

# Centralized and Distributed Lossy Source Coding of Densely Sampled Gaussian Data, with and without Transforms

David L. Neuhoff and S. Sandeep Pradhan  
EECS Department, University of Michigan  
Ann Arbor, MI 48109, USA  
{neuhoff, pradhanv}@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 spatially nonbandlimited, time IID, continuous-space, discrete-time Gaussian source by spatially sampling at a sufficiently high rate, applying the Karhunen-Loeve transform to sufficiently long blocks, and independently coding transform coefficients of each type at the first-order rate-distortion function of that type, with a distortion target chosen appropriately for that type. This paper compares and contrasts this classical result with several recently explored alternative schemes for encoding source samples taken at a high rate. The first scheme, which scalar quantizes the samples and then losslessly encodes the quantized samples at their entropy-rate, is known to have rate approaching infinity when distortion is held at  $D$ . Is such catastrophic behavior due to the scalar quantizer or to the distributed nature of the quantization? Recent results show that even without a transform, but with distributed vector quantization, it is possible to attain performance that differs from the rate-distortion function by only a finite constant. This suggests it was the scalar quantizer that caused the catastrophic behavior. The final recent result suggests the situation is more nuanced, because it shows that if in the classical scheme scalar quantizers with entropy coding replace the ideal coding of the coefficients at their first-order rate-distortion functions, then again performance differs from the rate-distortion function by a finite constant.

## I. INTRODUCTION

Consider the task of lossy encoding of a continuous-time, stationary, Gaussian random process  $X_t$ , either bandlimited or not, with the goal of minimizing the encoding rate in bits/sec, while attaining a target mean squared error (MSE) distortion  $d$ . In this paper we compare and contrast the encoding rates attainable with four schemes, each based on coding samples taken at a high rate. Two schemes use distributed encoding, whereas the other two use transforms, which cannot be implemented in a distributed fashion. Two schemes use scalar quantization, whereas the other two use block coding, i.e. vector quantization. We are interested in understanding the performance limitations due to scalar quantization, distributed coding and the inability to use a transform.

The first scheme can be considered to be classical (c.f. [1], pp. 116). It applies the Karhunen-Loeve Transform (KLT) to long blocks of samples followed by independent block coding of each type of transform coefficient. When the sampling rate is large and the block codes are suitably chosen, the encoding rate approaches the rate-distortion function  $\mathcal{R}(d)$  of the

continuous-time source  $X_t$ , i.e. the best possible performance, which can be expressed parametrically in the form

$$\mathcal{R}(\theta) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \max\left\{\frac{1}{2} \log_2 \frac{S(\Omega)}{\theta}, 0\right\} d\Omega \quad (1)$$

$$\mathcal{D}(\theta) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \min\{S(\Omega), \theta\} d\Omega \quad (2)$$

where  $S(\Omega)$  is the power spectral density of  $X_t$ ,  $\theta \geq 0$  is a free parameter, and  $\mathcal{R}(\theta)$  has units of bits/sec (bps). This parametric formula originated with Kolmogorov [2]. An alternate scheme for approaching  $\mathcal{R}(d)$ , which uses a continuous-time KLT without sampling [3], [1], will not be considered in this paper.

The second scheme is conceptually simpler. The samples are scalar quantized and losslessly encoded at the entropy-rate of the quantized samples. With distortion held constant, the entropy-rate of the quantized samples goes to zero due to increasing correlation. However, as shown by Marco et al. [4], [5], the rate in bits/sec, namely the entropy-rate times the sampling rate, tends to infinity as sampling rate increases. Such catastrophically bad performance contrasts sharply with the performance of the classical scheme, described in the previous paragraph, which can attain any target distortion  $d$  with a finite rate. It also contrasts with the fact that when coding a discrete-time process, scalar quantization plus entropy-rate coding attains performance close to the rate-distortion function of the discrete-time sampled process. For example, at low distortions, its encoding rate exceeds the rate-distortion function of the sampled process by only 0.255 bits/sample [6]. Evidently, when multiplied by the sampling rate, the excess encoding rate in bits/sample required by scalar quantization with entropy coding approaches infinity.

Aside from its conceptual simplicity, the focus on scalar quantization with lossless coding in [4], [5] was originally motivated by the problem of field-gathering with a wireless sensor networks, which required distributed lossy source encoding. In particular, the samples of a random process taken at one sampling instance must be encoded without knowledge of samples taken at other instances. (In the sensor context,  $t$  is a spatial variable.) Let us now suppose that there are independent sample functions  $X_{t,u}$ ,  $u \in \mathbb{Z}$ , of the process  $X_t$ . (In the sensor context,  $u$  is a temporal variable.) In this case

Slepian-Wolf coding [7] can be used to losslessly encode the quantized samples in a distributed fashion, at a rate equal to the entropy-rate of the quantized samples. Specifically, at any sampling times  $t_i, t_j$ , the quantized samples  $\{\widehat{X}_{t_i,u}\}_{u=-\infty}^{\infty}$  at  $t_i$  are encoded independently of those at  $t_j$ ,  $\{\widehat{X}_{t_j,u}\}_{u=-\infty}^{\infty}$ . It was indeed disappointing that the performance of scalar quantization with lossless coding should be so bad at high sampling rates.

It is natural to ask whether the poor performance of the second scheme is due to the use of scalar, rather than block coding, or is it a limitation of all distributed lossy coding schemes. Thus the third scheme to consider is ideal distributed lossy source coding of the samples taking at high rate, where the block coding is across  $u$ . The performance of such a scheme was analyzed recently by Kashyap et al. [8], who found an upper bound to the rate of ideal distributed lossy coding applied with distortion at most  $d$  that remains bounded as sampling rate increases. Thus ideal distributed lossy source coding does not have catastrophically bad behavior. Indeed, it was shown recently by the present authors [9] that its performance can often be close to  $\mathcal{R}(d)$ . One concludes that in the second scheme it is scalar quantization, rather than distributed coding, that causes the very poor performance.

Given the poor performance of scalar quantization in the second scheme, one wonders if performance would be equally bad if scalar quantization with entropy coding replaced block coding in the classic scheme – after the transform. Recently, the authors of the present paper [10] have shown that in the limit of high sampling rate, the encoding rate of such a scheme also remains finite. Thus there are no catastrophic effects to using scalar quantization when preceded by a transform.

In the remainder of this paper, we briefly review and compare the four schemes.

## II. TRANSFORM, BLOCK CODING

We begin by reviewing how performance approaching the rate-distortion function  $\mathcal{R}(d)$ , i.e. (1),(2), can be attained with the classical scheme [1]. Suppose  $X_t$  has mean zero, unit power, and power spectral density  $S(\Omega)$ . In the classical approach, one (a) samples the process at some rate  $N$  samples per second, where for convenience we take  $N$  to an integer; (b) applies the KLT to long blocks of, say,  $M$  successive samples, obtaining a block of  $M$  independent transform coefficients from each block; (c) collects the coefficients of each type from all blocks into a separate stream; (d) independently lossy block encodes each of the  $M$  streams with a target distortion  $d_{N,M,i}$  tailored specifically to the  $i$ th stream and with a rate equal to the first-order rate-distortion function<sup>1</sup>  $R_{N,M,i}(d_{N,M,i})$  of that stream at its target distortion; and (e) multiplexes the  $M$  streams into one stream of bits. To decode, one (a) demultiplexes the received stream of bits into  $M$  streams; (b) applies

<sup>1</sup>We use script fonts  $\mathcal{R}$  and  $\mathcal{D}$  to denote rates and distortions for continuous-time processes. We use ordinary  $R$  and  $D$  to denote rates and distortions for discrete-time processes. The  $i$ th stream is a stationary Gaussian random process with variance  $\lambda_{N,M,i}$ . In general, it is not IID. The first-order rate-distortion function is the rate-distortion function of an IID process with the same first-order distribution.

the corresponding lossy source decoder to each stream; (c) recreates blocks of  $M$  coefficients from the decoded streams; (d) applies the inverse KLT to each block, obtaining a block of  $M$  reconstructed samples; (e) concatenates the blocks to obtain a sequence of reconstructed samples; and (f) applies waveform reconstruction to the samples – for example, sample-and-hold reconstruction will suffice.

For the discrete-time samples, the distortion and rate attained with such a scheme are the averages of those for each stream, namely,

$$D_{N,M} = \frac{1}{M} \sum_{i=1}^M d_{N,M,i} \quad (3)$$

$$R_{N,M} = \frac{1}{M} \sum_{i=1}^M R_{N,M,i}(d_{N,M,i}) \text{ bits/sample.} \quad (4)$$

Let  $\lambda_{N,M,1} \geq \lambda_{N,M,2} \geq \dots \geq \lambda_{N,M,M}$  denote the eigenvalues, in decreasing order, of the covariance matrix  $C_{N,M}$  of  $M$  successive samples from the source  $X_t$  taken at intervals of  $1/N$ . They are also the variances of the  $M$  streams.

Given a target distortion  $d$ , to attain the performance given in (1,2), one chooses the  $d_{N,M,i}$ 's as follows. One first finds  $\phi_N \geq 0$  such that

$$\frac{1}{M} \sum_{i=1}^M \min\{\lambda_{N,M,i}, \phi_N\} = d.$$

Then one chooses

$$d_{N,M,i} = \begin{cases} \lambda_{N,M,i}, & \lambda_{N,M,i} \leq \phi_N \\ \phi_N, & \lambda_{N,M,i} > \phi_N \end{cases}$$

Substituting this into (3), the distortion in the reconstructed samples becomes

$$D_{N,M} = \frac{1}{M} \sum_{i=1}^M \min\{\lambda_{N,M,i}, \phi_N\} = d. \quad (5)$$

Substituting the well known form of the Gaussian rate-distortion function

$$R_{N,M,i}(d) = \max\left\{\frac{1}{2} \log_2 \frac{\lambda_{N,M,i}}{d}, 0\right\}$$

into (4), the rate in bits per sample becomes

$$R_{N,M} = \frac{1}{M} \sum_{i=1}^M \max\left\{\frac{1}{2} \log_2 \frac{\lambda_{N,M,i}}{\phi_N}, 0\right\}. \quad (6)$$

The performance in (1,2) is obtained by setting  $\phi_N = N\theta$  and letting  $M$  tend to infinity for fixed  $N$ , then letting  $N$  tend to infinity. Specifically, with  $N$  fixed, the Grenander-Szego asymptotic eigenvalue theorem [11] shows that for any function  $g$

$$\lim_{M \rightarrow \infty} \sum_{i=1}^M g(\lambda_{N,M,i}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\Phi_N(\omega)) d\omega$$

where  $\Phi_N(\omega)$  is the power spectral density of the discrete-time process obtained by sampling  $X_t$  at rate  $N$ .

Therefore, taking the limits of (6) and (5) yields rate (in bits per sample) and distortion

$$R_N(\phi_N) = \lim_{M \rightarrow \infty} R_{N,M}(\phi_N) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \max\left\{\frac{1}{2} \log_2 \frac{\Phi_N(\omega)}{\phi_N}, 0\right\} d\omega$$

$$D_N(\phi_N) = \lim_{M \rightarrow \infty} D_{N,M}(\phi_N) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \min\{\Phi_N(\omega), \phi_N\} d\omega$$

For the continuous-time process the rate in bits/second is

$$\mathcal{R}_N(\phi_N) = NR_N(\phi_N)$$

and with sample and hold reconstruction, the distortion  $\mathcal{D}_N(\phi_N)$  of the reconstructed waveform is approximately equal to the  $D_N(\phi_N)$ , the MSE of the discrete-time samples (before waveform reconstruction). More specifically,

$$\lim_{N \rightarrow \infty} (D_N(\phi_N) - \mathcal{D}_N(\phi_N)) = 0.$$

Now, fixing  $\theta \geq 0$ , making the change of variables  $\Omega = \omega N$  in the integrals, taking limits as  $N \rightarrow \infty$  with  $\phi_N = N\theta$ , and using the above and the fact that  $\Phi_N(\Omega/N)/N \rightarrow S(\Omega)$ , one obtains (1,2).

### III. NO TRANSFORM, SCALAR QUANTIZATION

The fact that with distortion held constant, scalar quantization with entropy coding has rate (in bits/sec) approaching infinity is due, essentially, to the fact that scalar quantization captures too much information. For example, with a fixed scalar quantizer, and consequently, with distortion of the reconstructed waveform tending to the distortion of the scalar quantizer, as sampling rate increases, one can recover the times at which the  $X_t$  crosses the quantizer thresholds with arbitrary accuracy. Since these times are random variables with infinite entropy, it must be that the entropy-rate of the quantized samples approaches infinity [5].

### IV. NO TRANSFORM, DISTRIBUTED BLOCK CODING

Next we consider bounds to the performance of ideal distributed lossy coding operating directly on the samples, without a transform. Block coding across the parameter  $u$  is used. Although the optimal performance of distributed lossy codes is not known, except in the case of two encoders, one can use the Berger-Tung [12] bound to attain an upper bound to rate attainable with distributed coding at a given target distortion  $d$ . This approach was used in [8] to find a finite upper bound to the rate attainable with distortion  $d$  when coding  $N$  successive samples taken at rate  $N$  over a unit interval with distortion target  $d$ . It was also used in [9] to bound the rate attainable when coding  $M$  samples taken at rate  $N$  over an interval of length  $M/N$  in the limit of large  $M$ . We focus now on the latter case, since it leads to better performance and simpler formulas.

With the Berger-Tung approach, one can straightforwardly show that for any  $\phi_N \geq 0$  the following rate (in bits per

sample) and distortion are attainable with distributed coding of  $M$  adjacent samples taken at rate  $N$ :

$$R_{N,M}^d(\phi_N) = \frac{1}{2M} \sum_{i=1}^M \log \left( \frac{\lambda_{N,M,i}}{\phi_N} + 1 \right)$$

$$D_{N,M}^d(\phi_N) = \frac{1}{M} \sum_{i=1}^M \frac{\lambda_{N,M,i}}{\frac{\lambda_{N,M,i}}{\phi_N} + 1}.$$

With  $N$  fixed, we let  $M \rightarrow \infty$ , in which case the Grenander-Szego eigenvalue theorem shows that the following limits are attainable:

$$R_N^d(\phi_N) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \log \left( \frac{\Phi_N(\omega)}{\phi_N} + 1 \right) d\omega$$

$$D_N^d(\phi_N) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left( \frac{\Phi_N(\omega)}{\frac{\Phi_N(\omega)}{\phi_N} + 1} \right) d\omega.$$

Now, just as in Section II, changing variables, letting  $\phi_N = \theta N$  for  $\theta > 0$ , and taking the limit as  $N \rightarrow \infty$  yields the following attainable rates (in bits per second) and distortions for the continuous time process:

$$\mathcal{R}^d(\theta) = \frac{1}{4\pi} \int_{-\infty}^{\infty} \log \left( \frac{S(\Omega)}{\theta} + 1 \right) d\Omega$$

$$\mathcal{D}^d(\theta) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left( \frac{S(\Omega)}{\frac{S(\Omega)}{\theta} + 1} \right) d\Omega$$

### V. TRANSFORM, SCALAR QUANTIZATION

The last scheme to consider replaces the ideal block coding of transform coefficients in the classical scheme with scalar quantization with entropy coding [10]. Let  $R_{sq}(d)$  denote the operational rate-distortion function of scalar quantization with entropy coding for a unit variance Gaussian variable, which is defined to be the least output entropy of any scalar quantizer applied to a unit variance Gaussian random variable. To simplify discussion, we assume either  $R_{sq}(d)$  is strictly convex, or time-sharing of scalar quantizers is permitted so as to make it convex<sup>2</sup>.

With an  $M$ -dimensional KLT and target distortion  $\{d_{N,M,i}\}_{i=1}^M$ , the following rates and distortions are attained

$$R_{N,M}^{sq} = \frac{1}{M} \sum_{i=1}^M R_{sq} \left( \frac{d_{N,M,i}}{\lambda_{N,M,i}} \right)$$

$$D_{N,M}^{sq} = \frac{1}{M} \sum_{i=1}^M d_{N,M,i}.$$

Using the Karush-Kuhn-Tucker conditions (c.f. [13]), one can show that the following rate and distortion are attainable

<sup>2</sup>To the authors' knowledge, it is not known if  $R_{sq}(d)$  is convex. The scalar quantizers considered here could be replaced by any family of lossy source codes with a convex operational rate-distortion function.

for any  $c_{N,M} < 0$ :

$$R_{N,M}^{sq}(c_{N,M}) = \frac{1}{M} \sum_{i=1}^M R_{sq}(\tilde{D}(c_{N,M}\lambda_{N,M,i}))$$

$$D_{N,M}^{sq}(c_{N,M}) = \frac{1}{M} \sum_{i=1}^M \lambda_{N,M,i} \min \left\{ 1, \tilde{D}(c_{N,M}\lambda_{N,M,i}) \right\}$$

where  $\tilde{D}(s)$  is the inverse of the first derivative of  $R_{sq}(d)$  for  $s \leq R'_{sq}(1)$ , and  $\tilde{D}(s) = 1$ , for  $R'_{sq}(1) \leq s \leq 0$ .

Again using the Grenander-Szego eigenvalue theorem [11], it can be shown that in the limit of large  $M$  the following rate-distortion pair is attainable:

$$R_N^{sq}(c_N) = \frac{1}{2\pi} \int_{-\pi}^{\pi} R_{sq}(\tilde{D}(c_N\Phi_N(\omega))) d\omega$$

$$D_N^{sq}(c_N) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_N(\omega) \min \left\{ 1, \tilde{D}(c_N\Phi_N(\omega)) \right\} d\omega$$

where  $c_{N,M} = c_N < 0$  for all  $M$ .

As in previous sections, changing variables, letting  $c_N = 1/\theta N$  for  $\theta < 0$ , multiplying rate in bits/sample by the sampling rate  $N$ , and taking the limit as  $N \rightarrow \infty$  yields the following attainable rates and distortions for the continuous time process:

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

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

## VI. COMPARISONS AND REMARKS

Let us first consider the first case of block coding with transform. When  $N$  and  $M$  are appropriately large, the coefficients corresponding to frequency  $\Omega$  are coded with rate  $\frac{1}{2} \log \frac{S(\Omega)}{\theta}$  incurring a distortion of  $\theta$  if the variance of the coefficient is greater than  $\theta$  or else it is discarded. In other words, the rate allocated to frequency  $\Omega$  is zero if the corresponding  $S(\Omega)$  is below a threshold  $\theta$ . Further, the distortion allocated to frequency  $\Omega$  is at a constant level  $\theta$  if the corresponding  $S(\Omega)$  is above the threshold. To illustrate this better let us consider an example of a Gauss-Markov source with continuous-time power spectral density given by  $S(\Omega) = 1/(1 + \Omega^2)$ . The distortion profile for the case of block coding with transform is depicted in Figure 1 for the case when  $D = 1/3$ .

For the second case of scalar quantization without a transform (or an identity transform), since the process is stationary, every sample is assigned the same quantizer. Since a scalar quantizer has a certain fixed performance gap with respect to that of the ideal block coding (vector quantizer), cumulatively, this has the effect of making the product of the entropy-rate and the sampling rate tend to a large value as sampling rate becomes high.

For the third case of distributed block coding (without transform), the encoders cannot decorrelate the source spatially, and hence can only add quantization noise which is white to the source. This results in a non-flat distortion profile

when minimum mean squared error (MMSE) estimation is performed at the decoder. This is illustrated in Figure 1. From the finiteness of the operational rate-distortion function in this case, one can infer that the catastrophic loss of scalar quantization in the absence of a transform is mitigated by the temporal block quantization of the samples. Further, one should note that the block quantizers assigned to different spatial samples of the source are different, a fact that often goes unnoticed in the details of the random coding argument used in the Berger-Tung inner bound. To elaborate on this, observe that in this random coding argument, a block quantizer for each spatial sample is chosen randomly and independently of that of the other samples. Hence with probability one block quantizers of different spatial samples will all be different. After evaluating the average performance of these random quantizers, the existence of a collection of good quantizers (one for each spatial sample) is shown.

In essence one can gather enough useful information about the source by using different quantizers of large blocklength for different spatial source samples without expending too much rate in bits per second. Note that as the block-length increases, the disparity among these quantizers also increases as the degrees of freedom in choosing the quantizers increases with dimension. This diversity in the block quantizers seems to be the key that keeps the operational rate-distortion function from blowing up for  $D > 0$ .

Now for the fourth case of scalar quantization with a transform, since the encoder has the freedom to decorrelate the source, the scheme is again amenable to the use of a diverse set of quantizers albeit scalar. Roughly speaking, transform coefficients with higher variance are allocated more rate than those with smaller variance. Further, one can show that the fraction of the coefficients which are allocated zero rate goes to one as  $N$  and  $M$  become large. This effect seems to keep the operational rate-distortion function finite for non-zero distortions. We can now compare the distortion profile of this scheme with those of the first and the third scheme. Observe from the operational rate-distortion function that those coefficients which correspond to frequency  $\Omega$  such that  $S(\Omega) \leq \theta R'_{sq}(1)$  are discarded, i.e., allocated distortion equal to  $S(\Omega)$ . For those coefficients which are allocated non-zero rate, the distortion allocated is given by  $S(\Omega)\tilde{D}(S(\Omega)/\theta)$ . Hence within the region  $\{\Omega : S(\Omega) \geq \theta R'_{sq}(1)\}$ , the distortion profile as a function of  $\Omega$  need not be flat in contrast to that of the first scheme. However, unlike that of the third scheme where non-zero rate is allocated for every  $\Omega$  for which  $S(\Omega) > 0$ , here for those  $\Omega$  where  $S(\Omega)$  is small, zero rate is allocated. Thus the distortion profile is a mix between that of the first and the third scheme.

## ACKNOWLEDGMENT

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

## REFERENCES

- [1] T. Berger, *Rate distortion theory: A mathematical basis for data compression*. Englewood Cliffs: Prentice Hall, 1971.

- [2] A. 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.
- [3] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 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)*, Palo Alto, CA., pp. 1-16, Apr. 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] H. Gish and J. N. Pierce, "Asymptotically efficient quantization," *IEEE Trans. Info. Theory*, vol. 14, pp. 676-683, Sep. 1968.
- [7] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Info. Theory*, vol. 19, pp. 471-480, Jul. 1973.
- [8] A. Kashyap, L. A. Lastras-Montano, C. Xia, and Z. Liu, "Distributed source coding in dense sensor networks," *Data Compression Conference (DCC)*, Snowbird, UT, Apr. 2005.
- [9] D. L. Neuhoff and S. S. Pradhan, "An Upper Bound to the Rate of Ideal Distributed Lossy Source Coding of Densely Sampled Data," *ICASSP*, Toulouse, May 2006.
- [10] S. S. Pradhan and D. L. Neuhoff, "Transform Coding of Densely Sampled Gaussian Data," submitted to *ISIT 2007*, Nice, France, Jan. 2007.
- [11] U. Grenander and G. Szego, *Toeplitz forms and their applications*. Berkeley, CA: University of California Press, 1958.
- [12] T. Berger, "Multiterminal Source Coding," *Information Theory Approach to Communication, (CISM Courses and Lecture Notes No. 229)*, G. Longo, Ed., Wien and New York: Springer-Verlag, 1977.
- [13] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 2003.

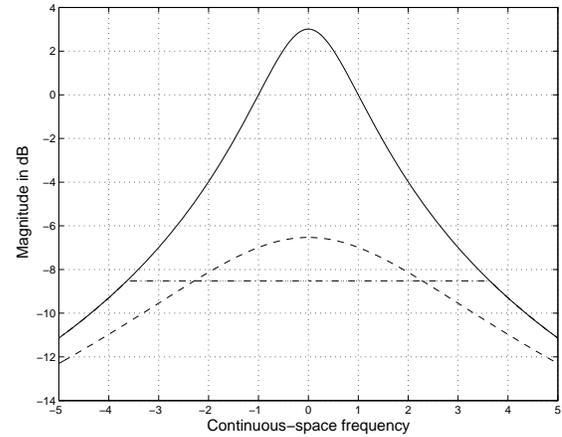


Fig. 1. Gauss-Markov source: The solid curve denotes the continuous-space power spectral density, the dash-dotted curve denotes the distortion profile due to the optimum block coding with transform scheme (inverse water pouring), and the dashed curve denotes the distortion profile due to distributed block coding (Scheme 3).