A ROBUST AND PRACTICAL METHOD TO SEPARATE PERIODIC SIGNALS FROM MEG DATA USING SECOND ORDER STATISTICS

Hidekazu Fukai
Department of Engineering, Gifu University, Yanagido 1-1, Gifu-shi, Gifu-ken, Japan

Keywords: Blind source separation, Magnetoencephalography, Periodic signals, Second-order statistics.

Abstract: The analyses of recordings of magnetoencephalography (MEG) and other imaging techniques may require the separation of periodic signals from the observed signals. Blind source separation (BSS) is widely used for the separation of specific signals these days. Though several algorithms based on the BSS scheme for the separation of periodic signals have been proposed, they usually assume the system to be well-posed, satisfactory results often cannot be obtained for practical recordings. In this study, we show that a method based on the joint approximate diagonalization of correlation matrices with several time delays (JADCM) is robust and good results can be obtained by choosing the time delays carefully, especially in practical ill-posed situations such as signal separation from MEG recordings. The performance of the proposed method is compared with that of Periodic BSS and JADCM using the conventional parameter set.

1 INTRODUCTION

The blind source separation (BSS) scheme is a powerful tool for the analysis of recordings of magnetoencepharography (MEG) and other imaging techniques to separate artifacts, noise, and brain signals. The potential of BSS for use in MEG data analysis has been confirmed in many studies (Hyvärinen et al., 2001).

Generally, we consider a linear and instantaneous superposition of independent source signals and multidimensional observations of the form

$$x(t) = A \cdot s(t)$$

for the BSS problem, where $\boldsymbol{x}(t) = (x_1(t), \cdots, x_m(t))^T$ denote the observations, $\boldsymbol{s}(t) = (s_1(t), \cdots, s_n(t))^T$ denote the unknown source signals, and the mixing matrix **A** is the transfer function between sources and sensors. BSS involves identifying **A** and retrieving the source signals $\boldsymbol{s}(t)$ without *a priori* information about the mixing matrix **A**.

Many types of BSS methods have been proposed thus far (Hyvärinen et al., 2001) (Theis and Inouye, 2006). Each of them require additional conditions for separation, so that has each suitable data types. We have to choose the methods and parameters carefully depending on the entities to be separated or extracted.

MEG data to which the BSS scheme is applied have several features such that (c1) most source signals have time structures; (c2) the number of source signals is larger than that of observations, i.e., the problem is ill-posed; and (c3) some sources are periodic.

The second-order statistics is useful for signal separation in case each source signal has specific autocorrelation and no cross-correlations. In this case, correlation matrices of source signals for arbitrary time delays are diagonal. Furthermore, if a correlation matrix for a time delay has distinct eigenvalues, we can separate the source signals by the diagonalization of the correlation matrix (Belouchrani et al., 1997) (Ziehe and Müller, 1998). Several methods for the separation of periodic signals have been proposed (Barros and Cichocki, 2001) (Jafari et al., 2006).

Joint diagonalizations of several correlation matrices with different time delays are effective for the case of correlation matrices having the same or almost the same eigenvalues (Belouchrani et al., 1997) (Ziehe and Müller, 1998). Several methods based on the joint approximated diagonalization have been proposed (Theis and Inouye, 2006) and applied to MEG data analysis (Ziehe et al., 2000).

In the joint diagonalization process, we have to choose the values and the total number of time delays for the correlation matrices. If the problem is well-posed, as assumed in most proposed methods, the selection of the time delays does not critically affect the results of the separation. On the other hand, in the case of practical application to MEG data, the number

Fukai H..
A ROBUST AND PRACTICAL METHOD TO SEPARATE PERIODIC SIGNALS FROM MEG DATA USING SECOND ORDER STATISTICS.
DOI: 10.5220/0003296103720377
In *Proceedings of the International Conference on Bio-inspired Systems and Signal Processing* (BIOSIGNALS-2011), pages 372-377
ISBN: 978-989-8425-35-5

of source signals is larger than that of observations, so that the problem is ill-posed (c2). In this case, the information has already been lost during observation, and generally, we can only get approximated source signals from any decomposing process. There is no matrix that diagonalizes all the correlation matrices simultaneously. The approximated results vary considerably depending on the selection of the values of time delays, but very few reports have discussed the selection of the values of time delays (Tang et al., 2005).

In order to observe and separate specific signals, we can use the features of the signals for the selection of parameters of the method. In this paper, we propose practical and effective values and numbers of time delays for the extraction of periodic signals (c3) from MEG data.

In this paper, we treat 60 Hz power supply noise as an example of separate and eliminate periodic signals from MEG data. The BSS-type method we use is based on the approximately joint diagonalization of several time-delayed correlation matrices without noise term (Ziehe and Müller, 1998). We use the Jacobi-like joint diagonalization method (Cardoso and Souloumiac, 1996). The performance of the proposed method is compared with that of conventional usage of JADCM and Periodic BSS (Jafari et al., 2006), which is a sequential algorithm based on second-order statistical information.

2 METHODS

2.1 Experimental Procedure

MEG data were obtained by using a whole-head helmet-shaped 64-channel SQUID sensor array (Model 100, CTF Systems Inc.) in a magnetically shielded room. Each MEG sensor has an axial gradiometer that approximately measures the magnetic field in the normal direction. One healthy volunteer (32-year-old male) participated this study. After explaining the nature of the study, informed consent was obtained from him. The experimental procedures were in accordance with the Declaration of Helsinki. The subject had no history of neurological and psychiatric disorders. During recording, the subject was seated relaxed under a helmet-shaped device. The sampling ratio was 625 Hz.

2.2 Blind Separation using Second-order Statistics

The problem of blind source separation (BSS) consists of recovering unknown source signals $\boldsymbol{s}(t) = (s_1(t), \ldots, s_n(t))^T$ from observations $\boldsymbol{x}(t) = (x_1(t), \ldots, x_m(t))^T$, generated by the mixing model

$$\boldsymbol{x}(t) = \mathbf{A}\boldsymbol{s}(t)$$

where the mixing operator $\mathbf{A} \in \mathbb{R}^{m\times n}$ is a fixed but unknown matrix and t is discrete time index (Ziehe and Müller, 1998).

Generally, we suppose well-posed problem, i.e., the matrix $\mathbf{A}$ has full column rank. In the case where the number of sources is equal to the number of measurements $m = n$, the model becomes exactly determined problem. The separating system then estimates the original source signals according to

$$\boldsymbol{y}(t) = \mathbf{B}\boldsymbol{x}(t) = \mathbf{B}\mathbf{A}\boldsymbol{s}(t) = \Lambda\mathbf{P}\boldsymbol{s}(t) \propto \boldsymbol{s}(t),$$

where $\boldsymbol{y}(t) = (y_1(t), \ldots, y_m(t))^T$ represents the recovered sources. Matrices $\mathbf{B}$, Λ and $\mathbf{P} \in \mathbb{R}^{m\times m}$ are separating, scaling and permutation matrix respectively. Because we cannot know the variances of sources, we assume the correlation matrix of $\boldsymbol{s}(t)$ as unit matrix and each $\boldsymbol{s}$(t) has mean zero without losing generality. In the joint diagonalization scheme, we consider correlation matrices $\mathbf{C}_x(\tau)$ of time-lagged mixed signals $\boldsymbol{x}(t)$,

$$\mathbf{C}_x(\tau) \stackrel{\text{def}}{=} E\{\boldsymbol{x}(t)\boldsymbol{x}^T(t+\tau)\}$$

where $E\{\cdot\}$ denotes statistical expectation over t and τ is a time-shift parameter. Practically the expectation is computed from the available data with a finite sample size N as a sample average

$$\hat{\mathbf{C}}_x(\tau_k) = \frac{1}{N}\sum_{t=1}^{N-\tau_k}(\boldsymbol{x}(t)\boldsymbol{x}^T(t+\tau_k)).$$

We see that the correlation of $\boldsymbol{x}(t)$ is related to the correlation of $\boldsymbol{s}(t)$ according to

$$\begin{aligned}\mathbf{C}_x(\tau) &= E\{(\mathbf{A}\boldsymbol{s}(t))(\mathbf{A}\boldsymbol{s}(t+\tau))^T\} \\ &= \mathbf{A}E\{\boldsymbol{s}(t)\boldsymbol{s}(t+\tau)^T\}\mathbf{A}^T \\ &= \mathbf{A}\mathbf{C}_s(\tau)\mathbf{A}^T\end{aligned}$$

due to the linearity of the expectation operator and the mixing model. Here, all cross-correlation terms of s which are the off-diagonal elements of $\mathbf{C}_s(\tau)$ are zero for independent signals and thus $\mathbf{C}_s(\tau)$ is a diagonal matrix. If $\mathbf{A}$ is invertible, this can also be written as

$$\mathbf{B}\mathbf{C}_x(\tau)\mathbf{B}^T = \mathbf{C}_s(\tau)$$

where the matrix $\mathbf{B} = \mathbf{A}^{-1}$ diagonalizes $\mathbf{C}_x(\tau)$.

We take two steps to identify **B**, i.e., *whitening step* and *joint-diagonalization step*.

The whitening step is achieved by applying to $\boldsymbol{x}(t)$ a whitening matrix **W** satisfying:

$$E\{\mathbf{W}\boldsymbol{x}(t)\boldsymbol{x}(t)^T\mathbf{W}^T\} = \mathbf{W}\mathbf{C}_x(0)\mathbf{W}^T = \mathbf{W}\mathbf{A}\mathbf{A}^T\mathbf{W}^T = \mathbf{I}.$$

In other words, whitening step correspond to principal component analysis (PCA) procedure which removes correlations between the observations $\boldsymbol{x}(t)$. This follows that if **W** is a whitening matrix, then there exists a orthogonal matrix **U** such that $\mathbf{WA} = \mathbf{U}$ for any whitening matrix **W**. With the orthogonal matrix **U**, the mixture matrix **A** can be determined as $\mathbf{A} = \mathbf{W}^{-1}\mathbf{U}$. The whitening step reduces the determination of the matrix **A** to that of a orthogonal matrix **U**.

The next joint-diagonalization step is achieved by finding **U** as a rotetion matrix which diagonalize the time-delayed correlation matrices $E\{\mathbf{W}\boldsymbol{x}(t)\boldsymbol{x}(t+\tau_k)^T\mathbf{W}^T\}$ simultaneously. Here, any rotating procedure does not change the correlation matrix of zero time delay.

Though only one additional time delayed correlation matrix is required theoretically (Tong et al., 1991)(Molgedey and Schuster, 1994), we may fail to find the diagonalization matrix **B** if $\mathbf{C}_x(\tau)$ does not have distinct eigenvalues. This can be avoided by simultaneous diagonalization of multiple $\mathbf{C}_x(\tau_k), 1 \leq k \leq K$ (Belouchrani et al., 1997)(Ziehe and Müller, 1998). We use Jacobi like algorithm proposed by Cardoso and Souloumiac (Cardoso and Souloumiac, 1996) based on several Givens rotations to find **U**.

2.3 Selection of Time Delays for Periodic Sources on Ill-posed Problem

In most theoretical discussions on the BSS problem, we treat exactly determined problem in which we can get exact solutions. In the joint-diagonalization scheme, not much focus on the selection of time delays is required if the problems are well-posed. Theoretically, in case each source signal $\boldsymbol{s}(t)$ has its auto-correlations and the correlation matrix for the time delay has distinct eigenvalues, only one time delay is required. Several time delays and their joint-diagonalizations are effective in case the correlation matrices have the same or almost the same eigenvalues (Belouchrani et al., 1997)(Ziehe and Müller, 1998).

However, the problems are ill-posed in most practical cases, especially in MEG recordings, and only approximated solutions can be obtained. There is no orthogonal matrix that makes all the correlation matrices with the time delay diagonal on the practical ill-posed situations. In this case, the joint-diagonalization algorithm minimizes the sum of the off-diagonal values to obtain the target orthogonal matrix.

Generally, in ill-posed problems, the separation performance depends on the trade-off between the source signals. In the case of joint-diagonalization, the results vary depending on the selection of the set of the time delay τ. If specific sources have to be extracted, we have to set the number and value of time delays depending on the condition and achive good separation performance.

In case the target signals have specific features in the auto-correlation function, we can select the values of time delays depending on that. If the signals are periodic with cycle T, the auto-correlation function has peaks on the time delay $kT, k = 1, 2, \cdots$ and we can get larger evaluation values to minimize the peaks by setting $\tau = kT, k = 1, 2, \cdots$.

Furthermore, if brain signals are not the target to be separated, it is important to choose the time delays to avoid the effect of brain signals. Because almost all brain signals have prominent auto-correlation values within 1 s of the time delay, it is effective to choose time delays over 1 s to avoid the effect of brain signals.

In conclusion, for the separation and elimination of power supply noise of 60 Hz, we set the time delay as $\tau = kT, k = 61, 62, \cdots$, and $T = 1/60$ (s). The total number of τ may be chosen such that it is sufficient for convergence.

3 RESULTS

As an example of the separation of periodic signals from MEG data, we tried to separate 60 Hz power supply noise and eliminate them from the data.

Figure 1 shows auto-correlation functions of separated signals (SS). We selected five signals from separated 64 signals, labeled SS1 to SS4, which have remarkable peaks showing time delays associated with 60 Hz cycle and its higher harmonics on auto-correlation functions. The peak values become constant over 0.05 s, which approximately indicates the noise originated power supply has little temporal fluctuations. On the other hand, the signals originate from the brain activities have temporal fluctuations, so that the values of auto-correlations decrease after a short period (SS9). This difference is a useful factor for the separation. The time delay parameter used here was $\tau = kT,\ k = 61, 62, \cdots, 80,\ T = 1/60$ (s) for 60 Hz power supply noise. This parameter set was chosen

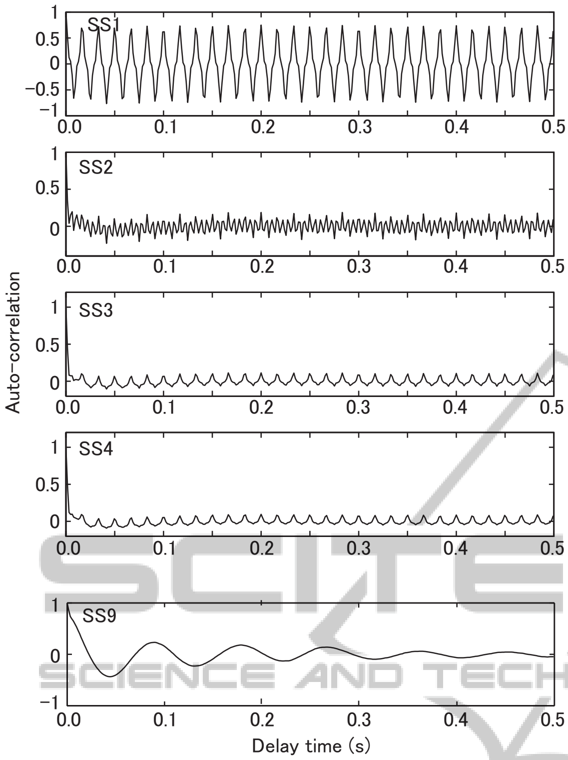

Figure 1: Auto-correlation functions of separated signals related to power supply noise (SS1–SS4) and brain signal (SS9).

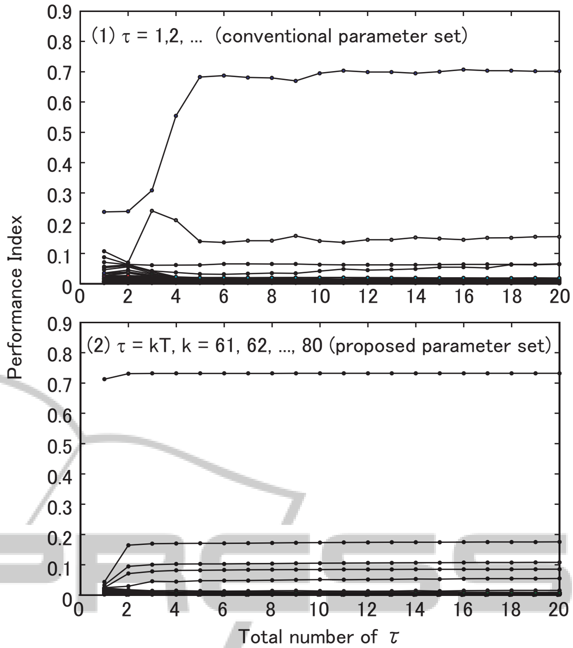

Figure 2: Performance Indexes of separation improved as total number of correlation matrix to diagonalize increase.

to utilize peaks of auto-correlation functions of target signals and circumvent the effect of auto-correlation values of signals originating from brain activities.

Figure 2 shows the performance indexes of separation to estimate a sufficient number of correlation matrices. The abscissas are total number of correlation matrices, used for joint-diagonalization, the number of τ. The performance index is defined as average of auto-correlation values of each separated signal $y(t)$ on the peaks correspond to 60Hz. The index values are calculated as

$$\frac{1}{K_2 - K_1} \sum_{k=K_1}^{K_2} \frac{1}{N} \sum_{t=1}^{N-kT} (y_j(t) y_j(t+kT)), \ \ j = 1, \cdots, 64,$$

where (K_1, K_2) was set as $(61, 80)$ after auto-correlation functions originating from other biological signals decrease. All index values calculated for 64 components are plotted.

Using conventional τ set, $\tau = 1, 2, \cdots$, we need over 10 correlation matrices until we get converged results (Fig.2 (1)). On the other hand, the index values converge at the total number of τ is about 3 or 4 with proposed τ set, $\tau = kT, k = 61, 62, \cdots, 80$. This smaller total number of correlation matrices for joint-diagonalization has an advantage in calculation time. Furthermore, the score of the performance index after convergence is also better on proposed τ set (Fig.2 (2)).

The separation performance is also compared with Periodic BSS (Jafari et al., 2006). Figure 3 shows the performance indexes of separated signals obtained by Periodic BSS (dashed), by JADCM with a conventional parameter set, $\tau = 1, 2, \cdots$ (dotted), and by JADCM with proposed parameter set $\tau = kT, k = 61, 62, \cdots, 80$ (solid). The separated signals are sorted according to the performance index, and on 10 from 64 signals are shown. In the case of an ill-posed system, the proposed method shows good performance.

To evaluate the noise reduction performance of the proposed method, we reconstructed the original MEG sensor signals from output signals after rejecting four components that were related to power supply noise corresponding to SS1 to SS4 (Fig.1). Figure 3 (1) shows the power spectral density function of the original row data from MEG channel L42. Prominent peaks can be confirmed at frequencies close to those of the power supply and also at higher harmonics. Even after noise reduction using Periodic BSS (Fig.3 (2)) nor JADCM with the ordinary parameter set, $\tau = 1, \cdots, 20$, (Fig.3 (3)) the peaks are still observed. After noise reduction using the proposed parameter set, $\tau = kT, k = 61, 62, \cdots, 80$, peaks related to power supply noise are hardly found.

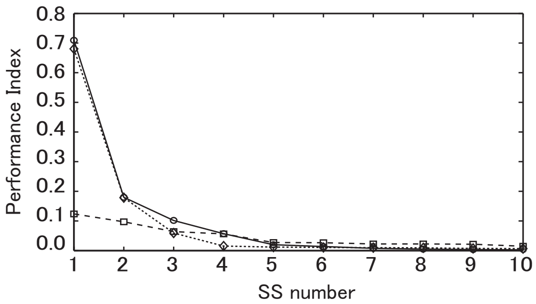

Figure 3: Performance indexes of SSs separated by Periodic BSS (dashed), JADCM with conventional parameter set $\tau = 1, 2, \cdots, 20$ (dotted), and by JADCM with proposed parameter set $\tau = kT, k = 61, 62, \cdots, 80$ (solid). The SSs are sorted according to the performance indexes.

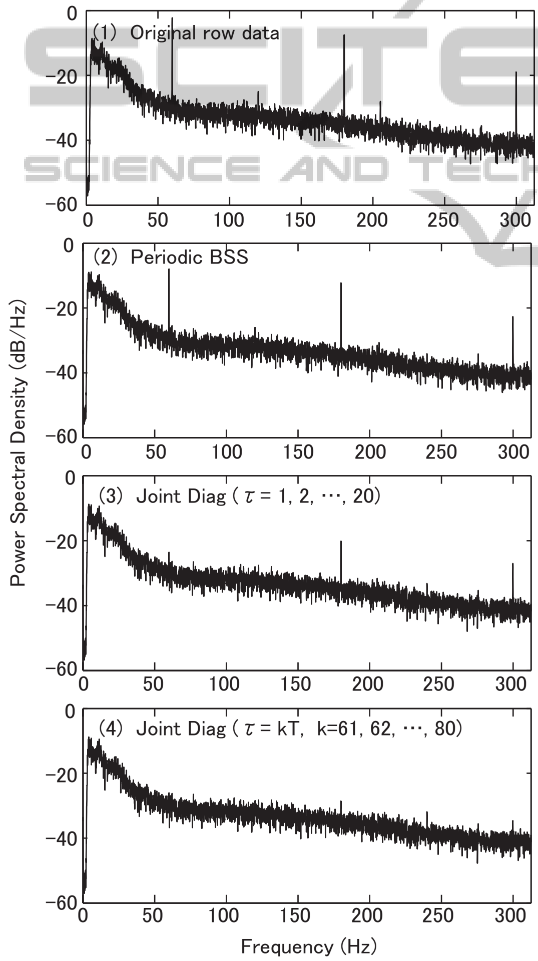

Figure 4: Power spectral density of MEG channel L42. (1) original row data, (2) after noise reduction using Periodic BSS, (3) after noise reduction using JADCM with conventional parameter set $\tau = 1, 2, \cdots, 20$, and (4) after noise reduction using JADCM with proposed parameter set $\tau = kT, k = 61, 62, \cdots, 80$.

4 CONCLUSIONS

Several methods to separate periodic signals from multichannel recordings have been proposed thus far. However, they usually assume the problem to be exactly determinable as a result of which their performance in practical signal processings is not good.

In this report, we proposed a robust and practical to those method to extract periodic signals of an ill-posed system. This method is based on the jointdiagonalization of several time-delayed correlation matrices. We showed the importance of selecting time delay parameters for the extraction of specific signals. We also verified the efficiency of the proposed method in reducing power supply noise in MEG recordings. We compared results with those of Periodic BSS method and JADCM with the conventional parameter set.

Because the power supply noise is artificial, it has little fluctuations in its period and auto-correlation functions are constant even for long time delays. The same feature is also found in periodic brain responses, indicating that the method is effective.

REFERENCES

Barros, A. and Cichocki, A. (2001). Extraction of specific signals with temporal structure. *Neural Computation*, 13:1995–2003.

Belouchrani, A., Abed-Meraim, K., Cardoso, J., and Moulines, E. (1997). A blind source separation technique using second-order statistics. *IEEE Trans. Signal Processing*, 45:434–443.

Cardoso, J. and Souloumiac, A. (1996). Jacobi angles for simultaneous diagonalization. *SIAM J. Math. Anal. Appl.*, 17:161–164.

Hyvärinen, A., Karhunen, J., and Oja, E. E. (2001). *Independent Component Analysis*. John Wiley & Sons Inc., New York.

Jafari, M., Wang, W., Chambers, J., Hoya, T., and Cichocki, A. (2006). Sequential blind source separation based exclusively on second-order statistics developed for a class of periodic signals. *IEEE Trans. Signal Process.*, 54(3):1028–1040.

Molgedey, L. and Schuster, H. (1994). Separation of a mixture of independent signals using time delayed correlations. *Phys. Rev. Let.*, 72:3634–3637.

Tang, A., Liu, J., and Sutherland, M. (2005). Recovery of correlated neuronal sources from eeg: The good and bad ways of using sobi. *NeuroImage*, 28:507–519.

Theis, F. and Inouye, Y. (2006). On the use of joint diagonalization in blind signal processing. In *Proc. ISCAS*, pages 3586–3589.

Tong, L., Liu, R., Soon, C., and Huang, Y. (1991). Indeterminacy and identifiability of blind identification. *IEEE Trans. Circuit and Sys.*, 38:499–509.

Ziehe, A. and Müller, K. (1998). Tdsep-and efficient algorithm for blind separation using time structure. In *Proc. Int. Conf. on Artificial Neural Networks (ICANN'98)*, pages 675–680.

Ziehe, A., Müller, K., Nolte, G., Mackert, B., and Curio, G. (2000). Artifact reduction in magnetoneurography based on time-delayed second-order correlations. *IEEE Trans. Biomed. Eng.*, 47:75–87.