COMPRESSED SENSING - PROBABILISTIC ANALYSIS OF A NULL-SPACE CHARACTERIZATION

Mihailo Stojnic, Weiyu Xu, and Babak Hassibi

California Institute of Technology, Pasadena, CA
e-mail: mihailo,weiyu,hassibi@systems.caltech.edu

ABSTRACT

It is well known that compressed sensing problems reduce to solving large under-determined systems of equations. To assure that the problem is well defined, i.e., that the solution is unique the vector of unknowns is of course assumed to be sparse. Nonetheless, even when the solution is unique, finding it in general may be computationally difficult. However, starting with the seminal work of [2], it has been shown that linear programming techniques, obtained from an l_1-norm relaxation of the original non-convex problem, can provably find the unknown vector in certain instances. In particular, using a certain restricted isometry property, [2] shows that for measurement matrices chosen from a random Gaussian ensemble, l_1 optimization can find the correct solution with overwhelming probability even when the number of non-zero entries of the unknown vector is proportional to the number of measurements (and the total number of unknowns). The subsequent paper [1] uses results on neighborly polytopes from [5] to give a "sharp" bound on what this proportionality should be in the Gaussian case. In the current paper, we observe that what matters is not so much the distribution from which the entries of the measurement matrix A are drawn, but rather the statistics of the null-space of A. Using this observation, we provide an alternative proof of the main result of [2] by analyzing matrices whose null-space is isotropic (of which i.i.d. Gaussian ensembles are a special case).

Index Terms: compressed sensing, l_1-optimization

1. INTRODUCTION

In this paper we are interested in compressed sensing problems. As is well known these problems are very easy to pose and very difficult to solve. Namely, we would like to find $\mathbf{x}$ such that

$$A\mathbf{x} = \mathbf{y} \tag{1}$$

where A is an $m \times n$ measurement matrix, $\mathbf{y}$ is $m \times 1$ measurement vector, and $\mathbf{x}$ is $n \times 1$ unknown k-sparse vector. In the rest of the paper we will assume that the number of the measurements is $m = \alpha n$ and the number of the non-zero components of $\mathbf{x}$ is $k = \beta n$, where $0 < \beta < 1$ and $0 < \alpha < 1$ are constants independent of n. This problem setup is quite natural and arises in many practical applications (see e.g. [15, 16] and references therein).

A particular way of solving (1) which recently generated a large amount of research is called l_1-optimization [2]. It proposes

solving the following problem

$$
\begin{aligned}
\min \quad & \|\mathbf{x}\|_1 \\
\text{subject to} \quad & A\mathbf{x} = \mathbf{y}.
\end{aligned}
\tag{2}
$$

Quite remarkably in [2] the authors were able to show that if the number of the measurements $m = \alpha n$, α, and n are given, the matrix A is given and satisfies a special property called the restricted isometry property (RIP), then any unknown vector $\mathbf{x}$ with no more than $k = \beta n$ (where β is an absolute constant which of course is a function of α, independent of n, and explicitly calculated in [2]) non-zero elements can be recovered by solving (2). As expected, this assumes that $\mathbf{y}$ was in fact generated by that $\mathbf{x}$ and given to us (more on the case when the available measurements are noisy versions of $\mathbf{y}$ interested reader can find in e.g. [13, 14]).

As can be immediately seen, the previous result heavily relies on the assumption that the measurement matrix A satisfies the RIP condition. What is indeed remarkable about [2] is the fact that for several specific classes of matrices the RIP holds with overwhelming probability. It happens that if the components of A are i.i.d. zero mean Gaussian or Bernoulli the RIP condition holds with overwhelming probability [2, 3, 4]. However, it should be noted that the RIP is only a sufficient condition for l_1-optimization to produce a solution of (1).

Instead of characterizing the $m \times n$ matrix A through the RIP condition, in [1] the authors assume that A constitutes a k-neighborly poly-tope. It turns out (as shown in [1]) that this characterization of the matrix A is in fact a necessary and sufficient condition for (2) to produce the solution of (1). Furthermore, using the results of [5], it can be shown that if the matrix A has i.i.d. zero-mean Gaussian entries with overwhelming probability it also constitutes a k-neighborly poly-tope. Of course, the precise relation between m and k in order for this to happen is characterized in [1] as well. It should also be noted that for a given value m i.e. for a given value of the constant α, the value of the constant β is significantly better in [1] than in [2]. Furthermore, the values of constants β obtained for different values of α in [1] approach the ones obtained by simulation as $n \longrightarrow \infty$.

In this paper we will make use of another characterization that guarantees l_1 optimization works for the matrix A. This characterization will be equivalent to the neighborly polytope characterization from [1] since it also constitutes both necessary and sufficient conditions which the matrix A should satisfy in order that (2) be the solution of (1). However, we will provide much simpler analysis that shows that a non-zero β is achievable and further provides a lower bound for it.

2. NULL-SPACE CHARACTERIZATION

In this section we introduce another useful characterization of the matrix A. The following theorem can be proved (see [10, 11, 12] for variation of this result).

Theorem 1 *Assume that an $m \times n$ measurement matrix A is given. Further, assume that $\mathbf{y} = A\mathbf{x}$ and $\mathbf{x}$ has at most k non-zero elements and that $\mathbf{w}$ is an $n \times 1$ vector. Let K be any subset of $\{1, 2, \ldots, n\}$ such that $|K| = k$ and let K_i denote the i-th element of K. Further, let $\bar{K} = \{1, 2, \ldots, n\}/K$. Then (2) will produce the solution of (1) if and only if*

$$(\forall \mathbf{w} \in \mathbf{R}^n | A\mathbf{w} = 0) \quad and \quad \forall K, \quad \sum_{i=1}^{k} |\mathbf{w}_{K_i}| \leq \sum_{i=1}^{n-k} |\mathbf{w}_{\bar{K}_i}|. \tag{3}$$

Proof 1 *Omitted [10, 11, 12].*

Remark: Of course, we need not to check (3) for all subsets K; checking the subset with the k largest (in absolute value) elements of $\mathbf{w}$ is sufficient. However, Theorem 1 will be more convenient for our subsequent analysis.

Now, let Z be a basis of the null space of A, i.e. let Z be a matrix such that $AZ = 0$. Clearly, Z is an $n \times (n - m)$ matrix. Furthermore, any $n \times 1$ vector $\mathbf{w}$ from the null-space of A can be represented as $Z\mathbf{v}$ where $\mathbf{v} \in \mathbf{R}^{n-m}$. Then the condition from the Theorem 1 can be transformed to

$$\sum_{i=1}^{k} |Z_{K_i}\mathbf{v}| \leq \sum_{i=1}^{n-k} |Z_{\bar{K}_i}\mathbf{v}| \quad \forall \mathbf{v} \in \mathbf{R}^{n-m}, \forall K \quad \text{s. t.} \quad |K| = k \tag{4}$$

where Z_i is the i-th row of the matrix Z. To facilitate writing let $I_{\mathbf{v}}$ denote the event $\sum_{i=1}^{k} |Z_{K_i}\mathbf{v}| \leq \sum_{i=1}^{n-k} |Z_{\bar{K}_i}\mathbf{v}|$. In the following section we will for a given value $\alpha = \frac{m}{n}$, determine the value of $\beta = \frac{k}{n}$ such that (4) is satisfied with overwhelming probability.

The standard results on compressed sensing assume that the matrix A has i.i.d. $\mathcal{N}(0,1)$ entries. In this case, the following lemma gives a characterization of the resulting null-space.

Lemma 1 *Let $A \in R^{m \times n}$ be a random matrix with i.i.d. $\mathcal{N}(0,1)$ entries. Then the following statements hold:*

- *The distribution of A is left-rotationally invariant, $P_A(A) = P_A(A\Theta)$, $\Theta\Theta^* = \Theta^*\Theta = I$*

- *The distribution of Z, any basis of the null-space of A is right-rotationally invariant. $P_Z(Z) = P_Z(\Theta^* Z)$, $\Theta\Theta^* = \Theta^*\Theta = I$*

- *It is always possible to choose a basis for the null-space such that $Z \in R^{n \times (n-m)}$ has i.i.d. $\mathcal{N}(0,1)$ entries.*

In view of Theorem 1 and Lemma 1 what matters is that the null-space of A be rotationally invariant. For any such A, the sharp bounds of ([1]), for example, apply. In this paper, we shall analyze the null-space directly. [It should be noted that we will present the result for the case when the matrix Z has real Gaussian entries; however it is straightforward to extend it to the case when the matrix Z is comprised of complex Gaussian entries.]

3. PROBABILISTIC ANALYSIS OF THE NULL-SPACE CHARACTERIZATION

In this section we probabilistically analyze the validity of (4). Before proceeding further, let us recall what exactly is the problem that we will solve in this section.

Assume that we are given an $n \times (n - m)$ matrix Z. Let Z_i be the i-th row of Z and let Z_{ij} be the i, j-th element of Z. Further, let Z_{ij} be i.i.d. zero-mean unit-variance Gaussian random variables. Let $\mathbf{v} \in \mathbf{R}^{n-m}$ be any real vector of length $(n - m)$. Let further $\alpha = \frac{m}{n}$ be a given constant independent of n. Then we will find $\beta = \frac{k}{n}$ such that

$$\lim_{n \to \infty} P(I_{\mathbf{v}} \; \forall \mathbf{v} \in \mathbf{R}^{n-m}, \forall K \subset \{1, 2, \ldots, n\}, |K| = k) = 1. \tag{5}$$

Proving (5) will of course be enough to prove that for all random matrices A which have isotropically distributed null-space, (2) with overwhelming probability solves (1). In order to prove (5) we will actually prove that

$$\lim_{n \to \infty} P_f = 0, \tag{6}$$

$P_f = P(\exists \mathbf{v} \in \mathbf{R}^{n-m}, \exists K \subset \{1, 2, \ldots, n\}, |K| = k \text{ s. t. } \bar{I}_{\mathbf{v}})$ and $\bar{I}_{\mathbf{v}}$ denotes the complement of $I_{\mathbf{v}}$, i.e. it denotes the event $\sum_{i=1}^{k} |Z_{K_i}\mathbf{v}| < \sum_{i=1}^{n-k} |Z_{\bar{K}_i}\mathbf{v}|$. Now, using the union bound we can write

$$P_f \leq \sum_{l=1}^{\binom{n}{k}} P(\exists \mathbf{v} \in \mathbf{R}^{n-m} \quad \text{s. t.} \quad \sum_{i=1}^{k} |Z_{K_i^{(l)}}\mathbf{v}| \geq \sum_{i=1}^{n-k} |Z_{\bar{K}_i^{(l)}}\mathbf{v}|) \tag{7}$$

where $K^{(l)}$ is a subset of $\{1, 2, \ldots, n\}$ and $|K^{(l)}| = k$. Clearly the number of these subsets is $\binom{n}{k}$ and hence the summation in (7) goes from 1 to $\binom{n}{k}$. Since the probability in (7) is insensitive to scaling of $\mathbf{v}$ by a constant we can restrict $\mathbf{v}$ to lie on the unit sphere (in l_2-norm). Furthermore, since the elements of the matrix Z are i.i.d. all $\binom{n}{k}$ terms in the first summation on the right hand side of (7) will then be equal. Therefore we can further write

$$P_f \leq \binom{n}{k} P(\exists \mathbf{v} \in \mathbf{R}^{n-m}, ||\mathbf{v}||_2 = 1 \text{ s. t. } \sum_{i=1}^{k} |Z_i\mathbf{v}| \geq \sum_{i=k+1}^{n} Z_i\mathbf{v}|). \tag{8}$$

The main difficulty in computing the probability on the right hand side of (8) is in the fact that the vector $\mathbf{v}$ (i.e. its components) is continuous. Our approach will be based on the discrete covering of the unit sphere. In order to do that we will use small spheres of radius ϵ. It can be shown [6, 7, 4] that $\epsilon^{-(n-m)}$ spheres of radius ϵ is enough to cover the surface of the $(n-m)$-dimensional unit sphere. Let the coordinates of the centers of these $\epsilon^{-(n-m)}$ small spheres be the vectors $\mathbf{z}_t, t = 1, \ldots, \epsilon^{-(n-m)}$. Clearly, $1 - \epsilon \leq ||\mathbf{z}_t||_2 \leq 1 + \epsilon$. Further, let $S_t, t = 1, \ldots, \epsilon^{-(n-m)}$ denote the $\epsilon^{-(n-m)}$ small spheres. Since they cover the unit sphere we have

$$(\mathbf{v} \in \mathbf{R}^{n-m} \;\; | \;\; ||\mathbf{v}||_2 = 1) \subset \bigcup_{t=1}^{\epsilon^{-(n-m)}} S_t.$$

Then using the union bound over the spheres we can further write

$$P_f \leq \binom{n}{k} P(\exists \mathbf{v} \in \bigcup_{t=1}^{\epsilon^{-(n-m)}} S_t \quad \text{s. t.} \quad \sum_{i=1}^{k} |Z_i\mathbf{v}| \geq \sum_{i=k+1}^{n} |Z_i\mathbf{v}|)$$

$$\leq \frac{\binom{n}{k}}{\epsilon^{(n-m)}} \max_t P(\exists \mathbf{v} \in S_t \quad \text{s. t.} \quad \sum_{i=1}^{k} |Z_i\mathbf{v}| \geq \sum_{i=k+1}^{n} |Z_i\mathbf{v}|). \tag{9}$$

Every vector $\mathbf{v}$ covered by the sphere S_t can be represented as $\mathbf{v} = \mathbf{z}_t + \mathbf{e}$ where $||\mathbf{e}||_2 \leq \epsilon$. Then we have

$$\max_t P(\exists \mathbf{v} \in S_t \quad \text{s. t.} \quad \sum_{i=1}^k |Z_i \mathbf{v}| \geq \sum_{i=k+1}^n |Z_i \mathbf{v}|)$$

$$= \max_t P(\exists \mathbf{e}, ||\mathbf{e}||_2 \leq \epsilon \text{ s. t.} \sum_{i=1}^k |Z_i(\mathbf{z}_t+\mathbf{e})| \geq \sum_{i=k+1}^n |Z_i(\mathbf{z}_t+\mathbf{e})|).$$

$$(10)$$

Given the symmetry of the problem it should be noted that without loss of generality we can assume $\mathbf{z}_t = [||\mathbf{z}_t||_2, 0, 0, \ldots, 0]$. Further, using the results from [9] we have that N^{n-m} points can be located on the sphere of radius $c\epsilon$ centered at $\mathbf{z}_t$ such that S_t (whose radius is ϵ) is inside a poly-tope determined by them and

$$c \leq \begin{cases} \frac{1}{(1-\ln(N))\sqrt{2\ln(N)-\frac{\ln(n-m)}{n-m}}} & \text{if } N < \sqrt{2} \\ \frac{1}{1-(1+\frac{1}{N^2})\frac{1}{2N^2}} & \text{otherwise.} \end{cases}$$

$$(11)$$

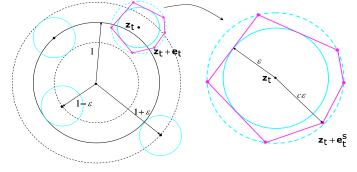To get a feeling what values N and c can take we refer to [8]

Fig. 1. Covering of the unit sphere

where it was stated that 3^{n-m} points can be located on the sphere of radius $\sqrt{\frac{9}{8}}\epsilon$ centered at $\mathbf{z}_t$ such that S_t is inside a poly-tope determined by them.

Let us call the poly-tope determined by N^{n-m} points P_t. Let $\mathbf{e}_t^s, s = 1, 2, \ldots, N^{n-m}$ be its N^{n-m} corner points. Since $|Z_i(\mathbf{z}_t + \mathbf{e})| \leq |Z_i \mathbf{z}_t| + |Z_i \mathbf{e}|$, $|Z_i \mathbf{z}_t| - |Z_i \mathbf{e}| \leq |Z_i(\mathbf{z}_t + \mathbf{e})|$, and $S_t \subset P_t$ we have

$$\max_t P(\exists \mathbf{e}, ||\mathbf{e}||_2 \leq \epsilon \quad \text{s. t.} \quad \frac{\sum_{i=k+1}^n |Z_i(\mathbf{z}_t+\mathbf{e})|}{\sum_{i=1}^k |Z_i(\mathbf{z}_t+\mathbf{e})|} \leq 1)$$

$$\leq \max_t P(\exists \mathbf{e}, (\mathbf{z}_t+\mathbf{e}) \in P_t \text{ s. t.} \frac{\sum_{i=k+1}^n(|Z_i\mathbf{z}_t|-|Z_i\mathbf{e}|)}{\sum_{i=1}^k(|Z_i\mathbf{z}_t|+|Z_i\mathbf{e}|)} \leq 1)$$

$$\leq \max_t P(\frac{\sum_{i=k+1}^n |Z_i\mathbf{z}_t| - \max_s \sum_{i=k+1}^n |Z_i\mathbf{e}_t^s|}{\sum_{i=1}^k |Z_i\mathbf{z}_t| + \max_s \sum_{i=1}^k |Z_i\mathbf{e}_t^s|} \leq 1).$$

$$(12)$$

where the second inequality follows from the property that the maximum of a convex function over a poly-tope is achieved at its corner points. Connecting (9), (10), and (12) we obtain

$$P_f \leq \frac{\binom{n}{k}}{\epsilon^{(n-m)}} \max_t P(\frac{\sum_{i=k+1}^n |Z_i\mathbf{z}_t| - \max_s \sum_{i=k+1}^n |Z_i\mathbf{e}_t^s|)}{\sum_{i=1}^k |Z_i\mathbf{z}_t| + \max_s \sum_{i=1}^k |Z_i\mathbf{e}_t^s|} \leq 1).$$

$$(13)$$

Using the union bound over s we further have

$$\max_t P(\frac{\sum_{i=k+1}^n |Z_i\mathbf{z}_t| - \max_s \sum_{i=k+1}^n |Z_i\mathbf{e}_t^s|}{\sum_{i=1}^k |Z_i\mathbf{z}_t| + \max_s \sum_{i=1}^k |Z_i\mathbf{e}_t^s|} \leq 1)$$

$$\leq \max_t \sum_{s'=1}^{N^{n-m}} P(\frac{\sum_{i=k+1}^n |Z_i\mathbf{z}_t| - \max_s \sum_{i=k+1}^n |Z_i\mathbf{e}_t^s|}{\sum_{i=1}^k |Z_i\mathbf{z}_t| + \sum_{i=1}^k |Z_i\mathbf{e}_t^{s'}|} \leq 1)$$

$$\leq \max_t \sum_{s'=1}^{N^{n-m}} \sum_{s''=1}^{N^{n-m}} P(\frac{\sum_{i=k+1}^n |Z_i\mathbf{z}_t| - \sum_{i=k+1}^n |Z_i\mathbf{e}_t^{s''}|}{\sum_{i=1}^k |Z_i\mathbf{z}_t| + \sum_{i=1}^k |Z_i\mathbf{e}_t^{s'}|} \leq 1).$$

$$(14)$$

Given that only the first component of $\mathbf{z}_t$ is not equal to zero we can write

$$\max_t \sum_{s'=1}^{N^{n-m}} \sum_{s''=1}^{N^{n-m}} P(\frac{\sum_{i=k+1}^n |Z_i\mathbf{z}_t| - \sum_{i=k+1}^n |Z_i\mathbf{e}_t^{s''}|}{\sum_{i=1}^k |Z_i\mathbf{z}_t| + \sum_{i=1}^k |Z_i\mathbf{e}_t^{s'}|} \leq 1) \leq N^{2(n-m)}$$

$$\times \max_{t,s',s''} P\left(\frac{\sum_{i=k+1}^n\left(|Z_{i1}(||\mathbf{z}_t||_2 - |(\mathbf{e}_t^{s''})_1|)| - |\sum_{j=2}^{n-m} Z_{ij}(\mathbf{e}_t^{s''})_j|\right)}{\sum_{i=1}^k\left(|Z_{i1}(||\mathbf{z}_t||_2 + |(\mathbf{e}_t^{s'})_1|)| + |\sum_{j=2}^{n-m} Z_{ij}(\mathbf{e}_t^{s'})_j|\right)} \leq 1\right)$$

$$(15)$$

where $(\mathbf{e}_t^{s'})_j$ and $(\mathbf{e}_t^{s''})_j$ denote j-th components of $\mathbf{e}_t^{s'}$ and $\mathbf{e}_t^{s''}$, respectively. Let $A_i = Z_{i1}(||\mathbf{z}_t||_2+|(\mathbf{e}_t^{s'})_1|)$, $B_i = \sum_{j=2}^{n-m} Z_{ij}(\mathbf{e}_t^{s'})_j$, $C_i = Z_{i1}(||\mathbf{z}_t||_2 - |(\mathbf{e}_t^{s''})_1|)$, and $D_i = \sum_{j=2}^{n-m} Z_{ij}(\mathbf{e}_t^{s''})_j$. Clearly, A_i, B_i, C_i, D_i are independent zero-mean Gaussian random variables. Furthermore it holds

$$\text{var}(A_i) = (||\mathbf{z}_t||_2 + |(\mathbf{e}_t^{s'})_1|)^2, \text{var}(B_i) = ||\mathbf{e}_t^{s'}||_2^2 - |(\mathbf{e}_t^{s'})_1|^2$$

$$\text{var}(C_i) = (||\mathbf{z}_t||_2 - |(\mathbf{e}_t^{s''})_1|)^2, \text{var}(D_i) = ||\mathbf{e}_t^{s''}||_2^2 - |(\mathbf{e}_t^{s''})_1|^2.$$

Then we can rewrite (15) as

$$\max_t \sum_{s'=1}^{N^{n-m}} \sum_{s''=1}^{N^{n-m}} P(\frac{\sum_{i=k+1}^n |Z_i\mathbf{z}_t| - \sum_{i=k+1}^n |Z_i\mathbf{e}_t^{s''}|}{\sum_{i=1}^k |Z_i\mathbf{z}_t| + \sum_{i=1}^k |Z_i\mathbf{e}_t^{s'}|} \leq 1) \leq$$

$$(N^2)^{n-m} \max_{t,s',s''} P\left(\sum_{i=1}^k (|A_i| + |B_i|) \geq \sum_{i=k+1}^n (|C_i| - |D_i|)\right).$$

$$(16)$$

Let E_i, F_i, G_i, H_i be independent zero-mean Gaussian random variables such that

$$\text{var}(E_i) = (1 + \epsilon + ||\mathbf{e}_t^{s'}||)^2, \text{var}(F_i) = ||\mathbf{e}_t^{s'}||_2^2$$

$$\text{var}(G_i) = (1 - \epsilon - ||\mathbf{e}_t^{s''}||_2)^2, \text{var}(H_i) = ||\mathbf{e}_t^{s''}||_2^2.$$

Since $\text{var}(E_i) \geq \text{var}(A_i)$, $\text{var}(F_i) \geq \text{var}(B_i)$, $\text{var}(G_i) \leq \text{var}(C_i)$, and $\text{var}(H_i) \geq \text{var}(D_i)$ we have from (16)

$$(N^2)^{n-m} \max_{t,s',s''} P\left(\sum_{i=1}^k (|A_i| + |B_i|) \geq \sum_{i=k+1}^n (|C_i| - |D_i|)\right)$$

$$\leq (N^2)^{n-m} \max_{t,s',s''} P\left(\sum_{i=1}^k (|E_i| + |F_i|) \geq \sum_{i=k+1}^n (|G_i| - |H_i|)\right).$$

$$(17)$$

Since $\|\mathbf{e}_t^{s'}\|_2$, $\|\mathbf{e}_t^{s''}\|_2$, and (N^2) don't depend on t, s' and s'', the outer maximization can be omitted. Furthermore, $\|\mathbf{e}_t^{s'}\|_2 = \|\mathbf{e}_t^{s''}\|_2 = c\epsilon$. Using the Chernoff bound we further have

$$(N^2)^{n-m} P\left(\sum_{i=1}^{k} (|E_i| + |F_i|) \geq \sum_{i=k+1}^{n} (|G_i| - |H_i|) \right)$$
$$\leq (N^2)^{n-m} (Ee^{\mu|E_1|})^k (Ee^{\mu|F_1|})^k (Ee^{-\mu|G_1|})^{n-k} (Ee^{\mu|H_1|})^{n-k}. \tag{18}$$

where μ is a positive constant. Connecting (13)-(18) we have

$$P_f \leq \binom{n}{k} \left(\frac{N^2}{\epsilon}\right)^{n-m} (Ee^{\mu|E_1|} Ee^{\mu|F_1|})^k \left(\frac{Ee^{-\mu|G_1|}}{(Ee^{\mu|H_1|})^{-1}} \right)^{n-k}. \tag{19}$$

After setting $k = \beta n$, $m = \alpha n$, and computing the expectations in (19) we finally obtain

$$\lim_{n\to\infty} P_f \leq \lim_{n\to\infty} \xi^n \tag{20}$$

where

$$\xi = \frac{(N^2/\epsilon)^{(1-\alpha)}}{e^{H(\beta) - \mu^2(c\epsilon)^2/2}} \left(e^{\mu^2(1+\epsilon+c\epsilon)^2/2} \left(\mathrm{erf}\left(\mu \frac{1+\epsilon+c\epsilon}{\sqrt{2}} \right) + 1 \right) \right)^{\beta}$$
$$\times \left(\mathrm{erf}\left(\frac{\mu c\epsilon}{\sqrt{2}} \right) + 1 \right) \left(e^{\mu^2(1-\epsilon-c\epsilon)^2/2} \mathrm{erfc}\left(\mu \frac{1-\epsilon-c\epsilon}{\sqrt{2}} \right) \right)^{(1-\beta)}. \tag{21}$$

Now, for a given α we can determine numerically the largest β so that there are $\mu \geq 0$, N, c, and $0 \leq \epsilon < 1$ such that $\xi < 1$ and $\lim_{n\to\infty} P_f = 0$ in (20). In fact using the well-known approximations of the erf and erfc functions it can be explicitly shown that this can always be done. Since this is fairly obvious we omit the derivation here. As a numerical example we say that for $\mu = 18$, $\epsilon = 0.011$, $N = \sqrt{\sqrt{2}}$, $c = 2.045$, $\alpha = 0.5$, $\beta = 0.0016$, we have $\xi = 0.99$. This means that if we have $\frac{n}{2}$ measurements we can recover by l_1-optimization vectors $\mathbf{x}$ provided that they have less than $0.0016n$ nonzero elements. Of course, the constant 0.0016 is not as good as the one from [1] $0.044708n$. However, it still shows that it is possible to recover signals $\mathbf{x}$ whose number of nonzero elements grows linearly with n.

We summarize the previous results in the following theorem.

Theorem 2 *Assume that we use (2) to solve (1). Further assume that the matrix A in (1) has isotropically distributed null-space and that the number of rows of the matrix A is $m = \alpha n$. Then it is possible to recover vector $\mathbf{x}$ in (1) provided that $\mathbf{x}$ has less than βn nonzero components. α and β are absolute constants and independent of n. Furthermore, for any given α the explicit value of β can be numerically determined as a maximal value of β so that the right hand side of (21) is less than 1.*

Proof 2 *Follows from the previous discussion.*

4. CONCLUSION

We analyzed a null-space characterization of the necessary and sufficient conditions for the success of the l_1-norm optimization in compressed sensing. Our analysis provided a somewhat new technique in proving the optimality of the l_1-norm optimization for measurement matrices with isotropically distributed null-space.

5. REFERENCES

[1] David Donoho and Jared Tanner, "Neighborliness of randomly-projected simplices in high dimensions", *Proc. National Academy of Sciences*, 102(27), pp. 9452-9457, 2005.

[2] Emmanuel Candes and Terence Tao, "Decoding by linear programming", *IEEE Trans. on Information Theory*, 51(12), pp. 4203 - 4215, December 2005.

[3] Richard Baraniuk, Mark Davenport, Ronald DeVore, and Michael Wakin, "A simple proof of the restricted isometry property for random matrices", *To appear in Constructive Approximation*, available online at http://www.dsp.ece.rice.edu/cs/.

[4] Mark Rudelson and Roman Vershynin, "Geometric approach to error correcting codes and reconstruction of signals", *International Mathematical Research Notices*, 64, pp. 4019 - 4041, 2005.

[5] A.M. Vershik and P.V. Sporyshev, "Asymptotic Behavior of the Number of Faces of Random Polyhedra and the Neighborliness Problem", *Selecta Mathematica Sovietica* vol. 11, No. 2 (1992).

[6] A.D. Wyner, "Random packings and coverings of the unit N-sphere", *Bell systems technical journal*, vol. 46, 2111-2118, 1967.

[7] I. Dumer, M.S. Pinsker, and V.V. Prelov, "On the thinest coverings of the spheres and ellipsoids", Information Transfer and Combinatorics, LNCS 4123, pp.883-910, 2006.

[8] I. Barany and N. Simanyi, "A note on the size of the largest ball inside a convex polytope", *Periodica Mathematica Hungarica*, vol. 51 (2), 2005, pp. 15-18.

[9] K. Boroczky, Jr. and G. Wintsche, "Covering the sphere by equal spherical balls", in Discrete and computational geometry: the Goodman-Pollach Festschrift (ed. by S. Basu et al.), 2003, 237-253.

[10] A. Feuer and A. Nemirovski, "On sparse representation in pairs of bases", *IEEE Transactions on Information Theory* 49(6): 1579-1581 (2003).

[11] N. Linial and I. Novik, "How neighborly can a centrally symmetric polytope be?", *Discrete and Computational Geometry*, 36(2006) 273-281.

[12] Y. Zhang, "When is missing data recoverable", available online at http://www.dsp.ece.rice.edu/cs/.

[13] J. Haupt and R. Nowak, "Signal reconstruction from noisy random projections", *IEEE Trans. on Information Theory*, 52(9), pp. 4036-4048, September 2006.

[14] M. J. Wainwright, "Sharp thresholds for high-dimensional and noisy recovery of sparsity", *Proc. Allerton Conference on Communication, Control, and Computing*, Monticello, IL, September 2006.

[15] J. Tropp, M. Wakin, M. Duarte, D. Baron, and R. Baraniuk, Random filters for compressive sampling and reconstruction. *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Toulouse, France, May 2006.

[16] Emmanuel Candes, "Compressive sampling", *Proc. International Congress of Mathematics*, 3, pp. 1433-1452, Madrid, Spain, 2006.