A FREQUENCY METHOD FOR BLIND SEPARATION OF AN ANECHOIC MIXTURE

Wendyam S. B. Ouedraogo[1], *Barbara Nicolas*[1], *Benoit Oudompheng*[1,2], *Jérôme I. Mars*[1], *Christian Jutten*[1]

[1]GIPSA-lab, UMR 5216 CNRS, 11 rue des mathématiques, BP 46, F-38402 Grenoble Cedex, France
[2]MicrodB, 28 Chemin du Petit Bois, 69134 Écully Cedex, France

ABSTRACT

This paper presents a new frequency method for blind separation of mixtures of scaled and delayed versions of sources. This kind of problem can occur in air and underwater acoustics. By assuming the mutual independence of the sources, we make use of the power spectral densities and the cross power spectral densities of mixed data to estimate the sources, the mixing coefficients, and the relative delays between a reference sensor and the other sensors. Simulations on synthetic data of sound radiated by a ship show the effectiveness of the proposed method.

***Index Terms*—** Blind source separation, Anechoic mixture, Relative delays, Power spectral density, Ship noise, Underwater acoustics.

1. INTRODUCTION

We deal with the problem of source separation when the mixed data are linear combinations of delayed versions of the sources. Such a model is known as an anechoic mixture [1] or a pure delays mixture [2], and this can occur in many areas, such as biomedical signal processing [3], and acoustic source demixing in nonreverberant environments [4]. The mixing model is given by (1):

$$x_k(t) = \sum_{l=1}^{L} a_{kl} s_l \left(t - \tau_{kl}\right) + b_k(t). \tag{1}$$

where x_k, $1 \leq k \leq K$, is the signal recorded at sensor k, s_l, $1 \leq l \leq L$, is the l-th source, a_{kl} is the mixing coefficient of source l at sensor k ($a_{kl} \geq 0$), τ_{kl} is the propagation delay between source l and sensor k ($\tau_{kl} \geq 0$), and b_k is additive noise on sensor k. The task is to estimate the source, the mixing coefficients, and the delays.
Many algorithms have been proposed for solving this problem (1), including temporal [3, 4], frequency [2], and time-frequency [5, 6] methods. By assuming that the sources are bandlimited and the delays are small (typically, $\forall\ k, l$ $\tau_{kl} << \frac{1}{\sqrt{2}\pi f_{max}}$ where f_{max} is the maximal frequency of the sources) [4], and by making use of the Taylor expansion of the exponential function, the temporal methods model the mixed data by an instantaneous linear combination of the sources and their derivatives. The new problem is solved by joint diagonalization of a set of matrices that are obtained from the autocovariance matrix of whitened mixed data [3,4]. Unfortunately, temporal methods require a number of sensors at least twice that of the number of sources, which is a limitation in the context of a small number of sensors. In addition, existing temporal methods do not explicitly take into account that the mixing matrix can be ill-conditioned.
Yeredor proposed a frequency method to solve (1), which consists of jointly diagonalizing the matrices of the power spectral densities (PSDs) and cross PSDs of mixed data, for the different frequencies [2]. However, his method is limited to the case where $K = L = 2$, and its extension to $L > 2$ is not obvious. Nion et al. proposed a time-frequency method that uses the short-time Fourier transform of mixed data and alternating least-squares estimation, coupled with enforcing a Vandermonde structure of the estimated mixing matrix across its frequency mode [5]. This method also does not take into account the case where the mixing matrix is ill-conditioned. By using the Wigner-Ville distributions Omlor and Giese proposed a two-step time-frequency method [6], which starts by estimating the mixing coefficients and the modulus of the source Fourier transforms, followed by iterative estimation of the delays and phases of the source Fourier transforms. We were inspired by this approach to design the the method we proposed here.

This paper presents a three-step frequency method for solving (1). In the first step, we estimate the mixing coefficients and the modulus of the Fourier transforms of the sources. In the second step, we estimate the relative delays between a reference sensor and the other sensors. Finally, in the third step, the phases of the Fourier transforms of the sources and the source temporal profiles are estimated. Our method has the same steps as that proposed in [6]. The main differences lie in the estimation of the delays and the phases of the source Fourier transforms, where we propose a noniterative method, and we explicitly take into account the mixing matrix conditioning. In addition, we do not use Wigner-Ville distributions.

This work was supported by the French DGA - MRIS (grant RAPID ARMADA).

SigmaPhy team from Gipsa-lab is part of Labex OSUG@2020 (ANR10 LABX56).

This paper is organized as follows: section 2 describes the three steps of the proposed method. In section 3, we show simulation results on synthetic data, while section 4 derives the conclusions and provides direction for future studies.

2. PROPOSED METHOD

For designing our method, we assume that:

- (H1): The sources are zero-mean random variables that are mutually independent and stationary.
- (H2): The number of sensors is greater than the number of sources: $K > L$.
- (H3): The noise components are Gaussian with zero-mean, mutually independent, and independent of sources.

2.1. Estimation of the mixing coefficients and modulus of the source Fourier transforms

The mixed data autocorrelation functions are defined by (2), where E is the expectation operator.

$$R_{x_k}(u) = E\left[x_k(t)x_k(t+u)\right]. \quad (2)$$

Given (1), and because of assumptions (H1) and (H2), it can easily be verified that:

$$R_{x_k}(u) = \sum_{l=1}^{L} a_{kl}^2 R_{s_l}(u) + R_{b_k}(u). \quad (3)$$

where $R_{s_l}(u)$ and $R_{b_k}(u)$ are the autocorrelation functions of the sources and the noise, respectively. By taking the Fourier transform of (3), we obtain (4):

$$\Gamma_{x_k}(f) = \sum_{l=1}^{L} a_{kl}^2 \Gamma_{s_l}(f) + \Gamma_{b_k}(f). \quad (4)$$

where $\Gamma_{s_l}(f)$, $\Gamma_{x_k}(f)$ and $\Gamma_{b_k}(f)$ are the PSDs of the sources, the mixed data and the noise at sensor k, respectively. Let us assume that there is access to frames of recorded data where the sources are not active, we can then compute an estimate of the PSD of b_k, $\hat{\Gamma}_{b_k}(f)$, and deduce the denoised PSD of x_k through (5).

$$\begin{aligned}\tilde{\Gamma}_{x_k}(f) &= \Gamma_{x_k}(f) - \hat{\Gamma}_{b_k}(f)\\ &\approx \sum_{l=1}^{L} a_{kl}^2 \Gamma_{s_l}(f).\end{aligned} \quad (5)$$

Note that $a_{kl}^2 > 0$ and $\Gamma_{s_l}(f) \geq 0$, so to make sure that $\tilde{\Gamma}_{x_k}(f) \geq 0$, we set $\tilde{\Gamma}_{x_k}(f) = max\left(\tilde{\Gamma}_{x_k}(f), 0\right)$.
In the remaining part of this section, we will use the denoised PSDs of mixed data that we group in vector $\tilde{\Gamma}_{\mathbf{x}}(f)$ through (6). We also group the PSDs of sources in vector $\Gamma_{\mathbf{s}}(f)$ through (7).

$$\tilde{\Gamma}_{\mathbf{x}}(f) = [\tilde{\Gamma}_{x_1}(f), \tilde{\Gamma}_{x_2}(f), \cdots, \tilde{\Gamma}_{x_K}(f)]^T. \quad (6)$$

$$\Gamma_{\mathbf{s}}(f) = [\Gamma_{s_1}(f), \Gamma_{s_2}(f), \cdots, \Gamma_{s_L}(f)]^T. \quad (7)$$

Equation (5) can be rewritten in a compact form as (8):

$$\tilde{\Gamma}_{\mathbf{x}}(f) \approx \mathbf{H}\Gamma_{\mathbf{s}}(f). \quad (8)$$

where $\mathbf{H}$ is a $K \times L$-size matrix where the kl-th entry is given by $h_{kl} = a_{kl}^2$. Equation (8) shows that the denoised PSDs of the mixed data are a linear instantaneous combination of the source PSDs. We can then estimate the matrix $\mathbf{H}$ and the source PSDs, using a classical algorithm of independent components analysis [7] or nonnegative matrix factorization [8]. The estimated components $\hat{h}_{kl}$ and $\hat{\Gamma}_{s_l}(f)$ are scaled versions of h_{kl} and $\Gamma_{s_l}(f)$ [7, 8]. It is possible to estimate the scaled versions of the modulus of the source Fourier transforms, and the mixing coefficients, through (9):

$$\hat{a}_{kl} = \sqrt{\hat{h}_{kl}} \text{ and } |\hat{S}_l(f)| = \sqrt{\hat{\Gamma}_{s_l(f)}}. \quad (9)$$

2.2. Estimation of the relative delays between a reference sensor and the other sensors

Let us set a reference sensor r. For $k \neq r$, the cross-correlation function between x_k and x_r is defined by:

$$R_{x_k x_r}(u) = E\left[x_k(t)x_r(t+u)\right]. \quad (10)$$

Using the working hypothesis, the cross-correlation function is given by:

$$R_{x_k x_r}(u) = \sum_{l=1}^{L} a_{kl} a_{rl} R_{s_l}(u + \tau_{kl} - \tau_{rl}). \quad (11)$$

Note that in $R_{x_k x_r}(u)$ we do not have a term that is related to the noise, as the noise components are assumed to be mutually independent and independent of the sources.
The Fourier transform of (11) leads to the relationship between the cross PSDs of the mixed data and the PSDs of the sources, as follows:

$$\Gamma_{x_k x_r}(f) = \sum_{l=1}^{L} a_{kl} a_{rl} \Gamma_{s_l}(f) e^{\imath 2\pi f[\tau_{kl} - \tau_{rl}]}. \quad (12)$$

We define the relative delays $\zeta_{kl} = \tau_{kl} - \tau_{rl}$ and the set of frequencies $\mathcal{F}$ by:

$$\mathcal{F} = \{f, \sqrt{2}\pi f \zeta_{kl} \ll 1\}. \quad (13)$$

Given the propagation velocity c, and the distance between the sensor k and the reference sensor r, d_{kr}, a frequency $f_l \in \mathcal{F}$ if

$$f_l \ll \frac{c}{\sqrt{2}\pi d_{kr}}. \quad (14)$$

For a frequency $f \in \mathcal{F}$, we can approximate $e^{\imath 2\pi f \zeta_{kl}}$ by its first-order Taylor expansion [4], such that:

$$e^{\imath 2\pi f \zeta_{kl}} \approx 1 + \imath 2\pi f \zeta_{kl}. \quad (15)$$

Equation (12) can be rewritten as (16):

$$\Gamma_{x_k x_r}(f) \approx \sum_{l=1}^{L} a_{kl} a_{rl} \Gamma_{s_l}(f) \left[1 + \imath 2\pi f \zeta_{kj}\right] \quad (16)$$

and we deduce that:

$$\frac{\Im\left[\Gamma_{x_k x_r}(f)\right]}{2\pi f} \approx \sum_{l=1}^{L} a_{kl} a_{rl} \Gamma_{s_l}(f) \zeta_{kl} \quad (17)$$

where $\Im\left[\Gamma_{x_k x_r}(f)\right]$ is the imaginary part of $\Gamma_{x_k x_r}(f)$. Let us choose L frequencies $f_1, f_2, \cdots, f_L$ that belong in $\mathcal{F}$, and set the vector $\Upsilon_{kr} = \left[\frac{\Im\left[\Gamma_{x_k x_r}(f_1)\right]}{2\pi f_1}, \cdots, \frac{\Im\left[\Gamma_{x_k x_r}(f_L)\right]}{2\pi f_L}\right]^T$, the vector $\zeta_k = \left[\zeta_{k1}, \cdots, \zeta_{kL}\right]^T$, and $\mathbf{U}_{kr}$ the $L \times L$ size matrix where the il-th entry is defined by $\left[\mathbf{U}_{kr}\right]_{il} = a_{kl} a_{rl} \Gamma_{s_l}(f_i)$, then:

$$\Upsilon_{kr} = \mathbf{U}_{kr} \zeta_k. \quad (18)$$

An estimate of $\mathbf{U}_{kr}$ is given by $\left[\hat{\mathbf{U}}_{kr}\right]_{il} = \hat{a}_{kl} \hat{a}_{rl} \hat{\Gamma}_{s_l}(f_i)$. Since $\left[\hat{\mathbf{U}}_{kr}\right]$ is square and well-conditionned, one can compute $\hat{\zeta}_k$ by:

$$\hat{\zeta}_k = \left[\hat{\mathbf{U}}_{kr}\right]^{-1} \Upsilon_{kr}. \quad (19)$$

It should be noted that when the purpose is localization, an estimation of the relative delays is generally sufficient.

2.3. Estimation of the phases of the Fourier transforms of the sources, and of their temporal profiles

The last step of our method is the estimation of the phases of the Fourier transforms of the sources, to compute their temporal profiles. For a fixed reference sensor r, we define y_l, a delayed version of s_l, through (20), where τ_{rl} is the delay in the arrival of the source s_l at sensor r:

$$y_l(t) = s_l(t + \tau_{rl}). \quad (20)$$

We seek here to estimate $y_l(t)$. Let $Y_l(f)$, respectively $S_l(f)$, be the Fourier transform of $y_l(t)$, respectively $s_l(t)$. We set $Y_l(f) = |Y_l(f)| e^{\imath \varphi_{y_l}(f)}$ and $S_l(f) = |S_l(f)| e^{\imath \varphi_{s_l}(f)}$, where $|Y_l(f)|$ and $\varphi_{y_l}(f)$ (respectively $|S_l(f)|$ and $\varphi_{s_l}(f)$) are the modulus and the phase, respectively, of the Fourier transform of $y_l(t)$ (respectively $s_l(t)$), at frequency f. Then:

$$|Y_l(f)| = |S_l(f)| \text{ and } \varphi_{y_l}(f) = \varphi_{s_l}(f) + 2\pi f \tau_{rl}. \quad (21)$$

When replacing s_l by y_l in (1), this gets:

$$\begin{aligned} x_k(t) &= \sum_{l=1}^{L} a_{kl} y_l\left(t + \tau_{rl} - \tau_{kl}\right) + b_k(t) \\ &= \sum_{l=1}^{L} a_{kl} y_l\left(t - \zeta_{kl}\right) + b_k(t). \end{aligned} \quad (22)$$

Fourier transforming (22) leads to the relationship between the Fourier transforms of mixed data x_k and the Fourier transforms of the delayed sources y_l, as given by (23):

$$\begin{aligned} X_k(f) &= \sum_{l=1}^{L} a_{kl} Y_l(f) e^{-\imath 2\pi f \zeta_{kl}} + B_k(f) \\ &= \sum_{l=1}^{L} a_{kl} |Y_l(f)| e^{\imath \varphi_{y_l}(f)} e^{-\imath 2\pi f \zeta_{kl}} + B_k(f). \end{aligned} \quad (23)$$

Equation (23) can also be expressed as:

$$\mathbf{X}(f) = \mathbf{M}\Phi(f) + \mathbf{B}(f) \quad (24)$$

where the vector $\mathbf{X}(f) = [X_1(f), X_2(f), \cdots, X_K(f)]^T$, the vector $\mathbf{B}(f) = [B_1(f), B_2(f), \cdots, B_K(f)]^T$, and the vector $\Phi(f) = [e^{\imath \varphi_{y_1}(f)}, e^{\imath \varphi_{y_2}(f)}, \cdots, e^{\imath \varphi_{y_L}(f)}]^T$. The matrix $\mathbf{M}$ is of size $K \times L$, and the kl-th entry of $\mathbf{M}$ is defined by:

$$m_{kl} = a_{kl} |Y_l(f)| e^{-\imath 2\pi f \zeta_{kl}}. \quad (25)$$

The entry m_{kl} can be estimated using (9), (19) and (21). Finally, the source Fourier transform phases can be estimated by solving the optimization of Equation (26).

$$\hat{\Phi}(f) = \arg\min_{\Phi(f)} \|\mathbf{X}(f) - \mathbf{M}\Phi(f)\|_2^2 + \lambda \|\Phi(f)\|_2^2. \quad (26)$$

The Tikhonov regularization term is added in this optimization, because the matrix $\mathbf{M}$ can be ill-conditionned, especially when the sensors are close to each other. The regularization parameter λ can be computed by the L-curve method [9] [10]. After estimation of the phases of the source Fourier transforms, at each frequency, we can estimate the temporal profile of $y_l(t)$ through (27), where $|\hat{Y}_l(f)| = |\hat{S}_l(f)|$.

$$\hat{y}_l(t) = IFFT\left[|\hat{Y}_l(f)| e^{\imath \hat{\varphi}_{y_l}(f)}\right] \quad (27)$$

It should be noted that each source is estimated up to an unknown scale and delay, and only the relative delays between the reference sensor and other sensors are estimated. This is not a fundamental limitation, as the unknown scale and delay are the intrinsic indeterminates related to the demixing of a convolutive mixture. Indeed, our mixture model given (1) is a special case of a convolutive mixture.

3. SIMULATION RESULTS

This section presents the simulation results on synthetic data of ship noise sources in underwater acoustics. The proposed method is compared to the temporal method proposed by [3] and to the time-frequency method proposed by [5].

The efficiency of the source estimation is quantified by the signal-to-interference ratio (SIR) which is defined in [11]. To compute the SIR, the estimated source $\hat{s}_l$ is decomposed as:

$$\hat{s}_l = s_{target} + e_{interf} + e_{noise} + e_{artifact}. \quad (28)$$

where s_{target} is a version of s_l modified by an allowed distortion, and where e_{interf}, e_{noise}, and $e_{artifact}$ are the interference, noise, and artifact error terms, respectively. The reader is referred to [11] for more details of the previous decomposition. The SIR for estimated source l is computed through (29). The larger the SIR, the better the separation, and it can be assumed that a source is correctly estimated if $SIR_l(dB) \geq 15$.

$$SIR_l = 10 \log_{10} \frac{\|s_{target}\|^2}{\|e_{interf}\|^2}. \tag{29}$$

The accuracy of the estimated relative delays of the source l is quantified using the normalized mean error (NME), as defined by (30). The smaller the NME, the better the estimation of the relative delays.

$$NME_l = \frac{\frac{1}{K}\sum_{k=1}^{K} |\hat{\zeta}_{kl} - \zeta_{kl}|}{\frac{1}{K}\sum_{k=1}^{K} |\zeta_{kl}|}. \tag{30}$$

Synthesis of mixed data

We consider the scenario where two omnidirectional point acoustic sources that are located in differents positions radiate in an underwater environment. The acoustic waves radiated by these sources propagate to a linear antenna of five hydrophones. The distance between two consecutive hydrophones is 50 cm. The not-to-scale Figure 1 illustrates the simulated scenario. The planes of the sources and the sensors are spaced at 10 m. The mixed data are generated by

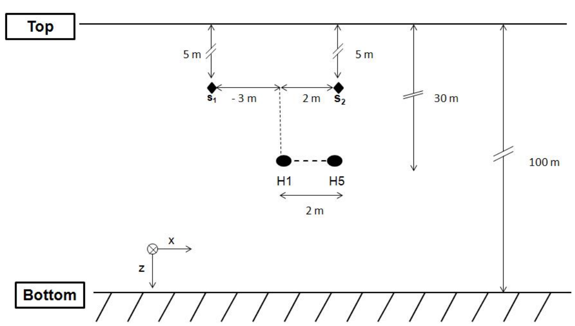
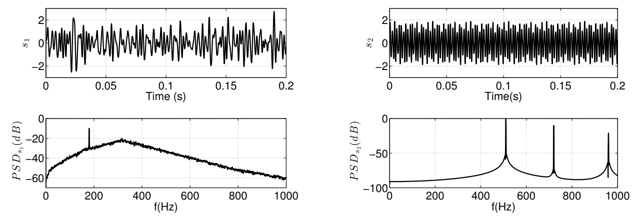


Fig. 1. Simulated scenario.

Equation (1), where the mixing coefficients and the delays are given by $a_{kl} = \frac{1}{D_{kl}}$ and $\tau_{kl} = \frac{D_{kl}}{c}$, respectively. D_{kl} is the distance between the source l and the sensor k, and c is the sound velocity. The reflections on the surface are not considered in this simulation. The first source is formed by a tone at frequency 180 Hz and a broadband component; the second source consists of three sinusoids at frequencies 510 Hz, 720 Hz and 960 Hz. The two sources roughly simulate the sound radiated by a propeller and the vibrations of the hull of a ship, respectively [12] [13]. The temporal profiles and the power spectral densities of the original sources are plotted in



(a) Original s_1 (b) Original s_2

Fig. 2. The original sources.

Figure 2. We set the sound velocity to $c = 1500\ m/s$, the sampling frequency to $f_s = 44100\ Hz$, and the duration to $T = 10\ s$. Each of the simulated noise-free mixed data is corrupted by a zero-mean additive white Gaussian, where the signal-to-noise ratio defined by (31) is set to 10 dB.

$$SNR_k = 10 \log_{10} \left(\frac{P_{\bar{x}_k}}{P_{b_k}} \right). \tag{31}$$

where $P_{\bar{x}_k}$ is the power of the noise-free mixed data given by (32), and P_{b_k} is the power of the noise.

$$\bar{x}_k(t) = \sum_{l=1}^{L} a_{kl} s_l \left(t - \tau_{kl}\right). \tag{32}$$

Figure 3 shows the estimated sources, while Tables 1 and 2 give the SIR and the NME for the three methods (i.e., proposed frequency method, temporal method, and time-frequency method).

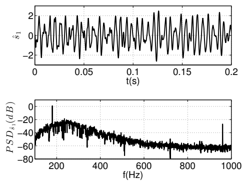
	Proposed	**Temporal**	**Time-frequency**
$\hat{s}_1$	28.89	18.73	-9.22
$\hat{s}_2$	23.71	14.62	9.49

Table 1. Signal-to-interference ratio (dB).

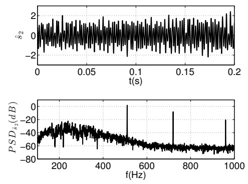
	Proposed	**Temporal**	**Time-frequency**
$\hat{s}_1$	0.12	0.76	1.54
$\hat{s}_2$	0.51	0.83	1.03

Table 2. Normalized mean error on the relative delays.

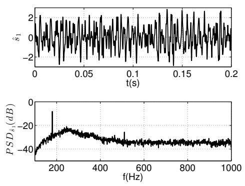
It can be seen from Figure 3 that the proposed method, and the temporal method, correctly estimates the temporal profiles and the PSDs of both of the sources. There is a residue of the estimation of each source in the second one, which can be seen in the PSDs of estimated sources. The time-frequency method appears to fail; this might be because it includes an alternating least square and the mixing matrix is ill-conditionned. Table 1 shows that the proposed method presents the best SIR for both source 1 and source 2. The small value of the SIR of estimated source 2 can be explained



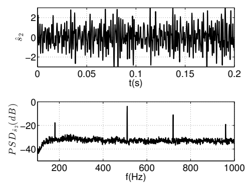
(a) $\hat{s}_1$ by proposed method



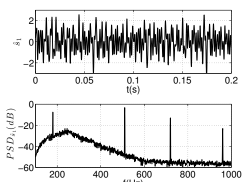
(b) $\hat{s}_2$ by proposed method



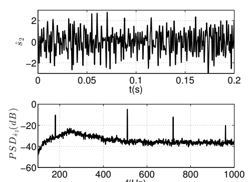
(c) $\hat{s}_1$ by temporal method



(d) $\hat{s}_2$ by temporal method



(e) $\hat{s}_1$ by time-frequency method



(f) $\hat{s}_2$ by time-frequency method

Fig. 3. Estimated sources.

because the energy of the broadband component of the residue of estimated source 1 is relatively high compared to the energy of the three tones of estimated source 2. From Table 2, we see that the proposed method has the lowest NME, which indicates that our method has the best estimates of the relative delays.

4. CONLUSION AND FUTURE STUDIES

In this paper, we have developed a frequency method for the blind separation of an anechoic mixture. Our method first estimates the mixing coefficients and the modulus of the source Fourier transforms, followed by estimation of the relative delays between a reference sensor and the other sensors. The last step consists of the estimation of the phases of the source Fourier transform and the temporal profiles of the sources. Simulations on synthetic data show that the proposed method outperforms some of the existing methods, especially when the mixing matrix is ill-conditioned. Future works will include an evaluation of this proposed method on real data, and the inclusion of reflected paths in the mixing model. The case of moving sources will also be investigated.

REFERENCES

[1] L. Omlor and M. A. Giese, "Anechoic blind source separation using wigner marginals," *Journal of Machine Learning Research*, vol. 12, pp. 1111–1148, 2011.

[2] A. Yeredor, "Blind source separation with pure delays mixture," in *ICA*, 2001.

[3] N. Jiang and D. Farina, "Covariance and time-scale methods for blind separation of delayed sources," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 3, pp. 550–556, 2011.

[4] G. Chabriel and J. Barrere, "An instantaneous formulation of mixtures for blind separation of propagating waves," *IEEE Transactions on Signal Processing*, vol. 54, no. 1, pp. 49–58, 2006.

[5] D. Nion, B. Vandewoestyne, S. Vanaverbeke, K. Van Den Abeele, H. De Gersem, and L. De Lathauwer, "A time-frequency technique for blind separation and localization of pure delayed sources," in *9th International Conference on Latent Variable Analysis and Signal Separation*, 2010, vol. 6365 of *Lecture Notes in Computer Science*, pp. 546–554.

[6] L. Omlor and M. A. Giese, "Blind source separation for over-determined delayed mixtures," *Advances in Neural Information Processing Systems*, vol. 19, pp. 1074–1081, 2007.

[7] J. F. Cardoso and A. Souloumiac, "Blind beamforming for non gaussian signal," *IEE Proceedings-F*, vol. 140, pp. 362–370, 1993.

[8] M. W. Berry, M. Browne, A. N. Langville, V. P. Pauca, and R. J. Plemmons, "Algorithms and applications for approximate nonnegative matrix factorization," *Computational Statistics and Data Analysis*, vol. 52, no. 1, pp. 155–173, 2007.

[9] P. C. Hansen, "The l-curve and its use in the numerical treatment of inverse problems," in *Computational Inverse Problems in Electrocardiology, ed. P. Johnston*, 2000, vol. 4 of *Advances in Computational Bioengineering*, pp. 119–142.

[10] D. Calvetti, S. Morigi, L. Reichel, and F. Sgallari, "Tikhonov regularization and the l-curve for large discrete ill-posed problems," *Journal of Computational and Applied Mathematics*, vol. 123, no. 12, pp. 423 – 446, 2000.

[11] E. Vincent, R. Gribonval, and C. Fevotte, "Performance measurement in blind audio source separation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, no. 4, pp. 1462–1469, 2006.

[12] P. T. Arveson and D. J. Vendittis, "Radiated noise characteristics of a modern cargo ship," *The Journal of the Acoustical Society of America*, vol. 107, pp. 118–129, 2000.

[13] S. C. Wales and R. M. Heitmeyer, "An ensemble source spectra model for merchant ship-radiated noise," *The Journal of the Acoustical Society of America*, vol. 111, pp. 1211–1231, 2002.