

A Learning-Based Framework for Quantized Compressed Sensing

Rabeeh Karimi Mahabadi*, Junhong Lin[†], Volkan Cevher[†]

*Idiap Research Institute

[†]Laboratory for Information and Inference Systems
École Polytechnique Fédérale de Lausanne

Abstract—Sparse recovery from undersampled random quantization measurements is a recent active research topic. Previous work asserts that stable recovery can be guaranteed via the basis pursuit dequantizer (BPDQ) if the measurements number is large enough, considering random sampling patterns. In this paper, we study a learning-based method for optimizing the sampling pattern, within the framework of sparse recovery via BPDQ from their uniformly quantized measurements. Given a set of representative training signals, the method finds the sampling pattern that performs the best on average over these signals. We compare our approach with the random sampling and other state-of-the-art sampling methods, which shows that it achieves superior reconstruction performance. We demonstrate that proper accounting for sampling and careful sampler design has a significant impact on the performance of quantized compressive sensing methods.

Index Terms—Compressed sensing, quantized measurements, learning-based method, greedy algorithm, sampling pattern.

I. INTRODUCTION

The basic goal of compressed sensing (CS) [1, 2, 3] is to recover a K -sparse or compressible signal $\mathbf{x} \in \mathbb{R}^N$ from a small number of linear measurements compared to the input dimension of the signal. Such a problem is ubiquitous in signal processing, having an extensive range of applications including medical resonance imaging (MRI) [4, 5], spectroscopy [6], radar [7], Fourier optics [8, 9], and computerized tomography [10, 11]. With the help of the restricted isometry property [2], classical CS theory [3] asserts that basis pursuit denoising (BPDN) can stably recover the target signal provided the measurement number is of order $K \log^c N$ for some value $c > 0$, considering sub-Gaussian, random partial Fourier, or random partial Hadamard measurement matrices [12, 13].

Although in traditional CS methods, the measurement matrix and vector are assumed to be real-valued, quantization is unavoidable in the real acquisition systems. In practice, analog measurements are quantized and mapped from real values to discrete values of quantization points. Sometimes, quantization bit-rate is enforced by the hardware. In addition, quantization strategies are preferred in real-world sensing systems since low-bit data acquisition is efficient and inexpensive. In these scenarios due to limited bandwidth and energy constraints, quantized measurements are beneficial. Moreover, having low-bit measurements helps in substantially decreasing the cost of converting the analog signals to the digital ones, and simplifying the complexity of the needed hardware. It has

been also shown that quantized CS strategies tend to be more robust to amplification and other errors [14]. These practical necessities and advantages motivate the quantization of measurements in CS.

To accommodate the distortions created by quantization in the measurements, one may view the quantization error as an l_2 -bounded noise and approximate the target signal via solving BPDN. However, theoretically, the quantization distortions are highly non-Gaussian, and treating the quantization error as alternative bounded noise is more preferable [15, 16]. Particularly, Jacques et al. [15] modeled the quantization error as an l_p -bounded noise with $p > 2$, and they theoretically showed that such a method can model the quantization distortion more faithfully than the conventional l_2 -bounded approach. Based on this, they proposed a convex optimization-based decoder titled BPDQ of moment p (BPDQ _{p}), and showed that BPDQ _{p} performs better than BPDN according to their numerical results considering random sampling patterns.

While considering random sampling patterns plays a key role for developing theoretical results of classical CS and quantized CS, in practice, selecting the sampling pattern in an advantaged/greedy way using extra available information would significantly improve the performance of the decoder. Indeed, the so-called dictionary learning [17, 18], adapted sampling [4, 19], learning-based methods [20, 21] are falling into this category. Particularly, the learning-based methods have been successfully applied to classical CS, either with a linear decoder [21] or with a BPDN decoder [20].

In this paper, we study quantized CS, i.e., sparse recovery via BPDQ _{p} from undersampled quantization measurements. We consider the measurement matrix of the form $\Phi = P_\Omega \Psi$, consisted of a subsampling operator P_Ω indexed with a subset $\Omega \subseteq \{1, \dots, N\}$ and an orthonormal matrix Ψ . We propose a learning-based method for optimizing the sampling pattern Ω . Given a set of representative training signals, the method finds the sampling pattern that performs the best on average over these signals using a greedy algorithm. We experimentally show that our proposed method efficiently reconstructs the signals and substantially improves the recovery performance. Recently, [22, 23] proposed a deep learning framework to jointly optimize for the binary measurement matrix with elements constraint in $\{-1, 1\}$ and a non-uniform quantizer. However, the focus of this paper is not on using a quantized measurement matrix, whereas efficiently computing

a mask indices by employing the learning-based approach. We investigated the effect of the sampling patterns in quantized CS scenarios.

The rest of this paper is structured as follows. In Section II, we briefly review the principles of quantized CS and BPDQ_p decoder. Section III introduces the greedy mechanism to optimize the sampling patterns. To demonstrate the power of sampling method, we provide experimental results on signal and image reconstruction problems in section IV. We finally conclude the paper in Section V.

II. QUANTIZED COMPRESSED SENSING

In this section, we briefly review the mathematical formulation and BPDQ decoder for quantized CS.

In CS [1, 24], signals are assumed to have a sparse representation in an orthogonal basis $\Psi' \in \mathbb{R}^{N \times N}$. Typically, the measurement matrix Φ has a special form $\Phi = P_\Omega \Psi$, where $P_\Omega : \mathbb{R}^N \rightarrow \mathbb{R}^m$ is a subsampling operator that selects rows of another orthonormal basis Ψ given in the index set $\Omega \subseteq \{1, \dots, N\}$, with $|\Omega| = m$ and $m < N$. We refer to Ω as the sampling pattern. Usually, the sampling indexes are chosen randomly in classical CS theory.

In classical CS, one assumes that the measurement matrix Φ and the measurement vector $\mathbf{y} = \Phi \mathbf{x} + \mathbf{n}'$ are given, where \mathbf{n}' is typically assumed to be a Gaussian noise. Here, for simplicity, we consider the noiseless case, i.e., $\mathbf{n}' = 0$. To further accommodate the quantization distortion occurring during the measurement of the continuous signals by digital devices, we consider the following uniform-quantization measurement model:

$$\mathbf{y}_q = Q_\alpha[\Phi \mathbf{x}] = \Phi \mathbf{x} + \mathbf{n}, \quad (1)$$

where $(Q_\alpha[\cdot])_i = \alpha \lfloor \frac{(\cdot)_i}{\alpha} \rfloor + \frac{\alpha}{2}$, is a uniform quantization operator in \mathbb{R}^m with bin width α , \mathbf{y}_q is the quantized measurement vector and \mathbf{n} is a noise vector due to quantization distortion.

Given the quantized measurements \mathbf{y}_q and the measurement matrix Φ indexed with a sampling pattern Ω , the goal of a decoder g is to find an estimate $\hat{\mathbf{x}}$ which is written as follows:

$$\hat{\mathbf{x}} = g(\Phi, \mathbf{y}_q). \quad (2)$$

With a slight abuse of notation, we write $g(\Phi, \mathbf{y}_q)$ as $g(\Omega, \mathbf{y}_q)$ in the follows assuming that the orthonormal basis Ψ is fixed throughout.

A variety of decoding techniques have been proposed for quantized CS. In what follows, we first review BPDN method, which models the quantization distortion as a Gaussian noise. We then continue with reviewing the BPDQ_p decoder which is one of the best-performing techniques for quantized CS. With a suitable data fidelity constraint, BPDQ_p decoders are able to model the quantization distortion more faithfully.

A. Basis pursuit denoise

Viewing the quantization noise \mathbf{n} as an l_2 -bounded noise, i.e., $\|\mathbf{n}\|_2 \leq \epsilon$ for some $\epsilon > 0$, then stable and robust recovery

of the signal \mathbf{x} from the noisy measurements $\mathbf{y}_q = \Phi \mathbf{x} + \mathbf{n}$ is possible using the BPDN:

$$\Delta(\mathbf{y}_q, \epsilon) = \arg \min_{\mathbf{u} \in \mathbb{R}^p} \|\mathbf{u}\|_1 \quad \text{s.t.} \quad \|\mathbf{y}_q - \Phi \mathbf{u}\|_2 \leq \epsilon. \quad (3)$$

Note that (3) is a convex quadratic problem, and thus it can be solved by many general solvers, such as interior point methods. For large-scale problems, many specialized methods that are faster than interior point methods have been proposed [25, 26]. In the noiseless case $\epsilon = 0$, problem (3) is called Basis Pursuit program and is addressed in [27].

B. Basis pursuit dequantizer

BPDQ_p program returns a sparse estimation of \mathbf{x} from the quantized measurements $\mathbf{y}_q = \Phi \mathbf{x} + \mathbf{n}$ given that the noise \mathbf{n} is bounded in l_p -norm, i.e., $\|\mathbf{n}\|_p^p = \sum_k |\mathbf{n}_k|^p \leq \epsilon^p$ for some $\epsilon > 0$ and $2 < p \leq \infty$. The program is defined as follows:

$$\Delta_p(\mathbf{y}_q, \epsilon) = \arg \min_{\mathbf{u} \in \mathbb{R}^N} \|\mathbf{u}\|_1 \quad \text{s.t.} \quad \|\mathbf{y}_q - \Phi \mathbf{u}\|_p \leq \epsilon. \quad (4)$$

The fidelity constraint expressed in the l_p -norm is now tuned to noises that follow a zero-mean generalized Gaussian distribution of shape parameter p , with the uniform noise case corresponding to $p \rightarrow \infty$. Thus, BPDQ_p with a suitable p outperforms classic BPDN in quantized CS empirically [15]. In fact, for measurement matrix Φ with i.i.d. standard Gaussian entries and $m \gtrsim (K \log(N/K))^{\frac{2}{p}}$ and assuming that the quantization distortion is uniformly distributed in each quantization bin, BPDN_p can stably recover a K -sparse signal with high probability [15]:

$$\|\hat{\mathbf{x}} - \mathbf{x}\|_2 \lesssim \frac{\alpha}{\sqrt{p+1}}.$$

This further justifies that BPDN_p can outperform BPDN theoretically in some special cases. As noted in [15], the objective and the constraint functions in BPDN_p are non-smooth and one can use monotone operator splitting proximal methods [26] to solve the program. More precisely, they are practically solved using the Douglas-Rachford splitting method which can be written as follows:

$$\mathbf{u}^{(t+1)} = (1 - \frac{\alpha_t}{2})\mathbf{u}^{(t)} + \frac{\alpha_t}{2}(2S_\gamma - I) \circ (2\mathcal{P}_{T_p(\epsilon)} - I)(\mathbf{u}^{(t)}), \quad (5)$$

where $\alpha_t \in (0, 2)$, $\forall t \in \mathbb{N}$, S_γ is the soft-thresholding operator applied component-wise with threshold $\gamma > 0$, I is the identity operator, and $\mathcal{P}_{T_p(\epsilon)}$ is the orthogonal projection operator on the set $T_p(\epsilon)$ defined as: $T_p(\epsilon) = \{\mathbf{u} \in \mathbb{R}^N : \|\mathbf{y}_q - \Phi \mathbf{u}\|_p \leq \epsilon\}$.

III. LEARNING-BASED FRAMEWORK

In this section, we discuss how to learn the optimal sampling pattern $\hat{\Omega}$ using a greedy method which optimizes the performance of a BPDQ_p decoder on the target signal \mathbf{x} . Our approach is motivated by the learning-based framework [20], which is outlined as follows.

- Let $\mathbf{x}_1, \dots, \mathbf{x}_M$ be a set of training signals and \mathbf{x} an unknown target signal with similar properties to the given training set.

- We assume that decoder, BPDQ_p, is given.
- Given a recovery performance measure η_{Ω} , e.g., SNR, the goal is to find a sampling mask $\hat{\Omega}$ with the maximum empirical average performance on the training signals:

$$\hat{\Omega} = \arg \max_{\Omega: |\Omega|=m} \frac{1}{M} \sum_{j=1}^M \eta_{\Omega}(\mathbf{x}_j) \quad (6)$$

The main idea is that since the training signals and the unknown target signal are similar, the obtained mask $\hat{\Omega}$ which performs the best on average over the training signals is expected to also perform well on the target signal \mathbf{x} .

- The problem (6) is generally computationally challenging. However, an approximate solution can be found using a parameter-free greedy method [20], which will be discussed in the next section.

A. Greedy method

In this section, we describe the details of the greedy method to determine the subsampling pattern. Our goal is to find a sampling pattern $\Omega \subseteq \mathbf{E}$ with $|\Omega| = m$ and $m < N$, where $\mathbf{E} = \{1, \dots, N\}$ is the index set of all rows of the measurement matrix. Algorithm 1 describes a parameter-free method to greedily select the sampling pattern. The algorithm starts with an empty set Ω , and at each iteration, adds one element to Ω . At each step, the greedy method considers all the remaining candidates e in \mathbf{E} , which are not yet included in the mask, and adds the element $e^* \in \mathbf{E}$ resulting in the maximum average performance η over the training signals. The method stops when m elements are included in the mask.

Algorithm 1: Greedy optimization of the mask

Input : Training data $\mathbf{x}_1, \dots, \mathbf{x}_M$, reconstruction rule g , sampling subsets \mathbf{E} , number of mask elements m

Output: Sampling pattern Ω satisfying $|\Omega| = m$

```

1  $\Omega \leftarrow \emptyset$ ;
2 while  $|\Omega| \leq m$  do
3   for  $e \in \mathbf{E}$  do
4      $\Omega' = \Omega \cup e$ ;
5     for  $j = 1, \dots, M$  do
6       set  $\mathbf{y}_{q_j} \leftarrow Q_{\alpha}[\mathbf{P}_{\Omega'} \Psi \mathbf{x}_j]$ ,  $\hat{\mathbf{x}}_j \leftarrow g(\Omega', \mathbf{y}_{q_j})$ 
7     end
8      $\eta(\Omega') \leftarrow \frac{1}{M} \sum_{j=1}^M \eta(\mathbf{x}_j, \hat{\mathbf{x}}_j)$ 
9   end
10  compute  $e^* = \arg \max_e \eta(\Omega \cup e) - \eta(\Omega)$ 
11  and set  $\Omega \leftarrow \Omega \cup e^*$ 
12 end

```

In this paper, we did not try to discuss the statistical guarantees of the above algorithm, but one should keep in mind that the theoretical results developed in [20] for learning based approach in classical CS setting should be also true in our setting. We left this to the interested readers.

B. Learning-based method with parametric approach

It is alternatively possible to apply the learning-based method to parametric variable-density approaches. In these approaches, a set of candidate masks $\Omega_1, \dots, \Omega_D$ by trying a variety of parameters are generated. The method selects the mask with maximum empirical performance on the training dataset. The details of this approach are shown in Algorithm 2.

Algorithm 2: Learning based idea for a set of parametric masks [20]

Input : Candidate masks $\Omega_1, \dots, \Omega_D$, training data $\mathbf{x}_1, \dots, \mathbf{x}_M$, reconstruction rule g , performance measure η

Output: Sampling pattern Ω

```

1 for  $i = 1, \dots, D$  do
2   for  $j = 1, \dots, M$  do
3     set  $\mathbf{y}_{q_j} \leftarrow Q_{\alpha}[\mathbf{P}_{\Omega_i} \Psi \mathbf{x}_j]$ ,  $\hat{\mathbf{x}}_j \leftarrow g(\Omega_i, \mathbf{y}_{q_j})$ 
4   end
5    $\eta(\Omega_i) \leftarrow \frac{1}{M} \sum_{j=1}^M \eta(\mathbf{x}_j, \hat{\mathbf{x}}_j)$ 
6 end
7 compute  $i^* = \arg \max_{i=1, \dots, D} \eta_i$  and set  $\Omega \leftarrow \Omega_{i^*}$ 

```

IV. EXPERIMENTS

In this section, we provide numerical experiments demonstrating the performance of the learning based framework on quantized CS and compare it with state-of-the-art sampling methods. We compared the performance of the greedy sampling method described in Algorithm 1 with the following baselines:

- **Parametric**: We considered parametric randomized variable density method [4]. This method has two parameters 1) the size or radius of a fully-sampled region at low frequencies (r); and 2) the polynomial rate of decay of sampling at higher frequencies (d). We find the optimal mask using learning-based framework described in algorithm 2. We used the implementation available online [4].
- **Single-image**: We considered the approach of [19]. This method considers a single training image, and generate a probability density function, in which probability is proportional to the energy of the signal. Then, mask indices are randomly selected from this probability distribution by random under-sampling of k-space according to its energy distribution.
- **Single-signal**: We applied the same method for generating single-image masks on 1-D signals.
- **Coherence-based**: We considered variable density method proposed in [4]. We used no training data, and optimize an incoherence criterion as suggested in [4]. The minimization is done using Monte Carlo method.

We considered Fourier operator as the orthonormal basis Ψ of the measurement matrix $\Phi = \mathbf{P}_{\Omega} \Psi$ through all the experiments.

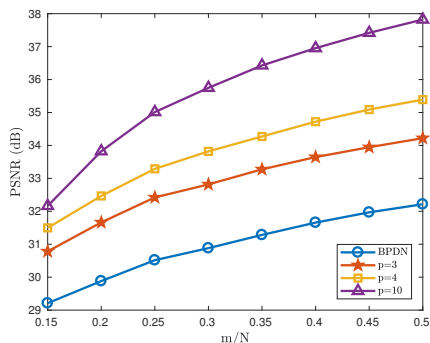


Fig. 1. Reconstruction performance using $BPDQ_p$ and greedy optimized sampling pattern for 1-D sparse random signals for different p values.

TABLE I
THE OBTAINED AVERAGED RECONSTRUCTION SNRS FOR DIFFERENT SAMPLING METHODS USING $BPDQ_p$ DECODER.

| $\frac{m}{N}$ \ Method | greedy | parametric | single-image | coherence-based |
|------------------------|--------------|------------|--------------|-----------------|
| 0.15 | 34.03 | 32.61 | 29.11 | 31.28 |
| 0.2 | 35.67 | 33.94 | 30.54 | 33.87 |
| 0.25 | 36.91 | 35.11 | 31.90 | 34.21 |
| 0.3 | 37.98 | 36.18 | 33.65 | 34.41 |
| 0.35 | 38.95 | 37.03 | 35.50 | 35.27 |
| 0.4 | 39.86 | 38.15 | 35.60 | 34.72 |
| 0.45 | 40.72 | 39.27 | 37.07 | 37.64 |
| 0.5 | 41.58 | 40.11 | 38.14 | 38.94 |

A. Synthetic 1-D signals

In the first experiment, we considered 1-D exactly sparse random signals with dimension $N = 512$ and the sparsity level of $K = 8$. We generated 500 signals. The nonzero elements of the signals are generated from the standard Gaussian distribution $\mathcal{N}(0, 1)$, positioned in the support drawn uniformly in $\{1, \dots, \lfloor 0.2N \rfloor\}$, where $\lfloor \cdot \rfloor$ is the floor function. We set the bin width to $\alpha = \frac{\|\Phi \mathbf{x}\|_\infty}{20}$ similar to [15]. The dataset is split into 30 signals as the training signals and 470 signals as the test signals. The quantized measurements for each sparse signal were recorded using the $\Phi = P_\Omega \Psi$ measurement matrix using the equation (1).

We recovered the signals using $BPDQ_p$ decoder for various $p \geq 2$ values. Figure 1 demonstrates the averaged reconstruction performance for $BPDQ_p$ decoder for various values of $p \geq 2$, and $m/N \in [0.15, 0.5]$. We measured the reconstruction performance in terms of SNR. The SNR for each signal is computed as $\text{SNR} = 20 \log_{10} \frac{\|\mathbf{x}\|}{\|\mathbf{x} - \hat{\mathbf{x}}\|}$, where \mathbf{x} is the real signal, and $\hat{\mathbf{x}}$ is the obtained reconstructed signal. As shown in figure 1, the $BPDQ_p$ optimized with greedy sampling pattern for higher oversampling factors (m/N), obtains higher SNR values.

B. Kenya dataset

We created an image dataset in 16-bit `tiff` format of 182 Kenya images of resolution of 512×512 previously collected in [21]. We split the dataset into a training set of the first 30 images, and a test set with the remaining 152 images. For $BPDQ_p$ decoder, we set $p = 10$, and $\alpha = 50$. Example reconstructed images are provided in supplementary material.

TABLE II
THE OBTAINED AVERAGED RECONSTRUCTION SNRS FOR DIFFERENT SAMPLING METHODS USING $BPDQ_p$ DECODER.

| $\frac{m}{N}$ \ Method | greedy | parametric | single-signal |
|------------------------|--------------|------------|---------------|
| 0.15 | 29.17 | 28.24 | 24.51 |
| 0.2 | 30.25 | 29.19 | 25.92 |
| 0.25 | 31.12 | 30.06 | 28.11 |
| 0.3 | 31.84 | 30.79 | 28.94 |
| 0.35 | 32.27 | 31.24 | 29.78 |
| 0.4 | 32.63 | 31.78 | 29.17 |
| 0.45 | 32.98 | 32.29 | 29.95 |
| 0.5 | 33.28 | 32.66 | 30.75 |

Table I shows the obtained SNR values for different sampling patterns. As shown, greedy method substantially improves the performance of recovery of signals.

C. iEEG dataset

This experiment is carried out on the *1001-P034-D01* dataset from *iEEG.org* portal. This data consists of approximately 1 day, 8 hours, and 10 minutes of recording at 5kHz, which is approximately 6×10^8 samples. We used the first 6×10^4 samples, and we extracted signals of channels 1-31. Then, we split the samples to 30 signals for training, and the rest for testing. The training data is used to learn the sampling pattern using greedy method and choosing the baseline masks obtaining the maximum performance.

Let each sequence of the signals be of dimension $N = 512$. For parametric approach we experimented with $r = \{\frac{4}{2N}, \frac{5}{2N}, \dots, \frac{16}{2N}\}$, and $d = \{1, 2, \dots, 75\}$. We generated 20 random draws for each set of parameters, and choose the one which performs the best over the training signals. For $BPDQ_p$ decoder, we set $\alpha = 50$, and $p = 10$.

Table II shows the obtained SNRs for varying sampling rates. Learning based methods (greedy and parametric) obtain superior performance compared to the baseline method. Especially, the greedy method achieves the highest reconstruction performance.

V. CONCLUSION

In this paper, we have developed a learning-based quantized CS method for recovery of sparse signals. We showed that our framework efficiently optimizes the sampling pattern by using the training signals. We experimentally demonstrated that our approach is effective in real-world datasets, and can improve the performance of quantized CS methods, and obtains superior performance compared to the previous state-of-the-art sampling patterns.

ACKNOWLEDGMENT

The research was supported by Hasler Foundation Program: Cyber Human Systems (project number 16066), and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement number 725594 - time-data). It was also sponsored by the Department of the Navy, Office of Naval Research (ONR) under a grant number N62909-17-1-2111.

REFERENCES

- [1] David L Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.
- [2] Emmanuel J Candes and Terence Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE Transactions on Information Theory*, 52(12):5406–5425, 2006.
- [3] Emmanuel J Candes, Justin K Romberg, and Terence Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59(8):1207–1223, 2006.
- [4] Michael Lustig, David Donoho, and John M Pauly. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*, 58(6):1182–1195, 2007.
- [5] Michael Lustig, David L Donoho, Juan M Santos, and John M Pauly. Compressed sensing MRI. *IEEE Signal Processing Magazine*, 25(2):72–82, 2008.
- [6] Jacob N Sanders, Semion K Saikin, Sarah Mostame, Xavier Andrade, Julia R Widom, Andrew H Marcus, and Alán Aspuru-Guzik. Compressed sensing for multidimensional spectroscopy experiments. *The Journal of Physical Chemistry Letters*, 3(18):2697–2702, 2012.
- [7] Joachim HG Ender. On compressive sensing applied to radar. *Signal Processing*, 90(5):1402–1414, 2010.
- [8] David J Brady, Kerkil Choi, Daniel L Marks, Ryoichi Horisaki, and Sehoon Lim. Compressive holography. *Optics Express*, 17(15):13040–13049, 2009.
- [9] Rebecca M Willett, Roummel F Marcia, and Jonathan M Nichols. Compressed sensing for practical optical imaging systems: a tutorial. *Optical Engineering*, 50(7):072601, 2011.
- [10] Hengyong Yu and Ge Wang. Compressed sensing based interior tomography. *Physics in Medicine & Biology*, 54(9):2791, 2009.
- [11] David Gross, Yi-Kai Liu, Steven T Flammia, Stephen Becker, and Jens Eisert. Quantum state tomography via compressed sensing. *Physical Review Letters*, 105(15):150401, 2010.
- [12] Mark Rudelson and Roman Vershynin. On sparse reconstruction from Fourier and Gaussian measurements. *Communications on Pure and Applied Mathematics*, 61(8):1025–1045, 2008.
- [13] Richard Baraniuk, Mark Davenport, Ronald DeVore, and Michael Wakin. A simple proof of the restricted isometry property for random matrices. *Constructive Approximation*, 28(3):253–263, 2008.
- [14] Hao-Jun Michael Shi, Mindy Case, Xiaoyi Gu, Shenyinying Tu, and Deanna Needell. Methods for quantized compressed sensing. In *Information Theory and Applications Workshop*, pages 1–9. IEEE, 2016.
- [15] Laurent Jacques, David K Hammond, and Mohamed-Jalal Fadili. Dequantizing compressed sensing: When oversampling and non-gaussian constraints combine. *IEEE Transactions on Information Theory*, 57(1):559–571, 2011.
- [16] Argyrios Zymnis, Stephen Boyd, and Emmanuel Candes. Compressed sensing with quantized measurements. *IEEE Signal Processing Letters*, 17(2):149–152, 2010.
- [17] Ivana Tomic and Pascal Frossard. Dictionary learning. *IEEE Signal Processing Magazine*, 28(2):27–38, 2011.
- [18] Saiprasad Ravishanker and Yoram Bresler. MR image reconstruction from highly undersampled k-space data by dictionary learning. *IEEE Transactions on Medical Imaging*, 30(5):1028, 2011.
- [19] Jaganathan Vellagoundar and Ramasubba Reddy Machireddy. A robust adaptive sampling method for faster acquisition of MR images. *Magnetic Resonance Imaging*, 33(5):635–643, 2015.
- [20] Baran Gözcü, Rabeeh Karimi Mahabadi, Yen-Huan Li, Efe Ilıcak, Tolga Cukur, Jonathan Scarlett, and Volkan Cevher. Learning-based compressive MRI. *IEEE Transactions on Medical Imaging*, 37(6):1394–1406, 2018.
- [21] Luca Baldassarre, Yen-Huan Li, Jonathan Scarlett, Baran Gözcü, Ilija Bogunovic, and Volkan Cevher. Learning-based compressive subsampling. *IEEE Journal of Selected Topics in Signal Processing*, 10(4):809–822, 2016.
- [22] Biao Sun and Hui Feng. Efficient compressed sensing for wireless neural recording: A deep learning approach. *IEEE Signal Processing Letters*, 24:863–867, 2017.
- [23] Biao Sun, Hui Feng, Kefan Chen, and Xinshan Zhu. A deep learning framework of quantized compressed sensing for wireless neural recording. *IEEE Access*, 4:5169–5178, 2016.
- [24] Emmanuel J Candes and Justin Romberg. Quantitative robust uncertainty principles and optimally sparse decompositions. *Foundations of Computational Mathematics*, 6(2):227–254, 2006.
- [25] Mohamed-Jalal Fadili and Jean-Luc Starck. Monotone operator splitting for fast sparse solutions of inverse problems. *SIAM Journal on Imaging Sciences*, pages 2005–2006, 2009.
- [26] Patrick L Combettes. Solving monotone inclusions via compositions of nonexpansive averaged operators. *Optimization*, 53(5-6):475–504, 2004.
- [27] Scott Shaobing Chen, David L Donoho, and Michael A Saunders. Atomic decomposition by basis pursuit. *SIAM Journal of Scientific Computing*, 43(1):129–159, 2001.