Mathematics of sparsity (and a few other things)

Emmanuel Candès*

Abstract. In the last decade, there has been considerable interest in understanding when it is possible to find structured solutions to underdetermined systems of linear equations. This paper surveys some of the mathematical theories, known as compressive sensing and matrix completion, that have been developed to find sparse and low-rank solutions via convex programming techniques. Our exposition emphasizes the important role of the concept of incoherence.

Mathematics Subject Classification (2010). Primary 00A69.

Keywords. Underdetermined systems of linear equations, compressive sensing, matrix completion, sparsity, low-rank-matrices, ℓ_1 norm, nuclear norm, convex programing, Gaussian widths.

1. Introduction

Many engineering and scientific problems ask for solutions to underdetermined systems of linear equations: a system is considered underdetermined if there are fewer equations than unknowns (in contrast to an overdetermined system, where there are more equations than unknowns). Examples abound everywhere but we immediately give two concrete examples that we shall keep as a guiding thread throughout the article.

- *A compressed sensing problem.* Imagine we have a signal $x(t)$, $t = 0, 1, \ldots, n-1$, with possibly complex-valued amplitudes and let $\hat{x}$ be the discrete Fourier transform (DFT) of x defined by

$$\hat{x}(\omega) = \sum_{t=0}^{n-1} x(t) e^{-i2\pi\omega t/n}, \quad \omega = 0, 1, \ldots, n-1.$$

  In applications such as magnetic resonance imaging (MRI), it is often the case that we do not have the time to collect all the Fourier coefficients so we only

*The author gratefully acknowledges support from AFOSR under grant FA9550-09-1-0643, NSF under grant CCF-0963835 and from the Simons Foundation via the Math + X Award. He would like to thank Mahdi Soltanolkotabi and Rina Foygel Barber for their help in preparing this manuscript.

sample $m \ll n$ of them. This leads to an underdetermined system of the form $y = Ax$, where y is the vector of Fourier samples at the observed frequencies and A is the $m \times n$ matrix whose rows are correspondingly sampled from the DFT matrix.[1] Hence, we would like to recover x from a highly incomplete view of its spectrum.

- *A matrix completion problem.* Imagine we have an $n_1 \times n_2$ array of numbers $x(t_1, t_2)$ perhaps representing users' preference for a collection of items as in the famous Netflix challenge; for instance, $x(t_1, t_2)$ may be a rating given by user t_1 (e.g. Emmanuel) for movie t_2 (e.g. "The Godfather"). We do not get to see many ratings as only a few entries from the matrix x are actually revealed to us. Yet we would like to correctly infer all the unseen ratings; that is, we would like to predict how a given user would rate a movie she has not yet seen. Clearly, this calls for a solution to an underdetermined system of equations.

In both these problems we have an n-dimensional object x and information of the form

$$y_k = \langle a_k, x \rangle, \quad k = 1, \dots, m, \tag{1}$$

where m may be far less than n. Everyone knows that such systems have infinitely many solutions and thus, it is apparently impossible to identify which of these candidate solutions is indeed the correct one without some additional information. In this paper, we shall see that if the object we wish to recover has a bit of structure, then exact recovery is often possible by simple convex programming techniques.

What do we mean by structure? Our purpose, here, is to discuss two types, namely, sparsity and low rank.

- *Sparsity.* We shall say that a signal $x \in \mathbb{C}^n$ is sparse, when most of the entries of x vanish. Formally, we shall say that a signal is s-sparse if it has at most s nonzero entries. One can think of an s-sparse signal as having only s degrees of freedom (df).

- *Low-rank.* We shall say that a matrix $x \in \mathbb{C}^{n_1 \times n_2}$ has low rank if its rank r is (substantially) less than the ambient dimension $\min(n_1, n_2)$. One can think of a rank-r matrix as having only $r(n_1 + n_2 - r)$ degrees of freedom (df) as this is the dimension of the tangent space to the manifold of rank-r matrices.

The question now is whether it is possible to recover a sparse signal or a low-rank matrix—both possibly depending upon far fewer degrees of freedom than their ambient dimension suggests—from just a few linear equations. The answer is in general negative. Suppose we have a 20-dimensional vector x that happens to be 1-sparse with all coordinates equal to zero but for the last component equal to one. Suppose we have 10 equations revealing the first 10 entries of x so that $a_k = e_k$, $k = 1, \dots, 10$, where throughout e_k is the kth canonical basis vector of $\mathbb{C}^n$

[1] More generally, x might be a two- or three-dimensional image.

or $\mathbb{R}^n$ (here, $n = 20$). Then $y = 0$ and clearly no method whatsoever would be able to recover our signal x. Likewise, suppose we have a 20×20 matrix of rank 1 with a first row equal to an arbitrary vector x and all others equal to zero. Imagine that we see half the entries selected completely at random. Then with overwhelming probability we would not see all the entries in the first row, and many completions would, therefore, be feasible even with the perfect knowledge that the matrix has rank exactly one.

These simple considerations demonstrate that structure is not sufficient to make the problem well posed. To guarantee recovery from $y = Ax$ by any method whatsoever, it must be the case that the structured object x is not in the null space of the matrix A. We shall assume an *incoherence property*, which roughly says that in the sparse recovery problem, while x is sparse, the rows of A *are not*, so that each measurement y_k is a weighted sum of all the components of x. A different way to put this is to say that the *sampling vectors* a_k do not correlate well with the signal x so that each measurement contains a little bit of information about the nonzero components of x. In the matrix completion problem, however, the sampling elements are sparse since they reveal entries of the matrix x we care to infer, so clearly the matrix x cannot be sparse. As explained in the next section, the right notion of incoherence is that sparse subsets of columns (resp. rows) cannot be singular or uncorrelated with all the other columns (resp. rows). A surprise is that under such a general incoherence property as well as some randomness, solving a simple convex program usually recovers the unknown solution exactly. In addition, the number of equations one needs is—up to possible logarithmic factors—proportional to the degrees of freedom of the unknown solution. This paper examines this curious phenomenon.

2. Recovery by Convex Programming

The recovery methods studied in this paper are extremely simple and all take the form of a norm-minimization problem

$$\text{minimize} \quad \|x\| \quad \text{subject to} \quad y = Ax, \tag{2}$$

where $\|\cdot\|$ is a norm promoting the assumed structure of the solution. In our two recurring examples, these are:

- The ℓ_1 *norm* for the compressed sensing problem. The ℓ_1 norm, $\|x\|_{\ell_1} = \sum_i |x_i|$, is a convex surrogate for the ℓ_0 counting 'norm' defined as $|\{i : x_i \neq 0\}|$. It is the best surrogate in the sense that the ℓ_1 ball is the smallest convex body containing all 1-sparse objects of the form $\pm e_i$.

- The *nuclear norm*, or equivalently, *Schatten-1 norm* for the matrix completion problem defined as the sum of the singular values of a matrix X. It is the best convex surrogate to the rank functional in the sense that the nuclear ball is the smallest convex body containing all rank-1 matrices with spectral norm at most equal to 1. This is the analogue to the ℓ_1 norm in the sparse

recovery problem above since the rank functional simply counts the number of nonzero singular values.

In truth, there is much literature on the empirical performance of ℓ_1 minimization [72, 67, 66, 26, 73, 41] as well as some early theoretical results explaining some of its success [55, 35, 37, 34, 75, 40, 46]. In 2004, starting with [16] and then [32] and [20], a series of papers suggested the use of random projections as means to acquire signals and images with far fewer measurements than were thought necessary. These papers triggered a massive amount of research spanning mathematics, statistics, computer science and various fields of science and engineering, which all explored the promise of cheaper and more efficient sensing mechanisms. The interested reader may want to consult the March 2008 issue of the *IEEE Signal Processing Magazine* dedicated to this topic and [49, 39]. This research is highly active today. In this paper, however, we focus on modern mathematical developments inspired by the three early papers [16, 32, 20]: in the spirit of compressive sensing, the sampling vectors are, therefore, randomized.

Let F be a distribution of random vectors on $\mathbb{C}^n$ and let $a_1, \ldots, a_m$ be a sequence of i.i.d. samples from F. We require that the ensemble F is complete in the sense that the covariance matrix $\Sigma = \mathbb{E}\, aa^*$ is invertible (here and below, a^* is the adjoint), and say that the distribution is *isotropic* if Σ is proportional to the identity. The *incoherence parameter* is the smallest number $\mu(F)$ such that if $a \sim F$, then

$$\max_{1 \le i \le n} \quad |\langle a, e_i \rangle|^2 \le \mu(F) \tag{3}$$

holds either deterministically or with high probability, see [14] for details. If F is the uniform distribution over scaled canonical vectors such that $\Sigma = I$, then the coherence is large, i.e. $\mu = n$. If $x(t)$ were a time-dependent signal, this sampling distribution would correspond to revealing the values of the signal at randomly selected time points. If, however, the sampling vectors are spread as when F is the ensemble of complex exponentials (the rows of the DFT) matrix, the coherence is low and equal to $\mu = 1$. When $\Sigma = I$, this is the lowest value the coherence parameter can take on since by definition, $\mu \ge \mathbb{E}\, |\langle a, e_i \rangle|^2 = 1$.

Theorem 2.1 ([14]). *Let $x^\star$ be a fixed but otherwise arbitrary s-sparse vector in $\mathbb{C}^n$. Assume that the sampling vectors are isotropic ($\Sigma = I$) and let $y = Ax^\star$ be the data vector and the ℓ_1 norm be the regularizer in* (2)*. If the number of equations obeys*

$$m \ge C_\beta \cdot \mu(F) \cdot \mathrm{df} \cdot \log n, \quad \mathrm{df} = s,$$

then $x^\star$ is the unique minimizer with probability at least $1 - 5/n - e^{-\beta}$. Further, C_β may be chosen as $C_0(1+\beta)$ for some positive numerical constant C_0.

Loosely speaking, Theorem 2.1 states that if the rows of A are diverse and incoherent (not sparse), then if there is an s-sparse solution, it is unique and ℓ_1 will find it. This holds as soon as the number of equations is on the order of $s \cdot \log n$. Continuing, one can understand the probabilistic guarantee as saying that most deterministic systems with diverse and incoherent rows have this property. Hence,

Theorem 2.1 is a fairly general result with minimal assumptions on the sampling vectors, and which then encompasses many signal recovery problems frequently discussed in practice, see [14] for a non-exhaustive list.

Theorem 2.1 is also sharp in the sense that for any reasonable values of (μ, s), one can find examples for which *any* recovery algorithm would fail when presented with fewer than a constant times $\mu(F) \cdot s \cdot \log n$ random samples [14]. As hinted, our result is stated for isotropic sampling vectors for simplicity, although there are extensions which do not require Σ to be a multiple of the identity; only that it has a well-behaved condition number [53].

Three important remarks are in order. The first, is that Theorem 2.1 extends the main result from [16], which established that a s-sparse signal can be recovered from about $20 \cdot s \cdot \log n$ random Fourier samples via minimum ℓ_1 norm with high probability (or equivalently, from almost all sets with at least this cardinality). Among other implications, this mathematical fact motivated MR researchers to speed up MR scan acquisition times by sampling at a lower rate, see [56, 78] for some impressive findings. Moreover, Theorem 2.1 also sharpens and extends another earlier incoherent sampling theorem in [9]. The second is that other types of Fourier sampling theorems exist, see [43] and [79]. The third is that in the case the linear map A has i.i.d. Gaussian entries, it is possible to establish more precise sampling theorems. Section 5 is dedicated to describing a great line of research on this subject.

We now turn to the matrix completion problem. Here, the entries X_{ij} of an $n_1 \times n_2$ matrix X are revealed uniformly at random so that the sampling vectors a are of the form $e_i e_j^*$ where (i, j) is uniform over $[n_1] \times [n_2]$ ($[n] = \{1, \dots n\}$). With this,

$$X_{ij} = \left\langle e_i e_j^*, X \right\rangle$$

where $\langle \cdot, \cdot \rangle$ is the usual matrix inner product. Again, we have an isotropic sampling distribution in which $\Sigma = (n_1 n_2)^{-1} I$. We now need a notion of incoherence between the sampling vectors and the matrix X, and define the *incoherence parameter* $\mu(X)$ introduced in [15], which is the smallest number $\mu(X)$ such that

$$\begin{aligned} \max_{1 \le i \le n_1} \quad & (n_1/r) \cdot \|\pi_{\mathrm{col}(X)} e_i\|_{\ell_2}^2 \le \mu(X) \\ \max_{1 \le j \le n_2} \quad & (n_2/r) \cdot \|\pi_{\mathrm{row}(X)} e_j\|_{\ell_2}^2 \le \mu(X), \end{aligned} \tag{4}$$

where r is the rank of X and $\pi_{\mathrm{col}(X)}$ (resp. $\pi_{\mathrm{row}(X)}$) is the projection onto the column (resp. row) space of X. The coherence parameter measures the overlap or correlation between the column/row space of the matrix and the coordinate axes. Since $\sum_i \|\pi_{\mathrm{col}(X)} e_i\|_{\ell_2}^2 = \mathrm{tr}(\pi_{\mathrm{col}(X)}) = r$, we can conclude that $\mu(X) \ge 1$. Conversely, the coherence is by definition bounded above by $\max(n_1, n_2)/r$. A matrix with low coherence has column and row spaces away from the coordinate axes as in the case where they assume a uniform random orientation.[2] Conversely, a matrix with high coherence may have a column (or a row space) well aligned

[2] If the column space of X has uniform orientation, then for each i, $(n_1/r) \cdot \mathbb{E}\, \|\pi_{\mathrm{col}(X)} e_i\|_{\ell_2}^2 = 1$.

with a coordinate axis. As should become intuitive, we can only hope to recover 'incoherent' matrices; i.e. matrices with relatively low-coherence parameter values.

Theorem 2.2. *Let $X^\star$ be a fixed but otherwise arbitrary matrix of dimensions $n_1 \times n_2$ and rank r. Let y in (2) be the set of revealed entries of $X^\star$ at randomly selected locations and $\|\cdot\|$ be the nuclear norm. Then with probability at least $1 - n^{-10}$, $X^\star$ is the unique minimizer to (2) provided that the number of samples obeys*

$$m \geq C_0 \cdot \mu(X) \cdot \mathrm{df} \cdot \log^2(n_1 + n_2), \quad \mathrm{df} = r(n_1 + n_2 - r),$$

for some positive numerical constant C_0.

We have adopted a formulation emphasizing the resemblance with the earlier sparse recovery theorem. Indeed just as before, Theorem 2.2 states that one can sample without any information loss the entries of a low-rank matrix at a rate essentially proportional to the coherence times its degrees of freedom. Moreover, the sampling rate is known to be optimal up to a logarithmic factor in the sense that for any reasonable values of the pair $(\mu(X), \mathrm{rank}(X))$, there are matrices that cannot be recovered from fewer than a constant times $\mu(X) \cdot \mathrm{df} \cdot \log(n_1 + n_2)$ randomly sampled entries [21].

The role of the coherence in this theory is also very natural, and can be understood when thinking about the prediction of movie ratings. Here, we can imagine that the complete matrix of ratings has (approximately) low rank because users' preferences are correlated. Now the reason why matrix completion is possible under incoherence is that we can exploit correlations and infer how a specific user is going to like a movie she has not yet seen, by examining her ratings and learning about her general preferences, and inferring how other users with such preferences have rated this particular item. Whenever we have users or small groups of users that are very singular in the sense that their ratings are orthogonal to those of all other users, it is not possible to correctly predict their missing entries. Such matrices have large coherence. (To convince oneself, consider situations where a few users enter ratings based on the outcome of coin tosses.) An amusing example of a low-rank and incoherent matrix may be the voting patterns of senators and representatives in the U. S. Congress.

A first version of this result appeared in [15], however, with one additional technical assumption concerning the approximate orthogonality between left- and right-singular vectors. This condition appears in all the subsequent literature except in unpublished work from Xiaodong Li and the author and in [27], so that Theorem 2.2, as presented here, holds. Setting $n = \max(n_1, n_2)$, [15] proved that on the order of $\mu(X) \cdot n^{6/5} r \cdot \log n$ sampled entries are sufficient for perfect recovery, a bound which was lowered to $\mu(X) \cdot nr \cdot \log^a n$ in [21], with $a \leq 6$ and sometimes equal to 2. Later, David Gross [47], using beautiful and new arguments, demonstrated that the latter bound holds with $a = 2$. (Interestingly, all three papers exhibit completely different proofs.) For a different approach to matrix completion, please see [52].

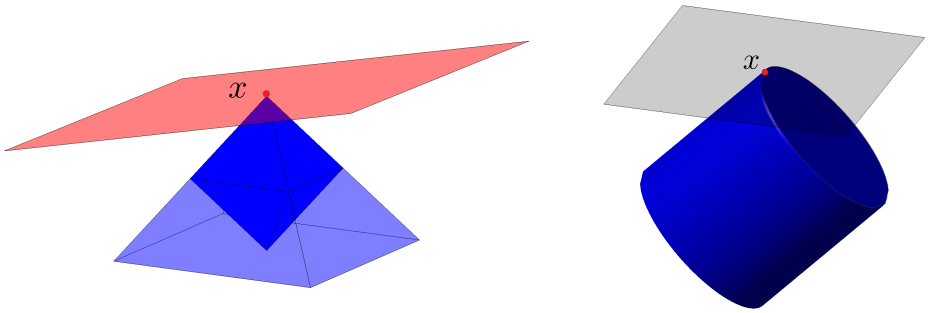
One can ask whether matrix completion is possible from more general random equations, where the sampling matrices may not have rank one, and are still

i.i.d. samples from some fixed distribution F. By now, one would believe that if the sampling matrices do not correlate well with the unknown matrix X, then matrix completion ought to be possible. This belief is correct. To give a concrete example, suppose we have an orthobasis of matrices $\mathcal{F} = \{B_j\}_{1 \le j \le n_1 n_2}$ and that we select elements from this family uniformly at random. Then [47] shows that if

$$\begin{aligned} \max_{B \in \mathcal{F}} \quad & (n_1/r) \cdot \|\pi_{\mathrm{col}(X)} B\|_F^2 \le \mu(X) \\ \max_{B \in \mathcal{F}} \quad & (n_2/r) \cdot \|B \pi_{\mathrm{row}(X)}\|_F^2 \le \mu(X), \end{aligned}$$

($\|\cdot\|_F$ is the Frobenius norm) holds along with another technical condition, Theorem 2.2 holds. Note that in the previous example where $B = e_i e_j^*$, $\|\pi_{\mathrm{col}(X)} B\|_F^2 = \|\pi_{\mathrm{col}(X)} e_i\|_{\ell_2}^2$ so that we are really dealing with the same notion of coherence.

3. Why Does This Work?



(a) ℓ_1 ball **(b)** Nuclear ball

Figure 1: Balls associated with the ℓ_1 and nuclear norms together with the affine feasible set for (2). The ball in (b) corresponds to 2×2 symmetric matrices—thus depending upon three parameters—with nuclear norm at most equal to that of x. When the feasible set is tangent to the ball, the solution to (2) is exact.

The results we have presented may seem surprising at first: why is it that with on the order of $s \cdot \log n$ random equations, ℓ_1 minimization will find the unique s-sparse solution to the system $y = Ax$? Our intent is to give an intuitive explanation of this phenomenon. Define the *cone of descent* of the norm $\|\cdot\|$ at a point x as

$$\mathcal{C} = \{h : \|x + ch\| \le \|x\| \text{ for some } c > 0\}. \tag{5}$$

This convex cone[3] is the set of non-ascent directions of $\|\cdot\|$ at x. In the literature on convex geometry, this object is known as the tangent cone. Now it is straightforward to see that a point x is the unique solution to (2) if and only if the null

[3] A cone is a set closed under positive linear combinations

space of A misses the cone of descent at x, i.e. $\mathcal{C} \cap \mathrm{null}(A) = \{0\}$. A geometric representation of this fact is depicted in Figure 1. Looking at the figure, we also begin to understand why minimizing the ℓ_1 and nuclear norms recovers sparse and low-rank objects: indeed, as the figure suggests, the tangent cone to the ℓ_1 norm is 'narrow' at sparse vectors and, therefore, even though the null space is of small codimension m, it is likely that if m is large enough, it will miss the tangent cone. A similar observation applies to the nuclear ball, which also appears pinched at low-rank objects.

As intuitive as it is, this geometric observation is far from accounting for the style of results introduced in the previous section. For instance, consider Theorem 2.1 in the setting of Fourier sampling: then we would need to show that a plane spanned by $n - m$ complex exponentials selected uniformly at random misses the tangent cone. For matrix completion, the null space is the set of all matrices vanishing at the locations of the revealed entries. There, the null space misses the cone of the nuclear ball at low-rank objects, which are sufficiently incoherent. It does not miss the cone at coherent low-rank matrices since the exact recovery property cannot hold in this case. So how do we go about proving these things?

Introduce the subdifferential of $\|\cdot\|$ at x, defined as the set of vectors

$$\partial\|x\| = \{w : \|x + h\| \geq \|x\| + \langle w, h\rangle \text{ for all } h\}. \tag{6}$$

Then x is a solution to (2) if and only if

$$\exists v \perp \mathrm{null}(A) \text{ such that } v \in \partial\|x\|.$$

For the ℓ_1 norm, letting T be the linear span of vectors with the same support as x and $T^\perp$ be its orthogonal complement (those vectors vanishing on the support of x),

$$\partial\|x\|_{\ell_1} = \{\mathrm{sgn}(x) + w : w \in T^\perp, \|w\|_{\ell_\infty} \leq 1\}, \tag{7}$$

where $\mathrm{sgn}(x)$ is the vector of signs equal to $x_i/|x_i|$ whenever $|x_i| \neq 0$ and to zero otherwise. If we would like x to be the unique minimizer, a sufficient (and almost necessary) condition is this: $T \cap \mathrm{null}(A) = \{0\}$ and

$$\exists v \perp \mathrm{null}(A) \text{ such that } v = \mathrm{sgn}(x) + w,\ w \in T^\perp,\ \|w\|_{\ell_\infty} < 1. \tag{8}$$

In the literature, such a vector v is called a *dual certificate*.

What does this mean for the Fourier sampling problem where we can only observe the Fourier transform of a signal $x(t)$, $t = 0, 1, \ldots, n-1$, at a few random frequencies $k \in \Omega \subset \{0, 1, \ldots, n-1\}$? The answer: a sparse candidate signal x is solution to the ℓ_1 minimization problem if and only if there exists a trigonometric polynomial with sparse coefficients $P(t) = \sum_{k\in\Omega} c_k \exp(i2\pi kt/n)$ obeying $P(t) = \mathrm{sgn}(x(t))$ whenever $x(t) \neq 0$ and $|P(t)| \leq 1$ otherwise. If there is no such polynomial, (2) must return a different answer. Moreover, if $T \cap \mathrm{null}(A) = \{0\}$ and there exists P as above with $|P(t)| < 1$ off the support of x, then x is the unique solution to (2).[4]

[4]The condition $T \cap \mathrm{null}(A) = \{0\}$ means that the only polynomial $P(t) = \sum_{0\leq k\leq n-1} c_k \exp(i2\pi kt/n)$, with $c_k = 0$ whenever $k \in \Omega$ and support included in that of x, is the zero polynomial $P = 0$.

Turning to the minimum nuclear norm problem, let $X = USV^*$ be a singular value decomposition. Then

$$\partial \|X\|_{S^1} = \{\operatorname{sgn}(X) + W : W \in T^\perp, \|W\|_{S^\infty} \leq 1\};$$

here, $\|\cdot\|_{S^1}$ and $\|\cdot\|_{S^\infty}$ are the nuclear and spectral norms, $\operatorname{sgn}(X)$ is the matrix defined as $\operatorname{sgn}(X) = UV^*$, and $T^\perp$ is the set of matrices with both column and row spaces orthogonal to those of X. With these definitions, everything is as before and X is the unique solution to (2) if $T \cap \operatorname{null}(A) = \{0\}$ and, swapping the ℓ_∞ norm for the spectral norm, (8) holds.

4. Some Probability Theory

We wish to show that a candidate solution $x^\star$ is solution to (2). This is equivalent to being able to construct a dual certificate, which really is the heart of the matter. Starting with [16], a possible approach is to study an ansatz, which is the solution v to:

$$\text{minimize} \quad \|v\|_{\ell_2} \quad \text{subject to} \quad v \perp \operatorname{null}(A) \text{ and } P_T v = \operatorname{sgn}(x^\star),$$

where P_T is the projection onto the linear space T defined above. If $\|\cdot\|^*$ is the norm dual to $\|\cdot\|$, then the property $\|P_{T^\perp} v\|^* < 1$ would certify optimality (with the proviso that $T \cap \operatorname{null}(A) = \{0\}$). The motivation for this ansatz is twofold: first, it is known in closed form and can be expressed as

$$v = A^* A_T (A_T^* A_T)^{-1} \operatorname{sgn}(x), \tag{9}$$

where A_T is the restriction of A to the subspace T; please observe that $A_T^* A_T$ is invertible if and only if $T \cap \operatorname{null}(A) = \{0\}$. Hence, we can study this object analytically. The second reason is that the ansatz is the solution to a least-squares problem and that by minimizing its Euclidean norm we hope to make its dual norm small as well.

At this point it is important to recall the random sampling model in which the rows of A are i.i.d. samples from a distribution F so that

$$A^* A = \sum_{k=1}^{m} a_k a_k^*$$

can be interpreted as an empirical covariance matrix. When the distribution is isotropic ($\Sigma = I$) we know that $\mathbb{E}\, A^* A = m\, I$ and, therefore, $\mathbb{E}\, A_T^* A_T = m\, I_T$. Of course, $A^* A$ cannot be close to the identity since it has rank $m \ll n$ but we can nevertheless ask whether its restriction to T is close to the identity on T. It turns out that under the stated assumptions of the theorems,

$$\frac{1}{2} I_T \preceq \frac{1}{m} A_T^* A_T \preceq \frac{3}{2} I_T, \tag{10}$$

meaning that $m^{-1} A_T^* A_T$ is reasonably close to its expectation. For our two running examples and presenting progress in a somewhat chronological fashion, [16]

and [21] established this property by combinatorial methods, following a strategy originating in the work of Eugene Wigner [81]. The idea is to develop bounds on moments of the difference between the sampled covariance matrix and its expectation,

$$H_T = I_T - m^{-1} A_T^* A_T.$$

Controlling the growth of $\mathbb{E}\,\mathrm{tr}(H_T^{2k})$ for large powers gives control of $\|H_T\|_{S^\infty}$. However, since the entries of A are in general not independent, it is not possible to invoke standard moment calculation methods, and this approach leads to delicate combinatorial issues involving statistics of various paths in the plane that can be interpreted as complicated variants of Dyck's paths.

Next, to show that the ansatz (9) is indeed a dual certificate, one can expand the inverse of $A_T^* A_T$ as a Neumann series and write it as

$$v = \sum_{j \ge 0} v_j, \quad v_j = m^{-1} A^* A_T \, H_T^j \, \mathrm{sgn}(x).$$

In the ℓ_1 problem, we would need to show that $\|P_{T^\perp} v\|_{\ell_\infty} < 1$; that is to say, for all t at which $x(t) = 0$, $|v(t)| < 1$. In [16], this is achieved by a combinatorial method bounding the size of each term $v_j(t)$ by controlling an appropriately large moment $\mathbb{E}\,|v_j(t)|^{2k}$. This strategy yields the $20 \cdot s \cdot \log n$ bound we presented earlier. In the matrix completion problem, each term v_j in the sum above is a matrix and we wish to bound the spectral norm of the random matrix $P_{T^\perp} v$. The combinatorial approach from [21] also proceeds by controlling moments of the form $\mathbb{E}\,\mathrm{tr}(z_j^* z_j)^k$, where z_j is the random matrix $z_j = P_{T^\perp} v_j$.

There is an easier way to show that the restricted sampled covariance matrix is close to its mean (10), which goes by means of powerful tools from probability theory such as the Rudelson selection theorem [64] or the operator Bernstein inequality [2]. The latter is the matrix-valued analog of the classical Bernstein inequality for sums of independent random variables and gives tail bounds on the spectral norm of a sum of mean-zero independent random matrices. This readily applies since both $I - A^* A$ and its restriction to T are of this form. One downside is that these general tools are unfortunately not as precise as combinatorial methods. Also, this is only one small piece of the puzzle, and it is not clear how one would use this to show that $\|P_{T^\perp} v\|^* < 1$, although [15] made some headway. We refer to [61] for a presentation of these ideas in the context of signal recovery.

A bit later, David Gross [47] provided an elegant construction of an inexact dual certificate he called the *golfing scheme*, and we shall dedicate the remainder of this section to presenting the main ideas behind this clever concept. To fix things, we will assume that we are working on the minimum ℓ_1 problem although all of this extends to the matrix completion problem. Our exposition is taken from [14]. To begin with, it is not hard to see that if (10) holds, then the existence of a vector $v \perp \mathrm{null}(A)$ obeying

$$\|P_T(v - \mathrm{sgn}(x))\|_{\ell_2} \le \delta \quad \text{and} \quad \|P_{T^\perp} v\|_{\ell_\infty} < 1/2, \tag{11}$$

with δ sufficiently small, certifies that x is the unique solution. This is interesting because by being a little more stringent on the size of v on $T^\perp$, we can

relax the condition $P_T v = \operatorname{sgn}(x)$ so that it only holds approximately. To see why this is true, take v as in (11) and consider the perturbation $v' = v - A^* A_T (A_T^* A_T)^{-1} P_T(\operatorname{sgn}(x) - v)$. Then $v' \perp \operatorname{null}(A)$, $P_T v' = \operatorname{sgn}(x)$ and

$$\|P_{T^\perp} v'\|_{\ell_\infty} \leq 1/2 + \|A_{T^\perp}^* A_T (A_T^* A_T)^{-1} P_T(\operatorname{sgn}(x) - v)\|_{\ell_\infty}.$$

Because the columns of A have Euclidean norm at most $\mu(F)\sqrt{m}$, then (10) together with Cauchy-Schwarz give that the second term in the right-hand side is bounded by $\delta \cdot \sqrt{2}\mu(F)$, which is less than $1/2$ if δ is sufficiently small.

Now partition A into row blocks so that from now on, A_1 are the first m_1 rows of the matrix A, A_2 the next m_2 rows, and so on. The ℓ matrices $\{A_j\}_{j=1}^{\ell}$ are independently distributed, and we have $m_1 + m_2 + \ldots + m_\ell = m$. The golfing scheme then starts with $v_0 = 0$, inductively defines

$$v_j = \frac{1}{m_j} A_j^* A_j P_T(\operatorname{sgn}(x) - v_{j-1}) + v_{j-1}$$

for $j = 1, \ldots, \ell$, and sets $v = v_\ell$. Clearly, v is in the row space of A, and thus perpendicular to the null space. To understand this scheme, we can examine the first step

$$v_1 = \frac{1}{m_1} A_1^* A_1 P_T \operatorname{sgn}(x),$$

and observe that it is perfect on the average since $\mathbb{E}\, v_1 = P_T \operatorname{sgn}(x) = \operatorname{sgn}(x)$. With finite sampling, we will not find ourselves at $\operatorname{sgn}(x)$ and, therefore, the next step should approximate $P_T(\operatorname{sgn}(x) - v_1)$, and read

$$v_2 = v_1 + \frac{1}{m_2} A_2^* A_2 P_T(\operatorname{sgn}(x) - v_1).$$

Continuing this procedure gives the golfing scheme, which stops when v_j is sufficiently close to the target. This reminds us of a golfer taking a sequence of shots to eventually put his ball in the hole, hence the name. This also has the flavor of an iterative numerical scheme for computing the ansatz (9), however, with a significant difference: at each step we use a fresh set of sampling vectors to compute the next iterate.

Set $q_j = P_T(\operatorname{sgn}(x) - v_j)$ and observe the recurrence relation

$$q_j = \Big(I_T - \frac{1}{m_j} P_T A_j^* A_j P_T\Big) q_{j-1}.$$

If the block sizes are large enough so that $\|I_T - m_j^{-1} P_T A_j^* A_j P_T\|_{S^\infty} \leq 1/2$ (this is again the property that the empirical covariance matrix does not deviate too much from the identity, compare (10)), then we see that the size of the error decays exponentially to zero since it is at least halved at each iteration.[5] We now examine

[5] Writing $H_j = I_T - m_j{}^{-1} P_T A_j^* A_j P_T$, note that we do not require that $\|H_j\|_{S^\infty} \leq 1/2$ with high probability, only that for a fixed vector $z \in T$, $\|H_j z\|_{\ell_2} \leq \|z\|_{\ell_2}/2$, since H_j and q_{j-1} are independent. This fact allows for smaller block sizes.

the size of v on $T^\perp$, that is, outside of the support of x, and compute

$$v = \sum_{j=1}^{\ell} \frac{1}{m_j} A_j^* A_j q_{j-1}.$$

The key point is that by construction, $A_j^* A_j$ and q_{j-1} are stochastically independent. In a nutshell, conditioned on q_{j-1}, $A_j^* A_j q_{j-1}$ is just a random sum of the form $\sum_k a_k \langle a_k, q_{j-1} \rangle$ and one can use standard large deviation inequalities to bound the size of each term as follows:

$$\frac{1}{m_j} \|P_T A_j^* A_j q_{j-1}\|_{\ell_\infty} \le t_j \|q_{j-1}\|_{\ell_2}$$

for some scalars $t_j > 0$, with inequality holding with large probability. Such a general strategy along with many other estimates and ideas that we cannot possibly detail in a paper of this scope, eventually yield proofs of the two theorems from Section 2. Gross' method is very general and useful, although it is generally not as precise as the combinatorial approach.

5. Gaussian Models

The last decade has seen a considerable literature, which is impressive in its achievement, about the special case where the entries of the matrix A are i.i.d. real-valued standard normal variables. As a result of this effort, the community now has a very precise understanding of the performance of both ℓ_1- and nuclear-norm minimization in this *Gaussian model*. We wish to note that [62] was the first paper to study the recovery of a low-rank matrix from Gaussian measurements, using ideas from restricted isometries.

The Gaussian model is very different from the Fourier sampling model or the matrix completion problem from Section 1. To illustrate this point, we first revisit the ansatz (9). The key point here is that when A is a Gaussian map,

$$P_{T^\perp} v = A_{T^\perp}^* q, \quad q = A_T (A_T^* A_T)^{-1} \operatorname{sgn}(x),$$

where q and $A_{T^\perp}^*$ are independent, no matter what T is [8].[6] Set d_T to be the dimension of T (this is the quantity we called degrees of freedom earlier on). Conditioned on q, $P_{T^\perp} v$ is then distributed as

$$\iota_{T^\perp} g,$$

where $\iota_{T^\perp}$ is an isometry from $\mathbb{R}^{n-d_T}$ onto $T^\perp$ and $g \sim \mathcal{N}(0, m^{-1}\|q\|_{\ell_2}^2 I)$. In the sparse recovery setting, this means that conditioned on q, the nonzero components of $P_{T^\perp} v$ are i.i.d. $\mathcal{N}(0, m^{-1}\|q\|_{\ell_2}^2)$. In addition,

$$\|q\|_{\ell_2}^2 = \left\langle \operatorname{sgn}(x), (A_T^* A_T)^{-1} \operatorname{sgn}(x) \right\rangle$$

[6] In the Gaussian model, $A_T^* A_T$ is invertible with probability one has long as m is greater or equal to the dimension of the linear space T.

and classical results in multivariate statistics assure us that up to a scaling factor, $\|q\|_{\ell_2}^2$ is distributed as an inverse chi-squared random variable with $m - d_T + 1$ degrees of freedom. From there, it is not hard to establish that just about $2s \log n$ samples taken from a Gaussian map are sufficient for perfect recovery of an s-sparse vector. Also, one can show that just about $3r(n_1 + n_2 - 5/3r)$ samples suffice for an arbitrary rank-r matrix. We refer to [8] for details and results concerning other structured recovery problems.

This section is not about these simple facts. Rather it is about the fact that under Gaussian maps, there are immediate connections between our recovery problem and deep ideas from convex geometry: as we are about to see, these connections enable to push the theory very far. Recall from Section 3 that x is the unique solution to (2) if the null space of A misses the cone of descent $\mathcal{C}$. What makes a Gaussian map special is that its null space is uniformly distributed among the set of all $(n - m)$-dimensional subspaces in $\mathbb{R}^n$. It turns out that Gordon [45] gave a precise estimate of the probability that a random uniform subspace misses a convex cone. To state Gordon's result, we need the notion of *Gaussian width* of a set $\mathcal{K} \subset \mathbb{R}^n$ defined as:

$$w(\mathcal{K}) := \mathbb{E}_g \sup_{z \in \mathcal{K} \cap \mathbb{S}^{n-1}} \langle g, z \rangle ,$$

where $\mathbb{S}^{n-1}$ is the unit sphere of $\mathbb{R}^n$ and the expectation is taken over $g \sim \mathcal{N}(0, I)$. To the best of the author's knowledge, Rudelson and Vershynin [65] were the first to recognize the importance of Gordon's result in this context.

Theorem 5.1 (Gordon's escape through the mesh lemma, [45]). *Let $\mathcal{K} \subset \mathbb{R}^n$ be a cone and A be a Gaussian map. If*

$$m \geq (w(\mathcal{K}) + t)^2 + 1,$$

then $\text{null}(A) \cap \mathcal{K} = \{0\}$ *with probability at least* $1 - e^{-t^2/2}$.

Hence, Gordon's theorem allows to conclude that slightly more than $w(\mathcal{C})$ Gaussian measurements are sufficient to recover a signal x whose cone of descent is $\mathcal{C}$. As we shall see later on, slightly fewer than $w(\mathcal{C})$ would not do the job.

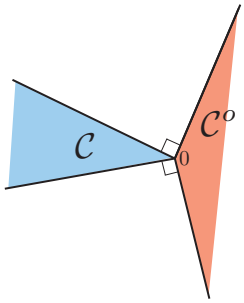


Figure 2: Schematic representation of the cone $\mathcal{C}$ and its polar $\mathcal{C}^o$.

For Theorem 5.1 to be useful we need tools to calculate these widths. One popular way of providing an upper bound on the Gaussian width of a descent cone is via polarity [68, 60, 24, 3, 76]. The *polar cone* to $\mathcal{C}$ is the set

$$\mathcal{C}^o = \{y : \langle y, z\rangle \le 0 \text{ for all } z \in \mathcal{C}\},$$

see Figure 2 for a schematic representation. For us, the cone polar to the cone of descent is the set of all directions $t \cdot w$ where $t > 0$ and $w \in \partial\|x\|$. With this, convex duality gives

$$w^2(\mathcal{C}) \le \mathbb{E}_g \min_{z\in\mathcal{C}^o} \|g - z\|_{\ell_2}^2, \tag{12}$$

where, once again, the expectation is taken over g. In words, the right-hand side is the average squared distance between a random Gaussian vector and the cone $\mathcal{C}^o$, and is called the *statistical dimension* of the descent cone denoted by $\delta(\mathcal{C})$ in [3]. (One can can check that $\delta(C) = \mathbb{E}_g \|\pi_{\mathcal{C}}(g)\|_{\ell_2}^2$ where π is the projection onto the convex cone $\mathcal{C}$.) The particular inequality (12) appears in [24] but one can trace this method to the earlier works [68, 60].[7] The point is that the statistical dimension of $\mathcal{C}$ is often relatively easy to calculate for some usual norms such as the ℓ_1 and nuclear norms, please see [24, 3] for other interesting examples. To make this claim concrete, we compute the statistical dimension of an 'ℓ_1 descent cone'.

Let $x \in \mathbb{R}^n$ be an s-sparse vector assumed without loss of generality to have its first s components positive and all the others equal to zero. We have seen that $\partial\|x\|_{\ell_1}$ is the set of vectors $w \in \mathbb{R}^n$ obeying $w_i = 1$, for all $i \le s$ and $|w_i| \le 1$ for $i > s$. Therefore,

$$\delta(\mathcal{C}) = \mathbb{E} \inf_{t\ge 0} \left\{ \sum_{j\le s} (g_j - t)^2 + \sum_{j>s} (|g_j| - t)_+^2 \right\}, \tag{13}$$

where $a_+ := \max(a, 0)$. Using $t = 2\log(n/s)$ in (13) together with some algebraic manipulations yield

$$\delta(\mathcal{C}) \le 2s\log(n/s) + 2s$$

as shown in [24]. Therefore, just about $2s\log(n/s)$ Gaussian samples are sufficient to recover an s-sparse signal by ℓ_1 minimization.

A beautiful fact is that the statistical dimension provides a sharp transition between success and failure of the convex program (2), as made very clear by the following theorem taken from Amelunxen, Lotz, McCoy and Tropp (please also see related works from Stojnic [71, 70, 69]).

Theorem 5.2 (Theorem II in [3]). *Let $x^\star \in \mathbb{R}^n$ be a fixed vector, $\|\cdot\|$ a norm, and $\delta(\mathcal{C})$ be the cone of descent at $x^\star$. Suppose A is a Gaussian map and let $y = Ax^\star$. Then for a fixed tolerance $\varepsilon \in (0, 1)$,*

$$\begin{aligned} m \le \delta(\mathcal{C}) - a_\varepsilon\sqrt{n} \quad &\Longrightarrow \quad (2) \text{ succeeds with probability} \le \varepsilon; \\ m \ge \delta(\mathcal{C}) + a_\varepsilon\sqrt{n} \quad &\Longrightarrow \quad (2) \text{ succeeds with probability} \ge 1 - \varepsilon. \end{aligned}$$

[7] There is an inequality in the other direction, $w^2(\mathcal{C}) \le \delta(C) \le w^2(C)+1$ so that the statistical dimension of a convex cone is a real proxy for its Gaussian width.

The quantity $a_\varepsilon = \sqrt{8 \log(4/\varepsilon)}$.

In other words, there is a phase transition of width at most a constant times root n around the statistical dimension. Later in this section, we discuss some history behind this result.

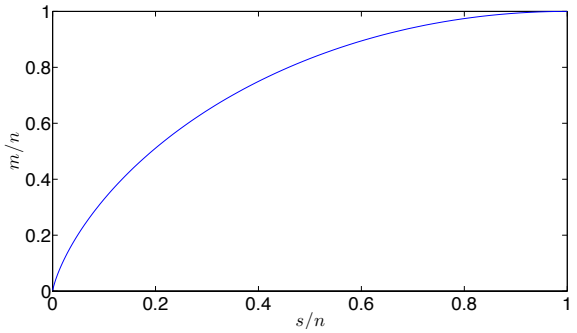


Figure 3: The curve $\psi(\rho)$.

It is possible to develop accurate estimates of the statistical dimension for ℓ_1- and nuclear-descent cones. For the the ℓ_1 norm, it follows from (13) that

$$\begin{aligned} \delta(\mathcal{C}) &\le \inf_{t \ge 0} \ \mathbb{E} \ \left\{ \sum_{j \le s} (g_j - t)^2 + \sum_{j > s} (|g_j| - t)_+^2 \right\} \\ &= \inf_{t \ge 0} \ \mathbb{E} \ \left\{ s \cdot (g_1 - t)^2 + (n - s) \cdot (|g_1| - t)_+^2 \right\} \\ &= n \cdot \psi(s/n), \end{aligned} \tag{14}$$

where the function $\psi : [0, 1] \to [0, 1]$ shown in Figure 3 is defined as

$$\psi(\rho) = \inf_{t \ge 0} \quad \left\{ \rho \cdot \mathbb{E}(g - t)^2 + (1 - \rho) \cdot \mathbb{E}(|g| - t)_+^2 \right\}, \quad g \sim \mathcal{N}(0, 1).$$

There is a connection to estimation theory: let $z \sim \mathcal{N}(\mu, 1)$ and consider the soft-thresholding rule defined as

$$\eta(z; \lambda) = \begin{cases} z - \lambda, & z > \lambda, \\ 0, & || \le \lambda, \\ z + \lambda, & z < -\lambda. \end{cases}$$

Define its risk or mean-square error at μ (when the mean of z is equal to μ) as

$$r(\mu, \lambda) = \mathbb{E}(z - \mu)^2.$$

Then with $r(\infty, \lambda) = \lim_{\mu \to \infty} r(\mu, \lambda) = (1 + \lambda^2)$,

$$\psi(\rho) = \inf_{\lambda \ge 0} \quad \left\{ \rho \cdot r(\infty, \lambda) + (1 - \rho) \cdot r(0, \lambda) \right\}.$$

Informally, in large dimensions the scalar t which realizes the minimum in (13) is nearly constant (it concentrates around a fixed value) so that the upper bound (14) is tight. Formally, [3, Proposition 4.5] shows that the statistical dimension of the ℓ_1 descent cone at an s-sparse point obeys

$$\psi(s/n) - \frac{2}{\sqrt{s \cdot n}} \le \frac{\delta(\mathcal{C})}{n} \le \psi(s/n).$$

Hence, the statistical dimension is nearly equal to the total mean-square error one would get by applying a coordinate-wise soft-thresholding rule, with the best parameter λ, to the entries of a Gaussian vector $z \sim \mathcal{N}(\mu, I)$, where $\mu \in \mathbb{R}^n$ is structured as follows: it has a fraction ρ of its components set to infinity while all the others are set to zero. For small values of s, the statistical dimension is approximately equal to $2s \log(n/s)$ and equal to the leading order term in the calculation from [24] we presented earlier. This value, holding when s is small compared to n is also close to the $2s \log n$ bound given by the ansatz.

There has been much work over the last few years with the goal of characterizing as best as possible the phase transition from Theorem 5.2. As far as the author knows, the transition curve ψ first appears in the work of Donoho [33] who studied the recovery problem in an asymptotic regime, where both the ambient dimension n and the number of samples m tend to infinity in a fixed ratio. He refers to this curve as the *weak* threshold. Donoho's approach relies on the polyhedral structure of the ℓ_1 ball known as the cross-polytope in the convex geometry literature. A signal x with a fixed support of size s and a fixed sign pattern belongs to a face $\mathcal{F}$ of dimension $s - 1$. The projection of the cross-polytope—its image through the Gaussian map—is a polytope and it is rather elementary to see that ℓ_1 minimization recovers x (and any signal in the same face) if the face is conserved, i.e. if the image of $\mathcal{F}$ is a face of the projected polytope. Donoho [33] and Donoho and Tanner [31] leveraged pioneering works by Vershik and Sporyshev and by other authors on polytope-angle calculations to understand when low-dimensional faces are conserved; they established that the curve ψ asymptotically describes a transition between success and failure (we forgo some subtleties cleared in [3]). [31] as well as related works [38] also study projections conserving all low-dimensional faces (the *strong* threshold).

One powerful feature about the approach based on Gaussian process theory described above, is that it is not limited to polytopes. Stojnic [68] used Gordon's work to establish empirically sharp lower bounds for the number of measurements required for the ℓ_1-norm problem. These results are asymptotic in nature and improve, in some cases, on earlier works. Oymak and Hassibi [60] used these ideas to give bounds on the number of measurements necessary to recover a low-rank matrix in the Gaussian model, see also [63]. In the square $n \times n$ case, for small rank, simulations in [60] show that about $4nr$ measurements may suffice for recovery (recall that the ansatz gives a nonasymptotic bound of about $6nr$). Chandrasekaran, Recht, Parrilo and Willsky [24] derived the first precise nonasymptotic bounds, and demonstrated how applicable the Gaussian process theory really is. Amelunxen et al. [3] bring definitive answers, and in some sense, this work

represents the culmination of all these efforts, even though some nice surprises continue to come around, see [42] for example. Finally, heuristic arguments from statistical physics can also explain the phase transition at ψ, see [36]. These heuristics have been justified rigorously in [5].

6. How Broad Is This?

Applications of sparse signal recovery techniques are found everywhere in science and technology. These are mostly well known and far too numerous to review. Matrix completion is a newer topic, which also comes with a very rich and diverse set of applications in fields ranging from computer vision [74] to system identification in control [57], multi-class learning in data analysis [1], global positioning—e.g. of sensors in a network—from partial distance information [7], and quantum-state tomography [48]. The list goes on and on, and keeps on growing. As the theory and numerical tools for matrix completion develop, new applications are discovered, which in turn call for even more theory and algorithms... Our purpose in this section is not to review all these applications but rather to give a sense of the breadth of the mathematical ideas we have introduced thus far; we hope to achieve this by discussing two examples from the author's own work.

Phase retrieval. Our first example concerns the fundamental phase retrieval problem, which arises in many imaging problems for the simple reason that photographic plates, CCDs and other light detectors can only measure the intensity of an electromagnetic wave as opposed to measuring its phase. For instance, consider X-ray crystallography, which is a well-known technique for determining the atomic structure of a crystal: there, a collimated beam of X-rays strikes a crystal; these rays then get diffracted by the crystal or sample and the intensity of the diffraction pattern is recorded. Mathematically, if $x(t_1, t_2)$ is a discrete two-dimensional object of interest, then to cut a long story short, one essentially collects data of the form

$$y(\omega_1, \omega_2) = \left| \sum_{t_1, t_2}^{n-1} x(t_1, t_2) e^{-i2\pi(\omega_1 t_1 + \omega_2 t_2)} \right|^2, \quad (\omega_1, \omega_2) \in \Omega, \tag{15}$$

where Ω is a sampled set of frequencies in $[0, 1]^2$. The question is then how one can invert the Fourier transform from phaseless measurements. Or equivalently, how can we infer the phase of the diffraction pattern when it is completely missing? This question arises in many fields ranging from astronomical imaging to speech analysis and is, therefore, of significance.

While it is beyond the scope of this paper to review the immense literature on phase retrieval, it is legitimate to ask in which way this is related to the topics discussed in this paper. After all, the abstract formulation of the phase retrieval problem asks us to solve a system of quadratic equations,

$$y_k = |\langle a_k, x \rangle|^2, \quad k = 1, \dots, m, \tag{16}$$

in which x is an n-dimensional complex or real-valued object; this is (15) with the a_k's being trigonometric exponentials. This is quite different from the underdetermined linear systems considered thus far. In passing, solving quadratic equations is known to be notoriously difficult (NP-hard) [6, Section 4.3].

As it turns out, the phase retrieval problem can be cast as a matrix completion problem [10], see also [22] for a similar observation in a different setup. To see this, introduce the $n \times n$ positive semidefinite Hermitian matrix variable $X \in \mathcal{S}^{n\times n}$ equal to xx^*, and observe that

$$|\langle a_k, x\rangle|^2 = \mathrm{tr}(a_k a_k^* x x^*) = \mathrm{tr}(A_k X), \quad A_k = a_k a_k^*. \tag{17}$$

By lifting the problem into higher dimensions, we have turned quadratic equations into linear ones! Suppose that (16) has a solution x_0. Then there obviously is a rank-one solution to the linear equations in (17), namely, $X_0 = x_0 x_0^*$. Thus the phase retrieval problem is equivalent to finding a rank-one matrix from linear equations of the form $y_k = \mathrm{tr}(a_k a_k^* X)$. This is a rank-one matrix completion problem! Since the nuclear norm of a positive definite matrix is equal to the trace, the natural convex relaxation called *PhaseLift* in [10] reads:

$$\text{minimize} \quad \mathrm{tr}(X) \quad \text{subject to} \quad X \succeq 0, \quad \mathrm{tr}(a_k a_k^* X) = y_k,\ k \in [m]. \tag{18}$$

Similar convex relaxations for optimizing quadratic objectives subject to quadratic constraints are known as Schor's semidefinite relaxations, see [6, Section 4.3] and [44] on the MAXCUT problem from graph theory. The reader is also encouraged to read [80] to learn about another convex relaxation.

Clearly, whatever the sampling vectors might be, we are very far from the Gaussian maps studied in the previous section.[8] Yet, a series of recent papers have established that PhaseLift succeeds in recovering the missing phase of the data (and, hence, in reconstructing the signal) in various stochastic models of sampling vectors, ranging from highly structured Fourier-like models to unstructured Gaussian-like models. In fact, the next theorem shows an even stronger result than necessary for PhaseLift, namely, that there is only one matrix satisfying the feasibility conditions of (18) and, therefore, PhaseLift must recover $x_0 x_0^*$ exactly with high probability.

Theorem 6.1. *Suppose the a_k's are independent random vectors uniformly distributed on the sphere—equivalently, independent complex-valued Gaussian vectors—and let $\mathcal{A} : \mathbb{C}^{n\times n} \to \mathbb{R}^m$ be the linear map $\mathcal{A}(X) = \{\mathrm{tr}(a_k a_k^* X)\}_{1\le k\le m}$. Assume that*

$$m \ge c_0\, n, \tag{19}$$

where c_0 is a sufficiently large constant. Then the following holds with probability at least $1 - O(e^{-\gamma m})$: for all x_0 in $\mathbb{C}^n$, the feasibility problem

$$\{X : X \succeq 0 \text{ and } \mathcal{A}(X) = \mathcal{A}(x_0 x_0^*)\}$$

[8] Under a Gaussian map, a sample is of the form $\langle W, X\rangle$, where W is a matrix with i.i.d. $\mathcal{N}(0,1)$ entries.

has a unique point, namely, $x_0 x_0^*$. Thus, with the same probability, PhaseLift recovers any signal $x_0 \in \mathbb{C}^n$ up to a global sign factor.

This theorem states that a convenient convex program—a semidefinite program (SDP)—can recover any n-dimensional complex vector from on the order of n randomized quadratic equations. The first result of this kind appeared in [18]. As stated, Theorem 6.1 theorem can be found in [11], see also [29]. Such results are not consequences of the general theorems we presented in Section 2.

Of course the sampling vectors from Theorem 6.1 are not useful in imaging applications. However, a version of this result holds more broadly. In particular, [13] studies a physically realistic setup where one can modulate the signal of interest and then collect the intensity of its diffraction pattern, each modulation thereby producing a sort of *coded diffraction pattern*. To simplify our exposition, in one dimension we would collect the pattern

$$y(\omega) = \left| \sum_{t=0}^{n-1} x(t) d(t) e^{-i 2\pi \omega t / n} \right|^2, \quad \omega = 0, 1, \ldots, n-1, \tag{20}$$

where $d := \{d(t)\}$ is a code or modulation pattern with random entries. This can be achieved by masking the object we wish to image or by modulating the incoming beam. Then [13] shows mathematically and empirically that if one collects the intensity of a few diffraction patterns of this kind, then the solution to PhaseLift is exact.

In short, convex programming techniques and matrix completion ideas can be brought to bear, with great efficiency, on highly nonconvex quadratic problems.

Robust PCA. We now turn our attention to a problem in data analysis. Suppose we have a family of n points belonging to a high-dimensional space of dimension d, which we regard as the columns of a $d \times n$ matrix M. Many data analysis procedures begin by reducing the dimensionality by projecting each data point onto a lower dimensional subspace. *Principal component analysis* (PCA) [51] achieves this by finding the matrix X of rank k, which is closest to M in the sense that it solves:

$$\text{minimize} \quad \|M - X\| \quad \text{subject to} \quad \operatorname{rank}(X) \leq k,$$

where $\|\cdot\|$ is either the Frobenius or the usual spectral norm. The solution is given by truncating the singular value decomposition as to retain the k largest singular values. When our data points are well clustered along a lower dimensional plane, this technique is very effective.

In many real applications, however, many entries of the data matrix are typically either unreliable or missing: entries may have been entered incorrectly, sensors may have failed, occlusions in image data may have occurred, and so on. The problem is that PCA is very sensitive to outliers and few errors can throw the estimate of the underlying low-dimensional structure completely off. Researchers have long been preoccupied with making PCA robust and we cannot possibly review the

literature on the subject. Rather, our intent is again to show how this problem fits together with the themes from this paper.

Imagine we are given a $d \times n$ data matrix

$$M = L_0 + S_0,$$

where L_0 has low rank and S_0 is sparse. We observe M but L_0 and S_0 are hidden. The connection with our problem is that we have a low-rank matrix that has been corrupted in possibly lots of places but we have no idea about which entries have been tampered with. Can we recover the low-rank structure? The idea in [12] is to de-mix the low-rank and the sparse components by solving:

$$\text{minimize} \quad \|L\|_{S_1} + \lambda \|S\|_{\ell_1} \quad \text{subject to} \quad M = L + S; \tag{21}$$

here, λ is a positive scalar and abusing notation, we write $\|S\|_{\ell_1} = \sum_{ij} |S_{ij}|$ for the ℓ_1 norm of the matrix S seen as an $n \times d$ dimensional vector. Motivated by a beautiful problem in graphical modeling, Chandresakaran et al. proposed to study the same convex model [25], see also [23]. For earlier connections on ℓ_1 minimization and sparse corruptions, see [19, 82, 55]. The surprising result from [12] is that if the low-rank component is incoherent and if the nonzero entries of the sparse components occur at random locations, then (21) with $\lambda = 1/\sqrt{\max(n,d)}$ recovers L_0 and S_0 perfectly! To streamline our discussion, we sketch the statement of Theorem 1.1 from [12].[9]

Theorem 6.2 (Sketch of Theorem 1.1 in [12]). *Assume without loss of generality that $n \geq d$, and let L_0 be an arbitrary $n \times d$ matrix with coherence $\mu(L_0)$ as defined in Section 2. Suppose that the support set of S_0 is uniformly distributed among all sets of cardinality m. Then with probability at least $1 - O(n^{-10})$ (over the choice of support of S_0), (L_0, S_0) is the unique solution to (21) with $\lambda = 1/\sqrt{n}$, provided that*

$$\text{rank}(L_0) \leq C_0 \cdot d \cdot \mu(L_0)^{-1} (\log n)^{-2} \quad \text{and} \quad m \leq C_0' \cdot n \cdot d. \tag{22}$$

Above, C_0 and C_0' are positive numerical constants.

Hence, if a positive fraction of the entries from an incoherent matrix of rank at most a constant times $d/\log^2 n$ are corrupted, the convex program (21) will detect those alterations and correct them automatically. In addition, the article [12] presents analog results when entries are both missing and corrupted but we shall not discuss such extensions here. For further results, see [50, 54] and [25] for a deterministic analysis.

Figure 4 shows the geometry underlying the exact de-mixing. The fact that we need incoherence should not be surprising. Indeed if L_0 is a rank-1 matrix with one row equal to x and all the others equal to y, there is no way any algorithm can detect and recover corruptions in the x vector.

[9] Technically, [12] requires the additional technical assumption discussed in Section 2 although it is probably un-necessary thanks to the sharpening from Li and the author, and from [27] discussed earlier.

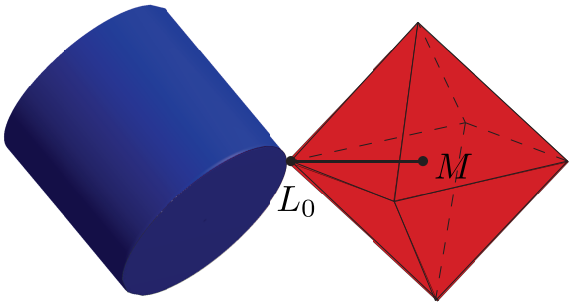


Figure 4: Geometry of the robust PCA problem. The blue body is the nuclear ball and the red the ℓ_1 ball (cross polytope). Since $S_0 = M - L_0$, $M - L_0$ is on a low-dimensional face of the cross polytope.

Finally, Figure 5 from [12] shows the practical performance of the convex programming approach to robust PCA on randomly generated problems: there is a sharp phase transition between success and failure. Looking at the numbers, we see that we can corrupt up until about 22.5% of the entries of a 400×400 matrix of rank 40, and about 37.5% of those of a matrix of rank 20.

7. Concluding Remarks

A paper of this length on a subject of this scope has to make some choices. We have certainly made some, and have consequently omitted to discuss other important developments in the field. Below is a partial list of topics we have not touched.

- We have not presented the theory based on the concept of restricted isometry property (RIP). This theory decouples the 'stochastic part' from the 'deterministic part'. In a nutshell, in the sparse recovery problem, once a sampling matrix obeys a relationship called RIP in [19] of the form (10) *for all* subspaces T spanned by at most $2s$ columns of A, then exact and stable recovery of all s-sparse signals occur [19, 17]; this is a deterministic statement. For random matrices, the stochastic part of the theory amounts to essentially showing that RIP holds [20, 4, 61]. For the matrix-completion analog, see [62].

- In almost any application the author can think of, signals are never exactly sparse, matrices do not have exactly low rank, and so on. In such circumstances, the solution to (2) continue to be accurate in the sense that if a signal is approximately sparse or a matrix has approximately low rank, then the recovered object is close. Ronald DeVore gave a plenary address at the 2006 ICM in Madrid on this topic as the theory started to develop. We refer to his ICM paper [30] as well as [28].

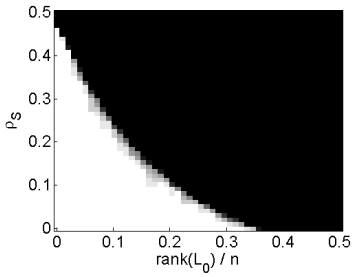


Figure 5: Fraction of correct recoveries across 10 trials, as a function of $\mathrm{rank}(L_0)$ (x-axis) and sparsity of S_0 (y-axis). Here, $n = d = 400$. In all cases, $L_0 = XY^*$ is a product of independent $n \times r$ i.i.d. $\mathcal{N}(0, 1/n)$ matrices, and $\mathrm{sgn}(S_0)$ is random. Trials are considered successful if $\|\hat{L} - L_0\|_F / \|L_0\|_F < 10^{-3}$. A white pixel indicates 100% success across trials, a black pixel 0% success, and a gray pixel some intermediate value.

- For the methods we have described to be useful, they need to be robust to noise and measurement errors. There are noise aware variants of (2) with excellent empirical and theoretical estimation properties—sometimes near-optimal. We have been silent about this, although many of the articles cited in this paper will actually contain results of this sort. To give two examples, Theorem 2.1 from Section 2 comes with variants enjoying good statistical properties, see [14]. The PhaseLift approach also comes with optimal estimation guarantees [11].

- We have not discussed algorithmic alternatives to convex programming. For instance, there are innovative greedy strategies, which can also have theoretical guarantees, e.g. under RIP see the works of Needell, Tropp, Gilbert and colleagues [77, 59, 58].

The author is thus guilty of a long string of omissions. However, he hopes to have conveyed some enthusiasm for this rich subject where so much is happening on both the theoretical and practical/empirical sides. Nonparametric structured models based on sparsity and low-rankedness are powerful and flexible and while they may not always be the best models in any particular application, they are often times surprisingly competitive.

References

[1] J. Abernethy, F. Bach, T. Evgeniou, and J.-P. Vert. Low-rank matrix factorization with attributes. Technical Report N24/06/MM, Ecole des Mines de Paris, 2006.

[2] R. Ahlswede and A. Winter. Strong converse for identification via quantum channels. *IEEE Trans. Inf. Theory*, 48(3):569–579, 2002.

[3] D. Amelunxen, M. Lotz, M. B. McCoy, and J. A. Tropp. Living on the edge: Phase transitions in convex programs with random data. 2013.

[4] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin. A simple proof of the restricted isometry property for random matrices. *Constructive Approximation*, 28(3):253–263, 2008.

[5] M. Bayati and A. Montanari. The dynamics of message passing on dense graphs, with applications to compressed sensing. *Information Theory, IEEE Transactions on*, 57(2):764–785, 2011.

[6] A. Ben-Tal and A. Nemirovski. *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*, volume 2. Society for Industrial and Applied Mathematics (SIAM), 2001.

[7] P. Biswas, T-C. Lian, T-C. Wang, and Y. Ye. Semidefinite programming based algorithms for sensor network localization. *ACM Trans. Sen. Netw.*, 2(2):188–220, 2006.

[8] E. Candès and B. Recht. Simple bounds for recovering low-complexity models. *Mathematical Programming*, 141(1-2):577–589, 2013.

[9] E. Candès and J. Romberg. Sparsity and incoherence in compressive sampling. *Inverse problems*, 23(3):969, 2007.

[10] E. J. Candès, Y. C Eldar, T. Strohmer, and V. Voroninski. Phase retrieval via matrix completion. *SIAM Journal on Imaging Sciences*, 6(1):199–225, 2013.

[11] E. J. Candès and X. Li. Solving quadratic equations via Phaselift when there are about as many equations as unknowns. *Foundations of Computational Mathematics* (to appear), 2012.

[12] E. J. Candès, X. Li, Y. Ma, and J. Wright. Robust principal component analysis? *Journal of the ACM (JACM)*, 58(3):11, 2011.

[13] E. J. Candès, X. Li, and M. Soltanolkotabi. Phase retrieval from coded diffraction patterns. *arXiv:1310.3240*, 2013.

[14] E. J. Candès and Y. Plan. A probabilistic and ripless theory of compressed sensing. *IEEE Transactions on Information Theory*, 57(11):7235–7254, 2011.

[15] E. J. Candès and B. Recht. Exact matrix completion via convex optimization. *Foundations of Computational Mathematics*, 9(6):717–772, 2009.

[16] E. J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *Information Theory, IEEE Transactions on*, 52(2):489–509, 2006.

[17] E. J. Candès, J. K. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on pure and applied mathematics*, 59(8):1207–1223, 2006.

[18] E. J Candès, T. Strohmer, and V. Voroninski. Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming. *Communications on Pure and Applied Mathematics*, 66(8):1241–1274, 2013.

[19] E. J. Candes and T. Tao. Decoding by linear programming. *Information Theory, IEEE Transactions on*, 51(12):4203–4215, 2005.

[20] E. J. Candès and T. Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *Information Theory, IEEE Transactions on*, 52(12):5406–5425, 2006.

[21] E. J. Candès and T. Tao. The power of convex relaxation: Near-optimal matrix completion. *Information Theory, IEEE Transactions on*, 56(5):2053–2080, 2010.

[22] A. Chai, M. Moscoso, and G. Papanicolaou. Array imaging using intensity-only measurements. *Inverse Problems*, 27(1), 2011.

[23] V. Chandrasekaran, P. A. Parrilo, and A. S. Willsky. Latent variable graphical model selection via convex optimization. *Annals of Statistics*, 40(4):1935–1967, 2012.

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

[25] V. Chandrasekaran, S. Sanghavi, P. A. Parrilo, and A. S. Willsky. Rank-sparsity incoherence for matrix decomposition. *SIAM Journal on Optimization*, 21(2):572–596, 2011.

[26] S .S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, 1998.

[27] Y. Chen. Incoherence-optimal matrix completion. *arXiv:1310.0154*, 2013.

[28] A. Cohen, W. Dahmen, and R. DeVore. Compressed sensing and best k-term approximation. *Journal of the American Mathematical Society*, 22(1):211–231, 2009.

[29] L. Demanet and P. Hand. Stable optimizationless recovery from phaseless linear measurements. *arXiv:1208.1803*, 2012.

[30] R. DeVore. Optimal computation. In *Proceedings oh the International Congress of Mathematicians: Madrid, August 22-30, 2006: invited lectures*, pages 187–215, 2006.

[31] D. Donoho and J. Tanner. Counting faces of randomly projected polytopes when the projection radically lowers dimension. *Journal of the American Mathematical Society*, 22(1):1–53, 2009.

[32] D. L. Donoho. Compressed sensing. *Information Theory, IEEE Transactions on*, 52(4):1289–1306, 2006.

[33] D. L. Donoho. High-dimensional centrally symmetric polytopes with neighborliness proportional to dimension. *Discrete & Computational Geometry*, 35(4):617–652, 2006.

[34] D. L. Donoho and X. Huo. Uncertainty principles and ideal atomic decomposition. *Information Theory, IEEE Transactions on*, 47(7):2845–2862, 2001.

[35] D. L. Donoho and B. F. Logan. Signal recovery and the large sieve. *SIAM Journal on Applied Mathematics*, 52(2):577–591, 1992.

[36] D. L. Donoho, A. Maleki, and A. Montanari. Message-passing algorithms for compressed sensing. *Proceedings of the National Academy of Sciences*, 106(45):18914–18919, 2009.

[37] D. L. Donoho and P. B. Stark. Uncertainty principles and signal recovery. *SIAM Journal on Applied Mathematics*, 49(3):906–931, 1989.

[38] D. L. Donoho and J. Tanner. Counting the faces of randomly-projected hypercubes and orthants, with applications. *Discrete & Computational Geometry*, 43(3):522–541, 2010.

[39] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, Ting Sun, K. F. Kelly, and R. G. Baraniuk. Single-Pixel Imaging via Compressive Sampling. *Signal Processing Magazine, IEEE*, 25(2):83–91, March 2008.

[40] M. Elad and A. M. Bruckstein. A generalized uncertainty principle and sparse representation in pairs of bases. *Information Theory, IEEE Transactions on*, 48(9):2558–2567, 2002.

[41] M. Elad, J.-L Starck, D. L. Donoho, and P. Querre. Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA). *ACHA*, 19(3):340–358, 2005.

[42] R. Foygel and L. Mackey. Corrupted sensing: Novel guarantees for separating structured signals. *Information Theory, IEEE Transactions on*, 60(2):1223–1247, 2014.

[43] A. Gilbert, S Muthukrishnan, and M Strauss. Improved time bounds for near-optimal sparse Fourier representations. In *Optics & Photonics 2005*, pages 59141A–59141A. International Society for Optics and Photonics, 2005.

[44] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42(6):1115–1145, 1995.

[45] Y. Gordon. *On Milman's inequality and random subspaces which escape through a mesh in* $\mathbb{R}^n$. Springer, 1988.

[46] R. Gribonval and M. Nielsen. Sparse representations in unions of bases. *Information Theory, IEEE Transactions on*, 49(12):3320–3325, 2003.

[47] D. Gross. Recovering low-rank matrices from few coefficients in any basis. *Information Theory, IEEE Transactions on*, 57(3):1548–1566, 2011.

[48] D. Gross, Y. Liu, S. T. Flammia, S. Becker, and J. Eisert. Quantum-state tomography via compressed sensing. *Physical Review Letters*, 105(15), 2010.

[49] B. Hayes. The best bits: A new technology called compressive sensing slims down data at the source. *American scientist*, 97(4):276, 2009.

[50] D. Hsu, S. M. Kakade, and T. Zhang. Robust matrix decomposition with sparse corruptions. *Information Theory, IEEE Transactions on*, 57(11):7221–7234, 2011.

[51] I. Jolliffe. *Principal component analysis*. Wiley Online Library, 2005.

[52] R. H. Keshavan, A. Montanari, and S. Oh. Matrix completion from a few entries. *Information Theory, IEEE Transactions on*, 56(6):2980–2998, 2010.

[53] R. Kueng and D. Gross. Ripless compressed sensing from anisotropic measurements. *Linear Algebra and its Applications*, 441:110–123, 2014.

[54] X. Li. Compressed sensing and matrix completion with constant proportion of corruptions. *Constructive Approximation*, 37(1):73–99, 2013.

[55] B. F. Logan. *Properties of high-pass signals*. PhD thesis, Columbia Univ., New York, 1965.

[56] M. Lustig, D. L. Donoho, and J. M. Pauly. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn. Reson. Med.*, 58(6):1192–1195, 2007.

[57] M. Mesbahi and G. P. Papavassilopoulos. On the rank minimization problem over a positive semidefinite linear matrix inequality. *IEEE Transactions on Automatic Control*, 42(2):239–243, 1997.

[58] D. Needell and J. A. Tropp. Cosamp: Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, 26(3):301–321, 2009.

[59] D. Needell and R. Vershynin. Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit. *Foundations of computational mathematics*, 9(3):317–334, 2009.

[60] S. Oymak and B. Hassibi. New null space results and recovery thresholds for matrix rank minimization. *arXiv preprint arXiv:1011.6326*, 2010.

[61] H. Rauhut. Compressive sensing and structured random matrices. *Theoretical foundations and numerical methods for sparse recovery*, 9:1–92, 2010.

[62] B. Recht, M. Fazel, and P. A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM review*, 52(3):471–501, 2010.

[63] B. Recht, W. Xu, and B. Hassibi. Null space conditions and thresholds for rank minimization. *Mathematical programming*, 127(1):175–202, 2011.

[64] M. Rudelson. Random vectors in the isotropic position. *J. Funct. Anal.*, 164(1):60–72, 1999.

[65] M. Rudelson and R. Vershynin. On sparse reconstruction from Fourier and Gaussian measurements. *Communications on Pure and Applied Mathematics*, 61(8):1025–1045, 2008.

[66] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992.

[67] F. Santosa and W. W. Symes. Linear inversion of band-limited reflection seismograms. *SIAM Journal on Scientific and Statistical Computing*, 7(4):1307–1330, 1986.

[68] M. Stojnic. Various thresholds for ℓ-optimization in compressed sensing. 2009.

[69] M. Stojnic. A framework to characterize performance of lasso algorithms. *arXiv:1303.7291*, 2013.

[70] M. Stojnic. A performance analysis framework for socp algorithms in noisy compressed sensing. *arXiv:1304.0002*, 2013.

[71] M. Stojnic. Upper-bounding l1-optimization weak thresholds. *arXiv:1303.7289*, 2013.

[72] H. L. Taylor, S. C. Banks, and J. F. McCoy. Deconvolution with the ℓ_1 norm. *Geophysics*, 44(1):39–52, 1979.

[73] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.

[74] C. Tomasi and T. Kanade. Shape and motion from image streams under orthography: a factorization method. *International Journal of Computer Vision*, 9(2):137–154, 1992.

[75] J. A. Tropp. Just relax: Convex programming methods for identifying sparse signals in noise. *Information Theory, IEEE Transactions on*, 52(3):1030–1051, 2006.

[76] J. A. Tropp. Convex recovery of a structured signal from independent random linear measurements. *To appear in Sampling Theory, a Renaissance*, 2014.

[77] J. A. Tropp and A. C. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *Information Theory, IEEE Transactions on*, 53(12):4655–4666, 2007.

[78] J. Trzasko and A. Manduca. Highly undersampled magnetic resonance image reconstruction via homotopic-minimization. *Medical Imaging, IEEE Transactions on*, 28(1):106–121, 2009.

[79] M. Vetterli, P. Marziliano, and T. Blu. Sampling signals with finite rate of innovation. *Signal Processing, IEEE Transactions on*, 50(6):1417–1428, 2002.

[80] I. Waldspurger, A. d'Aspremont, and S. Mallat. Phase recovery, maxcut and complex semidefinite programming. *arXiv:1206.0102*, 2012.

[81] E. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Ann. of Math.*, 62:548–564, 1955.

[82] J. Wright and Y. Ma. Dense error correction via-minimization. *Information Theory, IEEE Transactions on*, 56(7):3540–3560, 2010.

Departments of Mathematics and of Statistics, Stanford University, CA 94305, USA
E-mail: candes@stanford.edu