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2017, pp. 2666–2670.

[25] A. H. Khan, M. Taseska, E. A. Habets, A geometrically constrained independent vector analysis algorithm for online source extraction, in: International Conference on Latent Variable Analysis and Signal Separation, Springer, 2015, pp. 396–403.

[26] S. Bhinge, R. Mowakeaa, V. D. Calhoun, T. Adalı, Extraction of time-varying spatiotemporal networks using parameter-tuned constrained IVA, IEEE Transactions on Medical Imaging 38 (7) (2019) 1715–1725.

[27] F. Nesta, Z. Koldovskỳ, Supervised independent vector analysis through pilot dependent components, in: 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2017, pp. 536–540.

[28] Z. Koldovskỳ, P. Tichavskỳ, Gradient algorithms for complex non-Gaussian independent component/vector extraction, question of convergence, IEEE Transactions on Signal Processing 67 (4) (2018) 1050–1064.

[29] Z. Koldovskỳ, P. Tichavskỳ, N. Ono, Orthogonally-constrained extraction of independent non-Gaussian component from non-Gaussian background without ICA, in: International Conference on Latent Variable Analysis and Signal Separation, Springer, 2018, pp. 161–170.

[30] A. Hyvarinen, Fast and robust fixed-point algorithms for independent component analysis, IEEE Transactions on Neural Networks 10 (3) (1999) 626–634.

[31] D. P. Bertsekas, Constrained optimization and Lagrange multiplier methods, Academic Press, 2014.

[32] T. Kim, H. T. Attias, S.-Y. Lee, T.-W. Lee, Blind source separation exploiting higher-order frequency dependencies, IEEE Transactions on Audio, Speech, and Language Processing 15 (1) (2006) 70–79.

[33] T. Eltoft, T. Kim, T.-W. Lee, On the multivariate Laplace distribution, IEEE Signal Processing Letters 13 (5) (2006) 300–303.

[34] E. Moreau, O. Macchi, A one stage self-adaptive algorithm for source separation, in: Proceedings of ICASSP'94. IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 3, IEEE, 1994, pp. III–49.

[35] S. Choi, A. Cichocki, L. Zhang, S.-I. Amari, Approximate maximum likelihood source separation using the natural gradient, IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences 86 (1) (2003) 198–205.

[36] S. Bobbia, R. Macwan, Y. Benezeth, A. Mansouri, J. Dubois, Unsupervised skin tissue segmentation for remote photoplethysmography, Pattern Recognition Letters 124 (2019) 82–90.

[37] L. Albera, A. Ferréol, P. Chevalier, P. Comon, ICAR: a tool for blind source separation using fourth-order statistics only, IEEE Transactions on Signal Processing 53 (10) (2005) 3633–3643.