

Distributed Distortion-Rate Optimized Compressed Sensing in Wireless Sensor Networks

Markus Leinonen, *Student Member, IEEE*, Marian Codreanu, *Member, IEEE*,
and Markku Juntti, *Senior Member, IEEE*,

Abstract—This paper addresses lossy distributed source coding for acquiring correlated sparse sources via compressed sensing (CS) in wireless sensor networks. Noisy CS measurements are separately encoded at a finite rate by each sensor, followed by joint reconstruction of the sources at the decoder. We develop a novel complexity-constrained distributed variable-rate quantized CS method which minimizes a weighted sum between the mean square error signal reconstruction distortion, and the average encoding rate. The encoding complexity of each sensor is restrained by pre-quantizing the encoder input, i.e., the CS measurements, via vector quantization. Following the entropy-constrained design, each encoder is modeled as a quantizer followed by a lossless entropy encoder, and variable-rate coding is incorporated via rate measures of an entropy bound. For a two-sensor system, necessary optimality conditions are derived, practical training algorithms are proposed, and complexity analysis is provided. Numerical results show that the proposed method achieves superior compression performance as compared to baseline methods, and lends itself to versatile setups with different performance requirements.

Index Terms—Distributed lossy source coding, joint sparsity, entropy-constrained vector quantization, variable-length coding, rate measure.

I. INTRODUCTION

THE proliferation of wireless sensor networks (WSNs), especially due to the advent of the internet of things [1], engenders the need of energy-efficient communication techniques for resource-limited sensors in, e.g., environmental, industrial, healthcare, and military applications. In a typical monitoring task, geographically distributed sensors measure a correlated information source, encode the measurements separately into finite-rate bit sequences, and communicate the messages to the sink for joint decoding of the sources. Since wireless access consumes lots of battery energy [2], source compression [3]–[5] is prevalently used to minimize the number of transmissions to prolong the network lifetime. In particular, the decentralized network structure calls for distributed source coding (DSC) [5]–[7], commonly referred

to as Slepian-Wolf coding [8]. Furthermore, as many sensor signals are analog/continuous-valued, quantization [9]–[14] is inevitable in practice, and, thus, the compression becomes lossy [15].

An information theoretic foundation for lossy source coding is the rate-distortion theory [16, Ch. 2] [17, Ch. 10], i.e., source coding with a fidelity criterion [18, Ch. 9]. Whereas the theory is primarily applicable to performance analysis and benchmarking, vector quantization (VQ) [13], [14] is a practical compression method capable to achieve performance arbitrary close to the theoretical optimum [3], [4], [19]. The inevitable loss from a finite VQ dimension can be compensated for by variable-length coding [20], [21] [17, Ch. 5] that assigns shorter codewords to more frequent symbols to minimize the average transmission rate. This can be realized by entropy-constrained VQ (ECVQ) [22] that models the encoder as a VQ followed by a lossless entropy encoder for the symbols. In practice, VQs/ECVQs can be optimized by the Linde-Buzo-Gray algorithm [23], descending from the iterative Lloyd algorithm [24], [25].

In WSNs, standard source compression is altered by two peculiarities. The first is a DSC setting where multiple sensors separately measure, encode, and communicate multiple (correlated) sources to a sink for joint signal reconstruction. DSC originates from the pioneering work by Slepian and Wolf [8] for discrete sources, which is extended to Gaussian sources in [26]. Distributed quantizer design for compressing correlated sources appears in, e.g., [27]–[29], which incorporate entropy coding via rate measures of entropy bounds. The second modification arises from remote sensing: encoding at each sensor relies only on indirect observations of a source due to, e.g., a dedicated measurement device, or noise-corrupted sensor inputs. This necessitates remote source coding, introduced in the seminal work by Dobrushin and Tsybakov [30]. VQ design for compressing a remote source is addressed in [31]. Combining these two factors gives rise to distributed joint estimation and compression [32]. Distributed compression of a remote source is addressed in [33], [34], the CEO problem in [35], and DSC of multiple remote Gaussian sources in [36]. Related quantizer based studies include distributed estimation of a remote source in [34], [37], [38], and its extension to multiple remote sources in [39].

As an embodiment of remote source coding, compressed sensing (CS) [40]–[48] is a recently emerged joint sampling and compression paradigm which enables accurate reconstruction of sparse (or compressible) signals from a few linear measurements. Sparse signals are encountered in, e.g.,

Manuscript received July 6, 2017; revised November 13, 2017; accepted December 21, 2017. The associate editor coordinating the review of this paper and approving it for publication was Dr. M. Grangetto.

This work was presented in part in IEEE Global Conf. on Signal and Inform. Proc., Orlando, Florida, Dec. 14–16 2015, and IEEE Workshop on Sign. Proc. Adv. in Wirel. Commun., Edinburgh, UK, Jul. 3–7 2016.

This research was financially supported by the Academy of Finland, Emil Aaltonen Foundation, HPY Research Foundation, Infotech Oulu Doctoral Program, Nokia Foundation, Riitta ja Jorma J. Takanen Foundation, Tauno Tönnö Foundation, and Walter Ahlström Foundation.

The authors are with Centre for Wireless Communications – Radio Technologies, University of Oulu, P.O. Box 4500, 90014, University of Oulu, Finland. e-mail: {markus.leinonen,marian.codreanu,markku.juntti}@oulu.fi.

environmental monitoring [49]–[51], source localization [52], and cognitive radio applications like spectrum sensing and direction of arrival estimation [53]. In particular, distributed CS [49], [54] establishes a unified decentralized compression framework where intra- and inter-signal dependencies of multi-dimensional compressible signals are exploited via joint sparsity models. CS benefits from simple and universal encoding by shifting most of the computational load to the decoder [46].

Whereas the early era of CS operated exclusively on analog signals, the practical necessity of converting the measurements into finite-rate bit sequences initiated quantized CS (QCS) [55]–[58]. Consequently, numerous QCS algorithms that accommodate the non-linear impact of CS and quantization in the encoder/decoder to ameliorate the signal recovery performance have been developed. The works in [56], [59], [60] devised optimized scalar quantizers (SQs) for a fixed CS decoder, whereas quantization aware CS decoding algorithms for a fixed SQ encoder developed in [57], [58], [61]–[63]. These methods, however, are sub-optimal because they 1) optimize only either the encoder or decoder, 2) use SQ instead of VQ, and/or 3) minimize the measurement quantization distortion, which, due to non-linearity of the CS and quantization, does not in general minimize the signal reconstruction distortion [64, Sect. 3.2.3] [61]. To this end, the first joint designs of VQ based encoder-decoder pair(s) to minimize the signal reconstruction error in QCS setups are [64]–[67]. Shirazinia *et al.* derive necessary optimality conditions for acquiring sparse sources over noisy channels in a single-sensor case in [65], and in a two-sensor DSC setup in [66], [67]. However, the enhanced compression entails high encoding complexity: a sensor must reconstruct a minimum mean square error (MMSE) estimate – a task of exponential complexity [68]–[70]. As a countermeasure, QCS methods that moderate encoding complexity via pre-quantizing the encoder inputs were developed in a single-sensor case in [71], and in a DSC setup, albeit for ideal channels, in [72]. Besides algorithm development, information theoretic limits of QCS have been recently studied in, e.g., [73], [74].

Contributions: We consider DSC for finite-rate CS acquisition of correlated sparse sources in WSNs. We develop a novel distributed variable-rate QCS method under complexity-constrained encoding which minimizes a weighted sum between the mean square error (MSE) signal reconstruction distortion and the average encoding rate. The encoding complexity of each sensor is restrained via pre-quantizing the encoder input, i.e., the CS measurements, with a VQ. Following the ECVQ approach, each encoder is modeled as a quantizer followed by a lossless entropy encoder, and variable-rate coding is incorporated via the rate measures of an entropy bound. We confine to separate coding of sensors' message indices. Moreover, the code rates are merely approximated via rate measures, i.e., finding the optimal codebooks are outside the scope of this paper. We derive necessary optimality conditions in a two-sensor case, propose practical training algorithms, and provide complexity analysis. The proposed method is numerically demonstrated to efficiently utilize signal correlation, achieve high distortion-rate performance, and suit to setups with different performance criteria.

Related works: To the best of our knowledge, this is the first work on distortion-rate optimization in a (distributed) QCS setup. The CS model bases on the distributed CS in [49], [54]. The quantizer design is mostly related to [66], [67] which, differently from our work, 1) focus on channel-optimized VQs, 2) use fixed-rate VQs, and 3) involve complex encoders impeding the practical implementation. The entropy coding part utilizes the original ideas of [22], and those extended to multiple remote sources in, e.g., [39]. Apart from our initial works [71], [72], pre-quantization as a means to moderate encoding complexity, and the use of variable-rate coding in a QCS setup have not been addressed before. A philosophy similar to our pre-quantization appears in multiple prototype based approaches in, e.g., [75], [76]. Whereas [71] focuses on joint source-channel coding in a single-sensor setup, this paper takes a pure DSC approach similar to [72], and extends it by providing 1) more detailed derivations and model illustrations, 2) practical training algorithms, 3) complexity analysis, and 4) more insightful and extensive numerical experiments.

Organization: The system model is defined in Section II. In Section III, the novel distributed QCS method is developed, practical training algorithms are proposed, and algorithms' complexities are analysed. Numerical results are presented in Section IV. Section V summarizes the paper, and discusses future work.

Notations: Italic capital letters denote random variables (X); boldface non-italic capital letters denote random vectors (\mathbf{X}); boldface non-italic small letters denote realizations of random vectors (\mathbf{x}); boldface italic small and capital letters denote deterministic vectors (\mathbf{a}) and matrices (\mathbf{A}), respectively; calligraphy letters denote sets/alphabets (\mathcal{B}). The identity matrix of size $N \times N$ is denoted as \mathbf{I}_N , and a matrix with all entries zeros as $\mathbf{0}$. $(\cdot)^T$ denotes the matrix transpose, and $|\mathcal{S}|$ denotes the cardinality of set \mathcal{S} . $\|\cdot\|_0$ and $\|\cdot\|_2$ denote the ℓ_0 -norm and ℓ_2 -norm, respectively.

II. SYSTEM MODEL

Consider a distributed QCS system consisting of two¹ CS based sensors and one sink, as depicted in Fig. 1. By the DSC philosophy, each sensor acquires noisy CS measurements of its source, converts them into finite-rate bit sequences via *separate encoding* (i.e., without inter-sensor collaboration), and communicates the messages to the sink for *joint decoding* of both sources. As the main focus is on source coding, the transmissions from each sensor to the sink are assumed to be error-free.

A. CS Signal Acquisition

The correlated sources \mathbf{X}_1 and \mathbf{X}_2 are given by a joint sparsity model termed JSM-2 in [49], [54] as

$$\mathbf{X}_l = \bar{\Theta} + \Theta_l, \quad l = 1, 2, \quad (1)$$

¹The presented alternating optimization framework, where one optimizes one system block while keeping the others fixed, can be readily extended to a general multi-sensor system, although the computational and memory requirements rapidly grow intolerably high. Nevertheless, the main features of the considered scheme are more lucidly highlighted in the two-sensor setup while avoiding cumbersome derivations.

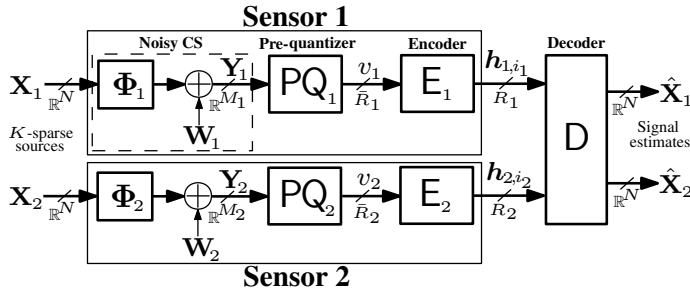


Fig. 1. Distributed variable-rate QCS acquisition of correlated sparse sources under complexity-constrained encoding.

where the common component $\bar{\Theta}$ and innovation component Θ_l are real-valued length- N random vectors which both are K -sparse, and share the same random (unknown) support, i.e., the set of indices of non-zero components. Consequently, each real-valued length- N source random vector \mathbf{X}_l is K -sparse as $\|\mathbf{x}_l\|_0 \leq K \leq N$, $l = 1, 2$. The vectors $\bar{\Theta}$, Θ_1 , and Θ_2 are assumed to be independent of each other. JSM-2 signals² are encountered in, e.g., a group of sensors monitoring an audio source or spectrum occupancy [49].

Let $\mathcal{T}_s \subseteq \{1, \dots, N\}$ be an index set representing the s th sparsity pattern with $|\mathcal{T}_s| = K$, $s = 1, \dots, \binom{N}{K}$. The $\binom{N}{K}$ index sets are different, i.e., $\mathcal{T}_s \setminus \mathcal{T}_{s'} \neq \emptyset$, $\forall s' \neq s = 1, \dots, \binom{N}{K}$. Each support \mathcal{T}_s is associated with the *a priori* probability $p(\mathcal{T}_s) \in [0, 1]$ with $\sum_{s=1}^{\binom{N}{K}} p(\mathcal{T}_s) = 1$.

Each sensor measures the source \mathbf{X}_l through a fixed (and known) CS measurement matrix $\Phi_l \in \mathbb{R}^{M_l \times N}$ as

$$\mathbf{Y}_l = \Phi_l \mathbf{X}_l + \mathbf{W}_l, \quad l = 1, 2, \quad (2)$$

where \mathbf{Y}_l is the length- M_l measurement random vector, and $\mathbf{W}_l \sim \mathcal{N}(\mathbf{0}, \sigma_{\mathbf{W}}^2 \mathbf{I}_{M_l})$ is the measurement noise random vector. Whereas $K \leq M_l \leq N$ is assumed for a conventional CS setup³, the design is not restricted to any particular range for M_l . It is worth emphasizing that as (2) models the physics of the sensing process, the encoder at each CS-based sensor $l = 1, 2$ has no access to \mathbf{X}_l , but only to \mathbf{Y}_l . Consequently, the compression scheme falls into remote source coding [30].

The structure of each Φ_l , $l = 1, 2$, has a significant impact on the CS signal recovery performance. In this respect, the restricted isometry property⁴ and the coherence of Φ_l , which establish guarantees on stable and accurate CS signal reconstruction (in a non-quantized case), are often used to assess the quality of each Φ_l in CS [48], [81] [80, Ch. 6] [82, Sect. 1.4.2 – 1.4.3] [83, Sect. 1.3.3]. Nonetheless, we need no such assumptions of Φ_1 and Φ_2 .

²Common sparse signals have been also studied under compressive support recovery and signal reconstruction problems with multiple measurement vectors in a variety of monitoring applications [77]–[79].

³Whereas $M_l \leq N$ complies with the fundamental CS theory, over-sampling (i.e., $M_l > N$) may be useful in QCS setups. Namely, given a quantization bit resolution of an A/D converter, over-sampling is a practical – and often cost-effective – way to improve the reconstruction accuracy [58].

⁴For instance, Φ_l with i.i.d. Gaussian entries satisfies the restricted isometry property with overwhelming probability if $M_l \geq C_l K \log(N/K)$, where C_l is a positive constant [80, Ch. 1].

B. Measurement Space Pre-Quantization

Prior to the actual source encoding, the input of the encoding system at each sensor $l = 1, 2$, i.e., the measurement random vector \mathbf{Y}_l of (2), is discretized with a VQ. Each pre-quantizer⁵ PQ_l is a focal block in restraining the encoding complexity of a sensor while providing high compression performance. As a by-product, it simplifies the optimization design by converting optimization over continuous random variables into that over discrete ones, and facilitates offline training by allowing pre-computation of required input quantities. A general description of each pre-quantizer is given next whereas their specific optimization is deferred until Section III-B.

Let $\mathcal{V}_l \triangleq \{1, \dots, |\mathcal{V}_l|\}$ be a set of *cell indices* $v_l \in \mathcal{V}_l$ with pre-quantization rate $\bar{R}_l = \log_2 |\mathcal{V}_l|$ bits/vector \mathbf{Y}_l for sensor $l = 1, 2$. Let $\mathcal{G}_l \triangleq \{\mathbf{g}_{l,1}, \dots, \mathbf{g}_{l,|\mathcal{V}_l|}\}$ be a *pre-quantization codebook* consisting of *codepoints* $\mathbf{g}_{l,v_l} \in \mathbb{R}^{M_l}$. Each pre-quantizer PQ_l is a $|\mathcal{V}_l|$ -level VQ that partitions the M_l -dimensional measurement vector space determined by (2) into *cells* $\mathcal{S}_{l,1}, \dots, \mathcal{S}_{l,|\mathcal{V}_l|}$, i.e., $\mathcal{S}_{l,v_l} \cap \mathcal{S}_{l,v'_l} = \emptyset$, $v_l \neq v'_l \in \mathcal{V}_l$, and $\bigcup_{v_l=1}^{|\mathcal{V}_l|} \mathcal{S}_{l,v_l} = \mathbb{R}^{M_l}$. Thus, PQ_l is a lossy mapping

$$\text{PQ}_l : \mathbb{R}^{M_l} \rightarrow \mathcal{V}_l, \quad l = 1, 2, \quad (3)$$

i.e., for a given measurement realization, it assigns a cell index as $\text{PQ}_l(\mathbf{y}_l) = v_l \in \mathcal{V}_l$, if $\mathbf{y}_l \in \mathcal{S}_{l,v_l}$.

Remark 1. The codepoints $\mathbf{g}_{l,1}, \dots, \mathbf{g}_{l,|\mathcal{V}_l|}$ are intermediate quantities for the actual encoding at E_l , i.e., they are solely used to determine in which cell \mathcal{S}_{l,v_l} each realization \mathbf{y}_l belongs to, $l = 1, 2$. This classification is made explicit by the encoding rule defined in Section III-B. In summary, the quantized version of random vector \mathbf{Y}_l is never reconstructed in the system.

C. Encoding and Decoding

The outputs of each PQ_l , i.e., the cell indices $v_l \in \mathcal{V}_l$, are fed to encoder E_l at each sensor $l = 1, 2$. Following a customary approach, E_l is modelled as the concatenation of a (lossy) quantizer, and a lossless entropy encoder [27], [28]. Accordingly, each sensor encodes the cell indices into message indices, and further, into binary source codewords. The sink uses the received pairs of codewords to jointly reconstruct estimates of \mathbf{X}_1 and \mathbf{X}_2 . Distributed quantizer blocks are described next; the treatment of entropy coding is elaborated in Section II-D.

1) *Separate Encoders:* Let $\mathcal{I}_l \triangleq \{1, \dots, |\mathcal{I}_l|\}$ be a set of *message indices* $i_l \in \mathcal{I}_l$, $|\mathcal{I}_l| \leq |\mathcal{V}_l|$, for sensor $l = 1, 2$. Let $\mathcal{H}_l \triangleq \{\mathbf{h}_{l,1}, \dots, \mathbf{h}_{l,|\mathcal{H}_l|}\}$, $|\mathcal{H}_l| = |\mathcal{I}_l|$, be a *source codebook* consisting of binary *codewords*. Each encoder is a composite mapping $E_l : \mathcal{V}_l \rightarrow \mathcal{I}_l \rightarrow \mathcal{H}_l$ as

$$E_l : \mathcal{V}_l \rightarrow \mathcal{I}_l \rightarrow \mathcal{H}_l, \quad l = 1, 2. \quad (4)$$

The first mapping $\pi_l : \mathcal{V}_l \rightarrow \mathcal{I}_l$, termed the *message index mapping*, maps each cell index $v_l \in \mathcal{V}_l$ into a message index $i_l \in \mathcal{I}_l$, i.e., $\pi_l = \{\pi_l(1), \dots, \pi_l(|\mathcal{V}_l|)\}$, where

⁵Pre-quantization has also a pragmatic aspect: prior to source encoding, the sensor inputs are necessarily discretized with an A/D converter in any digital sensor device.

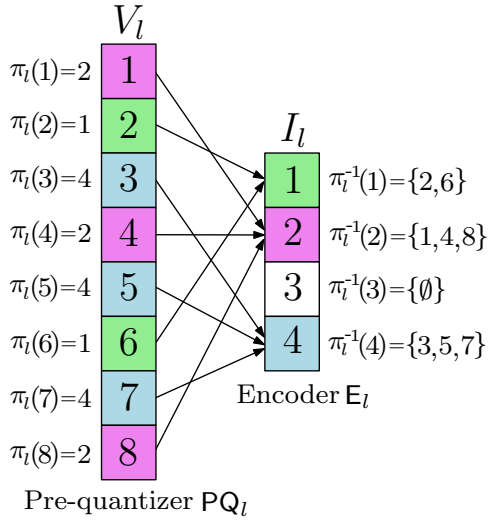


Fig. 2. Interconnections between the indices of PQ_l and E_l for $|\mathcal{V}_l| = 8$ and $|\mathcal{I}_l| = 4$, $l = 1, 2$. Note that $I_l = 3$ is assigned no cell index $v_l \in \mathcal{V}_l$.

$\pi_l(v_l) \in \mathcal{I}_l$. Given π_l , each $i_l \in \mathcal{I}_l$ is associated with a set of cell indices mapped to itself, i.e., the inverse image $\pi_l^{-1}(i_l) = \{v_l \in \mathcal{V}_l | \pi_l(v_l) = i_l\}$. Note that for $|\mathcal{I}_l| < |\mathcal{V}_l|$, π_l is a many-to-one mapping, i.e., it performs *lossy* compression of the cell index V_l . This accounts for the prefix "pre" for PQ_l , as each sensor has a concatenation of two quantizers: the VQ of PQ_l , and the message index mapping π_l in E_l . Interconnections between the indices of PQ_l and E_l are illustrated in Fig. 2.

Given an entropy code, the second mapping $\alpha_l : \mathcal{I}_l \rightarrow \mathcal{H}_l$ is a one-to-one *lossless* mapping from the set of message indices to binary source codewords $\mathbf{h}_{l,i_l} \in \mathcal{H}_l$, i.e., $\alpha_l(i_l) = \mathbf{h}_{l,i_l}$, $i_l \in \mathcal{I}_l$. Fixed-to-variable-length coding is assumed, i.e., each index $i_l \in \mathcal{I}_l$ is mapped to one codeword at a time, whereas the binary representations of $\mathbf{h}_{l,i_l} \in \mathcal{H}_l$ have, in general, different lengths.

2) *Joint Decoder*: The joint decoder comprises of two composite mappings $D : \{\beta_1 \circ \alpha_1^{-1}, \beta_2 \circ \alpha_2^{-1}\}$ as

$$D : (\mathcal{H}_1 \times \mathcal{H}_2) \rightarrow (\mathcal{I}_1 \times \mathcal{I}_2) \rightarrow \mathcal{C}_l, \quad l = 1, 2. \quad (5)$$

The first mappings $\alpha_l^{-1} : \mathcal{H}_l \rightarrow \mathcal{I}_l$, $l = 1, 2$, decode the entropy codes, i.e., the received codeword $\mathbf{h}_{l,i_l} \in \mathcal{H}_l$ is used to recover the (transmitted) message index $i_l \in \mathcal{I}_l$. By this procedure, we confine to separate coding⁶ of message indices I_1 and I_2 . Owing to the information lossless property, the pairs $(\alpha_1, \alpha_1^{-1})$ and $(\alpha_2, \alpha_2^{-1})$ constitute uniquely decodable codes.

As for the second mappings $\beta_l : (\mathcal{I}_1 \times \mathcal{I}_2) \rightarrow \mathcal{C}_l$, the decoded message indices are used to jointly reconstruct estimates of sources \mathbf{X}_1 and \mathbf{X}_2 as $\hat{\mathbf{x}}_1 := \beta_1(i_1, i_2) = \mathbf{c}_{1,i_1,i_2}$ and $\hat{\mathbf{x}}_2 := \beta_2(i_1, i_2) = \mathbf{c}_{2,i_1,i_2}$, where $\mathbf{c}_{l,i_1,i_2} \in \mathbb{R}^N$ is the *codeword* of a *reconstruction codebook* $\mathcal{C}_l \triangleq \{\mathbf{c}_{l,1,1}, \dots, \mathbf{c}_{l,|\mathcal{I}_1|,|\mathcal{I}_2|}\}$, $|\mathcal{C}_l| = |\mathcal{I}_1||\mathcal{I}_2|$, $l = 1, 2$. Thus, the decoder D performs a single operation to 1) invert the quantization/encoding steps applied to \mathbf{Y}_1 and \mathbf{Y}_2 , and 2) reconstruct the signal estimates of \mathbf{X}_1

⁶Separate coding enables low-complexity and low-delay coding as the pairs $(\alpha_l, \alpha_l^{-1})$, $l = 1, 2$, can be chosen to constitute two *instantaneous* lossless source codes.

and \mathbf{X}_2 , given the underlying distributed CS setup.

D. Entropy Coding

The pairs $(\alpha_1, \alpha_1^{-1})$ and $(\alpha_2, \alpha_2^{-1})$ can realize different entropy coding classes in the system. In practice, the choice for used source code may depend on, e.g., application requirements like reconstruction fidelity and maximum allowed delay (i.e., coding block length), and implementation factors such as sensors' computation and memory capabilities. A unified framework for subsuming entropy coding in the distributed quantizer design is introduced next.

1) *Average Rate*: Let R_l be the average encoding rate in bits/vector \mathbf{Y}_l for sensor $l = 1, 2$. Thus, the average sum rate of the sensors is

$$\underline{R} \triangleq R_1 + R_2. \quad (6)$$

In practice, R_l is determined by the average codeword length of the codebook \mathcal{H}_l , i.e.,

$$R_l \triangleq \mathbb{E}[\gamma(\cdot)] = \sum_{i_l=1}^{|\mathcal{I}_l|} p(i_l)\gamma(i_l), \quad l = 1, 2, \quad (7)$$

where $\gamma(i_l)$ is the length of codeword $\mathbf{h}_{l,i_l} \in \mathcal{H}_l$, and $p(i_l) \triangleq \Pr(I_l = i_l)$ is the probability of index $i_l \in \mathcal{I}_l$. An alternative rate definition is presented in the following.

2) *Rate Measure*: Following the approaches in [22], [27], [28] [29, Sect. 4.2], instead of using (7), the average rate is approximated via the *entropy bound* of a source code⁷. Let $r(p(i_l))$ be a *rate measure* which is a function of the message index probabilities $p(i_l)$, $i_l \in \mathcal{I}_l$, $l = 1, 2$. Accordingly, \underline{R} in (6) is given as the expectation of the rate measures, i.e.,

$$\underline{R} \triangleq \sum_{l=1}^2 \sum_{i_l=1}^{|\mathcal{I}_l|} p(i_l)r(p(i_l)). \quad (8)$$

As a major benefit, the rate definition in (8) permits flexible treatment of the entropy coding without tying the design to any particular source code.

Entropy bounds and their associated rate measures for various coding settings have been listed in, e.g., [27] [29, Sect. 4.2]. For the considered separate coding, $r_{\text{sep}}(p(i_l)) = -\log_2 p(i_l)$, $i_l \in \mathcal{I}_l$, and, thus, the average sum rate is approximated as $\underline{R}_{\text{sep}} = H(I_1) + H(I_2)$, where $H(I_l) = -\sum_{i_l=1}^{|\mathcal{I}_l|} p(i_l)\log_2 p(i_l)$ is the entropy of message index I_l , $l = 1, 2$.

III. DISTRIBUTED VARIABLE-RATE QCS METHOD

This section describes a distortion-rate optimization framework for the variable-rate communication system of Fig. 1. A novel distributed QCS method for efficient acquisition of the correlated sparse sources \mathbf{X}_1 and \mathbf{X}_2 under complexity-constrained encoding is developed. Practical training algorithms are proposed, and implementation and complexity aspects are discussed.

A. Problem Formulation

Let $\hat{\mathbf{X}}_l$ be a length- N random vector that represents the estimate of source \mathbf{X}_l , $l = 1, 2$, at the output of decoder D .

⁷The design of practical codes is beyond the scope of this paper.

The average sum MSE reconstruction distortion is defined as

$$\underline{D} \triangleq D_1 + D_2 = \sum_{l=1}^2 \mathbb{E}[\|\mathbf{X}_l - \hat{\mathbf{X}}_l\|_2^2], \quad (9)$$

where the expectation $\mathbb{E}[\cdot]$ is taken over the distributions of \mathbf{X}_l and \mathbf{W}_l , $l = 1, 2$. Furthermore, let $\mathcal{L}_\mu(\underline{D}, \underline{R})$ be a weighted distortion-rate cost function as

$$\mathcal{L}_\mu(\underline{D}, \underline{R}) \triangleq (1 - \mu)\underline{D} + \mu\underline{R}, \quad (10)$$

where $\mu \in [0, 1]$ is a weighting parameter for adjusting the distortion-rate trade-off, and \underline{R} is the average sum rate given in (8).

The objective is to minimize $\mathcal{L}_\mu(\underline{D}, \underline{R})$ in (10) for a given μ by optimizing the pre-quantizers PQ_l (i.e., the pre-quantization codebooks \mathcal{G}_l), the encoders E_l (i.e., the message index mappings π_l), and the decoder D (i.e., the reconstruction codebooks \mathcal{C}_l) such that the pre-quantization cells \mathcal{S}_{l,v_l} , $v_l \in \mathcal{V}_l$, at each PQ_l , $l = 1, 2$, satisfy a nearest-neighbor (NN) condition [84]. As will be elaborated later, this structural constraint is focal in restraining sensors' encoding complexity. Since the joint optimization of all these blocks is intractable, the design is split into two steps: first, each PQ_l is optimized under the NN constraints (Section III-B), followed by the optimization of E_1 , E_2 , and D for fixed pre-quantizers (Section III-C). Despite the sub-optimality, the approach is empirically shown to yield satisfactory performance (Section IV).

The rate term $\mu\underline{R}$ in (10) eliminates the need of constructing specific codebooks \mathcal{H}_1 and \mathcal{H}_2 in the optimization phase. As demonstrated in [22] (for a point-to-point case), decent performance is achievable by subsuming rate measures $r_{\text{sep}}(p(i_l)) = -\log_2 p(i_l)$, $i_l \in \mathcal{I}_l$, in the optimization loop, followed by a source code whose average codeword length is close to the index entropy [22]. These include Huffman codes [20] [17, Sect. 5.6], and arithmetic codes [85]. Note that $\mu = 0$ realizes a minimum distortion fixed-rate method with $R_l = \log_2 |\mathcal{I}_l|$, $l = 1, 2$.

B. Optimization of Pre-Quantizers

At each PQ_l , the cells \mathcal{S}_{l,v_l} and codepoints \mathbf{g}_{l,v_l} , $v_l \in \mathcal{V}_l$, of the $|\mathcal{V}_l|$ -level VQ are optimized to minimize the MSE distortion induced by discretizing the measurement vector space (see (2)), i.e., the distortion

$$D_l^{\text{PQ}} \triangleq \sum_{v_l=1}^{|\mathcal{V}_l|} p(v_l) \mathbb{E}[\|\mathbf{Y}_l - \mathbf{g}_{l,v_l}\|_2^2 | V_l = v_l], \quad l = 1, 2. \quad (11)$$

This approach inherently results in the required NN encoding (cf. (12)). Since finding the *globally* optimal partition and codebook of a quantizer is intractable, an alternating optimization technique [23]–[25], [27], [65], [66], [86]–[88] is adopted to derive the necessary optimality conditions. Such conditions serve as a practical means to train each PQ_l , $l = 1, 2$, via principles of the iterative Linde-Buzo-Gray (LBG) algorithm [23], elaborated in Section III-D.

For given codepoints \mathbf{g}_{l,v_l} , $v_l \in \mathcal{V}_l$, the optimal cells which minimize D_l^{PQ} satisfy the NN condition [23]

$$\mathcal{S}_{l,v_l}^* = \left\{ \mathbf{y}_l : \|\mathbf{y}_l - \mathbf{g}_{l,v_l}\|_2^2 \leq \|\mathbf{y}_l - \mathbf{g}_{l,v_l'}\|_2^2, \forall v_l' \neq v_l \right\}, \quad v_l \in \mathcal{V}_l, \quad (12)$$

i.e., the cells form a Voronoi partition [89, Ch. 5.1]. For given cells \mathcal{S}_{l,v_l} , $v_l \in \mathcal{V}_l$, the optimal codepoints satisfy the centroid condition [23]

$$\mathbf{g}_{l,v_l}^* = \mathbb{E}[\mathbf{Y}_l | V_l = v_l] = \frac{1}{p(v_l)} \int_{\mathbf{y}_l \in \mathcal{S}_{l,v_l}} \mathbf{y}_l f(\mathbf{y}_l) d\mathbf{y}_l, \quad v_l \in \mathcal{V}_l, \quad (13)$$

where $f(\mathbf{y}_l)$ is the probability density function of \mathbf{Y}_l .

C. Alternating Optimization of Encoders & Decoder

In this section, the encoders E_1 and E_2 , and decoder D are optimized to minimize $\mathcal{L}_\mu(\underline{D}, \underline{R})$ in (10) for fixed PQ_1 and PQ_2 (designed as described in Section III-B and III-D). The optimization involves six sets of optimization variables: the message index mappings π_1 and π_2 , the rate measures $r(p(i_1))$ and $r(p(i_2))$, and the codevectors \mathbf{c}_{1,i_1,i_2} and \mathbf{c}_{2,i_1,i_2} , $i_1 \in \mathcal{I}_1$, $i_2 \in \mathcal{I}_2$. Due to the intractability of joint optimization, the alternating optimization principles (see Section III-B) are applied, and the necessary optimality conditions for each variable set while keeping the others fixed are derived. The implementation perspective is detailed in Section III-D.

1) Message Index Mappings: Firstly, the minimization of $\mathcal{L}_\mu(\underline{D}, \underline{R})$ over the message index mapping π_l is independent of $r(p(i_{l'}))$, $i_{l'} \in \mathcal{I}_{l'}$, $l \neq l'$. Hence, for fixed $r(p(i_l))$, $i_l \in \mathcal{I}_l$, \mathcal{C}_1 , \mathcal{C}_2 , and $\pi_{l'}$, the optimal message index mapping π_l^* for sensor $l \neq l'$ is the one that minimizes $\mathcal{L}_\mu(\underline{D}, \underline{R})$. Let us start by reformulating the distortion term \underline{D} of $\mathcal{L}_\mu(\underline{D}, \underline{R})$ as

$$\begin{aligned} \underline{D} &= \sum_{l=1}^2 \sum_{v_1 \in \mathcal{V}_1} \sum_{v_2 \in \mathcal{V}_2} \int_{\mathbf{y}_1} \int_{\mathbf{y}_2} p(v_1, v_2 | \mathbf{y}_1, \mathbf{y}_2) \\ &\quad \mathbb{E}[\|\mathbf{X}_l - \mathbf{c}_{l,\pi_1(v_1),\pi_2(v_2)}\|_2^2 | V_1 = v_1, V_2 = v_2, \dots \\ &\quad \mathbf{Y}_1 = \mathbf{y}_1, \mathbf{Y}_2 = \mathbf{y}_2] f(\mathbf{y}_1, \mathbf{y}_2) d\mathbf{y}_1 d\mathbf{y}_2 \\ &\stackrel{(a)}{=} \sum_{l=1}^2 \sum_{v_1 \in \mathcal{V}_1} \sum_{v_2 \in \mathcal{V}_2} \int_{\mathbf{y}_1 \in \mathcal{S}_{1,v_1}} \int_{\mathbf{y}_2 \in \mathcal{S}_{2,v_2}} \\ &\quad \left\{ \mathbb{E}[\|\mathbf{X}_l\|_2^2 | \mathbf{Y}_1 = \mathbf{y}_1, \mathbf{Y}_2 = \mathbf{y}_2] + \|\mathbf{c}_{l,\pi_1(v_1),\pi_2(v_2)}\|_2^2 - \right. \\ &\quad \left. 2\mathbf{c}_{l,\pi_1(v_1),\pi_2(v_2)}^T \mathbb{E}[\mathbf{X}_l | \mathbf{Y}_1 = \mathbf{y}_1, \mathbf{Y}_2 = \mathbf{y}_2] \right\} \\ &\quad f(\mathbf{y}_1, \mathbf{y}_2) d\mathbf{y}_1 d\mathbf{y}_2 \\ &\stackrel{(b)}{=} \sum_{l=1}^2 \sum_{v_1 \in \mathcal{V}_1} \sum_{v_2 \in \mathcal{V}_2} p(v_1, v_2) \\ &\quad \left\{ \mathbb{E}[\|\mathbf{X}_l\|_2^2 | V_1 = v_1, V_2 = v_2] + \|\mathbf{c}_{l,\pi_1(v_1),\pi_2(v_2)}\|_2^2 - \right. \\ &\quad \left. 2\mathbf{c}_{l,\pi_1(v_1),\pi_2(v_2)}^T \mathbf{z}_{l,v_1,v_2} \right\} \end{aligned} \quad (14)$$

where (a) follows from i) the Markov properties $V_l \rightarrow \mathbf{Y}_l \rightarrow \mathbf{Y}_{l'} \rightarrow V_{l'}$, $\mathbf{X}_l \rightarrow (V_1, V_2) \rightarrow \mathcal{C}_l$, $\mathbf{X}_l \rightarrow \mathbf{Y}_l \rightarrow V_l$, and $\mathbf{X}_l \rightarrow \mathbf{Y}_{l'} \rightarrow V_{l'}$, $l \neq l'$, and ii) the fact that $p(v_l | \mathbf{y}_l) = 1$, if $\mathbf{y}_l \in \mathcal{S}_{l,v_l}$, and 0 otherwise, $l = 1, 2$; (b) follows from $p(v_1, v_2) = \int_{\mathbf{y}_1 \in \mathcal{S}_{1,v_1}} \int_{\mathbf{y}_2 \in \mathcal{S}_{2,v_2}} f(\mathbf{y}_1, \mathbf{y}_2) d\mathbf{y}_1 d\mathbf{y}_2$, and the definition of a vector $\mathbf{z}_{l,v_1,v_2} \in \mathbb{R}^N$ as

$$\begin{aligned} \mathbf{z}_{l,v_1,v_2} &\triangleq \mathbb{E}[\mathbf{X}_l | V_1 = v_1, V_2 = v_2], \quad l = 1, 2 \\ &= \frac{1}{p(v_1, v_2)} \int_{\mathbf{y}_1 \in \mathcal{S}_{1,v_1}} \int_{\mathbf{y}_2 \in \mathcal{S}_{2,v_2}} \\ &\quad \mathbb{E}[\mathbf{X}_l | \mathbf{Y}_1 = \mathbf{y}_1, \mathbf{Y}_2 = \mathbf{y}_2] f(\mathbf{y}_1, \mathbf{y}_2) d\mathbf{y}_1 d\mathbf{y}_2, \end{aligned} \quad (15)$$

which represents the *centroid* of the minimum mean square error (MMSE) estimates of source \mathbf{X}_l , $l = 1, 2$, given such

measurement realizations \mathbf{y}_1 and \mathbf{y}_2 that are pre-quantized to the cell index pair (v_1, v_2) , $v_1 \in \mathcal{V}_1$, $v_2 \in \mathcal{V}_2$ at PQ_1 and PQ_2 .

Remark 2. A closed-form expression of the MMSE estimate $\mathbb{E}[\mathbf{X}_l | \mathbf{Y}_1 = \mathbf{y}_1, \mathbf{Y}_2 = \mathbf{y}_2]$, $l = 1, 2$, in (15) has been derived for a similar signal setup in, e.g., [66, Proposition 1]. Nevertheless, as it will turn out later, computation of these complex estimates is obviated by means of the pre-quantizers both in the offline training and online communication phase.

Finally, the rate term \underline{R} in $\mathcal{L}_\mu(\underline{D}, \underline{R})$ can be extended as

$$\begin{aligned} \underline{R} &= \sum_{l=1}^2 \sum_{i_l \in \mathcal{I}_l} p(i_l) r(p(i_l)) \\ &= \sum_{l=1}^2 \sum_{v_l \in \mathcal{V}_l} p(v_l) r(p(\pi_l(v_l))). \end{aligned} \quad (16)$$

Combining (14) and (16), and dropping the terms $\mathbb{E}[\|\mathbf{X}_l\|_2^2 | V_1 = v_1, V_2 = v_2]$ not affecting the minimization, finding the optimal message index mapping $\pi_l^* = \{\pi_l^*(1), \dots, \pi_l^*(|\mathcal{V}_l|)\}$ for each sensor $l = 1, 2$ separates into $|\mathcal{V}_l|$ subproblems. Namely, the optimal message index for a v_1 th cell of sensor 1 is given by the minimization problem

$$\begin{aligned} \pi_1^*(v_1) &= \underset{\pi_1(v_1) \in \mathcal{I}_1}{\operatorname{argmin}} \left\{ (1 - \mu) \sum_{l=1}^2 \sum_{v_2 \in \mathcal{V}_2} p(v_1, v_2) \right. \\ &\quad \left. \left(\|\mathbf{c}_{l, \pi_1(v_1), \pi_2(v_2)}\|_2^2 - 2\mathbf{c}_{l, \pi_1(v_1), \pi_2(v_2)}^T \mathbf{z}_{l, v_1, v_2} \right) + \right. \\ &\quad \left. \mu p(v_1) r(p(\pi_1(v_1))) \right\}, \quad \forall v_1 \in \mathcal{V}_1. \end{aligned} \quad (17)$$

Each optimal message index $\pi_2^*(v_2)$, $v_2 \in \mathcal{V}_2$, for sensor 2 is similarly found by swapping the roles of the sensor indices.

2) *Rate Measures:* The rate measure update follows the procedures in [22], [27]. Accordingly, for fixed π_l , the optimal rate measures for each sensor $l = 1, 2$ are given as

$$\begin{aligned} r^*(p(i_l)) &= -\log_2 p(i_l), \quad \forall i_l \in \mathcal{I}_l, \\ &= -\log_2 \sum_{v_l \in \pi_l^{-1}(i_l)} p(v_l). \end{aligned} \quad (18)$$

By (18), the average sum rate (8) becomes $\underline{R} = H(I_1) + H(I_2)$, where $H(I_l) = -\sum_{i_l=1}^{|\mathcal{I}_l|} p(i_l) \log_2 p(i_l)$ is the entropy of I_l , $l = 1, 2$.

3) *Reconstruction Codebooks:* Firstly, the minimization of $\mathcal{L}_\mu(\underline{D}, \underline{R})$ in (10) with respect to the codebook \mathcal{C}_l is independent of \underline{R} , and $\mathcal{C}_{l'}$, $l' \neq l$. Accordingly, for fixed π_1 and π_2 , the optimal reconstruction codebook $\mathcal{C}_l^* = \{\mathbf{c}_{l,1,1}^*, \dots, \mathbf{c}_{l,|\mathcal{I}_1|,|\mathcal{I}_2|}^*\}$ for source $l = 1, 2$ is found from solving $|\mathcal{I}_1| |\mathcal{I}_2|$ separate optimization problems for each index pair (i_1, i_2) , $i_1 \in \mathcal{I}_1$, $i_2 \in \mathcal{I}_2$, as

$$\mathbf{c}_{l, i_1, i_2}^* = \underset{\mathbf{c}_{l, i_1, i_2} \in \mathbb{R}^N}{\operatorname{argmin}} \mathbb{E} [\|\mathbf{X}_l - \mathbf{c}_{l, i_1, i_2}\|_2^2 | I_1 = i_1, I_2 = i_2]. \quad (19)$$

From (19), each $\mathbf{c}_{l, i_1, i_2}^*$ is given by the MMSE estimate of source \mathbf{X}_l , $l = 1, 2$, given the message index pair (i_1, i_2) [66],

Algorithm 1 Training algorithm for PQ_l , $l = 1, 2$ (Offline).

Inputs: a) CS matrix Φ_l ; b) measurement training vectors $\{\mathbf{y}_l^{(1)}, \mathbf{y}_l^{(2)}, \dots\}$; c) pre-quantization rate \bar{R}_l .

Initialization: Initial codebook $\mathcal{G}_l = \{\mathbf{g}_{l,1}, \dots, \mathbf{g}_{l,|\mathcal{V}_l|}\}$.

Repeat

1) For given \mathcal{G}_l , find the optimal cells \mathcal{S}_{l, v_l}^* , $v_l \in \mathcal{V}_l$, by classifying the vectors $\{\mathbf{y}_l^{(1)}, \mathbf{y}_l^{(2)}, \dots\}$ according to (12).

2) For given \mathcal{S}_{l, v_l} , $v_l \in \mathcal{V}_l$, compute the optimal codepoints \mathbf{g}_{l, v_l}^* , $v_l \in \mathcal{V}_l$, as the conditional expectations in (13).

until convergence

Output: Pre-quantization codebook $\mathcal{G}_l = \{\mathbf{g}_{l,1}, \dots, \mathbf{g}_{l,|\mathcal{V}_l|}\}$.

i.e.,

$$\begin{aligned} \mathbf{c}_{l, i_1, i_2}^* &= \mathbb{E}[\mathbf{X}_l | I_1 = i_1, I_2 = i_2], \quad \forall i_1 \in \mathcal{I}_1, i_2 \in \mathcal{I}_2 \\ &= \sum_{v_1 \in \mathcal{V}_1} \sum_{v_2 \in \mathcal{V}_2} p(v_1, v_2 | i_1, i_2) \\ &\quad \mathbb{E}[\mathbf{X}_l | I_1 = i_1, I_2 = i_2, V_1 = v_1, V_2 = v_2] \\ &\stackrel{(a)}{=} \sum_{v_1 \in \mathcal{V}_1} \sum_{v_2 \in \mathcal{V}_2} \frac{p(v_1, v_2) p(i_1, i_2 | v_1, v_2)}{p(i_1, i_2)} \\ &\quad \mathbb{E}[\mathbf{X}_l | V_1 = v_1, V_2 = v_2] \\ &\stackrel{(b)}{=} \sum_{v_1 \in \pi_1^{-1}(i_1)} \sum_{v_2 \in \pi_2^{-1}(i_2)} \frac{p(v_1, v_2)}{p(i_1, i_2)} \\ &\quad \mathbb{E}[\mathbf{X}_l | V_1 = v_1, V_2 = v_2] \\ &\stackrel{(c)}{=} \frac{\sum_{v_1 \in \pi_1^{-1}(i_1)} \sum_{v_2 \in \pi_2^{-1}(i_2)} p(v_1, v_2) \mathbf{z}_{l, v_1, v_2}}{\sum_{v_1 \in \pi_1^{-1}(i_1)} \sum_{v_2 \in \pi_2^{-1}(i_2)} p(v_1, v_2)}, \end{aligned} \quad (20)$$

where (a) follows from the Markov properties $\mathbf{X}_l \rightarrow V_l \rightarrow I_l$ and $\mathbf{X}_l \rightarrow V_{l'} \rightarrow I_{l'}$, $l \neq l'$; (b) follows from the Markov property $I_l \rightarrow V_l \rightarrow V_{l'} \rightarrow I_{l'}$, $l \neq l'$, and the fact that $p(i_l | v_l) = 1$, if $\pi_l(v_l) = i_l$, and 0 otherwise; (c) follows from (15).

D. Algorithm Implementation

Implementation aspects of the proposed method, termed DQCS-PQ, are elaborated next. Two offline training algorithms, and the algorithm for online communication phase are proposed. Each algorithm's complexity in terms of computational and memory requirements is also analyzed.

1) *Offline Training Phase:* The proposed *offline* training algorithm for optimizing each PQ_l , $l = 1, 2$, is described in Algorithm 1, and the algorithm for optimizing E_1 , E_2 , and D is presented in Algorithm 2. The foundation of both algorithms is the iterative Lloyd algorithm [23], [25]: in Algorithm 1, each PQ_l , $l = 1, 2$, is optimized by successively applying the necessary optimality conditions (12) and (13) (Steps 1) and 2)). Similarly, the optimization of E_1 , E_2 , and D in Algorithm 2 relies on the six-step iteration loop⁸, in which the optimality conditions (17), (18), and (20) (Steps 1) – 3)) are first applied with respect to sensor 1, and then, with respect to sensor 2. Once Algorithm 2 has converged, the resulting message index probabilities $p(i_l)$, $i_l \in \mathcal{I}_l$, $l = 1, 2$, are used to generate the desired binary source codebooks \mathcal{H}_1 and \mathcal{H}_2 .

The requisite training data sets for Algorithms 1 and 2 are generated as follows. Source and noise samples $\{\mathbf{x}_l^{(1)}, \mathbf{x}_l^{(2)}, \dots\}$ and $\{\mathbf{w}_l^{(1)}, \mathbf{w}_l^{(2)}, \dots\}$ are created by sampling

⁸Other optimization orders within the sequential algorithm could be considered as well.

Algorithm 2 Training algorithm for E_1 , E_2 , and D (Offline).

Inputs: a) CS matrices Φ_l ; b) pre-quantization rates \bar{R}_l ; c) codebook sizes $|\mathcal{I}_l|$; d) rate measures $r(p(i_l))$, $i_l \in \mathcal{I}_l$; e) weight parameter $\mu \in [0, 1]$; f) MMSE estimate centroids \mathbf{z}_{l,v_1,v_2} , $v_1 \in \mathcal{V}_1$, $v_2 \in \mathcal{V}_2$, $l = 1, 2$.
Initializations: a) Initial codebooks $\mathcal{C}_l = \{\mathbf{c}_{l,1,1}, \dots, \mathbf{c}_{l,|\mathcal{I}_1|,|\mathcal{I}_2|}\}$, $l = 1, 2$; b) initial message index mapping $\pi_2 = \{\pi_2(1), \dots, \pi_2(|\mathcal{V}_2|)\}$.
Repeat
for $l = 1, 2$ **do**
 1) For given \mathcal{C}_1 , \mathcal{C}_2 , $\pi_{l'}$, $l' \neq l$, and $r(p(i_l))$, $i_l \in \mathcal{I}_l$, find the optimal message index mapping $\pi_l^* = \{\pi_l^*(1), \dots, \pi_l^*(|\mathcal{V}_l|)\}$ via (17).
 2) For given π_l , update the rate measures $r(p(i_l))$, $i_l \in \mathcal{I}_l$, according to (18).
 3) For given π_1 and π_2 , compute the optimal codevectors \mathbf{c}_{l,i_1,i_2}^* , $i_1 \in \mathcal{I}_1$, $i_2 \in \mathcal{I}_2$, as the conditional expectations in (20).
end for
until convergence
 Generate the binary source codebooks \mathcal{H}_l using the index probabilities $p(i_l)$, $i_l \in \mathcal{I}_l$, $l = 1, 2$.
Outputs: Message index mappings $\pi_l = \{\pi_l(1), \dots, \pi_l(|\mathcal{V}_l|)\}$; source codebooks $\mathcal{H}_l = \{\mathbf{h}_{l,1}, \dots, \mathbf{h}_{l,|\mathcal{H}_l|}\}$; reconstruction codebooks $\mathcal{C}_l = \{\mathbf{c}_{l,1,1}, \dots, \mathbf{c}_{l,|\mathcal{I}_1|,|\mathcal{I}_2|}\}$; encoder mappings α_l ; decoder mappings α_l^{-1} and β_l , $l = 1, 2$.

from their respective distributions. Then, the measurement training vectors $\{\mathbf{y}_l^{(1)}, \mathbf{y}_l^{(2)}, \dots\}$ generated according to (2) are used to optimize each PQ $_l$, $l = 1, 2$. The MMSE estimate centroids \mathbf{z}_{l,v_1,v_2} , $v_1 \in \mathcal{V}_1$, $v_2 \in \mathcal{V}_2$, for Algorithm 2 are pre-computed as the conditional expectations (15) using the source samples $\{\mathbf{x}_l^{(1)}, \mathbf{x}_l^{(2)}, \dots\}$, and the indices $\{v_l^{(1)}, v_l^{(2)}, \dots\}$ obtained from quantizing $\{\mathbf{y}_l^{(1)}, \mathbf{y}_l^{(2)}, \dots\}$ via PQ $_l$, $l = 1, 2$. The multi-dimensional integrals in (13) and (20) are similarly evaluated via Monte Carlo integration techniques.

Convergence: The convergence of the iterative descent algorithms in Algorithms 1 and 2 rests on the rationale behind the Lloyd and LBG algorithms [23]–[25]; at each iteration step, the objective function value either decreases or remains the same, and thus, the algorithm converges. However, since only necessary but not the sufficient optimality conditions are met, the resulting quantization system is, at best, locally optimal [22], [23], [28], [88]. LBG type algorithms are known to be sensitive to initialization, making them susceptible to poor local minima. This accentuates the importance of proper initialization of system blocks (i.e., π_2 , \mathcal{C}_1 , and \mathcal{C}_2). One mitigation technique is the splitting method [23], which is applied in the numerical experiments in Section IV. Alternative robust designs include a deterministic annealing method for fixed-rate VQ design in [90], and its extension to entropy-constrained design in [91].

Remark 3. The entropy-constrained design has a built-in tendency to reduce the codebook sizes, especially for large values of μ [22]. Namely, as Algorithm 2 proceeds, some message in-

Algorithm 3 Distributed variable-rate QCS method with complexity-constrained encoding (DQCS-PQ) (Online).

Inputs: a) CS matrices Φ_l ; b) PQ $_l$: codebook \mathcal{G}_l ; c) E_l : mappings π_l and α_l , and codebook \mathcal{H}_l ; d) D: mappings α^{-1} and β_l , and codebooks \mathcal{H}_l and \mathcal{C}_l , $l = 1, 2$.
Separate encoding at sensor $l = 1, 2$:
 (1) Acquisition of CS measurements \mathbf{y}_l according to (2).
 (2) PQ $_l$: Assignment of pre-quantization index v_l^* via (21) using \mathcal{G}_l .
 (3) E_l : a) Assignment of message index as $i_l^* = \pi_l(v_l^*)$;
 b) Assignment of binary codeword as $\mathbf{h}_{l,i_l^*} = \alpha_l(i_l^*)$.
Joint decoding at the sink:
 (4) D: a) Decoding of binary codewords as $(i_1^*, i_2^*) = (\alpha_1^{-1}(\mathbf{h}_{1,i_1^*}), \alpha_2^{-1}(\mathbf{h}_{2,i_2^*}))$; b) Joint reconstruction of source estimates as $\hat{\mathbf{x}}_l = \beta_l(i_1^*, i_2^*) = \mathbf{c}_{l,i_1^*,i_2^*}$ using \mathcal{C}_l , $l = 1, 2$.

dex, say $i_1 \in \mathcal{I}_1$, may become unpopulated (i.e., $\pi_1^{-1}(i_1) = \emptyset$) in Step 1). Consequently, the rate measure in Step 2) becomes $r^*(p(i_1)) = \infty$, and the codevectors $\mathbf{c}_{1,i_1,1}, \dots, \mathbf{c}_{1,i_1,|\mathcal{I}_2|}$ and $\mathbf{c}_{2,i_1,1}, \dots, \mathbf{c}_{2,i_1,|\mathcal{I}_2|}$ in Step 3) become undefined. It was conjectured in [22] that, unlike for a fixed-rate quantizer, there is no rationale to re-include such unassigned indices in the subsequent iterations.

Complexity: The complexity⁹ of optimizing the cells of each PQ $_l$, $l = 1, 2$, via (12) scales as the number of levels $|\mathcal{V}_l|$, and the size of the training data set; the complexity of optimizing the codepoints via (13) scales as $|\mathcal{V}_l|$. The complexity of computing the MMSE estimate centroids of (15) scales as $2|\mathcal{V}_1||\mathcal{V}_2|$, and the size of the training data set. Representing the most demanding training steps, the complexity of optimizing E_1 and E_2 via (17) scales as $|\mathcal{V}_1||\mathcal{V}_2||\mathcal{I}_1|$ and $|\mathcal{V}_1||\mathcal{V}_2||\mathcal{I}_2|$, respectively. The complexity of optimizing D via (20) scales as $2|\mathcal{V}_1||\mathcal{V}_2|$.

2) *Online Communication Phase:* The main operations executed for reconstructing a pair of source realizations \mathbf{x}_1 and \mathbf{x}_2 in the online phase of DQCS-PQ are summarized in Algorithm 3. At each sensor $l = 1, 2$, the measurement vector \mathbf{y}_l is mapped to the (optimal) cell index $v_l^* \in \mathcal{V}_l$ (i.e., the output of PQ $_l$) using $\mathcal{G}_l = \{\mathbf{g}_{l,1}, \dots, \mathbf{g}_{l,|\mathcal{V}_l|}\}$ as

$$v_l^* = \underset{v_l \in \mathcal{V}_l}{\operatorname{argmin}} \|\mathbf{y}_l - \mathbf{g}_{l,v_l}\|_2^2, \quad l = 1, 2. \quad (21)$$

For given $\pi_l = \{\pi_l(1), \dots, \pi_l(|\mathcal{V}_l|)\}$, α_l , and $\mathcal{H}_l = \{\mathbf{h}_{l,1}, \dots, \mathbf{h}_{l,|\mathcal{H}_l|}\}$, the index v_l^* is mapped to the message index $i_l^* = \pi_l(v_l^*)$, and further to the binary codeword $\mathbf{h}_{l,i_l^*} = \alpha_l(i_l^*)$ at E_l . The joint decoder D maps the received codewords to the index pair $(i_1^*, i_2^*) = (\alpha_1^{-1}(\mathbf{h}_{1,i_1^*}), \alpha_2^{-1}(\mathbf{h}_{2,i_2^*}))$, and reconstructs source estimates as $\hat{\mathbf{x}}_l = \beta_l(i_1^*, i_2^*) = \mathbf{c}_{l,i_1^*,i_2^*}$ using $\mathcal{C}_l = \{\mathbf{c}_{l,1,1}, \dots, \mathbf{c}_{l,|\mathcal{I}_1|,|\mathcal{I}_2|}\}$, $l = 1, 2$.

Complexity: Thanks to the imposed NN constraint for each PQ $_l$, $l = 1, 2$, the encoding complexity remains tolerable: a

⁹The computational and memory requirements are assessed through most influential factors by estimating the number of involved training vectors, summands, and variables (i.e., indices) in one optimization step of an algorithm.

sensor has to only calculate $|\mathcal{V}_l|$ distances in (21), followed by the table look-ups associated with π_l , α_l , and \mathcal{H}_l . Adjustment of the pre-quantization rate $\bar{R}_l = \log_2|\mathcal{V}_l|$ allows DQCS-PQ to trade-off between the encoding complexity and compression performance. Clearly, a typical achievable performance range depends on the scaling and operation point of the sensing setup, e.g., via source statistics, and parameters N , M_l , K , Φ_l , and $\sigma_{\mathbf{W}}^2$, $l = 1, 2$. Thus, the need to upscale \bar{R}_l in a high-dimensional setup may expand the VQ look-up table size $2^{\bar{R}_l}$ intolerably high. Consequently, the proposed DQCS-PQ is mainly applicable to setups with moderate signal dimensions. At the cost of performance, the complexity could be restrained by approximate methods, including SQ based approaches, and low-complexity VQ variants like tree-structured, multi-step, and lattice VQs [13].

Remark 4. Even though an MMSE estimate $\mathbb{E}[\mathbf{X}_l|\mathbf{Y}_1 = \mathbf{y}_1, \mathbf{Y}_2 = \mathbf{y}_2]$, $l = 1, 2$, is a constituent quantity in defining the optimal encoders and decoder, the novel pre-quantization technique relieves the sensor of reconstructing it in the online communication phase (see Algorithm 3). This is different from the related works [65], [66], where the necessity of forming these exponentially complex estimates [68]–[70] hinders the practical implementation. Recall that the pre-quantization also obviates the need of computing the estimates in the offline training phase (see Algorithm 2).

3) *Optimal Codeword Lengths:* As the rate approximation (8) merely steers and primes the quantizer design in subsuming a desired source code, one must eventually generate the respective codebooks \mathcal{H}_1 and \mathcal{H}_2 in practice. For the separate coding, the optimal codeword lengths $\gamma(i_l)$, $i_l \in \mathcal{I}_l$, satisfying the Kraft inequality are known. For instance, (instantaneous) Huffman coding [20] [17, Sect. 5.6], and arithmetic coding [85] [17, Sect. 13.3] perform close to the entropy bound $R_l \geq H(I_l)$, $l = 1, 2$ [22], [28].

Remark 5. Improved compression performance is achievable by DSC of the message indices which takes the inter-sensor correlation of I_1 and I_2 into account. The optimal codeword lengths, however, are unknown [28]. Typically, low-density parity-check codes, turbo codes, and syndrome-based codes are employed to approach the Slepian-Wolf boundary [5], [92], [93]. Unfortunately, as the coding involves (large) blocks of indices, the increased complexity and delay of such non-instantaneous codes may become an issue. Potential variants include the Slepian-Wolf type finite-dimensional codes for lossless/near-lossless coding introduced in [94], [95]. Nevertheless, these are outside the scope of this paper.

IV. NUMERICAL RESULTS

Numerical results are presented to illustrate the distortion-rate behavior of the proposed DQCS-PQ method (Algorithm 3), and provide comparison against baseline QCS methods.

A. Simulation Setup

Consider a setup with equal number of measurements $M \triangleq M_1 = M_2$, and equal support probabilities

$p(\mathcal{T}_s) = 1/\binom{N}{K}$, $\forall s = 1, \dots, \binom{N}{K}$. For a given \mathcal{T}_s , the non-zero parts of Θ and Θ_l , denoted as $\bar{\Theta}_{\mathcal{T}_s}$ and Θ_{l,\mathcal{T}_s} , are defined as i.i.d. Gaussian random variables $\bar{\Theta}_{\mathcal{T}_s} \sim \mathcal{N}(\mathbf{0}, \sigma_{\Theta}^2 \mathbf{I}_K)$, and $\Theta_{l,\mathcal{T}_s} \sim \mathcal{N}(\mathbf{0}, \sigma_{\Theta}^2 \mathbf{I}_K)$, $s = 1, \dots, \binom{N}{K}$, $l = 1, 2$. Thus, $\mathbf{X}_{l,\mathcal{T}_s} \sim \mathcal{N}(\mathbf{0}, (\sigma_{\Theta}^2 + \sigma_{\Theta}^2) \mathbf{I}_K)$. The (spatial) correlation between the sensors is adjusted by parameter $\rho^{\text{corr}} \triangleq \sigma_{\Theta}^2 / \sigma_{\Theta}^2$ with $\sigma_{\Theta}^2 = 1$. Two types of measurement matrices Φ_1 and Φ_2 are considered by 1) drawing the entries from Gaussian distribution $\mathcal{N}(0, 1/M)$, and normalizing the columns as $\|\cdot\|_2^2 = 1$, and 2) taking the first (last) M rows of an $N \times N$ DCT-matrix, and normalizing the columns as $\|\cdot\|_2^2 = 1$.

The splitting method¹⁰ [13], [23] is employed in optimizing both the pre-quantizers (Algorithm 1), and the encoders and decoder (Algorithm 2). Unless otherwise stated, $\bar{R}_1 = \bar{R}_2 = 10$ bits, and $\log_2|\mathcal{I}_1| = \log_2|\mathcal{I}_2| = \{1, \dots, 8\}$ bits. For DQCS-PQ, the codebooks \mathcal{H}_1 and \mathcal{H}_2 are generated via the Huffman algorithm, and the average rates are defined via (7). The weight parameter is set by a rule $\mu \triangleq \mu_c / \log_2|\mathcal{I}_l|$, where μ_c is a non-negative constant. When E_1 , E_2 , and D are optimized for multiple (ascending) values of μ_c , the quantities π_2 , \mathcal{C}_1 , and \mathcal{C}_2 obtained for the preceding (smaller) μ_c are used to initialize Algorithm 2 for the next μ_c .

Two baseline QCS methods are considered, which, conforming to the customary approaches, use fixed rate $R_l = \log_2|\mathcal{I}_l|$, and rely on a VQ encoder that minimize MSE quantization distortion:

- 1) VQ-E_S-D_S: a method with separate NN encoding, and separate decoding, where the VQ encoder of l th sensor minimizes $\sum_{i_l \in \mathcal{I}_l} p(i_l) \mathbb{E}[\|\mathbf{Y}_l - \hat{\mathbf{y}}_{l,i_l}\|_2^2 | I_l = i_l]$ with encoder codepoints $\hat{\mathbf{y}}_{l,i_l} \in \mathbb{R}^M$, and the decoder consists of the MSE-optimal decoder codevectors $\mathbf{c}_{l,i_l}^{\text{E-S-D}_S} = \mathbb{E}[\mathbf{X}_l | I_l = i_l]$, $i_l \in \mathcal{I}_l$, for the associated Voronoi regions, $l = 1, 2$.
- 2) VQ-E_S-D_J: a method with separate encoding identical to VQ-E_S-D_S, but with joint decoding via decoder codevectors $\mathbf{c}_{l,i_1,i_2}^{\text{E-S-D}_J} = \mathbb{E}[\mathbf{X}_l | I_1 = i_1, I_2 = i_2]$, $i_1 \in \mathcal{I}_1$, $i_2 \in \mathcal{I}_2$, $l = 1, 2$.

The distortion is measured as $D_{\text{ave}} \triangleq 10 \log_{10} (D_1 / \mathbb{E}[\|\mathbf{X}_1\|_2^2] + D_2 / \mathbb{E}[\|\mathbf{X}_2\|_2^2])$ (dB), where $\mathbb{E}[\|\mathbf{X}_l\|_2^2] = \sum_{s=1}^{\binom{N}{K}} p(\mathcal{T}_s) \mathbb{E}[\|\mathbf{X}_{l,\mathcal{T}_s}\|_2^2] = K(\sigma_{\Theta}^2 + \sigma_{\Theta}^2)$, $l = 1, 2$. The rate is measured as $R_{\text{ave}} \triangleq \bar{R}/2$ (bits). Due to the exponential complexity of quantizer parts in the offline training phase, the experiments are confined to moderate quantization rates and signal dimensions. Data sets of size 5×10^5 were used for training and performance evaluation.

B. An Illustrative Example

For visualizing the general operation of DQCS-PQ, consider a low-dimensional setup with $N = 3$, $M = 2$, $K = 1$,

¹⁰By the splitting procedure, a quantization system is first optimized for codebook sizes $|\mathcal{I}_l| = 2$, $l = 1, 2$ ($|\mathcal{V}_l|$ for PQ_l). Then, each resultant codevector is split through small perturbation, and the algorithm is run for $|\mathcal{I}_l| = 4$ using the new codevectors as initial ones (i.e., "warm start"). This gradual bifurcation is repeated until the desired codebook sizes are reached.

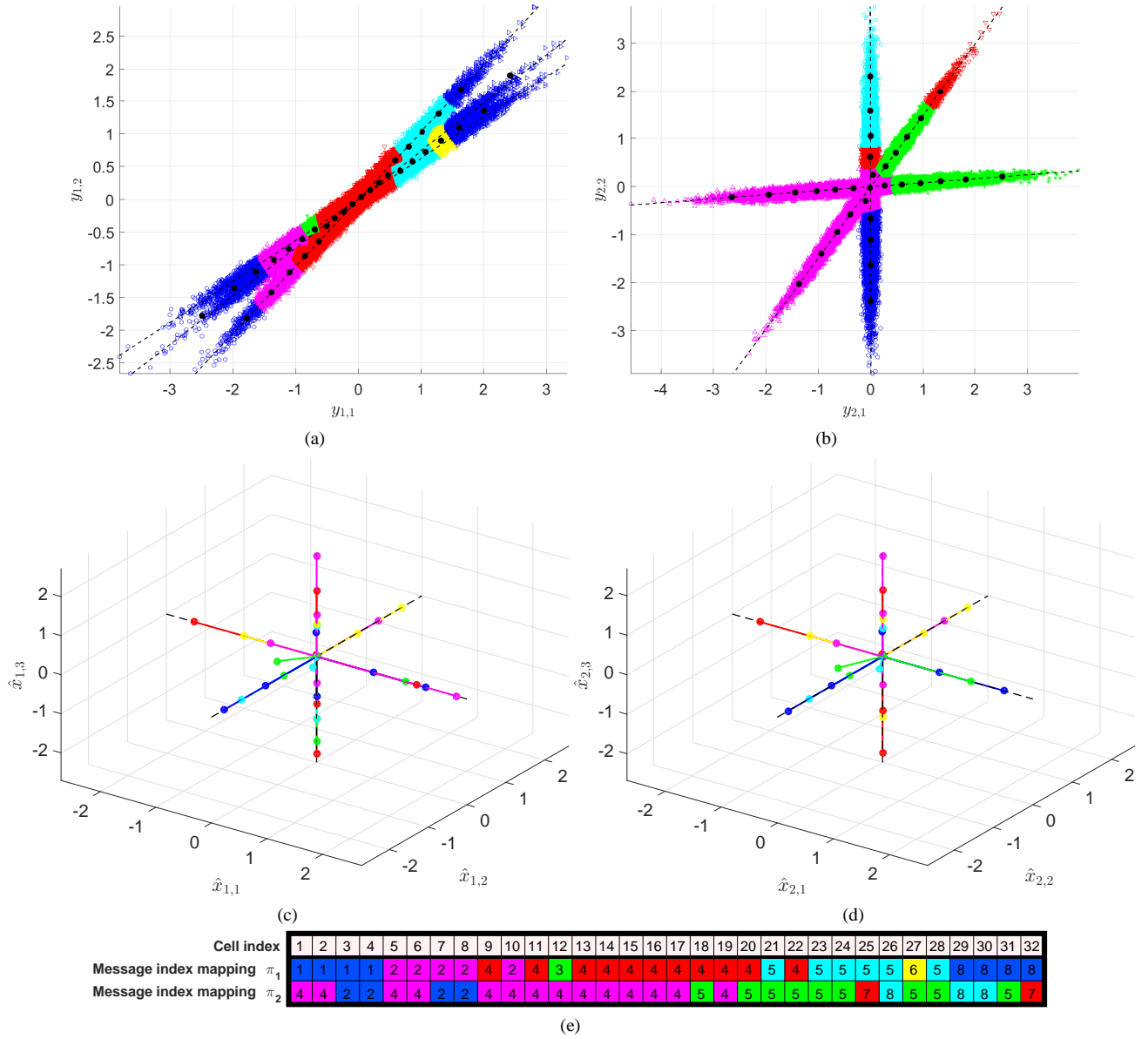


Fig. 3. Illustration of DQCS-PQ in terms of (a) pre-quantizer PQ_1 and encoder E_1 , (b) pre-quantizer PQ_2 and encoder E_2 , (c) reconstruction codebook C_1 , (d) reconstruction codebook C_2 , and (e) message index mappings π_1 and π_2 . In (a) and (b), the black dots represent the codepoints $g_{l,1}, \dots, g_{l,32}$ of PQ_1 and PQ_2 ; each dotted line represents the sparsity-dependent span $\phi_{l,n}X_{l,n}$, $n = 1, 2, 3$; each color indicates the set $\pi_l^{-1}(i_l)$ for message index $i_l \in \mathcal{I}_l$, $l = 1, 2$ (the colors are equivalent in (e)). In (c) and (d), the dots represent the codevectors $c_{l,1,1}, \dots, c_{l,8,8}$; the coordinate axes are shown as dashed lines.

$\sigma_{\mathbf{W}}^2 = 0.003$, $\rho^{\text{corr}} = 10^2$, Gaussian measurement matrices

$$\begin{aligned} \Phi_1 &= \begin{bmatrix} 0.8472 & 0.8044 & -0.6982 \\ 0.5312 & 0.5941 & -0.7159 \\ 0.9966 & -0.5589 & -0.0016 \end{bmatrix}, \\ \Phi_2 &= \begin{bmatrix} 0.0827 & -0.8293 & 1.0000 \end{bmatrix}, \end{aligned} \quad (22)$$

and quantization parameters $\mu_c = 0.5$, $|\mathcal{V}_l| = 32$, and $|\mathcal{I}_l| = 8$. The encoder/decoder sides optimized via Algorithms 1 and 2 are depicted in Fig. 3: (a) and (b) illustrate the outcomes of pre-quantization of \mathbf{Y}_1 and \mathbf{Y}_2 , and the subsequent message index mappings π_1 and π_2 , respectively; (c) and (d) depict the reconstruction codebooks C_1 and C_2 , respectively; (e) tabulates π_1 and π_2 . The source codebooks \mathcal{H}_1 and \mathcal{H}_2 are presented in Table I. To get a better understanding on the main assets of DQCS-PQ, VQ-ES-DJ is similarly illustrated in Fig. 4.

Firstly, the resulting performance figures for DQCS-PQ are $D_{\text{ave}} = -10.19$ dB and $R_{\text{ave}} = 1.83$ bits, and $D_{\text{ave}} = -4.475$ dB and $R_{\text{ave}} = 3$ bits for VQ-ES-DJ. This implies a striking 39 % reduction of rate, and 5.7 dB decrease of distortion in favor of DQCS-PQ.

The key function of PQ_1 and PQ_2 , "encoder shaping", can be perceived from Fig. 3(a) and (b). Take as examples the message index $I_1 = 2$ and $I_2 = 7$: the 5 cells assigned to $I_1 = 2$, i.e., $\pi_1^{-1}(2) = \{5, 6, 7, 8, 10\}$, form an irregularly shaped region of measurement vectors, whereas the region for $I_2 = 7$ consists of 2 discontinuous cells¹¹ by $\pi_2^{-1}(7) = \{25, 32\}$. This shaping feature allows to merge the cells $\mathcal{S}_{l,1}, \dots, \mathcal{S}_{l,32}$ (i.e., classify measurement (training) vectors) into message indices

¹¹This can be interpreted as an index reuse, characteristic to a distributed quantization setup [88].

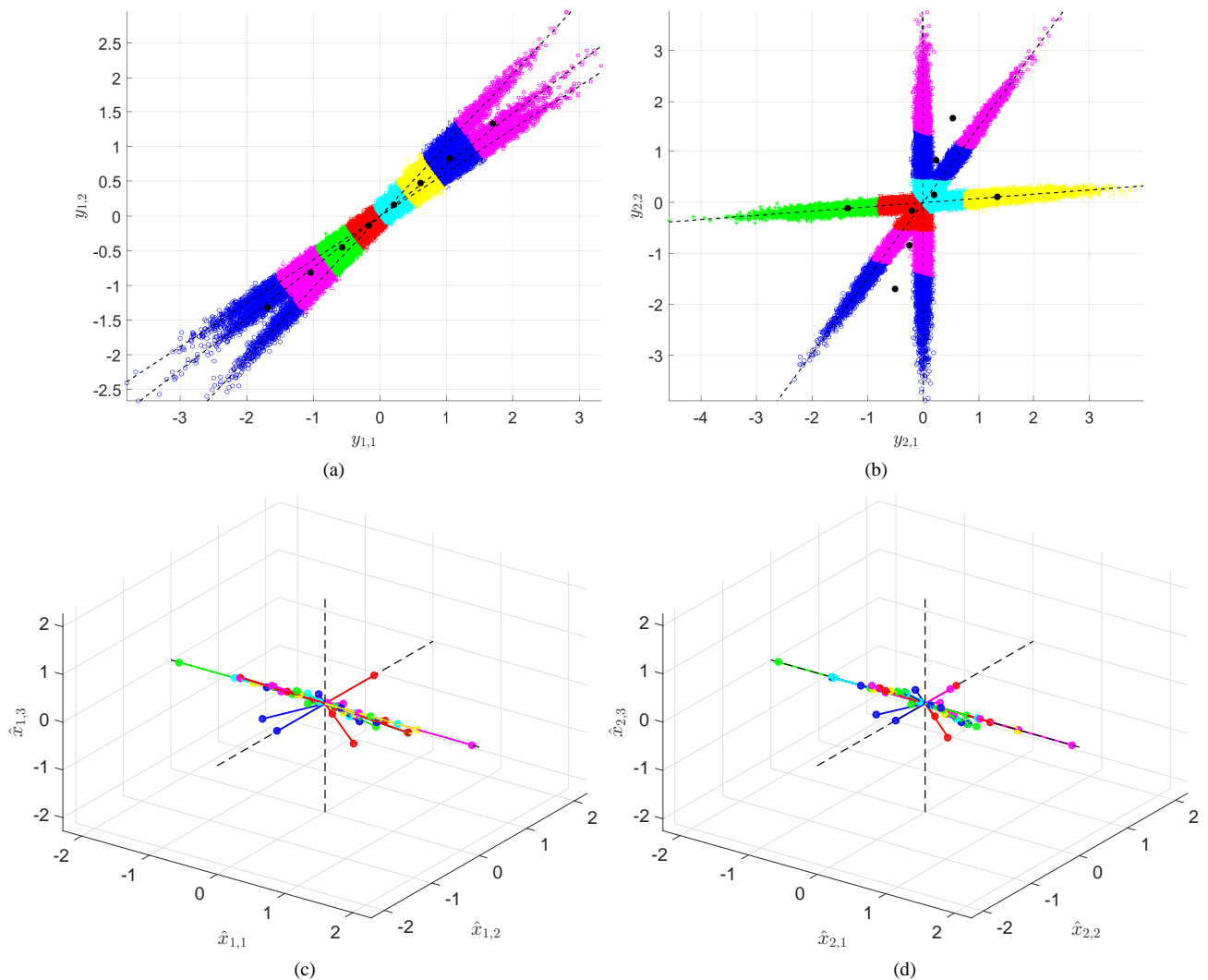


Fig. 4. Illustration of VQ-E_S-D_J in terms of (a) encoder at sensor 1, (b) encoder at sensor 2, (c) decoder codevectors for \mathbf{X}_1 , and (d) decoder codevectors for \mathbf{X}_2 . The lines, dots, and colors have the meanings equivalent to those in Fig. 3.

TABLE I
 HUFFMAN SOURCE CODEBOOKS \mathcal{H}_1 AND \mathcal{H}_2 OF DQCS-PQ USED IN SECTION IV-B

I_1/I_2	Codebook \mathcal{H}_1	Codebook \mathcal{H}_2
1	10010	-
2	101	101
3	10000	-
4	0	0
5	11	11
6	10011	-
7	-	1001
8	10001	1000

$\mathcal{I}_l = \{1, \dots, 8\}$, $l = 1, 2$, by taking the collective effect of the CS and DSC into account. As a consequence, the partitioning in Fig. 3(a) follows to some extent the sparsity-dependent linear projections $\phi_{l,n}X_{l,n} + \mathbf{W}_l$, where $\phi_{l,n} \in \mathbb{R}^M$ is n th column of Φ_l , and $X_{l,n}$ is n th element of \mathbf{X}_l , $n = 1, 2, 3$. Similarly, Fig. 3(c) and (d) present that nearly all reconstruction codevectors are (approximately) K -sparse vectors. As they are also rather evenly distributed in \mathbb{R}^3 , the codebooks \mathcal{C}_1 and \mathcal{C}_2 presumably contain accurate estimates for the K -sparse

sources \mathbf{X}_1 and \mathbf{X}_2 . On the contrary, the encoder regions of VQ-E_S-D_J in Fig. 4(a) are solely based on the probability mass of \mathbf{Y}_1 , and the codevectors in Fig. 4(c) and (d) are dispersed with no K -sparse structure. Such incognizance of the CS resulted in poor performance.

Table I illustrates the power of entropy coding in DQCS-PQ: the codebooks \mathcal{H}_1 and \mathcal{H}_2 contain only 7 and 5 codewords (instead of 8), respectively, and the expected codeword length is minimized by assigning the most frequent indices $I_1 = I_2 = 4$ the shortest codewords "0", the second most frequent indices $I_1 = I_2 = 5$ the codewords "11", and so forth. Note that even though \mathcal{H}_1 contains 5-bit codewords, R_{ave} for DQCS-PQ is 39 % lower (and D_{ave} 5.7 dB lower) than for VQ-E_S-D_J that uses fixed-length 3-bit codewords.

C. Distortion-Rate Performance

Fig. 5(a) depicts the average distortion D_{ave} versus the average rate R_{ave} for $N = 20$, $M = 8$, $K = 2$, $\sigma_{\mathbf{W}}^2 = 0.01$, $\rho^{\text{corr}} = 10^3$, and the Gaussian type Φ_1 and Φ_2 . By completely disregarding the signal correlation, VQ-E_S-D_S performs poorly. Significant gains are achieved by VQ-E_S-D_J

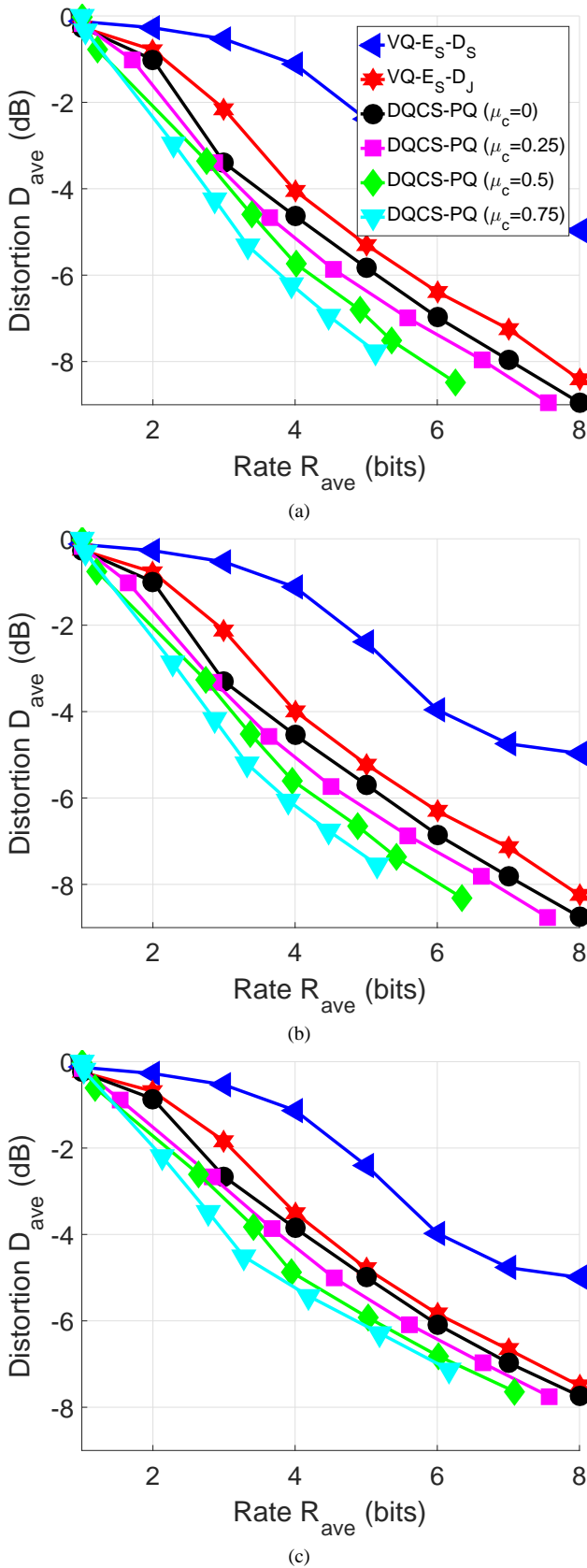


Fig. 5. Average distortion vs. average rate for $N = 20$, $M = 8$, $K = 2$, the Gaussian type Φ_1 and Φ_2 , and signal correlation parameter (a) $\rho^{corr} = 10^3$, (b) $\rho^{corr} = 10^2$, and (c) $\rho^{corr} = 10^1$. The colors and markers of the curves in (b) and (c) are equivalent to those in (a).

for which the joint decoding effectively expands the codebook size from $|\mathcal{I}_l|$ to $|\mathcal{I}_l|^2$ (especially for high correlation). However, since the encoders of both methods are blind to the peculiarities of the CS, and they use fixed rates, the proposed variable-rate CS-aware DQCS-PQ invariably obtains the best compression performance for all values of μ_c , including the fixed-rate version ($\mu_c = 0$). Similar performance trend carries over all the subsequent experiments as well.

It can be seen that VQ- E_S - D_S saturates around the distortion $D_{ave} = -5$ dB. Such an unavoidable constant error level is caused by the remote source coding setup in CS, and, thus, also occurs for the other methods for sufficiently high rates. Namely, regardless of the compression/coding method, the CS parts with additive measurement noise at the sensors prevent the perfect reconstruction of \mathbf{X}_1 and \mathbf{X}_2 at the decoder, even for the rates approaching infinity. For VQ- E_S - D_S , the error level is defined by the rate-independent MMSE estimation error of the form $\sum_{l=1}^2 \int \mathbb{E} [\|\mathbf{X}_l - \mathbb{E}[\mathbf{X}_l | \mathbf{Y}_l = \mathbf{y}_l]\|_2^2] f(\mathbf{y}_l) d\mathbf{y}_l$. For VQ- E_S - D_J and DQCS-PQ, the level is given by $\sum_{l=1}^2 \int \int \mathbb{E} [\|\mathbf{X}_l - \mathbb{E}[\mathbf{X}_l | \mathbf{Y}_1 = \mathbf{y}_1, \mathbf{Y}_2 = \mathbf{y}_2]\|_2^2] f(\mathbf{y}_1, \mathbf{y}_2) d\mathbf{y}_1 d\mathbf{y}_2$.

D. Influence of Weight Parameter μ

The impact of weight parameter μ in DQCS-PQ can be perceived from Fig. 5(a). By choosing an appropriate variable-rate code, DQCS-PQ can flexibly trade-off the distortion-rate performance, and thus, lends itself to varying compression scenarios. For instance, consider a rate-limited WSN application, where the maximum (source) rate is restricted to, say, $R_{ave} = 5$ bits due to, e.g., bad channel conditions, or congested data traffic. Via entropy coding of the message indices, the reconstruction performance is improved from $D_{ave} = -5.8$ dB ($\mu_c = 0$) to $D_{ave} = -7.6$ dB ($\mu_c = 0.75$). Alternatively, consider an application with minimum required reconstruction fidelity of, say, $D_{ave} = -7$ dB. With the aid of variable-length coding, a significant rate reduction of 1.5 bits from $R_{ave} = 6.0$ bits ($\mu_c = 0$) to $R_{ave} = 4.5$ bits ($\mu_c = 0.75$) is achieved.

E. Influence of Source Correlation

Fig. 5(a) – (c) show the compression performance for different signal correlation levels $\rho^{corr} = \{10^3, 10^2, 10^1\}$. As a first remark, DQCS-PQ utilizes the signal correlation efficiently for compressing the sparse sources. The higher the correlation, the greater the compression gain is in favor of DQCS-PQ. For $\rho^{corr} = 10^3$ (high correlation), the distortion $D_{ave} = -7$ dB is achieved at rates $R_{ave} = 6.7$ bits for VQ- E_S - D_J , and 4.5 bits for DQCS-PQ ($\mu_c = 0.75$). For $\rho^{corr} = 10^1$ (low correlation), the respective rates are $R_{ave} = 7.4$ and 6.0 bits. Note that whereas VQ- E_S - D_J and DQCS-PQ perform better with the increasing correlation, the indiscernible disparities of the curves of fully separate VQ- E_S - D_S are solely attributed to different signal powers $\mathbb{E}[\|\mathbf{X}_l\|_2^2]$.

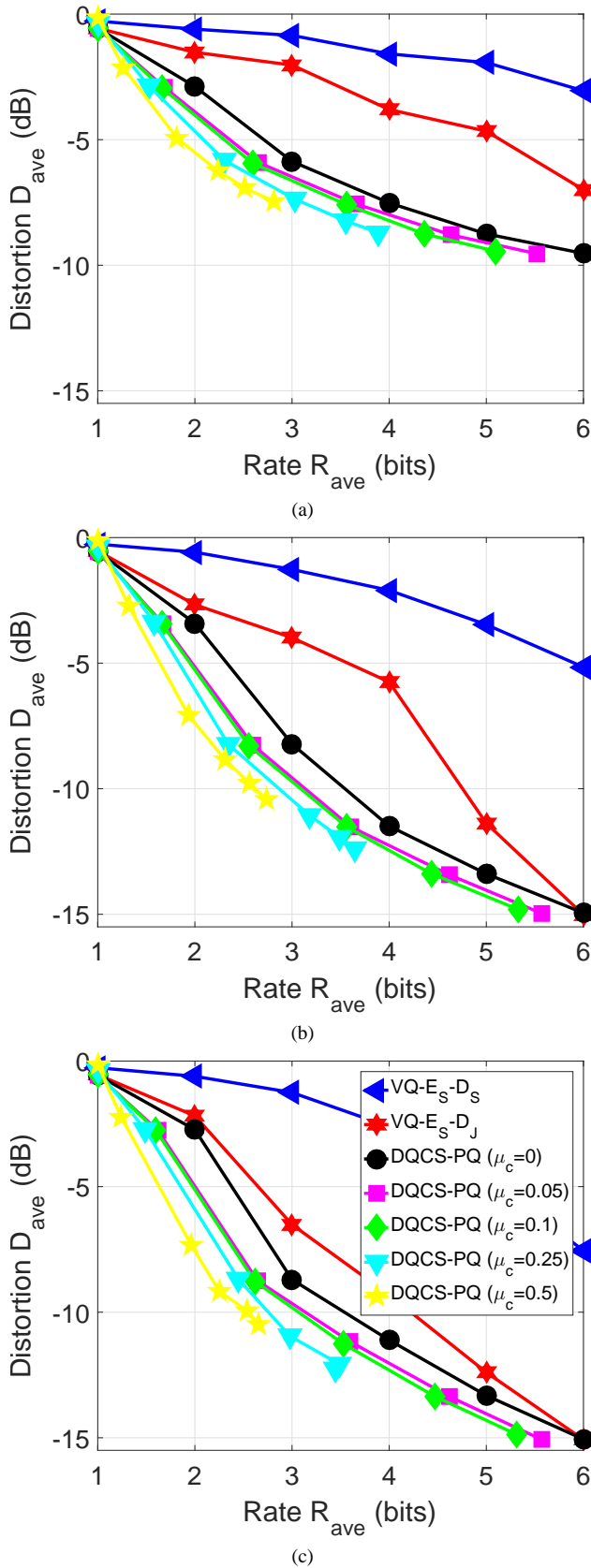


Fig. 6. Average distortion vs. average rate for $N = 10$, $K = 1$, $\rho^{\text{corr}} = 10^2$, the DCT type Φ_1 and Φ_2 , and number of CS measurements (a) $M = 2$, (b) $M = 3$, and (c) $M = 4$. The colors and markers of the curves in (a) and (b) are equivalent to those in (c).

F. Influence of Number of Measurements

Fig. 6 demonstrates the influence of different number of CS measurements $M = \{2, 3, 4\}$ in the setup with $N = 10$, $K = 1$, $\sigma_{\mathbf{W}}^2 = 0.01$, $\rho^{\text{corr}} = 10^2$, and the DCT type Φ_1 and Φ_2 . Unlike the other methods, DQCS-PQ achieves decent performance even in a very noisy scenario (i.e., for $M = 2$). By increasing the measurements to $M = 3$, and further to $M = 4$, the setup becomes less contaminated which allows each QCS method to compress the sparse sources more reliably. However, whereas VQ-E_S-D_S and VQ-E_S-D_J obtain considerable gain when switching from $M = 3$ to $M = 4$, the respective improvement for DQCS-PQ is negligible. This is congruent with the CS philosophy: when M is sufficient for successful support recovery, a further increase of M does not provide significant gains. Accordingly, given that M is at a satisfactory level, it is more cost-effective to improve the reconstruction quality by increasing the rate. Since acquiring more measurements can be expensive – or even infeasible – in practice, the capability to operate at low signal-to-noise ratios is indisputably an advantageous feature of DQCS-PQ.

V. CONCLUSIONS AND FUTURE WORK

A distortion-rate optimized variable-rate distributed QCS method was developed for efficient and complexity-constrained acquisition of correlated sparse sources in WSNs. The proposed DSC method minimizes a weighted sum of the average MSE signal reconstruction distortion, and the average rate. As a specific target to ameliorate the practical feasibility, the encoding complexity of each sensor is restrained by pre-discretizing the measurement vector space via VQ. Following the entropy-constrained design, the method incorporates a desired variable-rate code by means of rate measures of an entropy bound to enhance the compression performance. Alternating optimization was used to derive necessary optimality conditions and propose practical training algorithms for a two-sensor system. The complexity aspects of training and communication phases were discussed. Based on the numerical results, the proposed method has superior distortion-rate performance under varying signal correlation levels and signal-to-noise ratios. Moreover, the method is adaptable to distinct compression settings with, e.g., stringent rate or distortion requirements. As the key finding, efficient finite-rate acquisition of correlated sparse sources calls for 1) CS-awareness, 2) DSC design, and 3) entropy coding.

While VQ for pre-quantization is justifiable owing to the coding optimality [89, Sect. 3.4], its applicability is limited to moderate signal dimensions/quantization rates. As this is the first work on distortion-rate optimized QCS, sub-optimal quantization methods necessary for high-dimensional setups (see Section III-D) were left for future studies. Potential extensions to enhance the compression performance include DSC of message indices (see Remark 5), and designing the pre-quantizers jointly with the encoders and decoder. While JSM-2 inspired by timely WSN applications like collaborative spectrum sensing is used as the signal model, other signal types including JSM-1 and JSM-3 in [49], [54], and common sparse signals in [77]–[79] are straightforwardly applicable.

However, by restricting to JSM-2 signals, the conducted experiments provide distinct insights and benchmarks for future studies – both for experimental and theoretical ones. In fact, as the signal characteristics were not explicitly used in the derivations, the generality extends beyond sparse signals and CS-based sensors, i.e., the proposed design can be applied for a general distributed joint estimation and compression scheme.

As the design adheres solely to source coding, reliable communications of encoder outputs in a practical noisy channel setup necessitates channel coding. For a mere concatenation of the devised source encoders and channel encoders, a known downside is the catastrophic effect of channel errors [96]: a few erroneously received bits of a variable-length code induce error propagation. On the theoretical side, while separate source and channel coding is asymptotically optimal in a point-to-point setup owing to the celebrated Shannon's separation theorem [3, Theorem 21], the result does not hold in general multi-user networks [97, Sect. 1.4.4]. Accordingly, a viable solution may be obtained via distributed joint source-channel coding [96], [98]–[100].

REFERENCES

- [1] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of things for smart cities," *IEEE Internet of Things J.*, vol. 1, no. 1, pp. 22–32, Feb. 2014.
- [2] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, pp. 393–422, 2002.
- [3] C. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, Jul. 1948.
- [4] —, "Coding theorems for a discrete source with a fidelity criterion," *IRE Nat. Conv. Rec.*, vol. 7, pp. 142–163, 1959.
- [5] Z. Xiong, A. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Processing Mag.*, vol. 21, no. 5, pp. 80–94, Sep. 2004.
- [6] S. S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microsensor network," *IEEE Signal Processing Mag.*, vol. 19, no. 2, pp. 51–60, Mar. 2002.
- [7] P. L. Dragotti and M. Gastpar, *Distributed Source Coding: Theory, Algorithms and Applications*. Academic Press, 2009.
- [8] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. 19, no. 4, pp. 471–480, Jul. 1973.
- [9] W. Bennett, "Spectra of quantized signals," *Bell System Technical Journal*, vol. 27, no. 3, pp. 446–472, Jul. 1948.
- [10] B. Oliver, J. Pierce, and C. Shannon, "The philosophy of PCM," *Proceedings of the IRE*, vol. 36, no. 11, pp. 1324–1331, Nov. 1948.
- [11] H. Gish and J. Pierce, "Asymptotically efficient quantizing," *IEEE Trans. Inform. Theory*, vol. 14, no. 5, pp. 676–683, Sep. 1968.
- [12] A. Gersho, "Asymptotically optimal block quantization," *IEEE Trans. Inform. Theory*, vol. 25, no. 4, pp. 373–380, Jul. 1979.
- [13] R. Gray, "Vector quantization," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 1, no. 2, pp. 4–29, Apr. 1984.
- [14] R. Gray and D. Neuhoff, "Quantization," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.
- [15] T. Berger and J. Gibson, "Lossy source coding," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2693–2723, Oct. 1998.
- [16] T. Berger, *Rate-Distortion Theory: A Mathematical Basis for Data Compression*, ser. Prentice-Hall Series in Information and System Sciences. Prentice Hall, 1971.
- [17] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. New York, USA: John Wiley, 2006.
- [18] R. Gallager, *Information Theory and Reliable Communication*. New York, USA: John Wiley & Sons, Inc., 1968.
- [19] R. Gray, D. Neuhoff, and J. Omura, "Process definitions of distortion-rate functions and source coding theorems," *IEEE Trans. Inform. Theory*, vol. 21, no. 5, pp. 524–532, Sep. 1975.
- [20] D. Huffman, "A method for the construction of minimum-redundancy codes," *Proc. IRE*, vol. 40, no. 9, pp. 1098–1101, Sep. 1952.
- [21] J. Ziv and A. Lempel, "Compression of individual sequences via variable-rate coding," *IEEE Trans. Inform. Theory*, vol. 24, no. 5, pp. 530–536, Sep. 1978.
- [22] P. Chou, T. Lookabaugh, and R. Gray, "Entropy-constrained vector quantization," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 1, pp. 31–42, Jan. 1989.
- [23] Y. Linde, A. Buzo, and R. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. 28, no. 1, pp. 84–95, Jan. 1980.
- [24] J. Max, "Quantizing for minimum distortion," *Proc. IRE*, vol. 6, no. 1, pp. 7–12, Mar. 1960.
- [25] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inform. Theory*, vol. 28, no. 2, pp. 129–137, Mar. 1982.
- [26] A. Wagner, S. Tavildar, and P. Viswanath, "Rate region of the quadratic gaussian two-encoder source-coding problem," *IEEE Trans. Inform. Theory*, vol. 54, no. 5, pp. 1938–1961, May 2008.
- [27] D. Rebollo-Monedero, R. Zhang, and B. Girod, "Design of optimal quantizers for distributed source coding," in *Proc. IEEE Data Compr. Conf.*, Snowbird, UT, USA, Mar. 25–27 2003, pp. 13–22.
- [28] M. Fleming, Q. Zhao, and M. Effros, "Network vector quantization," *IEEE Trans. Inform. Theory*, vol. 50, no. 8, pp. 1584–1604, Aug. 2004.
- [29] D. Rebollo-Monedero, "Quantization and transforms for distributed source coding," Ph.D. dissertation, Stanford University, Dec. 2007.
- [30] R. Dobrushin and B. Tsybakov, "Information transmission with additional noise," *IRE Trans. Inform. Theory*, vol. 8, no. 5, pp. 293–304, Sep. 1962.
- [31] Y. Ephraim and R. M. Gray, "A unified approach for encoding clean and noisy sources by means of waveform and autoregressive model vector quantization," *IEEE Trans. Inform. Theory*, vol. 34, no. 4, pp. 826–834, Jul. 1988.
- [32] J.-J. Xiao, A. Ribeiro, Z.-Q. Luo, and G. Giannakis, "Distributed compression-estimation using wireless sensor networks," *IEEE Signal Processing Mag.*, vol. 23, no. 4, pp. 27–41, Jul. 2006.
- [33] H. Yamamoto and K. Itoh, "Source coding theory for multiterminal communication systems with a remote source," *Trans. IECE Japan*, vol. 63, no. 10, pp. 700–706, Oct. 1980.
- [34] T. Flynn and R. Gray, "Encoding of correlated observations," *IEEE Trans. Inform. Theory*, vol. 33, no. 6, pp. 773–787, Nov. 1987.
- [35] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem," *IEEE Trans. Inform. Theory*, vol. 42, no. 3, pp. 887–902, May 1996.
- [36] Y. Oohama, "Indirect and direct gaussian distributed source coding problems," *IEEE Trans. Inform. Theory*, vol. 60, no. 12, pp. 7506–7539, Dec. 2014.
- [37] W.-M. Lam and A. Reibman, "Design of quantizers for decentralized estimation systems," *IEEE Trans. Commun.*, vol. 41, no. 11, pp. 1602–1605, Nov. 1993.
- [38] J. A. Gubner, "Distributed estimation and quantization," *IEEE Trans. Inform. Theory*, vol. 39, no. 4, pp. 1456–1459, Jul. 1993.
- [39] D. Rebollo-Monedero and B. Girod, "Network distributed quantization," in *Proc. IEEE Inform. Theory Workshop*, Tahoe City, CA, USA, Sep. 2–6 2007, pp. 426–431.
- [40] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [41] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [42] E. Candes, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.
- [43] J. Haupt and R. Nowak, "Signal reconstruction from noisy random projections," *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 4036–4048, Sep. 2006.
- [44] E. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
- [45] E. Candes, "Compressive sampling," in *Proceedings of the International Congress of Mathematicians*, vol. 3, Madrid, Spain, Aug. 22–30 2006, pp. 1433–1452.
- [46] M. Duarte, M. Wakin, D. Baron, and R. Baraniuk, "Universal distributed sensing via random projections," in *Proc. IEEE Int. Symp. on Inform. Proc. in Sensor Networks*, New York, NY, USA, 2006, pp. 177–185.
- [47] R. Baraniuk, "Compressive sensing [lecture notes]," *IEEE Signal Processing Mag.*, vol. 24, no. 4, pp. 118–121, Jul. 2007.
- [48] E. Candes and M. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

- [49] D. Baron, M. Wakin, M. Duarte, S. Sarvotham, and R. Baraniuk, "Distributed compressed sensing," Dept. Elect. Comput. Eng., Rice University, Houston, TX, Tech. Rep., Nov. 2005.
- [50] G. Quer, R. Masiero, G. Pillonetto, M. Rossi, and M. Zorzi, "Sensing, compression, and recovery for WSNs: Sparse signal modeling and monitoring framework," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3447–3461, Oct. 2012.
- [51] M. Leinonen, M. Codreanu, and M. Juntti, "Sequential compressed sensing with progressive signal reconstruction in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1622–1635, Mar. 2015.
- [52] D. Malioutov, M. Çetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Processing*, vol. 53, no. 8, pp. 3010–3022, Aug. 2005.
- [53] S. K. Sharma, E. Lagunas, S. Chatzinotas, and B. Ottersten, "Application of compressive sensing in cognitive radio communications: A survey," *IEEE Commun. Surveys & Tutorials*, vol. 18, no. 3, pp. 1838–1860, Third quarter 2016.
- [54] M. Duarte, M. Wakin, D. Baron, S. Sarvotham, and R. Baraniuk, "Measurement bounds for sparse signal ensembles via graphical models," *IEEE Trans. Inform. Theory*, vol. 59, no. 7, pp. 4280–4289, Jul. 2013.
- [55] V. Goyal, A. Fletcher, and S. Rangan, "Compressive sampling and lossy compression," *IEEE Signal Processing Mag.*, vol. 25, no. 2, pp. 48–56, 2008.
- [56] J. Sun and V. Goyal, "Optimal quantization of random measurements in compressed sensing," in *Proc. IEEE Int. Symp. Inform. Theory*, Seoul, Korea, Jun. 28 – Jul.3 2009, pp. 6–10.
- [57] A. Zymnis, S. Boyd, and E. Candes, "Compressed sensing with quantized measurements," *IEEE Signal Processing Lett.*, vol. 17, no. 2, pp. 149–152, 2010.
- [58] L. Jacques, D. Hammond, and J. Fadili, "Dequantizing compressed sensing: When oversampling and non-Gaussian constraints combine," *IEEE Trans. Inform. Theory*, vol. 57, no. 1, pp. 559–571, 2011.
- [59] U. Kamilov, V. Goyal, and S. Rangan, "Message-passing dequantization with applications to compressed sensing," *IEEE Trans. Signal Processing*, vol. 60, no. 12, pp. 6270–6281, Dec. 2012.
- [60] A. Shirazinia, S. Chatterjee, and M. Skoglund, "Analysis-by-synthesis quantization for compressed sensing measurements," *IEEE Trans. Signal Processing*, vol. 61, no. 22, pp. 5789–5800, Nov. 2013.
- [61] W. Dai and O. Milenkovic, "Information theoretical and algorithmic approaches to quantized compressive sensing," *IEEE Trans. Commun.*, vol. 59, no. 7, pp. 1857–1866, 2011.
- [62] G. Coluccia, E. Magli, A. Roumy, and V. Toto-Zarasoa, "Lossy compression of distributed sparse sources: A practical scheme," in *Proc. European Sign. Proc. Conf.*, Barcelona, Spain, Aug. 29–Sep. 2 2011, pp. 422–426.
- [63] P. Boufounos, "Universal rate-efficient scalar quantization," *IEEE Trans. Inform. Theory*, vol. 58, no. 3, pp. 1861–1872, Mar. 2012.
- [64] A. Shirazinia, "Source and channel coding for compressed sensing and control," Ph.D. dissertation, KTH Royal Institute of Technology, Apr. 2014. [Online]. Available: <http://kth.diva-portal.org/smash/get/diva2:709596/FULLTEXT01.pdf>
- [65] A. Shirazinia, S. Chatterjee, and M. Skoglund, "Joint source-channel vector quantization for compressed sensing," *IEEE Trans. Signal Processing*, vol. 62, no. 14, pp. 3667–3681, Jul. 2014.
- [66] —, "Distributed quantization for measurement of correlated sparse sources over noisy channels," *IEEE Trans. Signal Processing, submitted*, p. 12, 2015, available at <http://arxiv.org/abs/1404.7640v2>.
- [67] —, "Distributed quantization for compressed sensing," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Florence, Italy, May 4–9 2014, pp. 6439–6443.
- [68] M. Elad and I. Yavneh, "A plurality of sparse representations is better than the sparsest one alone," *IEEE Trans. Inform. Theory*, vol. 55, no. 10, pp. 4701–4714, Oct. 2009.
- [69] M. Protter, I. Yavneh, and M. Elad, "Closed-form MMSE estimation for signal denoising under sparse representation modeling over a unitary dictionary," *IEEE Trans. Signal Processing*, vol. 58, no. 7, pp. 3471–3484, Jul. 2010.
- [70] J. S. Turek, I. Yavneh, and M. Elad, "On MMSE and MAP denoising under sparse representation modeling over a unitary dictionary," *IEEE Trans. Signal Processing*, vol. 59, no. 8, pp. 3526–3535, Aug. 2011.
- [71] M. Leinonen, M. Codreanu, and M. Juntti, "Channel-robust compressed sensing via vector pre-quantization in wireless sensor networks," in *Proc. IEEE Global Conf. on Signal and Inform. Proc.*, Orlando, Florida, USA, Dec. 14–16 2015, pp. 383–387.
- [72] —, "Distributed variable-rate quantized compressed sensing in wireless sensor networks (invited paper)," in *Proc. IEEE Works. on Sign. Proc. Adv. in Wirel. Commun.*, Edinburgh, UK, Jul. 3–7 2016, pp. 1–5.
- [73] M. Leinonen, M. Codreanu, M. Juntti, and G. Kramer, "Rate-distortion lower bound for compressed sensing via conditional remote source coding," in *Proc. IEEE Inform. Theory Workshop*, Cambridge, UK, Sep. 11–14 2016, pp. 275–279.
- [74] A. Kipnis, G. Reeves, Y. Eldar, and A. Goldsmith, "Compressed sensing under optimal quantization," in *Proc. IEEE Int. Symp. Inform. Theory*, Aachen, Germany, Jun. 25–30 2017, pp. 2148–2152.
- [75] A. Rao, D. Miller, K. Rose, and A. Gersho, "A generalized VQ method for combined compression and estimation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Atlanta, GA, USA, May 7–10 1996, pp. 2032–2035.
- [76] A. Saxena, J. Nayak, and K. Rose, "A global approach to joint quantizer design for distributed coding of correlated sources," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Toulouse, France, May 14–19 2006, pp. II–53–II–56.
- [77] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Processing*, vol. 53, no. 7, pp. 2477–2488, Jul. 2005.
- [78] T. Wimalajeewa and P. K. Varshney, "OMP based joint sparsity pattern recovery under communication constraints," *IEEE Trans. Signal Processing*, vol. 62, no. 19, pp. 5059–5072, Oct. 2014.
- [79] J. Zhu, D. Baron, and F. Krzakala, "Performance limits for noisy multimeasurement vector problems," *IEEE Trans. Signal Processing*, vol. 65, no. 9, pp. 2444–2454, May 2017.
- [80] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, ser. Applied and Numerical Harmonic Analysis. Springer New York, 2013.
- [81] E. Candes, "The restricted isometry property and its implications for compressed sensing," *C.R. Acad. Sci. Paris, Ser. I*, vol. 346, no. 9–10, pp. 589–592, May 2008.
- [82] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012.
- [83] A. Y. Carmi, L. S. Mihaylova, and S. J. Godsill, *Compressed Sensing & Sparse Filtering*, 1st ed., ser. Signals and Communication Technology. Springer-Verlag Berlin Heidelberg, 2014.
- [84] A. Gersho, "On the structure of vector quantizers," *IEEE Trans. Inform. Theory*, vol. 28, no. 2, pp. 157–166, Mar. 1982.
- [85] J. Rissanen and G. Langdon, "Arithmetic coding," *IBM Journal of Research and Development*, vol. 23, no. 2, pp. 149–162, Mar. 1979.
- [86] T. Fine, "Properties of an optimum digital system and applications," *IEEE Trans. Inform. Theory*, vol. 10, no. 4, pp. 287–296, Oct. 1964.
- [87] K. Zeger and A. Gersho, "Vector quantizer design for memoryless noisy channels," in *Proc. IEEE Int. Conf. Commun.*, Philadelphia, PA, USA, Jun. 1988, pp. 1593–1597.
- [88] N. Wernersson, J. Karlsson, and M. Skoglund, "Distributed quantization over noisy channels," *IEEE Trans. Commun.*, vol. 57, no. 6, pp. 1693–1700, Jun. 2009.
- [89] R. Gray, *Source Coding Theory*, ser. The Kluwer International Series in Engineering and Computer Science. Kluwer Academic Publishers, 1990, vol. 83.
- [90] K. Rose, E. Gurewitz, and G. C. Fox, "Vector quantization by deterministic annealing," *IEEE Trans. Inform. Theory*, vol. 38, no. 4, pp. 1249–1257, Jul. 1992.
- [91] K. Rose and D. Miller, "A deterministic annealing algorithm for entropy-constrained vector quantizer design," in *Proc. Annual Asilomar Conf. Signals, Syst., Comp.*, Pacific Grove, CA, Nov. 1–3 1993, pp. 1651–1655.
- [92] S. S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): Design and construction," *IEEE Trans. Inform. Theory*, vol. 49, no. 3, pp. 626–643, Mar. 2003.
- [93] Z. Liu, S. Cheng, A. D. Liveris, and Z. Xiong, "Slepian-Wolf coded nested lattice quantization for Wyner-Ziv coding: High-rate performance analysis and code design," *IEEE Trans. Inform. Theory*, vol. 52, no. 10, pp. 4358–4379, Oct. 2006.
- [94] Q. Zhao and M. Effros, "Low complexity code design for lossless and near-lossless side information source codes," in *Proc. IEEE Data Compr. Conf.*, Mar. 2003, pp. 3–12.
- [95] —, "Lossless and near-lossless source coding for multiple access networks," *IEEE Trans. Inform. Theory*, vol. 49, no. 1, pp. 112–128, Jan. 2003.
- [96] M. Fresia, F. Perez-Cruz, H. V. Poor, and S. Verdú, "Joint source and channel coding," *IEEE Signal Processing Mag.*, vol. 27, no. 6, pp. 104–113, Nov. 2010.

- [97] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. New York, USA: Cambridge University Press, 2012.
- [98] T. Cover, A. E. Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Trans. Inform. Theory*, vol. 26, no. 6, pp. 648–657, Nov. 1980.
- [99] J. Barros and S. D. Servetto, "Network information flow with correlated sources," *IEEE Trans. Inform. Theory*, vol. 52, no. 1, pp. 155–170, Jan. 2006.
- [100] D. Gündüz, E. Erkip, A. Goldsmith, and H. V. Poor, "Source and channel coding for correlated sources over multiuser channels," *IEEE Trans. Inform. Theory*, vol. 55, no. 9, pp. 3927–3944, Sep. 2009.



Markus Leinonen (S'11) received his B.Sc. (EE) and M.Sc. (EE) degrees from University of Oulu, Oulu, Finland in 2010 and 2011, respectively. In 2010, he joined Centre for Wireless Communications (CWC) at University of Oulu, where he is currently working toward Dr.Sc. (Tech.) degree. During 2013 and 2014, he has been a guest researcher at Technical University of Munich, Germany. His research interests include distributed cross-layer optimization, compressed sensing, data compression, and rate-distortion analysis in energy-efficient data gathering

wireless sensor networks.



Marian Codreanu (S'02-M'07) received the M.Sc. degree from the University Politehnica of Bucharest, Romania, in 1998, and the Ph.D. degree from University of Oulu, Finland, in 2007. From 1998 to 2002 he was a Teaching Assistant at the Telecommunications Department of Politehnica University of Bucharest. In 2002 he joined the Centre for Wireless Communications at University of Oulu where he is currently an Adjunct Professor and holds an Academy Research Fellow position. In 2008 he was a visiting postdoctoral researcher at the University

of Maryland, College Park, USA. His research interests include optimization, compressed sensing, information theory, and signal processing for wireless communication systems and networks.

Dr. Codreanu received the best doctoral thesis prize within the area of all technical sciences in Finland in 2007. In 2013 he was nominated Academy Research Fellow by the Academy of Finland. He was a Co-Chair of the Technical Program Committee (TPC) of The First Nordic Workshop on Cross-Layer Optimization in Wireless Networks in 2010 and a Co-Chair of the TPC of the Nordic Workshop on System and Network Optimization for Wireless (SNOW) in 2011, 2013, 2014 and 2016. Currently, he is serving as Vice Chair of the IEEE Finland Communications and Information Theory Joint Societies Chapter.



Markku Juntti (S'93-M'98-SM'04) received his M.Sc. (EE) and Dr.Sc. (EE) degrees from University of Oulu, Oulu, Finland in 1993 and 1997, respectively.

Dr. Juntti was with University of Oulu in 1992–98. In academic year 1994–95, he was a Visiting Scholar at Rice University, Houston, Texas. In 1999–2000, he was a Senior Specialist with Nokia Networks. Dr. Juntti has been a professor of communications engineering since 2000 at University of Oulu, Centre for Wireless Communications (CWC), where he

leads the Communications Signal Processing (CSP) Research Group. He also serves as Head of CWC – Radio Technologies (RT) Research Unit. His research interests include signal processing for wireless networks as well as communication and information theory. He is an author or co-author in some 450 papers published in international journals and conference records as well as in books *WCDMA for UMTS*, *Signal Processing Handbook* and *5G Wireless Technologies*. Dr. Juntti is also an Adjunct Professor at Department of Electrical and Computer Engineering, Rice University, Houston, Texas, USA.

Dr. Juntti is an Editor of IEEE TRANSACTIONS ON COMMUNICATIONS and was an Associate Editor for IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY in 2002–2008. He was Secretary of IEEE Communication Society Finland Chapter in 1996–97 and the Chairman for years 2000–01. He has been Secretary of the Technical Program Committee (TPC) of the 2001 IEEE International Conference on Communications (ICC 2001), and the Co-Chair of the Technical Program Committee of 2004 Nordic Radio Symposium and 2006 IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2006), and the General Chair of 2011 IEEE Communication Theory Workshop (CTW 2011). He has served as Co-Chair of the Signal Processing for Communications Symposium of Globecom 2014 Signal Processing for Communications Symposium and IEEE GlobalSIP 2016 Symposium on Transceivers and Signal Processing for 5G Wireless and mm-Wave Systems.