Efficient Reconstruction of Sparse Vectors from Quantized Observations

Amine Mezghani and Josef A. Nossek
Institute for Circuit Theory and Signal Processing
Munich University of Technology, 80290 Munich, Germany
E-Mail: {Mezghani, Nossek}@nws.ei.tum.de

***Abstract*— Compressive sensing is a recent technique for estimating a sparse vector from a reduced number of observations. Several algorithms have been developed and studied in this context. However, the reduction of number of samples (or the sampling rate) is one aspect of compressive sensing. In fact, the quantization of the observations, which is generally unavoidable in practice (due to the analog-to-digital conversion), could be also understood as another aspect of compressed sensing aiming at reducing the complexity of the sampling device. Moreover, reducing the ADC resolution might be much more beneficial in terms of circuit complexity and power consumption than decreasing the sampling rate. Therefore, we investigate this further aspect of compressive sensing related to the resolution of measurements instead of their number. We first present an efficient message-passing-like iterative algorithm for estimating a vector from quantized linear noisy observations. Contrary to the related work in [1], the algorithm does not require any prior information about the sparse input (such as distribution or norm) and can be applied for arbitrary quantizer resolution. Then, a state evolution analysis is carried out to study the dynamics of the iterative algorithm and can be used to optimize the quantizer characteristic. Finally, some experimental results are provided to demonstrate the validity and performance of the presented algorithm.**

I. Introduction

Most of the contributions on reconstructing sparse vectors assume that the observation vector is available with infinite precision. In practice, however, a quantizer (e.g. A/D-converter) is applied to the observed analog signal, so that the data can be processed in the digital domain. Recently, there are several works that deal with the problem of reconstructing a sparse vector from quantized input. In [2], a fixed point iterative reconstruction algorithm was proposed based on a maximum likelihood approach. A generalized approximate message passing was derived in [1] based on the work [3], and a quantizer optimization by means of a state evolution analysis was carried out. In fact, approximate message passing algorithms can achieve very good reconstruction performance while having a low complexity compared to other algorithms. However, the proposed message passing-like algorithm in [1] still require the knowledge of the input statistic since it is based on a Bayesian formulation.

There are many applications where severe quantization may be unavoidable or even preferred. In fact, the quantization of the observations which is generally unavoidable in practice (due to the analog-to-digital conversion) could be also understood as an other aspect of compressed sensing aiming at reducing the complexity of the sampling device. Moreover, reducing the ADC resolution might be much more beneficial in terms of circuit complexity and power consumption than decreasing the sampling rate. Therefore, we investigate this further aspect of compressive sensing related to the resolution of measurements instead of their number.

To this end, the problem of reconstructing a signal vector from quantized measurements is formulated as ℓ_1 and ℓ_2 regularized problems and a low complexity iterative algorithm is derived, which does not require information about the input distribution. The presented algorithm, can be considered as a generalization to the thresholding algorithm proposed in [4] for the unquantized case. In other words, we propose a general formulation and solution for the sparse signal estimation problem with quantized observations. Experimental results are provided to demonstrate the validity and performance of the presented algorithm quite well.

By simulation, we also study the effect of quantization and noise on the estimation performance of the algorithm. Additionally, we carry out a state evolution analysis similarly to [3], [1] and [5] to evaluate the compromise between resolution and the number of observations and study the possibility of optimizing the quantizer characteristic. We note that the state evolution analysis is not mathematically rigorous since it is based on some heuristic assumptions. Nevertheless, simulations demonstrate the validity of the state evolution analysis, since it predicts the performance of the proposed algorithm.

II. Problem Formulation

We consider the compressed sensing reconstruction problem, where we aim at estimating a sparse M-dimensional vector $\boldsymbol{x} \in \mathbb{R}^M$ according to the linear model

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\eta}, \tag{1}$$

where $\boldsymbol{y}$ is the noisy unquantized measurement vector, $\boldsymbol{A} \in \mathbb{R}^{N\times M}$ is the random measurement matrix, and $\boldsymbol{\eta}$ refers to Gaussian noise vector with covariance $\boldsymbol{R}_{\eta\eta} = \mathrm{E}[\boldsymbol{\eta}\boldsymbol{\eta}^{\mathrm{T}}] = \sigma_0^2\mathbf{I}$. Usually the random measurement matrix $\boldsymbol{A}$ is normalized to have unit column variance, which ensures that $\mathrm{E}_{\boldsymbol{A}}[\|\boldsymbol{A}\boldsymbol{x}\|^2 |\boldsymbol{x}] = \|\boldsymbol{x}\|^2$ for any input $\boldsymbol{x}$. Additionally, we define

$$\beta = \frac{N}{M}, \tag{2}$$

as the number of measurements per vector coefficient, i.e., the oversampling factor.

In a practical system, each receive signal component y_j, $1 \le j \le N$, is quantized by a b-bit resolution scalar quantizer

(A/D-converter). Thus, the resulting quantized signals read as

$$r_j = \mathcal{Q}(y_j), \tag{3}$$

where $\mathcal{Q}(\cdot)$ denotes the quantization operation. For the case that we use a uniform symmetric mid-riser type quantizer, the quantized receive alphabet for each dimension is given by

$$r_j \in \{(-\frac{2^b}{2} - \frac{1}{2} + k)\Delta;\ k = 1, \cdots, 2^b\} = \mathcal{R}, \tag{4}$$

where Δ is the quantizer step-size and b the number of quantizer bits, which are set the same for all the quantizers.

With these definitions, the conditional probability of the quantized output given an input $\boldsymbol{x}$ reads as

$$\Pr(\boldsymbol{r}|\boldsymbol{x}) = \prod_{j=1}^{N} \rho(r_j|\boldsymbol{a}_j^{\mathrm{T}}\boldsymbol{x}), \tag{5}$$

where $\boldsymbol{a}_j^{\mathrm{T}}$ is the j-th row of $\boldsymbol{A}$ and

$$\begin{aligned}\rho(r_j|\boldsymbol{a}_j^{\mathrm{T}}\boldsymbol{x}) &= \frac{1}{\sqrt{2\pi\sigma_0^2}} \int_{r_j^{\mathrm{lo}}}^{r_j^{\mathrm{up}}} \mathrm{e}^{-\frac{(y-\boldsymbol{a}_j^{\mathrm{T}}\boldsymbol{x})^2}{2\sigma_0^2}} dy \\ &= \Phi\left(\frac{r_j^{\mathrm{up}} - \boldsymbol{a}_j^{\mathrm{T}}\boldsymbol{x}}{\sigma_0}\right) - \Phi\left(\frac{r_j^{\mathrm{lo}} - \boldsymbol{a}_j^{\mathrm{T}}\boldsymbol{x}}{\sigma_0}\right),\end{aligned} \tag{6}$$

with $\Phi(x)$ represents the cumulative Gaussian distribution given by

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) \mathrm{dt}. \tag{7}$$

Hereby, the lower and upper quantization boundaries are (for uniform symmetric quantizers)

$$r_j^{\mathrm{lo}} = \begin{cases} r_j - \frac{\Delta}{2} & \text{for} \quad r_j \geq -\frac{\Delta}{2}(2^b - 2) \\ -\infty & \text{otherwise,} \end{cases}$$

and

$$r_j^{\mathrm{up}} = \begin{cases} r_j + \frac{\Delta}{2} & \text{for} \quad r_j \leq \frac{\Delta}{2}(2^b - 2) \\ +\infty & \text{otherwise.} \end{cases}$$

III. Approximative MAX-SUM Algorithm

To estimated the vector $\boldsymbol{x}$ from the quantized output vector $\boldsymbol{r}$, we aim at solving the following ℓ_p ($p \in \{1, 2\}$) regularized optimization problem

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{\operatorname{argmax}} \left[\frac{1}{\lambda'} \sum_{j=1}^{N} \ln(\rho(r_j|\boldsymbol{a}_j^{\mathrm{T}}\boldsymbol{x})) - \frac{1}{p}\|\boldsymbol{x}\|_p^p\right], \tag{8}$$

with a certain regularization parameter λ' that can be chosen depending on the noise variance and the sparsity of $\boldsymbol{x}$. Using the *generalized approximate message passing* (GAMP) approach [3], [6], the following iterative approximative MAX-SUM algorithm, which will be derived in the next subsections, has been obtained to solve both regularized problems:

Initialization: $\boldsymbol{z}^0 = \boldsymbol{0}$, $\boldsymbol{x}^0 = \boldsymbol{0}$, $\theta^0 = \text{const} \neq 0$
Repeat until convergence:
Output Step:

$$\boldsymbol{z}^t = g_t(-\boldsymbol{A}\boldsymbol{x}^{t-1} + \theta^{t-1} \cdot \boldsymbol{z}^{t-1}, \boldsymbol{r}) \tag{9}$$

$$\xi^t = \left\langle \dot{g}_t(-\boldsymbol{A}\boldsymbol{x}^{t-1} + \theta^{t-1} \cdot \boldsymbol{z}^{t-1}, \boldsymbol{r}) \right\rangle \tag{10}$$

Input Step:

$$\boldsymbol{x}^t = f_t(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{z}^t + \xi^t \boldsymbol{x}^{t-1}) \tag{11}$$

$$\theta^t = \beta \cdot \left\langle \dot{f}_t(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{z}^t + \xi^t \boldsymbol{x}^{t-1}) \right\rangle \tag{12}$$

The scalar functions $f_t(v)$ and $g_t(-u, r)$ introduced above are applied to their arguments component-by-component and they are given by the expressions

$$f_t(v) = \underset{x}{\operatorname{argmax}} \left[-\frac{1}{p}\|x\|_p^p - \frac{1}{2\xi^t}(\xi^t x - v)^2\right], \tag{13}$$

$$g_t(-u, r) = \frac{1}{\theta^t} \underset{w}{\operatorname{argmax}} \left[\frac{1}{\lambda'} \ln \rho(r|w) - \frac{1}{2\theta^t}(w - u)^2\right] - \frac{u}{\theta^t}, \tag{14}$$

while $\dot{f}_t(v)$, $\dot{g}_t(-u, r)$ are their respective derivative with respect to the first argument. Further, we have introduced the notation $\langle \boldsymbol{v} \rangle = \sum_{i=1}^{N} v_i/N$ for any vector $\boldsymbol{v} \in \mathbb{R}^N$.
Interestingly, the function $g_t(-u, r)$ in (14) is quite simple for the noiseless case and takes the following form regardless of the considered vector norm

$$g_t(-u, r)|_{\sigma_0=0} = \begin{cases} 0 & \text{if} \quad r^{\mathrm{lo}} \leq u \leq r^{\mathrm{up}} \\ \frac{1}{\theta^t}(r^{\mathrm{lo}} - u) & \text{if} \quad u \leq r^{\mathrm{lo}} \\ \frac{1}{\theta^t}(r^{\mathrm{up}} - u) & \text{if} \quad u \geq r^{\mathrm{up}}. \end{cases} \tag{15}$$

For the noisy case, there is no closed form expression for $g_t(-u, r)$. Nevertheless, we can use the following approximation to solve the optimization in (14) in closed form

$$\begin{aligned}&\frac{1}{\lambda'} \ln\left[\Phi(\frac{r^{\mathrm{up}} - x}{\sigma_0}) - \Phi(\frac{r^{\mathrm{lo}} - x}{\sigma_0})\right] \\ &\quad \approx \begin{cases} 0 & \text{if} \quad r^{\mathrm{lo}} + c_\delta \leq u \leq r^{\mathrm{up}} - c_\delta \\ -\frac{1}{2\lambda}(r^{\mathrm{lo}} + c_\delta - x)^2 & \text{if} \quad x \leq r^{\mathrm{lo}} + c_\delta, \\ -\frac{1}{2\lambda}(r^{\mathrm{up}} - c_\delta - x)^2 & \text{if} \quad x \geq r^{\mathrm{up}} - c_\delta, \end{cases}\end{aligned} \tag{16}$$

with appropriately chosen constant c_δ and λ. A low complexity approximation of the output function is then given by

$$g_t(-u, r) = \begin{cases} 0 & \text{if} \quad r^{\mathrm{lo}} + c_\delta \leq u \leq r^{\mathrm{up}} - c_\delta \\ \frac{1}{\theta^t + \lambda}(r^{\mathrm{lo}} + c_\delta - u) & \text{if} \quad u \leq r^{\mathrm{lo}} + c_\delta \\ \frac{1}{\theta^t + \lambda}(r^{\mathrm{up}} - c_\delta - u) & \text{if} \quad u \geq r^{\mathrm{up}} - c_\delta. \end{cases} \tag{17}$$

We have found by trying several values that choosing

$$c_\delta = \min\{1.8 \cdot \sigma_0, (r^{\mathrm{up}} - r^{\mathrm{lo}})/2\}, \tag{18}$$

provides good results (c.f. Fig. 1), while the regularization parameter λ has to be adapted depending on the noise variance σ_0^2 and the sparsity of the vector $\boldsymbol{x}$. Also, we see that (17) is equivalent to (15) when $\lambda = c_\delta = 0$.

On the other hand, the function $f_t(v)$ defined by the optimization problem (13) depends on the considered norm. For the ℓ_1 problem, it is given by

$$f_t(v) = \operatorname{sign}(v) \max\left(|v| - 1, 0\right)/\xi^t, \tag{19}$$

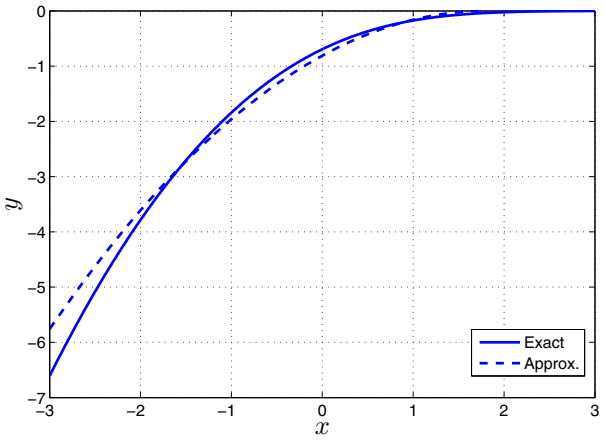

Fig. 1. Approximation of $\ln(\Phi(x))$ by $\frac{1}{4}(x-1.8)^2$ for $x \leq 1.8$.

whereas the ℓ_2 problem is solved by means of

$$f_t(v) = \frac{v}{1+\xi^t}. \tag{20}$$

The output step of the algorithm evaluates the error vector $\boldsymbol{z}$ (i.e. the uncertainty) between the estimated noiseless output and the observation. We notice that the term $\theta^t \cdot \boldsymbol{z}^t$ ensures in some sense that only extrinsic information is used from iteration to the next to avoid creating direct positive feedback from previous iterations, which improves the convergence behavior and thus the reconstruction performance. Based on this error vector, an update is carried out for the estimated vector using the input function $f_t(v)$ obtained according to the special norm regularization.

A. Factor Graph Representation

To derive the algorithm, and in analogy to [7], a factor graph representation of the quantized measurement system is introduced in Fig. 2. Each unknown coefficient x_i is represented by a circle, referred to symbol node, and each received quantized signal r_j corresponds to a square, called the signal node. Each edge connecting i and j represents the corresponding gain factor $a_{j,i}$, if $a_{j,i} \neq 0$. Ignoring the cycles in the graph, let us derive the so called "loopy" message passing algorithm (or MAX-SUM algorithm) from the factor graph representation.[1]

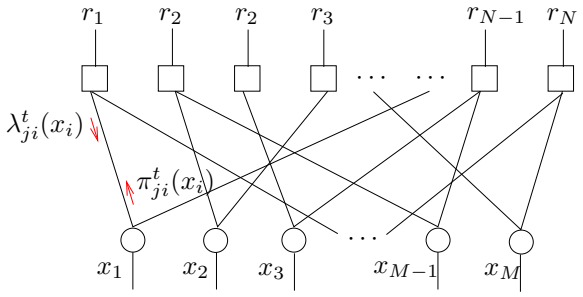

Fig. 2. Factor-graph representation of the quantized MIMO channel.

[1]We note that the MAX-SUM algorithm is optimal for cycle free graphs and performs nearly optimal in sparse graphs. In the case of dense, large enough, channel matrices, it may provide good approximate solution as we will see later.

B. Derivation of the approximate MAX-SUM algorithm

For the derivation of the presented algorithm, a message passing algorithm based on the MAX-SUM rules while ignoring cycles in the underlying factor Graph is considered. This approach is fairly similar to the derivation in [3]. It is made possible due the structure of the cost function (8) which consists of a sum of input (input vector norm) and output (output probabilities) terms. Each iteration t of this algorithm consists in, first sending messages from each signal node j to each symbol node i (output step), and then vice versa (input step). The output step message is given by

$$\lambda^t_{j,i}(x_i) = \max_{\boldsymbol{x}\setminus x_i} \frac{\ln \rho(r_j|\boldsymbol{a}_j^{\mathrm{T}}\boldsymbol{x})}{\lambda'} + \sum_{i'\neq i} \pi^t_{j,i'}(x_{i'}). \tag{21}$$

The input step message is then

$$\pi^t_{j,i}(x_i) = -\frac{1}{p}|x_i|^p + \sum_{j'\neq j} \lambda^t_{j',i}(x_i). \tag{22}$$

The messages are initialized at $t=0$ with $\lambda^t_{i,r}(x_j) = 0$, and the final estimate can be calculated as the maximizer of

$$\pi^t_i(x_i) = \frac{1}{p}|x_i|^p + \sum_{j} \lambda^t_{j,i}(x_j). \tag{23}$$

We now provide an approximation for the MAX-SUM algorithm, which become exact in the large system limit. It is based on the fact that the entries $a_{j,i}$ scales as $1/\sqrt{N}$. To this end, we introduce the following values

$$\begin{aligned} \hat{x}^t_i &= \operatorname*{argmax}_{x_i} \pi^t_i(x_i) \\ \hat{x}^t_{j,i} &= \operatorname*{argmax}_{x_i} \pi^t_{j,i}(x_i) \\ \frac{1}{\mu^t_i} &= -\frac{\partial^2}{\partial x_i^2}\pi^t_i(\hat{x}^t_i) \\ \frac{1}{\mu^t_{j,i}} &= -\frac{\partial^2}{\partial x_i^2}\pi^t_{j,i}(\hat{x}^t_{i,j}). \end{aligned} \tag{24}$$

Then, we can use the following second order approximation for the input step messages around its maximizer,

$$\pi^t_{j,i}(x_i) \approx \pi^t_{j,i}(\hat{x}^t_{j,i}) - \frac{1}{2\mu^t_i}(x_i - \hat{x}^t_{j,i})^2, \tag{25}$$

where we assume $\mu^t_{j,i} \approx \mu^t_i$ in the large system limit. Thus, the cost function in (21) becomes

$$\frac{\ln \rho(r_j|\boldsymbol{a}_j^{\mathrm{T}}\boldsymbol{x})}{\lambda'} + \sum_{i'\neq i} \pi^t_{j,i'}(\hat{x}^t_{j,i'}) - \frac{1}{2\mu^t_{i'}}(x_{i'} - \hat{x}^t_{j,i'})^2. \tag{26}$$

The motivation for such approximation is that the impact of x_i on the function $\ln \rho(r_j|\boldsymbol{a}_j^{\mathrm{T}}\boldsymbol{x})$ is asymptotically very small and that the maximizing value of x_i for the cost function (21) will not deviate much from $\hat{x}_{j,i}$. We solve the maximization in (21) in two steps

$$\max_{q_j} \max_{\boldsymbol{x}\setminus x_i,\ \text{s.t. } \boldsymbol{a}_j^{\mathrm{T}}\boldsymbol{x}=q_j} \rho(r_j|q_j) - \sum_{i'\neq i} \frac{1}{2\mu^t_{i'}}(x_{i'} - \hat{x}^t_{j,i'})^2. \tag{27}$$

The solution for the inner maximization leads to

$$\max_{q_j} \rho(r_j|q_j) - \frac{1}{2\theta^t_{j,i}}(q_j - \hat{u}_{j,i} - a_{j,i}x_i)^2, \tag{28}$$

where, we introduced the variables

$$q_j = a_{j,i}x_i + \sum_{i' \neq i} a_{j,i'}x_{i'}, \tag{29}$$

$$\hat{u}_{j,i}^t = \sum_{i' \neq i} a_{j,i'}\hat{x}_{j,i'}^t, \tag{30}$$

and

$$\theta_{j,i}^t = \sum_{i' \neq i} |a_{j,i'}|^2 \mu_{i'}^t \approx \sum_i |a_{j,i}|^2 \mu_i^t = \theta_j \ \forall i, \tag{31}$$

where the approximation holds for the large system limit. Furthermore, by defining

$$\hat{u}_j^t = \sum_i a_{j,i}\hat{x}_{j,i}^t, \tag{32}$$

we get

$$\max_{q_j} \frac{\ln \rho(r_j|q_j)}{\lambda'} - \frac{1}{2\theta_j^t}(q_j - \hat{u}_i^t - a_{j,i}\hat{x}_{j,i}^t - a_{j,i}x_i)^2. \tag{33}$$

These obtained output step messages can be then approximated as

$$\begin{aligned}
\lambda_{j,i}^t(x_i) &= \max_{q_j} \frac{\ln \rho(r_j|q_j)}{\lambda'} - \frac{1}{2\theta_j^t}(q_j - \hat{u}_j^t + a_{j,i}\hat{x}_{j,i}^t - a_{j,i}x_i)^2 \\
&= \max_{q_j} \frac{\ln \rho(r_j|q_j)}{\lambda'} - \frac{1}{2\theta_j^t}(q_j - \hat{u}_j^t - a_{j,i}(x_i - \hat{x}_{j,i}^t))^2 \\
&\approx \max_{q_j} \frac{\ln \rho(r_j|q_j)}{\lambda'} - \frac{1}{2\theta_j^t}(q_j - \hat{u}_j^t - a_{j,i}(x_i - \hat{x}_i^t))^2,
\end{aligned} \tag{34}$$

where the approximation is obtained by neglecting the terms of order $a_{j,i}^2$. Let us now define

$$\hat{q}_j^t = \operatorname*{argmax}_{q_j} \frac{\ln \rho(r_j|q_j)}{\lambda'} - \frac{1}{2\theta_j^t}(q_j - \hat{u}_j^t)^2, \tag{35}$$

and

$$z_j^t = \frac{1}{\theta_j^t}(\hat{q}_j^t - \hat{u}_j^t) \doteq g_t(-\hat{u}_j^t, r_j), \tag{36}$$

As done before for the input step message, we derive now a second order expansion of the output step message around $\hat{x}_i^t$. Evaluating the first derivative

$$\left.\frac{\partial \lambda_{j,i}^t(x_i)}{\partial x_i}\right|_{x_i = \hat{x}_i^t} = a_{j,i}z_j^t, \tag{37}$$

and the second derivative

$$\left.\frac{\partial^2 \lambda_{j,i}^t(x_i)}{\partial x_i^2}\right|_{x_i = \hat{x}_i^t} = a_{j,i}\frac{\partial z_j^t}{\partial x_i} = a_{j,i}^2 \dot{g}_t(-\hat{u}_j^t, r_j) \doteq a_{j,i}^2 \frac{1}{\mu_j^{s,t}}, \tag{38}$$

leads to

$$\lambda_{j,i}^t(x_i) \approx \text{const} \ - \frac{1}{2\mu_j^{s,t}}(\mu_j^{s,t}z_j^t - a_{j,i}(x_i - \hat{x}_i^t))^2. \tag{39}$$

The input step messages can be obtained now as (c.f. (22))

$$\pi_{j,i}^t(x_i) = -\frac{1}{p}|x_i|^p - \frac{1}{2\xi_{j,i}^t}(\hat{v}_{j,i}^t - \xi_{j,i}^t x_i)^2, \tag{40}$$

where we introduced the definitions

$$\begin{aligned}
\hat{v}_{j,i}^t &= \sum_{l \neq j}(a_{l,i}z_i^t + \frac{1}{\mu_l^{s,t}}a_{l,i}^2\hat{x}_i^t) \\
&= \xi_{j,i}^t\hat{x}_i^t + \xi_{j,i}^t \sum_{l \neq j} a_{l,i}z_l^t \\
&= \underbrace{\xi_{j,i}^t\hat{x}_i^t + \sum_l a_{l,i}z_l^t}_{\hat{v}_i^t} - a_{j,i}z_j^t,
\end{aligned} \tag{41}$$

Thereby, we used the substitution

$$\begin{aligned}
\xi_{j,i}^t &= \sum_{l \neq j} a_{l,i}^2 \frac{1}{\mu_l^{s,t}} \approx \sum_j a_{j,i}^2 \frac{1}{\mu_l^{s,t}} \doteq \xi_i^t \ \forall j \\
&\to \frac{1}{N}\sum_j \dot{g}_t(-\hat{u}_j^t, r_j) \doteq \xi^t.
\end{aligned} \tag{42}$$

We can proceed to compute $\hat{x}_{j,i}^t$

$$\begin{aligned}
\hat{x}_{j,i}^t &= \operatorname*{argmax}_{x_i} -\frac{1}{p}|x_i|^p - \frac{1}{2\xi_{i,j}^t}(\hat{v}_{i,j}^t - \xi_i^t x_i)^2 \\
&= \operatorname*{argmax}_{x_i} -\frac{1}{p}|x_i|^p - \frac{1}{2\xi_{i,j}^t}(\hat{v}_i^t - a_{j,i}z_j^t - \xi_i^t x_i)^2 \\
&\approx \hat{x}_i^t - \Gamma_i^t a_{j,i}z_j^t,
\end{aligned} \tag{43}$$

where we define

$$\begin{aligned}
\hat{x}_i^t &= \operatorname*{argmax}_{x_i} \pi_i^t(x_i) \\
&= \operatorname*{argmax}_{x_i} -\frac{1}{p}|x_i|^p - \frac{1}{2\xi_i^t}(\hat{v}_i^t - \xi_i^t x_i)^2 \\
&\doteq f_t(\hat{v}_i^t),
\end{aligned} \tag{44}$$

and

$$\Gamma_i^t = \frac{\partial \hat{x}_i^t}{\partial \hat{v}_i} = \dot{f}_t(\hat{v}_i^t). \tag{45}$$

Next, it can be shown due to the fact that the first derivative $\dot{\pi}_i^t(\hat{x}_i^{t+1}) = 0$ that $\Gamma_i^t = \mu_i^t$ (c.f. (24)). Therefore, (31) becomes

$$\theta_j = \sum_i |a_{j,i}|^2 \dot{f}_t(\hat{v}_i^t) \to \frac{\beta}{M}\sum_i \dot{f}_t(\hat{v}_i^t) = \theta \ \forall j. \tag{46}$$

Finally, (32) becomes using (43) and the fact that $\mu_i^t = \Gamma_i^t = \dot{f}_t(\hat{v}_i^t)$ as follows

$$\hat{u}_j = \sum_i a_{j,i}\hat{x}_i - \theta_j z_j^t. \tag{47}$$

It can be observed that the steps presented in the algorithm corresponds to Eqs (47), (36), (42), (41), (44) and (46).

IV. State Evolution Analysis

State evolution analysis (also known as density evolution for the case of sparse matrices) is a powerful tool to study the behavior of message passing algorithms in the large-system limit [8], [3]. The large-system limit means that we consider the limit when M and N go to infinity, while the ratio $\beta = M/N$ is kept fixed. Under the conjecture that the presented algorithm converges to a certain fixed point, this analysis would deliver useful theoretical results about its reconstruction performance under quantization. For the analysis, we assume

a random matrix $\boldsymbol{A}$, where the entries $\{a_{j,i}\}$ are i.i.d. with zero mean and variance $1/N$.

Such analysis has been used in coding theory for analyzing the behavior of iterative algorithms and was justified on the fact that the underlaying Graph is sparse. In general, the state evolution analysis is not mathematically rigorous and it is based on some heuristic assumptions. However it was shown in [6] that the state evolution analysis provide mathematically exact results for the case of dense graphs, where the entries of the matrix is drawn from a Gaussian distribution but when the observation vector is not quantized. For the case of quantized observations, a mathematical rigorous proof of its correctness is still missing. Nevertheless, simulations demonstrate the usefulness of this analysis, especially as an analytical way for the offline optimization of the system parameters, such as the quantizer characteristic, as done in [3].

The state evolution analysis is based on the assumption that the components of the estimated vector $\boldsymbol{u}^t = \boldsymbol{A}\boldsymbol{x}^t - \theta^t \boldsymbol{z}^t$ (c.f. (9)) at each iteration and the noiseless vector $\boldsymbol{w} = \boldsymbol{A}\cdot\boldsymbol{x}$ as well as the quantized observation vector $\boldsymbol{r}$ converge empirically to three random variables u, w and r given by the joint distribution

$$f(u,w,r) = \rho(r|w)\cdot\rho_G^t(u,w), \tag{48}$$

where $\rho(r|w)$ is given in (6) and $\rho_G^t(u,w)$ is the bivariate Gaussian probability density function with the covariance matrix

$$\mathrm{E}\left[\begin{bmatrix} w \\ u \end{bmatrix}^{\mathrm{T}} \cdot \begin{bmatrix} w \\ u \end{bmatrix}\right] = \boldsymbol{C}_{w,u}^t, \tag{49}$$

in other words,

$$(w,u) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{C}_{w,u}^t). \tag{50}$$

On the other hand, the components of the input vector $\boldsymbol{x}$, $\boldsymbol{v}^t = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{z}^t + \xi^t \boldsymbol{x}^{t-1}$ (c.f. (11)), as well as $\boldsymbol{x}^t$ at each iteration t converges empirically to a joint distribution described by the linear one-dimensional model

$$v^t = x + n \text{ with } n \sim \mathcal{N}(0, \nu^{t,2}), \tag{51}$$

and the nonlinear scalar estimation rule

$$x^t = f_t(v^t). \tag{52}$$

This one-dimensional equivalent model for the input step of the algorithm is illustrated in Fig. 3. In fact, the state evolution analysis implies that the performance can be fully described by the scalar one dimensional estimation problem, where the underlying observation channel is purely Gaussian and surprisingly linear and does not present any nonlinear effects except the estimation function $f_t(v)$. The state evolution equations of the input step (11) and (12) are then given by the expectations

$$\begin{aligned} c_x &= \mathrm{E}_x[x], \\ c_{x,\hat{x}}^t &= \mathrm{E}_{x,v}[x f_t(v)], \\ c_{\hat{x}}^t &= \mathrm{E}_{x,v}[f_t(v)^2], \\ \theta^t &= \beta\cdot\mathrm{E}_{x,v}[\dot{f}(v)]. \end{aligned} \tag{53}$$

On the other hand, the state evolution equations of the output step are given by [3], [5]

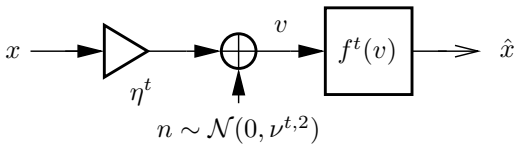

Fig. 3. Equivalent Scalar Model

$$\begin{aligned} \nu^{t,2} &= \mathrm{E}_{u,w,r}[g_t(-u,r)], \\ \xi^t &= \mathrm{E}_{u,w,r}[\dot{g}_t(-u,r)], \\ \eta^t &= \mathrm{E}_{u,w,r}[\frac{\dot{\rho}(r|w)}{\rho(r|w)} g_t(-u,r)], \end{aligned} \tag{54}$$

where $\dot{\rho}(r|w)$ is the derivative of $\rho(r|w)$ with respect to w. It can be shown that the covariance matrix $\boldsymbol{C}_{w,u}$ from (49) is given by [5]

$$(w,u^t) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{C}_{w,u}), \text{ with } \boldsymbol{C}_{w,u^t}^t = \beta \begin{bmatrix} c_x & c_{x,\hat{x}}^t \\ c_{x,\hat{x}}^t & c_{\hat{x}}^t \end{bmatrix},$$

and thus we have the following joint density function

$$\rho_G^t(w,u) = \frac{\mathrm{e}^{-\frac{(w-\frac{c_{x,\hat{x}}^t}{c_{\hat{x}}^t}u)^2}{2\beta(c_x - \frac{c_{x,\hat{x}}^{t,2}}{c_{\hat{x}}^t})}}}{\sqrt{2\pi\beta(c_x - \frac{c_{x,\hat{x}}^{t,2}}{c_{\hat{x}t}})}} \frac{\mathrm{e}^{-\frac{u^2}{2\beta c_{\hat{x}}^t}}}{\sqrt{2\pi\beta c_{\hat{x}}^t}}. \tag{55}$$

With these definitions we get the recursive formulas (54) characterizing the dynamics of the algorithm and initialized as $c_{\hat{x}}^0 = c_{\hat{x}x}^0 = 0$, $\theta^0 = \mathrm{const}$,

$$\begin{aligned} \nu^{t,2} &= \mathrm{E}[g_t(-u,r)^2] \\ &= \sum_r \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_t(-u,r)^2 \rho(r|w)\rho_G^t(w,u)\mathrm{d}u\mathrm{d}w \\ &= \sum_r \int_{-\infty}^{\infty} \bar{\rho}^t(r|\frac{c_{x,\hat{x}}^t}{c_{\hat{x}}^t}u) g_t(-u,r)^2 \cdot \frac{\mathrm{e}^{-\frac{u^2}{2\beta c_{\hat{x}}^t}}}{\sqrt{2\pi\beta c_{\hat{x}}^t}}\mathrm{d}u, \end{aligned} \tag{56}$$

and

$$\begin{aligned} \eta^t &= \sum_r \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_t(-u,r)\cdot\dot{\rho}(r|w)\rho_G^t(w,u)\mathrm{d}u\mathrm{d}w \\ &= \sum_r \int_{-\infty}^{\infty} \dot{\bar{\rho}}^t(r|\frac{c_{x,\hat{x}}^t}{c_{\hat{x}}^t}u) g_t(-u,r)\cdot\frac{\mathrm{e}^{-\frac{u^2}{2\beta c_{\hat{x}}^t}}}{\sqrt{2\pi\beta c_{\hat{x}}^t}}\mathrm{d}u, \end{aligned} \tag{57}$$

while

$$\begin{aligned} \xi^t &= \mathrm{E}[\dot{g}_t(-u,r)] \\ &= \sum_r \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \dot{g}_t(-u,r)\cdot\rho(r|w)\rho_G^t(w,u)\mathrm{d}u\mathrm{d}w \\ &= \sum_r \int_{-\infty}^{\infty} \bar{\rho}^t(r|\frac{c_{x,\hat{x}}^t}{c_{\hat{x}}^t}u)\rho(r|w)\cdot\dot{g}_t(-u,r)\cdot\frac{\mathrm{e}^{-\frac{u^2}{2\beta c_{\hat{x}}^t}}}{\sqrt{2\pi\beta c_{\hat{x}}^t}}\mathrm{d}u, \end{aligned} \tag{58}$$

where

$$\bar{\rho}^t(r|u\frac{c^{2,t}_{x,\hat{x}}}{c^t_{\hat{x}}}) = \Phi\left(\frac{r^{\mathrm{up}} - \frac{c^t_{x,\hat{x}}}{c^t_{\hat{x}}}u}{\sqrt{\beta(c_x - \frac{c^{2,t}_{x,\hat{x}}}{c^t_{\hat{x}}}) + \sigma_0^2}}\right) - \Phi\left(\frac{r^{\mathrm{lo}} - \frac{c^t_{x,\hat{x}}}{c^t_{\hat{x}}}u}{\sqrt{\beta(c_x - \frac{c^{2,t}_{x,\hat{x}}}{c^t_{\hat{x}}}) + \sigma_0^2}}\right). \tag{59}$$

We note that for the special case of unquantized measurements, above state evolution parameters takes the values [4]

$$\xi^t\big|_{\mathrm{unquantized}} = \eta^t\big|_{\mathrm{unquantized}} = 1, \tag{60}$$

and

$$\nu^{t,2}\big|_{\mathrm{unquantized}} = \sigma_0^2 + \beta \cdot (c_x + c^t_{\hat{x}} - 2c^t_{x,\hat{x}}). \tag{61}$$

On the other hand, the state evolution equations of the input step can be obtained by evaluating the expectations given in (53) for the special choice of the function $f_t(v)$ defined in (13) and the statistics of $\boldsymbol{x}$. Therefore, we consider the special case where we have that the elements of the vector $\boldsymbol{x}$ undergo the following statistics

$$x_i \sim \begin{cases} \mathcal{N}(0,1) & \text{w.p. } s \\ 0 & \text{w.p. } 1-s, \end{cases} \tag{62}$$

with s being the fraction of non-zero elements, i.e., it represents the sparsity of the vector $\boldsymbol{x}$. Then, for the ℓ_1 regularization function $f_t(v)$ defined in (19), we obtain the following results at each iteration t

$$\begin{aligned}
c_x &= \mathrm{E}[x^2] = s \\
c_{\hat{x}} &= \mathrm{E}[\hat{x}^{t,2}] \\
&= 2s \int_{-\infty}^{\infty}\int_{-\infty}^{-\frac{1}{\nu^t}} (\frac{\nu^t}{\xi^t}v + 1/\xi^t)^2 \cdot \frac{1}{2\pi}\mathrm{e}^{-\frac{x^2+(v-x\frac{\eta^t}{\nu^t})^2}{2}}\mathrm{d}v\mathrm{d}x \\
&\quad + 2(1-s)\int_{-\infty}^{-\frac{1}{\nu^t}} (\frac{\nu^t}{\xi^t}v + 1/\xi^t)^2 \cdot \frac{1}{2\pi}\mathrm{e}^{-\frac{v^2}{2}}\mathrm{d}v \\
&= 2s\left[\Phi(\frac{-1/\xi^t}{\sqrt{(b^t+1)\frac{d^t}{b^t}}})((b^t+1)\frac{d^t}{b^t} + \frac{1}{\xi^{t,2}}) \right. \\
&\quad \left. - \frac{\mathrm{e}^{-\frac{1}{2\xi^{t,2}(b^t+1)\frac{d^t}{b^t}}}}{\sqrt{2\pi}\xi^t}\sqrt{(b^t+1)\frac{d^t}{b^t}}\right] \\
&\quad + 2(1-s)\left[\Phi(\frac{-1}{\xi^t\sqrt{d^t}})(d^t + \frac{1}{\xi^{t,2}}) - \frac{\mathrm{e}^{-\frac{1}{2\xi^{t,2}d^t}}}{\sqrt{2\pi}\xi^t}\sqrt{d^t}\right],
\end{aligned} \tag{63}$$

and

$$\begin{aligned}
c^t_{x,\hat{x}} &= \mathrm{E}[x \cdot f_t(v)] \\
&= 2s\int_{-\infty}^{\infty}\int_{-\infty}^{-\frac{1}{\nu^t}} (\frac{\nu^t}{\xi^t}v + \frac{1}{\xi^t}) \cdot x \cdot \frac{1}{2\pi}\mathrm{e}^{-\frac{x^2+(v-x\frac{\eta^t}{\nu^t})^2}{2}}\mathrm{d}v\mathrm{d}x \\
&= 2s \cdot \sqrt{\frac{d^t}{b^t}} \cdot \Phi(\sqrt{\frac{b^t}{d^t}}\frac{-1}{\xi^t\sqrt{b^t+1}}),
\end{aligned} \tag{64}$$

while

$$\begin{aligned}
\mathrm{E}[\dot{f}(v)] &= \frac{2s}{\xi^t}\int_{-\infty}^{\infty}\int_{-\infty}^{-\frac{1}{\nu^t}} \frac{1}{2\pi}\mathrm{e}^{-\frac{x^2+(v-x\frac{\eta^t}{\nu^t})}{2}}\mathrm{d}v\mathrm{d}x \\
&\quad + \frac{2(1-s)}{\xi^t}\int_{-\infty}^{\infty}\int_{-\infty}^{-\frac{1}{\nu^t}} \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{v^2}{2}}\mathrm{d}v \\
&= \frac{2s}{\xi^t}\Phi(-\frac{1}{\xi^t\sqrt{(b^t+1)\frac{d^t}{b^t}}}) + \frac{2(1-s)}{\xi^t}\Phi(-\frac{1}{\xi^t\sqrt{d^t}}),
\end{aligned} \tag{65}$$

where

$$\begin{aligned}
d^t &= \frac{\nu^{t,2}}{\xi^{t,2}}, \\
b^t &= \frac{\nu^{t,2}}{\eta^{t,2}}.
\end{aligned} \tag{66}$$

On the other hand, for the ℓ_2 minimization, we obtain these parameters using the estimation function given in (20)

$$c^t_{x\hat{x}} = \frac{s\sqrt{\frac{d^t}{b}}}{1+\frac{1}{\xi^t}}, \tag{67}$$

$$c^t_{\hat{x}} = \frac{s\frac{d^t}{b} + d^t}{(1+\frac{1}{\xi^t})^2}, \tag{68}$$

$$\mathrm{E}[\dot{f}_t(v)] = \frac{1}{1+\xi^t}. \tag{69}$$

The six equations provided in (56), (57), (58) and (64), (63), (65) for the ℓ_1 or (67), (68), (69) for the ℓ_2 formulation provide the recursive formulas of the state evolution analysis, which can be used to track the dynamics of the presented algorithm and to calculate an analytical formula for the MSE after a certain number of iterations as

$$\mathrm{E}[||\boldsymbol{x} - \boldsymbol{x}^t||_2^2] = c_x + c^t_{\hat{x}} - 2 \cdot c^t_{x,\hat{x}}. \tag{70}$$

V. Numerical results

Let us consider a simulation setup with a vector $\boldsymbol{x}$ of length $M = 400$ with a sparsity ratio of s which is the average fraction of non-zero entries, which are i.i.d. and generated according to the distribution

$$x_i \sim \begin{cases} \mathcal{N}(0,1) & \text{w.p. } s \\ 0 & \text{w.p. } 1-s. \end{cases} \tag{71}$$

We apply the same scalar quantizers to the measurement vector $\boldsymbol{y} = \boldsymbol{A} \cdot \boldsymbol{x} + \boldsymbol{\eta}$, where the measurement matrix $\boldsymbol{A}$ has i.i.d. Gaussian distributed elements with variance $1/N$. In the single bit case, we use an asymmetric 1-bit quantizer since taking just sign-measurements (symmetric quantizer), extensively treated in many previous works, would not deliver information about the norm of the vector, whereas the asymmetric 1-bit quantizer exhibits this feature. This aspect of optimizing the quantizer thresholds for general resolution can be done deterministically based on the state evolution analysis. For instance, Fig. 4 shows for different values of the sparsity s and different value of β the analytical MSE (70) normalized by $c_x = s$ as function of the 1-bit threshold, assuming noiseless observations $\sigma_0 = 0$. We can observe from the Fig. 4 that choosing this threshold to correspond to the standard deviation of the noiseless output,

that is $\sqrt{\beta s}$, provides good estimation performance independently of the sparsity of the vector $\boldsymbol{x}$. That is

$$\text{Threshold}_{1-\text{bit}} = \sqrt{\text{E}[||\boldsymbol{A}\boldsymbol{x}||_2^2]/N} = \sqrt{\beta \cdot s}. \quad (72)$$

To evaluate now the effect of quantization on the reconstruction performance by simulations, we first consider the noiseless case, i.e., $\sigma_0 = 0$. For resolutions higher that 1-bit, uniform mid-riser quantizers are considered, the step-size of which can be optimized based on the state evolution results. To evaluate the effect of quantization on the reconstruction performance, we first consider the noiseless case, i.e., $\sigma_0 = 0$. In Fig. 5, we compare the performance in terms of relative mean squared error (i.e. $\text{E}[||\boldsymbol{x} - \hat{\boldsymbol{x}}||_2^2/||\boldsymbol{x}||_2^2]$) of the ℓ_1 and ℓ_2 regularized optimizations with noiseless ($\sigma_0 = 0$) 1-bit measurements and non-sparse Gaussian vector $\boldsymbol{x}$ ($s = 1$) as function of the oversampling ratio N/M, while maintaining $M = 400$ constant. 100 Monte Carlo trials have been carried out to generate the curves. Interestingly, there is hardly no difference between both formulations in that case.

Next, Fig. 6 shows for the same scenario but for a sparsity ratio of $s = 0.1$ the estimation performance in terms of the relative mean square error (MSE) of the proposed algorithm for the ℓ_1 and ℓ_2 regularized optimization as function of the oversampling ratio N/M. As expected the ℓ_1 outperforms in this case the ℓ_2 regularization, but in both cases the MSE decrease with the order of N^{-2} with respect to the number of measurements, which corresponds to the maximal possible decay predicted by the theory [9]. Furthermore, the analytical MSE curves found by the state evolution analysis of the iterative algorithm are also shown and they match closely to the experimental results.

We now concentrate on the ℓ_1 formulation and perform again for the same setting ($s = 0.1$, $M = 400$, $\sigma_0 = 0$) 100 Monte Carlo trials for varying ratio N/M and different bit resolutions (1-, 2-, 3- and 4-bit). The step sizes of the quantizers can be optimized, and the following rule-of-thumb is obtained[2]

$$\Delta_{\text{b}-\text{bit}} = 4.8 \cdot 2^{-b} \cdot \sqrt{\text{E}[||\boldsymbol{A}\boldsymbol{x}||_2^2]/N}. \quad (73)$$

It can be observed from Fig. 7 that increasing the resolution by one bit improve the MSE by roughly factor 4 (i.e. 6dB) in terms of relative MSE. This leads to the conclusion that the MSE decreases as 2^{-2b}, which is consistent to the fact that the step-size (in other words the uncertainty) decreases as 2^{-b}.

Let us now consider noisy measurements, in combination with 1-bit quantization. Again, we fix $s = 0.1$ and $M = 400$. Fig. 8 and 9 present the performance of the ℓ_1 based iterative algorithm in terms of the relative MSE for two different SNR$= \text{E}[||\boldsymbol{A}\boldsymbol{x}||_2^2]/N/\sigma_0^2$ values (i.e. the ratio of signal variance and noise variance at each output) are considered. For comparison the relative MSE in the ideal case with unquantized observations is also considered. The regularization parameter λ in the function (17) was adapted to the noise level (see the legend of the figure) by trial to get the lowest MSE. First of all, we see that the MSE decreases now inverse proportionally to the number of measurements ($\propto N^{-1}$), which means that at high oversampling ratios the estimation performance becomes noise limited. Additionally, it turns out that the high SNR value (20dB) the performance loss due to quantization is significant, while for the low SNR value (0dB), the performance loss becomes much small, where roughly tripling the number of measurements leads to the same MSE. This suggests the use of low resolution quantizers in the medium or low SNR regimes, which is a result that has been also observed in the context multi-antenna communications [10]. For comparison the analytical curves found by the state evolution algorithm are also plotted, where again we see that the performance is predicted very well.

[2]The prescalar 4.8 was obtained for $s = 0.1$ but for other sparsity levels, other values might provide better results.

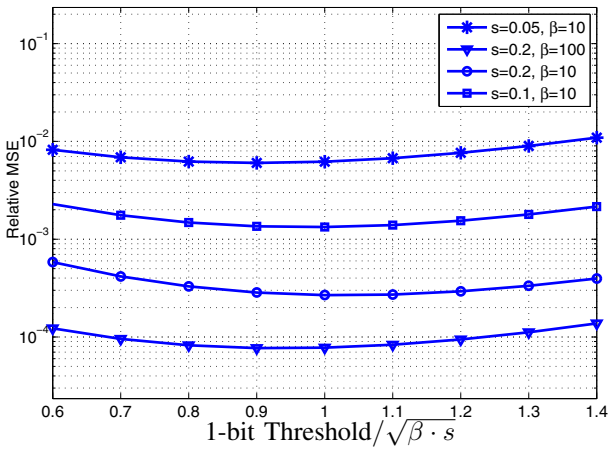

Fig. 4. State evolution results of optimizing the 1-bit threshold for different values of β and sparsity s with $\sigma_0 = 0$.

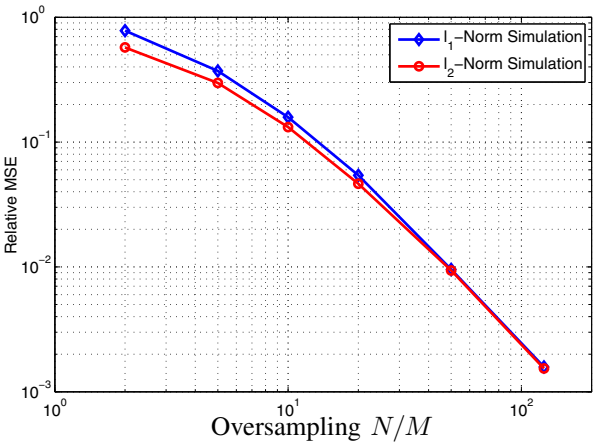

Fig. 5. MSE Performance of the ℓ_1 and ℓ_2 regularized iterative algorithms for estimating a $M = 400$ dimensional vector with no sparsity ($s = 1$) as function of the number of 1-bit noiseless measurements.

VI. Conclusion

We presented an low complexity algorithm for performing an ℓ_1 or ℓ_2 regularized estimation of a vector from quantized measurements assuming a random measurement matrix. As expected, the ℓ_1 outperforms the ℓ_2 fomulation when the vector is sparse. Contrary to previous works no prior information, in terms of density function or amplitude of the unknown vector $\boldsymbol{x}$ is need for the reconstruction. Simulations show that the iterative algorithm presents good reconstruction performance

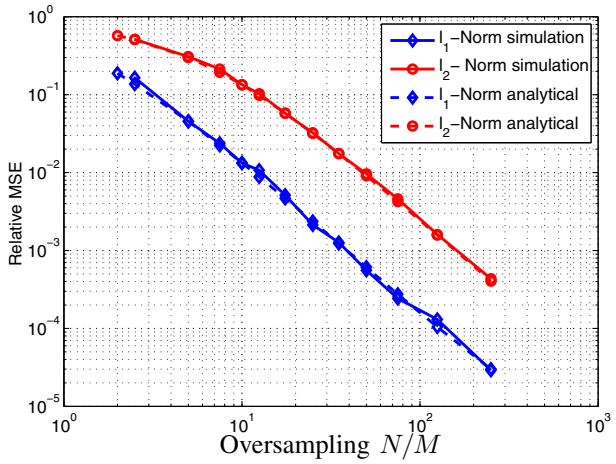

Fig. 6. MSE Performance of the ℓ_1 and ℓ_2 regularized iterative algorithms for estimating a $M = 400$ dimensional vector with sparsity ratio $s = 0.1$ as function of the number of 1-bit noiseless measurements.

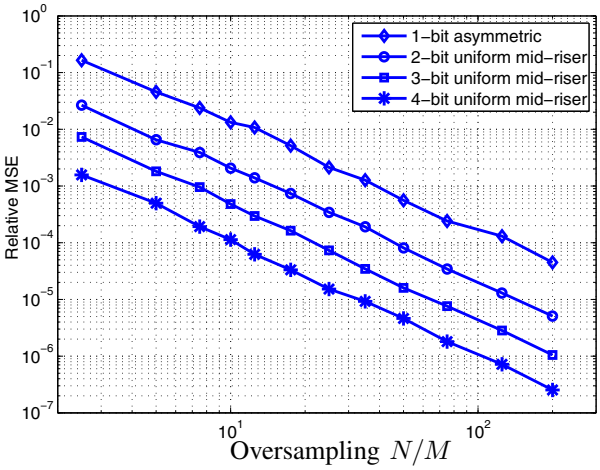

Fig. 7. Relative MSE of the ℓ_1 iterative algorithms for estimating a 400 dimensional vector with sparsity ratio $s = 0.1$ as function of the N/M and bit resolution ($\sigma_0 = 0$).

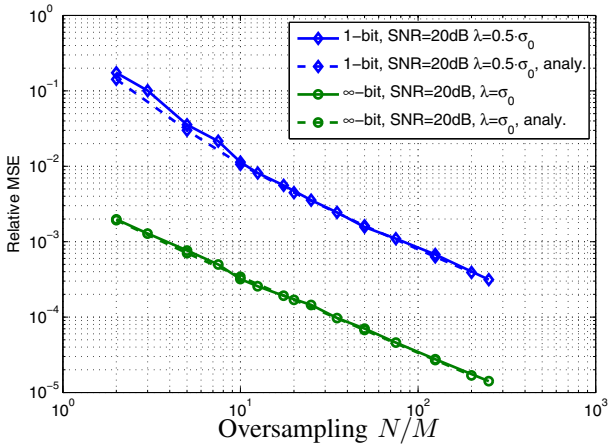

Fig. 8. MSE Performance of the ℓ_1 and ℓ_2 regularized iterative algorithms for estimating a $M = 400$ dimensional vector with sparsity ratio $s = 0.1$ as function of the number of 1-bit noisy measurements, where SNR$= \mathrm{E}[||\boldsymbol{Ax}||_2^2]/N/\sigma_0^2 \equiv 20$dB. For comparison, the performance without quantization is also shown.

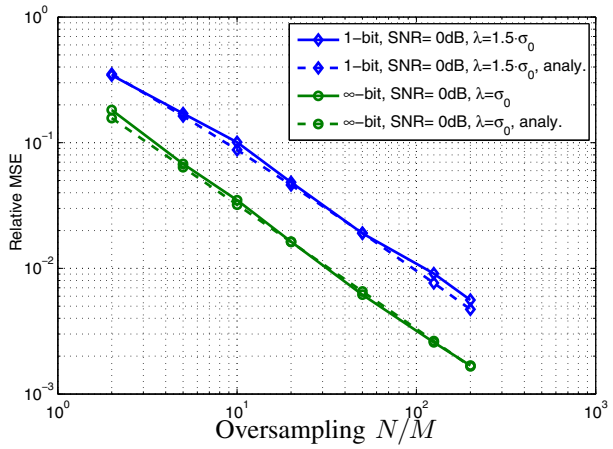

Fig. 9. MSE Performance of the ℓ_1 and ℓ_2 regularized iterative algorithms for estimating a $M = 400$ dimensional vector with sparsity ratio $s = 0.1$ as function of the number of 1-bit noisy measurements, where SNR$= \mathrm{E}[||\boldsymbol{Ax}||_2^2]/N/\sigma_0^2 \equiv 0$dB. For comparison, the performance without quantization is also shown.

while having significant low complexity. We note that the number of measurements needs to be higher than the vector dimension when the quantization is very coarse, e.g. 1-bit, which might be advantageous in some applications, since a 1-bit quantizer is a simple device and can operate at very high speed. Remarkably, when noise becomes significant, it turns out that the loss due to quantization compared to the ideal case become marginal, and the tradeoff between using finer quantization or increasing the number of observations becomes less effective. This suggests that the use of low resolution high speed sampling devices might reduce the number of measurement bits per estimated vector coefficient.

REFERENCES

[1] U. Kamilov, V. K Goyal, and S. Rangan, "Message-Passing Estimation from Quantized Samples," May 2011, arXiv:1105.6368v1.

[2] A. Zymnis, S. Boyd, and E. Candes, "Compressed sensing with quantized measurements," *IEEE Signal Process. Lett.*, vol. 17, no. 2, pp. 149–152, Feb. 2010.

[3] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," Oct. 2010, arXiv:1010.5141v1.

[4] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Nat. Acad. Sci.*, vol. 106, no. 45, pp. 18914 – 18919, Oct. 2009.

[5] A. Mezghani and J. A. Nossek, "Belief Propagation based MIMO Detection Operating on Quantized Channel Output," *IEEE International Symposium on Information Theory (ISIT), Seoul, South Korea*, June 2010.

[6] M. Bayati and A. Montanari, "The dynamics of message passing on dense graphs, with applications to compressed sensing," Jan. 2010, arXiv:1001.3448v4.

[7] D. Guo and C.-C. Wang, "Multiuser Detection of Sparsely Spread CDMA," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 3, pp. 421–431, April 2008.

[8] T. J. Richardson and R. L. Urbanke, "The capacity of low-density parity check codes under message-passing decoding," *IEEE Trans. Inform. Theory, vol.*, vol. 47, no. 2, pp. 599–618, Feb. 2001.

[9] N. T. Thao and M. Vetterli, "Reduction of the MSE in R-times oversampled A/D conversion from $O(1/R)$ to $O(1/R^2)$," *IEEE Trans. Signal Process.*, vol. 42, no. 1, pp. 200–203, Jan. 1994.

[10] A. Mezghani and J. A. Nossek, "On Ultra-Wideband MIMO Systems with 1-bit Quantized Outputs: Performance Analysis and Input Optimization," *IEEE International Symposium on Information Theory (ISIT), Nice, France*, June 2007.