Blind Spectral-GMM Estimation for Underdetermined Instantaneous Audio Source Separation

Simon Arberet[1], Alexey Ozerov[2,⋆], Rémi Gribonval[1], and Frédéric Bimbot[1]

[1] METISS Group, IRISA-INRIA
Campus de Beaulieu, 35042 Rennes Cedex, France
{simon.arberet,remi.gribonval,frederic.bimbot}@irisa.fr
[2] LTCI (TELECOM ParisTech & CNRS) - TSI Department
37-39 rue Dareau, 75014 Paris, France
alexey.ozerov@telecom-paristech.fr

Abstract. The underdetermined blind audio source separation problem is often addressed in the time-frequency domain by assuming that each time-frequency point is an independently distributed random variable. Other approaches which are not blind assume a more structured model, like the Spectral Gaussian Mixture Models (Spectral-GMMs), thus exploiting statistical diversity of audio sources in the separation process. However, in this last approach, Spectral-GMMs are supposed to be learned from some training signals. In this paper, we propose a new approach for learning Spectral-GMMs of the sources without the need of using training signals. The proposed blind method significantly outperforms state-of-the-art approaches on stereophonic instantaneous music mixtures.

1 Introduction

The problem of underdetermined Blind Source Separation (BSS) is to recover single-channel source signals $s_n(\tau)$, $1 \leq n \leq N$, from a multichannel mixture signal $x_m(\tau)$, $1 \leq m \leq M$, with $M < N$. Taking the Short Time Fourier Transform (STFT) $X_m(t,f)$ of each channel $x_m(t)$ of the mixture, the instantaneous mixing process is modeled in the time-frequency domain as:

$$\mathbf{X}(t,f) = \mathbf{A}\mathbf{S}(t,f) \tag{1}$$

where $\mathbf{X}(t,f)$ and $\mathbf{S}(t,f)$ denote respectively the column vectors $[X_m(t,f)]_{m=1}^{M}$ and $[S_n(t,f)]_{n=1}^{N}$, and $\mathbf{A}$ is the $M \times N$ real-valued mixing matrix.

The underdetermined BSS problem is often addressed in a two step approach where: first the mixing matrix is estimated, and then the source coefficients are estimated with the Maximum A Posteriori (MAP) criterion given a sparse source prior and the mixing matrix. Sources are then recovered using the inverse

⋆ A part of this work was done while A. Ozerov was with KTH, Stockholm, Sweden.

T. Adali et al. (Eds.): ICA 2009, LNCS 5441, pp. 751–758, 2009.

STFT. In the audio domain, sparse prior distributions are usually a Laplacian [1], a generalized Gaussian [2], a Student-t [3], or a mixture of two Gaussians [4].

This approach however suffers from the following issues:

1. In each time-frequency point, the maximum number of nonzero sources is usually assumed to be limited to the number M of channels [1,2].
2. The assumed nonzero sources are always the M neighboring directions which points toward the direction of the observed mixture [5].
3. Each time-frequency coefficient is estimated independently of the others without taking into account the structure of each source in the time-frequency domain. In other words, the signal redundancy and structure are not fully exploited.

In this paper we assume that $\mathbf{A}$ is known or has already been estimated [6], and the columns are pairwise linearly independent. Issues two and three have been addressed by the Statistically Sparse Decomposition Principle (SSDP)[5], which exploit the correlation between the mixture channels and more recently, the three issues have been addressed by the Local Gaussian Modeling (LGM) [7], where time-frequency coefficients are modeled via Gaussian priors with free variances. The third issue has been indeed partially addressed by SSDP and LGM, which exploit the neighborhood of the time-frequency points, in order to estimate the source distribution of the coefficients.

A more globally structured approach (to address these three issues) consists in assuming a spectral model of the sources via Spectral Gaussian Mixture Models (Spectral-GMMs) [8,9]. This approach has been successfully used to separate sources in the monophonic case ($M = 1$) [8,9], when sparse methods are unsuitable. However this approach is not blind because the models need to be learned from some training sources which should have characteristics similar to those of the sources to be separated. An EM algorithm could be used to learn GMMs directly from the mixture [10,11], but this approach suffers from two big issues. First, the number of Gaussians in the observation density grows exponentially with the number of sources, which often leads to an intractable algorithm. Second, the algorithm can be very slow and converges to a local maximum depending on the initial values.

In this paper, we propose a framework to blindly learn Spectral-GMMs with a linear complexity, provided that we have for each n-th source and for each time-frequency point (t, f) the following two estimates:

1. an estimate $\tilde{S}_n(t, f)$ of the source coefficient $S_n(t, f)$;
2. an estimate $\tilde{\sigma}^2_{n,t}(f)$ of the coefficient estimation squared error:

$$e^2_{n,t}(f) \triangleq \left| \tilde{S}_n(t, f) - S_n(t, f) \right|^2 . \tag{2}$$

The paper is organized as follows. In section 2, we describe the Spectral-GMM source estimation method assuming known models. In section 3, we recall the LGM source estimation method and show that, with this approach, we can also provide the two above-mentioned estimates required by the proposed framework.

In section 4, we describe our approach to blindly estimate Spectral-GMMs of the sources, with a linear complexity EM algorithm. Finally, we evaluate the performances of our approach on musical data in section 5.

2 Spectral-GMM Source Estimation

In this section, we briefly describe the principles of the Spectral-GMM source estimation methods [9], that we extend to the multichannel case. The short time Fourier spectrum $S(t) = [S(t,f)]_f$[1] of source S at time t is modeled by a multidimensional zero-mean complex valued $\mathcal{K}$-states GMM with probability density function (pdf) given by:

$$P(S(t)|\,\lambda) = \sum\nolimits_{k=1}^{\mathcal{K}} \pi_k\, N_c(S(t); \bar{0}, \Sigma_k) \tag{3}$$

where $N_c(S(t); \bar{0}, \Sigma_k) \triangleq \prod_f \frac{1}{\pi\sigma_k^2(f)} \exp\left[-\frac{|S(t,f)|^2}{\sigma_k^2(f)}\right]$, $\lambda \triangleq \{\pi_k, \Sigma_k\}_k$ is a Spectral-GMM of source S, π_k being a weight of Gaussian k of GMM λ, and $\Sigma_k \triangleq \mathrm{diag}([\sigma_k^2(f)]_f)$ is a diagonal covariance matrix of Gaussian k of GMM λ.

Provided that we know the Spectral-GMMs $\mathbf{\Lambda} = [\lambda_n]_{n=1}^N$ of the sources, the separation is performed in the STFT domain with the Minimum Mean Square Error (MMSE) estimator, which can be viewed as a form of adaptive Wiener filtering:

$$\widehat{\mathbf{S}}(t,f) = \sum_{\mathbf{k}} \gamma_{\mathbf{k}}(t) \mathbf{W}_{\mathbf{k}}(f) \mathbf{X}(t,f) \tag{4}$$

where $\mathbf{k} \triangleq [k_n]_{n=1}^N$, and $\gamma_{\mathbf{k}}(t)$ is the state probability at frame t ($\sum_{\mathbf{k}} \gamma_{\mathbf{k}}(t) = 1$):

$$\gamma_{\mathbf{k}}(t) \triangleq P(\mathbf{k}|\mathbf{X}(t); \mathbf{A}, \mathbf{\Lambda}) \propto \pi_{\mathbf{k}} \prod_f N_c\left(\mathbf{X}(t,f); \bar{0}, \mathbf{A}\mathbf{\Sigma}_{\mathbf{k}}(f)\mathbf{A}^T\right) \tag{5}$$

with $\mathbf{X}(t) \triangleq [\mathbf{X}(t,f)]_f$, $\mathbf{\Sigma}_{\mathbf{k}}(f) \triangleq \mathrm{diag}([\sigma^2_{n,k_n}(f)]_{n=1}^N)$, and the Wiener filter is given by:

$$\mathbf{W}_{\mathbf{k}}(f) \triangleq \mathbf{\Sigma}_{\mathbf{k}}(f)\mathbf{A}^T \left(\mathbf{A}\,\mathbf{\Sigma}_{\mathbf{k}}(f)\mathbf{A}^T\right)^{-1} \tag{6}$$

Thus, at each frame t, the source estimation is done in two steps:

1. *decoding step*, where the state probabilities $\gamma_{\mathbf{k}}(t)$ are calculated with equation (5);
2. *filtering step*, where the source coefficients are estimated by the weighted Wiener filtering of equation (4).

In such an approach the models λ_n are usually learned separately [8] by maximization of the likelihoods $P(\bar{S}_n|\,\lambda_n)$, where $\bar{S}_n$ is the STFT of the training signal for source s_n. This maximization is achieved via the Expectation Maximization (EM) algorithm [12] initialized by some clustering algorithm (e.g., K-means). As we will see in section 5, the performances of this method can be very good. However, it suffers from two big issues:

[1] The notation $[S(t,f)]_f$ means a column vector composed of elements $S(t,f), \forall f$.

– The approach requires availability of training signals, that are difficult to obtain in most realistic situations [9].
– As the mixture state $\mathbf{k}$ is a combination of all the source states $k_n, 1 \leq n \leq N$, the decoding step of equation (5) is of complexity $O(\mathcal{K}^N)$.

3 LGM Source Estimation

The LGM [7] is a method which consists in estimating local source variances $\mathbf{v}(t,f) = [\sigma_n^2(t,f)]_{n=1}^N$ in each time-frequency point and then estimating the source coefficients with the MMSE estimator given by the Wiener filter:

$$\tilde{\mathbf{S}}(t,f) = \mathbf{W}(t,f)\mathbf{X}(t,f) \tag{7}$$

where $\mathbf{W}(t,f) = \widehat{\mathbf{\Sigma}}(t,f)\mathbf{A}^T \left(\mathbf{A}\widehat{\mathbf{\Sigma}}(t,f)\mathbf{A}^T\right)^{-1}$, and $\widehat{\mathbf{\Sigma}}(t,f)$ is a diagonal matrix whose entries are the estimated source variances: $\widehat{\mathbf{\Sigma}}(t,f) = \mathrm{diag}\left(\widehat{\mathbf{v}}(t,f)\right)$.

The LGM is based on the empirical local covariance matrix in the time-frequency domain, which has already been used by mixing matrix estimation methods [6,13] so as to select time-frequency regions where only one source is supposed active, and which is defined by:

$$\widehat{\mathbf{R}}_x(t,f) = \sum_{t',f'} w(t-t', f-f')\mathbf{X}(t',f')\mathbf{X}^H(t',f') \tag{8}$$

where w is a bi-dimensional normalized window function which defines the neighborhood shape, and H denotes the conjugate transpose of a matrix. If we assume that each complex source coefficient $S_n(t,f)$ in a time-frequency neighborhood follows an independent (over time t and frequency f) zero-mean Gaussian distribution with variance $\sigma_n^2(t,f)$, then the mixture coefficients in that neighborhood follow a zero-mean Gaussian distribution with covariance matrix:

$$\mathbf{R}_x(t,f) = \mathbf{A}\mathbf{\Sigma}(t,f)\mathbf{A}^T. \tag{9}$$

The Maximum Likelihood (ML) estimate of source variances σ_n^2 (we drop the (t,f) index for simplicity) is obtained by minimization of the Kullback-Leibler (KL) divergence between the empirical and the mixture model covariances [14]:

$$\widehat{\mathbf{\Sigma}} = \underset{\mathbf{\Sigma}=\mathrm{diag}(\mathbf{v}),\mathbf{v}\geq 0}{\arg\min} KL(\widehat{\mathbf{R}}_x|\mathbf{R}_x), \tag{10}$$

where $KL(\widehat{\mathbf{R}}_x|\mathbf{R}_x)$ is defined as:

$$KL(\widehat{\mathbf{R}}_x|\mathbf{R}_x) = \frac{1}{2}\left(\mathrm{tr}\left(\widehat{\mathbf{R}}_x\mathbf{R}_x^{-1}\right) - \log\det\left(\widehat{\mathbf{R}}_x\mathbf{R}_x^{-1}\right) - M\right). \tag{11}$$

The LGM method [7] uses a global non-iterative optimization algorithm to solve the problem of equation (10), which roughly consists of estimating variances $\mathbf{v}$ by solving the linear system $\widehat{\mathbf{R}}_x \approx \mathbf{R}_x$ with a sparsity assumption on the variance vector $\mathbf{v}$.

The MMSE estimate of $(\tilde{\mathbf{S}}-\mathbf{S})(\tilde{\mathbf{S}}-\mathbf{S})^H$ is given by the covariance matrix of $\mathbf{S}$ given $\mathbf{X}$:

$$\mathbf{C} \stackrel{\triangle}{=} \mathbb{E}\left[(\tilde{\mathbf{S}}-\mathbf{S})(\tilde{\mathbf{S}}-\mathbf{S})^H \middle| \mathbf{X}\right] = (\mathbf{I}-\mathbf{WA})\widehat{\mathbf{\Sigma}}, \tag{12}$$

where $\mathbf{W}$ is defined just after equation (7).

So the MMSE estimate of $e^2_{n,t}(f)$ defined by equation (2) is given by the corresponding diagonal element of matrix $\mathbf{C}(t,f)$:

$$\tilde{\sigma}^2_{n,t}(f) = \mathbb{E}\left[\left|\tilde{S}_n(t,f)-S_n(t,f)\right|^2 \middle| \mathbf{X}\right] = \mathbf{C}(t,f)_{n,n}. \tag{13}$$

4 Spectral-GMM Blind Learning Framework

The aim of the proposed framework is to learn the Spectral-GMM λ_n for each source s_n, provided that at each time-frequency point (t,f), we have an estimate $\tilde{S}_n(t,f)$ of the source coefficient $S_n(t,f)$ together with an estimate $\tilde{\sigma}^2_{n,t}(f)$ of the squared error $e^2_{n,t}(f)$ defined by equation (2).

The learning step is done for each source independently, so in the following we drop the source's index n for simplicity. Let us denote the error of source estimation as $\tilde{E}(t,f) \stackrel{\triangle}{=} \tilde{S}(t,f) - S(t,f)$. Now we assume that $\tilde{E}(t) = [\tilde{E}(t,f)]_f$ is a realization of a Gaussian complex vector with zero mean and a diagonal covariance matrix $\tilde{\Sigma}_t = \mathrm{diag}([\tilde{\sigma}^2_t(f)]_f)$, i.e., $P(\tilde{E}(t)|\tilde{\Sigma}_t) = N_c(\tilde{E}(t); \bar{0}, \tilde{\Sigma}_t)$. The relation:

$$\tilde{S}(t,f) = S(t,f) + \tilde{E}(t,f) \tag{14}$$

can be interpreted as a single sensor source separation problem with mixture $\tilde{S}$ and sources S and $\tilde{E}$, where source $\tilde{E}$ is modeled by $\tilde{\mathbf{\Sigma}} = [\tilde{\Sigma}_t]_{t=1}^T$, which is fixed, and source S is modeled by GMM $\lambda = \{\pi_k, \Sigma_k\}_{k=1}^{\mathcal{K}}$ that we want to estimate in the ML sense, given the observed mixture $\tilde{S}$ and fixed model $\tilde{\mathbf{\Sigma}}$. Thus, we are looking for λ optimizing the following ML criterion:

$$\lambda = \arg\max_{\lambda'} p(\tilde{S}|\lambda', \tilde{\mathbf{\Sigma}}). \tag{15}$$

Algorithm 1 summarizes an Expectation-Maximization (EM) algorithm for optimization of criterion (15) (see [9] for derivation). Initialization is done by applying K-means clustering algorithm to the source estimate $\tilde{S}$.

Once we have learned the source models $\mathbf{\Lambda} = [\lambda_n]_{n=1}^N$, we could estimate the sources with the procedure of section 2, but in that case the decoding step at each frame t will still be of complexity $O(\mathcal{K}^N)$. In order to have a linear complexity method, we do not calculate all the $\mathcal{K}^N$ mixture state probabilities, but only the $\mathcal{K}$ state probabilities of each source using $\gamma^{(L+1)}_{k_n}(t)$, where $\gamma^{(L)}_{k_n}(t)$ (defined by equation (16) in algorithm 1) are the state probabilities of source S_n calculated during the last iteration of algorithm 1. The source coefficients are then estimated with the Wiener filter $\mathbf{W}_{\mathbf{k}^*(t)}(f)$ of equation (6), with $\mathbf{k}^*(t) = [k^*_n(t)]_{n=1}^N$ and where $k^*_n(t) = \arg\max_k P(q(t)=k|\tilde{S}_n, \lambda_n^{(L+1)}, \tilde{\mathbf{\Sigma}}_n)$ is the most likely state of model $\lambda_n^{(L+1)}$ at frame t, given $\tilde{S}_n(t,f)$ and $\tilde{\mathbf{\Sigma}}_n$.

Algorithm 1. EM Algorithm for source Spectral-GMM estimation in ML sense (index (l) in power denotes the parameters estimated at the l^{th} iteration of the algorithm)

1. Compute the weights $\gamma_k^{(l)}(t)$ satisfying $\sum_k \gamma_k^{(l)}(t) = 1$ and

$$\gamma_k^{(l)}(t) \triangleq P(q(t) = k|\tilde{S}, \lambda^{(l)}, \tilde{\mathbf{\Sigma}}) \propto \pi_k^{(l)} N_c(\tilde{S}(t); \bar{0}, \Sigma_k^{(l)} + \tilde{\Sigma}_t) \tag{16}$$

   where $q(t)$ is the current state of GMM λ at frame t.
2. Compute the expected Power Spectral Density (PSD) for state k

$$\left\langle |S(t,f)|^2 \right\rangle_k^{(l)} \triangleq \mathbb{E}_S \left[|S(t,f)|^2 \left| q(t) = k, \tilde{S}, \lambda^{(l)}, \tilde{\mathbf{\Sigma}} \right. \right] =$$

$$= \frac{\sigma_k^{2,(l)}(f)\tilde{\sigma}_t^2(f)}{\sigma_k^{2,(l)}(f) + \tilde{\sigma}_t^2(f)} + \left| \frac{\sigma_k^{2,(l)}(f)}{\sigma_k^{2,(l)}(f) + \tilde{\sigma}_t^2(f)} \tilde{S}(t,f) \right|^2 \tag{17}$$

3. Re-estimate Gaussian weights

$$\pi_k^{(l+1)} = \frac{1}{T} \sum_t \gamma_k^{(l)}(t) \tag{18}$$

4. Re-estimate covariance matrices

$$\sigma_k^{2,(l+1)}(f) = \frac{\sum_t \left\langle |S(t,f)|^2 \right\rangle_k^{(l)} \gamma_k^{(l)}(t)}{\sum_t \gamma_k^{(l)}(t)} \tag{19}$$

5 Experimental Results

We evaluate our method[2] over music mixtures, with the number of sources N varying from 3 to 6. For each N a mixing matrix was computed as described in [15], given an angle of $50 - 5N$ degrees between successive sources, and ten instantaneous mixtures were generated from different source signals of duration 10 s, sampled at 22.05 kHz. The STFT was computed with a sine window of length 2048 (93 ms). The performance measure used is the Signal-to-Distortion Ratio (SDR) defined in [16]. The bi-dimensional window w defining time-frequency neighborhoods of the LGM method was the outer product of two Hanning windows with length 3 as in [7]. The Spectral-GMMs were learned with 30 iterations of algorithm 1 using LGM parameters given by equations (7) and (13) as entries, and the number $\mathcal{K}$ of states per GMM was chosen equal to 8, because it yielded the best results in SDR. Figure 1 compares the average SDR achieved by the proposed Spectral-GMM method, the LGM method presented in Section 3 and the classical DUET [17]. The proposed algorithm outperforms DUET by more than 5 dB in the 3 sources case and LGM by at least 2 dB whatever the number of sources. We also plotted the performance of the (oracle) Spectral-GMM separation when models, with the same number $\mathcal{K}$ of states, are learned and

[2] This method was also submitted to the 2008 Signal Separation Evaluation Campaign.

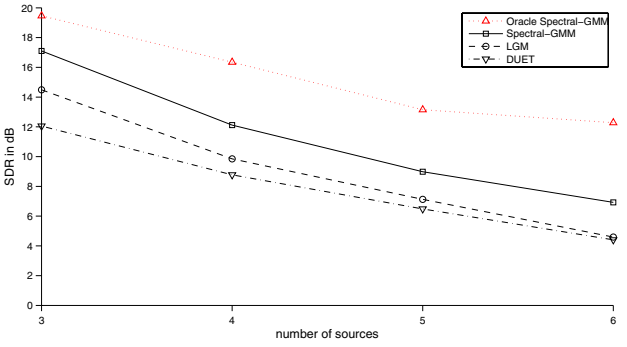

Fig. 1. Source separation performance over stereo instantaneous musical mixtures. STFT window length is 2048 (93 ms) and $\mathcal{K} = 8$.

decoded using the original sources. We can notice that the performance of the proposed method remains between 2 dB and 5 dB below the oracle performance, and that this gap increases with the number of sources, showing the difficulty to blindly learn Spectral-GMM when the number of sources is high. As for the computational load, the MATLAB implementation of the proposed algorithm on a 3.4 GHz CPU runs in 133 s in the 4 sources case, while it runs in 120 s for LGM and in 2 s for DUET.

6 Conclusion

In this paper, we proposed a new framework for the blind audio source separation problem in the multichannel instantaneous mixture case. In this framework Spectral-GMM models of sources were blindly learned, i.e. without using any other informations than the mixture and the mixing matrix, with an EM algorithm having a linear $O(N\mathcal{K})$ complexity, in contrast to some related state-of-the-art methods having an exponential $O(\mathcal{K}^N)$ complexity. As opposed to the other blind audio source separation methods, the proposed method exploits the structure of each source in the time-frequency domain. The proposed method outperforms the state-of-the-art methods tested by between 2 dB and 5 dB in SDR. Further work include an extension of the method to the anechoic and convolutive cases, evaluation of the robustness of the method by using mixing matrices which are not perfectly estimated, and improvement of the method to fill the gap between the blindly learned models and the oracle ones.

References

1. Zibulevsky, M., Pearlmutter, B.A., Bofill, P., Kisilev, P.: Blind source separation by sparse decomposition in a signal dictionary. In: Independent Component Analysis: Principles and Practice, pp. 181–208. Cambridge Press (2001)
2. Vincent, E.: Complex nonconvex l_p norm minimization for underdetermined source separation. In: Davies, M.E., James, C.J., Abdallah, S.A., Plumbley, M.D. (eds.) ICA 2007. LNCS, vol. 4666, pp. 430–437. Springer, Heidelberg (2007)

3. Fevotte, C., Godsill, S.: A bayesian approach for blind separation of sparse sources. IEEE Transactions on Audio, Speech, and Language Processing 14(6), 2174–2188 (2006)
4. Davies, M.E., Mitianoudis, N.: Simple mixture model for sparse overcomplete ICA. IEE Proceedings on Vision, Image and Signal Processing 151(1), 35–43 (2004)
5. Xiao, M., Xie, S., Fu, Y.: A statistically sparse decomposition principle for underdetermined blind source separation. In: Proc. Int. Symp. on Intelligent Signal Processing and Communication Systems (ISPACS), pp. 165–168 (2005)
6. Arberet, S., Gribonval, R., Bimbot, F.: A robust method to count and locate audio sources in a stereophonic linear instantaneous mixture. In: Rosca, J.P., Erdogmus, D., Príncipe, J.C., Haykin, S. (eds.) ICA 2006. LNCS, vol. 3889, pp. 536–543. Springer, Heidelberg (2006)
7. Vincent, E., Arberet, S., Gribonval, R.: Underdetermined audio source separation via gaussian local modeling. In: Proc. Int. Conf. on Independent Component Analysis and Blind Source Separation (ICA) (2009)
8. Benaroya, L., Bimbot, F.: Wiener based source separation with HMM/GMM using a single sensor. In: Proc. ICA, pp. 957–961 (2003)
9. Ozerov, A., Philippe, P., Bimbot, F., Gribonval, R.: Adaptation of bayesian models for single-channel source separation and its application to voice/music separation in popular songs. IEEE Transactions on Audio, Speech and Language Processing 15(5), 1564–1578 (2007); see also: IEEE Transactions on Speech and Audio Processing
10. Moulines, E., Cardoso, J.F., Gassiat, E.: Maximum likelihood for blind separation and deconvolution of noisy signals using mixture models. In: IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 5, pp. 3617–3620 (April 1997)
11. Attias, H.: Independent factor analysis. Neural Comput. 11(4), 803–851 (1999)
12. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the em algorithm
13. Arberet, S., Gribonval, R., Bimbot, F.: A robust method to count and locate audio sources in a stereophonic linear anechoic mixture. In: ICASSP 2007, vol. 3, pp. 745–748 (April 2007)
14. Pham, D.T., Cardoso, J.F.: Blind separation of instantaneous mixtures of non stationary sources. IEEE Trans. on Signal Processing 49(9), 1837–1848 (2001)
15. Pulkki, V., Karjalainen, M.: Localization of amplitude-panned virtual sources I: stereophonic panning. Journal of the Audio Engineering Society 49(9), 739–752 (2001)
16. Vincent, E., Gribonval, R., Févotte, C.: Performance measurement in blind audio source separation. IEEE Trans. on Audio, Speech and Language Processing 14(4), 1462–1469 (2006)
17. Yılmaz, O., Rickard, S.T.: Blind separation of speech mixtures via time-frequency masking. IEEE Trans. on Signal Processing 52(7), 1830–1847 (2004)