MMA Memo 220

Quantization Efficiency for Eight or more Sampling Levels

A. R. Thompson

July 9, 1998

Abstract A simplified method is used to derive the quantization efficiency (that is, the fractional loss in signal-to-noise ratio) resulting from quantization for eight or more levels. The method is based on the estimation of the increase in the variance of the signal waveform by the addition of quantization errors. The accuracy appears to be approximately 0.1% for eight-levels and should improve as the number of levels is increased.

Digital samples with three or more bits per sample (eight or more quantization levels) have been suggested as a possible format for IF signals that would be digitally filtered before input to the delays and correlator of the MMA. Since such signals would have to be truncated to no more than four levels before entering the correlator, use of more levels in the initial quantization would keep the overall loss of sensitivity similar to that for a single quantization without subsequent truncation. For consideration of such a scheme it is necessary to know the quantization efficiency factors for the larger numbers of levels, that is, the fractions by which the signal-to-noise ratio at the correlator output is degraded by the quantization. The quantization efficiencies for two, three, and four levels have been calculated by a number of authors: see e.g. Cooper (1970), Hagen and Farley (1973) or the review of methods and results in Thompson et al. (1986). These methods involve calculation of the mean and mean-squared outputs of the correlator and become somewhat cumbersome when applied to larger numbers of levels. This memorandum explores the possibility of using a simpler approximation for cases where the number of levels is large, and consequently the loss in sensitivity is small. The principle of the method is to calculate the fractional increase in the variance of a signal that results from the quantization. The signal-to-noise ratio at the correlator output is inversely proportional to this variance.

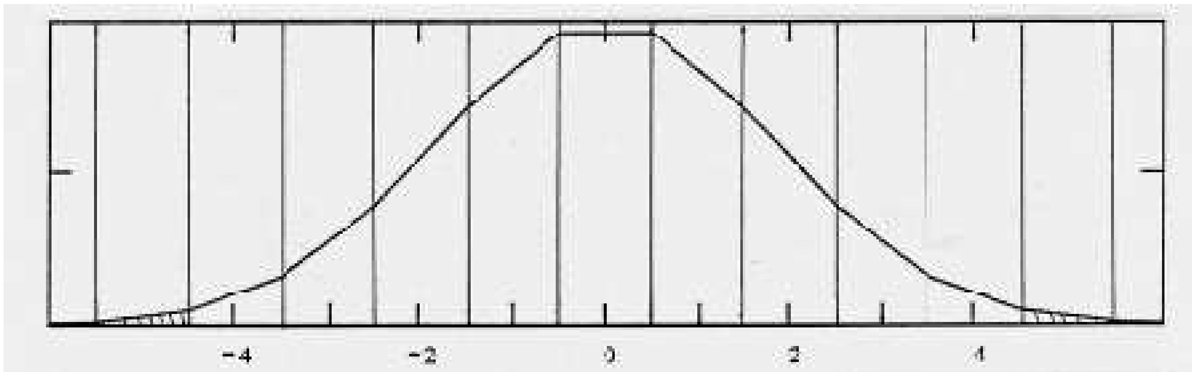

Figure 1. A representation of the Gaussian probability distribution of amplitude of the IF signal, with quantization thresholds for nine-level sampling shown by the vertical lines. The figures along the bottom indicate the digitized levels (weights) assigned to the signals that fall within the corresponding levels. The width of each quantization level is $\alpha\sigma$ within the range $\pm 4\frac{1}{2}\alpha\sigma$ of signal voltage. Signal levels outside of this range fall within the small shaded areas under the curve and are assigned quantized values of -4 or $+4$.

Figure 1 shows a piecewise-linear approximation of the Gaussian probability distribution of a signal from one antenna. This approximation simplifies the mathematics without significantly degrading the results. Quantization thresholds for nine-level representation of the signal are indicated by the vertical lines. The widths of the levels are $\alpha\sigma$ in voltage, that is, α in units of σ where σ^2 is the unquantized variance. We consider here the case where the number of levels is odd, and the central level is centered on zero volts. The probability that any one sample will fall between the two consecutive thresholds at $(m - \frac{1}{2})\alpha\sigma$ and $(m + \frac{1}{2})\alpha\sigma$, where m is an integer, is

$$\frac{1}{\sqrt{2\pi}\sigma}\left[e^{-(m-\frac{1}{2})^2\alpha^2/2} + e^{-(m+\frac{1}{2})^2\alpha^2/2}\right]\frac{\alpha}{2}. \tag{1}$$

Consider, for example, the level for which $m = 3$. Any voltage that falls within this level is encoded as amplitude 3, although it can be anywhere between $2\frac{1}{2}$ and $3\frac{1}{2}$ in units of $\alpha\sigma$. The mean increase in the variance resulting from this representation is

$$\frac{2}{\alpha\sigma}\int_0^{\alpha\sigma/2} x^2 dx = \frac{1}{3}\left(\frac{\alpha\sigma}{2}\right)^2. \tag{2}$$

Note that this is exact for the piecewise-linear probability curve in Fig. 1. The same increase in variance resulting from quantization applies to the range of signal levels from $-4\frac{1}{2}\alpha\sigma$ to $+4\frac{1}{2}\alpha\sigma$. The fraction of the area under the Gaussian probability curve that lies between these levels is

$$\frac{1}{\sqrt{2\pi}}\int_{-4.5\alpha\sigma}^{4.5\alpha\sigma} e^{-x^2/2\sigma^2} dx = \mathrm{erf}\left(\frac{4.5\alpha}{\sqrt{2}}\right), \tag{3}$$

Thus the variance resulting from quantization for signals between $-4\frac{1}{2}\alpha\sigma$ to $+4\frac{1}{2}\alpha\sigma$ is

$$\frac{1}{3}\Big(\frac{\alpha\sigma}{2}\Big)^2 \mathrm{erf}\Big(\frac{4.5\alpha}{\sqrt{2}}\Big). \tag{4}$$

We shall assume that the quantization error is essentially uncorrelated with the unquantized signal. In the extreme case of two-level sampling the quantization error is essentially the same as the unquantized signal, so the treatment used here would not apply. Consider, however, the case of multilevel quantization as in Fig. 1. If the signal voltage is increased steadily, the quantization error decreases from a maximum at each quantization threshold to zero when the voltage is equal to the mid point of two thresholds. At each threshold the quantization error changes sign and the cycle repeats. This behaviour effectively destroys any correlation between the quantization error and the signal waveform.

It is also necessary to take account of the effect of counting all signals below $-4\frac{1}{2}\alpha\sigma$ as level -4, and those above $+4\frac{1}{2}\alpha\sigma$ as $+4$. To make an approximate estimate of this effect we divide the range of signal level outside of $\pm 4\frac{1}{2}\alpha\sigma$ into intervals of width $\alpha\sigma$. Consider, for example, the interval centered on $6\alpha\sigma$. The probability of the signal falling within this level is equal to the corresponding area under the curve, which for the piecewise-linear approximation is

$$\tfrac{1}{2}\tfrac{\alpha}{\sqrt{2\pi}}\big[e^{-(5.5\alpha)^2/2} + e^{-(6.5\alpha)^2/2}\big]. \tag{5}$$

The square of the mean error resulting from quantization of the signal within this range is $[(6-4)\alpha\sigma]^2$, so the total variance of the quantization error for signals outside the range $\pm 4\frac{1}{2}\alpha\sigma$ is

$$\frac{\alpha^3\sigma^2}{\sqrt{2\pi}}\sum_{m=5}^{15}(m-4)^2\big[e^{-(m-\frac{1}{2})^2\alpha^2/2} + e^{-(m+\frac{1}{2})^2\alpha^2/2}\big]. \tag{6}$$

The upper limit of the summation in (6) is chosen to be large enough that increasing it does not significantly change the result. The quantization error resulting from the truncation of the signal values outside the range $\pm 4\frac{1}{2}\alpha\sigma$ clearly has some degree of correlation with the unquantized signal level. However, this is not large because the fraction of samples for which the signal lies outside $\pm 4\frac{1}{2}\alpha\sigma$ is less than 1.5% for nine-level quantization, and decreases as the number of quantization levels increases. We shall therefore treat the quantization error resulting from the truncation of the signal peaks as uncorrelated with the signal, but bear in mind that this assumption may introduce a small uncertainty into the calculation.

The variance of the quantized signal is equal to the variance of the unquantized signal (σ^2) plus the variance of the quantization errors in (4) and (6), that is,

$$\sigma^2 + \frac{1}{3}\Big(\frac{\alpha\sigma}{2}\Big)^2 \mathrm{erf}\Big(\frac{4.5\alpha}{\sqrt{2}}\Big) + \frac{\alpha^3\sigma^2}{\sqrt{2\pi}}\sum_{m=5}^{15}(m-4)^2\big[e^{-(m-\frac{1}{2})^2/2\alpha^2} + e^{-(m+\frac{1}{2})^2\alpha^2/2}\big]. \tag{7}$$

If the variance is the same for both signals at the correlator input, and if the correlation of the two signals is small, the signal-to-noise ratio at the correlator output is inversely proportional to the variance. Thus the quantization efficiency factor is

$$\eta_{(2N+1)} = \left\{1 + \frac{1}{3}\Big(\frac{\alpha}{2}\Big)^2 \mathrm{erf}\Big(\frac{(N+\frac{1}{2})\alpha}{\sqrt{2}}\Big) + \frac{\alpha^3}{\sqrt{2\pi}} \sum_{m=N+1}^{N+20} (m-N)^2 \big[e^{-(m-\frac{1}{2})^2\alpha^2/2} + e^{-(m-\frac{1}{2})^2\alpha^2/2}\big]\right\}^{-1}. \tag{8}$$

Here the equation has been generalized for for $2N+1$ levels. The result from Eq. (8) for nine-level quantization is $\eta_9 = 0.968$ for $\alpha = 0.547$. The value of α was chosen to maximize η_9. These values can be compared with results for nine-level quantization computed by F. R. Schwab using the more precise methods mentioned above. For nine levels this involves evaluation of $\sim 9^2/2$ probability integrals. Schwab obtained $\eta_9 = 0.969$ for $\alpha = 0.534$. The quantization efficiency varies only slowly with α, and the difference in the two α values above makes no significant change in η_9. The value of η_9 from Eq. (8) agrees with the value obtained by Schwab to within $\sim 0.1\%$, or $\sim 3\%$ in the degradation factor $(1-\eta_9)$. This agreement verifies the present method within these limits of accuracy.

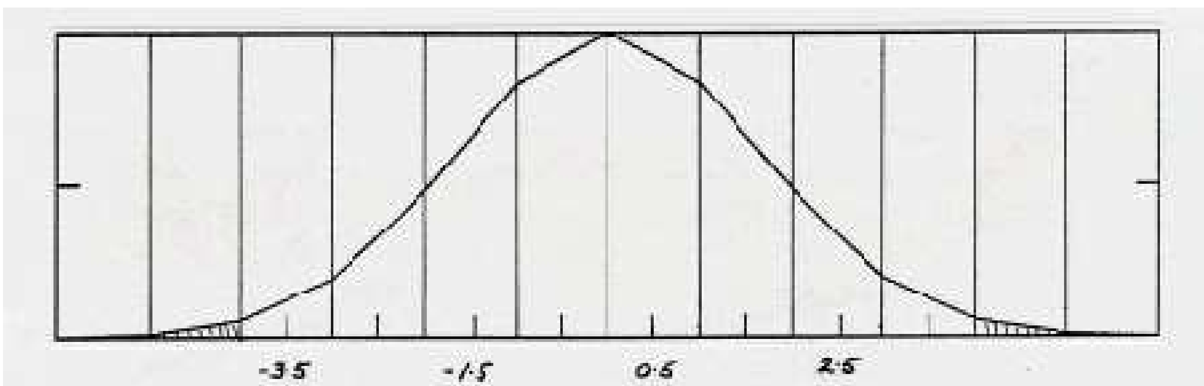

Figure 2. Representation of the Gaussian probability distribution of the amplitude of the IF signal, as in Fig. 1, but for an even number of quantization levels (eight shown in the figure). The quantization thresholds are shown by the vertical lines, and the zero level for signal input coincides with the central threshold, rather than with the center of a level as in Fig. 1. Again the figures along the bottom indicate the digitized levels (weights) assigned to the signals in that fall within the corresponding levels. In this case there is no zero level. Signal levels outside the range $\pm 4\frac{1}{2}\alpha\sigma$ fall within the the shaded areas.

For quantization with an even number of levels, as shown in Fig. 2, the central threshold level is at zero signal volts, and the the expression for the quantization efficiency is

$$\eta_{(2N)} = \left\{1 + \frac{1}{3}\Big(\frac{\alpha}{2}\Big)^2 \mathrm{erf}\Big(\frac{N\alpha}{\sqrt{2}}\Big) + \frac{\alpha^3}{\sqrt{2\pi}} \sum_{m=N+1}^{N+20} (m-N)^2 \big[e^{-(m-1)^2\alpha^2/2} + e^{-m^2\alpha^2/2}\big]\right\}^{-1}. \tag{9}$$

Results from Eqs. (8) and (9), evaluated using Mathcad, are given in Table 1 for 8, 9, 16, and 32 levels.

TABLE 1. Calculations of Quantization Efficiency and Other Parameters.

No. of Levels	N	α	P	η
8	4	0.604	0.016	0.960
9	4	0.547	0.014	0.968
16	8	0.399	0.0067	0.988
32	16	0.189	0.0025	0.996

The second column of Table 1 gives the value of N which indicates the range of quantized values. The third column of the Table gives P, which is the fraction of samples for which the signal amplitude is greater than $\pm(N + \frac{1}{2})\alpha\sigma$ for odd numbers of levels or greater than $\pm N\alpha\sigma$ for even numbers of levels. P is the fraction of signal samples that contribute to the variance in (6). The values of P are less than 1.6% in all cases listed in Table 1, and decrease as the number of levels increases.

With values of quantization efficiency approaching unity other effects, such as the departure of the bandpass responses from the ideal rectangular shape, become limiting factors. These should be borne in mind if the absolute value for η is critical. An interesting point about the treatment here is that it shows the connection between the quantization efficiency and the quantization errors rather more clearly than other methods.

Thanks are due to F. Schwab for permission to quote his unpublished results and for discussion of the memorandum in draft.

REFERENCES

Cooper, B. F. C., Correlators with Two-Bit Quantization, *Aust. J. Phys.*, **23**, 521-527, 1970.

Hagen, J. B. and D. T. Farley, Digital Correlation Techniques in Radio Science, *Radio Sci.*, **8**, 775-784, 1973.

Thompson, A. R., J. M. Moran, and G. W. Swenson, Interferometry and Synthesis in Radio Astronomy, Wiley, NY, 1986, and reprinted by Kreiger, Malabar, FL. 1991, 1994. (see Ch. 8).