

# Transform Coding of Densely Sampled Gaussian Data

S. Sandeep Pradhan  
EECS Department  
University of Michigan  
Ann Arbor, MI 48109  
pradhanv@eecs.umich.edu

David L. Neuhoff  
EECS Department  
University of Michigan  
Ann Arbor, MI 48109  
neuhoff@eecs.umich.edu

**Abstract**—With mean-squared error  $D$  as a goal, it is well known that one may approach the rate-distortion function  $R(D)$  of a nonbandlimited, continuous-time Gaussian source by sampling at a sufficiently high rate, applying the Karhunen-Loeve transform to sufficiently long blocks, and then independently coding the transform coefficients of each type. In particular, the coefficients of a given type are ideally encoded with performance attaining a suitably chosen point on the first-order rate-distortion function of that type of coefficient.

This paper considers a similar sample-and-transform coding scheme in which ideal coding of coefficients is replaced by coding with some specified family of quantizers, whose operational rate-distortion function is convex. A prime example is scalar quantization with entropy-coding and, if needed for convexity, time sharing. It is shown that when the sampling rate is large, the operational rate-distortion function of such a scheme comes within a finite constant of  $R(D)$ . Applied to the scalar quantization family, the finiteness of this bound contrasts with a recent result showing that direct scalar quantization of samples (without a transform) has unbounded rate when distortion is held constant and sampling rate becomes large, even when the quantized samples are compressed to their entropy-rate. Thus, at high sampling rates, the transform reduces the loss due to scalar quantization from something infinite to something finite.

## I. INTRODUCTION

Consider a continuous-time, nonbandlimited, stationary Gaussian random process  $X(t)$ . Suppose it is desired to encode and reconstruct this process with distortion measured by a single-letter mean squared error criterion. Then from Kolmogorov's derivation of the rate-distortion function  $R(D)$  [1] (c.f. [2]), which characterizes optimal performance, it is evident that one can approach the rate-distortion function by (a) sampling the source at a sufficiently high rate to obtain a discrete-time random process, (b) applying the Karhunen-Loeve transform [3] to long blocks of samples of this discrete-time process, and (c) independently coding the transform coefficients of each type with an ideal code operating on the first-order rate-distortion of that type. In doing so, one uses inverse "water-pouring" on the continuous-time power spectral density of the original process  $X(t)$  to determine, for each type of coefficient, the appropriate operating point on its first-order rate-distortion function.

This paper considers a similar sample-and-transform coding scheme in which ideal coding of coefficients is replaced by

coding with some specified family of quantizers, whose operational rate-distortion function is convex. A prime example is scalar quantization with entropy-coding and, also, time sharing, if the latter is needed to ensure convexity of the operational rate-distortion function. The principal result shows that when the sampling rate is large, the operational rate-distortion function of such a scheme comes within a finite constant of  $R(D)$ . Applied to the scalar quantization family, the finiteness of this bound contrasts with a recent result [4], [5] showing that direct scalar quantization of samples (without a transform) has unbounded rate when distortion is held constant and sampling rate becomes large, even when the quantized samples are compressed to their entropy-rate. Since the quantized samples become highly correlated as the sampling rate increases, their entropy-rate tends to zero. However, the product of entropy-rate and sampling rate which gives the rate in bits/second, was shown to grow without bound as sampling rate increases [4], [5]. Thus we see that at high sampling rates, the transform reduces the loss due to scalar quantization from something infinite to something finite.

## II. MAIN RESULTS

Let  $\{X(t) : -\infty < t < \infty\}$  be a zero-mean, unit variance, continuous-time, stationary Gaussian process. The goal is to map the sample functions of some finite length  $T$  seconds to indices coming from a set, so that a decoder having access to these random indices can create a reconstruction  $\hat{X}(t)$ . The fidelity of the reconstruction process is measured by the following average mean squared error (MSE):

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T E(X(t) - \hat{X}(t))^2 dt .$$

The rate of this mapping is given by the average number of bits per second produced by the encoder.

Subject to a constraint on MSE (defined by the above equation), we wish to find an upper bound to the rate (in bits per second) of a lossy source coding scheme that operates by first sampling the continuous-time source at sampling period  $1/N$  to produce  $M$  samples  $X(1/N), \dots, X(M/N)$ , and then encoding these discrete samples into bits that are transmitted error-free to a decoder. The decoder first reconstructs the  $M$  samples based on the bits received, and then obtains a

reconstruction of the continuous-time sample function of the source  $\{X(t) : 0 \leq t \leq M/N\}$ . The rate in bits per second expended for the continuous-time source is equal to the rate in bits/sample expended for encoding the  $M$ -length vector times the sampling rate,  $N$  samples per second. The distortion for the continuous-time source is approximately equal to the average distortion of the  $M$ -length vector when a sample-and-hold method is used. This approximation becomes better as sampling rate increases. The objective is to understand the behavior of the rate versus distortion trade-off in encoding and reconstructing a random process defined over an unbounded region with dense sampling. For tractability, we take the limit first as  $M$  becomes large, and then as  $N$  becomes large<sup>1</sup>.

In the following we put some more structure on our encoder. The encoder has two components. The first component involves an  $M \times M$  orthogonal linear transform  $\mathcal{T}^{N,M}$  that maps  $\mathcal{R}^M \rightarrow \mathcal{R}^M$ . The encoder applies this transform to the set of  $M$  samples given by  $(X(1/N), \dots, X(M/N))$  to obtain  $M$  random variables  $(Y(1), \dots, Y(M))$ . The second component involves scalar quantization with a structure described in the sequel. The decoder first reconstructs an approximation (inverse of quantization) to the  $M$ -length vector  $(Y(1), \dots, Y(M))$ , applies the inverse transform, and then uses sample-and-hold to get an approximation to the continuous-time sample function.

**Definition:** A collection  $\mathcal{Q}$  of scalar quantizers is called a *family* if for every  $0 < D < 1$ , there exists a scalar quantizer  $Q$  in the collection such that the mean squared error of  $Q$  when operating on a unit variance discrete-time Gaussian process is at most  $D$ .

The rate of a quantizer  $Q$  in  $\mathcal{Q}$  is given by the average number of bits per sample produced by the quantizer when operating on a unit variance independent identically distributed (IID) Gaussian source<sup>2</sup>. Then, with a family of quantizers  $\mathcal{Q}$ , one can associate a function  $R_{\mathcal{Q}}^*(D)$ , called the *operational rate-distortion function*, which specifies the minimum rate of all the quantizers in  $\mathcal{Q}$  with distortion  $D$  or less when operating on a unit variance independent identically distributed (IID) Gaussian source. A family  $\mathcal{Q}$  of scalar quantizers is said to be *good* if it satisfies the following conditions. The rate-distortion function  $R_{\mathcal{Q}}^*(D)$  of the family is decreasing, strictly convex, and finite except at  $D = 0$ . While it has not been shown that scalar quantization with entropy-coding is a good family, the fact that the slope of its operational rate-distortion function has the same slope as the rate-distortion function [7] suggests that it well may be. It turns out that showing convexity is a thorny issue. If this family is not good, i.e., its operational rate-distortion function is not convex, then

<sup>1</sup>We could have taken  $N$  to a large value first, then  $M$  to a large value. This could have modeled a coding scheme where time-limited portions of sample functions are coded successively with dense sampling. However, this is not tractable at the time of writing this paper.

<sup>2</sup>A deliberate vagueness in the definition is intended to make the proposed coding structure encompass a variety of quantization schemes. Examples include entropy-constrained scalar quantizers, uniform scalar quantizers with arithmetic coding, and Lloyd-Max scalar quantizers with fixed-rate coding.

a bigger family obtained by time-sharing among these scalar quantizers will be good.

In the second component, the encoder, for each integer  $1 \leq i \leq M$ , chooses a scalar quantizer from  $\mathcal{Q}$ , scales it by the standard deviation  $\sigma_i^{N,M}$  of  $Y(i)$ , and uses it to quantize the random variable  $Y(i)$ . Let  $R_1^{N,M}, \dots, R_M^{N,M}$  be the rates allocated to  $Y(1), \dots, Y(M)$  respectively, with distortions given respectively by  $D_1^{N,M}, \dots, D_M^{N,M}$ . Then  $R_i^{N,M} = R_{\mathcal{Q}}^*(D_i^{N,M}/(\sigma_i^{N,M})^2)$ .

Given a random process  $X(t)$  and a good family of scalar quantizers, we can now define the operational rate-distortion function  $R_{\mathcal{Q}}(M, N, D)$  of transform coding with family  $\mathcal{Q}$  as the minimum rate in bits/sec required to attain distortion  $D$  or less, where the minimization is over  $T^{N,M}$  and target distortions  $\{D_i^{N,M}\}_{i=1}^M$ . Further, we define the limiting operational rate-distortion

$$R_{\mathcal{Q}}(D) = \lim_{N \rightarrow \infty} \lim_{M \rightarrow \infty} R_{\mathcal{Q}}(M, N, D).$$

Note that the rate in bits per unit distance is simply the average of  $R_1^{N,M}, \dots, R_M^{N,M}$  (bits per sample) multiplied by the sampling rate  $N$  (samples per unit time). Second, it can be easily shown that there exists a sequence (indexed by  $N$ ) of reconstruction methods (e.g., sample and hold) such that in the limit of large  $N$ , the MSE for the continuous-time reproduction equals the limit (as  $N \rightarrow \infty$ ) of the average of  $D_1^{N,M}, \dots, D_M^{N,M}$ . As a result, we can find an upper bound to  $R_{\mathcal{Q}}(D)$  by finding an attainable tuple  $(R_1^{N,M}, \dots, R_M^{N,M}, D_1^{N,M}, \dots, D_M^{N,M})$ , averaging rates and distortions over  $M$ , and taking limits as  $M$  and  $N$  increase.

We pursue this upper bound to  $R_{\mathcal{Q}}(D)$  by letting  $\mathcal{T}^{N,M}$  be the Karhunen-Loeve transform corresponding to the random vector  $(X(1/N), \dots, X(M/N))$ . This results in the collection of independent random variables  $\{Y(i)\}_{i=1}^M$ , where the variance of  $Y(i)$  is equal to the  $i$ th eigenvalue  $\lambda_i^{N,M}$  of the covariance matrix of the random vector  $(X(1/N), \dots, X(M/N))$ . This results in the following rate per sample and average distortion pair:

$$\begin{aligned} R^{N,M} &= \frac{1}{M} \sum_{i=1}^M R_{\mathcal{Q}}^* \left( \frac{D_i^{N,M}}{\lambda_i^{N,M}} \right) \\ D^{N,M} &= \frac{1}{M} \sum_{i=1}^M D_i^{N,M}. \end{aligned}$$

Using the properties of  $R_{\mathcal{Q}}^*(D)$  and the Karush-Kuhn-Tucker conditions (c.f. [8]), we can perform optimal rate allocation and find that for any  $c < 0$ , the following rate-distortion pair is attainable:

$$\begin{aligned} R_{\mathcal{Q}}^{N,M}(c) &= \frac{1}{M} \sum_{i=1}^M R_{\mathcal{Q}}^* \left( \tilde{D}_{\mathcal{Q}} \left( c \lambda_i^{N,M} \right) \right) \\ D_{\mathcal{Q}}^{N,M}(c) &= \frac{1}{M} \sum_{i=1}^M \lambda_i^{N,M} \min \left\{ 1, \tilde{D}_{\mathcal{Q}} \left( c \lambda_i^{N,M} \right) \right\}, \end{aligned}$$

where  $\tilde{D}_{\mathcal{Q}}(s)$  is the inverse of the first derivative of  $R_{\mathcal{Q}}^*(D)$  for  $s \leq (R_{\mathcal{Q}}^*)'(1)$ , and  $\tilde{D}(s) = 1$  for  $(R_{\mathcal{Q}}^*)'(1) \leq s \leq 0$ , and we have used the fact that  $R_{\mathcal{Q}}^*(D) = 0$  for  $D \geq 1$ . From now on, to simplify notation, we let  $R'_{\mathcal{Q}}(D)$  denote  $(R_{\mathcal{Q}}^*)'(D)$ .

Now using the Grenander-Szego eigenvalue distribution theorem [9], as  $M \rightarrow \infty$ , we get the following achievable rate-distortion pair:

$$R_{\mathcal{Q}}^{N,\infty}(c) = \frac{1}{2\pi} \int_{-\pi}^{\pi} R_{\mathcal{Q}}^* \left( \tilde{D}_{\mathcal{Q}}(c\Phi_N(w)) \right) dw$$

$$D_{\mathcal{Q}}^{N,\infty}(c) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_N(w) \min \left\{ 1, \tilde{D}_{\mathcal{Q}}(c\Phi_N(w)) \right\} dw,$$

where  $\Phi_N(w)$  denotes the discrete-time power spectral density of the process that is obtained by sampling the original process with time-sampling interval  $1/N$ .

We have obtained a parametric formula for an achievable rate-distortion function for the discrete-time process. A rate per unit time is given by  $NR^{N,\infty}$ , and an average distortion for this discrete-time process is given by  $D^{N,\infty}$ . By taking the limit as  $N \rightarrow \infty$ , we obtain our principal result.

**Theorem 1:** An upper bound on the operational rate-distortion function  $R_{\mathcal{Q}}(D)$  function of  $X(t)$ , associated with a good family  $\mathcal{Q}$ , follows directly from the achievability of the following rate and distortion pairs: for  $\theta < 0$ ,

$$R_{\mathcal{Q}}(\theta) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R_{\mathcal{Q}}^* \left( \tilde{D}_{\mathcal{Q}} \left( \frac{S(\Omega)}{\theta} \right) \right) d\Omega \quad \text{bit/second}$$

$$D_{\mathcal{Q}}(\theta) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\Omega) \min \left\{ 1, \tilde{D}_{\mathcal{Q}} \left( \frac{S(\Omega)}{\theta} \right) \right\} d\Omega,$$

where  $S(\Omega)$  is the continuous-time power spectral density of the process  $X(t)$ .

Typical forms of  $R_{\mathcal{Q}}^*(D)$  and  $\tilde{D}_{\mathcal{Q}}(s)$  are shown in Figures 1 and 2.

**Proof:** With the goal of having  $NR^{N,\infty}$  approach a finite value as  $N \rightarrow \infty$ , we take  $c = 1/(\theta N)$  with  $\theta < 0$ . With this choice, we obtain the following statements:

$$R_{\mathcal{Q}}(\theta) \triangleq \lim_{N \rightarrow \infty} NR_{\mathcal{Q}}^{N,\infty} \left( \frac{1}{N\theta} \right)$$

$$= \lim_{N \rightarrow \infty} \frac{N}{2\pi} \int_{-\pi}^{\pi} R_{\mathcal{Q}}^* \left( \tilde{D}_{\mathcal{Q}} \left( \frac{1}{\theta N} \Phi_N(w) \right) \right) dw$$

$$= \lim_{N \rightarrow \infty} \frac{1}{2\pi} \int_{-\pi N}^{\pi N} R_{\mathcal{Q}}^* \left( \tilde{D}_{\mathcal{Q}} \left( \frac{\Phi_N(\frac{\Omega}{N})}{N\theta} \right) \right) d\Omega$$

$$= \lim_{N \rightarrow \infty} \frac{1}{2\pi} \int_{-\infty}^{\infty} I_{[-\pi N, \pi N]}(\Omega) R_{\mathcal{Q}}^* \left( \tilde{D}_{\mathcal{Q}} \left( \frac{\Phi_N(\frac{\Omega}{N})}{N\theta} \right) \right) d\Omega$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \lim_{N \rightarrow \infty} I_{[-\pi N, \pi N]}(\Omega) R_{\mathcal{Q}}^* \left( \tilde{D}_{\mathcal{Q}} \left( \frac{\Phi_N(\frac{\Omega}{N})}{N\theta} \right) \right) d\Omega$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} R_{\mathcal{Q}}^* \left( \tilde{D}_{\mathcal{Q}} \left( \frac{S(\Omega)}{\theta} \right) \right) d\Omega,$$

where  $I_{[-\pi N, \pi N]}(\Omega)$  equals one for  $|\Omega| \leq \pi N$  and zero if not, where the fifth equality uses the Lebesgue convergence theorem [10], while assuming that  $\frac{1}{N}\Phi_N(\Omega/N)$  is bounded

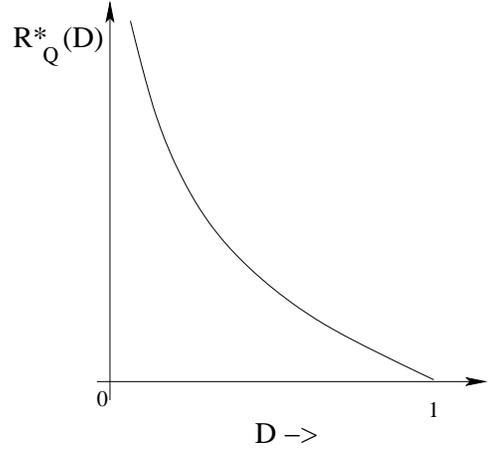


Fig. 1. Typical behavior of the operational rate-distortion function  $R_{\mathcal{Q}}^*(D)$ .

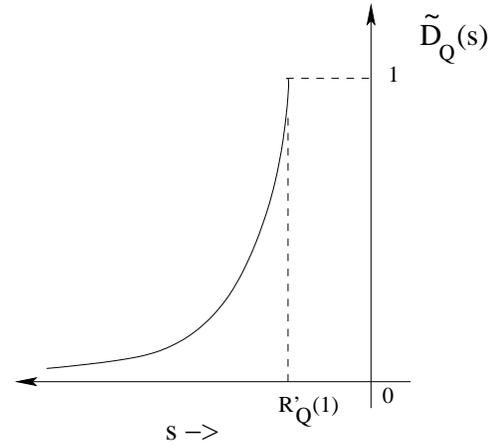


Fig. 2. Typical behavior of  $\tilde{D}(\cdot)$ , the inverse of the first derivative of the operational rate-distortion function, where  $R'_{\mathcal{Q}}(D)$  denotes the first derivative of  $R_{\mathcal{Q}}^*(D)$ .

almost everywhere, and where the sixth uses the easily demonstrated fact that  $\frac{1}{N}\Phi_N(\Omega/N)$  converges to  $S(\Omega)$ . Using similar methods, we can show

$$\hat{D}_{\mathcal{Q}}(\theta) \triangleq \lim_{N \rightarrow \infty} D_{\mathcal{Q}}^{N,\infty}(1/\theta N)$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\Omega) \min \left\{ 1, \tilde{D}_{\mathcal{Q}} \left( \frac{S(\Omega)}{\theta} \right) \right\} d\Omega .$$

Note that although  $\hat{D}_{\mathcal{Q}}(\theta)$  is the limit of the MSE in the reconstruction of the discrete-time random process, as mentioned earlier, it can be shown that there exists a sequence of methods (indexed by  $N$ ) for reconstructing the original continuous-time random process from the quantized samples of the corresponding discrete-time random process (obtained with sampling frequency  $N$ ) such that the limit  $D_{\mathcal{Q}}(\theta)$  of the MSE for the continuous-time process equals  $\hat{D}_{\mathcal{Q}}(\theta)$ . This finishes the proof that  $(R_{\mathcal{Q}}(\theta), D_{\mathcal{Q}}(\theta))$  is an achievable rate-distortion pair, and the theorem follows.  $\square$

To see that  $R_{\mathcal{Q}}(\theta)$  is finite for any  $D > 0$ , let us first observe

that the facts that  $R_{\mathcal{Q}}^*(D)$  is no smaller than the Shannon rate-distortion function of unit variance IID Gaussian source, and that the first derivative of the Shannon rate-distortion function at  $D = 1$  is  $-1/2$ , imply that  $R_{\mathcal{Q}}^*(D)$  must satisfy  $R'_{\mathcal{Q}}(1) \leq -1/2$ . Second, since  $X(t)$  is a unit variance Gaussian process, and  $R'_{\mathcal{Q}}(1) \leq -1/2$ , for  $\theta \neq 0$  the Lebesgue measure of the set  $\{\Omega : S(\Omega) \geq \theta R'_{\mathcal{Q}}(1)\}$  is finite. Third, let us rewrite the above expressions for the rate and distortion in the following form:

$$R_{\mathcal{Q}}(\theta) = \frac{1}{2\pi} \int_{\{\Omega: S(\Omega) \geq \theta R'_{\mathcal{Q}}(1)\}} R_{\mathcal{Q}}^* \left( \tilde{D}_{\mathcal{Q}} \left( \frac{S(\Omega)}{\theta} \right) \right) d\Omega$$

$$D_{\mathcal{Q}}(\theta) = \frac{1}{2\pi} \int_{\{\Omega: S(\Omega) \geq \theta R'_{\mathcal{Q}}(1)\}} S(\Omega) \left( \tilde{D}_{\mathcal{Q}} \left( \frac{S(\Omega)}{\theta} \right) \right) d\Omega$$

$$+ \frac{1}{2\pi} \int_{\{\Omega: S(\Omega) \leq \theta R'_{\mathcal{Q}}(1)\}} S(\Omega) d\Omega .$$

Since  $R_{\mathcal{Q}}^*(D)$  is bounded except at  $D = 0$ , and  $\tilde{D}_{\mathcal{Q}}(S(\Omega)/\theta)$  is bounded (because  $S(\Omega)$  is bounded), it follows that for any  $D > 0$ ,  $R_{\mathcal{Q}}(D)$  is finite.

### III. DISTRIBUTION OF EIGENVALUES

Now we contrast the proposed approach with that of [4], [5]. Both methods employ scalar quantization. In the latter approach, however, since there is no transform, and the process is stationary, identical scalar quantizers are allocated to every sample  $X(i/N)$  for  $i = 1, 2, \dots, M$ . Since a scalar quantizer has a certain performance gap with respect to the optimal Shannon rate-distortion function, cumulatively this effect contributes to making the product of the entropy-rate and the sampling rate approach infinity as the sampling rate tends to infinity. On the other hand, with transform coding, efficient allocation of quantizers to transform coefficients keeps the rate at a finite value. We would expect then that a significant fraction of the transform coefficients be allocated nearly zero rate. This is specially important in light of the result shown in [7] that scalar quantization is optimal at low rates for Gaussian sources with mean squared error. The following discussion confirms this intuition.

In the following we provide a characterization of the asymptotic distribution of eigenvalues of the autocovariance matrix of the random vectors  $(X(1/N), \dots, X(M/N))$ .

Let for any  $a > 0$

$$Z_{N,M} \triangleq \frac{1}{M} \sum_{i=1}^M I_{[0,aN]}(\lambda_i^{N,M}) .$$

**Theorem 2:**

$$\lim_{N \rightarrow \infty} \lim_{M \rightarrow \infty} Z_{N,M} = 1 .$$

**Proof:** Let  $Z_N = \lim_{M \rightarrow \infty} Z_{N,M}$ . Then

$$Z_N = \frac{1}{2\pi} \int_{-\pi}^{\pi} I_{[0,a]}(\Phi_N(w)/N) dw$$

$$= \frac{1}{2\pi N} \int_{-\pi N}^{\pi N} I_{[0,a]} \left( \frac{\Phi_N(\Omega/N)}{N} \right) d\Omega$$

$$= \frac{1}{2\pi N} |\{\Omega : 0 \leq A_N(\Omega) \leq a, -\pi N \leq \Omega \leq \pi N\}| .$$

where  $A_N(\Omega) = \frac{1}{N} \Phi_N(\Omega/N)$ . Now taking the limit as  $N \rightarrow \infty$  and using the fact that  $A_N(\Omega)$  converges to  $S(\Omega)$  as  $N \rightarrow \infty$ , we have,

$$\lim_{N \rightarrow \infty} Z_N = \lim_{N \rightarrow \infty} \frac{|\{\Omega : 0 \leq S(\Omega) \leq a, -\pi N \leq \Omega \leq \pi N\}|}{2\pi N}$$

$$= 1 .$$

□

The significance of this result can be appreciated by making the following observations. First note that, for large  $N$ ,  $\Phi_N(\Omega) \approx NS(\Omega/N)$ . Hence the maximum value of  $\Phi_N(\Omega)$  grows linearly with  $N$ . Second, using the arguments in the proof of the above theorem, one can also infer that for large  $N$ , for  $0 \leq b \leq c \leq \max \Phi_N(\Omega)$ , and  $|b-c|$  sufficiently small, the fraction of eigenvalues of the random vector  $(X(1/N), \dots, X(M/N))$  that lie in the interval  $(b, c)$  is approximately equal to

$$\frac{1}{2\pi N} |\{\Omega : b \leq A_N(\Omega) \leq c, -\pi N \leq \Omega \leq \pi N\}| .$$

Hence some of the eigenvalues will have values that increase linearly with  $N$ . But Theorem 2 says for any  $a > 0$ , almost all of them will have values less than  $aN$  for large  $N$ . Therefore most of the eigenvalues are very small and are coded to distortions that are nearly equal to their variances, and cumulatively incur insignificant performance loss.

*Remark:* Although, the preceding discussion concentrated on good families of scalar quantizers, the two results presented above can be easily extended to any good family of quantizers in a straightforward way.

### ACKNOWLEDGMENT

This work was supported by NSF Grants CCF-0329715 and CCF-0448115 (CAREER).

### REFERENCES

- [1] N. Kolmogorov, "On the Shannon theory of information transmission in the case of continuous signals," *IRE Trans. Inform. Theory*, vol. 2, pp. 102-108, 1956.
- [2] T. Berger, *Rate distortion theory: A mathematical basis for data compression*. Englewood Cliffs: Prentice Hall, 1971.
- [3] H. L. van Trees, *Detection estimation and modulation theory*. Wiley: New York, 1968.
- [4] D. Marco, E. J. Duarte-Melo, M. Liu, and D. L. Neuhoff, "On the many-to-one transport capacity of a dense wireless sensor network and the compressibility of its data," *Proc. Information processing in sensor networks (IPSN)*, 2003.
- [5] D. Marco and D.L. Neuhoff, "Entropy of quantized data at high sampling rates," *Proc. IEEE Int. Symp. Inform. Thy.*, Adelaide, Aus., Sept. 2005.
- [6] T. M. Cover and J. A. Thomas, *Elements of information theory*. New York: Wiley, 1991.
- [7] D. Marco and D.L. Neuhoff, "Low resolution scalar quantization for Gaussian sources and squared error," *IEEE Trans. Inform. Theory*, vol. 52, pp. 1689-1697, April 2006.
- [8] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 2003.
- [9] U. Grenander and G. Szegö, *Toeplitz forms and their applications*. Berkeley, CA: University of California Press, 1958.
- [10] H. L. Royden, *Real analysis*. Englewood Cliffs: Prentice Hall, 1988.