On the Cost of Finite Block Length in Quantizing Unbounded Memoryless Sources[1]

Tamás Linder[2] and Kenneth Zeger[3]

[2] Dept. of Telecommunications, Technical University of Budapest, Hungary.
email: linder@vma.bme.hu.

[3] Coordinated Science Lab., Dept. of Elect. and Comp. Engineering, University of Illinois, Urbana-Champaign, IL 61801
email: zeger@uiuc.edu.

Abstract — **The problem of fixed-rate block quantization of an unbounded real memoryless source is studied. It is proved that if the source has a finite sixth moment, then there exists a sequence of quantizers Q_n of increasing dimension n and fixed rate R such that the mean squared distortion $\Delta(Q_n)$ is bounded as $\Delta(Q_n) = D(R) + O(\sqrt{\log n/n})$, where $D(R)$ is the distortion-rate function of the source. Applications of this result include the evaluation of the distortion redundancy of fixed-rate universal quantizers, and the generalization to the non-Gaussian case of a result of Wyner on the transmission of a quantized Gaussian source over a memoryless channel.**

Shannon's source coding theorem with a fidelity criterion [1] showed that by increasing the blocklength n of a lossy source code, it is possible to have the mean squared error approach the distortion-rate lower bound arbitrarily closely. Pilc [3] showed that for finite alphabet sources the convergence of the mean squared error to the distortion-rate function occurs at a rate $O(\log n/n)$. It has recently been shown [2] that for bounded real memoryless sources and squared distortion this convergence occurs at a rate $O(\sqrt{\log n/n})$ This result was used in [2] to analyze the performance of a certain universal quantization scheme. On the other hand, the assumption of bounded support is sometimes a severe restriction in signal quantization, especially since some of the most popular source models have unbounded support, such as the Laplacian. The convergence rate results mentioned above also assume that binary information is transmitted across a lossless channel. In the present paper we eliminate the bounded support requirement and also consider transmission across a noisy channel. In addition we are able to obtain a rate of convergence result for universal lossy source coding.

Theorem 1 *Let $X_1, X_2, \ldots$ be a real i.i.d. source with $\mathbf{E}|X_1|^2 = M_2$ and $\mathbf{E}|X_1|^6 < \infty$. Let $0 < R_1 < R_2$ and assume that $D(R_2) > 0$. Then for any $R \in [R_1, R_2]$ there exists an n-dimensional quantizer Q_n with rate $r(Q_n) \le R$ such that*

$$\Delta(Q_n) \le D(R) + B\sqrt{\frac{\log n}{n}},$$

for all $n \ge 1$, where the constant B depends only on R_1, R_2, and the source distribution. Furthermore, the quantizers satisfy

$$\max_{x\in\mathbb{R}^n} \frac{1}{n}\|Q_n(x)\|^2 \le 2M_2.$$

In [4] Wyner proved that $D_n(R) - D(R) = O(\log n/n)$ for memoryless Gaussian sources, and in [5] showed that $D_n(R) - D(R) = O(\sqrt{\log n/n})$ for any correlated Gaussian source with a sufficiently well-behaved spectral density. Recently Zamir and Feder [6] showed that a $O(\log n/n)$ convergence rate is achievable for correlated Gaussian sources by means of a variable rate coding scheme using subtractive dither.

Corollary 1 *Suppose we are given a real memoryless source $X_1, X_2, \ldots$ with distortion-rate function $D(R)$, satisfying $\mathbf{E}|X|^6 < \infty$, and a discrete memoryless channel of capacity C, accepting one input per source output. Then there exists a source-channel coding scheme with delay n, such that denoting by $\hat{X}_1, \hat{X}_2, \ldots, \hat{X}_n$ the channel decoder output, we have*

$$\frac{1}{n}\mathbf{E}\left(\sum_{i=1}^{n}|X_i - \hat{X}_i|^2\right) \le D(C) + O\left(\sqrt{\frac{\log n}{n}}\right).$$

Corollary 2 *For any $R > 0$, $k > 8$, and $\epsilon > 2/(k-4)$ there exists a sequence of universal quantizers $\{Q_n\}$ such that*

$$r(Q_n) - R = O\left(\left(\frac{\log n}{n}\right)^{(1/2)-\epsilon}\right),$$

and for any memoryless real source with $\mathbf{E}|X_1|^k < \infty$

$$\Delta(Q_n) - D(R) = O\left(\left(\frac{\log n}{n}\right)^{(1/2)-\epsilon}\right).$$

References

[1] R. G. Gallager. *Information Theory and Reliable Communication.* Wiley, New York, 1968.

[2] T. Linder, G. Lugosi, and K. Zeger. Rates of convergence in the source coding thorem, in empirical quantizer design, and in universal lossy source coding. *IEEE Trans. Inform. Theory*, vol. 40, no. 6, pp. 1728-1740, November 1994.

[3] R. Pilc. The transmission distortion of a source as a function of the encoding block length. *Bell System Technical Journal*, 47:827–885, 1968.

[4] A. D. Wyner. Communication of analog data from a Gaussian source over a noisy channel. *Bell System Technical Journal*, pages 801–812, May-June 1968.

[5] A. D. Wyner. On the transmission of correlated Gaussian data over a noisy channel with finite encoding block length. *Information and Control*, 20:193–215, 1972.

[6] R. Zamir and M. Feder. Information rates of pre/post filtered dithered quantizers. submitted to *IEEE Trans. Inform. Theory*, 1993.

[1]The research was supported in part by the National Science Foundation under Grants No. NCR-92-96231 and INT-93-15271 and the Joint Services Electronics Program.