HIGH-RATE ANALYSIS OF CHANNEL-OPTIMIZED VECTOR QUANTIZATION

Chandra R. Murthy[1] *and Bhaskar D. Rao*[2]

[1]Beceem Communications
Embassy Star, 4th Floor, No. 8
Vasanth Nagar, Bangalore 560 052, India
cmurthy1@gmail.com

[2]Department of Electrical and Computer Engineering
University of California San Diego
La Jolla, CA 92093-0407, USA
brao@ucsd.edu

ABSTRACT

This paper considers the high-rate performance of channel optimized source coding for noisy discrete symmetric channels with random index assignment. Specifically, with Mean Squared Error (MSE) as the performance metric, an upper bound on the asymptotic (i.e., high-rate) distortion is derived by assuming a general structure on the codebook. This structure enables extension of the analysis of the channel optimized source quantizer to one with a *singular* point density: for channels with small errors, the point density that minimizes the upper bound is continuous, while as the error rate increases, the point density becomes singular. The extent of the singularity is also characterized. The accuracy of the expressions obtained are verified through Monte Carlo simulations.

Keywords: source coding, compression.

1. INTRODUCTION

It is well known that the performance of a source quantization scheme can be very sensitive to errors introduced when the codepoint index is transmitted over a noisy channel. 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 to practical communication.

The effect of channel errors on VQ can be modeled as an *index error*, that is, the index i corresponding to the current source instantiation is received as a possibly different index j. Classical source coding has devoted much effort to the problem of source compression for noisy channels, and two dominant approaches have emerged. The first is channel-optimized VQ, i.e., to replace the distortion measure used for optimizing the quantizer with the expected distortion over the noisy channel (e.g., [1]). The second approach involves *Index Assignment* (IA), i.e., to design the quantizer without considering channel errors and then map codewords to indices in such a way that codewords resulting in small inter-codepoint distortions are mapped to index pairs that correspond to channel symbols with large transition probability and vice versa (e.g., [2]).

In this sequel to [3], a channel optimized approach based on classical results from the source coding literature [4] - [7] is adopted to derive new results for the performance of source coding with Mean Squared Error (MSE) distortion for discrete symmetric channels with random IA. Random indexing results in a Simplex Error Channel (SEC), i.e., a channel for which the probability of receiving an index j when an index i is sent only depends on whether or not i and j are different. To analyze the performance, a general structure is assumed for the codebook of quantized vectors, which leads to an upper bound on the expected distortion with a channel-optimized codebook. It is assumed that some fraction of the codepoints are merged into the source centroid, while the remaining codepoints are distinct, as a way of building a joint source-channel code. Optimization of the extent of the singularity is also considered; and to the best of the authors' knowledge, this is the first time that the performance of the source coding scheme with channel errors has been extended to the class of *singular* point densities. Interesting results and insights are obtained on the design and performance of channel-optimized codebooks for a wide range of channel error rates. The hope is that the results presented here would eventually be extended to obtain insights into the design and analysis of joint source-channel codes for many other scenarios.

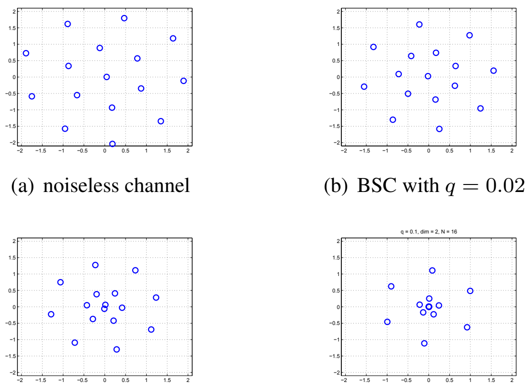
This work was supported in part by UC Discovery grant no. Com 02-10105 sponsored by Conexant Systems Inc.

2. PRELIMINARIES

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 $\mathbf{x}$. Without loss of generality, let the mean of the source be at the origin, denoted $\mathbf{0}$. 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 $\hat{\mathbf{x}}_i$. Whenever $\mathbf{x} \in \mathcal{R}_i$, the quantizer outputs index i, which is sent over a noisy channel. At the receiver, the index is received as an index j with probability $P_{j|i}$, upon which it outputs $\hat{\mathbf{x}}_j$ as its estimate of $\mathbf{x}$. The distortion resulting from representing $\mathbf{x}$ as $\hat{\mathbf{x}}$ is measured as $d(\mathbf{x}, \hat{\mathbf{x}}) \triangleq \|\mathbf{x} - \hat{\mathbf{x}}\|^2$.

It is well known that for any source, the Channel Optimized Vector Quantizer (COVQ) [1] satisfies two conditions: the *weighted nearest neighbor condition*, and the *weighted centroid condition*. Given the code book and a source instantiation $\mathbf{x}$, the optimum quantizer chooses the index i that minimizes the expected distortion after accounting for possible channel errors. That is, denoting the expected distortion after the index is transmitted over the noisy channel as $d_c(\mathbf{x}, \hat{\mathbf{y}}_i) \triangleq \sum_{k=1}^{N} d(\mathbf{x}, \hat{\mathbf{y}}_k) P_{k|i}$, the transmit optimized quantizer selects the index i that minimizes $d_c(\mathbf{x}, \hat{\mathbf{x}}_i)$. Likewise, given the quantization regions, the weighted centroid condition computes $\hat{\mathbf{x}}_i$ as a weighted sum of the geometrical centroids of all the quantization regions, with the weights determined by the a-posteriori channel transition probabilities. It can be shown that the weighted centroid condition yields the optimum code-point for any given a quantization region in the sense of minimizing the expected distortion given a particular received index, after accounting for possible channel errors. The COVQ thus simultaneously satisfies the weighted nearest neighbor condition and the weighted centroid condition.

In this paper, the channel is modeled as a discrete symmetric



(a) noiseless channel

(b) BSC with $q = 0.02$

(c) BSC with $q = 0.05$

(d) BSC with $q = 0.1$

Fig. 1. VQ codepoints for $N = 16$ level quantization of an $n = 2$ dimensional i.i.d. standard Gaussian source. The codebooks were generated using the generalized Lloyd algorithm in [1].

channel [8] with *random* IA, which leads to an SEC with transition probability given by

$$P_{j|i} = \epsilon(N) + (1 - N\epsilon(N))\,\delta(i, j) \tag{1}$$

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

An example for the SEC 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 $\epsilon(N) = \left(1 - (1-q)^B\right) / (N-1)$. Detailed justification and motivation for the random IA assumption can be found in [3]. Note that, in practical implementation, the effect of random IA can be achieved by employing different IAs (in a pre-specified pseudo-random pattern) over time.

Fig. 1 shows the channel-optimized codepoints obtained using the generalized Lloyd algorithm described in [1], for 16-level quantization of a 2-dimensional i.i.d. Gaussian source, when the index is mapped to a 4-bit symbol and sent over a Binary Symmetric Channel (BSC) with bit cross-over probability q. From the figure, it can be seen that as the channel deteriorates, more and more codepoints get merged at the origin, until, for a completely degenerate channel (i.e., one for which $P_{j|i} = 1/N$ for all i and j), the codebook contains just one distinct codepoint at the source centroid. This motivates the structure assumed for the Channel Optimized Vector Quantizer (COVQ) codebook in the sequel. In classical source coding, when the channel is noiseless, it is known that the distribution of codepoints often approximates a *continuous point density*. However, when the channel has errors, the codepoints of the optimum codebook initially shrink closer towards the centroid of the source distribution, and eventually, some of the codepoints collapse together and the point density becomes *singular*. This singular point density can be thought of as the sum of a continuous point density and one or more singular points (Dirac delta function).

Thus, of the total N code points, αN *distinct* points are drawn from a *continuous* density (it is assumed that αN is large), while the remaining $(1-\alpha)N$ points are at the origin, where $0 \leq \alpha \leq \frac{N-1}{N}$. Note that α can itself be a function of N and the channel transition probability, which then corresponds to tuning the quantizer to the channel at each specific N. Under this assumption, the code book can be represented as having $\alpha N + 1$ points $\{\hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_{\alpha N}, \mathbf{0}\}$, and the equivalent index transition probability matrix $P_{j|i}$ is given by the $(N\alpha + 1) \times (N\alpha + 1)$ matrix

$$P = \begin{bmatrix} \tilde{\epsilon} & \epsilon & \ldots & \epsilon & N(1-\alpha)\epsilon \\ \epsilon & \tilde{\epsilon} & \ddots & & N(1-\alpha)\epsilon \\ \vdots & & & \tilde{\epsilon} & \vdots \\ \epsilon & \ldots & & \epsilon & 1 - \alpha N\epsilon \end{bmatrix}, \tag{2}$$

where $\epsilon = \epsilon(N)$ and $\tilde{\epsilon} = 1 - (N-1)\epsilon(N)$. The above equivalent index transition probability is no longer simplex, because the point $\mathbf{0}$ actually consists of $N(1-\alpha)$ codepoints, and when the source $\mathbf{x}$ lies in the quantization region for the point $\mathbf{0}$ (denoted $\mathcal{R}_{\mathbf{0}}$), one of the $N(1-\alpha)$ indices corresponding to $\mathbf{0}$ is randomly picked and sent across the channel.

3. PERFORMANCE ANALYSIS

In this section, the high-rate performance of VQ for an SEC is stated. Due to the lengthy nature of the proof and space limitations, it is omitted and the reader is referred to [9, 10] for details. The expected distortion is obtained by taking a double expectation over the source density and the channel transition probabilities as follows:

$$E_d = \sum_{i,j=1}^{N} P_{j|i} \int_{\mathbf{x} \in \mathcal{R}_i} \|\mathbf{x} - \hat{\mathbf{x}}_j\|^2 f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}. \tag{3}$$

For the result to follow, the standard high-rate approximations in [4, 6], and the quantization cell approximation in [7] are employed. Then, the expected distortion can be shown to be given by

$$E_d = \int_{\mathbf{x}} E_{d,\mathbf{x}} f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \tag{4}$$

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

$$\begin{aligned} E_d \quad \approx \quad & N\epsilon(N)\sigma_{\mathbf{x}}^2 + \varphi\epsilon(N) \int \mathbf{y}^T \mathbf{y} \lambda(\mathbf{y}) d\mathbf{y} \\ & + g_n \varphi^{-\frac{2}{n}} \int \lambda^{-\frac{2}{n}}(\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}. \end{aligned} \tag{5}$$

In the above, $\varphi \triangleq \alpha N + 1$ (note that $1 \leq \varphi \leq N$), $g_n \triangleq n\kappa_n^{-\frac{2}{n}}/(n+2)$, κ_n is the volume of an n-dimensional unit sphere, and $\sigma_{\mathbf{x}}^2$ is the source variance. Also, $\lambda(\mathbf{x})$ is the *fractional point density*, defined as follows. The *specific point density* [5] is given by

$$\lambda_\varphi(\mathbf{x}) \triangleq \frac{1}{\varphi V(\mathcal{R}_i)}, \text{if } \mathbf{x} \in \mathcal{R}_i, \text{ for } i = 1, 2, \ldots, \varphi; \tag{6}$$

where $V(\mathcal{R}_i)$ is the volume of the region $\mathcal{R}_i$. Then, from [6], when φ is large, $\lambda_\varphi(\mathbf{x})$ asymptotically approximates a continuous non-negative density function $\lambda(\mathbf{x})$ having a unit integral. When $n = 1$ (scalar quantization) and $\varphi = N$ (continuous point density), the above expression reduces to similar expressions in [11, 12].

The high rate distortion expression in (5) is valid for any point density as long as the regions $\mathcal{R}_i$ get small as N gets large and the quantizer satisfies the nearest neighbor condition, i.e., it is not restricted to the optimum point density. It can be seen that the overall distortion splits as the sum of the channel error-induced distortion (the first two terms) and the source quantization-induced distortion (the last term). The last term in (5) is the high-rate distortion result for an noiseless channel (i.e., when $\epsilon(N) = 0$), and is therefore minimized when $\varphi = N$ and $\lambda(\mathbf{x}) = \lambda_{\text{conv}}(\mathbf{x})$, the conventional point density given in [5]. Also, when $\varphi = N$, if $\epsilon(N)$ is proportional to $N^{-(n+2)/n}$, the terms are *balanced*, since the last term varies with N as $N^{-2/n}$. Note that the first term is independent of both the point density and α, and thus the expected distortion with the given codebook structure is lower-bounded by $N\epsilon(N)\sigma_{\mathbf{x}}^2$. For a noisy channel, it may be possible to reduce the overall distortion by choosing a point density that has a smaller second moment than the conventional point density, as that would lead to a reduction in the second-moment term above. The above expression also shows the effect of choosing different values of $0 \leq \alpha \leq (N-1)/N$. Clearly, when $\epsilon(N) = 0$, $\alpha = (N-1)/N$ is optimum, which corresponds to employing a point density with no singularity at the origin. As the channel gets worse, $\epsilon(N)$ gets larger, and therefore the second term starts dominating the performance. In this case, one must employ a smaller value of α to balance the second and third terms.

While the above qualitative arguments provide an intuitive feel for the interplay between the different terms, the following analysis shows how to optimize $\lambda(\mathbf{x})$ as well as α. For the purposes of optimization, when N is large, φ can be considered to be a continuous variable, and it is reasonable to expect that the optimum discrete value of φ would be one of the nearby integers. Thus, taking the partial derivative with respect to φ and equating to zero yields

$$\varphi_{\text{opt}} = \left[\frac{2g_n \int \lambda^{-\frac{2}{n}}(\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}}{n\epsilon(N)\sigma_{\mathbf{y}}^2} \right]^{\frac{n}{n+2}} \tag{7}$$

It is easily verified the second partial derivative is positive, i.e., φ_{opt} given above is indeed a local minimizer. Note that the above equation is valid provided $1 \leq \varphi_{\text{opt}} \leq N$. Otherwise, the optimum value of φ is one of the end-points, either 0 or N. When φ_{opt} is given by (7), the expected distortion becomes

$$\begin{aligned} E_d \quad \approx \quad & \frac{n+2}{2} \left(\frac{2g_n}{n} \right)^{\frac{n}{n+2}} \epsilon^{\frac{2}{n+2}}(N) \left(\int \mathbf{y}^T \mathbf{y} \lambda(\mathbf{y}) d\mathbf{y} \right)^{\frac{2}{n+2}} \\ & \cdot \left(\int \lambda^{-\frac{2}{n}}(\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \right)^{\frac{n}{n+2}} + N\epsilon(N)\sigma_{\mathbf{x}}^2 \end{aligned} \tag{8}$$

In the above expression, the point density that minimizes the second moment (the first integral) is a delta-function, whereas the point density that minimizes the second integral is the conventional point density, and the optimum point density must find the right trade-off between the two. In addition, *provided φ_{opt} in* (7) *satisfies* $1 \leq \varphi_{opt} \leq N$*, the optimum point density is independent of N and $\epsilon(N)$ for large N*, since they appear only as multiplying factors. Now, finding the point density that minimizes (8) directly is hard, so an indirect method is adopted here. Using the calculus of variations, the point density $\lambda(\mathbf{x})$ that minimizes (5) subject to the constraints (positive, and integrates to unity) can be shown to be given by

$$\lambda_{\text{opt}}(\mathbf{x}) = \left[\frac{2g_n f_{\mathbf{x}}(\mathbf{x})}{n \left(\varphi^{\frac{n+2}{n}} \epsilon(N) \mathbf{x}^T \mathbf{x} + \mu \varphi^{\frac{2}{n}} \right)} \right]^{\frac{n}{n+2}} \tag{9}$$

where the normalization constant μ is chosen such that $\lambda_{\text{opt}}(\mathbf{x})$ integrates to 1. Again, it is easily verified that since the second partial derivative is positive, the above $\lambda_{\text{opt}}(\mathbf{x})$ is indeed a local minimizer.

Now, it was seen above that provided $1 \leq \varphi_{\text{opt}} \leq N$, the optimum point density is a function independent of N and $\epsilon(N)$. This is possible only if the value of φ varies with N such that

$$\frac{n\varphi^{\frac{n+2}{n}}\epsilon(N)}{2g_n} = K, \tag{10}$$

where K is independent of N, which implies that the optimum point density $\lambda_{\text{opt}}(\mathbf{x})$ is given by

$$\lambda_{\text{opt}}(\mathbf{x}) = \left[\frac{f_{\mathbf{x}}(\mathbf{x})}{K\mathbf{x}^T\mathbf{x} + M} \right]^{\frac{n}{n+2}}, \tag{11}$$

with M being a normalization constant.

The following lemmas establish critical values of the error probability, and then state that the value of M in the optimum point density above is in fact $M = 0$. The proofs can be found in [9, 10]; they are omitted here due to lack of space.

Lemma 1 *Define $\epsilon_{crit,1}$ as*

$$\epsilon_{crit,1} \triangleq N^{-\frac{n+2}{n}} \left[\frac{2g_n \int \lambda_{conv}^{-\frac{2}{n}}(\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}}{n \int \mathbf{x}^T \mathbf{x} \lambda_{conv}(\mathbf{x}) d\mathbf{x}} \right], \tag{12}$$

where $\lambda_{conv}(\mathbf{x})$ is the optimum point density for a noiseless channel (given in [5]). Then, if the index error probability satisfies $0 \leq \epsilon(N) \leq \epsilon_{crit,1}$, $\lambda_{opt}(\mathbf{x})$ is continuous*, and is given by* (9) *with $\varphi = N$.*

Lemma 2 *Define $\epsilon_{crit,2}$ as*

$$\epsilon_{crit,2} = N^{-\frac{n+2}{n}} \left[\int \left(\frac{2g_n f_{\mathbf{x}}(\mathbf{x})}{n\,\mathbf{x}^T\mathbf{x}} \right)^{\frac{n}{n+2}} d\mathbf{x} \right]^{\frac{n+2}{n}}. \tag{13}$$

Then, for $\epsilon_{crit,2} < \epsilon(N) \leq 1/N$, the optimum point density is singular and the optimum $\lambda(\mathbf{x})$ and φ are given by

$$\lambda_{opt}(\mathbf{x}) = \left(\frac{f_{\mathbf{x}}(\mathbf{x})}{K\mathbf{x}^T\mathbf{x}} \right)^{\frac{n}{n+2}}, \qquad \varphi_{opt} = \left[\frac{2Kg_n}{n\epsilon(N)} \right]^{\frac{n}{n+2}}, \tag{14}$$

where K is a normalization constant. Also, $\epsilon_{crit,2}$ is the largest $\epsilon(N)$ for which no codepoints are merged, i.e., $\varphi_{opt} = N$ is satisfied.

It can be verified that the above value of φ_{opt} and optimum point density in (14) jointly satisfy (7) and (9), thereby showing that they are indeed the local minimizers. Moreover, although K is independent of N or $\epsilon(N)$, both N and $\epsilon(N)$ affect the actual value of the overall distortion, as they should.

4. SIMULATION RESULTS

Consider a 2-dimensional i.i.d standard Gaussian distributed source vector. The channel is modeled as a BSC with bit transition probability q and random index assignment. The generalized Lloyd algorithm described in [1] is used to generate a COVQ codebook with MSE as the distortion metric. For training the Lloyd algorithm as well as for measuring performance, 50,000 random instantiations of $\mathbf{x}$ were employed. Figs. 2 and 3 plot the expected MSE versus the number of quantized bits B and the bit transition probability q, respectively. The plots show the improvement in MSE that can be obtained by using an optimum codebook (compared with the curves obtained using the conventional codebook, labeled "unopt-CB"). Also,

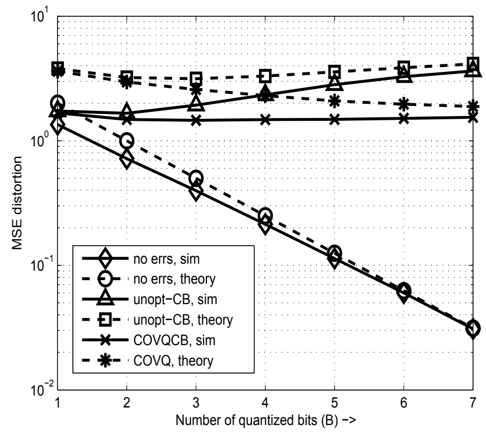


Fig. 2. MSE versus B, for a 2-d standard Gaussian distributed random vector and index sent over the BSC with $q = 0.005$.

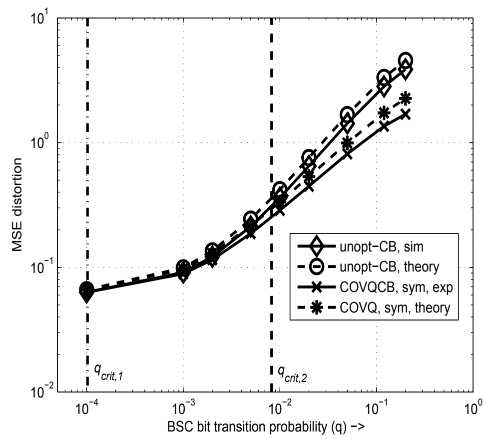


Fig. 3. MSE for a 2-d standard Gaussian distributed random vector with the conventional point density and $B = 6$. The index is sent over a BSC with bit transition probability q (the x-axis). The two vertical lines show the values of q corresponding to $\epsilon_{\text{crit},1}$ and $\epsilon_{\text{crit},2}$, the two critical values of $\epsilon(N)$, respectively.

in Fig. 3, the two values of q corresponding to $\epsilon(N) = \epsilon_{\text{crit},1}$ and $\epsilon(N) = \epsilon_{\text{crit},2}$ are plotted, to show that the simulation results agree with the theory over a wide range of values of q. Also, when $q > q_{\text{crit},2}$, the optimum point density is singular, i.e., $\varphi < N$.

Table 1 compares the theoretical and simulation-based values of φ for different BSC bit transition probabilities q and number of codepoints N, which shows that the theoretical and experimental values of φ match closely. For the experimental results, the φ was computed as the difference between N and the number of codepoints whose Voronoi cells were empty in the Lloyd algorithm. Also shown in the table is the observed MSE distortion.

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 case of the SEC, the asymptotic performance with MSE distortion functions was theoretically analyzed. It was demonstrated that the distortion is given by the sum of the distortion due to channel errors and the representation error in quantizing the source using a finite number of bits. The rate of decay of the two terms as the number of quantization levels N increases can be different, in which case, one of the two terms will

Table 1. Experimental and Theoretical Values of φ for different N and q. The tuples correspond to $(\varphi_{\text{exp}}, \varphi_{\text{theory}})$, for a 2-dimensional standard Gaussian distributed random vector. The number below the tuple is E_d, the expected distortion. φ_{theory} is computed from (14).

$N \backslash q$	0	0.0020	0.0050	0.0100
32	$(32, 32)$ 0.1136	$(32, 32)$ 0.1587	$(32, 32)$ 0.2216	$(32, 32)$ 0.3150
64	$(64, 64)$ 0.0600	$(64, 64)$ 0.1153	$(64, 64)$ 0.1869	$(63, 57)$ 0.2867
128	$(128, 128)$ 0.0309	$(128, 128)$ 0.0941	$(114, 107)$ 0.1674	$(82, 76)$ 0.2670
32	$(32, 31)$ 0.4753	$(24, 20)$ 0.8156	$(17, 15)$ 1.2092	$(16, 14)$ 1.3399
64	$(45, 41)$ 0.4472	$(31, 27)$ 0.8132	$(22, 20)$ 1.2320	$(20, 19)$ 1.3538
128	$(60, 52)$ 0.4337	$(37, 36)$ 0.8169	$(27, 27)$ 1.2547	$(22, 25)$ 1.3849

dominate as N gets large. Also, a novel theoretical analysis of the optimum *singular* point density that minimizes the overall distortion was derived, and its MSE performance evaluated. The accuracy of the theoretical results was illustrated through simulations.

5. 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] K. Zeger and A. Gersho, “Pseudo-Gray coding,” *IEEE Trans. Comm.*, vol. 38, no. 12, pp. 2147 – 2156, Dec. 1990.

[3] C. R. Murthy and B. D. Rao, “High-rate analysis of source coding for symmetric error channels,” in *Data Compression Conf. (DCC)*, Snowbird, UT, Mar 2006.

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

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

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

[7] 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.

[8] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, New York: Willy, 1991.

[9] E. R. Duni C. R. Murthy and B. D. Rao, “High-rate vector quantization for noisy channels with random index assignment,” *IEEE Trans. Info. Theory*, Submitted, 2006.

[10] Chandra R. Murthy, *Channel Estimation and Feedback Methods for Multiple-Antenna Communication*, Ph.D. thesis, Univ. of California, San Diego, 2006.

[11] S. W. McLaughlin and D. L. Neuhoff, “Asymptotic quantization for noisy channels,” in *Proc. IEEE Interational Symposium on Information Theory*, Jan. 1993, pp. 442–442.

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