

# A SIMPLE PERFORMANCE ANALYSIS OF $\ell_1$ OPTIMIZATION IN COMPRESSED SENSING

Mihailo Stojnic

Purdue University, West Lafayette, IN  
e-mail: mstojnic@purdue.edu

## ABSTRACT

It is well known that compressed sensing problems reduce to solving large under-determined systems of equations. If we choose the elements of the compressed measurement matrix according to some appropriate probability distribution and if the signal is sparse enough then the  $\ell_1$  optimization can recover it with overwhelming probability (see, e.g. [4], [6], [7]). In fact, [4], [6], [7] establish (in a statistical context) that if the number of measurements is proportional to the length of the signal then there is a sparsity of the unknown signal proportional to its length for which the success of the  $\ell_1$  optimization is guaranteed. In this paper we introduce a novel, very simple technique for proving this fact. Furthermore, in addition to being very simple the new technique provides very good values for proportionality constants. In some cases, the presented analysis, although very simple, provides the best currently known values for the proportionality constants.

**Index Terms:** compressed sensing,  $\ell_1$ -optimization

## 1. INTRODUCTION

In this paper we are interested in the mathematical background of certain compressed sensing problems (more on the compressed sensing problems and their importance the interested reader can find in excellent references [12, 4, 15]). These problems are very easy to pose and very difficult to solve. Namely, we would like to find  $\mathbf{x}$  such that

$$A\mathbf{x} = \mathbf{y} \quad (1)$$

where  $A$  is an  $m \times n$  measurement matrix and  $\mathbf{y}$  is an  $m \times 1$  measurement vector. In usual compressed sensing context  $\mathbf{x}$  is an  $n \times 1$  unknown  $k$ -sparse vector (this means that  $\mathbf{x}$  has only  $k$  nonzero components; more on the so-called approximately sparse signals the interested reader can find in e.g. [5, 20]). In the rest of the paper we will further assume that  $k = \beta n$  and  $m = \alpha n$  where  $\alpha$  and  $\beta$  are absolute constants independent of  $n$ .

A particular way of solving (1) which recently generated a large amount of research is called  $\ell_1$ -optimization [4] (more on different solving algorithms the interested reader can find in [1, 22, 17, 16, 14]). The basic  $\ell_1$ -optimization algorithm finds  $\mathbf{x}$  in (1) by solving the following problem

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

Quite remarkably, in [4] the authors were able to show that if  $\alpha$  and  $n$  are given, the matrix  $A$  is given and satisfies a special property called the restricted isometry property (RIP) (more on when  $A$

satisfies the RIP condition the interested reader can find in [4, 2]), then any unknown vector  $\mathbf{x}$  with no more than  $k = \beta n$  ( $\beta$  is an absolute constant dependent on  $\alpha$  and explicitly calculated in [4] to which we will refer throughout the paper as threshold) 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. The case when the available measurements are noisy versions of  $\mathbf{y}$  is also of interest [4, 3, 23]. Although that case is not the primary concern of the present paper we mention in passing that the recent popularity of  $\ell_1$ -optimization in compressed sensing is significantly due to its robustness with respect to noisy measurements.

Instead of characterizing the  $m \times n$  matrix  $A$  through the RIP condition, in [6, 7] the author establishes that solutions of (2) and (1) coincide if and only if the projection of the regular  $n$ -dimensional cross-polytope by matrix  $A$  is a  $k$ -neighborly polytope. For  $A$  chosen as a random ortho-projector the precise relation between  $m$  and  $k$  so that the solutions of (2) and (1) coincide is characterized in [7, 6] as well.

In (1) one can further restrict  $\mathbf{x}$  to have only non-negative values (in the rest of the paper we refer to the vector  $\mathbf{x}$  with all non-negative components as *non-negative* vector). Instead of (2) one can then use the following, more restrictive, optimization problem

$$\begin{aligned} \min \quad & \|\mathbf{x}\|_1 \\ \text{subject to} \quad & A\mathbf{x} = \mathbf{y} \\ & \mathbf{x}_i \geq 0, \quad 0 \leq i \leq n \end{aligned} \quad (3)$$

to recover  $\mathbf{x}$ . As expected, adding positivity constraints helps (3) to be even more successful than (2) in recovering  $\mathbf{x}$ . A set of attainable values for strong and weak thresholds (more on the strong and weak threshold the interested reader can find in [7, 8]) for the success of (3) in recovering non-negative  $\mathbf{x}$  was computed in [9, 8, 10] for different types of statistical matrices  $A$ .

In this paper we will provide a technique for computing the thresholds that is different from those in [4, 6, 9, 10] but is rather simple. Our technique will massively utilize the structure of the null-space of the matrix  $A$ .

## 2. NULL-SPACE CHARACTERIZATIONS

The following two theorems provide the null-space characterizations of the matrix  $A$  that guarantee success of  $\ell_1$  optimizations (2) and (3) in recovering general and non-negative  $\mathbf{x}$ , respectively. (for similar characterizations see also [11, 24, 14]).

**Theorem 1 (Non-negative  $\mathbf{x}$ )** *Assume that an  $m \times n$  measurement matrix  $A$  is given. Let  $\mathbf{x}$  be a  $k$ -sparse vector whose non-zero components are positive. Further, assume that  $\mathbf{y} = A\mathbf{x}$  and that  $\mathbf{w}$  is an  $n \times 1$  vector. Let  $K$  be any subset of  $\{1, 2, \dots, n\}$  such*

that  $|K| = k$  and let  $K_i$  denote the  $i$ -th element of  $K$ . Further, let  $\bar{K} = \{1, 2, \dots, n\} \setminus K$ . Then (3) will produce the solution of (1) if

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

$$-\sum_{i=1}^k \mathbf{w}_{K_i} < \sum_{i=1}^{n-k} |\mathbf{w}_{\bar{K}_i}|. \quad (4)$$

**Proof 1** Suppose that the vectors from the null-space of the matrix  $A$  satisfy (4). Let  $\hat{\mathbf{x}}$  and  $\mathbf{x}$  be the solutions of (3) and (1), respectively. Further, assume that  $\hat{\mathbf{x}}$  is different from  $\mathbf{x}$ , i.e. assume that  $\hat{\mathbf{x}} \neq \mathbf{x}$ . Set  $\mathbf{w} = \hat{\mathbf{x}} - \mathbf{x}$ . Since  $A\hat{\mathbf{x}} - A\mathbf{x} = 0$  we have  $\mathbf{A}\mathbf{w} = 0$ . The following line of inequalities is easy to establish as well

$$\sum_{i=1}^k \mathbf{x}_{K_i} = \|\mathbf{x}\|_1 \geq \|\hat{\mathbf{x}}\|_1 = \|\mathbf{x} + \mathbf{w}\|_1 = \|\mathbf{x}_K + \mathbf{w}_K\|_1 + \|\mathbf{x}_{\bar{K}} + \mathbf{w}_{\bar{K}}\|_1$$

$$\geq \sum_{i=1}^k (\mathbf{x}_{K_i} + \mathbf{w}_{K_i}) + \|\mathbf{x}_{\bar{K}} + \mathbf{w}_{\bar{K}}\|_1 = \sum_{i=1}^k \mathbf{x}_{K_i} + \sum_{i=1}^k \mathbf{w}_{K_i} + \|\mathbf{w}_{\bar{K}}\|_1.$$

Connecting beginning and end we have  $-\sum_{i=1}^k \mathbf{w}_{K_i} \geq \|\mathbf{w}_{\bar{K}}\|_1$ . However this contradicts (4). Therefore we have  $\mathbf{x} = \hat{\mathbf{x}}$ . This concludes the proof. ■

**Theorem 2 (General  $\mathbf{x}$ )** Assume the setup of Theorem 1. Additionally, assume that non-zero components of  $\mathbf{x}$  can be both positive or negative. Let further  $\mathbf{1}$  be a  $2^k \times k$  sign matrix. Each element of the matrix  $\mathbf{1}$  is either 1 or  $-1$  and there are no two rows that are identical. Let  $\mathbf{1}_j$  be the  $j$ -th row of the matrix  $\mathbf{1}$ . Then (2) will produce the solution of (1) if

$$(\forall \mathbf{w} \in \mathbf{R}^n | \mathbf{A}\mathbf{w} = 0) \quad \text{and} \quad \forall K, j \quad -\mathbf{1}_j \mathbf{w}_K < \sum_{i=1}^{n-k} |\mathbf{w}_{\bar{K}_i}|. \quad (5)$$

**Proof 2** Follows from the proof of Theorem 1 after considering all possible combinations of signs that the components of  $\mathbf{x}_K$  can take and realizing that now  $\mathbf{w}_{\bar{K}_i}, 1 \leq i \leq n-k$ , does not need to be non-negative. ■

In the rest of the paper we compute the values of  $\alpha$  and  $\beta$  so that for certain random matrices  $A$  (4) and (5) hold with overwhelming probability.

### 3. PROBABILISTIC ANALYSIS

In the first part of this section we consider random matrices  $A$  that have a basis of the null-space comprised of i.i.d. zero-mean Gaussian elements (we refer to these matrices as null-gaussian). In the second part of this section we readily generalize the results to random matrices  $A$  that have a basis of the null-space comprised of i.i.d. Bernoulli elements (we refer to these matrices as null-bernoulli).

#### 3.1. Gaussian null-space

We focus on deriving the strong thresholds for non-negative case. The results for weak thresholds/general  $\mathbf{x}$  will easily follow.

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. Assume that  $A$  is a matrix such that  $Z$  is a basis of its null space. It then holds  $AZ = 0$ . Furthermore, any  $n \times 1$  vector  $\mathbf{w}$  from the null-space of  $A$  can be represented as  $Z\mathbf{v}$  where  $\mathbf{v} \in \mathbf{R}^{n-m}$ . Further, let  $I_{\mathbf{v}}$  denote the event  $-\sum_{i=1}^k Z_{K_i} \mathbf{v} \leq \sum_{i=1}^{n-k} |Z_{\bar{K}_i} \mathbf{v}|$  and let  $C_{K, \mathbf{v}}$  denote the polyhedral cone  $Z_{\bar{K}_i} \mathbf{v} \geq 0, 1 \leq i \leq (n-k)$ . Essentially, for any given constant  $\alpha = \frac{m}{n}$  we will compute a constant  $\beta = \frac{k}{n}$  such that

$$\lim_{n \rightarrow \infty} P(I_{\mathbf{v}} \quad \forall K \subset \{1, 2, \dots, n\}, |K| = k, \forall \mathbf{v} \in C_{K, \mathbf{v}}) = 1. \quad (6)$$

In order to show that (6) holds for certain values of  $\alpha$  and  $\beta$  we will actually show that

$$\lim_{n \rightarrow \infty} P_f = 0, \quad (7)$$

$P_f = P(\exists K \subset \{1, 2, \dots, n\}, |K| = k, \exists \mathbf{v} \in C_{K, \mathbf{v}} \text{ s. t. } \bar{I}_{\mathbf{v}})$  and  $\bar{I}_{\mathbf{v}}$  denotes the complement of  $I_{\mathbf{v}}$ , i.e. it denotes the event  $-\sum_{i=1}^k Z_{K_i} \mathbf{v} \geq \sum_{i=1}^{n-k} |Z_{\bar{K}_i} \mathbf{v}|$ . In what follows we will repeatedly use  $P_f$ . Our goal will always be to show that  $\lim_{n \rightarrow \infty} P_f = 0$ . Now, using the union bound we can write

$$P_f \leq \sum_{l=1}^{\binom{n}{k}} P(\exists \mathbf{v} \in C_{K^{(l)}, \mathbf{v}}, \text{ s. t. } -\sum_{i=1}^k Z_{K_i^{(l)}} \mathbf{v} \geq \sum_{i=1}^{n-k} |Z_{\bar{K}_i^{(l)}} \mathbf{v}|) \quad (8)$$

where  $K^{(l)}$  is a subset of  $\{1, 2, \dots, n\}$  and  $|K^{(l)}| = k$ . Clearly the number of these subsets is  $\binom{n}{k}$  and hence the summation in (8) goes from 1 to  $\binom{n}{k}$ . 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 (8) will then be equal. Therefore we can further write

$$P_f \leq \binom{n}{k} P(\exists \mathbf{v} \in C_{\mathbf{v}}, \text{ s. t. } -\sum_{i=1}^k Z_i \mathbf{v} \geq \sum_{i=k+1}^n |Z_i \mathbf{v}|) \quad (9)$$

where  $C_{\mathbf{v}}$  is the polyhedral cone  $Z_i \mathbf{v} \geq 0, k+1 \leq i \leq n$ . Let  $E$  be the set of all extreme rays of  $C_{\mathbf{v}}$ . Clearly,  $|E| \leq \binom{n-k}{n-m}$ . The function  $f(\mathbf{v}) = -\sum_{i=1}^k Z_i \mathbf{v} - \sum_{i=k+1}^n |Z_i \mathbf{v}|$  is convex (in fact linear) over the cone  $C_{\mathbf{v}}$  and achieves the maximum (up to the scaling constant) on its extreme rays. Hence we have

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

$$\leq \binom{n}{k} P(\max_{\mathbf{v} \in E} (-\sum_{i=1}^k Z_i \mathbf{v} - \sum_{i=k+1}^n |Z_i \mathbf{v}|) \geq 0). \quad (10)$$

Using the union bound over  $\mathbf{v}$  we further obtain

$$P_f \leq \binom{n}{k} P(\max_{\mathbf{v} \in E} (-\sum_{i=1}^k Z_i \mathbf{v} - \sum_{i=k+1}^n |Z_i \mathbf{v}|) \geq 0)$$

$$\leq \binom{n}{k} \sum_{t=1}^{\binom{n-k}{n-m}} P(-\sum_{i=1}^k Z_i \mathbf{v}_t - \sum_{i=k+1}^n |Z_i \mathbf{v}_t| \geq 0). \quad (11)$$

where  $\mathbf{v}_t, 1 \leq t \leq \binom{n-k}{n-m}$  are the extreme rays of  $C_{\mathbf{v}}$  (see Figure 1). Let  $L$  be the set of all subsets of  $\{k+1, k+2, \dots, n\}$  of car-

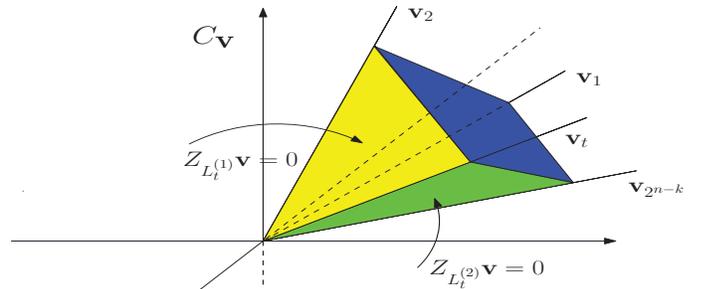


Fig. 1. Cone  $C_{\mathbf{v}}$ .

dinality  $n-m+1$ . Let  $L_t, 1 \leq t \leq \binom{n-k}{n-m}$  be the elements of the set  $L$  and let  $L_{t(l)}, 1 \leq l \leq n-m-1$  be the elements of the set  $L_t$ . Clearly,  $\mathbf{v}_t$  can be found as solutions to the systems of equations  $Z_{L_t} \mathbf{v}_t = 0$  where  $Z_{L_t}$  is a matrix obtained by selecting rows of  $Z$  indexed by the elements of  $L_t$  (We assume the worst case, i.e. that the matrices  $Z_{L_t}$  are nonsingular; moreover, as noted in [10], the matrices  $Z_{L_t}$  are nonsingular with overwhelming probability for each of the null-space distributions of interest in this paper, see [18]). Then from (11) we further have

$$P_f \leq \binom{n}{k} \sum_{t=1}^{\binom{n-k}{n-m}} P\left(-\sum_{i=1}^k Z_i \mathbf{v}_t - \sum_{i=k+1}^n |Z_i \mathbf{v}_t| \geq 0\right)$$

$$= \binom{n}{k} \sum_{t=1}^{\binom{n-k}{n-m}} P\left(\frac{-\sum_{i=1}^k Z_i \mathbf{v}_t}{\sum_{i=k+1, i \notin L_t}^n |Z_i \mathbf{v}_t|} \geq 1 | Z_{L_t}\right) P(Z_{L_t}). \quad (12)$$

Since  $\mathbf{v}_t$  only depends on  $Z_{L_t}$  it is independent of  $Z_i, 1 \leq i \leq n, i \notin L_t$ . Hence,  $\mathbf{v}_t$  on the right side of (12) can be treated as a constant vector. Furthermore, it is not difficult to see that the right side of (12) is independent of the index  $t$ . Using these facts without loss of generality we obtain

$$P_f \leq \binom{n}{k} \sum_{t=1}^{\binom{n-k}{n-m}} P\left(\frac{-\sum_{i=1}^k Z_i \mathbf{v}_t}{\sum_{i=k+1, i \notin L_t}^n |Z_i \mathbf{v}_t|} \geq 1 | Z_{L_t}\right) P(Z_{L_t})$$

$$= \binom{n}{k} \binom{n-k}{n-m} \times P\left(\frac{-\sum_{i=1}^k Z_i \mathbf{c}}{\sum_{i=k+1}^{m+1} |Z_i \mathbf{c}|} \geq 1, Z_i \mathbf{c} \geq 0, k+1 \leq i \leq m+1\right) \quad (13)$$

where  $\mathbf{c}$  is a deterministic vector. Since the norm of the vector  $\mathbf{c}$  is irrelevant we can further assume that  $\mathbf{a}_i = Z_i \mathbf{c}, 1 \leq i \leq m+1$  are i.i.d. zero mean Gaussian with variance 1. Hence we obtain

$$P_f \leq \binom{n}{k} \binom{n-k}{n-m} P\left(\frac{-\sum_{i=1}^k \mathbf{a}_i}{\sum_{i=k+1}^{m+1} |\mathbf{a}_i|} \geq 1, \mathbf{a}_i \geq 0, k+1 \leq i \leq m+1\right) \quad (14)$$

Removing the positivity constraint on  $\mathbf{a}_i, k+1 \leq i \leq m+1$ , and using the symmetry of Gaussian random variables it easily follows

$$P_f \leq 2^{k-m-1} \binom{n}{k} \binom{n-k}{n-m} P\left(\sum_{i=1}^k \mathbf{a}_i \geq \sum_{i=k+1}^{m+1} |\mathbf{a}_i|\right).$$

Using the Chernoff bound we further have

$$P_f \leq 2^{k-m-1} \binom{n}{k} \binom{n-k}{n-m} (E e^{\mu \mathbf{a}_1})^k (E e^{-\mu \mathbf{a}_1})^{m-k+1} \quad (15)$$

where  $\mu$  is a positive constant. After setting  $k = \beta n, m = \alpha n$ , and using the facts that  $\binom{n}{k} \approx e^{-nH(\beta)}$  and  $\binom{n-k}{n-m} \approx e^{-n(1-\beta)H(\frac{1-\alpha}{1-\beta})}$  ( $H$  is the base- $e$  entropy function) we finally obtain

$$P_f \leq (\xi_S)^n \quad (16)$$

where

$$\xi_S = \frac{2^{(\beta-\alpha)} e^{\frac{\beta\mu^2}{2}}}{e^{H(\beta)} e^{(1-\beta)H(\frac{1-\alpha}{1-\beta})}} \left( e^{\mu^2/2} \operatorname{erfc}\left(\frac{\mu}{\sqrt{2}}\right) \right)^{(\alpha-\beta)}. \quad (17)$$

Clearly, as long as  $\alpha \leq 1, \mu$ , and  $\beta$  are such that  $\xi_S < 1$ , (16) guarantees that (7) and subsequently (6) will hold.

### Theorem 3 (Strong threshold, Non-negative $\mathbf{x}$ , Null-gaussian)

Let  $A$  be an  $m \times n$  measurement matrix with a basis of null-space comprised of i.i.d.  $\mathcal{N}(0, 1)$  random variables. Let  $\alpha = \frac{m}{n}$ . Then, with overwhelming probability, (3) can recover any non-negative  $\mathbf{x}$  in (1) with sparsity no greater than  $\beta_S n$ . The value of  $\beta_S$  can be obtained by solving

$$\begin{aligned} \max \quad & \beta \\ \text{s.t.} \quad & \xi_S < 1, \mu > 0 \end{aligned} \quad (18)$$

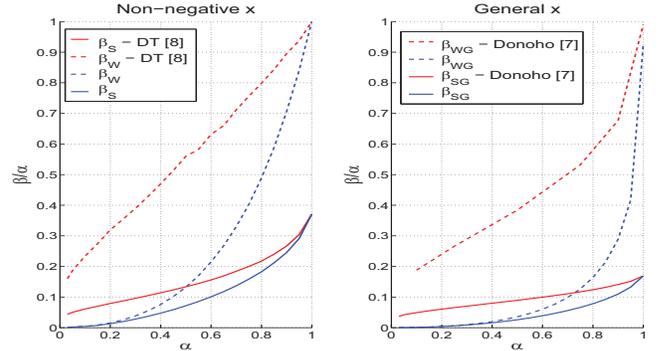
where  $\xi_S$  is as defined in (17).

**Proof 3** Follows from the previous discussion. ■

Corresponding value  $\beta_W$  for weak threshold can also be obtained by solving (18). However, instead of  $\xi_S$  one should rather use  $\xi_W = e^{H(\beta)} \xi_S$ .  $e^{H(\beta)}$  corresponds to a combinatorial term  $\binom{n}{k}$  that was used in the union bound in the above derivation. However, since that comes from a union bound over all possible signs of  $\mathbf{x}$  in the case of weak threshold it is not necessary to be included. The values for strong and weak thresholds in case of non-negative signals obtained based on the analysis presented here, as well as the best known ones [8], are displayed on the left side of Figure 2. To obtain the values for strong thresholds in case of general  $\mathbf{x}$   $\beta_{SG}$  one should solve (18) with (the details will be presented in the extended version of the paper [19])

$$\xi_{SG} = \frac{2^\beta e^{\frac{\beta\mu^2}{2}}}{e^{H(\beta)} e^{(1-\beta)H(\frac{1-\alpha}{1-\beta})}} \left( e^{\mu^2/2} \operatorname{erfc}\left(\frac{\mu}{\sqrt{2}}\right) \right)^{(\alpha-\beta)}. \quad (19)$$

Similarly as for non-negative signals, to obtain the values for weak thresholds  $\beta_{WG}$  in case of general  $\mathbf{x}$  one should solve (18) with  $\xi_{WG} = 2^{-\beta} e^{H(\beta)} \xi_{SG}$  [19]. The obtained values for strong and weak thresholds in case of general (non-necessarily positive) signals, as well as the best known ones [7], are displayed on the right side of Figure 2.



**Fig. 2.** Thresholds —  $A$  has Gaussian null-space; left) Non-negative  $\mathbf{x}$ , right) General  $\mathbf{x}$

### 3.2. Bernoulli null-space

We assume the complete setup of Section 3.1, except that now components of the matrix  $Z, Z_{ij}$  are assumed to be Bernoulli random variables, i.e.  $P(Z_{ij} = 1) = P(Z_{ij} = -1) = \frac{1}{2}$ . The derivation of the Section 3.1 can be repeated again. The only difference is that the value on the right side of (13) will not be independent of  $\mathbf{c}$ . However, it will be independent of the value  $\|\mathbf{c}\|_2$ . Hence, one can fix  $\|\mathbf{c}\|_2 = 1$  and instead of (15) we have [19]

$$P_f \leq \frac{\binom{n}{k} \binom{n-k}{n-m}}{2^{m-k}} \max_{\|\mathbf{c}\|_2=1} (E e^{\mu Z_1 \mathbf{c}})^k \max_{\|\mathbf{c}\|_2=1} (E e^{-\mu Z_1 \mathbf{c}})^{m-k+1} \quad (20)$$

where  $\mu$  is a positive constant and maximization over  $\mathbf{c}$  has a role of finding the worst case one among those that satisfy  $\|\mathbf{c}\|_2 = 1$ . We now derive upper bounds on terms on the right side of (20). Using the results from [13] we have

$$\max_{\|\mathbf{c}\|_2=1} E e^{\mu Z_1 \mathbf{c}} \leq E e^{\mu z} = e^{\frac{\mu^2}{2}} \quad (21)$$

where  $z$  is a zero-mean unit-variance Gaussian random variable. On the other hand using Taylor expansion we have

$$\begin{aligned} \max_{\|\mathbf{c}\|_2=1} E e^{-\mu Z_1 \mathbf{c}} &\leq 1 - \mu \max_{\|\mathbf{c}\|_2=1} E(Z_1 \mathbf{c}) + \frac{\mu E(Z_1 \mathbf{c})^2}{2} \\ &= 1 - \mu \max_{\|\mathbf{c}\|_2=1} E(Z_1 \mathbf{c}) + \frac{\mu^2}{2}. \end{aligned} \quad (22)$$

The Khintchine inequality [21] states that

$$\max_{\|c\|_2=1} E(Z_1 c) \geq \frac{1}{\sqrt{2}}. \quad (23)$$

Replacing the upper bound from (23) in (22) we obtain

$$\max_{\|c\|_2=1} E e^{-\mu Z_1 c} \leq 1 - \frac{\mu}{\sqrt{2}} + \frac{\mu^2}{2}. \quad (24)$$

Combining (20), (21), and (24) we finally have

$$P_f \leq \frac{\binom{n}{k} \binom{n-k}{n-m}}{2^{m-k}} e^{\frac{k\mu^2}{2}} \left(1 - \frac{\mu}{\sqrt{2}} + \frac{\mu^2}{2}\right)^{m-k}.$$

Following the procedure after (15) one can define

$$\xi_S^{bern} = \frac{2^{(\beta-\alpha)} e^{\frac{\beta\mu^2}{2}}}{e^{H(\beta)} e^{(1-\beta)H(\frac{1-\alpha}{1-\beta})}} \left(1 - \frac{\mu}{\sqrt{2}} + \frac{\mu^2}{2}\right)^{(\alpha-\beta)} \quad (25)$$

and formulate the following theorem.

**Theorem 4 (Strong threshold, Non-negative x, Null-bernoulli)**  
Let  $A$  be an  $m \times n$  measurement matrix with a basis of null-space comprised of i.i.d. Bernoulli random variables. Let  $\alpha = \frac{m}{n}$ . Then, with overwhelming probability, (3) can recover any non-negative  $x$  in (1) with sparsity no greater than  $\beta_S^{bern} n$ . The value of  $\beta_S^{bern}$  can be obtained by solving

$$\begin{aligned} \max \quad & \beta \\ \text{s.t.} \quad & \xi_S^{bern} < 1, \mu > 0. \end{aligned}$$

where  $\xi_S^{bern}$  is defined in (25).

**Proof 4 Omitted.** ■

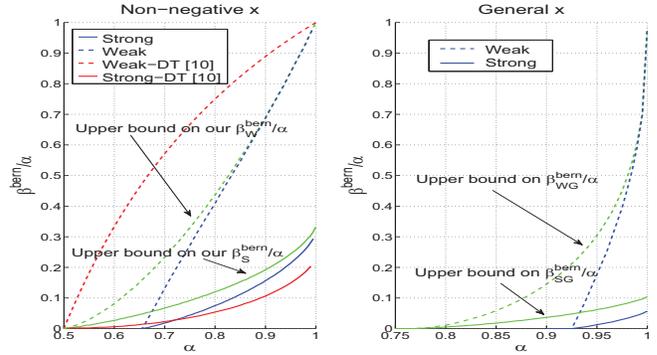
Quantities  $\beta_W^{bern}$ ,  $\beta_{WG}^{bern}$ , and  $\beta_{SG}^{bern}$  that would denote weak threshold for non-negative, weak threshold for general, and strong threshold for general signals, respectively can be determined in the same way that they were determined in the previous subsection (see [19]). Obtained values for  $\beta_S^{bern}$ ,  $\beta_W^{bern}$ ,  $\beta_{WG}^{bern}$ ,  $\beta_{SG}^{bern}$  are displayed in blue on Figure 3 (on the left side of Figure 3 are threshold values for non-negative  $x$ ; on the right side are threshold values for general  $x$ ). The best known values for  $\beta_S^{bern}$ ,  $\beta_W^{bern}$  are those from [10] and they are shown in red on the left side of Figure 3. It is not known to the author that there are corresponding results for  $\beta_{SG}^{bern}$ ,  $\beta_{WG}^{bern}$  available in the literature. Also, it should be noted that in green, the upper limits (obtained by computing the left side of (22) for a fixed  $c$  rather than upper-bounding it in general) of our analysis are shown as well [19].

#### 4. CONCLUSION

Our analysis provided a somewhat new technique in proving the optimality of the  $l_1$ -norm optimization in compressed sensing problems with measurement matrices that have Gaussian and Bernoulli distributed null-spaces. The technique is very simple, yet it produces very good values of thresholds.

#### 5. REFERENCES

- [1] M. Akcakaya and V. Tarokh. Performance bounds on sparse representations using redundant frames. 2007. available online at arXiv:cs/0703045.
- [2] R. Baraniuk, M. Davenport, R. DeVore, and M. 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/>.
- [3] E. Candes, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. on Information Theory*, 52:489–509, December 2006.



**Fig. 3.** Thresholds —  $A$  has Bernoulli null-space; left) Non-negative  $x$ , right) General  $x$

- [4] E. Candes and T. Tao. Decoding by linear programming. *IEEE Trans. on Information Theory*, 51:4203–4215, Dec. 2005.
- [5] A. Cohen, W. Dahmen, and R. DeVore. Compressed sensing and best k-term approximation. 2006. available online at <http://www.dsp.ece.rice.edu/cs/>.
- [6] D. Donoho. Neighborly polytopes and sparse solutions of underdetermined linear equations. 2004. Technical report, Department of Statistics, Stanford University.
- [7] D. Donoho. High-dimensional centrally symmetric polytopes with neighborliness proportional to dimension. *Disc. Comput. Geometry*, 35(4):617–652, 2006.
- [8] D. Donoho and J. Tanner. Neighborliness of randomly-projected simplices in high dimensions. *Proc. National Academy of Sciences*, 102(27):9452–9457, 2005.
- [9] D. Donoho and J. Tanner. Sparse nonnegative solutions of underdetermined linear equations by linear programming. *Proc. National Academy of Sciences*, 102(27):9446–9451, 2005.
- [10] D. Donoho and J. Tanner. Counting the face of randomly projected hypercubes and orthants with application. 2008. available online at <http://www.dsp.ece.rice.edu/cs/>.
- [11] D. L. Donoho and X. Huo. Uncertainty principles and ideal atomic decompositions. *IEEE Trans. Inform. Theory*, 47(7):2845–2862, November 2001.
- [12] M. Duarte, M. Davenport, D. Takhar, J. Laska, T. Sun, K. Kelly, and R. Baraniuk. Single-pixel imaging via compressive sampling. *IEEE Signal Processing Magazine*, 25(2), 2008.
- [13] M. L. Eaton. A note on symmetric Bernoulli random variables. *The annals of mathematical statistics*, 41(4):1223–1226, 1970.
- [14] R. Gribonval and M. Nielsen. Sparse representations in unions of bases. *IEEE Trans. Inform. Theory*, 49(12):3320–3325, December 2003.
- [15] M. Mishali and Y. Eldar. Reduce and boost: Recovering arbitrary sets of jointly sparse vectors. 2007. available online at <http://www.dsp.ece.rice.edu/cs/>.
- [16] D. Needell and J. A. Tropp. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. 2008. available online at <http://www.dsp.ece.rice.edu/cs/>.
- [17] F. Parvaresh and B. Hassibi. Explicit measurements with almost optimal thresholds for compressed sensing. *IEEE ICASSP*, Mar-Apr 2008.
- [18] M. Rudelson and R. Vershynin. The smallest singular value of a rectangular random matrix. 2008. preprint.
- [19] M. Stojnic. A simple performance analysis of  $l_1$  optimization in compressed sensing. 2008. preprint, available from the author.
- [20] M. Stojnic, W. Xu, and B. Hassibi. Compressed sensing of approximately sparse signals. *ISIT, International symposium on information theory*, July 2008.
- [21] S. J. Szarek. On the best constants in the Khintchine inequality. *Studia Math.*, 58:197–208, 1976.
- [22] J. Tropp and A. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Trans. on Information Theory*, 53(12):4655–4666, 2007.
- [23] M. J. Wainwright. Sharp thresholds for high-dimensional and noisy recovery of sparsity. *Proc. Allerton Conference on Communication, Control, and Computing*, September 2006.
- [24] Y. Zhang. When is missing data recoverable. available online at <http://www.dsp.ece.rice.edu/cs/>.