

# Practical Compression Methods for Quantized Compressed Sensing

Markus Leinonen\*, Marian Codreanu<sup>†</sup>, and Markku Juntti\*

\*Centre for Wireless Communications – Radio Technologies, FI-90014, University of Oulu, Finland

<sup>†</sup>Department of Science and Technology, Linköping University, Sweden

e-mail: markus.leinonen@oulu.fi, marian.codreanu@liu.se, markku.juntti@oulu.fi

**Abstract**—In order to save energy of low-power sensors in Internet of Things applications, minimizing the number of bits to compress and communicate real-valued sources with a pre-defined distortion becomes crucial. In such a lossy source coding context, we study rate-distortion (RD) performance of various single-sensor quantized compressed sensing (QCS) schemes for compressing sparse signals via quantized/encoded noisy linear measurements. The paper combines and refines the recent advances of QCS algorithm designs and theoretical analysis. In particular, several practical symbol-by-symbol quantizer based QCS methods of different complexities relying on 1) compress-and-estimate, 2) estimate-and-compress, and 3) support-estimation-and-compress strategies are proposed. Simulation results demonstrate the RD performances of different schemes and compare them to the information-theoretic limits.

## I. INTRODUCTION

In the near future, there will be a burgeoning demand for the deployment of smart low-power sensors to serve the myriad of Internet of Things (IoT) applications in, e.g., environmental, industrial, healthcare, and military monitoring tasks [1–3]. By 2020, the number of IoT devices is envisioned to reach hundreds of billions with market value being trillions of dollars [2, 3]. Key enablers of diverse IoT applications are wireless sensor networks (WSNs), where sensors are deployed to collect intelligent data of an underlying application to, e.g., improve system performance and reduce maintenance costs.

In a typical WSN, geographically distributed, battery-powered sensors separately observe and encode one or multiple information sources to be communicated to the sink for joint signal reconstruction. Limited batteries – often non-rechargeable or irreplaceable – dictate the lifespan of a WSN. The main contributors to a sensor’s energy consumption are wireless communications [4], and in some setups, the sensing/sampling part [5]. Consequently, a well-designed sensor acquires only a few data samples of a physical phenomenon (e.g., temperature, humidity, or light), which, after compressed at a minimal bit rate, are sufficient to reconstruct the signal with a pre-defined fidelity at the sink.

Compressed sensing (CS) [6, 7] is an effective joint sampling and compression paradigm to conserve sensors’ energy in gathering sparse signals in, e.g., environmental monitoring [8, 9], source localization [10], and biomedical sensing [11]. While the early CS works dealt with analog signals, the practical necessity of converting the measurements into bit sequences initiated quantized CS (QCS) [12]. Existing QCS

works can be classified as a) quantization-aware algorithm designs, b) performance analysis under symbol-by-symbol quantizers, and c) information-theoretic analysis, which, due to indirect observations, falls into remote source coding [13].

This paper deals with classes a) and c) above by addressing the rate-distortion (RD) performance of various single-sensor QCS setups, where a sparse source is communicated to the sink by quantized/encoded noisy compressed measurements. The main contribution of the paper is to combine and refine several recent advances in QCS algorithm designs and theoretical analysis into a unified framework with empirical performance comparison. In particular, several practical QCS methods of different complexities relying on 1) compress-and-estimate, 2) estimate-and-compress, and 3) support-estimation-and-compress strategies are proposed. In particular, the strategy 3) contains several novel methods which are empirically shown to achieve competent RD performance with low complexity, beneficial to energy-limited IoT sensor applications. Simulation results demonstrate the RD performances of different practical QCS schemes and compare them to the compression limit – the remote rate-distortion function – which is evaluated through an analytical lower bound and a modified Blahut-Arimoto algorithm proposed in [14].

The paper is organized as follows. Section II defines the system model. Section III presents quantization preliminaries which are needed in developing the different practical QCS methods detailed in Section IV. Section V addresses the theoretical limits of the QCS setup. Simulation results are provided in Section VI, and Section VII concludes the paper.

## II. QUANTIZED COMPRESSED SENSING FRAMEWORK

We study the RD performance of a QCS framework depicted in Fig. 1. The encoder E observes the information source  $\mathbf{X}$  via noisy linear measurements  $\mathbf{Y}$ , followed by finite-rate quantization/encoding and error-free communication to the decoder D for signal reconstruction.

Let  $\{\mathbf{X}_n\}_{n=1}^{\infty}$  be a discrete-time memoryless vector source sequence. Each vector  $\mathbf{X}_n = [X_{n,1} \cdots X_{n,N}]^T$  is  $K$ -sparse, i.e.,  $\|\mathbf{x}\|_0 = K \leq N$ . Each  $\mathbf{X}_n$  is generated from the memoryless sequence of tuples  $\{(\mathbf{G}_n, \mathbf{B}_n)\}_{n=1}^{\infty}$  as  $\mathbf{X}_n = \mathbf{G}_n \odot \mathbf{B}_n$ , where  $\odot$  denotes the Hadamard product;  $\mathbf{G}_n$  is a length- $N$  zero mean Gaussian random vector  $\mathbf{G}_n \sim \mathcal{N}(\mathbf{0}, \Sigma_{\mathbf{G}})$  with positive definite covariance matrix  $\Sigma_{\mathbf{G}} \in \mathbb{S}_{++}^N$ , where  $\mathbb{S}_{++}^N$  is the set of symmetric positive definite  $N \times N$  matrices;  $\mathbf{B}_n$

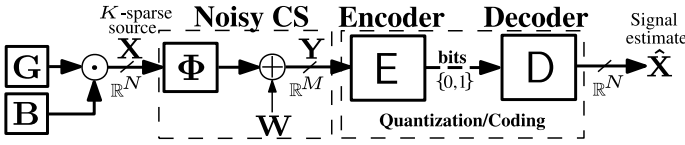


Fig. 1. Acquisition of a sparse source via quantized/encoded noisy compressed measurements.

is a length- $N$  binary support random vector, independent of  $\mathbf{G}_n$ , with a discrete alphabet  $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_{|\mathcal{B}|}\}$ ,  $|\mathcal{B}| = \binom{N}{K}$ . Each sparsity pattern  $\mathbf{b}_s = [b_{s,1} \dots b_{s,N}]^T \in \mathcal{B}$  has  $K$  ones and  $N - K$  zeros, and is associated with the *a priori* probability  $p(\mathbf{b}_s) \in [0, 1]$  so that  $\sum_{s=1}^{|\mathcal{B}|} p(\mathbf{b}_s) = 1$ .

The sensor (i.e., the encoder) observes  $\{\mathbf{X}_n\}_{n=1}^{\infty}$  indirectly [6, 7] as

$$\mathbf{Y}_n = \Phi \mathbf{X}_n + \mathbf{W}_n, \quad n = 1, 2, \dots \quad (1)$$

where  $\Phi \in \mathbb{R}^{M \times N}$  is a fixed and known measurement matrix performing dimensionality reduction,  $K \leq M \leq N$ ,  $\mathbf{W}_n \sim \mathcal{N}(\mathbf{0}, \Sigma_{\mathbf{W}})$  with  $\Sigma_{\mathbf{W}} \in \mathbb{S}_{++}^M$  is a length- $M$  noise random vector independent of  $\mathbf{X}_n$ , and each  $\mathbf{Y}_n$  is a length- $M$  measurement random vector. Note that while there are no restrictions on  $\Phi$ , it plays a key role in the CS signal recovery performance, which can be assessed by the restricted isometry property (RIP) and coherence of  $\Phi$  [15, Ch. 5, 6].

The encoder-decoder pair E-D can take various implementations, each having different encoding complexity, memory requirements, compression performance etc. Accordingly, we study the achievable signal reconstruction performance 1) by different *practical* QCS methods relying on symbol-by-symbol quantization having moderate encoding complexity, and 2) in an *information-theoretic* context using encoder/decoders of excessively high complexity. The end-to-end mean square error (MSE) is used as the distortion criterion, i.e.,

$$D = \mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}\|_2^2], \quad (2)$$

where  $\hat{\mathbf{X}}$  represents a (method-dependent) length- $N$  reproduction random vector at the output of decoder D.

### III. QUANTIZATION PRELIMINARIES

This section introduces three general-purpose quantizers that are needed in designing the QCS methods in Section IV.

Let  $\mathbf{U}$  be a length- $L$  random vector with PDF  $f(\mathbf{u})$ . The vector  $\mathbf{U}$  is the input of a quantizer Q determined by 1) encoder regions  $\mathcal{S}_i$ ,  $i \in \mathcal{I} = \{1, \dots, |\mathcal{I}|\}$ , which partition its input as  $\mathcal{S}_i \subseteq \mathbb{R}^L$ ,  $\mathcal{S}_i \cap \mathcal{S}_{i'} = \emptyset$ , for any  $i \neq i'$ , and  $\bigcup_{i=1}^{|\mathcal{I}|} \mathcal{S}_i = \mathbb{R}^L$ , and 2) a reconstruction codebook  $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_{|\mathcal{I}|}\}$  with codevectors  $\mathbf{c}_i \in \mathbb{R}^L$ . For a realization  $\mathbf{u}$ , quantizer encoder  $Q: \mathbb{R}^L \rightarrow \mathcal{I}$  produces an index  $i = Q(\mathbf{u})$  if  $\mathbf{u} \in \mathcal{S}_i$ . For a received index  $i$ , quantizer decoder  $Q^{-1}: \mathcal{I} \rightarrow \mathcal{C}$  produces an estimate of  $\mathbf{u}$  as  $\hat{\mathbf{u}} = Q^{-1}(i) = \mathbf{c}_i \in \mathcal{C}$ .

1) *A fixed-rate Lloyd-Max quantizer* [16, 17]  $Q^{\text{lm}}$  of rate  $R'$  bits/U optimizes the  $|\mathcal{I}| = 2^{R'}$  quantization regions and the codebook to minimize the MSE distortion

$$D' = \sum_{i=1}^{2^{R'}} p(i) \mathbb{E}[\|\mathbf{U} - \mathbf{c}_i^{\text{lm}}\|_2^2 | I = i], \quad (3)$$

---

#### Algorithm 1 Lloyd-Max quantizer training (offline).

---

**Input:** a) Sequence  $\{\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \dots\}$  sampled from  $f(\mathbf{u})$ ; b) quantization rate  $R'$ .

**Initialization:** Reconstruction codebook  $\mathcal{C}^{\text{lm}}$ .

**Repeat until convergence**

1) For a given codebook, find the optimal regions by classifying  $\{\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \dots\}$  as

$$\mathcal{S}_i^{\text{lm}} = \left\{ \mathbf{u} : \|\mathbf{u} - \mathbf{c}_i^{\text{lm}}\|_2^2 \leq \|\mathbf{u} - \mathbf{c}_{i'}^{\text{lm}}\|_2^2, \forall i' \neq i \right\}, \quad i \in \mathcal{I}. \quad (4)$$

2) For given regions, find the optimal codevectors as the conditional expectations

$$\mathbf{c}_i^{\text{lm}} = \mathbb{E}[\mathbf{U} | I = i] = \frac{1}{p(i)} \int_{\mathbf{u} \in \mathcal{S}_i^{\text{lm}}} \mathbf{u} f(\mathbf{u}) d\mathbf{u}, \quad i \in \mathcal{I}. \quad (5)$$

**Output:** Reconstruction codebook  $\mathcal{C}^{\text{lm}} = \{\mathbf{c}_1^{\text{lm}}, \dots, \mathbf{c}_{2^{R'}}^{\text{lm}}\}$ .

---



---

#### Algorithm 2 Fixed-rate quantization of an input $\mathbf{u}$ (online).

---

**Input:** Reconstruction codebook  $\mathcal{C}^{\text{fix}}$  ( $\mathcal{C}^{\text{lm}}$  or  $\mathcal{C}^{\text{u}}$ ).

**Encoding:** 1) Find the encoding index  $i^* = \operatorname{argmin}_{i \in \mathcal{I}} \|\mathbf{u} - \mathbf{c}_i^{\text{fix}}\|_2^2$ ; 2) communicate the  $R'$ -bit representation of  $i^*$  to the decoder.

**Decoding:** Given the received index  $i^*$ , obtain an estimate of  $\mathbf{u}$  as  $\hat{\mathbf{u}} = \mathbf{c}_{i^*}^{\text{fix}} \in \mathcal{C}^{\text{fix}}$ .

---

where  $p(i) \triangleq \Pr(I = i)$  and  $I$  represents the quantization index. Following the iterative Lloyd and Linde-Buzo-Gray (LBG) algorithms [16–18],  $Q^{\text{lm}}$  can be trained via the alternating optimization presented in Algorithm 1. The online compression phase is presented in Algorithm 2.

2) *A variable-rate quantizer*  $Q^{\text{vr}}$  minimizes a cost function

$$(1 - \mu')D' - \mu' \sum_{i \in \mathcal{I}} p(i) \log(p(i)) \quad (6)$$

where  $D'$  is given in (3), and  $\mu' \in [0, 1]$  is an RD trade-off parameter. Following the principles of entropy-constrained scalar/vector quantization (ECSQ/ECVQ) [19–21],  $Q^{\text{vr}}$  can be trained by the three-step iterative loop presented in Algorithm 3. The online phase is presented in Algorithm 4. In Algorithm 3, the source codebook  $\mathcal{H} \triangleq \{\mathbf{h}_1, \dots, \mathbf{h}_{|\mathcal{H}|}\}$ , which, in general, contains *variable-length* binary codewords, can be generated by, e.g., the Huffman coding [22]. The average rate of  $Q^{\text{vr}}$  is  $\sum_{i=1}^{|\mathcal{H}|} p(i) l(\mathbf{h}_i)$  bits/U, where  $l(\cdot)$  is a function that computes the length of a binary codeword.

3) *A uniform scalar quantizer* (USQ)  $Q^{\text{u}}$  of rate  $R'$  bits/U consists of fixed-length intervals as presented in Algorithm 5. The online phase is equivalent to Algorithm 2.

### IV. QUANTIZED COMPRESSED SENSING ALGORITHMS

This section presents the design of three classes of practical QCS methods, as depicted in Fig. 2. The devised methods rely on the different quantization schemes described in Section III.

#### A. Compress-And-Estimate QCS Methods

A compress-and-estimate (C&E) scheme depicted in Fig. 2(a) consists of two stages: 1) a compression stage where

---

**Algorithm 3** Entropy-constrained quantizer training (offline).

**Input:** a) Sequence  $\{\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \dots\}$  sampled from  $f(\mathbf{u})$ ; b) quantization levels  $2^{R'}$ ; c) weight parameter  $\mu' \in [0, 1]$ .

**Initialization:** i) Reconstruction codebook  $\mathcal{C}^{\text{vr}}$ , and ii) index probabilities  $p(i) = 1/2^{R'}$ ,  $i \in \mathcal{I}$ .

**Repeat until convergence**

1) For a given codebook and rate measures, find the optimal regions by classifying  $\{\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \dots\}$  as

$$S_i^{\text{vr}} = \left\{ \mathbf{u} : (1 - \mu') \|\mathbf{u} - \mathbf{c}_i^{\text{vr}}\|_2^2 - \mu' \log(p(i)) \leq (1 - \mu') \|\mathbf{u} - \mathbf{c}_{i'}^{\text{vr}}\|_2^2 - \mu' \log(p(i')), \forall i' \neq i \right\}, i \in \mathcal{I}. \quad (7)$$

2) Update the rate measures  $-\log(p(i))$ ,  $i \in \mathcal{I}$ , given the new regions.

3) For given regions, find the optimal codevectors  $\mathbf{c}_i^{\text{vr}}$ ,  $i \in \mathcal{I}$ , equivalently as in (5).

**Output:** a) Reconstruction codebook  $\mathcal{C}^{\text{vr}} = \{\mathbf{c}_1^{\text{vr}}, \dots, \mathbf{c}_{2^{R'}}^{\text{vr}}\}$ ; b) index probabilities  $p(i)$ ,  $i \in \mathcal{I}$ ; c) source codebook  $\mathcal{H} = \{\mathbf{h}_1, \dots, \mathbf{h}_{|\mathcal{H}|}\}$  generated using  $p(i)$ ,  $i \in \mathcal{I}$ .

---



---

**Algorithm 4** Variable-rate quantization of an input  $\mathbf{u}$  (online).

**Input:** a) Reconstruction codebook  $\mathcal{C}^{\text{vr}}$ ; b) index probabilities  $p(i)$ ,  $i \in \mathcal{I}$ ; c) weight parameter  $\mu' \in [0, 1]$ ; d) source codebook  $\mathcal{H}$ .

**Encoding:** 1) Find the encoding index  $i^* = \operatorname{argmin}_{i \in \mathcal{I}} (1 - \mu') \|\mathbf{u} - \mathbf{c}_i^{\text{vr}}\|_2^2 - \mu' \log(p(i))$ ; 2) communicate the binary codeword  $\mathbf{h}_{i^*}$  to the decoder.

**Decoding:** Given the received codeword  $\mathbf{h}_{i^*}$ , find the corresponding index  $i^*$  and obtain an estimate of  $\mathbf{u}$  as  $\hat{\mathbf{u}} = \mathbf{c}_{i^*}^{\text{vr}} \in \mathcal{C}^{\text{vr}}$ .

---



---

**Algorithm 5** Uniform scalar quantizer (offline).

**Input:** a) Quantization rate  $R'$ ; b) input range parameters  $q^{\text{hi}} \geq q^{\text{lo}}$ .

**Output:** Reconstruction codebook  $\mathcal{C}^{\text{u}} = \{c_1^{\text{u}}, \dots, c_{2^{R'}}^{\text{u}}\}$  with codepoints  $c_1^{\text{u}} = q^{\text{lo}}$ ;  $c_{2^{R'}}^{\text{u}} = q^{\text{hi}}$ ; and  $c_i^{\text{u}} = c_{i-1}^{\text{u}} + (q^{\text{hi}} - q^{\text{lo}})/(2^{R'} - 1)$ , for all  $2 \leq i \leq 2^{R'} - 1$ .

---

the encoder quantizes  $\mathbf{Y}$  under an MSE distortion criterion that depends only on  $\mathbf{Y}$  (not on  $\mathbf{X}$ ), and 2) an estimation stage where the decoder estimates  $\mathbf{X}$  from the decoded quantized measurements  $\hat{\mathbf{Y}} = \mathbf{Q}^{-1}(\mathbf{Q}(\mathbf{Y}))$ . Let  $\hat{\mathbf{X}}^{\text{C\&E}}$  denote a length- $N$  reproduction random vector at the decoder output for C&E. The end-to-end MSE distortion can be written as

$$D^{\text{C\&E}} = \mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}^{\text{C\&E}}\|_2^2] = \mathbb{E}[\|\mathbf{X} - \mathbf{R}_D[\mathbf{Q}^{-1}(\mathbf{Q}(\mathbf{Y}))]\|_2^2], \quad (8)$$

where  $\mathbf{R}_D : \hat{\mathbf{Y}} \rightarrow \hat{\mathbf{X}}^{\text{C\&E}}$  is a signal reconstruction algorithm at D. A C&E principle underlies many early QCS algorithms in, e.g., [23, 24]. An information-theoretic study of C&E based QCS can be found in [25].

We consider four C&E algorithms that use the following quantizers: 1) USQ, 2) Lloyd-Max SQ (referred to as "SQ"

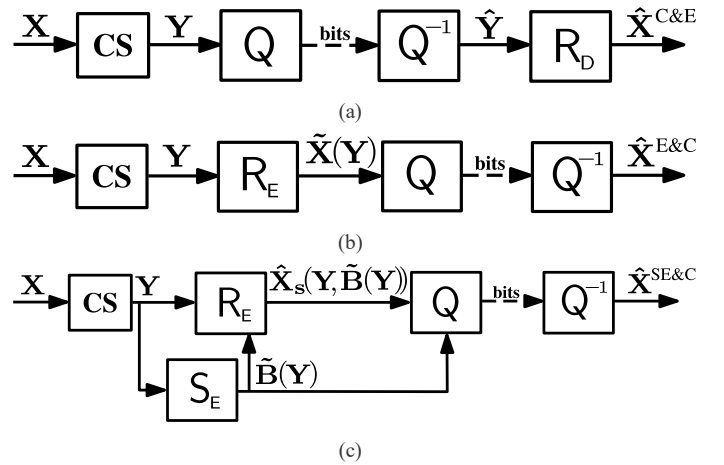


Fig. 2. QCS setups relying on (a) compress-and-estimate, (b) estimate-and-compress, and (c) support-estimation-and-compress strategies.

in the QCS method names henceforth), 3) ECSQ, and 4) VQ. Table I summarizes the considered methods and details the main parameters for implementing the involved quantizer using the algorithms described in Section III. Note that the total average rate  $R_{\text{tot}} = NR$  bits/ $\mathbf{X}$  is reported in the last column, where the rate  $R$  is defined as the bits/entry of  $\mathbf{X}$ . For entropy-constrained methods,  $\gamma(\cdot)$  computes the average codeword length of a binary codebook  $\mathcal{H}$ .

Summarizing the basic operations of the SQ based methods 1)–3) above, the decoder receives  $M$  indices, forms  $\hat{\mathbf{y}} = [c_{1, i_1^*}^{\text{C\&E}} \dots c_{M, i_M^*}^{\text{C\&E}}]^T$ , and estimates the source as  $\hat{\mathbf{x}} = \mathbf{R}_D(\hat{\mathbf{y}})$ . For C&E<sub>vq</sub>, the MSE-optimal output vectors of  $\mathbf{R}_D(\cdot)$  that minimize the distortion in (8) are given as

$$\hat{\mathbf{x}}(\mathbf{c}_i^{\text{C\&E}_{vq}}) = \mathbb{E}[\mathbf{X} | I = i] = \frac{1}{p(i)} \int_{\mathbf{y} \in S_i^{\text{C\&E}_{vq}}} \hat{\mathbf{x}}(\mathbf{y}) f(\mathbf{y}) d\mathbf{y}, i \in \mathcal{I}, \quad (9)$$

where  $\hat{\mathbf{x}}(\mathbf{y}) \in \mathbb{R}^N$  is the minimum mean square error (MMSE) estimate of  $\mathbf{X}$  given a measurement realization  $\mathbf{Y} = \mathbf{y}$ , i.e.,

$$\hat{\mathbf{x}}(\mathbf{y}) \triangleq \mathbb{E}[\mathbf{X} | \mathbf{Y} = \mathbf{y}]. \quad (10)$$

Let  $\hat{\mathbf{X}}(\mathbf{Y})$  denote the corresponding MMSE estimator.

### B. Estimate-and-Compress QCS Methods

An estimate-and-compress (E&C) scheme depicted in Fig. 2(b) consists of two stages: 1) an estimation stage of  $\mathbf{X}$  from  $\mathbf{Y}$  at the encoder<sup>1</sup>, and 2) a compression stage to quantize the resulting signal estimate under an MSE distortion criterion. The end-to-end MSE distortion reads as

$$D^{\text{E\&C}} = \mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}^{\text{E\&C}}\|_2^2] = \mathbb{E}[\|\mathbf{X} - \mathbf{Q}^{-1}[\mathbf{Q}(\mathbf{R}_E(\mathbf{Y}))]\|_2^2], \quad (11)$$

where  $\hat{\mathbf{X}}^{\text{E\&C}}$  is the length- $N$  reproduction random vector at the decoder output,  $\mathbf{R}_E : \mathbf{Y} \rightarrow \tilde{\mathbf{X}}(\mathbf{Y})$  is a reconstruction algorithm at E, and  $\tilde{\mathbf{X}}(\mathbf{Y})$  denotes the estimator of  $\mathbf{X}$  from

<sup>1</sup>By modeling that E&C and SE&C schemes access real-valued  $\mathbf{Y}$ , we neglect the quantization error induced by the sensor's analog-to-digital converter assuming that it operates at a relatively high rate as compared to the actual (low-rate) source encoder/quantizer.

TABLE I  
DIFFERENT QCS METHODS AND THEIR QUANTIZATION PARAMETERS

Class	Method	Training/ Online algorithm	$\mathbf{U}$	$L$	$R_{\text{tot}} = NR$
Compress-and-Estimate	C&E <sub>usq</sub>	Algorithm 5 / 2	$Y_m$	1	$MR'$
	C&E <sub>sq</sub>	Algorithm 1 / 2	$Y_m$	1	$MR'$
	C&E <sub>ecsqsq</sub>	Algorithm 3 / 4	$Y_m$	1	$\sum_{m=1}^M \gamma(\mathcal{H}_m^{\text{C\&E}})$
	C&E <sub>vq</sub>	Algorithm 1 / 2	$\mathbf{Y}$	$M$	$R'$
Estimate-and-Compress	E&C <sub>usq</sub>	Algorithm 5 / 2	$[\tilde{\mathbf{X}}(\mathbf{Y})]_n$	1	$NR'$
	E&C <sub>sq</sub>	Algorithm 1 / 2	$[\tilde{\mathbf{X}}(\mathbf{Y})]_n$	1	$NR'$
	E&C <sub>ecsqsq</sub>	Algorithm 3 / 4	$[\tilde{\mathbf{X}}(\mathbf{Y})]_n$	1	$\sum_{n=1}^N \gamma(\mathcal{H}_n^{\text{E\&C}})$
	E&C <sub>vq</sub>	Algorithm 1 / 2	$\tilde{\mathbf{X}}(\mathbf{Y})$	$N$	$R'$
	E&C <sub>ecvq</sub>	Algorithm 3 / 4	$\tilde{\mathbf{X}}(\mathbf{Y})$	$N$	$\gamma(\mathcal{H}^{\text{E\&C}})$
Support-Estimation-and-Compress	SE&C <sub>usq</sub>	Algorithm 5 / 2	$[\tilde{\mathbf{X}}_s(\mathbf{Y}, \tilde{\mathbf{B}}(\mathbf{Y}))]_n$	1	$\gamma(\mathcal{H}_{\text{supp}}^{\text{SE\&C}}) + KR'$
	SE&C <sub>sq</sub>	Algorithm 1 / 2	$[\tilde{\mathbf{X}}_s(\mathbf{Y}, \tilde{\mathbf{B}}(\mathbf{Y}))]_n$	1	$\gamma(\mathcal{H}_{\text{supp}}^{\text{SE\&C}}) + KR'$
	SE&C <sub>ecsqsq</sub>	Algorithm 3 / 4	$[\tilde{\mathbf{X}}_s(\mathbf{Y}, \tilde{\mathbf{B}}(\mathbf{Y}))]_n$	1	$\gamma(\mathcal{H}_{\text{supp}}^{\text{SE\&C}}) + \sum_{k=1}^K \gamma(\mathcal{H}_k^{\text{SE\&C}})$
	SE&C <sub>vq</sub>	Algorithm 1 / 2	$\tilde{\mathbf{X}}_s(\mathbf{Y}, \tilde{\mathbf{B}}(\mathbf{Y}))$	$K$	$\gamma(\mathcal{H}_{\text{supp}}^{\text{SE\&C}}) + R'$

$\mathbf{Y}$ . The E&C scheme is the optimal encoding structure for remote source coding [26]. E&C based QCS algorithms have been devised in, e.g., [14, 21, 27–30]. The related information-theoretic studies include [14, 31, 32].

Five E&C algorithms with 1) USQ, 2) SQ, 3) ECSQ, 4) VQ, and 5) ECVQ are considered (see Table I). For SQ based methods 1) – 3), the decoder uses the  $N$  received indices to estimate the source as  $\hat{\mathbf{x}} = [c_{1,i_1}^{\text{E\&C}} \dots c_{N,i_N}^{\text{E\&C}}]^T$ . For VQ based methods 4) – 5), the decoder operates simply as  $\hat{\mathbf{x}} = \mathbf{c}_{i_s}^{\text{E\&C}}$ .

### C. Support-Estimation-and-Compress QCS Methods

A support-estimation-and-compress (SE&C) scheme depicted in Fig. 2(c) consists of three stages: 1) a support estimation stage where the encoder estimates  $\mathbf{B}$  from  $\mathbf{Y}$ , 2) an estimation stage of  $\mathbf{X}$  given  $\mathbf{Y}$  and the support estimator at E, and 3) a two-phase compression stage of the resulting source estimate applying i) lossless compression for the estimated support, and ii) lossy compression for its non-zero part. The end-to-end MSE distortion can be expressed as

$$D^{\text{SE\&C}} = \mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}^{\text{SE\&C}}\|_2^2] = \mathbb{E}[\|\mathbf{X} - \mathbf{Q}^{-1}[\mathbf{Q}(\mathbf{S}_E(\mathbf{Y}) + \mathbf{R}_E(\mathbf{Y}|\mathbf{S}_E(\mathbf{Y})))]\|_2^2], \quad (12)$$

where  $\hat{\mathbf{X}}^{\text{SE\&C}}$  is the length- $N$  reproduction random vector at the decoder output,  $\mathbf{S}_E: \mathbf{Y} \rightarrow \tilde{\mathbf{B}}(\mathbf{Y})$  is a support estimation algorithm taking values on alphabet  $\mathcal{B}$  at E, and  $\tilde{\mathbf{B}}(\mathbf{Y})$  denotes the estimator of  $\mathbf{B}$  from  $\mathbf{Y}$ . The sparsity  $K$  is assumed to be known by  $\mathbf{S}_E(\cdot)$ . QCS schemes akin to SE&C have been considered in, e.g., [12, 27].

Let  $\mathbf{X}_s$  be a length- $K$  random vector representing the non-zero part of  $\mathbf{X}$  for a given support  $\mathbf{B} = \mathbf{b}_s$ ,  $\mathbf{b}_s \in \mathcal{B}$ . The optimal  $\mathbf{R}_E(\cdot)$  in (12) is the one that minimizes the MSE distortion  $\mathbb{E}[\|\mathbf{X}_s - \mathbf{R}_E(\cdot)\|_2^2]$ . This is given as  $\mathbf{R}_E(\mathbf{Y}|\mathbf{S}_E(\mathbf{Y})) = \tilde{\mathbf{X}}_s(\mathbf{Y}, \tilde{\mathbf{B}}(\mathbf{Y}))$ , i.e., the MMSE estimator of  $\mathbf{X}_s$  given  $\mathbf{Y}$  and  $\tilde{\mathbf{B}}(\mathbf{Y})$ , with realizations

$$\hat{\mathbf{x}}_s(\mathbf{y}, \tilde{\mathbf{b}}_s) \triangleq \mathbb{E}[\mathbf{X}_s | \mathbf{Y} = \mathbf{y}, \tilde{\mathbf{B}}(\mathbf{Y}) = \tilde{\mathbf{b}}_s] = \Sigma_{\mathbf{X}_s} \Phi_s^T (\Phi_s \Sigma_{\mathbf{X}_s} \Phi_s^T + \Sigma_{\mathbf{W}})^{-1} \mathbf{y}, \quad (13)$$

where  $\Phi_s \in \mathbb{R}^{M \times K}$  extracts the  $K$  columns of  $\Phi$  with indices  $\{k | \tilde{b}_{s,k} = 1, k = 1, \dots, N\}$ .

Four SE&C algorithms with 1) USQ, 2) SQ, 3) ECSQ, and 4) VQ are considered (see Table I). Quantizers for the estimated non-zero parts operate equivalently as in the corresponding E&C schemes. The support estimator  $\tilde{\mathbf{B}}(\mathbf{Y})$  is communicated losslessly using a binary source codebook  $\mathcal{H}_{\text{supp}}^{\text{SE\&C}}$ , which can be generated by, e.g., the Huffman coding.

### V. RATE-DISTORTION PERFORMANCE LIMITS OF QUANTIZED COMPRESSED SENSING

For the QCS setup in Fig 1, the compression limit, i.e., the minimum achievable rate  $R$  for a given distortion  $D$ , is given by the remote rate-distortion function (RDF) of source  $\mathbf{X}$ , denoted as  $R_{\mathbf{X}}^{\text{rem}}(D)$ . While the closed-form solution of  $R_{\mathbf{X}}^{\text{rem}}(D)$  remains open, the RD performance of QCS can be assessed by the techniques derived in [14]. Thus, we establish information-theoretic benchmarks for the proposed QCS algorithms of Section IV by evaluating:

1) The analytical lower bound to  $R_{\mathbf{X}}^{\text{rem}}(D)$ , termed the conditional remote RDF  $R_{\mathbf{X}|\mathbf{B}}^{\text{rem}}(D)$ , given in [14, Theorem 1].  $R_{\mathbf{X}|\mathbf{B}}^{\text{rem}}(D)$  is based on having support side information  $\mathbf{B}$  at the encoder and decoder. Since the encoder of a practical QCS method can at best obtain an estimate of  $\mathbf{B}$ , each method is expected to have a considerable gap to  $R_{\mathbf{X}|\mathbf{B}}^{\text{rem}}(D)$ .

2) The numerical approximation technique for  $R_{\mathbf{X}}^{\text{rem}}(D)$  represented by the modified Blahut-Arimoto (BA) algorithm in [14, Algorithm 1]. Such an approximate curve, termed  $R_{\mathbf{X},\text{ba}}^{\text{rem}}(D)$ , can be obtained by running the BA algorithm [14, Algorithm 1] using a VQ that is trained via Algorithm 1 with  $\mathbf{U} = \tilde{\mathbf{X}}(\mathbf{Y})$ , where  $\tilde{\mathbf{X}}(\mathbf{Y})$  is the MMSE estimator (see (10)).  $R_{\mathbf{X},\text{ba}}^{\text{rem}}(D)$  serves as a meaningful lower bound to practical QCS methods because, as demonstrated in Section VI, it can be approached using VQ and entropy coding – at the cost of increased encoding complexity.

### VI. NUMERICAL RESULTS

Consider a setup with  $N = 30$ ,  $M = \{8, 12, 16, 20\}$ ,  $K = 2$ ,  $\Sigma_{\mathbf{G}} = \mathbf{I}_N$ ,  $\Sigma_{\mathbf{W}} = 0.01 \mathbf{I}_M$ ,  $p(\mathbf{b}_s) = 1/|\mathcal{B}|$ ,  $\forall \mathbf{b}_s \in \mathcal{B}$ ,

and  $\Phi$  generated by taking the first  $M$  rows of an  $N \times N$  discrete cosine transform matrix and normalizing the columns  $\|\cdot\|_2 = 1$ . Basis pursuit denoising (BPDN)<sup>2</sup> [34] is used as  $R_D$  in C&E,  $R_E$  in E&C, and  $S_E$  in SE&C;  $S_E$  forms the support estimate from the indices of the  $K$  largest magnitudes of the BPDN output. For C&E<sub>usq</sub> and SE&C<sub>usq</sub>,  $q^{lo}$  and  $q^{hi}$  are set as the minimum and maximum codepoints of the corresponding Lloyd-Max SQs. The same procedure is first used for E&C<sub>usq</sub>, after which each codepoint is added a constant shift so that E&C<sub>usq</sub> contains the codepoint of E&C<sub>sq</sub> that is closest to zero. ECSQs/ECVQs are run with  $\mu' = 0.1/\log(|I|)$ . All source codebooks  $\mathcal{H}$  are generated via the Huffman coding. The quantization algorithms are run with  $R' = \{2^1, \dots, 2^{10}\}$ . The BA algorithm [14, Algorithm 1] is run with  $2^{10}$  quantization levels.

Fig. 3 depicts the average normalized distortion  $10\log_{10}(\mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}\|_2^2]/\mathbb{E}[\|\mathbf{X}\|_2^2])$  (dB) versus the average rate  $R$  (bits/entry of  $\mathbf{X}$ ) of different QCS schemes for  $M = \{8, 12, 16, 20\}$ . The performances of the SQ based methods can be ranked as E&C < C&E < SE&C, and the VQ based methods as SE&C < C&E < E&C. Let us elaborate each of the three algorithm classes in more detail.

C&E: C&E<sub>usq</sub> and C&E<sub>sq</sub> perform nearly identically. The entropy coding in C&E<sub>ecsq</sub> can only slightly improve the performance. For large values of  $M$ , C&E<sub>vq</sub> slightly outperforms E&C<sub>vq</sub> even though E&C<sub>vq</sub> uses the optimal E&C structure. This is caused by the fact that  $R_E$  in E&C is approximated as the BPDN instead of the MMSE estimator in (10); while C&E<sub>vq</sub> uses a suboptimal C&E strategy, it uses the MSE-optimal outputs in (9) for a given VQ.

E&C: Among all QCS methods, entropy coding has the most significant impact on E&C<sub>sq</sub>  $\rightarrow$  E&C<sub>ecsq</sub>. This stems from the fact that while E&C<sub>sq</sub> quantizes the frequently occurring near zero inputs inefficiently using a fixed-length codeword, E&C<sub>ecsq</sub> intelligently assigns a very short codeword to them. The best practical QCS method is the E&C<sub>ecvq</sub> method proposed in [14, Sect. V] which approaches the compression limit of any practical QCS method, i.e.,  $R_{\mathbf{X},ba}^{rem}(D)$ .

SE&C: It can be seen that the proposed adaptive compression of the support set and estimated magnitudes is an effective strategy as the SE&C algorithms outperform all SQ based C&E and E&C methods. Thus, the SE&C method even with a simple USQ is a very competent, low-complexity alternative. Since all SE&C curves nearly coincide, the support recovery performance of chosen  $S_E$  is the limiting factor in SE&C.

As  $M$  increases, the QCS methods perform better and the gap between the analytical MMSE estimation floor with support side information (the lower horizontal line) and the error floor of the BPDN reconstruction (the upper horizontal line) decreases. Note that only C&E and E&C methods that rely on the plain BPDN outputs saturate to the upper floor; because SE&C removes the noise outside the (estimated)

<sup>2</sup>Note that better performance is attainable if, e.g.,  $R_E$  is set as the MMSE estimator in (10), yet at the cost of exponential encoding complexity [33] for a resource-limited sensor. Similarly, more sophisticated support recovery algorithms  $S_E$  could be considered as well.

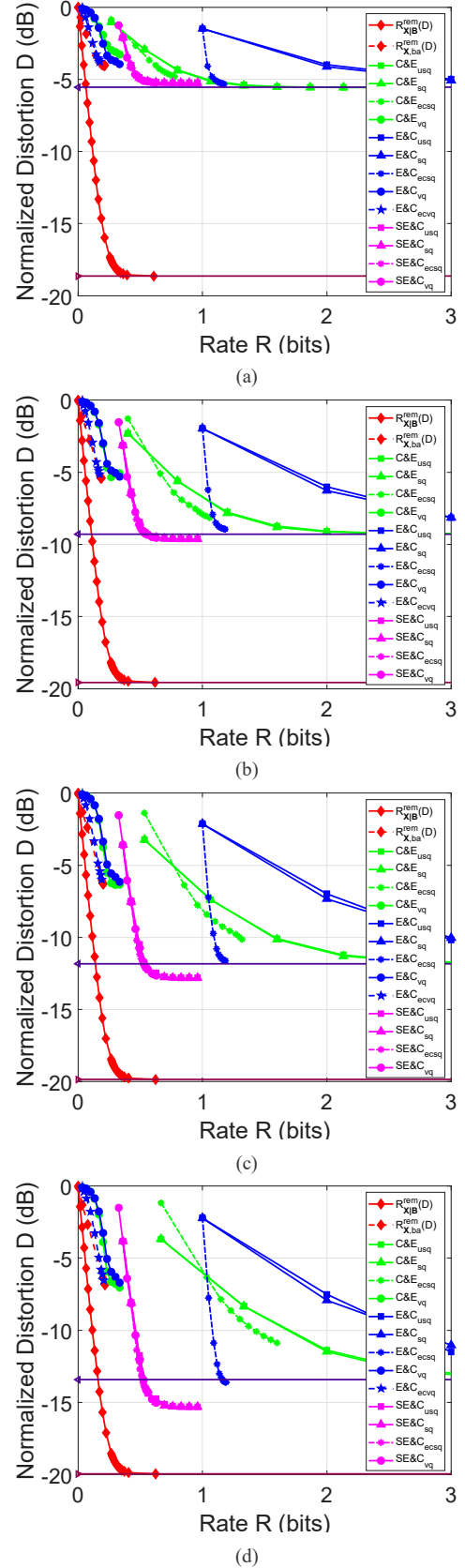


Fig. 3. RD performance of QCS schemes for  $N = 30$ ,  $K = 2$ , and (a)  $M = 8$ , (b)  $M = 12$ , (c)  $M = 16$ , and (d)  $M = 20$ .

support of each BPDN output, it can surpass the floor, as illustrated by Figs. 3(b)–(d).

## VII. CONCLUSIONS

The paper studied the RD performance of various practical and theoretical QCS schemes. Practical QCS algorithms relying on 1) compress-and-estimate, 2) estimate-and-compress, and 3) support-estimation-and-compress strategies were proposed. The information-theoretic limit of the setup – the remote RDF – was evaluated by an analytical lower bound and a numerical BA approximation method. Simulation results showed that when SQ is used, the proposed adaptive compression of the support set and estimated magnitudes is an effective strategy. When VQ/ECVQ is used, the estimate-and-compress strategy, as supported by the theory, is the best one. Accordingly, the ECVQ based estimate-and-compress method was numerically shown to approach the remote RDF.

## ACKNOWLEDGEMENT

The work of M. Leinonen and M. Codreanu has been financially supported by Infotech Oulu. M. Codreanu would like to acknowledge the support of the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 793402 (COMPRESS NETS). The research has also been financially supported by the Academy of Finland, and Academy of Finland 6Genesis Flagship (grant 318927).

## REFERENCES

- [1] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, “Internet of Things for smart cities,” *IEEE Internet of Things J.*, vol. 1, no. 1, pp. 22–32, Feb. 2014.
- [2] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, “Internet of Things: A survey on enabling technologies, protocols, and applications,” *IEEE Commun. Surveys & Tutorials*, vol. 17, no. 4, pp. 2347–2376, Fourth quarter 2015.
- [3] Y. Mehmood, F. Ahmad, I. Yaqoob, A. Adnane, M. Imran, and S. Guizani, “Internet-of-Things-based smart cities: Recent advances and challenges,” *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 16–24, Sep. 2017.
- [4] V. Raghunathan, C. Schurgers, S. Park, and M. B. Srivastava, “Energy-aware wireless microsensor networks,” *IEEE Signal Processing Mag.*, vol. 19, no. 2, pp. 40–50, Mar. 2002.
- [5] G. Anastasi, M. Conti, M. D. Francesco, and A. Passarella, “Energy conservation in wireless sensor networks: A survey,” *Ad Hoc Netw.*, vol. 7, no. 3, pp. 537–568, May 2009.
- [6] E. J. Candés, J. Romberg, and T. Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [7] D. L. Donoho, “Compressed sensing,” *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [8] D. Baron, M. Wakin, M. Duarte, S. Sarvotham, and R. Baraniuk, “Distributed compressed sensing,” Dept. Elect. Comput. Eng., Rice University, Houston, TX, Tech. Rep. ECE-0612, Nov. 2005. [Online]. Available: <http://webee.technion.ac.il/people/drorb/pdf/DCS112005.pdf>
- [9] M. Leinonen, M. Codreanu, and M. Juntti, “Sequential compressed sensing with progressive signal reconstruction in wireless sensor networks,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1622–1635, Mar. 2015.
- [10] D. Malioutov, M. Çetin, and A. S. Willsky, “A sparse signal reconstruction perspective for source localization with sensor arrays,” *IEEE Trans. Signal Processing*, vol. 53, no. 8, pp. 3010–3022, Aug. 2005.
- [11] A. M. R. Dixon, E. G. Allstot, D. Gangopadhyay, and D. J. Allstot, “Compressed sensing system considerations for ECG and EMG wireless biosensors,” *IEEE Trans. Biomed. Circ. Syst.*, vol. 6, no. 2, pp. 156–166, Apr. 2012.
- [12] V. Goyal, A. Fletcher, and S. Rangan, “Compressive sampling and lossy compression,” *IEEE Signal Processing Mag.*, vol. 25, no. 2, pp. 48–56, 2008.
- [13] R. Dobrushin and B. Tsybakov, “Information transmission with additional noise,” *IRE Trans. Inform. Theory*, vol. 8, no. 5, pp. 293–304, Sep. 1962.
- [14] M. Leinonen, M. Codreanu, M. Juntti, and G. Kramer, “Rate-distortion performance of lossy compressed sensing of sparse sources,” *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4498–4512, Oct. 2018.
- [15] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, ser. Applied and Numerical Harmonic Analysis. Springer New York, 2013.
- [16] J. Max, “Quantizing for minimum distortion,” *Proc. IRE*, vol. 6, no. 1, pp. 7–12, Mar. 1960.
- [17] S. Lloyd, “Least squares quantization in PCM,” *IEEE Trans. Inform. Theory*, vol. 28, no. 2, pp. 129–137, Mar. 1982.
- [18] Y. Linde, A. Buzo, and R. Gray, “An algorithm for vector quantizer design,” *IEEE Trans. Commun.*, vol. 28, no. 1, pp. 84–95, Jan. 1980.
- [19] R. Wood, “On optimum quantization,” *IEEE Trans. Inform. Theory*, vol. 15, no. 2, pp. 248–252, Mar. 1969.
- [20] P. Chou, T. Lookabaugh, and R. Gray, “Entropy-constrained vector quantization,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 1, pp. 31–42, Jan. 1989.
- [21] M. Leinonen, M. Codreanu, and M. Juntti, “Distributed distortion-rate optimized compressed sensing in wireless sensor networks,” *IEEE Trans. Commun.*, vol. 66, no. 4, pp. 1609–1623, Apr. 2018.
- [22] D. Huffman, “A method for the construction of minimum-redundancy codes,” *Proc. IRE*, vol. 40, no. 9, pp. 1098–1101, Sep. 1952.
- [23] L. Jacques, D. Hammond, and J. Fadili, “Dequantizing compressed sensing: When oversampling and non-Gaussian constraints combine,” *IEEE Trans. Inform. Theory*, vol. 57, no. 1, pp. 559–571, 2011.
- [24] W. Dai and O. Milenkovic, “Information theoretical and algorithmic approaches to quantized compressive sensing,” *IEEE Trans. Commun.*, vol. 59, no. 7, pp. 1857–1866, 2011.
- [25] A. Kipnis, G. Reeves, and Y. C. Eldar, “Single letter formulas for quantized compressed sensing with Gaussian codebooks,” in *Proc. IEEE Int. Symp. Inform. Theory*, Vail, CO, USA, Jun. 17–22 2018, pp. 71–75.
- [26] J. Wolf and J. Ziv, “Transmission of noisy information to a noisy receiver with minimum distortion,” *IEEE Trans. Inform. Theory*, vol. 16, no. 4, pp. 406–411, Jul. 1970.
- [27] A. Shirazinia, S. Chatterjee, and M. Skoglund, “Joint source-channel vector quantization for compressed sensing,” *IEEE Trans. Signal Processing*, vol. 62, no. 14, pp. 3667–3681, Jul. 2014.
- [28] —, “Distributed quantization for measurement of correlated sparse sources over noisy channels,” Jul. 2015, available at <http://arxiv.org/abs/1404.7640v2>.
- [29] M. Leinonen, M. Codreanu, and M. Juntti, “Channel-robust compressed sensing via vector pre-quantization in wireless sensor networks,” in *Proc. IEEE Global Conf. on Signal and Inform. Proc.*, Orlando, Florida, USA, Dec. 14–16 2015, pp. 383–387.
- [30] —, “Distributed variable-rate quantized compressed sensing in wireless sensor networks,” in *Proc. IEEE Works. on Sign. Proc. Adv. in Wirel. Comms.*, invited paper, Edinburgh, UK, Jul. 3–7 2016, pp. 1–5.
- [31] A. Kipnis, G. Reeves, Y. Eldar, and A. Goldsmith, “Compressed sensing under optimal quantization,” in *Proc. IEEE Int. Symp. Inform. Theory*, Aachen, Germany, Jun. 25–30 2017, pp. 2148–2152.
- [32] M. Leinonen, M. Codreanu, M. Juntti, and G. Kramer, “Rate-distortion lower bound for compressed sensing via conditional remote source coding,” in *Proc. IEEE Inform. Theory Workshop*, Cambridge, UK, Sep. 11–14 2016, pp. 275–279.
- [33] M. Elad and I. Yavneh, “A plurality of sparse representations is better than the sparsest one alone,” *IEEE Trans. Inform. Theory*, vol. 55, no. 10, pp. 4701–4714, Oct. 2009.
- [34] S. Chen, D. Donoho, and M. Saunders, “Atomic decomposition by basis pursuit,” *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.