1-Bit Compressive Sensing: From the Perspective of Detection and Estimation Theory

Term Paper: Detection and Estimation Theory (EE623A),
Presented by: **Vipul Gupta** (11810) and **Anirudh K. Agrawal** (11098)
Supervisor: Prof. Aditya K. Jagannatham
Email: {vipgupta, kanirudh}@iitk.ac.in

***Abstract*—Compressive sensing is a new signal acquisition technology with the potential to reduce the number of measurements required to acquire signals that are sparse or compressible in some basis. Compressive sensing reconstruction has been shown to be robust to multi-level quantization of the measurements, in which the reconstruction algorithm is modified to recover a sparse signal consistent to the quantization measurements. In this term paper, we discuss the limiting case of 1-bit measurements, which preserve only the sign information of the random measurements. Here, we formulate a convex optimization problem by treating the 1-bit measurements as sign constraints and discuss various cost functions associated to the optimization problem which have been proven to work well in literature. We will draw comparison between several methods of estimating the sparse signal using 1-bit quantized measurements in terms of performance, complexity and number of measurements required. The methods we discuss here include maximum likelihood estimation, regularized least squares method, sign-sketch procedure and estimation by constraining the problem on unit sphere.**

***Index Terms*—Compressive Sensing, maximum-likelihood estimation, least squares estimation, quantization, support recovery.**

I. Introduction

Compressed Sensing (CS) [1] is a new paradigm which enables the reconstruction of compressible or sparse signals with far fewer samples using a universal sampling procedure compared to that with traditional sampling methods. In this framework, a small collection of linear random projections of a sparse signal contains sufficient information for signal recovery. The fundamental premise is that certain classes of signals, such as natural images or communications signals, have a representation in terms of a sparsity inducing basis (or sparsity basis for short) where most of the coefficients are zero or small and only a few are large. For example, smooth signals and piecewise smooth signals are sparse in a Fourier and wavelet basis, respectively.

The quantization of CS measurements has been studied recently and it has been shown that accurate and stable signal acquisition is possible even when each measurement is quantized to only one single bit which is termed 1-bit Compressive Sensing. In the following paper we discuss algorithms for solving the 1-bit compressive sensing problem in presence of uncertainty. In [2], the authors use two different approaches to solve the problem. The first is the l1-regularised Maximum Likelihood (ML) approach and other a more naive method based on-regularized least squares. It has been demonstrated that reconstruction from 1-bit compressive sensing measurements can be significantly improved if the appropriate measurement model is used in the reconstruction [3]. However, their approach is mostly applicable in the case where measurements are inexpensive whereas precision quantization is expensive. They have imposed the constraint $\|x\|_2 = 1$, otherwise a minimization based reconstruction algorithm that only requires consistency with the measurements will drive the solution to x = 0. However in practice we always encounter some noise in the observations,so we discuss a method using Adaptive Outlier Pursuits for Robust 1-bit compressive sensing [5] which uses an iterative procedure to estimate the the sparse vector.

The problem of support recovery has received little attention to date in the CS literature when corrupted by non-Gaussian uncertainties. Majority of the work till now has treated the uncertainty as noise. The methods discussed below attempts the problem when measurements are corrupted by outliers in addition to Gaussian noise and even subsequently quantized. In [4] we discuss a method called Sign-Sketch procedure, which is shown to be a robust and sufficiently accurate.

II. Background

A. Compressive Sensing

Signals which are sparse or compressible in some sense can be reconstructed using the newly developed method known as compressive sensing [1], [6]. Let us assume that $\mathbf{x} \in \mathbf{R}^N$. If there are at most K non-zero coefficients α_i in the expansion of the basis $x = \sum_i \alpha_i \mathbf{b}_k$ which can be denoted as $\mathbf{B}\alpha$, the signal is said to be K-sparse in sparsity-inducing basis $\mathbf{b}_i$. Similarly, the signal is called K-compressible if it can be well represented by the K-most significant coefficients in the expansion.

M measurements are taken of the signal using measurement vectors $\phi_i, i = 1, 2, ..., M$ such that:

$$y_i = \langle x, \phi_i \rangle$$

Above can be compactly denoted by $y = \Phi x = \Phi \mathbf{B}\alpha$ where y is the measurement vector and Φ is the measurement operator modelling the measurement system.

For all further discussion, signal is assumed to be sparse or compressible in canonical basis, i.ie., $\mathbf{B} = I$. Now if the signal is sparse in any basis, we can easily apply the subsequent results by substituting$\bar{\Phi} = \Phi\mathbf{B}$ as the measurement system and treating α, instead of x, as the sparse signal to be reconstructed.

According to the classical sampling theory,the set ϕ_i should form a Riesz basis or a frame for robust linear reconstruction of any signal x which makes it necessary to have N measurements for recovery of signals. Whereas compressive sensing allows us to use only $M = O(Klog(N/K))$ non-adaptive measurements which is much less than N, to reconstruct K-sparse or K-compressible signals linearly.

The reconstruction of the signal x from y leads to determining the sparsest signal that can explain the measurements y. Although the strictest measure of sparsity is the l_0 pseudonorm of the signal which is defined as the number of non-zero coefficients of the signal, compresssive enforcing ensures sparsity by minimizing the l_1 norm of the the reconstructed signal, $||x||_1 = \sum_i |x_i|$ due to the combinatorially complexity of the l_0 pseudonorm.In the case of classical compressive sensing reconstruction methods including many other cases, minimizing the L_1 norm has been theoretically proven to be equivalent of minimizing the L_0 pseudonorm of the signal.

Thus we obtain reconstruction from compressive sensing measurements by solving the following minimization problem:

$$\hat{x} = \arg\min_x ||x_1|| s.t. \mathbf{y} = \Phi x$$

Signal can be exactly recovered using the above equation if proper Φ is used.To be more specific, measurement vectors ϕ_i should be sufficiently incoherent with the sparsity basis b_i for exact recovery. The minimum number of measurements required to ensure recovery using a random measurement system is $M = \mathcal{O}(K \log N/K)$

B. Measurement Quantization

Quantization is modelled as measurement value added with measurement noise denoted by **n**.

$$\mathbf{y} = \mathcal{Q}(\Phi\mathbf{x}) = \Phi\mathbf{x} + \mathbf{n},$$

where $\mathcal{Q}(.)$ is the quantizer and n is dependent on quantization accuracy and energy-bounded as follows:

$$||n||_2 = \left(\sum_i ||n_i||^2\right)^{1/2} \leq \epsilon.$$

For a uniform linear quantizer with quantization interval $\Delta, \epsilon \leq \sqrt{M\Delta^2/12}$

In the case when measurement is limited by norm such as quantization,robust reconstruction can be obtained by solving:

$$\hat{\mathbf{x}} = \arg\min_x ||x||_1 \quad s.t. \quad ||\mathbf{y} - \Phi\mathbf{x}||_2 \leq \epsilon$$

In this approach, the reconstruction error is limited by $||\mathbf{x} - \hat{\mathbf{x}}||_2 \leq \mathcal{C}\epsilon$, where C is constant and dependent only on the properties of the measurement system Φ.

Above optimization problem is often relaxed by following problem which is more efficient to solve:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} ||\mathbf{x}||_1 + \frac{\lambda}{2}||\mathbf{y} - \Phi\mathbf{x}||_2^2$$

C. Consistent Reconstruction

Consistent Reconstruction emphasized the intutive idea that solution obtained must be consistent with the prior knowledge of the signal and its measurement process. This implies that if reconstructed signal is measured using the same measurement process and quantized using same quantization, we will get the same measured value as was obtained to reconstruct the signal.

Reconstruction is not consistent in the case when measurement noise is due to quantization noise.

Specifically, when we use uniform linear quantization, all noise components have magnitude $|n_i| \leq \Delta/2$, which implies that consistent reconstruction produces a signal which satisfies:

$$|(\Phi\hat{\mathbf{x}} - \mathbf{y})_i| \leq \frac{\Delta}{2}$$

For 1-bit quantization, the quantizer is often implemented as a comparator to a voltage level *l*,which is usually zero. In our method, consistent reconstruction requires that reconstructed signal measurements is on the same side of the voltage level as the measurements obtained from the measurement system which can be mathematically expressed as:

$$sign\left((\Phi\hat{\mathbf{x}})_i - l\right) = y_i.$$

III. 1-BIT COMPRESSIVE SENSING MEASUREMENTS

A. Measurement Model

As it was explained in the above section, each measurement is simply the sign of the inner product of the sparse signal with a measurement vector ϕ_i as follows:

$$y_i = sign\left(\langle\phi_i, \mathbf{x}\rangle\right).$$

Clearly each quantized measurement multilied with the measurement is always non-negative:

$$y_i sign\left(\langle\phi_i, \mathbf{x}\rangle\right) \geq 0.$$

Above two equations can also be expressed using matrix and vector notations as:

$$\mathbf{y} = sign\left(\Phi\mathbf{x}\right) \quad \text{and}$$

$$Y\Phi\mathbf{x} \geq 0.$$

where Y = diag(y) and the inequality is applied element-wise.

B. Consistent Reconstruction

For consistent reconstruction using 1-bit measurements,the measurements are seen as sign constraints which are enforced in the reconstruction to recover the signal.l_1 norm is enforced as a sparsity measure in the reconstruction.

If **x** is consistent with the measurements then so is $a\mathbf{x}$ for all $0 \leq a < 1$. Since $||a\mathbf{x}_1 = a\|x\|_1 < \|x\|_1$, a minimization based reconstruction algorithm that only requires consistency with the measurements will drive the solution to $\mathbf{x} = 0$. To enforce reconstruction at a non-trivial solution we need to artificially resolve the amplitude ambiguity. Thus, we impose an energy constraint that the reconstructed signal lies on the unit l_2-sphere:

$$\|x\|_2 = \left(\sum_i x_i^2\right)^{1/2} = 1 \tag{1}$$

Note that this constraint significantly reduces the optimization search space. This reduction plays an important role in improving the reconstruction performance.

The sparsest signal on the unit sphere that is consistent with the measurements is the solution to:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} ||\mathbf{x}||_1 \quad \text{s.t. } \mathbf{Y}\Phi\mathbf{x} \geq 0 \quad \text{and } \|\mathbf{x}\|_2 = 1. \tag{2}$$

To enforce the constraint we relax the problem using a cost function $f(x)$ that is positive for $x < 0$ and zero for $x \geq 0$ and a relaxation parameter λ:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} ||\mathbf{x}||_1 + \lambda \sum_i f((\mathbf{Y}\Phi\mathbf{x})_i) \quad \text{s.t. } \|\mathbf{x}\|_2 = 1. \tag{3}$$

Assuming that the original problem (2) is feasible, as λ tends to infinity (2) and (3) have the same solution. The algorithm we introduce by the authors in [2] minimizes problem (2) for $f(x) = \frac{x^2}{2}u(x)$, where $u(x)$ is the unit step function. The convexity and smoothness of function $f(x)$ allows the use of gradient descent and fixed-point methods to perform the minimization.

For notational convenience, in the remainder of this paper we use $g(x) = \|x\|_1$ to denote the $l-1$ norm part of the cost function, and $\bar{f}(\mathbf{Y}\Phi\mathbf{x})$ to denote the one-sided quadratic penalty

$$\bar{f}(x) = \sum_i f(x_i)f(xi)$$

such that the cost function is equal to:

$$Cost(x) = g(x) + \lambda \bar{f}(\mathbf{Y}\Phi\mathbf{x}). \tag{4}$$

C. Reconstruction Algorithm

The authors in [3] employ a variation of the fixed point continuation (FPC) algorithm [7]. Specifically, we introduce two modifications. The first modifies the computation of the gradient descent step such that it computes the gradient of the one-sided quadratic penalty projected on the unit sphere $\|x\|_2 = 1$. The second introduces a renormalization step after each iteration of the algorithm to enforce the constraint that the solution lies on the unit sphere. These modifications are introduced to stabilize the reconstruction of sparse signal from their zero crossings. The similarity is not coincidental. Both sign measurements and zero crossings information eliminate amplitude information from the signal. The main difference between the two problems is that measurements of zero crossings are signal-dependent, whereas compressive measurements are signal-independent. Although it is possible to reformulate the reconstruction from zero-crossings as reconstruction from 1- bit measurements, this reformulation is beyond the scope of this paper. The algorithm computes and follows the gradient of the cost function in (4). If the minimization is not constrained on the sphere, then the gradient of the cost at the minimum is 0:

$$Cost'(\mathbf{x}) = 0 = g'(\mathbf{x}) + \lambda(\mathbf{Y}\Phi)^T \bar{f}'(\mathbf{Y}\Phi\mathbf{x}) \tag{5}$$

$$\Rightarrow \frac{g'(\mathbf{x})}{\lambda} = -(\mathbf{Y}\Phi)^T \bar{f}'(\mathbf{Y}\Phi\mathbf{x}). \tag{6}$$

Algorithm 1 Renormalized Fixed Point Iteration

1) **Initialization:**
   Seed: $\widehat{\mathbf{x}}_0$ s.t. $\|\widehat{\mathbf{x}}_0\|_2 = 1$,
   Descent Step Size: δ
   Counter: $k \leftarrow 0$
2) **Counter Increase:**
   $k \leftarrow k + 1$
3) **One-sided Quadratic Gradient:**
   $\bar{\mathbf{f}}_k \leftarrow (\mathrm{Y}\Phi)^T \bar{f}'(\mathrm{Y}\Phi\mathbf{x}_{k-1})$
4) **Gradient Projection on Sphere Surface:**
   $\widetilde{\mathbf{f}}_k \leftarrow \bar{\mathbf{f}}_k - \langle \bar{\mathbf{f}}_k, \mathbf{x}_{k-1}\rangle \mathbf{x}_{k-1}$
5) **One-sided Quadratic Gradient Descent:**
   $\mathbf{h} \leftarrow \widehat{\mathbf{x}}_{k-1} - \delta\widetilde{\mathbf{f}}_k$
6) **Shrinkage (ℓ_1 gradient descent):**
   $(\mathbf{u})_i \leftarrow \operatorname{sign}((\mathbf{h})_i) \max\left\{|(\mathbf{h})_i| - \frac{\delta}{\lambda}, 0\right\}$, for all i,
7) **Normalization:**
   $\widehat{\mathbf{x}}_k \leftarrow \frac{\mathbf{u}}{\|\mathbf{u}\|_2}$
8) **Iteration:** Repeat from 2 until convergence.

It follows that if the sphere constraint is introduced then the gradient of the cost function at the minimum is orthogonal to the sphere. Thus, a gradient descent algorithm followed by renormalization has the minimum of (4) on the unit sphere as a fixed point.

The iterative steps to reconstruct the signal are presented in Algorithm 1. The algorithm is seeded with an initial signal estimate $\hat{\mathbf{x}}_0$ and a gradient descent step size δ/λ. At every iteration the algorithm computes the gradient of the one-sided quadratic in Step 3, projects it on the sphere in Step 4 and descends on that gradient in Step 5. Step 6 is a shrinkage step using the soft threshold shrinkage function shown in the solid line in Figure 1. Step 7 renormalizes the estimate to have unit magnitude and the algorithm iterates from Step 2 until the solution converges.

The shrinkage Step 6 is interpreted as a gradient descent on the $l1$-norm component of the cost function. Specifically, for $|x_i| \geq \delta/\lambda$ the magnitude of the coefficient is reduced by δ/λ, which is the expected behavior of a gradient descent. For $|x_i|\delta/\lambda$ the discontinuity at zero makes the gradient descent set the coefficient to 0.

The reconstruction algorithm should be executed with λ large enough such that the relaxed minimization (3) converges to the constrained minimization in (2). Unfortunately, the larger the value of λ, the smaller the descent step δ/λ.

Furthermore, the value of λ that is sufficiently large is not known in advance of the algorithm.

Both issues are resolved by wrapping the algorithm in an outer iteration loop that executes the algorithm using a small value λ_0 until convergence and then restarts the algorithm with a higher value $\lambda_i = c\lambda_{i-1}, c > 1$ using the previous estimate as a seed for the next execution. The outer loop terminates once the solution from the current iteration is not significantly different from the solution of the previous iteration. Since the minimization is performed on the unit sphere, the problem is not convex. However, a good estimate of the solution can be found using the algorithm.

IV. Compressive Sensing with Quantized Measurements

In this section we present the work by authors in [2]. We start with noise corrupted $z = Ax + v$ where $A \in \mathbf{R}^{m\times n}$ is the unquantized measurement vector and $z \in R^m$. v is IID $\mathcal{N}(0, \sigma^2)$ noise. $\mathcal{Q}_i : R \rightarrow \mathcal{Y}_i$ functions as the quantizer for z_i where $\mathcal{Y}_i$ is a finite set of pre-defined key-words.
We get quantized measurement as follow:

$$y_i = \mathcal{Q}_i(z_i), i = 1, ..., m.$$

which is same as $z_i \in \mathcal{Q}_i^{-1}(y_i)$.

We will work with the case in which $\mathcal{Q}_i^{-1}(y_i) \in [l_i, u_i)$ i.e., considering only the lower limit. The values l_i and u_i are thresholds associated with the quantized measurement y_i and they can take the values $-\infty$ and ∞ respectively depending on the interval.

Thus, for all our measurements,

$$\mathbf{l} \leq \mathbf{Ax} + \mathbf{v} \geq \mathbf{u}$$

where **l** and **u** are upper and lower limits respectively of the corresponding keywords.

A. Method 1: l_1 regularized Maximum Likelihood

The conditional probability of the codeword y_i corresponding to some x is given by:

$$p(y_i|x) = \Phi\left(\frac{-a_i^T x + u_i}{\sigma_i}\right) - \Phi\left(\frac{-a_i^T x + l_i}{\sigma_i}\right)$$

where a_i^T is the i^{th} row of A and

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp\frac{-t^2}{2}\,\mathrm{d}t$$

is the cumulative distribution function of the standard normal distribution. The negative log-likelihood of given is given by

$$-\sum_{i=1}^{m} \log\left(\Phi\left(\frac{-a_i^T x + u_i}{\sigma_i}\right) - \Phi\left(\frac{-a_i^T x + l_i}{\sigma_i}\right)\right)$$

which we can express as $f_{ml}(Ax)$, where

$$f_{ml}(z) - \sum_{i=1}^{m} \log\left(\Phi\left(\frac{-z_i + u_i}{\sigma_i}\right) - \Phi\left(\frac{-z_i + l_i}{\sigma_i}\right)\right)$$

The negative log-likelihood function is a smooth convex function. We add l_1 regularization parameter in the f_{ml} expression to incorporate sparsity prior and then minimize $f_{ml}(Ax) + \lambda||x||_1$ to get ML estimate of x and adjusting λ to get desired sparsity in x, .

B. Method 2: l_1 regularized Least Squares

This method ignores the quantization. In this method, unquantized real value is used for each quantization interval and this real value is noise corrupted measurement.Corresponding to each measurement y_i, we assign $\hat{y} \in \mathbf{R}$ some value such as centroid or mid point of the interval $[l_i, u_i)$. The centroid is

$$\hat{y} = \frac{\int_{l_i}^{u_i} wp(w)\,\mathrm{d}w}{\int_{l_i}^{u_i} p(w)\,\mathrm{d}w}$$

Measurement can thus be expressed as $z = \hat{y} + q$ where q is the quantization error.

Despite q being a function of Ax+v, we consider it to be a random variable with zero mean and variance

$$\hat{\sigma}^2 = \frac{\int_{l_i}^{u_i} (w - \hat{y})^2 p(w)\,\mathrm{d}w}{\int_{l_i}^{u_i} p(w)\,\mathrm{d}w}$$

$\hat{\sigma}^2 = (u_i - l_i)^2/12$ for the case of a uniform (assumed) distribution on z_i.Now pretending that q is gaussian, we have $\hat{y} = Ax + \hat{v}$, where $v_i \sim \mathcal{N}(0, \sigma^2 + \hat{\sigma}^2)$. We can now use least-squares to estimate x, by minimizing the (convex quadratic) function $f_{ls}(Ax)$, where

$$f_{ls}(z) = -\sum_{i=1}^{m} \frac{1}{2} \frac{(z_i - \hat{y_i})^2}{\sigma^2 + \hat{\sigma}^2}$$

Finally to obtain sparse estimate, we add l_1 regularization as in method 1 and minimize $f_{ls}(z) + \lambda||x||_1$.

V. Sign-Sketch Method

The *Sign-Sketch* procedure employs sparse random measurement matrices, and utilizes a computationally efficient support recovery procedure that is a variation of a technique from the sketching literature.

A. Problem Definition

The support S(x) of a $x \in \mathbb{R}^n$ is defined as the set of positions in x where it is non-zero. To say that $\bar{x}$ is k-sparse means that $|S| = k$. For each vector $\bar{x}$ we acquire 'm' measurements of the form 'Ax', where A is a m x n matrix. The authors have considered two scenarios for inclusion of uncertainty in the measurements

1) *Corrupted by Gaussian Noise and Outliers*
Measurements are of the form
$$y = Ax + w + o \quad (7)$$
$w \sim \mathcal{N}(0, \sigma^2 I_m)$ and $o \in \Re^m$ is a sparse vector (whose values may take arbitrary large values if non-zero)
2) *Quantized Noisy Measurements*
Measurements are of the form
$$y = sgn(Ax + w + o) \quad (8)$$

here w and o have the same definition as above and sgn(z) denotes a scalar sign function defined as

$$sgn(z) = \begin{cases} -1 & z < 0 \\ 0 & z = 0 \\ +1 & z > 0 \end{cases} \quad (9)$$

We aim to calculate $\hat{S}(y, A)$ which is an accurate estimate of the true unknown signal support *S*.

B. Sparse Estimation

We propose to use the same structured random matrices, similar to those proposed in the count-sketch procedure in [9]. Matrices $\mathcal{A} \sim \mathcal{A}(R, T, n,)$ are composed of the vertical concatenation of T individual random matrices, denoted A_t for t = 1, . . . , T, each having R rows and n columns. Each A_t is a sparse matrix containing exactly n non-zero values, one per column, where the location of the non-zero component in each column is chosen uniformly at random (with replacement) from the set $\{1, 2, ..., R\}$, and the non-zero component takes the value $\pm\alpha$ with probability $\frac{1}{2}$. For a given realization, let $h_{t,1}, ..., h_{t,n}, n \in \{1, ..., R\}$ denote which entry of the corresponding column of A_t is non-zero, and we let $s_{t,1}, ..., s_{t,n} \in \alpha, +\alpha$ be the corresponding values, for t = 1, . . . , T. All random quantities are assumed independent in the construction of A_t.

They start the estimation by forming "estimates" $\tilde{x}_t \in \mathbb{R}^n$, where $\tilde{x}_{t,i} = s_{t,i} y_{t,h_{t,i}}$. The procedure leverages the observation that for the two uncertainty models described above, the majority of the $\{\tilde{x}_{t,i}\}_{t=1}^{T}$ may have the same *sign* as x_i for indices $i \in S$, and their signs will otherwise be equally likely for $i \in S^c$. The first step in the *Sign-Sketch* procedure is to estimate the following

$$\hat{x} = \frac{1}{T} \sum sgn(\tilde{x_t})$$

here $\tilde{x}$ are as defined above and the support is estimated using a threshold τ as

$$\hat{\mathcal{S}} = \{i \in \{1, 2, ..., n\} : |\tilde{x}_i| > \tau\} \quad (10)$$

C. Important Results

The first results ensures the accuracy of the estimation process of the support vector where observation might be corrupted by Gaussian Noise and outliers.

Theorem 1 : Given that measurements of x are obtained according to the model (7) where $\mathcal{A} \sim \mathcal{A}(R, T, n,)$ for some specified $R, T \in \mathbb{N}$ and $\alpha > 0$, $w \sim \mathcal{N}(0, \sigma 2 I_m)$, and $o \in \mathbb{R}^m$ is a vector of outliers whose entries take some unspecified (and possibly large) value independently with probability q, and 0 otherwise.

Let the number of non-zero entries of x be k, and let x_{min} denote the minimum amplitude of the non-zero components of x (i.e., $x_{min} = min_{i \in S}|x_i|$). Define the quantities

$$\tilde{p} := \frac{k-1}{R} + \frac{1}{2} exp(-\frac{\alpha^2 x_{min}^2}{2\sigma^2}) + q$$

If the following a true:

- $\tilde{p} < 1/2$
- τ is chosen such that $0 < \tau < 1 - 2\tilde{p}$
- T satisfies for any $\lambda > 0$
$$T \geq max\{\frac{2}{(\tau - (1-2\tilde{p}))^2} \log(4k(n-k)^{\lambda}, \frac{2}{\tau^2} \log(4(n-k)^{\lambda+1})\}$$

then the estimate $\hat{\mathcal{S}}$ as defined earlier satisfies the following $Pr(\hat{\mathcal{S}} \neq \mathcal{S}) \leq (n-k)^{-\lambda}$

We can deduce a few salient points from the above theorem. First, the requirement that $\tilde{p} \leq 1/2$ implies that the following are strictly necessary conditions:

1) for each matrix A_t, number of rows $R > 2(k-1)$
2) the minimum signal amplitude x_{min} must be $\Omega(\sigma/\alpha)$
3) the probability of outliers must satisfy $q < 1/2$

Secondly, we can significantly adjust certain parameters to offset the impact of other whilst still succeeding at recovering the support. For example, provided that the parameters remain within the allowable ranges (so that $\tilde{p} < 1/2$), doubling R, and thus the total number of measurements, offsets the effect of a doubling of the outlier probability q, consistent with intuition. Also, increasing R can permit recovery of signals with weaker features (i.e., smaller values of x_{min}), and so on. Finally, note that for a given q and x_{min} such that

$$\frac{1}{2} exp(-\frac{\alpha^2 x_{min}^2}{2\sigma^2}) + q < \varrho < 1/2$$

the procedure will succeed provided $R > \frac{k-1}{1/2-\varrho}$, implying total number of measurements $RT = \mathcal{O}(max\{k \log(n-k), k \log k\})$.

The robustness of Sign-Sketch model can be inferred from the fact that it utilizes only the sign of each measurement. Hence, we can conclude that the results from Theorem 1 hold for the case when observations are comprised of sign measurements only i.e the Quantized Noisy Measurements.

VI. Robust 1-bit Compressive Sensing using Adaptive Outlier Pursuit

The problem definition followed here is the same as given in background section. Binary iterative hard thresholding (BIHT or BIHT-l_2) in [8] is a high performance algorithm for solving the 1-bit case when it is noise free, however when However when there are a lot of sign flips, the performance of BIHT and BIHT-l_2 is worsened by the noisy measurements. There is no method to detect the sign flips in the measurements, but adaptively finding the sign flips and reconstructing the signals can be combined together as in [22] to obtain better performance. Let us assume first that the noise level (the ratio of the number of sign flips over the number of measurements for 1-bit compressive sensing) is provided. Based on this information, we can choose a proper integer L such that at most L elements of the total measurements are wrongly detected (having sign flips). For measurements $y_i \in +1, -1^m$ $\Lambda \in \mathcal{R}^m$ is a binary vector denoting the correct data:

$$\Lambda_i = \begin{cases} 1 & \text{if} \quad y_i \quad \text{is correct} \\ 0 & \text{else} \end{cases}$$

According to the assumption $\sum_{i=1}^{M} \Lambda_i \leq L$. Introducing Λ_i into the BIHT framework we get the following minimization problem

$$\underset{x,\Lambda}{\text{minimize}} \qquad \sum_{i=1}^{M} \Lambda_i \phi(y_i, (\phi x)_i)$$

$$st. \sum_{i=1}^{M} \Lambda_i \leq L$$

$$\Lambda_i \in \{0,1\} i = 1, 2, ..., M$$

$$||x||_2 = 1, ||x||_0 \leq K$$

The above model can also be interpreted in the following way. Let us consider the noisy measurements y as the signs of ϕx with additive unknown noise n, i.e.$y = sign(\phi x + n)$. Although the binary measurement is robust to noise as long as the sign does not change, there exist some n_is such that the corresponding measurements change. In our problem, only a few measurements are corrupted, and only these corresponding n_i's are important. Therefore, can be considered as sparse noise with non-zero entries at these locations.

Note the problem defined above is non-convex and has both continuous and discrete variables. It is difficult to find $(x\Lambda)$ together, thus we use alternating minimization method, which separates the energy minimization over x and Λ into two steps.

The algorithm is described below

Algorithm 1: AOP

Input: $\Phi \in \mathbf{R}^{M\times N}$, $y \in \{-1,1\}^M$, $K > 0$, $L \geq 0$, $\alpha > 0$, Miter > 0

Initialization: $x^0 = \frac{\Phi^T y}{\|\Phi^T y\|}$, $k = 0$, $\Lambda = \mathbf{1} \in \mathbf{R}^M$, Loc $= 1 : M$, tol $= \inf$, TOL $= \inf$.

while $k \leq$ Miter and $L \leq$ tol **do**

Compute $\beta^{k+1} = x^k + \alpha\Phi(\text{Loc}, :)^T(y(\text{Loc}) - \text{sign}(\Phi(\text{Loc}, :)x^k))$.

Update $x^{k+1} = \eta_K(\beta^{k+1})$,

Set tol $= \|y - A(x^{k+1})\|_0$.

if tol $\leq$ TOL **then**,

Compute Λ with (10).

Update Loc to be the location of 1-entries of Λ.

Set TOL $=$ tol.

end if

$k = k + 1$.

end while

return $\frac{x^k}{\|x^k\|}$.

VII. Conclusion

Results in [3] demonstrate that reconstruction from 1-bit compressive sensing measurements can be significantly improved if the appropriate measurement model is used in the reconstruction. Specifically, 1-bit measurements eliminate amplitude information, and therefore the signal can only be recovered within a positive scalar factor. Constraining the reconstruction to be on the unit sphere resolves this ambiguity and significantly reduces the reconstruction search space. Similarly, treating each measurement as a constraint instead of a value to be matched in a mean-squared sense allows exploiting consistent reconstruction principles. Results in [3] also demonstrate that both contributions significantly improve the reconstruction performance from 1-bit measurements.

[2] considers the problem of estimating a sparse signal from a set of quantized, Gaussian noise corrupted measurements, where each measurement corresponds to an interval of values. The authors in [2] give two methods for solving this problem, each based on minimizing a differentiable convex function plus an regularization term. According to the paper, using presented methods, compressed sensing can be carried out even when the quantization is very coarse, e.g., 1 or 2 bits per measurement. Numerical simulations show that both methods work relatively well, with the first method outperforming the second one for coarsely quantized measurements.

In addition to above we discussed a method robust method [5] which proposes an iterative method for detecting the sign flips in measurements and recovering the signals from correct measurements, this method is shown through numerical simulations to obtain better results in both finding the noisy measurements and recovering the signals, even when there are a lot of sign flips in the measurements. We also discussed a support estimation technique [4] using sketching for sparse vectors which more general in its modelling of uncertainties. It took into account the possibility of outliers along with gaussian noise. Also the models used by the authors can be extended to other noise specifications and also to the case of 'missing data'.

References

[1] D. Donoho, Compressed Sensing, IEEE Trans. Info. Theory , vol. 52,no. 4, pp. 12891306, Sep. 2006

[2] Zymnis, A.; Boyd, S.; Candes, E., "Compressed Sensing With Quantized Measurements," Signal Processing Letters, IEEE , vol.17, no.2, pp.149,152, Feb. 2010.

[3] Boufounos, P.T.; Baraniuk, R.G., "1-Bit compressive sensing," Information Sciences and Systems, 2008. CISS 2008. 42nd Annual Conference on , vol., no., pp.16,21, 19-21 March 2008.

[4] Haupt, J.; Baraniuk, R., "Robust support recovery using sparse compressive sensing matrices," Information Sciences and Systems (CISS), 2011 45th Annual Conference on , vol., no., pp.1,6, 23-25 March 2011.

[5] Ming Yan; Yi Yang; Osher, S., "Robust 1-bit Compressive Sensing Using Adaptive Outlier Pursuit," Signal Processing, IEEE Transactions on , vol.60, no.7, pp.3868,3875, July 2012

[6] E. J. Cand'es, Compressive sampling, in Proc. International Congress of Mathematicians, vol. 3, Madrid, Spain, 2006, pp. 14331452.

[7] E. T. Hale, W. Yin, and Y. Zhang, A fixed-point continuation method for '1-regularized minimization with applications to compressed sensing. Rice University, CAAM Tech. Report TR07-07, July 2007. [

[8] L. Jacques, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk, Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors, Apr. 2011 [Online]. Available: http://arxiv.org/abs/1104.3160.

[9] M. Charikar, K. Chen, and M. Farach-Colton, Finding frequent items in data streams, in Proceedings of the 29th International Colloquium on Automata, Languages and Programming, 2002.