

Rate Loss in Distributed Functional Source Coding

John Z. Sun

Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, MA 02139, USA
Email: johnsun@mit.edu

Vivek K Goyal

Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, MA 02139, USA
Email: v.goyal@ieee.org

Abstract—For point-to-point and distributed communication of continuous sources, high-resolution quantization theory provides an achievable rate–distortion trade-off that is simple to compute and motivates practical compression architectures. Moreover, high-resolution analysis gives good inner bounds for the Shannon rate–distortion region when a more general characterization is difficult. In this paper, we analyze the sum-rate gap between coded nonuniform scalar quantization and the Shannon rate–distortion region for a system that requires fidelity in a computation applied to the source variables. We find that the loss can be as low as 0.255 bits/sample, which has previously been observed in the point-to-point setting, and it is achieved using a simple architecture of nonuniform quantization followed by Slepian–Wolf coding.

I. INTRODUCTION

Distributed functional scalar quantization (DFSQ) was proposed as a framework to characterize the performance of lossy compression in a distributed setting when the goal is recovery of a computation applied to source realizations rather than recovery of the source variables themselves [1]. In other previous work (see Section II-C), high-resolution quantization theory was used to determine closed-form expressions for achievable rate regions for a wide range of source models and fidelity criteria in situations where the large blocklength analysis is unwieldy. In this paper, we determine the minimum sum rate R_{sum}^* of the Shannon rate–distortion region for coding over distributed architectures such as in Fig. 1 and characterize the rate loss resulting from using the simple implementation of DFSQ for small distortion. We find the loss is small under many scenarios, reinforcing the usefulness of high-resolution quantization theory in understanding source coding in distributed systems.

This work is motivated by the well-known fact that, in the high-resolution (low-distortion) limit, scalar quantization followed by entropy coding has redundancy of only 0.255 bits/sample over more sophisticated vector quantizers that achieve the rate–distortion function in point-to-point communication. The fast convergence of the performance of scalar quantizers at finite rates to the low-distortion asymptotic underlies the importance of this theory to real-world applications, a point highlighted by Berger and Gibson in their survey of lossy source coding [2]. In a network scenario, the importance of constraints to blocklength and coding complexity is magnified due to resource limitations in practical implementations.

In this paper, we refer to the expression that characterizes

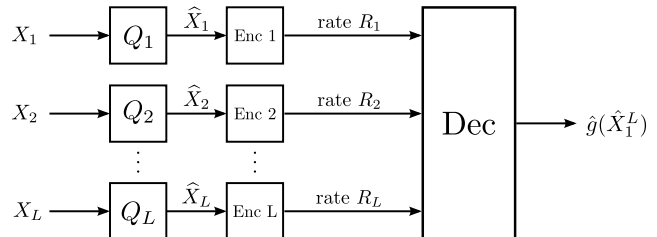


Fig. 1. A distributed computation network, where each of L spatially-separated sources generate a scalar X_i . The scalars are encoded and communicated over rate-limited links to a central decoder without interaction between encoders. The decoder computes an estimate of the function $g(X_1^n) = g(X_1, X_2, \dots, X_n)$ from the received data using $\hat{g}(\hat{X}_1^n)$. Each encoder is allowed transmission rate R_i for a total sum rate of R_{sum} .

the optimal trade-off between R_{sum}^* and allowed distortion in the functional setting as the *fMSE rate–distortion function*. The main contribution of this work is a precise statement on when nonuniform scalar quantization followed by Slepian–Wolf coding can be within 0.255 bits/sample of the fMSE rate–distortion function. When the coding after quantization is constrained, the loss can be substantially larger. Our analysis borrows heavily from source coding results for nondifference distortion measures [3], which were inspired by fidelity criteria that model human perception better than MSE. Although the contexts of these problems are different, the analyses have many similarities [1, Section IV-E].

The rest of the paper will be organized as follows. We review key results in network quantization and source coding in Section II. Then in Section III, we characterize the rate–distortion region for the distributed network under a weighted MSE distortion measure. Next, we relate DFSQ performance to the above model and analyze the rate loss in Section IV. Finally, we address future directions of this work in Section V.

II. PRELIMINARIES

In this section, we introduce the problem model and summarize previous results in rate–distortion theory and high-resolution quantization theory that are pertinent to the low-distortion regime. An extensive analysis of the complementary nature of the two asymptotics is given in [4, Section IV-I]. A comprehensive set of works related to DFSQ is presented in [1].

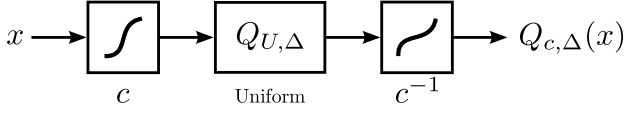


Fig. 2. A block diagram for companding as a constructive method for nonuniform scalar quantization. The notation $Q_{U,\Delta}$ is used to describe the canonical uniform quantizer with step size Δ in the granular region $[0, 1]$. The resulting quantizer is denoted $Q_{c,\Delta}$.

A. Problem Model

Consider the distributed compression problem in Fig. 1, where L correlated sources X_1^L are observed at separate sensors and communicated to a central decoder. We assume that the sources are memoryless over time and can be specified by a joint density $f_{X_1^L}$. If the sources are continuous, reliable communication leads to a loss in fidelity, which is quantified by a distortion measure $d(X_1^L, \hat{X}_1^L)$, where \hat{X}_i is the compressed version of X_i taken from a codebook \mathcal{C}_i . The central goals of lossy source coding are to design a codebook that minimizes both the communication rate and distortion $\mathbb{E}[d(X_1^L, \hat{X}_1^L)]$, and to study their fundamental trade-off.

For continuous-valued sources, we generally discretize the amplitude using a quantizer. Although uniform quantization is common in practice, nonuniform quantizers can lead to better compression performance. Fig. 2 presents the compander model as a method for generating nonuniform scalar quantizers from a uniform quantizer with step size $\Delta = 1/|\mathcal{C}|$. In this model, the scalar source is transformed using a nondecreasing and smooth compressor function $c: \mathbb{R} \rightarrow [0, 1]$, then quantized using a uniform quantizer with step size Δ on $[0, 1]$, and finally passed through the expander function c^{-1} . Compressor functions are defined such that $\lim_{x \rightarrow -\infty} c(x) = 0$ and $\lim_{x \rightarrow \infty} c(x) = 1$. A companding scalar quantizer is therefore uniquely specified by the companding function c and step size Δ ; it is denoted $Q_{c,\Delta}$.

The compander model is useful because it leads to precise characterization of distortion behavior as Δ becomes small, and calculus methods can optimize c to determine the best scalar quantizer. Quantizers designed in this manner are *asymptotically optimal*, meaning that the quantizer optimized over c has distortion that approaches the performance of the best scalar quantizer found by any means as $\Delta \rightarrow 0$ [5]–[7].

B. Distributed Functional Scalar Quantization

In the DFSQ problem, the goal is to design L scalar quantizers and a reconstruction function \hat{g} under a total rate constraint to minimize the functional MSE (fMSE) with respect to a scalar computation g :

$$D_{\text{fmse}}(c_1^L, \Delta_1^L) = \mathbb{E} \left[\left| g(X_1^L) - \hat{g}(\hat{X}_1^L) \right|^2 \right], \quad (1)$$

where c_1^L and Δ_1^L are vectors of the parameters that specify the L scalar quantizers.

To understand how a quantizer affects fMSE, it is convenient to first define how a computation locally affects distortion.

Definition 1: The i th functional sensitivity profile of a multivariate function g is defined as

$$\gamma_i(x) = \left(\mathbb{E} \left[|g^{(i)}(X_1^L)|^2 \mid X_i = x \right] \right)^{1/2}, \quad (2)$$

where $g^{(i)}(x)$ is the partial derivative of g with respect to its i th argument evaluated at the point x .

Given the functional sensitivity profile, the main result of [1] says

$$D_{\text{fmse}}(c_1^L, \Delta_1^L) \simeq \sum_{i=1}^L \frac{\Delta_i^2}{12} \mathbb{E} \left[\left(\frac{\gamma_i(X_i)}{c'_i(X_i)} \right)^2 \right], \quad (3)$$

where the reconstruction is the joint centroid

$$\hat{g}_{\text{jc}}(x_1^L) = \mathbb{E} \left[g(X_1^L) \mid \hat{X}_1^L = Q_{c_1^L, \Delta_1^L}(x_1^L) \right]$$

and \simeq indicates the ratio of expressions approach one as Δ decreases. The result requires the following technical conditions on $f_{X_1^L}$, g and c_1^L :

FQ1. The function g is Lipschitz continuous and twice differentiable in every argument except possibly on a set of Jordan measure zero.

FQ2. The source pdf $f_{X_1^L}$ is continuous, bounded, and supported on $[0, 1]^L$.

FQ3. The function g and companding functions c_1^L allow $\mathbb{E}[(\gamma_i(X_i)/c'_i(X_i))^2]$ to be defined and finite for all i .

A different set of conditions generalizes DFSQ theory to densities of unbounded support and a simpler decoder [8] but the same optimal distortion–rate trade-off holds.

In the DFSQ setting, the communication rate can be constrained in three ways based on limitations to encoder complexity. In the first, called fixed-rate scalar quantization (FRSQ), no further coding is performed after quantization, leading to asymptotically optimal rate–distortion performance

$$R_{\text{frsq}}(D) \simeq -\frac{L}{2} \log \left(\frac{12D}{LC_{\text{fr}}} \right), \quad (4)$$

where

$$C_{\text{fr}} = \left(\prod_{i=1}^L \|\gamma_i^2 f_{X_i}\|_{1/3} \right)^{1/L}.$$

In the second, called entropy-constrained scalar quantization (ECSQ), entropy coding is performed at each encoder and decoded separately by the central decoder, leading to asymptotically optimal performance

$$R_{\text{ecsq}}(D) \simeq -\frac{L}{2} \log \left(\frac{12D}{L} \right) + \sum_{i=1}^L (h(X_i) + \mathbb{E}[\log \gamma_i(X_i)]), \quad (5)$$

where $h(X)$ is the differential entropy function. Finally, we can notice that the output of the companding quantizers are generally correlated if the sources themselves are correlated and apply Slepian-Wolf coding to the quantizer outputs, called

scalar quantization with Slepian–Wolf coding (SWSQ). This type of coding yields asymptotically optimal performance

$$R_{\text{swsq}}(D) \simeq -\frac{L}{2} \log \left(\frac{12D}{L} \right) + h(X_1^L) + \sum_{i=1}^L \mathbb{E}[\log \gamma_i(X_i)]. \quad (6)$$

The results shown above are for the sum rate after applying bit allocation under a Lagrangian optimization that assumes all rates are positive and allowed to take on real numbers [9]. This is a reasonable approximation in the low-distortion regime, especially if the sources have similar variances.

C. Rate–Distortion Theory for Small Distortion

For vector quantization of L -dimensional sources, the fundamental trade-off between communication and fidelity is given by the Shannon rate–distortion function:¹

$$R(D) = \inf_{\mathbb{E}[d(X_1^L, \hat{X}_1^L)] \leq D} I(X_1^L, \hat{X}_1^L).$$

In general, the rate–distortion function is difficult to characterize, but the Shannon lower bound (SLB) is a useful bounding tool when d is a difference distortion measure, i.e. $d(X_1^L, \hat{X}_1^L) = \rho(X_1^L - \hat{X}_1^L)$ [10]. For MSE, the SLB is [11]

$$R(D) \geq h(X) - \frac{1}{2} \log(2\pi e D).$$

In the point-to-point case, it has been shown that this bound becomes tight as D becomes small [12], [13]. In the multiterminal setting, the SLB is also shown to be tight for difference distortion measures when D is small, achieved using lattice quantization followed by Slepian–Wolf coding [14].

In this paper, we are concerned with fMSE, which is generally not a difference distortion measure. The fMSE is closely related to the weighted mean-squared error (W-MSE) criterion:

$$d(X_1^L, \hat{X}_1^L) = \|W(X_1^L)(X_1^L - \hat{X}_1^L)\|^2, \quad (7)$$

where W is the input-dependent weighting matrix. Using a similar approach to the SLB, Linder and Zamir show that, in the joint-encoder case for W-MSE distortion [3], [15],

$$\lim_{D \rightarrow 0} \left(R(D) + \frac{L}{2} \log(2\pi e D/L) \right) = h(X_1^L) + \mathbb{E}[\log |W(X_1^L)|]. \quad (8)$$

We will use this result to find the minimum sum rate of the rate–distortion region in the distributed version of this problem with the same distortion measure. This then becomes a building block in determining the fMSE rate–distortion function.

¹In this work, we will focus solely on the informational definition rather than the operational one.

III. RATE–DISTORTION BEHAVIOR OF DISTRIBUTED SYSTEMS FOR W-MSE

We first analyze the compression performance of the distributed system in Fig. 1 for the W-MSE measure (7) under the following regularity conditions on W and the joint density of X_1^L :

WC1. $W(x_1^L)$ is strictly nonnegative for all inputs x_1^L and equal to zero at most on a set of Jordan measure zero.

WC2. $W(x_1^L)$ is diagonal, and the i th diagonal entry, denoted $w_i(x_1^L)$, depends only on x_i .

WC3. $W(x_1^L)$ is piecewise continuously differentiable with respect to x_1^L .

WC4. X_1^L has finite differential entropy and second moment

WC5. $W(x_1^L)$ and $f_{X_1^L}$ satisfy

$$\mathbb{E}[\log |\det W(X_1^N)|] < \infty$$

and

$$\mathbb{E} \left[(\text{tr}(W^{-1}(X_1^N)))^3 \right] < \infty.$$

Then the minimum sum rate of the achievable rate region is characterized asymptotically by the following theorem:

Theorem 1: Consider L correlated sources that are encoded separately and an allowed W-MSE distortion D . If the joint source density $f_{X_1^L}$ and input-dependent weighting matrix W satisfy Conditions WC1–WC5, then the minimum sum rate R_{sum}^* of the achievable rate region under the W-MSE distortion satisfies:

$$\begin{aligned} \lim_{D \rightarrow 0} \left(R_{\text{sum}}^*(D) + \frac{L}{2} \log(2\pi e D/L) \right) \\ = h(X_1^L) + \sum_{i=1}^L \mathbb{E}[\log w_i(X_i)]. \end{aligned} \quad (9)$$

Proof: The outer bound follows directly from (8), which requires Conditions WC3–WC5 hold [3].² We now devise an achievability scheme and sketch out a proof to confirm that the outer bound becomes tight in the low-distortion limit. We emphasize that the encoders of the L sources are separate and can use long blocks in this setup.

Let us define monotonic scalar functions \tilde{g}_i satisfying

$$\tilde{g}_i'(x) = w_i(x).$$

Since $w_i(x)$ is positive except on a set of Jordan measure zero (Condition WC1), \tilde{g}_i is unique up to a constant and strictly increasing. Consider a set of transformed sources $\{Y_i\}_{i=1}^L$, defined as $Y_i = \tilde{g}_i(X_i)$. In general, the set $\{Y_i\}_{i=1}^L$ is correlated if $\{X_i\}_{i=1}^L$ is correlated. We now solve a modified problem, which is to determine an achievable rate region for distributed compression of Y_1^L under the MSE fidelity criterion. Using the direct-coding proof of Zamir and Berger [14, Theorem 2], we see that the sum rate can asymptotically achieve

$$R_{\text{sum}}^*(D_1^L) \lesssim h(Y_1^L) - \frac{1}{2} \sum_{i=1}^L \log(2\pi e D_i) + \varepsilon(D_1^L),$$

²WC5 is a technical condition of [3].

where D_i is the allowable distortion between Y_i and \hat{Y}_i and $\varepsilon(D_1^L) \rightarrow 0$ as $D_1^L \rightarrow 0$. The notation \lesssim indicates the ratio of the left and right expressions is asymptotically less than one as D decreases. Using Taylor's theorem to expand \tilde{g}_i around X_i , we can approximate the MSE of compressing Y_i as:

$$\begin{aligned} D_i(Y_i, \hat{Y}_i) &= \mathbb{E}[|Y_i - \hat{Y}_i|^2] \\ &= \mathbb{E}[|\tilde{g}_i(X_i) - \tilde{g}_i(\hat{X}_i)|^2] \\ &\approx \mathbb{E}[|\tilde{g}'_i(X_i)|^2 (X_i - \hat{X}_i)^2]. \end{aligned}$$

We see that this becomes equivalent to the fidelity requirements of the W-MSE distortion as the distortion decreases. Setting $D_i = D/L$ for every i yields the achievable sum rate

$$R_{\text{sum}}^*(D) \lesssim h(Y_1^L) - \frac{L}{2} \log(2\pi e D/L) + \varepsilon(D_1^L).$$

In general, if $Y_1^L = u(X_1^L)$, then

$$h(Y_1^L) \leq h(X_1^L) + \mathbb{E}[\log |\partial u / \partial X_1^L|]$$

with equality if and only if u is bijective [11]. Since each \tilde{g}_i is strictly increasing and hence invertible, equality holds for this choice of u and we can rewrite the asymptotic achievable rate as

$$\begin{aligned} R_{\text{sum}}^*(D) &\lesssim \\ h(X_1^L) &- \frac{L}{2} \log(2\pi e D/L) + \sum_{i=1}^L \mathbb{E}[\log w_i(X_i)] + \varepsilon(D_1^L). \end{aligned}$$

This matches the outer bound as D becomes small, providing a tight characterization of the sum rate ■

Presented here is only the sum-rate characterization of the rate-distortion region. As in all multiterminal source coding problems, we must also bound all individual rates as well as the sums of all subsets of rates. For brevity, we do not present these calculations here, but they are easily found following the same analysis.

The achievability scheme of transforming the source before applying Slepian-Wolf coding is highly reminiscent of companding quantizers. Hence, one nice interpretation is that fine nonuniform quantization followed by block coding can achieve the optimal sum rate in a distributed setting for the W-MSE distortion with diagonal weighting matrices, provided that the target distortions for all sources are small.

IV. RATE LOSS OF DFSQ

We recall the fidelity criterion of interest is fMSE as defined in (1) and the performance of nonuniform scalar quantizers is given in (4)–(6). In this section, we will understand the achievable rate region with respect to fMSE using Theorem 1 and compare the sum rate to the DFSQ results; this determines the rate loss from using scalar quantizers.

We require the following conditions on the source density $f_{X_1^L}$ and computation g , recalling $g^{(i)}(x)$ denotes the partial derivative of g with respect to its i th argument evaluated at the point x .

FC1. For all $i \in \{1, \dots, L\}$, $g^{(i)}(X_1^L)$ is nonzero except on a set of Jordan measure 0.

FC2. $g(X_1^L)$ is twice differentiable with respect to X_1^L .

FC3. X_1^L has finite differential entropy and second moment

FC4. For all $i \in \{1, \dots, L\}$, $\mathbb{E}[\log(|g^{(i)}(X_1^L)|^2)] < \infty$.

Moreover,

$$\mathbb{E} \left[\left(\sum_{i=1}^L |g^{(i)}(X_1^L)|^{-2} \right)^3 \right] < \infty.$$

Then the fMSE rate-distortion function is characterized by the following theorem:

Theorem 2: Consider L correlated sources that are encoded separately and an allowed fMSE distortion D . If the joint source density $f_{X_1^L}$ and function g satisfy Conditions FC1–FC4, then the minimum sum rate R_{sum}^* of the rate-distortion region under the fMSE satisfies:

$$\begin{aligned} \lim_{D \rightarrow 0} \left(R_{\text{sum}}^*(D) + \frac{L}{2} \log(2\pi e D/L) \right) \\ = h(X_1^L) + \sum_{i=1}^L \mathbb{E}[\log \gamma_i(X_i)], \quad (10) \end{aligned}$$

where γ_i is defined in (2).

Proof: In the first step of the proof sketch, we demonstrate that the fMSE can be approximated with a W-MSE distortion in the small-distortion regime. Using Taylor's theorem, we rewrite fMSE as:

$$\begin{aligned} \mathbb{E}[|g(X_1^L) - g(\hat{X}_1^L)|^2] &= \\ \mathbb{E} \left[\left| \sum_{i=1}^L g^{(i)}(X_1^L) (X_i - \hat{X}_i) + R_1(X_i - \hat{X}_i) \right|^2 \right], \quad (11) \end{aligned}$$

where $R_1(x)$ is a remainder term of order $\mathcal{O}(|x - \hat{x}|^2)$. As the distortion decreases, the terms in the product that contain R_1 decay quickly and become negligible in the small-distortion limit (see [8]). Moreover, all terms containing $(X_i - \hat{X}_i)(X_j - \hat{X}_j)$ for $i \neq j$ also become negligible as quantization errors become white in the high-resolution limit [16]. Hence, the resulting fMSE distortion can be simplified to

$$\begin{aligned} \mathbb{E}[|g(X_1^L) - g(\hat{X}_1^L)|^2] &\approx \sum_{i=1}^L \mathbb{E} \left[|g^{(i)}(X_1^L)|^2 (X_i - \hat{X}_i)^2 \right] \\ &= \sum_{i=1}^L \mathbb{E} \left[\gamma_i^2(X_i) (X_i - \hat{X}_i)^2 \right], \quad (12) \end{aligned}$$

where the relative error in the approximation can be made arbitrarily small with increasing codebook size. We see that this distortion corresponds to the diagonal W-MSE distortion studied in Section III. In particular, the i th diagonal entry of the weighting matrix W is $\gamma_i^2(X_i)$, ensuring it is nonnegative. Conditions FC1–FC4 then satisfy the sufficient conditions for Theorem 1, thereby giving the minimum sum rate of the achievable rate region. ■

We can now determine the gap between the minimum sum rate using nonuniform scalar quantization followed by Slepian-Wolf coding (SWSQ) and the optimum solution shown in Theorem 2:

Corollary 1: The sum rate from using *optimal* nonuniform scalar quantization followed by Slepian–Wolf coding becomes exactly $(L/2)\log(\pi e/6)$ bits worse than the sum rate R_{sum}^* of the fMSE rate–distortion function in the limit of small distortion. This corresponds to a rate loss of 0.255 bits/sample, which matches the result in the point-to-point setting.

We can also bound the rate loss of ECSQ:

Corollary 2: The sum rate from using *optimal* nonuniform scalar quantization followed by entropy coding becomes exactly $(L/2)\log(\pi e/6) + I(X_1; \dots; X_L)$ bits worse than the sum rate R_{sum}^* of the fMSE rate–distortion function in the limit of small distortion, where $I(X_1; \dots; X_L)$ denotes the multivariate mutual information.

We can similarly derive a rate gap for FRSQ, but the closed-form expression is less intuitive because of dependencies on the source and computation.

A. Remarks

1. Conditions FC1–FC4 and FQ1–FQ3 are not the most general that can be derived. In particular, it is possible to generalize the results to infinite-support densities provided the differential entropy is finite. The precise conditions will parallel the analysis in [8].

2. The results in this section match existing results when the problem is constrained in certain ways. For example, when sources are independent, the problem reduces to the point-to-point SLB with a scalar W-MSE distortion that is a function of the functional sensitivity profile. When the computation is linear, i.e. the functional sensitivity profile is flat, the model reduces to the case studied by Zamir and Berger [14]. When $L = 1$, the result is the scalar W-MSE result [3].

3. Although the gap between DFSQ and the rate–distortion function can be small when the correlation between the quantized observations is coded properly, the multivariate mutual information gap can be quite large. Proper classification of this gap between ECSQ and SWSQ helps motivate engineering choices when designing sensor networks.

4. In (4)–(6), the optimal sum rate is achieved using bit allocation. For a given distortion, there usually exists a unique point in the achievable region where the sum rate is minimized, meaning this is unique set of rates that will achieve the rate gap. This is in comparison to the Slepian–Wolf result, where the minimum sum rate can be achieved by many sets of rates.

V. FUTURE WORK

In this paper, we have characterized the sum-rate gap between coded scalar quantization and the achievable rate region when the fidelity criterion is fMSE. This result is especially encouraging in the distributed setting because it demonstrates that simple compression structures are efficient, even with strict latency constraints and power considerations. This bridge between Shannon source coding theory and high-resolution quantization theory can impact compression design for general network architectures and fidelity criteria. In particular, it motivates the need to design efficient entropy codes and Slepian–Wolf codes on top of scalar-quantizer outputs.

It may also be useful to compare DFSQ to other known achievable results that use coding over long blocks for specific models where the analysis is tractable, e.g. quadratic Gaussian [17]. Additionally, it is interesting to consider extensions to remote sources [18].

In prior work, we demonstrated how intersensor interaction can affect the design of scalar quantizers in distributed architectures [19]. One extension is to determine the rate-loss bounds for this generalization, building off rate–distortion results in [20].

ACKNOWLEDGMENT

This material is based upon work supported by the National Science Foundation under Grant No. 1115159.

REFERENCES

- [1] V. Misra, V. K. Goyal, and L. R. Varshney, “Distributed scalar quantization for computing: High-resolution analysis and extensions,” *IEEE Trans. Inform. Theory*, vol. 57, no. 8, pp. 5298–5325, Aug. 2011.
- [2] T. Berger and J. D. Gibson, “Lossy source coding,” *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2693–2723, Oct. 1998.
- [3] T. Linder and R. Zamir, “High-resolution source coding for non-difference distortion measures: The rate–distortion function,” *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 533–547, Mar. 1999.
- [4] R. M. Gray and D. L. Neuhoff, “Quantization,” *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.
- [5] J. A. Bucklew and G. L. Wise, “Multidimensional asymptotic quantization theory with r th power distortion measures,” *IEEE Trans. Inform. Theory*, vol. IT-28, no. 1, pp. 239–247, Mar. 1982.
- [6] S. Cambanis and N. L. Gerr, “A simple class of asymptotically optimal quantizers,” *IEEE Trans. Inform. Theory*, vol. IT-29, no. 5, pp. 664–676, Sep. 1983.
- [7] T. Linder, “On asymptotically optimal companding quantization,” *Prob. Contr. Inform. Theory*, vol. 20, no. 6, pp. 475–484, 1991.
- [8] J. Z. Sun, V. Misra, and V. K. Goyal, “Distributed functional scalar quantization simplified,” *IEEE Trans. Signal Process.*, 2013, accepted.
- [9] J. J. Y. Huang and P. M. Schultheiss, “Block quantization of correlated Gaussian random variables,” *IEEE Trans. Commun. Syst.*, vol. CS-11, no. 3, pp. 289–296, Sep. 1963.
- [10] C. E. Shannon, “Coding theorems for a discrete source with a fidelity criterion,” *IRE Int. Conv. Rec., part 4*, vol. 7, pp. 142–163, 1959.
- [11] T. Berger, *Rate Distortion Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1971.
- [12] Y. N. Linkov, “Evaluations of epsilon entropy of random variables for small epsilon,” *Prob. Inform. Transm.*, vol. 1, pp. 12–18, 1965.
- [13] T. Linder and R. Zamir, “On the asymptotic tightness of the Shannon lower bound,” *IEEE Trans. Inform. Theory*, vol. 40, no. 6, pp. 2026–2031, Nov. 1994.
- [14] R. Zamir and T. Berger, “Multiterminal source coding with high resolution,” *IEEE Trans. Inform. Theory*, vol. 45, no. 1, pp. 106–117, Jan. 1999.
- [15] T. Linder, R. Zamir, and K. Zeger, “High-resolution source coding for non-difference distortion measures: Multidimensional companding,” *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 548–561, Mar. 1999.
- [16] H. Viswanathan and R. Zamir, “On the whiteness of high-resolution quantization errors,” *IEEE Trans. Inform. Theory*, vol. 47, no. 5, pp. 2029–2038, Jul. 2001.
- [17] Y. Yang, Y. Zhang, and Z. Xiong, “On the sum-rate loss of quadratic Gaussian multiterminal source coding,” *IEEE Trans. Inform. Theory*, vol. 57, no. 9, pp. 5588–5614, Sep. 2011.
- [18] H. Feng, M. Effros, and S. A. Savari, “Functional source coding for networks with receiver side information,” in *Proc. 42nd Annu. Allerton Conf. Commun. Control Comput.*, Sep. 2004, pp. 1419–1427.
- [19] J. Z. Sun and V. K. Goyal, “Chatting in distributed quantization networks,” in *Proc. 50th Ann. Allerton Conf. on Commun., Control and Comp.*, Monticello, IL, Oct. 2012.
- [20] A. H. Kaspi and T. Berger, “Rate–distortion for correlated sources with partially separated encoders,” *IEEE Trans. Inform. Theory*, vol. IT-28, no. 6, pp. 828–840, Nov. 1982.