

Sharif University of Technology Scientia Iranica Transactions D: Computer Science & Engineering and Electrical Engineering http://scientiairanica.sharif.edu

 $Research \ Note$

Bayesian hypothesis testing for one bit compressed sensing with sensing matrix perturbation

H. Zayyani^{a,*}, M. Korki^b, and F. Marvasti^c

a. Department of Electrical and Computer Engineering, Qom University of Technology, Qom, Iran.

b. School of Software and Electrical Engineering, Swinburne University of Technology, Hawthorn, 3122 Australia.

 $c. \ Department \ of \ Electrical \ Engineering, \ Sharif \ University \ of \ Technology, \ Tehran, \ Iran.$

Received 9 January 2016; received in revised form 16 November 2016; accepted 11 March 2017

KEYWORDS

Compressed sensing; One bit measurements; Bayesian hypothesis test; ML estimator. **Abstract.** This paper proposes a low-computational Bayesian algorithm for noisy sparse recovery in the context of one-bit compressed sensing with sensing matrix perturbation. The proposed algorithm which is called BHT-MLE comprises a sparse support detector and an amplitude estimator. The support detector utilizes Bayesian hypothesis test, while the amplitude estimator uses an ML estimator obtained by solving a convex optimization problem. Simulation results show that Bayesian hypothesis testing in combination with the ML estimator has more reconstruction accuracy than that of only an ML estimator and also has less computational complexity.

© 2018 Sharif University of Technology. All rights reserved.

1. Introduction

The one-bit Compressed Sensing (CS) which is the extreme case of quantized compressed sensing [1] has been extensively investigated recently [2-11]. According to CS theory, a sparse signal can be reconstructed from a number of linear measurements which could be much smaller than the signal dimension [12-13]. Classical CS neglects the quantization process, and assumes that the measurements are real continuous valued. However, in practice, the measurements should be quantized to some discrete levels. This is known as quantized compressed sensing [1]. In the extreme case, there are only two discrete levels. This is called one-bit compressed sensing and has gained much attention in the research community these years [2-11]. In the one-

doi: 10.24200/sci.2017.4374

bit compressed sensing framework, it is proved that accurate and stable recovery can be achieved using only the sign of linear measurements [5].

Many algorithms have been developed for onebit compressed sensing. A Renormalized Fixed-Point Iteration (RFPI) algorithm, which is based on l^1 norm minimization, has been presented in [2]. Also, a Matching Sign Pursuit (MSP) algorithm has been proposed in [3]. A Binary Iterative Hard Thresholding (BIHT) algorithm has been introduced in [5], shown to have better performance than MSP. Moreover, a Restricted-Step Shrinkage (RSS) algorithm which has been devised in [4] has provable convergence guarantees.

In addition to noise-free settings, there may be noisy sign measurements. In this case, we may be encountered with sign flips which will worsen the performance. In [6], an Adaptive Outlier Pursuit (AOP) algorithm is developed to detect the sign flips and reconstruct the signals with very high accuracy even when there are a large number of sign flips [6]. Moreover, Noise-Adaptive RFPI (NARFPI) algorithm combines the idea of RFPI and AOP [7]. In addition

^{*.} Corresponding author. Tel.: +98 253 2805366 E-mail addresses: zayyani@qut.ac.ir (H. Zayyani); mkorki@swin.edu.au (M. Korki); marvasti@sharif.edu (F. Marvasti).

Plan and Vershynin, [8] propose a convex approach to solve the problem. Recently, a one-bit Bayesian compressed sensing [14] and a MAP approach [15] have been developed for solving the problem. Also, Zhu et al. [16] focus on the ML estimation of a vector parameter from sign measurements with sensing matrix perturbation.

In this paper, similar to some sparse recovery algorithms (see e.g., [17-18]), we propose a two-step approach to one-bit compressed sensing with sensing matrix perturbation. The first step is the support detection, and the second step is the amplitude es-In the first step, inspired from [18], a timation. Bayesian hypothesis test is used to detect the active samples of sparse vector. On the other hand, for amplitude recovery of active samples of the sparse vector, similar to [16], we utilize an ML estimator. Compared to [16] which ignores the sparsity, the main advantage of our proposed algorithm is to exploit the sparsity of the sparse vector. Moreover, detecting the active samples in the first step reduces the complexity of the optimization problem of ML estimator. Our simulation results verify that exploiting the sparsity enhances the reconstruction performance of the sparse vector.

The rest of this paper is organized as follows. Section 2 states the problem formulation. Section 3 introduces our proposed algorithm, including the two steps of support and amplitude recovery. Simulation results are presented in Section 4. Finally, conclusions are drawn in Section 5.

2. Problem formulation

Consider a sparse vector \boldsymbol{s} which is observed via a corrupted sensing matrix as follows:

$$\boldsymbol{x} = (\boldsymbol{A} + \boldsymbol{E})^T \boldsymbol{s} + \boldsymbol{n},\tag{1}$$

where $A \in \mathbb{R}^{m \times N}$ is a known sensing matrix, \boldsymbol{E} is an error random matrix whose elements are i.i.d. with $e_{ij} \sim \mathcal{N}(0, \sigma_e^2)$, where σ_e^2 is viewed as the perturbation strength, and \boldsymbol{n} is the additive noise vector (independent of \boldsymbol{E}) with $\boldsymbol{n} \sim \mathcal{N}(0, \sigma_n^2 \boldsymbol{I})$. In the usual CS setting, one requires $N \ll m$, whereas in the 1-bit CS setting, $N \ge m$ is allowed.

In one-bit compressed sensing with sensing matrix perturbation, we aim to estimate sparse vector \boldsymbol{s} based on the sign of linear measurements, $\boldsymbol{y} = \text{sign}(\boldsymbol{x})$, which is:

$$\boldsymbol{y} = \operatorname{sign}(\boldsymbol{A}^T \boldsymbol{s}) + \boldsymbol{z}, \tag{2}$$

where $\boldsymbol{z} = \boldsymbol{E}^T \boldsymbol{s} + \boldsymbol{n}$ is called equivalent noise which is the sum of a multiplicative noise and an additive noise [16]. It can be simply shown that the variance of

noise
$$\boldsymbol{z} \sim \mathcal{N}(0, \sigma_n^2 \boldsymbol{I})$$
 is equal to [16]

$$\sigma_z^2 = ||\mathbf{s}||_2^2 \sigma_e^2 + \sigma_n^2.$$
(3)

Sparse vector \boldsymbol{s} is assumed to have a Bernoulli-Gaussian (BG) distribution, i.e., $s_j = q_j r_j$ where q_j is the activity of the *j*'th element of sparse vector and r_j is the amplitude of the element. Similar to some sparse recovery algorithms [17-18], the sparse recovery is equivalent to estimating both activity vector, $\boldsymbol{q} = [q_1, q_2, ..., q_m]^T$, and amplitude vector, $\boldsymbol{r} = [r_1, r_2, ..., r_m]^T$.

3. The proposed algorithm

The proposed algorithm is divided into two steps. In the first step, activity vector \boldsymbol{q} is estimated, which is called support detection. The second step of our proposed algorithm is called amplitude recovery. In this step, amplitude vector \boldsymbol{r} is estimated.

3.1. Support detection using Bayesian hypothesis testing

For determining the activity of the j'th element of sparse vector \mathbf{s} , two hypotheses are considered. The first hypothesis, H_{1j} , assumes that s_j is inactive, and the second hypothesis, H_{2j} , considers that s_j is active. Thus, we have:

$$H_{1j}: \begin{cases} y_1 = \operatorname{sign}(\boldsymbol{a}_1^T \boldsymbol{s}_{(-j)} + z_1), \\ y_2 = \operatorname{sign}(\boldsymbol{a}_2^T \boldsymbol{s}_{(-j)} + z_2), \\ \vdots \\ y_N = \operatorname{sign}(\boldsymbol{a}_N^T \boldsymbol{s}_{(-j)} + z_N), \end{cases}$$
(4)

 and

$$H_{2j}:\begin{cases} y_1 = \operatorname{sign}(\boldsymbol{a}_1^T \boldsymbol{s} + z_1), \\ y_2 = \operatorname{sign}(\boldsymbol{a}_2^T \boldsymbol{s} + z_2), \\ \vdots \\ y_N = \operatorname{sign}(\boldsymbol{a}_N^T \boldsymbol{s} + z_N), \end{cases}$$
(5)

where \mathbf{s}_{-j} is sparse vector \mathbf{s} with $s_j = 0$, \mathbf{a}_i^T is the *i*'th measurement vector, and y_i is the *i*'th sign measurement element. The Bayesian hypothesis test is:

$$q_{j} = \begin{cases} 0 & p(H_{1j} | \boldsymbol{y}) \ge p(H_{2j} | \boldsymbol{y}), \\ 1 & p(H_{1j} | \boldsymbol{y}) < p(H_{2j} | \boldsymbol{y}), \end{cases}$$
(6)

where $\boldsymbol{y} = [y_1, y_2, ..., y_N]^T$ is the sign measurement vector. Using the MAP rule, the activity rule based on the hypothesis test is:

$$pr(H_{1j})p(\boldsymbol{y}|H_{1j}) < pr(H_{2j})p(\boldsymbol{y}|H_{2j}).$$
 (7)

Assume that the prior probabilities are $pr(H_1j) = 1 - pr$ and $pr(H_{2j}) = pr$, where pr is the activity

probability. Also, assuming the independency of measurements y_i , we have:

$$p(y|H_{1j}) = \prod_{i=1}^{N} p(y_i|H_{1j}),$$

and:

$$p(\boldsymbol{y}|H_{2j}) = \prod_{i=1}^{N} p(y_i|H_{2j}).$$

The logarithm of both sides of Relation (5) with some manipulations leads to the following activity rule:

$$\sum_{i=1}^{N} \ln p(y_i | H_{2j}) - \sum_{i=1}^{N} \ln p(y_i | H_{1j})$$

> $Th = \ln\left(\frac{1 - Pr}{Pr}\right).$ (8)

In addition, we have:

$$p(y_i|H_{1j}) = \begin{cases} p(\boldsymbol{a}_i^T \boldsymbol{s}_{-j} + z_i > 0), & y_i = 1, \\ p(\boldsymbol{a}_i^T \boldsymbol{s}_{-j} + z_i \le 0), & y_i = -1, \end{cases}$$
(9)

which is equal to:

$$p(y_i|H_{1j}) = \Phi\left(\frac{y_i \boldsymbol{a}_i^T \boldsymbol{s}_{-j}}{\sigma_z}\right), \qquad (10)$$

where $\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{x^2}{2}} dx$ is the cumulative distribution function of the standard Gaussian distribution. Similarly, we have:

$$p(y_i|H_{2j}) = \Phi\left(\frac{y_i \boldsymbol{a}_i^T \boldsymbol{s}}{\sigma_z}\right).$$

Hence, the overall hypothesis test for determining the activity of the j'th element is as follows:

$$\sum_{i=1}^{N} \ln \left(\frac{\Phi(\frac{y_i \boldsymbol{a}_i^T \boldsymbol{s}}{\sigma_z})}{\Phi(\frac{y_i \boldsymbol{a}_i^T \boldsymbol{s}_{-j}}{\sigma_z})} \right) > Th, \tag{11}$$

where $Th = \operatorname{In}\left(\frac{1-pr}{pr}\right)$. Therefore, the overall support detection consists of (multiple) binary hypothesis testing instead of a single composite hypothesis testing. Replacing *m* binary hypothesis testing for determining the activity of each element renders a substantially lower computational complexity than using a composite hypothesis testing to search activity vector \boldsymbol{q} over 2^m possible activity vectors.

3.2. Amplitude recovery: An ML estimation

Having detected the support vector, the amplitude of sparse vector is estimated. If we remove the inactive locations of sparse vector \boldsymbol{s} , we have an amplitude vector $\boldsymbol{w} = [w_1, w_2, ..., w_K]^T$, where K is the number of active locations. If I_a represents a set of indexes corresponding to active locations, we have the following relations:

$$\begin{cases} y_1 = \operatorname{sign}(\boldsymbol{h}_1^T \boldsymbol{w} + z_1), \\ y_2 = \operatorname{sign}(\boldsymbol{h}_2^T \boldsymbol{w} + z_2), \\ \vdots \\ y_N = \operatorname{sign}(\boldsymbol{h}_N^T \boldsymbol{w} + z_N), \end{cases}$$
(12)

where $\mathbf{h}_i = \mathbf{a}_i(I_a)$ is the reduced measurement vector, and $I_a = [0 \cdots 1 \cdots 01 \cdots 0 \cdots]^T$ is the index of active elements of \mathbf{s} . Now, we have the problem of estimating non-sparse vector \mathbf{w} based on the sign measurements in Relation (12). This problem has been investigated in [16] and an ML estimator has been introduced. The likelihood of measurement data can be obtained as follows [16]:

$$p(\boldsymbol{y}; \boldsymbol{w}) = \prod_{i=1}^{N} \Phi\left(\frac{y_i \boldsymbol{h}_i^T \boldsymbol{w}}{\sigma_z}\right), \qquad (13)$$

and the log-likelihood is:

$$l(\boldsymbol{y}; \boldsymbol{w}) = \sum_{i=1}^{N} \ln \Phi \left(\frac{y_i \boldsymbol{h}_i^T \boldsymbol{w}}{\sigma_z} \right).$$
(14)

Maximizing the log-likelihood is equivalent to minimizing the negative log-likelihood. Consequently, the ML estimator is the solution to the following optimization problem:

$$P_{1}: \operatorname{Minimize}_{\boldsymbol{w} \in R^{K}} - \sum_{i=1}^{N} \ln \Phi\left(\frac{y_{i}\boldsymbol{h}_{i}^{T}\boldsymbol{w}}{\sqrt{||\boldsymbol{w}||_{2}^{2}\sigma_{e}^{2} + \sigma_{n}^{2}}}\right),$$
(15)

where the above optimization problem is non-convex and cannot be solved by steepest-descent or Newton's method [16]. In [16], it is shown that the above nonconvex optimization problem P_1 is equivalent to the following convex optimization problem:

$$P_{2} : \text{Minimiz } e_{v \in R^{K}} - \sum_{i=1}^{N} \ln\left(y_{i}\boldsymbol{h}_{i}^{T}\boldsymbol{v}\right)$$

Subject to $||\boldsymbol{v}||_{2}^{2} < \frac{1}{\sigma_{e}^{2}},$ (16)

where the optimal solution is obtained by:

$$\boldsymbol{w} = \frac{\sigma_n}{\sqrt{1 - \sigma_e^2 \left| \left| \boldsymbol{v} \right| \right|_2^2}} \boldsymbol{\nu}.$$
 (17)

Moreover, in [13], it is proved that if the unconstrained optimization problem be considered as follows:

$$P_3: \operatorname{Minimize}_{\boldsymbol{v} \in R^K} - \sum_{i=1}^N \ln \Phi(y_i \boldsymbol{h}_i^T \boldsymbol{v}), \qquad (18)$$

then the optimal point of problem P_1 in Relation (15) exists if and only if optimal point \boldsymbol{v}^* of problem

 P_3 satisfies constraint $||\boldsymbol{v}^*||_2^2 < \frac{1}{\sigma_c^2}$ [16]. Therefore, first, the unconstrained optimization problem P_3 should be solved, and then whether it satisfies the above-mentioned constraint must be checked to ensure whether the original ML estimation problem P_1 in Relation (15) has an optimal point [16] (for further details see [16]).

4. Simulation results

This section presents the simulation results. In the simulations, unknown sparse vector \mathbf{s} is drawn from a BG distribution with activity probabilities pr = 0.1 and pr = 0.2, and with variance of active, samples $\sigma_r^2 = 1$. To ensure that the norm is finite, the sparse vector is normalized to have unit norm. The size of the sparse vector is assumed to be m = 200. The sensing matrix elements are obtained from a standard Gaussian distribution with $a_{ij} \sim \mathcal{N}(0, 1)$. The error matrix or perturbation matrix elements are considered to be $e_{ij} \sim \mathcal{N}(0, \sigma_e^2)$ with $\sigma_e = 0.1$. Additive noise \mathbf{n} is regarded as Gaussian random variable with distribution $n_i \sim \mathcal{N}(0, \sigma_n^2)$, where $\sigma_n = 0.1$.

The proposed BHT-MLE method is compared with the ML estimation method [16], which is denoted as MLE. For the initialization of BHT-MLE and MLE, $\hat{s}_0 = A^{\dagger}y$ is used, where A^{\dagger} is the pseudo-inverse of matrix A. To calculate the threshold $Th = \ln\left(\frac{1-pr}{pr}\right)$ in Relation (11), pr needs to be estimated. Similar to parameter estimation methods in [19],

$$\hat{pr} = rac{\operatorname{Card}\left\{i \mid \mid s_i \mid \mid \alpha \sqrt{var\left(s\right)}\right\}}{m},$$

is used, where Card{.} is the cardinality operator, showing the number of elements of a set and α is a parameter that controls the false alarm. For initial iterations, we overestimate pr by choosing small value for $\alpha = 0.5$. At final iterations, we choose $\alpha \approx 3$. Extensive experimental studies demonstrate that α converges to its optimal value within 10 iterations. Hence, we use a linear increase of α as $\alpha^{(k)} = 1.2\alpha^{(k-1)}$, where k is the index of iteration. Similar to [16], we assume that variances σ_e^2 and σ_n^2 are known in advance. For solving unconstrained optimization problem, P_3 , similar to [16], the MATLAB fminunc function is used.

The Normalized Mean Square Error (NMSE) is utilized as a performance metric, which is defined as follows:

$$\text{NMSE} \triangleq 20 \log_{10} \left(\frac{||\boldsymbol{s} - \hat{\boldsymbol{s}}||_2}{||\boldsymbol{s}||_2} \right), \tag{19}$$

where \hat{s} is the estimate of true sparse vector s.

All the NMSEs are averaged over 100 Monte Carlo (MC) simulations. The number of binary measurements varies between 400 and 800. Figure 1 shows the NMSE performance versus the number of sign measurements for both BHT-MLE and MLE methods with pr = 0.1 and pr = 0.2. It is seen that the proposed BHT-MLE method, which utilizes Bayesian hypothesis testing, outperforms the conventional ML estimator (MLE) method by at least 5 dB gain because it exploits the activity information of the sparse vector provided by the first step of algorithm. Figure 2 depicts the average simulation time with respect to the number of sign measurements when pr = 0.1. It shows that the proposed method has lower computational complexity than MLE at least by a factor of two because it reduces the dimension of the optimization problem from m =200 to K = round(pr.m) = 20.



N: Number of sign measurments

Figure 1. Normalized mean square error of reconstructed sparse vector versus number of sign measurements.



Figure 2. Average simulation time versus number of sign measurements.

5. Conclusion

A new BHT-MLE algorithm was proposed for the noisy sparse signal recovery in one-bit compressed sensing with the sensing matrix perturbation framework. BHT-MLE algorithm consists of a Bayesian Hypothesis Testing (BHT) for support detection and an ML Estimator (MLE) for the amplitude estimation. In the support detection step, the proposed BHT-MLE algorithm uses a sequence of binary hypothesis tests. Simulation results in a special case verify that this BHT-based support detection approach improves the sparse reconstruction accuracy by at least 5 dB gain and reduces the computational complexity by a factor of two.

Acknowledgment

The first author wishes to thank Iran National Elite's Foundation for financial support.

References

- Zymnis, A., Boyd, S., and Candes, E. "Compressed sensing with quantized measurements", *IEEE Signal Processing Letters*, **17**(2), pp. 149-152 (2010).
- BouFounos, P. and Baraniuk, R. "1-bit compressive sensing", Proceeding 42nd Annu. Conf. Inf. Sci. Sys. (2008).
- BouFounos, P. "Greedy sparse signal reconstruction from sign measurements", *Proceeding 43rd Asilomar.* Conf. Signals, Syst., Comput. (Asilomar'09) (2009).
- Laska, J.N., Wen, Z., Yin, W., and Baraniuk, R. "Trust, but verify: fast and accurate signal recovery from 1-bit compressive measurements", *IEEE Trans* Signal Processing, 59(11), pp. 5289-5301 (2011).
- Jacques, L., Laska, J., Boufounos, P., and Baraniuk, R. "Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors", *IEEE Trans Inf. The*ory, 59(4), pp. 2082-2102 (2013).
- Yan, M., Yang, Y., and Osher, S. "Robust 1-bit compressive sensing using adaptive outlier pursuit", *IEEE Trans Signal Processing*, **60**(7), pp. 3868-3875 (2012).
- Movahed, A., Panahi, A. and Durisi, G. "A robust RFPI-based 1-bit compressive sensing reconstruction algorithm", Proceeding IEEE ITW (2012).
- Plan, Y. and Vershynin, R. "Robust 1-bit compressed sensing and sparse logistic regression: A convex programming approach", *IEEE Trans Inf. Theory*, 59(1), pp. 482-494 (2013).
- Zayyani, H., Korki, M., and Marvasti, F. "Dictionary learning for blind one bit compressed sensing", *IEEE* Signal Processing Letters, 23(2), pp. 187-191 (2016).

- Zayyani, H., Haddadi, F., and Korki, M. "Double detector for sparse signal detection from one-bit compressed sensing measurements", *IEEE Signal Process*ing Letters, 23(11), pp. 1637-1641 (2016).
- Zayyani, H., Korki, M., and Marvasti, F. "A distributed 1-bit compressed sensing algorithm robust to impulsive noise", *IEEE Communications Letters*, **20**(6), pp. 1132-1135 (2016).
- Candes, E.J. and Tao, T. "Near-optimal signal recovery from random projections: universal encoding strategies?", *IEEE Trans Inf. Theory*, **52**(12), pp. 5406-5425 (2006).
- Donoho, D.L. "Compressed sensing", *IEEE Trans Inf. Theory*, **52**(4), pp. 1289-1306 (2006).
- Li, F., Fang, J., Li, H., and Huang, L. "Robust onebit Bayesian compressed sensing with sign-flip errors", *IEEE Signal Processing Letters*, **22**(7), pp. 857-861 (2015).
- Dong, X. and Zhang, Y. "A MAP approach for 1-bit compressive sensing in synthetic aperture radar imaging", *IEEE Geoscience and Remote Sensing Letters*, 12(6), pp. 1237-1241 (2015).
- Zhu, J., Wang, X., Lin, X., and Ju, Y. "Maximum likelihood estimation from sign measurements with sensing matrix perturbation", *IEEE Trans Signal Processing*, 62(15), pp. 3741-3753 (2014).
- Zayyani, H., Babaie-zadeh, M., and Jutten, C. "An iterative Bayesian algorithm for sparse component analysis (SCA) in presence of noise", *IEEE Trans* Signal Processing, 57(11), pp. 4378-4390 (2009).
- Zayyani, H., Babaie-zadeh, M., and Jutten, C. "Bayesian pursuit algorithm for sparse representation", *Proceeding ICASSP2009* (2009).
- Zayyani, H. and Babaie-zadeh, M. "Thresholded smoothed L0 (SL0) dictionary learning for sparse representations", *Proceeding ICASSP2009* (2009).

Biographies

Hadi Zayyani received his BSc, MSc, and PhD degrees all in Electrical Engineering and all in the communication field from Sharif University of Technology in 2000, 2003, and 2010, respectively. From 2012 to now, he is an Assistant Professor of Qom University of Technology (QUT). His research interests include statistical signal processing, sparse signal processing, adaptive filter theory, radar signal processing and distributed signal processing.

Mehdi Korki received his BSc degree in Electrical Engineering from Shiraz University in 2001 and his ME degree in Electrical Engineering from Swinburne University of Technology in 2012, and the PhD degree in Electrical Engineering from Swinburne University of Technology in 2016. From 2001 to 2009, he was employed by Telecommunication Company of Booshehr (T.C.B), Iran, as an optical systems engineer. His research interests include statistical signal processing, machine learning, source localization, Power Line Communication (PLC), and smart grid communications.

Farokh Marvasti received his BS, MS, and PhD degrees all from Renesselaer Polytechnic Institute in 1970, 1971, and 1973, respectively. He has worked, consulted, and taught in various industries and academic institutions since 1972. Among which are Bell Labs, University of California Davis, Illinois Institute of

Technology, University of London, King's College. He was one of the Editors and Associate Editors of IEEE Trans on Communications and Signal Processing from 1990-1997. He has about 60 journal publications and has written several reference books. His last book is on Nonuniform Sampling: Theory and Practice by Kluwer in 2001. He is also a guest editor on Special Issue on Nonuniform Sampling for the Sampling Theory & Signal and Image Processing Journal. Dr. Marvasti is currently a Professor at Sharif University of Technology and the Director of Advanced Communications Research Institute (ACRI).