



HAL
open science

Compressed sensing and the synthesis formulation

Claire Boyer, Jonas Kahn, Maximilian März, Pierre Weiss

► **To cite this version:**

Claire Boyer, Jonas Kahn, Maximilian März, Pierre Weiss. Compressed sensing and the synthesis formulation. SPARS, Jul 2019, TOULOUSE, France. hal-02374485

HAL Id: hal-02374485

<https://hal.science/hal-02374485>

Submitted on 21 Nov 2019

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.

Compressed sensing and the synthesis formulation

Claire Boyer

Sorbonne Université & ENS Paris
claire.boyer@sorbonne-universite.fr

Jonas Kahn

CNRS, IMT & ITAV Toulouse
jonas.kahn@math.univ-toulouse.fr

Maximilian März

TU Berlin
maxmaerz@gmail.com

Pierre Weiss

CNRS, IMT & ITAV Toulouse
pierre.armand.weiss@gmail.com

Abstract—There is clear empirical evidence that the use of redundant dictionaries instead of orthogonal bases can significantly improve the recovery results in compressed sensing. The theoretical understanding of this phenomenon is however still partial. The objective of this work is to better understand the so-called synthesis formulation in the framework of random Gaussian measurements. We derive new sharp guarantees for robust recovery of coefficient representations and of signals using the machinery of conic Gaussian widths.

I. INTRODUCTION

Consider a signal $u \in \mathbb{R}^n$ and let $A \in \mathbb{R}^{m \times n}$ be a random Gaussian sensing matrix. Given some linear measurements $y = Au + e$, where $e \in \mathbb{R}^m$ models measurement noise with $\|e\|_2 \leq \eta$, we wish to recover an approximation of u by solving the ℓ^1 problem

$$\hat{X} \stackrel{\text{def.}}{=} \operatorname{argmin}_{x \in \mathbb{R}^d} \|x\|_1 \quad \text{s.t.} \quad \|A\Phi x - y\|_2 \leq \eta. \quad (1)$$

In this equation, $\Phi = [\phi_1, \dots, \phi_d] \in \mathbb{R}^{n \times d}$ is a given surjective matrix (dictionary) with $d \geq n$. Given a solution $\hat{x} \in \hat{X}$, we can form an estimate \hat{u} of u as $\hat{u} = \Phi \hat{x}$.

When dealing with redundant dictionaries, two formulations are possible: the synthesis and the analysis formulations [1]. Problem (1) is a *synthesis problem* since a signal u is ‘synthesized’ as a linear combination of atoms (columns) of the dictionary. Most of the recent theoretical activity revolved around the analysis formulation [2]–[5], while both approaches yield similar results with well chosen dictionaries. Surprisingly, we are only aware of a single paper devoted to the synthesis formulation [6]. The authors study the uniform recovery of all s -sparse vectors using the notion of coherence of the dictionary $\mu = \max_{i,j} |\langle \phi_i, \phi_j \rangle|$. Put roughly, their main result states that for all s -sparse x_0 with $s = O(\frac{1}{\mu})$, the condition $m \gtrsim s \log(d/s)$ is enough to ensure the stable recovery of x_0 and hence u . This result suffers from several problems:

- The condition $s = O(\mu^{-1})$ is restrictive: in particular, any dictionary composed of correlated atoms (e.g. translation invariant wavelet transform) is ruled out. In addition, it implicitly suffers from the square-root bottleneck [7] which imposes $s = O(\sqrt{n})$.
- Recovering a specific s -sparse vector instead of all s -sparse vectors may lead to a better understanding of the problem’s geometry.
- Recovering the coefficients x_0 is often less relevant than the signal u itself. This may lead to very different conditions.

In what follows we address the above limitations by providing tight bounds on the number of measurements needed for the recovery of a specific signal u and of its coefficients in the case of Gaussian measurements.

II. MAIN RESULTS

The set of minimal ℓ^1 representations of u is defined by

$$X_{\ell^1} \stackrel{\text{def.}}{=} \operatorname{argmin}_{x \in \mathbb{R}^d} \|x\|_1 \quad \text{s.t.} \quad \Phi x = u. \quad (2)$$

Following [8], we let $\mathcal{D}(x)$ denote the descent cone of $\|\cdot\|_1$ at x and $\sigma_{\min}(\Phi, \mathcal{D}(x))$ denote the minimal conic singular value of Φ , restricted on $\mathcal{D}(x)$. The Gaussian width of a cone C is defined as

$$w_c(C) \stackrel{\text{def.}}{=} \mathbb{E} \left[\sup_{x \in C \cap \mathbb{S}^{d-1}} \langle g, x \rangle \right], \quad (3)$$

where g is a standard normal vector of \mathbb{R}^d and \mathbb{S}^{d-1} denotes the unit-sphere in \mathbb{R}^d . Our main results are given below. The proofs follow the methodologies in [8], [9].

Theorem II.1 (Coefficient recovery). *Assume that $X_{\ell^1} = \{x_{\ell^1}\}$ is a singleton. Let $\kappa = \sigma_{\min}(\Phi, \mathcal{D}(x_{\ell^1}))$. Every $\hat{x} \in \hat{X}$ satisfies $\|\hat{x} - x_{\ell^1}\|_2 = O(\eta)$ with probability larger than $1 - \varepsilon$, if*

$$m \gtrsim \frac{\|\Phi\|_{2 \rightarrow 2}^2}{\kappa^2} s \log\left(\frac{1}{\varepsilon}\right). \quad (4)$$

Now, we let $x_{\ell^1}^+$ denote the element in X_{ℓ^1} of maximal support and $s^+ = |\operatorname{supp}(x_{\ell^1}^+)|$ denote the cardinality of its support.

Theorem II.2 (Signal recovery). *Every $\hat{u} \in \Phi \hat{X}$ satisfies $\|\hat{u} - u\|_2 = O(\eta)$ with probability larger than $1 - \varepsilon$, if*

$$m \gtrsim w_c(\Phi \mathcal{D}(x_{\ell^1}^+))^2 \log\left(\frac{1}{\varepsilon}\right) \quad (5)$$

There is no hope to recover the coefficients of u if X_{ℓ^1} is not a singleton. In that case, Theorem II.1 states that the key quantity to control the conic singular value $\sigma_{\min}(\Phi, \mathcal{D}(x_{\ell^1}))$. The result is sharp in the sense that X_{ℓ^1} is a singleton, if and only if $\sigma_{\min}(\Phi, \mathcal{D}(x_{\ell^1})) > 0$. In particular, if Φ is orthogonal, we retrieve the usual bound on the number of measurements. Theorem (II.2) involves the Gaussian width of the cone $C = \Phi \mathcal{D}(x_{\ell^1}^+)$. We propose various upper-bounds on this quantity by using the following methodology:

- 1) Show that C can be described as a conic hull of the form

$$C = \operatorname{cone}(\{\pm s^+ \phi_i - \Phi \cdot \operatorname{sign}(x_{\ell^1}^+), i = 1, \dots, d\})$$

- 2) Decompose this polyhedral cone into a subspace L_C of dimension s^+ (e.g. its lineality space) and its orthogonal complement, i.e.

$$C = L_C \oplus (C \cap L_C)^\perp.$$

- 3) Embed $(C \cap L_C)^\perp$ into a circular cone of angle α .

We can then show the bound

$$w_c^2(C) \lesssim \dim(L_C) + \frac{\log(d)}{(\cos \alpha)^2}. \quad (6)$$

If the circular cone in 3) is chosen adequately, the term $(\cos \alpha)^{-2}$ can be shown to be of the same order as $\dim(L_C) = s^+$. This approach allows to take highly correlated atoms (e.g. duplicate atoms) and still requires a low number of measurements. In case the paper gets accepted, we will discuss these phenomena and compare them to the analysis formulation [5].

REFERENCES

- [1] M. Elad, P. Milanfar, and R. Rubinstein, "Analysis versus synthesis in signal priors," *Inverse Problems*, vol. 23, p. 947, 04 2007.
- [2] E. J. Candès, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Applied and Computational Harmonic Analysis*, vol. 31, no. 1, pp. 59–73, 2011.
- [3] S. Nam, M. E. Davies, M. Elad, and R. Gribonval, "The cosparsity analysis model and algorithms," *Applied and Computational Harmonic Analysis*, vol. 34, no. 1, pp. 30–56, 2013.
- [4] M. Kabanava and H. Rauhut, "Analysis ℓ^1 -recovery with frames and gaussian measurements," *Acta Applicandae Mathematicae*, vol. 140, no. 1, pp. 173–195, 2015.
- [5] M. Genzel, G. Kutyniok, and M. März, " ℓ^1 -analysis minimization and generalized (co-) sparsity: When does recovery succeed?" *arXiv preprint arXiv:1710.04952*, 2017.
- [6] H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed sensing and redundant dictionaries," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 2210–2219, May 2008.
- [7] P. Kuppinger, G. Durisi, and H. Boelcskei, "Where is randomness needed to break the square-root bottleneck?" 07 2010, pp. 1578 – 1582.
- [8] D. Amelunxen, M. Lotz, M. B. McCoy, and J. A. Tropp, "Living on the edge: Phase transitions in convex programs with random data," *Information and Inference: A Journal of the IMA*, vol. 3, no. 3, pp. 224–294, 2014.
- [9] C. Liaw, A. Mehrabian, Y. Plan, and R. Vershynin, "A simple tool for bounding the deviation of random matrices on geometric sets," in *Geometric aspects of functional analysis*. Springer, 2017, pp. 277–299.