A fixed-point blind source extraction algorithm and its application to ECG data analysis

Hongjuan Zhang
Department of Mathematics
Shanghai University
Shanghai,200444,P R China
Email: zhanghongjuan@shu.edu.cn

Zikai Wu
Business School
University of Shanghai for Science and Technology
Shanghai,200093,P R China
Email: zkwu2011@gmail.com

Shuxue Ding
Department of Computer Software
The University of Aizu
Tsuruga, Ikki-Machi, Aizu-Wakamatsu City
Fukushima,965-8580,Japan
Email: sding@u-aizu.ac.jp

Luonan Chen
Key Laboratory of Systems Biology
Shanghai Institutes for Biological Sciences,
Chinese Academy of Sciences
Shanghai,200031,P R China
lnchen@sibs.ac.cn

***Abstract*—Generalized autocorrelations and complexity pursuit are two recently developed methods for extracting interesting component from time series. They are the extensions of projection pursuit to time series data. In this paper, a fixed-point blind source extraction (BSE) algorithm for generalized autocorrelations and complexity pursuit of the desired signals is presented. The fixed-point algorithm inherits the advantages of the well-known FastICA algorithm of ICA, which is very simple, converges fast, and does not need to choose any learning step sizes. Numerical experiments on electrocardiogram (ECG) data indicate its better performance.**

I. Introduction

Independent component analysis (ICA) is an active research field that has attracted great interest in the field of biomedical signal analysis and processing, geophysical data processing, data mining, speech analysis and image recognition, etc[1], [2], [3], [4], [5], [6], [7]. The model of ICA consists of mixing independent random variables, usually linearly. In many applications, however, what are mixed are some time-dependent signals, or time series, which may or may not be random variables. When the desired source signals are periodic or quasi-periodic, one convenient way to exploit these signals is to employ the time-dependent information along. Many source extraction algorithms have considered such a case. For example, Barros and Cichocki [8] provided a simple algorithm (simplified with ”BCBSE” algorithm) which can quickly extract a desired source signal with a specific period. This algorithm, in all the cases, could extract the desired sources, no matter they are colored or not, as long as they are decorrelated and have temporal structures. However, this method carries out the constrained minimization of the mean squared error only, which can not well describe the probability distribution of the innovations of the signals. In addition, it needs *a prior* information about the optimal time delay and is very sensitive to the estimation error of the time delay. To overcome these drawbacks, Shi et al. [9] developed a blind source extraction algorithm (simplified with ”SemiBSE” algorithm), which is based on the non-Gaussianity and the autocorrelation of the source signal and contains either the mean squared error objective function presented by Barros and Cichocki [8]. This method can improve the performance of BCBSE algorithm, and its tolerance to large estimated errors of the period makes that the desired signal can be extracted robustly. However, it must be noted that its better tolerance lies the fact that the choice of initial weight is not random but is determinant, i.e., the unit vector. If the weight is initialized randomly, SemiBSE algorithm becomes more sensitive to the estimation error of time delay. An alternative approach, for extracting desired source signal with linear or nonlinear autocorrelations, was first introduced in literature [10]. Based on the generalized (i.e., linear or nonlinear) autocorrelations of the primary sources, the authors proposed a blind source extraction (BSE) algorithm (called ”GABSE” algorithm). It has been shown that this method has good stability and linear convergence speed. Furthermore, the convergence is global except for a zero measure region. This method only assumes the sources are decorrelated with each other and every source has different temporal structures, but does not necessarily have to be statistically independent. GABSE algorithm has been applied to many cases directly and the performance is satisfying in a certain extent. Whereas, it also suffers some disadvantages; for example, its tolerance to estimated errors of time delay is not very robust. Based on literatures [9], [10], Zhang et al [11] developed an object function, which incorporates the generalized autocorrelations and the complexity pursuit of the desired signals, for extraction of the desired signal. Based on these priori special characteristics, a gradient BSE algorithm (simplified with ”GACP” algorithm) for the approximate optimization of proposed objective function was given. Numerical computation and theoretical analysis showed that GACP algorithm both can improve the performance of

978-1-4673-4398-5/12/$31.00 ©2012 IEEE

existing algorithms and has better robustness to the estimated error of time delay. However, we all know that the gradient algorithm need to choose some learning step sizes, which increases the number of parameter in GACP algorithm, and degrade its convergence properties.

In this paper, a fixed-point BSE algorithm for generalized autocorrelations and complexity pursuit of the desired signals is presented. This fixed-point algorithm inherits the advantages of the well-known FastICA algorithm for ICA, which is very simple, converges fast, and does not need to choose any learning step sizes. Numerical experiments on electrocardiogram (ECG) data indicate its better performance.

This manuscript is organized as follows. The next section describes the objective function and the proposed algorithm. Section 3 demonstrates the present technique with experiments using ECG data. The final section provides some discussions and conclusions.

II. Proposed algorithm

A. Objective function

Denote the observed sensor signals $\mathbf{x}(t) = (x_1(t), \cdots, x_n(t))^T$ described by matrix equation

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

where $\mathbf{A}$ is an $n \times n$ unknown mixing matrix and $\mathbf{s}(t) = (s_1(t), \cdots, s_n(t))^T$ is a vector of unknown temporally correlated sources. We assume that desired source signal s_i has specific temporal structures—linear or nonlinear autocorrelations. And it can be modeled by a linear autoregressive model, which has just one predicting term as

$$s_i(t) = b_i s_i(t-\tau) + \delta_i(t), \tag{2}$$

where $\delta_i(t)$ is a zero-mean, independent identically distributed ($i.i.d$) time series called innovation[12], b_i is a coefficient and τ is a delay in time.

Because we want to extract only a desired source signal, for this purpose we design a single neural processing unit described as

$$\tilde{y}(t) = \mathbf{w}^T\tilde{\mathbf{x}}(t), \tag{3}$$
$$\tilde{y}(t-\tau) = \mathbf{w}^T\tilde{\mathbf{x}}(t-\tau), \tag{4}$$

where $\tilde{y}(t)$ and $\tilde{y}(t-\tau)$ are the extracted signals at time t and $(t-\tau)$, respectively. $\mathbf{w} = (\mathrm{w}_1, \cdots, \mathrm{w}_n)^T$ is the weight vector. Here, assume that the measured sensor signals $\mathbf{x}$ have already been followed with an $n \times n$ whitening filter $\mathbf{V}$ so that the components of $\tilde{\mathbf{x}}(t) = \mathbf{V}\mathbf{x}(t) = \mathbf{V}\mathbf{A}\mathbf{s}(t) = \tilde{\mathbf{A}}\mathbf{s}(t)$ are of unit variance and uncorrelated, where the new mixing matrix $\tilde{\mathbf{A}}$ is orthogonal[13].

When extracting a desired signal, we are interested in its temporal characteristics (e.g. autocorrelation) of the interesting signal and the probability distribution of its innovations (e.g. non-Gaussianity). Then, the extraction of desired signal can be formulated as the following constrained problem by maximizing the convex combination between generalized autocorrelations and the complexity pursuit of the desired signals.

$$\begin{aligned}\max_{\|\mathbf{w}\|=1} \Psi(\mathbf{w}) &= \lambda E\{G(\tilde{y}(t))G(\tilde{y}(t-\tau))\} \\ &\quad -(1-\lambda)E\{F(\tilde{y}(t) - b\tilde{y}(t-\tau))\},\end{aligned} \tag{5}$$

where $(\tilde{y}(t) - b\tilde{y}(t-\tau))$ represents an innovation of the extractions, b is a coefficient. λ is a scalar parameter between 0 and 1 for controlling the balance between the generalized autocorrelations and the complexity pursuit of the desired signals. For instance, only the generalized autocorrelations are considered in the extraction for $\lambda = 1$, while only the probability distributions of innovations are optimized for $\lambda = 0$. Generally, the parameter is $0 < \lambda < 1$, depending on the requirement of extraction. The function G is a differentiable function which measures the autocorrelation degree of the desired signal. Examples of choices are $G(u) = u$, $G(u) = u^2$ and $G(u) = \log\cosh(u)$. F is a differentiable function too, which should be determined by the probability distribution of innovations [12], [14], [15].

B. Learning algorithm

To find the maxima of the objective function (5), we can use a fixed-point iteration along a similar line of the FastICA algorithm for maximizing the nongaussianity[6]. The fixed-point algorithm can be found using an approximate Newton method. For simplicity of the equations the time index t is omitted in the following, i.e., $\tilde{\mathbf{x}} = \tilde{\mathbf{x}}(t)$, $\tilde{\mathbf{x}_\tau} = \tilde{\mathbf{x}}(t-\tau)$, $\tilde{y} = \tilde{y}(t)$ and $\tilde{y}_\tau = \tilde{y}(t-\tau)$. To derive the approximate Newton method, we first denote by

$$\tilde{\mathbf{z}} = \tilde{\mathbf{x}} - b\tilde{\mathbf{x}}_\tau \tag{6}$$

According to the Lagrange conditions, the optima of $\Psi(\mathbf{w})$ under the constraint $\|\mathbf{w}\| = 1$ are obtained at points where the gradient of the Lagrangian is zero:

$$\begin{aligned}&\lambda E\{g(\tilde{y})G(\tilde{y}_\tau)\tilde{\mathbf{x}} + G(\tilde{y})g(\tilde{y}_\tau)\tilde{\mathbf{x}}_\tau\} \\ &-(1-\lambda)E\{\tilde{\mathbf{z}}f(\mathbf{w}^T\tilde{\mathbf{z}})\} + \beta\mathbf{w} = 0,\end{aligned} \tag{7}$$

where β is some constant and function g is the derivative of G. Now let us try to solve this equation by Newton method, which is equivalent to finding the optima of the Lagrangian by Newton's method. Denoting the function on the left-hand side of (7) by K, we obtain its Jacobian matrix $JK(\mathbf{w})$ as

$$\begin{aligned}\frac{\partial K(\mathbf{w})}{\partial \mathbf{w}} &= \lambda E\{g'(\tilde{y})G(\tilde{y}_\tau)\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T + G(\tilde{y})g'(\tilde{y}_\tau)\tilde{\mathbf{x}}_\tau\tilde{\mathbf{x}}_\tau^T \\ &+g(\tilde{y})g(\tilde{y}_\tau)(\tilde{\mathbf{x}}\tilde{\mathbf{x}}_\tau^T + \tilde{\mathbf{x}}_\tau\tilde{\mathbf{x}}^T)\} - (1-\lambda)E\{\tilde{\mathbf{z}}\tilde{\mathbf{z}}^T f'(\mathbf{w}^T\tilde{\mathbf{z}})\} \\ &+\beta\mathbf{I}.\end{aligned} \tag{8}$$

978-1-4673-4398-5/12/$31.00 ©2012 IEEE

We obtain the following approximate Newton iteration:

$$\begin{aligned}
&\mathbf{w} \leftarrow \mathbf{w} - K \Big/ \frac{\partial K(\mathbf{w})}{\partial \mathbf{w}} \\
&= \mathbf{w} - \Big[\lambda E\{g(\tilde{y})G(\tilde{y}_\tau)\tilde{\mathbf{x}} + G(\tilde{y})g(\tilde{y}_\tau)\tilde{\mathbf{x}}_\tau\} \\
&-(1-\lambda)E\{\tilde{\mathbf{z}}f(\mathbf{w}^T\tilde{\mathbf{z}})\} + \beta\mathbf{w}\Big] \Big/ \Big[\lambda E\{g'(\tilde{y})G(\tilde{y}_\tau)\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T \\
&+G(\tilde{y}) \cdot g'(\tilde{y}_\tau)\tilde{\mathbf{x}}_\tau\tilde{\mathbf{x}}_\tau^T + g(\tilde{y})g(\tilde{y}_\tau)[\tilde{\mathbf{x}}\tilde{\mathbf{x}}_\tau^T + \widetilde{\mathbf{x}_\tau}\tilde{\mathbf{x}}^T]\} \\
&-(1-\lambda)E\{\tilde{\mathbf{z}}\tilde{\mathbf{z}}^T f'(\mathbf{w}^T\tilde{\mathbf{z}})\} + \beta\mathbf{I}\Big].
\end{aligned} \tag{9}$$

This algorithm can be further simplified by multiplying both sides of (9) by $-\{\lambda E\{g'(\tilde{y})G(\tilde{y}_\tau)\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T + G(\tilde{y})g'(\tilde{y}_\tau)\tilde{\mathbf{x}}_\tau\tilde{\mathbf{x}}_\tau^T + g(\tilde{y})g(\tilde{y}_\tau)(\tilde{\mathbf{x}}\tilde{\mathbf{x}}_\tau^T + \tilde{\mathbf{x}}_\tau\tilde{\mathbf{x}}^T)\} - (1-\lambda)E\{\tilde{\mathbf{z}}\tilde{\mathbf{z}}^T f'(\mathbf{w}^T\tilde{\mathbf{z}})\} + \beta\mathbf{I}\}$. After straightforward algebraic simplification, this gives:

$$\begin{aligned}
\mathbf{w} \leftarrow &\Big[\lambda E\{g(\tilde{y})G(\tilde{y}_\tau)\tilde{\mathbf{x}} + G(\tilde{y})g(\tilde{y}_\tau)\tilde{\mathbf{x}}_\tau\} \\
&-(1-\lambda)E\{\tilde{\mathbf{z}}f(\mathbf{w}^T\tilde{\mathbf{z}})\}\Big] \\
&-\Big[\lambda E\{g'(\tilde{y})G(\tilde{y}_\tau)\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T + G(\tilde{y})g'(\tilde{y}_\tau)\tilde{\mathbf{x}}_\tau\tilde{\mathbf{x}}_\tau^T \\
&+g(\tilde{y})g(\tilde{y}_\tau)(\tilde{\mathbf{x}}\tilde{\mathbf{x}}_\tau^T + \tilde{\mathbf{x}}_\tau\tilde{\mathbf{x}}^T)\} \\
&-(1-\lambda)E\{\tilde{\mathbf{z}}\tilde{\mathbf{z}}^T f'(\mathbf{w}^T\tilde{\mathbf{z}})\}\Big]\mathbf{w}.
\end{aligned} \tag{10}$$

To simplify the resulting algorithm we use the following approximations.

$$E\{g'(\tilde{y})G(\tilde{y}_\tau)\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T\} \approx E\{g'(\tilde{y})G(\tilde{y}_\tau)\}E\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T\} \tag{11}$$

$$E\{G(\tilde{y})g'(\tilde{y}_\tau)\widetilde{\mathbf{x}_\tau}\tilde{\mathbf{x}}_\tau^T\} \approx E\{G(\tilde{y})g'(\tilde{y}_\tau)\}E\{\tilde{\mathbf{x}}_\tau\tilde{\mathbf{x}}_\tau^T\} \tag{12}$$

$$\begin{aligned}
&E\{g(\tilde{y})g(\tilde{y}_\tau)(\tilde{\mathbf{x}}\tilde{\mathbf{x}}_\tau^T + \widetilde{\mathbf{x}_\tau}\tilde{\mathbf{x}}^T)\} \\
&\approx E\{g(\tilde{y})g(\tilde{y}_\tau)\}E\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}_\tau^T + \tilde{\mathbf{x}}_\tau\tilde{\mathbf{x}}^T)\}
\end{aligned} \tag{13}$$

$$E\{\tilde{\mathbf{z}}\mathbf{z}^T f'(\mathbf{w}^T\tilde{\mathbf{z}})\} \approx E\{\tilde{\mathbf{z}}\tilde{\mathbf{z}}^T\}E\{f'(\mathbf{w}^T\tilde{\mathbf{z}})\}. \tag{14}$$

First, since $\tilde{\mathbf{x}}$ are whitened, we have,

$$E\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T\} = \mathbf{I} \tag{15}$$

$$E\{\widetilde{\mathbf{x}_\tau}\tilde{\mathbf{x}}_\tau^T\} = \mathbf{I}. \tag{16}$$

Then

$$E\{g'(\tilde{y})G(\tilde{y}_\tau)\}E\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T\} = E\{g'(\tilde{y})G(\tilde{y}_\tau)\}\mathbf{I} \tag{17}$$

$$E\{G(\tilde{y})g'(\tilde{y}_\tau)\}E\{\tilde{\mathbf{x}}_\tau\tilde{\mathbf{x}}_\tau^T\} = E\{G(\tilde{y})g'(\tilde{y}_\tau)\}\mathbf{I}. \tag{18}$$

Second, the innovation as the components of processes $\tilde{\mathbf{s}} = \mathbf{s} - b\mathbf{s}_\tau$ are independent each other, then

$$E\{\tilde{\mathbf{s}}\tilde{\mathbf{s}}^T\} = \mathbf{\Lambda}. \tag{19}$$

That is,

$$\begin{aligned}
&E\{\mathbf{s}\mathbf{s}^T\} + b^2 E\{\mathbf{s}_\tau\mathbf{s}_\tau^T\} - b(E\{\mathbf{s}\mathbf{s}_\tau^T + \mathbf{s}_\tau\mathbf{s}^T\}) \\
&= \mathbf{\Lambda},
\end{aligned} \tag{20}$$

where $\mathbf{\Lambda}$ denotes a diagonal matrix with $n \times n$. Due to source $\mathbf{s}$ is assumed of unit variance, the above equation (20) can be written as

$$(1 + b^2)\mathbf{I} - b(E\{\mathbf{s}\mathbf{s}_\tau^T + \mathbf{s}_\tau\mathbf{s}^T\}) = \mathbf{\Lambda}. \tag{21}$$

Multiplying both sides of (21) by $\mathbf{A}$ and $\mathbf{A}^T$, respectively, we obtain

$$E\{\mathbf{x}\mathbf{x}_\tau^T + \mathbf{x}_\tau\mathbf{x}^T\} = \frac{1}{b}[(1 + b^2)\mathbf{I} - E\{\tilde{\mathbf{z}}\tilde{\mathbf{z}}^T\}]. \tag{22}$$

Then

$$\begin{aligned}
&E\{g(\tilde{y})g(\tilde{y}_\tau)\}E\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}_\tau^T + \tilde{\mathbf{x}}_\tau\tilde{\mathbf{x}}^T\} \\
&= E\{g(\tilde{y})g(\tilde{y}_\tau)\}\frac{1}{b}[(1 + b^2)\mathbf{I} - E\{\tilde{\mathbf{z}}\tilde{\mathbf{z}}^T\}].
\end{aligned} \tag{23}$$

Therefore, a reasonable approximation about (10) is

$$\begin{aligned}
\mathbf{w} \leftarrow &\Big[\lambda E\{g(\tilde{y})G(\tilde{y}_\tau)\tilde{\mathbf{x}} + G(\tilde{y})g(\tilde{y}_\tau)\tilde{\mathbf{x}}_\tau\} \\
&-(1-\lambda)E\{\tilde{\mathbf{z}}f(\mathbf{w}^T\tilde{\mathbf{z}})\}\Big] \\
&-\Big[\lambda E\{g'(\tilde{y})G(\tilde{y}_\tau) + G(\tilde{y})g'(\tilde{y}_\tau) + g(\tilde{y})g(\tilde{y}_\tau)\gamma\} \\
&-(1-\lambda)E\{\tilde{\mathbf{z}}\tilde{\mathbf{z}}^T\}E\{f'(\mathbf{w}^T\tilde{\mathbf{z}})\}\Big]\mathbf{w},
\end{aligned} \tag{24}$$

where γ is chosen as $\frac{1}{b}[(1 + b^2)\mathbf{I} - E\{\tilde{\mathbf{z}}\tilde{\mathbf{z}}^T\}]$, g and f are the derivatives of the function G and F respectively. Notice that f should be chosen according to the probability distribution of the innovation. If it is super-Gaussian, $f(u) = \mathrm{sign}(u)$ is suitable, which can be approximated by a smoother function $f(u) = \tanh(au)$, where $a \geq 1$. For sub-Gaussian innovation, one can choose one of $f(u) = u - \tanh(u)$ or $f(u) = u^3$, for example. Note that the coefficient b can be learned by a least-squares method as

$$b = \mathbf{w}^T E\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}_\tau^T\}\mathbf{w}. \tag{25}$$

Specifically, the proposed fixed-point algorithm, which is simplified as FastGACP, can be described as follows:

Step 1. Center each observed signal $\mathbf{x}$ to make it mean zero and whiten them as $\widetilde{\mathbf{x}}$.

Step 2. Choose parameter λ, τ and randomly initialize a value of unit norm for $\mathbf{w}$; set $b = 1$.

Step 3. Updata $\mathbf{w}$ and b are by (24) and (25) respectively.

Step 4.Let $\mathbf{w} \leftarrow \mathbf{w}/||\mathbf{w}||$

Step 5. If not converged, go back to Step 3.

III. Numerical Experiments

In order to verify the efficiency of our algorithm, we made many numerical experiments with artificial ECG data and real-world ECG data[17]. Moreover, we compared the proposed method with several existing techniques– BCBSE algorithm[8], SemiBSE algorithm[9], GABSE algorithm[10] and GACP algorithm[11]. In these comparisons, the performance of algorithms for estimating the desired signal is measured by the performance index (PI), which is defined as follows

$$\mathrm{PI} = \sum_{j=1}^{n} \frac{|\ p_j\ |}{\max_k |\ p_k\ |} - 1, \quad k = 1, \ldots, n \tag{26}$$

where p_j denotes the j element of the global vector $\mathbf{p} = \mathbf{w}^T\mathbf{VA}$. PI is zero when the desired signal is perfectly extracted. Besides, the accuracy of the extracted signal compared

978-1-4673-4398-5/12/$31.00 ©2012 IEEE

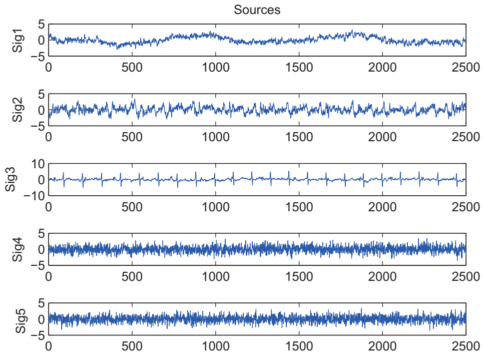

Fig. 1. Five artificial ECG signals. (Sig1) Breathing artifact. (Sig2) Electrode artifact. (Sig3) FECG. (Sig4-Sig5) Two $i.i.d.$ Gaussian signals.

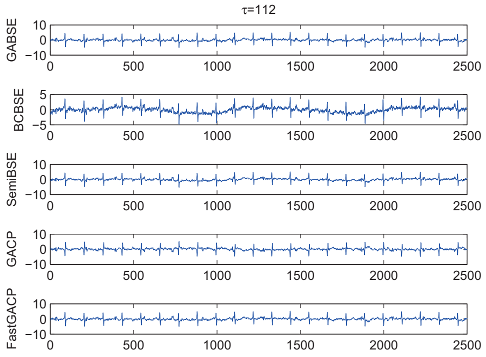

Fig. 2. Extractions for artificial ECG signals with $\tau = 112$. From top to bottom, the extracted FECGs by GABSE, BCBSE, SemiBSE, GACP and FastGACP algorithm.

to the true source is expressed using the signal-to-noise ratio (SNR) in dB given by

$$\mathrm{SNR} = 10\log_{10}(s^2/\mathrm{MSE}), \tag{27}$$

where s^2 denotes the variance of the source signal, MSE denotes the mean square error between the original signal and the extracted signal. The higher SNR is, the better performance is.

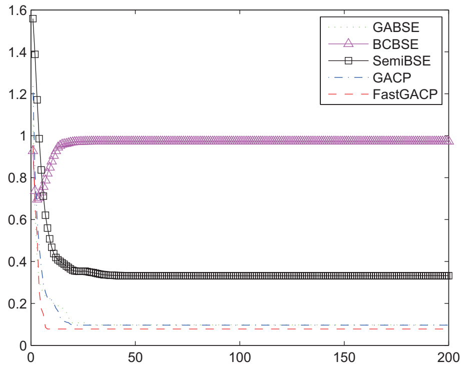

Fig. 3. Comparisons of average PIs over 100 independent runs for five algorithms with $\tau = 112$ on artificial ECG data.

A. Experiments on artificial ECG data

We adopted five zero-mean and unit-variance source signals (2500 samples), which are shown in Fig.1. From the top to down, they are, one breathing artifact, one electrode artifact, one fetal ECG (FECG) ($\tau = 112$), and two $i.i.d.$ Gaussian signals. The observed signals are generated by a 5×5 random mixing matrix. The extracted FECG signals are shown in Fig.2. The accuracies of them are 25.4808 dB(GABSE), 3.2434 dB (BCBSE), 9.9631 dB (SemiBSE), 26.3605 dB (GACP) and 26.9363 dB (FastGACP). The extracted FECG signal by FastGACP algorithm is the best, and those extracted by GACP and GABSE algorithm are also satisfying. However, the extractions by SemiBSE and BCBSE algorithm are the worst, in which there are still some or lots of respiration noise remaining in. Moreover, this experiment is independently repeated 100 times and the averaged SNRs are 25.4808 dB(GABSE), 3.2434 dB (BCBSE), 9.9630 dB (SemiBSE), 26.4522 dB (GACP) and 26.6406 dB (FastGACP) respectively. It is worthy of noting that for the proposed algorithm has the highest SNR index among the five algorithms. For the comparison of performance index (26) at $\tau = 112$, we test five algorithms in artificial ECG experiment. The performance is estimated as the averaged PI values of 100 independent trials. At every trial, five algorithms are run with 200 iterations, which seem to be always enough for convergence. Here $\mathbf{A}$ and $\mathbf{w}$ are initialized randomly. The results are depicted in Fig.3. Obviously, FastGACP algorithm performs more efficiently than the other algorithms and its convergence is the best. It can converge, on average, within 10 iterations. Note that the nonlinear functions in SemiBSE and GABSE algorithm are chosen as $\log\cosh(x)$ and x^3 respectively. In GACP and FastGACP algorithm, $G(x)$ and $F(x)$ are selected as x^2 and $\log\cosh(u)$ respectively. The learning rates of SemiBSE algorithm are set to be 0.5 ($\mu_{\mathbf{w}}$) and 0.0005 (μ_b), and those of GACP algorithm are 1 (i.e. $\mu_{\mathbf{w}} = 1, \mu_b = 1$). Moreover, in GACP and FastGACP algorithm, the parameter λ all aims to balance generalized autocorrelation and the complexity pursuit of the desired signal. $\lambda = 0.15$ is used in GACP algorithm[11]. In FastGACP, we obtain an optimal solution when $\lambda = 0.05$ by adjusting this parameter. And such a result is actually robust with the parameter perturbations, e.g. we can obtain the same corresponding result with another parameter, such as $\lambda = 0.1, 0.15, \cdots, 0.95, 1$. Fig.4 gives the comparison of PI with various values of λ for artificial ECG data. It is shown that the performance of FastGACP is robust for the majority of the parameter values between 0 and 1. Therefore, we can use a value between 0 and 1, which corresponds to the lowest PI value and in this experiment $\lambda = 0.05$ is used. Furthermore, Fig.4 also shows that FastGACP as the improvement of GACP has lower PI values, better convergence and performance than original algorithm when λ varies between 0 and 1. Note that the performance is estimated as the mean of the PI values

978-1-4673-4398-5/12/$31.00 ©2012 IEEE

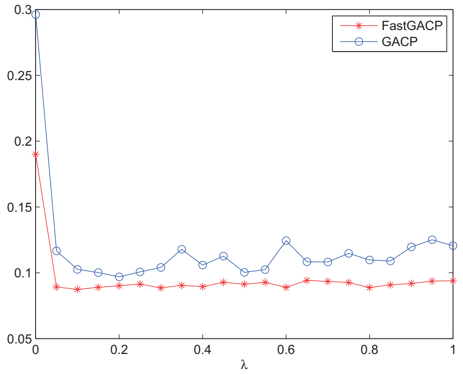

Fig. 4. Comparisons of average PIs over 100 independent runs for two algorithms with λ at $\tau = 112$ on artificial ECG data.

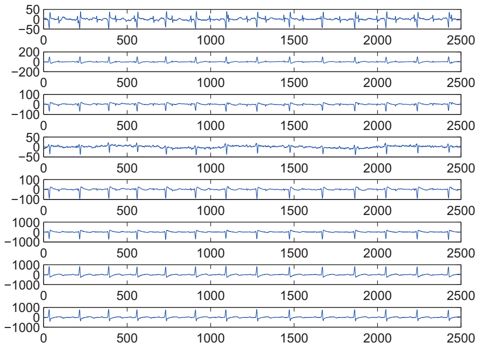

Fig. 5. The 8-channel of ECG recording obtained from a pregnant woman.

of 100 independent trials for each value of λ. In each trial, two algorithms are run with 200 iterations, which seems to be always enough for convergence.

B. Experiments on real ECG data

We have also performed experiments with real-world ECG data which is distributed by De Moor [17]. This data is a famous ECG measured from a pregnant woman (in Fig.5). One can see the heart beating of both the mother (stronger and slower) and the fetus (weaker and faster). Note that the fetal influence is stronger in the first channel of Fig.7. The ECG measurements are recorded over 10s and sampled at 250Hz (although in De Moor's homepage he claims the sampling frequency is 500Hz, Barros et al. [8] assure it is 250Hz). By using prior information about FECG frequency, we can estimate the optimal time delay $\tau = 112$. The choice of nonlinear functions, the learning rates and other parameters of five algorithms are the same as the foregoing artificial ECG simulations. Fig.6 provides the extracted FECGs by GABSE, BCBSE, SemiBSE, GACP and FastGACP algorithm, respectively. It shows that the desired FECGs are well extracted by all algorithms except BCBSE and GABSE algorithms, which include much respiration noises. It should be noticed that, since the mixing matrix $\mathbf{A}$ and the pure FECG signal are

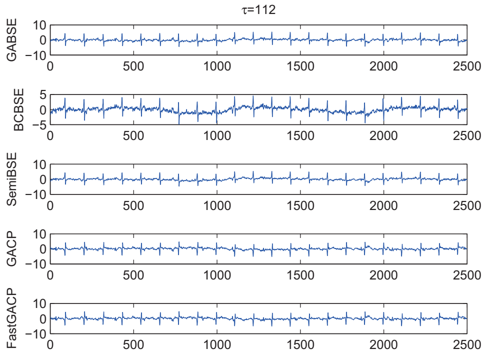

Fig. 6. Comparisons of extracted FECGs by BGABSE, BCBSE, SemiBSE, GACP and FastGACP algorithm with the optimal time delay $\tau = 112$ for real-world ECG data.

not available, the performance index, such as SNR and PI cannot be computed as above. But we can perceive distinctly the quality of extracted FECG through experience.

IV. Conclusion

In this paper, we have presented a fixed-point algorithm based on the generalized autocorrelations and complexity pursuit of desired signals. The fixed-point algorithm is obtained by using the similar lines as the FastICA algorithm. Moreover, in contrast to the gradient-based algorithms, the proposed algorithm does not need to choose any learning step sizes, which reduces the number of parameters in the GACP algorithm, and improves its convergence properties. It can converge within 10 iterations on average. It is noting that the parameter λ of FastGACP algorithm, which balances the generalized autocorrelation of the desired signal and the non-Gaussianity of its innovations, is actually robust if the parameter disturbs between 0 and 1, which does not increase the consumption of calculation of the algorithm and facilitates its application in practice. As demonstrated in this article, the proposed method is only applied to ECG data, such as artificial ECG and real world ECG data. So, as part of the future work, the applications for other type data will be studied.

Acknowledgment

The work was supported by the National Natural Science Foundation of China under the Grants 11126057, Shanghai Leading Academic Discipline Project (J50101), key Disciplines of Shanghai Municipality (Operations Research and Cybernetics, S30104), PhD start funds of University of Shanghai for Science and Technology No 1D-10-303-003, Foundation for the Youth Scholars by Educational Commission of Shanghai(No 51-12-303-105) and Grants-In-Aid for Scientific Research, Ministry of Education, Culture, Sports, Science and Technology, Japan, Project (No. 24500280).

References

[1] S. Amari and A. Cichocki, *Adaptive blind signal processing-neural network approaches*. Proc. IEEE 86 (10) (1998) 2016-2048.

[2] K. Anand, G. Mathew and V. Reddy, *Blind separation of multiple co-channel BPSK signals arriving at an antenna array*. IEEE Signal Proc. Letters 2 (9) (1995) 176-178.

[3] S. Boudet, L. Peyrodie, P. Gallois and C. Vasseur, *Filtering by optimal projection and application to automatic artifact removal from EEG*. Signal Process. 87 (8)(2007) 1978-1992.

[4] E. Chaumette, P. Comon and D. Muller, *ICA-based technique for radiating sources estimation: application to airport surveillance*. IEE Proceedings-F 140 (6)(1993) 395-401.

[5] A. Cichocki, T. Rutkowski, A. K. Barros and S. H. Oh, *A blind extraction of temporally correlated but statistically dependent acoustic signals*. In: Neural Networks for Signal Processing X: Proceedings of the 2000 IEEE Signal Processing Society Workshop (NNSP2000), IEEE Signal Processing Society, (Sidney, Australia, 2000), 455-464.

[6] K. E. Hild, H. T. Attias, S. Comani and S. S. Nagarajan, *Fetal cardiac signal extraction from magnetocardiographic data using a probabilistic algorithm*, Signal Process. 87 (8)(2007) 1993-2004.

[7] L. De Lathauwer, B. De Moor and J. Vandewalle, *Fetal electrocardiogram extraction by source subspace separation*. In: Proc. HOS'95, (Aiguablava, Spain, 1995) 134-138.

[8] A. K. Barros and A. Cichocki, *Extraction of specific signals with temporal structure*. Neural Comput. 13 (2001) 1995-2003.

[9] Z. Shi and C. Zhang, *Semi-blind source extraction for fetal electrocardiogram extraction by combining non-Gaussianity and time-correclation*. Neurocomput. 70 (2007) 1574-1581.

[10] Z. Shi and C. Zhang, *Blind source extraction using generalized autocorrelations*, IEEE Trans. Neural Networks 18 (5) (2007) 1516-1524.

[11] H. Zhang, Z. Shi and C. Guo, *Blind source extraction based on generalized autocorrelations and complexity pursuit*. Neurocomput. 72 (2009) 2556-2562.

[12] A. Hyvärinen, *Complexity pursuit: separating interesting components from time-series*. Neural Comput. 13 (4) (2001) 883-898.

[13] A. Hyvärinen and E. Oja, *A fast fixed-point algorithm for independent component analysis*. Neural comput. 9(7) (1997) 1483-1492.

[14] A. Hyvärinen, J. Karhunen and E. Oja, *Independent Component Analysis*. New York, Wiley, 2001.

[15] Z. Shi, H. Tang and Y. Tang, *A fast fixed-point algorithm for complexity pursuit*. Neurocomput. 64 (2005) 529-536.

[16] A. Hyvärinen, *Independent component analysis for time-dependent stochastic processes*. In: Proceedings of the InternationalConference on Artificial NeuralNetworks (ICANN98), Skövde, Sweden, 1998, pp. 135-140.

[17] D. De Moor(Ed.) *Daisy: database for the identification of systems*. available online at: http://www.esat.kuleuven.ac.be/sista/daisy.

978-1-4673-4398-5/12/$31.00 ©2012 IEEE