One-Bit Compressed Sensing Using Smooth Measure of ℓ_0 Norm

Sina Alemohammad[†], and Arash Amini[†]
[†]Sharif Universiy of Technology
Emails: alemohammad_s@sharif.edu, aamini@sharif.edu

***Abstract*—Quantization of signals and parameters happens in all digital data acquisition devices. It is commonly regarded as a non-ideality of the system, and shall be taken into account when designing or analyzing a system. The topic of one-bit compressed sensing studies the effect of quantization in the extreme case where the samples are quantized with only one bit, i.e., the sign bit. The recovery of a sparse signal based on one-bit measurements is widely accomplished via thresholding methods or variants of ℓ_1-minimization techniques. In this paper, we introduce a recovery method arising from smoothing directly the ℓ_0 pseudo-norm. While we numerically verify the superior performance of the proposed method compared to the state-of-the-art techniques in our simulations, we briefly discuss the convergence analysis of this method.**

***Index Terms*—Compressed Sensing, One-bit measurements, Quantization.**

I. Introduction

The framework of compressed sensing investigates the sampling and reconstruction of high-dimensional signals $\mathbf{x} \in \mathbb{R}^N$ (or $\mathbb{C}^N$) using low-dimensional linear measurements $\mathbf{y}_{M\times 1} = \mathbf{\Phi}_{M\times N}\mathbf{x}_{N\times 1}$ with $M < N$ [1], [2]. The under-determined system of equations might have a unique K-sparse solution $\mathbf{x} \in \Sigma_K$ where $\Sigma_K := \{\mathbf{x} \in \mathbb{R}^N : \|\mathbf{x}\|_0 = |\mathrm{supp}(\mathbf{x})| \leq K\}$, if $\mathbf{\Phi}$ satisfies certain constraints. The sparsest solution of this system can be ideally recovered via

$$\mathbf{x}^* = \underset{\mathbf{u}\in\mathbb{R}^N}{\arg\min}\ \|\mathbf{u}\|_0\ \ s.t.\ \mathbf{y} = \mathbf{\Phi}\mathbf{u}. \tag{1}$$

In conventional compressed sensing, the measurements are assumed to be continuous and real-valued. In practical digital systems, however, a quantization stage is always present, that maps real-valued quantities into a finite set. In recent years, some theoretical and practical results in quantized compressed sensing are achieved [3], [4].

An extreme case of quantization is the one-bit quantization scenario, in which a sparse signal $\mathbf{x} \in \Sigma_K$ is encoded as

$$\bar{\mathbf{y}} = \mathrm{sign}(\mathbf{\Phi}\mathbf{x}) \in \{\pm 1\}^M \tag{2}$$

where $\mathbf{\Phi} \in \mathbb{R}^{M\times N}$ is the measurements matrix, and the observed vector $\bar{\mathbf{y}}$ is a signed version of the continuous-valued measurements. On one hand, as the sign measurements carry less information compared to continuous-valued measurements, generally, larger number of sign measurements are required for similar reconstruction quality. One the other hand, sign measurements are more robust against nonlinear distortions. Thus, distributing the measurement budget into multiple sign measurements could be justifiable.

The study of one-bit compressed sensing was initiated in [5] and gained more attention in [6] and [7]. The known reconstruction techniques from one-bit measurements are either greedy methods (e.g., BIHT) or ℓ_1 minimization techniques adapted to sign measurements. In this paper, we introduce a new method based on smoothed ℓ_0-norm. The rest of the paper is organized as follows: In Section II we explain the model in this paper. In Section III, we propose our algorithm, and in Section IV, we briefly discuses its convergence. The simulation results are reported in Section V.

II. Problem Model and Related Works

In one-bit compressed sensing, we would like to recover a sparse signal $\mathbf{x} \in \Sigma_K$ from the sign of a set of linear measurements (2). Therefor, we are looking for the solution of the following problem:

$$\mathbf{x}^* = \underset{\mathbf{u}\in\mathbb{R}^N}{\arg\min}\ \|\mathbf{u}\|_0\ \ s.t.\ \bar{\mathbf{y}} = \mathrm{sign}(\mathbf{\Phi}\mathbf{u}). \tag{3}$$

The obvious drawback of one-bit compressed sensing is the loss of norm (amplitude) information. With slightly different sampling strategies, the extraction of norm is also possible [8], [9]. For the purpose of this paper, we restrict our signal $\mathbf{x}$ to the unit ball

$$S^{N-1} := \{\mathbf{x} \in \mathbb{R}^N : \|\mathbf{x}\|_2 = 1\} \tag{4}$$

and solve the following optimization problem

$$\mathbf{x}^* = \underset{\mathbf{u}\in S^{N-1}}{\arg\min}\ \|\mathbf{u}\|_0\ \ s.t.\ \bar{\mathbf{y}} = \mathrm{sign}(\mathbf{\Phi}\mathbf{u}). \tag{5}$$

Here, we search for $\mathbf{u}$ on the unit ball that has the same sign information as the original $\mathrm{sign}(\mathbf{\Phi}\mathbf{x})$. The first reconstrcution gaurantee is provided in [6] when $\mathbf{\Phi}$ is a Gaussian ensemble:

Theorem 1. Fix $0 \leq \beta \leq 1$ and $\varepsilon > 0$. If the entries of the measurement matrix $\mathbf{\Phi}$ follow the Gaussian distribution $\mathcal{N}(0,1)$ and the number of measurements M satisfies

$$M > \tfrac{2}{\varepsilon}\big(2K\log(N) + 4K\log(\tfrac{17}{\varepsilon}) + \log(\tfrac{1}{\beta})\big), \tag{6}$$

then, for all $\mathbf{x}, \widehat{\mathbf{x}} \in \Sigma_K^* := \Sigma_K \cap S^{N-1}$ we have that

$$\|\mathbf{x} - \widehat{\mathbf{x}}\|_2 > \varepsilon \implies \mathrm{sign}(\mathbf{\Phi}\mathbf{x}) \neq \mathrm{sign}(\mathbf{\Phi}\widehat{\mathbf{x}}) \tag{7}$$

with probability greater that $1-\beta$.

This result implies that the solution to (5) with sufficiently large number of measurements ($M \geq \mathcal{O}(\frac{K\log(N)}{\varepsilon})$) is controllably close to the original signal. But the optimization problem (5) is NP-Hard in general. For a convex relaxation, it is common to replace $\|\mathbf{x}\|_0$ by $\|\mathbf{x}\|_1$. In [5], a renormalized fixed point iteration (RFPI) algorithm based on the following minimization problem is proposed:

$$\mathbf{x}^* = \underset{\mathbf{u}\in\mathbb{R}^N}{\arg\min}\ \|\mathbf{u}\|_1 + \lambda\|(\bar{\mathbf{Y}}\mathbf{\Phi}\mathbf{u})_-\|_2^2\ \ s.t.\ \|\mathbf{u}\|_2 = 1 \tag{8}$$

where in order to enforce consistency for $\text{sign}(\mathbf{\Phi}\mathbf{x}) = \bar{\mathbf{y}}$, the term $\lambda\|(\bar{\mathbf{Y}}\mathbf{\Phi}\mathbf{u})_-\|_2^2$ is added. Here, $(t)_-$ stands for $\frac{|t|-t}{2}$ that only retains the negative values. Besides, $\bar{\mathbf{Y}} = \text{diag}(\bar{\mathbf{y}})$. The approach of [5] for solving (8) is based on gradient decent and the fixed point method. In [6], the reconstruction is reformulated as

$$\mathbf{x}^* = \underset{\mathbf{u}\in\mathbb{R}^N}{\arg\min}\ \|(\bar{\mathbf{Y}}\mathbf{\Phi}\mathbf{u})_-\|_1\ \ s.t.\ \|\mathbf{u}\|_2 = 1, \|\mathbf{u}\|_0 = K. \tag{9}$$

To find the minimizer, a Binary Iterative Hard Thresholding technique is proposed in [6]. In each iteration of IBHT, we move in the direction of the gradient of $\|(\bar{\mathbf{Y}}\mathbf{\Phi}\mathbf{u})_-\|_1$, and then, project the result onto Σ_K^*. While BIHT is a fast and simple method, due to its greedy nature, it might not converge to the optimal solution of (5).

The first computationally tractable method with provable reconstruction guarantees was proposed in [7] by solving

$$\mathbf{x}^* = \underset{\mathbf{u}\in\mathbb{R}^N}{\arg\min}\ \|\mathbf{u}\|_1\ s.t.\ \|\mathbf{\Phi}\mathbf{u}\|_1 = M, \bar{\mathbf{Y}}\mathbf{\Phi}\mathbf{u} \geq 0. \tag{10}$$

Unlike RFPI and BHIT, the optimization problem (10) is convex and can be cast as a linear programming problem. However, the guaranteed performance of this method in [7] requires high sample complexity of $M = \mathcal{O}(\frac{K\log N}{\varepsilon^5})$.

III. The Proposed Algorithm

A. Main Idea

In this paper, we introduce a new technique to solve the problem in (5) without convex relation. Our technique borrows some ideas from the smoothed ℓ_0-norm minimization (Sℓ_0) which was first introduced in [10] for solving ((1)); therefore, we call out method One-Bit Sℓ_0. This approach is based on approximating the ℓ_0 norm with a differentiable and continuous function so that we can apply the gradient descent method to find its minimizer. Consider a set of functions $H_\delta : \mathbb{R}^N \longrightarrow \mathbb{R}^+$ that approximate the ℓ_0-norm; here, δ determines the approximation quality: the smaller the δ, the better the H_δ as a substitute for the ℓ_0 norm. For the sake of simplicity, we consider a separable form for H_δ as

$$H_\delta(\mathbf{x}) = \sum_{i=1}^{N} h_\delta(x_i), \tag{11}$$

where $\mathbf{x} = [x_1, \dots, x_N]^T \in \mathbb{R}^N$, and

$$\lim_{\delta\to 0} h_\delta(x) = \begin{cases} 1, & x \neq 0, \\ 0, & x = 0. \end{cases} \tag{12}$$

This property shows that

$$\lim_{\delta\to 0} H_\delta(\mathbf{x}) = \|\mathbf{x}\|_0. \tag{13}$$

We further restrict H_δ by imposing conditions on h_δ. The point is to facilitate the convergence of the gradient descent method in the minimization task.

Assumption 1. Let the one-dimensional function $f : \mathbb{R} \longrightarrow [0, 1]$ be such that

1) f is symmetric and unimodal,
2) $f(x) = 1 \iff x = 0$,
3) $f'(0) = 0$,
4) $f''(0) < 0$, and
5) $\lim_{|x|\to\infty} f(x) = 0$.

Then, we say that the family $\{f_\delta\}_{\delta\in\mathbb{R}^+}$ with $f_\delta(x) = f(\frac{x}{\delta})$ satisfies Assumption 1.

It is straightforward to conclude that if $\{f_\delta\}_{\delta\in\mathbb{R}^+}$ satisfies Assumption 1, then, f_δ converges to the Kronecker delta function as $\delta \longrightarrow 0$:

$$\lim_{\delta\to 0} f_\delta(x) = \begin{cases} 0, & x \neq 0, \\ 1, & x = 0. \end{cases} \tag{14}$$

It is easy to check that the class of Gaussian functions

$$f_\sigma(x) = \exp(-\frac{x^2}{2\delta^2}) \tag{15}$$

satisfies Assumption 1. Indeed, we use this class in this paper. However, the same approach can be implemented using other choices that satisfy Assumption 1.

If we define $h_\delta(x) = 1 - f_\delta(x)$, we have that

$$\lim_{\delta\to 0} h_\delta(x) = 1 - \lim_{\delta\to 0} f_\delta(x) = \begin{cases} 1, & x \neq 0, \\ 0, & x = 0, \end{cases} \tag{16}$$

which establishes the ℓ_0-norm approximating property

$$\|\mathbf{x}\|_0 \approx \sum_{i=1}^{N} h_\delta(x_i) = N - \sum_{i=1}^{N} f_\delta(x_i) \tag{17}$$

Now, the optimization problem in (5) can be smoothed as

$$\mathbf{x}_\delta^* = \underset{u\in\mathbb{R}^N}{\arg\min}\ N - \sum_{i=1}^{i=N} f_\delta(u_i)\ \ s.t.\ \bar{\boldsymbol{y}} = \text{sign}(\mathbf{\Phi}\mathbf{u}), \|\mathbf{u}\|_2 = 1, \tag{18}$$

or equivalently, by defining $F_\delta : \mathbb{R}^N \longrightarrow \mathbb{R}^+$ as $F_\delta(\mathbf{x}) = \sum_{i=1}^{i=N} f_\delta(x_i)$, the problem (18) can be formulated as

$$\mathbf{x}_\delta^* = \underset{u\in\mathbb{R}^N}{\arg\max}\ F_\delta(\mathbf{u})\ \ s.t.\ \bar{\boldsymbol{y}} = \text{sign}(\mathbf{\Phi}\mathbf{u}), \|\mathbf{u}\|_2 = 1. \tag{19}$$

The differentiability of f_δ, and in turn $F_\delta(\mathbf{u})$, allows us to use the gradient ascent method to find a local maximizer of (19), while it was not possible for the non-smooth cost in (5). Nevertheless, the result of (19) depends on δ. As δ decreases, H_δ provides a better approximation of $\|\mathbf{x}\|_0$ while the non-convexity of F_δ and the number of its local maxima increases.

The idea to avoid local maxima is to use a technique for optimizing none-convex functions known as Graduate None-Convexity (GNC) [10]. In this technique we use a sequence $\delta_1, \delta_2, ..., \delta_L$ for the approximation parameters of the smoothed functions. We initialize δ with large value δ_1 so that F_δ becomes convex. The maximizer of (19) with this choice is

Algorithm 1

Input:
Measurement Matrix $\mathbf{\Phi}$
Quatized measurements $\bar{y}$
The number external and internal loop iterations, I and J
Constance c for making decreasing sequence of δ
Step size μ
Output:
The sparse estimation x^*
procedure ONE-BIT Sℓ_0($\mathbf{\Phi}$,$\bar{y}$,I,J,c,μ)
  Initialize $\mathbf{u}^0$ by solving (20) for a random $\tilde{\mathbf{x}} \in \mathbb{R}^N$
  for $\{i = 1, ..., I\}$ **do**
    $\delta_i = c\delta_{i-1}$
    $\mathbf{u}^0_{\delta_i} = \mathbf{u}^J_{\delta_{i-1}}$
    for $\{j = 1, ..., J\}$ **do**
      $\widehat{\mathbf{u}}^j_{\delta_i} = \mathbf{u}^{j-1}_{\delta_i} - \frac{\mu}{2}[x_1 \exp(-\frac{x_1^2}{2\delta_i^2}), ..., x_N \exp(-\frac{x_N^2}{2\delta_i^2})]^T$
      $\bar{\mathbf{u}}^j_{\delta_i} = \underset{\bar{\mathbf{u}}^j_{\delta_i} \in R^N}{\arg\min} \|\bar{\mathbf{u}}^j_{\delta_i} - \widehat{\mathbf{u}}^j_{\delta_i}\|_2^2 \;\; s.t. \; \bar{Y}\mathbf{\Phi}\bar{\mathbf{u}}^j_{\delta_i} \geq 0$
      $\mathbf{u}^j_{\delta_i} = \frac{\bar{\mathbf{u}}^j_{\delta_i}}{\|\bar{\mathbf{u}}^j_{\delta_i}\|_2}$
    end for
  end for
  return $\mathbf{u}^J_{\delta_I}$
end procedure

$\mathbf{x}_{\delta_1}$. Next, we gradually decrease δ to δ_i ($i > 1$) to improve the approximation of $\|\mathbf{x}\|_0$ in (19). For each δ_i, we initialize the maximization problem of F_{δ_i} with the maximizer of (19) for δ_{i-1} (i.e., $\mathbf{x}_{\delta_{i-1}}$) and obtain $\mathbf{x}_{\delta_i}$. As we gradually shift from a convex problem to a non-convex problem, the method might succeed in avoiding the local maxima.

B. Gradient Projection

For each δ, we apply the Gradient Projection [11] method to solve (19). In each iteration of GP, the direction of gradient ascent is projected back onto the feasible set determined by constraints in (19). Mathematically, if $\tilde{\mathbf{x}} = \mathbf{x} + \mu_j \nabla F_{\delta_i}(\mathbf{x})$, the projection onto the feasible set is achieved by

$$\mathbf{x} = \underset{u \in \mathbb{R}^N}{\arg\min} \; \|\mathbf{u} - \tilde{\mathbf{x}}\|_2^2 \;\; s.t. \; \bar{Y}\mathbf{\Phi}\mathbf{u} \geq 0, \|\mathbf{u}\|_2 = 1. \tag{20}$$

Unfortunately, (20) is a non-convex problem. Instead, we drop the constraint $\|\mathbf{u}\|_2 = 1$ to achieve convex relaxation as

$$\mathbf{x} = \underset{u \in \mathbb{R}^N}{\arg\min} \; \|\mathbf{u} - \tilde{\mathbf{x}}\|_2^2 \;\; s.t. \; \bar{Y}\mathbf{\Phi}\mathbf{u} \geq 0 \tag{21}$$

We solve (21), and then, apply the constraint $\|\mathbf{u}\|_2 = 1$ by projecting the outcome onto the unit ball.

For the completion of the GP method, we have to derive the gradient of $F_\delta(\mathbf{x})$ with respect to $\mathbf{x}$, which is

$$\nabla F_\delta(\mathbf{x}) = -\frac{1}{2\delta^2}\left[x_1 \exp(-\tfrac{x_1^2}{2\delta^2}), ..., x_N \exp(-\tfrac{x_N^2}{2\delta^2})\right]^T \tag{22}$$

C. Initialization

As explained earlier, we initialize the algorithm with a large δ value. Since f_δ satisfies Assumption 1, for $\delta \gg 1$ we can write that

$$F_\delta(\mathbf{x}) = \sum_{i=1}^{N} f_\delta(x_i) \approx \sum_{i=1}^{N}(1 + f''(0)\tfrac{x_i^2}{\delta^2}) = N + \tfrac{\gamma}{\delta^2}\|\mathbf{x}\|_2^2, \tag{23}$$

where we used a Taylor series representation of $f_\delta(\mathbf{x})$ and neglected the higher order terms. Here γ stands for $f''(0)$. As $\|\mathbf{x}\|_2 = 1$ in our feasible set, the initialization of the algorithm becomes the answer to the following problem:

$$\mathbf{x} = \underset{u \in \mathbb{R}^N}{\arg\min} \; N + \frac{\gamma}{\delta^2} \;\; s.t. \; \bar{Y}\mathbf{\Phi}\mathbf{u} \geq 0, \|\mathbf{u}\|_2 = 1, \tag{24}$$

which is simply any $\mathbf{x}_0 \in \mathbb{R}^N$ in the feasible set. Alternatively, $\mathbf{x}_0$ can be set as the solution to (20) for any random $\tilde{\mathbf{x}} \in \mathbb{R}^N$.

D. The Final Algorithm

A formal description of the One-Bit Sℓ_0 is provided in Algorithm 1 which consists of two loops. The external loop is designed to iterate over δ_i for using the GNC method. The internal loop is aimed for finding the maximizer of $F_\delta(\mathbf{u})$ using the GP technique. In the following, based on Theorem 1, we assume a sample complexity of $M = \mathcal{O}(\frac{K \log N}{\varepsilon})$ for the proposed Algorithm 1.

Remark 1. The decreasing sequence of δ can be set as $\delta_i = c^{(i-1)}\delta_{i-1}$ for $i \geqslant 1$. δ_0 should be chosen large enough so that it behaves like ∞. As $\mathbf{x}_0$ is bound to be on the unit sphere, we have that $\|\mathbf{x}_0\|_\infty \leq 1$; therefore, $\exp(\frac{\mathbf{x}_{0_i}^2}{2\delta_0^2}) \geq \exp(\frac{1}{2\delta_0^2})$. By setting $\delta_0 = 5$, we have that $\exp(\frac{1}{2\delta_0^2}) = \exp(\frac{1}{50}) > 0.98 \approx 1$. Hence, $\delta_0 = 5$ is roughly a good initialization choice.

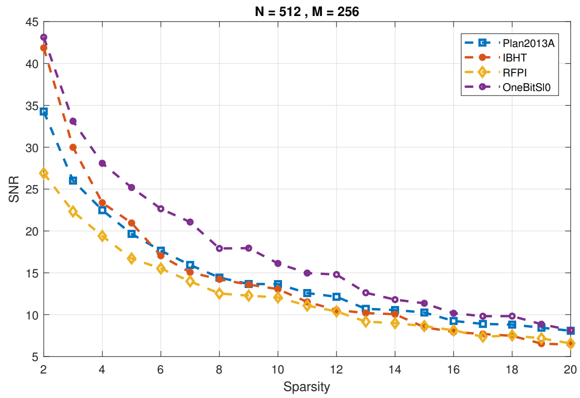
Remark 2. In the gradient ascent method, μ_i should be chosen small enough in order to prevent the algorithm from divergence caused by the variation of $F_{\delta_i}(\mathbf{x})$ with respect to the variation of $\mathbf{x}$. As δ_i decreases in the external loop of the algorithm, the variation of the function $F_{\delta_i}(\mathbf{x})$ grow larger. Consequently, a smaller μ_i should be used. A suitable choice that is well adapted to the behavior of $F_{\delta_i}(\mathbf{x})$ with respect to δ_i is $\mu_i = \delta_i^2\mu$, where μ is a constant given by user. As a result, the gradient ascent of the algorithm simplifies to

$$\mathbf{u}^j_{\delta_i} = \mathbf{u}^{j-1}_{\delta_i} - \frac{\mu}{2}[x_1 \exp(-\tfrac{x_1^2}{2\delta_i^2}), ..., x_N \exp(-\tfrac{x_N^2}{2\delta_i^2})]^T \tag{25}$$

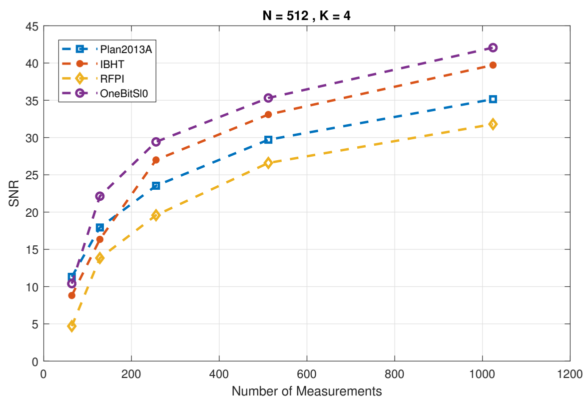
IV. Convergence Analysis

In this section, we discuss the convergence of our algorithm. Note that in the procedure of the algorithm, we use a decreasing sequence of δ in the external loop to gradually create a good approximation of $\|\mathbf{x}\|_0$ and each δ results in $\mathbf{x}_\delta$. For the convergence analysis, we assume that each $\mathbf{x}_\delta$ is the global maximizer of the $F_\delta(\mathbf{x})$ and the gradient ascent algorithm in the internal loop has not been trapped in the local maxima. Assume that $\mathbf{x}^* \in \Sigma_K$ is the solution of the original problem (5) and $\widehat{\mathbf{x}}$ is the solution of (19) for $\delta \longrightarrow 0$. As $\mathbf{x}_\delta$ is the global maximizer of (19), we have

$$\lim_{\delta \to 0} F_\delta(\widehat{\mathbf{x}}) \geq \lim_{\delta \to 0} F_\delta(\mathbf{x}^*) \tag{26}$$



(a)



(b)

Fig. 1: SNR values for different methods, number of measurements and sparsity levels

By considering the definition of $F_\delta(\mathbf{x})$

$$N - \|\widehat{\mathbf{x}}\|_0 \geq N - \|\mathbf{x}^*\|_0 \Rightarrow \|\widehat{\mathbf{x}}\|_0 \leq \|\mathbf{x}^*\|_0 = K \tag{27}$$

Since $\mathbf{x}^*$ is the solution of (5), any $\mathbf{x}$ in the feasible set has a sparsity more than K. As a result

$$\|\widehat{\mathbf{x}}\|_0 \geq K \tag{28}$$

Combining (27) and (28), we will have

$$\|\widehat{\mathbf{x}}\|_0 = K \tag{29}$$

Based on the above conclusion, $\widehat{\mathbf{x}} \in \Sigma_K$ and it is in the feasible set, as it is the solution of (19) and satisfies the constraint. As a result, $\widehat{\mathbf{x}}$ is the solution of original problem (5) by definition.

V. Numerical Experiments

In this section, the performance of the One-Bit $S\ell_0$ is evaluated empirically through simulation, and is compared to RFPI [5], IBHT [6] and Plan2013A [7] algorithms. The perfomance of the algorithms are compared by calculating $\text{SNR} = 20\log(\frac{\|\mathbf{x}\|_2}{\|\mathbf{x}-\mathbf{x}^*\|_2})$, where $\mathbf{x}^*$ is the output of algorithms as the estimation of $\mathbf{x}$. For the first experiment, We have generated measurement matrix $\mathbf{\Phi} \in \mathcal{N}^{256\times512}(0,1)$ and for different sparsity levels $2 \leq K \leq 20$. The results of this experiments is shown at Fig.1a.

In the second experiment, we have fixed $N = 256$ and $K = 4$ and we have generated different measurement matrix $\mathbf{\Phi} \in \mathcal{N}^{M\times256}(0,1)$ for number of measurements $M = [64, 128, 256, 512, 1024]$.The results of this experiments is shown at Fig.1b.

The reported SNR is the mean of calculated SNR for the algorithms over 100 randomly generated signals of $\mathbf{x}$ for each sparsity level in Fig.1a and each number of measurements in Fig.1b. In both experiments; our algorithm outperforms the state-of-the-art algorithms regarding reconstruction SNR.

VI. Conclusion

In this paper, we propose a new approach to recover a sparse vector from a set of one-bit quantized linear measurements by using a soft measure of ℓ_0 norm through maximizing a non-convex function as an estimation of ℓ_0 norm which requires $M = \mathcal{O}(\frac{K\log N}{\varepsilon})$ measurements. In the numerical experiments, we illustrated that our method surpasses existing state-of-the-art methods in terms of SNR.

References

[1] D. L. Donoho, "Compressed sensing," *IEEE Transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
[2] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE signal processing magazine*, vol. 25, no. 2, pp. 21–30, 2008.
[3] A. Zymnis, S. Boyd, and E. Candes, "Compressed sensing with quantized measurements," *IEEE Signal Processing Letters*, vol. 17, no. 2, pp. 149–152, 2010.
[4] J. N. Laska, P. T. Boufounos, M. A. Davenport, and R. G. Baraniuk, "Democracy in action: Quantization, saturation, and compressive sensing," *Applied and Computational Harmonic Analysis*, vol. 31, no. 3, pp. 429–443, 2011.
[5] P. T. Boufounos and R. G. Baraniuk, "1-bit compressive sensing," in *Information Sciences and Systems, 2008. CISS 2008. 42nd Annual Conference on.* IEEE, 2008, pp. 16–21.
[6] L. Jacques, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk, "Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors," *IEEE Transactions on Information Theory*, vol. 59, no. 4, pp. 2082–2102, 2013.
[7] Y. Plan and R. Vershynin, "One-bit compressed sensing by linear programming," *Communications on Pure and Applied Mathematics*, vol. 66, no. 8, pp. 1275–1297, 2013.
[8] K. Knudson, R. Saab, and R. Ward, "One-bit compressive sensing with norm estimation." *IEEE Trans. Information Theory*, vol. 62, no. 5, pp. 2748–2758, 2016.
[9] U. S. Kamilov, A. Bourquard, A. Amini, and M. Unser, "One-bit measurements with adaptive thresholds," *IEEE Signal Processing Letters*, vol. 19, no. 10, pp. 607–610, 2012.
[10] A. Blake and A. Zisserman, *Visual reconstruction*. MIT press, 1987.
[11] D. P. Bertsekas, *Nonlinear programming*. Athena scientific Belmont, 1999.
[12] H. Mohimani, M. Babaie-Zadeh, and C. Jutten, "A fast approach for overcomplete sparse decomposition based on smoothed ℓ^0 norm." *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 289–301, 2009.