UNDERDETERMINED BLIND SEPARATION OF AUDIO SOURCES FROM THE TIME-FREQUENCY REPRESENTATION OF THEIR CONVOLUTIVE MIXTURES

Abdeldjalil Aïssa-El-Bey, Karim Abed-Meraim and Yves Grenier

ENST-Paris, TSI Department, 46 rue Barrault 75634, Paris Cedex 13, France
{elbey, abed, grenier}@tsi.enst.fr

ABSTRACT

This paper considers the blind separation of nonstationary sources in the underdetermined convolutive mixture case. We introduce two methods based on the sparsity assumption of the sources in the time-frequency (TF) domain. The first one assumes that the sources are disjoint in the TF domain; i.e. there is at most one source signal present at a given point in the TF domain. In the second method, we relax this assumption by allowing the sources to be TF-nondisjoint to a certain extent. In particular, the number of sources present (active) at a TF point should be strictly less than the number of sensors. In that case, the separation can be achieved thanks to subspace projection which allows us to identify the active sources and to estimate their corresponding time-frequency distribution (TFD) values.

Index Terms— Separation, deconvolution, time-frequency analysis, identification.

1. INTRODUCTION

The blind source separation of more sources than sensors (referred to as UBSS for underdetermined blind source separation) is still a challenging problem especially in the convolutive mixtures case. In the instantaneous mixture case, some methods exploiting the sparseness of the sources in certain transform domain have been proposed for UBSS [1–4]. These methods proceed 'roughly' as follows: The mixtures are first transformed to an appropriate representation domain; the transformed sources are then estimated using their sparseness, and finally one recovers their time waveforms by source synthesis (for more information, see the recent survey work [5]).

UBSS methods for *nonstationary sources* have been proposed, given that these sources are sparse in the time-frequency (TF) domain [2, 3]. The main assumption used in these methods is that the sources are TF-disjoint. In other words, there is at most one source present at any point in the TF domain. This assumption is rather restrictive, though the methods have also showed that they worked well under a quasi sparseness condition, i.e. sources are TF-almost-disjoint.

In this paper we focus on the UBSS in convolutive mixture case and target the relaxation of the TF-disjointness condition by allowing the sources to be *nondisjoint* in the TF domain; that is, multiple sources are possibly present at any point in the TF domain. This case has been considered in [4] for the separation of instantaneous mixtures, in [6] for the deconvolution of single-path channels with non-zero delays and in [7] where binary TF-masking and ICA technique are jointly used. The main contribution of this paper consists in two new algorithms (TF-CUBSS for Time-frequency convolutive underdetermined blind source separation) for solving the UBSS in the TF domain; the first one uses vector clustering while the other uses subspace projection.

2. PROBLEM FORMULATION

2.1. Data model

Let $s_1(t), \ldots, s_N(t)$ be the desired sources to be recovered from the convolutive mixtures $x_1(t), \ldots, x_M(t)$ given by:

$$\mathbf{x}(t) = \sum_{k=0}^{K} \mathbf{H}(k)\mathbf{s}(t-k) + \mathbf{w}(t) \tag{1}$$

where $\mathbf{s}(t) = [s_1(t), \ldots, s_N(t)]^T$ is the source vector, $\mathbf{x}(t) = [x_1(t), \ldots, x_M(t)]^T$ is the mixture vector (with $M < N$), $\mathbf{w}(t)$ is the observation noise, and $\mathbf{H}(k)$ are $M \times N$ matrices for $k \in [0, K]$ representing the impulse response coefficients of the channel that satisfies:

Assumption 1 *The channel is such that each column vector of*

$$\mathbf{H}(z) \stackrel{\text{def}}{=} \sum_{k=0}^{K} \mathbf{H}(k) z^{-k} \stackrel{\text{def}}{=} [\mathbf{h}_1(z), \ldots, \mathbf{h}_N(z)]$$

is irreducible, i.e. the entries of $\mathbf{h}_i(z)$ *denoted* $h_{ji}(z)$, $j = 1, \ldots, M$, *have no common zeros* $\forall i$. *Moreover, any* M *column vectors of* $\mathbf{H}(z)$ *form a polynomial matrix* $\tilde{\mathbf{H}}(z)$ *that it full rank over the unit-circle, i.e. rank*$(\tilde{\mathbf{H}}(f)) = M \ \forall f$.

2.2. TF conditions on the sources

In order to deal with UBSS, one often seeks for a sparse representation of the sources [1]. In other words, if the sources can be sparsely represented in some domain, then their separation can be carried out in that domain by exploiting their sparseness.

2.2.1. *TF-disjoint sources*

Recently, there have been several UBSS methods, notably those in [2] and [3], in which the TF domain has been chosen to be the underlaying sparse domain. These two papers have based their solutions on the assumption that the sources are disjoint in the TF domain. Mathematically, if Ω_1 and Ω_2 are the TF supports of two sources $s_1(t)$ and $s_2(t)$ then the sources are said TF-disjoint if $\Omega_1 \cap \Omega_2 = \emptyset$. However, this is a rather strict assumption. A more practical assumption is that the sources are almost-disjoint in the TF domain [2], allowing some small overlapping in the TF domain, for which the above two methods also worked.

2.2.2. *TF-nondisjoint sources*

In this paper, we want to relax the TF-disjoint condition by allowing the sources to be nondisjoint in the TF domain.
Therefore, we will allow the sources to be *nondisjoint* in the TF domain; that is, multiple sources are allowed to be present at any point in the TF domain. However, instead of being inevitably nondisjoint, we limit ourselves by making the following constraint:

Assumption 2 *The number of active sources (i.e. sources that overlap) at any TF point is strictly less than the number of sensors.*

In other words, for the configuration of M sensors, there exists at most $(M-1)$ overlapping sources at any point in the TF domain. For the special case when $M=2$, Assumption 2 reduces to the disjoint condition.

3. TF-CUBSS ALGORITHM

In order to solve the UBSS problem in the convolutive case, we propose to identify first the impulse response of the channels. This problem in overdetermined case is very difficult and becomes almost impossible in the underdetermined case without side information on the considered sources. In this work and similarly to [8], we exploit the sparseness property of the audio sources by assuming that from time to time only one source is present. In other words, we consider the following assumption:

Assumption 3 *There exists, periodically, time intervals where only one source is present in the mixture. This occurs for all source signals of the considered mixtures.*

To detect these time intervals, we propose to use information criteria based testing for the estimation of the number of sources present in the signal (see Section 3.1 for more details).

3.1. Channel estimation

Based on assumption 3, we propose here to apply SIMO (Single Input Multiple Output) based techniques to blindly estimate the channel impulse response. Regarding the problem at hand, we have to solve three different problems: first, we have to select time intervals where only one source signal is effectively present; then, for each selected time interval one should apply an appropriate blind SIMO identification technique to estimate the channel parameters; finally, the way we proceed, the same channel may be estimated several times and hence one has to group together (cluster) the channel estimates into N classes corresponding to the N source channels.

3.1.1. Source number estimation

Let define the spatio-temporal vector:

$$\mathbf{x}_d(t) = [\mathbf{x}^T(t), \ldots, \mathbf{x}^T(t-d+1)]^T = \sum_{k=1}^{N} \mathbf{H}_k \mathbf{s}_k(t) + \mathbf{w}_d(t), \quad (2)$$

where $\mathbf{H}_k$ are block-Sylvester matrices of size $dM \times (d+K)$, $\mathbf{s}_k(t) \stackrel{\text{def}}{=} [s_k(t), \ldots, s_k(t-K-d+1)]^T$ and d is a chosen processing window size. Under the data model assumption and for large window sizes (see [9] for more details), matrices $\mathbf{H}_k$ are full column rank. Hence, in the noiseless case, the rank of the data covariance matrix $\mathbf{R} \stackrel{\text{def}}{=} E[\mathbf{x}_d(t)\mathbf{x}_d^H(t)]$ is equal to $\min(p(d+K), dM)$ where p is the number of sources present in the considered time interval over which the covariance matrix is estimated. In particular, for $p=1$, one has the minimum rank value equal to $(d+K)$. Therefore, our approach consists in estimating the rank of the sample averaged covariance matrix $\mathbf{R}$ over several time slots (intervals) and select those corresponding to the smallest rank value $r = d+K$. The estimation of the rank value is done here by Akaike's criterion [9] according to:

$$r = \arg\min_k \left[-2\log \left(\frac{\prod_{i=k+1}^{Md} \lambda_i^{1/(Md-k)}}{\frac{1}{Md-k}\sum_{i=k+1}^{Md} \lambda_i} \right)^{(Md-k)T_s} + 2k(2Md-k) \right], \quad (3)$$

where $\lambda_1 \geq \ldots \geq \lambda_{Md}$ represent the eigenvalues of $\mathbf{R}$ and T_s is the time slot size. Note that it is not necessary at this stage, to know exactly the channel degree K as long as $d > K$ (i.e. an over-estimation of the channel degree is sufficient) in which case the presence of one signal source is characterized by:

$$d < r < 2d \, .$$

3.1.2. Blind channel identification

To perform the blind channel identification, we have used in this paper the Cross-Relation (CR) technique described in [10]. This method is used on the time slots, where only one source signal is active. The latter are selected using the previously described Akaike's criterion. Note that there exist an improved, but more expensive, version of the CR method exploiting the quasi-sparse nature of acoustic impulse response [11] which can be used as well at this stage.

3.1.3. Clustering of channel vector estimates

The first step of our channel estimation method consists in detecting the time slots where only one single source signal is 'effectively' present. However, the same source signal s_i may be present in several time intervals leading to several estimates of the same channel vector $\mathbf{h}_i \stackrel{\text{def}}{=} [h_{1i}(0) \ldots h_{Mi}(0) \ldots h_{1i}(K) \ldots h_{Mi}(K)]^T$.
We end up, finally, with several estimates of each source channel that we need to group together into N classes. This is done by clustering the estimated vectors using k-means algorithm [12]. The i^{th} channel estimate is evaluated as the centroid of the i^{th} class.

3.2. UBSS algorithm with TF-disjoint assumption

In this section, we propose a new *cluster-based TF-CUBSS algorithm* using the STFT (Short Time Fourier Transform) for convolutive mixture case. After transformation into the TF domain using the STFT, the model in (1) becomes (in the noiseless case):

$$\mathcal{S}_{\mathbf{x}}(t,f) = \mathbf{H}(f)\mathcal{S}_{\mathbf{s}}(t,f), \quad (4)$$

where $\mathcal{S}_{\mathbf{x}}(t,f)$ is the mixture STFT vector, $\mathcal{S}_{\mathbf{s}}(t,f)$ is the source STFT vector and $\mathbf{H}(f) = [\mathbf{h}_1(f) \ldots \mathbf{h}_N(f)]$ is the channel Fourier transform matrix. Under the assumption that all sources are disjoint in the TF domain, (4) reduces to

$$\mathcal{S}_{\mathbf{x}}(t,f) = \mathbf{h}_i(f)\mathcal{S}_{s_i}(t,f), \quad \forall (t,f) \in \Omega_i, \forall i \in \mathcal{N}, \quad (5)$$

where $\mathcal{N} = \{1, \ldots, N\}$. Consequently, two TF points (t_1, f_1) and (t_2, f_2) belonging to the same region Ω_i (i.e. corresponding to the source signal s_i) are 'associated' with the same channel $\mathbf{h}_i$.
This latter observation is used next to cluster together the TF points

of a given source signal. More precisely the algorithm proceeds as follows: First, we compute the STFT of the mixtures according to:

$$\mathcal{S}_{x_i}(t,f) = \int_{-\infty}^{\infty} x_i(\tau) w(\tau - t) e^{-j2\pi f \tau} d\tau, \quad i = 1, \dots, M,$$
$$\mathcal{S}_{\mathbf{x}}(t,f) = [\mathcal{S}_{x_1}(t,f), \dots, \mathcal{S}_{x_M}(t,f)]^T.$$

Then, we apply a noise thresholding procedure which mitigates the noise effect but also reduces the computational cost as only the selected TF points are further treated by our algorithm. In particular, for each frequency-slice (t, f_p) of the TFD representation, we apply the following criterion for all the time points t_k belonging to this frequency-slice

$$\text{If} \quad \frac{\|\mathcal{S}_{\mathbf{x}}(t_k, f_p)\|}{\max_t \{\|\mathcal{S}_{\mathbf{x}}(t, f_p)\|\}} > \epsilon, \quad \text{then keep } (t_k, f_p), \tag{7}$$

where ϵ is a small threshold (typically, $\epsilon = 0.01$). Then, the set of all selected points, Ω, is expressed by $\Omega = \bigcup_{i=1}^{N} \Omega_i$, where Ω_i is the TF support of source s_i. Note that, the effects of spreading the noise energy while localizing the source energy in the time-frequency domain amounts to increasing the robustness of the proposed method with respect to noise. Hence, by equation (7), we would keep only time-frequency points where the signal energy is non-negligible, the other time-frequency points are rejected, i.e. not further processed, since considered to represent noise contribution only. Also, due to the noise energy spreading, the contribution of the noise in the source time-frequency points is relatively, negligible at least for moderate and high SNRs. Finally,the clustering procedure can be done as follows: For each TF point, we obtain the spatial direction vectors by:

$$\mathbf{v}(t,f) = \frac{\mathcal{S}_{\mathbf{x}}(t,f)}{\|\mathcal{S}_{\mathbf{x}}(t,f)\|}, \quad (t,f) \in \Omega, \tag{8}$$

and force them, without loss of generality, to have the first entry real and positive.These vectors are clustered into N classes $\{C_i \,|\, i \in \mathcal{N}\}$ by minimizing the criterion:

$$\mathbf{v}(t,f) \in C_i \Longleftrightarrow i = \arg\min_k \left\| \mathbf{v}(t,f) - \frac{\widehat{\mathbf{h}}_k(f) e^{-j\theta_k}}{\|\widehat{\mathbf{h}}_k(f)\|} \right\|^2 \tag{9}$$

where $\widehat{\mathbf{h}}_k(f)$ is the Fourier Transform of the k^{th} channel vector estimate and θ_k is the phase argument of $\widehat{h}_{1k}(f)$ (this is to force the first entry to be real positive as for $\mathbf{v}(t,f)$).
The collection of all points, whose vectors belong to the class C_i, form the TF support Ω_i of source s_i. Therefore, we can estimate the STFT of each source s_i by:

$$\widehat{\mathcal{S}}_{s_i}(t,f) = \begin{cases} \frac{\widehat{\mathbf{h}}_i^H(f)}{\|\widehat{\mathbf{h}}_i(f)\|^2} \mathcal{S}_{\mathbf{x}}(t,f), & \forall (t,f) \in \Omega_i, \\ 0, & \text{otherwise,} \end{cases} \tag{10}$$

since, from (5), we have

$$\frac{\widehat{\mathbf{h}}_i^H(f)}{\|\widehat{\mathbf{h}}_i(f)\|^2} \mathcal{S}_{\mathbf{x}}(t,f) = \frac{\widehat{\mathbf{h}}_i^H(f)\mathbf{h}_i(f)}{\|\widehat{\mathbf{h}}_i(f)\|^2} \mathcal{S}_{s_i}(t,f) \approx \mathcal{S}_{s_i}(t,f), \quad \forall (t,f) \in \Omega_i.$$

3.3. UBSS algorithm with TF-nondisjoint assumption

We have seen the cluster-based TF-CUBSS method, using the STFT. This method relies on the assumption that the sources are TF-disjoint, which led to the TF-transformed structure in (5). The latter is no longer valid, when the sources are nondisjoint in the TF domain. Under the TF-nondisjointness condition, stated in Assumption 2, we propose in this section an alternative method using subspace projection. Recall that the first two steps of the cluster-based quadratic TF-CUBSS algorithm do not rely on the assumption of TF-disjoint sources. Therefore, we can reuse these steps to obtain the channel estimates and all the TF points of the sources, i.e. Ω. Under the TF-nondisjointness condition, consider a TF point $(t,f) \in \Omega$ at which there are $\mathcal{J} < M$ sources[1] $s_{\alpha_1}(t), \dots, s_{\alpha_\mathcal{J}}(t)$ present, where $\alpha_1, \dots, \alpha_\mathcal{J} \in \mathcal{N}$ denote the indices of the active sources at (t,f). Our goal is to identify the sources that are present at (t,f), i.e. $\alpha_1, \dots, \alpha_\mathcal{J}$, and to estimate the STFT of each of these contributing sources. We define the following:

$$\tilde{\mathbf{s}} = [s_{\alpha_1}(t), \dots, s_{\alpha_\mathcal{J}}(t)]^T, \tag{11a}$$
$$\tilde{\mathbf{H}}_\alpha(f) = [\mathbf{h}_{\alpha_1}(f), \dots, \mathbf{h}_{\alpha_\mathcal{J}}(f)]. \tag{11b}$$

Then, (4) is reduced to the following

$$\mathcal{S}_{\mathbf{x}}(t,f) = \tilde{\mathbf{H}}_\alpha(f) \mathcal{S}_{\tilde{\mathbf{s}}}(t,f). \tag{12}$$

Let $\tilde{\mathbf{H}}_\beta(f) = [\mathbf{h}_{\beta_1}(f), \dots, \mathbf{h}_{\beta_\mathcal{J}}(f)]$ and $\mathbf{Q}_\beta(f)$ be the orthogonal projection matrix onto the noise subspace of $\tilde{\mathbf{H}}_\beta(f)$ expressed by:

$$\mathbf{Q}_\beta(f) = \mathbf{I} - \tilde{\mathbf{H}}_\beta(f) \left(\tilde{\mathbf{H}}_\beta^H(f) \tilde{\mathbf{H}}_\beta(f) \right)^{-1} \tilde{\mathbf{H}}_\beta^H(f). \tag{13}$$

We have the following observation:

$$\begin{cases} \mathbf{Q}_\beta(f)\mathbf{h}_i(f) = 0, & i \in \{\beta_1, \dots, \beta_\mathcal{J}\} \\ \mathbf{Q}_\beta(f)\mathbf{h}_i(f) \neq 0, & i \in \mathcal{N} \backslash \{\beta_1, \dots, \beta_\mathcal{J}\} \end{cases}. \tag{14}$$

Consequently, as $\mathcal{S}_{\mathbf{x}}(t,f) \in \text{Range}\{\tilde{\mathbf{H}}_\alpha(f)\}$, we have

$$\begin{cases} \mathbf{Q}_\beta(f)\mathcal{S}_{\mathbf{x}}(t,f) = 0, & \text{if } \{\beta_1, \dots, \beta_\mathcal{J}\} = \{\alpha_1, \dots, \alpha_\mathcal{J}\} \\ \mathbf{Q}_\beta(f)\mathcal{S}_{\mathbf{x}}(t,f) \neq 0, & \text{otherwise} \end{cases}. \tag{15}$$

Since $\mathbf{H}(f)$ has already been estimated by the method presented in Section 3.1, then this observation gives us the criterion to detect the indices $\alpha_1, \dots, \alpha_\mathcal{J}$ and hence, the contributing sources at the considered TF point (t,f). In practice, to take into account noise, one detects the column vectors of $\tilde{\mathbf{H}}_\alpha(f)$ by minimizing:

$$\{\alpha_1, \dots, \alpha_\mathcal{J}\} = \arg \min_{\beta_1, \dots, \beta_\mathcal{J}} \{\|\mathbf{Q}_\beta(f)\mathcal{S}_{\mathbf{x}}(t,f)\|\}. \tag{16}$$

Next, TFD values of the $\mathcal{J}$ sources at the considered TF point are estimated by:

$$\widehat{\mathcal{S}}_{\tilde{\mathbf{s}}}(t,f) \approx \tilde{\mathbf{H}}_\alpha^{\#}(f) \mathcal{S}_{\mathbf{x}}(t,f), \tag{17}$$

where the superscript $(\cdot)^{\#}$ represents the Moore-Penrose's pseudo-inversion operator.
In the simulation, the optimization problem of (16) is solved using exhaustive search. This is computationally tractable for small vector array sizes but would be prohibitive if M is very large.

4. SIMULATION RESULTS

In the simulations, we have considered an array of $M = 3$ sensors, that receives signals from $N = 4$ independent speech sources. The filter coefficients are chosen randomly and the channel order

[1] To avoid the difficult problem of estimating the number of active sources at each TF point, we have chosen in this paper to set $\mathcal{J}$ to a fixed value in the range $1 < \mathcal{J} < M$.

is $K = 6$. The sample size is $T = 8192$ samples (corresponding approximately to 1 second recording of speech signals sampled at 8 KHz). The separation quality is measured by the normalized mean squares estimation errors (NMSE) of the sources evaluated over $N_r = 200$ Monte-Carlo runs and defined as:

$$\mathrm{NMSE}_i \stackrel{\mathrm{def}}{=} \frac{1}{N_r}\sum_{r=1}^{N_r} \min_{\alpha}\left(\frac{\|\alpha\widehat{\mathbf{s}}_{i,r} - \mathbf{s}_i\|^2}{\|\mathbf{s}_i\|^2}\right) \quad (18)$$

$$\mathrm{NMSE}_i = \frac{1}{N_r}\sum_{r=1}^{N_r} 1 - \left(\frac{\widehat{\mathbf{s}}_{i,r}\mathbf{s}_i^H}{\|\widehat{\mathbf{s}}_{i,r}\|\|\mathbf{s}_i\|}\right)^2 \quad (19)$$

$$\mathrm{NMSE} = \frac{1}{N}\sum_{i=1}^{N} \mathrm{NMSE}_i \,. \quad (20)$$

where $\mathbf{s}_i \stackrel{\mathrm{def}}{=} [s_i(0), \ldots, s_i(T-1)]$, $\widehat{\mathbf{s}}_{i,r}$ (defined similarly) represents the r^{th} estimate of source $\mathbf{s}_i$ and α is a scalar factor that compensates for the scale indeterminacy of the BSS problem. In Fig. 1, we compare the separation performance obtained by the subspace-based algorithm with $\mathcal{J} = 2$ and the cluster-based algorithm. It is observed that subspace-based algorithm provides much better separation results than those obtained by the cluster-based algorithm. This is mainly due to the high occurrence of overlapping sources in the TF domain for this type of signals so that the 'TF-disjointness' assumption used by the TF-CUBSS algorithm is poorly satisfied. The plot in Fig. 2 presents the separation performance of the subspace method when using the exact matrix $\mathbf{H}$ compared to that obtained with the proposed estimate $\widehat{\mathbf{H}}$. The observed performance loss is due to the channel estimation error which is relatively high for low SNRs and becomes negligible for high SNRs.

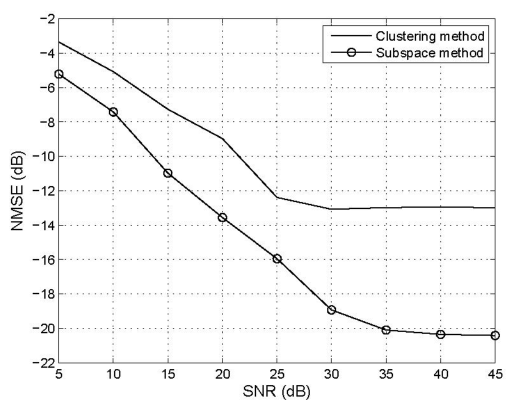

Fig. 1. Comparison between subspace-based and cluster-based TF-CUBSS algorithms: normalized MSE (NMSE) versus SNR for 4 speech sources and 3 sensors.

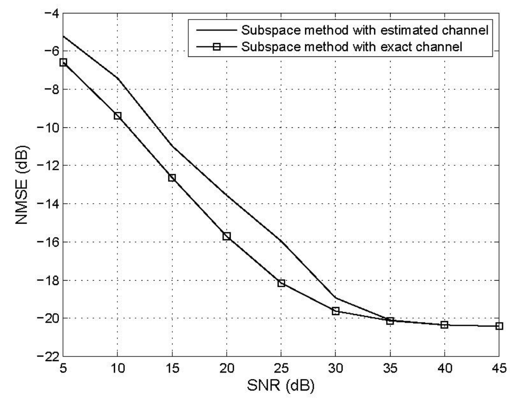

Fig. 2. Comparison, for the subspace-based TF-CUBSS algorithm, when the mixing channel $\mathbf{H}$ is known or unknown: NMSE of the source estimates.

5. CONCLUSION

This paper introduces new methods for the UBSS of TF-disjoint and TF-nondisjoint nonstationary sources in the convolutive mixture case using their time-frequency representations. The first proposed method has the advantage of simplicity while the second uses a weaker assumption on the source 'sparseness', i.e. the sources are not necessarily TF-disjoint, and proposes an explicit treatment of the overlapping points using subspace projection, leading to significant performance improvements.

6. REFERENCES

[1] P. Bofill and M. Zibulevsky, "Underdetermined blind source separation using sparse representations," *Signal Processing*, vol. 81, no. 11, pp. 2353–2362, Nov. 2001.

[2] N. Linh-Trung, A. Belouchrani, K. Abed-Meraim, and B. Boashash, "Separating more sources than sensors using time–frequency distributions," *EURASIP Journal Applied Signal Processing (also in Proc. ISSPA 2001)*, vol. 2005, no. 17, pp. 2828–2847, 2005.

[3] O. Yilmaz and S. Rickard, "Blind separation of speech mixtures via time-frequency masking," *IEEE Transaction on Signal Processing*, vol. 52, no. 7, pp. 1830–1847, July 2004.

[4] N. Linh-Trung, A. Aïssa-El-Bey, K. Abed-Meraim, and A. Belouchrani, "Underdetermined blind source separation of non-disjoint nonstationary sources in time-frequency domain," in *Proc. ISSPA*, Sydney, Australia, Aug. 2005, vol. 1, pp. 46–49.

[5] P. O'Grady, B. Pearlmutter, and S. Rickard, "Survey of sparse and non-sparse methods in source separation," *International Journal Imaging Sys. Tech.*, vol. 15, no. 1, pp. 18–33, 2005.

[6] J. Rosca, C. Borss, and R. Balan, "Generalized sparse signal mixing model and application to noisy blind source separation," in *Proc. ICASSP*, 2004, vol. 3, pp. 877–880.

[7] S. Araki, S. Makino, A. Blin, R. Mukai, and H. Sawada, "Blind separation of more speech than sensors with less distortion by combining sparseness and ICA," in *Proc. IWAENC*, Kyoto, Japan, Sept. 2003, pp. 271–274.

[8] Y. Huang, J. Benesty, and J. Chen, "A blind channel identification-based two-stage approach to separation and dereverberation of speech signals in a reverberant environment," *IEEE Transaction on Speech & Audio Processing*, vol. 13, pp. 882–895, Sept. 2005.

[9] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Transaction on Acoustic, Speech & Signal Processing*, vol. ASSP-33, pp. 387–392, Feb. 1985.

[10] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Transaction on Signal Processing*, vol. 43, pp. 2982–2993, Dec. 1995.

[11] R. Ahmad, A. W. H. Khong, and P. Naylor, "Proportionate frequency domain adaptive algorithms for blind channel identification," in *Proc. ICASSP*, Toulouse, France, May 2006, vol. 5, pp. 29–32.

[12] I. E. Frank and R. Todeschini, *The data analysis handbook*, Elsevier, Sci. Pub. Co., 1994.