FROM TIME-DOMAIN SEPARATION OF STATIONARY TEMPORALLY CORRELATED SOURCES TO FREQUENCY-DOMAIN SEPARATION OF NONSTATIONARY SOURCES

Shahram HOSSEINI, Yannick DEVILLE, Hicham SAYLANI

Laboratoire d'Astrophysique de Toulouse-Tarbes
Université Paul Sabatier , Observatoire Midi-Pyrénées
14 Av. Edouard Belin, 31400 Toulouse, France
shosseini or ydeville or hsaylani@ast.obs-mip.fr

ABSTRACT

This paper demonstrates and exploits some interesting frequency-domain properties of nonstationary signals. Considering these properties, a new method for blind separation of linear instantaneous mixtures of mutually uncorrelated, nonstationary, real sources is proposed which is based on spectral decorrelation of the sources. It allows the existing time-domain algorithms developed for stationary, temporally correlated sources to be applied to nonstationary, temporally uncorrelated sources just by mapping the mixtures into the frequency domain. The method sets no constraint on the piecewise stationarity of the sources, unlike most of previously reported methods.

1. INTRODUCTION

Linear instantaneous blind source separation consists in recovering unobserved source signals from several observed signals which are supposed to be linear instantaneous mixtures of these source signals. It has been shown that this goal can be achieved by exploiting nonGaussianity, time correlation or nonstationarity [1], leading to numerous algorithms [2].

In this paper, our goal is to propose a new approach using the nonstationarity of the sources. A few authors have studied this problem [3]-[6]. Many of these works use a statistical framework and take advantage of the assumed nonstationarity of the variance of the sources. In [3], separation of nonstationary signals is achieved by computing output components which are uncorrelated *at every time point*, using a recurrent neural network. In [4], the observed signals are divided in two subintervals. Then, the joint diagonalization of two covariance matrices, estimated on the two subintervals, allows one to separate the sources. Another approach, presented in [5], is based on the maximization of the nonstationarity, measured by the cross-cumulant, of a linear combination of the observed mixtures. In [6], the authors develop novel approaches based on the principles of maximum likelihood and minimum mutual information.

The third author is also with Laboratoire d'Automatique et de Micro-Ondes, Université Ibn Tofail, Kénitra, Morocco.

In most of these works, the estimation of the considered statistics requires that they do not change within some intervals. This means that the nonstationary sources are supposed piecewise stationary with respect to the considered statistics, while this hypothesis may not be realistic for many real-world signals.

The statistical, frequency-domain method proposed in the present paper is based on spectral decorrelation of the signals. It results from some interesting frequency-domain properties of nonstationary signals, and may be used for separating linear instantaneous mixtures of Gaussian or nonGaussian nonstationary, mutually uncorrelated signals. The piecewise stationarity hypothesis is not required for the proposed method.

We should mention that frequency-domain methods have been used for separating convolutive mixtures of nonstationary sources (see for example [7]), but in a totally different context which consists in transforming a convolutive time-domain mixture into an instantaneous frequency-domain mixture. Moreover, there exist several methods exploiting the time-frequency diversity of the nonstationary sources for separating them [8]-[11].

2. PROBLEM STATEMENT

In a general framework (without noise and with the same numbers of mixtures and sources), the blind separation of instantaneous linear mixtures can be formulated as follows. Suppose N samples of K instantaneous mixtures of K unknown *discrete-time* sources[1] are available. The mixing model is given by:

$$\mathbf{x}(n) = \mathbf{A}\mathbf{s}(n) \qquad (1)$$

where $\mathbf{x}(n) = [x_1(n), x_2(n), \cdots, x_K(n)]^T$ and $\mathbf{s}(n) = [s_1(n), s_2(n), \cdots, s_K(n)]^T$ are, respectively, the observation and the source vectors, and $\mathbf{A}$ is an *unknown* mixing matrix. We suppose the sources are zero-mean, real signals, and the mixing matrix $\mathbf{A}$ is real and nonsingular. The goal is to find an

[1] In this paper, we consider discrete-time signals because in practice one deals usually with these signals. However, it must be emphasized that the methods proposed in this paper can be used also for processing continuous-time signals.

estimate of the matrix $\mathbf{A}$ (or its inverse, the separating matrix) up to a permutation and a diagonal matrix. In the following, we suppose also that the components of the source vector $\mathbf{s}(n)$ in (1), *i.e.* the source signals $s_i(n)$, are mutually uncorrelated. In other words, we suppose that $E[s_i(n_1)s_j(n_2)] = 0 \quad \forall i \neq j, \forall n_1, n_2$.

Let's denote the Fourier transforms[2] of $s_i(n)$ and $x_i(n)$ by $S_i(\omega)$ and $X_i(\omega)$, and define $\mathbf{S}(\omega) = [S_1(\omega), S_2(\omega), \cdots, S_K(\omega)]^T$ and $\mathbf{X}(\omega) = [X_1(\omega), X_2(\omega), \cdots, X_K(\omega)]^T$. Taking the Fourier transform of (1), we obtain:

$$\mathbf{X}(\omega) = \mathbf{A}\mathbf{S}(\omega). \tag{2}$$

Proposition 1: $E[S_i(\omega_1)S_j^*(\omega_2)] = 0 \quad \forall i \neq j, \forall \omega_1, \omega_2$, where $S^*(\omega)$ is the complex conjugate of $S(\omega)$.

Proof: See Appendix A.

The following corollary, results from Proposition 1, and will be used in our method.

Corollary 1: The matrix $\mathbf{P}_S(\omega, \upsilon) = E[\mathbf{S}(\omega+\upsilon)\mathbf{S}^H(\omega)]$, where $\mathbf{S}^H$ denotes the Hermitian transpose of $\mathbf{S}$, is diagonal for every value of υ.

3. METHOD

Our method is based on the following theorem.

Theorem 1: If $s(n)$ is a temporally uncorrelated, real, zero-mean signal with a nonstationary variance $q(n)$, *i.e.* if $E[s(n_1)s(n_2)] = q(n_1)\delta(n_1 - n_2)$, where $\delta(n)$ is the unit impulse, then its Fourier transform, $S(\omega)$ is a wide-sense stationary, autocorrelated process with autocorrelation $Q(\upsilon)$, which is the Fourier transform of $q(n)$, *i.e.*

$$E[S(\omega + \upsilon)S^*(\omega)] = Q(\upsilon) = \sum_{n=-\infty}^{\infty} q(n)e^{-j\upsilon n}. \tag{3}$$

Proof: See Appendix B.

Hence, if we suppose that the mutually uncorrelated sources $s_i(n)$ are real, zero-mean, temporally uncorrelated and nonstationary with respect to their variances, then Proposition 1, Theorem 1 and Equation (2) entail that $X_i(\omega)$ are linear mixtures of mutually uncorrelated, wide-sense stationary and autocorrelated frequency-domain processes $S_i(\omega)$. Many algorithms have been proposed for separating such mixtures [13]-[18]. Although these algorithms were originally developed for time-domain wide-sense stationary, time-correlated processes, nothing prohibits us from applying them to frequency-domain wide-sense stationary, frequency-correlated processes. Thus, only by mapping the nonstationary temporally uncorrelated observed signals in the frequency domain, the source separation can be achieved using one of the numerous methods developed previously for time-correlated stationary mixtures.

A simple BSS algorithm which may be considered as a frequency-domain variant of the time-domain AMUSE algorithm [14] consists in jointly diagonalizing the two matrices $\mathbf{P}_X(\omega, 0) = E[\mathbf{X}(\omega)\mathbf{X}^H(\omega)]$ and $\mathbf{P}_X(\omega, \upsilon_1) = E[\mathbf{X}(\omega + \upsilon_1)\mathbf{X}^H(\omega)]$ for some frequency lag υ_1. The joint diagonalization may be achieved using generalized eigenvalue decomposition as the following theorem suggests.

Theorem 2: Suppose $s_i(n)$ are K mutually uncorrelated zero-mean signals. Suppose also there is a constant υ_1 such that $\forall i \neq j$

$$\frac{E[S_i(\omega + \upsilon_1)S_i^*(\omega)]}{E[|S_i(\omega)|^2]} \neq \frac{E[S_j(\omega + \upsilon_1)S_j^*(\omega)]}{E[|S_j(\omega)|^2]}. \tag{4}$$

If $\mathbf{V}$ is a matrix whose columns are the eigenvectors of $\mathbf{P}_X(\omega, 0)^{-1}\mathbf{P}_X(\omega, \upsilon_1)$, *i.e.* if

$$\mathbf{P}_X(\omega, 0)^{-1}\mathbf{P}_X(\omega, \upsilon_1) = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1} \tag{5}$$

where $\mathbf{\Lambda}$ is a diagonal matrix[3], then $\mathbf{V} = \mathbf{D}\mathbf{P}\mathbf{A^T}^{-1}$, where $\mathbf{D}$ is a diagonal matrix and $\mathbf{P}$ is a permutation matrix.

Proof: See Appendix C.

An inverse theorem can be formulated as follows.

Theorem 3: Suppose $\exists i \neq j$ such that

$$\frac{E[S_i(\omega + \upsilon_1)S_i^*(\omega)]}{E[|S_i(\omega)|^2]} = \frac{E[S_j(\omega + \upsilon_1)S_j^*(\omega)]}{E[|S_j(\omega)|^2]} \tag{6}$$

for a given frequency ω and a constant υ_1. Then, the eigenvalue decomposition (5) at υ_1 does not give the matrix $\mathbf{A}$ up to a permutation and a diagonal matrix.

Proof: See Appendix D.

Following Theorem 1, if two sources $s_i(n)$ and $s_j(n)$ are temporally-uncorrelated, real, zero-mean signals with nonstationary variances $q_i(n)$ and $q_j(n)$, then the numerators and the denominators in (4) are the Fourier transforms of these variances at the frequencies υ_1 and zero. Thus, the sources may be separated only if they have different variance profiles.

Since the processes $S_i(\omega)$ and therefore $X_i(\omega)$ are wide-sense stationary, we can hope they are also wide-sense ergodic, so that the expected values can be estimated by frequency averages. In this case, the proposed BSS algorithm reduces to jointly diagonalizing the two sample covariance matrices $\hat{\mathbf{P}}_X(\omega, 0) = \sum_\omega \mathbf{X}(\omega)\mathbf{X}^H(\omega)$ and $\hat{\mathbf{P}}_X(\omega, \upsilon_1) = \sum_\omega \mathbf{X}(\omega + \upsilon)\mathbf{X}^H(\omega)$.

[2]The Fourier transform of a discrete-time stochastic process $u(n)$ is a stochastic process $U(\omega)$ given by $U(\omega) = \sum_{n=-\infty}^{\infty} u(n)e^{-j\omega n}$ [12].

[3]In fact, the diagonal entries of $\mathbf{\Lambda}$ are the *generalized eigenvalues* of the two matrices $\mathbf{P}_X(\omega, 0)$ and $\mathbf{P}_X(\omega, \upsilon_1)$, and the columns of $\mathbf{V}$ are the *generalized eigenvectors* of these two matrices, because $\mathbf{P}_X(\omega, \upsilon_1)\mathbf{V} = \mathbf{P}_X(\omega, 0)\mathbf{V}\mathbf{\Lambda}$.

Note also that, similar to the time domain algorithm, the diagonalization may be done serially, *i.e.* by first whitening data which is equivalent to diagonalizing $\mathbf{P}_X(\omega, 0)$ and then by computing $\mathbf{P}_X(\omega, \upsilon_1)$ on the whitened data and diagonalizing it using a unitary matrix.

The constant υ_1 may be chosen by plotting the empirical autocorrelations of the sequences $X_i(\omega)$ and by choosing a frequency lag ensuring (4). Unlike in the temporal method, the choice $\upsilon_1 = 1$ is not always the best. A good idea is to choose a nonzero value of ω maximizing the autocorrelation function. An extended version of the proposed method, which may improve the separation performance too, is to simultaneously diagonalize several covariance matrices corresponding to several frequency lags, which can be considered as a frequency domain equivalent of the SOBI algorithm [13]. We will come back to these points in the following section.

4. SIMULATION RESULTS

In the first experiment, we considered the two sources shown in Figure 1, which were obtained by multiplying two independent Gaussian i.i.d. signals, respectively by a sinus and by a periodical triangle, both of frequency $f_0 = 8Hz$. The mixing matrix is $\mathbf{A} = \begin{pmatrix} 1 & 0.9 \\ 0.8 & 1 \end{pmatrix}$.

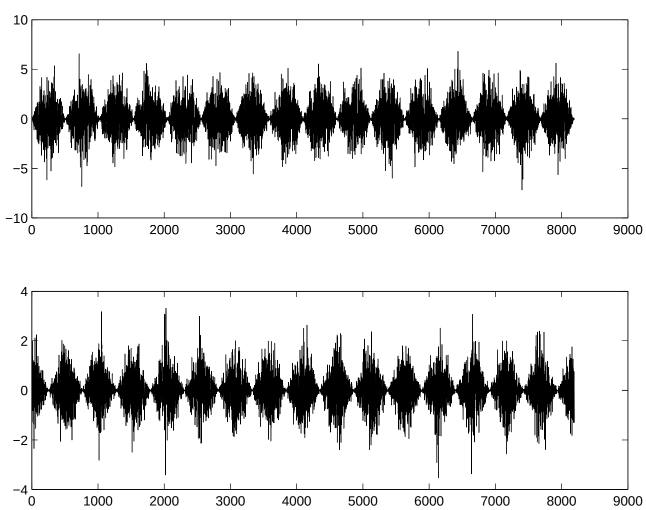

Fig. 1. The two sources used in the first experiment.

The experiment was done using 1 second of the two sources containing 8192 samples. The absolute value of the autocorrelation function of $X_1(\omega)$ is shown in Figure 2 which presents three main peaks at $\upsilon = 0$ and $\upsilon = \pm 16Hz$ (this can be demonstrated easily by computing the autocorrelations of the two sources and by using the result of Theorem 1). The separating matrix may be estimated by applying the method mentioned in the previous section choosing $\upsilon_1 = 16Hz$.

We used a modified version of the AMUSE algorithm [14] for this purpose. This simple and fast algorithm, originally

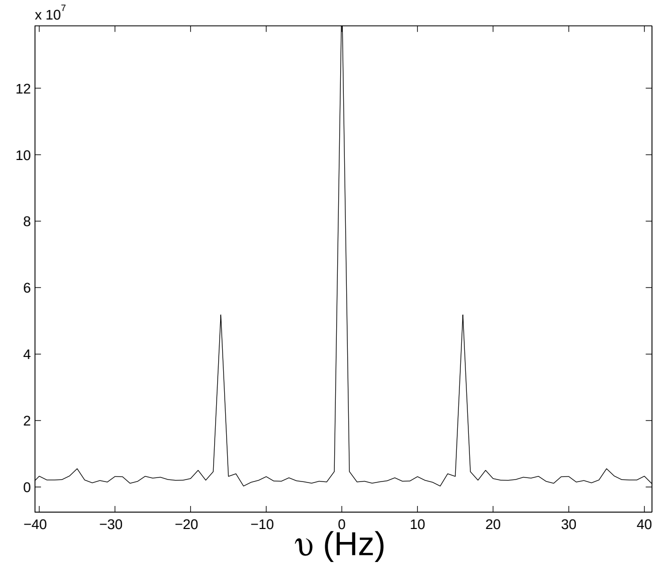

Fig. 2. Absolute value of the Autocorrelation function of $X_1(\omega)$.

developed for separating time-correlated stationary sources in the time domain, here works as follows. (a) Spatially whiten the data $\mathbf{X}(\omega)$ to obtain $\mathbf{Z}(\omega)$. (b) Compute the eigenvalue decomposition of the symmetric matrix $\overline{\mathbf{C}_{\upsilon_1}^{\mathbf{Z}}} = \frac{1}{2}[\mathbf{C}_{\upsilon_1} + \mathbf{C}_{\upsilon_1}^T]$, where $\mathbf{C}_{\upsilon_1} = E[\mathbf{Z}(\omega + \upsilon_1)\mathbf{Z}^*(\omega)]$ is the covariance matrix corresponding to lag υ_1. (c) The rows of the separating matrix $\mathbf{B}$ are given by the eigenvectors of $\overline{\mathbf{C}_{\upsilon_1}^{\mathbf{Z}}}$.

The experiment was repeated 100 times corresponding to 100 different seed values of the random variable generator. For each experiment, the output Signal to Noise Ratio (in dB) was computed by $SNR = 0.5\sum_{i=1}^{2} 10\log_{10}\frac{E[s_i^2]}{E[(\hat{s}_i - s_i)^2]}$, after normalizing the estimated sources, $\hat{s}_i(n)$, so that they have the same variances and signs as the source signals, $s_i(n)$. The mean and the standard deviation of SNR on the 100 experiments were 51.8 dB and 5.9 dB .

Other experiments with different profiles of nonstationary variance for the sources $s_1(n)$ and $s_2(n)$ led to similar results.

In the second experiment, the above algorithm based on AMUSE was used for separating mixtures of speech signals. Three tests using three couples of 44100-sample speech signals led to an average SNR of 40.6 dB. A modified algorithm aiming at joint diagonalizing several covariance matrices corresponding to several frequency lags (which may be considered as a frequency equivalent of the SOBI algorithm) was also used for separating the same speech signals, and led to an average SNR of 46.7 dB.

This experiment shows that although Theorem 1 is derived for temporally uncorrelated signals, the proposed method works well also for temporally correlated signals.

5. CONCLUSION

A major objective of this paper was to demonstrate and exploit some theoretically interesting frequency-domain properties of signals which are nonstationary in the time domain. These properties provide sufficient second-order constraints in the frequency domain for separating instantaneous linear mixtures of nonstationary sources.

A separating method was proposed based on these properties. This method is very simple and powerful because it allows the time-domain algorithms developed for stationary time-correlated signals to be applied to temporally uncorrelated sources which are nonstationary in the time domain, just by mapping the signals in the frequency domain. It should be remarked that this algorithm does not require the variance of the sources to be constant over subintervals, while this hypothesis is necessary in the majority of the source separation algorithms based on the nonstationarity of variance which have been reported in the literature.

A. PROOF OF PROPOSITION 1

Consider two mutually uncorrelated zero-mean real signals $s_i(n)$ and $s_j(n)$, with Fourier transforms $S_i(\omega)$ and $S_j(\omega)$. We can write:

$$E[S_i(\omega_1)S_j^*(\omega_2)] = \sum_{n_1=-\infty}^{\infty}\sum_{n_2=-\infty}^{\infty} E[s_i(n_1)s_j(n_2)]e^{-j(\omega_1 n_1 - \omega_2 n_2)} = 0$$

because $E[s_i(n_1)s_j(n_2)] = 0 \ \ \forall n_1, n_2$.

B. PROOF OF THEOREM 1

A process $S(\omega)$ is wide-sense stationary if its expected value is constant *i.e.* $E[S(\omega)] = \eta$ and if its autocorrelation $E[S(\omega+\upsilon)S^*(\omega)]$ is not a function of ω *i.e.* $E[S(\omega+\upsilon)S^*(\omega)] = Q(\upsilon)$.

1) $s(n)$ is supposed zero-mean. Hence

$$E[S(\omega)] = \sum_{n_1=-\infty}^{\infty} E[s(n)]e^{-j\omega n} = 0$$

2) If $E[s(n_1)s(n_2)] = q(n_1)\delta(n_1 - n_2)$, where $\delta(n_1 - n_2)$ is a unit impulse, then

$$\begin{aligned} & E[S(\omega+\upsilon)S^*(\omega)] \\ &= \sum_{n_1=-\infty}^{\infty}\sum_{n_2=-\infty}^{\infty} E[s(n_1)s(n_2)]e^{-j(\omega+\upsilon)n_1}e^{j\omega n_2} \\ &= \sum_{n_1=-\infty}^{\infty}\sum_{n_2=-\infty}^{\infty} q(n_1)\delta(n_1-n_2)e^{-j\omega(n_1-n_2)}e^{-j\upsilon n_1}. \end{aligned}$$

Since $\delta(n_1 - n_2)e^{-j\omega(n_1-n_2)} = \delta(n_1 - n_2)$,

$$\begin{aligned} & E[S(\omega+\upsilon)S^*(\omega)] \\ &= \sum_{n_1=-\infty}^{\infty} q(n_1)e^{-j\upsilon n_1} \sum_{n_2=-\infty}^{\infty} \delta(n_1-n_2) \\ &= \sum_{n_1=-\infty}^{\infty} q(n_1)e^{-j\upsilon n_1} = Q(\upsilon). \end{aligned}$$

C. PROOF OF THEOREM 2

From (2), we have

$$\mathbf{P}_X(\omega,\upsilon_1) = \mathbf{A}\mathbf{P}_S(\omega,\upsilon_1)\mathbf{A}^H = \mathbf{A}\mathbf{P}_S(\omega,\upsilon_1)\mathbf{A}^T \qquad (7)$$

and

$$\mathbf{P}_X(\omega,0) = \mathbf{A}\mathbf{P}_S(\omega,0)\mathbf{A}^H = \mathbf{A}\mathbf{P}_S(\omega,0)\mathbf{A}^T \qquad (8)$$

because $\mathbf{A}$ is real. If $\mathbf{P}_S(\omega,0)$ is nonsingular, *i.e.* if $E[|S_i(\omega)|^2] \neq 0 \ \forall i$, then left multiplying (7) by the inverse of (8) yields

$$\mathbf{P}_X^{-1}(\omega,0)\mathbf{P}_X(\omega,\upsilon_1) = \mathbf{A}^{T^{-1}}\mathbf{P}_S^{-1}(\omega,0)\mathbf{P}_S(\omega,\upsilon_1)\mathbf{A}^T. \qquad (9)$$

Since according to Corollary 1 $\mathbf{P}_S^{-1}(\omega,0)\mathbf{P}_S(\omega,\upsilon_1)$ is a diagonal matrix, the above equation is nothing but an eigenvalue decomposition of the matrix $\mathbf{P}_X^{-1}(\omega,0)\mathbf{P}_X(\omega,\upsilon_1)$. If the K eigenvalues are distinct (*i.e.* if the algebraic multiplicity of each eigenvalue equals one), then the dimension of the eigenspace corresponding to each eigenvalue equals one (see [19]-page 58). In other words, if $\mathbf{v}$ and $\mathbf{u}$ are two eigenvectors corresponding to the same eigenvalue λ, then $\mathbf{u} = \alpha\mathbf{v}$ where α is a (complex) scalar. Moreover, it is clear that the eigenvalues may be arranged as diagonal entries of a diagonal matrix in an arbitrary order.

Hence, if the matrix $\mathbf{P}_X^{-1}(\omega,0)\mathbf{P}_X(\omega,\upsilon_1)$ has K distinct eigenvalues, *i.e.* if $\frac{E[S_i(\omega+\upsilon_1)S_i^*(\omega)]}{E[|S_i(\omega)|^2]} \neq \frac{E[S_j(\omega+\upsilon_1)S_j^*(\omega)]}{E[|S_j(\omega)|^2]} \ \forall i \neq j$, and if $\mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$ is an eigenvalue decomposition of $\mathbf{P}_X^{-1}(\omega,0)\mathbf{P}_X(\omega,\upsilon_1)$, then the columns of $\mathbf{V}$ are equal to the columns of $\mathbf{A}^{T^{-1}}$ up to a scaling factor and a permutation, so that $\mathbf{V} = \mathbf{D}\mathbf{P}\mathbf{A}^{\mathbf{T}^{-1}}$, where $\mathbf{D}$ is a diagonal matrix and $\mathbf{P}$ is a permutation matrix.

D. PROOF OF THEOREM 3

If $\lambda = \frac{E[S_i(\omega+\upsilon_1)S_i^*(\omega)]}{E[|S_i(\omega)|^2]} = \frac{E[S_j(\omega+\upsilon_1)S_j^*(\omega)]}{E[|S_j(\omega)|^2]}$, then $\mathbf{P}_X^{-1}(\omega,0)$ $\mathbf{P}_X(\omega,\upsilon_1)$ has two identical eigenvalues λ. Since $\mathbf{A}$ is supposed nonsingular, the columns of $\mathbf{A}^{T^{-1}}$ (which represent the eigenvectors of $\mathbf{P}_X^{-1}(\omega,0)\mathbf{P}_X(\omega,\upsilon_1)$) are linearly independent. Hence, the eigenspace corresponding to λ, and spanned by two columns of $\mathbf{A}^{T^{-1}}$, is of dimension 2. It is well known that every nonzero element of this eigenspace is an eigenvector corresponding to λ (see [19]-section 1.4). Therefore, the two columns of $\mathbf{A}^{T^{-1}}$ corresponding to λ can not be identified up to a permutation and a scaling factor using the eigenvalue decomposition (5).

E. REFERENCES

[1] J.-F. Cardoso, The three easy routes to independent component analysis: contrast and geometry, in *Proc. ICA2001*, San Diego, 2001, pp. 1-6.

[2] A. Hyvarinen, J. Karhunen, E. Oja, *Independent component analysis*, John Wiley and Sons, 2001.

[3] K. Matsuoka, M. Ohya, and Mitsuru Kawamoto, A neural net for blind separation of nonstationary signals, *Neural Networks*, vol. 8, no. 3, pp. 411-419, 1995.

[4] A. Souloumiac, Blind source detection and separation using second-order non-stationarity, in *Proc. ICASSP*, 1995, pp. 1912-1915.

[5] A. Hyvarinen, Blind source separation by nonstationarity of variance: a cumulant based approach, *IEEE Trans. on Neural Networks*, 12(6), pp. 1471-1474, 2001.

[6] D.-T. Pham, and J.-F. Cardoso, Blind separation of independent mixtures of nonstationary sources, *IEEE Trans. on Signal Processing*, 49(9), 2001.

[7] L. Parra, and C. Spence, Convolutive blind separation of non-stationary sources, *IEEE Trans. on Speech and Audio Processing*, 8(3), 2000.

[8] A. Belouchrani, and M. G. Amin, Blind source separation based on time-frequency signal representation, *IEEE Trans. on Signal Processing*, 46(11), November 1998.

[9] F. Abrard, Y. Deville, P. White, From blind source separation to blind source cancellation in the underdetermined case: a new approach based on time-frequency analysis, in *Proc. ICA2001*, pp. 734-739, San Diego, USA, 2001.

[10] Y. Deville, Temporal and time-frequency correlation-based blind source separation methods, in *Proc. ICA2003*, pp. 1059-1064, Nara, Japan, 2003.

[11] B. Albouy, Y. Deville, A time-frequency blind source separation method based on segmented coherence function, in *Proc. IWANN2003*, vol. 2, pp. 289-296, Mao, Menorca, Spain, 2003.

[12] A. Papoulis, and S. U. Pillai, *Probability, random variables and stochastic processes*, 4th Ed., McGraw-Hill, 2002.

[13] A. Belouchrani, K. Abed Meraim, J.-F. Cardoso, and E. Moulines, A blind source separation technique based on second order statistics, *IEEE Trans. on Signal Processing*, vol. 45, pp. 434-444, Feb. 1997.

[14] L. Tong, R. Liu, V. Soon, and Y. Huang, Indeterminacy and identifiability of blind identification, *IEEE Trans. Circuits Syst.*, vol. 38, pp. 499-509, May 1991.

[15] L. Molgedey, and H. G. Schuster, Separation of a mixture of independent signals using time delayed correlation, *Physical Review Letters*, vol. 72, pp. 3634-3636, 1994.

[16] A. Ziehe, and K. R. Muller, TDSEP - an efficient algorithm for blind separation using time structure, in *Proceedings of Int. Conf. on Artificial Neural Networks*, Skovde, Sweden, pp. 675-680, 1998.

[17] S. Degerine, and R. Malki, Second order blind separation of sources based on canonical partial innovations, *IEEE Trans. on Signal Processing*, vol. 48, pp. 629-641, 2000.

[18] S. Hosseini, C. Jutten, and D.-T. Pham, Markovian source separation, *IEEE Trans. on Signal Processing*, vol. 51, pp. 3009-3019, 2003.

[19] R. A. Horn, and C. R. Johnson, *Matrix analysis*, Cambridge University Press, 1999.