FINITE SAMPLE CORRELATIONS OF QUANTIZED GAUSSIANS

By T. Cole*

[*Manuscript received February* 27, 1968]

Summary

A theory is developed for the treatment of quantized signals and for the estimate of correlation from quantized signals of only finite length. Different quantization systems can be compared on the accuracy of estimation of correlation for the same record length. For Gaussian statistics it is shown that with low values of correlation the efficiencies of one, two, and three bit quantizations are respectively 64, 89, and 95% with respect to continuous multiplication. For highly correlated signals the one bit system becomes the most efficient.

I. Introduction

In practical systems of correlation measurement one has only a finite record length of the random variables $X(t)$ and $Y(t)$. The normalized correlation ρ^T will hence be only an approximation to the true correlation ρ obtained from an infinite record length. If ergodicity is assumed, different records of the same length T will give ρ^T, which will fluctuate about ρ with some standard deviation σ_T. It is convenient then, to consider ρ as the "signal" and σ_T as the "noise" in a definition of output signal to noise ratio.

The signals $X(t)$ and $Y(t)$ would have to be sampled in time and quantized in amplitude if a computer were to be used to determine the correlation. We define $\{\hat{X}(i)\}$ and $\{\hat{Y}(i)\}$ as the quantized and sampled forms of $X(t)$ and $Y(t)$. The normalized and averaged product of $\hat{X}(i)$ and $\hat{Y}(i)$ over N samples is denoted by $\hat{\rho}^N$. However, ρ is not, in general, a linear function in $\hat{\rho}$. For the quantized signals the estimate of correlation is given in terms of the function A relating ρ to $\hat{\rho}$ as

$$\bar{\rho}^N = A(\hat{\rho}^N). \tag{1}$$

By a knowledge of the statistics of $\hat{X}(t)$ and $\hat{Y}(t)$ and hence those of $\hat{\rho}^N$, the output signal to noise for the record length N and quantization system chosen is given by $\rho/\bar{\sigma}_N$ where $\bar{\sigma}_N$ is the standard deviation of the value of $\bar{\rho}^N$ about ρ. The signals $X(t)$ and $Y(t)$ could be quantized in an infinite number of ways and each system would produce its own estimate of correlation and output signal to noise ratio. By defining efficiency as the decrease in output signal to noise of the estimate of ρ in one quantization system with respect to continuous multiplication for the same signals and record length, one can compare the various quantization systems.

It is well known that the output noise is inversely proportional to the square root of the record length, so a knowledge of the statistics, and hence efficiency, for only one sample length is sufficient for all record lengths.

* Observatoire de Meudon, Meudon 92, France; present address: Cavendish Laboratories, University of Cambridge, Cambridge, England.

Systems to which such a theory is applicable include the square-law detector ($X(t) = Y(t)$), the polarity coincidence correlator (one bit correlation), and the more complex forms of correlation detection. The problem of correlation of quantized signals but with infinite records has been treated before, and after a review of the quantization process the later sections develop an approach to the less elegant problem of finite record length. The results are from an application of this to four types of quantizations of Gaussian signals.

II. Quantization Process

Bonnet (1962) considers a general quantization function, but it is only necessary here to consider the one which is shown by him to maintain a zero mean and which gives a finite correlation output for very weak signals. The quantizing system consists of bands of width q such that, if $X(t)$ lies within this band, it takes on the value of the midpoint of the band. In general we consider the origin at a transition between bands and an equal number $(K-1)$ of bands on either side of the origin. That is, for integers K, k, and q,

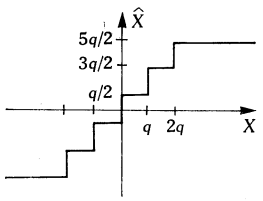

Fig. 1.—Finite quantization system with $K = 3$ and bands of width q in which X is transformed to $\hat{X}$.

$$\hat{X}(i) = \left.\begin{array}{ll} (-K+\tfrac{1}{2})q & \text{if } X(t) < (-K+1)q, \\ (k-\tfrac{1}{2})q & \text{if } (k-1)q \leqslant X(t) < kq, \\ (K-\tfrac{1}{2})q & \text{if } X(t) \geqslant (K-1)q, \end{array}\right\} \tag{2}$$

Figure 1 illustrates the case where $K = 3$. The above definition is general in that it includes infinite clipping ($K = 1$), infinite quantization ($K \to \infty$, q finite), the linear limiter ($q \to 0$, $K \to \infty$, Kq constant), and the identity ($Kq \to \infty$).

III. Finite Quantization

Given the statistical properties of the signals $X(t)$ and $Y(t)$, it is possible to give formulae for the estimate $\hat{\rho}^N$, the standard deviation $\hat{\sigma}^N$ of $\hat{\rho}^N$ about $\hat{\rho}$, and the function A relating $\hat{\rho}$ to ρ.

If $p(X)$ is the probability distribution (pdf) of X, we define a quantity P_j by

$$\left.\begin{array}{ll} P_K = \displaystyle\int_{(K-1)q}^{\infty} p(X)\,\mathrm{d}X, & \\ & (-K+2) \leqslant j \leqslant (K-1), \\ P_j = \displaystyle\int_{(j-1)q}^{jq} p(X)\,\mathrm{d}X, & \\ & j \text{ integer}. \\ P_{(-K+1)} = \displaystyle\int_{-\infty}^{(-K+1)q} p(X)\,\mathrm{d}X, & \end{array}\right\} \tag{3}$$

Then $p(\hat{X})$, the probability density distribution of $\hat{X}$, is

$$p(\hat{X}) = \sum_{-K+1}^{K} P_j\, \delta(\hat{X}-(j-\tfrac{1}{2})q)\,, \tag{4}$$

where $\delta(x)$ is the Dirac delta function.

In two dimensions consider functions X and Y with correlation coefficient ρ and joint pdf of $p(X, Y; \rho)$. If both X and Y undergo the same quantization to $\hat{X}$ and $\hat{Y}$ we can express the joint distribution of $\hat{X}$ and $\hat{Y}$ in terms of ${}^2P_{jk}$, where

$$\left.\begin{aligned} {}^2P_{jk} = \int_{a_j}^{b_j}\int_{c_k}^{d_k} p(X, Y; \rho)\, \mathrm{d}Y\, \mathrm{d}X\,, \qquad & (-K+1) \leqslant j, k \leqslant K\,, \\ & j, k \text{ integers.} \end{aligned}\right\} \tag{5}$$

The limits of integration are defined as:

$$a_j = \begin{cases} (j-1)q & \text{if } (-K+2) \leqslant j \leqslant K\,, \\ -\infty & \text{if } j = (-K+1)\,, \end{cases}$$

$$b_j = \begin{cases} \infty & \text{if } j = K\,, \\ jq & \text{if } (-K+1) \leqslant j \leqslant K-1\,, \end{cases}$$

$$c_k = \begin{cases} (k-1)q & \text{if } (-K+2) \leqslant k \leqslant K\,, \\ -\infty & \text{if } j = (-K+1)\,, \end{cases}$$

$$d_k = \begin{cases} \infty & \text{if } k = K\,, \\ kq & \text{if } (-K+1) \leqslant k \leqslant K-1\,. \end{cases}$$

The joint pdf is then $p(\hat{X}, \hat{Y}; \hat{\rho})$ and is given by

$$p(\hat{X}, \hat{Y}; \hat{\rho}) = \sum_{-K+1}^{K} \sum_{-K+1}^{K} {}^2P_{jk}\, {}^2\delta(\hat{X}-(j-\tfrac{1}{2})q, \hat{Y}-(k-\tfrac{1}{2})q)\,, \tag{6}$$

where ${}^2\delta(x, y)$ is the two-dimensional Dirac delta function. For the case of interest here, we desire an expression for the pdf of the random variable z equal to the product $\hat{X}\hat{Y}$. Evidently $p(z)$ is a discrete distribution given by

$$p(z) = \sum_{-K+1}^{K} \sum_{-K+1}^{K} {}^2P_{jk}\, \delta(z-(k-\tfrac{1}{2})(j-\tfrac{1}{2})q^2) \qquad j, k \text{ integers.} \tag{7}$$

The mean of z is

$$\bar{z} = \int_{-\infty}^{\infty} z\, p(z)\, \mathrm{d}z = \sum_j \sum_k {}^2P_{jk}\, (k-\tfrac{1}{2})(j-\tfrac{1}{2})q^2 \tag{8}$$

and the mean square is

$$\overline{z^2} = \int_{-\infty}^{\infty} z^2\, p(z)\, \mathrm{d}z\,. \tag{9}$$

When $\bar{z}$ is normalized to the value at $\rho = 1$, it is identical with $\hat{\rho}$ while $\hat{\sigma}^1$ is the standard deviation of $p(z)$ normalized in the same way, i.e.

$$\hat{\sigma}^1 = \left(\overline{z^2} - (\bar{z})^2\right)^{\frac{1}{2}} / (\bar{z})_{\rho=1}. \tag{10}$$

IV. Calculation

Another way of considering the transformation is via a plot of $p(X, Y; \rho)$ on the two-dimensional X, Y plane cut into squares of side Q. The distribution $p(\hat{X}, \hat{Y}; \hat{\rho})$ is a two-dimensional distribution of delta functions on the centres of the squares and with values ${}^2P_{ij}$, where ${}^2P_{ij}$ is the integral of $p(X, Y; \rho)$ over that particular square (see Fig. 2). That is, $p(\hat{X}, \hat{Y}; \hat{\rho})$ can be represented as the two-dimensional array of numbers $A(I, J) = {}^2P_{ij}$ in which an increment of unity in the index represents a shift of q on the $\hat{X}$ or $\hat{Y}$ axis. When the product $\hat{X}\hat{Y}$ is formed, a one-dimensional array of numbers $B(M)$ is obtained in which M represents multiples of $\frac{1}{4}Q^2$. The multiplication can thus be considered as a transformation on the indices rather than an operation on the actual values of P_{ij}.

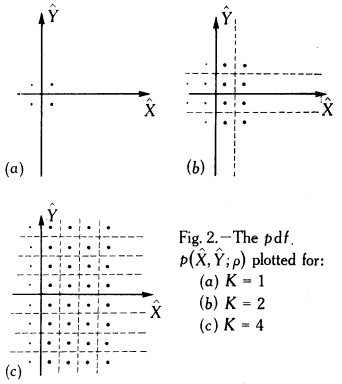

Fig. 2.—The pdf. $p(\hat{X}, \hat{Y}; \rho)$ plotted for: (a) $K = 1$ (b) $K = 2$ (c) $K = 4$

Henceforth consider a symmetrical pdf so that we need treat only the first and fourth quadrants, say. In the first quadrant let the array $A(I, J) = {}^2P_{ij}$ be the representation of $p(\hat{X}, \hat{Y}; \hat{\rho})$, where I and J are positive integers by definition. A symmetrical pdf implies that $p(\hat{X}, \hat{Y}; \hat{\rho})$ is symmetrical about the 45° axis and so, for the case $I \neq J$, XY takes the value $(I-\frac{1}{2})(J-\frac{1}{2})$ in four positions on the X, Y plane:

$$(I-\tfrac{1}{2})(J-\tfrac{1}{2}) = (J-\tfrac{1}{2})(I-\tfrac{1}{2}) = (-I+\tfrac{1}{2})(-J+\tfrac{1}{2}) = (-J+\tfrac{1}{2})(-I+\tfrac{1}{2}).$$

Therefore if $I \neq J$ and $I > J$, $B(M) = 4A(I, J)$, where $M = (2I-1)(2J-1)$ is a multiple of $\frac{1}{4}Q^2$.

On the 45° axis $I = J$ and there are only two positions where $\hat{X}\hat{Y}$ takes the value $(I-\frac{1}{2})^2$. Hence $B(M) = 2A(I, I)$ in the first quadrant. Thus the positive axis of $B(M)$ can be found:

$$B(M) = \sum_{I=1}^{K} \left(\sum_{J=1}^{I-1} 4A(I, J)\, \delta\big(M-(2I-1)(2J-1)\big) \right) + \sum_{I=1}^{K} 2A(I, I)\, \delta\big(M-(2I-1)^2\big). \tag{11}$$

In the fourth quadrant the same arguments apply and $p(\hat{X}, \hat{Y}; \hat{\rho})$ is represented by the array $C(N, L) = P_{n,-l+1}$ and N and L are integers $1 \ldots +K$. Taking the product $\hat{X}\hat{Y}$ gives the negative axis of $B(M)$, represented by $B(W)$,

$$B(W) = \sum_{N=1}^{K} \left(\sum_{L=1}^{N-1} 4C(N, L)\, \delta\big(W-(2N-1)(2L-1)\big) \right) + \sum_{N=1}^{K} 2C(N, N)\, \delta\big(W-(2N-1)^2\big). \tag{12}$$

Evidently the mean value of $\hat{X}\hat{Y}$ is

$$\bar{z} = \left(\sum_{M=1}^{\infty} M\,B(M)\right)\frac{Q^2}{4} - \left(\sum_{W=1}^{\infty} W\,B(W)\right)\frac{Q^2}{4} \tag{13}$$

and the mean square is

$$\overline{z^2} = \left(\sum_{M=1}^{\infty} M^2 B(M)\right)\frac{Q^4}{16} + \left(\sum_{W=1}^{\infty} W^2 B(W)\right)\frac{Q^4}{16}. \tag{14}$$

In general it is very difficult to determine the values of P_{ij}. Numerical integration seems the only solution. For example, if X and Y are Gaussian variables, the case $\rho = 1$ is the only one expressible in closed form (and then in terms of the error function). A computer programme can perform the operations specified by the above equations. It would give results enabling the function $\rho = A(\hat{\rho})$ and the value of $\hat{\sigma}^1$ to be plotted for various values of ρ and quantization systems.

V. Product of Two Gaussians

The primary purpose of the calculations is to compare the various quantization systems with the signal to noise ratio of continuous multiplication of two Gaussian signals of finite length. The pdf $p(z)$ has to be found where $z = XY$, the product of the two Gaussians X and Y, and where $p(z)\,\mathrm{d}z$ is the probability that z is between z and $z+\mathrm{d}z$. On a plot of the two-dimensional pdf, $p(x, y; \rho)$, $p(z)\,\mathrm{d}z$ is thus the integral of the areas between the curves $z = xy$ and $z+\mathrm{d}z = xy$, that is, the area C,

$$p(z)\,\mathrm{d}z = \iint_C p(x, y; \rho)\,\mathrm{d}x\,\mathrm{d}y\,. \tag{15}$$

The result is one that is well known in statistics and is given, for example, in more general form by Wishart and Bartlett (1932). In an information-theory context it is derived by Lampard (1956), and in the present notation it is the less general expression

$$p(z) = \frac{1}{\pi(1-\rho)^{\frac{1}{2}}}\exp\left(\frac{\rho z}{1-\rho^2}\right) K_0\left(\frac{z}{1-\rho^2}\right), \tag{16}$$

where $K_0(x)$ is the modified Bessel function of the first kind of zero order.

As $p(z)$ is a proper probability density, it normalizes and so, integrating (16) from $-\infty$ to $+\infty$,

$$1 = \int_{-\infty}^{\infty} p(z)\,\mathrm{d}z = \frac{1}{\pi(1-\rho)^{\frac{1}{2}}}\int_{-\infty}^{\infty} \exp\left(\frac{\rho z}{1-\rho^2}\right) K_0\left(\frac{z}{1-\rho^2}\right)\mathrm{d}z\,. \tag{17}$$

With the new variable $u = z/(1-\rho^2)$,

$$1 = \pi^{-1}(1-\rho^2)^{\frac{1}{2}}\int_{-\infty}^{\infty} \exp(\rho u)\,K_0(u)\,\mathrm{d}u\,, \tag{18}$$

and so

$$(1-\rho^2)^{-\frac{1}{2}} = \pi^{-1}\int_{-\infty}^{\infty} \exp(\rho u)\,K_0(u)\,\mathrm{d}u\,. \tag{19}$$

This can be differentiated on both sides with respect to ρ to give

$$-\tfrac{1}{2}(1-\rho^2)^{-3/2}(-2\rho) = \pi^{-1}\int_{-\infty}^{\infty} u\exp(\rho u)\,K_0(u)\,\mathrm{d}u\,, \tag{20}$$

which is just

$$\rho = \pi^{-1}(1-\rho^2)^{3/2}\int_{-\infty}^{\infty} u\exp(\rho u)\,K_0(u)\,\mathrm{d}u \tag{21}$$

$$= \int_{-\infty}^{\infty} z\,p(z)\,\mathrm{d}z\,. \tag{22}$$

Hence the mean of z is

$$\bar{z} = \rho\,. \tag{23}$$

This is also clear from first principles.

Equation (20) may be differentiated again to give

$$3\rho^2(1-\rho^2)^{-5/2}+(1-\rho^2)^{-3/2} = \pi^{-1}\int_{-\infty}^{\infty} u^2\exp(\rho u)\,K_0(u)\,\mathrm{d}u\,, \tag{24}$$

and so

$$1+2\rho^2 = \pi^{-1}(1-\rho^2)^{5/2}\int_{-\infty}^{\infty} u^2\exp(\rho u)\,K_0(u)\,\mathrm{d}u \tag{25}$$

$$= \int_{-\infty}^{\infty} z^2\,p(z)\,\mathrm{d}z\,. \tag{26}$$

Hence

$$\overline{z^2} = 1+2\rho^2 \tag{27}$$

and the variance is

$$\overline{z^2}-(\bar{z})^2 = (1+2\rho^2)-\rho^2 = 1+\rho^2\,. \tag{28}$$

The output noise to signal ratio for continuous multiplication is hence

$$\sigma/\rho = (1+\rho^2)^{\frac{1}{2}}/\rho\,. \tag{29}$$

Obviously the higher moments could be found by continuation of the differentiation process.

VI. Results

By using the formulae of Section IV, the output noise to signal ratio could be calculated for Gaussian statistics and for values of K from 1 to 4 with correlation coefficient ρ and quantization level Q (fraction of standard deviation) as parameters. The mean of $\hat{X}\hat{Y}$ represents $\hat{\rho}$ when it is normalized to the value for the case $\rho = 1$. Hence it is possible to trace curves of $\rho = A(\hat{\rho})$ for various K and Q as shown in Figure 3. As expected, for K increasing, the system behaves more closely in character to the continuous multiplication, while for a given value of K there is a value of Q at which the quantization best approximates the continuous case. $K = 1$ represents the infinite clipping or one bit case and the curve is the $(2/\pi)\sin^{-1}\rho$ curve given by Van Vleck (1943).

It is also possible to plot the ratio of $\hat{\sigma}^1/\hat{\rho}$ against ρ, K, and Q. Denoting this ratio by N/S, Figure 4 also shows that, for a given K and ρ, an optimum Q exists for least noise.

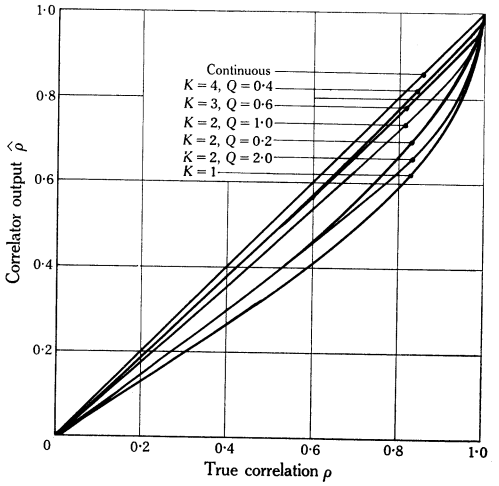

Fig. 3.—Plots of output correlation coefficient $\hat{\rho}$ against input ρ for various quantization systems and operating conditions.

To obtain the relative efficiencies of the quantization systems, the values of $\bar{\sigma}^1$ must be found through the use of the curves of $\hat{\rho}$ versus ρ. A quantized correlation measure on one sample would give an estimate $\bar{\rho}^1 = A(\hat{\rho}^1)$, and $\bar{\sigma}^1$ is the standard

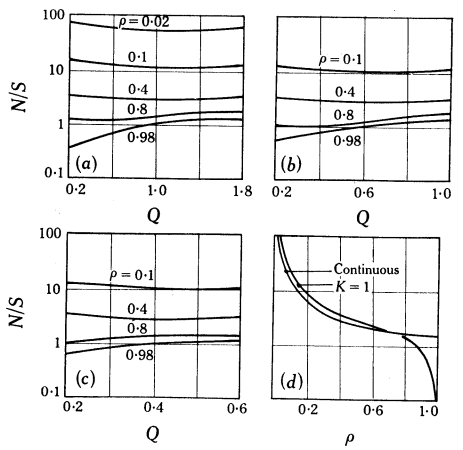

Fig. 4.—Plots of output noise to signal ratio N/S against step width Q (fraction of standard deviation) for various values of ρ and
(a) $K = 2$,
(b) $K = 3$,
(c) $K = 4$,
and against ρ for
(d) $K = 1$ and continuous.

deviation of $\bar{\rho}^1$ about ρ. The σ's are related since errors in the measurement of $\hat{\rho}$ and ρ are related; $\bar{\sigma}^1$ is merely the value $\hat{\sigma}^1$ divided by the slope of the appropriate curve in Figure 3 at the relevant value of ρ. The ratio $\bar{\sigma}^1/\rho$ is hence the noise to signal ratio of the estimate of correlation with that particular quantization system and record length. This ratio is plotted in Figure 5 as a function of ρ for various

quantization types. The following section interprets these curves in terms of the two practical uses: correlation and power measurement.

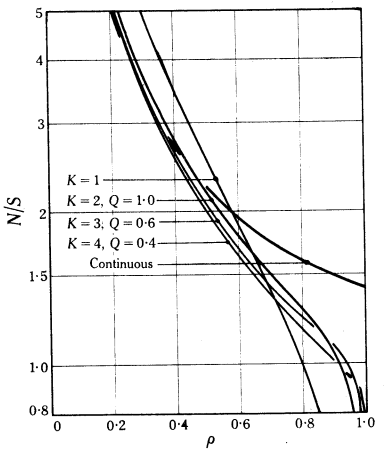

Fig. 5.—Output noise to signal ratio N/S plotted against ρ for various systems to indicate relative efficiencies.

VII. Practical Measurement

The determination of correlation between two slightly correlated signals is important in radio astronomy for the study of polarization and also spectral lines. In these cases the correlation is usually small and, as Figures 3, 4, and 5 indicate, linearity is approached. By normalizing the output signal to noise ratio of the various

Table 1

EFFICIENCIES RELATIVE TO A CONTINUOUS SYSTEM

ρ	Continuous	System $K = 1$	2 $Q = 1\cdot0$	3 $0\cdot6$	4 $0\cdot4$
0·1	100	64·2	88·5	93·6	94·6
0·2	100	67	90	95	96·5
0·3	100	71	93·5	99	100
0·4	100	77·5	98	103	106
0·5	100	87	103	109	112
0·6	100	102	110	115	120
0·7	100	121	113	119	125
0·8	100	168	125	131	138
0·9	100	282	142	138	148
0·95	100	480	160	151	155

systems to that of continuous multiplication at each value of ρ, Figure 5 is re-represented in Table 1 as a list of relative efficiencies. This table confirms the relative efficiencies of 0·64 for one bit and 0·88 for two bit as given by Weinreb (1963) and Clark (1967) (for his $n = 9$ and $m = 3$), in the limiting case where ρ tends to zero.

However, against these curves and efficiencies must be placed the practical complexities of implementing a quantization system. For a rapid many-point determination of a correlation function, many multipliers are needed and in this case the one bit system is simplest. The operations for a one bit system would be binary arithmetic so that the only source of error could be in determination of the point where the signals change sign. The electronics are simple and work rapidly (5 MHz), as has been shown, for example, by Weinreb (1963).

Calibration is simply a count of N, the record length. This gives the value which would be obtained in the case $\rho = 1$ and which enables an immediate determination of $\bar{\rho}^N$ as $\sin(\frac{1}{2}\pi\hat{\rho}^N)$. Systems of two, three, or higher bits involve several levels apart from zero that divide the amplitude of the signal into bands of width Q. It is evidently important to consider the effect of absolute level and stability of level on the system performance. Determination at the zero level of amplitude is extremely accurate by the use of amplifiers and limiters. The levels above and below zero must be generated in the quantizer and held stable. The theory in this paper uses levels as a fraction of the standard deviation of the signal, which is generally an unknown. In other words, the absolute level as a fraction of the standard deviation would not be known extremely accurately. The theory would indicate the region of maximum efficiency and only calibration could determine the actual function $\rho = A(\hat{\rho})$. Fortunately Figure 4 indicates that efficiency is a slowly varying function of the level Q.

Consequently, the important point with multiple bit systems is the stability of the quantizer. The quantizer as well as the level of the signal would need to be stable between calibrations to a level commensurate with the level of correlation coefficient being measured. Calibration would be more complicated as it not only requires a knowledge of the output for two completely correlated signals but another test signal is required to determine the level Q and hence the relation $\rho = A(\hat{\rho})$.

In other applications (particularly in radio astronomy for correlation receivers, for compound arrays, and for polarization measurements) the actual correlation coefficient between two signals is not required. In these cases the actual output of the correlator is used as an indication of power of the correlated component. Calibration is then achieved by insertion of a signal of known power.

The two applications are related in that, if σ_s^2 is the power of the correlated part and σ_n^2 the power in the uncorrelated part of each of the assumed equal channels, the value σ_s^2 can be expressed in terms of the correlation coefficient of the two signals as

$$\sigma_s^2 = \{\rho/(1-\rho)\}\sigma_n^2 . \tag{30}$$

For small excursions around the operating point, the relative efficiencies will be the same as before. Again the result emerges that for operating points at high values of ρ the one bit system is the most efficient.

VIII. Conclusions

This paper has presented an approach to the general problem of correlation measurements on quantized Gaussians of only finite length. The results for Gaussian

statistics and one, two, and three bit quantizers indicate that for ρ greater than $0 \cdot 3$, three bit is more efficient than continuous, while for ρ greater than $0 \cdot 6$ or so, one bit is the most efficient. For the other extreme of low values of ρ, one bit is only 65% efficient and a large increase in efficiency to 89% can be obtained by the small increase in complexity of two bit representation. The rewards of three bit representation are only another 6% greater than two bit.

In general the efficiencies and optimum operating positions are very much a function of the system parameters and one would need to know the approximate system parameters before any ideal operating conditions could be chosen.

IX. Acknowledgments

This work was performed while the author was under a Gledden Research Fellowship from the University of Western Australia. The author would like to thank Dr. E. J. Blum of the Observatory at Meudon for suggesting the topic, and the referee, who suggested the derivation in Section V and provided the author with references to the product of continuous Gaussians.

X. References

Bonnet, G. (1962).—*C. r. hebd. Séanc. Acad. Sci., Paris* **255**, 825–7.

Clark, B. G. (1967).—Two bit correlator. Rep. Natn. Radioastron. Obs., Greenbank, U.S.A. (unpublished).

Lampard, D. G. (1956).—*I.R.E. Trans. Inf. Theory* **IT2**, No. 1, 4–11.

Van Vleck, J. H. (1943).—Radio Res. Lab. Harvard Univ. Rep. No. 51.

Weinreb, S. (1963).—M.I.T. Res. Lab. of Electronics Tech. Rep. No. 412.

Wishart, J., and Bartlett, M. F. (1932).—*Proc. Camb. phil. Soc. math. phys. Sci.* **28**, 455.