THE STABILITY OF REGULARIZED ORTHOGONAL MATCHING PURSUIT ALGORITHM

Nam H. Nguyen, Trac D. Tran

Department of Electrical and Computer Engineering
The Johns Hopkins University
{nam,trac}@jhu.edu

ABSTRACT

This paper studies a fundamental problem that arises in sparse representation and compressed sensing community: can greedy algorithms give us a stable recovery from incomplete and contaminated observations ? Using the Regularized Orthogonal Matching Pursuit (ROMP) algorithm, a modified version of Orthogonal Matching Pursuit (OMP) [1], which was recently introduced by D.Needell and R.Vershynin [2], we assert that ROMP is stable and guarantees approximate recovery of non-sparse signals, as good as the Basis Pursuit algorithm [16]. We also will set up criterions at which the algorithm halts, and the upper bounds for the reconstructed error. It will be proved in the paper that these upper bounds are proportional to the noise energy.

***Index Terms*—** Compressed Sensing, Sparse Representation, Orthogonal Matching Pursuit, Basis Pursuit, Greedy Algorithm.

1. INTRODUCTION

In the last two years, there has been numerous works on the emerging field of compressed sensing [3], [4], [5], [6], [7]. The basis idea is how to reconstruct exactly a signal from vastly incomplete observations. Mathematically, consider a problem of recovering a sparse signal $v \in R^n$ which has only few non-zero coefficients. All we know about the signal is m non-adaptive measurements x which is observed from

$$x = \Phi v \tag{1}$$

Where Φ is a $R^{m\times n}$ matrix, called sensing matrix. Of course when the measurements m is equal to length of the signal, we can easily recover the signal perfectly, provided the matrix is invertible. In the case $m \ll n$, we also wish to recover exactly the signal v from these few incomplete measurements.

The two appealing algorithmic approaches to signal recovery called basis pursuit (also named $L1$-minimization) [8], [4], [9] and greedy pursuit such as Orthogonal Matching Pursuit (OMP) [10], [9], [11], [12] have been received much attention. Using $L1$-minimization approach, Candes and Tao proved a beautiful result in their paper [13] that one can reconstruct the signal precisely by solving the linear program

$$(L_1) \qquad \min \|v\|_1 \qquad \text{subject to} \quad y = \Phi v$$

provided that the sensing matrix Φ obeys a *Restricted Isometry Condition* which is defined as follow

Definition 1. *(Restricted Isometry Condition for sensing matrix): let Φ be a measurement matrix. Φ satisfies Restricted Isometry Condition (RIC) if there exists a constant number $\delta_S \in (0,1)$ such that*

$$(1-\delta_\wedge)\|v\|_2^2 \le \|\Phi v\|_2^2 \le (1+\delta_\wedge)\|v\|_2^2 \tag{2}$$

With every T-sparse vector $v \in R^n$ and $|T| \le |\wedge|$

It was shown in [13] by the authors that if $\delta_{3\wedge} < 1/3$, then solving L_1 recovers perfectly every sparse signals with support T satisfying $|T| \le |\wedge|$.

Although $L1$-minimization has strong guarantees of exact recovery, it has disadvantages in computational cost and implementation complexity. As a result, another line of research that seem valuable to explore is OMP algorithm. This recovery scheme is especially simpler to implement and potentially faster than basis pursuit (BP).

From stochastic point of view, Tropp and Gilbert [14] analyzed the performance of OMP algorithm for a set of $m \times n$ random matrices that if pleasing four properties (admissible measurement matrices), then OMP can recover perfectly the signal v with high probability provided that number of measurements is proportional to the sparsity T of signals: $m = C \cdot |T| \cdot log(n)$. Nonetheless, their claim is only appropriate for some sets of signals which are statistically independent from rows of sensing matrix.

Taking advantage of fast implementation of Fourier transform, Rauhut [15] investigated the performance of OMP for the case of Fourier matrix. By numerous experiments, they suggested that $O(Tlog(n))$ measurements are sufficient for OMP to recover precisely the T-sparse signal v. They also mathematically explained their empirical works for the first iteration of OMP, they showed that at the first iteration, OMP can identify a correct column from the sensing matrix, provided the measurements $m \sim Tlog(n)$. Unfortunately, because of subtle stochastic dependency between columns of the matrix, it is difficult to analyze subsequent iterations of the OMP algorithm.

Recently, D.Needell and R.Vershynin [2] introduced a novel algorithm called Regularized OMP which is a variant of OMP. They showed that for any matrix with the Restricted Isometry Constant proportional to $1/\sqrt{log(T)}$, ROMP can identify exactly every T-sparse signals. This RIC is equivalent to the number of measurements $m \sim Tlog^k(n)$ where the constant k is dependent on the choice of sensing matrix, for instant, $k = 2$ if the sensing matrix is i.i.d Gaussian or Bernoulli, and $k = 5$ with partial Fourier matrix.

Unfortunately, in real scenarios, signals are usually not likely to be truly sparse, but compressible or can be compressible under some bases. It implies that coefficients of signals or transformed signals decay with power law. Furthermore, we could not assume that measurements are completely clean. In this paper, we will concentrate on the issue of using ROMP to stably reconstruct compressible signals from a set of noisy measurements. To be concise, we get the measurements of the form

$$y = \Phi v + e = x + e \tag{3}$$

Where Φ is the sensing matrix, v is an unknown compressible signal, and e is an arbitrary unknown noise vector bounded by a known energy $\|e\|_2 \leq \eta$. In the case of basis pursuit, Candes and Tao [16] stated that if sensing matrix Φ obeys a RIC with $\delta_{4\wedge} < 1/2$, then the solution of L_1^*-minimization

$$(L_1^*) \qquad \min \|v\|_1 \qquad \text{subject to} \quad \|y - \Phi v\|_2 \leq \eta$$

give us a stable result, in the sense that small changes in the observations should results in small changes of the recovery. More precisely, they showed that the solution $v^\sharp$ of L_2 is within a noise level: $\|v^\sharp - v\|_2 \leq C_{1,T} \cdot \eta + C_{2,T} \cdot \|v - v_T\|_1/\sqrt{T}$, where v_T is a vector of length n with T biggest absolute entries of v. A question is now still remained is whether ROMP also achieve stability as good as L_1^*-minimization ? Our paper will answer for this doubt.

The remainder of the paper is organized as follows, section 2 describes some notations and matrix tools that will be used throughout the paper. Section 3 will restate the ROMP algorithm of [2]. In section 4, we give a quick look at the result of D.Needell and R.Vershynin [2] when signal is exactly sparse and measurements are not polluted by noise. The two subsequence sections 5 and 6 are the main components of this paper. In section 5, we will introduce two important theorems that guarantees stable recovery of the algorithm. The proofs will be showed in section 6 . Section 7 will support our claims by numerous experiments.

2. NOTATION AND MATRIX TOOLS

In this section, we provide a few notations of sub-vectors and sub-matrices and introduce some useful operator norm formulas which were simply verified by Tropp [17]

Suppose that S is a set $\in \{1, ..., n\}$, then $|S|$ is the cardinality of set S. If a vector v has only $|S|$ non-zero entries, then we call v has support S.

Define $v_{|S}$ as a sub-vector of length $|S|$ whose entries are extracted from $|S|$ entries of vector v, and Φ_S is a sub-matrix whose columns are withdrawn from $|S|$ columns of matrix Φ. The pseudo-inverse of the matrix Φ_S is denoted by $\Phi_S^\dagger$, and it may be calculated using the formula $\Phi_S^\dagger = (\Phi_S^*\Phi_S)^{-1}\Phi_S^*$. The matrix P_S will denotes the orthogonal projection onto the span of $|S|$ atoms of the matrix Φ. This projector may be expressed using pseudo-inverse: $P_S = \Phi_S\Phi_S^\dagger$.

Define $\|A\|_{2,2}$ as a operator norm. For ease, we will simplify $\|\cdot\| = \|\cdot\|_{2,2}$. We have two useful lower norm bounds which is simply proved by Tropp [17]

Proposition 1. *For every matrix A,*

$$\min_{\substack{x \in R(A^*) \\ x \neq 0}} \frac{\|Ax\|_2}{\|x\|_2} \geq \|A^\dagger\|^{-1} \tag{4}$$

If A is invertible, this result implies

$$\min_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} \geq \|A^{-1}\|^{-1} \tag{5}$$

The symbol $R(A)$ denotes the range of matrix A.

3. ALGORITHM

ROMP is an iterative algorithm which is fundamentally based on OMP with some modifications. Let us start with the ROMP algorithm.

Algorithm

INPUT:

- *A sensing matrix $\Phi \in R^{m \times n}$*
- *The measurement vector $x \in R^m$*
- *The support T of signal v*

OUTPUT:

- *A index set $I_t \in \{1, ..., n\}$*
- *A residual vector $r_t \in R^m$*
- *A reconstructed signal $v^\sharp \in R^n$*

PROCEDURE:

1. ***Initialize**: Let the residual vector $r_t = y$, the index set $I_t = \emptyset$, and start the iteration counter with t = 1.*

2. ***Identify***: *Choose a set J of $|T|$ biggest absolute values of the observation vector $u = \Phi^* r_t$.*

3. ***Regularize***: *Divide the set J into subsets J_k which satisfies*

$$|u(i)| \leq 2 \cdot |u(j)| \quad \textit{for all } i, j \in J_k$$

and choose the subset J_0 with the maximum energy of $\|u_{|J_0}\|$

4. ***Update***: *Set $I_t = I_{t-1} \cup J_0$*
Calculate the new output approximation by solving the least square equation

$$v_t = \underset{c}{\operatorname{argmin}} \|y - \Phi_{I_t} c\|_2$$

Update the residual: $r_t = x - \Phi_{I_t} v_t$

5. ***Stopping***: *Check the stopping criterion, if not, then keep increasing $t = t + 1$*

The main difference between OMP and Regularized OMP algorithm is the identification and regularization steps. In stead of choosing only one biggest correlation between the residual and columns of the matrix at each iteration as in OMP, ROMP choose a set of $|J_0|$ coefficients from $|J|$ biggest absolute coefficients of $\Phi^* r_t$. By this way, ROMP can recover signals perfectly without going through all $|T|$ iterations. As a result, ROMP performs much faster that OMP. Numerous experiments has proved these observations.

4. RECOVERY OF SPARSE SIGNALS

In the paper of D.Needell and R.Vershynin [2], they showed that in the situations that measurements are not corrupted by noise and signals are truly sparse, then at most after $|T|$ iteration, ROMP will identify correctly $|T|$ columns of sensing matrix that corresponds to $|T|$ nonzero entries of the signal v, provided that sensing matrix Φ obey a Uniform Uncertainty Principle (UUP). The claim based on the theorem 3.1 of their paper [2]

Theorem 1. *Assume Φ satisfies the Restricted Isometry Condition with $\delta_{2T} \leq 0.03/\sqrt{log(T)}$, then at each iteration of ROMP, after the regularization step, at leat 50% of newly selected set J_0 belongs to the support of v. If we define $J_1 = J_0 \cap T$, $J2 = J_0/T$, then*

$$|J_1| \geq |J_2| \tag{6}$$

The theorem implies that after at most $|T|$ iterations, ROMP will select the set I_t which has cardinality less than $2|T|$ and covers the entire support T: $|I_t| \leq 2|T|$ and $T \in I_t$. Therefore, the recovery signal can be easily obtained by solving $x = \Phi_{I_t} v$. This linear system has more observations than unknown and the matrix Φ_{I_t} is invertible since the Restricted Isometry Constant $\delta_{2T} < 1$. As a result, v is retrieved exactly: $v = \Phi_{I_t}^{\dagger} x$

5. RECOVERY OF COMPRESSIBLE SIGNALS FROM NOISY OBSERVATIONS

Naturally, we could not expect to recover exact a signal when measurements are polluted by noise. What we want to consider is how good the recovery signal will be ? The answer is dependent on the choices of appropriate stopping criterions for the algorithm which is used to reconstruct the signal. Obviously, a good stopping criterion should base on the noise energy η.

At first, for simplicity, we set up the problem for the case signals are truly sparse. Afterward, theorem 3 will answer to the question of compressible signals.

5.1. With sparse signals

Theorem 2. *If Φ is a measurement matrix with restricted isometry constant $\delta_{2T} \in (0, C_T/(2C_T+1))$. After iteration t, the Regularized Orthogonal Matching Pursuit algorithm halts if:*

$$\|r_t\|_2 < K_{1,T} \times \eta \tag{7}$$

Where $K_{1,T}$ is defined by:

$$K_{1,T} = \frac{C_T(1-\delta_{2T})}{C_T(1-2\delta_{2T}) - \delta_{2T}} \quad \textit{with} \quad C_T = \frac{1}{5.6\sqrt{log|T|}}$$

If the algorithm terminates at the end of iteration t, then we will conclude that:

- *The algorithm has selected at least t indices from the optimal set T.*
- *The solution $v^\sharp$ of the algorithm will obey*

$$\|v - v^\sharp\|_2 < K_{2,T} \times \eta \tag{8}$$

Where $K_{2,T} = \sqrt{2} \cdot (K_{1,T} + 1)$

In particular, if until number entries of the set I_t start greater than $2|T|$, formula (7) is still not satisfied, then the algorithm halts when $|I_t| > 2|T|$ and $v^\sharp$ obeys

$$\|v - v^\sharp\|_2 < K_{3,T} \times \eta \tag{9}$$

Where $K_{3,T} = \sqrt{3/2} \cdot \{(2+\delta)(K_{1,T} + K_{1,T}/C_T) + 2\}$

From the theorem, we observe that the smaller energy the noise is, the better recovery signals the algorithm can find, and as $\eta = 0$, ROMP recover exact signals. It is the case that D.Needle and R.Vershynin showed in their paper.

At the first glance, it seems that the stopping criterion is complicated. However, the value $K_{1,T}$ can be simplified. Define the restricted isometry constant as:

$$\delta_{2T} = \frac{C_T}{a(2C_T+1)} = \frac{1}{a(2+5.6\sqrt{log|T|})} \quad \textit{with} \quad a > 1 \tag{10}$$

$K_{1,T}$ is then represented by

$$K_{1,T} = 1 + \frac{C_T + 1}{(2C_T+1)(a-1)} < 1 + \frac{1}{a-1} = \frac{a}{a-1}$$

Corollary 1. *Suppose the restricted isometry constant of matrix Φ has the form as in (10). If after iteration t*

$$\|r_t\|_2 < \frac{a \cdot \eta}{(a-1)} \tag{11}$$

Then the algorithm selects at least t indices from the set T and terminates at the end of iteration t. Also, the solution $v^\sharp$ of the algorithm will obey

$$\|v - v^\sharp\|_2 < \frac{(2a-1)\sqrt{2}}{a-1} \cdot \eta \tag{12}$$

The corollary 1 obtains easily from an obvious result: $\delta_{2T} < 1/2$. Furthermore, suppose the factor $a = 2$, then the recovery error of the algorithm is only $3\sqrt{2} \cdot \eta$. This result for ROMP is quite competitive to $L1$-minimization approach [**?**], yet the price for outperforming is a much smaller restricted isometry constant.

Corollary 2. *Suppose formula (11) is still not assured when number entries of the set I_t start greater than $2|T|$, then the algorithm halts when $|I_t| > 2|T|$ and the recovery error obeys*

$$\|v - v^\sharp\|_2 < \sqrt{\frac{3}{2}} \cdot (7 + 28\sqrt{log|T|}) \cdot \frac{a}{a-1} \cdot \eta \tag{13}$$

The upper bound of error approximation in inequality (13) seem to be quite big, and is proportional to $\sqrt{log(T)}$. However, our numerous experiments demonstrate that in practice, this bound is much smaller, and it is equivalent to the bound in (12). In the best situations, when the algorithm selected all T indices in the support of v as $|I_t|$ start greater than $2|T|$. Then the recovery vector $v^\sharp$ will be the projection of measurement y onto the span of $|I_t|$ columns of sensing matrix Φ.

$$v^\sharp = \Phi^\dagger_{I_t} y = \Phi^\dagger_{I_t}(\Phi_{I_t} v + e) = v + \Phi^\dagger_{I_t} e \tag{14}$$

Recovery error will be optimally bounded using inequalities in corollary 4 (in section 6)

$$\begin{aligned} \|v^\sharp - v\|_2 = \|\Phi^\dagger_{I_t} e\|_2 &\leq \|\Phi^*_{I_t}\Phi_{I_t}\|_{2,2}\|\Phi^*_{I_t}\|_{2,2}\|e\|_2 \\ &\leq \frac{\sqrt{1+\delta_{2T}}}{1-\delta_{2T}} \cdot \eta \leq 2\sqrt{2}\eta \end{aligned} \tag{15}$$

The stability of the ROMP algorithm implies that ROMP can identify every sufficient large coefficients of vector v. When the noise energy is small enough, all support T will be identified, and the approximation error is as (15). Corollary 3 will develop such a condition on the noise energy.

Corollary 3. *ROMP will select all the indices of set T if the noise level is bounded by:*

$$\eta \leq \frac{C_T(1-2\delta) - \delta}{(1-\delta)\sqrt{1+\delta}(C_T+1)} \times \min_{i \in T} |v(i)| \tag{16}$$

Value δ should be constant with every sparsity T of signal, and should be as close to 1 as good since the number of measurements are proportional to reciprocal of the Restricted Isometry Constant. Figure 1 describes the ratio $C_T/(2C_T+1)$ with different support T, we realize that as $|T|$ goes to infinity, this ratio is almost constant, then δ is not dependent on T. However, δ is quite small as the support T is big. By numerous experiments, we conjecture that $\sqrt{log(T)}$ in the denominator of δ_{2T} can be disappeared.

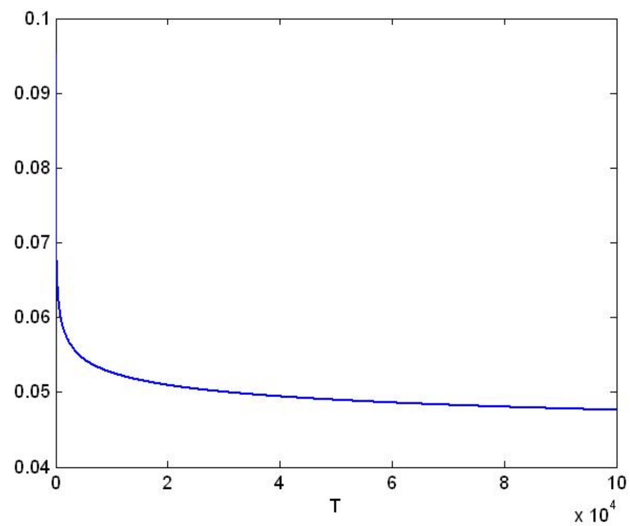


Fig. 1. Ratio $C_T/(2C_T+1)$ with different supports T's

Suppose $a = 4$, we demonstrate values $K_{2,T}$ and $K_{3,T}$ with different supports T on figure 2. From the figure, we see that as T goes to infinity, $K_{3,T}$ comes to 67. So that when $\delta_{2T} = 0.012$, the maximum recovery error is $67 \cdot \eta$.

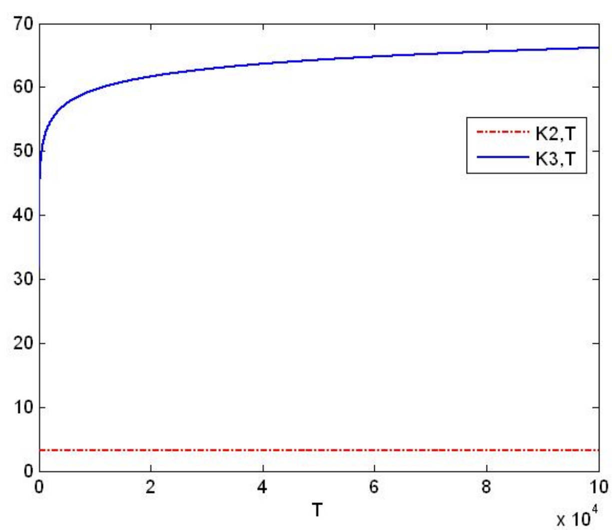


Fig. 2. Values of $K_{2,T}$ and $K_{3,T}$ with different T's

5.2. With compressible signals

If not only the measurements are corrupted by noise, the signal is compressible. the ROMP algorithm can also have stable recovery. Suppose v be a signal whose coefficients follows the power law, that is if we arrange entries of v in decreasing order of magnitude $|v|_{(1)} \geq |v|_{(2)} \geq ... \geq |v|_{(n)}$, we say

$$|v|_{(i)} \leq R \cdot i^{-1/p}$$

Where p controls the decay rate of coefficients.
Suppose v_T is a vector of T biggest absolute entries of v, then

$$\|v - v_T\|_2 \leq K_p \cdot R \cdot |T|^{1/2-1/p}$$

Here, K_p is a constant that only depends on p.

Let Φ_{opt} be a subset of columns of measurement matrix which is corresponding to $2T$ nonzero entries of the original signal v, x_{opt} be the optimal approximation over Φ_{opt}. We have the theorem 3

Theorem 3. *If Φ is a measurement matrix with restricted isometry constant $\delta_{4T} \in (0, C_{2T}/(2C_{2T}+1))$. After iteration t, the Regularized Orthogonal Matching Pursuit algorithm halts if:*

$$\|r_t\|_2 < K_{1,2T} \times (\|y - x_{opt}\|_2 + \eta) \quad (17)$$

If the algorithm terminates at the end of iteration t, then we will conclude that:

- *The algorithm has selected at least t indices from the optimal set T.*
- *The solution $v^\sharp$ of the algorithm will obey*

$$\|v - v^\sharp\|_2 < K_{2,2T} \cdot \eta + \sqrt{1+\delta}(K_{2,2T}+1)\frac{\|v - v_T\|_1}{\sqrt{|T|}} \quad (18)$$

In particular, if until number entries of the set I_t start greater than $4|T|$, formula (17) is still not satisfied, then the algorithm halts when $|I_t| > 4|T|$ and $v^\sharp$ obeys

$$\|v - v^\sharp\|_2 < K_{3,2T} \cdot \eta + \sqrt{1+\delta}(K_{3,2T}+1)\frac{\|v - v_T\|_1}{\sqrt{|T|}} \quad (19)$$

Where $K_{1,2T}$, $K_{2,2T}$ and $K_{3,2T}$ are defined as in theorem 2

The theorem states that ROMP can recover stably $|T|$-largest entries of the unknown signal v. This result can be comparable with $L1$-minimization [16].

We can derive results that are similar to corollary 1 and 2 from theorem 3 with one additional term. For instant, if $a = 4$ and T goes to infinity, δ_{4T} will comes to 0.0115, $K_{3,2T}$ reach 68, and the maximum recovery error will be

$$\|v - v^\sharp\|_2 < 68 \cdot \eta + 69.5 \cdot \frac{\|v - v_T\|_1}{\sqrt{T}}$$

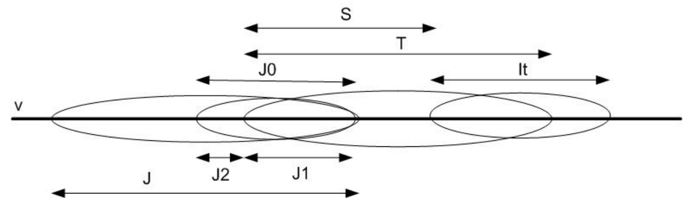


Fig. 3. Simple relation between subsets

6. PROOFS OF THEOREMS 2 AND 3

6.1. Theorem 2

Before proving the theorem, let us show some simple results that can imply directly from the definition of RIC.

Corollary 4. *Let Φ_S and Φ_J be two subset of columns of sensing matrix Φ where S and J are two disjoint sets in $\{1, ..., n\}$ and $|S|, |J| \leq |T|$, then*

$$\|\Phi^*_{S\cup J}\Phi_{S\cup J} - I\| \leq \delta_{2T}$$

$$\|\Phi^*_J \Phi_S\| \leq \delta_{2T}$$

$$\|\Phi^\dagger_J \Phi_S\| \leq \frac{\delta_{2T}}{1-\delta_{2T}}$$

Those results are quite straight forward, we leave the proof for appendix. The theorem will be more clarified by having the lemma 1 at hand.

Lemma 1. *After iteration t, the ROMP algorithm still identify at least half of the correct subset T if*

$$\|\Phi^*_{J_2} r_t\|_2 \geq \frac{C_T(1-\delta)\sqrt{1+\delta}}{C_T(1-2\delta)-\delta} \times \eta \quad (20)$$

Before proving the lemma 1, let us restate the lemma 3.7 in the paper of N.Deanna and R.Vershynin [2] which will be used for our proof.

Lemma 2. *Let z be a vector $\in R^d$. Then there exists a subset $A \in \{1, 2, ..., k\}$ such that entries of z in A obey*

$$|z(i)| \leq 2 \cdot |z(j)| \qquad \forall i, j \in A$$

And the total energy is satisfied

$$\|z_{|A}\|_2 \geq \frac{1}{2.5\sqrt{log(d)}}\|z\|_2$$

We are now going to prove lemma 1

Proof. From now on, for simplicity, we write down δ and consider it as δ_{2T}. Define $S = T/I_t$ (figure 3 describe roughly the relation between subsets that we are considering). At the iteration t, the residual r_t is computed by

$$\begin{aligned} r_t &= y - P_{I_t}y = x + e - P_{I_t}x - P_{I_t}e \\ &= \Phi_S v_{|S} - P_{I_t}\Phi_S v_{|S} + e - P_{I_t}e \end{aligned} \quad (21)$$

We assume that at iteration t, more than half of subset J_0 are outside the correct support T, it indicates that $|J_2| > |J_1|$. Using this hypothesis, we find the upper bound of $\|\Phi_{J_2}^* r_t\|_2$. Lemma 1 will be stated from the negation of this upper bound.

From the regularization step, since $|u(i)| \leq 2 \cdot |u(j)|$ with $\forall i, j \in J_0$, the assumption $|J_2| > |J_1|$ conduct a lower bound of energy of $u_{|J_2}$

$$\|u_{|J_2}\|_2 \geq \frac{1}{\sqrt{5}} \|u_{|J_0}\|_2 \tag{22}$$

Furthermore, we also obtain a lower bound for energy of $u_{|J_0}$ from lemma 2

$$\|u_{|J_0}\|_2 \geq \frac{1}{2.5\sqrt{log|T|}} \|u_{|J}\|_2 \tag{23}$$

We now start finding a lower bound of $\|u_J\|_2$. Since J is a subset corresponding to $|J|$ biggest coordinates of vector u and from (21), we get

$$\begin{aligned} \|u_{|J}\|_2 &\geq \|u_{|S}\|_2 = \|\Phi_S^* r_t\|_2 \\ &= \|\Phi_S^*(I - P_{I_t})\Phi_S v_{|S} + \Phi_S^*(I - P_{I_t})e\|_2 \\ &\geq \|\Phi_S^*(I - P_{I_t})\Phi_S v_{|S}\|_2 - \|\Phi_S^*\| \|I - P_{I_t}\| \|e\|_2 \\ &\geq \|Q^{-1}\|^{-1} \|v_{|S}\|_2 - \eta\sqrt{1+\delta} \end{aligned} \tag{24}$$

Where $Q = \Phi_S^*(I - P_{I_t})\Phi_S$ and the last inequality of (24) is followed from (5) since Q is invertible. Analyzing $\|Q^{-1}\|$

$$\begin{aligned} \|Q^{-1}\| &= \|[I - (I - Q)]^{-1}\| = \|I + (I - Q) + ...\| \\ &\leq 1 + \|I - Q\| + \|I - Q\|^2 + ... \end{aligned}$$

And therefore,

$$\|Q^{-1}\|^{-1} \geq \frac{1}{1 + \|I - Q\| + \|I - Q\|^2 + ...} = 1 - \|I - Q\|$$

Here

$$\begin{aligned} \|I - Q\| &= \|I - \Phi_S^*(I - P_{I_t})\Phi_S\| \\ &\leq \|I - \Phi_S^*\Phi_S\| + \|\Phi_S^* P_{I_t} \Phi_S\| \\ &\leq \|I - \Phi_S^*\Phi_S\| + \|\Phi_S^*\Phi_{I_t}\| \|\Phi_{I_t}^\dagger \Phi_S\| \\ &\leq \delta + \delta \cdot \frac{\delta}{1-\delta} = \frac{\delta}{1-\delta} \end{aligned}$$

Where the last inequality are followed from corollary 4.
Substitute above inequalities to (24), we get

$$\begin{aligned} \|u_{|J}\|_2 &\geq (1 - \frac{\delta}{1-\delta})\|v_{|S}\|_2 - \eta\sqrt{1+\delta} \\ &= \frac{1-2\delta}{1-\delta}\|v_{|S}\|_2 - \eta\sqrt{1+\delta} \end{aligned} \tag{25}$$

The upper bound $\|u_{|J_2}\|$ is developed as follows

$$\begin{aligned} &\|u_{|J_2}\|_2 = \|\Phi_{J_2}^* r_t\|_2 = \|\Phi_{J_2}^*[(I - P_{I_t})\Phi_S v_{|S} + (I - P_{I_t})e]\|_2 \\ &\leq \|\Phi_{J_2}^*(I - P_{I_t})\Phi_S v_{|S}\|_2 + \|\Phi_{J_2}^*(I - P_{I_t})e]\|_2 \\ &\leq \|\Phi_{J_2}^*\Phi_S v_{|S}\|_2 + \|\Phi_{J_2}^* P_{I_t}\Phi_S v_{|S}\|_2 + \|\Phi_{J_2}^*\| \|I - P_{I_t}\| \|e\|_2 \end{aligned} \tag{26}$$

Since entries of vector $v_{|J_2}$ are zeros, the first term at the left-hand side of the last formula of (26) can be represented

$$\begin{aligned} \|\Phi_{J_2}^*\Phi_S v_{|S}\|_2 &= \|\Phi_{J_2}^*(\Phi_S v_{|S} + \Phi_{I_2} v_{|J_2}) - v_{|J_2}\|_2 \\ &= \|\Phi_{J_2}^*\Phi_{J_2\cup S} v_{|J_2\cup S} - v_{|J_2}\|_2 \end{aligned}$$

Since J_2 and S are two disjoint subsets, so

$$\begin{aligned} \|\Phi_{J_2}^*\Phi_S v_{|S}\|_2 &\leq \|\Phi_{J_2\cup S}^*\Phi_{J_2\cup S} v_{|J_2\cup S} - v_{|J_2\cup S}\|_2 \\ &= \|[\Phi_{J_2\cup S}^*\Phi_{J_2\cup S} - I] v_{|J_2\cup S}\|_2 \\ &\leq \delta\|v_{|J_2\cup S}\|_2 = \delta\|v_{|S}\|_2 \end{aligned}$$

Terms 2^{nd} and 3^{rd} of the last inequality of (26) can be simply bounded based on the corollary 4

$$\begin{aligned} \|\Phi_{J_2}^* P_{I_t}\Phi_S v_{|S}\|_2 &\leq \|\Phi_{J_2}^*\Phi_{I_t}\| \|\Phi_{I_t}^\dagger \Phi_S\| \|v_{|S}\|_2 \\ &= \delta \cdot \frac{\delta}{1-\delta}\|v_{|S}\|_2 \end{aligned}$$

And

$$\|\Phi_{J_2}^*\| \|I - P_{I_t}\| \|e\|_2 \leq \eta\sqrt{1+\delta}$$

Thus, the upper bound of $\|u_{|J_2}\|_2$ will be

$$\|u_{|J_2}\|_2 \leq \frac{\delta}{1-\delta}\|v_{|S}\|_2 + \eta\sqrt{1+\delta} \tag{27}$$

Combining (22), (23), (25) and (27), the condition $|J_2| > |J_1|$ lead to

$$\frac{\delta}{1-\delta}\|v_{|S}\|_2 + \eta\sqrt{1+\delta} > C_T\frac{1-2\delta}{1-\delta}\|v_{|S}\|_2 - C_T\eta\sqrt{1+\delta}$$

Where C_T is defined by:

$$C_T = \frac{1}{5.6\sqrt{log|T|}}$$

It is concluded that

$$\|v_{|S}\|_2 < \frac{(1-\delta)\sqrt{1+\delta}(C_T + 1)}{C_T(1-2\delta) - \delta} \times \eta \tag{28}$$

From (28), we can understand the correlation between power noise and energy of entries of v that are not being selected by the algorithm at iterations $1 \to t$. If the power noise is small, the energy of unidentified entries of v is small. It implies that the recovery signal $v^\sharp$ is a good approximation of v.

It is also noted that the sufficient condition become worthless when the denominator is negative, so $C_T(1-2\delta)-\delta$ must be strictly greater than zero, which yields $\delta < C_T/(2C_T+1)$. This is a very essential requirement of restricted isometry constant δ to guarantee the theorem meaningful.

In addition, following from inequality (27)

$$\|v_{|S}\|_2 > \frac{1-\delta}{\delta}(\|\Phi_{J_2}^* r_t\|_2 - \eta\sqrt{1+\delta})$$

So, in order for the inequality (28) to be satisfied, we need

$$\frac{1-\delta}{\delta}(\|\Phi^*_{J_2}r_t\|_2 - \eta\sqrt{1+\delta}) < \frac{(1-\delta)\sqrt{1+\delta}(C_T+1)}{C_T(1-2\delta)-\delta}\eta$$

Thus

$$\|\Phi^*_{J_2}r_t\|_2 < \frac{C_T(1-\delta)\sqrt{1+\delta}}{C_T(1-2\delta)-\delta} \times \eta$$

□

We are now ready to complete the theorem 2

Proof of theorem 2 :

From lemma 1, the algorithm halts if the inequality (20) fails. However, this condition does not seem practical since the subset J_2 is unknown at each iteration. We investigate three possible cases.

Case 1: If at a iteration t

$$\|r_t\|_2 < \frac{C_T(1-\delta)}{C_T(1-2\delta)-\delta} \times \eta \tag{29}$$

Then, since $\|\Phi^*_{J_2}r_t\|_2 \leq \sqrt{1+\delta}\|r_t\|_2$, the inequality (20) fails. Therefore, we set one criterion that the algorithm will stop at iteration t if (29) occurs.
Since the new signal estimate $v^\sharp$ is obtained by solving the least-square problem $v^\sharp = \mathrm{argmin}_c \|y - \Phi_{I_t}c\|_2$, so the residual

$$\begin{aligned}\|r_t\|_2 &= \|y - \Phi_{I_t}v^\sharp_{I_t}\|_2 = \|\Phi_{I_t}v^\sharp_{I_t} - \Phi_T v_T - e\|_2 \\ &\geq \|\Phi_{T\cup I_t}v^\sharp_{T\cup I_t} - \Phi_{T\cup I_t}v_{T\cup I_t}\|_2 - \|e\|_2 \\ &\geq \|\Phi^\dagger_{T\cup I_t}\|^{-1}\|v^\sharp_{T\cup I_t} - v_{T\cup I_t}\|_2 - \eta\end{aligned}$$

Where the last inequality of the above formula come from (4), and $\|\Phi^\dagger_{T\cup I_t}\|$ is the reciprocal of minimum singular value of matrix $\Phi_{T\cup I_t}$, which is smaller than $1/\sqrt{1-\delta}$. This yields

$$\|r_t\|_2 \geq \sqrt{1-\delta}\|v^\sharp_{T\cup I_t} - v_{T\cup I_t}\|_2 - \eta = \sqrt{1-\delta}\|v - v^\sharp\|_2 - \eta$$

Substituting the upper bound of $\|r_t\|_2$ in (29) to the above formula and note that $\delta < 1/2$, we obtain

$$\|v - v^\sharp\|_2 < \sqrt{2}\cdot\{\frac{C_T(1-\delta)}{C_T(1-2\delta)-\delta} + 1\} \times \eta$$

Case 2: If the inequality (20) still hold at each iteration until at a iteration t, the cardinality of set I_t start greater than $2|T|$. Thus we set another stopping criterion to be $|I_t| > 2|T|$. In this case, the lemma 1 says at every iterations, ROMP identify at least haft of correct support of T until $|I_t| > 2|T|$. As a result, the support T must belong to I_t: $T \in I_t$, and the recovery error is optimal and computed as in (15).

$$\|v^\sharp - v\|_2 \leq 2\sqrt{2}\eta$$

Case 3: If after t iterations t, $|I_t|$ start greater than $2|T|$, but (20) fails and (29) is not satisfied. In this case, the selected set I_t of the algorithm may not cover the support of signal as in case 2 (since lemma 1 may fail).

Now, from the proof of lemma 1, (20) fails lead to inequality (28), we will develop the upper bound for the recovery signal

$$\begin{aligned}v^\sharp_{|I_t} = \Phi^\dagger_{I_t}y &= \Phi^\dagger_{I_t}(\Phi_S v_{|S} + \Phi_{T\cap I_t}v_{|T\cap I_t} + e) \\ &= \Phi^\dagger_{I_t}(\Phi_S v_{|S} + \Phi_{I_t}v_{|I_t} + e)\end{aligned}$$

Therefore,

$$\begin{aligned}\|v^\sharp_{|I_t} - v_{|I_t}\|_2 &= \|\Phi^\dagger_{I_t}(\Phi_S v_{|S} + e)\|_2 \\ &\leq \|\Phi^\dagger_{I_t}\Phi_S\|\|v_{|S}\|_2 + \|\Phi^\dagger_{I_t}\|\|e\|_2 \\ &\leq (1+\delta)\|v_{|S}\|_2 + \eta\frac{\sqrt{1+\delta}}{1-\delta}\end{aligned}$$

The recovery bound is then followed

$$\begin{aligned}\|v^\sharp - v\|_2 &= \|v^\sharp_{|T\cup I_t} - v_{|T\cup I_t}\|_2 \\ &\leq \|v^\sharp_{|I_t} - v_{|I_t}\|_2 + \|v_{|S}\|_2 \\ &\leq (2+\delta)\|v_{|S}\|_2 + \eta\frac{\sqrt{1+\delta}}{1-\delta}\end{aligned} \tag{30}$$

Substitute (28) for (30) and using $\delta < 1/2$, we get the inequality (19) of theorem 2, and the theorem is now proved.

The corollary 3 can be derived from the claim of inequality (28) in the proof of lemma 1. When (28) fails at every iterations until $|I_t| > 2|T|$, then from lemma 1, ROMP algorithm can identify all support T of signal v.

6.2. Theorem 3

Proof. Let v_{2T} be a vector of length n with support $2T$ taken from $2|T|$ biggest (in magnitude) entries of v, and $v_{(2T)^c}$ be a vector of length n: $v_{(2T)^c} = v - v_{2T}$. We have

$$\begin{aligned}y = \Phi v + e &= \Phi v_{2T} + \Phi v_{(2T)^c} + e \\ &= x_{2T} + x_{(2T)^c} + e\end{aligned} \tag{31}$$

Where $(2T)^c$ is denoted the complement set of $2T$. We can consider $\Phi v_{(2T)^c}$ as an another noise, v_{2T} as truly $2T$-sparse signal that need to be reconstructed. So the problem of stably recovering compressible signal become the problem of recovering truly sparse signal. Following the lines of proof as the section 6.1, one of the criterion to halt the algorithm is

$$\begin{aligned}\|r_t\|_2 &< K_{1,2T}\cdot(\|x_{(2T)^c}\|_2 + \|e\|_2) \\ &= K_{1,2T}\cdot(\|y - x_{2T}\|_2 + \|e\|_2)\end{aligned} \tag{32}$$

Note that $K_{1,2T}$ and the following $K_{2,2T}$, $K_{3,2T}$ are defined as in theorem 2.

We can have another stopping criteria. From the definition of x_{opt} in section 5.2

$$x_{opt} = \Phi_{opt} \cdot argmin_c \|y - \Phi_{opt} c\|_2$$

In the words, x_{opt} is the projection of y onto the space spanned by Φ_{opt}. So,

$$\|y - x_{opt}\|_2 \le \|y - x_{2T}\|_2$$

Then we can stop the ROMP algorithm if

$$\|r_t\|_2 < K_{1,2T} \times (\|y - x_{opt}\|_2 + \eta) \quad (33)$$

As a result, the solution $v^\sharp$ of ROMP will obey

$$\begin{aligned} \|v_{2T} - v^\sharp\|_2 &< K_{2,2T} \cdot (\|y - x_{2T}\|_2 + \eta) \\ &= K_{2,2T} \cdot (\|\Phi(v - v_{2T})\|_2 + \eta) \end{aligned}$$

Applying the triangular inequality: $\|v - v^\sharp\|_2 \le \|v_{2T} - v^\sharp\|_2 + \|v - v_{2T}\|_2$ to the above formula, we get the bound of recovery error

$$\|v - v^\sharp\|_2 < K_{2,2T} \cdot \eta + (K_{2,2T} + 1) \cdot \|\Phi(v - v_{2T})\|_2 \quad (34)$$

If (33) is not satisfied when $|I_t|$ start greater than $4|T|$, then we halt the algorithm when $|I_t| > 4|T|$. In this case, the recovery error will be derived as the formula (19) of theorem 2

$$\|v - v^\sharp\|_2 < K_{3,2T} \cdot \eta + (K_{3,2T} + 1) \cdot \|\Phi(v - v_{2T})\|_2 \quad (35)$$

The theorem 3 will be completed by applying the following lemma 3 with $a = v - v_{2T}$ to inequalities (34) and (35).

Lemma 3. *Suppose the sensing matrix Φ has a Restricted Isometry Constant $\delta_{4T} \in (0, 1/2)$, then with any vector $a \in R^n$*

$$\|\Phi a\|_2 < \sqrt{1+\delta} \cdot (\|a_{|T}\|_2 + \frac{\|a\|_1}{\sqrt{T}})$$

Where $a_{|T}$ is $|T|$ biggest absolute entries of a.

Proof. We begin by partitioning the vector a into vectors a_{T_1}, a_{T_2}, ... a_{T_l} in decreasing order of magnitude (except probably a_{T_l}). Subsets T_1, T_2, ..., T_l with length $|T|$ are chosen such that they are all disjointed. We do this because the restricted isometry condition give us control over the small set of columns of matrix Φ.

From the definition of T_i ($i = 1, ..., l$), for each $p \in T_{i-1}$, $q \in T_i$, we have $a(q) \le a(p)$. Thus

$$|a(q)| \le \|a_{T_{i-1}}\|_1 / |T|$$

Then

$$\|a_{T_i}\|_2 \le \|a_{T_{i-1}}\|_1 / \sqrt{|T|}$$

Therefore

$$\begin{aligned} \|\Phi a\|_2 &= \|\sum_{i=1}^{l} \Phi_{T_i} a_{T_i}\|_2 \le \sum_{i=1}^{l} \|\Phi_{T_i} a_{T_i}\|_2 \\ &\le \sum_{i=1}^{l} \sqrt{1+\delta} \|a_{T_i}\|_2 \\ &< \sqrt{1+\delta} \cdot (\|a_{T_1}\|_2 + \sum_{i=2}^{l} \|a_{T_i}\|_2) \\ &< \sqrt{1+\delta} \cdot (\|a_{T_1}\|_2 + \sum_{i=2}^{l} \frac{\|a_{T_{i-1}}\|_1}{\sqrt{T}}) \\ &\le \sqrt{1+\delta} \cdot (\|a_T\|_2 + \frac{\|a\|_1}{\sqrt{T}}) \end{aligned}$$

□

Applying the lemma 3 with $a = v - v_{2T}$, we have

$$\|\Phi(v - v_{2T})\|_2 < \sqrt{1+\delta} \cdot (\|(v - v_{2T})_{|T}\|_2 + \frac{\|v - v_{2T}\|_1}{\sqrt{T}})$$

Since each entry in the subset T of vector $(v - v_{2T})$ is smaller than every entry in the subset T of vector $(v - v_T)$: $|(v - v_{2T})_{|T}(q)| \le |(v - v_T)_{|T}(p)|$ with $q \in (v - v_{2T})_{|T}$, $p \in (v - v_T)_{|T}$, thus

$$|(v - v_{2T})_{|T}(q)| \le \|(v - v_T)_{|T}\|_1 / |T|$$

Then

$$\|(v - v_{2T})_{|T}\|_2 \le \|(v - v_T)_{|T}\|_1 / \sqrt{|T|}$$

So that

$$\begin{aligned} \|\Phi(v - v_{2T})\|_2 &< \sqrt{1+\delta} \cdot (\frac{\|(v - v_T)_{|T}\|_1}{\sqrt{|T|}} + \frac{\|v - v_{2T}\|_1}{\sqrt{|T|}}) \\ &= \sqrt{1+\delta} \cdot \|v - v_T\|_1 / \sqrt{|T|} \end{aligned}$$

The theorem 3 is now proved. □

7. NUMERICAL EXPERIMENTS

In this section, we will illuminate the theorems by some demonstrations. Our simulations say that, in practice, the theorem bounds are quite optimistic, and those can be even reduced by a factor of $\sqrt{log|T|}$, that is what we conjectured.

We first do experiments with a series of truly sparse signals. In each experiment, a signal of length 1024 with different sparsities is sensed by a Gaussian matrix of size 300×1024. The measurements are then contaminated by white Gaussian noise: $y = \Phi v + e$, with $e \sim N(0, \sigma^2)$. The variance σ will be changed during the experiments in order to change the noise level.

Figure 4 shows the results of recovery errors at different noise levels and different sparsity $|T|$. The setup of our experiments is described as follows. At each noise level, we perform 1000 independent trials. At a given trial, the signal is generated by choosing $|T|$ coefficients uniformly at random, then set each coefficient a value $+1$ or -1 with equal probability. An example of such a signal is shown in figure 6a. A 300×1024 Gaussian matrix is generated and fixed during experiments. ROMP will reconstruct the signals from 300 noisy measurements. The average recovery errors over 1000 trials is lined out in figure 4

Stopping criterions of the algorithm is setup as in the theorem. It bases on two conditions: when r_t is smaller than $K_{1,T} \cdot \eta$ or when $|I_t|$ is strictly greater than $2|T|$. Suppose the constant $a = 3$, it is equivalent to say the restricted isometry constant has value $\delta_{2T} = 0.028$. Then considering again the formula (11), we have one stopping criterion: $\|r_t\|_2 < 1.5\eta$, and the smallest theorem constant bound will happen to be 3.53. This result seems to be reasonable comparing to the simulations.

We show one demonstration of effectiveness of the algorithm in figure 6a, ROMP is used to recover a 50-sparse signal from 300 noisy measurements with noise level $\sigma = 0.05$ (noise energy $\eta = 0.71$). Only after 3 iterations, the algorithm stops as $\|r_t\|_2 < 1.5\eta$, and the reconstructed vector has error approximation $\|v^* - v\|_2 = 1.57$, which is roughly twice the noise energy η.

We secondly run experiments with compressible signals with decay factor 7/9. Signals of length 1024 are generated from the equation

$$u_{sort}(n) = 5 \cdot n^{-8/9}$$

Those signals are then randomly permuted and multiplied with a uniformly random sign sequence. Following the same setup as the first case, with instead of sparse signals, we use compressible signals with different $|T|$ largest absolute coefficients. The recovery error with different noise levels are showed in figure 5. This result one again assert the stability of ROMP algorithm.

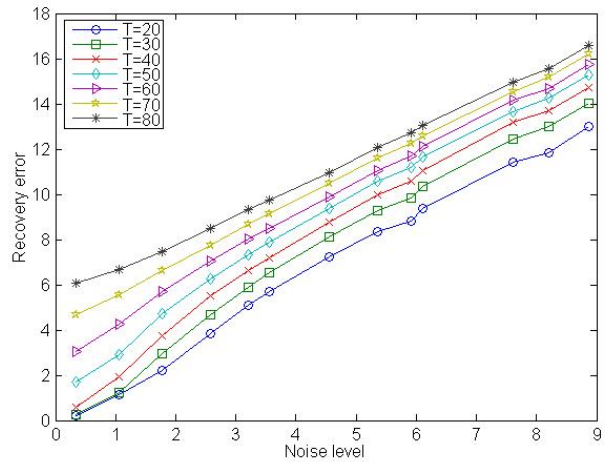


Fig. 4. Recovery error of truly T-sparse signals with different noise levels

8. APPENDIX

We are now considering the proof of corollary 4.

PART 1:

$$\begin{aligned}\|\Phi_{S\cup J}^*\Phi_{S\cup J} - I\| &= \max_{\|x\|=1} x^*(\Phi_{S\cup J}^*\Phi_{S\cup J} - I)x \\ &= \max_{\|x\|=1} x^*\Phi_{S\cup J}^*\Phi_{S\cup J}x - x^*x \\ &= \|\Phi_{S\cup J}^*\Phi_{S\cup J}\| - 1 \\ &\leq (1+\delta_{2T}) - 1 = \delta_{2T}\end{aligned}$$

PART 2:

$$\begin{aligned}&\|\Phi_{S\cup J}^*\Phi_{S\cup J} - I\| = \max_{\|x\|=1} \|(\Phi_{S\cup J}^*\Phi_{S\cup J} - I)x\| \\ &\geq \max_{\|y\|=1} \left\| \begin{pmatrix} \Phi_S^*\Phi_S - I & \Phi_S^*\Phi_J \\ \Phi_J^*\Phi_S & \Phi_J^*\Phi_J - I \end{pmatrix} \begin{pmatrix} y \\ 0 \end{pmatrix} \right\| \\ &= \max_{\|y\|=1} \left\| \begin{pmatrix} \Phi_S^*\Phi_S - I \\ \Phi_J^*\Phi_S \end{pmatrix} y \right\| \\ &\geq \max_{\|y\|=1} \|\Phi_J^*\Phi_S y\| = \|\Phi_J^*\Phi_S\|\end{aligned}$$

Therefore, from part 1

$$\|\Phi_J^*\Phi_S\| \leq \|\Phi_{S\cup J}^*\Phi_{S\cup J} - I\| \leq \delta_{2T}$$

PART 3: The statement is easily followed from part 2 and definition of RIC

$$\|\Phi_J^\dagger\Phi_S\| \leq \|(\Phi_J^*\Phi_J)^{-1}\|\|\Phi_J^*\Phi_S\| \leq \frac{\delta_{2T}}{1-\delta_{2T}}$$

9. REFERENCES

[1] Y.C. Pati, R. Rezaiifar, and P.S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," *in Proc. 27th Annu. Asilomar Conf. Signals, Systems and Computers*, Nov. 1993.

[2] D. Needell and R. Vershynin, "Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit," Submitted, 2007.

[3] E.J. Candes, J.Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from

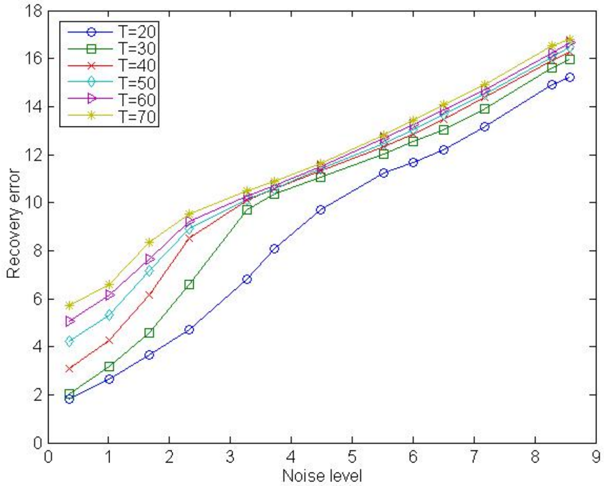


Fig. 5. Recovery error of compressible signs with different noise levels

highly incomplete frequency information," *IEEE Trans. Info. Theory*, vol. 52, pp. 489–509, 2006.

[4] E.J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies," *IEEE Trans. Info. Theory*, vol. 52, pp. 5406–5425, 2006.

[5] D.L. Donoho, "Comresses sensing," *IEEE Trans. Info. Theory*, vol. 52, pp. 1289–1306, 2006.

[6] M. Rudelson and R. Vershynin, "On sparse reconstruction from fourier and gaussian measurements," to appear.

[7] "Compressed sensing website," http://wwww.dsp.ece.rice.edu/cs.

[8] D.L. Donoho and X. Huo, "Uncertainty principke and ideal atomic decomposition," *IEEE Trans. Info. Theory*, vol. 47, pp. 2845–2862, 2001.

[9] D.L. Donoho, M. Elad, and V.N. Temlyakov, "Stable recovery of sparse overcomplete representations in the present of noise," *IEEE Trans. Info. Theory*, vol. 52, pp. 6–18, 2006.

[10] J.A. Tropp and A.C. Gilbert, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Info. Theory.*, vol. 50, pp. 2231–2242, 2004.

[11] H. Rauhut, "Stability results for random sampling of sparse triginometric polynomials," To appear.

[12] H. Rauhut, "On the impossibility of uniform sparse reconstruction using greedy methods," To appear.

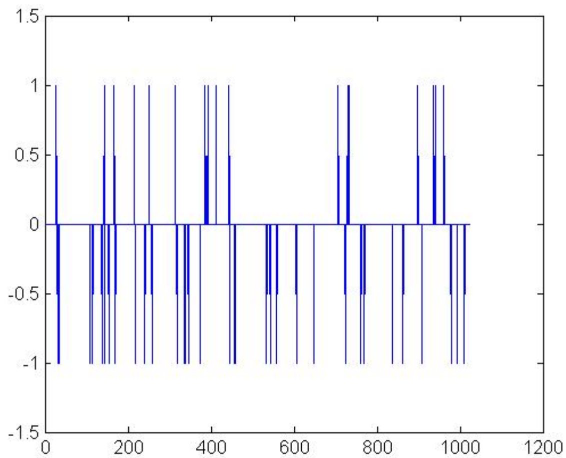
[13] E.J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Info. Theory*, vol. 51, pp. 4203–4215, 2005.

[14] J.A. Tropp and A.C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Info. Theory.*, 2007.

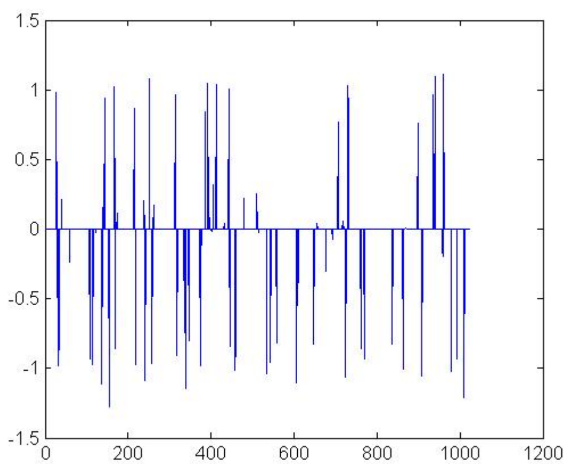
[15] S. Kunis and H. Rauhut, "Random sampling of sparse trigonometric polynomials ii - orthogonal matching pursuit versus basis pursuit," Submitted, 2006.

[16] E.J. Candes, J.Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math.*, vol. 59, pp. 1207–1223, 2006.

[17] J.A. Tropp, "Just relax: Comvex programming methods for subset selection and sparse approximation," ICES Report 04-04, Univ. Texas at Austin, Feb. 2004.



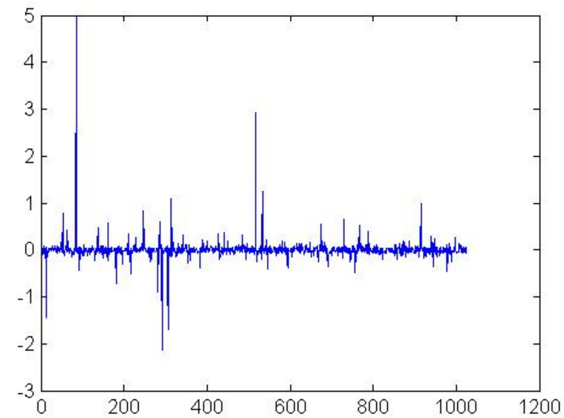
(a) Input signal



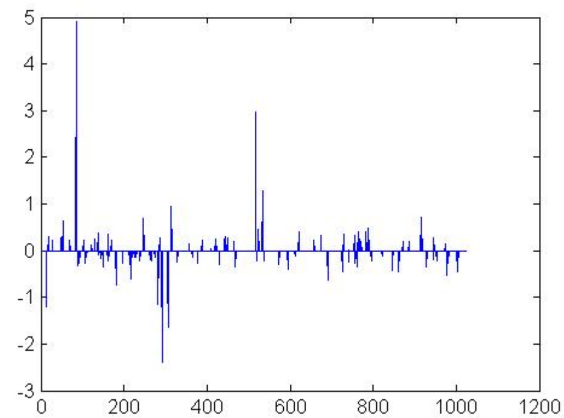
(b) Recovery signal

Fig. 6. Example of recovery signal with size = 1024, sparsity $T = 50$, noise level $\sigma = 0.05$ and so noise energy $\eta = 0.71$, number of iterations = 3, recovery error = 1.57.

.



(a) Input signal



(b) Recovery signal

Fig. 7. Example of recovery compressible signal with size = 1024, $T = 50$ biggest coefficients, noise level $\sigma = 0.05$ and so noise energy $\eta = 0.72$, number of iteration = 5, recovery error = 3.04.

.