

# Empirical Quantizer Design in the Presence of Source Noise or Channel Noise

Tamás Linder<sup>1</sup> Gábor Lugosi<sup>2</sup> Kenneth Zeger<sup>3</sup>

<sup>1,3</sup>Dept. of Electrical and Computer Engineering, University of California, San Diego,  
MC 0407, La Jolla, CA 92093, USA, ({linder,zeger}@code.ucsd.edu).

<sup>2</sup>Dept. of Economics, Pompeu Fabra University, Ramon Trias Fargas 25-27, 08005 Barcelona, Spain (lugosi@upf.es).

**Abstract** — The problem of vector quantizer empirical design from training vectors is studied for noisy channels and for noisy sources. It is shown that global empirical error minimization for designing quantizers for transmission over a discrete noisy channel have the same performance (convergence rate) as in ordinary quantizer design. For noisy source quantization an appropriate analogue of empirical error minimization is developed. Consistency and convergence rates are proved under appropriate regularity conditions in this case.

## I. NOISY CHANNEL QUANTIZERS

Let  $X$  be a  $k$ -dimensional random vector and consider a discrete noisy channel with  $N$  input and output symbols. An  $N$ -level noisy-channel vector quantizer is defined via its encoder  $Q_C$  which maps  $\mathbb{R}^k$  into  $\{1, \dots, N\}$ , and its decoder  $Q_D$  which maps  $\{1, \dots, N\}$  onto the set of codewords  $\{y_1, \dots, y_N\} \subset \mathbb{R}^k$ . The rate of the quantizer is  $(1/k) \log N$  bits per source symbol. The index  $I = Q_C(X)$  is transmitted through a noisy channel, and the decoder receives the index  $J \in \{1, \dots, N\}$ , whose conditional distribution given  $I$  is  $\{p(j|i)\}$ , the channel transition probabilities. The output of the quantizer is  $\hat{X} = Q_D(J) = y_J$  and we will use the notation  $\hat{X} = Q(X)$ . The mean-squared distortion  $\mathbb{E}\|X - \hat{X}\|^2$  is given by  $\sum_{i=1}^N \int_{R_i} \left( \sum_{j=1}^N \|y_j - x\|^2 p(j|i) \right) P_X(dx)$  where  $R_i = \{x : Q_C(x) = i\}$ . We call an  $N$ -level quantizer optimal if it minimizes the distortion over all  $N$ -level quantizers. Let us denote the distortion of such an optimal quantizer  $Q^*$  by

$$D^* = \mathbb{E}[\|Q^*(X) - X\|^2] = \min_{Q_C, Q_D} \mathbb{E}[\|Q(X) - X\|^2].$$

The task is to design an  $N$ -level noisy channel vector quantizer when the channel is known but the statistics of  $X$  are not. Rather,  $n$  training vectors  $X_1, \dots, X_n$ , independently drawn from the distribution of  $X$ , are given. The empirical distortion of the noisy channel quantizer is defined by  $D_n = n^{-1} \sum_{l=1}^n d_Q(X_l)$ , where

$$d_Q(x) = \sum_{i=1}^N \mathbf{1}_{\{x \in R_i\}} \left( \sum_{j=1}^N \|y_j - x\|^2 p(j|i) \right).$$

The empirical error minimization procedure produces a quantizer  $Q_n^*$  which minimizes the empirical distortion over the training data

$$Q_n^* = \arg \min_Q \frac{1}{n} \sum_{l=1}^n d_Q(X_l).$$

<sup>0</sup>This work was partially supported by the National Science Foundation and by OTKA Grant F 014174. <sup>1</sup>On leave from the Technical University of Budapest, Hungary.

The following result gives an upper bound on how much the expected distortion of the quantizer designed from  $n$  training vectors exceeds the distortion of the optimal quantizer.

**Theorem 1** Assume that the source  $X \in \mathbb{R}^k$  is bounded as  $\mathbb{P}(\|X\|^2 \leq B) = 1$  for some  $B > 0$ . Then

$$\mathbb{E}[\|Q_n^*(X) - X\|^2] \leq D^* + c \sqrt{\frac{\log n}{n}} + O(n^{-1/2}).$$

## II. NOISY SOURCE QUANTIZERS

Let  $Y = X + \nu$  be a version of  $X$  degraded by the independent additive noise  $\nu$ . An  $N$ -level noisy source quantizer  $Q$  is an ordinary  $N$ -level vector quantizer with general encoding regions  $\{R_i\}$ .  $Y$  is quantized by  $Q$ , and the distortion, which is measured with respect to the clean source  $X$ , is given by  $D(Q) = \mathbb{E}[\|X - Q(Y)\|^2]$ . It is well known that

$$D(Q) = \mathbb{E}[\|X - M(Y)\|^2] + \mathbb{E}[\|M(Y) - Q(Y)\|^2],$$

where  $M(y) = \mathbb{E}[X|Y = y]$ . In fact, the optimal quantizer  $Q^*$  can always be written as  $Q^*(y) = \hat{Q}^*(M(y))$ , where  $\hat{Q}^*$  is a nearest neighbor quantizer. The task is to design a noisy source quantizer when the statistics of the noise  $\nu$  and the noisy training vectors and  $Y_1, \dots, Y_n$  are given. The lack of the clean source samples corresponding to the  $Y_i$  require that we appropriately redefine the notion of empirical distortion. Let  $M_n(y)$  be an estimator of the regression function  $M(y)$  based on the first half of the noisy training samples. Define  $\hat{Q}_n^*$  as the nearest neighbor quantizer which minimizes the empirical error over the  $n/2$  modified training vectors  $M_n(Y_{n/2+1}), \dots, M_n(Y_n)$ . Then the output of the procedure is the noisy source quantizer given by  $Q_n^*(y) = \hat{Q}_n^*(M_n(y))$ . We have the following result on the excess distortion of  $Q_n^*$  over the distortion  $D^*$  of the optimal quantizer  $Q^*$ .

**Theorem 2** Assume that  $\mathbb{P}(\|X\|^2 \leq B) = 1$  and suppose that  $a_n$ , the  $L_2$  error of the estimator  $M_n$ , converges to zero as  $n \rightarrow \infty$ . Then

$$\mathbb{E}[\|X - Q_n^*(Y)\|^2] \leq D^* + c \sqrt{\frac{\log n}{n}} + O(n^{-1/2}) + O(a_n^{1/2}).$$

Based on deconvolution techniques, we can construct a consistent estimator  $M_n$  if the noise has a density whose characteristic function is nonzero a.e. Thus in this case the above procedure gives a consistent quantizer design method. It can also be proved if the clean training vectors are given instead of the noisy ones, the convergence rate of Theorem 1 is achievable, which is almost optimal in the minimax sense [1]

## REFERENCES

- [1] P. Bartlett, T. Linder, and G. Lugosi, "The minimax distortion redundancy in empirical quantizer design," *IEEE International Symposium on Information Theory*, Ulm, Germany, 1997.