Signal Recovery From 1-Bit Quantized Noisy Samples via Adaptive Thresholding

Shahin Khobahi* and Mojtaba Soltanalian
Department of Electrical and Computer Engineering
University of Illinois at Chicago
Chicago, USA

Abstract

This paper is concerned with the problem of signal recovery from 1-bit noisy measurements. We present an efficient method to obtain an estimation of the signal of interest in the presence of colored or white noise. To the best of our knowledge, the proposed framework is the first approach in the area of one-bit sampling and signal recovery that can deal with the presence of colored noise. The proposed method is based on a convex quadratically constrained quadratic program (QCQP) utilizing an adaptive quantization thresholding approach. Due to its adaptive nature, the proposed method can recover both fixed and time-varying parameters (e.g., coming from a sinusoidal signal source) from their quantized one-bit samples.

Index Terms

Parameter estimation, quantization, 1-bit sampling, one-bit quantization threshold design, wireless sensor networks, compressive sensing, quadratic programming, convex optimization

I. Introduction

Wireless sensor networks (WSNs) present significant potential for usage in spatially wide-scale detection and estimation due to their many advantageous characteristics such as an inherent distributed structure. On the other hand, due to the natural constraints imposed on the nodes in a WSN such as low power consumption, low cost manufacturing, and more importantly, limited computational and transmission capabilities, it is desirable to develop a distributed estimation framework with which the nodes can reliably communicate with the fusion center (FC) while they are satisfying the power and computational constraints. Analog-to-digital converters (ADCs) are a key component in most of the modern digital systems in that they are bridging the gap between the analog world and digital systems [1]. Yet, a key gridlock of the power consumption constraint in a WSN or in many other *Internet of Things (IoT)* applications is ADCs. In addition, sampling at high data rates with high resolution ADCs would dramatically increase the manufacturing cost of these electronic components. Note that, in most of the digital signal processing schemes, an essential step after sampling the signal is to quantize the samples accordingly. That is, the samples of the signal is rounded to one of the predefined quantization levels; hence, an immediate solution is to reduce the bit resolution and sampling rate of the ADCs. More importantly, a fundamental constraint in a WSN is that the bandwidth of the system is limited and so it is of essence to use a proper quantization scheme to reduce the transmitted bits prior to transmission to meet this constraint. So, it is of essence to develop new estimation algorithms that can deal with low-resolution samples for signal recovery purposes.

In this work we consider the extreme case of 1-bit quantization and develop a novel algorithm for one-bit signal recovery using adaptive thresholding scheme. Our proposed general estimation framework can properly track and estimate an unknown parameter from its noisy one-bit measurements under both scenarios of white and colored noise.

II. Problem Formulation and System Model

We consider a wireless sensor network with N spatially distributed single-antenna nodes each of which observing an unknown (but deterministic) parameter $\theta \in \mathbb{R}$, at the time index k, according to the following linear observation model,

$$z_i^{(k)} = \theta^{(k)} + v_i^{(k)} \tag{1}$$

where i denotes the sensor index, and $v_i^{(k)}$ is the additive Gaussian observation noise. The observed signal at all nodes can be compactly formulated as,

$$\mathbf{z}^{(k)} = \theta^{(k)}\mathbb{1} + \mathbf{v}^{(k)} \tag{2}$$

where $\mathbb{1} = [1, \ldots, 1]^T$ denotes the all-one vector. As it was mentioned earlier, in order to satisfy the inherent bandwidth and power budget constraints in WSNs, we assume that each node utilizes a pre-defined *1-bit quantization* scheme to encode

* Corresponding author (e-mail: *skhoba2@uic.edu*). This work was supported in part by U.S. National Science Foundation Grant CCF-1704401.

its observation into 1 bit of information, and then transmits this data to the fusion center. Namely, the ith node applies the following quantization function $Q(.)$ on its observed data z_i,

$$r_i^{(k)} = Q(z_i^{(k)}, \tau_i^{(k)}) := sgn(z_i^{(k)} - \tau_i^{(k)}) = \begin{cases} +1, & \text{if } z_i^{(k)} > \tau_i^{(k)} \\ -1, & \text{if } z_i^{(k)} \leq \tau_i^{(k)} \end{cases}. \tag{3}$$

where $\tau_i^{(k)}$ denotes the ith node quantization threshold at the time index k. Assuming that each node can reliably transmit 1 bit of information to the FC without distortion, the aggregated received data from all nodes at the FC can be expresses as,

$$\mathbf{y} = \mathbf{r} = sgn\left(\mathbf{z}^{(k)} - \boldsymbol{\tau}^{(k)}\right) = sgn\left(\theta^{(k)}\mathbb{1} + \mathbf{v}^{(k)} - \boldsymbol{\tau}^{(k)}\right) \tag{4}$$

where $\mathbf{v} = [v_1, \ldots, v_N^{(k)}]^T$ is the combined observation noise vector with covariance matrix $\mathbf{\Sigma} = \mathbb{E}\{\mathbf{v}\mathbf{v}^H\}$, and $\boldsymbol{\tau}^{(k)} = [\tau_1^{(k)}, \ldots, \tau_N^{(k)}]$ is the quantization threshold vector. Now, our goal here is to develop an inference approach to estimate the unknown parameter θ, based on the N quantized 1-bit observations. Note that, the vector of 1-bit measurements $\{r_i^{(k)}\}_{i=1}^N$ given the quantization thresholds $\{\tau_i^{(k)}\}_{i=1}^N$ represents a limitation on the *geometric* location of the unquantized information $\mathbf{z}$; particularly, we can capture this geometric information on $\mathbf{z}$ through the following linear inequality:

$$\mathbf{\Omega}^{(k)}\left(\mathbf{z}^{(k)} - \boldsymbol{\tau}^{(k)}\right) \succeq 0 \tag{5}$$

where $\mathbf{\Omega} = \mathit{Diag}\{\mathbf{r}\} = \mathit{Diag}\{sgn\left(\mathbf{z}^{(k)} - \boldsymbol{\tau}^{(k)}\right)\}$, and $\mathit{Diag}\{.\}$ is the diagonalization operator. The set of scalar inequalities in (5) defines a set of half-planes based on the vector of 1-bit observations and the quantization thresholds presenting the subspace at which the unquantized information $\mathbf{z}$ can reside on.

We lay the ground for our 1-bit statistical and inference model using the weighted least square (WLS) method. Clearly, if the unquantized information vector $\mathbf{z}$ was available at the FC, then the maximum likelihood (ML) estimation of the unknown parameter θ given $\mathbf{z}$ can be expressed as,

$$\hat{\theta}^{(k)} = [\mathbb{1}^T\mathbf{\Sigma}^{-1}\mathbb{1}]^{-1}\mathbb{1}^T\mathbf{\Sigma}^{-1}\mathbf{z}^{(k)} \tag{6}$$

where the variance of the estimation is given by $\mathrm{Var}(\hat{\theta}^{(k)}) = [\mathbb{1}^T\mathbf{\Sigma}^{-1}\mathbb{1}]^{-1}$. Alternatively, one can obtain the maximum-likelihood estimator of the unknown parameter by minimizing the following WLS criterion,

$$\mathcal{Q}(\mathbf{z}, \theta) := ||\mathbf{z} - \mathbb{1}\theta||^2_{\mathbf{\Sigma}^{-1}} = \left(\mathbf{z} - \mathbb{1}\theta\right)^H \mathbf{\Sigma}^{-1}\left(\mathbf{z} - \mathbb{1}\theta\right), \tag{7}$$

where $\mathcal{Q}(\mathbf{z}, \theta)$ is our objective function to be minimized over the parameters $(\mathbf{z}, \theta)$. Henceforth, a natural approach to obtain an estimation of θ using the 1-bit quantization vector $\mathbf{r}$, is to use an alternating optimization approach and further exploit the limitation on the geometric location of $\mathbf{z}$ imposed by (5) to first obtain an estimate of $\mathbf{z}$ by fixing the variable θ and then estimating the unknown parameter θ using (6) which is the explicit solution of $\mathit{argmin}_\theta \mathcal{Q}(\mathbf{z}, \theta)$. Interestingly, for a fixed parameter $\mathbf{z}$, the optimal θ^* that minimizes $\mathcal{Q}(\mathbf{z}, \theta)$ coincides with that of the MLE of θ given in (6). Therefore, we can substitute the optimal $\hat{\theta}^{(k)}$ of (6) into (7) to further simplify the objective function $\mathcal{Q}(\mathbf{z}, \theta)$ in terms of the parameter $\mathbf{z}$, i.e.,

$$\mathcal{Q}(\mathbf{z}) = \mathcal{Q}(\mathbf{z}, \hat{\theta}^{(k)}) = \mathbf{z}^H(\mathbf{I} - \mathbb{1}\boldsymbol{\eta}^H)^H\mathbf{\Sigma}^{-1}(\mathbf{I} - \mathbb{1}\boldsymbol{\eta}^H)\mathbf{z}, \tag{8}$$

where we define $\boldsymbol{\eta}$ as,

$$\boldsymbol{\eta} \triangleq \frac{\mathbf{\Sigma}^{-1}\mathbb{1}}{\mathbb{1}^T\mathbf{\Sigma}^{-1}\mathbb{1}}. \tag{9}$$

Namely, we cast the problem of estimating the *unquantized* vector $\mathbf{z}$, as the solution to the following quadratically constrained quadratic program (QCQP),

$$\hat{\mathbf{z}}^{(k)} = \underset{\mathbf{z}}{\mathit{argmin}} \quad \mathbf{z}^H(\mathbf{I} - \mathbb{1}\boldsymbol{\eta}^H)^H\mathbf{\Sigma}^{-1}(\mathbf{I} - \mathbb{1}\boldsymbol{\eta}^H)\mathbf{z}, \tag{10}$$

$$\mathit{subject\ to} \quad \mathbf{\Omega}^{(k)}(\mathbf{z} - \boldsymbol{\tau}^{(k)}) \succeq \mathbf{0}. \tag{11}$$

Note that the scalar constraints in (11) ensures the consistency between the received 1-bit data $\mathbf{r}^{(k)}$ (incorporated in the matrix $\mathbf{\Omega}^{(k)}$) and the solution $\hat{\mathbf{z}}^{(k)}$. Due to the fact that the matrix $\mathbf{M} := (\mathbf{I} - \mathbb{1}\boldsymbol{\eta}^H)^H\mathbf{\Sigma}^{-1}(\mathbf{I} - \mathbb{1}\boldsymbol{\eta}^H)$ is positive definite and so the program of (10)-(11) is convex and admits a unique global solution and can be efficiently solved using the interior point methods. After solving (10) and fixing the solution $\hat{\mathbf{z}}^{(k)}$, we then minimize $\mathcal{Q}(\hat{\mathbf{z}}^{(k)}, \theta)$ over θ. Namely, we estimate the unknown parameter $\theta^{(k)}$ using,

$$\hat{\theta}^{(k)} = \underset{\theta}{\mathit{argmin}}\ \mathcal{Q}(\hat{\mathbf{z}}^{(k)}, \theta), \tag{12}$$

which has the closed form solution of the form,

$$\hat{\theta}^{(k)} = [\mathbb{1}^T\mathbf{\Sigma}^{-1}\mathbb{1}]^{-1}\mathbb{1}^T\mathbf{\Sigma}^{-1}\hat{\mathbf{z}}^{(k)}. \tag{13}$$

Table 1 The Proposed Adaptive Signal Recovery and Threshold Design Approach for 1-bit Sampled Data.

Initialization: Initialize the thresholds vector $\boldsymbol{\tau}^{(0)}$ with a random vector in $\mathbf{R}^N$, and initialize the threshold design variance σ_τ^2.
Step 1 - Quantization: Each node uses the received quantization threshold $\tau_i^{(k)}$ from the FC, and performs the quantized measurement $r_i^{(k)} = \text{sgn}(z_i^{(k)} - \tau_i^{(k)})$ accordingly, and then transmits the quantized information $r_i^{(k)}$ to the FC (i denotes the node number).
Step 2: FC constructs the quantized matrix $\boldsymbol{\Omega}^{(i)} = \text{Diag}\{\mathbf{r}^{(i)}\}$, based on the received 1-bit measurements vector $\mathbf{r}^{(k)} = [r_1^{(i)}, \ldots, r_N^{(k)}]^T$.
Step 4 - Parameter Estimation: FC establishes $\hat{\mathbf{z}}^{(k)}$ via solving the proposed QCQP in (10). Namely,

$$\hat{\mathbf{z}}^{(k)} = \text{argmin}_{\mathbf{z}}\ \mathcal{Q}(\mathbf{z}) \quad \text{s.t. } \boldsymbol{\Omega}^{(k)}(\mathbf{z} - \boldsymbol{\tau}^{(k)}) \succeq \mathbf{0}.$$

Step 5: Given $\hat{\mathbf{z}}^{(i)}$ from the previous step, FC estimates the unknown parameter θ via solving (12), i.e.,

$$\hat{\theta}^{(k)} = \left[\mathbb{1}^T \boldsymbol{\Sigma}^{-1} \mathbb{1}\right]^{-1} \mathbb{1}^T \boldsymbol{\Sigma}^{-1} \hat{\mathbf{z}}^{(k)}$$

Step 6 - Threshold Design:
The fusion center chooses the next quantization threshold for each node according to the following model,

$$\boldsymbol{\tau}_k^{(k+1)} = \hat{\theta}^{(k)} \mathbb{1} + \mathbf{w}_\tau^{(k)}$$

where $\mathbf{w}_\tau^{(k)} = [w_1^{(k)}, \ldots, w_N^{(k)}]^T$ denotes N realizations of the random variable $w_\tau \sim \mathcal{N}(0, \sigma_\tau^2)$.
Finally, FC transmits the thresholds $\boldsymbol{\tau}^{(i+1)}$, back to the nodes and waits for the next observation-estimation cycle (repeat steps 1-6).

A. Threshold Design

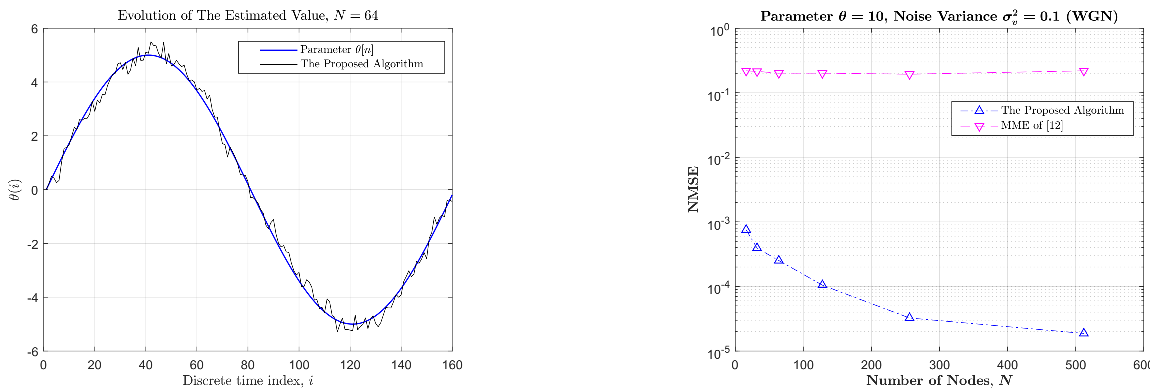
Now, we propose our threshold design scheme for each observation/estimation cycle. It can be shown that the optimal threshold given the Bernoulli observations $\{r_i^{(k)}\}_{i=1}^N$ is indeed equal to the unknown parameter θ; i.e., $\tau_{opt}^{(k)} = \theta^{(k)}$. Clearly, we cannot use the optimal threshold because it is a function of the unknown parameter. However, we can exploit our current knowledge of the unknown parameter to set the next quantization thresholds at each node in that $\hat{\theta}^{(k)}$ is our best estimate of the real value of the parameter $\theta^{(k)}$, at the time index k. This motivates us to use an evolutionary programming (EP) [8]–[11] and dynamic stochastic selection (DSS) [11] approach to establish our threshold design scheme. Briefly, having the current estimation of the unknown parameter $\hat{\theta}^{(k)}$, the FC samples the next threshold for each node from a normal distribution with mean $\hat{\theta}^{(k)}$, and variance σ_τ^2; in other words, fusion center adds N realizations of a zero-mean random variable $w \sim \mathcal{N}(0, \sigma_\tau^2)$ to the current estimate of the parameter to obtain the next thresholds to be transmitted back to the nodes. Namely, after obtaining $\hat{\theta}^{(k)}$ at the kth cycle, the FC chooses and transmits the next quantization thresholds $\boldsymbol{\tau}^{(k+1)}$ according to the following model,

$$\boldsymbol{\tau}^{(k+1)} = \hat{\theta}^{(k)} \mathbb{1} + \mathbf{w}_\tau^{(k)}. \tag{14}$$

where $\mathbf{w}_\tau^{(k)} = [w_1^{(k)}, \ldots, w_N^{(k)}]$ are N independent samples drawn from the normal distribution of $\mathcal{N}(0, \sigma_\tau^2)$. Furthermore, the variance of the random variable w can be chosen arbitrarily or according to the observation noise variance. More generally, we can model $\boldsymbol{\tau}^{(k+1)}$ as a multivariate Gaussian random vector with mean vector $\boldsymbol{\mu}^{(k+1)} = \theta^{(k)}\mathbb{1}$, and covariance matrix $\boldsymbol{\Sigma}_\tau$, e.g., $\boldsymbol{\tau}^{(k+1)} \sim \mathcal{N}(\boldsymbol{\mu}^{(k+1)}, \boldsymbol{\Sigma}_\tau)$. Next, FC transmits back the new quantization thresholds to the nodes and the next observation-estimation cycle begins. The proposed signal recovery and threshold design method is summarized in Table 1.

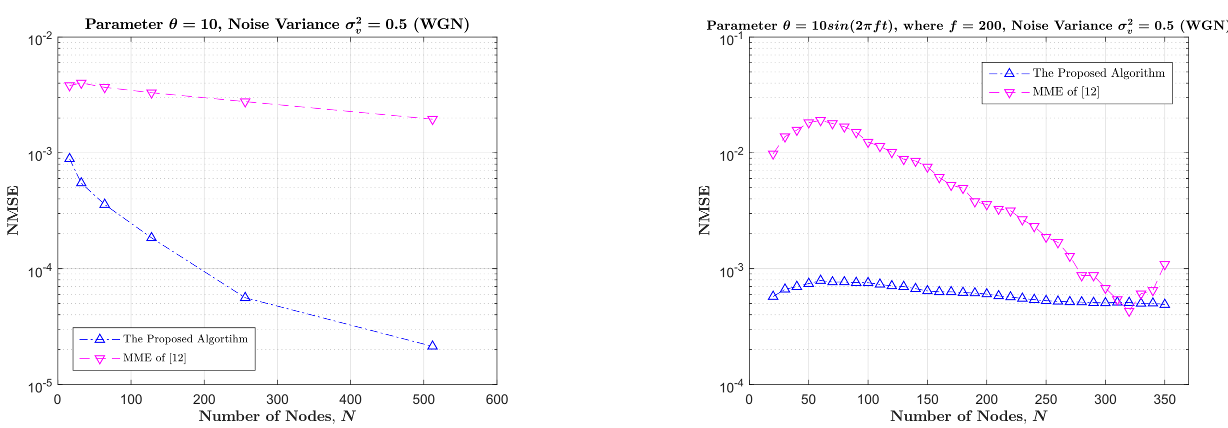
III. Simulation Results

In this section we provide a number of simulation results for evaluating the performance of the proposed algorithm (Table 1). We compare the performance of our proposed method in the presence of white Gaussian noise (WGN) with the modified mean estimator (MME) of [12]. It must be noted that our algorithm can handle the task of parameter estimation in the presence of the both colored and white noise; however, the MME method of [12] only works in the presence of *white* Gaussian noise. Also, we use the normalized mean square error (NMSE) as the performance metric in our simulations. Fig. 1(a) illustrates the evolution of our proposed estimation method in time for $N = 64$, and when the unknown parameter $\theta^{(k)} = 5sin(2\pi fk)$, where $f = 50\,Hz$. In addition, Fig. 1(b) and Fig. 2(a) show the NMSE vs. total number of nodes for the WGN scenario when $\theta = 10$, and for different values of the noise variance. In Fig. 2(b), we set $\theta^{(k)} = 10sin(2\pi fk)$ where $f = 200\,Hz$. Moreover, Figs. 3(a)-(b) demonstrates the NMSE of our proposed algorithm for the colored Gaussian noise versus the total power of the noise $P_{tot} = Tr(\boldsymbol{\Sigma})$ for the fixed parameter $\theta = 10$, and the time-varying case of $\theta^{(k)} = 10sin(2\pi fk)$ with $f = 50\,Hz$. Finally, Fig. 4(a)-(b) show the NMSE versus the total number of nodes assuming that the $P_{tot} = 5$ for the fixed parameter $\theta = 10$, and the time-varying signal $\theta^{(k)} = 10sin(2\pi fk)$, with $f = 50\,Hz$. It can be observed from the simulations that as the number of nodes (1-bit information) increases the accuracy of the proposed method improves significantly and the NMSE in most of the cases attains values that are virtually zero (note that we used logarithmic scale in our simulations).



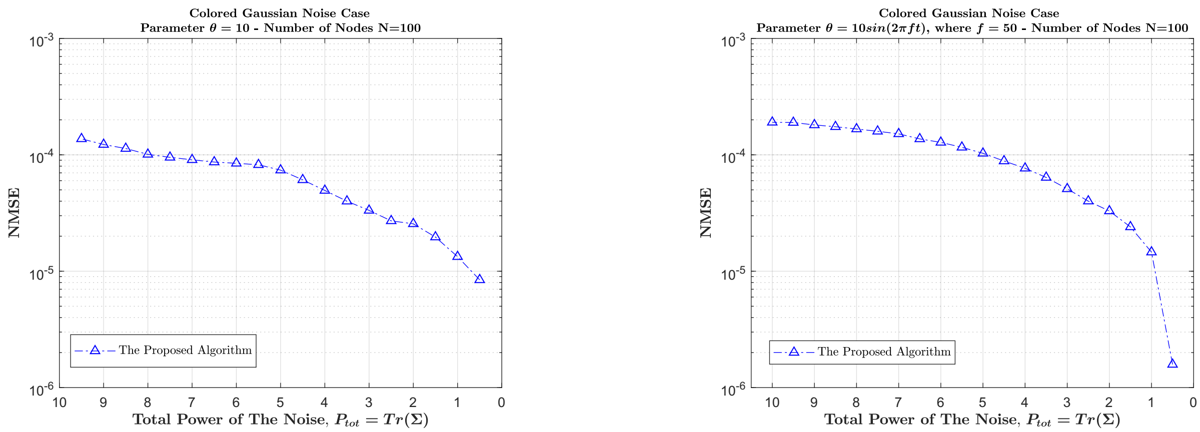
(a) (b)

Fig. 1. (a) Demonstrates the time evolution of the proposed estimation method in tracking the unknown parameter $\theta^{(k)}$ when $N = 64$, and the unknown parameter is a sinusoidal signal $\theta^{(k)} = 5sin(2\pi ft)$, where $f = 50$. Clearly, our method can accurately track and estimate a time-varying signal. Moreover, (b) shows the NMSE vs. the toal number of nodes N, in the presence of white Gaussian observation noise with a variance of $\sigma_v^2 = 0.1$. Our proposed quadratic programming approach (Table 1) shows significantly superior performance than that of the MME of [12].



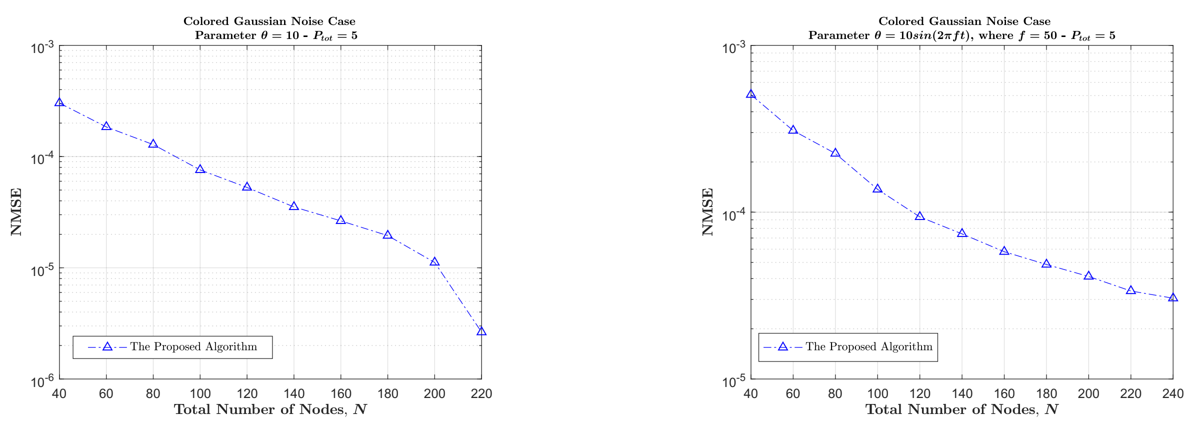
(a) (b)

Fig. 2. (a) demonstrates the performance of our proposed algorithm when the noise variance $\sigma_v^2 = 0.5$ and the unknown parameter is fixed in time $\theta = 10$. Furthermore, (b) demonstrates the case of a time-varying unknown parameter with high deviation frequency as $\theta^{(k)} = 10sin(2\pi fk)$, where $f = 200$, and the noise variance $\sigma_v^2 = 0.5$. Clearly, our quadratic programming approach (Table 1) shows significantly better performance in both cases and assumes a very low NMSE showing the high accuracy of our proposed estimation method (note that we used a logarithmic scale for the simulations).



(a) (b)

Fig. 3. NMSE versus the total power of the *colored* Gaussian noise $P_{tot} = Tr(\mathbf{\Sigma})$ for a network with $N = 100$, when (a) the unknown parameter is fixed in time $\theta = 10$, and (b) when it is a time-varying sinusoidal signal $\theta^{(k)} = 10sin(2\pi ft)$, where $f = 50$. Clearly, our proposed QP approach (Table 1) can accurately recover the signal in both cases of time-varying and fixed parameter when the observations are corrupted by colored Gaussian noise.



(a) (b)

Fig. 4. NMSE versus the total number of nodes N, for the scenario of (a) estimating a fixed parameter $\theta = 10$, and (b) when the unknown parameter $\theta^{(k)} = 10sin(2\pi ft)$, where $f = 50$. In both cases we set the total power of the observation colored Gaussian noise $P_{tot} = 5$. Obviously, as the total number of nodes (1-bit information) increases, the accuracy of our proposed algorithm significantly improves and assumes virtually zero NMSE.

REFERENCES

[1] B. Le, T. W. Rondeau, J. H. Reed, and C. W. Bostian, "Analog-to-digital converters," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 69–77, 2005.

[2] 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.

[3] 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.

[4] J. N. Laska, Z. Wen, W. Yin, and R. G. Baraniuk, "Trust, but verify: Fast and accurate signal recovery from 1-bit compressive measurements," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5289–5301, 2011.

[5] P. T. Boufounos, "Greedy sparse signal reconstruction from sign measurements," in *Signals, Systems and Computers, 2009 Conference Record of the Forty-Third Asilomar Conference on*. IEEE, 2009, pp. 1305–1309.

[6] A. Gupta, R. Nowak, and B. Recht, "Sample complexity for 1-bit compressed sensing and sparse classification," in *Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on*. IEEE, 2010, pp. 1553–1557.

[7] M. Yan, Y. Yang, and S. Osher, "Robust 1-bit compressive sensing using adaptive outlier pursuit," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3868–3875, 2012.

[8] D. B. Fogel, *Evolutionary computation: toward a new philosophy of machine intelligence*. John Wiley & Sons, 2006, vol. 1.

[9] C. A. C. Coello, G. B. Lamont, D. A. Van Veldhuizen *et al.*, *Evolutionary algorithms for solving multi-objective problems*. Springer, 2007, vol. 5.

[10] A. E. Eiben, J. E. Smith *et al.*, *Introduction to evolutionary computing*. Springer, 2003, vol. 53.

[11] T. Back, *Evolutionary algorithms in theory and practice: evolution strategies, evolutionary programming, genetic algorithms*. Oxford university press, 1996.

[12] T. Wu and Q. Cheng, "Distributed estimation over fading channels using one-bit quantization," *IEEE Transactions on Wireless Communications*, vol. 8, no. 12, 2009.