



Beyond sparsity : recovering structured representations by ℓ^1 -minimization and greedy algorithms.

Rémi Gribonval, Morten Nielsen

► To cite this version:

Rémi Gribonval, Morten Nielsen. Beyond sparsity : recovering structured representations by ℓ^1 -minimization and greedy algorithms.. Advances in Computational Mathematics, 2008, 28 (1), pp.23–41. 10.1007/s10444-005-9009-5 . inria-00544767

HAL Id: inria-00544767

<https://inria.hal.science/inria-00544767>

Submitted on 6 Feb 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

BEYOND SPARSITY : RECOVERING STRUCTURED REPRESENTATIONS BY ℓ^1 MINIMIZATION AND GREEDY ALGORITHMS.

*Rémi Gribonval **

IRISA-INRIA
Campus de Beaulieu
F-35042 Rennes Cedex, France

Morten Nielsen

Department of Mathematical Sciences
Aalborg University
Fredrik Bajers Vej 7 G
DK-9220 Aalborg East, Denmark

ABSTRACT

Finding a sparse approximation of a signal from an arbitrary dictionary is a very useful tool to solve many problems in signal processing. Several algorithms, such as Basis Pursuit (BP) and Matching Pursuits (MP, also known as greedy algorithms), have been introduced to compute sparse approximations of signals, but such algorithms *a priori* only provide sub-optimal solutions. In general, it is difficult to estimate how close a computed solution is from the optimal one. In a series of recent results, several authors have shown that both BP and MP can successfully recover a sparse representation of a signal provided that it is sparse enough, that is to say if its support (which indicates where are located the nonzero coefficients) is of sufficiently small size.

In this paper we define identifiable structures that support signals that can be recovered exactly by ℓ^1 minimization (Basis Pursuit) and greedy algorithms. In other words, if the support of a representation belongs to an identifiable structure, then the representation will be recovered by BP and MP. In addition, we obtain that if the output of an arbitrary decomposition algorithm is supported on an identifiable structure, then one can be sure that the representation is optimal within the class of signals supported by the structure.

As an application of the theoretical results, we give a detailed study of a family of multichannel dictionaries with a special structure (corresponding to the representation problem $X = \mathbf{A}S\Phi^T$) often used in, e.g., under-determined source separation problems or in multichannel signal processing. An identifiable structure for such dictionaries is defined using a generalization of Tropp's Babel function which combines the coherence of the mixing matrix \mathbf{A} with that of the time-domain dictionary Φ , and we obtain explicit structure conditions which ensure that both ℓ^1 minimization and a multichannel variant of Matching Pursuit can recover structured multichannel representations. The multichannel Matching Pursuit algorithm is described in detail and we conclude with a discussion of some implications of our results in terms of blind source separation based on sparse decompositions.

1. INTRODUCTION

Approximating a signal with a sparse linear expansion from a dictionary of atoms is a useful tool to solve many problems in signal processing. However, finding the best approximation of a signal from an arbitrary dictionary with a prescribed number of atoms is an NP-hard problem [1] so, in general, one settles for sub-optimal approximations. Several algorithms have been introduced that try to decompose a signal in a sparse way. We mention Matching Pursuit [2, 3], Basis Pursuit [4, 5], but there are many more. A significant problem with such algorithms is that it is

*This work, which was completed while the first author was visiting the Dept. Mathematical Sciences at the University of Aalborg, is supported in part by the Danish Technical Science Foundation, Grant no. 9701481 (WAVES), and by the European Union's Human Potential Programme, under contract HPRN-CT-2002-00285 (HASSIP).

difficult to know how close a computed solution is to the representation which minimizes the approximation error under the sparsity constraint.

Several recent papers [6, 4, 7, 8, 9, 10] have identified situations where algorithms such as Basis Pursuit actually compute an optimal representation of a given signal, in the sense that they solve the best approximation problem under a sparsity constraint. In this paper we build on these results and study identifiable classes of representations that can be recovered exactly using constructive algorithms such as Basis Pursuit or greedy algorithms.

First let us fix the notation that will be used throughout this paper. Let F and G be two finite index sets. Let $\mathcal{H} = \mathbb{C}^G$ be the signal space. A dictionary for \mathcal{H} is a linear map $\mathcal{D}: \mathbb{C}^F \rightarrow \mathbb{C}^G$ from the coefficient space \mathbb{C}^F onto the signal space (note that the theory developed in this paper is also valid if we replace \mathbb{C}^F and \mathbb{C}^G with \mathbb{R}^F and \mathbb{R}^G). The atoms associated with \mathcal{D} are the columns of the matrix representation of \mathcal{D} wrt. the canonical bases for \mathbb{C}^F and \mathbb{C}^G , i.e., $\mathcal{D} = [\mathbf{g}_i]_{i \in F}$. We will always assume that the dictionary is normalized with respect to the ℓ^2 norm, i.e., that $\|\mathbf{g}_i\| = 1$, for $i \in F$. The support of a coefficient sequence $S = (s_i)_{i \in F} \in \mathbb{C}^F$ is defined as $\text{supp}(S) = \{i \in F : s_i \neq 0\} \subseteq F$. With this notations the sparse approximation problem can be expressed as

$$\min_S \|X - \mathcal{D}(S)\| \quad \text{subject to } |\text{supp}(S)| \leq m \quad (1)$$

where $X = (x_i)_{i \in G} \in \mathbb{C}^G$ and $|I|$ denotes the cardinal of the set I .

We generalize this problem by considering, for any family \mathcal{S} of subsets of F , the following *structured approximation problem*, or approximation with structure constraint \mathcal{S}

$$\min_S \|X - \mathcal{D}(S)\| \quad \text{subject to } \text{supp}(S) \in \mathcal{S}. \quad (2)$$

A particular instance of the structured approximation problem is the sparse approximation problem (1), which corresponds to the family $\mathcal{S}_m = \{I \subseteq F : |I| \leq m\}$. That is to say, we simply put as a constraint a bound on the allowed number of nonzero coefficients in S . However, in many cases it also makes sense to consider families \mathcal{S} taking into account not only the sparsity of I but also properties that may be related to the “geometry” of F and G . A typical example that we will consider in details since it appears in multichannel signal processing and blind source separation problems is when $F = [1, N] \times [1, K]$, $G = [1, M] \times [1, T]$ and \mathbb{C}^F (resp. \mathbb{C}^G) can be identified with a linear space of $N \times K$ (resp. $M \times T$) matrices and the action of the dictionary is $\mathcal{D}(S) = \mathbf{A}S\mathbf{\Phi}$. The $M \times N$ matrix \mathbf{A} is usually called the *mixing matrix* while the $K \times T$ matrix $\mathbf{\Phi}$ is itself a dictionary of atomic waveforms. Often, the size of the class \mathcal{S} grows exponentially with the length $|I|$ of its largest elements, and the optimization problem (2), which is combinatorial in nature, is hard to approach directly. To study this problem more closely, we introduce the concept of *identifiable structures*.

Definition 1 A family \mathcal{S} of subsets of F is called an identifiable structure if $\mathcal{D}(S) = \mathcal{D}(S')$ with $\text{supp}(S), \text{supp}(S') \in \mathcal{S}$ implies that $S = S'$.

Notice that for \mathcal{S} an identifiable structure, any subfamily $\mathcal{S}' \subseteq \mathcal{S}$ is also an identifiable structure.

The significance of Definition 1 is the following.

1. if a signal X satisfies the model $X = \mathcal{D}(S)$ with S supported on an identifiable structure \mathcal{S} , then the representation S is the *unique* representation of X supported on \mathcal{S} , and it can be recovered as the unique solution of the optimization problem (2).
2. if an *algorithm* (supposedly computationally efficient) provides *some* representation $X = \mathcal{D}(S_{\text{alg}})$ where S_{alg} is supported on an identifiable structure \mathcal{S} , then one can be sure that this representation is optimal within the class of representations supported by \mathcal{S} , thus bypassing the (generally hard) combinatorial optimization in (2).

Based on existing results [9, 6], it is known that for a general dictionary \mathcal{D} with coherence $\mu(\mathcal{D}) := \max_{k \neq \ell} |\langle \mathbf{g}_k, \mathbf{g}_\ell \rangle|$, the family \mathcal{S}_m with $m < \lfloor (1 + 1/\mu)/2 \rfloor$ is identifiable.

The structure of the paper is as follows. In Section 2 we define two special (abstract) identifiable structures $\mathcal{S}_{LP} \supset \mathcal{S}_G$ which have the following additional interesting property: if a signal X satisfies the *model* $X = \mathcal{D}(S)$ with S supported on an \mathcal{S}_{LP} (resp. \mathcal{S}_G), then the representation S can be recovered by Basis Pursuit (resp. Basis Pursuit and Matching Pursuit(s)). A more explicit (but also smaller) identifiable structure \mathcal{S}_B is defined in Section 2.3 using a generalized version of Tropp's Babel function, see [8]. We conclude Section 2 with a concrete example (based on the results in [9, 8]) of identifiable structure for dictionaries made up of a union of incoherent orthonormal bases. Section 3 contains an application of the theory to a study of the identifiability of structured multichannel representations, with an application to the analysis of sparse underdetermined source separation. We obtain several explicit identifiability conditions, one of which is expressed as follows

Theorem 1 Consider \mathbf{A} and Φ two matrices and let $\mu(\mathbf{A})$ and $\mu(\Phi)$ be their coherence. If a matrix X has two representations $\mathbf{A}S\Phi^T$ and $\mathbf{A}S'\Phi^T$ where

- each column of the matrices S and S' has at most one nonzero entry,
- both S and S' have no more than C columns with nonzero entries,
- the integer C satisfies

$$C < \frac{1}{2} \left(1 + \frac{1}{\mu(\Phi)} \right) - \max \left(0, \frac{\mu(\mathbf{A})}{\mu(\Phi)} - 1 \right), \quad (3)$$

then $S = S'$. Moreover, this unique representation of $X = \mathbf{A}S\Phi^T$ with at most C columns and at most one nonzero entry per column is recovered by the Basis Pursuit and Matching Pursuit(s) algorithms.

We conclude the section with a complete description of the multichannel Matching Pursuit algorithm which can be used to recover structured multichannel representations under the conditions of our theorems.

In Section 4 we conclude the paper by considering identifiable structures that support some signals which cannot be recovered by the Basis Pursuit algorithm. In particular, we give examples of a structure supporting signals that can only be recovered by combinatorial optimization.

2. ABSTRACT IDENTIFIABLE STRUCTURES

In this section we consider identifiable structures as defined in Definition 1. The main goal is to define identifiable structures which support signals that can be recovered by constructive algorithms. We fix the dictionary $\mathcal{D}: \mathbb{C}^F \rightarrow \mathbb{C}^G$ throughout this section.

2.1. Recovery by linear programming

The first identifiable structure we define is not only identifiable : it also allows signal recovery using the Basis Pursuit algorithm based on linear programming. For $I \subset F$, we define

$$P_1(I) = \sup_{Z \in \text{Ker}(\mathcal{D}), Z \neq 0} \frac{\sum_{i \in I} |z_i|}{\|Z\|_1},$$

with $\|Z\|_1 := \sum_{i \in F} |z_i|$ and $\text{Ker}(\mathcal{D})$ the null space of \mathcal{D} .

The following was proven by the authors [10, Corollary 1].

Lemma 1 Let $X = \mathcal{D}(S)$. Let I be such that $\text{supp}(S) \subseteq I$, and suppose

$$P_1(I) := \sup_{Z \in \text{Ker}(\mathcal{D}), Z \neq 0} \frac{\sum_{i \in I} |z_i|}{\|Z\|_1} < \frac{1}{2}. \quad (4)$$

Then S is the unique solution of $\min \sum_{i \in F} |s'_i|$ subject to $X = \mathcal{D}(S')$.

Lemma 1 motivates the following definition.

Definition 2 We define the class \mathcal{S}_{LP} by

$$\mathcal{S}_{LP} = \{I \subseteq F : P_1(I) < \frac{1}{2}\}.$$

The following theorem is an immediate consequence of Lemma 1.

Theorem 2 The structure class \mathcal{S}_{LP} is identifiable: if a signal X has two representations S and S' satisfying $\text{supp}(S), \text{supp}(S') \in \mathcal{S}_{LP}$, then $S = S'$. Moreover, the unique representation of $X = \mathcal{D}(S)$ with $\text{supp}(S) \in \mathcal{S}_{LP}$ is the solution of the ℓ^1 minimization problem

$$\min \sum_{i \in F} |s'_i| \quad \text{subject to } X = \mathcal{D}(S'),$$

and it can therefore be recovered by the Basis Pursuit algorithm.

Proof. Since $\text{supp}(S) \in \mathcal{S}$, S is the unique solution of the ℓ^1 minimization problem, and the same holds for S' . Being both the unique solution of the same ℓ^1 minimization problem, S and S' must coincide. \square

One problem with \mathcal{S}_{LP} is its rather abstract definition: it is not easy to check the condition given by (4). In the following sections we will introduce subclasses of \mathcal{S}_{LP} with more transparent definitions.

2.2. Recovery by greedy algorithms

Now we consider an identifiable structure supporting signals that can be recovered by both Basis Pursuit and Matching Pursuit. The definition of the structure is inspired by the work of Tropp [8]. For $I \subseteq F$ we define (in matrix notations) $\mathcal{D}_I = [\mathbf{g}_i]_{i \in I}$, the restriction of \mathcal{D} to the atoms in I . We let \mathcal{D}_I^+ denote the pseudo-inverse of \mathcal{D}_I , and recall that $\mathcal{D}_I^+ = (\mathcal{D}_I^* \mathcal{D}_I)^{-1} \mathcal{D}_I^*$ (whenever $\mathcal{D}_I^* \mathcal{D}_I$ is invertible) with \mathcal{D}_I^* the adjoint of \mathcal{D}_I . Tropp proved the following result.

Lemma 2 ([8]) Let $X = \mathcal{D}(S)$ and suppose that $\text{supp}(S) = I$ satisfies the Exact Recovery Condition

$$\max_{i \notin I} \|\mathcal{D}_I^+ \mathbf{g}_i\|_1 < 1. \quad (5)$$

Then $P_1(I) < 1/2$, and both Basis Pursuit and Orthogonal Matching Pursuit exactly recover the representation S of X .

In [11], one of the authors extended this result (in some slightly weaker sense) to plain Matching Pursuit and other variants of Matching Pursuit. Guided by Lemma 2 we have the following definition.

Definition 3 We define the class $\mathcal{S}_G \subset \mathcal{S}_{LP}$ by

$$\mathcal{S}_G = \{I \subseteq F : \max_{i \notin I} \|\mathcal{D}_I^+ \mathbf{g}_i\|_1 < 1\}.$$

Note that for $I \in \mathcal{S}_G$ and $J \subseteq I$, it is not clear whether we have $J \in \mathcal{S}_G$, but we do have $J \in \mathcal{S}_{LP}$.

An application of Lemma 2 and [11, Theorem 1] gives.

Theorem 3 The structure class \mathcal{S}_G is identifiable: if a signal X has two representations S and S' satisfying $\text{supp}(S), \text{supp}(S') \in \mathcal{S}_G$, then $S = S'$. The unique representation of $X = \mathcal{D}(S)$ with $\text{supp}(S) \in \mathcal{S}_G$ is recovered by the Basis Pursuit and Matching Pursuit(s) algorithms.

2.3. Identifiable structures defined using the Babel function

It is not always easy to estimate $\max_{i \notin I} \|\mathcal{D}_I^+ \mathbf{g}_i\|_1$ directly, which makes it hard to apply Lemma 2. Below we consider a more explicit condition based on the so-called Babel function. For an individual set I we define the *setwise Babel function*

$$\mu_1(\mathcal{D}, I) := \max_{i \notin I} \sum_{j \in I} |\langle \mathbf{g}_i, \mathbf{g}_j \rangle|. \quad (6)$$

For \mathcal{S} a family of subsets of F , we define the *structured Babel function*

$$\mu_1(\mathcal{D}, \mathcal{S}) := \sup_{I \in \mathcal{S}} \mu_1(\mathcal{D}, I). \quad (7)$$

The structured Babel function $\mu_1(\mathcal{D}, \mathcal{S})$ generalizes the Babel function $\mu_1(\mathcal{D}, m)$ introduced by Tropp in [8]. In fact, let $\mathcal{S}_m = \{I \subseteq F : |I| = m\}$, $m = 1, 2, \dots, |F|$. Then $\mu_1(\mathcal{D}, m) = \mu_1(\mathcal{D}, \mathcal{S}_m)$.

The Babel function can be used to get a sufficient condition implying the Exact Recovery Condition (5). We have the following lemma.

Lemma 3 Suppose $I \subseteq F$ is such that

$$\mu_1(\mathcal{D}, I) + \max_{\ell \in I} \mu_1(\mathcal{D}, I \setminus \{\ell\}) < 1. \quad (8)$$

Then the Exact Recovery Condition $\max_{i \notin I} \|\mathcal{D}_I^+ \mathbf{g}_i\|_1 < 1$ is satisfied and $P_1(I) < 1/2$.

The proof of Lemma 3 follows [8], but we include it for the sake of completeness.

Proof. We have $\mathcal{D}_I^+ = (\mathcal{D}_I^* \mathcal{D}_I)^{-1} \mathcal{D}_I^*$ so

$$\max_{i \notin I} \|\mathcal{D}_I^+ \mathbf{g}_i\|_1 \leq \|(\mathcal{D}_I^* \mathcal{D}_I)^{-1}\|_{\ell^1 \rightarrow \ell^1} \{\max_{i \notin I} \|\mathcal{D}_I^* \mathbf{g}_i\|_1\} \quad (9)$$

We notice that

$$\max_{i \notin I} \|\mathcal{D}_I^* \mathbf{g}_i\|_1 = \max_{i \notin I} \sum_{j \in I} |\langle \mathbf{g}_i, \mathbf{g}_j \rangle| = \mu_1(\mathcal{D}, I). \quad (10)$$

To estimate $\|(\mathcal{D}_I^* \mathcal{D}_I)^{-1}\|_{\ell^1 \rightarrow \ell^1}$ we follow [8] and write $\mathcal{D}_I^* \mathcal{D}_I = Id + A$, where A contains the off-diagonal part of $\mathcal{D}_I^* \mathcal{D}_I$, which is made of inner products $\langle \mathbf{g}_\ell, \mathbf{g}_j \rangle$, $\ell, j \in I$, $\ell \neq j$. Notice that by (8) and a standard estimate of the $\|\cdot\|_{\ell^1 \rightarrow \ell^1}$ operator norm,

$$\|A\|_{\ell^1 \rightarrow \ell^1} = \max_{\ell \in I} \sum_{j \in I \setminus \{\ell\}} |\langle \mathbf{g}_\ell, \mathbf{g}_j \rangle| \leq \max_{\ell \in I} \mu_1(\mathcal{D}, I \setminus \{\ell\}) < 1,$$

so we can expand $(Id + A)^{-1}$ in a Neumann series

$$\begin{aligned} \|(\mathcal{D}_I^* \mathcal{D}_I)^{-1}\|_{\ell^1 \rightarrow \ell^1} &= \|(Id + A)^{-1}\|_{\ell^1 \rightarrow \ell^1} \\ &= \left\| \sum_{k=0}^{\infty} (-A)^k \right\|_{\ell^1 \rightarrow \ell^1} \\ &\leq \sum_{k=0}^{\infty} \|A\|_{\ell^1 \rightarrow \ell^1}^k \\ &= \frac{1}{1 - \|A\|_{\ell^1 \rightarrow \ell^1}} \\ &\leq \frac{1}{1 - \max_{\ell \in I} \mu_1(\mathcal{D}, I \setminus \{\ell\})}. \end{aligned} \quad (11)$$

Hence, combining (8), (9), (10), and (11) gives the estimate

$$\max_{i \notin I} \|\mathcal{D}_I^+ \mathbf{g}_i\|_1 \leq \frac{\mu_1(\mathcal{D}, I)}{1 - \max_{\ell \in I} \mu_1(\mathcal{D}, I \setminus \{\ell\})} < 1.$$

The conclusion then follows directly from Lemma 2. \square

Guided by Lemma 3 we introduce yet a new class.

Definition 4 The class $\mathcal{S}_B \subseteq \mathcal{S}_G$ is defined by

$$\mathcal{S}_B = \{I \subseteq F : \mu_1(\mathcal{D}, I) + \max_{\ell \in I} \mu_1(\mathcal{D}, I \setminus \{\ell\}) < 1\}.$$

Lemma 3 gives the following Theorem.

Theorem 4 The structure class \mathcal{S}_B is identifiable: if a signal X has two representations S and S' satisfying $\text{supp}(S), \text{supp}(S') \in \mathcal{S}_B$, then $S = S'$. The unique representation of $X = \mathcal{D}(S)$ with $\text{supp}(S) \in \mathcal{S}_B$ is recovered by the Basis Pursuit and Matching Pursuit(s) algorithms.

2.4. An example of an explicit identifiable structure

In [9] the authors considered the special case of dictionaries made up of a union of mutually incoherent orthonormal bases $\mathcal{B}_\ell = [\mathbf{g}_{\ell,i}]_{1 \leq i \leq N}$, that is to say $\mathcal{D} = [\mathbf{g}_{\ell,i}]_{(\ell,i) \in F}$ with $F = [1, L] \times [1, N]$. Using the overall mutual coherence $\mu := \max_{(\ell,i) \neq (\ell',j)} |\langle \mathbf{g}_{\ell,i}, \mathbf{g}_{\ell',j} \rangle|$ between atoms coming from different bases, they proved the following result [9, Theorem 2], which we have reworded here using the concepts introduced in the present paper.

Lemma 4 Let $I = \bigcup_{\ell=1}^L (\{\ell\} \times I_\ell) \subseteq F$ be an index set (I_ℓ indexes the atoms from the ℓ -th basis \mathcal{B}_ℓ which belong to I) and denote $|I_{\ell^*}| = \max_{1 \leq \ell \leq L} |I_\ell|$. If

$$\sum_{\ell=1}^L \frac{\mu |I_\ell|}{1 + \mu |I_\ell|} < \frac{1}{2(1 + \mu |I_{\ell^*}|)} + \frac{\mu |I_{\ell^*}|}{1 + \mu |I_{\ell^*}|} = \frac{1 + 2\mu |I_{\ell^*}|}{2(1 + \mu |I_{\ell^*}|)} \quad (12)$$

then $P_1(I) < 1/2$.

Tropp showed [8] that the Exact Recovery Condition (5) also holds when the condition (12) is satisfied. In the special case $L = 2$, i.e., the case of a union of two orthonormal bases, condition (12) can be restated more simply as

$$2\mu^2 |I_1| |I_2| + \mu \min(|I_1|, |I_2|) < 1, \quad (13)$$

which is the major result of [7].

Defining the class¹ $\mathcal{S}_{\text{UONB}} \subseteq \mathcal{S}_G$ of index sets $I = \bigcup_{\ell=1}^L \{\ell\} \times I_\ell \subseteq F$ which satisfy (12) with $|I_{\ell^*}| = \max_{1 \leq \ell \leq L} |I_\ell|$, we conclude from the previous machinery that this class is identifiable. In the simple case $L = 2$, this implies in particular that if some algorithm provides a representation $X = \mathcal{B}_1 S_1 + \mathcal{B}_2 S_2$ of a signal where $I_\ell = \text{supp}(S_\ell)$, $\ell = 1, 2$ satisfy (13), then this representation is the only one with this property and it is the one which minimizes the ℓ^1 norm over all possible representations of X .

The problem of finding an optimal decomposition of a signal in a union of orthonormal bases arises in audio compression: multilayer “hybrid” representations with a tonal part and a transient part (plus a residual) were proposed in [12], where the tonal part has a sparse representation in a local cosine basis (MDCT) while the transient part is sparse in a wavelet basis (DWT). The considered bases have a rather high overall mutual coherence, so the theory of identifiable representations in incoherent bases only yields trivial results (typically, if I satisfies (13) then $I_{\ell^*}^*$ must

¹It is not hard to show that there exist constants c, C such that $\mathcal{S}_{[c/\mu]} \subseteq \mathcal{S}_{\text{UONB}} \subseteq \mathcal{S}_{[C/\mu]}$.

correspond to a single atom!). However, one can group the elements of both bases into time-frequency blocks of 2^j elements which essentially live in the time interval $[n2^j, (n+1)2^j]$ and the frequency band $[2^j, 2^{j+1}]$, and the coherence within such a block is of the order of $2^{-j/2}$. If an audio signal has a hybrid representation where on each time-frequency block the numbers $|I_{\text{MDCT}}|$ and $|I_{\text{DWT}}|$ of elements coming from each basis satisfy a condition of the type (13), then it seems reasonable to extrapolate from the theoretical results that Basis Pursuit and Matching Pursuit will stably (almost) recover it. The authors are currently investigating a possible formal proof of this result based on an estimate of the generalized Babel function for sets I with the described structure.

3. IDENTIFIABILITY OF STRUCTURED MULTICHANNEL REPRESENTATIONS

In this section we investigate a particular type of overcomplete dictionary \mathcal{D} which corresponds to problems often encountered in under-determined source separation problems or multichannel signal processing. Typically, the blind source separation (BSS) problem –in its linear instantaneous instantiation– consists in finding a factorization of a matrix X of observed data (which rows $x_m(t)$ are signals) as $X = \mathbf{A}C$ where \mathbf{A} is an unknown mixing matrix and the rows $c_n(t)$ of C are the unknown source signals, generally assumed to be realizations of independent random variables. The traditional approach to BSS is based on Independent Component Analysis [13] but Pearlmutter and Zibulevsky [14] introduced a new approach based on sparse representations: the sources $c_n(t)$ are modeled as sparse expansions from a dictionary Φ of atoms $\phi_k(t)$, that is to say we assume $c_n(t) = \sum_k s_{n,k} \phi_k(t)$, or in matrix form $C = S\Phi^T$ with S a matrix of sparse coefficients. Overall, BSS is performed by jointly optimizing \mathbf{A} and S to get maximally sparse coefficients S in the representation $X = \mathbf{A}S\Phi^T$. Even though the mixing matrix \mathbf{A} is unknown, a common approach is to perform either sequentially, or iteratively

- a *learning step* where an estimate of \mathbf{A} is computed or updated;
- an *inference step* where the source coefficients are computed using the current estimate of \mathbf{A} .

Our goal in this section is to show that, under sufficient structure of the coefficient matrix S , the inference step can be successfully performed with Basis Pursuit or Matching Pursuit(s) if the mixing matrix was correctly estimated during the learning step. Looking back at our formalism from the previous section, the data vector X is “folded” into a matrix $X = (x_{mt})$ with M rows which we wish to represent as

$$X = \mathbf{A}S\Phi^T, \quad (14)$$

with $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_N]$ a $M \times N$ *mixing matrix*, $S = (s_{nk})$ a $N \times K$ matrix of coefficients and $\Phi = [\phi_1, \dots, \phi_K]$ a *monochannel dictionary* of K *monochannel atoms*. In other words, we consider $F = [1, N] \times [1, K]$. In the unfolded world where the data X and (especially) the coefficients S are considered rather as vectors than matrices we have an “unfolded” *multichannel dictionary*

$$\mathcal{D}(S) := \mathbf{A}S\Phi^T = \sum_{n=1}^N \sum_{k=1}^K s_{nk} \mathbf{a}_n \phi_k^T \quad (15)$$

which is made of *multichannel atoms* $\mathbf{g}_{n,k} = \mathbf{a}_n \phi_k^T$. Measuring the approximation error between X and $\mathcal{D}(C)$ with the Frobenius norm $\|X - \mathcal{D}(C)\|_{\text{Frob}}$ corresponds to considering X as living in a Hilbert space equipped with the inner product $[X, Y] = \text{Trace}(XY^H)$. It is easy to check that for this inner product we have

$$[\mathbf{g}_{n,k}, \mathbf{g}_{n',k'}] = \langle \mathbf{a}_n, \mathbf{a}_{n'} \rangle \cdot \langle \phi_k, \phi_{k'} \rangle \quad (16)$$

with $\langle \cdot, \cdot \rangle$ the standard inner product between vectors.

Without any originality we will assume that the columns \mathbf{a}_n of the mixing matrix, just as the atoms ϕ_k of the dictionary, are normalized. In BSS terms, this is one way of fixing the gain indeterminacy of the source estimation problem, and it

will make it possible to estimate the generalized Babel function $\mu_1(\mathcal{D}, I)$ of the multichannel dictionary in terms of the standard Babel functions $\mu_1(\mathbf{A}, R)$ and $\mu_1(\Phi, C)$, where R and C are related to the number (or the size) of rows / columns of I . We will use this estimate together with the theory of the previous section to obtain identifiability results for certain structured representations of multichannel data. We will then proceed to give a sufficiently explicit description of multichannel Matching Pursuit, which is able to recover such representations.

3.1. Estimate of the Babel function

For any index set $I \subset [1, M] \times [1, K]$ we denote

$$\begin{aligned} \text{col}_k(I) &:= \{n, (n, k) \in I\}, & \text{Cols}(I) &:= \{k, \text{col}_k(I) \neq \emptyset\}, \\ \text{row}_n(I) &:= \{k, (n, k) \in I\}, & \text{Rows}(I) &:= \{n, \text{row}_n(I) \neq \emptyset\}. \end{aligned}$$

We have the following lemma:

Lemma 5 *Let I be an index set with at most C non-empty columns, each of size at most R , that is to say assume $|\text{Cols}(I)| \leq C$ and $\max_k |\text{col}_k(I)| \leq R$. Then we have*

$$\mu_1(\mathcal{D}, I) \leq \max \left(\mu_1(\Phi, C) \cdot (1 + \mu_1(\mathbf{A}, R - 1)), \mu_1(\mathbf{A}, R) + \mu_1(\Phi, C - 1) \cdot (1 + \mu_1(\mathbf{A}, R - 1)) \right). \quad (17)$$

Proof. First, we write

$$\sum_{(n', k') \in I} |[\mathbf{g}_{n', k'}, \mathbf{g}_{n, k}]| = \sum_{k' \in \text{Cols}(I)} |\langle \phi_k, \phi_{k'} \rangle| \sum_{n' \in \text{col}_{k'}(I)} |\langle \mathbf{a}_n, \mathbf{a}_{n'} \rangle|.$$

Then, for $(n, k) \notin I$, there are only two possible cases : either $k \notin \text{Cols}(I)$, or $k \in \text{Cols}(I)$ with $n \notin \text{col}_k(I)$. We can estimate the supremum in the first case

$$\begin{aligned} \sup_{k \notin \text{Cols}(I)} \sup_n \sum_{(n', k') \in I} |[\mathbf{g}_{n', k'}, \mathbf{g}_{n, k}]| &\leq \sup_{k \notin \text{Cols}(I)} \sum_{k' \in \text{Cols}(I)} |\langle \phi_k, \phi_{k'} \rangle| \cdot \sup_n \sum_{n' \in \text{col}_{k'}(I)} |\langle \mathbf{a}_n, \mathbf{a}_{n'} \rangle| \\ &\leq \sup_{k \notin \text{Cols}(I)} \sum_{k' \in \text{Cols}(I)} |\langle \phi_k, \phi_{k'} \rangle| \cdot (1 + \mu_1(\mathbf{A}, R - 1)) \\ &= \mu_1(\Phi, C) \cdot (1 + \mu_1(\mathbf{A}, R - 1)) \end{aligned}$$

and in the second case

$$\begin{aligned} \sup_{k \in \text{Cols}(I)} \sup_{n \notin \text{col}_k(I)} \sum_{(n', k') \in I} |[\mathbf{g}_{n', k'}, \mathbf{g}_{n, k}]| &\leq \sup_{k \in \text{Cols}(I)} \sum_{k' \in \text{Cols}(I)} |\langle \phi_k, \phi_{k'} \rangle| \cdot \sup_{n \notin \text{col}_k(I)} \sum_{n' \in \text{col}_{k'}(I)} |\langle \mathbf{a}_n, \mathbf{a}_{n'} \rangle| \\ &\leq \sup_{k \in \text{Cols}(I)} \left(|\langle \phi_k, \phi_k \rangle| \cdot \sup_{n \notin \text{col}_k(I)} \sum_{n' \in \text{col}_k(I)} |\langle \mathbf{a}_n, \mathbf{a}_{n'} \rangle| \right. \\ &\quad \left. + \sum_{k' \in \text{Cols}(I) \setminus \{k\}} |\langle \phi_k, \phi_{k'} \rangle| \cdot \sup_{n \notin \text{col}_k(I)} \sum_{n' \in \text{col}_{k'}(I)} |\langle \mathbf{a}_n, \mathbf{a}_{n'} \rangle| \right) \\ &\leq \sup_{k \in \text{Cols}(I)} \left(\mu_1(\mathbf{A}, R) + \sum_{k' \in \text{Cols}(I) \setminus \{k\}} |\langle \phi_k, \phi_{k'} \rangle| \cdot (1 + \mu_1(\mathbf{A}, R - 1)) \right) \\ &= \mu_1(\mathbf{A}, R) + \mu_1(\Phi, C - 1) \cdot (1 + \mu_1(\mathbf{A}, R - 1)) \end{aligned}$$

Since the supremum over $(n, k) \notin I$ is the maximum of the suprema corresponding to the two possible cases, overall we obtain the claimed result (17). \square

3.2. Simplified identifiability conditions for structured multichannel representations

In many underdetermined audio BSS applications where the problem of representing $X = \mathbf{A}S\Phi^T$ arose in practice, the dictionary Φ^T is simply chosen to be a (nonredundant) orthonormal Modified Discrete Cosine Transform (MDCT) basis [14]. In this case, the Babel function $\mu_1(\Phi, C)$ is zero and Lemma 5 gives the estimate $\mu_1(\mathcal{D}, I) \leq \mu_1(\mathbf{A}, R)$ whenever $\max_k |\text{col}_k(I)| \leq R$. Applying the general results from the previous section we get the sufficient recovery condition $\mu_1(\mathbf{A}, R) + \mu_1(\mathbf{A}, R - 1) < 1$ of Tropp [15]: S can be recovered provided that each of its columns is sparse enough.

When Φ is overcomplete, $\mu_1(\Phi, C)$ is no longer zero and it has to be taken into account. In BSS, it is common [16] to assume that at most one source $c_n(t)$ uses each given atom ϕ_k in its representation. The heuristic argument is that the sources have mutually independent and sparse representations, so the probability that two sources activate the same atom is small [17]. This corresponds to considering index sets I in Lemma 5 with $R = 1$. Theorem 1 is simply the consequence of Lemma 5 together with the general results from the previous section in this special case. To state it in its simple form, we replaced the Babel functions $\mu_1(\mathbf{A}, R)$ and $\mu_1(\Phi, C)$ with their upper estimates $\mu_1(\mathbf{A}, R) \leq R \cdot \mu(\mathbf{A})$ and $\mu_1(\Phi, C) \leq C \cdot \mu(\Phi)$ (see [8]) in terms of the *coherence*

$$\begin{aligned}\mu(\mathbf{A}) &:= \mu_1(\mathbf{A}, 1) = \max_{n \neq n'} |\langle \mathbf{a}_n, \mathbf{a}_{n'} \rangle| \\ \mu(\Phi) &:= \mu_1(\Phi, 1) = \max_{k \neq k'} |\langle \phi_k, \phi_{k'} \rangle|.\end{aligned}$$

For the convenience of the reader, let us state Theorem 1 again right now.

Theorem 5 *If a multichannel signal X has two representations $\mathbf{A}S\Phi^T$ and $\mathbf{A}S'\Phi^T$ where each column of S and S' has at most one nonzero entry and both S and S' have no more than C columns with nonzero entries where the integer C satisfies*

$$C < \frac{1}{2} \left(1 + \frac{1}{\mu(\Phi)} \right) - \max \left(0, \frac{\mu(\mathbf{A})}{\mu(\Phi)} - 1 \right) \quad (18)$$

then $S = S'$. Moreover, this unique representation of $X = \mathbf{A}S\Phi^T$ with at most C columns and at most one nonzero entry per column is recovered by the Matching Pursuit(s) and Basis Pursuit algorithms.

Note that in statistical words, Basis Pursuit corresponds to the Maximum Likelihood estimator of S under a Laplacian model, and we have just showed that it will recover the true S provided that it has the right *structure*: each atom can be activated by at most one source, and the total number of atoms activated by the sources altogether cannot exceed a certain limit. In the rare practical cases where the coherence $\mu(\mathbf{A})$ of the mixing matrix would be smaller than that of the dictionary $\mu(\Phi)$, the limit on the number of activated atoms matches the sparsity limit $(1 + 1/\mu(\Phi))/2$ that guarantees the recovery of a representation of a monochannel signal with Φ [6, 9]. As could be expected, the limit becomes more restrictive when the mixing matrix has a larger coherence than the dictionary.

Proof. Consider the structure class

$$\mathcal{S}_{1,C}^{\text{col}} := \left\{ I, |\text{Cols}(I)| \leq C, \max_k |\text{col}_k(I)| \leq 1 \right\} \quad (19)$$

of index sets I with at most C non-empty columns, each of size at most 1. For any $I \in \mathcal{S}_{1,C}^{\text{col}}$ by Lemma 5

$$\begin{aligned}\mu_1(\mathcal{D}, I) &\leq \max \left(\mu_1(\Phi, C), \mu_1(\mathbf{A}, 1) + \mu_1(\Phi, C - 1) \right) \leq \max \left(C \cdot \mu(\Phi), (C - 1) \cdot \mu(\Phi) + \mu(\mathbf{A}) \right) \\ &\leq C \cdot \mu(\Phi) + \max \left(0, \mu(\mathbf{A}) - \mu(\Phi) \right).\end{aligned}$$

Moreover, for $\ell \in I$ we have $I \in \mathcal{S}_{1,C-1}^{\text{col}}$, so by using a similar estimate for $\mu_1(\mathcal{D}, I \setminus \{\ell\})$ we get

$$\mu_1(\mathcal{D}, I) + \max_{\ell \in I} \mu_1(\mathcal{D}, I \setminus \{\ell\}) \leq (2C - 1) \cdot \mu(\Phi) + 2 \cdot \max \left(0, \mu(\mathbf{A}) - \mu(\Phi) \right).$$

Under the assumption (18) we obtain $\mu_1(\mathcal{D}, I) + \max_{\ell \in I} \mu_1(\mathcal{D}, I \setminus \{\ell\}) < 1$ and it follows that $\mathcal{S}_{1,C}^{\text{col}} \subseteq \mathcal{S}_G$ is an identifiable structure. \square

3.3. General multichannel identifiability results

Based on Lemma 5 and the general theory developed in the previous sections, we can also get more general theorems (not restricted to $R = 1$ nonzero entry per columns of S) on the recovery of structured multichannel representations by ℓ^1 minimization and greedy algorithms.

Theorem 6 *If a multichannel signal X has two representations $\mathbf{A}S\Phi^T$ and $\mathbf{A}S'\Phi^T$ where the pair $(R, C) := (\max_k |\text{col}_k(I)|, |\text{Cols}(I)|)$ satisfies*

$$\max \left(\mu_1(\Phi, C) \cdot (1 + \mu_1(\mathbf{A}, R - 1)), \mu_1(\mathbf{A}, R) + \mu_1(\Phi, C - 1) \cdot (1 + \mu_1(\mathbf{A}, R - 1)) \right) < \frac{1}{2} \quad (20)$$

for both $I = \text{support}(S)$ and $I = \text{support}(S')$, then $S = S'$. The unique representation of $X = \mathbf{A}S\Phi^T$ for which the pair (R, C) satisfies (20) is recovered by the Matching Pursuit(s) and Basis Pursuit algorithms.

Proof. Consider the structure class

$$\mathcal{S}_{R,C}^{\text{col}} := \left\{ I, |\text{Cols}(I)| \leq C, \max_k |\text{col}_k(I)| \leq R \right\} \quad (21)$$

of index sets I with at most C non-empty columns, each of size at most R . If (R, C) satisfies (20), then by Lemma 5 we have for any $I \in \mathcal{S}_{R,C}^{\text{col}}$

$$\mu_1(\mathcal{D}, I) + \max_{\ell \in I} \mu_1(\mathcal{D}, I \setminus \{\ell\}) \leq 2 \mu_1(\mathcal{D}, I) < 1.$$

It follows that $\mathcal{S}_{R,C}^{\text{col}} \subseteq \mathcal{S}_G$ is an identifiable structure. Considering E^{col} the set of all pairs (R, C) which satisfy (20) we have indeed

$$\mathcal{S}_{\text{Mult}}^{\text{col}} := \bigcup_{(R,C) \in E^{\text{col}}} \mathcal{S}_{R,C}^{\text{col}} \subseteq \mathcal{S}_G$$

which means that $\mathcal{S}_{\text{Mult}}^{\text{col}}$ itself is an identifiable structure and each representation S supported in $\mathcal{S}_{\text{Mult}}^{\text{col}}$ is unique and can be recovered by the Matching Pursuit(s) algorithm and the Basis Pursuit algorithm, the latter corresponding to the ℓ^1 minimization problem

$$\min \sum_{n,k} |s'_{nk}| \quad \text{subject to} \quad X = \mathbf{A}S'\Phi^T \quad (22)$$

\square

Exchanging the role of rows and columns we immediately get a similar result

Corollary 1 *If a multichannel signal X has two representations $\mathbf{A}S\Phi^T$ and $\mathbf{A}S'\Phi^T$ where the pair $(R, C) := (|\text{Rows}(I)|, \max_n |\text{row}_n(I)|)$ satisfies*

$$\max \left(\mu_1(\mathbf{A}, R) \cdot (1 + \mu_1(\Phi, C - 1)), \mu_1(\Phi, C) + \mu_1(\mathbf{A}, R - 1) \cdot (1 + \mu_1(\Phi, C - 1)) \right) < \frac{1}{2} \quad (23)$$

for both $I = \text{support}(S)$ and $I = \text{support}(S')$, then $S = S'$. The unique representation of $X = \mathbf{A}S\Phi^T$ for which the pair (R, C) satisfies (20) is recovered by the Matching Pursuit(s) and Basis Pursuit algorithms.

Notice that this corresponds to the identifiability of $\mathcal{S}_{\text{Mult}}^{\text{row}} := \bigcup_{(R,C) \in E^{\text{row}}} \mathcal{S}_{R,C}^{\text{row}} \subseteq \mathcal{S}_G$ where the definition of E^{row} and $\mathcal{S}_{R,C}^{\text{row}}$ mimics those of E^{col} and $\mathcal{S}_{R,C}^{\text{col}}$. It follows immediately that the structure $\mathcal{S}_{\text{Mult}}^{\text{row}} \cup \mathcal{S}_{\text{Mult}}^{\text{col}} \subseteq \mathcal{S}_G$ is identifiable and can be recovered by the Matching Pursuit(s) and Basis Pursuit algorithms. We get the corollary:

Corollary 2 *Consider two representations $\mathbf{A}\mathbf{S}\Phi^T$ and $\mathbf{A}\mathbf{S}'\Phi^T$ of the same multichannel signal X , and let $I = \text{support}(S)$, $I' = \text{support}(S')$. If $(\max_k |\text{col}_k(I)|, |\text{Cols}(I)|)$ satisfies (20) or $(|\text{Rows}(I)|, \max_n |\text{row}_n(I)|)$ satisfies (23), and if the same holds for I' , then*

- $S = S'$;
- S can be recovered by the Matching Pursuit(s) and Basis Pursuit algorithms.

3.4. Multichannel greedy algorithms

It might not be obvious what multichannel greedy algorithms look like, and since they are proved to perform well on the structured representations we have just described, it is worth describing them in more details. In the multichannel setting, the greedy algorithm starts with $X^{(0)} = X$ and iteratively computes residuals $X^{(J)}$ as follows.

1. Compute the inner product

$$[X^{(J)}, \mathbf{g}_{n,k}] = \text{Trace}(X^{(J)} \overline{\phi_k} \mathbf{a}_n^T) = \overline{\mathbf{a}_n^T} X^{(J)} \overline{\phi_k}, \quad 1 \leq n \leq N, 1 \leq k \leq K. \quad (24)$$

2. Select an atom \mathbf{g}_{n_J, k_J} that reaches the maximum absolute value of this inner product.
3. Compute the new residual. This step has essentially two flavors

- (a) Matching Pursuit (MP):

$$X^{(J+1)} = X^{(J)} - [X^{(J)}, \mathbf{g}_{n_J, k_J}] \cdot \mathbf{g}_{n_J, k_J} = X^{(J)} - [X^{(J)}, \mathbf{g}_{n_J, k_J}] \cdot \mathbf{a}_{n_J} \phi_{k_J}^T; \quad (25)$$

- (b) Orthonormal Matching Pursuit (OMP):

$$X^{(J+1)} = X - P_J(X) \quad (26)$$

where P_J is the orthogonal projector onto the linear span of the atoms \mathbf{g}_{n_j, k_j} , $0 \leq k \leq J$.

4. Test if a stopping criterion is reached, else increment J and go back to Step 1.

We will not document here the multichannel orthonormal projection (26) for OMP, rather we give a short but explicit description of standard Matching Pursuit in the multichannel framework, which should make it fairly easy to implement. Writing the multichannel residual $X^{(J)}$ as a collection of M rows, each of which is a signal $x_m^{(J)}$, we observe that the multichannel inner products $[X^{(J)}, \mathbf{g}_{n,k}]$ are equal to

$$[X^{(J)}, \mathbf{g}_{n,k}] = \langle y_n^{(J)}, \phi_k \rangle, \quad 1 \leq n \leq N, 1 \leq k \leq K. \quad (27)$$

with $y_n^{(J)} := \sum_{m=1}^M \overline{a_{m,n}} x_m^{(J)}$ $1 \leq n \leq N$. As a result, multichannel MP can be implemented as follows:

1. Compute the signals (which correspond to the rows of the multichannel data $Y := \mathbf{A}^H X$)

$$y_n^{(0)} := \sum_{m=1}^M \overline{a_{m,n}} x_m^{(0)} \quad 1 \leq n \leq N. \quad (28)$$

2. Compute the inner products $[X^{(J)}, \mathbf{g}_{n,k}]$ using (27) and select a pair (n_J, k_J) which maximizes $|[X^{(J)}, \mathbf{g}_{n,k}]|$.
3. Update the residuals channel-wise as

$$y_n^{(J+1)} = y_n^{(J)} - [X^{(J)}, \mathbf{g}_{n_J, k_J}] \cdot \left(\sum_{m=1}^M \overline{a_{m,n}} a_{m, n_J} \right) \cdot \phi_{k_J}, \quad 1 \leq n \leq N \quad (29)$$

$$x_m^{(J+1)} = x_m^{(J)} - [X^{(J)}, \mathbf{g}_{n_J, k_J}] \cdot a_{m, n_J} \cdot \phi_{k_J}, \quad 1 \leq m \leq M. \quad (30)$$

4. Test if a stopping criterion is reached, else increment J and go back to Step 2.

One should note that the multichannel Matching Pursuit described here is different from greedy algorithms for *simultaneous approximation* of several channels [18, 19, 17, 15, 20]: the latter are not even aware of a mixing matrix, and their goal is to represent the data as $X = C\Phi^T$ with an $M \times K$ coefficient matrix C rather than $X = \mathbf{A}S\Phi^T$. In these algorithms, the atom ϕ_{k_J} which is selected at each step is the one which reaches the maximum value of the *multichannel energy*

$$\sum_{m=1}^M |\langle x_m^{(J)}, \phi_k \rangle|^2$$

(or more generally of some multichannel norm $\sum_{m=1}^M |\langle x_m^{(J)}, \phi_k \rangle|^p$). As for the residual, it is updated channel-wise as

$$x_m^{(J+1)} = x_m^{(J)} - \langle x_m^{(J)}, \phi_{k_J} \rangle \phi_{k_J}.$$

Even though both versions of “multichannel Matching Pursuit” enjoy the usual convergence properties (the residual $\|X^{(J)}\|_{\text{Frob}}$ tends to zero [3, 21, 18, 19]), our recovery results only hold for the version we have described since the other is not even *aware* of a mixing matrix. Greedy algorithms for simultaneous sparse approximation also enjoy other types of recovery properties, see [15].

4. BEYOND ℓ^1 -IDENTIFIABLE STRUCTURES

The largest identifiable structure \mathcal{S}_{LP} introduced in Section 2 is characterized by the fact that it supports signals that can be recovered exactly by ℓ^1 minimization. Below we will consider other identifiable structures supporting signals that are not necessarily recoverable by ℓ^1 minimization.

4.1. ℓ^0 -identifiability without ℓ^1 -identifiability

Let \mathcal{D} be a dictionary, and recall that $\mathcal{S}_m = \{I \subseteq F : |I| \leq m\}$. Let $m_0(\mathcal{D})$ be the largest integer m satisfying: for all $S \in \mathbb{C}^F$ with $\text{supp}(S) \in \mathcal{S}_m$, S is the unique minimizer of $\min_{C \in \mathbb{C}^F} \|C\|_0$ subject to $\mathcal{D}(C) = \mathcal{D}(S)$, where $\|C\|_0 := |\text{supp}(C)|$. Similarly, let $m_1(\mathcal{D})$ be the largest integer m satisfying: for all $S \in \mathbb{C}^F$ with $\text{supp}(S) \in \mathcal{S}_m$, S is the unique minimizer of $\min_{C \in \mathbb{C}^F} \sum_{i \in F} |C_i|$ subject to $\mathcal{D}(C) = \mathcal{D}(S)$. A detailed discussion of the numbers $m_0(\mathcal{D})$ and $m_1(\mathcal{D})$ can be found in [10]. In particular, it was shown in [10] that $m_0(\mathcal{D}) \geq m_1(\mathcal{D})$ and that $m_0(\mathcal{D}) = \lfloor Z_0(\mathcal{D})/2 \rfloor$, where $Z_0(\mathcal{D}) := \min_{z \in \text{Ker}(\mathcal{D}), z \neq 0} \|z\|_0$ is the so-called spark of \mathcal{D} [6]. Using the same reasoning as in the proof of Theorem 2, we deduce that $\mathcal{S}_{m_1(\mathcal{D})} \subseteq \mathcal{S}_{m_0(\mathcal{D})}$ are identifiable structures. We also notice for dictionaries with $m_0(\mathcal{D}) > m_1(\mathcal{D})$, there are elements supported on $\mathcal{S}_{m_0(\mathcal{D})}$ which *cannot* be recovered exactly by ℓ^1 minimization, but they are recoverable by ℓ^0 minimization (i.e., by a costly combinatorial optimization).

4.2. Example: union of bases

Let us consider a specific example where $m_0(\mathcal{D}) \gg m_1(\mathcal{D})$. An example is given in [22] of a pair of orthonormal bases $\mathcal{B}_1, \mathcal{B}_2$ for \mathbb{R}^N , $N = 2^{2j+1}$, for which the coherence $\mu([\mathcal{B}_1, \mathcal{B}_2]) = 1/\sqrt{N}$ is minimum and

$$m_1([\mathcal{B}_1, \mathcal{B}_2]) = \lfloor (\sqrt{2} - 1/2)/\mu \rfloor = \lfloor (\sqrt{2} - 1/2)\sqrt{N} \rfloor \approx 0.914\sqrt{N}.$$

However, for a general union of two orthonormal bases with mutual coherence $\mu = 1/\sqrt{N}$,

$$m_0([\mathcal{B}_1, \mathcal{B}_2]) \geq \lfloor 1/\mu \rfloor = \lfloor \sqrt{N} \rfloor > 0.914\sqrt{N},$$

see [7]. For such a dictionary, there is a large family of signals that can be recovered exactly by ℓ^0 minimization but not by ℓ^1 minimization.

4.3. Multichannel case

Returning to the multichannel case, let us see what can be said when condition (20) is weakened.

Theorem 7 *Keep the notations of Theorem 6 and assume that, instead of (20), we have*

$$\max \left(\mu_1(\Phi, 2C) \cdot (1 + \mu_1(\mathbf{A}, 2R - 1)), \mu_1(\mathbf{A}, 2R) + \mu_1(\Phi, 2C - 1) \cdot (1 + \mu_1(\mathbf{A}, 2R - 1)) \right) < 1 \quad (31)$$

Then, X has no other representation S' with $|\text{Cols}(I')| \leq C$ and $\max_k |\text{col}_k(I')| \leq R$, where $I' = \text{supp}(S')$.

Note that this result is much weaker than what we have with the stronger requirement (20):

1. we do not claim that ℓ^1 minimization or a greedy algorithm will recover S .
2. S is only unique within the representations constrained by $|\text{Cols}(I')| \leq C$ and $\max_k |\text{col}_k(I')| \leq R$, with $R := \max_k |\text{col}_k(I)|$, $C := |\text{Cols}(I)|$ and $I := \text{supp}(S)$. This does not rule out the fact that there could be another representation S' for which $I' := \text{supp}(S')$, $R' := \max_k |\text{col}_k(I')|$ and $C' := |\text{Cols}(I')|$ also satisfy (31).
3. neither can we mix as simply as before the role of columns and rows to get a result similar to Corollary 2.

This is only natural, since condition (31) is indeed (although not obviously ...) a weaker requirement than (20).

Proof. Consider S' such a representation and let J be the support of $S - S'$. From the assumptions, we easily get that $J \subset I \cup I' \in \mathcal{S}_{2R, 2C}^{\text{col}}$ as defined in (21). Based on Lemma 5 and the hypothesis (31) we conclude that $\mu_1(\mathcal{D}, J) < 1$. This implies (by an easy adaptation of [15, Lemma 2.3]) that for any sequence Z supported in J , $\|\mathcal{D}(Z)\|_{\text{Frob}}^2 \geq (1 - \mu_1(\mathcal{D}, J))\|Z\|_{\text{Frob}}^2$. Since $Z = S - S'$ is supported in J and $\mathcal{D}(Z) = \mathcal{D}(S) - \mathcal{D}(S') = 0$ we conclude that $Z = 0$, that is to say $S' = S$. \square

5. CONCLUSION

We have introduced the notion of an identifiable structure for a general overcomplete dictionary of atoms. Several (abstract) identifiable structures supporting signals that can be recovered exactly by ℓ^1 minimization and greedy algorithms are considered. For signals supported on such an identifiable structure, the Basis Pursuit/Matching Pursuit algorithms recover the unique sparsest representation of the signal.

A particular family of structured dictionaries corresponding to multichannel representations is studied in detail. Such dictionaries are often used in under-determined source separation problems or in multichannel signal processing. A corresponding identifiable structure is defined using a generalized version of Tropp's Babel function. Signals

supported on this multichannel structure can be recovered exactly by both the Basis Pursuit and the multichannel greedy algorithm.

We are currently investigating several extensions or consequences of the results presented in this paper. Besides estimating the generalized Babel function for unions of MDCT and wavelet bases, our main goal is now to define and study structures that are *robustly identifiable* in the presence of noise / approximation error, since it seems a necessary step to make identifiability results usable in practice. Similarly, to better understand sparsity based source separation, it is worth understanding the robustness of the sparse representation estimation under errors in the mixing matrix estimate.

6. REFERENCES

- [1] G. Davis, S. Mallat, and M. Avellaneda, “Adaptive greedy approximations,” *Constr. Approx.*, vol. 13, no. 1, pp. 57–98, 1997.
- [2] Geoffrey Davis, Stephane Mallat, and Zhifeng Zhang, “Adaptive time-frequency approximations with matching pursuits,” in *Wavelets: theory, algorithms, and applications (Taormina, 1993)*, vol. 5 of *Wavelet Anal. Appl.*, pp. 271–293. Academic Press, San Diego, CA, 1994.
- [3] Lee K. Jones, “On a conjecture of Huber concerning the convergence of projection pursuit regression,” *Ann. Statist.*, vol. 15, no. 2, pp. 880–882, 1987.
- [4] David L. Donoho and Xiaoming Huo, “Uncertainty principles and ideal atomic decomposition,” *IEEE Trans. Inform. Theory*, vol. 47, no. 7, pp. 2845–2862, 2001.
- [5] Scott Shaobing Chen, David L. Donoho, and Michael A. Saunders, “Atomic decomposition by basis pursuit,” *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61 (electronic), 1998.
- [6] David L. Donoho and Michael Elad, “Optimally sparse representation in general (nonorthogonal) dictionaries via l^1 minimization,” *Proc. Natl. Acad. Sci. USA*, vol. 100, no. 5, pp. 2197–2202 (electronic), 2003.
- [7] Michael Elad and Alfred M. Bruckstein, “A generalized uncertainty principle and sparse representation in pairs of bases,” *IEEE Trans. Inform. Theory*, vol. 48, no. 9, pp. 2558–2567, 2002.
- [8] J. A. Tropp, “Greed is good: algorithmic results for sparse approximation,” *IEEE Trans. Inform. Theory*, vol. 50, no. 10, pp. 2231–2242, 2004.
- [9] Rémi Gribonval and Morten Nielsen, “Sparse representations in unions of bases,” *IEEE Trans. Inform. Theory*, vol. 49, no. 12, pp. 3320–3325, 2003.
- [10] R. Gribonval and M. Nielsen, “Highly sparse representations from dictionaries are unique and independent of the sparseness measure,” Tech. Rep., Aalborg University, Aalborg, 2003.
- [11] R. Gribonval and P. Vandergheynst, “On the exponential convergence of Matching Pursuits in quasi-incoherent dictionaries,” Tech. Rep. PI-1619, IRISA, Rennes, France, Apr. 2004, submitted to *IEEE Trans. Inf. Th.*
- [12] L. Daudet and B. Torr sani, “Hybrid representations for audiophonic signal encoding,” *Signal Processing, special issue on Image and Video Coding Beyond Standards*, vol. 82, no. 11, pp. 1595–1617, 2002.
- [13] Jean-Fran ois Cardoso, “Blind signal separation: statistical principles,” *Proceedings of the IEEE. Special issue on blind identification and estimation*, vol. 9, no. 10, pp. 2009–2025, Oct. 1998.
- [14] M. Zibulevsky and B.A. Pearlmutter, “Blind source separation by sparse decomposition in a signal dictionary,” *Neural Computations*, vol. 13, no. 4, pp. 863–882, 2001.

- [15] J. A. Tropp, A. C. Gilbert, and M. J. Strauss., “Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit,” Tech. Rep., University of Michigan, Ann Arbor, 2004.
- [16] A. Jourjine, S. Rickard, and O. Yilmaz, “Blind separation of disjoint orthogonal signals: Demixing n sources from 2 mixtures,” in *Proc. Int. Conf. Acoust. Speech Signal Process. (ICASSP’00)*, Istanbul, Turkey, June 2000, vol. 5, pp. 2985–2988.
- [17] R. Gribonval, “Piecewise linear source separation,” in *Proc. SPIE ’03*, M.A. Unser, A. Aldroubi, and A.F. Laine, Eds., San Diego, CA, Aug. 2003, vol. 5207 Wavelets: Applications in Signal and Image Processing X, pp. 297–310.
- [18] R. Gribonval, “Sparse decomposition of stereo signals with matching pursuit and application to blind separation of more than two sources from a stereo mixture,” in *Proc. Int. Conf. Acoust. Speech Signal Process. (ICASSP’02)*, Orlando, Florida, USA, 2002.
- [19] D. Leviatan and V. N. Temlyakov, “Simultaneous approximation by greedy algorithms,” *Advances in Comp. Math.*, (to appear).
- [20] K. Engan S. F. Cotter, B. D. Rao and K. Kreutz-Delgado, “Sparse solutions to linear inverse problems with multiple measurement vectors,” *IEEE Transactions on Signal Processing*, 2005, to appear.
- [21] V.N. Temlyakov, “Weak greedy algorithms,” *Advances in Computational Mathematics*, vol. 12, no. 2,3, pp. 213–227, 2000.
- [22] Arie Feuer and Arkadi Nemirovski, “On sparse representation in pairs of bases,” *IEEE Trans. Inform. Theory*, vol. 49, no. 6, pp. 1579–1581, 2003.