BLIND SEPARATION OF SECOND-ORDER NONSTATIONARY AND TEMPORALLY COLORED SOURCES

Seungjin Choi [§†], *Andrzej Cichocki* [†], *Adel Belouchrani* [‡]

[§] Department of Computer Science & Engineering, POSTECH, Korea
seungjin@postech.ac.kr
[†] Laboratory for Advanced Brain Signal Processing, Brain Science Institute, RIKEN, Japan
cia@brain.riken.go.jp
[‡] Department of Electrical Engineering, Ecole Nationale Polytechnique, Algeria
belouchrani@hotmail.com

ABSTRACT

This paper presents a method of blind source separation that jointly exploits the nonstationarity and temporal structure of sources . The method needs only multiple time-delayed correlation matrices of the observation data, each of which is evaluated at different time-windowed data frame, to estimate the demixing matrix. We show that the method is quite robust with respect to the spatially correlated but temporally white noise. We also discuss the extension of some existing second-order blind source separation methods. Extensive numerical experiments confirm the validity of the proposed method.

1. INTRODUCTION

Blind source separation (BSS) is a fundamental problem that is encountered in many practical applications such as telecommunications, image/speech processing, and biomedical signal analysis where multiple sensors are involved. In its simplest form, the m-dimensional observation vector $\boldsymbol{x}(t) \in \mathbb{R}^m$ is assumed to be generated by

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{v}(t), \tag{1}$$

where $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is the unknown mixing matrix, $\boldsymbol{s}(t)$ is the n-dimensional source vector (which is also unknown and $n \leq m$), and $\boldsymbol{v}(t)$ is the additive noise vector that is statistically independent of $\boldsymbol{s}(t)$.

A variety of methods/algorithms for BSS have been developed for last decade (for example, see [1] and references therein). Although many different BSS algorithms are available, their principles can be categorized by three distinctive methods which are based on (1) the non-Gaussianity of source [2], (2) the temporal structure of source [3], and (3) the nonstationarity of source [4].

In this paper we present methods that jointly exploits the nonstationarity and temporal structure of sources to estimate the mixing matrix (or the demixing matrix) in the presence of spatially correlated but temporally white noise (not necessarily Gaussian). Thus our methods works even for the case where multiple Gaussian sources with no temporal correlations exit as long as their variances are slowly time-varying. Moreover, we show that if we use just time-delayed correlations of the observation data, we can find a robust estimate of the demixing matrix. To this end, we introduce a method of robust whitening and present the Second-Order Nonstationary source Separation (SEONS) method. We also present the extension of some existing second-order BSS methods which are (1) the extended matrix pencil method and (2) the extended Pham-Cardoso method.

Throughout this paper, the following assumptions are made:

(AS1) The mixing matrix $\boldsymbol{A}$ is of full column rank.

(AS2) Sources are spatially uncorrelated but are temporally correlated (colored) stochastic signals with zero mean.

(AS3) Sources are second-order nonstationary signals in the sense that their variances are time varying.

(AS4) Additive noises $\{v_i(t)\}$ are spatially correlated but temporally white, i.e.,

$$E\{\boldsymbol{v}(t)\boldsymbol{v}^T(t-\tau)\} = \delta_\tau \boldsymbol{\Gamma}, \tag{2}$$

where δ_τ is the Kronecker symbol and $\boldsymbol{\Gamma}$ is an arbitrary $m \times m$ matrix.

2. ROBUST WHITENING

The whitening (or data sphering) is an important pre-processing step in a variety of BSS methods. The conventional whitening exploits the equal-time correlation matrix of the data $\boldsymbol{x}(t)$, so that the effect of additive noise can not be removed. The idea of robust whitening lies in utilizing the time-delayed correlation matrices that are not sensitive to the additive white noise. The robust whitening method is explained for the case of stationary signals.

It follows from the assumptions (AS2) and (AS4) that the time-delayed correlation matrix of the observation data $\boldsymbol{x}(t)$ has the form

$$\begin{aligned} \boldsymbol{R}_x(\tau) &= E\{\boldsymbol{x}(t)\boldsymbol{x}^T(t-\tau)\} \\ &= \boldsymbol{A}\boldsymbol{R}_s(\tau)\boldsymbol{A}^T, \end{aligned} \tag{3}$$

for $\tau \neq 0$. One can easily see that the transformation $\boldsymbol{R}_x^{-\frac{1}{2}}(\tau)$ whiten the data $\boldsymbol{x}(t)$ without the effect of the noise vector $\boldsymbol{v}(t)$. It reduces the noise effect and project the data onto the signal subspace, in contrast to the conventional whitening transformation $\boldsymbol{R}_x^{-\frac{1}{2}}(0)$. Some source separation methods already employ this robust whitening transformation [5, 6, 7, 8].

In general, however, the matrix $\boldsymbol{R}_x(\tau)$ is not always positive definite, so the whitening transformation $\boldsymbol{R}_x^{-\frac{1}{2}}(\tau)$ may not be valid for some time-lag τ. The idea of the robust whitening is to consider a linear combination of several time-delayed correlation matrices, i.e.,

$$\boldsymbol{C}_x = \sum_{i=1}^{K} \alpha_i \boldsymbol{M}_x(\tau_i), \tag{4}$$

where

$$\boldsymbol{M}_x(\tau_i) = \frac{1}{2}\left\{\boldsymbol{R}_x(\tau_i) + \boldsymbol{R}_x^T(\tau_i)\right\}. \tag{5}$$

A proper choice of $\{\alpha_i\}$ may result in a positive definite matrix $\boldsymbol{C}_x$. For example, the FSGC method [9] can be used to find a set of coefficients $\{\alpha_i\}$ such that the matrix $\boldsymbol{C}_x$ is positive definite.

The matrix $\boldsymbol{C}_x$ has the eigen-decomposition

$$\boldsymbol{C} = [\boldsymbol{U}_1, \boldsymbol{U}_2] \begin{bmatrix} \boldsymbol{D}_1 & \\ & \boldsymbol{0} \end{bmatrix} [\boldsymbol{U}_1, \boldsymbol{U}_2]^T, \tag{6}$$

where $\boldsymbol{U}_1 \in \mathbb{R}^{m\times n}$ and $\boldsymbol{D}_1 \in \mathbb{R}^{n\times n}$. Then the robust whitening transformation matrix is given by $\boldsymbol{Q} = \boldsymbol{D}_1^{-\frac{1}{2}}\boldsymbol{U}_1^T$. The transformation $\boldsymbol{Q}$ project the data onto n-dimensional signal subspace as well as whitening.

Let us denote the whitened n-dimensional data by $\boldsymbol{z}(t)$

$$\begin{aligned} \boldsymbol{z}(t) &= \boldsymbol{Q}\boldsymbol{x}(t) \\ &= \boldsymbol{B}\boldsymbol{s}(t) + \boldsymbol{Q}\boldsymbol{v}(t), \end{aligned} \tag{7}$$

where $\boldsymbol{B} \in \mathbb{R}^{n\times n}$. The whitened data $\boldsymbol{z}(t)$ is a unitary mixture of sources with additive noise, i.e., $\boldsymbol{B}\boldsymbol{B}^T = \boldsymbol{I}$.

3. SECOND-ORDER NONSTATIONARY SOURCE SEPARATION

This section describes our main method, SEONS, as well as some extensions such as the extended matrix pencil method and the extended Pham-Cardoso method.

Now we consider the case where sources are second-order nonstationary and have non-vanishing temporal correlations. It follows from the assumptions (AS1)-(AS4) that we have

$$\boldsymbol{M}_x(t_k, \tau_i) = \boldsymbol{A}\boldsymbol{M}_s(t_k, \tau_i)\boldsymbol{A}^T, \tag{8}$$

for $\tau_i \neq 0$ and the index t_k is for time since we deal with nonstationary sources. In practice $\boldsymbol{M}_x(t_k, \tau_i)$ is computed using the samples in the kth time-windowed data frame, i.e.,

$$\boldsymbol{R}_x(t_k, \tau_i) = \frac{1}{N_k}\sum_{t\in\mathcal{N}_k} \boldsymbol{x}(t)\boldsymbol{x}^T(t-\tau_i), \tag{9}$$

$$\boldsymbol{M}_x(t_k, \tau_i) = \frac{1}{2}\left\{\boldsymbol{R}_x(t_k, \tau_i) + \boldsymbol{R}_x^T(t_k, \tau_i)\right\}, \tag{10}$$

where $\mathcal{N}_k$ contains the data points in the kth time-windowed frame and N_k is the number of data points in $\mathcal{N}_k$.

The matrix pencil method [4] was applied to the blind separation of temporally colored sources. In general, however, the pencil that consists of two time-delayed correlation matrices is not symmetric definite pencil, which may cause some numerical problems in calculating generalized eigenvectors. The extended matrix pencil method (which is described below) employs a symmetric definite pencil.

Algorithm Outline: Extended Matrix Pencil Method (nonstationary case)

1. We partition the observation data into two non-overlapping blocks, $\{\mathcal{N}_1, \mathcal{N}_2\}$.
2. Compute $\boldsymbol{M}_x(t_2, \tau_2)$ for some time-lag $\tau_2 \neq 0$ using the data points in $\mathcal{N}_2$.
3. Calculate the matrix $\boldsymbol{C}_1(t_1) = \sum_{i=1}^{J} \alpha_i \boldsymbol{M}_x(t_1, \tau_i)$ by the FSGC method using the data points in $\mathcal{N}_1$.
4. Find the generalized eigenvector matrix $\boldsymbol{V}$ of the pencil $\boldsymbol{M}_x(t_2, \tau_2) - \lambda \boldsymbol{C}_1(t_1)$ which satisfies
$$\boldsymbol{M}_x(t_2, \tau_2)\boldsymbol{V} = \boldsymbol{C}_1(t_1)\boldsymbol{V}\boldsymbol{\Lambda}. \tag{11}$$
5. The demixing matrix is given by $\boldsymbol{W} = \boldsymbol{V}^T$.

In order to improve the statistical efficiency, we can employ a joint approximate diagonalization method [10], as in the JADE [11] and SOBI [3]. The joint approximate diagonalization method in [10] finds an unitary transformation that jointly diagonalizes several matrices (which do not have to be symmetric nor positive definite). The method SEONS is based on this joint approximate diagonalization. In this sense the SEONS includes the SOBI as its special case (if sources are stationary). The algorithm is summarized below.

Algorithm Outline: SEONS

1. The robust whitening method (described in Section 2) is applied to obtain the whitened vector $\boldsymbol{z}(t) = \boldsymbol{Q}\boldsymbol{x}(t)$. In the robust whitening step, we used the whole available data points.
2. Divide the whitened data $\{\boldsymbol{z}(t)\}$ into K non-overlapping blocks and calculate $\boldsymbol{M}_z(t_k, \tau_j)$ for $k = 1, \ldots, K$ and $j = 1, \ldots, J$. In other words, at each time-windowed data frame, we compute J different time-delayed correlation matrices of $\boldsymbol{z}(t)$.
3. Find a unitary joint diagonalizer $\boldsymbol{V}$ of $\{\boldsymbol{M}_z(t_k, \tau_j)\}$ using the joint approximate diagonalization method in [10], which satisfies
$$\boldsymbol{V}^T\boldsymbol{M}_z(t_k, \tau_j)\boldsymbol{V} = \boldsymbol{\Lambda}_{k,j}, \tag{12}$$
where $\{\boldsymbol{\Lambda}_{k,j}\}$ is a set of diagonal matrices.
4 The demixing matrix is computed as $\boldsymbol{W} = \boldsymbol{V}^T\boldsymbol{Q}$.

Recently Pham [12] developed a joint approximate diagonalization method where non-unitary joint diagonalizer of several Hermitian positive matrices is computed by a way similar to the classical Jacobi method. Second-order nonstationarity was also exploited in [13], but only noise-free data was considered. In order to extend the Pham-Cardoso algorithm into the case of noisy data, we employ a linear combination of multiple time-delayed correlation matrices which is ensured to be positive definite, at each data block. The method is referred to as the extended Pham-Cardoso (which is summarized below). One advantage of the extended Pham-Cardoso is that it does not require the whitening step because the joint approximate diagonalization method in [13] finds a non-unitary joint diagonalizer. However, its drawback lies in the fact that it requires the set of matrices to be Hermitian and positive definite, so we need to find a linear combination of time-delayed correlation matrices that is positive definite at each data frame, which increase the computational complexity.

Algorithm Outline: Extended Pham-Cardoso

1. Divide the data $\{\boldsymbol{x}(t)\}$ into K non-overlapping blocks and calculate $\boldsymbol{M}_x(t_k, \tau_j)$ for $k = 1, \ldots, K$ and $j = 1, \ldots, J$.
2. At each data frame, we compute

$$\boldsymbol{C}_k = \sum_{i=1}^{J} \alpha_i^{(k)} \boldsymbol{M}_x(t_k, \tau_i) \tag{13}$$

   by the FSGC method for $k = 1, \ldots, K$. Note that $\{\boldsymbol{C}_k\}$ is symmetric and positive definite.

3. Find a non-unitary joint diagonalizer $\boldsymbol{V}$ of $\{\boldsymbol{C}_k\}$ using the joint approximate diagonalization method in [12], which satisfies

$$\boldsymbol{V}\boldsymbol{C}_k\boldsymbol{V}^T = \boldsymbol{\Lambda}_k, \tag{14}$$

   where $\{\boldsymbol{\Lambda}_k\}$ is a set of diagonal matrices.

4. The demixing matrix is computed as $\boldsymbol{W} = \boldsymbol{V}$.

4. NUMERICAL EXPERIMENTS

Several numerical experimental results are presented to evaluate the performance of our method (SEONS) and to compare it with some existing methods such as JADE [11], SOBI [3], matrix pencil methods [4], and Pham-Cardoso [13]. Through numerical experiments, we confirm the useful behavior of the proposed method, SEONS, in two cases: (1) the case where several nonstationary Gaussian sources exist and each Gaussian source has no temporal correlation; (2) the case where additive noises are spatially correlated but temporally white Gaussian processes.

In order to measure the performance of algorithms, we use the performance index (PI) defined by

$$\begin{aligned}\mathrm{PI} \;=\; & \frac{1}{n(n-1)} \sum_{i=1}^{n} \left\{ \left(\sum_{k=1}^{n} \frac{|g_{ik}|}{\max_j |g_{ij}|} - 1 \right) \right. \\ & \left. + \left(\sum_{k=1}^{n} \frac{|g_{ki}|}{\max_j |g_{ji}|} - 1 \right) \right\}, \end{aligned} \tag{15}$$

where g_{ij} is the (i, j)-element of the global system matrix $\boldsymbol{G} = \boldsymbol{W}\boldsymbol{A}$ and $\max_j g_{ij}$ represents the maximum value among the elements in the ith row vector of $\boldsymbol{G}$, $\max_j g_{ji}$ does the maximum value among the elements in the ith column vector of $\boldsymbol{G}$. When the perfect separation is achieved, the performance index is zero. In practice, the value of performance index around 10^{-3} gives quite a good performance.

4.1. Experiment 1

The first experiment was designed to evaluate the effectiveness of the proposed method in the presence of several Gaussian signals. In this experiment, we used three speech signals that are sampled at 8 kHz and two Gaussian signals (with no temporal correlations) whose variances are slowly varying. These 5 sources were mixed using a randomly generated 5×5 mixing matrix to generate 5-dimensional observation vector with 10000 data points. No measurement noise was added.

In this experiment, we compared the SEONS with JADE, SOBI, and Pham-Cardoso [13]. It is expected that the performance of JADE and SOBI is degraded because of the presence of two white Gaussian sources. The result is shown in Fig. 1 in which the Hinton diagram of the global system matrix $\boldsymbol{G}$ is plotted. In Hinton diagram, each square's area represents the magnitude of the element of the matrix and each square's color represents the sign of the element (red for negative value and green for positive value). For successful separation, each row and column has only one dominant square (regardless of its color). Small squares contribute performance degradation. One can observe that SEONS and Pham-Cardoso work well even in the presence of nonstationary Gaussian sources (see (a) and (b) in Fig. 1), compared to JADE and SOBI (see (c) and (d) in Fig. 1). For the case of JADE, the first and last row of $\boldsymbol{G}$ has a relatively big square besides the dominant square, which verifies that the two white Gaussian sources are difficult to be separated out. The SOBI gives slightly better performance than JADE, but its performance is not comparable to SEONS (see the first and fourth row of $\boldsymbol{G}$, (d) in Fig. 1).

The following parameters were used in this experiment:

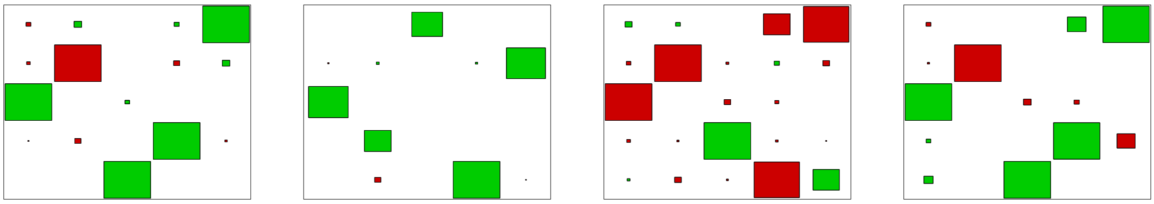
- In SEONS and Pham-Cardoso, we partitioned the whole data (10000 data points) into 100 different frames of data (each frame contains 100 data points) to calculate 100 different equal-time correlation matrices. These matrices were used to estimate the demixing matrix.
- In SOBI, we used 20 different time-delayed correlation matrices to estimate the demixing matrix.

4.2. Experiment 2

The second experiment was designed to show the robustness of the SEONS in the presence of spatially correlated but temporally white noise. We used 3 digitized voice signals and 2 music signals, all of which were sampled at 8 kHz. The mixing matrix $\boldsymbol{A} \in \mathbb{R}^{5\times 5}$, all the elements of which were drawn from standardized Gaussian distribution (i.e., zero mean and unit variance). As in the experiment 1, the whole data has 10000 samples.

The algorithms that are tested in this experiment, include the extended matrix pencil method (Extended MP), SEONS, extended Pham-Cardoso, JADE, SOBI, and SOBI with robust whitening method [8] (see Fig. 2). In SEONS, we partitioned the data into 50 no overlapping blocks (each frame has 200 data points). The robust whitening was performed using a combination of 5 time-delayed correlation matrices (with time-lags $\{1, 2, \ldots, 5\}$). In each data frame, we computed 5 time-delayed correlation matrices. The joint approximate diagonalizer of 250 correlation matrices (5 of each blocks = 5×50) was computed to estimate the demixing matrix.

At high SNR, most of algorithms worked very well, except for the extended MP method since it uses only two matrices. At low SNR, one can observe that the SOBI with robust whitening outperforms the SOBI without whitening. The SEONS gives slightly better performance than the SOBI with robust whitening in most of ranges of SNR. In the range between 0 and 6 dB, the SEONS is worse than the SOBI with robust whitening. It might result from the fact that the SEONS takes only 200 data points to calculate the time-delayed correlation matrices, so the temporal whiteness of the noise vector is not really satisfied. One can use less number of blocks (so more data points for each block) to reduce this drawback. The advantage of SEONS over SOBI with robust whitening



(a) (b) (c) (d)

Figure 1: Hinton diagrams of global system matrices: (a) SEONS; (b) Pham-Cardoso; (c) JADE; (d) SOBI with PI .001, .001, .05, .01, respectively.

lies in the fact that the first method works even for the case of nonstationary sources with identical spectra shape, whereas the latter does not (see the result of Experiment 1).

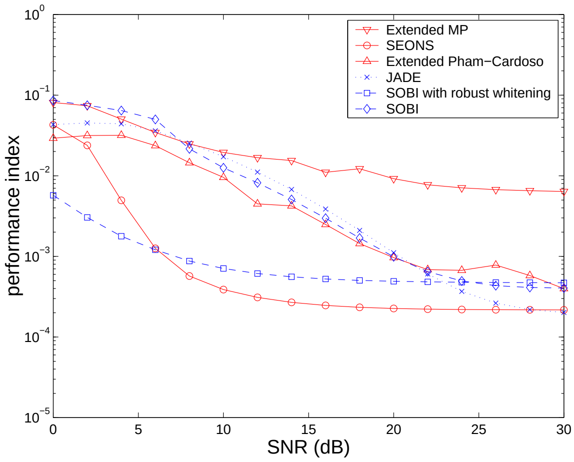


Figure 2: The performance comparison for SEONS, SOBI, SOBI with robust whitening, extended MP, JADE, and extended Pham-Cardoso.

5. CONCLUSION

In this paper we have presented a BSS method that jointly exploits the nonstationarity and temporal structure of sources. We have shown that our method, SEONS, was robust with respect to the temporally white noise and worked well even for the case of several nonstationary Gaussian sources (with no temporal correlations).

6. ACKNOWLEDGMENT

This work was supported by KOSEF 2000-2-20500-009-5 and by Korea Ministry of Science and Technology under an International Cooperative Research Project and by ETRI.

7. REFERENCES

[1] S. Haykin, *Unsupervised Adaptive Filtering: Blind Source Separation*. Prentice-Hall, 2000.

[2] S. Amari and A. Cichocki, "Adaptive blind signal processing - neural network approaches," *Proc. of the IEEE, Special Issue on Blind Identification and Estimation*, vol. 86, no. 10, pp. 2026–2048, Oct. 1998.

[3] A. Belouchrani, K. Abed-Merain, J. F. Cardoso, and E. Moulines, "A blind source separation technique using second order statistics," *IEEE Trans. Signal Processing*, vol. 45, pp. 434–444, Feb. 1997.

[4] C. Chang, Z. Ding, S. F. Yau, and F. H. Y. Chan, "A matrix-pencil approach to blind separation of colored nonstationary signals," *IEEE Trans. Signal Processing*, vol. 48, pp. 900–907, Mar. 2000.

[5] K. R. Müller, P. Philips, and A. Ziehe, "JADE$_{TD}$: Combining higher-order statistics and temporal information for blind source separation (with noise)," in *Proc. ICA'99*, (Aussois, France), pp. 87–92, 1999.

[6] S. Choi and A. Cichocki, "Blind separation of nonstationary sources in noisy mixtures," *Electronics Letters*, vol. 36, pp. 848–849, Apr. 2000.

[7] S. Choi and A. Cichocki, "Blind separation of nonstationary and temporally correlated sources from noisy mixtures," in *Proc. IEEE Workshop on Neural Networks for Signal Processing*, (Sidney, Austrailia), pp. 405–414, 2000.

[8] A. Belouchrani and A. Cichocki, "Robust whitening procedure in blind source separation context," *Electronics Letters*, vol. 36, pp. 2050–2051, Nov. 2000.

[9] L. Tong, Y. Inouye, and R. Liu, "A finite-step global convergence algorithm for the parameter estimation of multichannel MA processes," *IEEE Trans. Signal Processing*, vol. 40, pp. 2547–2558, Oct. 1992.

[10] J. F. Cardoso and A. Souloumiac, "Jacobi angles for simultaneous diagonalization," *SIAM J. Mat. Anal. Appl.*, vol. 17, pp. 161–164, Jan. 1996.

[11] J. F. Cardoso and A. Souloumiac, "Blind beamforming for non Gaussian signals," *IEE Proceedings-F*, vol. 140, no. 6, pp. 362–370, 1993.

[12] D. T. Pham, "Joint approximate diagonalization of positive definite hermitian matrices," Tech. Rep. LMC/IMAG, University of Grenoble, France, 2000.

[13] D. T. Pham and J. F. Cardoso, "Blind separation of instantaneous mixtures of nonstationary sources," in *Proc. ICA*, (Helsinki, Finland), pp. 187–192, 2000.