ABSTRACT

We analyze the noise robustness of sparse signal reconstruction based on the compressive sensing equivalent of a list-decoding algorithm for Reed Solomon codes - the Coppersmith-Sudan algorithm. We use results from the perturbation analysis of singular subspaces of matrices to prove the existence of bounds for the noise levels (in the measurements) below which the error in the recovered signal (with respect to the original sparse signal) will be guaranteed to be upper bounded. Numerical simulations have been presented which compare the experimental recovery probability to the theoretical lower bound.

Index Terms—Compressed Sensing, Reed Solomon decoding.

1. INTRODUCTION

The typical problem one looks to solve in compressive sensing (CS) is the following: retrieve \( x \in \mathbb{C}^n \) from \( y \in \mathbb{C}^m \) where \( y = Ax \) for some \( A \in \mathbb{C}^{m \times n} \) (\( m < n \)), under the assumption that the number of nonzero entries of \( x \) is at most \( s \), for some \( s \ll n \), i.e. \( ||x||_0 \leq s \). One naive way to arrive at a solution of the above problem is to do an exhaustive search over all \( \binom{n}{s} \) possible support sets of \( x \). However, for some classes of matrices \( A \), one can do better in terms of solution complexity.

One relaxation of the CS problem is obtained by replacing the \( \ell_0 \)-norm with the closest convex norm, namely \( \ell_1 \)-norm

\[
\min_{y = Bx_0} ||x_0||_1 \quad (1)
\]

It is well known that under certain classes of conditions on \( A \), known as Restricted Isometry Properties (RIP) or neighborly polytope conditions, the relaxed problem in (1) gives the same result as the original exhaustive search (e.g. [1, 2]). Furthermore, there exist results in the literature which show that if \( A \) is drawn randomly from certain ensembles of measurement matrices, then with high probability, those conditions hold ([3, 4]), for certain regimes of problem dimensions.

Another line of research in CS has been dedicated to the explicit construction of the measurement matrix \( A \), with the motivation that upon careful construction, the resulting achievable recovery thresholds can be higher than those of random matrices. Furthermore, certain classes of explicit matrices are amenable to novel recovery algorithms different from the linear programming decoding, and potentially faster. A few examples of recent work along this line are [5, 6] that propose certain greedy algorithms for high quality expander graph based measurement, [7, 8] which suggest sublinear time recovery algorithms for certain deterministic constructions, and [9, 10] that use explicit construction using ideas from Reed Solomon (RS) list decoding [11]. Our focus in this paper is on the latter classes of algorithms, namely List decoding algorithms. These algorithms are attractive because of their potentially higher error correction capabilities. For example it has been shown in [10] that under the assumption of synchronized error locations, RS decoding approaches the weak threshold of \( \ell_1 \) minimization. However, unlike typical \( \ell_1 \) decoding [1], theoretical guarantees for the robustness of Reed Solomon decoding algorithms are not very well studied. This work is an attempt to fill in that gap.

The remaining paper is organized as follows. In Section 2 we present the recovery algorithm. In Sections 3 and 4, we establish sufficient conditions for noisy signal recovery. Numerical validation is presented in Section 5.

2. THE PROBLEM MODEL

2.1. The measurement matrix

Let \( \mathbb{C} \) be the complex field. Let \( f : \mathbb{C} \to \mathbb{C} \) be a polynomial of degree lesser than \( K \). Let \( \lambda \) be a primitive \( N^{th} \) root of unity. Let \( \omega_i = \lambda^i \). Then \( (f(\omega_0), f(\omega_1), \ldots, f(\omega_{N-1}))^T \in \mathbb{C}^N \) is a codeword. All the valid codewords form a vector space \( RS(N, K) \) of dimension \( K \). Let the subspace orthogonal to that be denoted by \( Synd(N, K) \). Then if \( \Psi \in \mathbb{C}^{N \times K} \) denote the linear operator which projects vectors in \( \mathbb{C}^K \) to \( RS(N, K) \), then define \( \Phi \in \mathbb{C}^{(N-K) \times N} \) to be such that \( \Phi \Psi = 0 \). The measurement \( y \in \mathbb{C}^{N-K} \) is defined by

\[
y = \Phi x \quad (2)
\]

The noisy measurement is then given by

\[
\tilde{y} = \Phi x + \nu \quad (3)
\]
where the entries of \( \nu \) are drawn from an iid noise process.

2.2. Equivalence of sparse signal recovery to RS decoding

We have \( \Phi^H y \in \text{Synd}(N, K) \). Also, any \( x \in \mathbb{C}^n \) can be written as
\[
x = (I - \Phi^H \Phi) x + \Phi^H \Phi x
\]
But the first term is a Reed Solomon code hence, \( \Phi^H y = x - (I - \Phi^H \Phi) x \) represents a Reed Solomon codeword "corrupted" at \( s \) locations if \( x \) is \( s \)-sparse.

2.3. The decoding algorithm

The following description (Section 3, [10]) has been modified to take into account noisy measurements. First we introduce some notations.

Let \( l \) and \( p \) be positive integers. Let \( f_{x, y} \), be a column vector whose entries correspond to all monomials \( x^a y^b \), such that \( a + (K - 1) b < l \). Also let \( f_{x, y}^{e,d} \) stand for a column vector whose entries correspond to the following term \( \frac{\partial}{\partial x^{a} \partial y^{b}} x^a y^b |_{(x, y)} \). Define \( S_p = \{ (e, d) : e + d < p \} \).

Then define \( B_j \) to be the \( M_j \times |S_p| \) matrix such that each column of \( B_j \) corresponds to \( f_{x, y}^{e,d} \). Please note that each column of \( B_j \) is uniquely indexed by a \( (e, d) \in S_p \). Define the \( M_j \times N |S_p| \) size matrix \( H \) as \( \{ B_1, B_2, \ldots, B_N \} \). We note that any element of the right null-space vector \( b \) of \( H \), can be indexed by \( (i, (e, d)) \) where \( 1 \leq i \leq N_j, (e, d) \in S_p \). Let \( m_b \in \mathbb{R}^N \), defined for the above vector \( b \) be such that \( m_{b_i} = \max_{(e, d) \in S_p} |b_{(i, (e, d))}| \), where \( m_{b_i} \) refers to the \( i \)-th element of the vector \( m_b \). The algorithm is formally described in Algorithm 1.

Algorithm 1 The Coppersmith Sudan algorithm for list decoding over the complex field (noisy case)

- **Input:** A sequence of pairs \( \{ (\alpha_i, \beta_i) \}_{i=0}^{N-1} \), an agreement parameter \( t \), an integer \( K \), integers \( l \) and \( p \).
- **Output:** If it exists, a polynomial \( p(x) \) of degree lesser than \( K \) such that
  \[
  \beta_j = p(\alpha_j), \quad \forall j \in J \subseteq \{0, 1, \ldots, N - 1\} \text{ and } |J| \geq t
  \]
1. Find out any right singular vector \( b \) corresponding to the smallest singular value of \( H \), i.e. \( b = \arg \min_{|z|=1} ||Hx|| \)
2. Let \( J \) represent the index set of the \( K \) largest elements of \( m_b \)
3. Fit a polynomial \( p(x) \) of degree lesser than \( K \), such that \( \sum_{i \in J} ||p(\alpha_i) - \beta_i||^2 \) is minimized.
4. Output \( p(x) \)

3. DERIVATION OF A SUFFICIENT CONDITION FOR NOISE ROBUSTNESS

The above algorithm has two phases:

- Recovery of the complement of the support set of the original \( s \)-sparse signal
- Polynomial fit through the entries in the above set

We next establish sufficient conditions for the correct recovery of the support set.

Let \( H \in \mathbb{C}^{m \times n} \) \( (m \geq n \geq k, \text{rank}(H) = k) \), and let \( \tilde{H} = H + E \) where \( E \) represents a small perturbation matrix. Let \( ||E||_2 \) and \( ||E||_F \) stand for the spectral norm and Frobenius norms of \( E \) respectively. The singular value decomposition of \( H \) is given by
\[
H = (U_1 \ U_2 \ U_3) \begin{pmatrix}
\Sigma_1 & 0 \\
0 & \Sigma_2 \\
0 & 0
\end{pmatrix} \begin{pmatrix}
V_1^H \\
V_2^H \\
V_3^H
\end{pmatrix}
\]

where \( \Sigma_1 = \text{diag}(\sigma_1, \ldots, \sigma_k), \Sigma_2 = \text{diag}(\sigma_{k+1}, \ldots, \sigma_n), \sigma_i \leq \sigma_j \ \forall i \geq j, \text{and} \ \sigma_k > 0 \) and \( \sigma_{k+1} = 0 \). Let \( \tilde{H} = U_1 \Sigma_1 V_1^H \). The singular value decomposition for \( \tilde{H} \) is similarly defined as
\[
\tilde{H} = (\tilde{U}_1 \ \tilde{U}_2 \ \tilde{U}_3) \begin{pmatrix}
\Sigma'_1 & 0 \\
0 & \Sigma'_2 \\
0 & 0
\end{pmatrix} \begin{pmatrix}
\tilde{V}_1^H \\
\tilde{V}_2^H \\
\tilde{V}_3^H
\end{pmatrix}
\]

Here \( \Sigma'_1 = \text{diag}(\sigma'_1, \ldots, \sigma'_k), \Sigma'_2 = \text{diag}(\sigma'_{k+1}, \ldots, \sigma'_n), \sigma'_i \leq \sigma'_j \ \forall (i \geq j) \). Also, define \( \tilde{H}_2 = \tilde{U}_2 \Sigma'_2 \tilde{V}_2^H \). We first state a theorem, due to Wedin [12], about the deviation of the singular values due to small perturbations.

**Theorem.** \( |\sigma_i - \sigma'_i| \leq ||E||_2 \ \forall i \in \{1, 2, \ldots, N\} \)

We next state a theorem due to Wedin (Section 5, [13]) about the stability of singular subspaces corresponding to small singular values.

**Definition 1.** For matrices \( M_1 \) and \( M_2 \) we define
\[
\sin(M_1, M_2) = (I - P_{M_1}) P_{M_2}
\]
where \( P_M \) stands for the projection matrix onto the range space of \( M \).

**Theorem (Wedin’s theorem).** If there exists an interval \([\beta, \alpha]\) and \( \epsilon > 0 \) such that all the singular values of \( H_1 \) lie in \([\beta, \alpha]\) and such that singular values of \( H_2 \) lie outside of \([\beta - \epsilon, \alpha + \epsilon]\), then
\[
\max(||\sin(\tilde{V}_2, V_2)||_F, ||\sin(\tilde{U}_2, U_2)||_F) \leq \frac{\sqrt{2} ||E||_F}{\epsilon}
\]

We now establish a sufficient condition for the correct recovery in Step 1 of the Algorithm 1 in the presence of noise. Due to lack of space the proofs have been omitted.
Lemma 1. If \( ||\sin(\tilde{v}, v)|| \leq \epsilon \), then
\[
\min(||\tilde{v} - v||, ||\tilde{v} + v||) \leq \frac{\epsilon}{\sqrt{2}}
\] (8)

We now prove an analogous fact for subspaces.

Theorem 1. Let \( V_2 \) and \( \tilde{V}_2 \) be as defined earlier. If
\[
||\sin(V_2, \tilde{V}_2)|| \leq \mu
\] (9)

Then \( \forall v \in \tilde{V}_2 \) such that \( ||v|| = 1 \), there exists an unit vector \( v_0 \in V_2 \) such that
\[
\min(||v_0 - v||, ||v_0 + v||) \leq \frac{\mu}{\sqrt{2}}
\] (10)

We next use the above result to arrive at a condition when the support set would be correctly recovered even in the presence of noise. We have the following fact.

Lemma 2. If for a vector space \( V \), \( \forall v \in V \) such that \( ||v|| = 1 \), the support set of \( m_v \) is at least \( K \), then
\[
\min_{v \in V} \left| m^*_v(K) \right| > \gamma \text{ for some } \gamma > 0
\]

where \( m^*_v(i) \) refers to the \( i^{th} \) largest entry (by absolute value) of \( m_v \).

We now state a sufficient condition guaranteeing the recovery of the correct support set. Here \( \tilde{H} \) is as constructed in Algorithm 1 with \( \beta = \Phi^H \hat{y} \) (\( \hat{y} \) is defined in Eqn 3), \( \alpha_i \) being the \( i^{th} \) power of a primitive root of unity. \( \delta \) is the smallest non zero singular value of \( H (\beta = \Phi^H y \text{ in Algorithm 1 with } y \text{ defined in Eqn 2}) \) and \( \gamma \) is the constant in Lemma 2 with \( V \) being the right null space of \( H \).

Theorem 2. If for \( E = \tilde{H} - H \)
\[
1. 2||E||_F < \delta
2. \frac{2||E||_F}{s} \leq \frac{\gamma}{\sqrt{2}}
\]

then a subset of the complement of the support set of sparse signal \( x \) is recovered in Algorithm 1.

4. SOME RESULTS FOR A MORE SPECIFIC CONSTRUCTION

We compute lower bounds for the above threshold for a special case of the general construction in Algorithm 1. More specifically, let \( \tilde{H} \) be defined as (Using \( p = 1 \) in algorithm 1)
\[
\tilde{H} = \begin{pmatrix}
1 & 1 & \ldots & 1 \\
\alpha_0 & \alpha_1 & \ldots & \alpha_{N-1} \\
\vdots & \vdots & \ddots & \vdots \\
\alpha_0^{l-1} & \alpha_1^{l-1} & \ldots & \alpha_{N-1}^{l-1} \\
\beta_0 & \beta_1 & \ldots & \beta_{N-1} \\
\alpha_0 \beta_0 & \alpha_1 \beta_1 & \ldots & \alpha_{N-1} \beta_{N-1} \\
\vdots & \vdots & \ddots & \vdots \\
\alpha_0^{l-K} \beta_0 & \alpha_1^{l-K} \beta_1 & \ldots & \alpha_{N-1}^{l-K} \beta_{N-1}
\end{pmatrix}
\] (11)

Here \( \beta \) is the noisy “codeword” obtained from the noisy measurement in Equation 3 as follows
\[
\beta = \Phi^H \hat{y}
\]

\( H \) is similarly defined with \( \beta \) in Eq 11 replaced by \( \beta = \Phi^H y \).

Thus
\[
E = \begin{pmatrix}
0 & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0^{l-1} & 0^{l-1} & \ldots & 0^{l-1} \\
\mu_0 & \mu_1 & \ldots & \mu_{N-1} \\
\alpha_0 \mu_0 & \alpha_1 \mu_1 & \ldots & \alpha_{N-1} \mu_{N-1} \\
\vdots & \vdots & \ddots & \vdots \\
\alpha_0^{l-K} \mu_0 & \alpha_1^{l-K} \mu_1 & \ldots & \alpha_{N-1}^{l-K} \mu_{N-1}
\end{pmatrix}
\]

where \( \mu = \Phi^H \nu \). We note that the first \( l \) rows of \( H \) are Vandermonde and hence are full rank \( l \). With each non zero entry of \( x \), the row rank increases by one. Thus the number of linearly independent rows is \( \text{rank}(H) = l + s \). We also note that \( l + s < N \) and \( 2l - K > N \), for a right null space vector to exist. Then we would like to obtain a lower bound on the \( K^{th} \) largest entry of \( b \) such that \( Hb = 0 \) and \( ||b|| = 1 \). We have
\[
Hb = 0 \Rightarrow b^H Hb = 0 \quad \forall c \in C^{2l-K}
\] (12)

Now, we know that almost surely, the indices corresponding to the support set of the \( s \)-sparse signal \( x \) cannot figure in a linear dependency. Hence let us consider any collection of \( l + 1 \) columns involving measurements which do not lie on the support set of the \( s \) sparse \( x \). Denote that matrix by \( B \).

The rank of \( B \) is exactly \( l \), thus the support set of any null space vector is \( l + 1 \). We now proceed to show that the right null space can be totally characterized in terms of the \( \alpha_i \)s. We have
\[
b^H B = (r(\alpha_0), r(\alpha_1), r(\alpha_2), \ldots, r(\alpha_{N-1}))
\] (13)

where \( r \) is a degree \( l \) polynomial whose coefficients are determined by \( c \) which is arbitrary. Then, from Equation 12 for all choices of \( r \) we need to have \( \sum_{i=0}^{l} b_i r(\alpha_i) = 0 \). Let
\( r_{0i}(\alpha) = \prod_{1 \leq j \leq l, j \neq i} (\alpha - \alpha_j). \) Let \( r_{0i} \) be chosen for all \( i \in \{1, 2, \ldots, l\}. \) This will allow us to find out all the entries of \( b \) in terms of \( b_0. \) More specifically,

\[
b = b_0 \left( -1, \frac{r_{01}(a_0)}{r_{01}(a_1)}, \frac{r_{02}(a_0)}{r_{02}(a_2)}, \ldots, \frac{r_{0l}(a_0)}{r_{0l}(a_l)} \right)^T
\]

We can characterize the nullspace of \( H \) by finding out \( N - l - s \) (i.e. rank of the space) distinct support sets of size \( l + 1. \) For each such set, one can define a \( b \) as above. Thus, we can find \( N - l - s \) linearly independent vectors in the null space of \( H. \) Clearly, then \( \gamma \) or the \( K^{th} \) largest(in magnitude) entry for any vector in this subspace is a function of the roots of unity, and can be precomputed, knowing \( N \) and \( l. \)

We now compute a lower bound on probability of correct recovery \( P_{cr} \) of a subset of the complement of the support set of \( x \) conditioned on the input \( x. \) From Theorem 2, we have

\[
P_{cr} \geq P(\mu^H \mu \leq \frac{(\gamma \delta(x))^2}{8(l-K)^2})
\]

However, by construction of \( \Phi, \mu^H \mu = \nu^H \Phi \Phi^H \nu = \nu^H \nu. \) Assuming \( \nu \sim N(0, \sigma^2 I_{l-K}) \), we see that

\[
\frac{\nu^H \nu}{\sigma^2} \sim \chi^2_{l-K}
\]

Thus if \( F(x, N-K) \) represents the cumulative distribution function of \( \chi^2_{l-K}(x) \), then

\[
P_{cr}(x) \geq F(\frac{(\gamma \delta(x))^2}{8(l-K)\sigma^2}, N-K) \quad (14)
\]

5. NUMERICAL SIMULATIONS

We plot (Fig. 1) the empirical probability of correct recovery and the associated theoretical lower bound (Eq. 14) as a function of \( \sigma \) - the noise standard deviation. The non zero locations in the sparse signal \( x \) were uniformly selected and the entries were subsequently randomly drawn from \( N(0,1). \) The empirical and theoretical values were averaged over 1000 samples of \( x. \) \( \gamma \) was obtained by minimizing the \( K^{th} \) largest entry of vectors lying on the unit sphere in the right null space of \( H \) (using a simulated annealing routine and a subsequent Nelder Mead minimization). We see that the obtained sufficient conditions are somewhat pessimistic with respect to the observed simulation results.

6. ACKNOWLEDGMENT

We would like to thank Matthew Thill and Yash Deshpande for helpful discussions.

7. CONCLUSION

Sufficient conditions guaranteeing boundedness of the reconstruction error in a list decoding algorithm for Compressed Sensing have been derived.

8. REFERENCES


