

Research Article

Fast Extraction for Skewed Source Signals Using Conditional Expectation

Qiao Su , Yimin Wei, Changliang Deng, and Yuehong Shen

College of Communication Engineering, Army Engineering University of PLA, Nanjing 210007, China

Correspondence should be addressed to Qiao Su; qiaosu810@foxmail.com

Received 14 December 2017; Revised 12 June 2018; Accepted 2 July 2018; Published 22 July 2018

Academic Editor: Fanli Meng

Copyright © 2018 Qiao Su et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The extraction of the stochastic source signals whose probability density functions (PDFs) are skewed is very important in many applications such as biomedical signal processing and mechanical fault diagnosis. This paper shows that the skewed source signal with the maximal absolute value of skewness can be fast extracted by a proposed algorithm using conditional expectation. Compared with the existing conditional expectation-based algorithms, the proposed one possesses two main advantages. One is that it does not require the prior knowledge of the positive support of the desired source, namely the time indices where the source of interest is positive. The other is that it can be employed both in the determined and underdetermined cases. Furthermore, the proposed algorithm is mainly based on the first- and second-order statistics and does not need the preprocessing so that the computational cost is significantly low. Simulation results show the superiority of the proposed algorithm over the existing methods and indicate that the proposed algorithm also performs well in the underdetermined case when the number of sensors is slightly less than that of sources.

1. Introduction

The target of independent source extraction is to estimate a specific source from the observations mixed by the source signals, where the source signals are mutually independent. It can be applied in various areas such as speech and image processing, biomedical signal processing, mechanical fault diagnosis, and wireless communication. Due to these wide application fields, independent source extraction has gained much attention in the past few decades [1, 2]. In some applications especially biomedical signal processing and mechanical fault diagnosis, it is of significance to extract the stochastic source signals with the skewed probability density functions (PDFs). For instance, the mechanical vibrations derived from defective bearings, which are desired to be extracted in vibration analysis, may have asymmetric PDF [3], and the fetal electrocardiogram (FECG) signal has quite different skewness compared to the maternal electrocardiogram (MECG) signals [4], where the FECG requires to be estimated from the mixtures of the FECG and the MECG. The existing extraction algorithms mainly employ the

second- and higher-order statistics by exploiting the statistical independence of the sources, and some of them need to preprocess the mixture data such as FastICA [5]. Recently, Zarzoso et al. [6] and Xu et al. [7] and Xu and Shen [8] proposed a class of more computationally efficient algorithms for independent source extraction based on the first-order statistics as well as the conditional expectation. But before performing this class of algorithm, it requires to know the time indices where the source of interest is positive, that is, the positive support of the desired source, which reduces the practicability of this class of algorithm.

In this paper, we propose a new extraction algorithm using the conditional expectation for the skewed source signal with the maximal absolute value of skewness. It does not require the prior knowledge of the positive support of the desired source and can be applied both in the determined and underdetermined cases. It should be noted that after obtaining the source signal with the maximal absolute value of skewness, the source signal with the second maximal absolute value of skewness can also be extracted by the proposed algorithm from the mixtures which subtract the component

of the estimated source (e.g., through linear regression as in [6]). Likewise, if required, the other skewed source signals with different skewness can be gained sequentially. For simplicity, this paper assumes that the skewed source signal with the maximal absolute value of skewness is the desired source. The proposed algorithm obtains the desired column vector of the mixing matrix corresponding to the source of interest by the conditional expectation and retrieves the desired source by the minimum mean-squared error-based (MMSE) beamforming approach [9]. Through several iterations with the initial value of the desired column vector of the mixing matrix gotten by the estimated approximately purely positive or negative interval, the estimation of the desired source can be derived accurately. The proposed algorithm is rather cost-effective, since it is mainly based on the first- and second-order statistics and does not require the preprocessing. Simulation results validate the superiority of the proposed algorithm over the existing methods and show that the proposed algorithm even performs well in the underdetermined case when the number of sensors is close to that of sources.

2. Data Model

Consider the instantaneous linear mixture shown by

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \quad (1)$$

where $\mathbf{s}(t) = [s_1(t), \dots, s_N(t)]^T$ is composed of N source signals which are mutually independent, $\mathbf{x}(t) = [x_1(t), \dots, x_M(t)]^T$ consists of M mixtures received by sensors, \mathbf{A} is the unknown $M \times N$ mixing matrix, and the superscript T represents the transpose operator. In this paper, we consider that the source signals are stochastic with the unimodal continuous PDFs, since many signals such as some vibrational signals [10] and ECG signals [11] possess these characteristics. For convenience, we further assume that the source signals are stationary with zero mean and unit variance. Note that the assumption of stationary source signals is reasonable, since some non-stationary sources can also be divided to be several stationary blocks and we can tackle these stationary blocks separately. For example, some ECG signals can be regarded to be stationary in the duration of one heartbeat, despite they are nonstationary signals [12]. Our goal is to extract the skewed source signal which has the maximal absolute value of skewness from the M accessible mixtures.

3. The Proposed Algorithm

Since the first-order statistics-based algorithms [6–8] generally extract the source of interest under the condition that the mixing matrix is a unitary matrix, they can be only used in the determined case. In this paper, we remove this condition and obtain the column vector of the mixing matrix corresponding to the desired source based on the conditional expectation shown by

$$\mathbf{a}_d = \frac{E\{\mathbf{x}(t)|_{s_d(t)>0}\}}{\alpha}, \quad (2)$$

or $\mathbf{a}_d = \frac{E\{\mathbf{x}(t)|_{s_d(t)<0}\}}{\beta},$

where $E\{\cdot\}$ denotes the expectation operator, $s_d(t)$ is the desired source, \mathbf{a}_d is the d th column vector of the mixing matrix, $\alpha = E\{s_d(t) |_{s_d(t)>0}\}$, and $\beta = E\{s_d(t) |_{s_d(t)<0}\}$. Obviously, α and β are the constants according to the stationary assumption of the sources. The proof of (2) can be easily deduced from the assumptions of the sources, that is,

$$\frac{E\{\mathbf{x}(t)|_{s_d(t)>0}\}}{\alpha} = \frac{\mathbf{A}E\{\mathbf{s}(t)|_{s_d(t)>0}\}}{\alpha} = \mathbf{A}\mathbf{e}_d = \mathbf{a}_d, \quad (3)$$

where \mathbf{e}_d is the unit vector in which the d th entry is 1 and the other entries are 0. The proof of the other equation when $s_d(t) < 0$ in (2) is similar with (3). Then, we estimate the desired source by the MMSE beamforming approach, which is

$$\widehat{s}_d(t) = (\mathbf{C}_x^{-1}\alpha_d)^T \mathbf{x}(t), \quad (4)$$

where $\widehat{s}_d(t)$ is the estimation of $s_d(t)$, $\mathbf{C}_x = E\{\mathbf{x}^T(t)\mathbf{x}(t)\}$ is the covariance matrix of the mixtures, and the superscript -1 stands for the inversion operator. Thus, when α and β and the positive or negative support of the desired source are provided, the desired source can be estimated by (2) and (4). Actually, α and β can be easily figured out when the PDF of the desired source is known. The values of α for some normalized distributions are shown in [6]. However, when the PDF of the desired source is unknown, α and β are not obtainable. In this case, we can only get the direction of the vector \mathbf{a}_d by (2) which is the same with the direction of $E\{\mathbf{x}(t) |_{s_d(t)>0}\}$ or $E\{\mathbf{x}(t) |_{s_d(t)<0}\}$. Since the estimated \mathbf{a}_d has the correct direction and unknown size, it will lead to the ambiguous amplitude of the estimated $s_d(t)$ by (4). Fortunately, this indeterminacy of amplitude for estimating the desired source is allowable in many applications. For simplicity, we set $\alpha = \beta = 1$, and then (2) denotes the direction of \mathbf{a}_d . Unless stated otherwise, in the rest of this section, estimating \mathbf{a}_d refers to estimate the direction of the vector \mathbf{a}_d .

In practical applications, the complete information about the positive or negative support of the desired source is extremely hard to be acquired. However, it is more possible to get a subset of the samples of the desired source in which the positive samples are more than the negative samples obviously, or it is the opposite. We can see that this subset is close to a purely positive or negative set. We define the correct index classification ratio $r = N_1/N$ as in [6], where N_1 is the number of the positive samples in the subset and N is the total number of the samples in the subset. It was suggested in [6] that when r is close to 0 or 1, the subset can be also used to estimate the desired source and the estimation performance is only slightly worse than that of employing the complete information about the positive or negative support of the desired source. Similarly, if we get the subset like this, we

can use the information about this subset to roughly estimate \mathbf{a}_d . Fortunately, for a signal with unimodal continuous skewed distribution, we can utilize the asymmetry of its PDF to get the subset which is close to be purely positive or negative.

Figure 1 shows the PDF of unimodal continuous skewed distribution, where ξ denotes the skewness defined by $\xi = E\{(y - \mu)^3 / \sigma^3\}$ in which μ and σ are the mean and standard deviation of the random variable y , respectively. $\xi > 0$, $\xi < 0$, and $\xi = 0$ in Figure 1 represent the positively skewed distribution, the negatively skewed distribution, and the symmetric distribution, respectively. Since we consider the case when y is zero mean and unit variance, the definition of skewness is reduced to be $\xi = E\{y^3\}$. When y is subject to a positively skewed distribution ($\xi > 0$), we take a finite set of samples, Θ , generated by the distribution of y into account. As the skewness reflects the asymmetry of a PDF, a larger absolute value of skewness means stronger asymmetry of a PDF. Thus, it is easily deduced that most samples in Θ are smaller than zero, and the proportion of the negative samples increases when ξ rises. When randomly extracting some samples from Θ to form a subset, we can find that the subset may get close to a purely negative set. And with the increase of ξ , it will be more likely that the samples in this subset are completely negative. Likewise, in the case with $\xi < 0$, we can obtain the similar results. These provide the possibility to extract the source with the maximal absolute value of skewness based on the conditional expectation. Nevertheless, the random extraction is unstable, because this method is probabilistic. Instead, we separate the whole samples into several equivalent intervals and test all the intervals to find out the one closest to be purely positive or negative. Note that the size of each interval should be appropriate. If the size of each interval is too large, there will be high probability to contain the positive and negative values simultaneously. On the contrary, too small size of each interval cannot present the skewness characteristic of the skewed signal, since exhibiting the skewness requires a certain quantity of samples. The proper size should be adjusted depending on the nature of the signal, mainly its PDF. In the following section, we show that this proper size and the number of the intervals can be gotten by the simulations.

According to the analysis above, we divide the mixtures into K equivalent intervals, and the size of each interval is $L = \text{floor}(T/K)$, where T is the number of samples and $\text{floor}(\cdot)$ represents the floor function. Meanwhile, the inaccessible desired source is divided into K intervals accordingly. Assume that the i th interval of the desired source has c_i positive samples. Then, in the i th interval, the correct index classification ratio r_i is c_i/L . The objective is to find the optimal interval in which the samples of the desired source are closest to be purely positive or negative. Thus, \mathbf{a}_d can be roughly estimated by (2) through employing the information of this optimal interval. Assume that this optimal described by b interval is the k th interval which can be mathematically expressed as

$$\hat{k} = \arg \max |r_i - 0.5|_{i=1, \dots, K}. \quad (5)$$

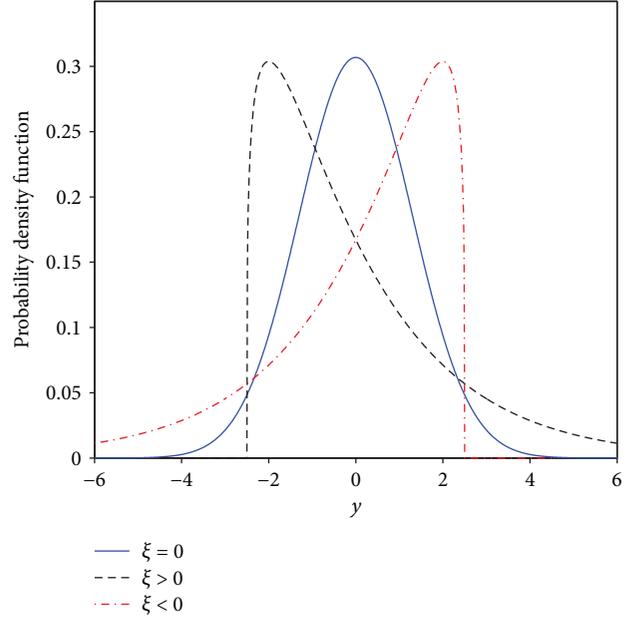


FIGURE 1: The PDF of unimodal continuous skewed distribution, where ξ denotes the skewness and $\xi > 0$, $\xi < 0$, and $\xi = 0$ indicate the positively skewed distribution, the negatively skewed distribution, and the symmetric distribution, respectively.

Then, we propose a new method to find the index \hat{k} . We suppose that the q th interval of the desired source is purely positive or negative, namely, $r_q = 0$ or 1 . According to (2), \mathbf{a}_d is estimated by

$$\hat{\mathbf{a}}_{d_q} = \frac{1}{L} \sum_{t \in \Omega_q} \mathbf{x}(t), \quad (6)$$

where Ω_q is a set constituted by the indices in the q th interval. Nevertheless, r_q is almost impossible to be 0 or 1, which leads to the inaccurate result by (6). We further rewrite the right side of (6) as

$$\frac{1}{L} \sum_{t \in \Omega_q} \mathbf{x}(t) = \frac{1}{L} \mathbf{A} \sum_{t \in \Omega_q} s(t). \quad (7)$$

According to (3), if $\sum_{t \in \Omega_q} s(t)/L$ is the integer multiple of e_d , $\hat{\mathbf{a}}_{d_q}$ will be the accurate estimation of \mathbf{a}_d . Since the desired source has the maximal absolute value of skewness, the \hat{k} th interval of the desired source is closer to a pure positive or negative region and has much greater absolute value of the sum of the samples than the other sources' \hat{k} th interval, namely,

$$\left| \sum_{t \in \Omega_{\hat{k}}} s_d(t) \right| \gg \left| \sum_{t \in \Omega_{\hat{k}}} s_d(t) \right|_{j \in \{1, \dots, N\}, j \neq d}. \quad (8)$$

Then, we have $\sum_{t \in \Omega_{\hat{k}}} \mathbf{s}(t)/L \approx a \mathbf{e}_d$, where a is an integer. Therefore, by performing (6) with the indices in the \hat{k} th interval, we can obtain the approximate estimation $\hat{\mathbf{a}}_{d_{\hat{k}}}$ of

\mathbf{a}_d . Obviously, the estimation accuracy of $\hat{a}_{d_{\hat{k}}}$ is better than that of \hat{a}_{d_p} ($p \in \{1, \dots, K\}, p \neq \hat{k}$) which is the estimation of \mathbf{a}_d by employing the other intervals. Using (4), we get the estimations of the desired source written as $\hat{s}_{d_{\hat{k}}}(t)$ and $\hat{s}_{d_p}(t)$ by applying $\hat{a}_{d_{\hat{k}}}$ and \hat{a}_{d_p} , respectively. It is obvious that $\hat{s}_{d_{\hat{k}}}(t)$ has higher estimation precision than $\hat{s}_{d_p}(t)$, so the absolute value of the skewness of $\hat{s}_{d_{\hat{k}}}(t)$, $|\xi_{\hat{k}}|$ is larger than the absolute value of the skewness of $\hat{s}_{d_p}(t)$, $|\xi_p|$. Given all that, the index \hat{k} of objective interval is the index corresponding to the maximum of $|\xi_i|_{i=1, \dots, K}$.

By means of the above method, \mathbf{a}_d can be roughly estimated as $\hat{a}_{d_{\hat{k}}}$ by the information of the \hat{k} th interval. In order to improve the estimate accuracy, we propose an iteration method which is implemented by iterating (4) and (2) with the initial value $\hat{a}_{d_{\hat{k}}}$ until convergence. The outline of the proposed algorithm is shown below.

- (1) Calculate \mathbf{C}_x and its inverse.
- (2) Divide the mixtures into K intervals, compute \hat{a}_{d_i} by (6), and then obtain $\hat{s}_{d_i}(t)$ by (4) and the corresponding $|\xi_i|$.
- (3) Select the index \hat{k} corresponding to the maximum of $|\xi_i|_{i=1, \dots, K}$ and get the initial value $\hat{a}_{d_{\hat{k}}}$ by (6).
- (4) Achieve $\hat{s}_d(t)$ by (4) and acquire the positive or negative support of $\hat{s}_d(t)$.
- (5) Obtain the estimation of \mathbf{a}_d by (2) through employing the knowledge about the positive or negative support in the step 4.
- (6) Iterate the step 4 and the step 5 until convergence.

4. Simulation Results

We consider five skewed source signals and fifteen symmetric distributed source signals ($N = 20$). The five skewed source signals are generated as follows. Firstly, we generate five skewed signals based on the noncentral t -distributions [13] whose degrees of freedom and noncentrality parameters are (10, 2), (10, 1), (12, -1), (15, -2), (20, 1), respectively. Then, we centralize and standardize these skewed signals so that they are zero mean and unit variance. The values of skewness of the five skewed source signals are 0.6701, 0.4880, -0.5532, -0.3527, and 0.1486, respectively. The symmetric distributed source signals are generated based on the standard normal distribution. The $M \times N$ mixing matrix is generated randomly, whose entries are subject to the standard normal distribution. We set the number of samples to 600 and rate to 300. We aim at extracting the skewed source signal with the maximal absolute value of skewness from the M mixtures. FastICA [5] and the algorithm in [6] will be compared with the proposed algorithm.

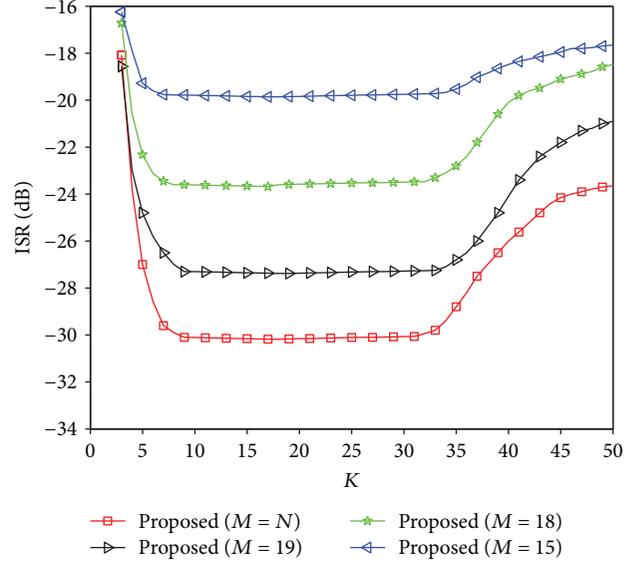


FIGURE 2: The average ISR versus K obtained by the proposed algorithm over 200 Monte Carlo runs when $M = N$, $M = 19$, $M = 18$, and $M = 15$.

We denote the algorithm in [6] as the first-order algorithm (FOA for short).

Figure 2 illustrates the average interference-to-signal ratio (ISR) (defined in [6]) versus K obtained by the proposed algorithm over 200 Monte Carlo runs when $M = N$, $M = 19$, $M = 18$, and $M = 15$. It can be observed from Figure 2 that the proposed algorithm has the low extraction performance when K is selected to be too small or large. Since each interval should not be too short or long according to the discussion above, K needs to be set properly. Based on Figure 2, we choose K in (8, 33) empirically and set $K = 15$ in the following experiments. In Figure 3, we show the average loci of the ISR versus the iteration number obtained by FastICA, FOA with $r = 0.6$ and $r = 0.8$ when $M = N$, and the proposed algorithm when $M = N$, $M = 19$, $M = 18$, and $M = 15$ over 200 Monte Carlo runs, where r is the correct index classification ratio defined in [6]. As depicted in Figure 3, the proposed algorithm can extract the desired source successfully after several iterations both in the determined and underdetermined cases. When $M = N$, the proposed algorithm has better extraction performance and faster convergence rate than FastICA. Moreover, the proposed algorithm possesses a faster convergence rate than FOA with $r = 0.6$ when $M = N$ and the proposed algorithm and FOA with $r = 0.6$ and $r = 0.8$ when $M = N$ own the same performance of the steady state. Although the performance of the convergence rate of FOA with $r = 0.8$ is better than that of the proposed algorithm, FOA is hard to get a priori knowledge satisfying $r = 0.8$. It can also be seen from Figures 2 and 3 that the extraction performance of the proposed algorithm deteriorates with the decrease of the number of sensors, M . However, the proposed algorithm still performs well when M is slightly less than N in the underdetermined case.

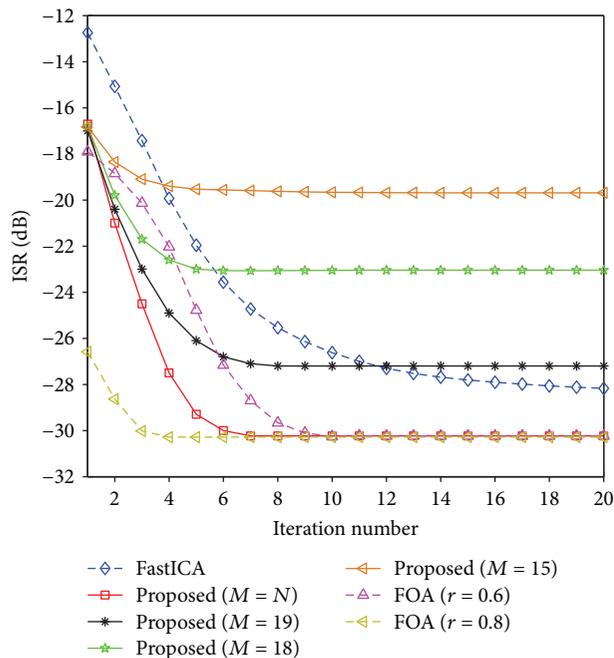


FIGURE 3: The average ISR versus the iteration number obtained by FastICA, FOA with $r=0.6$ and $r=0.8$ when $M=N$, and the proposed algorithm when $M=N$, $M=19$, $M=18$, and $M=15$ over 200 Monte Carlo runs.

TABLE 1: The average computational cost of fastICA, FOA with $r=0.6$, when $M=N$, and the proposed algorithm when $M=N$, $M=19$, and $M=15$ over 200 Monte Carlo runs.

	FastICA	Proposed ($M=N$)	Proposed ($M=19$)	Proposed ($M=15$)	FOA ($r=0.6$)
Time (s)	0.1601	0.0104	0.0092	0.0088	0.0032

Additionally, we compare the average computational cost of FastICA, FOA with $r=0.6$ when $M=N$, and the proposed algorithm when $M=N$, $M=19$, and $M=15$ over 200 Monte Carlo runs as shown in Table 1. We can see from Table 1 that the proposed algorithm costs less in computations than FastICA and the proposed algorithm with less M has slightly lower computational cost. It can also be observed from Table 1 that FOA with $r=0.6$ possesses lower computational cost than the proposed algorithm. This is because FOA is only based on the first-order statistics. However, FOA requires some prior knowledge of the desired source which is difficult to get and cannot be applied into the underdetermined case. So, the proposed algorithm is more practical than FOA.

5. Conclusion

In this paper, we proposed a cost-effective algorithm based on the conditional expectation for the extraction of the skewed source signal with the maximal absolute value of skewness. Simulation results testify the superiority of the proposed algorithm and validate that it performs well even in the

underdetermined case when the sensor number is close to the source number. Future research of this work should extend to study more complicated and more practical mixture modes such as the convolutive mixture and the nonlinear mixture.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant nos. 61172061 and 61201242 and the National Natural Science Foundation of Jiangsu Province in China under Grant no. BK2012057.

References

- [1] R. Phlypo, V. Zarzoso, and I. Lemahieu, "Source extraction by maximizing the variance in the conditional distribution tails," *IEEE Transactions on Signal Processing*, vol. 58, no. 1, pp. 305–316, 2010.
- [2] A. K. Barros and A. Cichocki, "Extraction of specific signals with temporal structure," *Neural Computation*, vol. 13, no. 9, pp. 1995–2003, 2001.
- [3] A. K. Ovacikli, P. Paajarvi, and J. P. Leblanc, "Skewness as an objective function for vibration analysis of rolling element bearings," in *2013 8th International Symposium on Image and Signal Processing and Analysis (ISPA)*, pp. 462–466, Trieste, Italy, September 2013.
- [4] F. Jafari, M. A. Tinati, and B. Mozaffari, "A new fetal ECG extraction method using its skewness value which lies in specific range," in *2010 18th Iranian Conference on Electrical Engineering*, pp. 30–34, Isfahan, Iran, May 2010.
- [5] A. Hyvarinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural Computation*, vol. 9, no. 7, pp. 1483–1492, 1997.
- [6] V. Zarzoso, R. Martín-Clemente, and S. Hornillo-Mellado, "Independent component analysis based on first-order statistics," *Signal Processing*, vol. 92, no. 8, pp. 1779–1784, 2012.
- [7] P. C. Xu, Y. H. Shen, H. Li, J. G. Wang, and K. Wu, "Independent component analysis of complex valued signals based on first-order statistics," *Radioengineering*, vol. 22, no. 4, pp. 1194–1201, 2013.
- [8] P. C. Xu and Y. H. Shen, "Convolutive independent component analysis based on first-order statistics for complex-valued source extraction," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 9, no. 1, pp. 245–264, 2016.
- [9] Z. Koldovský, P. Tichavský, A. H. Phan, and A. Cichocki, "A two-stage MMSE beamformer for underdetermined signal separation," *IEEE Signal Processing Letters*, vol. 20, no. 12, pp. 1227–1230, 2013.
- [10] F. Honarvar and H. R. Martin, "New statistical moments for diagnostics of rolling element bearings," *Journal of Manufacturing Science and Engineering*, vol. 119, no. 3, pp. 425–432, 1997.
- [11] Y. Yin M. Ye et al., "FECG extraction using Bayesian inference and neural networks approximation," *Journal of Computational Information Systems*, vol. 6, no. 6, pp. 1769–1778, 2010.

- [12] P. Tichavsky and Z. Koldovsky, "Weight adjusted tensor method for blind separation of underdetermined mixtures of nonstationary sources," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 1037–1047, 2011.
- [13] M. Evans, N. Hastings, and B. Peacock, *Statistical Distributions*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2nd edition, 1993.



Hindawi

Submit your manuscripts at
www.hindawi.com

