Proceedings Paper:
Zhang, Z orcid.org/0000-0003-0204-3867, Li, H and Mandic, D (2016) Blind source separation and artefact cancellation for single channel bioelectrical signal. In: 13th International Conference on Wearable and Implantable Body Sensor Networks (BSN 2016). BSN 2016, 14-17 Jun 2016, San Francisco, USA. IEEE , pp. 177-182. ISBN 9781509030873

https://doi.org/10.1109/BSN.2016.7516255

Blind Source Separation and Artefact Cancellation for Single Channel Bioelectrical Signal

Zhiqiang Zhang[1], Huihui Li[2] and Danilo Mandic[3]

***Abstract*— Bioelectrical signal analysis is gaining significant interests from both academics and industries due to its capability for improved diagnosis and therapy of chronic diseases. In practice, different bio-signals, such as EEG, ECG, EOG and EMG, are usually contaminating each other, and the measured signal is the linear combination of them. It is critical to separate them since analysis of one type or several of them separately is of more interest. In the case of multichannel recording, several blind source separation methods are available to extract its original components. However, for single channel scenarios, the problem has yet to be well studied. Therefore in this paper, we explore blind source separation and artefact cancellation for a single channel signal by combining signal decomposition method singular spectrum analysis (SSA) with different blind source separation methods, such as principal component analysis (PCA), maximum noise fraction (MNF), independent component analysis (ICA) and canonical correlation analysis (CCA). We also systematically compare the separation performance by combing different decomposition methods (wavelet transform (WT), ensemble empirical mode decomposition (EEMD) and SSA) with blind source separation methods (PCA, MNF ICA and CCA). The good simulation results have demonstrated the effectiveness and efficiency of the proposed method.**

I. INTRODUCTION

Last decade has witnessed the measurement of human physiological signals transiting from a hospital-centric devices toward ambulatory wearable systems, which are often operated by patients themselves at home environment [1] [2] [3]. Although such systems are capable of continuously monitoring for better diagnosis and therapy of chronic diseases, it also increases the likelihood of getting poor quality signal by taking the measurements system out of clinical environments. In particular, different bioelectrical signals, such as EEG, ECG, EOG and EMG, can contaminate each other, and the measured signal is the linear combination of them [4] [5]. Therefore, in order to analyse signal of interest and extract useful information from it, it is crucial to separate the underlying sources.

For this purpose, blind source separation (BSS) methods have been widely explored. A number of methods, such as principal component analysis (PCA), maximum noise fraction (MNF), independent component analysis (ICA) and canonical correlation analysis (CCA), have been proposed so far [6]. In the case of multichannel recording, all these methods can efficiently unmix the given signal into their constituent sources, but all of them implicitly assumes the number of underlying sources is equal or less than the number of signal channels. However, in ambulatory settings, the wearable systems normally only provides one channel signal, thus the aforementioned BSS methods will not be applicable.

Thus far, there are also several methods aiming for single channel source separation. For example, Davies et al. [7] presented an adaptation of ICA to single-channel signals, called single-channel ICA (SCICA), and it required sources were stationary, which might not always hold in practical applications. Alternative approach is to construct multichannel data from a single channel signal, and then apply multiple channel source separation methods to extract the underlying components. For instance, Lin et al. [8] and Abbaspour et al. [9] demonstrated combining Wavelet transform (WT) and ICA for source separation and artefact removal. Mijović et al. [10] explored source separation from single-channel recordings by integrating empirical mode decomposition (EMD) and ICA. Han et al. [11], Sweeney et al. [12] and Chen et al. [13] all investigated the combination of EMD and CCA. Despite the wide application of WT and EMD, there are also some constraints. The extracted components vary significantly with different mother wavelet, and how to determine the proper mother wavelet is challenging for WT. The advantage of EMD, compared to WT, is that the EMD is a data driven algorithm, which does not use any predefined function in the decomposition stage, but it is extremely computationally expensive, which is particularly true for ensemble empirical mode decomposition (EEMD).

Singular spectrum analysis (SSA) is another powerful method to decompose real-valued time series, which overcomes the drawbacks of WT and EMD [14]. Therefore, in this paper, we propose to combine SSA with different source separation methods, such as PCA, MNF, ICA and CCA for blind source separation and artefact cancellation for a single channel signal. We also systematically compare the separation performance by combing different decomposition methods (WT, EEMD and SSA) and blind source separation methods (PCA, MNF ICA and CCA). The good simulation results have demonstrated the effectiveness and efficiency of the proposed method.

The rest of the paper is organized as follows. In section II, all the methods and algorithms used in the paper are briefly explained. In section III, we systematically compare the

[1]Zhiqiang Zhang is with school of Electronic and Electrical Engineering and school of Mechanical Engineering, University of Leeds, UK, LS2 9JT. Email: z.zhang3@leeds.ac.uk

[2]Huihui Li is with Shenzhen Institute of Advanced Technology, Chinese Academt of Sciences, ShenZhen, China. Email: hh.li@siat.ac.cn

[3]Danilo Mandic is with department of Electrical and Electronic Engineering, Imperial College London, UK, SW7 2AZ. Email: d.mandic@impreial.ac.uk

performance for different combinations using simulations. Section IV concludes the paper.

II. Methods

In this section, we briefly explain the various techniques employed in this paper by outlining different signal decomposition methods and blind source separation methods.

A. Signal Decomposition Methods

For any real-valued time series, $\mathbf{x} = (x_1, x_2, \cdots, x_N)$ of length N can be decomposed into M different components of length N, and the summation of all these components equals to $\mathbf{x}$. In this section, we will elaborate different methods achieving this goal.

1) Wavelet transform: Multi-resolution wavelet decomposition splits a signal into high-scale (low-frequency components) approximations and low-scale (high-frequency components) details. The decomposition process can be iterated, with successive approximations being decomposed in turn. The decomposition of $\mathbf{x}$ can be written as

$$\begin{aligned}\mathbf{x} &= \mathbf{x}_{A_0} \\ &= \mathbf{x}_{A_1} + \mathbf{x}_{D_1} \\ &= \mathbf{x}_{A_2} + \mathbf{x}_{D_2} + \mathbf{x}_{D_1} \\ &= \mathbf{x}_{A_{M-1}} + \mathbf{x}_{D_{M-1}} + \cdots + \mathbf{x}_{D_1}\end{aligned} \tag{1}$$

where

$$\begin{aligned}\mathbf{x}_{A_j} &= \mathbf{x}_{A_{j+1}} + \mathbf{x}_{D_{j+1}} \\ \mathbf{x}_{A_{j+1}} &= \sum_{k=-\infty}^{\infty} h(2n-k)\mathbf{x}_{A_j} \\ \mathbf{x}_{D_{j+1}} &= \sum_{k=-\infty}^{\infty} g(2n-k)\mathbf{x}_{A_j}.\end{aligned} \tag{2}$$

It is obvious the decomposition is only associated with the filter h and g, which are determined by mother wavelet.

2) Ensemble empirical mode decomposition: EMD is a method to separate an arbitrary signal $\mathbf{x}$ into a series of intrinsic mode functions (IMFs) through a sifting process. The IMFs are functions that satisfy two properties: 1) over the length of the dataset, the number of extrema and the number of zero-crossings must either be equal or differ at most by one. 2) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero. The procedure of the sifting process include:

1. For $\mathbf{x}$, find all the local maximum and minima.
2. Let μ be the mean of its upper and lower envelopes determined from a cubic-spline interpolation of local maxima and minima.
3. Define $\mathbf{h} = \mathbf{x} - \mu$ as the first proto-IMF
4. Repeat the step 1 to 3 until the resulting proto-IMF, which is called IMF1 $\mathbf{x_1}$, satisfying IMF properties.
5. Subtract the IMF1 from the original signal to get the residual $\mathbf{r_1} = \mathbf{x} - \mathbf{x_1}$.
6. Repeat step one to five to obtain other IMFs $\mathbf{x_2}, \cdots, \mathbf{x_{M-1}}$ until the residual $\mathbf{r_{M-1}}$ is a monotonic function.

Thus, $\mathbf{x}$ can be decomposed as:

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_{M-1} + \mathbf{r_{M-1}}. \tag{3}$$

where $M-1$ is the number of IMFs.

The EMD algorithm is very sensitive to noise; therefore, an EEMD method was proposed to overcome this issue by adding independent, identically distributed white noise into $\mathbf{x}$ with several trials. EEMD repeats the EMD decomposition for each trial, but with different noise, which obtains the means of corresponding IMFs as the optimum choice of IMFs.

3) Singular spectrum analysis: SSA is a model-free technique for signal decomposition. Given a integer $L(1 < L < N)$ as the window size, SSA decomposition for $\mathbf{x}$ can be described as:

- Embedding: The first step is to create a $L \times (N-L+1)$ matrix $\mathbf{X}$ using a delayed version of $\mathbf{x}$:

$$\mathbf{X} = \begin{pmatrix} x_1 & x_2 & \cdots & x_{N-L+1} \\ x_2 & x_3 & \cdots & x_{N-L+2} \\ \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+1} & \cdots & x_N \end{pmatrix} \tag{4}$$

  It is evident that $\mathbf{X}$ is a Hankel matrix since the elements of all the anti-diagonals $i+j = constant$ are the same.
- SVD: The second step is to perform the SVD of the trajectory matrix $\mathbf{X}$. Let $\mathbf{S} = \mathbf{X}\mathbf{X^T}$ and assume $\lambda_1, \lambda_2, \cdots, \lambda_N$ are eigenvalues of $\mathbf{S}$ in the decreasing order of magnitude $(\lambda_1 \geq \lambda_2 \geq, \cdots, \geq \lambda_N \geq 0)$ and the corresponding eigenvectors are $\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_L}$. We can thus write the trajectory matrix as

$$\mathbf{X} = \mathbf{X_1} + \mathbf{X_2} + \cdots + \mathbf{X_d} \tag{5}$$

  where $d = \underset{i}{\operatorname{argmax}}\{\lambda_i > 0\}$, $X_i = \sqrt{\lambda_i}\mathbf{u_i}\mathbf{v_i^T}$ and $\mathbf{v_i} = \mathbf{X^T}\mathbf{u_i}/\sqrt{\lambda_i}$. Here the collection $(\lambda_i,\ \mathbf{u_i},\ \mathbf{v_i})$ is called the ith eigentriple of SVD.
- Eigentriple grouping and diagonal averaging: The next step is to regroup the elementary matrices into several submatrices

$$X = \sum_{m=1}^{M} \hat{\mathbf{X}}_m \tag{6}$$

  where M is the number of groups, index m refers to mth subgroup of Eigentriple, and $\hat{\mathbf{X}}_m$ indicates the sum of $\mathbf{X}_i$ within group m. Each matrix $\hat{\mathbf{X}}_m$ is hankelized, which can thus transformed into a new series $\mathbf{x}_m$ of length N using the one-to-one correspondence between Hankel matrices and time series. Thus $\mathbf{x}$ can thus be decomposed as:

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_M. \tag{7}$$

For simplicity and integrity, we rewrite the decomposed elements of signal $\mathbf{x}$ using WT, EEMD and SSA all into $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_M$, thus we can construct a M channel signal $\mathbf{Y}$ with length L as

$$\mathbf{Y} = [\mathbf{x}_1^T,\ \mathbf{x}_2^T, \cdots, \mathbf{x}_M^T] \tag{8}$$

Here, $\mathbf{Y}$ is a $L \times M$ matrix, each column containing one channel signal.

B. Blind source separation methods

For any real-valued multiple channel signal $\mathbf{Y}$, we can find an unmixing matrix $\mathbf{\Psi}$ to determine the unknown original signal $\mathbf{S}$ as $\mathbf{S} = \mathbf{Y\Psi}$. In this section, we will elaborate different techniques to determine $\mathbf{\Psi}$ and $\mathbf{S}$.

1) Principal component analysis : In the context of artefact removal and blind source separation, it is widely to view PCA as an optimization problem, which tries to determine a linear transformation that maximizes the variance of the transformed variables subject to orthogonality constraints. The transformation can write as the following optimization problem:

$$\underset{\mathbf{\Psi}}{\operatorname{argmax}} \|\mathbf{Y\Psi}\| \tag{9}$$

subject to

$$\mathbf{\Psi^T\Psi} = \mathbf{I} \tag{10}$$

where $\mathbf{I}$ is the identity matrix. The solution to this problem are the eigenvectors of $\mathbf{Y^TY}$, which is equivalent to calculate the SVD for $\mathbf{Y}$ as:

$$\mathbf{Y} = \mathbf{\Sigma_1 \Lambda_1 \Psi_1^T} \tag{11}$$

where the columns of $\mathbf{\Sigma_1}$ contain the eigenvectors of $\mathbf{YY^T}$, the columns of $\mathbf{\Psi_1}$ contain the eigenvectors of $\mathbf{Y^TY}$, and the diagonal of $\mathbf{\Lambda_1}$ indicates the singular values of $\mathbf{Y}$. Thus, $\mathbf{\Psi} = \mathbf{\Psi_1}$, and the transformation $\mathbf{S}_{pca} = \mathbf{Y\Psi}$ can generate a new set of orthogonal variables ordered by decreasing variance as the unknown original signal.

2) Maximum noise fraction: It assume that the signal $\mathbf{Y}$ consists of real signal matrix $\hat{\mathbf{Y}}$ and noise $\mathbf{N}$

$$\mathbf{Y} = \hat{\mathbf{Y}} + \mathbf{N} \tag{12}$$

where $\hat{\mathbf{Y}}$ and $\mathbf{N}$ are orthogonal to each other. Thus we can have the following optimization

$$\underset{\mathbf{\Psi}}{\operatorname{argmax}} \frac{\|\hat{\mathbf{Y}}\mathbf{\Psi}\|}{\|\mathbf{N\Psi}\|} + 1 = \underset{\mathbf{\Psi}}{\operatorname{argmax}} \frac{\|\mathbf{Y\Psi}\|}{\|\mathbf{N\Psi}\|} \tag{13}$$

Similar to PCA, the solution to the above optimization is equivalent to generalized SVD:

$$\mathbf{Y^TY\Psi} = \lambda \mathbf{N^TN\Psi}. \tag{14}$$

In practice, it may not easy to acquire the covariance matrix of the noise, which thus can be estimated by computing the covariance of the difference. Define the ith row of matrix $d\mathbf{Y}$ as the difference of the ith and i+1 rows of the matrix $\mathbf{Y}$, thus we can make the approximation:

$$d\mathbf{Y^T}d\mathbf{Y} = \mathbf{N^TN}. \tag{15}$$

Therefore, the unmixing matrix can be calculated as:

1. Calculate the SVD for the covariance of $d\mathbf{Y}$:

$$d\mathbf{Y^T}d\mathbf{Y} = \mathbf{\Sigma_2 \Lambda_2 \Sigma_2^T} \tag{16}$$

2. Whiten signal $\mathbf{Y}$ as

$$\tilde{\mathbf{Y}} = \mathbf{Y\Sigma_2\Lambda_2^{-1/2}} \tag{17}$$

3. Calculate the SVD for the covariance of $\tilde{\mathbf{Y}}$

$$\tilde{\mathbf{Y}}^\mathbf{T}\tilde{\mathbf{Y}} = \mathbf{\Sigma_3 \Lambda_3 \Sigma_3^T} \tag{18}$$

3. Compute the maximum noise fraction basis vectors as the unknown original signal

$$\mathbf{S}_{mnf} = \mathbf{Y\Psi_2} \tag{19}$$

where $\mathbf{\Psi} = \mathbf{\Psi_2} = \mathbf{\Sigma_2\Lambda_2^{-1/2}\Sigma_3^T}$.

3) Independent component analysis: Similar to PCA and MNF, ICA also models BSS as a optimization by maximizing the statistical independence of the estimated components. In general, there are two ways defining independence: 1)minimization of mutual information or 2) maximization of non-Gaussianity. The minimization of mutual information family of ICA algorithms uses measures like Kullback-Leibler Divergence and maximum entropy, while non-Gaussianity family of ICA algorithms, motivated by the central limit theorem, uses kurtosis and negentropy. Regardless of the independence definition, the aim of ICA is to determine an unmixing matrix $\mathbf{\Psi_3}$ to derive the unknown independent component as the original source:

$$\mathbf{S}_{ica} = \mathbf{Y\Psi_3}. \tag{20}$$

Thus far, there are a number of different algorithms available, e.g. Bell-Sejnowski algorithm [15], extended ICA [16], and JADE [14]. In our implementation, the fast ICA algorithm [17] was applied to extract $\mathbf{\Psi_3}$ from $\mathbf{Y}$ due to its high efficiency.

4) Canonical correlation analysis: CCA solves the BSS problem by forcing the sources to be maximally autocorrelated and mutually uncorrected. Given any two matrix signals $\mathbf{Y_1}$ with size $N \times K_1$ and $\mathbf{Y_2}$ with size $N \times K_2$ (K_1 and K_2 indicate the number of channels), the canonical correlation can be given as the following optimization problem:

$$\underset{\mathbf{\Phi_1},\mathbf{\Phi_2}}{\operatorname{argmax}} \frac{\mathbf{\Phi_1^T Y_1^T Y_2 \Phi_2}}{\sqrt{\mathbf{\Phi_1^T Y_1^T Y_1 \Phi_1}}\sqrt{\mathbf{\Phi_2^T Y_2^T Y_2 \Phi_2}}} \tag{21}$$

subject to

$$\begin{cases} \mathbf{\Phi_1} \neq 0 \\ \mathbf{\Phi_2} \neq 0. \end{cases} \tag{22}$$

Conduct QR decomposition for $\mathbf{Y_1}$ and $\mathbf{Y_2}$ as $\mathbf{Y_1} = \mathbf{Q_1R_1}$ and $\mathbf{Y_2} = \mathbf{Q_2R_2}$, and SVD for $\mathbf{Q_1^TQ_2}$ as $\mathbf{Q_1^TQ_2} = \mathbf{\Sigma_4\Lambda_4\Upsilon^T}$, the canonical correlations can thus be given by the diagonal elements of $\mathbf{\Lambda_4}$. The transformation $\mathbf{\Phi_1}$ and $\mathbf{\Phi_2}$ can also be taken as $\mathbf{\Phi_1} = \mathbf{R_1^{-1}\Sigma_4}$ and $\mathbf{\Phi_2} = \mathbf{R_2^{-1}\Upsilon}$. In blind source separation applications, we can set $\mathbf{Y_1} = \mathbf{Y}$ and $\mathbf{Y_2}$ as a temporally delayed version of $\mathbf{Y}$. Therefore, underlying source signal can be derived as $\mathbf{S}_{cca} = \mathbf{Y\Psi}$, where $\mathbf{\Psi} = \mathbf{\Phi_1}$.

III. Simulation Results

In order to evaluate the performance of different ways of combing signal decomposition methods with blind source separation methods, we carried out the following essential experiments. Since it is quite challenging to find the ground

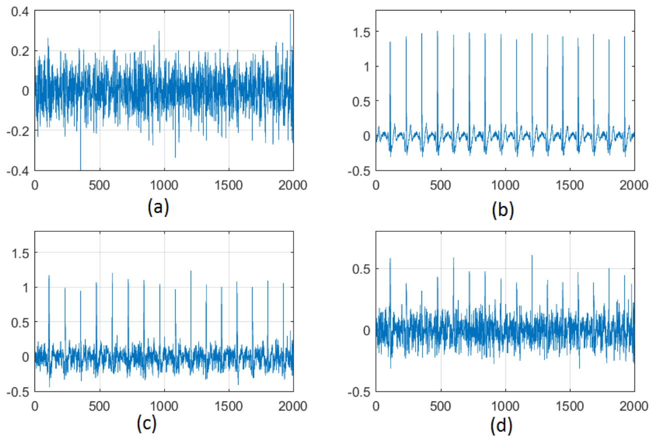

Fig. 1. The signal used in the simulation. (a) the EEG signal $\mathbf{x}_a$; (b) the ECG signal $\mathbf{x}_b$ as the artefacts; (c) the mixture $\mathbf{x}$ for SNR=0.5; and (d) $\mathbf{x}$ for SNR=1

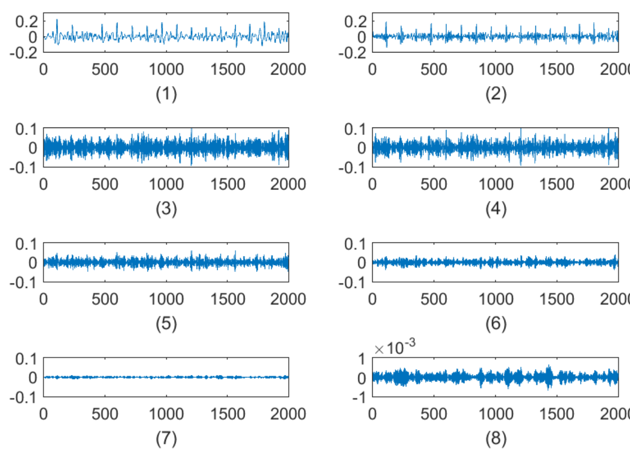

Fig. 2. The 8 SSA components of the synthetic signal $\mathbf{x}$ at SNR=1.

truth for each component in a contaminated signal in practice, we resort to simulation study with known parameters. In our simulation, we always mixed two signals: the signal of interest $\mathbf{x}_a$ and artefacts $\mathbf{x}_b$ in the following way:

$$\mathbf{x} = \mathbf{x}_a + \beta \mathbf{x}_b \tag{23}$$

where β is a proportion factor representing the contribution of artefacts. The signal to noise ratio, as a important indicator, can be adjusted by changing the value of β:

$$\mathrm{SNR} = \frac{\mathrm{RMS}(\mathbf{x}_a)}{\mathrm{RMS}(\beta \mathbf{x}_b)} \tag{24}$$

where the root mean square (RMS) is defined as:

$$\mathrm{RMS}(\mathbf{x}) = \sqrt{\frac{\mathbf{x}^{\mathbf{T}}\mathbf{x}}{N}}. \tag{25}$$

The relative root mean squire error (RRMSE) is applied to describe the separation performance:

$$\mathrm{RRMSE} = \frac{\mathrm{RMS}(\mathbf{x}_a - \hat{\mathbf{x}}_a)}{\mathrm{RMS}(\mathbf{x}_a)} \tag{26}$$

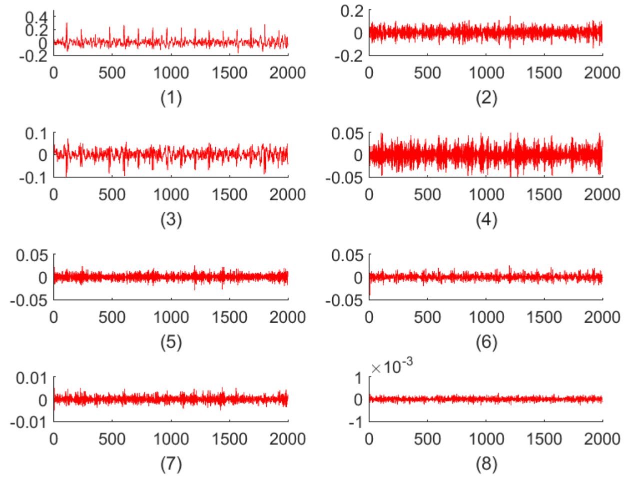

Fig. 3. The principal components of the SSA elements given in Fig. 2

where $\hat{\mathbf{x}}_a$ is the reconstruction of the signal of interest after removing the artefacts. The sources related to artefacts can be removed by setting the corresponding column of the source signal matrix $\mathbf{S}$ to zero, and the artefact-free multichannel signals can thus be reconstructed using the updated source matrix and the mixing matrix $\mathbf{\Psi}$. Therefore, the estimated signal of interest $\hat{\mathbf{x}}_a$ can be determined by simply summing artefact-free multichannel signals. Meanwhile, in our implementation, we also used correlation coefficient as the second criteria to evaluate the separation performance.

In practice, the EEG signal is always contaminated by other bioelctrical signals, such as ECG, EOG and EMG. There in our simulation, as an exemplar, we mixed EEG signal with ECG artefact, as shown in the Fig. 1. The frequency of the background brain electric activity ranges from 0 to 40Hz, while the ECG is a narrow band signal between 5 and 30 Hz, so the spectra of $\mathbf{x}_a$ and $\mathbf{x}_b$ are overlapping. In our implementation, we applied Daubechies 4 as the mother wavelet to decompose the $\mathbf{x}$ into 8 scales. For EEMD, the number of ensembles was set to 2, and 8 to 12 IMFs were usually produced by EMD, while for SSA, the window length L was set to 8, and each eig-triple was taken as an eig-triple group.

Fig. 2 and 3 illustrate the source separation process by combining SSA and PCA as the example: 1) the synthetic signal $\mathbf{x}$ was decomposed into 8 different elements by SSA, 2) PCA was then applied to extracted the principle components, and 3) reconstructed the signal of interest $\hat{\mathbf{x}}_a$ by setting the ECG related principle components 1 and 3 to 0. Fig. 4 shows the artefact removal and signal reconstruction performance by combining different signal decomposition methods and blind source separation method for SNR=1. To better illustrate the reconstruction results, only the first 200 samples were presented in the figure. It is obvious that all the algorithms can remove the ECG artefacts successfully. The reconstructed EEG signals can all follow the general trend of the original signal, similar reconstruction performance has been achieved for all the combinations of signal

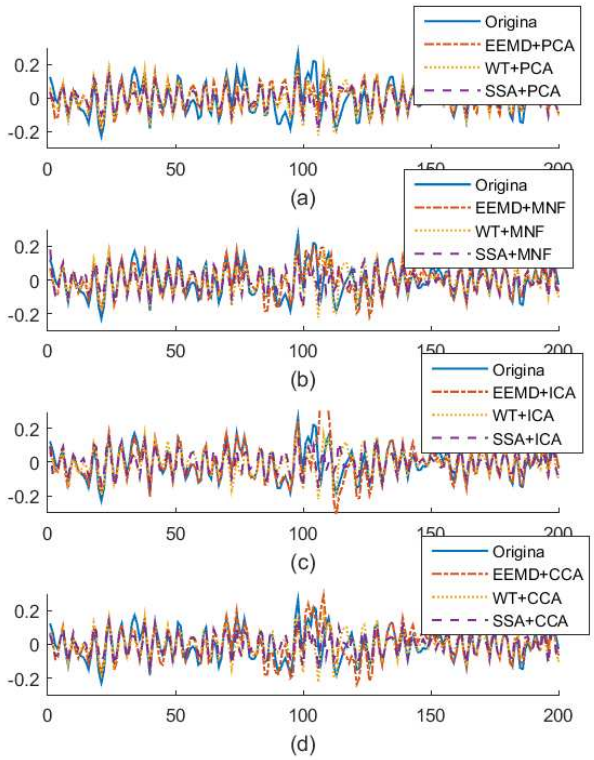

Fig. 4. The first 200 samples of the reconstruction signal by removing ECG artefacts at SNR=1. (a) PCA as the source separation method; (b) MNF as the source separation method; (c) ICA as the source septation method; and (d) CCA as the source separation method.

decomposition methods and blind source separation method. It is also worthy nothing to mention that in our simulation, the EEMD decomposition of the synthetic signal $\mathbf{x}$ into IMFs takes 3-5 minutes, while SSA and WT decomposition only take less than 0.1 second. We also noticed that the performance of the WT based artefact removal methods varies a lot with regards to different selections of mother wavelet, while there is no need to selection such function for SSA based methods. In summary, although all the algorithms have achieved comparable accuracy, the SSA based methods have shown some implementation advantages over EEMD-based and WT-based methods.

In addition to the qualitative analysis of the artefacts removal performance, the quantitatively results in terms of RRMSE and CC are shown in Fig. 5 and 6 to better illustrate the reconstruction performance by combining different signal decomposition methods and blind source separation methods. In general, when SNR increases, the artefact gets weaker and it is easier to recover the signal of interest from the contaminated ones with higher accuracy. As we can see from the figures, all the algorithms follow this rule with RRMSE raising and CC declining, which means that their performances get significant improvement when the SNR increases from 0.25 to 3. However, there are also some differences among these algorithms. The combinations of SSA with different blind source separation methods are

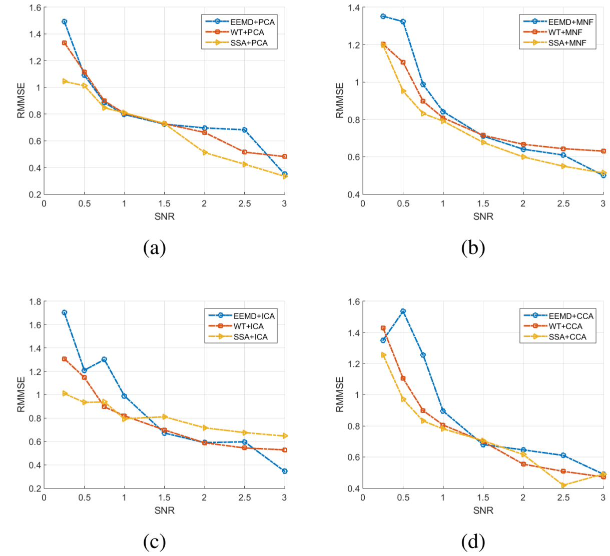

Fig. 5. Comparison of algorithms performance in RMMSE by combining different signal decomposition methods and blind source separation methods. (a) PCA as the source separation method; (b) MNF as the source separation method; (c) ICA as the source septation method; and (d) CCA as the source separation method.

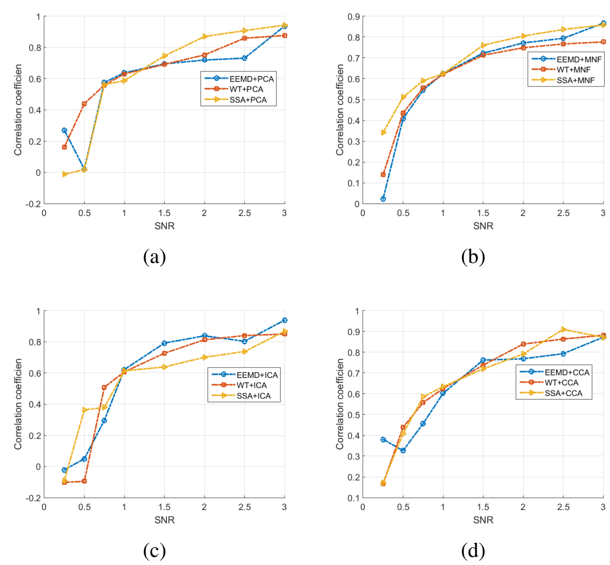

Fig. 6. Comparison of algorithms performance in correlation coefficient by combining different signal decomposition methods and blind source separation methods. (a) PCA as the source separation method; (b) MNF as the source separation method; (c) ICA as the source septation method; and (d) CCA as the source separation method.

slightly better than WT and EEMD based combinations, and smaller RRMSE and higher CC were achieved by SSA combinations regardless of SNR value. The only exception is the combination of SSA and ICA. Although combinations of EEMD and ICA is slightly better, the computational cost for EEMD is significantly more expensive than the SSA based methods, which prevent its routine usage on the ambulatory wearable systems in home environment.

In the SSA based algorithms, the only variable to be determined in advance in the window size L for SSA de-

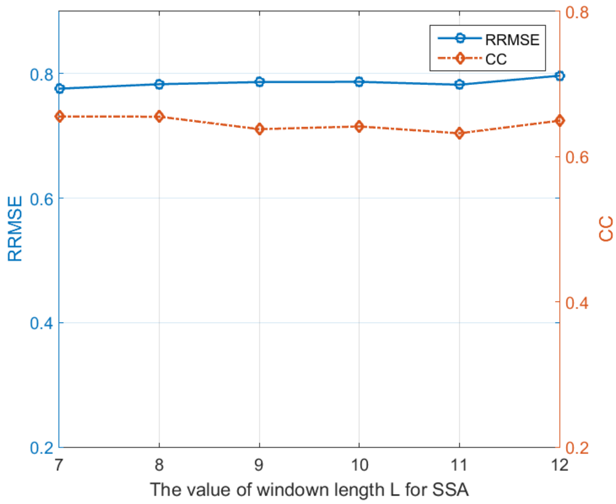

Fig. 7. The performance of artefact removal by combing SSA and MNF at SNR=1, where the window length L for SSA was set to different values.

composition. To further evaluate the effects of L on artefact removal, another simulation was conducted when L was set to different values. As an example, Fig. 7 shows the performance variations when L was set to different values. In the figure, the combination of SSA and MNF was applied, and SNR was set to 1. It is clear that neither RRMSE nor CC changes too much as the value of L increase from 7 to 12, which means that the performance is resilient to SSA decomposition window size L, and the value of L has very limited effects on the artefact removal. We also noticed that the combinations of SSA and other blind source separation methods also have the similar observations at different SNRs: MMSRE and CC are almost constant for different values of L. It once again illustrates the SSA-based method is hardly affected by the selection of L; therefore, in our implementation, we set $L = 8$ to get comparable components with regard to EEMD and WT decompositions.

IV. Conclusion

In this paper, we explored the source separation and artefacts removal performance for single channel recordings. Singular spectrum analysis (SSA) was applied to decompose single channel signal into multiple channels, which was then combined with different blind source separation methods, such as PCA, ICA, MNF and CCA to detect and remove artefacts. Other ways of combination among different different decomposition methods (WT and EEMD) and blind source separation methods (PCA, MNF, ICA and CCA) were also investigated for systematic comparison purpose. Although all the algorithms have achieved comparable accuracy, the SSA based methods have shown some implementation advantages over EEMD-based and WT-based methods.

In the future, although it is difficult to find the ground truth for each components for a combined signal in practice, real data experiments will be still carried out to further illustrate the reconstruction performance of different combinations of signal decomposition methods and blind sources separation methods. Proper separation performance evaluation methods will also be investigated.

References

[1] S. Patel, H. Park, P. Bonato, L. Chan, M. Rodgers *et al.*, "A review of wearable sensors and systems with application in rehabilitation," *J Neuroeng Rehabil*, vol. 9, no. 12, pp. 1–17, 2012.

[2] S. Ha, C. Kim, Y. M. Chi, A. Akinin, C. Maier, A. Ueno, and G. Cauwenberghs, "Integrated circuits and electrode interfaces for noninvasive physiological monitoring," *Biomedical Engineering, IEEE Transactions on*, vol. 61, no. 5, pp. 1522–1537, 2014.

[3] A. Lanata, G. Valenza, M. Nardelli, C. Gentili, and E. P. Scilingo, "Complexity index from a personalized wearable monitoring system for assessing remission in mental health," *Biomedical and Health Informatics, IEEE Journal of*, vol. 19, no. 1, pp. 132–139, 2015.

[4] N. W. Willigenburg, A. Daffertshofer, I. Kingma, and J. H. van Dieën, "Removing ecg contamination from emg recordings: A comparison of ica-based and other filtering procedures," *Journal of electromyography and kinesiology*, vol. 22, no. 3, pp. 485–493, 2012.

[5] J. Hu, C.-s. Wang, M. Wu, Y.-x. Du, Y. He, and J. She, "Removal of eog and emg artifacts from eeg using combination of functional link neural network and adaptive neural fuzzy inference system," *Neurocomputing*, vol. 151, pp. 278–287, 2015.

[6] W. A. W. Azlan and Y. F. Low, "Feature extraction of electroencephalogram (eeg) signal-a review," in *Biomedical Engineering and Sciences (IECBES), 2014 IEEE Conference on*. IEEE, 2014, pp. 801–806.

[7] M. E. Davies and C. J. James, "Source separation using single channel ica," *Signal Processing*, vol. 87, no. 8, pp. 1819–1832, 2007.

[8] J. Lin and A. Zhang, "Fault feature separation using wavelet-ica filter," *NDT & E International*, vol. 38, no. 6, pp. 421–427, 2005.

[9] S. Abbaspour, H. Gholamhosseini, and M. Linden, "Evaluation of wavelet based methods in removing motion artifact from ecg signal," in *16th Nordic-Baltic Conference on Biomedical Engineering*. Springer, 2015, pp. 1–4.

[10] B. Mijović, M. De Vos, I. Gligorijević, J. Taelman, and S. Van Huffel, "Source separation from single-channel recordings by combining empirical-mode decomposition and independent component analysis," *Biomedical Engineering, IEEE Transactions on*, vol. 57, no. 9, pp. 2188–2196, 2010.

[11] M. Hassan, S. Boudaoud, J. Terrien, B. Karlsson, and C. Marque, "Combination of canonical correlation analysis and empirical mode decomposition applied to denoising the labor electrohysterogram," *Biomedical Engineering, IEEE Transactions on*, vol. 58, no. 9, pp. 2441–2447, 2011.

[12] K. T. Sweeney, S. F. McLoone, and T. E. Ward, "The use of ensemble empirical mode decomposition with canonical correlation analysis as a novel artifact removal technique," *Biomedical Engineering, IEEE Transactions on*, vol. 60, no. 1, pp. 97–105, 2013.

[13] X. Chen, A. Liu, H. Peng, and R. K. Ward, "A preliminary study of muscular artifact cancellation in single-channel eeg," *Sensors*, vol. 14, no. 10, pp. 18 370–18 389, 2014.

[14] J. B. Elsner and A. A. Tsonis, *Singular spectrum analysis: a new tool in time series analysis*. Springer Science & Business Media, 2013.

[15] E. S. Barriga, M. Pattichis, D. Tso, M. Abramoff, R. Kardon, Y. Kwon, and P. Soliz, "Independent component analysis using prior information for signal detection in a functional imaging system of the retina," *Medical image analysis*, vol. 15, no. 1, pp. 35–44, 2011.

[16] M. Girolami, "An alternative perspective on adaptive independent component analysis algorithms," *Neural Computation*, vol. 10, no. 8, pp. 2103–2114, 1998.

[17] M. Chen and P. Zhou, "A novel framework based on fastica for high density surface emg decomposition," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2015.