A SIMPLER APPROACH TO WEIGHTED ℓ_1 MINIMIZATION

Anilesh K. Krishnaswamy

Indian Institute of Technology Madras
Chennai, India – 600036

Samet Oymak and Babak Hassibi

California Institute of Technology
Pasadena, CA – 91125

ABSTRACT

In this paper, we analyze the performance of weighted ℓ_1 minimization over a non-uniform sparse signal model by extending the "Gaussian width" analysis proposed in [1]. Our results are consistent with those of [7] which are currently the best known ones. However, our methods are less computationally intensive and can be easily extended to signals which have more than two sparsity classes. Finally, we also provide a heuristic for estimating the optimal weights, building on a more general model presented in [11]. Our results reinforce the fact that weighted ℓ_1 minimization is substantially better than regular ℓ_1 minimization and provide an easy way to calculate the optimal weights.

Index Terms— weighted ℓ_1 minimization, compressed sensing, Gaussian measurements, recovery threshold, Gaussian width

1. INTRODUCTION

In recent years, sparse recovery problems have been the subject of great interest([4,6,10]) due to their importance in various applications. It is now well known that there are algorithms which, under certain conditions, can provably recover the underlying sparse solution of a system. ℓ_1 minimization is arguably the most popular among these as the ℓ_1 norm is the tightest convex surrogate of the sparsity function ("ℓ_0" norm). The most common form of the ℓ_1 minimization problem is as follows:

$$\min \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{y} \tag{1}$$

where $\mathbf{A} \in \mathbb{R}^{m\times n}$ is the measurement matrix, $\mathbf{s} \in \mathbb{R}^n$ is the k-sparse solution, $\mathbf{y} = \mathbf{A}\mathbf{s} \in \mathbb{R}^m$ is the corresponding measurement and we want the solution $\mathbf{x}^*$ of (1) to be equal to $\mathbf{s}$.

Clearly, this problem is interesting when the system is underdetermined and when certain properties of $\mathbf{A}$, such as restricted isometry [4], can guarantee that $\mathbf{x}^* = \mathbf{s}$. An important problem is the characterization of the performance of the minimization (1). In particular, the relation between the sparsity of $\mathbf{s}$, the number of measurements m and the ambient dimension n has been studied in great detail for Gaussian measurements ([1,3,5]).

In this work, we investigate a modified ℓ_1 minimization algorithm, which significantly improves over regular ℓ_1 minimization, by using prior information about the structure of the "non-uniformly sparse" solution as described in Definition 2.1.

Contributions: In [7], Khajehnejad et al. use a Grassman angle approach to calculate the optimal weights. We should emphasize that by optimal weights we mean the weights that minimize the number of Gaussian measurements required for weighted ℓ_1 minimization to succeed in recovering the sparse signal. The approach of [7] requires extensive analytical work and the numerical procedure for optimal weight calculation is inefficient for more than two blocks. In this paper, we determine the optimal weights by extending the methods of [1]. In [1], Stojnic develops a new and relatively simpler method to obtain the minimum number of measurements m to recover a sparsity of k and finds the phase transition curve which matches with the results of [3,5], which are known to be *exact*. Consequently, without using complicated techniques, we are able to provide a method for obtaining optimal weights and because the analysis in [7] is based on [3,5], our results are consistent with [7]. Further, the resulting numerical procedure for optimal weight computation is cheaper and more clear. We also provide a heuristic method for estimating the optimal weight based on the more general "atomic decomposition" model presented in [11].

The structure of this paper is as follows. In the next section we define the "non-uniform sparse" model and weighted ℓ_1 minimization and give other supplementary defintions. We then state the main theorems that will be used in this paper and proceed to the "Gaussian width" analysis of the weighted ℓ_1 minimization.

2. SIGNAL MODEL

First, we introduce our model and algorithm.

Definition 2.1 (Non-uniformly Sparse Signal) *Let* $K = \{K_1, K_2, ..., K_u\}$ *be a partition of* $\{1, 2, ..., n\}$*, i.e.,* $K_i \cap K_j = \emptyset$ *for* $i \neq j$*,* $\cup_{i=1}^u K_i = \{1, 2, ..., n\}$ *and let* $|K_i| = n_i = \gamma_i n$ *for* $1 \leq i \leq u$*. Let* $B = \{\beta_1, \beta_2, ..., \beta_u\}$ *be a set of positive numbers in* $[0, 1]$*. An* $n \times 1$ *vector* $\mathbf{x} = (x_1, x_2, ..., x_n)^T$ *is said to be a non-uniformly sparse* *vector with sparsity fractions B over K if $k_i = \beta_i n_i$ entries of x in K_i are non-zero. Finally, the overall sparsity of the signal is denoted by $k = \sum_{i=1}^{u} k_i$ and $\mathbf{x}^i \in \mathbb{R}^{n_i}$ is the vector that corresponds to the entries of $\mathbf{x}$ in the i'th block for $1 \le i \le u$.*

Observe that, we don't have any information about the entries inside a particular block hence it is natural to use the same weight for all entries of the same block. We use the following weighted ℓ_1 minimization scheme:

$$\min_{\mathbf{Ax}=\mathbf{y}} \|\mathbf{x}\|_{\mathbf{w},\mathbf{1}} = \min_{\mathbf{Ax}=\mathbf{y}} \sum_{i=1}^{u} w_i \|\mathbf{x}^i\|_1. \qquad (2)$$

where $\mathbf{w} = \{w_1, w_2, ..., w_u\}$ is the weight vector we use. As argued in [9], one would expect to associate larger weights to the sparser blocks according to the so called "punish the sparser" scheme. By determining the optimum weight vector, we will show that this scheme can significantly improve regular ℓ_1 minimization.

Similar to relevant previous works, it is assumed that measurement matrix is $m \times n$ Gaussian. Also, we assume the so-called "linear regime" where $\{\gamma_i\}_{i=1}^u, \{\beta_i\}_{i=1}^u$ and $\alpha = \frac{m}{n}$ are constant. Our results will be true asymptotically with high probability (a.w.h.p.), i.e. as $n \to \infty$.

Finally, the "recovery threshold" of a weighted algorithm is the minimum α that can guarantee the successful recovery of a signal obeying the non-uniform model a.w.h.p. Naturally, the recovery threshold is a function of $\{\beta_i, \gamma_i, w_i\}_{i=1}^u$.

3. MAIN THEOREMS

We now state a seminal result from [8] which is the most important theorem used in this paper.

Theorem 3.1 (Escape through a mesh) *Let $\mathcal{S}$ be a subset of the unit Euclidean sphere $\mathcal{S}^{n-1}$ in $\mathbb{R}^n$. Let $\mathcal{Y}$ be a random $(n-m)$-dimensional subspace of $\mathbb{R}^n$, distributed uniformly in the Grassmanian with respect to the Haar measure. Let the **Gaussian width** of $\mathcal{S}$, denoted by $g(\mathcal{S})$, be defined as*

$$g(\mathcal{S}) = E \sup_{\mathbf{v}\in\mathcal{S}} (\tilde{\mathbf{h}}^T \mathbf{v}) \qquad (3)$$

where $\tilde{\mathbf{h}}$ is a random column vector in R^n with i.i.d. $\mathcal{N}(0,1)$ components. Assume that $g(\mathcal{S}) < (\sqrt{m} - \frac{1}{4\sqrt{m}})$. Then

$$P(\mathcal{Y} \cap \mathcal{S} = \emptyset) > 1 - 3.5 e^{-\frac{(\sqrt{m} - \frac{1}{4\sqrt{m}} - g(S))^2}{18}} \qquad (4)$$

Let $\mathbf{A}$ be a $k \times n$ "Gaussian Matrix" and $\mathcal{N}(\mathbf{A})$ denote the null space of $\mathbf{A}$, i.e. $\mathbf{t} \in \mathcal{N}(\mathbf{A}) \iff \mathbf{At} = 0$. As mentioned in [10], it is a well known fact that $\mathcal{N}(\mathbf{A})$ is distributed uniformly in the Grassmanian of $(n-k)$-dimensional subspaces of $\mathbb{R}^n$, with respect to the Haar measure.

Since we use Gaussian measurements $\mathbf{A}$, Theorem 3.1 helps one to analyze the null space of $\mathbf{A}$. This idea was first introduced by Stojnic, in [1], and was used to find various thresholds for regular ℓ_1 minimization [1], [12]. In this paper, we show that same method can be extended to analyze weighted ℓ_1 minimization. To do this, we need to characterize the undesired null space elements because whenever $\mathcal{N}(\mathbf{A})$ does not contain them weighted ℓ_1 minimization gives us the desired sparse solution.

It is known that (see, e.g., [1], [3]), the particular locations and signs of the non-zero elements of the underlying sparse solution $\mathbf{s}$ are irrelevant to the analysis of the recovery threshold. The following theorem gives a necessary and sufficient condition for the success of weighted ℓ_1 minimization ([7,9]):

Theorem 3.2 (Null Space Condition) *Assume $\mathbf{A} \in \mathbb{R}^{m\times n}$ is Gaussian and let $\mathbf{s}$ be a k-sparse vector as described in **Definition 2.1**. Assume its non-zero components are negative. Also, let the first n_1 components be in K_1, the next n_2 components in K_2, and so on. Let the j'th component of the i'th block $\mathbf{s}^i$ be denoted by s_j^i for $1 \le i \le u$ and $1 \le j \le n_i$. Also let $s_1^i = s_2^i = ... = s_{n_i-k_i}^i = 0$. Then subject to the observations $\mathbf{y} = \mathbf{As}$, minimization (2) will return $\mathbf{s}$ as the unique solution iff for any $\mathbf{t} \in \mathcal{N}(\mathbf{A})$ we have:*

$$\sum_{i=1}^{u} \left(w_i \sum_{j=n_i-k_i+1}^{n_i} t_j^i \right) < \sum_{i=1}^{u} \left(w_i \sum_{j=1}^{n_i-k_i} |t_j^i| \right) \qquad (5)$$

The next section is dedicated to the analysis of the weighted recovery threshold using Theorems 3.1 and 3.2.

4. ANALYSIS OF THE RECOVERY THRESHOLD

For simplicity, we'll focus on the case $u = 2$ and without loss of generality assume $w_1 = 1$ and $w_2 = w$.

Let $\mathcal{S}_\mathbf{w}$ be the set of vectors which have unit ℓ_2 norm and which do not satisfy (**??**). For a given $\mathbf{w}$, we will compute a tight upper bound on $g(\mathcal{S}_\mathbf{w})$ and then equate that to $\left(\sqrt{m} - \frac{1}{4\sqrt{m}}\right)$. We set $g(\tilde{\mathbf{h}}, \mathcal{S}_\mathbf{w}) = \max_{\mathbf{v}\in\mathcal{S}_\mathbf{w}} (\tilde{\mathbf{h}}^T \mathbf{v})$ and, as a first step, we'll determine an upper bound $B_\mathbf{w}$ on $g(\tilde{\mathbf{h}}, \mathcal{S}_\mathbf{w})$ for a fixed $\tilde{\mathbf{h}}$.

In the following analysis, due to insufficient space, we will omit some of the intermediate steps and will try to emphasize the main points instead. For a complete discussion, reader is referred to the longer version of this paper [14].

4.1. Upper bound $B_\mathbf{w}$ on $g(\tilde{\mathbf{h}}, \mathcal{S}_\mathbf{w})$

Our analysis will mostly follow that of [1]. Let $\tilde{\mathbf{h}}_{1:(n_i-k_i)}^i = (\tilde{h}_1^i, \tilde{h}_2^i, ..., \tilde{h}_{n_i-k_i}^i)^T$. Further, let $|\tilde{h}|_{(j)}^i$ be the j'th smallest magnitude of elements of $\tilde{\mathbf{h}}_{1:(n_i-k_i)}^i$. Let $\mathbf{h} = (\mathbf{h}^1, \mathbf{h}^2)$ where

$$\mathbf{h}^i = (|\tilde{h}|_{(1)}^i, |\tilde{h}|_{(2)}^i, ..., |\tilde{h}|_{(n_i-k_i)}^i, \tilde{h}_{n_i-k_i+1}^i, ..., \tilde{h}_n^i)^T \qquad (6)$$

By using elementary arguments, $g(\tilde{\mathbf{h}}, \mathcal{S}_{\mathbf{w}})$ can be expressed as the outcome of the following optimization ([1]):

$$\max_{\mathbf{y}\in\mathbb{R}^n} \quad (\mathbf{h}^1)^T\mathbf{y}^1 + (\mathbf{h}^2)^T\mathbf{y}^2$$

subject to

i) $y_j^i \geq 0$ for all $i = 1, 2$ and $0 \leq j \leq (n_i - k_i)$

ii) $$\sum_{i=1}^{2}\left(w_i \sum_{j=n_i-k_i+1}^{n_i} y_j^i\right) \geq \sum_{i=1}^{2}\left(w_i \sum_{j=1}^{n_i-k_i} |y_j^i|\right)$$

iii) $$\sum_{i=1}^{2}\left(\sum_{j=1}^{n_i}(y_j^i)^2\right) \leq 1.$$

Let $\mathbf{z} = (\mathbf{z}^1, w\mathbf{z}^2)$ where $\mathbf{z}^i \in R^{n_i}$ such that $z_j^i = 1, 1 \leq j \leq n_i - k_i$ and $z_j^i = -1, n_i - k_i + 1 \leq j \leq n_i$. The next step is changing the above maximization problem to a minimization problem by writing out its dual which will provide us an upper bound $g_{up}(\tilde{\mathbf{h}}, \mathcal{S}_{\mathbf{w}})$:

$$\min_{\gamma,\nu,\boldsymbol{\lambda}} \max_{\mathbf{y}} \quad \mathbf{h}^T\mathbf{y} - \gamma\|\mathbf{y}\|_2^2 + \gamma - \nu\mathbf{z}^T\mathbf{y} + \boldsymbol{\lambda}^T\mathbf{y}$$

subject to $\nu \geq 0, \gamma \geq 0$

$\lambda_j^i \geq 0$ for $i = 1, 2$ and $0 \leq j \leq n_i - k_i$.

After trivially solving the minimization over $\mathbf{y}$ and maximization over γ we end up with:

$$g_{up}(\tilde{\mathbf{h}}, \mathcal{S}_{\mathbf{w}}) = \min_{\nu,\boldsymbol{\lambda}} \quad ||\boldsymbol{\lambda} + \mathbf{h} - \nu\mathbf{z}||_2 \tag{7}$$

subject to $\nu \geq 0$

$\lambda_j^i \geq 0$ for $i = 1, 2$ and $0 \leq j \leq n_i - k_i$.

To tackle this problem, we deal with the square of the objective function in (6), namely $\|\boldsymbol{\lambda} + \mathbf{h} - \nu\mathbf{z}\|_2^2$. Now let $\boldsymbol{\lambda}_1 = (\lambda_1^1, \lambda_2^1, ..., \lambda_{c_1}^1, 0, 0, ..., 0)^T$ where $c_1 \leq (n_1 - k_1)$ and $\boldsymbol{\lambda}_2 = (\lambda_1^2, \lambda_2^2, ..., \lambda_{c_2}^2, 0, 0, ..., 0)^T$ where $c_2 \leq (n_2 - k_2)$. Parameters c_1 and c_2 will be determined later. This makes the optimization over ν simpler which can be done by differentiating the objective with respect to ν. Assuming ν is positive, λ_j^i's can be determined by further differentiation and proceeding in a similar way to [1]. Eventually, we end up with the following lemma for fixed c_1, c_2.

Lemma 4.1 *Let $c' = c_1 + w^2c_2$, $n' = n_1 + w^2n_2$ and*

$$f(\mathbf{h}, \mathbf{z}, w, c_1, c_2) = \frac{\left(\mathbf{h}^T\mathbf{z} - \sum_{j=1}^{c_1} h_j^1 - w\sum_{k=1}^{c_2} h_k^2\right)}{n' - c'} \tag{8}$$

Assume

$$f(\mathbf{h}, \mathbf{z}, w, c_1, c_2) \geq \max\left\{h_{c_1}^1, h_{c_2}^2\right\} \tag{9}$$

Then $g(\tilde{\mathbf{h}}, \mathcal{S}_{\mathbf{w}})$ is upper bounded by $g_{up}(\tilde{\mathbf{h}}, \mathcal{S}_{\mathbf{w}}, c_1, c_2) =$

$$\sqrt{\sum_{j=c_1+1}^{n_1}(h_j^1)^2 + \sum_{k=c_2+1}^{n_2}(h_k^2)^2 - (n' - c')f(\mathbf{h}, \mathbf{z}, w, c_1, c_2)^2}$$

Let $F_\alpha(\cdot)$ be the magnitude distribution of a unit variance Gaussian RV. For any $\epsilon > 0$, if c_1, c_2 are chosen such that

$$(1 - \epsilon)\mathbb{E}[f(\mathbf{h}, \mathbf{z}, w, c_1, c_2)] = F_\alpha^{-1}\left(\frac{c_1(1 + \epsilon)}{n_1(1 - \beta_1)}\right) \tag{10}$$

$$w(1 - \epsilon)\mathbb{E}[f(\mathbf{h}, \mathbf{z}, w, c_1, c_2)] = F_\alpha^{-1}\left(\frac{c_2(1 + \epsilon)}{n_2(1 - \beta_2)}\right) \tag{11}$$

it can be shown that as $n \to \infty$, (8) will be satisfied except for an exponentially small probability. Consequently,

$$\mathbb{E}[g(\tilde{\mathbf{h}}, \mathcal{S}_{\mathbf{w}})] \leq \mathbb{E}[g_{up}(\tilde{\mathbf{h}}, \mathcal{S}_{\mathbf{w}}, c_1, c_2)] + o(1) \tag{12}$$

$$\leq \sqrt{\mathbb{E}[(g_{up}(\tilde{\mathbf{h}}, \mathcal{S}_{\mathbf{w}}, c_1, c_2))^2]} + o(1) \tag{13}$$

Hence, we have the following theorem:

Theorem 4.1 *Let $\tilde{\mathbf{h}}$ be an i.i.d. Gaussian vector with entries having unit variance. $\mathbf{h}$ is function of $\tilde{\mathbf{h}}$ as given in (5) and n', c' are same as in Lemma 4.1. Then, for a given weight $\mathbf{w}$*

$$\mathbb{E}\left[\sum_{i=1}^{2}\sum_{j=c_i+1}^{n_i}(h_j^i)^2\right] - (n' - c')\mathbb{E}[f(\mathbf{h}, \mathbf{z}, w, c_1, c_2)]^2 \tag{14}$$

measurements are sufficient for recovering a signal that obeys Definition 2.1 w.h.p. Here c_1, c_2 are solutions of (9,10).

Next, letting $\epsilon \to 0$ in (9,10) and using the fact that $\tilde{\mathbf{h}}$ is Gaussian, we restate Theorem 4.2 for the more general case of $u \geq 2$ in a way that is accessible for numerical calculations.

Theorem 4.2 *Let η_i, κ, μ be functions of $\{\beta_i, \gamma_i, \theta_i\}_{i=1}^u$ defined as follows*

$$\eta_i = \textit{erfinv}\left(\frac{1 - \theta_i}{1 - \beta_i}\right) \textit{ for } 0 \leq i \leq u$$

$$\kappa = \left(\sum_{i=1}^{u} w_i^2\gamma_i\theta_i\right)^{-1}$$

$$\mu = \sum_{i=1}^{u} w_i\gamma_i(1 - \beta_i)\sqrt{\frac{2}{\pi}}\exp(-\eta_i^2)$$

where $\theta_i = \left(1 - \frac{c_i}{n_i}\right)$.

Step 1: *For $0 \leq i \leq u$ let θ_i be the solution of:*

$$\sqrt{2}\eta_i = w_i\kappa\mu \tag{15}$$

Step 2: *Then, an* **upper bound** *to the threshold of weighted ℓ_1 minimization with weight $\mathbf{w}$ is called $\alpha(\mathbf{w})$ and is given by*

$$\sum_{i=1}^{u}\gamma_i\left\{(1 - \beta_i)\left(1 + \frac{2\eta_i}{\sqrt{\pi}\exp(\eta_i^2)} - \textit{erf}(\eta_i)\right) + \beta_i\right\} - \kappa\mu^2$$

Hence $\alpha(\mathbf{w}) \times n$ measurements are sufficient to recover a "nonuniformly sparse signal" given in Definition 2.1 w.h.p.

Observe that **Steps 1 and 2** in Theorem 4.2 corresponds to (9,10) and (**??**) respectively.

Finally, in order to determine the **optimal weight**, $\alpha(\mathbf{w})$ is minimized over $\mathbf{w}$ which can be done numerically and will be the topic of the next section.

5. RESULTS

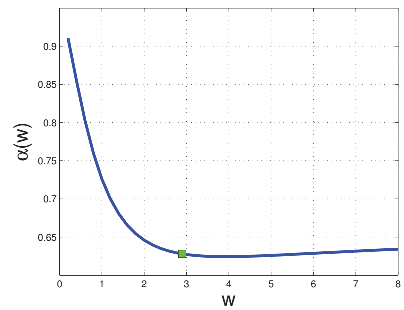
Comparison with previous work: Our results are consistent with [7] as simulations suggest we obtain the same optimal weighting given $\{\beta_i, \gamma_i\}_{i=1}^{u}$. Since [7] is based on [3] and exact, this shows our analysis yields the true optimal weight by using an alternative method.

On the other hand, our method is more accessible: Theorem 4.2 is concise and our numerical computations are much less intensive . For example, the weight curve in Figure 5 is smooth unlike those of [7]. Curves in [7] were not calculated with high density because of the high complexity of the optimal weight calculation [13].

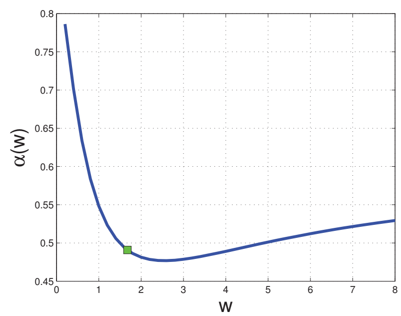
A heuristic to estimate the optimal weight
In [11], a new framework is provided to determine the approximate thresholds using Gaussian width analysis using which it can be shown that regular ℓ_1 requires $\alpha \geq 2(1+e^{-1})\beta\left(\log((1-\beta)/\beta)+1.5\right)$ for success. Keeping this as the reference point, one can come up with the following closed form approximation for the optimal weight in the case of weighted ℓ_1 minimization:

$$w = \sqrt{\frac{1+\log(\frac{1-\beta_2}{\beta_2})}{1+\log(\frac{1-\beta_1}{\beta_1})}} \tag{16}$$

It should be emphasized that (12) is independent of the block sizes γ_1, γ_2.



(a) $\beta_1 = 0.65, \beta_2 = 0.1$



(b) $\beta_1 = 0.4, \beta_2 = 0.05$

Fig. 1. $\alpha(\mathbf{w})$ vs. w for $\gamma_1 = \gamma_2 = 0.5$

5.1. Simulation results and Conclusion

Figure 5 presents the threshold value $\alpha(\mathbf{w})$ for different values of $w = \frac{w_2}{w_1}$ given the values of $\beta_1, \beta_2, \gamma_1, \gamma_2$. The value of the recovery threshold corresponding to the weight given by the heuristic in (12) is highlighted.

In general, we see a substantial decrease in the recovery threshold from regular ℓ_1 minimization. The heuristic, although not very satisfactory, gives us approximately the optimal weight.

6. REFERENCES

[1] M. Stojnic. “Various thresholds for ℓ_1-optimization in compressed sensing”. Preprint, 2009. Available at arXiv:0907.3666.

[2] A. Barvinok and A. Samorodnitsky. “Random weighting, asymptotic counting, and inverse isoperimetry”. Israel Journal of Mathematics, 158:159-191, 2007.

[3] D. Donoho. “High-dimensional centrally symmetric polytopes with neighborliness proportional to dimension”. Disc. Comput. Geometry, 35(4):617652, 2006.

[4] E.J. Candès and T. Tao. “Decoding by Linear Programming”. IEEE Trans. Inform. Theory, 51 4203-4215.

[5] D. Donoho and J. Tanner. “Neighbourliness of randomly-projected simplices in high dimensions”. Proc. Natl. Acad. Sci. USA, 102(27):9452-9457, 2005

[6] D. Donoho and J. Tanner. “Sparse nonnegative solutions of underdetermined linear equations by linear programming”. Proc. Natl. Acad. Sci. USA, 102(27):9446-9451, 2005

[7] M.A. Khajehnejad, W. Xu, S. Avestimehr, and B. Hassibi. “Analyzing Weighted ℓ_1 Minimization for Sparse Recovery With Nonuniform Sparse Models”. IEEE Tran. on Signal Proc. 59(5): 1985-2001 (2011)

[8] Y. Gordon. “On Milmans inequality and random subspaces which escape through a mesh in R^n. Geometric aspects of functional analysis, Isr. Semin. 1986-87, Lect. Notes Math. 1317, 84-106 (1988).

[9] S. Oymak, M.A. Khajehnejad, and B. Hassibi. “Weighted Compressed Sensing and Rank Minimization”. ICASSP 2011: 3736-3739.

[10] E. Vershynin and M.Rudelson, “On sparse reconstruction using Fourier and Gaussian Measurements”.

[11] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A.S. Willsky. “The Convex Geometry of Linear Inverse Problems”. Revised 2011. Available at arXiv:1012.0621v2.

[12] M. Stojnic. “Block-length dependent thresholds in block-sparse compressed sensing”. Preprint, 2009. Available at arXiv:0907.3679v1.

[13] M.A. Khajehnejad. Personal communication. 2011

[14] A.K. Krishnaswamy, S. Oymak, and B. Hassibi. “A simpler approach to weighted ℓ_1 minimization”. Complete report preprint, 2011. Available at `www.ee.iitm.ac.in/~ee08b017`