Abstract—Compressed sensing (CS) studies the recovery of a high dimensional signal from its low dimensional linear measurements under a sparsity prior. This paper is focused on the CS problem with quantized measurements. An algorithm is proposed based on a Bayesian perspective that treats measurement noises and quantization errors separately and allows data saturation. It is shown to improve the recovery accuracy in comparison with existing approaches by numerical simulations.

I. INTRODUCTION

Conventional sampling systems rely on the Shannon-Nyquist sampling theorem, which states that for its recovery a signal must be sampled at least at the Nyquist rate, a rate twice its bandwidth. The recently proposed compressed sensing (CS) framework [1], [2] breaks through such a constraint and has attracted much attention. In CS, the information content of a signal is not measured by its bandwidth but the sparsity. A signal is shown to improve the recovery accuracy in comparison with deterministic approaches, e.g., BPDN.

The conventional CS framework is mainly focused on the sparse signal recovery from the real-valued measurement $y$. The number of measurements $M$ is mainly studied for the recovery accuracy [10]–[13]. Since quantization is necessary for practical considerations, e.g., data storage and transmission, we study the sparse signal recovery from quantized measurements in this paper. During the quantization process, each continuous-valued measurement is quantized into some value in a finite set. A new challenge is thus the existence of quantization errors. A uniform unsaturated quantizer is considered in [3] where a naive BPDN method is adopted that treats the quantization errors as Gaussian noises. Since quantization errors are bounded in such a case and quite different from Gaussian noises, Jacques et al. [14] study the noise free case and show that BPDN is generally not optimal in recovering $x$ by characterizing the quantization errors as independent variables uniformly distributed in a common interval. They propose a family of decoders named as basis pursuit dequantizer of moment $p$ (BPDQ$_p$) with $p \geq 2$. One shortcoming of BPDQ$_p$ is that the optimal $p$ cannot be explicitly given in practice. It should be noted that both BPDN and BPDQ$_p$ are inappropriate in the case of a saturated quantizer since data saturation may lead to large

Accurate Signal Recovery in Quantized Compressed Sensing

Zai Yang Lihua Xie
EXQUISITUS, Centre for E-City
School of Electrical and Electronic Engineering
Nanyang Technological Technology
Singapore, 639798
Email: {yang0248, elhxie}@ntu.edu.sg

Cishen Zhang
Faculty of Engineering and Industrial Sciences
Swinburne University of Technology
Hawthorn VIC 3122, Australia
Email: cishenzhang@swin.edu.au
or even unbounded quantization errors that deteriorate their performance. To deal with the data saturation, Laska et al. [15] propose two modified versions of BPDN, including saturation rejection and consistency approaches, denoted by BPDN-SR and BPDN-SC respectively. The saturated measurements are simply rejected in BPDN-SR while they are incorporated in the signal recovery process in BPDN-SC. While quantization errors and measurement noises are coupled in most existing methods (some methods, e.g., BPDQ, consider only the noise free case to avoid such a problem), they are separately studied by Zymnis et al. [16] where the authors seek to find a signal estimate that maximizes the likelihood of the quantized measurements while \( \ell_1 \) norm is exploited to promote the signal sparsity. The resulting algorithm is quoted as \( \ell_1 \)-regularized maximum likelihood (L1RML).

In this paper, we introduce the Bayesian approach into quantized CS. The proposed algorithm, quoted as Q-VMP, is based on VMP and has guaranteed convergence property. In the proposed algorithm, we treat the quantization errors and measurement noises separately and seek to estimate the quantization errors along with the signal estimation process. A three-layer hierarchical prior recently proposed in [4] is adopted as the sparse signal prior. The proposed Bayesian framework allows data saturation. Numerical simulations show that Q-VMP can reduce the signal estimation error up to an amplitude in comparison with BPDN and L1RML.

An extreme case of quantized CS is so-called 1-bit CS [17] where each quantized measurement uses just one bit and keeps only the sign information of the real measurement. Since the amplitude information of the signal is lost, the 1-bit CS problem is quite different from the “multi-bit” case studied in this paper. Interested readers are referred to [18] for the extension of Q-VMP to such a case.

Notations used in this paper are as follows. Bold-face letters are reserved for vectors and matrices. For ease of exposition, we do not distinguish a random variable and the value it takes. \( x_i \) is the \( i \)th entry of a vector \( x \). \( A_i \) is the \( i \)th column of a matrix \( A \). \( \| x \|_1 \) and \( \| x \|_2 \) denote the \( \ell_1 \) and \( \ell_2 \) norms of a vector \( x \) respectively. \( (g(x))_{p(x)} \) is an expectation of a function \( g(x) \) with respect to a random variable \( x \) whose probability density function is \( p(x) \).

The rest of the paper is organized as follows. Section II introduces our observation model and sparse Bayesian formulation for quantized CS. Section III introduces the proposed Q-VMP algorithm. Section IV presents numerical simulations to illustrate the performance of the proposed Q-VMP algorithm in comparison with existing ones. Section V concludes the paper and highlights some future works.

II. Observation Model and Bayesian Formulation for Quantized CS

A. Observation Model

In quantized CS, the observed samples are noisy linear measurements of the original signal after quantization:

\[
z = Q(y), \quad y = Ax + n \tag{3}
\]

where \( x \) is the signal of interest, \( A \) is the sensing matrix, \( n \) is the measurement noise vector, \( y \) is the pre-quantized noisy measurement vector, \( Q \) denotes a quantizer and \( z \) is the observation. A (regular) quantizer \( Q(v) \) for a scalar \( v \in \mathbb{R} \) is defined as

\[
Q(v) = \begin{cases} 
0, & \text{if } v \in (u_0, u_1), \\
1, & \text{if } v \in [u_1, u_2), \\
\cdots, & \text{if } v \in [u_L-1, u_L), \\
L, & \text{if } v \in [u_L, u_{L+1}) \end{cases} \tag{4}
\]

where \( L \) denotes the quantization level and typically satisfies \( L = 2^B \) with \( B \) denoting the bit-depth (bits per quantized measurement), \( u_0 < u_1 < \cdots < u_L \), and \( v_i \in [u_i, u_{i+1}) \) for \( i = 0, \cdots, L-1 \). The quantizer \( Q(v) \) is called unsaturated if \((u_0, u_L)\) is a finite interval, or saturated otherwise. For a vector \( v \), \( Q(v) \) operates elementwise.

Denote \( D_y \) the domain of \( y \). Then we have

\[
D_y = Q^{-1}(z) := \{ y \in \mathbb{R}^M | Q(y) = z \}. \tag{5}
\]

We introduce an auxiliary variable \( e \) denoting the quantization error:

\[
e = z - y \tag{6}
\]

with its domain

\[
D_e = z - D_y := \{ z - y | y \in D_y \}. \tag{7}
\]

Note that \( D_e \) is unbounded as data saturation occurs. The model in (3) can be rewritten into

\[
z = Ax + e + n, \quad e \in D_e, \tag{8}
\]

which is the observation model to be used in this paper to recover \( x \).

Remark 1: In the observation model (8), \( y \) is written into two variables \( z \) and \( e \) that are dependent. As an example, the statistical information of \( e \) can be obtained given that of \( x \) and \( n \) for a fixed quantizer \( Q \). For the sake of computational ease, such dependence will not be exploited in the following signal estimation process, i.e., we assume that there is no a priori information of \( e \) but its domain \( D_e \).

B. Sparse Bayesian Formulation

In this subsection we formulate the observation model in (8) from a Bayesian perspective. The joint probability density function (PDF) \( p(z, x, e) \) is decomposed as

\[
p(z, x, e) = p(z|x, e)p(x)p(e). \tag{9}
\]

In the following, we define the three distributions on the right hand side.

1) Noise model: Under an assumption of white Gaussian measurement noise, i.e., \( n \sim \mathcal{N}(0, \sigma_n^2 I) \) where \( \sigma_n^2 \) is the noise variance and \( I \) denotes an identity matrix of proper dimension, we have

\[
p(z|x, e; \sigma_n^2) = \mathcal{N}(z|Ax + e, \sigma_n^2 I). \tag{10}
\]

Remark 2: In this paper, we assume that \( \sigma_n^2 \) is known. Though it can be estimated by assuming an inverse Gamma
Gaussian-Gamma-Gamma, hierarchical prior for sparse signal [19].

Fig. 1. A factor graph that encodes the joint PDF in (17) of our model, where $f_z = p(z|x,e)$, $f_x = p(x|\alpha)$, $f_\alpha = p(\alpha|\eta)$, $f_\eta = p(\eta)$ and $f_e = p(e)$.

prior for it in the case whether it is unknown, its estimate is inaccurate due to an “identifiability issue” as addressed in [19].

2) Sparse signal model: A sparse prior is needed for the sparse signal $x$ of interest. Inspired by [4], a three-layer, Gaussian-Gamma-Gamma, hierarchical prior for $x$ is exploited as

$$p(x; \epsilon, c, d) = \int \int p(x|\alpha) p(\alpha|\eta, \epsilon) p(\eta; c, d) \, d\alpha \, d\eta$$

where

$$p(x|\alpha) = N(x|0, \Lambda),$$

$$p(\alpha|\eta, \epsilon) = \prod_{i=1}^{\epsilon} \Gamma(\alpha_i|\eta, \epsilon),$$

$$p(\eta; c, d) = \Gamma(\eta|c, d)$$

with $\Lambda = \text{diag}(\alpha)$ and constants $\epsilon$, $c$, $d$. For a Gamma distributed variable $u \sim \Gamma(c, d)$, its PDF is

$$\Gamma(u|c, d) = \frac{du}{\Gamma(c)} u^{c-1} e^{-u}$$

with $\Gamma(c)$ being the Gamma function. By [4] the constants $\epsilon$, $c$, $d$ satisfy that $0 \leq \epsilon \leq 1$, $c, d \geq 0$. In this paper, we adopt $c = 1$, $d = 0$ to make the prior for $\eta$ in (14) noninformative (flat on $\mathbb{R}_+$). Further, we choose $\epsilon = 0$ since a smaller $\epsilon$ leads to a sparser prior and an estimator that approximates a hard-thresholding rule according to [4]. Readers are referred to [4] for more properties of the Gaussian-Gamma-Gamma prior and its relations with other sparse estimation techniques.

3) Quantization error model: We assume a uniform, non-informative prior for $e$ according to Remark 1:

$$e \sim U(D_e)$$

since the only information of $e$ that we use is $e \in D_e$.

As a result, we have the joint PDF of the observation model (8):

$$p(z, x, e, \alpha, \eta) = p(z|x, e) p(x|\alpha) p(\alpha|\eta) p(\eta) p(e)$$

with the distributions on the right hand side as defined respectively by (10), (12), (13), (14) and (16). A factor graph [20] that encodes the factorization of the joint PDF in (17) is shown in Fig. 1.

III. Q-VMP ALGORITHM

A. Variational Inference Using VMP

As is known that the Bayesian inference is based on the posterior distribution

$$p(x, e, \alpha, \eta|z) = \frac{p(z, x, e, \alpha, \eta)}{p(z)}$$

However, such an exact posterior distribution is intractable since

$$p(z) = \int \cdots \int p(z, x, e, \alpha, \eta) \, dx \, de \, d\alpha \, d\eta$$

cannot be expressed explicitly.

A variational inference approach [21] is adopted in this paper. Denote $V = \{x, e, \alpha, \eta\}$ the set of all unknown variables to be estimated. The goal in variational inference is to find a tractable distribution $q(V)$ that closely approximates the true posterior distribution $p(V|z)$. To do this, some family of distributions that has enough flexibility is firstly chosen to represent $q(V)$. Then the task is to find a member of the family that minimizes the Kullback-Leibler (KL) divergence between the true posterior $p(V|z)$ and the variational approximation $q(V)$. A commonly used variational distribution $q(V)$ is such that disjoint groups of variables are independent, i.e., $q(V)$ has a factorized form

$$q(V) = q(x) q(e) q(\alpha) q(\eta).$$

Variational message passing (VMP) is proposed in [8] for the variational inference using a message passing procedure on a graphical model. In VMP, the variational distributions $q(x)$, $q(e)$, $q(\alpha)$, $q(\eta)$ are iteratively updated to monotonically decrease the KL divergence and thus has guaranteed convergence. For a random variable $v \in V$, the variational distribution $q(v)$ is updated as the product of incoming messages from its neighboring factor nodes, i.e.,

$$q(v) \propto \prod_{f \in S_v} m_{f\rightarrow v}$$

where $S_v$ denotes the set of factor nodes neighboring $v$, $m_{f\rightarrow v}$ denotes the message from a factor node $f$ to $v$ and

$$m_{f\rightarrow v} = \text{exp} \left\{ \langle \ln f \rangle_{q(v)} \right\}$$

with $\langle \cdot \rangle_{q(v)}$ denoting an expectation with respect to all variables except $v$. Readers are referred to [8] for more details. The updates of $q(x)$, $q(\alpha)$, $q(\eta)$ are similar to those in [4] because of the similarity between quantized CS and conventional CS.

1) Update of $q(x)$: According to (22) we have

$$q(x) \propto m_{f_x\rightarrow x} m_{f_v\rightarrow x} \propto \text{exp} \left\{ \langle \ln p(z|x,e) \rangle_{q(e)} \right\} \text{exp} \left\{ \langle \ln p(x|\alpha) \rangle_{q(\alpha)} \right\} \propto \text{exp} \left\{ -\frac{1}{\sigma_n^2} \|z - \langle e \rangle_{q(e)} - A x \|^2_2 \right\}$$

$$-\frac{1}{2} x^T \langle \Lambda^{-1} \rangle_{q(\alpha)} x \right\}.$$
Thus, \( q(\mathbf{x}) \) is a Gaussian distribution \( \mathcal{N}(\mathbf{x} | \mu, \Sigma) \) with mean \( \mu \) and covariance \( \Sigma \):

\[
\mu = (x)_{q(x)} = \sigma_n^{-2} \Sigma A^T \left( z - (e)_{q(e)} \right), \quad (24)
\]

\[
\Sigma = \left( \sigma_n^{-2} A^T A + (\Lambda^{-1})_{q(\alpha)} \right)^{-1}. \quad (25)
\]

2) Update of \( q(\alpha) \):

\[
q(\alpha) \propto m_{f_{\alpha}} \rightarrow \alpha m_{f_{\alpha}} \rightarrow \alpha
\]

\[
\propto \prod_{n=1}^{N} \alpha_n^{-2} \exp \left\{ -\frac{1}{2} \alpha_n^{-1} \langle x_n^2 \rangle_{q(x)} - \alpha_n \langle \eta \rangle_{q(\eta)} \right\}
\]

where \( \langle x_n^2 \rangle_{q(x)} = \mu_n^2 + \Sigma_{nn} \). The expression on the right hand side is the product of generalized inverse Gaussian (GIG) PDFs [22]. The moments of the GIG distribution are given in closed form for any \( i \in \mathbb{R} \) [22]:

\[
\langle \alpha_i \rangle_{q(\alpha)} = \left( \frac{2}{\langle \eta \rangle_{q(\eta)}} \right)^{\frac{i}{2}} \frac{K_{\nu + i - \frac{1}{2}} \left( \sqrt{2 \langle \eta \rangle_{q(\eta)} \langle x_n^2 \rangle_{q(x)}} \right)}{K_{\nu - \frac{1}{2}} \left( \sqrt{2 \langle \eta \rangle_{q(\eta)} \langle x_n^2 \rangle_{q(x)}} \right)}
\]

where \( K_{\nu}(\cdot) \) is the modified Bessel function of the second kind and order \( \nu \in \mathbb{R} \). The case of \( i = -1 \) in (27) gives the evaluation of \( (\Lambda^{-1})_{q(\alpha)} \) used in (25), and the case of \( i = 1 \) gives the calculation of \( q(\alpha) \) in the later expression in (29).

3) Update of \( q(\eta) \): The update of \( q(\eta) \) can be shown to be

\[
q(\eta) = \Gamma \left( \eta | N \epsilon + c \sum_{n=1}^{N} \langle \alpha_n \rangle_{q(\alpha)} + d \right).
\]

The first moment of \( \eta \) used in (27) is given as

\[
\langle \eta \rangle_{q(\eta)} = \frac{N \epsilon + c}{\sum_{n=1}^{N} \langle \alpha_n \rangle_{q(\alpha)} + d}.
\]

4) Update of \( q(e) \): We finally derive the expression of the update of \( q(e) \):

\[
q(e) \propto m_{f_{\epsilon}} \rightarrow \epsilon m_{f_{\epsilon}} \rightarrow \epsilon
\]

\[
\propto \exp \left\{ \langle \ln p(z|x, e) \rangle_{q(x)} \right\} p(e)
\]

\[
\propto \exp \left\{ -\frac{1}{2} \sigma_n^{-2} \langle \|z - Ax\|_2^2 \rangle_{q(x)} \right\} I_e(D_e)
\]

\[
\propto \exp \left\{ -\frac{1}{2} \sigma_n^{-2} \langle \|e - (z - A \mu)\|_2^2 \rangle \right\} I_e(D_e)
\]

where \( I_e(D_e) \) is an indicator function that equals to 1 if \( e \in D_e \), or 0 otherwise. Hence, \( q(e) \) is the product of PDFs of truncated Gaussian distributions, i.e., for each \( m = 1, \ldots, M \), \( q(e_m) \) is the PDF of a truncated Gaussian distribution. As a result, the first moment of \( e_m \), \( m = 1, \ldots, M \), used in (24) can be given in closed form using the PDF \( \phi(\cdot) \) and cumulative distribution function (CDF) \( \Phi(\cdot) \) of a standard Gaussian distribution:

\[
\langle e_m \rangle_{q(e)} = \sigma_n \frac{\phi(l_{e_m}) - \phi(u_{e_m})}{\Phi(u_{e_m}) - \Phi(l_{e_m})} + \mu_{e_m}.
\]

where \( \mu_{e_m} = (z - A \mu)_{m} \), \( l_{e_m} \) and \( r_{e_m} \) satisfy that \( D_{e_m} = \{ \sigma_n l_{e_m} + \mu_{e_m}, \sigma_n r_{e_m} + \mu_{e_m} \} \) with \( D_{e_m} \) denoting the domain of \( e_m, \phi(u) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{u^2}{2} \right\} \) and \( \Phi(u) = \int_{-\infty}^{u} \phi(t) \; dt \) for \( u \in \mathbb{R} \).

The resulting algorithm is summarized in Algorithm 1, named as variational message passing with quantization (Q-VMP).

Algorithm 1: Q-VMP

Input: sensing matrix \( A \), quantized measurement \( z \), domain of quantization error \( D_\epsilon \), and noise variance \( \sigma_n^2 \).

1. initialize \( \langle \alpha_n^{-1} \rangle_{q(\alpha)}, n = 1, \ldots, N, \langle \eta \rangle_{q(\eta)} \) and \( \langle e \rangle_{q(\epsilon)} \);

2. while not converged do

3. update \( \Sigma \) according to (25);

4. update \( \mu \) according to (24);

5. update \( \langle \alpha_n^{-1} \rangle_{q(\alpha)} \) and \( \langle \alpha_n \rangle_{q(\alpha)} \), \( n = 1, \ldots, N \), according to (27);

6. update \( \langle \eta \rangle_{q(\eta)} \) according to (29);

7. update \( \langle e \rangle_{q(\epsilon)} \) according to (31);

8. end while

Output: recovered signal \( \hat{x} = \mu \).

B. Pruning a Basis Function

The most difficult computation of Q-VMP is the calculation of \( \Sigma \) that is the inverse of an \( N \times N \) matrix. Using the Woodbury matrix identity, we have

\[
\Sigma = (\Lambda^{-1})_{q(\alpha)} - (\Lambda^{-1})_{q(\alpha)} A^T C^{-1} A (\Lambda^{-1})_{q(\alpha)}^{-1}
\]

(32)

with \( C = \sigma_n^2 I + A (\Lambda^{-1})_{q(\alpha)} A^T \) being an \( M \times M \) matrix. Hence, to calculate \( \Sigma \) needs \( O \left( \min \{ N^3, N^2 M \} \right) \) operations. It is noted that if Q-VMP produces some \( \langle \alpha_n^{-1} \rangle_{q(\alpha)} \rightarrow +\infty \) with \( n \in \{1, \ldots, N\} \), then the corresponding parameter \( \alpha_n \) can be removed from the model. To further speed up Q-VMP, we prune a basis \( A_n \) from the model (to reduce \( N \)) when the corresponding parameter \( \langle \alpha_n^{-1} \rangle_{q(\alpha)} \) is larger than a certain threshold \( \tau_{pruning} \). Similar basis pruning approaches have been used in [4], [23].

IV. NUMERICAL SIMULATIONS

In this section, we study the empirical performance of the proposed Q-VMP algorithm in comparison with existing ones by numerical simulations. We set the free parameters \( \epsilon = 0, c = 1 \) and \( d = 0 \) in the Gaussian-Gamma-Gamma prior model as discussed in Subsection II-B.

A. The Unsaturated Quantizer Case

We consider the case of a uniform (unsaturated) quantizer in this subsection where \( u_0, u_1, \ldots, u_L \) are equispaced and \( v_i = (u_i + u_{i+1}) / 2, i = 1, \ldots, L - 1 \), in (4). In our experiment, we set \( N = 500, K = 10 \), the bit depth (bits per measurement) \( B = 4 \), and vary the bit budget (total bits of all measurements) in \{50, 100, \ldots, 1000\}. We set the signal
to noise ratio SNR = 30dB. In each trial, a K-sparse signal of length N is generated with Gaussian distributed nonzero locations and then scaled to unit norm. Entries of the sensing matrix A are generated independently according to a Gaussian distribution N(0, M\(^{-1}\)). Thus the noise free measurement \(y^0 = Ax\) has unit norm in expectation. Then a measurement noise \(n\) is added whose entries are independently Gaussian \(N(0, \sigma_n^2 I)\) with noise variance \(\sigma_n^2 = M^{-1} 10^{-6/10}\) leading to the desired SNR. To get a uniform quantizer, we define simply the quantizer in the range \((-\|y\|_\infty, \|y\|_\infty)\) with the quantization level \(L = 2^B\). The quantized measurement \(z = Q(y)\) is preserved for the following signal recovery. Besides Q-VMP, we also use BPDN [3] and L1RML [16] to recover the signal for comparison. Three metrics are recorded, including reconstruction SNR (RSNR), sparsity (support size) of the recovered signal and speed. RSNR is defined as

\[
\text{RSNR} = -20 \log_{10} \|x - \hat{x}\|_2
\]

where \(\hat{x}\) denotes the recovered signal of \(x\). Speed is measured by the CPU time consumption. All results are averaged over 200 trials for each bit budget. The three algorithms are implemented as follows.

**Q-VMP:** We initialize \(\langle \alpha^{-1}_n \rangle_{q(\alpha)} = 1/\|A_n^T z\|\), \(n = 1, \ldots, N\), \(\langle \eta \rangle_{q(\eta)} = 1\) and \(\langle e \rangle_{q(e)} = 0\). We set \(\tau_{\text{pruning}} = 10^4\). Q-VMP is terminated if \(\|\alpha^j - \hat{\alpha}^j\|_2 < 10^{-5}\) or the maximum number of iterations, set to 2000, is reached, where \(\hat{\alpha} = [\langle \alpha^{-1}_1 \rangle_{q(\alpha)}, \ldots, \langle \alpha^{-1}_N \rangle_{q(\alpha)}]^T\) and the superscript \(j\) indicates the iteration.

**BPDN:** BPDN solves the problem in (2) with \(y\) replaced by \(z\) and is implemented using \(\ell_1\)-magic [24]. We let \(\epsilon = \|z - Ax\|_2\) in our implementation to achieve the best result though it is unavailable in practice.

**L1RML:** L1RML solves the problem

\[
\min_{\hat{x}} \{ \lambda \|\hat{x}\|_1 - \log f_{ML}(A\hat{x}) \}
\]

where \(f_{ML}(\cdot)\) is the likelihood function of the observation with \(-\log f_{ML}(A\hat{x})\) being a convex function of \(\hat{x}\), and \(\lambda > 0\) is a regularization parameter. In general, a larger \(\lambda\) leads to a sparser solution. Since there are no good approaches to the choice of \(\lambda\) so far, we choose \(\lambda\) to our best such that its recovered signal has the optimal RSNR in our implementation. Additionally, we set \(\tau = \sigma_n^2 / A_n^2\), \(\epsilon = 10^{-4}\) and \(\beta = 0.5\). Readers are referred to [16] for the meanings of the parameters.

The experimental results are shown in Fig. 2, where red solid lines denote Q-VMP, black dashed dot lines denote BPDN, and blue dashed lines denote L1RML. Fig. 2(a) depicts the averaged reconstruction SNRs of the three algorithms. A significant improvement of the reconstruction SNR can be observed using the proposed Q-VMP. It is over 6dB in comparison with L1RML and about an amplitude for BPDN. In addition, Fig. 2(b) shows that the recovered signal of Q-VMP is sparser in comparison with those of BPDN and L1RML. It is noted that L1RML can produce a sparser solution by setting a larger value of \(\lambda\) as in [16] but at the cost of a lower RSNR since it is implemented to produce the optimal RSNR in our implementation. Fig. 2(c) shows that the speed of Q-VMP is comparable with that of BPDN and L1RML. Implemented with the basis pruning approach, Q-VMP is faster when more measurements are acquired since in such a case the basis pruning process tends to be faster.

**B. The Saturated Quantizer Case**

We next consider the case of a saturated quantizer. We adopt the same experimental setup but a saturated quantizer where a noisy measurement falls in each quantization interval with the same probability. By that both sensing matrix and measurement noise are Gaussian in the experiment, the noisy measurements are i.i.d. Gaussian \(N(0, M^{-1} + \sigma_n^2)\). Then it is easy to get the quantizer. As a result, 12.5% of the measurements are saturated in expectation. BPDN is inappropriate in such a case. We compare Q-VMP only with L1RML. The averaged reconstruction SNRs of the proposed Q-VMP and L1RML are presented in Fig. 3 (red solid lines). Q-VMP obtains a RSNR of about 10dB higher than L1RML when sufficient measurements are acquired. The performance of the two algorithms on support size and speed is similar to that in the uniform quantizer case and omitted.
Though the experiment above is carried out to illustrate the performance of Q-VMP in the case of a saturated quantizer, it may shed light on the optimal quantizer design for Q-VMP. By comparing the performance of Q-VMP in the two scenarios, it can be seen from Fig. 3 that the saturated quantizer outperforms the uniform unsaturated one when more measurements are taken for Q-VMP while it is not so clear for L1RML. We pose the problem of the optimal quantizer design for Q-VMP as a future work.

V. CONCLUSION AND FUTURE WORK

We studied the sparse signal recovery problem from quantized noisy compressive measurements in this paper. We proposed an algorithm that incorporates measurement noises and applies to a saturated quantizer. Numerical simulations were carried out, showing that the proposed algorithm has improved accuracy in comparison with existing ones.

One drawback of Q-VMP is its high computational complexity due to an inversion of a high dimensional matrix at each iteration though it has been greatly alleviated with a basis pruning approach adopted in this paper. Q-VMP may suffer from computational issues in the case of very high dimensional problems though such issues have not been encountered in simulations presented in this paper. One future work is to develop fast alternatives to the current implementation. Since the signal recovery accuracy in quantized CS is quite different when a different quantizer is adopted, as shown in the present paper and in [15], the optimal quantizer design that minimizes the signal recovery error is another future work.

REFERENCES