For Most Large Underdetermined Systems of Linear Equations the Minimal ℓ^1-norm Solution is also the Sparsest Solution

David L. Donoho
Department of Statistics
Stanford University

September 16, 2004

Abstract

We consider linear equations $y = \Phi\alpha$ where y is a given vector in $\mathbf{R}^n$, Φ is a given n by m matrix with $n < m \leq An$, and we wish to solve for $\alpha \in \mathbf{R}^m$. We suppose that the columns of Φ are normalized to unit ℓ^2 norm 1 and we place uniform measure on such Φ. We prove the existence of $\rho = \rho(A)$ so that for large n, and for all Φ's except a negligible fraction, the following property holds: *For every y having a representation $y = \Phi\alpha_0$ by a coefficient vector $\alpha_0 \in \mathbf{R}^m$ with fewer than $\rho \cdot n$ nonzeros, the solution α_1 of the ℓ^1 minimization problem*

$$\min \|x\|_1 \text{ subject to } \Phi\alpha = y$$

is unique and equal to α_0.

In contrast, heuristic attempts to sparsely solve such systems – greedy algorithms and thresholding – perform poorly in this challenging setting.

The techniques include the use of random proportional embeddings and almost-spherical sections in Banach space theory, and deviation bounds for the eigenvalues of random Wishart matrices.

Key Words and Phrases. Solution of Underdetermined Linear Systems. Overcomplete Representations. Minimum ℓ^1 decomposition. Almost-Euclidean Sections of Banach Spaces. Eigenvalues of Random Matrices. Sign-Embeddings in Banach Spaces. Greedy Algorithms. Matching Pursuit. Basis Pursuit.

Acknowledgements. Partial support from NSF DMS 00-77261, and 01-40698 (FRG) and ONR. Thanks to Emmanuel Candès for calling my attention to a blunder in a previous version, to Noureddine El Karoui for discussions about singular values of random matrices, and to Yaacov Tsaig for the experimental results mentioned here.

1 Introduction

Many situations in science and technology call for solutions to underdetermined systems of equations, i.e. systems of linear equations with fewer equations than unknowns. Examples in array signal processing, inverse problems, and genomic data analysis all come to mind. However, any sensible person working in such fields would have been taught to agree with the statement: "you have a system of linear equations with fewer equations than unknowns. There are infinitely many solutions." And indeed, they would have been taught well. However, the intuition imbued by that teaching would be misleading.

On closer inspection, many of the applications ask for *sparse* solutions of such systems, i.e. solutions with few nonzero elements; the interpretation being that we are sure that 'relatively few' of the candidate sources, pixels, or genes are turned 'on', we just don't know *a priori* which ones those are. Finding sparse solutions to such systems would better match the real underlying situation. It would also in many cases have important practical benefits, i.e. allowing us to install fewer antenna elements, make fewer measurements, store less data, or investigate fewer genes.

The search for sparse solutions can transform the problem completely, in many cases making unique solution possible (Lemma 2.1 below, see also [7, 8, 16, 14, 26, 27]). Unfortunately, this only seems to change the problem from an impossible one to an intractable one! Finding the sparsest solution to an general underdetermined system of equations is NP-hard [21]; many classic combinatorial optimization problems can be cast in that form.

In this paper we will see that for 'most' underdetermined systems of equations, when a sufficiently sparse solution exists, it can be found by convex optimization. More precisely, for a given ratio m/n of unknowns to equations, there is a threshold ρ so that most large n by m matrices generate systems of equations with two properties:

(a) If we run convex optimization to find the ℓ^1-minimal solution, and happen to find a solution with fewer than ρn nonzeros, then this is the unique sparsest solution to the equations; and

(b) If the result does not happen to have ρn nonzeros, there is no solution with $< \rho n$ nonzeros.

In such cases, if a sparse solution would be *very desirable* – needing far fewer than n coefficients – it may be found by convex optimization. If it is *of relatively small value* – needing close to n coefficients – finding the optimal solution requires combinatorial optimization.

1.1 Background: Signal Representation

To place this result in context, we describe its genesis.

In recent years, a large body of research has focused on the use of overcomplete signal representations, in which a given signal $S \in R^n$ is decomposed as $S = \sum \alpha_i \phi_i$ using a dictionary of $m > n$ atoms. Equivalently, we try to solve $S = \Phi\alpha$ for Φ and n by m matrix. Overcompleteness implies that $m > n$, so the problem is underdetermined. The goal is to use the freedom this allows to provide a sparse representation.

Motivations for this viewpoint were first obtained empirically, where representations of signals were obtained using in the early 1990's, eg. combinations of several orthonormal bases by Coifman and collaborators [4, 5] and combinations of several frames in by Mallat and Zhang's work on Matching Pursuit [19], and by Chen, Donoho, and Saunders in the mid 1990's [3].

A theoretical perspective showing that there is a sound mathematical basis for overcomplete representation has come together rapidly in recent years, see [7, 8, 12, 14, 16, 26, 27]. An early

result was the following: suppose that Φ is the concatenation of two orthobases, so that $m = 2n$. Suppose that the *coherence* - the maximal inner product between any pair of columns of Φ - is at most M. Suppose that $S = \Phi\alpha_0$ has at most N nonzeros. If $N < M^{-1}$, α_0 provides the unique optimally sparse representation of S. Consider the solution α_1 to the problem

$$\min \|\alpha\|_1 \text{ subject to } S = \Phi\alpha.$$

If $N \leq (1 + M^{-1})/2$ we have $\alpha_1 = \alpha_0$. In short, we can recover the sparsest representation by solving a convex optimization problem.

As an example, a signal of length n which is a superposition of no more than $\sqrt{n}/2$ total spikes and sinusoids is uniquely representable in that form and can be uniquely recovered by ℓ^1 optimization, (in this case $M = 1/\sqrt{n}$). The sparsity bound required in this result, comparable to $1/\sqrt{n}$, is disappointingly small, however, it was surprising at the time that any such result was possible. Many substantial improvements on these results have since been made [12, 8, 14, 16, 26, 27].

It was mentioned in [7] that the phenomena proved there represented only the tip of the iceberg. Computational results published there showed that for randomly-generated systems Φ one could get unique recovery even with as many as about $n/5$ nonzeros in a 2-fold overcomplete representation. Hence, empirically, even a mildly sparse representation could be exactly recovered by ℓ^1 optimization.

Very recently, Candès, Romberg and Tao [2] showed that for partial Fourier systems, formed by taking n rows *at random* from an m-by-m standard Fourier matrix, the resulting n by m matrix with overwhelming probability allowed exact equivalence between (P_0) and (P_1) in all cases where the number N of nonzeros was smaller than $cn/\log(n)$. This very inspiring result shows that equivalence is possible with a number of nonzeros almost proportional to n. Furthermore, [2] showed empirical examples where equivalence held with as many as $n/4$ nonzeros.

1.2 This Paper

In previous work, equivalence between the minimal ℓ^1 solution and the optimally sparse solution required that the sparse solution have an asymptotically negligible fraction of nonzeros. The fraction $O(n^{-1/2})$ could be accommodated in results of [7, 12, 8, 14, 26], and $O(1/\log(n))$ in [2].

In this paper we construct a large class of examples where equivalence holds even when the number of nonzeros is proportional to n. More precisely we show that there is a constant $\rho(A) > 0$ so that all but a negligible proportion of large n by m matrices Φ with $n < m \leq An$, have the following property: *for every system $S = \Phi\alpha$ allowing a solution with fewer than ρn nonzeros, ℓ^1 minimization uniquely finds that solution.* Here 'proportion of matrices' is taken using the natural uniform measure on the space of matrices with columns of unit ℓ^2 norm.

In contrast, greedy algorithms and thresholding algorithms seem to fail in this setting.

An interesting feature of our analysis is its use of techniques from Banach space theory, in particular quantitative extensions of Dvoretsky's almost spherical sections theorem, (by Milman, Kashin, Schechtman, and others), and other related tools exploiting randomness in high-dimensional spaces, including properties of the minimum eigenvalue of Wishart matrices.

Section 2 gives a formal statement of the result and the overall proof architecture; Sections 3-5 prove key lemmas; Section 6 discusses the failure of Greedy and Thresholding Procedures; Section 7 describes a geometric interpretation of these results. Section 8 discusses a heuristic that correctly predicts the empirical equivalence breakdown point. Section 9 discusses stability and well-posedness.

2 Overview

Let $\phi_1, \phi_2, \ldots, \phi_m$ be random points on the unit sphere $\mathbf{S}^{n-1}$ in $\mathbf{R}^n$, independently drawn from the uniform distribution. Let $\Phi = [\phi_1 \ldots \phi_m]$ be the matrix obtained by concatenating the resulting vectors. We denote this $\Phi_{n,m}$ when we wish to emphasize the size of the matrix.

For a vector $S \in \mathbf{R}^n$ we are interested in the sparsest possible representation of S using columns of Φ; this is given by:

$$(P_0) \qquad \min \|\alpha\|_0 \text{ subject to } \Phi\alpha = S,$$

It turns out that, if (P_0) has *any* sufficiently sparse solutions, then it will typically have a unique sparsest one.

Lemma 2.1 *On an event E having probability 1, the matrix Φ has the following* **unique sparsest solution** *property:*

> *For every vector α_0 having $\|\alpha_0\|_0 < n/2$ the vector $S = \Phi\alpha_0$ generates an instance of problem (P_0) whose solution is uniquely α_0.*

Proof. With probability one, the ϕ_i are in general position in $\mathbf{R}^n$. If there were two solutions, both with fewer than $n/2$ nonzeros, we would have $\Phi\alpha_0 = \Phi\alpha_1$ implying $\Phi(\alpha_1 - \alpha_0) = 0$, a linear relation involving n conditions satisfied using fewer than n points, contradicting general position. QED

In general, solving (P_0) requires combinatorial optimization and is impractical. The ℓ^1 norm is in some sense the convex relaxation of the ℓ^0 norm. So consider instead the minimal ℓ^1-norm representation:

$$(P_1) \qquad \min \|\alpha\|_1 \text{ subject to } \Phi\alpha = S,$$

This poses a convex optimization problem, and so in principle is more tractable than (P_0). Surprisingly, when the answer to (P_0) is sparse, it can be the same as the answer to (P_1).

Definition 2.1 *The* **Equivalence Breakdown Point** *of a matrix Φ, $EBP(\Phi)$, is the maximal number N such that, for every α_0 with fewer than N nonzeros, the corresponding vector $S = \Phi\alpha_0$ generates a linear system $S = \Phi\alpha$ for which problems (P_1) and (P_0) have identical unique solutions, both equal to α_0.*

Using known results, we have immediately that the EBP typically exceeds $c\sqrt{n/\log(m)}$.

Lemma 2.2 *For each $\eta > 0$,*

$$Prob\{EBP(\Phi_{n,m}) > \sqrt{\frac{n}{(8+\eta)\log(m)}}\} \to 1, \quad n \to \infty.$$

Proof. The mutual coherence $M = \max_{i \neq j} |\langle \phi_i, \phi_j \rangle|$ obeys $M < \sqrt{\frac{2\log(m)}{n}}(1 + o_p(1))$, compare calculations in [7, 8]. Applying [8], (P_0) and (P_1) have the same solution whenever $\|\alpha_0\|_0 < (1 + M^{-1})/2$. QED.

While it may seem that $O(\sqrt{n/\log(m)})$ is already surprisingly large, more than we 'really deserve', more soberly, this is asymptotically only a vanishing *fraction* of nonzeros. In fact, the two problems have the same solution over even a much broader range of sparsity $\|\alpha_0\|_0$, extending up to a *nonvanishing* fraction of nonzeros.

Theorem 1 *For each $A > 1$, there is a constant $\rho^*(A) > 0$ so that for every sequence (m_n) with $m_n \le An$*

$$Prob\{n^{-1}EBP(\Phi_{n,m}) \ge \rho^*(A)\} \to 1, \qquad n \to \infty.$$

In words, the overwhelming majority of n by m matrices have the property that ℓ^1 minimization will exactly recover the sparsest solution whenever it has at most $\rho^* n$ nonzeros. An explicit lower bound for ρ^* can be given based on our proof, but it is exaggeratedly small. As we point out later, empirical studies observed in computer simulations set $(3/10)n$ as the empirical breakdown point when $A = 2$, and a heuristic based on our proof quite precisely predicts the same breakdown point – see Section 8 below.

The space of $n \times m$ matrices having columns with unit norm is, of course,

$$\overset{\leftarrow m \text{ terms} \rightarrow}{\mathbf{S}^{n-1} \times \cdots \times \mathbf{S}^{n-1}}.$$

Now the probability measure we are assuming on our random matrix Φ is precisely the canonical uniform measure on this space. Hence, the above result shows that having $EBP(\Phi) \ge \rho^* n$ is a *generic* property of matrices, experienced on a set of nearly full measure.

2.1 Proof Outline

Let $S = \Phi\alpha_0$ and let $I = supp(\alpha_0)$. Suppose there is an alternate decomposition

$$S = \Phi(\alpha_0 + \delta)$$

where the perturbation δ obeys $\Phi\delta = 0$. Partitioning $\delta = (\delta_I, \delta_{I^c})$, we have

$$\Phi_I \delta_I = -\Phi_{I^c}\delta_{I^c}.$$

We will simply show that, on a certain event $\Omega_n(\rho, A)$

$$\|\delta_I\|_1 < \|\delta_{I^c}\|_1 \tag{2.1}$$

uniformly over every I with $|I| < \rho n$ and over every $\delta_I \ne 0$. Now

$$\|\alpha_0 + \delta\|_1 - \|\alpha_0\|_1 \ge \|\delta_{I^c}\|_1 - \|\delta_I\|_1.$$

It is then always the case that any perturbation $\delta \ne 0$ increases the ℓ^1 norm relative to the unperturbed case $\delta = 0$. In words, every perturbation hurts the ℓ^1 norm more off the support of α_0 than it helps the norm on the support of α_0, so it hurts the ℓ^1 norm overall, so every perturbation leads away from what, by convexity, must therefore be the global optimum.

It follows that the ℓ^1 minimizer is unique whenever $|I| < \rho n$ and the event $\Omega_n(\rho, A)$ occurs.

Formally, the event $\Omega_n(\rho, A)$ is the intersection of 3 subevents Ω_n^i, $i = 1, 2, 3$. These depend on positive constants η_i and ρ_i to be chosen later. The subevents are:

Ω_n^1 The minimum singular value of Φ_I exceeds η_1, uniformly in I with $|I| < \rho_1 n$

Ω_n^2 Denote $v = \Phi_I\delta_I$. The ℓ^1 norm $\|v\|_1$ exceeds $\eta_2\sqrt{n}\|v\|_2$, uniformly in I with $|I| < \rho_2 n$.

Ω_n^3 Let δ_{I^c} obey $v = -\Phi_{I^c}\delta_{I^c}$. The ℓ^1 norm $\|\delta_{I^c}\|_1$ exceeds $\eta_3\|v\|_1$ uniformly in I with $|I| < \rho_3 n$.

Lemmas 3.1,4.1, and 5.1 show that one can choose the ρ_i and η_i so that the complement of each of the Ω_n^i, $i = 1, 2, 3$ tends to zero exponentially fast in n. We do so. It follows, with $\rho_4 \equiv \min_i \rho_i$, that the intersection event $E_{\rho_4,n} \equiv \cap_i \Omega_n^i$ is overwhelmingly likely for large n.

When we are on the event $E_{\rho_4,n}$, Ω_n^1 gives us

$$\begin{aligned} \|\delta_I\|_1 &\leq \sqrt{|I|} \cdot \|\delta_I\|_2 \\ &\leq \sqrt{|I|}\|v\|_2/\lambda_{min}^{1/2}(\Phi_I^T\Phi_I) \\ &\leq \eta_1^{-1}|I|^{1/2}\|v\|_2. \end{aligned}$$

At the same time, Ω_n^2 gives us

$$\|v\|_1 \geq \eta_2\sqrt{n}\|v\|_2.$$

Finally, Ω_n^3 gives us

$$\|\delta_{I^c}\|_1 \geq \eta_3\|v\|_1,$$

and hence, provided

$$|I|^{1/2} < \eta_1 \cdot \eta_2 \cdot \eta_3 \cdot \sqrt{n},$$

we have (2.1), and hence ℓ^1 succeeds. In short, we just need to bound the fraction $|I|/n$.

Now pick $\rho^* = \min(\rho_4, (\eta_1 \cdot \eta_2 \cdot \eta_3)^2)$, and set $\Omega_n(\rho^*, A) = E_{\rho_4,n}$; we get $EBP(\Phi) \geq \rho^* n$ on $\Omega_n(\rho^*, A)$.

It remains to prove Lemmas 3.1, 4.1, and 5.1 supporting the above analysis.

3 Controlling the Minimal Eigenvalues

We first show there is, with overwhelming probability, a uniform bound $\eta_1(\rho, A)$ on the minimal singular value of every matrix Φ_I constructible from the matrix Φ with $|I| < \rho n$. This is of independent interest; see Section 9.

Lemma 3.1 *Let $\lambda < 1$. Define the event*

$$\Omega_{n,m,\rho,\lambda} = \{\lambda_{min}(\Phi_I^T\Phi_I) \geq \lambda, \qquad \forall |I| < \rho \cdot n\}.$$

There is $\rho_1 = \rho_1(\lambda, A) > 0$ so that, along sequences (m_n) with $m_n \leq An$,

$$P(\Omega_{n,m_n,\rho_1,\lambda}) \to 1, n \to \infty.$$

The bound $\eta_1(\rho, A) > 0$ is implied by this result; simply invert the relation $\lambda \mapsto \rho_1(\lambda, A)$ and put $\eta_1 = \lambda^{1/2}$.

3.1 Individual Result

We first study $\lambda_{min}(\Phi_I^T\Phi_I)$ for a single fixed I.

Lemma 3.2 *Let $\rho > 0$ be sufficiently small. There exist $\eta = \eta(\rho) > 0$, $\beta(\rho) > 0$ and $n_1(\rho)$, so that for $k = |I| \leq \rho n$ we have*

$$P\{\lambda_{min}(\Phi_I^T\Phi_I) \leq \eta^2\} \leq \exp(-n\beta), \qquad n > n_1.$$

Effectively our idea is to show that Φ_I is related to matrices of iid Gaussians, for which such phenomena are already known.

Without loss of generality suppose that $I = \{1, \ldots, k\}$. Let R_i, $i = 1, \ldots, k$ be iid random variables distributed $\chi_n/\sqrt{n}$, where χ_n denotes the χ_n distribution. These can be generated by taking iid standard normal RV's Z_{ij} which are independent of (ϕ_i) and setting

$$R_i = (n^{-1} \sum_{j=1}^{n} Z_{ij}^2)^{1/2}. \tag{3.1}$$

Let $x_i = R_i \cdot \phi_i$; then the x_i are iid $N(0, \frac{1}{n} I_n)$, and we view them as the columns of X. With $R = diag((R_i)_i)$, we have $\Phi_I = XR^{-1}$, and so

$$\lambda_{min}(\Phi_I^T \Phi_I) = \lambda_{min}((R^{-1})^T X^T X R^{-1}) \geq \lambda_{min}(X^T X) \cdot (\max_i R_i)^{-2}. \tag{3.2}$$

Hence, for a given $\eta > 0$ and $\epsilon > 0$, the two events

$$E = \{\lambda_{min}(X^T X) \geq (\eta + \epsilon)^2\} \qquad F = \{\max_i R_i < 1 + \epsilon/\eta\}$$

together imply

$$\lambda_{min}(\Phi_I^T \Phi_I) \geq \eta^2.$$

The following lemma will be proved in the next subsection:

Lemma 3.3 *For $u > 0$,*

$$P\{\max_i R_i > 1 - u\} \leq \exp\{-nu^2/2\}. \tag{3.3}$$

There we will also prove:

Lemma 3.4 *Let X be an n by k matrix of iid $N(0, \frac{1}{n})$ Gaussians, $k < n$. Let $\lambda_{min}(X^T X)$ denote the minimum eigenvalue of $X^T X$. For $\epsilon > 0$ and $k/n \leq \rho$,*

$$P\{\lambda_{min}(X^T X) < (1 - \sqrt{\rho} - \epsilon - t)^2\} \leq \exp(-nt^2/2), \qquad n > n_0(\epsilon, \rho). \tag{3.4}$$

Pick now $\eta > 0$ with $\eta < 1 - \sqrt{\rho}$, and choose ϵ so $2\epsilon < 1 - \sqrt{\rho} - \eta$; finally, put $t = 1 - \sqrt{\rho} - 2\epsilon - \eta$. Define $u = \epsilon/\eta$. Then by Lemma 3.4

$$P(E^c) \leq \exp(-nt^2/2), \qquad n > n_0(\epsilon, \rho),$$

while by Lemma 3.3

$$P(F^c) \leq \exp(-nu^2/2).$$

Setting $\beta < \min(t^2/2, u^2/2)$, we conclude that, for $n_1 = n_1(\epsilon, \rho, \beta)$,

$$P\{\lambda_{min}(\Phi_I^T \Phi_I) < \eta^2\} \leq \exp(-n\beta), \qquad n > n_1(\epsilon, \rho, \beta).$$

QED

3.2 Invoking Concentration of Measure

We now prove Lemma 3.3. Now (3.1) exhibits each R_i as a function of n iid standard normal random variables, Lipschitz with respect to the standard Euclidean metric, with Lipschitz constant $1/\sqrt{n}$. Moreover $\max_i R_i$ itself is such a Lipschitz function. By concentration of measure for Gaussian variables [18], (3.3) follows.

The proof of Lemma 3.4 depends on the observation – see Szarek [25], Davidson-Szarek [6] or El Karoui [13] – that the singular values of Gaussian matrices obey concentration of measure:

Lemma 3.5 *Let X be an n by k matrix of iid $N(0, \frac{1}{n})$ Gaussians, $k < n$. Let $s_\ell(X)$ denote the ℓ-th largest singular value of X, $s_1 \geq s_2 \geq \ldots$. Let $\sigma_{\ell;k,n} = Median(s_\ell(X))$ Then*

$$P\{s_\ell(X) < \sigma_{\ell;k,n} - t\} \leq \exp(-nt^2/2).$$

The idea is that a given singular value, viewed as a function of the entries of a matrix, is Lipschitz with respect to the Euclidean metric on $\mathbf{R}^{nk}$. Then one applies concentration of measure for scaled Gaussian variables.

As for the median $\sigma_{k;k,n}$ we remark that the well-known Marcenko-Pastur law implies that, if $k_n/n \to \rho$

$$\sigma_{k_n;k_n,n} \to 1 - \sqrt{\rho}, \qquad n \to \infty.$$

Hence, for given $\epsilon > 0$ and all sufficiently large $n > n_0(\epsilon, \rho)$, $\sigma_{k_n;k_n,n} > 1 - \sqrt{\rho} - \epsilon$. Observing that $s_k(X)^2 = \lambda_{min}(X^T X)$, gives the conclusion (3.4).

3.3 Proof of Lemma 3.1

We now combine estimates for individual I's obeying $|I| \leq \rho n$ to obtain the simultaneous result.

We need a standard combinatorial fact, used here and below:

Lemma 3.6 *For $p \in (0,1)$, let $H(p) = p\log(1/p) + (1-p)\log(1/(1-p))$ be Shannon entropy. Then*

$$\log \binom{N}{\lfloor pN \rfloor} = NH(p)(1 + o(1)), \qquad N \to \infty.$$

Now for a given $\lambda \in (0,1)$, and each index set I, define the event

$$\Omega_{n,I;\lambda} = \{\lambda_{min}(\Phi_I^T \Phi_I) \geq \lambda\}$$

Then

$$\Omega_{n,m,\rho,\lambda} = \cap_{|I| \leq \rho n} \Omega_{n,I;\lambda}.$$

By Boole's inequality,

$$P(\Omega_{n,m,\rho,\lambda}^c) \leq \sum_{|I| \leq \rho n} P(\Omega_{n,I;\lambda}^c),$$

so

$$\log P(\Omega_{n,m,\rho,\lambda}^c) \leq \log \#\{I : |I| \leq \rho n\} + \log P(\Omega_{n,I;\lambda}^c), \tag{3.5}$$

and we want the right-hand side to tend to $-\infty$. By Lemma 3.6,

$$\#\{I : |I| \leq \rho n\} = \log \binom{m_n}{\lfloor \rho n \rfloor} = AnH(\rho/A)(1 + o(1)).$$

Invoking now Lemma 3.2 we get a $\beta > 0$ so that for $n > n_0(\rho, \lambda)$, we have

$$\log P(\Omega^c_{n,I;\lambda}) \le -\beta n.$$

We wish to show that the $-\beta n$ in this relation can outweigh $AnH(\rho/A)$ in the preceding one, giving a combined result in (3.5) tending to $-\infty$. Now note that the Shannon entropy $H(p) \to 0$ as $p \to 0$. Hence for small enough ρ, $AH(\rho/A) < \beta$. Picking such a ρ – call it ρ_1 – and setting $\beta_1 = \beta - AH(\rho_1/A) > 0$ we have for $n > n_0$ that

$$\log(P(\Omega^c_{n,m,\rho_1,\lambda})) \le AnH(\rho_1/A)(1 + o(1)) - \beta n,$$

which implies an n_1 so that

$$P(\Omega^c_{n,m,\rho,\lambda}) \le \exp(-\beta_1 n), \qquad n > n_1(\rho, \lambda).$$

QED

4 Almost-Spherical Sections

Dvoretsky's theorem [10, 22] says that every infinite-dimensional Banach space contains very high-dimensional subspaces on which the Banach norm is nearly proportional to the Euclidean norm. This is called the spherical sections property, as it says that slicing the unit ball in the Banach space by intersection with an appropriate finite dimensional linear subspace will result in a slice that is effectively spherical. We need a quantitative refinement of this principle for the ℓ^1 norm in $\mathbf{R}^n$, showing that, with overwhelming probability, every operator Φ_I for $|I| < \rho n$ affords a spherical section of the ℓ^1_n ball. The basic argument we use derives from refinements of Dvoretsky's theorem in Banach space theory, going back to work of Milman and others [15, 24, 20]

Definition 4.1 *Let $|I| = k$. We say that Φ_I offers an ϵ-isometry between $\ell^2(I)$ and ℓ^1_n if*

$$(1 - \epsilon) \cdot \|\alpha\|_2 \le \sqrt{\frac{\pi}{2n}} \cdot \|\Phi_I \alpha\|_1 \le (1 + \epsilon) \cdot \|\alpha\|_2, \quad \forall \alpha \in \mathbf{R}^k. \tag{4.1}$$

Remarks: 1. The scale factor $\sqrt{\frac{\pi}{2n}}$ embedded in the definition is reciprocal to the expected ℓ^1_n norm of a standard iid Gaussian sequence. 2. In Banach space theory, the same notion would be called an $(1 + \epsilon)$-isometry [15, 22].

Lemma 4.1 Simultaneous ϵ-isometry. *Consider the event $\Omega^2_n (\equiv \Omega^2_n(\epsilon, \rho))$ that every Φ_I with $|I| \le \rho \cdot n$ offers an ϵ-isometry between $\ell^2(I)$ and ℓ^1_n. For each $\epsilon > 0$, there is $\rho_2(\epsilon) > 0$ so that*

$$P(\Omega^2_n(\epsilon, \rho_2)) \to 1, \qquad n \to \infty.$$

4.1 Proof of Simultaneous Isometry

Our approach is based on a result for individual I, which will later be extended to get a result for every I. This individual result is well known in Banach space theory, going back to [24, 17, 15]. For our proof, we repackage key elements from the proof of Theorem 4.4 in Pisier's book [22]. Pisier's argument shows that for *one specific* I, there is a *positive probability* that Φ_I offers an ϵ-isometry. We add extra 'bookkeeping' to find that the probability is actually overwhelming and later conclude that there is *overwhelming* probability that *every* I with $|I| < \rho n$ offers such isometry.

Lemma 4.2 Individual ϵ-isometry. *Fix $\epsilon > 0$. Choose δ so that*

$$(1-3\delta)(1-\delta)^{-1} \geq (1-\epsilon)^{1/2} \text{ and } (1+\delta)(1-\delta)^{-1} \leq (1+\epsilon)^{1/2}. \tag{4.2}$$

Choose $\rho_0 = \rho_0(\epsilon) > 0$ so that

$$\rho_0 \cdot (1 + 2/\delta) < \delta^2 \frac{2}{\pi},$$

and let $\beta(\epsilon)$ denote the difference between the two sides. For a subset I in $\{1, \ldots, m\}$ let $\Omega_{n,I}$ denote the event $\{$ Φ_I offers an ϵ-isometry to ℓ_n^1 $\}$. Then as $n \to \infty$,

$$\max_{|I| \leq \rho_1 n} P(\Omega_{n,I}^c) \leq 2\exp(-\beta(\epsilon)n(1+o(1))).$$

This lemma will be proved in Section 4.2. We first show how it implies Lemma 4.1.

With $\beta(\epsilon)$ as given in Lemma 4.2, we choose $\rho_2(\epsilon) < \rho_0(\epsilon)$ and satisfying

$$AH(\rho_2/A) < \beta(\epsilon),$$

where $H(p)$ is the Shannon entropy, and let $\gamma > 0$ be the difference between the two sides. Now

$$\Omega_n^2 = \cap_{|I| < \rho_2 n} \Omega_{n,I}.$$

It follows that

$$P((\Omega_n^2)^c) \leq \#\{I : |I| \leq \rho_2 n\} \cdot \max_{|I| \leq \rho n} P(\Omega_{n,I}^c).$$

Hence

$$\log(P((\Omega_n^2)^c)) \leq n[AH(\rho_2/A) - \beta(\epsilon)](1+o(1)) = -\gamma n \cdot (1+o(1)) \to -\infty.$$

4.2 Proof of Individual Isometry

We temporarily Gaussianize our dictionary elements ϕ_i. Let R_i be iid random variables distributed $\chi_n/\sqrt{n}$, where χ_n denotes the χ_n distribution. This can be generated by taking iid standard normal RV's Z_{ij} which are independent of (ϕ_i) and setting

$$R_i = (n^{-1} \sum_{j=1}^{n} Z_{ij}^2)^{1/2}. \tag{4.3}$$

Let $x_i = R_i \cdot \phi_i \cdot \sqrt{\frac{\pi}{2n}}$. Then x_i are iid n-vectors with entries iid $N(0, \frac{\pi}{2n^2})$. It follows that

$$E\|x_i\|_1 = 1.$$

Define, for each $\alpha \in \mathbf{R}^k$, $f_\alpha(x_1, \ldots, x_k) = \|\sum_i \alpha_i x_i\|_1$. If $\alpha \in \mathbf{S}^{k-1}$, the distribution of $\sum_i \alpha_i x_i$ is $N(0, \frac{\pi}{2n} I_n)$, hence $Ef_\alpha = 1$ for all $\alpha \in \mathbf{S}^{k-1}$. More transparently:

$$E\|\sum \alpha_i x_i\|_1 = \|\alpha\|_2, \qquad \forall \alpha \in \mathbf{R}^k.$$

In words, there is exact isometry between the ℓ^2 norm and the expectation of the ℓ^1 norm. We now show that over individual realizations there is approximate isometry, i.e. individual realizations are close to their expectations.

We need two standard lemmas in Banach space theory [15, 24, 17, 20]; we simplify versions in Pisier [22, Chapter 4]:

Lemma 4.3 *Let $x_i \in \mathbf{R}^n$. For each $\epsilon > 0$, choose δ obeying (4.2). Let $\mathcal{N}_\delta$ be a δ-net for $\mathbf{S}^{k-1}$ under ℓ_k^2 metric. The validity on this net of norm equivalence,*

$$1 - \delta \le \| \sum_i \alpha_i x_i \|_1 \le 1 + \delta, \qquad \forall \alpha \in \mathcal{N}_\delta,$$

implies norm equivalence on the whole space:

$$(1-\epsilon)^{1/2} \|\alpha\|_2 \le \| \sum_i \alpha_i x_i \|_1 \le (1+\epsilon)^{1/2} \|\alpha\|_2, \qquad \forall \alpha \in \mathbf{R}^k.$$

Lemma 4.4 *There is a δ-net $\mathcal{N}_\delta$ for $\mathbf{S}^{k-1}$ under ℓ_k^2 metric obeying*

$$\log(\#\mathcal{N}_\delta) \le k(1 + 2/\delta).$$

So, given $\epsilon > 0$ in the statement of our Lemma, invoke Lemma 4.3 to get a workable δ, and invoke Lemma 4.4 to get a net $\mathcal{N}_\delta$ obeying the required bound. Corresponding to each element α in the net $\mathcal{N}_\delta$, define now the event

$$E_\alpha = \{1 - \delta \le \| \sum_i \alpha_i x_i \|_1 \le 1 + \delta\}.$$

On the event $E = \cap_{\alpha \in \mathcal{N}_\delta} E_\alpha$, we may apply Lemma 4.3 to conclude that the system $(x_i : 1 \le i \le k)$ gives ϵ-equivalence between the ℓ_2 norm on $\mathbf{R}^k$ and the ℓ_1 norm on $Span(x_i)$.

Now $E_\alpha \equiv \{|f_\alpha - Ef_\alpha| > \delta\}$. We note that f_α may be viewed as a function g_α on kn iid standard normal random variables, where g_α is a Lipschitz function on $\mathbf{R}^{kn}$ with respect to the ℓ^2 metric, having Lipschitz constant $\sigma = \sqrt{\pi/2n}$. By concentration of measure for Gaussian variables [18, Section 1.2-1.3],

$$P\{|f_\alpha - Ef_\alpha| > t\} \le 2 \exp\{-t^2/2\sigma^2\}.$$

Hence

$$P(E_\alpha^c) \le 2 \exp\{-\delta^2 \cdot n \cdot \frac{2}{\pi}\}.$$

From Lemma 4.4 we have

$$\log \#\mathcal{N}_\delta \le k(1 + 2/\delta)$$

and so

$$\log(P(E^c)) \le k \cdot (1 + 2/\delta) + \log 2 - \delta^2 \cdot n \cdot \frac{2}{\pi} < \log(2) - n\beta(\epsilon).$$

We conclude that the x_i give a near-isometry with overwhelming probability.

We now de-Gaussianize. We argue that, with overwhelming probability, we also get an ϵ-isometry of the desired type for Φ_I. Setting $\gamma_i = \alpha_i \cdot \sqrt{\frac{2n}{\pi}} \cdot R_i^{-1}$, observe that

$$\sum_i \alpha_i x_i = \sum_i \gamma_i \phi_i. \tag{4.4}$$

Pick η so that

$$(1+\eta) < (1-\epsilon)^{-1/2}, \qquad (1-\eta) > (1+\epsilon)^{-1/2}. \tag{4.5}$$

Consider the event

$$G = \{(1-\eta) < R_i < (1+\eta) : i = 1, \ldots, n\},$$

On this event we have the isometry

$$(1-\eta)\cdot\|\alpha\|_2 \le \sqrt{\frac{2n}{\pi}}\cdot\|\gamma\|_2 \le (1+\eta)\cdot\|\alpha\|_2.$$

It follows that on the event $G\cap E$, we have:

$$\begin{aligned}\frac{(1-\epsilon)^{1/2}}{(1+\eta)}\cdot\sqrt{\frac{2n}{\pi}}\|\gamma\|_2 &\le (1-\epsilon)^{1/2}\|\alpha\|_2\\ &\le \|\sum_i \alpha_i x_i\|_1 \quad (=\|\sum_i \gamma_i\phi_i\|_1 \text{ by (4.4)})\\ &\le (1+\epsilon)^{1/2}\|\alpha\|_2 \le \frac{(1-\epsilon)^{1/2}}{(1-\eta)}\cdot\sqrt{\frac{2n}{\pi}}\|\gamma\|_2.\end{aligned}$$

taking into account (4.5), we indeed get an ϵ-isometry. Hence, $\Omega_{n,I}\subset G\cap E$.

Now

$$P(G^c) = P\{\max_i |R_i-1| > \eta\}.$$

By (4.3), we may also view $|R_i-1|$ as a function of n iid standard normal random variables, Lipschitz with respect to the standard Euclidean metric, with Lipschitz constant $1/\sqrt{n}$. This gives

$$P\{\max_i |R_i-1| > \eta\} \le 2m\exp\{-n\eta^2/2\} = 2m\exp\{-n\beta_G(\epsilon)\}. \tag{4.6}$$

Combining these we get that on $|I|<n\rho$,

$$P(\Omega_{n,I}^c) \le P(E^c)+P(G^c) \le 2\exp(-\beta(\epsilon)n) + 2m\exp(-\beta_G(\epsilon)n).$$

We note that $\beta_G(\epsilon)$ will certainly be larger than $\beta(\epsilon)$. QED.

5 Sign-Pattern Embeddings

Let I be any collection of indices in $\{1,\dots,m\}$; $Range(\Phi_I)$ is a linear subspace of $\mathbf{R}^n$, and on this subspace a subset Σ_I of possible *sign patterns* can be realized, i.e. sequences in $\{\pm1\}^n$ generated by

$$\sigma(k) = \mathrm{sgn}\big\{\sum_I \alpha_i\phi_i(k)\big\}, \quad 1\le k\le n.$$

Our proof of Theorem 1 needs to show that for every $v\in Range(\Phi_I)$, some approximation y to $\mathrm{sgn}(v)$ satisfies $|\langle y,\phi_i\rangle|\le 1$ for $i\in I^c$.

Lemma 5.1 Simultaneous Sign-Pattern Embedding. *Positive functions $\delta(\epsilon)$ and $\rho_3(\epsilon;A)$ can be defined on $(0,\epsilon_0)$ so that $\delta(\epsilon)\to 0$ as $\epsilon\to 0$, yielding the following properties. For each $\epsilon<\epsilon_0$, there is an event $\Omega_n^3(\equiv\Omega_{n,\epsilon}^3)$ with*

$$P(\Omega_n^3)\to 1, \qquad n\to\infty.$$

On this event, for every subset I with $|I|<\rho_3 n$, for every sign pattern $\sigma\in\Sigma_I$, there is a vector $y\,(\equiv y_\sigma)$ with

$$\|y-\epsilon\sigma\|_2 \le \epsilon\cdot\delta(\epsilon)\cdot\|\sigma\|_2, \tag{5.1}$$

and

$$|\langle\phi_i,y\rangle| \le 1, \qquad i\in I^c. \tag{5.2}$$

In words, a small multiple $\epsilon\sigma$ of any sign pattern σ almost lives in the dual ball $\{x : |\langle \phi_i, x\rangle| \le 1\}$. The key aspects are the *proportional dimension* of the constraint ρn and the *proportional distortion* required to fit in the dual ball.

Before proving this result, we indicate how it supports our claim for Ω_n^3 in the proof of Theorem 1; namely, that if $|I| < \rho_3 n$, then

$$\|\delta_{I^c}\|_1 \ge \eta_3 \|v\|_1, \tag{5.3}$$

whenever $v = -\Phi_{I^c}\delta_{I^c}$. By the duality theorem for linear programming the value of the primal program

$$\min \|\delta_{I^c}\|_1 \text{ subject to } \Phi_{I^c}\delta_{I^c} = -v \tag{5.4}$$

is at least the value of the dual

$$\max\langle v, y\rangle \text{ subject to } |\langle \phi_i, y\rangle| \le 1, \quad i \in I^c.$$

Lemma 5.1 gives us a supply of dual-feasible vectors and hence a lower bound on the dual program. Take $\sigma = \mathrm{sgn}(v)$; we can find y which is dual-feasible and obeys

$$\langle v, y\rangle \ge \langle v, \epsilon\sigma\rangle - \|y - \epsilon\sigma\|_2 \|v\|_2 \ge \epsilon\|v\|_1 - \epsilon\delta(\epsilon)\|\sigma\|_2\|v\|_2;$$

picking ϵ sufficiently small and taking into account the spherical sections theorem, we arrange that $\delta(\epsilon)\|\sigma\|_2\|v\|_2 \le \frac{1}{4}\|v\|_1$ uniformly over $v \in V_I$ where $|I| < \rho_3 n$; (5.3) follows with $\eta_3 = 3\epsilon/4$.

5.1 Proof of Simultaneous Sign-Pattern Embedding

The proof introduces a function $\beta(\epsilon)$, positive on $(0, \epsilon_0)$, which places a constraint on the size of ϵ allowed. The bulk of the effort concerns the following lemma, which demonstrates approximate embedding of a *single* sign pattern in the dual ball. The β-function allows us to cover many individual such sequences, producing our result.

Lemma 5.2 Individual Sign-Pattern Embedding. *Let $\sigma \in \{-1, 1\}^n$, let $\epsilon > 0$, and $y_0 = \epsilon\sigma$. There is an iterative algorithm, described below, producing a vector y as output which obeys*

$$|\langle \phi_i, y\rangle| \le 1, \quad i = 1, \ldots, m. \tag{5.5}$$

Let $(\phi_i)_{i=1}^m$ be iid uniform on $\mathbf{S}^{n-1}$; there is an event $\Omega_{\sigma,\epsilon,n}$ described below, having probability controlled by

$$Prob(\Omega^c_{\sigma,\epsilon,n}) \le 2n \exp\{-n\beta(\epsilon)\}, \tag{5.6}$$

for a function $\beta(\epsilon)$ which can be explicitly given and which is positive for $0 < \epsilon < \epsilon_0$. On this event,

$$\|y - y_0\|_2 \le \delta(\epsilon) \cdot \|y_0\|_2, \tag{5.7}$$

where $\delta(\epsilon)$ can be explicitly given and has $\delta(\epsilon) \to 0$ as $\epsilon \to 0$.

In short, with overwhelming probability (see (5.6)), a single sign pattern, "shrunken" appropriately, obeys (5.5) after a slight modification (indicated by (5.7)). Lemma 5.2 will be proven in a section of its own. We now show that it implies Lemma 5.1.

Lemma 5.3 *Let $V = Range(\Phi_I) \subset \mathbf{R}^n$. The number of different sign patterns σ generated by vectors $v \in V$ obeys*

$$\#\Sigma_I \le \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{|I|}.$$

Proof. This is known to statisticians as a consequence of the Vapnik-Chervonenkis VC-class theory. See Pollard [23, Chapter 4]. QED

Let again $H(p) = p\log(1/p) + (1-p)\log(1/(1-p))$ be the Shannon entropy. Notice that if $|I| < \rho n$, then

$$\log(\#\Sigma_I) \le nH(\rho)(1+o(1)),$$

while also

$$\log\#\{I : |I| < \rho n, I \subset \{1,\dots,m\}\} \le n \cdot A \cdot H(\rho/A) \cdot (1+o(1)).$$

Hence, the total number of all sign patterns generated by all operators Φ_I obeys

$$\log\#\{\sigma : \sigma \in \Sigma_I, |I| < \rho n\} \le n(H(\rho) + AH(\rho/A))(1+o(1)).$$

Now the function $\beta(\epsilon)$ introduced in Lemma 5.2 is positive, and $H(p) \to 0$ as $p \to 0$. hence, for each $\epsilon \in (0, \epsilon_0)$, there is $\rho_3(\epsilon) > 0$ obeying

$$H(\rho_3) + AH(\rho_3/A) < \beta(\epsilon).$$

Define

$$\Omega_n^3 = \cap_{|I|<\rho_3 n} \cap_{\sigma\in\Sigma_I} \Omega_{\sigma,I},$$

where $\Omega_{\sigma,I}$ denotes the instance of the event (called $\Omega_{\sigma,\epsilon,n}$ in Lemma 5.2) generated by a specific σ, I combination. On the event Ω_n^3, *every* sign pattern associated with *any* Φ_I obeying $|I| < \rho_3 n$ is almost dual feasible. Now

$$\begin{aligned} P((\Omega_n^2)^c) &\le \sum_{|I|<\rho_3 n} \sum_{\sigma\in\Sigma_I} P(\Omega_{\sigma,I}^c) \\ &\le \exp\{n(H(\rho_3) + AH(\rho_3/A))(1+o(1))\} \cdot \exp\{-n\beta(\epsilon)(1+o(1))\} \\ &= \exp\{-n(\beta(\epsilon) - (H(\rho_3) + AH(\rho_3/A)))(1+o(1))\} \to 0, \qquad n \to \infty. \end{aligned}$$

5.2 Proof of Individual Sign-Pattern Embedding

5.2.1 An Embedding Algorithm

We now develop an algorithm to create a dual feasible point y starting from a nearby almost-feasible point y_0. It is an instance of the successive projection method for finding feasible points for systems of linear inequalities [1].

Let I_0 be the collection of indices $1 \le i \le m$ with

$$|\langle \phi_i, y_0\rangle| > 1/2,$$

and then set

$$y_1 = y_0 - P_{I_0} y_0,$$

where P_{I_0} denotes the least-squares projector $\Phi_{I_0}(\Phi_{I_0}^T\Phi_{I_0})^{-1}\Phi_{I_0}^T$. In effect, we identify the indices where y_0 exceeds half the forbidden level $|\langle\phi_i, y_0\rangle| > 1$, and we "kill" those indices. Repeat the process, this time on y_1, and with a new threshold $t_1 = 3/4$. Let I_1 be the collection of indices $1 \le i \le m$ where

$$|\langle \phi_i, y_1\rangle| > 3/4,$$

and set

$$y_2 = y_0 - P_{I_0\cup I_1} y_0,$$

again "killing" the "offending" subspace. Continue in the obvious way, producing y_3, y_4, etc., with stage-dependent thresholds $t_\ell \equiv 1 - 2^{-\ell-1}$ successively closer to 1. Set

$$I_\ell = \{i : |\langle \phi_i, y_\ell \rangle| > t_\ell\},$$

and, putting $J_\ell \equiv I_0 \cup \cdots \cup I_\ell$,

$$y_{\ell+1} = y_0 - P_{J_\ell} y_0.$$

If I_ℓ is empty, then the process terminates, and set $y = y_\ell$. Termination must occur at stage $\ell^* \leq n$. (In simulations, termination often occurs at $\ell = 1, 2$, or 3). At termination,

$$|\langle \phi_i, y \rangle| \leq 1 - 2^{-\ell^*-1}, \quad i = 1, \ldots, m.$$

Hence y is definitely dual feasible. The only question is how close to y_0 it is.

5.2.2 Analysis Framework

In our analysis of the algorithm, we will study

$$\alpha_\ell = \|y_\ell - y_{\ell-1}\|_2,$$

and

$$|I_\ell| = \#\{i : |\langle \phi_i, y_\ell \rangle| > 1 - 2^{-\ell-1}\}.$$

We will propose upper bounds $\delta_{\ell,\epsilon,n}$ and $\nu_{\ell,\epsilon,n}$ for these quantities, of the form

$$\delta_{\ell,\epsilon,n} = \|y_0\|_2 \cdot \omega^\ell(\epsilon),$$

$$\nu_{\ell,\epsilon,n} = n \cdot \lambda_0 \cdot \epsilon^2 \cdot \omega^{2\ell+2}(\epsilon)/4;$$

here λ_0 can be taken in $(0, 1)$, for example as $1/2$; this choice determines the range $(0, \epsilon_0)$ for ϵ, and restricts the upper limit on ρ. $\omega(\epsilon) \in (0, 1/2)$ is to be determined below; it will be chosen so that $\omega(\epsilon) \to 0$ as $\epsilon \to 0$. We define sub-events

$$E_\ell = \{\alpha_j \leq \delta_j, \quad j = 1, \ldots, \ell, \quad |I_j| \leq \nu_j, j = 0, \ldots, \ell - 1\};$$

Now define

$$\Omega_{\sigma,\epsilon,n} = \cap_{\ell=1}^n E_\ell;$$

this event implies

$$\|y - y_0\|_2 \leq (\sum \alpha_\ell^2)^{1/2} \leq \|y_0\|_2 \cdot \omega(\epsilon)/(1 - \omega^2(\epsilon))^{1/2},$$

hence the function $\delta(\epsilon)$ referred to in the statement of Lemma 5.2 may be defined as

$$\delta(\epsilon) \equiv \omega(\epsilon)/(1 - \omega^2(\epsilon))^{1/2},$$

and the desired property $\delta(\epsilon) \to 0$ as $\epsilon \to 0$ will follow from arranging for $\omega(\epsilon) \to 0$ as $\epsilon \to 0$. We will show that, for $\beta(\epsilon) > 0$ chosen in conjunction with $\omega(\epsilon) > 0$,

$$P(E_{\ell+1}^c | E_\ell) \leq 2 \exp\{-\beta(\epsilon) n\}. \tag{5.8}$$

This implies

$$P(\Omega_{\sigma,\epsilon,n}^c) \leq 2n \exp\{-\beta(\epsilon) n\},$$

and the Lemma follows. QED

5.2.3 Transfer To Gaussianity

We again Gaussianize. Let φ_i denote random points in $\mathbf{R}^n$ which are iid $N(0, \frac{1}{n} I_n)$. We will analyze the algorithm below as if the φ's rather than the ϕ's made up the columns of Φ.

As already described in Section 4.2, there is a natural coupling between Spherical ϕ's and Gaussian φ's that justifies this transfer. As in Section 4.2 let R_i, $i = 1, \ldots, m$ be iid random variables independent of (ϕ_i) and which are individually $\chi_n/\sqrt{n}$. Then define

$$\varphi_i = R_i \phi_i, \qquad i = 1, \ldots, m.$$

If the ϕ_i are uniform on $\mathbf{S}^{n-1}$ then the φ_i are indeed $N(0, \frac{1}{n} I_n)$. The R_i are all quite close to 1 for large n. According to (4.6), for fixed $\eta > 0$,

$$P\{1 - \eta < R_i < 1 + \eta, i = 1, \ldots, m\} \geq 1 - 2m \exp\{-n\eta^2/2\}.$$

Hence it should be plausible that the difference between the ϕ_i and the φ_i is immaterial. Arguing more formally, we notice the equivalence

$$|\langle \phi_i, y \rangle| < 1 \Leftrightarrow |\langle \varphi_i, y \rangle| < R_i.$$

Running the algorithm using the φ's instead of the ϕ's, with thresholds calibrated to $1 - \eta$ via $t_0 = (1 - \eta)/2$, $t_1 = (1 - \eta) \cdot 3/4$, etc. will produce a result y obeying $|\langle \varphi_i, y \rangle| < 1 - \eta$, $\forall i$. Therefore, with overwhelming probability, the result will also obey $|\langle \phi_i, y \rangle| < 1$ $\forall i$.

However, such rescaling of thresholds is completely equivalent to rescaling of the input y_0 from $\epsilon\sigma$ to $\epsilon'\sigma$, where $\epsilon' = \epsilon(1 - \eta)$. Hence, if we can prove results with functions $\delta(\epsilon)$ and $\beta(\epsilon)$ for the Gaussian φ's, the same results are proven for the Spherical ϕ's with functions $\delta'(\epsilon) = \delta(\epsilon') = \delta(\epsilon(1 - \eta))$ and $\beta'(\epsilon) = \min(\beta(\epsilon'), \eta^2/2)$.

5.2.4 Adapted Coordinates

It will be useful to have coordinates specially adapted to the analysis of the algorithm. Given y_0, y_1, ..., define ψ_0, ψ_1, ... by Gram-Schmidt orthonormalization. In terms of these coordinates we have the following equivalent construction: Let $\alpha_0 = \|y_0\|_2$, let ξ_i, $1 \leq i \leq m$ be iid vectors $N(0, \frac{1}{n} I_n)$. We will sequentially construct vectors φ_i, $i = 1, \ldots, m$ in such a way that their joint distribution is iid $N(0, \frac{1}{n} I_n)$, but so that the algorithm has an explicit trajectory.

At stage 0, we realize m scalar Gaussians $Z_i^0 \sim^{iid} N(0, \frac{1}{n})$, threshold at level t_0, say, and define I_0 to be the set of indices so that $|\alpha_0 Z_i^0| > t_0$. For such indices i only, we define

$$\varphi_i = Z_i^0 \psi_0 + P_{\psi_0}^{\perp} \xi_i, \quad i \in I_0.$$

For all other i, we retain Z_i^0 for later use. We then define $y_1 = y_0 - P_{I_0} y_0$, $\alpha_1 = \|y_1 - y_0\|_2$ and ψ_1 by orthonormalizing $y_1 - y_0$ with respect to ψ_0.

At stage 1, we realize m scalar Gaussians $Z_i^1 \sim^{iid} N(0, \frac{1}{n})$, and define I_1 to be the set of indices not in I_0 so that $|\alpha_0 Z_i^0 + \alpha_1 Z_i^1| > t_1$. For such indices i only, we define

$$\varphi_i = Z_i^0 \psi_0 + Z_i^1 \psi_1 + P_{\psi_0, \psi_1}^{\perp} \xi_i, \quad i \in I_1.$$

For i neither in I_0 nor I_1, we retain Z_i^1 for later use. We then define $y_2 = y_0 - P_{I_0 \cup I_1} y_0$, $\alpha_2 = \|y_2 - y_1\|_2$ and ψ_2 by orthonormalizing $y_2 - y_1$ with respect to ψ_0 and ψ_1.

Continuing in this way, at some stage ℓ^* we stop, (i.e. I_{ℓ^*} is empty) and we define φ_i for all i not in $I_0 \cup \ldots \cup I_{\ell^*-1}$ (if there are any such) by

$$\varphi_i = \sum_{j=0}^{\ell^*-1} Z_i^j \psi_j + P^{\perp}_{\psi_0,\ldots,\psi_{\ell^*-1}} \xi_i, \quad i \notin I_0 \cup \ldots I_{\ell^*-1}$$

We claim that we have produced a set m of iid $N(0, \frac{1}{n} I_n)$'s for which the algorithm has the indicated trajectory we have just traced. A proof of this fact repeatedly uses independence properties of orthogonal projections of standard normal random vectors.

It is immediate that, for each ℓ up to termination, we have expressions for the key variables in the algorithm in terms of the coordinates. For example:

$$y_\ell - y_0 = \sum_{j=1}^{\ell} \alpha_j \psi_j; \qquad \|y_\ell - y_0\|_2 = (\sum_{j=1}^{\ell} \alpha_j^2)^{1/2}$$

5.2.5 Control on α_ℓ

We now develop a bound for

$$\alpha_{\ell+1} = \|y_{\ell+1} - y_\ell\|_2 = \|P_{I_\ell}(y_{\ell+1} - y_\ell)\|_2.$$

Recalling that

$$P_{I_\ell} v = \Phi_{I_\ell} (\Phi_{I_\ell}^T \Phi_{I_\ell})^{-1} \Phi_{I_\ell}^T v,$$

and putting $\lambda(I_\ell) = \lambda_{min}(\Phi_{I_\ell}^T \Phi_{I_\ell})$, we have

$$\|P_{I_\ell}(y_{\ell+1} - y_\ell)\|_2 \le \lambda(I_\ell)^{-1/2} \|\Phi_{I_\ell}^T (y_{\ell+1} - y_\ell)\|_2.$$

But $\Phi_{I_\ell} y_{\ell+1} = 0$ because $y_{\ell+1}$ is orthogonal to every φ_i, $i \in I_\ell$ by construction. Now for $i \in I_\ell$.

$$|\langle \varphi_i, y_\ell \rangle| \le |\langle \varphi_i, y_\ell - y_{\ell-1} \rangle| + |\langle \varphi_i, y_{\ell-1} \rangle| \le |\alpha_\ell Z_i^\ell| + t_\ell$$

and so

$$\|\Phi_{I_\ell}^T y_\ell\|_2 \le t_\ell |I_\ell|^{1/2} + \alpha_\ell \left(\sum_{i \in I_\ell} (Z_i^\ell)^2 \right)^{1/2} \tag{5.9}$$

We remark that

$$\{i \in I_\ell\} \Rightarrow \{|\langle \varphi_i, y_\ell \rangle| > t_\ell, |\langle \varphi_i, y_{\ell-1} \rangle| < t_{\ell-1}\} \Rightarrow \{|\langle \varphi_i, y_\ell - y_{\ell-1} \rangle| \ge t_\ell - t_{\ell-1}\};$$

putting $u_\ell = 2^{-\ell-1}/\alpha_\ell$ this gives

$$\sum_{i \in I_\ell} (Z_i^\ell)^2 \le \sum_{i \in J_{\ell-1}^c} (Z_i^\ell)^2 1_{\{|Z_i^\ell| > u_\ell\}}.$$

We conclude that

$$\alpha_{\ell+1}^2 \le 2 \cdot \lambda(I_\ell)^{-1} \big[|I_\ell| + \alpha_\ell^2 \big(\sum_{i \in J_{\ell-1}^c} (Z_i^\ell)^2 1_{\{|Z_i^\ell| > u_\ell\}} \big) \big]. \tag{5.10}$$

5.2.6 Large Deviations

Define the events

$$F_\ell = \{\alpha_\ell \le \delta_{\ell,\epsilon,n}\}, \qquad G_\ell = \{|I_\ell| \le \nu_{\ell,\epsilon,n}\},$$

so that

$$E_{\ell+1} = F_{\ell+1} \cap G_\ell \cap E_\ell.$$

Put

$$\rho_0(\epsilon) = \lambda_0 \epsilon^2.$$

On the event E_ℓ, $|J_\ell| \le \rho_0(\epsilon) n$. Recall the quantity $\eta_1(\rho, A)$ from Lemma 3.1. For some ϵ_1, $\eta_1(\rho_0(\epsilon), A)^2 \ge \lambda_0$ for all $\epsilon \in (0, \epsilon_1]$; we will restrict ourselves to this range of ϵ. On E_ℓ, $\lambda_{min}(I_\ell) > \lambda_0$. Also on E_ℓ, $u_j = 2^{-j-1}/\alpha_j > 2^{-j-1}/\delta_j = v_j$ (say) for $j \le \ell$. Now

$$P\{G_\ell^c | E_\ell\} \le P\{\sum_i 1_{\{|Z_i^\ell| > v_\ell\}} > \nu_\ell\},$$

and

$$P\{F_{\ell+1}^c | G_\ell, E_\ell\} \le P\{2 \cdot \lambda_0^{-1}\big[\nu_\ell + \delta_\ell^2\big(\sum_i \Big(Z_i^\ell\Big)^2 1_{\{|Z_i^\ell| > v_\ell\}}\big)\big] > \delta_{\ell+1}^2\}.$$

We need two simple large deviations bounds.

Lemma 5.4 *Let Z_i be iid $N(0,1)$, $k \ge 0$, $t > 2$.*

$$\frac{1}{m} \log P\{\sum_{i=1}^{m-k} Z_i^2 1_{\{|Z_i| > t\}} > m\Delta\} \le e^{-t^2/4} - \Delta/4,$$

and

$$\frac{1}{m} \log P\{\sum_{i=1}^{m-k} 1_{\{|Z_i| > t\}} > m\Delta\} \le e^{-t^2/2} - \Delta/4.$$

Applying this,

$$\frac{1}{m} \log P\{F_{\ell+1}^c | G_\ell, E_\ell\} \le e^{-\tau_\ell^2/4} - \Delta_\ell/4,$$

where

$$\tau_\ell = n \cdot v_\ell^2 = 2^{-2\ell-2}/\epsilon^2 \omega^{2\ell}(\epsilon),$$

and

$$\Delta_\ell = (\lambda_0 \delta_{\ell+1}^2/2 - \nu_\ell)/\delta_\ell^2 = \lambda_0 \omega^2(\epsilon)/4.$$

By inspection, for small ϵ and $\omega(\epsilon)$, the term of most concern is at $\ell = 0$; the other terms are always better. Putting

$$\beta(\epsilon) \equiv \beta(\epsilon; \omega) = \lambda_0 \omega^2(\epsilon)/4 - e^{-1/(16\epsilon^2 \omega^2(\epsilon))},$$

and choosing ω well, we get $\beta > 0$ on an interval $(0, \epsilon_2)$, and so

$$P\{F_{\ell+1}^c | G_\ell, E_\ell\} \le \exp(-n\beta(\epsilon)).$$

A similar analysis holds for the G_ℓ's. We get ϵ_0 in the statement of the lemma taking $\epsilon_0 = \min(\epsilon_1, \epsilon_2)$. QED

Remark: The large deviations bounds stated in Lemma 5.4 are far from best possible; we merely found them convenient in producing an explicit expression for β. Better bounds would be helpful in deriving reasonable estimates on the constant $\rho^*(A)$ in Theorem 1.

6 Geometric Interpretation

Our result has an appealing geometric interpretation. Let B_n denote the absolute convex hull of $\phi_1, \ldots, \phi_m$;

$$B_n = \{x \in \mathbf{R}^n : x = \sum_i \alpha(i)\phi_i, \quad \sum_i |\alpha(i)| \le 1\}.$$

Equivalently, B is exactly the set of vectors where $val(P_1) \le 1$. Similarly, let the octahedron $O_m \in \mathbf{R}^m$ be the absolute convex hull of the standard Kronecker basis $(e_i)_{i=1}^m$:

$$O_m = \{\alpha \in \mathbf{R}^m : \alpha = \sum_i \alpha(i)e_i, \quad \sum_{i=1}^m |\alpha(i)| \le 1\}.$$

Note that each set is polyhedral, and it is almost true that the vertices $\{\pm e_i\}$ of O_m map under Φ into vertices $\{\pm\phi_i\}$ of B_n. More precisely, the vertices of B_n are among the image vertices $\{\pm\phi_i\}$; because B_n is a convex hull, there is the possibility that for some i, ϕ_i lies strictly in the interior of B_n.

Now if ϕ_i were strictly in the interior of B_n, then we could write

$$\phi_i = \Phi\alpha_1, \quad \|\alpha_1\|_1 < 1,$$

where $i \notin supp(\alpha_1)$. It would follow that a singleton α_0 generates ϕ_i through $\phi_i = \Phi\alpha_0$, so α_0 necessarily solves (P_0), but, as

$$\|\alpha_0\|_1 = 1 > \|\alpha_1\|_1.$$

α_0 is not the solution of (P_1). So, when any ϕ_i is strictly in the interior of B_n, (P_1) and (P_0) are inequivalent problems.

Now on the event $\Omega_n(\rho^*, A)$, (P_1) and (P_0) have the same solution whenever (P_0) has a solution with $k = 1 < \rho^* n$ nonzeros. We conclude that on the event $\Omega_n(\rho^*, A)$, *the vertices of B_n are in one-one correspondence with the vertices of O_m.* Letting $Skel_0(C)$ denote the set of vertices of a polyhedral convex set C, this correspondence says:

$$Skel_0(B_n) = \Phi[Skel_0(O_m)].$$

Something much more general is true. By $(k-1)$-*face* of a polyhedral convex set C with vertex set $v = \{v_1, \ldots, \}$, we mean a $(k-1)$-simplex

$$\Sigma(v_{i_1}, \ldots, v_{i_k}) = \{x = \sum_j \alpha_j v_{i_j},\ \alpha_j \ge 0,\ \sum \alpha_j = 1\}.$$

all of whose points are extreme points of C. By $(k-1)$-*skeleton* $Skel_{k-1}(C)$ of a polyhedral convex set C, we mean the collection of all $(k-1)$-faces.

The 0-skeleton is the set of vertices, the 1-skeleton is the set of edges, etc. In general one can say that the $(k-1)$-faces of B_n form a subset of the images under Φ of the $(k-1)$-faces of O_n:

$$Skel_{k-1}(B_n) \subset \Phi[Skel_{k-1}(O_m)], \qquad 1 \le k < n.$$

Indeed, some of the image faces $\Phi\Sigma(\pm e_{i_1}, \ldots, \pm e_{i_k})$ could be at least partially interior to B_n, and hence they could not be part of the $(k-1)$-skeleton of B_n

Our main result says that much more is true; Theorem 1 is equivalent to this geometric statement:

Theorem 2 *There is a constant $\rho^* = \rho^*(A)$ so that for $n < m < An$, on an event $\Omega_n(\rho^*, A)$ whose complement has negligible probability for large n,*

$$Skel_{k-1}(B_n) = \Phi[Skel_{k-1}(O_m)], \qquad 1 \le k < \rho^* \cdot n.$$

In particular, with overwhelming probability, *the topology of every $(k-1)$-skeleton of B_n is the same as that of the corresponding $(k-1)$-skeleton of O_m*, even for k proportional to n. The topology of the skeleta of O_m is of course obvious.

7 Other Algorithms Fail

Several algorithms besides ℓ^1 minimization have been proposed for the problem of finding sparse solutions [21, 26, 9]. In this section we show that two standard approaches fail in the current setting, where ℓ^1 of course succeeds.

7.1 Subset Selection Algorithms

Consider two algorithms which attempt to find sparse solutions to $S = \Phi\alpha$ by selecting subsets I and then attempting to solve $S = \Phi_I \alpha_I$.

The first is *simple thresholding*. One sets a threshold t, selects a subset $\hat{I}$ of terms 'highly correlated with S':

$$\hat{I} = \{i : |\langle S, \phi_i \rangle| > t\},$$

and then attempts to solve $\hat{S} = \Phi_{\hat{I}} \alpha_{\hat{I}}$. Statisticians have been using methods like this on noisy data for many decades; the approach is sometimes called "subset selection by preliminary significance testing from univariate regressions".

The second is *greedy subset selection*. One selects a subset iteratively, starting from $R_0 = S$ and $\ell = 0$ and proceeding stagewise, through stages $\ell = 0, 1, 2, \ldots$. At the 0-th stage, one identifies the best-fitting single term:

$$i_0 = \mathrm{argmax}_i |\langle R_0, \phi_i \rangle|,$$

and then, putting $\alpha_{i_0} = \langle R_0, \phi_{i_0} \rangle$, subtracts that term off

$$R_1 = R_0 - \alpha_{i_0} \phi_{i_0};$$

at stage 1 one behaves similarly, getting $i_1$and R_2, etc. In general,

$$i_\ell = \mathrm{argmax}_i |\langle R_{\ell-1}, \phi_i \rangle|,$$

and

$$R_\ell = S - P_{i_1,\ldots,i_\ell} S.$$

We stop as soon $R_\ell = 0$. Procedures of this form have been used routinely by statisticians since the 1960's under the name stepwise regression; the same procedure is called Orthogonal Matching Pursuit in signal analysis, and called greedy approximation in the approximation theory literature. For further discussion, see [9, 26, 27].

Under sufficiently strong conditions, both methods can work.

Theorem 3 *(Tropp [26]) Suppose that the dictionary Φ has coherence $M = \max_{i \neq j} |\langle \phi_i, \phi_j \rangle|$. Suppose that α_0 has $k \le M^{-1}/2$ nonzeros, and run the greedy algorithm with $S = \Phi\alpha_0$. The algorithm will stop after k stages having selected at each stage one of the terms corresponding to the k nonzero entries in α_0, at the end having precisely found the unique sparsest solution α_0.*

A parallel result can be given for thresholding.

Theorem 4 *Let $\eta \in (0,1)$. Suppose that α_0 has $k \le \eta M^{-1}/2$ nonzeros, and that the nonzero coefficients obey $|\alpha_0(i)| \ge \frac{\eta}{\sqrt{k}}\|\alpha\|_2$ (thus, they are all about the same size). Choose a threshold so that exactly k terms are selected. These k terms will be exactly the nonzeros in α_0 and solving $S = \Phi_{\hat{I}}\alpha_{\hat{I}}$ will recover the underlying optimal sparse solution α_0.*

Proof. We need to show that a certain threshold which selects exactly k terms selects only terms in I. Consider the preliminary threshold $t_0 = \frac{\eta}{2\sqrt{k}}\|\alpha_0\|_2$. We have, for $i \in I$,

$$\begin{aligned} |\langle \phi_i, S\rangle| &= |\alpha_i + \sum_{j\neq i} \alpha_0(j)\langle \phi_i, \phi_j\rangle| \\ &\ge |\alpha_i| - M\sum_{j\neq i} |\alpha_0(j)| \\ &> |\alpha_i| - M\sqrt{k}\|\alpha_0\|_2 \\ &\ge \|\alpha_0\|_2 \cdot (\eta/\sqrt{k} - M\sqrt{k}) \\ &\ge \|\alpha_0\|_2 \cdot \eta/2\sqrt{k} = t_0 \end{aligned}$$

Hence for $i \in I$, $|\langle \phi_i, S\rangle| > t_0$. On the other hand, for $j \notin I$

$$\begin{aligned} |\langle \phi_j, S\rangle| &= |\sum_i \alpha_0(i)\langle \phi_i, \phi_j\rangle| \\ &\le M\sqrt{k}\|\alpha_0\|_2 = t_0 \end{aligned}$$

Hence, for small enough $\delta > 0$, the threshold $t_\delta = t_0 + \delta$ selects exactly the terms in I. QED

7.2 Analysis of Subset Selection

The present article considers situations where the number of nonzeros is proportional to n. As it turns out, this is far beyond the range where previous general results about greedy algorithms and thresholding would work. Indeed, in this article's setting of a random dictionary Φ, we have (see Lemma 2.2) coherence $M \sim \sqrt{2\log(m)}/\sqrt{n}$. Theorems 3 and 4 therefore apply only for $|I| = o(\sqrt{n}) \ll \rho n$. In fact, it is not merely that the theorems don't apply; the nice behavior mentioned in Theorems 3 and 4 is absent in this more challenging setting.

Theorem 5 *Let n, m, A, and ρ^* be as above. On an event Ω_n having overwhelming probability, there is a vector S with unique sparsest representation using at most $k < \rho^* n$ nonzero elements, for which the following are true:*

- *The ℓ^1-minimal solution is also the optimally sparse solution.*
- *The thresholding algorithm can only find a solution using n nonzeros.*
- *The greedy algorithm makes a mistake in its first stage, selecting a term not appearing in the optimally sparse solution.*

Proof. The statement about ℓ^1 minimization is of course just a reprise of Theorem 1. The other two claims depend on the following.

Lemma 7.1 *Let n,m,A, and ρ^* be as in Theorem 1. Let $I = \{1, \ldots, k\}$, where $\rho^*/2n < k < \rho^* n$. There exists $C > 0$ so that, for each $\eta_1, \eta_2 > 0$, for all sufficiently large n, with overwhelming probability some $S \in Range(\Phi_I)$ has $\|S\|_2 = \sqrt{n}$, but*

$$|\langle S, \phi_i \rangle| < C, \qquad i \in I,$$

and

$$\min_{i \in I} |\langle S, \phi_i \rangle| < \eta_2.$$

The Lemma will be proved in the next subsection. Let's see what it says about thresholding. The construction of S guarantees that it is a random variable independent of ϕ_i, $i \notin I$. With R_i as introduced in (4.3), the coefficients $\langle S, \phi_i \rangle R_i$ $i \in I^c$, are iid with standard normal distribution; and by (4.6) these differ trivially from $\langle S, \phi_i \rangle$. This implies that for $i \in I^c$, the coefficients $\langle S, \phi_i \rangle$ are iid with a distribution that is nearly standard normal. In particular, for some $a = a(C) > 0$, with overwhelming probability for large n, we will have

$$\#\{i \in I^c : |\langle S, \phi_i \rangle| > C\} > a \cdot m,$$

and, if η_2 is the parameter used in the invocation of the Lemma above, with overwhelming probability for large n we will also have

$$\#\{i \in I^c : |\langle S, \phi_i \rangle| > \eta_2\} > n.$$

Hence, thresholding will actually select $a \cdot m$ terms not belonging to I *before* any term belonging to I. Also, if the threshold is set so that thresholding selects $< n$ terms, then some terms from I will *not* be among those terms (in particular, the terms where $|\langle \phi_i, S \rangle| < \eta_2$ for η_2 small).

With probability one, the points ϕ_i are in general position. Because of Lemma 2.1, we can only obtain a solution to the original equations if one of two things is true:

- We select *all* terms of I;
- We select n terms (and then it doesn't matter which ones).

If *any* terms from I are omitted by the selection $\hat{I}$, we cannot get a sparse representation. Since with overwhelming probability *some of* the k terms appearing in I are not among the n best terms for the inner product with the signal, thresholding does not give a solution until n terms are included, and there must be n nonzero coefficients in the solution obtained.

Now let's see what the Lemma says about greedy subset selection. Recall that the $\langle S, \phi_i \rangle R_i$ $i \in I^c$, are iid with standard normal distribution; and these differ trivially from $\langle S, \phi_i \rangle$. Combining this with standard extreme value theory for iid Gaussians, we conclude that for each $\delta > 0$, with overwhelming probability for large n,

$$\max_{i \in I^c} |\langle S, \phi_i \rangle| > (1 - \delta)\sqrt{\rho^* \log m}.$$

On the other hand, with overwhelming probability

$$\max_{i \in I} |\langle S, \phi_i \rangle| < C.$$

It follows that with overwhelming probability for large n, the first step of the greedy algorithm will select a term not belonging to I. QED

Not proved here, but strongly suspected, is that there exist S so that greedy subset selection cannot find any exact solution until it has been run for at least n stages.

7.3 Proof of Lemma 7.1

Let $V = Range(\Phi_I)$; pick any orthobasis (ψ_i) for V, and let $Z_1, \ldots Z_k$ be iid standard Gaussian $N(0,1)$. Set $v = \sum_i Z_j \psi_j$. Then for all $i \in I$, $\langle \phi_i, v \rangle \sim N(0,1)$.

Let now ξ_{ij} be the array defined by $\xi_{ij} = \langle \phi_i, \psi_j \rangle$. Note that the ξ_{ij} are independent of v and are approximately $N(0, \frac{1}{k})$. (More precisely, with R_i the random variable defined earlier at (4.3), $R_i \xi_{ij}$ is exactly $N(0, \frac{1}{k})$).

The proof of Lemma 5.2 shows that the Lemma actually has nothing to do with signs; it can be applied to any vector rather than some sign pattern vector σ. Make the temporary substitutions $n \leftrightarrow k$, $\sigma \leftrightarrow (Z_j)$, $\phi_i \leftrightarrow (\xi_{ij})$, and choose $\epsilon > 0$. Apply Lemma 5.2 to σ. Get a vector y obeying

$$|\langle \xi_i, y \rangle| \le 1, \qquad i = 1, \ldots, k. \tag{7.1}$$

Now define

$$S \equiv \frac{\sqrt{n}}{\|y\|_2} \cdot \sum_{j=1}^{k} y_j \psi_j.$$

Lemma 5.2 stipulated an event, E_n on which the algorithm delivers

$$\|y - \epsilon v\|_2 \le \epsilon \delta(\epsilon) \|v\|_2.$$

This event has probability exceeding $1 - \exp\{-\beta n\}$. On this event

$$\|y\|_2 \ge \epsilon(1 - \delta(\epsilon)) \|v\|_2.$$

Arguing as in (4.6), the event $F_n = \{\|v\|_2 \ge (1-\eta)\sqrt{k}\}$ has

$$P(F_n^c) \le 2 \exp\{-k\eta^2/2\}.$$

Hence on an event $E_n \cap F_n$,

$$\|y\|_2 \ge \epsilon(1 - \delta(\epsilon))\sqrt{k}.$$

We conclude using (7.1) that, for $i = 1, \ldots, k$,

$$|\langle \phi_i, S \rangle| = \frac{\sqrt{n}}{\|y\|_2} |\langle \xi_i, y \rangle| \le \frac{1}{\epsilon(1 - \delta(\epsilon))(1-\eta)\sqrt{\rho^*/2}} \cdot 1 \equiv C, \text{ say.}$$

This is the first claim of the lemma.

For the second claim, notice that this would be trivial if $\langle S, \phi_i \rangle$ were iid $N(0,1)$. This is not quite true, because of conditioning involved in the algorithm underlying Lemma 5.2. However, this conditioning only makes the indicated event even more likely than for an iid sequence. QED

8 Breakdown Point Heuristics

It can be argued that, in any particular application, we want to know whether we have equivalence for the one specific I that supports the specific α_0 of interest. Our proof suggests an accurate heuristic for predicting the sizes of subsets $|I|$ where this restricted notion of equivalence can hold.

Definition 8.1 *We say that* **local equivalence** *holds at a specific subset I if, for all vectors $S = \Phi\alpha_0$ with $supp(\alpha_0) \in I$, the minimum ℓ^1 solution to $S = \Phi\alpha$ has $\alpha_1 = \alpha_0$.*

It is clear that in the random dictionary Φ, the probability of the event "local equivalence holds for I" depends only on $|I|$.

Definition 8.2 *Let $E_{k,n} = \{$ local equivalence holds at $I = \{1, \ldots, k\}\}$. The events $E_{k,n}$ are decreasing with increasing k. The* **Local Equivalence Breakdown Point** *$LEBP_n$ is the smallest k for which event $E_{k,n}^c$ occurs.*

Clearly $EBP_n \le LEBP_n$.

8.1 The Heuristic

Let x be uniform on $\mathbf{S}^{n-1}$ and consider the random ℓ^1-minimization problem

$$(RP_1(n,m)) \qquad \min \|\alpha\|_1 \quad \Phi\alpha = x.$$

Here Φ is, as usual, iid uniform on $\mathbf{S}^{n-1}$. Define the random variable $V_{n,m} = val(RP_1(n,m))$. This is effectively the random variable at the heart of the event Ω_n^3 in the proof of Theorem 1. A direct application of the "Individual Sign-Pattern" Lemma shows there is a constant $\eta(A)$ so that, with overwhelming probability for large n,

$$V_{n,m} \ge \eta\sqrt{n}.$$

It follows that for the median

$$\nu(n,m) = med\{V_{n,m}\};$$

we have

$$\nu(n, An) \ge \eta\sqrt{n}.$$

There is numerical evidence, described below, that

$$\nu(n, An) \cdot \sqrt{\frac{\pi}{2n}} \to \nu_0(A), \qquad n \to \infty. \tag{8.1}$$

where ν_0 is a decreasing function of A.

Heuristic for Breakdown of Local Equivalence. *Let $\rho^+ = \rho^+(A)$ solve the equation*

$$\frac{\sqrt{\rho}}{(1-\sqrt{\rho})} = \nu_0(A - \rho).$$

Then we anticipate that

$$LEBP_n/n \to_P \rho^+, \quad n \to \infty.$$

8.2 Derivation

We use the notation of Section 2.1. We derive $LEBP_n/n \le \rho^+$. Consider the specific perturbation δ_I given by the eigenvector e_{min} of $G_I = \Phi_I^T\Phi_I$ with smallest eigenvalue. This eigenvector will be a random uniform point on $\mathbf{S}^{k-1}$ and so

$$\|\delta_I\|_1 = \sqrt{\frac{2}{\pi}}\sqrt{|I|}\|\delta_I\|_2(1 + o_p(1)).$$

It generates $v_I = \Phi_I \delta_I$ with

$$\|v_I\|_2 = \lambda_{min}^{1/2} \|\delta_I\|_2.$$

Letting $\rho = |I|/n$, we have [11]

$$\lambda_{min} = (1 - \rho^{1/2})^2 \cdot (1 + o_p(1)).$$

Now v_I is a random point on $\mathbf{S}^{n-1}$, independent of ϕ_i for $i \in I^c$. Considering the program

$$\min \|\delta_{I^c}\|_1 \text{ subject to } \Phi_{I^c} \delta_{I^c} = -v_I$$

we see that it has value $V_{n,m-|I|} \cdot \|v_I\|_2$. Now if $\rho > \rho^+$, then

$$\begin{aligned} \|\delta_I\|_1 &\sim \sqrt{\frac{2}{\pi}} \sqrt{|I|} \|\delta_I\|_2 = \sqrt{\frac{2}{\pi}} \sqrt{\rho n} \|\delta_I\|_2 \\ &> \sqrt{\frac{2}{\pi}} \sqrt{n} \cdot \nu_0(A - \rho) \cdot (1 - \rho^{1/2}) \|\delta_I\|_2 \\ &\sim \nu(n, m) \|v_I\|_2 \sim \|\delta_{I^c}\|_1. \end{aligned}$$

Hence, for a specific perturbation,

$$\|\delta_I\|_1 > \|\delta_{I^c}\|_1. \tag{8.2}$$

If we pick α_0 supported in I with a specific sign pattern $\text{sgn}(\alpha_0)(i)$, $i \in I$, this equation implies that a small perturbation in the direction of δ can reduce the ℓ^1 norm below that of α_0. Hence local equivalence breaks down.

With work we can also argue in the opposite direction, that this is approximately the worst perturbation, and it cannot cause breakdown unless $\rho > \rho^+$. Note this is all conditional on the limit relation (8.1), which seems an interesting topic for further work.

8.3 Empirical Evidence

Yaakov Tsaig of Stanford University performed several experiments showing the heuristic to be quite accurate. He studied the behavior of $\nu(n, An)/\sqrt{n}$ as a function of A, finding that $\nu_0(A) = A^{-.42}$ fits well over a range of problem sizes. Combined with our heuristic, we get that, for $A = 2$, ρ^+ is nearly .3 – i.e. local equivalence can hold up to about 30% of nonzeros. Tsaig performed actual simulations in which a vector α_0 was generated at random with specific $|I|$ and a test was made to see if the solution of (P_1) with $S = \Phi\alpha_0$ recovered α_0. It turns out that breakdown in local equivalence does indeed occur when $|I|$ is near $\rho^+ n$.

8.4 Geometric Interpretation

Let $B_{n,I}$ denote the $|I|$-dimensional convex body $\{\sum_{i \in I} \alpha(i)\phi_i : \|\alpha\|_1 \leq 1\}$. This is the image of a $|I|$-dimensional octahedron by Φ_I. Note that

$$B_{n,I} \subset Range(\Phi_I) \cap B_n,$$

however, the inclusion can be strict. Local Equivalence at I happens precisely when

$$B_{n,I} = Range(\Phi_I) \cap B_n.$$

This says that the faces of O_m associated to I all are mapped by Φ to faces of B_n.

Under our heuristic, for $|I| > (\rho^* + \epsilon)n$, $\epsilon > 0$, each event $\{B_{n,I} = Range(\Phi_I) \cap B_n\}$ typically fails. This implies that most fixed sections of B_n by subspaces $Range(\Phi_I)$ have a different topology than that of the octahedron $B_{n,I}$.

9 Stability

Skeptics may object that our discussion of sparse solution to underdetermined systems is irrelevant because the whole concept is not stable. Actually, the concept is stable, as an implicit result of Lemma 3.1. There we showed that, with overwhelming probability for large n, all singular values of every submatrix Φ_I with $|I| < \rho n$ exceed $\eta_1(\rho, A)$. Now invoke

Theorem 6 *(Donoho, Elad, Temlyakov [9]) Let Φ be given, and set*

$$\eta(\rho, \Phi) = \min\{\lambda_{min}^{1/2}(\Phi_I^T \Phi_I) : |I| < \rho n\}.$$

Suppose we are given the vector $Y = \Phi\alpha_0 + z$, $\|z\|_2 \le \epsilon$, $\|\alpha_0\|_0 \le \rho n/2$. Consider the optimization problem

$$(P_{0,\epsilon}) \qquad \min \|\alpha\|_0 \text{ subject to } \|Y - \Phi\alpha\|_2 \le \epsilon.$$

and let $\hat{\alpha}_{0,\epsilon}$ denote any solution. Then:

- $\|\hat{\alpha}_{0,\epsilon}\|_0 \le \|\alpha_0\|_0 \le \rho n/2$; *and*
- $\|\hat{\alpha}_{0,\epsilon} - \alpha_0\|_2 \le 2\epsilon/\eta$, *where* $\eta = \eta(\rho, \Phi)$.

Applying Lemma 3.1, we see that the problem of obtaining a sparse approximate solution to noisy data is a stable problem: if the noiseless data have a solution with at most $\rho n/2$ nonzeros, then an error of size $\le \epsilon$ in measurements can lead to a reconstruction error of size $\le 2\epsilon/\eta_1(\rho, A)$. We stress that we make no claim here about stability of the ℓ^1 reconstruction; only that stability *by some method* is in principle possible. Detailed investigation of stability is being pursued separately.

References

[1] H.H. Bauschke and J.M. Borwein (1996) On projection algorithms for solving convex feasibility problems, *SIAM Review* **38**(3), pp. 367-426.

[2] E.J. Candès, J. Romberg and T. Tao. (2004) Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information. Manuscript.

[3] Chen, S. , Donoho, D.L., and Saunders, M.A. (1999) Atomic Decomposition by Basis Pursuit. *SIAM J. Sci Comp.*, **20**, 1, 33-61.

[4] R.R. Coifman, Y. Meyer, S. Quake, and M.V. Wickerhauser (1990) Signal Processing and Compression with Wavelet Packets. in *Wavelets and Their Applications*, J.S. Byrnes, J. L. Byrnes, K. A. Hargreaves and K. Berry, eds. 1994,

[5] R.R. Coifman and M.V. Wickerhauser. Entropy Based Algorithms for Best Basis Selection. *IEEE Transactions on Information Theory*, **32**:712–718.

[6] K.R. Davidson and S.J. Szarek (2001) Local Operator Theory, Random Matrices and Banach Spaces. *Handbook of the Geometry of Banach Spaces, Vol. 1* W.B. Johnson and J. Lindenstrauss, eds. Elsevier.

[7] Donoho, D.L. and Huo, Xiaoming (2001) Uncertainty Principles and Ideal Atomic Decomposition. *IEEE Trans. Info. Thry.* **47** (no. 7), Nov. 2001, pp. 2845-62.

[8] Donoho, D.L. and Elad, Michael (2002) Optimally Sparse Representation from Overcomplete Dictionaries via ℓ^1 norm minimization. *Proc. Natl. Acad. Sci. USA* March 4, 2003 **100** 5, 2197-2002.

[9] Donoho, D., Elad, M., and Temlyakov, V. (2004) Stable Recovery of Sparse Overcomplete Representations in the Presence of Noise. Submitted. URL: http://www-stat.stanford.edu/~donoho/Reports/2004.

[10] A. Dvoretsky (1961) Some results on convex bodies and Banach Spaces. *Proc. Symp. on Linear Spaces.* Jerusalem, 123-160.

[11] A. Edelman, Eigenvalues and condition numbers of random matrices, SIAM J. Matrix Anal. Appl. 9 (1988), 543-560

[12] M. Elad and A.M. Bruckstein (2002) A generalized uncertainty principle and sparse representations in pairs of bases. *IEEE Trans. Info. Thry.* **49** 2558-2567.

[13] Noureddine El Karoui (2004) *New Results About Random Covariance Matrices and Statistical Applications.* Ph.D. Thesis, Stanford University.

[14] J.J. Fuchs (2002) On sparse representation in arbitrary redundant bases. Manuscript.

[15] T. Figiel, J. Lindenstrauss and V.D. Milman (1977) The dimension of almost-spherical sections of convex bodies. *Acta Math.* **139** 53-94.

[16] R. Gribonval and M. Nielsen. Sparse Representations in Unions of Bases. To appear *IEEE Trans Info Thry.*

[17] W.B. Johnson and G. Schechtman (1982) Embedding ℓ_p^m into ℓ_1^n. *Acta Math.* **149**, 71-85.

[18] Michel Ledoux. *The Concentration of Measure Phenomenon.* Mathematical Surveys and Monographs **89**. American Mathematical Society 2001.

[19] S. Mallat, Z. Zhang, (1993). "Matching Pursuits with Time-Frequency Dictionaries," IEEE Transactions on Signal Processing, 41(12):3397–3415.

[20] V.D. Milman and G. Schechtman (1986) *Asymptotic Theory of Finite-Dimensional Normed Spaces.* Lect. Notes Math. **1200**, Springer.

[21] B.K. Natarajan (1995) Sparse Approximate Solutions to Linear Systems. *SIAM J. Comput.* **24**: 227-234.

[22] G. Pisier (1989) *The Volume of Convex Bodies and Banach Space Geometry.* Cambridge University Press.

[23] D. Pollard (1989) *Empirical Processes: Theory and Applications.* NSF - CBMS Regional Conference Series in Probability and Statistics, Volume 2, IMS.

[24] Gideon Schechtman (1981) Random Embeddings of Euclidean Spaces in Sequence Spaces. *Israel Journal of Mathematics* **40**, 187-192.

[25] Szarek, S.J. (1990) Spaces with large distances to ℓ_∞^n and random matrices. *Amer. Jour. Math.* **112**, 819-842.

[26] J.A. Tropp (2003) Greed is Good: Algorithmic Results for Sparse Approximation To appear, *IEEE Trans Info. Thry.*

[27] J.A. Tropp (2004) Just Relax: Convex programming methods for Subset Sleection and Sparse Approximation. Manuscript.