# A WIDEBAND DOUBLY-SPARSE APPROACH FOR MITO SPARSE FILTER ESTIMATION

*Simon Arberet*[⋆]*, Prasad Sudhakar*[†] *and Rémi Gribonval*[†]

[⋆]Institute of Electrical Engineering
École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland
simon.arberet@epfl.ch

[†]METISS Team, Centre de recherche INRIA-Rennes-Bretagne Atlantique
Rennes CEDEX 35042, France
{prasad.sudhakar,remi.gribonval}@inria.fr

## ABSTRACT

We propose an approach for the estimation of sparse filters from a convolutive mixture of sources, exploiting the time-domain sparsity of the mixing filters and the sparsity of the sources in the time-frequency (TF) domain. The proposed approach is based on a wideband formulation of the cross-relation (CR) in the TF domain and on a framework including two steps: (a) a clustering step, to determine the TF points where the CR is valid; (b) a filter estimation step, to recover the set of filters associated with each source. We propose for the first time a method to blindly perform the clustering step (a) and we show that the proposed approach based on the wideband CR outperforms the narrowband approach and the GCC-PHAT approach by between 5 dB and 20 dB.

***Index Terms***— Blind filter estimation, sparsity, convex optimisation, cross-relation, source separation

## 1. INTRODUCTION

Blind source separation (BSS) has applications in several fields such as speech and music processing, wireless communications or biomedical signal processing. In a general setting, we consider $M$ mixtures $x_i(t), i = 1 \dots M$ of $N$ source signals $s_j(t), j = 1 \dots N$, given by the convolutive model

$$x_i(t) = \sum_{j=1}^{N} (a_{ij} * s_j)(t) + v_i(t) \tag{1}$$

where $a_{ij}(t)$ is a filter of length $L$ that models the impulse response between the $j^{th}$ source and the $i^{th}$ microphone, and $v_i(t)$ is the noise at the $i^{th}$ microphone. For brevity, we denote the sources, filters, noise and mixtures by $s_j, a_{ij}, v_i$ and $x_i$ respectively, by dropping the time index.

BSS systems attempt to estimate the sources given only the mixtures. This is often done in two stages: the mixing filters are estimated first, and subsequently they are used for source estimation. In case of instantaneous and anechoic mixtures, the filters are simply scalars or time-delayed scalars, and several methods (see [1] and references therein) have been proposed to estimate the mixing parameters.

The authors acknowledge the support of the EU FP7 FET-Open program, SMALL project, under grant no. 225913.

The problem gets complicated with convolutive mixtures. Frequency domain techniques transform convolutive mixtures into multiple complex-valued instantaneous mixtures, under the narrowband approximation. However, this approach suffers from the ambiguities of arbitrary scaling and permutations of the sources, and solving these ambiguities is a challenging problem in itself [2].

On the other hand, when there is only one source, the problem of blindly estimating filters from the filtered versions of the source is well studied [3]. The main approach, described in section 2, exploits a *cross relation* (CR) in the time-domain and recast the filter estimation problem as a linear inverse problem. In addition, if the filters are sparse, e.g. in underwater acoustic communication, the problem can be regularized with a sparsity promoting norm (e.g. $\ell_p$ norm with $p \leq 1$) [4]. With that a priori, good reconstruction [5] is possible with observations $x_i(t)$ of small duration compared to the filter length $L$.

Recently a framework has been proposed for Sparse Filter Estimation (SFE) when multiple sources are active simultaneously. It assumes that the filters are sparse in the time domain and the sources are sparse in the frequency domain [6]. This SFE framework is composed of two steps: a first step where for each source, the set of frequencies where only this source is active is determined; a second step where the filters are estimated by solving a convex optimization problem. Note that to the opposite of most methods for MIMO channels identification in communication such as the subspace methods [7, 8], we assume that the filters are sparse and can be quite long[1]. As opposed to the subspace methods which require the number $N$ of sources to be less than the number $M$ of mixtures, our proposed framework can also deal with determined and underdetermined mixtures (i.e. mixtures where $N \geq M$).

In this paper, we extend the SFE framework for sparse filter estimation by introducing a new formulation of the CR called *wideband CR*. Unlike the narrowband CR, the wideband CR which we present in section 3, is exact as long as there is no interference of the other sources in the selected time-frequency regions. Unlike our previous work [6], where the selection of these TF regions was done with an oracle estimator, we propose in Section 4, a simple method to blindly select the TF regions in the case where we have a mixture of *one* sparsely filtered source with one or more instantaneously filtered sources. In Section 5 we evaluate the performance of the proposed approach to recover the convolutive filters in a scenario where we have two audio sources mixed on two channels.

[1]The length of the filters is more than 250 samples in our case while it is typically between 0 and 3 samples in MIMO subspace methods as in [7, 8].

## 2. SPARSE FILTER ESTIMATION USING THE CROSS RELATION (CR)

Before we present our contributions, let us first describe some existing work on blind estimation of sparse filters in single and multiple sources settings.

### 2.1. CR in the time domain:

Assume that there is only one source active $s$, and two outputs $x_1$ and $x_2$. This is the single-input-two-output (SITO) case and we have:

$$x_1(t) = (a_1 * s)(t), \quad x_2(t) = (a_2 * s)(t).$$

Let $T$ be the length of $s$ and $L$ the length of the filters, then the length of $x_i$ will be $T + L - 1$. We have the following cross-relation (CR)[9]:

$$(x_2 * a_1)(t) = (x_1 * a_2)(t) = (a_1 * a_2 * s)(t). \tag{2}$$

For convenience, let us associate the signal $a_i$ to the column vector $\mathbf{a}_i = [a_i(t)]_{t=1}^{L}$ and likewise $s$ to $\mathbf{s}$ and $x_i$ to $\mathbf{x}_i$.

The convolution $(x_i * a_i)(t)$ is associated to the multiplication between the Toeplitz matrix

$$\mathcal{X}_i = \begin{bmatrix} x_i(1) & 0 & \cdots & 0 \\ x_i(2) & x_i(1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ x_i(L) & x_i(L-1) & \cdots & x_i(1) \\ x_i(L+1) & x_i(L) & \cdots & x_i(2) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_i(T+L-1) \end{bmatrix} \tag{3}$$

and the vector $\mathbf{a}_i$. Define also the two-channel data matrix $\mathbf{B}_{\mathtt{td}} = \begin{bmatrix} \mathcal{X}_2, -\mathcal{X}_1 \end{bmatrix}$, where the subscript $\mathtt{td}$ stands for *time-domain*, and the two-channel filter vector $\mathbf{a} = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{bmatrix}$, we can write the time-domain CR as:

$$\mathbf{B}_{\mathtt{td}} \cdot \mathbf{a} = \mathbf{0}. \tag{4}$$

This relation has inspired several methods (named CR methods) to estimate the filters blindly from the observations [9, 3]. These methods generally do not assume anything about the nature of the filters, however in scenarios such as underwater/wide band wireless communications, the filters that model the channels are often sparse in the time domain.

### 2.2. Sparsity of the filters

With the additional sparsity assumption, the SITO filter estimation problem can be formulated as the following $\ell_1$ minimization problem [4, 6]:

$$\text{minimize} \quad \|\mathbf{a}\|_1 \quad \text{s.t.} \quad \|\mathbf{B} \cdot \mathbf{a}\|_2 \leq \epsilon \quad \text{and} \quad \mathbf{a}_1(t_0) = 1 \quad (P_1)$$

The notation $\mathbf{B}$ is our general notation for the CR matrix. A particular case is the time domain CR matrix $\mathbf{B}_{\mathtt{td}}$ we defined in Section 2.1. The normalisation $a_1(t_0) = 1$ (where $t_0$ is an arbitrarily chosen time index, as mentioned in [6]) is to avoid the trivial zero-vector solution. The resulting problem is a noise-aware variant of Basis Pursuit [10, 5] and can be solved using any standard convex optimization algorithm. We chose to use the `CVX` software package [11].

### 2.3. CR in the TF domain: the narrowband approach

When dealing with multiple sources, i.e. in the multiple inputs two outputs (MITO) case, the CR formulation (2) cannot be directly used without further assumptions. The SITO approach has been extended to $N$ sources [4], by assuming that it is possible to identify time segments where only one source contributes to the mixtures. Then a SITO problem can be formulated locally at such segments and solved to obtain the filters for that particular source.

However, in general, the sources may overlap in time. Hence this approach might not be suitable for the filter estimation task, even if the filters are sparse. Instead of disjoint time supports, it is a common assumption to consider sources almost disjoint in the TF domain [12]. This fact motivates the following formulation of the CR in the TF domain:

Let $\hat{\mathbf{x}}_i$ be the short-time Fourier transform (STFT) of the vector $\mathbf{x}_i$, and $\hat{\mathbf{a}}_i$ be the Fourier transform of $\mathbf{a}_i$ (appropriately zero padded) such that $\hat{\mathbf{x}}_i = [\hat{\mathbf{x}}_i(\tau, f)]_{\tau,f}$ and $\hat{\mathbf{a}}_i = [\hat{\mathbf{a}}_i(f)]_f$. Let us consider STFT frames $1 \leq \tau \leq \mathcal{T}$. Using the narrowband approximation, the CR in the TF domain can be expressed by:

$$\hat{\mathbf{x}}_2(\tau, f) \cdot \hat{\mathbf{a}}_1(f) \simeq \hat{\mathbf{x}}_1(\tau, f) \cdot \hat{\mathbf{a}}_2(f), \quad \forall(\tau, f). \tag{5}$$

Defining $\widehat{\mathcal{X}}_{i,\tau} = \operatorname{diag}\Big([\hat{\mathbf{x}}_i(\tau, f)]_f\Big)$, the CR in the STFT domain will be

$$\mathbf{B}_{\mathtt{nb}} \cdot \mathbf{a} \simeq \mathbf{0} \text{ with } \mathbf{B}_{\mathtt{nb}} = \begin{bmatrix} \widehat{\mathcal{X}}_{2,1}, & -\widehat{\mathcal{X}}_{1,1} \\ \vdots & \vdots \\ \widehat{\mathcal{X}}_{2,\mathcal{T}}, & -\widehat{\mathcal{X}}_{1,\mathcal{T}} \end{bmatrix} \begin{bmatrix} \mathbf{F}^* & \mathbf{0} \\ \mathbf{0} & \mathbf{F}^* \end{bmatrix}, \tag{6}$$

where $\mathbf{F}^*$ is the Fourier matrix of appropriate size. The optimization problem $(P_1)$ with $\mathbf{B} := \mathbf{B}_{\mathtt{nb}}$ can be solved to obtain the filters. This defines the narrowband CR approach.

### 2.4. A two-step framework for MITO

In the case of multiple sources, the approach in the TF domain is better than in the time-domain because the sources are often sparser in the TF domain, and thus it is more likely that the CR holds in some regions of that domain.

A two-step approach for sparse filter estimation was proposed in [6]. Its principle is that if we can identify a set $\Omega_j$ of TF points $(\tau, f)$ for each source $j$ where the CR holds, then these sets can be used to build the matrices $\mathbf{B}^{\Omega_j}$ such that $\mathbf{B}^{\Omega_j} \cdot \mathbf{a}^{(j)} \approx 0$, by selecting rows of $\mathbf{B}_{\mathtt{nb}}$ indexed by the points in $\Omega_j$. Then we can estimate the filters $\mathbf{a}^{(j)}$ by solving $(P_1)$ with $\mathbf{B} := \mathbf{B}^{\Omega_j}$. The algorithm we present in section 4 is based on this framework with a new definition of the CR which is introduced in section 3.

## 3. WIDEBAND FORMULATION OF THE CROSS-RELATION

To build the narrowband CR [6] we first performed a STFT and then formulated the CR based on the narrowband approximation. Thus, the narrowband CR is intrinsically approximate. The first main contribution of this paper is to propose an accurate wideband expression of the CR where we formulate the CR in the time domain and then take the transformation. It is based on the following lemma:

**Lemma 3.1.** *Let $x(t)$ be a bounded signal, $a(t)$ and $\psi(t)$ two finite support signals, and let $\langle f, g\rangle = \sum_{t=-\infty}^{+\infty} f(t)g^*(t)$ be the inner product between two signals $f$ and $g$.*
*Then: $\langle x * a, \psi\rangle = \langle a, \bar{x} * \psi\rangle$ with $\bar{x}(t) = x^*(-t)$.*

*Proof.* This equality is verified using Fubini's Theorem and a change of variable[2]. □

Using this lemma, the projection of the time-domain CR (2) on a signal or atom $\psi$, that is $\langle x_2 * a_1 - x_1 * a_2, \psi \rangle = 0$ can be written as:

$$\langle \bar{x}_2 * \psi, a_1 \rangle = \langle \bar{x}_1 * \psi, a_2 \rangle. \quad (7)$$

Unlike the narrowband CR of (5), we have in (7) a perfect equality if the time domain CR (2) holds. To capture the TF regions where the CR holds, we can use a dictionary $\mathcal{D}$ of atoms $\psi$ adapted to mixture the signal, e.g. for audio signal we can use a Gabor dictionary $\mathcal{D} = \{\psi_{\tau,f}\}_{\tau,f}$ where $\psi_{\tau,f}$ is a Gabor atom [13] with time shift $\tau$ and frequency $f$. Note that the time-domain CR given by (2) is the particular case of the wideband CR given by (7) with a dictionary $\mathcal{D} = \{\delta_\tau\}_\tau$ composed of translated Dirac $\delta_\tau(\cdot) = \delta(\cdot - \tau)$.

Let us express the wideband CR (7) more explicitly. As the filters $a_i$ are real-valued with support $[0, L-1]$, the inner product in (7) can be written as

$$\langle \bar{x}_i * \psi, a_j \rangle = \sum_{\tau=0}^{L-1} \langle x_i, \psi_{-\tau} \rangle a_j(\tau)$$

with $\psi_\tau(\cdot) = \psi(\cdot - \tau)$. Thus the row of $\mathbf{B}_{\text{wb}}$ corresponding to the atom $\psi$ is given by concatenating the vector $\varphi^L_{x_2,\psi} = [\langle x_2, \psi_{-\tau} \rangle]_{\tau=0}^{L-1}$ with $-\varphi^L_{x_1,\psi}$ (defined likewise).

For example, if $\mathcal{D}$ is a highly redundant STFT dictionary with an overlapping shift of one sample, then the vector $\varphi^L_{x_i,\psi_{\tau,f}}$ corresponding to the TF point $(\tau, f)$ is:

$$\varphi^L_{x_i,\psi_{\tau,f}} = [\hat{\mathbf{x}}_i(\tau, f), \hat{\mathbf{x}}_i(\tau - 1, f), \ldots, \hat{\mathbf{x}}_i(\tau - L + 1, f)]^T.$$

## 4. BLIND METHOD TO IDENTIFY THE SET $\Omega_j$

Previous work [6] addressed the estimation of the filters when the TF regions $\Omega_j$ are given by some oracle. In this paper, we propose a practical way (summarized in algorithm 1) to blindly estimate $\Omega_j$ from a mixture, in a scenario where all sources but one are associated to linear instantaneous filters. Let $k$ be the index of the "convolutive" source of the mixture. The length of the instantaneous filters is $L_j = 1, \forall j \neq k$, and each such source is associated to an *intensity parameter* (IP) $\theta_j = \tan^{-1}(a_{2j}(0)/a_{1j}(0))$.

Existing algorithms such as DEMIX [1] can both estimate the IP $\theta_j$ of the instantaneous sources and the TF regions $\Omega_j$ where they are prominently active based on a spatial criterion:

$$\Omega_{j \neq k} = \left\{ (\tau, f) : \left| \tan^{-1}\left( |(\hat{\mathbf{x}}_2(\tau, f)/\hat{\mathbf{x}}_1(\tau, f)| \right) - \theta_j \right| \leq \eta \right\}. \quad (8)$$

We build the set $\Omega^d_j, \forall j \neq k$ containing all the TF points close in time or in frequency to points in $\Omega_j$ using a dilation operation applied on $\Omega_j$ in the TF plan, and obtain the set $\Omega_k$ corresponding to the convolutive source as the complement of $\cup_{j \neq k} \Omega^d_j$.

Fig. 1 shows an example of STFT of two sources (where $k = 2$), and Fig. 2(a) displays the STFT of one of the mixture $x_1$ (black corresponds to high energy, white to low energy). Fig. 2(b) illustrates the set $\Omega_2$ obtained with the described approach, and we can see that as expected $\Omega_2$ only contains TF points where source $s_1$ has a very small energy.

[2]It is possible to extend this lemma into other specific spaces where $a(t)$ and $\psi(t)$ have an infinite support.

**Algorithm 1** for estimating the sparse filter $\mathbf{a}^{(k)}$ and $\theta_j, \forall j \neq k$.

1. Compute the STFT $\hat{\mathbf{x}}_i$ of the mixtures $\mathbf{x}_i, \forall i \in \{1, 2\}$.
2. Estimate the IP $\theta_j$ of the instantaneous source(s) using DEMIX [1].
3. Build the sets $\Omega^d_j$ of the instantaneous sources using (8) and a dilation operation applied on $\Omega_j$ in the TF plan.
4. Build the set $\Omega_k$ of the convolutive source: $\Omega_k = \overline{\cup_{j \neq k} \Omega^d_j}$.
5. Build the wideband matrix $\mathbf{B}^{\Omega_k}_{\text{wb}}$ as explained in section 3.
6. Solve the convex optimization problem $(P_1)$ with $\mathbf{B} := \mathbf{B}^{\Omega_k}_{\text{wb}}$.

## 5. EXPERIMENTS ON AUDIO SOURCES

We compared the new wideband CR method with the narrowband CR method in a scenario where two sources are simultaneously active in the time domain. As we consider only one "convolutive" source per mixture, we dropped the source index $(j)$ from the filter for the sake of legibility. The performance measure was the filter SNR corrected from unknown arbitrary time shift and scaling, defined by:

$$\text{SNR} = 10 \log_{10} \left( \frac{\sum_{i,t} \|a_i(t)\|_2^2}{\min_{t',\mu} \sum_{i,t} \left( \|a_i(t) - \mu \cdot \tilde{a}_i(t - t')\|_2^2 \right)} \right)$$

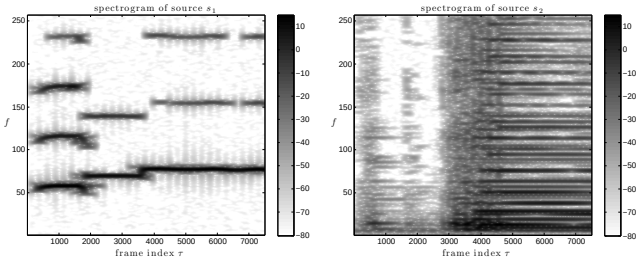
where $\tilde{a}_j$ is the estimated filter.

Note that this implies that our goal is to recover the filters up to a time shift and scaling, and not up to a global filter as targeted by convolutive ICA. To our knowledge, there is no other method that can address this problem when the sparsity $K = \|\mathbf{a}_i\|_0$ of the filters, i.e. the number of nonzero coefficients of $\mathbf{a}_i$, is larger than one[3]. However, in the anechoic case (i.e. when $K = 1$), GCC PHAT [14] performs state-of-the-art results to estimate the delay between the two channels. Then the magnitudes of the peaks can be estimated by averaging the IP of all the TF points of $\Omega$ [1].

The experiment was conducted using the two audio sources whose STFT is displayed on Fig. 1: $s_1$ is a flute sound, while $s_2$ is a guitar sound. The sources are mixed on two channels, $s_1$ with a IP $\theta_1 = 0.2$ radian, $s_2$ with two sparse filters $a_{12}(t)$ and $a_{22}(t)$. We varied the filter sparsity from $K = 1$ to $K = 8$, for a total filter length $L = 256$. For each sparsity, 20 random sparse filters were generated as in [6]. The STFT was computed using a Blackman-Harris window of 512 samples with a one-sample shift between each frame. We blindly built the sets $\Omega_1$ and $\Omega_2$ with $\eta = 0.1$ as explained in Section 4. Fig. 2(b) represents the selected TF points $\Omega_2$. As the number of points in $\Omega_2$ is very large, we only kept the points between $\tau = 4000$ and $\tau = 5000$, which is a segment where the two sources are simultaneously active. We reduced the number of rows of the matrix $\mathbf{B}$ by merging rows corresponding to the same frequency bin $f$. This merging was done by averaging for each frequency bin $f$ the normalized rows of $\mathbf{B}$ corresponding to $f$. The filters were then estimated by solving $(P_1)$ for both the wideband matrix $\mathbf{B}_{\text{wb}}$ and the narrowband matrix $\mathbf{B}_{\text{nb}}$, with the parameter[4] $\epsilon = 0.0006$. Results of the experiment showing the average SNR

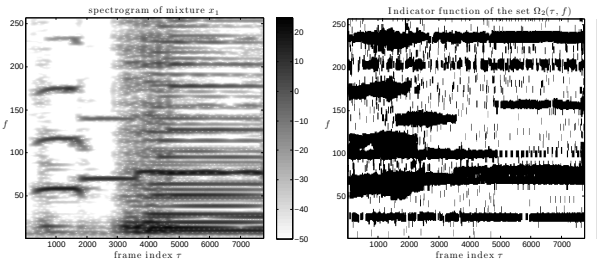
[3]Note that the sparsity of $\mathbf{a}$, which is the concatenation of $\mathbf{a}_1$ and $\mathbf{a}_2$, is thus $2K$.

[4]This value of $\epsilon$ was tuned empirically on two examples of the database, one with a small sparsity and one with a larger sparsity. The same value was then used for all sparsities and all draws of filters.




(a) (b)

**Fig. 1**. Spectrograms of the two sources: (a) source $s_1$ which is a flute playing and (b) source $s_2$ which is a guitar playing.




(a) (b)

**Fig. 2**. (a) Spectrogram of mixture $x_1$. (b) TF mask. The white pixels correspond to the points of the set $\Omega_2$

are plotted in Fig. 3. The wideband approach outperforms the narrowband one by between 5 and 20 dB. The GCC PHAT worked better than the wideband approach in the anechoic case, but if we perform a "debiasing" (DB) step [15] after solving $(P_1)$, i.e. after selecting the support of the $K$ largest coefficients of the estimated filters we solve a least squares problem to readjust those coefficients, then the wideband CR with DB improves the performance by 20 dB for $K < 4$ and outperforms GCC PHAT by more than 10 dB in the anechoic case.

## 6. CONCLUSION

We proposed a method to blindly estimate the sparse filters of a convolutive mixtures of sources. Our approach assumes that the filters are sparse and is based on a wideband formulation of the cross-relation. The approach contains a clustering step followed by a filter estimation step. We illustrated on a stereo mixture that this method is able to blindly estimate the convolutive sparse filters even when the sources overlap in the time-domain. Future work includes an extension of the clustering step so as to deal with more complex mixtures composed of several convolutive sources and the use of a fast solver (instead of `CVX`) for sparse recovery so as to deal with more points and longer filters. We also want to do more extensive experiments to assess the performance of the approach in various audio scenarios.

## 7. REFERENCES


[1] S Arberet, R Gribonval, and F Bimbot, "A robust method to count and locate audio sources in a multichannel underdetermined mixture," *IEEE Trans. on Signal Proc.*, vol. 58, no. 1, pp. 121–133, 2010.

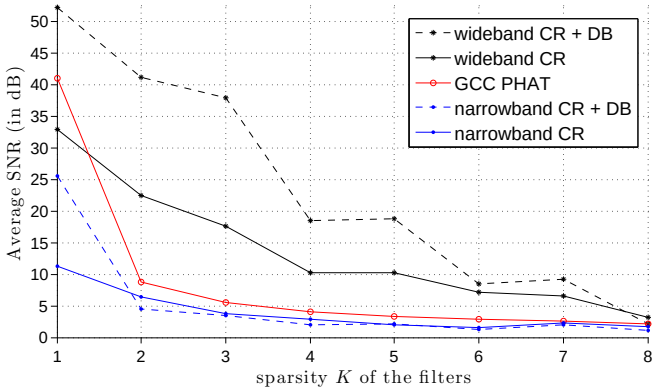




**Fig. 3**. Filters estimation results.


[2] P Comon and C Jutten, *Handbook of Blind Source Separation: Independent component analysis and applications*, Academic Press, 2010.

[3] P.A. Naylor, *Speech dereverberation*, Springer, 2010.

[4] A Aïssa-El-Bey and K Abed-Meraim, "Blind simo channel identification using a sparsity criterion," in *SPAWC*, 2008.

[5] J.J. Fuchs, "Recovery of exact sparse representations in the presence of bounded noise," *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3601–3608, 2005.

[6] P Sudhakar, S Arberet, and R Gribonval, "Double sparsity: Towards blind estimation of multiple channels," in *Proc 9th Int. Conf. LVA/ICA*, 2010.

[7] A. Gorokhov and P. Loubaton, "Subspace-based techniques for blind separation of convolutive mixtures with temporally correlated sources," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 44, no. 9, pp. 813–820, 2002.

[8] S. An, Y. Hua, J.H. Manton, and Z. Fang, "Group decorrelation enhanced subspace method for identifying FIR MIMO channels driven by unknown uncorrelated colored sources," *IEEE Trans. on Signal Proc.*, vol. 53, no. 12, pp. 4429–4441, 2005.

[9] G Xu, H Liu, L Tong, and T Kailath, "A least-squares approach to blind channel identification," *IEEE Trans. on Signal Proc.*, vol. 43, no. 12, pp. 2982 – 2993, 1995.

[10] S. Chen, D.L. Donoho, and M.A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, Jan. 1999.

[11] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 1.21," `http://cvxr.com/cvx`, Feb. 2011.

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

[13] Stéphane Mallat, *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*, Academic Press, 3rd edition, 2008.

[14] C Knapp and G Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. on Acoust., Speech and Signal Proc.*, vol. 24, no. 4, pp. 320 – 327, 1976.

[15] E. Candes and T. Tao, "The Dantzig selector: Statistical estimation when p is much larger than n," *Annals of Statistics*, vol. 35, no. 6, pp. 2313–2351, 2007.