HAL
open science

A review of cosparse signal recovery methods applied to sound source localization

Srđan Kitić, Nancy Bertin, Rémi Gribonval

To cite this version:

Srđan Kitić, Nancy Bertin, Rémi Gribonval. A review of cosparse signal recovery methods applied to sound source localization. Le XXIVe colloque Gretsi, Sep 2013, Brest, France. hal-00838080v3

HAL Id: hal-00838080

https://inria.hal.science/hal-00838080v3

Submitted on 12 Sep 2013

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.

A review of cosparse signal recovery methods applied to sound source localization *

Srđan Kitić[1], Nancy Bertin[2], Rémi Gribonval[1]

[1]Inria, Centre Inria Rennes - Bretagne Atlantique
Campus de Beaulieu, F-35042 Rennes Cedex, Rennes, France

[2]IRISA - CNRS UMR 6074
Campus de Beaulieu, F-35042 Rennes Cedex, Rennes, France

srdjan.kitic@inria.fr, nancy.bertin@irisa.fr,
remi.gribonval@inria.fr

Résumé – L'objectif de cet article est de comparer plusieurs méthodes de l'état de l'art pour la reconstruction coparcimonieuse de signaux, dans le contexte de la localisation de sources sonores. Nous évaluons les performances de cinq algorithmes de reconstruction coparcimonieuse : l'algorithme de "Greedy Analysis Structured Pursuit", les minimisations ℓ_1 et $\ell_{1,2}$ jointe, ainsi que les algorithmes "Structured Analysis Iterative Hard Thresholding" et "Structured Analysis Hard Thresholding Pursuit". Nous comparons également ces algorithmes à l'approche de parcimonie à la synthése, que nous résolvons par la minimisation jointe $\ell_{1,2}$ correspondante. Nous illustrons nos résultats dans le cadre d'une application à la localisation de sources sonores, réalise sur des simulations de mesures de champs de pression acoustique.

Abstract – This work aims at comparing several state-of-the-art methods for cosparse signal recovery, in the context of sound source localization. We assess the performance of five cosparse recovery algorithms: Greedy Analysis Structured Pursuit, ℓ_1 and joint $\ell_{1,2}$ minimization, Structured Analysis Iterative Hard Thresholding and Structured Analysis Hard Thresholding Pursuit. In addition, we evaluate the performance of these methods against the sparse synthesis paradigm, solved with corresponding joint $\ell_{1,2}$ minimization method. For this evaluation, the chosen applicative showcase is sound source localization from simulated measurements of the acoustic pressure field.

1 Introduction

Most real-life signals exhibit a so-called *compressible structure*, which means that, if represented in a suitable domain, most of the information they contain is inherently low-dimensional. This premise has been extensively used in inverse problems such as compressed sensing [1], where we seek to restore the original signal $\boldsymbol{x} \in \mathbb{R}^n$ from a set of undersampled and/or noisy linear measurements $\boldsymbol{y} = \boldsymbol{M}\boldsymbol{x} + \boldsymbol{\varepsilon} \in \mathbb{R}^m$ with $m < n$.

This underdetermined problem can be regularized by a sparse prior, where we assume that the signal $\boldsymbol{x}$ may be represented as a linear combination of a few column vectors extracted from a *dictionary* $\boldsymbol{\Psi} \in \mathbb{R}^{n\times d}$, $n \leq d$, weighted by some vector $\boldsymbol{\alpha} \in \mathbb{R}^d$ which has a small number k of non-zero coefficients compared to its dimension d. A significant amount of theoretical and empirical research has confirmed the usefulness of this, so-called *sparse synthesis* approach.

Beside sparse synthesis, an alternative approach termed *sparse analysis* or *cosparse* signal model [2] has emerged more recently. In this model, the product $\boldsymbol{z} = \boldsymbol{\Omega}\boldsymbol{x}$, where $\boldsymbol{\Omega} \in \mathbb{R}^{p\times n}$ ($p \geq n$), is assumed to be sparse. The matrix $\boldsymbol{\Omega}$ is called the *analysis operator*, and may be, for example, the finite difference operator. Although equivalent when $\boldsymbol{\Omega}$ is an invertible square matrix ($\boldsymbol{\Omega} = \boldsymbol{\Psi}^{-1}$), the two models differ as soon as $p \neq d$ [3].

In order to emphasize the distinction, we say that the signal $\boldsymbol{x}$ is *l-cosparse* if $l = p - \|\boldsymbol{\Omega}\boldsymbol{x}\|_0 < d$. The cosparse model has received less attention than the synthesis model so far. A foundation for cosparse signal modeling and counterparts of most common sparse synthesis methods can be found in [2, 4].

When the underlying sparse signal has a specific structure (*e.g.*, when non-zeros of $\boldsymbol{\alpha}$ follow some particular pattern), it has been shown [5] that exploiting this *structured sparsity* contributes to better signal recovery. *Joint sparsity* is one of these exploitable structure. In this case, the complete signal $\boldsymbol{\alpha}$ can be described as a concatenation of sparse subsignals ($\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2 \ldots \boldsymbol{\alpha}_s$), where each $\boldsymbol{\alpha}_i$ is of the dimension d/s, and which all share the same *support* (*i.e.*, the indices of nonzero elements are the same for all $\boldsymbol{\alpha}_i$).

*This work was supported in part by the European Research Council, PLEASE project (ERC-StG-2011-277906).

Similarly, we assume the signal $\boldsymbol{x}$ to be *jointly cosparse*, if the indices of zero-elements (called *cosupport*) in all $\boldsymbol{z}_i$ (subvectors of $\boldsymbol{z} = \boldsymbol{\Omega x}$) are equal.

Our aim is to exploit the cosparse framework for solving a source localization problem, described in the following section.

2 Sound source localization in a 2D acoustic pressure field

We wish to localize one or more sound sources, in a bounded two-dimensional space, from some linear measurements of the acoustic pressure field $p(\vec{r}, t)$. At all non-boundary locations, the acoustic pressure will approximatively obey the acoustic wave equation:

$$\Delta p(\vec{r}, t) - \frac{1}{c^2}\frac{\partial^2 p(\vec{r}, t)}{\partial t^2} = \begin{cases} 0, \text{ if no source at } \vec{r} \\ f(\vec{r}, t), \text{ if source at } \vec{r} \end{cases} \quad (1)$$

Using the Finite Difference Method and second-order approximations, the equation (1) can be discretized straightforwardly. By taking into account the boundary and the initial conditions, we can then design an appropriate operator $\boldsymbol{\Omega} \in \mathbb{R}^{d \times d}$ such that $\boldsymbol{\Omega p} = 0$ for all source-free positions. Note that the operator $\boldsymbol{\Omega}$ is very sparse, which is a convenient property in terms of computational cost.

Since the number of sources (s) is usually much smaller than the dimensionality of the discretized space, the signal $\boldsymbol{p}$ is cosparse, and even joint-cosparse if we assume that sources are not moving. Hence, localizing the sources amounts to the noiseless cosparse recovery problem:

$$\begin{aligned} &\text{recover } \boldsymbol{p} \text{ from } \boldsymbol{y} = \boldsymbol{M p}, \text{ s.t. } \|\boldsymbol{\Omega p}\|_0 \ll d, \\ &\qquad\qquad \text{and } \boldsymbol{p} \text{ is jointly cosparse.} \end{aligned} \quad (2)$$

We can also note that in this square case, a simple matrix inversion results in an equivalent synthesis model where the dictionary $\boldsymbol{\Psi} = \boldsymbol{\Omega}^{-1}$ are the *Green's functions*. We will discuss the algorithmic disadvantages of usage of Green's functions later in the text.

3 Algorithms and methods

In general, finding an exact solution for the cosparse regularization is known to be NP-hard [2].To circumvent this, different *approximative* methods have been proposed, out of which we consider the following in our study[1]:

ℓ_1 (analysis) minimization Standard ℓ_1 norm minimization has been widely used for its property to produce sparse solutions and its convenient convexity [6]. In the context of the analysis model, it is used in the following manner: $\hat{\boldsymbol{p}} = \arg\min_{\boldsymbol{p}} \|\boldsymbol{\Omega p}\|_1$, s.t. $\boldsymbol{y} = \boldsymbol{M p}$.

[1]The optimization algorithms are provided as a public domain software (under *GNU* licence) at http://people.rennes.inria.fr/Srdan.Kitic/

Joint $\ell_{1,2}$ (analysis) minimization The minimization of the $\ell_{1,2}$ norm [7] is equivalent to a minimization of ℓ_1 norm on *groups* $\boldsymbol{g}_{i,j}$ of $\boldsymbol{\Omega p}$, which gather the variables corresponding to each spatial index (i, j). The principle is to promote a small number of non-zero normed groups $\boldsymbol{g}_{i,j}$ *in space*, while not imposing a restriction within a given group, *in time*. Spatial grouping of variables leads to the following formulation of the method:

$$\begin{aligned} \hat{\boldsymbol{p}} = \arg\min_{\boldsymbol{p}} \|\boldsymbol{Z}(\boldsymbol{p})\|_{1,2} \text{ s.t. } \boldsymbol{y} = \boldsymbol{M p} \\ \text{where } \boldsymbol{Z}(\boldsymbol{p}) = \text{devect}^2(\boldsymbol{\Omega p}). \end{aligned} \quad (3)$$

GRASP As in the traditional sparse framework, cosparse recovery can also be addressed by greedy approaches. We retained here the use of GRASP - GReedy Analysis Structured Pursuit [8], an algorithm which is partially based on Greedy Analysis Pursuit (GAP) [2]. It can be intuitively seen as an approximation of the mixed norm minimization described in previous paragraph (for the detailed description, the reader may consult the referenced papers).

Structured AIHT and AHTP Structured Analysis Iterative Hard Thresholding and Structured Analysis Hard Thresholding Pursuit are modified versions of Analysis Iterative Hard Thresholding and Analysis Hard Thresholding Pursuit algorithms presented in [2, 4]. We introduced a change in the joint cosupport selection step, very similar to the one used in GRASP, in order to make them structure-aware and allow a fairer comparison with the previous. Simultaneous HTP has been already developed for the synthesis model [9], but, to our best knowledge, this remained to be done for analysis IHT and HTP.

Joint $\ell_{1,2}$ (synthesis) minimization For comparison purpose, we include in our benchmark an algorithm based on the synthesis paradigm for joint-sparse recovery, which minimizes the mixed $\ell_{1,2}$ norm. As noted in section 1, the synthesis and the analysis models should here produce the equivalent solutions ; however, the dictionary $\boldsymbol{\Psi}$ is not sparse, which vastly increases the computational cost in large-scale setting.

4 Experiments

The solution of a problem (2) depends on different parameters. In order to be comprehensive, we have tested the recovery methods for various settings.

4.1 Simulation settings

The experiments were conducted in an artificial two dimensional space, modeled as an uniform rectangular grid

[2]*devect* reshapes the input vector $\boldsymbol{u}$ into a matrix $\boldsymbol{U}$ such that rows of $\boldsymbol{U}$ correspond to a spatial dimension, while columns represent different time instances.

of size 20×20. The boundaries were modeled by Neumann or Dirichlet boundary conditions ($\frac{\mathrm{d}\vec{n}}{\mathrm{d}\vec{r}} = 0$ or $p = \text{const}$, respectively). The time span of the experiments (the "acquisition" time) was either $T_{\mathrm{a}1} = 40$ or $T_{\mathrm{a}2} = 100$, whereas the wave propagation speed was set to a constant value $c = 0.2$ (homogeneous medium).

For the physical settings described above, the appropriate finite difference operators $\mathbf{\Omega}$ were designed. The corresponding Green's function dictionaries $\mathbf{\Psi}$ were generated by taking the matrix inverse of the analysis operators. As expected, the numerical density (the number of non-zeros) of these matrices is very different: the dictionaries contain approximately 20% of non-zero elements (depending on the boundary conditions) whereas for the operators this percentage is around 0.004%.

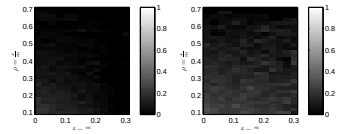
The first series of experiments compared the performance of recovery methods for the unit magnitude monochromatic sources. Each source was assigned the random phase $\varphi \in [0, 2\pi]$. The second series of experiments investigated the localization of wideband sources. These were modeled as white Gaussian zero-mean processes of variance $\sigma_2^2 = 0.04$.

We present the results in the form of empirical probability of accurate source localization, i.e. the number of correctly identified sources given the total number of sources in the experiment. The numbers of microphones m and sources s were varied relative to each other, and both were placed randomly in the spatial field. The x-axis indicates the ratio between the number of microphones and the total *spatial* dimension n, while the y-axis indicates the ratio between the number of sources and the number of microphones. The source locations were recovered after estimating the pressure signal $\boldsymbol{p}$, by localizing s largest (in ℓ_2-norm sense) spatial positions of $\mathbf{\Omega}\boldsymbol{p}$.

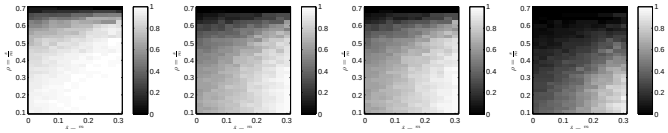
4.2 Results

The first six diagrams in Figure 1 depict the recovery of a lower frequency signal for the short acquisition time $T_{\mathrm{a}1}$. The presented diagrams are for the Neumann boundary condition, but we obtained almost identical results for the Dirichlet boundary as well. This trend continued throughout all the experiments, meaning that the performance was approximatively the same regardless the boundary condition applied. The results (a) and (b) clearly indicate the inferiority of SAIHT and SAHTP methods which, in addition, exhibited problems with the convergence. For concision purpose, these algorithms were omitted in the other experiments.

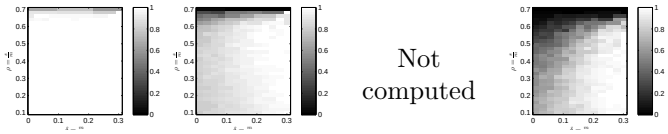
The diagrams (g), (h) and (i) show the significant improvement gained by extending the acquisition time to $T_{\mathrm{a}2}$. This is *only* feasible in the cosparse case, since the dictionary of the appropriate size is extremely memory consuming (over 12Go in this setting). Hence, the results for the synthesis algorithm are not presented.



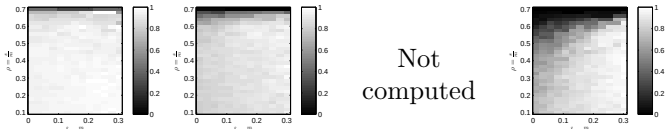
(a) SAIHT $T_{\mathrm{a}1} = 40$ (b) SAHTP $T_{\mathrm{a}1} = 40$



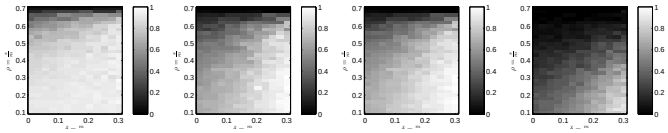
(c) ℓ_1 analysis $T_{\mathrm{a}1} = 40$ (d) $\ell_{1,2}$ analysis $T_{\mathrm{a}1} = 40$ (e) $\ell_{1,2}$ synthesis $T_{\mathrm{a}1} = 40$ (f) GRASP $T_{\mathrm{a}1} = 40$



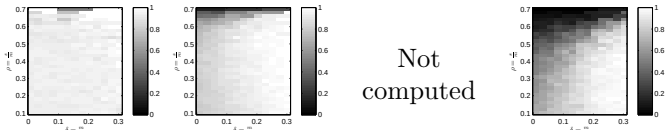
(g) ℓ_1 analysis $T_{\mathrm{a}2} = 100$ (h) $\ell_{1,2}$ analysis $T_{\mathrm{a}2} = 100$ Not computed (i) GRASP $T_{\mathrm{a}2} = 100$



(j) ℓ_1 analysis $T_{\mathrm{a}2} = 100$ (k) $\ell_{1,2}$ analysis $T_{\mathrm{a}2} = 100$ Not computed (l) GRASP $T_{\mathrm{a}2} = 100$

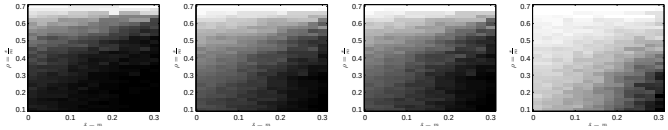


(m) ℓ_1 analysis $T_{\mathrm{a}1} = 40$ (n) $\ell_{1,2}$ analysis $T_{\mathrm{a}1} = 40$ (o) $\ell_{1,2}$ synthesis $T_{\mathrm{a}1} = 40$ (p) GRASP $T_{\mathrm{a}1} = 40$



(q) ℓ_1 analysis $T_{\mathrm{a}2} = 100$ (r) $\ell_{1,2}$ analysis $T_{\mathrm{a}2} = 100$ Not computed (s) GRASP $T_{\mathrm{a}2} = 100$

Figure 1: Probability of correct source localization for the low frequency signal ((a) - (i)), the high frequency signal ((j) - (l)) and the wideband signal ((m) - (s))



(a) ℓ_1 analysis $T_{\mathrm{a}1} = 40$ (b) $\ell_{1,2}$ analysis $T_{\mathrm{a}1} = 40$ (c) $\ell_{1,2}$ synthesis $T_{\mathrm{a}1} = 40$ (d) GRASP $T_{\mathrm{a}1} = 40$

Figure 2: Average norm of the signal estimation error

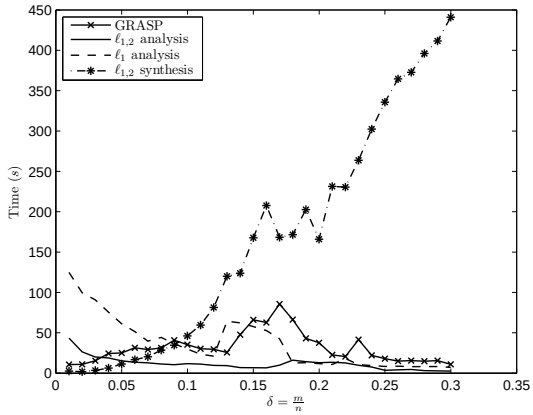


Figure 3: Average computation time

The following series ((j) - (l)) depicts the difference in localization performance when the signal frequency is higher (f_2). It has become evident that increasing the frequency leads to a decrease in performance. The final series of experiments, (m) - (s), represents the algorithm's performance in the case of a wideband signal.

Important fact is that the source localization performance is highly correlated with the success of a signal recovery (the underlying signal's estimation error). This is illustrated in Figure 2 for the source localization experiments 1(c) - 1(f).

Figure 3 illustrates the average computation time versus the relative number of samples for the four recovery methods. All cosparse methods are dramatically faster than the synthesis version of joint $\ell_{1,2}$ minimization. We observed that the more data becomes available, the slower the synthesis version gets, due to more linear operations needed per iteration. On the other hand, the analysis version gets faster, since the overall number of iterations becomes smaller.

5 Conclusion

We benchmarked several cosparse and one sparse recovery algorithms, for the particular case of synthetic wave signals.

Our experiments show that in the majority of cases, standard ℓ_1 norm minimization is the most robust recovery method, whereas joint $\ell_{1,2}$ minimization exhibits the fastest convergence while being the second best in source localization. GRASP is a moderately fast algorithm, but its accuracy is considerably lower than the former two. The structure-aware AIHT and AHTP algorithms are the worst performing both in accuracy and processing time, which is related to the inexact cosparse projection step [10], based on the hard thresholding.

The synthesis and the analysis versions of joint $\ell_{1,2}$ minimization perform the same in terms of accuracy, but the cosparse version converges much faster due to the sparsity of analysis operator. Moreover, the cosparse methods allow for longer acquisition time, which improves the recovery performance.

The signal frequency, the propagation speed and the discretization are closely related and impact the performance. The phenomenon is both physical and numerical, as a consequence of the discretization method's limitations. Their interaction is out of the scope of this paper and remains to be investigated.

References

[1] D. Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289–1306, 2006.

[2] S. Nam, M. E. Davies, M. Elad, and R. Gribonval, "The Cosparse Analysis Model and Algorithms," *Applied and Computational Harmonic Analysis*, vol. 34, no. 1, pp. 30–56, 2013. Preprint available on arXiv since 24 Jun 2011.

[3] M. Elad, P. Milanfar, and R. Rubinstein, "Analysis versus synthesis in signal priors," tech. rep., 2005.

[4] R. Giryes, S. Nam, M. Elad, R. Gribonval, and M. E. Davies, "Greedy-Like Algorithms for the Cosparse Analysis Model," *arXiv preprint arXiv:1207.2456*, 2013.

[5] M. Yuan, M. Yuan, Y. Lin, and Y. Lin, "Model selection and estimation in regression with grouped variables," *Journal of the Royal Statistical Society, Series B*, vol. 68, pp. 49–67, 2006.

[6] S. S. Chen, D. L. Donoho, Michael, and A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, pp. 33–61, 1998.

[7] R. Jenatton, J. Audibert, and F. Bach, "Structured variable selection with sparsity-inducing norms," *arXiv preprint arXiv:0904.3523*, 2009.

[8] S. Nam and R. Gribonval, "Physics-driven structured cosparse modeling for source localization," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pp. 5397–5400, IEEE, 2012.

[9] S. Foucart, "Recovering jointly sparse vectors via hard thresholding pursuit," *Proc. Sampling Theory and Applications (SampTA)],(May 2-6 2011)*, 2011.

[10] A. M. Tillmann, R. Gribonval, and M. E. Pfetsch, "Projection onto the k-cosparse set is np-hard," *CoRR*, vol. abs/1303.5305, 2013.