VLB ARRAY MEMO No. 332

# MINIMIZING STORAGE REQUIREMENTS FOR QUANTIZED NOISE
## (or, how to use less tape in VLBI)

Larry R. D'Addario
28 February 1984

## Introduction

It is well known (Hagen and Farley, 1973) that correlation receivers for Gaussian noise signals can use coarsely quantized versions of the signals with little loss of sensitivity. For example, under commonly used assumptions the relative signal-to-noise ratios for quantization to n=2, 3, 4, and infinity levels are $R_n$=0.64, 0.81, 0.88, and 1.00, respectively. If the signals are sampled at a rate that produces independent samples, then the number of samples that must be correlated to achieve a given SNR is proportional to $(1/R_n)^2$.

While the number of samples needed for a given sensitivity decreases monotonically with n, the information contained in each sample generally increases with n. (For n>2, other parameters of the system come into play, including the choice of quantization thresholds and multiplier weighting; we shall consider these details shortly.) Therefore, there is a choice of n which minimizes the total information which must be collected in order to achieve a given sensitivity. When the information needs to be stored before correlation (as in VLBI), this can be important. We shall see that the optimum n from this point of view is n=3.

## Calculations

The information per sample is given by

(1) $$I_n = -\sum_{i=1}^{n} p_i \log(p_i)$$

where $p_i$ is the probability that any one sample is quantized to level i. For small correlation coefficients and independent samples, the SNR per sample has been calculated for n=2 (Weinreb 1961) and for n=3,4 (Cooper 1970):

(2) $$R_2 = 2/\pi$$

(3) $$R_3 = \frac{2}{\pi} \frac{E^2}{(1-\phi)}$$

(4) $$R_4 = \frac{2}{\pi} \frac{WE^2 + 2E(1-E)}{\sqrt{W^2(1-\phi)^2 + 2\phi(1-\phi)}}$$

where (4) is based on the use of a simplified multiplier in which the products of smallest magnitude are set to zero. In these expressions, $\pm W$ are the weights of the largest-magnitude levels;

(5) $$E = \exp(-\alpha^2/2);$$

(6) $$\phi = \mathrm{erf}(\alpha/\sqrt{2});$$

and $\pm\alpha$ are the non-zero quantization thresholds divided by the rms signal amplitude.

The storage requirements are then

(7) $$S_n = I_n/R_n^2.$$

The latter quantity is plotted in Fig. 1 against $\alpha$ for $n=3$ and $n=4$ (simplified multiplier). Also plotted is $R_n$. The usual practice is to chose $\alpha$ to maximize $R_n$, but it is apparent that a different choice is needed to minimize $S_n$. Nevertheless, for either choice, $S_3 < S_4$ and $S_3 < S_2 = 2.468$ (since $I_2=1$ bit).

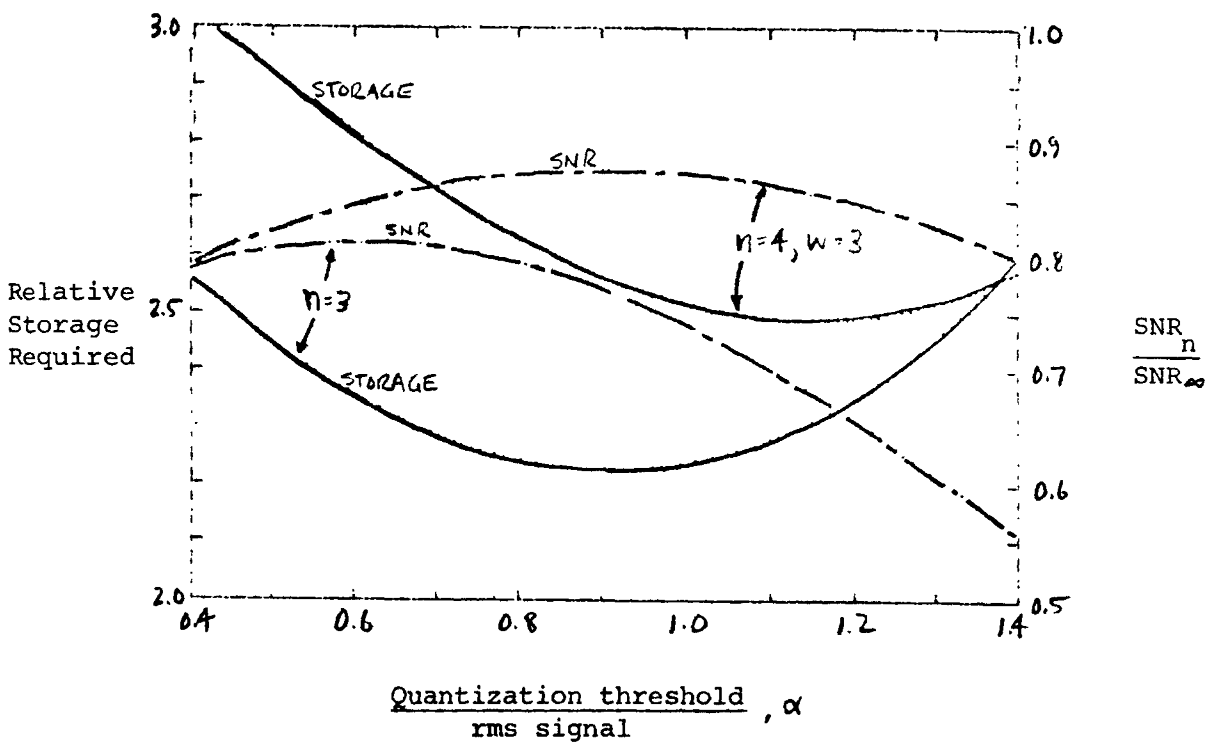


Figure 1: Signal to noise ratio and storage requirements vs. quantization threshold for 3-level and 4-level quantization.

These results can be seen in more detail in Table 1. The minimum storage required for 3-level quantization is about 12% less than for either the 2-level or 4-level scheme.

---

TABLE 1: Results

| | | | | | | |
|---|---|---|---|---|---|---|
| Quantization: | $n=$ | 2 | 3 | 3 | 4 | 4 |
| | $\alpha=$ | | .612 | .90 | .90 | 1.16 |
| Multiplier: | $W=$ | | | | 3 | 3 |
| # of products= | | 2 | 3 | 3 | 5 | 5 |
| Information/sample, bits: | $I_n=$ | 1 | 1.537 | 1.317 | 1.949 | 1.804 |
| $SNR_n/SNR_\infty$: | $R_n=$ | .637 | .810 | .769 | .872 | .852 |
| Relative storage, min. required: | | 2.468 | 2.343 | 2.225 | 2.561 | 2.486 |
| Relative storage, 1-sample codes: | | 2.47 | 2.36 | 2.31 | 2.63 | 2.76 |
| Bits/sample: | | 1 | 1.55 | 1.37 | 2 | 2 |
| Relative storage, fixed-length, multi-sample codes: | | 2.468 | 2.44 | 2.70 | 2.63 | 2.76 |
| Samples/word: | $L=$ | any | 5 | 5 | any | any |
| Bits/word: | | L | 8 | 8 | 2L | 2L |
| Variable-length, multi-sample codes: | | -----------not studied-------------- | | | | |

---

## Encoding the Samples

To achieve the minimum storage using the optimum quantization ($n=3$, $\alpha=.90$), the samples must be encoded so that the average sample occupies $I_3=1.317$ bits (see Table 1). Are there practical codes which approach this? Restricting attention to practical codes, could there be other quantizations (say, $n=4$) which require less storage?

First consider encoding one sample at a time. For n=2 the trivial code (1 bit per sample) is optimum. For n=3 the simplest code of 2 bits per sample (one unused code) is far from optimum. But for n=4, 2 bits per sample is the only reasonable possibility, and it is closer to the optimum of 1.804. A reasonable code for n=3 is the following:

| | | |
|---|---|---|
| sample=-1 | gives | code=10 (binary) |
| 0 | | 0 |
| +1 | | 11 |

This gives 1.370 bits per sample on the average. Such variable-length codes can be difficult to decode, since the decoder must figure out where each sample begins. But in this case it is easy: 2-bit codes always start with 1, and the 1-bit code is a 0. The storage requirements for these codes are given in Table 1; for n=4, with the code fixed at 2 bits/sample, the minimum storage is achieved for $\alpha$ = .90, since this gives best SNR. But n=3 and $\alpha$ = .90 requires about 14% less storage.

We can also consider codes which work on blocks of samples rather than one sample at a time. For n=3, fixed-length codes can achieve some advantage over 2 bits/sample; e.g., encoding every 5 samples in 8 bits gives 1.6 b/sample, with a small number of unused codes. For n=4, variable-length codes are needed. For sufficiently long block size, it should be possible to design decodable, variable-length codes which approach the optimum storage as closely as desired. This is achieved at the cost of complexity in the encoder and decoder. I have not studied this in any further detail.

## References

Cooper, B.F.C., 1970, "Correlators with two bit quantization." <u>Aust. J. Phys</u>., vol. 23, pp. 521-527.

Hagen, J.B. and D.T. Farley, 1973, "Digital correlation in radio science." <u>Radio Science</u>, vol. 8, pp. 775-784.

Weinreb, S., 1961, "Digital radiometer." <u>Proc. IEEE</u>, vol. 49, p. 1099.