PAPER • OPEN ACCESS

Optimality of 1-norm regularization among weighted 1-norms for sparse recovery: a case study on how to find optimal regularizations

View the article online for updates and enhancements.

Optimality of 1-norm regularization among weighted 1-norms for sparse recovery: a case study on how to find optimal regularizations

Yann Traonmilin[a] and Samuel Vaiter[b]

[a]CNRS, Institut de Mathématiques de Bordeaux, 33400 Talence, France; [b]CNRS, Institut de Mathématiques de Bourgogne, 9 avenue Alain Savary, 21000, Dijon

E-mail: [a]yann.traonmilin@math.u-bordeaux.fr, [b]samuel.vaiter@u-bourgogne.fr

Abstract. The 1-norm was proven to be a good convex regularizer for the recovery of sparse vectors from under-determined linear measurements. It has been shown that with an appropriate measurement operator, a number of measurements of the order of the sparsity of the signal (up to log factors) is sufficient for stable and robust recovery. More recently, it has been shown that such recovery results can be generalized to more general low-dimensional model sets and (convex) regularizers. These results lead to the following question: to recover a given low-dimensional model set from linear measurements, what is the "best" convex regularizer? To approach this problem, we propose a general framework to define several notions of "best regularizer" with respect to a low-dimensional model. We show in the minimal case of sparse recovery in dimension 3 that the 1-norm is optimal for these notions. However, generalization of such results to the n-dimensional case seems out of reach. To tackle this problem, we propose looser notions of best regularizer and show that the 1-norm is optimal among weighted 1-norms for sparse recovery within this framework.

1. Introduction

We consider the observation model in a Hilbert space $\mathcal{H}$ (with associated norm $\|\cdot\|_{\mathcal{H}}$):

$$y = Mx_0 \tag{1}$$

where M is an under-determined linear operator, y is a m-dimensional vector and x_0 is the unknown. We suppose that x_0 belongs to a low-dimensional model Σ (a union of subspaces). We consider the following minimization program.

$$x^* \in \arg \min_{Mx=y} R(x) \tag{2}$$

where R is a regularization function. A huge body of work gives practical regularizers ensuring that $x^* = x_0$ for several low-dimensional models (in particular sparse and low rank models, see [5] for a most complete review of these results). The operator M is generally required to satisfy some property (e.g., the restricted isometry property) to guarantee recovery. In this work, we aim at finding the "best" regularizer for exact recovery of $x_0 \in \Sigma$. Ideally we would like to set $R = \iota_\Sigma$ (the characteristic function of Σ) but it is not practical in many cases (sparse and low rank recovery) as it is generally not convex, and even NP-hard to compute as a combinatorial problem. Consequently, we restrict the search for the best regularizer to a class of interesting regularizers $\mathcal{C}$. In our examples, the set $\mathcal{C}$ is a subset of the set of convex functions. Other interesting classes might be considered, such as partly smooth functions [9].

1.1. Best regularizer with respect to a low dimensional model

Defining what is the "best" regularizer in $\mathcal{C}$ for recovery is not immediate. Ideally, to fit to the inverse problem, we must define a compliance measure that depends on both the kind of unknown and measurement operator we consider. If we have some knowledge that $M \in \mathcal{M}$ where $\mathcal{M}$ is a set of linear operators, we want to define a compliance measure $A_{\Sigma,\mathcal{M}}(R)$ that tells us if a regularizer is good in these situations, and maximize it. Such maximization might yield a function R^* that depends on M (e.g., in [7], when looking for tight continuous relaxation of the ℓ^0 penalty a dependency on M appears). We aim for a more *universal* notion of optimal regularizer that does not depend on M. Hence, we look for a compliance measure $A_{\Sigma}(R)$ and its maximization

$$R^* \in \arg\max_{R \in \mathcal{C}} A_{\Sigma}(R). \tag{3}$$

In the sparse recovery example studied in this article, the existence of a maximum of $A_{\Sigma}(R)$ is verified. However, we could ask ourselves what conditions on $A_{\Sigma}(R)$ and $\mathcal{C}$ are necessary and sufficient for the existence of a maximum, which is out of the scope of this article.

1.2. Compliance measures

When studying recovery with a regularization function R, two types of guarantees are generally used: uniform and non-uniform. To describe these recovery guarantees, we use the following definition of descent vectors.

Definition 1.1 (Descent vectors). *For any $x \in \mathcal{H}$, the collection of descent vectors of R at x is*

$$\mathcal{T}_R(x) := \{z \in \mathcal{H} : R(x + z) \leq R(x)\}. \tag{4}$$

We write $\mathcal{T}_R(\Sigma) := \bigcup_{x \in \Sigma} \mathcal{T}_R(x)$. Recovery is characterized by descent vectors (recall that x^* is the result of minimization (2)):

- Uniform recovery: Let M a linear operator. Then "for all $x_0 \in \Sigma$, $x^* = x_0$" is equivalent to $\mathcal{T}_R(\Sigma) \cap \ker M = \{0\}$.
- Non-uniform recovery: Let M a linear operator and $x_0 \in \Sigma$. Then $x^* = x_0$ is equivalent to $\mathcal{T}_R(x_0) \cap \ker M = \{0\}$.

Hence, a regularization function R is "good" if $\mathcal{T}_R(\Sigma)$ leaves a lot of space for $\ker M$ to not intersect it (trivially). In dimension n, if there is no orientation prior on the kernel of M, the amount of space left can be quantified by the "volume" of $\mathcal{T}_R(\Sigma) \cap S(1)$ where $S(1)$ is the unit sphere with respect to $\| \cdot \|_{\mathcal{H}}$. Hence a compliance measure for uniform recovery can be defined as

$$A_{\Sigma}^U(R) := 1 - \frac{\mathrm{vol}\,(\mathcal{T}_R(\Sigma) \cap S(1))}{\mathrm{vol}(S(1))}. \tag{5}$$

More precisely, here, the volume $\mathrm{vol}(E)$ of a set E is the measure of E with respect to the uniform measure on the sphere $S(1)$ (i.e. the $n-1$-dimensional Haussdorf measure of $\mathcal{T}_R(\Sigma) \cap S(1)$). When maximizing this compliance measure with convex regularizers (proper, coercive and continuous), it has been shown that we can limit ourselves to atomic norms with atoms included in the model [8, Lemma 2.1]. When looking at non-uniform recovery for random Gaussian measurements, the quantity $\frac{\mathrm{vol}(\mathcal{T}_R(x_0) \cap S(1))}{\mathrm{vol}(S(1))}$ represents the probability that a randomly oriented kernel of dimension 1 intersects (non trivially) $\mathcal{T}_R(x_0)$. The highest probability of intersection with respect to x_0 quantifies the lack of compliance of R, hence we can define:

$$A_{\Sigma}^{NU}(R) := 1 - \sup_{x \in \Sigma} \frac{\mathrm{vol}\,(\mathcal{T}_R(x) \cap S(1))}{\mathrm{vol}(S(1))} \tag{6}$$

Note that this can be linked with the Gaussian width and statistical dimension theory of sparse recovery [3, 1].

Remark 1.1. *In infinite dimension, the volume of the sphere $S(1)$ vanishes, making the measures above uninteresting. However, [8] and [6] show that we can often come back to a low-dimensional recovery problem in an intermediate finite (potentially high dimensional) subspace of $\mathcal{H}$. Adapting the definition of $S(1)$ to this subspace allows to extend these compliance measures.*

Another possibility for the n-dimensional case, which we develop in this article, is to use recovery results based on the restricted isometry property (RIP). They have been shown to be adequate for multiple models [8], to be tight in some sense [4] for sparse and low rank recovery, to be necessary in some sense [2] and to be well adapted to the study of random operators [6]. In particular, it has been shown that if M has a RIP with constant $\delta < \delta_\Sigma(R)$ on the secant set $\Sigma - \Sigma$, with $\delta_\Sigma(R)$ being fully determined by Σ and R [8], then stable recovery is possible. Hence, by taking

$$A_\Sigma^{RIP}(R) := \delta_\Sigma(R), \tag{7}$$

the larger A_Σ^{RIP} is, the less stringent are RIP recovery conditions for recovery of elements of Σ with R. We develop these ideas more precisely in Section 3 and perform the maximization with such measure in the sparse recovery case.

1.3. Optimality results
Within this framework we show the following results:

(i) In Section 2, when $\mathcal{H} = \mathbb{R}^3$, $\Sigma = \Sigma_1$ the set of 1-sparse vectors and $\mathcal{C}$ is the set of (symmetric) atomic norms with atoms included in the model. Both $A_\Sigma^U(R)$ and $A_\Sigma^{NU}(R)$ are uniquely maximized by multiples of the ℓ^1-norm (Theorem 2.1). In this case, it is possible to exactly compute and maximize the compliance measure. While this study gives a good geometrical insight of the quantities at hand, extending these exact calculations to the general k-sparse recovery in dimension n seems out of reach.

(ii) In Section 3, we describe precisely how compliance measures based on RIP recovery can be defined. We then study two of these measures (based on state of the art recovery guarantees) for $\mathcal{H} = \mathbb{R}^n$, $\Sigma = \Sigma_k$ the set of k-sparse vectors (vectors with at most k non zero coordinates) and $\mathcal{C}$ the set of weighted ℓ^1-norms. We show that both of these measures are uniquely maximized by multiples of the ℓ^1-norm (Theorem 3.1 and Theorem 3.2).

2. The case of 1-sparse vectors in 3D
We investigate in detail the case of 3-dimensional vectors which are 1 sparse, i.e. the simplest interesting case (uniform recovery of 1-sparse vectors in 2 dimensions is impossible if M is not invertible). Here, we have $\Sigma = \Sigma_1$ and $\mathcal{H} = \mathbb{R}^3$. We consider weighted ℓ^1-norms, i.e., atomic norms of the form $\mathcal{C} = \{\|\cdot\|_w : \|z\|_w = w_1|z_1| + w_2|z_2| + w_3|z_3|\} = \{\|\cdot\|_\mathcal{A} : \mathcal{A} = \{\pm w_i^{-1}e_i\}_{i=1,3}, w_i > 0, \max|w_i| = 1\}$ (where $\|\cdot\|_\mathcal{A}$ is the atomic norm generated by atoms $\mathcal{A}$). To maximize A_Σ^U (respectively A_Σ^{NU}), we have to compute the surface of the intersections of 3 descent cones with the unit sphere $S(1)$ (respectively the surface of the biggest possible intersection of a descent cone with $S(1)$).

As shown in Figure 1, with symmetries, we need to compute the intersection of 3 descent cones of a weighted ℓ^1-norm at 1-sparse vectors. This is the object of Lemma 2.1. Note that in the case of asymmetrical atoms, we would then need to calculate 6×4 intersections of the sphere with a tetrahedron. To ease the notations, we introduce the following quantities which represent the cosine of the angles of a tetrahedron of interest. For $i \neq j$, set $\mu_i = w_i^{-1}$, $\beta_{ij}(\mu) = \cos(\alpha_{ij}) = \left(1 + (\mu_j/\mu_i)^2\right)^{-\frac{1}{2}}$ and $\beta_{ii} = \prod_{i \neq j} \beta_{ij}$.
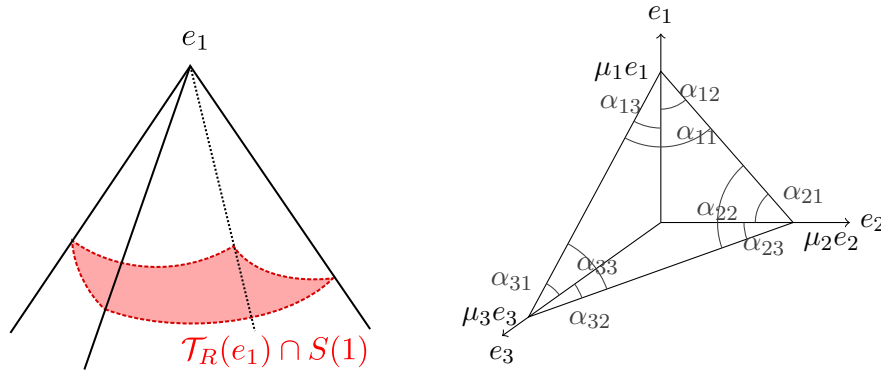
Figure 1. Left: a representation of the surface of interest. Right: parameterization of the angles of a tetrahedron.

Lemma 2.1. *Let $i \in \{1,2,3\}$, $x \in \mathbb{R}e_i$ a 1-sparse vector and $\mu_1, \mu_2, \mu_3 > 0$ three positive real numbers. Then,*

$$\mathrm{vol}\left(\mathcal{T}_{\|\cdot\|_w}(x) \cap S(1)\right) = 4\tan^{-1}\left(\frac{1}{1 + c_i(\mu)}\right) \tag{8}$$

where

$$c_i(\mu) = 1 + \sum_{j \neq i} \beta_{ij}(\mu) + \prod_{j \neq i} \beta_{ij}(\mu). \tag{9}$$

With this Lemma 2.1, we can prove that the ℓ^1-norm is the best regularizer for 1-sparse signals both for uniform and non-uniform recovery among all weighted ℓ^1-norms.

Theorem 2.1. *Let $\Sigma = \Sigma_1$ the set of 1-sparse vectors in $\mathbb{R}^3$ and $\mathcal{C} = \{\|\cdot\|_w : \|z\|_w = w_1|z_1| + w_2|z_2| + w_3|z_3|, \max|w_i| = 1, w_i > 0\}$. The ℓ^1-norm is the best regularizer among the class $\mathcal{C}$ for Σ, both for the uniform and non-uniform case,*

$$\arg\max_{R \in \mathcal{C}} A_\Sigma^U(R) = \arg\max_{R \in \mathcal{C}} A_\Sigma^{NU}(R) = \|\cdot\|_1$$

We expect a similar result if $\mathcal{C}$ is generalized to asymmetrical atomic norms. To extend these exact calculations to the n-dimensional case, we would need to be able to compare intersections of spheres and descent cones of R without an analytic formula, which appears to be a very difficult task.

3. Compliance measures based on the RIP

In this Section, we detail how we can use RIP recovery conditions to build compliance measures that are more easily managed. We start by recalling definitions and results about RIP recovery guarantees then apply our methodology. We also give Lemma that emphasize the relevant quantity (depending on the geometry of the regularizer and the model) to optimize.

Definition 3.1 (RIP constant). *In a Hilbert space $\mathcal{H}$, let Σ a union of subspaces. Let M a linear map, the RIP constant of M is defined as*

$$\delta(M) = \inf_{x \in \Sigma - \Sigma} \left| \frac{\|Mx\|_\mathcal{H}^2}{\|x\|_\mathcal{H}^2} - 1 \right|, \tag{10}$$

where $\Sigma - \Sigma$ (differences of elements of Σ) is called the secant set.

In [8], an explicit constant $\delta_\Sigma^{suff}(R)$ is given, such that $\delta(M) < \delta_\Sigma^{suff}(R)$ guarantees exact recovery of elements of Σ by minimization (2). This constant is only sufficient (and sharp in some sense for sparse and low rank recovery). An ideal RIP based compliance measure would be to use a sharp RIP constant $\delta_\Sigma^{sharp}(R)$ (which is not explicit, it is an open question to derive closed form formulations of this constant for sparsity and other low-dimensional models) defined as:

$$\delta_\Sigma^{sharp}(R) := \inf_{M:\ker M \cap \mathcal{T}_R(\Sigma) \neq \{0\}} \delta(M). \tag{11}$$

It is the best RIP constant of measurement operators where uniform recovery fail. The lack of analytic expressions for $\delta_\Sigma^{sharp}(R)$ limits the possibilities of exact optimization with respect to R. We propose to look at two compliance measures:

- Measures based on necessary RIP conditions [4] which yields sharp recovery constants for particular set of operators, e.g.,

$$\delta_\Sigma^{nec}(R) := \inf_{z \in \mathcal{T}_R(\Sigma) \setminus \{0\}} \delta(I - \Pi_z). \tag{12}$$

  where Π_z is the orthogonal projection onto the one-dimensional subspace $\mathrm{span}(z)$ (other intermediate necessary RIP constants can be defined). Another open question is to determine whether $\delta_\Sigma^{nec}(R) = \delta_\Sigma^{sharp}(R)$ generally or for some particular models.

- Measures based on sufficient RIP constants for recovery (e.g., $\delta_\Sigma^{suff}(R)$ from [8]).

Note that we have the relation

$$\delta_\Sigma^{suff}(R) \leq \delta_\Sigma^{sharp}(R) \leq \delta_\Sigma^{nec}(R). \tag{13}$$

To summarize, in the following, instead of considering the most natural RIP based compliance measure (based on $\delta_\Sigma^{sharp}(R)$), we use the best known bounds of this measure.

3.1. Compliance measures based on necessary RIP conditions
In this case, instead of working with actual RIP constants, it is easier to use (equivalently) the restricted conditioning.

Definition 3.2 (Restricted conditioning)**.**

$$\gamma(M) := \frac{\sup_{x \in (\Sigma - \Sigma) \cap S(1)} \|Mx\|_\mathcal{H}^2}{\inf_{x \in (\Sigma - \Sigma) \cap S(1)} \|Mx\|_\mathcal{H}^2}. \tag{14}$$

The RIP constant $\delta(M)$ is increasing with respect to $\gamma(M)$. In the following, we consider the compliance measure

$$A_\Sigma^{RIP,nec}(R) := \gamma_\Sigma(R) = \inf_{z \in \mathcal{T}_R(\Sigma) \setminus \{0\}} \gamma(I - \Pi_z). \tag{15}$$

When maximizing $A_\Sigma^{RIP,nec}(R)$, we look at optimal regularizers for recovery with tight frames with a kernel of dimension 1.

From here, we specialize to $\Sigma = \Sigma_k$ and $\mathcal{H} = \mathbb{R}^n$. Hence $\Sigma - \Sigma = \Sigma_{2k}$ with $k \geq 1$ and $n \geq 3$ (for $n < 3$ uniform recovery is not possible for non-invertible M). To show optimality of the ℓ^1-norm, we use the following characterization of $A_\Sigma^{RIP,nec}(R)$.

Lemma 3.1. *Let* $\Sigma = \Sigma_k$ *and* $\mathcal{H} = \mathbb{R}^n$. *Then*

$$A_\Sigma^{RIP,nec}(R) = \frac{1}{1 - \inf_{z \in \mathcal{T}_R(\Sigma) \setminus \{0\}} \sup_{x \in (\Sigma - \Sigma) \cap S(1)} \frac{\langle x, z \rangle^2}{\|z\|_\mathcal{H}^2}}. \tag{16}$$

We consider the set $\mathcal{C} = \{R : R(x) = \sum w_i|x_i| = \|x\|_w, \max |w_i| = 1, w \in (\mathbb{R}_+^*)^n\}$. Note that

$$\arg \max_{R \in \mathcal{C}} A_\Sigma^{RIP,nec}(R) = \arg \min_{R \in \mathcal{C}} \sup_{z \in \mathcal{T}_R(\Sigma)\backslash\{0\}} \frac{\|z_{T_2^c}\|_2^2}{\|z_{T_2}\|_2^2} = \arg \min_{R \in \mathcal{C}} B_\Sigma(R) \tag{17}$$

where T_2 is a notation for the support of $2k$ biggest coordinates in z, i.e. for all $i \in T_2, j \in T_2^c$, we have $|z_i| \geq |z_j|$.

Studying the quantity $B_\Sigma(R) := \sup_{z \in \mathcal{T}_R(\Sigma)\backslash\{0\}} \frac{\|z_{T_2^c}\|_2^2}{\|z_{T_2}\|_2^2}$ for $R \in \mathcal{C}$ permits to conclude.

Theorem 3.1. *Let $\Sigma = \Sigma_k$, $\mathcal{H} = \mathbb{R}^n$ and $\mathcal{C} = \{R : R(x) = \sum w_i|x_i| = \|x\|_w, \max |w_i| = 1, w \in (\mathbb{R}_+^*)^n\}$. Suppose $n \geq 2k$, then*

$$\|\cdot\|_1 = \arg \max_{R \in \mathcal{C}} A_\Sigma^{RIP,nec}(R). \tag{18}$$

In the next section, we will see that the optimization of the sufficient RIP constant leads to very similar expressions.

3.2. Compliance measures based on sufficient RIP conditions

In [8], it was shown for Σ a union of subspaces and an arbitrary regularizer R, that an explicit RIP constant $\delta_\Sigma^{suff}(R)$ is sufficient to guarantee reconstruction. From here we set $\Sigma := \Sigma_k$, $\mathcal{C} = \{R : R(x) = \sum w_i|x_i| = \|x\|_w, \max |w_i| = 1, w \in (\mathbb{R}_+^*)^n\}$.

Proposition 3.1. *Let $\Sigma = \Sigma_k$ the set of k-sparse vectors. Consider the constant $\delta_\Sigma^{suff}(R)$ from [8][Eq. (5)], we have:*

$$\delta_\Sigma^{suff}(R) = \frac{1}{\sqrt{\sup\limits_{z \in \mathcal{T}_R(\Sigma)\backslash\{0\}} \frac{\|z_{T^c}\|_\Sigma^2}{\|z_T\|_2^2} + 1}} \tag{19}$$

where T denotes the support of the k biggest coordinates of z and $\|\cdot\|_\Sigma$ is the atomic norm generated by $\Sigma \cap S(1)$ (the convex gauge induced by the convex envelope of $\Sigma \cap S(1)$).

In the following, we use

$$A_\Sigma^{RIP,suff}(R) := \delta_\Sigma^{suff}(R). \tag{20}$$

It must be noted that characterization of Proposition 3.1 was used as a lower bound on δ_Σ^{suff} in [8] (hence this part of the proof of the control of δ_Σ^{suff} in [8] is exact).

Similarly to the necessary case, from Proposition 3.1, we have

$$\arg \max_{R \in \mathcal{C}} A_\Sigma^{RIP,suff}(R) = \arg \min_{R \in \mathcal{C}} \sup_{z \in \mathcal{T}_R(\Sigma)\backslash\{0\}} \frac{\|z_{T^c}\|_\Sigma^2}{\|z_T\|_2^2} = \arg \min_{R \in \mathcal{C}} D_\Sigma(R) \tag{21}$$

where T denotes the support of the k biggest coordinates of z and $D_\Sigma(R) := \sup_{z \in \mathcal{T}_R(\Sigma)\backslash\{0\}} \frac{\|z_{T^c}\|_\Sigma^2}{\|z_T\|_2^2}$.

Remark the similarity between the fundamental quantity to optimize for the necessary case and the sufficient case $B_\Sigma(R)$ and $D_\Sigma(R)$. Studying the quantity $D_\Sigma(R)$ for $R \in \mathcal{C}$ leads to the result.

Theorem 3.2. *Let $\Sigma = \Sigma_k$, $\mathcal{H} = \mathbb{R}^n$ and $\mathcal{C} = \{R : R(x) = \sum w_i|x_i| = \|x\|_w, \max |w_i| = 1, w \in (\mathbb{R}_+^*)^n\}$. Suppose $n \geq 2k$, then*

$$\|\cdot\|_1 = \arg \sup_{R \in \mathcal{C}} A_\Sigma^{RIP,suff}(R). \tag{22}$$

4. Discussion and future work

We have shown that, not surprisingly, the ℓ^1-norm is optimal among weighted ℓ^1-norms for sparse recovery for several notions of compliance. This result had to be expected due to symmetries of the problem. However, the important point is that we could explicitly quantify the notion of good regularizer. This is promising for the search of optimal regularizers for more complicated low-dimensional models such as "sparse and low rank" models or hierarchical sparse models. For the case of sparsity, we expect to be able to generalize the optimality of the ℓ^1-norm to the more general case of atomic norms with atoms included in the model. We also expect similar result for low-rank recovery and the nuclear norm as technical tools are very similar.

It must be noted that for RIP compliance measures, we did not use a constructive proof (we exhibited the maximum of the compliance measure). A constructive proof, i.e. an exact calculation and optimization of the quantities $B_\Sigma(R)$ and $D_\Sigma(R)$ would be more satisfying as it would not require the knowledge of the optimum, which is our ultimate objective.

We used compliance measures based on (uniform) RIP recovery guarantees to give results for the general sparse recovery case, it would be interesting to do such analysis using (non-uniform) recovery guarantees based on the statistical dimension or Gaussian width of the descent cones [3, 1]. One would need to precisely lower and upper bound these quantities, similarly to our approach with the RIP, to get satisfying results.

Finally, while these compliance measures are designed to make sense with respect to known results in the area of sparse recovery, one might design other compliance measures tailored for particular needs, in this search for optimal regularizers.

Acknowledgements

References

[1] 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*, 3(3):224–294, 2014.

[2] A. Bourrier, M. Davies, T. Peleg, P. Perez, and R. Gribonval. Fundamental performance limits for ideal decoders in high-dimensional linear inverse problems. *Information Theory, IEEE Transactions on*, 60(12):7928–7946, 2014.

[3] V. Chandrasekaran, B. Recht, P. Parrilo, and A. Willsky. The convex geometry of linear inverse problems. *Foundations of Computational Mathematics*, 12(6):805–849, 2012.

[4] M. E. Davies and R. Gribonval. Restricted isometry constants where ℓ^p sparse recovery can fail for $0 < p \le 1$. *Information Theory, IEEE Transactions on*, 55(5):2203–2214, 2009.

[5] S. Foucart and H. Rauhut. *A mathematical introduction to compressive sensing*. Springer, 2013.

[6] G. Puy, M. E. Davies, and R. Gribonval. Recipes for stable linear embeddings from hilbert spaces to $\mathbb{R}^m$. *arXiv preprint arXiv:1509.06947*, 2015.

[7] E. Soubies, L. Blanc-Féraud, and G. Aubert. A continuous exact ℓ_0 penalty (cel0) for least squares regularized problem. *SIAM Journal on Imaging Sciences*, 8(3):1607–1639, 2015.

[8] Y. Traonmilin and R. Gribonval. Stable recovery of low-dimensional cones in Hilbert spaces: One RIP to rule them all. *Applied And Computational Harmonic Analysis, In Press*, 2016.

[9] S. Vaiter, M. Golbabaee, J. Fadili, and G. Peyré. Model selection with low complexity priors. *Information and Inference: A Journal of the IMA*, 4(3):230–287, 2015.