ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ARCHIVIO ISTITUZIONALE
DELLA RICERCA

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

Lossy Compression of Noisy Sparse Sources Based on Syndrome Encoding

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:
Ahmed Elzanaty, A.G. (2019). Lossy Compression of Noisy Sparse Sources Based on Syndrome Encoding. IEEE TRANSACTIONS ON COMMUNICATIONS, 67(10), 7073-7087 [10.1109/TCOMM.2019.2926080].

Availability:
This version is available at: https://hdl.handle.net/11585/731774 since: 2020-02-25

Published:
DOI: http://doi.org/10.1109/TCOMM.2019.2926080

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/).
When citing, please refer to the published version.

(Article begins on next page)

12 October 2024

This is the final peer-reviewed accepted manuscript of:

A. Elzanaty, A. Giorgetti and M. Chiani, "Lossy Compression of Noisy Sparse Sources Based on Syndrome Encoding," in *IEEE Transactions on Communications*, vol. 67, no. 10, pp. 7073-7087, Oct. 2019.

The final published version is available online at: https://doi.org/10.1109/TCOMM.2019.2926080

Rights / License:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/)

When citing, please refer to the published version.

Lossy Compression of Noisy Sparse Sources Based on Syndrome Encoding

Ahmed Elzanaty, *Member, IEEE*, Andrea Giorgetti, *Senior Member, IEEE*, and Marco Chiani, *Fellow, IEEE*

***Abstract*—Data originating from devices and sensors in Internet of Things scenarios can often be modeled as sparse signals. In this paper, we provide new source compression schemes for noisy sparse and non-strictly sparse sources, based on channel coding theory. Specifically, nonlinear excision filtering by means of model order selection or thresholding is first used to detect the support of the non-zero elements of sparse vectors in noise. Then, the sparse sources are quantized and compressed using syndrome based encoders. The theoretical performance of the schemes is provided, accounting for the uncertainty in the support estimation. In particular, we derive the operational distortion-rate and operational distortion-energy of the encoders for noisy Bernoulli-uniform and Bernoulli-Gaussian sparse sources. It is found that the performance of the proposed encoders approaches the information-theoretic bounds for sources with low sparsity order. As a case study, the proposed encoders are used to compress signals gathered from a real wireless sensor network for environmental monitoring.**

***Index Terms*—Internet of Things, source coding, sparse and compressible sources, error correcting codes, syndrome encoding, compressed sensing, rate-distortion.**

I. Introduction

The classic problem of designing efficient lossy compression schemes for sources is gaining increasing interest, supported by the tremendous increase of data generated by the Internet of Things (IoT) [1], [2]. The data size can be significantly decreased by exploiting some of their structures. One of these structures is the sparsity/compressibility, i.e., the ability to describe/approximate signals with a fewer number of coefficients compared to their dimension in some domains, e.g., time, frequency, discrete cosine transform (DCT), and Wavelet. Most of the signals of interest such as image, audio, video, and IoT data are compressible [3]–[6]. The key objectives regarding the compression of sparse sources in IoT scenarios are the design of energy-efficient source encoders with low computational complexity, and their performance analysis in terms of operational distortion-rate (ODR) and operational distortion-energy (ODE) in the presence of noise [2], [5]–[8].

This work was supported in part by MIUR under the program "Dipartimenti di Eccellenza (2018-2022)," and in part by the EU project eCircular (EIT Climate-KIC). This paper was presented in part at IEEE Globecom 2016 and IEEE Globecom 2017.

A. Elzanaty is with the Computer, Electrical and Mathematical Science and Engineering Division, King Abdullah University of Science and Technology, Thuwal 23955, Saudi Arabia (e-mail: ahmed.elzanaty@kaust.edu.sa). He was with the University of Bologna, Italy.

A. Giorgetti and M. Chiani are with the Department of Electrical, Electronic, and Information Engineering "Guglielmo Marconi" (DEI), University of Bologna, Italy (e-mail: {andrea.giorgetti, marco.chiani}@unibo.it).

To illustrate the problem, let us start by considering a simple model for strictly sparse signals in noise-free environments, where a signal $\boldsymbol{x} \in \mathbb{R}^N$ is emitted by a discrete-time continuous-valued source such that only k_0 elements of $\boldsymbol{x}$ are non-zero, i.e., $\boldsymbol{x}$ is a sparse vector with sparsity order k_0. The locations of the non-zero elements are chosen uniformly at random among all the $\binom{N}{k_0}$ subsets of $\{1, 2, \cdots, N\}$ with cardinality k_0. The target is to encode $\boldsymbol{x}$ at the minimum rate which still guarantees a predefined distortion. The first intuitive approach, which we name address coding (AC), is to quantize separately the k_0 non-zero components using a uniform scalar quantizer with b bits/sample, then encode each of their locations (addresses) with a fixed number of bits $\lceil \log_2 N \rceil$. The total number of required bits per sample (i.e., the rate) is calculated as $r_{\text{AC}} = \big(k_0 \lceil \log_2 N \rceil + k_0\, b\big)/N$. This approach is simple, and it requires the transmission of both the values and addresses separately. However, there are more efficient ways to encode the signal support. One is to assign a binary codeword to each possible support configuration using a Huffman code. Nevertheless, this would require all codewords to be stored along with their associated support in a look-up table with $\binom{N}{k_0}$ rows. This approach becomes impractical with increasing complexity for values of N and k_0 of interest, because the number of possible configurations $\binom{N}{k_0}$ is exponentially large in N.

Another method is based on compressed sensing (CS), where $M < N$ linear real-valued observations are collected from $\boldsymbol{x}$, i.e., $\boldsymbol{y} = \boldsymbol{A}\,\boldsymbol{x}$, through an $M \times N$ measurement matrix $\boldsymbol{A}$ [9], [10]. Then, in order to transmit the information with a finite rate, a quantization for the measurement vector $\boldsymbol{y}$ is needed, introducing a distortion [6]. One important advantage of CS is that the knowledge about the domain at which the signal is sparse is required only at the decoder side and not at the encoder. Moreover, CS has been proved to be stable with respect to non-strictly sparse (compressible) signals, and robust to noise [9]–[13]. On the other hand, the number of required measurements for sparse recovery is still considerably higher than the signal sparsity, e.g., $M \geq 2\,k_0\,\log(e\,N/k_0)$ is a sufficient condition for perfect recovery of a k_0-sparse vector from $\boldsymbol{y}$ with high probability [14, (9.24)]. Also, it is not easy to implement multiplication in the analog domain between the random measurement matrix and the signal, except for Rademacher and Bernoulli matrices [15], [16].

The research directions regarding the compression of sparse sources can be divided into three main categories as follows.

1) Information distortion-rate (IDR) of sparse sources: Several works address the problem of deriving information theoretic bounds in terms of the IDR for directly observed

sparse sources, i.e., when $\boldsymbol{x}$ is known at the source encoder. In particular, upper and lower bounds on the IDR for a class of sparse signals (i.e., Bernoulli-Gaussian (BG) sources) are provided in [5], [17], [18]. These bounds are asymptotically tight in the high-rate low-distortion region. Nevertheless, expressing the IDR of BG sources in closed-form is still an open problem.

2) Remote IDR of sparse sources: The remote source coding problem appears when the encoder does not have direct access to the source [19], [20]. This is the case when the encoder observes a noisy source, or when linear measurements of the signal are available as in CS. Lower and upper bounds on the remote IDR for noisy continuous distributed (not-sparse) sources were given in [21]. A lower bound on the remote IDR for CS has been derived in [22], assuming that both the encoder and decoder have the signal support as a side information. A numerical approximation and single letter formulas of the remote IDR for CS have been given in [22]–[24].

3) Practical encoders of sparse sources: In order to design source coding schemes with ODR close to the IDR, long sequences of the signal should be discretized using vector quantization methods with jointly-optimized quantization regions and codebooks. This process is usually computationally heavy, introducing significant complexity, power consumption, and latency, which may not fit many IoT applications. Instead, a simpler scheme for source encoders involves a scalar uniform quantizer followed by an efficient lossless encoder [25, Section 5.8]. In this regard, the ODR of scalar quantizers has been analyzed for non-sparse sources with uniform, generalized Gaussian, and exponential distributions [26]–[29]. For sparse generalized BG sources, an approximation for the optimal ODR of uniform quantizers without considering the saturation effect was provided in [30], while asymptotic formulas have been given in [31].

Source encoders based on channel coding theory have been utilized for the compression of discrete non-sparse sources. Specifically, syndrome based encoders for binary sources using block and polar channel codes have been proposed in [32] and [33], respectively. In [34], a lossy source encoder for discrete memoryless sources was proposed based on a nonlinear channel code constructed from a sparse graph. In a quite different context, syndrome encoding has been used to approach the Slepian-Wolf/Wyner-Ziv bounds with lossless/lossy distributed source coding, assuming side information at the receiver [35]–[37]. For example, a compression scheme for continuous sources considering that the decoder has side information, i.e., a noisy version of the source, is proposed in [38]. Several works have discussed designing codes for similar problems, e.g., multi-terminal source coding [39]–[42], successive refinement source compression [43], and joint source channel coding [44], [45].

Regarding CS, the ODR with scalar uniform quantization was analyzed in the high rate region in [46], considering that the support is known at the decoder and that the measurements are noise-free. The ODR of a CS based encoder for wireless sensor network (WSN) has been numerically investigated in [47], while prototypes for the encoder have been designed in [15]. Quantized CS based encoders and recovery algorithms have been investigated in several works, e.g., [48]–[52]. The connection between CS and syndrome channel coding over real fields has been discussed in [53]–[55].

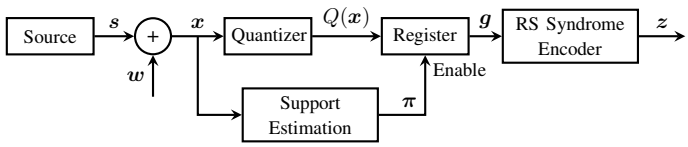
In this paper, we provide two practical approaches for source compression of noisy sparse sources. Our contribution can be summarized as follows:

- We provide two alternatives for nonlinear excision filtering to estimate the signal support at the encoder, i.e., a blind estimator based on model order selection theory and a thresholding estimator. After support estimation, the signal is denoised and quantized.
- We propose novel schemes for the lossless compression of this quantized sparse signal. The basic idea is to consider a sparse signal as an error vector summed to an all-zeros codeword, for a chosen error correcting code (e.g., RS). The compressed version of the sparse vector, computed at the encoder, is here the associated syndrome. From the syndrome, the decoder can reconstruct the sparse signal as the smallest weight (sparsest) error vector compatible with the syndrome.
- We provide the end-to-end performance analysis of the proposed schemes in terms of the ODR, for noisy Bernoulli-uniform and Bernoulli-Gaussian sources. In particular, closed-form expression for the ODR is derived accounting for finite rates, additive noise, quantization error, and imperfect support estimation.
- We derive operational distortion-energy design curves, considering a typical power profile for an implementation of the proposed encoders.
- As a case study, we address the compression of signals gathered from a real WSN, deployed to monitor landslides. When compared to a CS based encoder, it is found that syndrome based compression can increase the network lifetime and the accuracy of the recovered data.

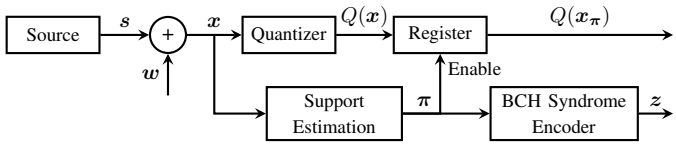
Throughout this paper, we denote the ℓ_0 quasi-norm with $\|\cdot\|_0$, the ℓ_2 norm with $\|\cdot\|$, the Dirac-delta function with $\delta(\cdot)$, vectors with bold small letters, matrices with bold capital letters, the N-dimensional identity matrix with $\boldsymbol{I}_N$, the Galois field of order q with $\mathbb{F}_q$, the multivariate Gaussian distribution having mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{C}$ with $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{C})$, the univariate Gaussian distribution having mean μ and variance σ^2 truncated between a and b and renormalized with $\mathcal{N}(\mu, \sigma^2, a, b)$, the uniform distribution between a and b with $\mathcal{U}(a, b)$, random variables (r.v.s) with calligraphic letters, e.g., $\mathcal{S}$, the probability density function (PDF) of $\mathcal{S}$ with $f_{\mathcal{S}}(s)$, random vectors with bold calligraphic letters, the error function with $\mathrm{erf}(\cdot)$, the complementary error function with $\mathrm{erfc}(\cdot)$, and the indicator function that equals one when the condition A is satisfied and zero otherwise with $\mathbf{1}_{\{A\}}$.

II. Mixed Distributed Sparse Sources

Let us consider a memoryless sparse source that emits independent, identically distributed (i.i.d.) symbols arranged into a vector $\boldsymbol{\mathcal{S}} \in \mathbb{R}^N$. Each source symbol $\mathcal{S}$ is generated as a multiplication of two r.v.s, i.e., $\mathcal{S} = \mathcal{Z}\,\mathcal{Y}$, where $\mathcal{Z} \in \{0, 1\}$ is a Bernoulli r.v. with $\mathbb{P}\{\mathcal{Z} = 1\} = p$, while $\mathcal{Y}$ is drawn



(a) RS based source encoder.



(b) BCH based source encoder.

Fig. 1: The block diagrams of the proposed compression schemes for noisy sparse sources.

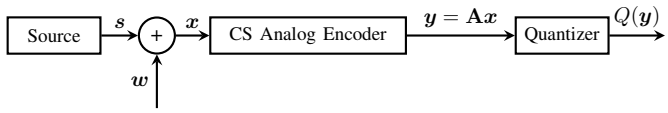


Fig. 2: The block diagram of the CS based source encoder [6], [15].

from some continuous distribution $f_{\mathcal{Y}}(y)$.[1] The corresponding PDF of the mixed distributed source is given by $f_{\mathcal{S}}(s) = p\, f_{\mathcal{Y}}(s) + (1-p)\delta(s)$. This model is referred as Bernoulli-Gaussian (BG) and Bernoulli-uniform (BU), considering the continuous distribution as Gaussian and uniform, respectively.

The acquisition device may add noise to the input signal, thus the acquired noisy vector can be represented as

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{S}} + \boldsymbol{\mathcal{W}} \tag{1}$$

where $\boldsymbol{\mathcal{W}} \sim \mathcal{N}(\mathbf{0}, \sigma_{\mathrm{n}}^2\, \boldsymbol{I}_N)$ represents the additive noise. This model fits also the case of compressible sources, as the vector $\boldsymbol{\mathcal{W}}$ can account for the negligible components [5]. Moreover, it is usually adopted within the framework of sparse representation for images, sounds, medical data, and sensor signals in appropriate transform domains [5], [56], [57], facilitating the design and analysis of several statistical signal processing algorithms for various applications, e.g., the compression and denoising of images, videos, and speech signals [58]–[60].

III. Compression Schemes based on Syndrome Encoding

In this section, we describe two novel schemes for efficient lossy compression of sparse sources, shown in Fig. 1, which will be compared to CS (Fig. 2). Let us consider a random realization of the noiseless source, $\boldsymbol{s}$, with k_0 non-zero elements. At first, the locations of the non-zero elements is estimated from $\boldsymbol{x}$. Then, the data is compressed by calculating the syndromes using the parity-check matrix of a RS or BCH code.

In the following, we will separately illustrate each part of the proposed scheme.

[1]Note that p represents the average sparsity ratio of the source.

A. Support Detection and Signal Denoising

We propose two methods that can be used for support estimation. In particular, the first approach exploits model order selection based on generalized information criterion (GIC), while the second method relies on an excision thresholding filter. The GIC does not require the knowledge of the statistical distribution of the source nor the noise power (i.e., it is universal), and its theoretical performance analysis is quite involved [61]. On the contrary, the excision filter needs the statistical distributions of the source and the noise to be known; however, the design and analysis of this filter is easier. In the following these techniques are discussed in detail.

1) Support Detection by Model Order Selection: A novel estimator is derived for the number and locations of the non-zero elements in the noisy sparse signal, based on model order selection theory [61]–[64]. A powerful solution to model estimation is based on information-theoretic criterion, where the model order is determined by minimizing a penalized likelihood [62], [65]. In particular, we consider the GIC because of its versatility in controlling the estimation accuracy [61], [62], [64].

Given a realization of the sparse signal, $\boldsymbol{s}$, considered unknown, let us define a family of models to fit the measured data, with the kth model representing the case where the vector $\boldsymbol{s}$ has k non-zero elements, with locations identified by the set $\boldsymbol{\pi} = \{\pi_1, \pi_2, \cdots, \pi_k\}$.[2] The positions of the zero elements are thus given by the complement set $\boldsymbol{\pi}^c \triangleq \{1, 2, \cdots, N\} \setminus \boldsymbol{\pi}$.

The PDF of the vector $\boldsymbol{\mathcal{X}}$ can be expressed as

$$f(\boldsymbol{x}; \boldsymbol{\Theta}^{(k)}) \;=\; \left(2\pi\sigma_{\mathrm{n}}^2\right)^{-N/2} e^{-\frac{\sum_{i\in\boldsymbol{\pi}}(x_i - s_i)^2 + \sum_{i\in\boldsymbol{\pi}^{\mathrm{c}}} x_i^2}{2\sigma_{\mathrm{n}}^2}} \tag{2}$$

where $\boldsymbol{\Theta}^{(k)} \triangleq (\sigma_{\mathrm{n}}^2, \boldsymbol{s}_{\boldsymbol{\pi}}, \boldsymbol{\pi})$ is the vector of the $2k+1$ unknown parameters, and $\boldsymbol{s}_{\boldsymbol{\pi}} \triangleq \{s_{\pi_1}, s_{\pi_2}, \cdots, s_{\pi_k}\}$. A common solution for the estimation of the parameters is by maximizing the likelihood function as $\widehat{\boldsymbol{\Theta}}^{(k)} = \arg\max_{\boldsymbol{\Theta}^{(k)}} \log f(\boldsymbol{x}; \boldsymbol{\Theta}^{(k)})$. However, a direct likelihood maximization would yield biased and inconsistent estimator for the noise variance, because of the incidental parameter problem [66], [67]. More precisely, the likelihood function depends on two types of parameters: the incidental (nuisance) parameters $\boldsymbol{s}_{\boldsymbol{\pi}}$ that appear only in k samples; and the structural parameters σ_{n}^2 and $\boldsymbol{\pi}$, common to all samples. Hence, the observations are partially consistent, leading to inconsistent estimation of the common parameters [66], [67]. In order to find consistent estimators for the parameters, following the procedure indicated in [66], we factorize the PDF in (2) as

$$\begin{aligned} f(\boldsymbol{x}; \boldsymbol{\Theta}^{(k)}) = & \underbrace{\left(2\pi\sigma_{\mathrm{n}}^2\right)^{\frac{k-N}{2}} \exp\left(-\frac{\sum_{i\in\boldsymbol{\pi}^{\mathrm{c}}} x_i^2}{2\sigma_{\mathrm{n}}^2}\right)}_{f_1(\boldsymbol{x}|\boldsymbol{x}_{\boldsymbol{\pi}};\sigma_{\mathrm{n}}^2;\boldsymbol{\pi}^{\mathrm{c}})} \\ & \times \underbrace{\left(2\pi\sigma_{\mathrm{n}}^2\right)^{-k/2} \exp\left(-\frac{\sum_{i\in\boldsymbol{\pi}}(x_i - s_i)^2}{2\sigma_{\mathrm{n}}^2}\right)}_{f_2\left(\boldsymbol{x}_{\boldsymbol{\pi}};\boldsymbol{\Theta}^{(k)}\right)} \end{aligned} \tag{3}$$

[2]To keep light the notation we will avoid to explicitly write the dependence of this and the following sets on k.

where the distribution of the observations conditional on $\boldsymbol{x}_{\boldsymbol{\pi}}$ does not depend on the incidental parameters. Thus, the locations of the zero elements and the variance can be estimated for the kth model by maximizing the partial likelihood function [66]

$$\left(\hat{\sigma}^2, \widehat{\boldsymbol{\pi}}^{\mathrm{c}}\right) = \arg\max_{\sigma_{\mathrm{n}}^2, \boldsymbol{\pi}^{\mathrm{c}}} f_1\left(\boldsymbol{x}|\boldsymbol{x}_{\boldsymbol{\pi}}; \sigma_{\mathrm{n}}^2; \boldsymbol{\pi}^{\mathrm{c}}\right). \tag{4}$$

From (3), we see that the maximum in (4) is obtained for the locations $\widehat{\boldsymbol{\pi}}^{\mathrm{c}}$ corresponding to the $N-k$ elements in $\boldsymbol{x}$ with the smallest square values, and for $\hat{\sigma}^2 = \frac{1}{N-k}\sum_{i\in\widehat{\boldsymbol{\pi}}^{\mathrm{c}}} x_i^2$. The values of the non-zero elements are estimated by maximizing $f_2\left(\boldsymbol{x}_{\boldsymbol{\pi}}; \boldsymbol{\Theta}^{(k)}\right)$, resulting in $\widehat{\boldsymbol{s}}_{\boldsymbol{\pi}} = \boldsymbol{x}_{\widehat{\boldsymbol{\pi}}}$, where $\widehat{\boldsymbol{\pi}}$, which is the complement of the set $\widehat{\boldsymbol{\pi}}^{\mathrm{c}}$, contains the positions of the k samples with the largest modulus. In other words, the estimate of the support in the kth model implies that the smallest (in module) $N-k$ samples are due to noise only.

Therefore, the log-likelihood function of the kth model can be written from (2) by substituting the parameters with their estimates as

$$\log f\left(\boldsymbol{x}; \widehat{\boldsymbol{\Theta}}^{(k)}\right) = -\frac{N}{2}\log\left(\frac{2\pi}{N-k}\sum_{i=k+1}^{N} x_{[i]}^2\right) - \frac{N-k}{2}$$

where $x_{[1]}, x_{[2]}, \ldots, x_{[N]}$ are the signal samples sorted according to their absolute values, i.e., $|x_{[1]}| \geq |x_{[2]}| \ldots \geq |x_{[N]}|$. Finally, the estimated sparsity order can be found by the GIC [62, (75)] as

$$\hat{k}_0 = \underset{k\in[0,k_{\max}]}{\arg\min}\left\{N\,\log\left(\sum_{i=k+1}^{N} x_{[i]}^2\right) - N\log(N-k) + k\,(2\,\nu-1)\right\}$$

where $k_{\max} < N$ is the maximum sparsity order, ν is the GIC penalty factor, and all terms that do not depend on k are omitted. Once we have $\hat{k}_0$, the estimated support is then the position of the $\hat{k}_0$ largest (in modulus) samples. Indicating this set as $\hat{\boldsymbol{\pi}}(\hat{k}_0)$, the corresponding filtered sparse signal is

$$x_i^* = \begin{cases} x_i, & i \in \hat{\boldsymbol{\pi}}(\hat{k}_0) \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

We will show in Section VII that this blind estimator performs well for high signal-to-noise ratio (SNR). As for most algorithms based on model selection, its theoretical performance analysis is quite complex, and the parameter ν can be optimized numerically, as investigated in Section VII.

2) Support Detection using Excision Filter: Here we present an alternative method for support recovery, useful when the distributions of $\mathcal{S}$ and $\mathcal{W}$ along with their parameters are known a-priori. To this aim, the support detection problem of the embedded signal in noise is reformulated as a hypothesis testing problem. At first, the detector goal is to discriminate, for each noisy source sample x, between two hypothesis: $\mathcal{H}_0 : x = w$ and $\mathcal{H}_1 : x = y + w$.[3] The PDFs of the noisy source, $\mathcal{X}$, given the hypothesis $\mathcal{H}_1$ and $\mathcal{H}_0$ are $f_{\mathcal{X}|\mathcal{H}_1}(x|\mathcal{H}_1)$ and $f_{\mathcal{X}|\mathcal{H}_0}(x|\mathcal{H}_0)$, respectively.

[3]Since the source is i.i.d., the subscript i is dropped here.

Regarding BG sources with $\mathcal{Y} \sim \mathcal{N}(0, \sigma_{\mathrm{s}}^2)$, the log-likelihood ratio test can be written as

$$\log\Lambda(x) = \log\left(\frac{\sigma_{\mathrm{n}}}{\sigma_{\mathrm{x}}}\right) + x^2\,\frac{\sigma_{\mathrm{x}}^2 - \sigma_{\mathrm{n}}^2}{2\,\sigma_{\mathrm{x}}^2\,\sigma_{\mathrm{n}}^2} \underset{\mathcal{D}_0}{\overset{\mathcal{D}_1}{\gtrless}} \lambda$$

with $\sigma_{\mathrm{x}} \triangleq \sqrt{\sigma_{\mathrm{n}}^2 + \sigma_{\mathrm{s}}^2}$, leading to

$$x^2 \underset{\mathcal{D}_0}{\overset{\mathcal{D}_1}{\gtrless}} \eta \tag{6}$$

where the threshold η is chosen to minimize the ODR, as will be illustrated later. Therefore, the support estimation turns into an energy detection problem to declare the presence/absence of the non-zero elements. Then, as illustrated before, the encoder nulls out the samples declared as $\mathcal{D}_0$ to obtain the denoised signal $\boldsymbol{x}^*$. The two important metrics that determine the performance of such detector are the probability of miss-detection, $\mathrm{P_{MD}}$, and the probability of false-alarm, $\mathrm{P_{FA}}$, given by

$$\mathrm{P_{MD}}(\eta) = \mathbb{P}\left\{\mathcal{X}^2 < \eta|\,\mathcal{H}_1\right\} = \mathrm{erf}\left(\sqrt{\frac{\eta}{2\sigma_{\mathrm{x}}^2}}\right) \tag{7}$$

$$\mathrm{P_{FA}}(\eta) = \mathbb{P}\left\{\mathcal{X}^2 \geq \eta|\,\mathcal{H}_0\right\} = \mathrm{erfc}\left(\sqrt{\frac{\eta}{2\sigma_{\mathrm{n}}^2}}\right). \tag{8}$$

For BU sources with $\mathcal{Y} \sim \mathcal{U}(-A, A)$, we have

$$f_{\mathcal{X}|\mathcal{H}_1}(x|\mathcal{H}_1) = \left[\mathrm{erf}\left(\frac{A-x}{\sqrt{2}\sigma_{\mathrm{n}}}\right) + \mathrm{erf}\left(\frac{A+x}{\sqrt{2}\sigma_{\mathrm{n}}}\right)\right]/4A,$$

leading to the log-likelihood ratio test

$$\log\Lambda(x) = \log\frac{\sqrt{2\pi}}{4A}\sigma_{\mathrm{n}} + \frac{x^2}{2\sigma_{\mathrm{n}}^2} + \log\left[\mathrm{erf}\left(\frac{A-x}{\sqrt{2}\sigma_{\mathrm{n}}}\right) + \mathrm{erf}\left(\frac{A+x}{\sqrt{2}\sigma_{\mathrm{n}}}\right)\right] \underset{\mathcal{D}_0}{\overset{\mathcal{D}_1}{\gtrless}} \eta. \tag{9}$$

This test metric can be approximated using the Taylor series expansion around $x = 0$ as

$$\log\Lambda(x) \simeq \log\left[\frac{\sigma_{\mathrm{n}}}{\sqrt{2}A}\,\mathrm{erf}\left(\frac{A}{\sqrt{2}\sigma_{\mathrm{n}}}\right)\right] + x^2\left(\frac{1}{2\sigma_{\mathrm{n}}^2} - \frac{Ae^{-\frac{A^2}{2\sigma_{\mathrm{n}}^2}}}{\sqrt{2\pi}{\sigma_{\mathrm{n}}}^3\,\mathrm{erf}\left(\frac{A}{\sqrt{2}\sigma_{\mathrm{n}}}\right)}\right)$$

which is equivalent to the energy detection in (6) with the same false-alarm (8). The probability of miss-detection is given as

$$\mathrm{P_{MD}}(\eta) = 2\int_0^{\sqrt{\eta}} f_{\mathcal{X}|\mathcal{H}_1}(x|\mathcal{H}_1)dx = \frac{\sigma_{\mathrm{n}}}{\sqrt{2\pi}A}e^{-\frac{(A+\sqrt{\eta})^2}{2\sigma_{\mathrm{n}}^2}} \times\left(1 - e^{\frac{2A\sqrt{\eta}}{\sigma_{\mathrm{n}}^2}}\right) + \frac{1}{2A}\left[\left(\sqrt{\eta} - A\right)\mathrm{erf}\left(\frac{A-\sqrt{\eta}}{\sqrt{2}\sigma_{\mathrm{n}}}\right) + \left(A + \sqrt{\eta}\right)\mathrm{erf}\left(\frac{A+\sqrt{\eta}}{\sqrt{2}\sigma_{\mathrm{n}}}\right)\right]. \tag{10}$$

B. Scalar Uniform Quantizer

Due to the large number of zero elements in $\boldsymbol{x}^*$, we consider a scalar mid-tread uniform quantizer, whose zero-valued level prevents the introduction of additional quantization noise out of the signal support [68]. This quantizer maps each element

x_i^* of $\boldsymbol{x}^*$ to an integer from a discrete set, $Q : \mathbb{R} \to \{0, 1, \cdots, 2^b - 2\}$, where b is the quantization depth indicating the number of bits per sample, and the step size between levels is

$$\Delta = \frac{2A}{2^b - 1}. \tag{11}$$

Note that the number of levels, $2^b - 1$, is odd in mid-tread uniform quantizers and $[-A, A]$ is the supported range beyond which the output is saturated.[4] Considering $q \triangleq 2^b$, the vector of the quantized signal is $\boldsymbol{g} \triangleq Q(\boldsymbol{x}^*) = (Q(x_1^*), Q(x_2^*), \cdots, Q(x_N^*)) \in \mathbb{F}_q^N$.

C. Syndrome Based Source Encoder

We propose two source encoders for the quantized sparse signal $\boldsymbol{g}$, based on the syndromes of RS (Fig. 1a) and BCH (Fig. 1b) codes.

Let us first consider the dual channel coding problem where the transmitter sends a codeword, $\boldsymbol{c} \in \mathbb{F}_q^N$, from the k_0-error-correcting RS code. If the channel changes k_0 symbols, then the received vector can be represented as $\boldsymbol{r} = \boldsymbol{c} + \boldsymbol{g}$, where $\boldsymbol{g} \in \mathbb{F}_q^N$ is the error vector with sparsity order k_0 (i.e., $\|\boldsymbol{g}\|_0 = k_0$), and the summation is in $\mathbb{F}_q$. Since the number of introduced errors is equal to the correcting capability of the code, the receiver can estimate the error vector $\boldsymbol{g}$ from the syndromes of the received signal, computed at the decoder.

Regarding our source coding problem, the main idea is to interpret the sparse signal as an error vector summed to the all-zeros codeword from a RS code. The quantized k_0-sparse source that is required to be encoded (i.e., $\boldsymbol{g}$) is compressed by calculating its syndrome vector at the encoder, through the parity check matrix of the k_0-error-correcting RS code. Consequently, the decoder can perfectly reconstruct $\boldsymbol{g}$ from the syndromes, because the sparsity order is equal to the correcting capability of the code. In particular, $\boldsymbol{g}$ corresponds to the smallest weight (sparsest) vector satisfying the syndromes.

More precisely, the syndrome (i.e., the compressed vector), $\boldsymbol{z} \in \mathbb{F}_q^{2\hat{k}_0}$, is computed by the source encoder as $\boldsymbol{z} = \boldsymbol{g}\,\boldsymbol{H}^T$ where all the operations are performed in $\mathbb{F}_q$, and

$$\boldsymbol{H} = \begin{bmatrix} 1 & \alpha & \alpha^2 & \cdots & \alpha^{N-1} \\ 1 & \alpha^2 & (\alpha^2)^2 & \cdots & (\alpha^2)^{N-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \alpha^{2\hat{k}_0} & (\alpha^{2\hat{k}_0})^2 & \cdots & (\alpha^{2\hat{k}_0})^{N-1} \end{bmatrix}$$

is the $2\,\hat{k}_0 \times N$ parity-check matrix for the $\hat{k}_0$-error-correcting RS code with $\hat{k}_0 \leq N/2$, $N = 2^b - 1$, and $b \geq 3$, while α is a primitive element in $\mathbb{F}_q$ [69], [70]. The resulting rate required for encoding the sparse vector using RS syndrome coding is

$$r_{\text{RS}} = 2\,b\,\frac{\hat{k}_0}{N} = 2\,\log_2(N+1)\,\frac{\hat{k}_0}{N} \qquad \text{[bits/sample]}. \tag{12}$$

A further compression gain can be achieved by separately sending the $\hat{k}_0$ quantized non-zero elements, then compressing the binary vector which determines their locations using the syndrome of a BCH code (Fig. 1b). In fact, since for BCH code the number of parity check bits $m(N, \hat{k}_0)$ satisfies $m \leq \hat{k}_0\,\log_2(N+1)$ [71, Section 6.1], the required rate is

$$\begin{aligned} r_{\text{BCH}} &= \frac{1}{N}\left(m + \hat{k}_0\,b\right) \\ &\leq \frac{\hat{k}_0}{N}\,\log_2(N+1) + b\,\frac{\hat{k}_0}{N} \qquad \text{[bits/sample]} \end{aligned} \tag{13}$$

where m is calculated from the design table of the BCH code, for a given sparsity order (i.e., error correcting capability) and dimension [71, Appendix C]. Note that the right hand side of (13) is equal to r_{RS} for $N = 2^b - 1$. Clearly, from (12) and (13) the rate of this scheme is upper-bounded by the RS based approach, but the non-zero values and the syndrome vector should be transmitted separately.

For the RS decoder, the sparse vector $\boldsymbol{g}$ can be recovered at the receiver from the syndrome vector $\boldsymbol{z}$ using Berlekamp's iterative algorithm [72]. Due to the minimum distance properties of the RS code and the maximum sparsity order of $\boldsymbol{g}$, the vector of the quantization numbers $\boldsymbol{g}$ is exactly recovered. Finally, the mapper $Q^{-1} : \{0, 1, \cdots, 2^b - 2\} \to \{i\,\Delta\}_{i=-2^{b-1}+1}^{2^{b-1}-1}$ reconstructs the quantized signal from its indexes, and the reconstructed signal is $\hat{\boldsymbol{x}} = Q^{-1}(\boldsymbol{g})$. For the BCH decoder, the binary location vector can be recovered using the Berlekamp's algorithm, and the non-zero entries are reconstructed from the quantized vector. The complexity of the syndrome based compression is low, as there are efficient devices for encoding and decoding the signal [70], [71].

We finally note that our scheme works in general with any channel code that allows to reconstruct the error pattern from the error syndrome, and in this regard other codes (LDPC, Polar, etc.) can also be used. The advantage in using RS or BCH is that we know by design their error correction capability, therefore the decoder can perfectly reconstruct the sparse vector from the syndrome, as long as the sparsity order is smaller than the maximum number of errors that the code can correct.

In the following, the proposed schemes in Fig. 1a and Fig. 1b will be referred to as RS based source coding (RSSC) and BCH based source coding (BCHSC), respectively.

IV. Operational Distortion-Rate with Known Support at the Encoder

In order to derive the operational distortion-rate (ODR) of the proposed schemes in Fig. 1, we start by assuming that the signal support is known only at the encoder, while the decoder does not have any side information. For example, in the high SNR regime, the excision filter can provide reliable support information with low probabilities of miss-detection and false-alarm. This assumption is relaxed in Section V, where the ODR with estimated support will be analyzed.

The ODR, $D(R)$, is a function that maps the expected rate R at which the system is working (i.e., the average number of bits required to describe a single source sample) to the average distortion [73]. More precisely, the distortion at the decoder output is

$$D(R) = \mathbb{E}_{\mathcal{S},\mathcal{W}}\left\{\left(\hat{\mathcal{S}}(\mathcal{S}, \mathcal{W}, R) - \mathcal{S}\right)^2\right\} = \int_{-\infty}^{\infty} \tau^2\, f_{\mathcal{T}}(\tau)\, d\tau \tag{14}$$

[4]For BU sources A is selected to match the support of $\mathcal{Y}$, while it can be chosen as a multiple of σ_{s} for BG [68].

where $\mathcal{S}$, $\mathcal{W}$, $\hat{\mathcal{S}}$, and $\mathcal{T} \triangleq \hat{\mathcal{S}} - \mathcal{S}$ are r.v.s representing the source output, additive noise, decoder output, and error due to both the quantization and the noise, respectively. Hence, finding the PDF of the error, $f_{\mathcal{T}}(\tau)$, is essential for deriving the ODR. With perfect denoising outside the signal support, the decoder output $\hat{\mathcal{S}}$ can be obtained from (1) and (5) as

$$\hat{\mathcal{S}} = Q^{-1}(Q(\mathcal{X}^*)) = Q^{-1}(Q(\mathcal{Z}\mathcal{Y} + \mathcal{Z}\mathcal{W})) = \mathcal{S} + \mathcal{T} \quad (15)$$

where $\mathcal{X}^*$ is the r.v. representing the filtered signal, $\mathcal{Z}$ is the Bernoulli r.v. defined in Section II, $\mathcal{W} \sim \mathcal{N}(0, \sigma_{\rm n}^2)$, and $Q^{-1}(Q(x))$ accounts for the scalar mid-tread uniform quantizer with bounded range described in Section III-B, i.e.,

$$Q^{-1}(Q(x)) = \begin{cases} \frac{\Delta}{2} - A, & x < -A \\ i\,\Delta, & (i - \frac{1}{2})\Delta \le x < (i + \frac{1}{2})\Delta \\ & \forall i \in I \triangleq \{i_{\min}, i_{\min}+1, ..., i_{\max}\} \\ A - \frac{\Delta}{2}, & x \ge A \end{cases} \quad (16)$$

where $i_{\max} = -i_{\min} = 2^{b-1} - 1 = A/\Delta - 1/2$.

The PDF of the error $\mathcal{T}$ can be written as

$$f_{\mathcal{T}}(\tau) = p\, f_{\mathcal{T}|\mathcal{Z}}(\tau|1) + (1-p)\,\delta(\tau) \quad (17)$$

where the error vanishes for $\mathcal{Z} = 0$, as mid-tread quantizers do not introduce distortion to the zero valued source symbol. Hence, only the non-zero elements are subject to distortion. Given that the source symbol is non-zero and conditioned on the noise $\mathcal{W} = w$, the function $g(\cdot)$ which maps the source $\mathcal{Y}$ into the error $\mathcal{T}$ can be written from (15) and (16) as

$$g(y) = \begin{cases} \Delta/2 - A - y, & y < -w - A \\ i\Delta - y, & \left(i - \frac{1}{2}\right)\Delta - w \le y < \left(i + \frac{1}{2}\right)\Delta - w \\ A - \Delta/2 - y, & y \ge A - w. \end{cases}$$

The PDF of the error $f_{\mathcal{T}|\mathcal{Z},\mathcal{W}}(\tau|1, w)$ can be represented from [74, (5-16)] as

$$f_{\mathcal{T}|\mathcal{Z},\mathcal{W}}(\tau|1, w) = \sum_j \frac{f_{\mathcal{Y}}(\widetilde{y}_j)}{|g'(\widetilde{y}_j)|}$$

where $\{\widetilde{y}_1, \widetilde{y}_2, \cdots\}$ indicate the solutions of $g(y) = \tau$ and $g'(\widetilde{y}_j)$ is the derivative of $g(y)$ at $\widetilde{y}_j$. Considering that the solutions depend on the value of τ and that $|g'(\widetilde{y}_j)| = 1$, then

$$\begin{aligned} f_{\mathcal{T}|\mathcal{Z},\mathcal{W}}(\tau|1, w) = & \sum_{i=i_{\min}}^{i_{\max}} f_{\mathcal{Y}}(i\Delta - \tau)\, \mathrm{rect}((\tau - w)/\Delta) \\ & + f_{\mathcal{Y}}(\Delta/2 - A - \tau)\, u(\tau - w - \Delta/2) \\ & + f_{\mathcal{Y}}(A - \Delta/2 - \tau)\, u(w - \tau - \Delta/2) \end{aligned} \quad (18)$$

where $\mathrm{rect}(x) = 1$ for $x \in [-1/2, 1/2]$ and zero otherwise, and $u(x)$ is the unit step function. Therefore, the PDF of the error can be derived by averaging over $\mathcal{W}$, i.e.,

$$\begin{aligned} f_{\mathcal{T}|\mathcal{Z}}(\tau|1) = & \mathbb{E}_{\mathcal{W}}\left\{f_{\mathcal{T}|\mathcal{Z},\mathcal{W}}(\tau|1, w)\right\} = \phi(\tau) \\ & \times f_{\mathcal{Y}}(A - \Delta/2 - \tau) + \Phi(\tau) \sum_{i=i_{\min}}^{i_{\max}} f_{\mathcal{Y}}(i\Delta - \tau) \\ & + \phi(-\tau)\, f_{\mathcal{Y}}(\Delta/2 - A - \tau) \end{aligned} \quad (19)$$

where

$$\Phi(x) \triangleq \frac{1}{2}\left[\mathrm{erf}\left(\frac{x + \Delta/2}{\sqrt{2}\,\sigma_{\rm n}}\right) - \mathrm{erf}\left(\frac{x - \Delta/2}{\sqrt{2}\,\sigma_{\rm n}}\right)\right]$$

and

$$\phi(x) \triangleq \frac{1}{2}\left[1 - \mathrm{erf}\left(\frac{x + \Delta/2}{\sqrt{2}\,\sigma_{\rm n}}\right)\right].$$

Altogether, (17)-(19) give the error density function for an arbitrary distribution of the non-zero elements, $f_{\mathcal{Y}}(y)$, as

$$\begin{aligned} f_{\mathcal{T}}(\tau) = (1-p)\,\delta(\tau) + p\Bigg[& \underbrace{\phi(\tau)\, f_{\mathcal{Y}}(A - \Delta/2 - \tau)}_{f_{\rm L}(\tau)} + \\ & \underbrace{\phi(-\tau)\, f_{\mathcal{Y}}\left(\frac{\Delta}{2} - A - \tau\right)}_{f_{\rm U}(\tau)} + \underbrace{\Phi(\tau) \sum_{i=i_{\min}}^{i_{\max}} f_{\mathcal{Y}}(i\Delta - \tau)}_{f_{\rm M}(\tau)}\Bigg] \end{aligned} \quad (20)$$

where the functions $f_{\rm L}(\tau)$, $f_{\rm U}(\tau)$ are related to the lower and upper saturation errors, and $f_{\rm M}(\tau)$ accounts for the error in the middle levels. We now specialize (20) for two important distributions of the source.

For BU sources, from the definitions of $f_{\rm L}(\tau)$ and $f_{\rm U}(\tau)$ in (20) we have

$$f_{\rm L}(\tau) = \begin{cases} \frac{1}{2A}\,\phi(\tau), & -\frac{\Delta}{2} \le \tau < 2A - \frac{\Delta}{2} \\ 0, & \text{otherwise} \end{cases} \quad (21)$$

$$f_{\rm U}(\tau) = \begin{cases} \frac{1}{2A}\,\phi(-\tau), & -2\,A + \frac{\Delta}{2} \le \tau < \frac{\Delta}{2} \\ 0, & \text{otherwise.} \end{cases} \quad (22)$$

Regarding $f_{\rm M}(\tau)$, the first step is to compute the summation as a function of τ. We define the number of non-zero terms as $N(\tau) = \{\#\, i \text{ s.t. } f_{\mathcal{Y}}(i\Delta - \tau) \neq 0\}$. After some manipulation, it results that

$$\begin{aligned} N(\tau) = & \min\{\lfloor (A + \tau)/\Delta \rfloor, A/\Delta - 1/2\} \\ & - \max\{\lceil (\tau - A)/\Delta \rceil, 1/2 - A/\Delta\} + 1. \end{aligned}$$

Therefore, the function $f_{\rm M}(\tau)$ can be written as

$$f_{\rm M}(\tau) = \frac{\Phi(\tau)}{2A} \cdot \begin{cases} \left\lfloor \frac{\tau}{\Delta} + \frac{1}{2} \right\rfloor + \frac{2A}{\Delta}, & -2\,A + \frac{\Delta}{2} \le \tau < -\frac{\Delta}{2} \\ \frac{2\,A}{\Delta}, & -\frac{\Delta}{2} \le \tau < \frac{\Delta}{2} \\ \left\lfloor \frac{1}{2} - \frac{\tau}{\Delta} \right\rfloor + \frac{2A}{\Delta}, & \frac{\Delta}{2} \le \tau < 2\,A - \frac{\Delta}{2} \\ 0, & \text{otherwise.} \end{cases} \quad (23)$$

By substituting (21), (22), and (23) in (20), the PDF of the error can be written as

$$\begin{aligned} f_{\mathcal{T}}^{\rm BU}(\tau) = & (1-p)\delta(\tau) + \frac{p}{2A} \\ & \times \begin{cases} 1 + \Phi(\tau)\left(\frac{2A}{\Delta} - 1\right), & |\tau| < \frac{\Delta}{2} \\ \Phi(\tau)\left(\left\lfloor \frac{1}{2} - \frac{|\tau|}{\Delta} \right\rfloor + \frac{2A}{\Delta}\right) + \phi(|\tau|), & \frac{\Delta}{2} \le |\tau| < 2A - \frac{\Delta}{2} \\ 0, & \text{otherwise} \end{cases} \end{aligned} \quad (24)$$

where $\phi(-\tau) + \phi(\tau) = 1 - \Phi(\tau)$, and $f_{\mathcal{T}}^{\rm BU}(\tau)$ is an even function. The PDF (24) is a mixed distribution with a Dirac delta function at zero with weight $1 - p$ representing a perfect recovery of the zero elements at the decoder.

For BG sources, the PDF of noise plus quantization error, $f_{\mathcal{T}}^{\text{BG}}(\tau)$, can be obtained, as before, by substituting the PDF of the Gaussian distribution into (20). For this source, further simplification of (20) seems quite involved.

After deriving the PDF of error, we now want to find the ODR for BU and BG sources. First, the relation between the distortion and the quantization step size, Δ, is analyzed. The mean squared distortion from (14) is

$$D^{\text{BU/BG}}(\Delta) = \int_{-\infty}^{\infty} \tau^2 \, f_{\mathcal{T}}^{\text{BU/BG}}(\tau) \, d\tau. \tag{25}$$

The exact distortion can be calculated by simple numerical integration. For closed-form analysis, we report an approximation of the distortion for BU sources in Appendix A by approximating the floor function in (24) as $\lfloor x \rfloor \simeq x - 1/2$ and calculating the integration (25). This approximation will be used in the numerical results. A simpler but less accurate expression valid in the low-distortion region, i.e., $A \gg \sigma_{\text{n}}$ and $\Delta \ll A$ or $\Delta \ll \sigma_{\text{s}}$, is given by $D^{\text{BU/BG}}(\Delta) \simeq \left(\frac{\Delta^2}{12} + \sigma_{\text{n}}^2\right) p$, assuming the quantization error uniformly distributed and independent from both the source and the additive noise [75].

In order to find the ODR, the distortion is expressed in terms of the average rate. For RSSC, the expected rate is calculated from (12) as

$$R = \mathbb{E}\left[\mathcal{R}_{\text{RS}}\right]/N = 2\,b\,\mathbb{E}\left[\mathcal{K}_0\right]/N = 2\,b\,p \text{ [bits/sample]} \tag{26}$$

where $\mathcal{K}_0$ is a binomial r.v. with mean pN representing the number of non-zero elements. Similarly, the average rate for BCHSC can be found from (13). Finally, the ODR can be written from (11), (13), (25), and (26) as

$$D_{\text{RS/BCH}}^{\text{BU/BG}}(R) = D^{\text{BU/BG}}\left(\frac{2A}{2^{b(R)} - 1}\right) \tag{27}$$

where

$$b(R) = \lfloor R/(2\,p) \rfloor, \text{ and } b(R) = \lfloor (N\,R - m(N,\,pN))/(N\,p) \rfloor$$

for RSSC and BCHSC.

As a benchmark for noiseless BU and BG sources, the ODR of the uniform quantizer followed by the optimal entropy encoder are derived in Appendix B.

V. Operational Distortion-Rate with Support Estimation

In this section, the ODR of the proposed schemes for BG and BU sources is analyzed when the signal support is estimated by thresholding. The optimal threshold, η, that minimizes the ODR is then derived. In fact, if the signal support is not estimated perfectly, i.e., $\text{P}_{\text{MD}} > 0$ and $\text{P}_{\text{FA}} > 0$, there are three causes of distortion:

1) over-estimation distortion when some noise samples are not filtered (due to false-alarm);
2) under-estimation distortion when some signal samples are filtered (due to miss-detection);
3) distortion due to noise and quantization error in the correctly detected samples.

Taking into account these contributions, the overall distortion for a given threshold η is

$$\begin{aligned} D(\Delta,\eta) &= \sum_{i=0}^{1}\sum_{j=0}^{1} \mathbb{P}_{\mathcal{D}|\mathcal{H}}\left\{\mathcal{D}_j|\mathcal{H}_i\right\} \mathbb{E}\left\{\left(\mathcal{S} - \hat{\mathcal{S}}\right)^2 \middle| \mathcal{H}_i, \mathcal{D}_j\right\} \\ &\times \mathbb{P}_{\mathcal{H}}\left\{\mathcal{H}_i\right\} = (1-p)\,\text{P}_{\text{FA}}(\eta)\mathbb{E}\left\{\left[Q^{-1}\left(Q\left(\mathcal{W}\right)\right)\right]^2 \middle| \mathcal{W}^2 \geq \eta\right\} \\ &+ p\,\text{P}_{\text{MD}}(\eta)\,\mathbb{E}\left\{\mathcal{Y}^2 \middle| (\mathcal{Y}+\mathcal{W})^2 < \eta\right\} + p\,\left(1 - \text{P}_{\text{MD}}(\eta)\right) \\ &\times \mathbb{E}\left\{\left[Q^{-1}\left(Q\left(\mathcal{Y}+\mathcal{W}\right)\right) - \mathcal{Y}\right]^2 \middle| (\mathcal{Y}+\mathcal{W})^2 \geq \eta\right\} \end{aligned} \tag{28}$$

where $(i = 0, j = 1)$, $(i = 1, j = 0)$, and $(i = 1, j = 1)$ correspond to over-estimation distortion 1), under-estimation distortion 2), and the distortion due to quantization and noise error 3), respectively. Closed-form expressions for these quantities are provided below.

Fo the over-estimation distortion, the encoder treats noise samples as if they belong to the sparse signal. The probability mass function of the quantized noise given that over-estimation event occurred is

$$\begin{aligned} &\mathbb{P}\left\{Q^{-1}\left(Q\left(\mathcal{W}\right)\right) = i\Delta | \mathcal{W}^2 \geq \eta\right\} = \left(2\,\text{erfc}\left(\sqrt{\frac{\eta}{2\sigma_{\text{n}}^2}}\right)\right)^{-1} \\ &\times \begin{cases} 2\,\text{erfc}\left(\frac{\sqrt{\eta}}{\sqrt{2}\sigma_{\text{n}}}\right) - 2\,\text{erfc}\left(\frac{\Delta\left(|i| + \frac{1}{2}\right)}{\sqrt{2}\sigma_{\text{n}}}\right), \\ \qquad \text{for } |i| = i_{\min}^{*} \text{ and } \eta \leq \Delta^2/4 \\ \text{erfc}\left(\frac{\sqrt{\eta}}{\sqrt{2}\sigma_{\text{n}}}\right) - \text{erfc}\left(\frac{\Delta\left(|i| + \frac{1}{2}\right)}{\sqrt{2}\sigma_{\text{n}}}\right), \\ \qquad \text{for } |i| = i_{\min}^{*} \text{ and } \Delta^2/4 < \eta < (A-\Delta)^2 \\ \text{erfc}\left(\frac{\Delta\left(|i| - \frac{1}{2}\right)}{\sqrt{2}\sigma_{\text{n}}}\right) - \text{erfc}\left(\frac{\Delta\left(|i| + \frac{1}{2}\right)}{\sqrt{2}\sigma_{\text{n}}}\right), \\ \qquad \text{for } i_{\min}^{*} + 1 \leq |i| < i_{\max} \\ \text{erfc}\left(\frac{\max\{A - \Delta, \sqrt{\eta}\}}{\sqrt{2}\sigma_{\text{n}}}\right), \qquad \text{for } |i| = i_{\max} \end{cases} \end{aligned}$$

where $i_{\min}^{*} \triangleq \left\lceil \sqrt{\eta}/\Delta - 1/2 \right\rceil$, for each integer i such that $\min\left\{i_{\min}^{*}, i_{\max}\right\} \leq |i| \leq i_{\max}$. The over-estimation distortion can be given now for BG and BU sources as

$$\begin{aligned} \mathbb{E}\left\{\left[Q^{-1}\left(Q\left(\mathcal{W}\right)\right)\right]^2 | \mathcal{W}^2 \geq \eta\right\} = &\sum_{i=\min\left\{i_{\min}^{*}, i_{\max}\right\}}^{i_{\max}} (i\Delta)^2 \\ &\times 2\,\mathbb{P}\left\{Q^{-1}\left(Q\left(\mathcal{W}\right)\right) = i\Delta | \mathcal{W}^2 \geq \eta\right\}. \end{aligned} \tag{29}$$

In the low distortion region, i.e., $b \geq 6$ bits/sample, $A \gg \sigma_{\text{n}}$, and $\eta < (A - \Delta)^2$, equation (29) can be approximated in closed-form considering a continuous un-quantized noise model. In fact, the distribution of $|\widetilde{\mathcal{W}}|$, where $\widetilde{\mathcal{W}} \triangleq \left\{\mathcal{W} | \mathcal{W}^2 \geq \eta\right\}$, follows that of a truncated Gaussian, i.e.,

$|\widetilde{\mathcal{W}}| \sim \mathcal{N}(0, \sigma_{\rm n}^2, \sqrt{\eta}, \infty)$; hence

$$\mathbb{E}\left\{\left[Q^{-1}\left(Q\left(\mathcal{W}\right)\right)\right]^2 |\mathcal{W}^2 \geq \eta\right\} \simeq \mathbb{E}\left\{\widetilde{\mathcal{W}}^2\right\} = \sigma_{\rm n}^2 + \sigma_{\rm n} e^{-\frac{\eta}{2\sigma_{\rm n}^2}} \sqrt{\frac{2\eta}{\pi}} \frac{1}{\operatorname{erfc}\left(\sqrt{\frac{\eta}{2\sigma_{\rm n}^2}}\right)}.$$

For the under-estimation distortion, we have

$$\mathbb{E}\left\{\mathcal{Y}^2 | (\mathcal{Y}+\mathcal{W})^2 < \eta\right\} = \frac{\mathbb{E}\left\{\mathcal{Y}^2 \cdot \mathbf{1}_{\{(y+w)^2<\eta\}}\right\}}{\mathbb{P}\left\{(\mathcal{Y}+\mathcal{W})^2 < \eta\right\}} = \frac{1}{\mathrm{P}_{\rm MD}(\eta)} \iint_{\rm S} y^2 f_{\mathcal{Y}}(y) f_{\mathcal{W}}(w)\, dy\, dw \quad (30)$$

where $\mathrm{S} \triangleq \left\{(y,w) \in \mathbb{R}^2 \text{ such that } (y+w)^2 < \eta\right\}$. Clearly, (30) depends on the distribution of the source. Regarding BU, we have

$$\mathbb{E}\left\{\mathcal{Y}^2 | (\mathcal{Y}+\mathcal{W})^2 < \eta\right\} = \frac{1}{6A\,\mathrm{P}_{\rm MD}(\eta)} \Bigg(\sqrt{\frac{2}{\pi}} \sigma_{\rm n} e^{-\frac{(A+\sqrt{\eta})^2}{2\sigma_{\rm n}^2}} \times \Big[-e^{\frac{2A\sqrt{\eta}}{\sigma_{\rm n}^2}} \left(A^2 + A\sqrt{\eta} + \eta + 2\sigma_{\rm n}^2\right) + A^2 - A\sqrt{\eta} + \eta + 2\sigma_{\rm n}^2 \Big] + \operatorname{erf}\left(\frac{A-\sqrt{\eta}}{\sqrt{2}\sigma_{\rm n}}\right) \left[\sqrt{\eta}\left(\eta + 3\,\sigma_{\rm n}^2\right) - A^3\right] + \operatorname{erf}\left(\frac{A+\sqrt{\eta}}{\sqrt{2}\sigma_{\rm n}}\right) \times \left[A^3 + \sqrt{\eta}\left(\eta + 3\,\sigma_{\rm n}^2\right)\right] \Bigg) \quad (31)$$

where $\mathrm{P}_{\rm MD}(\eta)$ is given by (10). For BG, from (7) and (30) we have

$$\mathbb{E}\left\{\mathcal{Y}^2 | (\mathcal{Y}+\mathcal{W})^2 < \eta\right\} = \mathbb{E}_{\mathcal{W}} \Bigg\{ \frac{1}{2}\sigma_{\rm s}^2 \left[\operatorname{erf}\left(\frac{\sqrt{\lambda}+w}{\sqrt{2}\sigma_{\rm s}}\right) - \operatorname{erf}\left(\frac{w-\sqrt{\lambda}}{\sqrt{2}\sigma_{\rm s}}\right)\right] + \sqrt{\frac{2}{\pi}} \sigma_{\rm s} e^{-\frac{\lambda+w^2}{2\sigma_{\rm s}^2}} \left[w \sinh\left(\frac{\sqrt{\lambda} w}{\sigma_{\rm s}^2}\right) - \sqrt{\lambda} \times \cosh\left(\frac{\sqrt{\lambda} w}{\sigma_{\rm s}^2}\right)\right] \Bigg\} = \sigma_{\rm s}^2 - \frac{\sigma_s^4}{\sigma_x^3} \sqrt{\frac{2\,\lambda}{\pi}} \frac{e^{-\frac{\lambda}{2\sigma_{\rm x}^2}}}{\operatorname{erf}\left(\sqrt{\frac{\lambda}{2\sigma_{\rm x}^2}}\right)}. \quad (32)$$

The distortion due to the noise and quantization for the correctly detected samples can be approximated from (25) when $\mathbb{P}\left\{(\mathcal{Y}+\mathcal{W})^2 \geq \eta\right\} \triangleq 1 - \mathrm{P}_{\rm MD}(\eta) \simeq 1$ as

$$\mathbb{E}\left\{\left[Q^{-1}\left(Q\left(\mathcal{Y}+\mathcal{W}\right)\right) - \mathcal{Y}\right]^2 \Big| (\mathcal{Y}+\mathcal{W})^2 \geq \eta\right\} = \frac{\mathbb{E}\left\{\left[Q^{-1}\left(Q\left(\mathcal{Y}+\mathcal{W}\right)\right) - \mathcal{Y}\right]^2 \cdot \mathbf{1}_{\{(y+w)^2\geq\eta\}}\right\}}{\mathbb{P}\left\{(\mathcal{Y}+\mathcal{W})^2 \geq \eta\right\}} \simeq \mathbb{E}\left\{\left[Q^{-1}\left(Q\left(\mathcal{Y}+\mathcal{W}\right)\right) - \mathcal{Y}\right]^2\right\} = D^{\rm BG/BU}(\Delta)/p. \quad (33)$$

The probability of miss-detection in (7) and (10) can be small for high SNR, with the proper design of η, leading to an increased accuracy for the approximation (33).

Finally, the distortion is approximated by substituting (29), (33), and (31) or (32) into (28).

The expected rate for RSSC considering the imperfect support estimation is

$$R_{\rm RS}(b,\eta) = \frac{2\,b}{N} \mathbb{E}\left\{\hat{\mathcal{K}}_0\right\} \quad (34)$$

where $\hat{\mathcal{K}}_0$ is a r.v. representing the number of non-zero elements at the excision filter output with

$$\mathbb{E}\left\{\hat{\mathcal{K}}_0\right\} = N\,p\,\left[1 - \mathrm{P}_{\rm MD}(\eta)\right] + N\,(1-p)\,\mathrm{P}_{\rm FA}(\eta).$$

Similarly, the average rate for BCH can be derived as $R_{\rm BCH}(b,\eta) = \mathbb{E}\left\{b\,\hat{\mathcal{K}}_0 + m(N,\,\hat{\mathcal{K}}_0)\right\}/N$.

The corresponding ODR, accounting for imperfect support recovery using thresholding, can be approximated accurately from (28), (29), (33), (31), (32), and (34) as

$$D_{\rm RS/BCH}(R,\eta) = D\left(\frac{2A}{2^{b(R)}-1}, \eta\right) \quad (35)$$

where

$$b(R) = \left\lfloor \frac{R}{2\,p\,\left[1 - \mathrm{P}_{\rm MD}(\eta)\right] + 2\,(1-p)\,\mathrm{P}_{\rm FA}(\eta)} \right\rfloor$$

and

$$b(R) = \left\lfloor \frac{NR - m\left(N,\,\mathbb{E}\left\{\hat{\mathcal{K}}_0\right\}\right)}{\mathbb{E}\left\{\hat{\mathcal{K}}_0\right\}} \right\rfloor$$

for RSSC and BCHSC, respectively.

The threshold η is crucial for the encoder performance, because as η increases the rate decreases, but at the same time the distortion could increase due to missing significant samples. Therefore, the optimal values of η and b such that the distortion is minimized for a fixed rate R can be calculated from (35) as

$$\left(\widehat{\eta}, \widehat{b}\right) = \underset{\eta\in\mathbb{R}^+,\, b\in\mathbb{N}^+}{\arg\min}\; D\left(\frac{2A}{2^b-1}, \eta\right) \quad \text{subject to} \quad R_{\rm RS/BCH}(b,\eta) = R\,. \quad (36)$$

Note that finding the optimal threshold requires simple numerical optimization.

VI. Operational Distortion-Energy Performance

The ODR is not always sufficient to fully compare various encoders, exhibiting different computational complexity and power consumption, e.g., the proposed solution and the CS scheme in Fig. 2. In fact, for many IoT applications the energy spent during acquisition, compression, and transmission is a critical aspect, and the operational distortion-energy (ODE) represents another key performance indicator [76], [77].

Let us start with a typical power model similar to that in [78] and [79, Section IV], where the total energy consumption in the sensor node can be calculated as a sum of the computational and communication energy $E_t = E_{\rm comp} + E_{\rm comm}$. The computational energy $E_{\rm comp} = E_{\rm acq} + E_{\rm sp} + E_{\rm bck}$ accounts for the acquisition, signal processing, and background energy, while the communication energy $E_{\rm comm} = E_{\rm tx} + E_{\rm rx}$ represents the consumption during transmission and reception.

For the proposed BCHSC, the signal is first sparsified: for a concrete example we assume that the DCT is used.[5] Then, it is compressed using the encoder in Fig. 1b. The energy spent in the acquisition is given by $E_{\rm acq} = N p_{\rm acq} T_{\rm ins}$, where $p_{\rm acq}$ is the

[5]Real signals from WSNs may not be sparse in time, but rather in other domains such as DCT.

power spent for acquisition and T_{ins} is the instruction execution time. The energy consumed for signal processing due to vector reading, DCT [80], support detection by thresholding, and syndrome calculation is

$$E_{\text{sp}} = \overbrace{N\epsilon_{\text{rd}}}^{\text{Reading}} + \overbrace{N(\epsilon_{\text{cmp}} + \epsilon_{\text{wr}})}^{\text{Thresholding}} + \overbrace{\left(N\log_2(N) - \frac{3N}{2} + 4\right)\epsilon_{\text{mul}} + \left(\frac{3N}{2}\left[\log_2(N) - 1\right] + 2\right)\epsilon_{\text{add}}}^{\text{DCT}} + \overbrace{m\left(N, \mathbb{E}\left\{\hat{\mathcal{K}}_0\right\}\right)\left[\epsilon_{\text{add}}\left(\mathbb{E}\left\{\hat{\mathcal{K}}_0\right\} - 1\right) + \mathbb{E}\left\{\hat{\mathcal{K}}_0\right\}\epsilon_{\text{mul}} + \epsilon_{\text{wr}}\right]}^{\text{Syndrome Calculation}} \quad (37)$$

with energy per instruction consumed during reading, addition, multiplication, comparing, and writing denoted by ϵ_{rd}, ϵ_{add}, ϵ_{mul}, ϵ_{cmp}, and ϵ_{wr}, respectively. Similarly to [79, (6)], the background power consumption can be calculated as

$$E_{\text{bck}} = p_{\text{bck}}\, T_{\text{ins}} \times \overbrace{\left[2.5\, N\log_2(N) + 2\,\mathbb{E}\left\{\hat{\mathcal{K}}_0\right\} m\left(N, \mathbb{E}\left\{\hat{\mathcal{K}}_0\right\}\right) + 6\right]}^{n_{\text{o}}} \quad (38)$$

where p_{bck} is the background power consumption, and n_{o} is the number of operations performed. The energy consumed during transmission can be found as

$$E_{\text{comm}}(b) = \mathbb{E}\left\{b\,\hat{\mathcal{K}}_0 + m(N, \hat{\mathcal{K}}_0)\right\}\left[(1 + h_{\text{oh}})\, E_{\text{tx}} + h_{\text{ack}} E_{\text{rx}}\right] \quad (39)$$

where h_{oh} and h_{ack} represent the percentage of overhead and acknowledgment with respect to the payload length. Then, the overall consumed energy from (37), (38), and (39) is

$$E_{\text{BCH}}(b) = E_{\text{acq}} + E_{\text{sp}} + E_{\text{bck}} + E_{\text{comm}}(b). \quad (40)$$

Now, the ODE for the syndrome encoding is derived from (28) and (40) as

$$D_{\text{BCH}}(E) = D\left(\frac{2A}{2^{b(E)} - 1}, \eta\right) \quad (41)$$

where $b(E)$ can be found from (40). Similarly, the ODE of RSSC can be calculated.

Regarding the CS scheme in Fig. 2, the total consumed energy is

$$E_{\text{CS}}(b) = E_{\text{acq}} + N\epsilon_{\text{rd}} + M\left[(N-1)\epsilon_{\text{add}} + N\epsilon_{\text{mul}}\right] + M\epsilon_{\text{wr}} + p_{\text{bck}} T_{\text{ins}}\left(N + 2NM\right) + Mb\left[(1 + h_{\text{oh}})\, E_{\text{tx}} + h_{\text{ack}} E_{\text{rx}}\right] \quad (42)$$

where it has been assumed that the power consumption of the analog CS encoder matches its digital counterpart. For the ODE of CS, the required energy can be calculated from (42), and the distortion can be found through Monte Carlo simulation, as illustrated in the numerical results.[6]

[6]Note that the analytical performance of CS encoders is beyond the scope of this paper. For more relevant analysis, one may refer to [6], [46], [52] and the references therein.

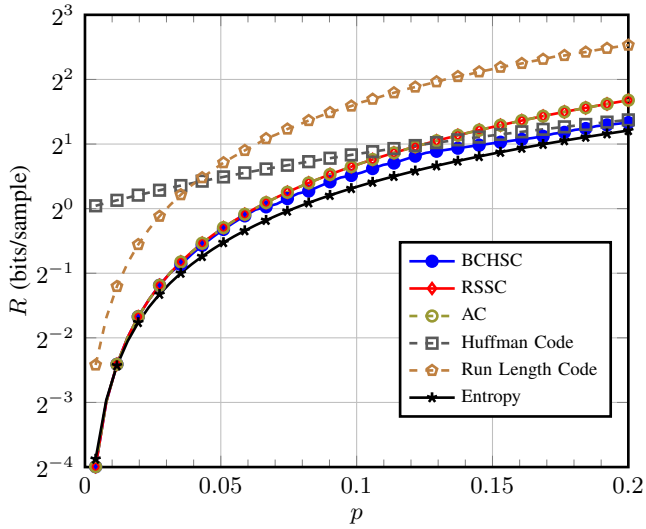


Fig. 3: Average rate in bits/sample as a function of the average sparsity ratio p, for BU sources quantized with $b = 8$ bits/sample.

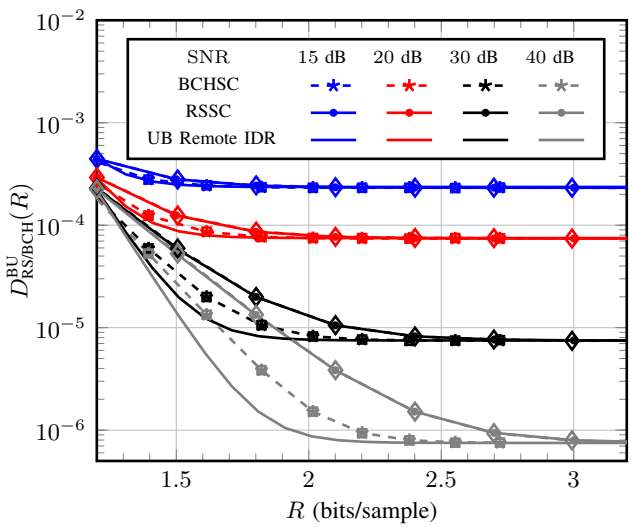


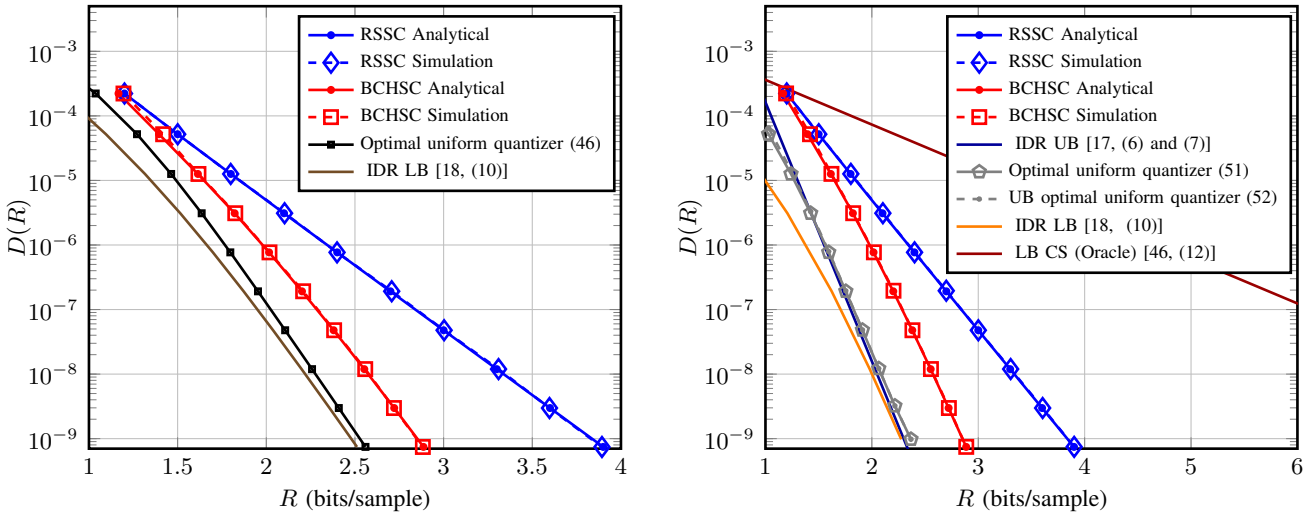
Fig. 4: The ODR of BCHSC (dashed) and RSSC (solid) and the remote IDR bound, with $b \in \{4, 5, ..., 12\}$. The approximation (44) is used for the distortion. The analytical performance ($*, \bullet$) coincides with simulation ($\square$, $\diamond$).

VII. Numerical Results

A. Performance Analysis

In this section, numerical results and Monte Carlo simulations are presented to illustrate the performance of the proposed RSSC and BCHSC schemes. The SNR is defined as the signal power to the additive noise power at the encoder, i.e., $\text{SNR} \triangleq \mathbb{E}\left\{\boldsymbol{\mathcal{S}}\boldsymbol{\mathcal{S}}^T\right\}/\mathbb{E}\left\{\boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{W}}^T\right\}$. In all numerical results $p = 0.15$, $A = 1$, and $N = 2^b - 1$, unless otherwise stated.

We first compare the syndrome encoding schemes with some classical variable length encoders like run length code and Huffman code applied on each source sample. In Fig. 3, we report the average rate needed to encode the noiseless



(a) Bernoulli-uniform.

(b) Bernoulli-Gaussian.

Fig. 5: The ODR of BCHSC and RSSC along with upper bounds on the IDR, with $A = 1$, $\sigma_{\rm s} = 1/5$, $b \in \{2, 3, ..., 13\}$, and $\mathrm{SNR} = \infty$.

quantized BU source using RSSC (12), BCHSC (13), AC, Huffman encoder, and run length encoding, plus the Shannon's lower bound (45), as a function of the average sparsity ratio p, for $b = 8$ bits/sample. It is noted that the rates indicated by both RSSC and AC are coincident, while the BCHSC can achieve a higher compression gain (up to 15% compared to RSSC). Also, the proposed approaches are superior to run length coding and Huffman coding for $p < 0.13$, with a small gap between the entropy bound and syndrome encoders for low sparsity ratios. As expected, both Huffman and run length codes perform poorly for sparse sources.

Fig. 4 illustrates the ODR of RSSC and BCHSC, (27), for BU sources. These ODRs are compared with the upper bound on the remote IDR

$$D(R) \leq \left(\frac{\sigma_{\rm s}^2}{\sigma_{\rm s}^2/\sigma_{\rm n}^2 + 1} + \frac{\sigma_{\rm s}^2}{\sigma_{\rm n}^2/\sigma_{\rm s}^2 + 1} 2^{-2[R - H_{\rm bern}(p)]/p} \right) p, \quad \text{for } R > H_{\rm bern}(p)$$

obtained by extending the upper bound on the remote IDR for continuous non-sparse sources in [21, (10)] to account for BG and BU sources, considering that the support is optimally encoded with a rate indicated by the Bernoulli source entropy $H_{\rm bern}(p) \triangleq -p \log_2(p) - (1-p)\log_2(1-p)$. It is clear that $D_{\rm BCH}(R) \leq D_{\rm RS}(R)$, since $R_{\rm BCH} \leq R_{\rm RS}$ from (12) and (13). The proposed schemes perform close to the bound on the remote IDR for low rates. Also, the distortion tends approximately to $p\,\sigma_{\rm n}^2$ as $b \to \infty$, exhibiting a floor at high rates. For SNRs equal to $30\,$dB and $40\,$dB, the distortion significantly decreases by increasing the rate from 1.2 to 2 bits/sample. On the other hand, the distortion saturates for low SNRs where it is dominated by the additive noise component.

In Fig. 5, we compare the ODR of proposed BCHSC and RSSC with that of the following: the optimal ODR of uniform quantizers for BU and BG sources, derived in Appendix B; the lower bound on the IDR for BU and BG sources in [18, (20)]; the converse upper bound on the IDR for BG sources in [17, (6) and (7)]; the asymptotic lower bound on the ODR of a CS based encoder with uniform quantization derived in [46, (12)], assuming that the support is known at the decoder. In this case, $D_{\rm CS}(R) = \frac{p}{\mu - 1}\,\sigma_{\rm s}^2\, 2^{-2\,R/(\mu\,p)}$ where $\mu \triangleq M/k_0$.[7] We can see that the gap between the BCHSC and optimal uniform quantizers in (46) and (51) is small for BU and BG sources, especially in the low-rate region. Also, the upper bound (52) on the ODR of the optimal uniform quantizer for BG sources is tight. Although the lower bound on the ODR of CS in [46] does not account for the rate budget required to convey the support information to the decoder, the proposed schemes are superior for $R > 1.2$ bits/sample. It is noted also that the lower and the upper bounds on the IDR for BG in [18] and [17] are tight for high rates.

The ODR of RSSC for BU with estimated support, (35)), is depicted in Fig. 6 and compared with the case of known support only at the encoder (genie-aided), (27). The locations of the non-zero elements are estimated by the following methods: the excision filter (6) using the optimal threshold $\widehat{\eta}$, (36); the excision filter (6) with a threshold designed through the Neyman-Pearson criterion for $\mathrm{P_{FA}} = 0.01$; the GIC (5) with $\nu = 5$ for $\mathrm{SNR} = 15\,$dB, and $\nu = 6$ for $\mathrm{SNR} = 30\,$dB. The simulation results agree with the analytical expressions for the ODR. The proposed optimal threshold achieves a better performance compared to those based on constant false-alarm and GIC, especially for low rate and high SNR. For example, the optimal threshold achieves about 30% reduction in the distortion, for $R = 1.8$ bits/sample and $\mathrm{SNR} = 30\,$dB. Moreover, the GIC performs well in the high rate region. On this point we would like to remark that the GIC is completely blind, while thresholding requires the knowledge of the SNR. Note that in the low-rate region the optimal threshold is superior to the genie-aided method with

[7]Note that for the bound on the ODR of CS, we used $M = 2\,k_0\,\log(e/p)$ [14, (9.24)].

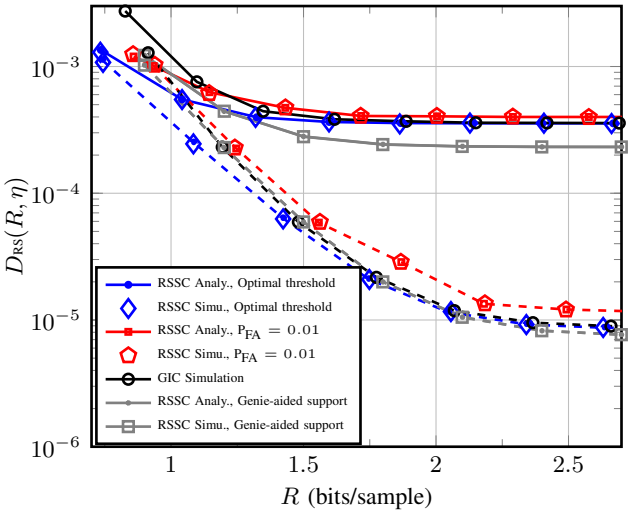


Fig. 6: The ODR for noisy BU sources with estimated support through the thresholding filter (6) using the optimal threshold (36) and Neyman-Pearson criteria (8), and through the GIC (5), for $\text{SNR} = 15\,\text{dB}$ (solid) and $\text{SNR} = 30\,\text{dB}$ (dashed).

perfect support knowledge. The rationale behind this is that even when the signal support is known at the encoder, it is not optimal to spend the rate budget to encode all the non-zero samples along with their locations. In fact, the increase in the distortion due to missing a small source sample can be compensated by the corresponding rate reduction. On the contrary, the support knowledge increases the performance for higher rates, where there is enough rate budget for coding all source samples.

B. Case Study: IoT for Environmental Monitoring

In this section, the proposed BCHSC scheme in Fig. 1 is utilized to encode signals acquired by a WSN for landslides monitoring [81], and compared with the CS encoder shown in Fig. 2. In particular, we encoded real data (temperature, humidity, and rain level) collected for over a year with a rate of one sample every 15 min. Thus, $40 \cdot 10^3$ samples from each sensor are grouped into vectors $\boldsymbol{x}_{\text{c}} \in \mathbb{R}^N$ with length $N = 127$.

The signal is first represented in the DCT domain, then the support of the significant samples is estimated using an excision filter with a threshold designed such that a predefined fraction, $\lambda \in [0, 1]$, of the total signal energy is preserved. For the humidity data, only $\hat{k}_0 = 25$ significant coefficients (out of $N = 127$) are sufficient to guarantee that more than 99.9% (i.e., $\lambda = 0.999$) of the signal energy, $\|\boldsymbol{x}_{\text{c}}\|^2$, is preserved. The corresponding results for the temperature and rain data are $\hat{k}_0 = 17$ and $\hat{k}_0 = 4$, respectively, indicating higher sparsity than the humidity signal.

In this setting, the distortion is due to the sparse approximation and the quantization of the non-zero elements.[8] More precisely, the mean squared error distortion for the lossy encoding of the compressible source can be written as

$$D(\Delta;\lambda) = \frac{1}{N}\,\mathbb{E}\left\{\left\|\boldsymbol{\mathcal{X}}_{\text{c}} - \boldsymbol{F}^{-1}\,Q^{-1}\left(Q\left(\boldsymbol{M}\,\boldsymbol{F}\boldsymbol{\mathcal{X}}_{\text{c}}\right)\right)\right\|^2\right\}$$

where $\boldsymbol{F} \in \mathbb{R}^{N\times N}$ is the DCT matrix and $\boldsymbol{M} = \text{diag}(\widehat{\boldsymbol{\pi}}) \in \mathbb{R}^{N\times N}$ is a selection matrix with $\widehat{\boldsymbol{\pi}} \in \{0,1\}^N$ indicating the estimated locations of the $\hat{k}_0$ significant elements. In order to calculate the distortion, we notice that the significant and non-significant data entries follow uniform and Gaussian distributions, respectively, in accordance to a noisy BU source model. The parameters of the distributions are then estimated from the data set. Since the DCT is an orthonormal basis, and the quantization noise, for large b, can be considered uniformly distributed with support $[-\Delta/2, \Delta/2]$, the distortion can be approximated as

$$D(\Delta;\lambda) \simeq \widehat{p}(\lambda)\frac{\Delta^2}{12} + [1-\widehat{p}(\lambda)]\;\widehat{\sigma}_n^2 \tag{43}$$

where $\widehat{p}(\lambda)$ is the estimated sparsity ratio and $\widehat{\sigma}_n^2$ is the estimated variance of the non-significant elements threated as noise.

For compressed sensing, the signals are acquired through a Rademacher measurement matrix [15], and are reconstructed from the quantized measurements $Q^{-1}\left(Q\left(\boldsymbol{y}\right)\right)$ using a sparse recovery algorithm. In particular, basis pursuit denoising algorithm is used for signal reconstruction [82], [83], which accounts for noisy measurements, e.g., due to quantization.

Then, the performance of the encoders is assessed through the ODR and ODE, for the gathered signals. More precisely, the distortion of the proposed scheme is approximated using (43), and the rate is identified by (13). On the other hand, the distortion of the CS encoder is evaluated through Monte Carlo simulation, and the rate is $r_{\text{CS}} = M\,b/N$. The energy consumption of both schemes is calculated by (40) and (42) in Section VI. In particular, we consider the power model of the MICA2 platform which embeds the Atmega128L processor and the CC1000 radio [78], [79]. Each data packet consists of a payload of up to 1872 bits and 168 additional bits as header ($h_{\text{oh}} = 9\%$ payload), while the acknowledgment packet length is 160 bits ($h_{\text{ack}} = 8.5\%$ payload). This platform has $p_{\text{acq}} = 15.01\,\text{mW}$, $T_{\text{ins}} = 0.133\,\mu\text{sec}$, $p_{\text{bck}} = 9.6\,\text{mW}$, $\epsilon_{\text{add}} = 3.3\,\text{nJ}$, $\epsilon_{\text{mul}} = 9.9\,\text{nJ}$, $\epsilon_{\text{cmp}} = 3.3\,\text{nJ}$, $\epsilon_{\text{rd}} = 0.26\,\text{nJ}$, $\epsilon_{\text{wr}} = 4.3\,\text{nJ}$, $E_{\text{rx}} = 0.923\,\mu\text{J/bit}$, and $E_{\text{tx}} = 1.135\,\mu\text{J/bit}$, using a transmit power corresponding to a maximum range of 52 m [79, Table 1].

The performance of the proposed BCHSC and CS schemes are depicted in Fig. 7. More precisely, the figure reports the convex hull (minimum) of the ODR and ODE, varying $\lambda \in [0.5, 0.9999]$ for BCHSC and $M \in [3, N-1]$ for CS. As can be seen, the proposed encoder indicates higher performance with respect to the CS based scheme, in terms of the required rate and the consumed energy. For example, BCHSC provides about 55% and 50% reduction in the required rate and consumed energy, respectively, for compressing the temperature signal with distortion $D(\Delta;\lambda) = 10^{-3}$.

VIII. Conclusion

This paper provided two novel schemes for efficient encoding of noisy sparse sources based on the syndromes associated

[8] Unlike the sparse signal model in Section II, the vector $\boldsymbol{w}$ is considered as a part of the signal.

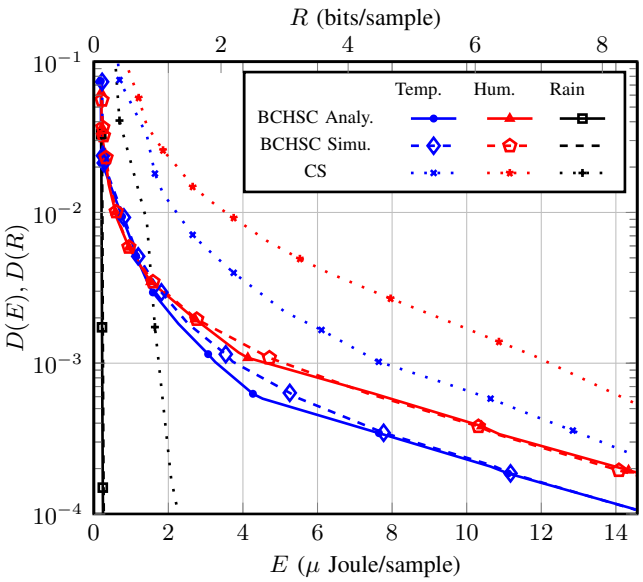


Fig. 7: The ODR and ODE of the proposed BCHSC (13) and CS for real data acquired from a WSN deployed in Torgiovannetto (Italy), for $b \in \{3, 5, \ldots, 13\}$ and $N = 127$.

with channel codes. Two sparse signal denoising techniques are proposed; the first is a blind estimator based on model order selection, while the second is a thresholding filter requiring prior knowledge of the signal model and of the noise power. The theoretical performance of the proposed encoders, derived in terms of the ODR and ODE, approaches entropy based bounds for sources with low sparsity order. Finally, we tested the encoders to compress real data gathered by a WSN for environmental monitoring, obtaining significant energy savings when compared to a compressed sensing encoder.

Appendix A
Quantizer Distortion for Noisy BU Sources

In order to derive the distortion due to both the quantization error and the additive noise for BU sources, $D^{\mathrm{BU}}(\Delta)$, the integration in (25) should be performed. Unfortunately, the floor function in the error distribution (24) complicates the analysis. Thus, a way to simplify the integral is to approximate the PDF (24) considering that $\lfloor x \rfloor \simeq x - 1/2$, i.e.,

$$f_{\mathcal{T}}^{\mathrm{BU}}(\tau) \simeq \tilde{f}_{\mathcal{T}}^{\mathrm{BU}}(\tau) \triangleq (1-p)\delta(\tau) + \frac{p}{2A} \times \begin{cases} 1 + \Phi(\tau)\left(\frac{2A}{\Delta} - 1\right), & |\tau| < \frac{\Delta}{2} \\ \Phi(\tau)\left(-\frac{|\tau|}{\Delta} + \frac{2A}{\Delta}\right) + \phi(|\tau|), & \frac{\Delta}{2} \le |\tau| < 2A - \frac{\Delta}{2} \\ 0, & \text{otherwise}. \end{cases}$$

Then, the distortion can be approximated as

$$D^{\mathrm{BU}}(\Delta) \simeq \widehat{D}^{\mathrm{BU}}(\Delta) \triangleq \int_{-\infty}^{\infty} \tau^2 \, \tilde{f}_{\mathcal{T}}^{\mathrm{BU}}(\tau)\, d\tau = \frac{p}{A} \times \left(\int_0^{\frac{\Delta}{2}} \tau^2 \left[1 + \Phi(\tau)\left(\frac{2A}{\Delta} - 1\right)\right] d\tau + \int_{\frac{\Delta}{2}}^{2A-\frac{\Delta}{2}} \tau^2 \times \left[\Phi(\tau)\left(\frac{2A}{\Delta} - \frac{\tau}{\Delta}\right) + \phi(\tau)\right] d\tau \right) = \frac{p\,(I_1 + I_2)}{48\sqrt{2\pi}\, A\, \Delta} \quad (44)$$

where

$$I_1 \triangleq 2\sigma_{\mathrm{n}} e^{-\frac{2A^2}{\sigma_{\mathrm{n}}^2}} \left[16A^3 - 32A^2\Delta + 18A\Delta^2 - 4\sigma_{\mathrm{n}}^2(A - 2\Delta) - 3\Delta^3\right] - 12\Delta e^{-\frac{\Delta^2}{2\sigma_{\mathrm{n}}^2}} \sigma_{\mathrm{n}}^3 + 4\sigma_{\mathrm{n}} e^{-\frac{(\Delta-2A)^2}{2\sigma_{\mathrm{n}}^2}} \left[3\Delta\sigma_{\mathrm{n}}^2 - 8A^3 - 4A^2\Delta + A\left(\Delta^2 + 2\sigma_{\mathrm{n}}^2\right)\right] + 2\sqrt{2\pi} A\Delta \,\mathrm{erfc}\left(\sqrt{2}A/\sigma_{\mathrm{n}}\right) \times \left[32A^2 - 18A\Delta + 3\Delta^2\right] + 2\Delta\sigma_{\mathrm{n}}\left[3\Delta^2 - 8\sigma_{\mathrm{n}}^2\right]$$

$$I_2 \triangleq \sqrt{2\pi}\,\mathrm{erf}\left(\frac{\sqrt{2}A}{\sigma_{\mathrm{n}}}\right)\left[32A^4 + 3\Delta\sigma_{\mathrm{n}}^2(8A - \Delta) + 18\sigma_{\mathrm{n}}^4\right] - \sqrt{2\pi}\,\mathrm{erf}\left(\frac{2A - \Delta}{\sqrt{2}\sigma_{\mathrm{n}}}\left[32A^4 + 3\Delta\sigma_{\mathrm{n}}^2(3\Delta - 8A) + 18\sigma_{\mathrm{n}}^4\right]\right) + 2\sqrt{2\pi}A\Delta^2\left[\Delta - 6A\right]\mathrm{erfc}\left(\frac{2A - \Delta}{\sqrt{2}\sigma_{\mathrm{n}}}\right) + 4\sqrt{2\pi}A\Delta^3 - 9\sqrt{2\pi}\sigma_{\mathrm{n}}^2\left[\Delta^2 + 2\sigma_{\mathrm{n}}^2\right]\mathrm{erf}\left(\Delta/(\sqrt{2}\sigma_{\mathrm{n}})\right).$$

This approximation is accurate as shown in the numerical results.

Appendix B
ODR of Optimal Uniform Quantizers for Noiseless BU and BG Sources

In order to derive the ODR of optimal uniform quantizers (i.e., uniform quantizer followed by an encoder with a rate equal to the entropy of the quantized source), we start with the entropy of the quantized source. The entropy of $\widetilde{\mathcal{S}} \triangleq Q^{-1}(Q(\mathcal{S}))$ at the output of the uniform quantizer is given, from (16), as

$$H(\widetilde{\mathcal{S}}) = -\sum_{i=i_{\min}}^{i_{\max}} \mathbb{P}\left\{\widetilde{\mathcal{S}} = i\,\Delta\right\} \log_2 \mathbb{P}\left\{\widetilde{\mathcal{S}} = i\,\Delta\right\}.$$

For quantized BU signals, the entropy is

$$H^{\mathrm{BU}}(p, b) = -\left[1 - p + \frac{p}{2^b - 1}\right] \log_2\left(1 - p + \frac{p}{2^b - 1}\right) - p\,\frac{2^b - 2}{2^b - 1} \log_2 \frac{p}{2^b - 1}. \quad (45)$$

Consequently, the rate-distortion of the optimal uniform quantizer can be found by minimizing the entropy of the quantized source such that the distortion is bounded below D [31], [73]. More precisely, the ODR is given by

$$R^{\mathrm{BU}}(p, D, A) = \inf_{\{b \in \mathbb{N}^+ \,|\, D^{\mathrm{BU}}(2A/(2^b-1)) \le D\}} H^{\mathrm{BU}}(p, b) = H^{\mathrm{BU}}\left(p, \left\lceil \log_2\left(1 + A\sqrt{\frac{p}{3\,D}}\right)\right\rceil\right) \quad (46)$$

where for noiseless BU sources the exact distortion, from (17) and [25, pg. 150], is simply $D^{\mathrm{BU}}(\Delta) = p\,\Delta^2/12$.

For quantized BG sources, the entropy is

$$H^{\mathrm{BG}}(p,b,\widetilde{A}) = \left[p\,\mathrm{erfc}\left(\frac{\widetilde{A}}{2^b-1}\right)-1\right] \times \log_2\left(1-p\,\mathrm{erfc}\left(\frac{\widetilde{A}}{2^b-1}\right)\right) - p\,\mathrm{erfc}\left(\widetilde{A}-\frac{2\widetilde{A}}{2^b-1}\right) \times \log_2\left(p\,\mathrm{erfc}\left(\widetilde{A}-\frac{2\widetilde{A}}{2^b-1}\right)\right) - \mathbf{1}_{\{b\geq 3\}}\left[\sum_{i=1}^{2^{b-1}-2} p \times \left[\mathrm{erf}\left(\frac{\widetilde{A}(2i+1)}{2^b-1}\right)-\mathrm{erf}\left(\frac{\widetilde{A}(2i-1)}{2^b-1}\right)\right] \log_2\left(p\left[\mathrm{erf}\left(\frac{\widetilde{A}(2i+1)}{2^b-1}\right)-\mathrm{erf}\left(\frac{\widetilde{A}(2i-1)}{2^b-1}\right)\right]\right)\right] \quad (47)$$

where $\widetilde{A} \triangleq A/(\sqrt{2}\sigma_{\mathrm{s}})$. The distortion, considering the saturation effect, can be calculated as

$$D^{\mathrm{BG}}(\Delta) = 2\,p\left[\int_0^{\Delta/2} x^2 f(x;\sigma_{\mathrm{s}}^2)\,dx + \int_{A-\Delta}^{\infty}(x-A+\Delta/2)^2 f(x;\sigma_{\mathrm{s}}^2)\,dx + \mathbf{1}_{\{A/\Delta\geq 7/2\}} \times \underbrace{\sum_{i=1}^{\frac{A}{\Delta}-\frac{3}{2}}\int_{(i-1/2)\Delta}^{(i+1/2)\Delta}|x-i\Delta|^2 f(x;\sigma_{\mathrm{s}}^2)\,dx}_{\Xi}\right] = 2p\left[-\frac{\Delta\sigma_{\mathrm{s}}e^{-\frac{\Delta^2}{8\sigma_{\mathrm{s}}^2}}}{2\sqrt{2\pi}} + \frac{1}{2}\sigma_{\mathrm{s}}^2\mathrm{erf}\left(\frac{\Delta}{2\sqrt{2}\sigma_{\mathrm{s}}}\right) - \frac{A\sigma_{\mathrm{s}}e^{-\frac{(A-\Delta)^2}{2\sigma_{\mathrm{s}}^2}}}{\sqrt{2\pi}} + \frac{1}{8}\left[(\Delta-2A)^2+4\sigma_{\mathrm{s}}^2\right] \times\mathrm{erfc}\left(\frac{A-\Delta}{\sqrt{2}\sigma_{\mathrm{s}}}\right) + \mathbf{1}_{\{A/\Delta\geq 7/2\}}\,\Xi\right] \quad (48)$$

where $f(x;\sigma_{\mathrm{s}}^2)$ is the PDF of a zero mean Gaussian r.v. with variance σ_{s}^2. The term Ξ can be bounded, since $f(x;\sigma_{\mathrm{s}}^2)$ is a decreasing function for $x>0$, as

$$\Xi < \sum_{i=1}^{\frac{A}{\Delta}-\frac{3}{2}} f\left((i-1/2)\Delta;\sigma_{\mathrm{s}}^2\right)\int_{(i-1/2)\Delta}^{(i+1/2)\Delta}|x-i\Delta|^2\,dx = \frac{\Delta^3}{12}\sum_{i=1}^{\frac{A}{\Delta}-\frac{3}{2}} f\left((i-1/2)\Delta;\sigma_{\mathrm{s}}^2\right) < \frac{\Delta^3}{12}\left(f\left(\Delta/2;\sigma_{\mathrm{s}}^2\right) + \frac{1}{2\,\Delta}\left[\mathrm{erf}\left(\frac{A-2\,\Delta}{\sqrt{2}\sigma_{\mathrm{s}}}\right)+\mathrm{erf}\left(\frac{\Delta}{2\sqrt{2}\sigma_{\mathrm{s}}}\right)\right]\right) \quad (49)$$

where (49) is due to bounding the summation of a decreasing function with the proper integration [31, (81)], i.e.,

$$\int_{(i-3/2)\Delta}^{(i-1/2)\Delta} f\left(x;\sigma_{\mathrm{s}}^2\right)dx > \Delta\, f\left([i-1/2]\Delta;\sigma_{\mathrm{s}}^2\right) \text{ for } i\geq 2\,.$$

Substituting (49) into (48) yields

$$D^{\mathrm{BG}}(\Delta) < \widetilde{D}^{\mathrm{BG}}(\Delta) \triangleq \frac{p}{12}\left[\sigma_{\mathrm{s}}\sqrt{\frac{2}{\pi}}\left(-12Ae^{-\frac{(A-\Delta)^2}{2\sigma_{\mathrm{s}}^2}} -6\Delta e^{-\frac{\Delta^2}{8\sigma_{\mathrm{s}}^2}}\right) + 12\,\sigma_{\mathrm{s}}^2\,\mathrm{erf}\left(\frac{\Delta}{2\sqrt{2}\sigma_{\mathrm{s}}}\right) + 3\left[(\Delta-2A)^2+4\sigma_{\mathrm{s}}^2\right] \times\mathrm{erfc}\left(\frac{A-\Delta}{\sqrt{2}\sigma_{\mathrm{s}}}\right) + \mathbf{1}_{\{A/\Delta\geq 7/2\}}\cdot\left(\sqrt{\frac{2}{\pi}}\frac{\Delta^3}{\sigma_{\mathrm{s}}}e^{-\frac{\Delta^2}{8\sigma_{\mathrm{s}}^2}} +\Delta^2\left[\mathrm{erf}\left(\frac{A-2\Delta}{\sqrt{2}\sigma_{\mathrm{s}}}\right)-\mathrm{erf}\left(\frac{\Delta}{2\sqrt{2}\sigma_{\mathrm{s}}}\right)\right]\right)\right]. \quad (50)$$

Finally, the ODR of the optimal uniform quantizer is derived from (47), (48), and (50) as

$$R^{\mathrm{BG}}(p,D,A) = H^{\mathrm{BG}}\left(p,\left\lceil\log_2\left(1+\frac{2A}{\Delta(D)}\right)\right\rceil,\widetilde{A}\right) \quad (51)$$

$$< H^{\mathrm{BG}}\left(p,\left\lceil\log_2\left(1+\frac{2A}{\widetilde{\Delta}(D)}\right)\right\rceil,\widetilde{A}\right) \quad (52)$$

where $\Delta(D)$ and $\widetilde{\Delta}(D)$ are the inverse of $D^{\mathrm{BG}}(\Delta)$, (48), and $\widetilde{D}^{\mathrm{BG}}(\Delta)$, (50), respectively.

REFERENCES

[1] L. Ang, K. Seng, A. Zungeru, and G. Ijemaru, "Big sensor data systems for smart cities," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1259–1271, Oct 2017.

[2] A. Biason, C. Pielli, A. Zanella, and M. Zorzi, "Access control for IoT nodes with energy and fidelity constraints," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 3242–3257, May 2018.

[3] M. Vetterli, "Wavelets, approximation, and compression," *IEEE Signal Process. Mag.*, vol. 18, no. 5, pp. 59–73, Sep. 2001.

[4] S. Mallat, *A wavelet tour of signal processing: the sparse way*. Academic press, 2008.

[5] C. Weidmann and M. Vetterli, "Rate distortion behavior of sparse sources," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 4969–4992, Aug 2012.

[6] V. Goyal, A. Fletcher, and S. Rangan, "Compressive sampling and lossy compression," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 48–56, March 2008.

[7] A. Ortega and K. Ramn, "Rate-distortion methods for image and video compression," *IEEE Signal Process. Mag.*, vol. 15, no. 6, pp. 23–50, Nov. 1998.

[8] M. Kaaniche, A. Fraysse, B. Pesquet-Popescu, and J. Pesquet, "A bit allocation method for sparse source coding," *IEEE Trans. Image Process.*, vol. 23, no. 1, pp. 137–152, Jan. 2014.

[9] E. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[10] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289 –1306, April 2006.

[11] A. Elzanaty, A. Giorgetti, and M. Chiani, "Limits on sparse data acquisition: RIC analysis of finite Gaussian matrices," *IEEE Trans. Inf. Theory*, vol. 65, no. 3, pp. 1578–1588, March 2019.

[12] G. Tang and A. Nehorai, "Performance analysis for sparse support recovery," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1383–1399, March 2010.

[13] N. Lee, "MAP support detection for greedy sparse signal recovery algorithms in compressive sensing," *IEEE Trans. Signal Process.*, vol. 64, no. 19, pp. 4987–4999, Oct 2016.

[14] S. Foucart and H. Rauhut, *A mathematical introduction to compressive sensing*. Springer, 2013.

[15] F. Chen, A. Chandrakasan, and V. Stojanovic, "Design and analysis of a hardware-efficient compressed sensing architecture for data compression in wireless sensors," *IEEE J. Solid-State Circuits*, vol. 47, no. 3, pp. 744–756, March 2012.

[16] O. Abari, F. Lim, F. Chen, and V. Stojanović, "Why analog-to-information converters suffer in high-bandwidth sparse signal applications," *IEEE Trans. Circuits Syst. I*, vol. 60, no. 9, pp. 2273–2284, 2013.

[17] C. Chang, "On the rate distortion function of bernoulli gaussian sequences," in *Proc. IEEE Int. Symp. on Inf. Theory*, Austin, TX, USA, June 2010, pp. 66–70.

[18] L. Palzer and R. Timo, "A lower bound for the rate-distortion function of spike sources that is asymptotically tight," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Cambridge, UK, Sept 2016, pp. 101–105.

[19] R. Dobrushin and B. Tsybakov, "Information transmission with additional noise," *IRE Trans. on Inf. Theory*, vol. 8, no. 5, pp. 293–304, Sept. 1962.

[20] J. Wolf and J. Ziv, "Transmission of noisy information to a noisy receiver with minimum distortion," *IEEE Trans. Inf. Theory*, vol. 16, no. 4, pp. 406–411, July 1970.

[21] M. Gastpar, "A lower bound to the AWGN remote rate-distortion function," in *Proc. IEEE 13th Workshop on Statistical Signal Processing*, July 2005, pp. 1176–1181.

[22] 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.

[23] A. Kipnis, G. Reeves, Y. C. Eldar, and A. J. Goldsmith, "Compressed sensing under optimal quantization," in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, June 2017, pp. 2148–2152.

[24] A. Kipnis, G. Reeves, and Y. C. Eldar, "Single letter formulas for quantized compressed sensing with gaussian codebooks," in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, June 2018, pp. 71–75.

[25] R. M. Gray, *Source coding theory*. Springer Science & Business Media, 2012, vol. 83.

[26] G. Myers, "Quantization of a signal plus random noise," *IRE Trans. Instrum.*, vol. PGI-5, pp. 181–186, June 1956.

[27] A. Gyorgy and T. Linder, "Optimal entropy-constrained scalar quantization of a uniform source," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2704–2711, Nov. 2000.

[28] A. Fraysse, B. Pesquet-Popescu, and J. Pesquet, "Rate-distortion results for Generalized Gaussian distributions," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Las Vegas, USA, March 2008, pp. 3753–3756.

[29] G. Sullivan, "Efficient scalar quantization of exponential and Laplacian random variables," *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1365–1374, Sep. 1996.

[30] M. Kaaniche, A. Fraysse, B. Pesquet-Popescu, and J. Pesquet, "Accurate rate-distortion approximation for sparse Bernoulli-Generalized Gaussian models," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Florence, Italy, May 2014, pp. 2020–2024.

[31] A. Fraysse, B. Pesquet-Popescu, and J. Pesquet, "On the uniform quantization of a class of sparse sources," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3243–3263, July 2009.

[32] T. Ancheta, "Syndrome-source-coding and its universal generalization," *IEEE Trans. Inf. Theory*, vol. 22, no. 4, pp. 432–436, July 1976.

[33] S. B. Korada and R. L. Urbanke, "Polar codes are optimal for lossy source coding," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1751–1768, April 2010.

[34] A. Gupta and S. Verdu, "Nonlinear sparse-graph codes for lossy compression," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 1961–1975, May 2009.

[35] A. Wyner, "Recent results in the Shannon theory," *IEEE Trans. Inf. Theory*, vol. 20, no. 1, pp. 2–10, Jan. 1974.

[36] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 1–10, Jan. 1976.

[37] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 471–480, July 1973.

[38] S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): design and construction," *IEEE Trans. Inf. Theory*, vol. 49, no. 3, pp. 626–643, March 2003.

[39] V. Stankovic, A. D. Liveris, Z. Xiong, and C. N. Georghiades, "On code design for the Slepian-Wolf problem and lossless multiterminal networks," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1495–1507, April 2006.

[40] E. Magli, M. Barni, A. Abrardo, and M. Grangetto, "Distributed source coding techniques for lossless compression of hyperspectral images," *EURASIP J. Appl. Signal Process.*, vol. 2007, no. 1, pp. 24–24, Jan 2007.

[41] F. Bassi, M. Kieffer, and C. Dikici, "Multiterminal source coding of Bernoulli-Gaussian correlated sources," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, April 2009, pp. 2501–2504.

[42] F. Bassi, M. Kieffer, and C. Weidmann, "Source coding with intermittent and degraded side information at the decoder," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, March 2008, pp. 2941–2944.

[43] C. Tian and S. Diggavi, "Side-information scalable source coding," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5591–5608, Dec 2008.

[44] A. Gabay, M. Kieffer, and P. Duhamel, "Joint source-channel coding using real BCH codes for robust image transmission," *IEEE Trans. Image Process.*, vol. 16, no. 6, pp. 1568–1583, June 2007.

[45] M. Vaezi and F. Labeau, "Distributed source-channel coding based on real-field BCH codes," *IEEE Trans. Signal Process.*, vol. 62, no. 5, pp. 1171–1184, March 2014.

[46] G. Coluccia, A. Roumy, and E. Magli, "Operational rate-distortion performance of single-source and distributed compressed sensing," *IEEE Trans. Commun.*, vol. 62, no. 6, pp. 2022–2033, June 2014.

[47] F. Chen, F. Lim, O. Abari, A. Chandrakasan, and V. Stojanovic, "Energy-aware design of compressed sensing systems for wireless sensors under performance and reliability constraints," *IEEE Trans. Circuits Syst.*, vol. 60, no. 3, pp. 650–661, March 2013.

[48] J. Z. Sun and V. K. Goyal, "Optimal quantization of random measurements in compressed sensing," in *Proc. IEEE Int. Symp. on Inf. Theory*, June 2009, pp. 6–10.

[49] L. Jacques, D. Hammond, and J. Fadili, "Dequantizing compressed sensing: When oversampling and non-gaussian constraints combine," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 559–571, Jan 2011.

[50] G. Coluccia, E. Magli, A. Roumy, and V. Toto-Zarasoa, "Lossy compression of distributed sparse sources: A practical scheme," in *Proc. European Signal Proces. Conf.*, Aug 2011, pp. 422–426.

[51] U. S. Kamilov, V. K. Goyal, and S. Rangan, "Message-passing dequantization with applications to compressed sensing," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6270–6281, Dec 2012.

[52] P. T. Boufounos, L. Jacques, F. Krahmer, and R. Saab, "Quantization and compressive sensing," in *Compressed Sensing and its Applications: MATHEON Workshop 2013*, H. Boche, R. Calderbank, G. Kutyniok, and J. Vybíral, Eds. Birkhäuser, Cham: Springer International Publishing, 2015, ch. 7, pp. 193–237.

[53] F. Zhang and H. D. Pfister, "Compressed sensing and linear codes over real numbers," in *Proc. Information Theory and Applications Workshop*, Jan 2008, pp. 558–561.

[54] M. Shirvanimoghaddam, Y. Li, B. Vucetic, J. Yuan, and P. Zhang, "Binary compressive sensing via analog fountain coding," *IEEE Trans. Signal Process.*, vol. 63, no. 24, pp. 6540–6552, Dec 2015.

[55] A. G. Dimakis, R. Smarandache, and P. O. Vontobel, "LDPC codes for compressed sensing," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 3093–3114, May 2012.

[56] F. Abramovich, T. Sapatinas, and B. Silverman, "Wavelet thresholding via a Bayesian approach," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 60, no. 4, pp. 725–749, 1998.

[57] V. Stanković, L. Stanković, and S. Cheng, "Compressive image sampling with side information," in *16th IEEE Inter. Conf. on Image Process. (ICIP)*, Cairo, Egypt, Nov 2009, pp. 3037–3040.

[58] G. Yu, G. Sapiro, and S. Mallat, "Solving inverse problems with piecewise linear estimators: From Gaussian mixture models to structured sparsity," *IEEE Transactions on Image Processing*, vol. 21, no. 5, pp. 2481–2499, May 2012.

[59] J. Yang, X. Yuan, X. Liao, P. Llull, D. J. Brady, G. Sapiro, and L. Carin, "Video compressive sensing using Gaussian mixture models," *IEEE Transactions on Image Processing*, vol. 23, no. 11, pp. 4863–4878, Nov 2014.

[60] K. Y. Lee, B. . Lee, I. Song, and S. Ann, "On Bernoulli-Gaussian process modeling of speech excitation source," in *International Conference on Acoustics, Speech, and Signal Processing*, April 1990, pp. 217–220 vol.1.

[61] A. Mariani, A. Giorgetti, and M. Chiani, "Model order selection based on information theoretic criteria: Design of the penalty," *IEEE Trans. Signal Process.*, vol. 63, no. 11, pp. 2779–2789, June 2015.

[62] P. Stoica and Y. Selen, "Model-order selection: a review of information criterion rules," *IEEE Signal Process. Mag.*, vol. 21, no. 4, pp. 36–47, July 2004.

[63] M. Chiani and M. Win, "Estimating the number of signals observed by multiple sensors," in *Proc. IEEE IAPR Workshop on Cognitive Inform. Process. (CIP)*, Elba Island, Italy, Jun. 2010, pp. 156–161.

[64] A. Mariani, A. Giorgetti, and M. Chiani, "Wideband spectrum sensing by model order selection," *IEEE Trans. Wireless Commun.*, vol. 14, no. 12, pp. 6710–6721, Dec. 2015.

[65] A. Giorgetti and M. Chiani, "Time-of-arrival estimation based on information theoretic criteria," *IEEE Trans. Signal Process.*, vol. 61, no. 8, pp. 1869–1879, April 2013.

[66] T. Lancaster, "The incidental parameter problem since 1948," *Journal of econometrics*, vol. 95, no. 2, pp. 391–413, 2000.
[67] J. Neyman and E. Scott, "Consistent estimates based on partially consistent observations," *Econometrica*, vol. 16, no. 1, pp. 1–32, 1948.
[68] Y. You, *Audio Coding: Theory and Applications*. Springer Science & Business Media, 2010.
[69] I. Reed and G. Solomon, "Polynomial codes over certain finite fields," *SIAM. J. Appl. Math.*, vol. 8, no. 2, pp. 300–304, 1960.
[70] W. Ryan and S. Lin, *Channel codes: classical and modern*. Cambridge University Press, 2009.
[71] S. Lin and D. Costello, *Error control coding, Fundamentals and applications*. Prentice Hall, Inc. Englewood Cliffs, New Jersey, 1983.
[72] E. Berlekamp, *Algebraic coding theory*. McGraw-Hill, 1968.
[73] T. Cover and J. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.
[74] A. Papoulis and S. Pillai, *Probability, Random Variables, and Stochastic Processes*, ser. McGraw-Hill series in electrical engineering: Communications and signal processing. McGraw-Hill, 2002.
[75] B. Widrow, I. Kollar, and M.-C. Liu, "Statistical theory of quantization," *IEEE Trans. Instrum. Meas.*, vol. 45, no. 2, pp. 353–361, Apr. 1996.
[76] C. Luna, Y. Eisenberg, R. Berry, T. Pappas, and A. Katsaggelos, "Joint source coding and data rate adaptation for energy efficient wireless video streaming," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 10, pp. 1710–1720, Dec 2003.
[77] A. Mämmelä, J. Riekki, A. Kotelba, and A. Anttonen, "Multidisciplinary and historical perspectives for developing intelligent and resource-efficient systems," *IEEE Access*, vol. PP, no. 99, pp. 1–41, March 2018.
[78] J. Vales-Alonso, E. Egea-López, A. Martínez-Sala, P. Pavón-Mariño, M. Bueno-Delgado, and J. García-Haro, "Performance evaluation of MAC transmission power control in wireless sensor networks," *Computer Networks*, vol. 51, no. 6, pp. 1483–1498, April 2007.
[79] C. Karakus, A. Gurbuz, and B. Tavli, "Analysis of energy efficiency of compressive sensing in wireless sensor networks," *IEEE Sensors J.*, vol. 13, no. 5, pp. 1999–2008, May 2013.
[80] E. Feig and S. Winograd, "Fast algorithms for the discrete cosine transform," *IEEE Trans. Signal Process.*, vol. 40, no. 9, pp. 2174–2193, Sept 1992.
[81] A. Giorgetti, M. Lucchi, E. Tavelli, M. Barla, G. Gigli, N. Casagli, M. Chiani, and D. Dardari, "A robust wireless sensor network for landslide risk analysis: System design, deployment, and field testing," *IEEE Sensors J.*, vol. 16, no. 16, pp. 6374–6386, Aug. 2016.
[82] Y. Zhang, "User's guide for yall1: Your algorithms for ℓ_1 optimization," Department of CAAM, Rice University, Houston, Texas, Tech. Rep., 2009.
[83] J. Yang and Y. Zhang, "Alternating direction algorithms for ℓ_1-problems in compressive sensing," *SIAM Journal on Scientific Computing*, vol. 33, no. 1, pp. 250–278, 2011.