the case for more conventional channels. This turns out to be only partly true with the gains more modest than anticipated.

The Steiner systems, used here simply because of their regular intersection or distance properties, appear to offer savings over the orthogonal systems at the lower bandwidths. Even the simple DBL system performs slightly better than the orthogonal case for all bandwidths considered. At the higher bandwidths 512 and 1024 (not reported on here), however, the improvement, while perhaps significant, may not justify the added complexity. Nonetheless, the tradeoffs among the various parameters make it an interesting study and, for a given bandwidth, nonorthogonal schemes of the type considered here should be of use.

ACKNOWLEDGMENT

The author would like to thank the reviewers for their valuable comments on the first version of this paper.

REFERENCES

[1] A. R. Cohen, J. A. Heller, and A. J. Viterbi, "A new coding technique for asynchronous multiple access communication," *IEEE Trans. Commun. Technol.*, vol. COM-19, pp. 849–855, 1971.

[2] A. J. Viterbi, "Convolutional codes and their performance in communication systems," *IEEE Trans. Commun. Technol.*, vol. COM-19, pp. 751–772, 1971.

[3] E. S. Kramer and D. M. Mesner, "Admissible parameters for Steiner systems $S(t, k, v)$ with a table for all $(v - t) \leq 498$," *Utilitas Math.*, vol. 7, pp. 211–222, 1975.

[4] M. Hall, *Combinatorial Theory*. Waltham, MA: Blaisdell, 1967.

A Note on the Computation of Optimal Minimum Mean-Square Error Quantizers

J. A. BUCKLEW AND N. C. GALLAGHER, JR.

Abstract—**This paper considers the problems associated with computing optimal minimum mean-square error quantizers. Most computational methods in current use are iterative. These iterative schemes are extremely sensitive to initial conditions. Various methods of obtaining good initial conditions are presented and discussed.**

I. INTRODUCTION

In his classic paper of 1960, Max presents an iterative scheme for the computation of one-dimensional minimum mean-squared error quantization characteristics [1]. In addition, he solves for the optimum Gaussian quantizer for up to 36 output levels. In [2], Gallagher uses Max's method in the computation of optimum Rayleigh quantizer parameters, and in [3] Paez and Glisson use the same method to compute the optimum Laplacian quantizer later recomputed by Adams and Giesler [4]. Max's algorithm is very simple to program into a digital computer, and we view this simplicity as a good reason for using his method. However, one problem that arises with this algorithm is its failure to always converge to the optimum solution when the number of quantizer output levels is large. The reason for this is that the initial guess for starting the iteration must be increasingly precise as the number of quantizer levels becomes large. So, for a 64-level quantizer, Max's algorithm will not converge to the optimum solution unless the initial guess for the first output level is very close to the true value. This difficulty has prompted others to employ more sophisticated optimization methods in the solution for optimum quantizers. For example, Pearlman and Senge [5] use a vector space optimization technique that is a combination of the steepest descent and Newton-Raphson methods to solve for the optimum Rayleigh quantizer. It is not our purpose to detract from this and similar methods that do work well, but in our view, if the starting point problem can be solved, Max's method is the preferred method of solution. In Section II we discuss several methods for choosing the iteration's initial condition very accurately, and we have demonstrated convergence of Max's algorithm for at least 10 000 output levels and present numerical examples in Section III.

II. THE COMPUTATION OF OPTIMUM ONE-DIMENSIONAL QUANTIZERS

A common method for implementing one-dimensional quantizers is the companding method as discussed by Smith [6]. The companding method is straightforward: the input signal x with probability density $p(x)$ first enters the invertible nonlinearity $g(x)$, called the compressor; then it goes into a uniform quantizer over the range [0, 1], and upon reconstruction it passes through the expansion nonlinearity $g^{-1}(x)$. For minimum mean-squared error quantization, the asymptotically optimum compressor function is given by

$$g(x) = \left[\int_{-\infty}^{\infty} [p(y)]^{1/3} dy\right]^{-1} \int_{-\infty}^{x} [p(y)]^{1/3}\, dy. \quad (1)$$

In Max's classic 1960 paper an iterative method is presented whereby the exact quantizer parameters can be computed for finite N.

Max's algorithm provides a method for the solution of the equations

$$e_i = (y_i + y_{i-1})/2, \qquad i = 2, \cdots, N \quad (2a)$$

and

$$\int_{e_i}^{e_{i+1}} (x - y_i) p(x)\, dx = 0, \qquad i = 1, \cdots, N \quad (2b)$$

where the output levels of the quantizer are denoted y_1, y_2, $\cdots$, y_N and the internal breakpoints as e_1, e_2, $\cdots$, e_{N+1}. Typically, endpoint values e_1 and e_{N+1} are known *a priori* and the first step of Max's procedure is to choose a value for y_1 with which to solve (2b) for the value e_2. We then use this value in (2a) to find y_2 and use this to find e_3 in (2b), and so on. The last integral over (e_N, e_{N+1}) can be used to determine

Paper approved by the Editor for Data Communication Systems of the IEEE Communications Society for publication without oral presentation. Manuscript received January 5, 1981; revised April 27, 1981. This work was supported by the Air Force Office of Scientific Research under Grant AFOSR 78-3605.

J. A. Bucklew is with the Department of Electrical and Computer Engineering, University of Wisconsin, Madison, WI 53706.

N. C. Gallagher, Jr. is with the School of Electrical Engineering, Purdue University, West Lafayette, IN 47907.

0090-6778/82/0100-0298$00.75

the accuracy of the initial guess for y_1. If the last integral is zero within a specified error, we use the computed parameters to specify the quantizer; if not, we make a new guess for y_1 and begin the procedure again. Details on how to modify the initial guess for y_1 are not specified by Max.

We have computed quantizers using Max's method for several densities. It has been our observation that the convergence properties of Max's algorithm are greatly dependent on the initial guess for y_1. Let y_{1N} denote the first output level for an optimum N level quantizer. Intuitively, if the first guess at y_{1N} (call it $\hat{y}_{1N}$) is very close to $y_{1(N+1)}$, then Max's algorithm tries to converge to the $N+1$ level quantizer. A consideration of Max's method indicates that the first N steps of the algorithm are the same for the N or $N+1$ level quantizers. Although never reported in the literature, it is our understanding that this phenomenon has been widely observed [7].

As an aside, we remark that the conditions presented in (2) are not sufficient conditions to specify the optimum quantizer; they are only necessary. However, in 1965 Fleisher [8] showed that if

$$\frac{d^2}{dx^2}[\ln p(x)] < 0$$

then the expressions in (2) are both necessary and sufficient for the specification of the minimum mean-squared error quantizer, and their solution provides us with the unique optimum quantizer.

We now describe two similar methods for generating a good initial condition. First, note that the initial condition can be a guess at the value for y_1 or a guess for the value of any y_i, $i = 1, \cdots, N$ wherever we choose to begin the iteration. The first method is a modified version of an estimation method by Panter and Dite [9] and Roe [10]. The second method employs a companding model to produce the iteration starting point. Both methods grow more precise as the number of quantization levels N increases. Each method, however, requires computation to generate an initial value; the complexity of this computation varies depending on the distribution of the variable to be quantized.

In the first method we use the asymptotic level density $\lambda(x)$ for the minimum mean-squared error quantizer. $\lambda(x)\Delta x$ is approximately the ratio of the number of output levels in a region Δx about x to the total number of output levels N. This function is the first derivative of the compressor function $g(x)$ in (1):

$$\lambda(x) = g'(x) = [p(x)]^{1/3}\left[\int_{-\infty}^{\infty}[p(y)]^{1/3}\,dy\right]^{-1}. \tag{3}$$

Smith [6] shows that this function has the property that for adjacent output levels y_i and y_{i+1},

$$y_{i+1} - y_i \cong \frac{1}{N\lambda(y)}, \qquad \text{for } y \in [y_i, y_{i+1}] \tag{4}$$

when the number of output levels is large. As an aside, we remark that our compressors always have unity range. Smith allows more generality in his formulas. The best way to illustrate the use of (4) is through an example. Suppose that $p(x)$ is a zero-mean symmetric density (no Dirac delta functions), that N is even, and that a unique optimum quantizer exists. The initial condition for the Max iteration is a guess for first output level greater than zero. We will call this level $y_{N/2}$. We first make the observation that the output levels must be symmetric about the origin. Also, for large N, the distance between the breakpoint at zero and $y_{N/2}$ approximately equale

$$y_{\frac{N}{2}} \cong \frac{1}{2N\lambda(y_{\frac{N}{2}})} \tag{5}$$

The solution of this equation provides the initial guess for $y_{N/2}$. This basic procedure can be used with modifications for N even or odd with most common probability densities. Some numerical examples are provided in the next section.

The second method uses the companding function to work backwards from the known uniform quantizer over [0, 1] in order to estimate the initial output level. In fact, the method provides a reasonable approximation to the entire quantizer. An N level uniform quantizer on [0, 1] has output levels

$$\hat{y}_i = \frac{2i-1}{2N}, \qquad i = 1, \cdots, N. \tag{6}$$

Therefore, the compander approximation is simply

$$y_i \cong g^{-1}(\hat{y}_i) = g^{-1}\left(\frac{2i-1}{2N}\right). \tag{7}$$

For the purpose of identification, we will refer to the first method of (5) as the λ-approximation and the second as the g-approximation. In hindsight these two methods seem obvious; however, they have apparently not been widely used.

III. NUMERICAL EXAMPLES

In this section we provide some examples using the λ- and g-approximations to estimate the initial input interval endpoint of a Max quantizer. The asymptotically optimum mean-square error companding characteristic is given by

$$\frac{\int_{-\infty}^{x} p(y)^{1/3}\,dy}{\int_{-\infty}^{\infty} p(y)^{1/3}\,dy} = g(x)$$

where $p(y)$ is our input probability density.

The first example we consider is when $p(y)$ is the Gaussian unit variance, zero mean, probability density: $g(x)$ is then given by $\frac{1}{2}(1 + \text{erf}\,(x/\sqrt{6}))$; hence, $g^{-1}(y) = \sqrt{6}\ \text{erf}^{-1}(2y-1)$. Using this equation, our expression for the initial positive input interval endpoint of an N output level quantizer is $x_{1\lambda} = \sqrt{6}\ \text{erf}^{-1}(2(N/2+1)-1)$.

The λ-approximation requires us to solve the equations (using a standard Newton-Raphson search)

$$x_{1\lambda} = \frac{1}{N\lambda(x_{1\lambda})} \qquad \text{for } N \text{ even}$$

$$x_{1\lambda} = \frac{1}{2N\lambda(x_{1\lambda})} \qquad \text{for } N \text{ odd}$$

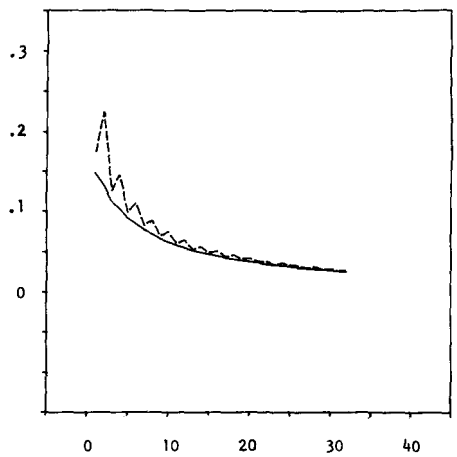

Fig. 1. P_g (solid line) and P_λ (dotted line) plotted as a function of N for the Gaussian density.

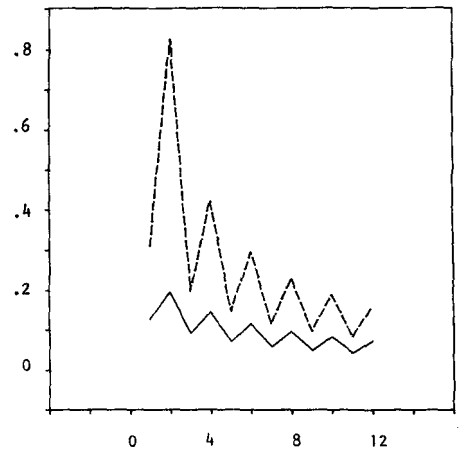

Fig. 2. P_g (solid line) and P_λ (dotted line) plotted as a function of N for the Laplacian density.

where

$$\lambda(x_{1\lambda}) = \frac{(2\pi)^{1/6}}{(6\pi)^{1/2}}\, p(x_{1\lambda})^{1/3}.$$

Since Max tabulated the actual values of the input interval endpoints, we may compute the quantities

$$P_g \triangleq \left| \frac{x_{1g} - x_{\text{act}}}{x_{\text{act}}} \right|$$

and

$$P_\lambda \triangleq \left| \frac{x_{1\lambda} - x_{\text{act}}}{x_{\text{act}}} \right|$$

for various values of N where x_{act} is the actual tabulated value.

In Fig. 1 we see P_g (solid line) and P_λ (dotted line) plotted as a function of N for values of N from 5 to 36. As may be seen from the figure, the g-approximation is better for all these vaues of N. Furthermore, the λ-approximation does not have a solution for $N = 4$, which is an additional drawback of using this approximation in low N regions.

We now perform the same computations for the Laplacian ($p(y) = \exp\{-|y|\}/2$) and Rayleigh ($p(y) = y \exp\{-y^2/2\}$) probability densities. In Fig. 2 we plot P_g (solid line) and P_λ (dotted line) for values of N from 5 to 16 for the Laplacian density. Again, the g-approximation is best for all values of N and, furthermore, the λ-approximation has no solution when $N = 4$.

In Fig. 3 we see plots of P_g (solid line) and P_λ (dotted line) for values of N from 2 to 36 for the Rayleigh distribution. For every value except $N = 2$, the g-approximation is better than the λ-approximation. The plot of P_g is noisy because calculation of x_{1g} for this density required a large numerical integration which was very sensitive to the number of samples used in the summation.

We should note that Max quantizers have been computed for the Rayleigh and the Gaussian densities using both $x_{1\lambda}$ and x_{1g} as the estimate for the initial interval endpoint. With no convergence problems, quantizers of 10 000 and 200 output

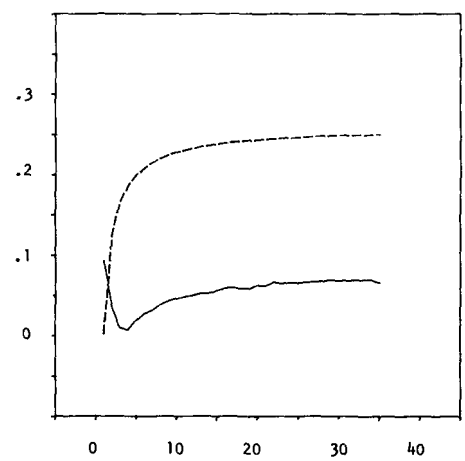

Fig. 3. P_g (solid line) and P_λ (dotted line) plotted as a function of N for the Rayleigh density.

levels have been computed for the Gaussian and Rayleigh probability densities, respectively. In practice, we find that both methods give sufficiently good estimates to allow quick convergence to the correct quantizer. A typical value is 200 iterations for a 1000 level Gaussian quantizer with the last level specified to 10^{-5} accuracy. We conclude that the x_{1g} estimate is a better approximation in most cases, but the $x_{1\lambda}$ estimate is often substantially easier to compute.

REFERENCES

[1] J. Max, "Quantizing for minimum distortion," *IRE Trans. Inform. Theory*, vol. IT-6, pp. 7–12, Mar. 1960.
[2] N. C. Gallagher, "Optimum quantization in digital holography," *Appl. Opt.*, vol. 17, pp. 109–115, Jan. 1, 1978.
[3] M. D. Paez and T. H. Glisson, "Minimum mean-squared error quantization in speech PCM and DPCM systems," *IEEE Trans. Commun.*, vol. COM-20, pp. 225–230, Apr. 1972.
[4] W. C. Adams and C. E. Giesler, "Quantizing characteristic for signals having Laplacian amplitude probability density function," *IEEE Trans. Commun.*, vol. COM-26, pp. 1295–1297, Aug. 1978.
[5] W. A. Pearlman and G. H. Senge, "Optimal quantization of the Rayleigh probability distribution," *IEEE Trans. Commun.*, vol. COM-27, pp. 101–112, Jan. 1979.
[6] B. Smith, "Instantaneous companding of quantized signals," *Bell Syst. Tech. J.*, vol. 36, pp. 653–709, May 1957.
[7] E. Delp and J. A. Bucklew, *mutual correspondence*, 1977.
[8] P. E. Fleisher, "Sufficient conditions for achieving minimum distortion in a quantizer," in *IEEE Int. Conv. Rec.*, 1964, pp. 104–111.

[9] P. F. Panter and W. Dite, "Quantization distortion in pulse-count modulation with nonuniform spacing of levels," *Proc. IRE*, vol. 39, pp. 44–48, Jan. 1951.
[10] G. M. Roe, "Quantizing for minimum distortion," *IRE Trans. Inform. Theory*, vol. IT-10, pp. 384–385, Oct. 1964.

A Note on the Distribution of Atmospherically Ducted Signal Power Near the Earth's Surface

E. J. DUTTON

Abstract—**Interference fields caused by low-level atmospheric ducts are of concern at UHF frequencies and above. This note develops a long-term median distribution of received signal power during ducting conditions from observations taken on a 175 km, 5 GHz path in New Jersey in 1966. A simple worst case estimate for the received signal power during ducting is also derived.**

I. INTRODUCTION

During an average year, ducting layers that occur at or near the earth's surface can be expected between about 1 and 15 percent of the time, depending upon one's location in the United States [1]–[3]. If these layers are of sufficient horizontal homogeneity, and any UHF and higher frequency radio links are in the vicinity, the potential is high for enhanced signal power reception at either a desired or undesired cochannel receiver for part or most of the time the ducting layers are present. There is, however, to the knowledge of this author, only extremely sparse information and data on the conditional distribution of received signal level (RSL) in the presence of ducting layers. There appears to be equally little information on how this distribution fits into the determinations of the unconditional distribution of RSL on a particular link. We shall restrict ourselves here to the conditional distribution, and present a procedure for obtaining the distribution and its parameters based, in part, on data taken during the POPSI experiment [4]. The POPSI experiment observed 5-min median signal levels once per hour for one week of observations, August 2–August 9, 1966. During this time there were several observations of on-path ducting. The ducting layer apparently was an elevated layer with some undulations near the earth's surface [4]. The observations were made at 5 GHz over a 175 km path between Highlands and Wildwood, NJ, which we shall call the "HW path" for brevity.

II. LONG-TERM DISTRIBUTION ESTIMATION

Fig. 1, which was presented by Crane [4], shows the distribution of the RSL's for the ducting cases on the HW path.

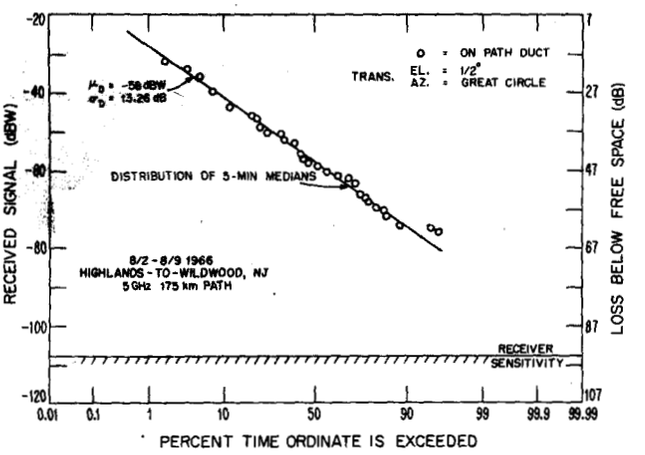

Fig. 1. Distribution of received signal power for a set of data observations.

The line drawn through these data is an "eyeball" fit.[1] These data are, according to Crane [4], 5-min median RSL's observed once per hour. Hence, Fig. 1 represents the "long-term" distribution of the median power x_m for the 5-min observation period. Thus, the straight line in Fig. 1 should be a lognormal distribution [12] of the form, conditional on the presence of ducting D:

$$P(X_m \leq x_m \mid D) = \frac{4.3429}{\sigma_p x_m \sqrt{2\pi}} \cdot \int_{-\infty}^{x_m} \exp\left[\frac{-(10 \log_{10} x_m - \mu_D)^2}{2\sigma_D^2}\right] dx_m. \quad (1)$$

In the case of Fig. 1, we can obtain $\mu_D = -58$ dBW and $\sigma_D = 13.26$ dB. Equation (1) is not advocated as the only long-term distribution applicable to median observations, although the lognormal form tends to be commonly used for many applications [5], [6]. Crane [4] also considers RSL data taken on the HW path during a different time period (August 19–September 1, 1966) in which further ducting, as well as troposcatter results, were obtained. Without going into detail (because the mathematical results would not be particularly clear without an accompanying long and distracting derivation), the troposcatter observations, which are also available for the August 2–August 9, 1966 period, can be combined with the ducting observations of Fig. 1, as described by (1). Also, a combined ducting-troposcatter conditional distribution can be obtained to represent August 19–September 1, 1966 HW path RSL data. Finally, the two distributions can be compared, as shown in Fig. 2, and it is noted that the maximum difference between the two independently derived distributions is about 3.5 dB. This adds a little credence to the use of the lognormal distribution (1). Note that neither

Paper approved by the Editor for Radio Communication of the IEEE Communications Society for publication without oral presentation. Manuscript received May 7, 1981; revised July 14, 1981.

The author is with the Propagation Predictions and Model Development Group, National Telecommunciations and Information Administration Institute for Telecommunication Sciences, U.S. Department of Commerce, Boulder, CO 80303.

[1] There is an admitted lack of rigor in an "eyeball" fit as opposed to use of certain statistical procedures to test goodness of fit. This was done because the values for the data points were unavailable and because this correspondence is intended primarily to illustrate methodology.