HIGH-RATE ANALYSIS OF VECTOR QUANTIZATION FOR NOISY CHANNELS

Chandra R. Murthy, Ethan R. Duni and Bhaskar D. Rao

Department of Electrical and Computer Engineering
University of California San Diego
La Jolla, CA 92093-0407
Email: {cmurthy, eduni, brao}@ucsd.edu

ABSTRACT

In this paper, the sensitivity of the high-rate performance of conventional source coding to symmetric channel errors (i.e., a channel where all index errors are equally likely) with arbitrary distortion measures is analyzed. It is shown that, in general, the overall distortion due to source quantization and channel errors cannot be expressed as the sum of the distortion due to the finite bit representation of the source and the distortion due to channel errors. An exception to this is when the distortion is measured as the mean-squared error. The binary symmetric channel with random index assignment is a special case of the analysis, and as the number of code-points gets large, the performance approaches a nonzero constant. Finally, the framework is applied to the wideband speech spectrum quantization problem, where it correctly predicts the channel error rate permissible for operation at a particular distortion level.

1. INTRODUCTION

The performance of a source quantization scheme can be very sensitive to errors introduced when the codepoint index is transmitted over a noisy channel, because the quantization process typically involves removing the redundancy in the source and encoding only the non-redundant part. For example, speech is typically compressed using a highly efficient vector quantization (VQ) scheme prior to transmission over a noisy channel, and the resulting indices could be very sensitive to errors in the channel over which they are transmitted. Hence, the performance of VQ when the index is sent over a noisy channel is pertinent, if they are to be used in practical communication systems.

Past works on VQ for noisy channels have adopted one of two approaches. The first is to replace the distortion measure used for optimizing the quantizer with the expected distortion over the noisy channel (e.g., [1] - [3]). The second approach involves *index assignment*, i.e., designing the quantizer without considering channel errors and then coding the indices to ensure that codeword pairs resulting in small (large) distortions are mapped to index pairs with large (small) transition probability (e.g., [4], [5]). Other recent works on source quantization for noisy channels include [6] and [7]. Most of the works in the literature employ the mean-squared error as the distortion function. However, more general distortion measures are considered here (such as Log Spectral Distortion in wideband speech spectrum quantization) and the performance of source compression under channel errors is examined.

This work was supported in part by UC Discovery grant no. Com 02-10105 sponsored by Conexant Systems Inc.; and by MICRO Grants 03-073, 04-074 and 05-033, sponsored by Qualcomm Inc.

In the sequel, based on classical results from the source coding literature [8] - [11], a novel technique is developed to analyze the effect of errors on the performance of source coding with arbitrary distortion measures for *symmetric error channels* (channels for which all index errors are equally likely). Clearly, it would be convenient if the overall distortion could be decomposed as the sum the distortion due to the source encoder and the distortion induced by channel errors. It is shown that while this decomposition is possible when the distortion is measured as the mean-squared error, in general there is an interdependence between the source and channel errors. The analysis presented in this paper characterizes this interdependence for the case of symmetric error channels.

2. PRELIMINARIES

The system model considered in this paper is as follows. Let $\mathbf{x} \in \mathcal{D}_{\mathbf{x}} \subset \mathbb{R}^n$ be a random source with pdf $f_{\mathbf{x}}(\mathbf{x})$, where $\mathcal{D}_{\mathbf{x}}$ is the domain of $\mathbf{x}$. The non-negative, twice continuously differentiable function $d(\mathbf{x}, \hat{\mathbf{x}}_i)$ measures the distortion resulting from representing $\mathbf{x}$ as $\hat{\mathbf{x}}_i$. A VQ encoder is described by N partition regions $\mathcal{R}_i, 1 \leq i \leq N$ that tile $\mathcal{D}_{\mathbf{x}}$. Associated with each partition region $\mathcal{R}_i$ is a code-vector $\hat{\mathbf{x}}_i$. In the case of the Lloyd centroid quantizers [10] considered in this paper, $\hat{\mathbf{x}}_i$ is the centroid of the random vector $\mathbf{x}$ conditioned on $\mathbf{x} \in \mathcal{R}_i$ under the distortion measure of interest; and $\mathcal{R}_i = \{\mathbf{x} : d(\mathbf{x}, \hat{\mathbf{x}}_i) \leq d(\mathbf{x}, \hat{\mathbf{x}}_j), 1 \leq j \leq N\}$. Note that the definitions of $\hat{\mathbf{x}}_i$ and $\mathcal{R}_i$ depend on each other, which means that, in practice, iterative algorithms are employed to generate a locally optimal codebook and a corresponding set of quantization regions.

Whenever $\mathbf{x} \in \mathcal{R}_i$, the quantizer outputs index i, which is mapped to a point in some constellation and sent over a noisy channel. At the receiver, a possibly different index j is received with probability $P_{j|i}$, and the receiver outputs $\hat{\mathbf{x}}_j$.

2.1. Symmetric Error Channels

In this paper, the channel is modelled as a symmetric error channel, i.e., a channel where all index errors are equally likely. This channel model leads to tractable, closed-form expressions for the overall performance, thus enabling one to compare (and trade off) the distortion due to the source quantization and the distortion due to channel errors. In case of the symmetric error channel, the *transition probability* $P_{j|i}$, which is the probability that the transmitted index i is received as index j, is given by

$$P_{j|i} = \left\{ \begin{array}{cc} a(N), & j \neq i \\ 1 - (N-1)a(N), & j = i \end{array} \right. , \qquad (1)$$

where $0 \leq a(N) \leq 1/(N-1)$ is the probability that index i is received as a different index j. From a practical perspective, it is also reasonable to expect $a(N) < \frac{1}{N}$, which ensures that $P_{i|i} = \max_{1 \leq j \leq N} P_{j|i}$ holds.

Note that under this set-up, $P_{i|i} = 1-(N-1)a(N)$ is the probability of correct reception. This implicitly assumes that as N is increased, more (or less) energy is used to transmit the symbol in order to maintain the probability of correct reception $P_{i|i}$. An example for the symmetric error channel is when the index is sent using orthogonal modulation over an AWGN channel. Another example is when the channel is a binary symmetric channel (BSC) with bit cross-over probability q and the assignment of the indices to the $B = \log_2(N)$ bit words is *random*. It can be shown that after averaging over all possible index assignments, the probability of correct reception is $P_{i|i} = (1-q)^B$, and thus $a(N) = \left(1-(1-q)^B\right)/(N-1)$.

3. PERFORMANCE ANALYSIS

In this section, the high-rate performance of VQ for the case of symmetric-error channels is stated; the proof is relegated to [12] as it is rather lengthy. The expected distortion is given by

$$E_d = \sum_{i,j=1}^{N} P_{j|i} \int_{\mathbf{x} \in \mathcal{R}_i} d(\mathbf{x}, \hat{\mathbf{x}}_j) f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \tag{2}$$

For the result to follow, the standard high-rate approximations in [8, 10], and the quantization cell approximation in [11] are employed. The high-rate performance expression is in terms of the continuous *point density function* $\lambda(\mathbf{x})$, which is defined such that $\lambda(\mathbf{x})\Delta\mathbf{x}$ is approximately the fraction of codepoints in a small region $\Delta\mathbf{x}$ around $\mathbf{x}$. For an optimal VQ, the quantization cell $\mathcal{R}_i$ is well approximated by a corresponding n-dimensional hyper-ellipsoid with the volume $(N\lambda(\hat{\mathbf{x}}_i))^{-1}$, and whose shape is determined by the Hessian of the distortion function at $\hat{\mathbf{x}}_i$. Under these approximations, it is possible to show that

$$E_d = \int_{\mathbf{x}} E_{d,\mathbf{x}} f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \tag{3}$$

where, $E_{d,\mathbf{x}}$, the expected distortion conditioned on the source instantiation $\mathbf{x}$, is given by

$$\begin{aligned} E_{d,\mathbf{x}} \approx Na(N) \Bigg\{ & \int_{\mathbf{y}} d(\mathbf{x},\mathbf{y}) \lambda(\mathbf{y}) d\mathbf{y} \\ & + \frac{N^{\frac{-2}{n}} \kappa_n^{\frac{-2}{n}}}{2(n+2)} \lambda^{\frac{-2}{n}}(\mathbf{x}) \left|D(\mathbf{x},\mathbf{x})\right|^{\frac{1}{n}} \\ & \cdot \operatorname{tr}\left(D^{-1}(\mathbf{x},\mathbf{x}) \left[\int_{\mathbf{y}} \left(D(\mathbf{x},\mathbf{y}) - D(\mathbf{x},\mathbf{x})\right) \lambda(\mathbf{y}) d\mathbf{y} \right] \right) \Bigg\} \\ & + \frac{nN^{\frac{-2}{n}} \kappa_n^{\frac{-2}{n}}}{2(n+2)} \lambda^{\frac{-2}{n}}(\mathbf{x}) \left|D(\mathbf{x},\mathbf{x})\right|^{\frac{1}{n}}, \end{aligned} \tag{4}$$

where κ_n is the volume of an n-dimensional unit sphere, and $D(\mathbf{x},\mathbf{y})$ is an $n \times n$ sensitivity matrix with j,k-th element $D_{k,j}(\mathbf{x},\mathbf{y}) = \left.\frac{\partial^2 d(\hat{\mathbf{x}},\mathbf{y})}{\partial \hat{x}_j \partial \hat{x}_k}\right|_{\hat{\mathbf{x}}=\mathbf{x}}$. Note that the last term is the asymptotic distortion in the absence of channel errors (i.e., when $a(N) = 0$). Also, the first term inside the curly braces represents the distortion caused by the channel errors only, and the second term characterizes interdependence of the channel errors and the source quantization errors. Thus, unless $\int_{\mathbf{y}} \left(D(\mathbf{x},\mathbf{y}) - D(\mathbf{x},\mathbf{x})\right) \lambda(\mathbf{y}) d\mathbf{y} = 0$, the distortion cannot be split as the sum of the distortion due to source encoding and the distortion due to channel errors.

3.1. Mean-Squared Error Distortion

Consider the special case where $d(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|^2$ is the mean-squared error function. Then, $D(\mathbf{x},\mathbf{y}) = 2I_n$ regardless of $\mathbf{x}$ and $\mathbf{y}$, thus, the second term in (4) vanishes. Let $(m_{\mathbf{x}}, \sigma_{\mathbf{x}}^2)$ and $(m_{\mathbf{y}}, \sigma_{\mathbf{y}}^2)$ be the mean vector and the trace of the covariance matrix of a random vector with probability density $f_{\mathbf{x}}(\mathbf{x})$ and $\lambda(\mathbf{x})$ respectively. When $d(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|^2$, it is easy to show that

$$\int_{\mathbf{x}} d(\mathbf{x},\mathbf{y}) f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} = \sigma_{\mathbf{x}}^2 + \|\mathbf{y} - m_{\mathbf{x}}\|^2. \tag{5}$$

A similar expression can be obtained by replacing $f_{\mathbf{x}}(\mathbf{x})$ in the above expression by $\lambda(\mathbf{x})$ and making corresponding changes in the right hand side. Substituting in (4) and simplifying,

$$\begin{aligned} E_d &\approx Na(N)\left(\sigma_{\mathbf{x}}^2 + \sigma_{\mathbf{y}}^2 + \|m_{\mathbf{y}} - m_{\mathbf{x}}\|^2\right) \\ &+ \frac{nN^{\frac{-2}{n}} \kappa_n^{\frac{-2}{n}}}{n+2} \int_{\mathbf{x}} \lambda^{\frac{-2}{n}}(\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}. \end{aligned} \tag{6}$$

Thus, with MSE as the performance metric, the overall distortion splits as two terms: the first term measures the distortion arising purely due to channel errors, and the second term measures the distortion due to the source encoding. This is in agreement with the result in [13], where the authors proved, for a scalar source, that the distortion can be separated into two terms as above.

3.2. Conventional Source Coding with Channel Errors

In this subsection, the sensitivity of conventional source coding to errors caused by the symmetric-error channel is analyzed. Conventionally (i.e., in the absence of channel errors), the point density function $\lambda(\mathbf{x})$ is chosen to minimize the last term in (6), i.e.,

$$\begin{aligned} \lambda_{\text{conv}}(\mathbf{x}) &= \arg\min_{\lambda(x)} \frac{nN^{\frac{-2}{n}} \kappa_n^{\frac{-2}{n}}}{n+2} \int_{\mathbf{x}} \lambda^{\frac{-2}{n}}(\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \\ &= \frac{f_{\mathbf{x}}^{\frac{n}{n+2}}(\mathbf{x})}{\int_{\mathbf{x}} f_{\mathbf{x}}^{\frac{n}{n+2}}(\mathbf{x}) d\mathbf{x}} \end{aligned} \tag{7}$$

The last equation is obtained by applying Hölder's inequality the optimization problem, and can be found in many sources, e.g., [9].

With the above point density, the expected MSE distortion can be obtained by substituting for $\lambda_{\text{conv}}(\mathbf{x})$ in (6) as

$$\begin{aligned} E_d &\approx Na(N)\left(\sigma_{\mathbf{x}}^2 + \sigma_{\mathbf{y}}^2 + \|m_{\mathbf{y}} - m_{\mathbf{x}}\|^2\right) \\ &+ \frac{nN^{\frac{-2}{n}} \kappa_n^{\frac{-2}{n}}}{n+2} \left(\int_{\mathbf{x}} f_{\mathbf{x}}^{\frac{n}{n+2}}(\mathbf{x})\, d\mathbf{x} \right)^{\frac{n+2}{n}}. \end{aligned} \tag{8}$$

It is interesting to note that, as N increases, the distortion is dominated by the behavior of $Na(N)$ relative to $N^{\frac{-2}{n}}$. That is,

1. If $Na(N) = o\left(N^{\frac{-2}{n}}\right)$, channel errors play an insignificant role in the asymptotic distortion.
2. If $N^{\frac{-2}{n}} = o\left(Na(N)\right)$, the error is dominated by the channel errors (i.e., the first term above). For example, when the index is sent over a BSC with bit error probability q, for large N, $Na(N) \to 1$, and hence the error approaches the non-zero constant $E_d \approx \left(\sigma_{\mathbf{x}}^2 + \sigma_{\mathbf{y}}^2 + \|m_{\mathbf{y}} - m_{\mathbf{x}}\|^2\right)$. Thus, the source and point densities only appear in the asymptotic distortion through its means and variances.

3. If $Na(N) = \Theta\left(N^{\frac{-2}{n}}\right)$, then both terms decrease at the same rate as N increases, so neither component dominates at high rates.

As an example, consider the case where $\mathbf{x}$ is an n-dimensional i.i.d. uniformly distributed random vector with each entry uniformly distributed on $[-0.5, 0.5)$. Then, $m_{\mathbf{x}} = \underline{\mathbf{0}}$ where $\underline{\mathbf{0}}$ is a vector of zeroes, and $\sigma_{\mathbf{x}}^2 = n/12$. Also, it can be verified that $\lambda_{\text{conv}}(\mathbf{x})$ is an n-dimensional i.i.d. uniformly distributed density with each entry uniformly distributed on $[-0.5, 0.5)$. Then,

$$E_d \approx \frac{nNa(N)}{6} + \frac{nN^{\frac{-2}{n}}\kappa_n^{\frac{-2}{n}}}{n+2} \tag{9}$$

When $n = 1$ (scalar quantization) and the channel is a BSC, the above expression agrees with an asymptotic result derived in [7].

4. SIMULATION RESULTS

For simplicity, first consider a source $\mathbf{x}$ that is i.i.d. and uniformly distributed on $[-0.5, 0.5)$, with mean squared error as the performance metric. Assume that the channel is a BSC with bit cross-over probability q and that the index assignment is random. To verify the sensitivity of the performance of VQ to channel errors, the conventionally optimized codebook is generated using the Lloyd algorithm [10] for different values of n (number of dimensions) and B in the absence of channel errors. For training the Lloyd algorithm, as well as for evaluating the performance, 10,000 independent random instantiations of $\mathbf{x}$ were employed.

Figure 1 shows the MSE distortion versus the number of quantized bits B for the uniformly distributed random vector with dimension $n = 1, 2, 3$ and 4. The theoretical distortion is given by (9), with $a(N) = \left(1 - (1-q)^B\right)/(N-1)$, and bit transition probability $q = 10^{-3}$. From the bottom curve ($n = 1$), notice that as B is increased, the overall distortion initially decreases due to better source quantization, and later starts to increase again as the effect of channel errors increases. Asymptotically, the distortion approaches the fixed constant $\sigma_{\mathbf{x}}^2 + \sigma_{\mathbf{y}}^2 = 1/6$. Figure 2 shows the MSE distortion versus q, with $B = 5$, or $N = 32$ quantization levels. For small values of q, the distortion with $N = 32$ is dominated by the distortion caused by quantization errors (i.e., the second term in (9)), whereas, as q increases, the first term gets larger and the distortion finally approaches $n(1 - 0.5^B)/6$. Figure 3 shows the distortion caused by channel errors only, denoted $E_d^{(1)}$, which from (9), is given by

$$E_d^{(1)} \approx \frac{nN\left(1 - (1-q)^B\right)}{6\,(N-1)}. \tag{10}$$

The above equation gives the rate at which $E_d^{(1)}$ approaches $n/6$ for a given q, as B increases. It is also instructive to observe the number of bits B at which the high-rate approximations become accurate. A good rule-of-thumb in source coding is that the high-rate results apply at about 2-3 bits per dimension when the channel is error-free. This is seen, for example, by examining the error-free curves plotted in figure 1. With channel errors, however, both the $E_d^{(1)}$ and the $E_d^{(3)}$ terms must converge to their theoretical values as B increases. From figure 3, a good rule-of-thumb for the convergence of the $E_d^{(1)}$ term is about 3-4 bits per dimension, slightly higher than the corresponding number for the $E_d^{(3)}$ term.

Next, an experiment is performed on wideband speech spectrum coding, under the Log Spectral Distortion (LSD) measure. The goal

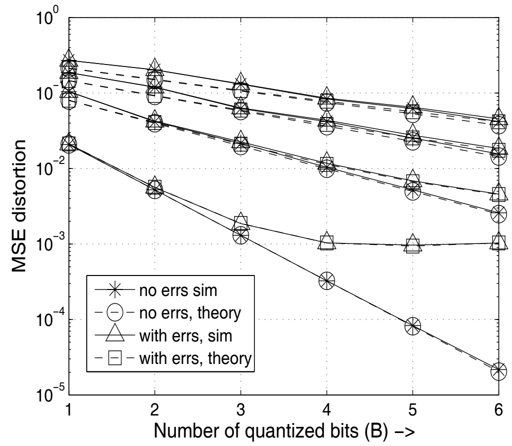

Fig. 1. MSE distortion versus number of quantized bits B, for a uniformly distributed random vector and index sent over the BSC with $q = 10^{-3}$. The curves correspond to $n = 1, 2, 3, 4$, from bottom to top.

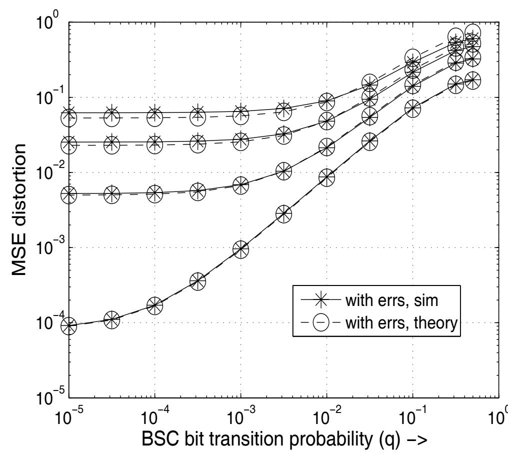

Fig. 2. MSE distortion for a uniformly distributed random vector with the conventional point density and $B = 5$. The index is sent over a BSC with bit transition probability q (the x-axis). The curves correspond to $n = 1, 2, 3, 4$ from bottom to top.

in this context is to achieve an average distortion of 1dB^2. This experiment is designed to determine at what error rates this goal is feasible. A database of 16-dimensional wideband speech spectrum vectors is gathered, and their sensitivities are evaluated using the method described in [11]. For sources with such a large dimension, the codebook sizes become very large (around 50 bits must be used). This rules out the use of full-search quantizers, and structured systems must instead be used to reduce the complexity. To this end, the Gaussian Mixture Model (GMM) based VQ described in [14] is employed. This system is able to operate with a low, rate-independent complexity, although it is suboptimal in the sense that its cells are not ellipsoidal; and as a result there is a small gap between the theoretical and experimental distortion curves.

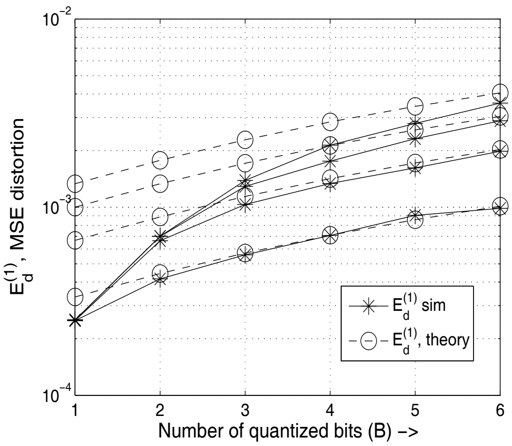

Fig. 3. The MSE distortion term $E_d^{(1)}$ for a uniformly distributed random vector versus B with $q = 10^{-3}$. The curves correspond to $n = 1, 2, 3, 4$ from bottom to top, and the theoretical curves are generated using (10).

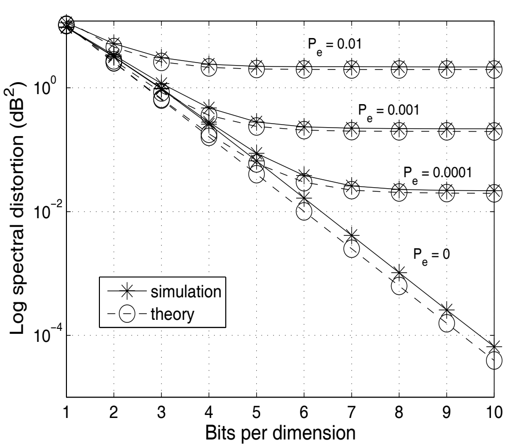

Fig. 4. Log Spectral Distortion on Wideband Speech LSF vectors versus B. Both predicted and actual distortions are shown for several values of P_e, the total probability of an index error.

Since the source density is unknown, the expectations in (3) and (4) must be evaluated via a Monte-Carlo method. The point density of the quantizer is itself a GMM, with parameters specified through the training process. Thus, the integral in (4) over $\mathbf{y}$ can be approximated by averages over data $\mathbf{y}$ drawn randomly according to the point density. A database of 65000 source vectors $\mathbf{x}_i$ is employed, along with a database of 65000 "error" vectors $\mathbf{y}_i$, drawn according to $\lambda(\mathbf{y})$ for the Monte-Carlo estimation. From figure 4, it can be seen that the theory is good at predicting the true high-rate distortion. Observe that for high values of the error probability, it is impossible to perform high-quality quantization of this source, with the error levelling off at around 3dB^2. For moderate error probabilities, 1dB^2 LSD can be achieved, although a few extra bits will be required to compensate for channel errors. At low values of the error probability, there is no penalty, as the channel effects do not become significant until well beyond the desired 1dB^2 operating point. Thus, a channel error probability between 0.001 and 0.0001 is judged to be permissible for the wideband speech spectrum quantization problem.

In conclusion, this paper considered the source quantization problem when the index is sent over a noisy channel before being used to reproduce the source at the receiver. For the special case of the symmetric-error channel, the asymptotic performance with arbitrary distortion functions was theoretically analyzed. Further, when the distortion is measured as the MSE, it was demonstrated that the distortion is given by the sum of two terms, the first representing the distortion arising purely due to channel errors and the second being the error in quantizing the source using a finite number of bits. The accuracy of the theoretical results were illustrated through Monte-Carlo simulation. Finally, the framework was applied to the problem of wideband speech spectrum quantization under the LSD measure, and seen to accurately characterize the effects of the source, the quantizer and the channel.

5. ACKNOWLEDGEMENT

The authors thank Jun Zheng for several insightful discussions and helpful comments.

6. REFERENCES

[1] N. Farvardin, "A study of vector quantization for noisy channels," *IEEE Trans. Info. Theory*, vol. 36, no. 4, pp. 799 – 809, July 1990.

[2] A. Gersho and K. A. Zeger, "Vector quantizer design for memoryless noisy channels," in *Proc. IEEE Int. Conf. on Commun. (ICC) 1988*, Philadelphia, PA, USA, June 1998, pp. 1593–1597.

[3] K. Zeger and V. Manzella, "Asymptotic bounds on optimal noisy channel quantization via random coding," *IEEE Trans. Info. Theory*, vol. 40, no. 6, pp. 1926 – 1938, Nov. 1994.

[4] J. R. B. DeMarca and N. S. Jayant, "An algorithm for assigning binary indices to the codevectors of multidimensional quantizers," in *Proc. IEEE Int. Conf. on Commun. (ICC) 1987*, June 1987, pp. 1128–1132.

[5] D. J. Goodman and T. J. Moulsley, "Using simulated annealing to design transmission codes for analogue sources," *Electronics Letters*, vol. 24, pp. 617 – 618, May 1988.

[6] B. Hochwald, "Tradeoff between source and channel coding on a gaussian channel," *IEEE Trans. Info. Theory*, vol. 44, no. 7, pp. 3044 – 3055, Nov. 1998.

[7] G. Ben-David and D. Malah, "Bounds on the performance of vector-quantizers under channel errors," *IEEE Trans. Info. Theory*, vol. 51, no. 6, pp. 2227 – 2235, June 2005.

[8] W. R. Bennett, "Spectra of quantized signals," *Bell Syst. Tech. Journal*, vol. 27, pp. 446 – 472, July 1948.

[9] A Gersho, "Asymptotically optimum block quantization," *IEEE Trans. Info. Theory*, vol. IT-25, no. 4, pp. 373–380, July 1979.

[10] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Info. Theory*, vol. IT-28, no. 2, pp. 129–137, Mar. 1982.

[11] W. R. Gardner and B. D. Rao, "Theoretical analysis of high-rate vector quantization of LPC parameters," *IEEE Trans. on Speech and Audio Proc.*, vol. 3, no. 5, pp. 367–381, Sept. 1995.

[12] C. R. Murthy and B. D. Rao, "Performance of vector quantizers with channel errors," *in preparation*, 2005.

[13] R. Totty and G. Clark, "Reconstruction error in waveform transmission," *IEEE Trans. Info. Theory*, vol. 13, no. 2, pp. 336 – 338, Apr. 1967.

[14] A. D. Subramaniam and B. D. Rao, "PDF Optimized Parametric Vector Quantization of Speech Line Spectral Frequencies" IEEE Trans. on Speech and Audio Proc., Vol. 11, (no.2), March 2003.