Network Distributed Quantization

David Rebollo-Monedero and Bernd Girod
Information Systems Laboratory, Department of Electrical Engineering
Stanford University, Stanford, CA 94305
{drebollo,bgirod}@stanford.edu

***Abstract*— We investigate the design of rate-distortion optimal quantizers for distributed compression in a network with multiple senders and receivers. In such network, several noisy observations of one or more unseen sources are separately encoded by each sender and the quantization indices transmitted to a number of receivers, which jointly decode all available transmissions with the help of side information locally available. The joint statistics of the source data, the noisy observations and the side information are known, and exploited in the quantizer design. A variety of lossless coders for the quantization indices, including ideal multiple-source Slepian-Wolf coders, are allowed.**

We present the optimality conditions such quantizers must satisfy, together with an extension of the Lloyd algorithm for a locally optimal design. Experimental results for distributed quantization of Gaussian sources confirm the high-rate quantization theory established in our previous work.

I. Introduction

Consider the sensor network depicted in Fig. 1, where sensors obtain noisy readings of some unseen data of interest which must be transmitted to a central unit. The central unit

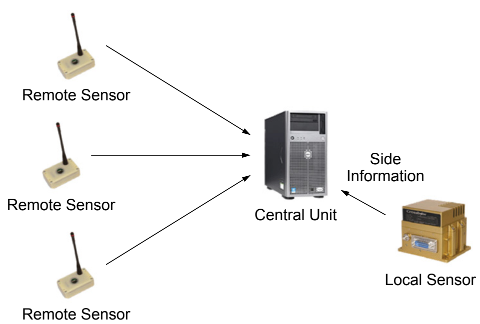


Fig. 1: Distributed source coding in a sensor network.

has access to side information, for instance archived data or readings from local sensors. At each sensor, neither the noisy observations of the other sensors nor the side information is available. Nevertheless, the statistical dependence among the unseen data, the noisy readings and the side information may be exploited in the design of each of the individual sensor encoders and the joint decoder at the central unit to optimize the rate-distortion performance. Clearly, if all the noisy readings and the side information were available at a single location, traditional joint denoising and encoding techniques could be used to reduce the transmission rate as much as possible, for a given distortion. However, since each of the noisy readings must be individually encoded without access to the side information, practical design methods for efficient distributed coders of noisy sources are needed.

This work was supported by NSF under Grant No. CCR-0310376.

The first attempts to design quantizers for Wyner-Ziv (WZ) coding, i.e., lossy source coding of a single instance of directly observed data with side information at the decoder, were inspired by the information-theoretic proofs. Zamir and Shamai [1] proved that, under certain circumstances, linear codes and high-dimensional nested lattices may approach the WZ rate-distortion function, in particular if the source data and side information are jointly Gaussian. This idea was further developed and applied by Pradhan *et al.* [2], [3], and Servetto [4], who published heuristic designs and performance analysis focusing on the Gaussian case, based on nested lattices, with either fixed-rate coding or entropy coding of the quantization indices.

A different approach was followed by Fleming and Effros [5], who generalized the Lloyd algorithm [6] for locally optimal, fixed-rate WZ quantizer design. Later, Fleming, Zhao and Effros [7] included rate-distortion optimized quantizers in which the rate measure is a function of the quantization index, for example, a codeword length. Unfortunately, vector quantizer dimensionality and entropy code block length coincide in their formulation, and thus the resulting quantizers either lack in performance or are prohibitively complex. An efficient algorithm for finding globally optimal quantizers among those with contiguous code cells was provided in [8], although, regrettably, it has been shown that code cell contiguity precludes optimality in general [9].

It may be concluded from the proof of the converse to the WZ rate-distortion theorem [10] that there is no asymptotic loss in performance by considering block codes of sufficiently large length, which may be thought of as vector quantizers, followed by fixed-length coders. This suggests a convenient implementation of WZ coders as quantizers, possibly preceded by transforms, followed by Slepian-Wolf coders, analogously to the implementation of nondistributed coders. The quantizer divides the signal space into cells, which, however, may consist of noncontiguous subcells mapped into the same quantizer index. Cardinal and Van Asche [11] considered Lloyd quantization for ideal symmetric Slepian-Wolf coding, without side information. Cardinal [12], [13] has also studied the problem of quantization of the side information itself for lossless distributed coding of discrete source data (which is not quantized).

A more general extension of the Lloyd algorithm appeared in our own work [14], which considered a variety of coding settings in a unified framework, including the important case of ideal Slepian-Wolf distributed coding of one or several sources with side information, i.e., when the rate is the joint conditional entropy of the quantization indices given the side information. A theoretic characterization of WZ quantizers with Slepian-Wolf coding at high rates was presented also in our own work [15]. Xiong *et al.* [16], [17] implemented a WZ encoder as a nested lattice quantizer followed by a Slepian-Wolf coder, and in [18], a trellis-coded quantizer was used instead (see also [19]).

As for quantization of a noisy observation of an unseen source, the nondistributed case was studied by Dobrushin, Tsybakov, Wolf, Ziv, Ephraim, Gray and others in [20]–[22]. Most of the operational work on distributed coding of noisy sources, i.e., for a fixed dimension, deals with quantizer design for a variety of settings, as in the work by Lam and Reibman [23], [24], and Gubner [25], but does not consider entropy coding or the characterization of such quantizers at high rates or transforms. The problem of optimal noisy WZ (NWZ) quantizer design with side information at the decoder and ideal Slepian-Wolf coding has only been studied under the assumptions of high rates and particular statistical conditions by our research group [26].

The object of this paper is to investigate the design of rate-distortion optimal quantizers for distributed compression in a network with multiple senders and receivers. In such network, several noisy observations of one or more unseen sources are separately encoded by each sender and the quantization indices transmitted to a number of receivers, which jointly decode all available transmissions with the help of side information locally available. We present a unified framework in which a variety of coding settings is allowed, including ideal Slepian-Wolf coding, along with an extension of the Lloyd algorithm for "locally" optimal design. This work builds upon the distributed coding problem studied in [27], by considering several decoders, and summarizes part of the research detailed in [28], where mathematical technicalities and proofs are provided.

Sec. II contains the formulation of the problem studied. A theoretic analysis for optimal quantizer design is presented in Sec. III, illustrated with the example of broadcast with side information in Sec. IV. Experimental results for NWZ coding of jointly Gaussian data are shown in Sec. V.

II. Formulation of the Problem

Throughout the paper, the measurable space in which a random variable (r.v.) takes on values will be called an alphabet. All alphabets are assumed to be Polish spaces to ensure the existence of regular conditional probabilities. We shall follow the convention of using uppercase letters for r.v.'s, and lowercase letters for particular values they take on. Probability density functions (PDF) and probability mass functions (PMF) are denoted by p and subindexed by the corresponding r.v.

We study the design of optimal quantizers for network distributed coding of noisy sources with side information. Fig. 2 depicts a network with several lossy encoders communicating with several lossy decoders. Let $m, n \in \mathbb{Z}^+$ represent the number of encoders and decoders, respectively. Let X, $Y = (Y_j)_{j=1}^n$ and $Z = (Z_i)_{i=1}^m$ be r.v.'s defined on a common probability space, statistically *dependent* in general, taking values in arbitrary, possibly different alphabets, respectively $\mathscr{X}$, $\mathscr{Y}$ and $\mathscr{Z}$. For each $i = 1, \dots, m$, Z_i represents an *observation* statistically related to some *source data* X of interest, available only at encoder i. For instance, Z_i might be an image corrupted by noise, a feature extracted from the source data such as a projection or a norm, the pair consisting of the source data itself together with some additional encoder side information, or any type of correlated r.v., and X might also be of the form $(X_i)_{i=1}^m$, where X_i could play the role of the source data from which Z_i is originated. For each $j = 1, \dots, n$, Y_j represents some *side information*, for example, previously decoded data, or an additional, local noisy observation, available at decoder j only. For each i, a quantizer $q_{i\cdot}(z_i) = (q_{ij}(z_i))_{j=1}^n$, which can also be regarded as a family of quantizers, is applied to the observation Z_i, obtaining the quantization indices $Q_{i\cdot} = (Q_{ij})_{j=1}^n$. Each quantization index Q_{ij} is losslessly encoded at encoder i, and transmitted to decoder j, where is losslessly decoded. We shall see that both encoding and decoding of quantization indices may be joint or separate. For each j, all quantization indices received by decoder j, $Q_{\cdot j} = (Q_{ij})_{i=1}^m$, and the side information Y_j are used jointly to estimate the unseen source data X. Let $\hat{X}_j$ represent this estimate, obtained with a measurable function $\hat{x}_j(q_{\cdot j}, y_j)$, called *reconstruction function*, in an alphabet $\hat{\mathscr{X}}_j$ possibly different from $\mathscr{X}$. Define $Q = (Q_{ij})_{i=1,j=1}^{m,n}$, $\hat{X} = (\hat{X}_j)_{j=1}^n$, and $\hat{x}(q, y) = (\hat{x}_j(q_{\cdot j}, y_j))_{j=1}^n$. $\mathscr{Q}$ and $\hat{\mathscr{X}}$ will denote the alphabets of Q and $\hat{X}$, respectively. *Partially connected networks* in which encoder i does *not* communicate with decoder j can easily be handled by redefining $Q_{i\cdot}$, $Q_{\cdot j}$ and Q not to include Q_{ij}.

We now extend the concept of rate function presented in [14] by means of a much more general definition, accompanied by a characteristic property which will play a major role in the extension of the Lloyd algorithm for optimized quantizer design, called 'update property'. Even though the definition of rate function and its update property may seem rather abstract and complex at this point, it possesses great generality and it is applicable to a wide range of problems. The terms 'cost function', 'distortion function' and 'rate function', formally defined to be equivalent, will be used to emphasize different connotations in the context of applications, usually a Lagrangian cost, a transmission rate, and the distortion introduced by lossy decoding, respectively.

A *cost function* is a nonnegative measurable function of the form $c(q, x, \hat{x}, y, z)$, possibly defined in terms of the joint probability distribution of (Q, X, Y, Z)[(a)]. Its *associated expected cost* is defined as $\mathcal{C} = \mathrm{E}\, c(Q, X, \hat{X}, Y, Z)$. Furthermore, a cost function will be required to satisfy the following

[(a)]Rigorously, a cost function may take as arguments probability distributions (measures) and probability functions, i.e., it is a function of a function. We shall become immediately less formal and call the evaluation of the cost function for a particular probability distribution or function, cost function as well.

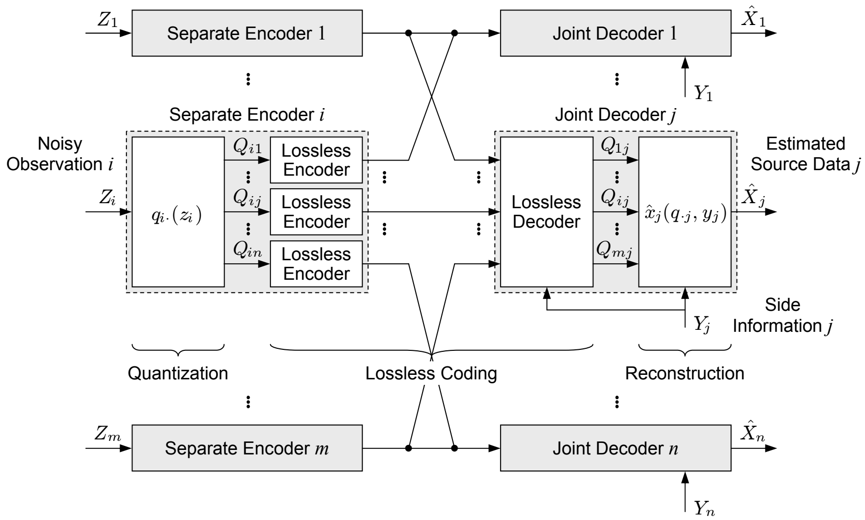


Fig. 2: Distributed quantization of noisy sources with side information in a network with m encoders and n decoders.

update property. For any modification of the joint probability distribution of (Q, X, Y, Z) preserving the marginal distribution of (X, Y, Z), there must exist an induced cost function c', consistent with the original definition almost surely (a.s.) but expressed in terms of the modified distribution, satisfying $\mathrm{E}\, c'(Q, X, Y, Z) \geqslant \mathcal{C}$, where the expectation is taken with respect to the original distribution.

For example, consider $r(q, y) = -\log p_{Q|\,Y}(q|\,y)$. The rate function corresponding to any other PMF $p'_{Q|\,Y}$ would be $r'(q, y) = -\log p'_{Q|\,Y}(q, y)$. The modified associated rate would then be $\mathcal{R}' = \mathrm{E}\, r'(Q, Y)$, where r' is defined in terms of the new PMF but the expectation is taken with respect to the original one. Clearly, $\mathcal{R}' - \mathcal{R} = \mathrm{D}(p_{Q|\,Y} \| p'_{Q|\,Y}) \geqslant 0$, hence r is indeed a rate function. In fact, it was observed in [14] that for this example of rate function, the associated rate $\mathcal{R} = \mathrm{H}(Q|Y)$ is precisely that introduced by an ideal Slepian-Wolf coder. Consequently, $r(q, y) = -\log p_{Q|\,Y}(q|\,y)$ was used to model *both* this type of coder and a conditional coder with access to the side information at the encoder. It can be seen that all rate functions defined in [14] satisfy the above definition, thereby making the corresponding coding settings applicable to this framework.

If the alphabets of X, Y, Z and $\hat{X}$ were equal to some common normed vector space, then an example of a distortion function would be $d(x, \hat{x}, y, z) = \alpha\|x - \hat{x}\|^2 + \beta\|y - \hat{x}\|^2 + \gamma\|z - \hat{x}\|^2$, for any $\alpha, \beta, \gamma \in [0, \infty)$. An example of a cost function suitable for distributed source coding, the main focus of this work, is $c(q, x, \hat{x}, y) = d(x, \hat{x}) + \lambda\, r(q, y)$, where λ is a non-negative real number determining the rate-distortion tradeoff in the Lagrangian cost $\mathcal{C} = \mathcal{D} + \lambda\,\mathcal{R}$.

Given a suitable cost function $c(q, x, \hat{x}, y, z)$, we address the problem of finding quantizers $q_{i\cdot}(z_i)$, and reconstruction functions $\hat{x}_j(q_{\cdot j}, y_j)$, minimizing the associated expected cost $\mathcal{C}$. The choice of the cost function leads to a particular noisy distributed source coding system, including a model for lossless coding.

III. Optimal Quantizer Design

We now establish necessary conditions for the optimal quantization and reconstruction functions, analogous to the nearest neighbor and centroid condition found in conventional, nondistributed quantization. Each necessary condition may be interpreted as the solution to a Bayesian decision problem. They will be expressed in terms of conditional cost functions, defined below, which play the role of conditional risks in Bayesian decision theory. A detailed, rigorous derivation of these conditions is presented in [28].

The *conditional cost function for encoder* i is defined as

$$\tilde{c}_i(q_{i\cdot}, z_i) = \mathrm{E}\left[[c(Q, X, \hat{x}(Q, Y), Y, Z)]_{Q_{i\cdot} = q_{i\cdot}} | z_i \right],$$

$[\text{expression}]_{\text{substitution}}$ denotes substitution in an expression, i.e., the expression of the definition is evaluated at $q_{i\cdot}$ and conditioned on the event $\{Z_i = z_i\}$. Observe that the conditional cost function is completely determined by the joint distribution of X, Y and Z, the cost function $c(q, x, \hat{x}, y, z)$, the quantization at other encoders $p_{Q_{i'\cdot}|Z_{i'}}(q_{i'\cdot}, z_{i'})$, and all the reconstruction functions, grouped as $\hat{x}(q, y)$. Using the fact that $Q_i = q_i(Z_i)$ and iterated expectation, it is easy to see that for *each* sender i, $\mathrm{E}\,\tilde{c}_i(Q_{i\cdot}, Z_i) = \mathcal{C}$. In terms of Bayesian decision theory, the expectation of conditional risks gives the overall Bayes risk. This key property leads to the necessary optimality condition for the quantization function at sender i, which may be informally expressed as:

$$q^*_{i\cdot}(Z_i) \overset{\text{a.s.}}{=} \arg\min_{q_{i\cdot}} \tilde{c}_i(q_{i\cdot}, Z_i), \tag{1}$$

provided that a minimum exists.

Similarly to the sender cost functions, the *conditional cost function for decoder* j is

$$\tilde{\tilde{c}}_j(q_{\cdot j}, \hat{x}_j, y_j) = \mathrm{E}[[c(Q, X, \hat{X}, Y, Z)]_{\hat{X}_j = \hat{x}_j} | q_{\cdot j}, y_j].$$

Observe that the defining expression is evaluated at $\hat{x}_j$ and conditioned on the event $\{Q_{\cdot j} = q_{\cdot j}, Y_j = y_j\}$, and it is completely determined by the joint distribution of X, Y and Z, the cost function $c(q, x, \hat{x}, y, z)$, all quantizers, and the reconstruction functions at other decoders $\hat{x}_{j'}(q_{\cdot j'}, y_{j'})$. Like in the case of the sender cost functions, it can be shown that for each j, $\mathrm{E}\,\tilde{\tilde{c}}_j(Q_{\cdot j}, \hat{X}_j, Y_j) = \mathcal{C}$. From this key property it follows that an optimal reconstruction function must satisfy

$$\hat{x}_j^*(Q_{\cdot j}, Y_j) \overset{\text{a.s.}}{=} \arg\min_{\hat{x}_j} \tilde{\tilde{c}}_j(Q_{\cdot j}, \hat{x}_j, Y_j), \qquad (2)$$

provided that a minimum exists, for each pair $(q_{\cdot j}, y_j)$ satisfying $p_{Q_{\cdot j}|Y_j}(q_{\cdot j}|y_j) > 0$ ($\hat{x}_j^*(q_{\cdot j}, y_j)$ can be arbitrarily defined elsewhere).

The necessary optimality conditions (1) and (2), together with the rate update property defining rate functions, suggest an alternating optimization algorithm that extends the Lloyd algorithm to the quantizer design problem considered in this work:

1. Initialization: For each $i = 1, \ldots, m$, $j = 1, \ldots, n$, choose initial quantization functions $(q_{i\cdot}^{(1)}(z_i))_i$, and initial reconstruction functions $(\hat{x}_j^{(1)}(q_{\cdot j}, y_j))_j$. Set $k = 1$ and $\mathcal{C}^{(0)} = \infty$.
2. Cost function update: Update the cost function $c^{(k)}(q, x, \hat{x}, y, z)$, completely determined by probability distributions involving $Q^{(k)}$, X, $\hat{X}$, Y and Z.
3. Convergence check: Compute the expected cost $\mathcal{C}^{(k)}$ associated with the current quantizers $(q_{i\cdot}^{(k)}(z_i))_i$, reconstruction functions $\hat{x}_j^{(k)}(q_{\cdot j}, y_j)$ and cost function $c^{(k)}(q, x, \hat{x}, y, z)$. Depending on its value with respect to $\mathcal{C}^{(k-1)}$, continue or stop.
4. Quantization update: For each i, obtain the next optimal quantization function $(q_{i\cdot}^{(k+1)}(z_i))_i$, given the most current quantization functions with index $i' \neq i$, all reconstruction functions $\hat{x}_j^{(k)}(q_{\cdot j}, y_j)$ and the cost function $c^{(k)}(q, x, \hat{x}, y, z)$.
5. Reconstruction update: For each j, obtain the next optimal reconstruction function $\hat{x}_j^{(k+1)}(q_{\cdot j}, y_j)$, given the most current version of the reconstruction functions with index $j' \neq j$, all quantization functions $(q_{i\cdot}^{(k+1)}(z_i))_i$, and the cost function $c^{(k)}(q, x, \hat{x}, y, z)$.
6. Increment k and go back to 2.

It can be proved that the sequence of costs $\mathcal{C}^{(k)}$ in the above algorithm is non-increasing, and since it is non-negative, it converges. In addition, any quantizer satisfying the optimality conditions (1) and (2), without ambiguity in any of the minimizations involved, is a fixed point of the algorithm. Even though these properties do not imply per se that the cost converges to a local or global minimum, the experimental results in the next section show good convergence properties, especially when the algorithm is combined with genetic search for initialization.

Many variations on the algorithm are possible, such as constraining the reconstruction function, for instance imposing linearity in Y for each Q when the distortion is the mean-squared error of the estimate $\hat{X}$, or any of the variations mentioned in [14].

IV. Example: Broadcast with Side Information

Even though the original motivation and purpose of our work is the network distributed coding depicted in Fig. 2, a number of problems of apparently different nature can in fact be unified within the theoretic framework developed [28]. An example is the problem of broadcast with side information, which we discuss next. Additional cases are quantization of side information, quantization with encoder side information, distributed classification, Gauss mixture modeling, the Blahut-Arimoto algorithm for WZ coding, and the bottleneck method.

In the problem of broadcast with side information, a single noisy observation Z is quantized, the quantization index Q is losslessly encoded, and the single, resulting bit stream is broadcast to two (or more) decoders, which have access to different side information, Y_1 and Y_2, in order to obtain the reconstructions $\hat{X}_1$ and $\hat{X}_2$ of some data of interest X. This problem, depicted in Fig. 3, corresponds to the case of network distributed quantization with $m = 1$ encoder and $n = 2$ decoders, with the additional constraint that $Q_{11} = Q_{12}$. Define $Q = Q_{11} = Q_{12}$. According to [29], the lowest

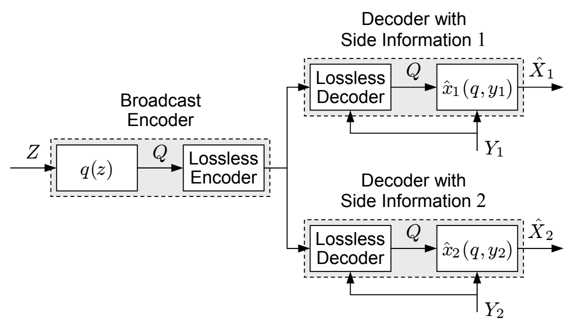


Fig. 3: Broadcast with side information. This corresponds to the network distributed setting with $m = 1$ encoder and $n = 2$ decoders, with the additional constraint $Q_{11} = Q_{12}$.

achievable rate corresponding to the common, broadcast bit stream coding Q, is $\mathcal{R} = \max\{\mathrm{H}(Q|Y_1), \mathrm{H}(Q|Y_2)\}$, in other words, the maximum of the rates required by Slepian-Wolf coders allowed to transmit a different bit stream to each receiver. The objective is to find the quantizer and the two reconstruction functions that minimize the distortion $\mathcal{D} = \mathrm{E}\,d(Q, X, \hat{X}, Y, Z)$ with a maximum rate constraint, $\mathcal{R} \leqslant \mathcal{R}_{\max}$. This includes the case when only one decoder has access to side information, studied from an information-theoretic perspective in [30], [31], where the motivational application is source coding "when side information may be absent".

The problem of designing quantizers and reconstruction functions to minimize a distortion given a rate constraint is usually handled via a distortion-rate Lagrangian cost. Precisely, points of the operational distortion-rate function $\mathcal{D}(\mathcal{R}_{\max}) = \inf_{\mathcal{R} \leqslant \mathcal{R}_{\max}} \mathcal{D}$, are found by minimizing the Lagrangian cost $\mathcal{C} = \mathcal{D} + \lambda\,\mathcal{R}$ instead, for different values of λ. If the operational distortion-rate function is convex, then the complete distortion-rate curve will be swept. In general, a lower convex envelope is obtained.

Define $\mathcal{R}_j = \mathrm{H}(Q|Y_j)$. To tackle the broadcast problem we observe that the rate constraint $\mathcal{R} = \max\{\mathcal{R}_1, \mathcal{R}_2\} \leqslant \mathcal{R}_{\max}$ is equivalent to the set of constraints $\mathcal{R}_1 \leqslant \mathcal{R}_{\max}$, $\mathcal{R}_2 \leqslant \mathcal{R}_{\max}$. In terms of the operational distortion-rate function,

$$\mathcal{D}(\mathcal{R}_{\max}) = \inf_{\mathcal{R} \leqslant \mathcal{R}_{\max}} \mathcal{D} = \inf_{\mathcal{R}_1, \mathcal{R}_2 \leqslant \mathcal{R}_{\max}} \mathcal{D}.$$

This leads to proposing the Lagrangian cost function $c = d + \lambda_1 r_1 + \lambda_2 r_2$, where $r_j(q, y_j) = -\log p_{Q|Y_j}(q|y_j)$, with associated Lagrangian expected cost $\mathcal{C} = \mathcal{D} + \lambda \mathcal{R}_1 + \lambda_2 \mathcal{R}_2$, which corresponds to the coding setting shown in Fig. 4. The

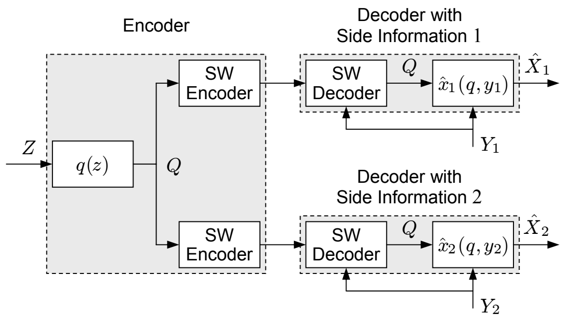


Fig. 4: Quantization setting equivalent to the problem of broadcast with side information when the Slepian-Wolf rates satisfy a common rate constraint, i.e., $\mathrm{H}(Q|Y_1), \mathrm{H}(Q|Y_2) \leqslant \mathcal{R}_{\max}$.

rate-distortion points of interest correspond to those multipliers λ_1, λ_2 such that $\mathcal{R}_1, \mathcal{R}_2 \leqslant \mathcal{R}_{\max}$. Of course, just as in any Lagrangian optimization, only the lower convex envelope of the operational distortion-rate function $\mathcal{D}(\mathcal{R}_{\max})$ will be found.

V. Experimental Results

In this section, we illustrate the theoretic analysis with experimental results for a simple, intuitive case of distributed source coding, represented in Fig. 5. In this case, X_1, X_2 are k i.i.d. samples drawn according to jointly Gaussian statistics, precisely, $\mathcal{N}\left(0, \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)\right)$, with $|\rho| < 1$. We wish to design distributed quantizers minimizing the Lagrangian cost $\mathcal{C} = \mathcal{D} + \lambda \mathcal{R}$, where the distortion is the mean-squared error $\mathcal{D} = \frac{1}{k} \mathrm{E} \|X - \hat{X}\|^2$, and the rate is that required by an ideal Slepian-Wolf coder, $\mathcal{R} = \frac{1}{k} \mathrm{H}(Q_1, Q_2)$, both normalized per sample.

The high-rate approximation theory for network distributed quantization [28] implies that

$$\mathcal{D} \simeq 2\pi e M_k \sqrt{1 - \rho^2}\, 2^{-2\mathcal{R}},$$

where M_k is the minimum normalized moment of inertia of the convex polytopes tessellating $\mathbb{R}^k$. Replacing M_k by its limit as $k \to \infty$, i. e., $1/2\pi e$, and solving for $\mathcal{R}$, we obtain the approximation $\mathcal{R} \simeq \frac{1}{2} \log \frac{\sqrt{1-\rho^2}}{\mathcal{D}}$ to the information-theoretic rate-distortion function. The exact information-theoretic rate-distortion function can be obtained by application of a recent result [32][b]:

$$\mathcal{R} = \frac{1}{4} \log \frac{(1 - \rho^2) \left(1 + \sqrt{1 + \frac{4\rho^2 \mathcal{D}^2}{(1-\rho^2)^2}}\right)}{2\mathcal{D}^2},$$

[b] According to our definition of $\mathcal{D}$ and $\mathcal{R}$, $\mathcal{D} = \frac{1}{2}(\mathcal{D}_1 + \mathcal{D}_2)$, $\mathcal{R} = \frac{1}{2}(\mathcal{R}_1 + \mathcal{R}_2)$, where the infimum $\mathcal{R}$ is given by [32, §2] in terms of $\mathcal{D}_1$ and $\mathcal{D}_2$. Show that the distortion allocation minimizing $\mathcal{R}$ leads to $\mathcal{D}_1 = \mathcal{D}_2$, for example using the Pareto optimality condition $\frac{\partial \mathcal{R}}{\partial \mathcal{D}_1} = \frac{\partial \mathcal{R}}{\partial \mathcal{D}_2}$.

which as expected matches our approximation for small $\mathcal{D}$.

For $\rho = 1/2$, dimensions $k = 1, 2$, and several values of λ, quantizers and reconstruction functions were designed using the extension of the Lloyd algorithm presented in this work. A simple genetic search method was combined with the Lloyd algorithm to select initial quantizers based on their cost after convergence. The corresponding distortion-rate points are shown in Fig. 6, along with the information-theoretic distortion-rate function, the high-rate approximation, and the distortion bound $\mathcal{D}_0 = \sigma^2_{X_i} = 1$. The results obtained

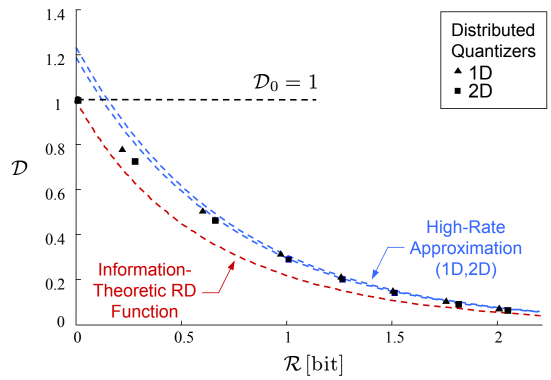
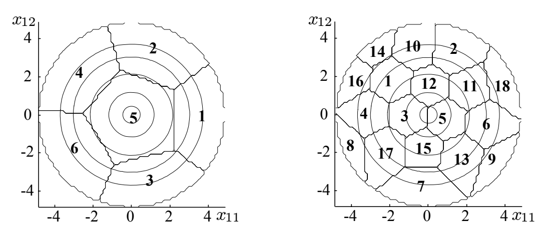


Fig. 6: Distortion-rate performance of optimized symmetric distributed quantizers with Slepian-Wolf coding for $k = 1, 2$. $\left(\begin{smallmatrix} X_1 \\ X_2 \end{smallmatrix}\right) \sim \mathcal{N}\left(0, \left(\begin{smallmatrix} 1 & 1/2 \\ 1/2 & 1 \end{smallmatrix}\right)\right)$.

confirm the theory developed in [28] and the usefulness of the algorithm presented, for the statistics of the example.

For $k = 1$ and high rates, all quantizers obtained experimentally where uniform without index repetition. Two of the 2-dimensional quantizers designed are depicted in Fig. 7. According to their rate-distortion performance compared to



(a) $\mathcal{R} \simeq 0.27$, $\mathcal{D} \simeq 0.73$.

(b) $\mathcal{R} \simeq 1.26$, $\mathcal{D} \simeq 0.21$.

Fig. 7: 2-dimensional symmetric distributed quantizers with Slepian-Wolf coding found by the Lloyd algorithm. $\left(\begin{smallmatrix} X_1 \\ X_2 \end{smallmatrix}\right) \sim \mathcal{N}\left(0, \left(\begin{smallmatrix} 1 & 1/2 \\ 1/2 & 1 \end{smallmatrix}\right)\right)$. Circles enclose sets of probability 0.1, 0.5, 0.9, 0.99 and 0.999.

the high-rate approximation in Fig. 6, quantizers (a) and (b) can be thought of as a low-rate and a high-rate quantizer respectively. Quantizer (b) is a uniform hexagonal tessellation, as predicted by [28]. In all our experiments, the two encoders had very similar quantizers for a given λ. The algorithm was run on a fine discretization of the PDF, which for $k = 2$ used approximately $5.0 \cdot 10^6$ samples in a 4-dimensional ellipsoid of probability of probability $1 - 10^{-4}$, giving 3657 points for each X_i. Due to this discretization, the edges of the quantization regions appear somewhat jagged.

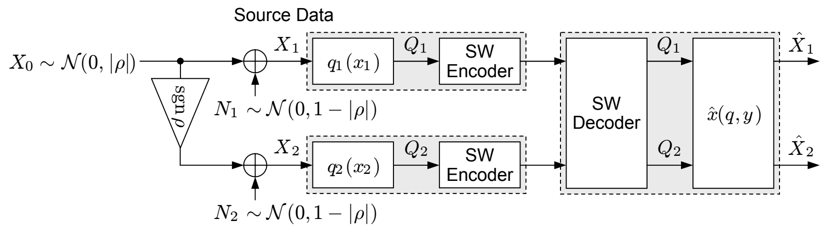


Fig. 5: Example of clean symmetric distributed quantization with Gaussian statistics. X_0, N_1, N_2 are independent, thus $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim \mathcal{N}\left(0, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$.

VI. Conclusion

We have established necessary optimality conditions for network distributed quantization of noisy sources and extended the Lloyd algorithm for its design. The concept of cost function enables us to model the lossless coding method used for the quantization indices, particularly ideal Slepian-Wolf coding. In addition, cost functions are the key to a very general formulation that unifies a number of related problems, such as broadcast with side information, distributed classification, or quantization of side information.

Experimental results confirm the high-rate approximation theory for distributed quantization, and the recent information-theoretic analysis of the two-terminal quadratic-Gaussian problem. In addition, they suggest that the convergence properties of the extended Lloyd algorithm are similar to those of the classical one, and can benefit from a genetic search algorithm for initialization.

References

[1] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1250–1276, June 2002.

[2] S. S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): Design and construction," in *Proc. IEEE Data Compression Conf. (DCC)*, Snowbird, UT, Mar. 1999, pp. 158–167.

[3] S. S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microsensor network," *IEEE Signal Processing Mag.*, vol. 19, pp. 51–60, Mar. 2002.

[4] S. D. Servetto, "Lattice quantization with side information," in *Proc. IEEE Data Compression Conf. (DCC)*, Snowbird, UT, Mar. 2000, pp. 510–519.

[5] M. Fleming and M. Effros, "Network vector quantization," in *Proc. IEEE Data Compression Conf. (DCC)*, Snowbird, UT, Mar. 2001, pp. 13–22.

[6] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 129–137, Mar. 1982.

[7] M. Fleming, Q. Zhao, and M. Effros, "Network vector quantization," *IEEE Trans. Inform. Theory*, vol. 50, no. 8, pp. 1584–1604, Aug. 2004.

[8] D. Muresan and M. Effros, "Quantization as histogram segmentation: Globally optimal scalar quantizer design in network systems," in *Proc. IEEE Data Compression Conf. (DCC)*, Snowbird, UT, Apr. 2002, pp. 302–311.

[9] M. Effros and D. Muresan, "Codecell contiguity in optimal fixed-rate and entropy-constrained network scalar quantizers," in *Proc. IEEE Data Compression Conf. (DCC)*, Snowbird, UT, Apr. 2002, pp. 312–321.

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

[11] J. Cardinal and G. V. Asche, "Joint entropy-constrained multiterminal quantization," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Lausanne, Switzerland, June 2002, p. 63.

[12] J. Cardinal, "Compression of side information," in *Proc. IEEE Int. Conf. Multimedia, Expo (ICME)*, vol. 2, Baltimore, MD, July 2003, pp. 569–572.

[13] ——, "Quantization of side information," unpublished.

[14] D. Rebollo-Monedero, R. Zhang, and B. Girod, "Design of optimal quantizers for distributed source coding," in *Proc. IEEE Data Compression Conf. (DCC)*, Snowbird, UT, Mar. 2003, pp. 13–22.

[15] D. Rebollo-Monedero, A. Aaron, and B. Girod, "Transforms for high-rate distributed source coding," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, vol. 1, Pacific Grove, CA, Nov. 2003, pp. 850–854, invited paper.

[16] Z. Xiong, A. Liveris, S. Cheng, and Z. Liu, "Nested quantization and Slepian-Wolf coding: A Wyner-Ziv coding paradigm for i.i.d. sources," in *Proc. IEEE Workshop Stat. Signal Processing (SSP)*, St. Louis, MO, Sept. 2003, pp. 399–402.

[17] Z. Liu, S. Cheng, A. D. Liveris, and Z. Xiong, "Slepian-wolf coded nested quantization (SWC-NQ) for Wyner-Ziv coding: Performance analysis and code design," in *Proc. IEEE Data Compression Conf. (DCC)*, Snowbird, UT, Mar. 2004, pp. 322–331.

[18] Y. Yang, S. Cheng, Z. Xiong, and W. Zhao, "Wyner-Ziv coding based on TCQ and LDPC codes," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, vol. 1, Pacific Grove, CA, Nov. 2003, pp. 825–829.

[19] Z. Xiong, A. D. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Processing Mag.*, vol. 21, no. 5, pp. 80–94, Sept. 2004.

[20] R. L. Dobrushin and B. S. Tsybakov, "Information transmission with additional noise," *IRE Trans. Inform. Theory*, vol. IT-8, pp. S293–S304, 1962.

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

[22] Y. Ephraim and R. M. Gray, "A unified approach for encoding clean and noisy sources by means of waveform and autoregressive vector quantization," *IEEE Trans. Inform. Theory*, vol. IT-34, pp. 826–834, July 1988.

[23] W. M. Lam and A. R. Reibman, "Quantizer design for decentralized estimation systems with communication constraints," in *Proc. Annual Conf. Inform. Sci. Syst. (CISS)*, Baltimore, MD, Mar. 1989.

[24] ——, "Design of quantizers for decentralized estimation systems," *IEEE Trans. Inform. Theory*, vol. 41, no. 11, pp. 1602–1605, Nov. 1993.

[25] J. A. Gubner, "Distributed estimation and quantization," *IEEE Trans. Inform. Theory*, vol. 39, no. 4, pp. 1456–1459, July 1993.

[26] D. Rebollo-Monedero, S. Rane, and B. Girod, "Wyner-Ziv quantization and transform coding of noisy sources at high rates," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, vol. 2, Pacific Grove, CA, Nov. 2004, pp. 2084 – 2088.

[27] D. Rebollo-Monedero and B. Girod, "Design of optimal quantizers for distributed coding of noisy sources," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)*, vol. 5, Philadelphia, PA, Mar. 2005, pp. 1097–1100, invited paper.

[28] D. Rebollo-Monedero, "Quantization and transforms for distributed source coding," Ph.D. dissertation, Stanford Univ., 2007.

[29] A. Sgarro, "Source coding with side information at several decoders," *IEEE Trans. Inform. Theory*, vol. 23, no. 2, pp. 179–182, Mar. 1997.

[30] C. Heegard and T. Berger, "Rate distortion when side information may be absent," *IEEE Trans. Inform. Theory*, vol. IT-31, no. 6, pp. 727–734, Nov. 1985.

[31] A. H. Kaspi, "Rate-distortion function when side information may be present at the decoder," *IEEE Trans. Inform. Theory*, vol. 40, no. 6, pp. 2031–2034, Nov. 1994.

[32] A. B. Wagner, S. Tavildar, and P. Viswanath, "Rate region of the quadratic Gaussian two-encoder source-coding problem," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Seattle, WA, July 2006, pp. 1404–1408.