Under-determined source separation via mixed-norm regularized minimization

Matthieu Kowalski, Emmanuel Vincent, Rémi Gribonval

UNDER-DETERMINED SOURCE SEPARATION VIA MIXED-NORM REGULARIZED MINIMIZATION

Matthieu Kowalski[1], *Emmanuel Vincent*[2] *and Rémi Gribonval*[2]

[1]LATP, Université de Provence
39, rue F. Joliot-Curie 13453 Marseille cedex France
phone: +(33/0)4 91 05 47 43, fax: +(33/0)4 91 05 47 42, email: kowalski@cmi.univ-mrs.fr

[2]IRISA-INRIA
Campus de Beaulieu 35042 Rennes cedex, France
phone: (+33/0) 299 842 {269,506}, fax: (+33/0) 299 847 {171,532}, email: {emmanuel.vincent,remi.gribonval}@irisa.fr

ABSTRACT

We consider the problem of extracting the source signals from an under-determined convolutive mixture assuming that the mixing filters are known. We wish to exploit the sparsity and approximate disjointness of the time-frequency representations of the sources. However, classical time-frequency masking techniques cannot be directly applied due to the convolutive nature of the mixture. To address this problem, we first formulate it as the minimization of a functional combining a classical ℓ_2 discrepancy term between the observed mixture and the mixture reconstructed from the estimated sources and a sparse regularization term defined in terms of mixed ℓ_2/ℓ_1 norms of source coefficients in a time-frequency domain. The minimum of the functional is then obtained by a thresholded Landweber iteration algorithm. Preliminary results are discussed for two synthetic audio mixtures.

1. UNDER-DETERMINED SOURCE SEPARATION

We consider the source separation problem for convolutive mixtures of the form

$$\mathbf{x} = A \star \mathbf{s} \;, \tag{1}$$

with $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_M)^T$, $\mathbf{x}_m \in \mathbb{R}^{T_x}$ $\forall m$, the channels of the observed mixture, $\mathbf{s} = (\mathbf{s}_1, \dots, \mathbf{s}_N)^T$, $\mathbf{s}_n \in \mathbb{R}^{T_s}$ $\forall n$, the unknown source signals, and A the mixing filter system. In the case of an under-determined mixture ($M < N$), this problem is generally split into two distinct steps, namely the estimation of the mixing system A and the extraction of the source signals $\mathbf{s}$ given A and $\mathbf{x}$. Both steps are equally challenging. In the following, we focus on the estimation of $\mathbf{s}$, assuming that A is known.

Since the linear system (1) is under-determined, additional hypotheses are needed in order to "invert it" and estimate the original sources. A classical assumption is to suppose that only few sources are simultaneously active. This assumption is often roughly satisfied after a properly chosen transform. In the case of audio sources, this is generally the case after switching to the time-frequency domain using the short-time Fourier transform (STFT). Taking the STFT of both hand sides of (1) yields the narrowband approximation

$$\hat{\mathbf{x}}(f,t) \approx \hat{A}(f)\hat{\mathbf{s}}(f,t) \;, \tag{2}$$

where $\hat{\mathbf{x}}$ (resp. $\hat{\mathbf{s}}$) is the STFT transform of $\mathbf{x}$ (resp. $\mathbf{s}$) and $\hat{A}$ is the Fourier transform of A. The sparsity assumption means that for each time-frequency point (f,t), few sources are simultaneously active. A classical approach [1, 2] to identify their contributions is to solve for each (f,t) the optimization problem

$$\min_{\hat{\mathbf{s}}} \|\hat{\mathbf{s}}(f,t)\|_p^p = \min_{\hat{\mathbf{s}}} \sum_{n=1}^{N} |\hat{s}_n(t,f)|^p \tag{3}$$
$$\text{subject to} \quad \hat{\mathbf{x}}(f,t) = \hat{A}(f)\hat{\mathbf{s}}(f,t) \;,$$

where the ℓ_p norm is a measure of sparsity for $p \leq 1$. Algorithms such as FOCUSS [3] or the study made in [4] can be used to solve (3). While this approach results in M active sources per time-frequency point for small p [4], a single active source can be selected instead using a time-frequency masking algorithm such as DUET [5]. A variable number of active sources between 0 and M can also be determined via ℓ_1-regularized minimization [6] or via a probabilistic approach [7].

The above source separation techniques cannot easily be applied to convolutive mixtures, or even anechoic mixtures with rather long delays, because the approximation (2) becomes to crude. Indeed, even if the sources happen to have sparse and almost disjoint time-frequency representations, their contributions $a_n \star s_n$ to the mixture (where a_n is the n-th column of the filter matrix A) may have significantly more overlapping time-frequency representations. In this paper, we give a general approach to estimate $\mathbf{s}$ from (1), avoiding the approximation made in (2). Our approach is based on the minimization of a suitable functional, discussed in section 2. We then derive a thresholded Landweber iteration algorithm that minimizes this functional in section 3. Section 4 provides some results for audio data.

2. SOURCE SEPARATION BY MINIMIZATION OF A SUITABLE FUNCTIONAL

The goal of this section is to provide, under the form of an optimization problem, a simple formulation to estimate the original sources. The problem of source separation (1) can be viewed as a linear inverse problem. A classical estimate for $\mathbf{s}$ is then given by minimizing the discrepancy

$$\|\mathbf{x} - A \star \mathbf{s}\|_2^2 \;, \tag{4}$$

where the ℓ_2 norm is the Froebenius norm (the sum of the energy across all channels). In a Bayesian point of view, it corresponds to a Gaussian prior on the residual.

However, minimizing the discrepancy term (4) is often insufficient and does not provide a good separation of the sources. In particular, when the mixture is under-determined, there are infinitely many solutions and one must introduce additional knowledge about the sources, which can be done by choosing a suitable regularization term.

Usually, the source signals are modeled by their expansion $\mathbf{s} = \sum_k \underline{s}_k \phi_k^T = \underline{\mathbf{s}}\Phi^T$ in a dictionary Φ associated with a time-frequency transform, such as the modified discrete cosine transform (MDCT) [8], and the regularization term is defined in terms of the transformed coefficients $\underline{\mathbf{s}}$. This leads to the following functional

$$\min_{\underline{\mathbf{s}}} \left[\left\| \mathbf{x} - A \star \left(\underline{\mathbf{s}}\Phi^T \right) \right\|_2^2 + \lambda E(\underline{\mathbf{s}}) \right] , \tag{5}$$

where E is the regularization term, and $\lambda \in \mathbb{R}^+$ a free penalty factor. The choice of the regularization term E depends on the assumptions made about the sources.

2.1 ℓ_1 regularization

One possible assumption is that the signals are sparse in the time-frequency domain [1]. This assumption has been used in particular by the Multichannel Morphological Component Analysis (MMCA) [9] algorithm, which relies on the following functional for instantaneous over-determined source separation

$$\min_{\underline{\mathbf{s}}} \left[\left\| \mathbf{x} - A \left(\underline{\mathbf{s}}\Phi^T \right) \right\|_2^2 + \lambda \|\underline{\mathbf{s}}\|_1 \right] . \tag{6}$$

where Φ is chosen as a union of frames (or bases) adapted to different components of the sources.

However, additional experiments showed us that this functional was not adapted to the under-determined case. In this case, the functional tends to threshold out the sources at high frequencies, and keep them all at low frequencies since they have higher energy. In other words, the sources are distorted at high frequencies and remain mixed together at low frequencies.

2.2 $\ell_{1,2}$ regularization

In order to achieve better separation, we follow the usual assumption made in the under-determined case: only few sources are active for each time-frequency index. However we would like to avoid "over-sparsifying", that is thresholding out time-frequency regions which have less energy. The mixed norm defined hereafter is well adapted to this aim.

Definition 1 *Let $p \geq 1$ and $q \geq 1$. Let $\mathbf{x} \in \mathbb{R}^L$ be labeled by a double index (n,k) such that $L = N \times K$. We call mixed norm of $\mathbf{x}$, the norm $\ell_{p,q}$ defined by*

$$\|\mathbf{x}\|_{p,q} = \left(\sum_{k=1}^{K} \left(\sum_{n=1}^{N} |x_{n,k}|^p \right)^{q/p} \right)^{1/q} . \tag{7}$$

The cases $p = +\infty$ and $q = +\infty$ can be obtained by replacing the corresponding norm by the supremum.

This type of norm was studied in the context of functional spaces (see e.g [10] and references therein). Such norms allow the "structuring" of sparsity on the coefficients: contrary to the simple ℓ_p norms, the coefficients are not considered to be independently distributed. The $\ell_{p,1}$ norm was used in the context of multichannel signal processing [11–13] and under the name of joint-sparsity [14–16]. The $\ell_{2,1}$ norm is used in the statistical community for the group-lasso [17] estimate.

In the context of source separation, the $\ell_{1,2}$ norm can be used to model the sparsity through the channels. Indeed, let k be a time-frequency index and n be the channel index, and consider the following $\ell_{1,2}$ norms on the sources:

$$\|\underline{\mathbf{s}}\|_{1,2}^2 = \sum_k \left(\sum_n |\underline{s}_{n,k}| \right)^2 .$$

Minimizing such a quantity will enforce sparsity across channels for each time-frequency index k, but not necessarily sparsity across time-frequency indices: for a given time-frequency index, we hope to keep only the coefficients corresponding to the most significant channels, but we expect some coefficients to be kept for most time-frequency indices. This behavior will be enlightened at the end of section 3.

Using this mixed norm as the regularization term in the functional (5), we obtain the following optimization problem

$$\min_{\underline{\mathbf{s}}} \left[\Psi(\underline{\mathbf{s}}) := \left\| \mathbf{x} - A \star \left(\underline{\mathbf{s}}\Phi^T \right) \right\|_2^2 + \lambda \|\underline{\mathbf{s}}\|_{1,2}^2 \right] . \tag{8}$$

Such a functional remains convex and can be minimized by a thresholded Landweber iteration following [18]. This is detailed in the next section.

3. ITERATIVE MINIMIZATION ALGORITHM

For the sake of simplicity, we introduce the following linear operator, which reconstructs a mixture from a representation of the sources by applying first the source reconstruction operator $\cdot\Phi^T$, then the filter matrix A:

$$\begin{aligned} \mathscr{F} : \mathbb{R}^{N \times T_s} &\to \mathbb{R}^{M \times T_x} \\ \underline{\mathbf{s}} &\mapsto \mathscr{F}(\underline{\mathbf{s}}) = A \star \left(\underline{\mathbf{s}}\Phi^T \right) . \end{aligned} \tag{9}$$

Then, one can apply the thresholded Landweber iteration algorithm described in [19], to minimize (8). We recall the iteration hereafter in the context of source separation.

If we differentiate the functional Ψ of (8) with respect to $\underline{s}_{n,k}$, the resulting system of variational equations is not readily solvable. Indeed these equations are coupled due to the presence of convolution products. In order to decouple the variational equations, we introduce the surrogate functional:

$$\begin{aligned} \Psi^{\text{sur}}(\underline{\mathbf{s}}, \underline{\mathbf{z}}) = \|\mathbf{x} - \mathscr{F}(\underline{\mathbf{s}}))\|_2^2 + C\|\underline{\mathbf{s}} - \underline{\mathbf{z}}\|_2^2 - \|\mathscr{F}(\underline{\mathbf{s}}) - \mathscr{F}(\underline{\mathbf{z}})\|_2^2 \\ + \lambda \|\underline{\mathbf{s}}\|_{1,2}^2 , \end{aligned} \tag{10}$$

where C is chosen greater than the square of the operator norm of $\mathscr{F}$, i.e. such that $\|\mathscr{F}(\underline{\mathbf{s}})\|_2^2 < C\|\underline{\mathbf{s}}\|_2^2$ for all $\underline{\mathbf{s}}$. Let us introduce the following notations: $\tilde{\underline{\mathbf{x}}} := (A^* \star \mathbf{x})\Phi$ and $\tilde{\underline{\mathbf{z}}} := \{A^* \star (A \star (\underline{\mathbf{z}}\Phi^T))\}\Phi$, where we denoted by A^* the adjoint operator of A, that is obtained by transposition of the channel indices and the source indices, and time reversal of A. Then, if we denote by

$$\mathbf{y}_{n,k} = \frac{\tilde{\underline{x}}_{n,k} + C\underline{s}_{n,k} - \tilde{\underline{z}}_{n,k}}{C} ,$$

the variational equations associated with Ψ^{sur} are the following

$$\forall(n,k) \quad \underline{s}_{n,k} = y_{n,k} - \text{sgn}(\underline{s}_{n,k})\frac{\lambda}{C}\|\underline{\mathbf{s}}_k\|_1 \ , \tag{11}$$

where $\underline{\mathbf{s}}_k := (\underline{s}_{1,k},\dots,\underline{s}_{N,k})$. Applying the theorem 3 of [19], the solution of this equation and then the argmin $\underline{\hat{\mathbf{s}}}$ of $\Psi^{\text{sur}}(\underline{\mathbf{s}},\underline{\mathbf{z}})$ with respect to $\underline{\mathbf{s}}$, is given coordinatewise by [1]

$$\underline{\hat{s}}_{n,k} = \text{sgn}\left(y_{n,k}\right)\left(\left|y_{n,k}\right| - \frac{\lambda/C}{1+L_k\frac{\lambda}{C}}|||\mathbf{y}_k|||\right)^+ \quad \forall(n,k) \ , \tag{12}$$

with $|||\mathbf{y}_k||| = \sum_{l=1}^{L_k}\check{y}_{k,l}$, where $\check{y}_{k,l}$ denote the coefficients $|\mathbf{y}_{k,l}|$ ordered by descending order. The quantity L_k is the number such that

$$\check{y}_{k,L_k+1} \leq \lambda \sum_{\ell=1}^{L_k+1}\left(\check{y}_{k,\ell} - \check{y}_{k,L_k+1}\right) \text{ and } \check{y}_{k,L_k} > \lambda\sum_{\ell=1}^{L_k}\left(\check{y}_{k,\ell}-\check{y}_{k,L_k}\right) \ . \tag{13}$$

Knowing this new "generalized thresholding operator", one can apply the thresholded Landweber iteration which is, starting from any $\underline{\mathbf{s}}^{(0)}$,

$$\underline{\mathbf{s}}^{(m+1)} = \underset{\underline{\mathbf{s}}}{\text{argmin}}\,\Psi^{\text{sur}}\left(\underline{\mathbf{s}},\underline{\mathbf{s}}^{(m)}\right) \ .$$

These iterations lead then to the following practical algorithm for source separation, with the mixing filter system A known.

Algorithm 1 **Let** $\underline{\tilde{\mathbf{x}}} = (\tilde{\mathbf{x}}_1,\dots,\tilde{\mathbf{x}}_N) := (A^* \star \mathbf{x})\Phi$,

$\underline{\mathbf{s}}^{(0)} = \mathbf{0}$

m = 0

do

$\underline{\tilde{\mathbf{s}}}^{(m)} = \left(\underline{\tilde{\mathbf{s}}}_1^{(m)},\dots,\underline{\tilde{\mathbf{s}}}_N^{(m)}\right) := \{A^*\star(A\star(\underline{\mathbf{s}}^{(m)}\Phi^T))\}\Phi$,

for *each time-frequency index k and each source index n*

$$y_{n,k}^{(m)} = \frac{\underline{\tilde{x}}_{n,k} + C\underline{s}_{n,k}^{(m)} - \underline{\tilde{s}}_{n,k}^{(m)}}{C}$$

$$\underline{s}_{n,k}^{(m+1)} = sgn\left(y_{n,k}^{(m)}\right)\left(\left|y_{n,k}^{(m)}\right| - \frac{\lambda/C}{1+L_k^{(m)}\frac{\lambda}{C}}\left|\left|\left|\mathbf{y}_k^{(m)}\right|\right|\right|\right)^+$$

with $L_k^{(m)}$ *and* $\left|\left|\left|\mathbf{y}_k^{(m)}\right|\right|\right|$ *as defined in* (13) *and* (12).

endfor

until *convergence*

It has been shown in [19] that this algorithm converges to a global minimum of the original functional Ψ of (8).

We shall stress that in the extreme case $\lambda = +\infty$, for a given index k, at least one coefficient will be kept among $(y_{1,k}^{(m)},\dots,y_{N,k}^{(m)})$ at each iterations. Furthermore, equation (11) assures that $\mathbf{y}_k^{(m)} = \mathbf{0}$ if and only if $\underline{\mathbf{s}}_k^{(m+1)} = \mathbf{0}$. This remark guaranties that once the algorithm converged, at least one coefficient is non zero for each time-frequency index if $\underline{\tilde{\mathbf{x}}}_k \neq \mathbf{0}$.

[1]The operator $^+$ is defined as follow: $x^+ = x$ if $x \geq 0$, and 0 if $x < 0$.

4. SOURCE SEPARATION RESULTS

We assessed the source separation performance of our algorithm on an anechoic audio mixture and a convolutive audio mixture obtained with simulated room impulse responses. We compared the performance of our algorithm with an implementation of DUET based on the narrowband approximation (2), using a STFT with a Gaussian window and defining $\hat{A}(f)$ as the Fourier transform of the known mixing filters. The window size for DUET has an influence on the performance: the window must be long enough in order for (2) to hold, but a too long window results in greater overlap of the sources. 1024 samples seems a good compromise for both examples. The resulting audio signals can be downloaded from [20] for listening comparisons.

We would like to stress again that the mixing filter system A is supposed to be known.

4.1 Anechoic mixture

The anechoic mixture consisted of four musical sources. Each source was sampled at 44100 Hz and had a duration of 2^{17} samples (about 3 s). The mixing system was obtained by combining the following instantaneous matrix (available on [21]) with delays between 0 and 512 samples

$$\begin{pmatrix} 0.3420 & 0.6428 & 0.7934 & 0.9239 \\ 0.9397 & 0.7660 & 0.6088 & 0.3827 \end{pmatrix}.$$

The chosen dictionary Φ was a MDCT basis with a sine window of 2048 samples (about 46 ms). Note that, similarly as above, a different window length could affect the results but that the chosen length seems a good compromise.

We display in Figure 1 the SDR (Signal to Distortion Ratio), SIR (Signal to Interference Ratio) and SAR (Signal to Artifact Ratio) [22], averaged over the four sources, as a function of the penalty factor λ. A larger value of SDR/SIR/SAR means a better quality of the separation. As one could expect, the SIR of DUET is always better than the SIR of our algorithm: since DUET uses only one source for each time-frequency index, it avoids as much as possible interference between sources. This comes to the price of a larger mismatch between the mixture and its reconstruction from the estimated sources. On the opposite, the SAR obtained by our algorithm is systematically better, because of the discrepancy term.

The SDR curve gives an idea of the global performance, taking into account both artifacts and interferences. One can see that, for well chosen λ, our iterative algorithm performs better than DUET. Having to choose the penalty factor λ is certainly one of the main drawbacks of the functional approach (5). On Figure 2 we plot the percentage of coefficients which are set to zero, and observe that the algorithm performs best when about 50 % coefficients are zeroed. Here, this means that on the average, at each time-frequency point two sources are considered active and two are set to zero. This seems a good compromise for a stereo mixture, and we believe that such a heuristic could be used as an adaptive strategy to tune the penalty factor in other settings.

4.2 Convolutive mixture

The convolutive mixture was generated by simulation of a recording in a meeting room with 250 ms reverberation time using a pair of omnidirectional microphones spaced by 1 m.

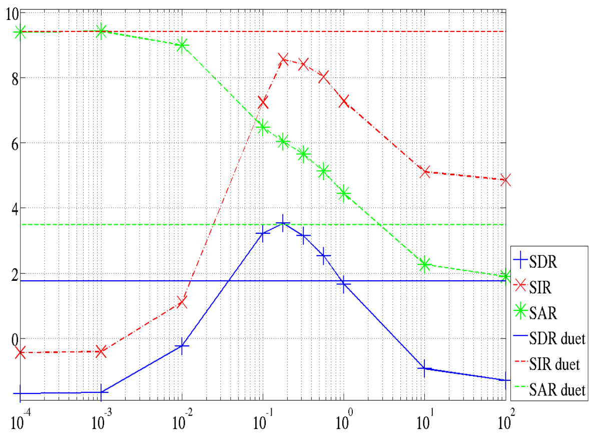

Figure 1: SDR, SIR, SAR of the proposed iterative algorithm and DUET as a function of the penalty factor λ (log scale).

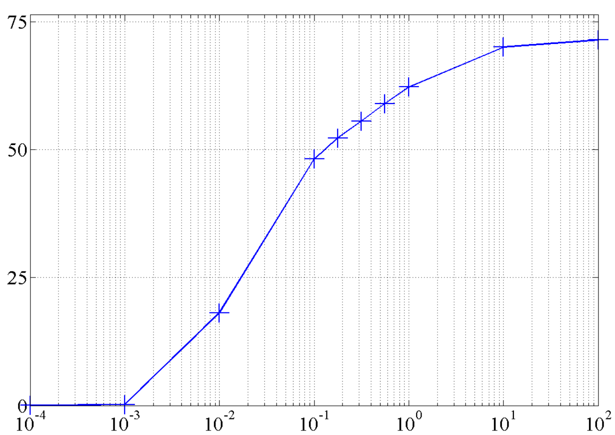

Figure 2: Percentage of coefficients set to zero as a function of the penalty factor λ (log scale).

The mixing filters were simulated by the image method and had a length of 4000 samples. The sources were four female speech recordings, sampled at 16000 Hz, and 2^{17} samples long (about 8 s). The matrix Φ corresponded to a MDCT basis, with a window of 512 samples (about 32 ms). All the data (sources and mixing system) are available on [21].

As for the anechoic mixture, we display in Figure 3 the the SDR, SIR and SAR obtained by our algorithm, averaged over four sources, as a function of the penalty factor. These are compared to the ones obtained by DUET. Figure 4 shows the percentage of coefficients which are set to zero.

In this case too, one can remark that the iterative algorithm performs best when about 50 % of coefficients are set to zero, even if the point $\lambda = 10^4$ seems to perform better than the others in terms of SAR (and correspond to 75 % of coefficients set to zero). Indeed, these curves are made in average, and, by chance, the first and the second sources are well estimated for this value of λ. These two sources are the best estimated ones by DUET, and that can explain why the iterative algorithm performs well. Further experiments showed that, in general, the iterative algorithm performs best when 50 % of coefficients are set to zero[2].

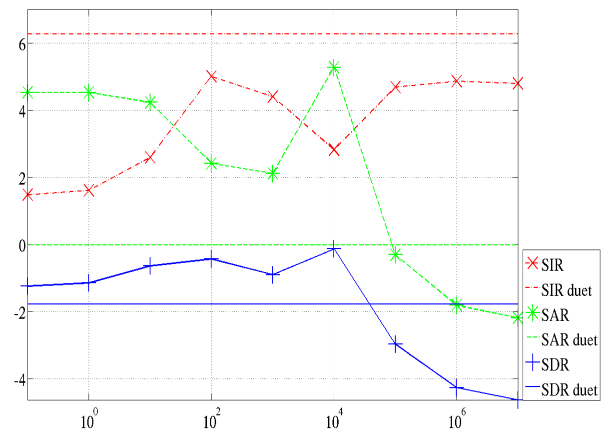

Figure 3: SDR, SIR, SAR of the proposed iterative algorithm and DUET as function of the penalty factor λ (log scale).

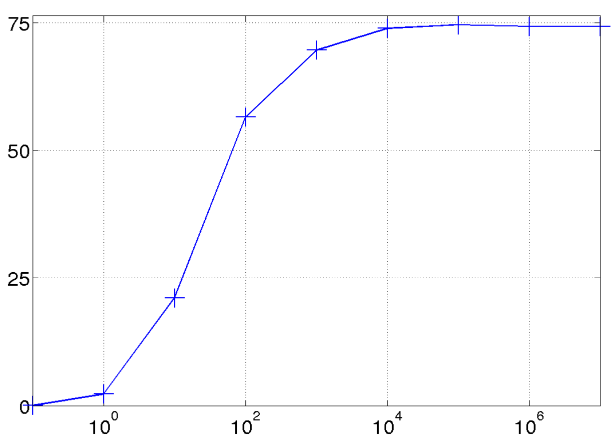

Figure 4: Percentage of coefficients set to zero as a function of the penalty factor λ (log scale).

We also compared briefly the results obtained by our minimization to the results given by an approach such (3), with $p = 1$ and $p = 10^{-3}$. It appears that the Landweber iteration gives better results in term of SDR/SIR/SAR in average on the sources, but can provide lower SDR/SIR on particular sources (typically, on the third source of the used mixture).

4.3 Computational efficiency

The computation time of the proposed algorithm depends mainly on the efficiency of the time-frequency transform and that of the convolution by the mixing system and its adjoint. The convergence of the thresholded Landwebver iterative algorithm itself is known to be quite slow, but potential speed-ups are proposed in the literature [23, 24]. In the above experiments, the overall computation time was about 5 min in

[2]For a stereo mixture of four sources.

the anechoic case and between 1 h for small values of λ and 5 min for very large values of λ in the convolutive case.

5. CONCLUSION

In this paper we developed an iterative optimization approach to separate convolutive mixtures using sparse source models in a time-frequency dictionary, when the mixing filter system is supposed to be known. Instead of the classical ℓ_1 penalty, we showed that the $\ell_{1,2}$ mixed norm appears to be well adapted to respect the sparse hypothesis over the channels: for each time-frequency index, only few sources are active. The functional we proposed can be easily minimized by a thresholded Landweber iteration and leads to a simple algorithm that one can quickly implement.

We showed that this approach gives interesting results compared to DUET: despite an expected lower SIR, the SAR is always better, and the SDR can be improved too. Moreover, this algorithm works for any number of mixture and any number of sources to estimate, i.e. for any $M,N \in \mathbb{N}$ with $M < N$.

One of the major benefits of the proposed approach, is that one can make additional assumptions on the sources, and take them into account in the regularization term. Indeed, the next step should be to try to incorporate some structures on each source, as, for example, harmonicity for music sources.

Acknowledgments

We wish to thank B. Torrésani for interesting discussions and his interest to this work.

REFERENCES

[1] M. Zibulevsky and B. Pearlmutter, "Blind source separation by sparse decomposition in a signal dictionary," *Neural Computation*, vol. 13, no. 4, pp. 863–882, 2001.

[2] P. Bofill and M. Zibulevsky, "Underdetermined blind source separation using sparse representations," *Signal Processing*, vol. 81, pp. 2353–2362, 2001.

[3] B. Rao and K. Kreutz-Delgado, "An affine scaling methodology for best basis selection," *IEEE Transactions On Signal Processing*, vol. 47, no. 1, pp. 187–200, January 1999.

[4] E. Vincent, "Complex nonconvex l_p norm minimization for underdetermined source separation," in *Proc. Int. Conf. on Independent Component Analysis and Blind Source Separation (ICA)*, 2007, pp. 430–437.

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

[6] M. Togami, T. Sumiyoshi, and A. Amano, "Sound source separation of overcomplete convolutive mixture using generalized sparseness," in *Proc. Int. Workshop on Acoustic Echo and Noise Control (IWAENC)*, 2006.

[7] L. Vielva, D. Erdogmus, and J. C. Príncipe, "Underdetermined blind source separation using a probabilistic source sparsity model," in *Proc. Int. Conf. on Independent Component Analysis and Blind Source Separation (ICA)*, 2001.

[8] S. Mallat, *A Wavelet Tour of Signal Processing*. San Diego, CA: Academic Press, 1998.

[9] J. Bobin, J.-L. Starck, J. Fadili, and Y. Moudden, "Sparsity and morphological diversity in blind source separation," *IEEE Transactions on Image Processing*, vol. 16, no. 11, pp. 2662–2674, November 2007.

[10] S. Samarah, S. Obeidat, and R. Salman, "A Schur test for weighted mixed-norm spaces," *Analysis Mathematica*, vol. 31, pp. 277–289, 2005.

[11] R. Gribonval, "Piecewise linear source separation," in *Proc. SPIE '03*, vol. 5207 Wavelets: Applications in Signal and Image Processing X, San Diego, CA, 2003, pp. 297–310.

[12] R. Gribonval, B. Mailhé, H. Rauhut, K. Schnass, and P. Vandergheynst, "Average case analysis of multichannel thresholding," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2007.

[13] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2477–2488, July 2005.

[14] G. Teschke and R. Ramlau, "An iterative algorithm for nonlinear inverse problems with joint sparsity constraints in vector valued regimes and an application to color image inpainting," 2007, Preprint.

[15] M. Fornasier and H. Rauhut, "Recovery algorithm for vector-valued data with joint sparsity constraints," *SIAM Journal on Numerical Analysis (to appear)*, 2007.

[16] M. Fornasier and H. Rauhut, "Iterative thresholding algorithms," *Applied and Computational Harmonic Analysis*, 2007, doi:10.1016/j.acha.2007.10.005.

[17] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *Journal of the Royal Statistical Society Series B*, vol. 68, no. 1, pp. 49–67, 2006.

[18] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413 – 1457, August 2004.

[19] M. Kowalski, "Sparse regression using mixed norms," 2007, Preprint. [Online]. Available: http://hal.archives-ouvertes.fr/hal-00202904/

[20] [Online]. Available: http://www.cmi.univ-mrs.fr/~kowalski/EUSIPCO08_bss.html

[21] [Online]. Available: http://sassec.gforge.inria.fr/

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

[23] I. Daubechies, M. Fornasier, and L. Loris, "Accelerated projected gradient methods for linear inverse problems with sparsity constraints," *Journal of Fourier Analysis and Application*, 2007, to appear.

[24] M. Fornasier, "Domain decomposition methods for linear inverse problems with sparsity constraints," *Inverse Problems*, vol. 23, pp. 2505–2526, 2007.