Technical Report No. 32-772

Demonstration of a Quantile System for Compression of Data from Deep Space Probes

Tage O. Anderson

Isidore Eisenberger

Warren A. Lushbaugh

Edward C. Posner

Richard Goldstein

R. Goldstein, Manager
Communications Systems Research Section

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

December 1, 1965

CONTENTS

FIGURES

FIGURES (Cont'd)

ABSTRACT

$\cancel{14629}$

The theory and design of an advanced engineering prototype of a quantile system of data compression for space telemetry is described. The basic idea is to transmit a few *quantiles* (or percentage points, as they are sometimes called) of a histogram of experimental values formed aboard a spacecraft. Only these quantiles are transmitted to Earth, and yet a large part of the information that was contained in the original histogram can be reconstructed on Earth. Compression ratios on the order of 100 to 1 are obtainable at 100% efficiency with a simple device that performs no on-board arithmetic operations.

After summarizing the theoretical background governing the use of quantiles, what the design considerations of such a system should be are considered. Data compression ratios are computed for a typical application of the quantile system. The detailed description of the advanced engineering quantile system that has been built at JPL is discussed; in addition, a block diagram is given and analyzed. An important feature of the quantile system is its self-adaptive features; these are studied. Examples are given whereby quantile systems for compressing telemetry data can find application in civilian technology. Finally, experimental results using the constructed system are given.

Author

I. THE THEORETICAL BACKGROUND OF THE QUANTILER

In circumstances where the locality of experiments does not coincide with the locality of the statistical computations performed on the experimental data, the sample data must be transmitted through a communications channel from the point of origin to the processing center. If many experiments are being performed simultaneously and the same communications channel is used to transmit the sample values of each experiment, there is a limit to the total number of observations that can be sent in a given time. This limits the number of experi-

ments that can be performed for given sample sizes. One way of effecting "data compression" is by transmitting a small number of sample quantiles, instead of all the sample values, and extracting the desired statistical information from these statistics. Following is a brief discussion of the use of sample quantiles for this purpose.

Let $x_1, x_2, \ldots, x_n$ be n independent sample values taken from a population with density function $g(x)$ and distribution function $G(x)$. If $g(x)$ is continuous, then ζ_p

is said to be the (population) quantile of order p, or the pth quantile of the distribution, if ζ_p is the (unique) solution of the equation $G(\zeta) = p$. Equivalently,

$$p = \int_{-\infty}^{\zeta_p} g(x)\,dx, \; (0 < p < 1)$$

Similarly, a sample quantile of order p, Z_p can be defined as follows: Arrange the observations in ascending order of magnitude

$$x_{(1)} \leq x_{(2)} \leq, \ldots, \leq x_{(n)}$$

Then

$$Z_p = x_{([np]+1)}$$

where $[np]$ denotes the greatest integer $\leq np$. This means that $100p\%$ of the sample values are smaller than Z_p. For example, $Z_{1/2}$ is well-known as the *median* of the sample distribution.

Now the sample quantiles, being random quantities, have probability distributions of their own. Moreover, subject to the condition that the parent density function possesses a continuous derivative in some neighborhood of each quantile value considered, the joint distribution of any (fixed) number of quantiles has the useful property of approaching the multivariate normal distribution as the sample size $n \to \infty$. The mean and variance of the limiting distribution of Z_p are given by

$$E(Z_p) = \zeta_p$$

$$Var(Z_p) = \frac{p(1-p)}{ng^2(\zeta_p)}$$

and the correlation ρ_{12} between Z_{p_1} and Z_{p_2}, $p_1 < p_2$, is

$$\rho_{12} = \left[\frac{p_1(1-p_2)}{p_2(1-p_1)}\right]^{1/2}$$

Thus for a sufficiently large sample size, if the limiting normal distribution of the sample quantiles is assumed when statistical analyses are based on sample quantile values, the error involved in making this assumption will be small.

The importance of being able to assume the normality of the sample quantiles lies in the fact that normal distributions are amenable to many types of statistical analyses. One such analysis is that necessary to determine the relative reliability of statistical information obtained using quantiles when compared to the same information obtained from the entire sample.

For example, estimates of the parameters of a normal population have been obtained using quantiles. A measure of the reliability of such an estimate is its *variance*. The relative reliability, called the efficiency of the estimate, is then defined to be the ratio of the variance of the best estimate using the entire sample to that of the estimate using quantiles.

Another type of statistical information obtained using quantiles is concerned with *tests of simple hypotheses*. Here a measure of reliability is the *power* of the test; that is, the probability of rejecting the *null hypothesis* when it is false. The relative reliability, also called the efficiency of the test, is defined to be the ratio of the power of the test using quantiles to that of the best test using all the sample values. The task of determining the efficiencies of the estimators and tests is simplified considerably by the assumption of the normality of the quantiles.

Such use of a small number of quantiles to obtain statistical information means that only a few observations need be transmitted instead of all the sample values. The result is a high *data compression ratio*. This compression ratio would be of little value, however, if the relative reliability of the information thus acquired were proportional to the ratio of the number of quantiles used to the total number of observations. That the relative reliability is not proportional to this ratio but depends only upon the number of quantiles used and how they are chosen, accounts for the advocacy of the use of quantiles for data compression.

The primary function of a "quantiler" is to choose a number of quantiles of specified orders from a set of samples and transmit these values to Earth where the desired information can then be extracted. Since for a fixed number of quantiles the relative reliability of a given type of statistical information depends critically upon the choice of the quantiles used, it is natural that one should wish to specify the orders of the quantiles which maximize the relative reliability. However, there are certain restrictions in making this optimal choice.

First of all, the orders of the quantiles are to be specified in advance and remain fixed throughout the flight. Secondly, the orders of the quantiles which maximize the relative reliability of one type of information are not

2

necessarily those which maximize the relative reliability of another type. Hence, in order to use the same set of quantiles for all types of statistical information that may be desired and which can be obtained with a high relative reliability using a small number of quantiles, a compromise must be made in the choice of their orders. A brief theoretical discussion will be given relative to the optimal choice of quantiles for estimating the mean μ and standard deviation σ of a normal parent population, which will illustrate the necessity of a compromise as well as propose one. Further details are found in Ref. 1.

Assuming that both the mean μ and standard deviation σ are unknown, an estimator of μ using $k \geq 2$ quantiles of distinct arbitrary orders of the form

$$\hat{\mu} = \sum_{i=1}^{k} C_i Z_i$$

is *unbiased* (has expected value μ) if the C_i are determined such that

$$\sum_{i=1}^{k} C_i = 1$$

and

$$\sum_{i=1}^{k} C_i \zeta_i^* = 0$$

where each Z_i is of order p_i and ζ_i^* is the population quantile of order p_i of the standard normal distribution. For, taking the expected value of $\hat{\mu}$, one has

$$E(\hat{\mu}) = \sum_{i=1}^{k} C_i \zeta_i = \sum_{i=1}^{k} C_i (\mu + \sigma \zeta_i^*)$$

$$= \mu + \sigma \sum_{i=1}^{k} C_i \zeta_i^* = \mu$$

It is obvious that for $k = 2$, unless Z_1 and Z_2 are symmetric quantiles (that is, unless $p_2 = 1 - p_1$, implying that $\zeta_1^* + \zeta_2^* = 0$), $\hat{\mu}$ will not be unbiased for any choice of C_1 and C_2. Moreover, for any even k, it has been shown that, in order for $\hat{\mu}$ to have minimum variance,

pairs of symmetric quantiles should be used. Under this restriction, it is easily seen that the estimate of μ of the form

$$\hat{\mu} = \sum_{i=1}^{\frac{k}{2}} \alpha_i (Z_i + Z_{k-i+1}), \qquad p_i + p_{k-i+1} = 1$$

will be unbiased if $\alpha_i > 0$, $i = 1, 2, \ldots, k/2$ and

$$\sum_{i=1}^{\frac{k}{2}} \alpha_i = \frac{1}{2}$$

By the method of maximum likelihood the coefficients α_i can be determined to provide an unbiased estimate of μ with minimum variance for $k/2$ arbitrarily chosen pairs of symmetric quantiles. Finally, by varying the p_i in the expression for the variance of $\hat{\mu}$, the orders of the particular set of $k/2$ pairs of symmetric quantiles as well as the coefficients α_i for which $\mathrm{Var}(\hat{\mu})$ is a minimum can be determined.

By using the same technique, unbiased estimators of σ from $k/2$ pairs of symmetric quantiles can be obtained, which are of the form

$$\hat{\sigma} = \sum_{i=1}^{\frac{k}{2}} \frac{\beta_i}{\zeta_{k-i+1}^*} (Z_{k-i+1} - Z_i)$$

where $\beta_i > 0$,

$$\sum_{i=1}^{\frac{k}{2}} \beta_i = \frac{1}{2}$$

For this case

$$E(\hat{\sigma}) = \sum_{i=1}^{\frac{k}{2}} \frac{\beta_i}{\zeta_{k-i+1}^*} (\mu + \sigma \zeta_{k-i+1}^* - \mu - \sigma \zeta_i^*)$$

$$= 2\sigma \sum_{i=1}^{\frac{k}{2}} \beta_i = \sigma$$

since $\zeta_i^* = -\zeta_{k-i+1}^*$

By again varying the p_i in the expression for the variance of the maximum-likelihood estimator of σ, the values of the p_i and β_i that minimize $Var(\hat{\sigma})$ can be determined. Only symmetric quantiles have been used to estimate σ (it has not yet been proved, but only conjectured, that this is the optimum procedure to adopt). We can say at present about the $\hat{\sigma}$ obtained in this manner that this $\hat{\sigma}$ is the maximum-likelihood minimum-variance unbiased estimator of σ using $k/2$ pairs of symmetric quantiles.

Further mathematical details are given in Ref. 1, where estimators of μ for $k = 1, 2, 3, 4, 6, \ldots, 20$ and estimators of σ for $k = 2, 4, 6, \ldots, 20$ are also given, having been derived by the method described above. The efficiency (relative reliability) of $\hat{\mu}$, as defined previously, increases from 0.637 for $k = 1$ to 0.994 for $k = 20$. The efficiency of $\hat{\sigma}$ increases from 0.612 for $k = 2$ to 0.984 for $k = 20$. Since μ is assumed unknown, an estimator for σ cannot be obtained using one quantile. The median is the only possible estimator of μ using one quantile if σ is unknown, since $\zeta^* \neq 0$ for values of p other than $p = 0.5$. In fact, the median is the minimum-variance estimator of μ, even if σ is known. If μ is known, however, an estimator of σ can be had using one quantile. Both $Z(0.9424)$, the quantile of order 0.9424, and $Z(0.0576)$, the quantile of order 0.0576, provide unbiased minimum-variance estimators, which are given by

$$\hat{\sigma_1} = 0.635\ [Z(0.9424) - \mu]$$

$$\hat{\sigma_2} = -0.635\ [Z(0.0576) - \mu]$$

The efficiency of both estimators is 0.304.

The attempt to achieve a high data compression ratio by the use of quantiles is encouraged by the fact that the best estimators of μ and σ, using as few as four quantiles, attain reasonably high efficiencies. With four quantiles, $Eff(\hat{\mu}) = 0.920$, $Eff(\hat{\sigma}) = 0.824$ and the estimators are given by

$$\hat{\mu} = 0.192\ [Z(0.1068) + Z(0.8932)] + 0.308\ [Z(0.3512)$$
$$+ Z(0.6488)]$$

$$\hat{\sigma} = 0.116\ [Z(0.9770) - Z(0.0230)] + 0.236\ [Z(0.8729)$$
$$- Z(0.1271)]$$

It can be seen, however, that the orders of the optimum quantiles for estimating μ are quite different from those for estimating σ, and similar disparities are apparent for all values of k. It is true that for small deviations from the optimum values, the loss in efficiency of either estimator will be insignificant; this fortunate occurrence is a consequence of the flatness of the surface representing $Var(\hat{\mu})$ and $Var(\hat{\sigma})$ about the optimum quantile values.

However, a more or less random choice of quantiles can result in a serious reduction in efficiency. For example, if the optimum four quantiles for estimating σ are used to estimate μ, the efficiency of $\hat{\mu}$ is reduced to 0.732, a loss in efficiency of 25.7%. And if the optimum four quantiles for estimating μ are used to estimate σ, the efficiency of $\hat{\sigma}$ is reduced to 0.665, a loss of 19.2%. The problem, then, becomes one of establishing a criterion of optimality on the basis of which suboptimum sets of k quantiles could be determined that would provide unbiased estimators of both μ and σ, using the same quantiles.

The decision was made to use as this criterion the linear combination

$$Var(\hat{u}) + b\ Var(\hat{\sigma}), \qquad b = 1, 2, 3, \ldots$$

and the set of $k/2$ pairs of symmetric quantiles which minimized this linear combination was defined as being suboptimum. For $b = 1, 2, 3$ and for $k = 2, 4, 6, \ldots, 20$, the suboptimum sets of quantiles, in the above sense, were determined, the estimators constructed, and the efficiencies computed. Most of these results are given in Ref. 1.

The results for $k = 4$ are as follows:

For $b = 1$,

$$\hat{\mu} = 0.141\ [Z(0.0668) + Z(0.9332)] + 0.359\ [Z(0.2912)$$
$$+ Z(0.7088)]$$

$$\hat{\sigma} = 0.258\ [Z(0.9332) - Z(0.0668)] + 0.205\ [Z(0.7088)$$
$$- Z(0.2912)]$$

$$Eff(\hat{\mu}) = 0.908, \qquad Eff(\hat{\sigma}) = 0.735$$

For $b = 2$,

$$\hat{\mu} = 0.106\ [Z(0.0434) + Z(0.9566)] + 0.394\ [Z(0.2381)$$
$$+ Z(0.7619)]$$

$\hat{\sigma} = 0.196\ [Z(0.9566) - Z(0.0434)] + 0.232\ [Z(0.7619) - Z(0.2381)]$

$$Eff(\hat{\mu}) = 0.876, \qquad Eff(\hat{\sigma}) = 0.779.$$

For $b = 3$,

$\hat{\mu} = 0.097\ [Z(0.0389) + Z(0.9611)] + 0.403\ [Z(0.2160) + Z(0.7840)]$

$\hat{\sigma} = 0.179\ [Z(0.9611) - Z(0.0389)] + 0.235\ [Z(0.7890) - Z(0.2160)]$

$$Eff(\hat{\mu}) = 0.857, \qquad Eff(\hat{\sigma}) = 0.792.$$

It is readily seen that as b increases, $Eff(\hat{\mu})$ decreases and $Eff(\hat{\sigma})$ increases. Since larger values of b mean that greater weight is being given to $Var(\hat{\sigma})$ this is not a surprising result. The point is, however, that the value of b should logically be determined by the relative importance of the two estimators. If $\hat{\sigma}$ is the important consideration, increasing the value of b gives different suboptimum sets of quantiles for which the efficiency of $\hat{\sigma}$ is improved. If $\hat{\mu}$ is of paramount importance, b should be small. It should be noted, perhaps, that five choices of b are at present available (including $b = 0$ and, in effect, $b = \infty$).

The need for suboptimum estimators arises from the desire to achieve maximum data compression. However, for an entirely different reason one may also prefer suffering a loss in efficiency rather than using optimum quantiles. For increasing values of k, the two optimum extreme quantile values, those of order p_1 and $p_k = 1 - p_1$, move farther out on the tails of the distribution. Although from a theoretical viewpoint this fact is of little consequence, from practical considerations there are two major objections to this behavior of p_1 and p_k. First, since n is never infinite, the true distributions of the sample quantiles are only approximately normal and, more important, the deviation from normality becomes more pronounced the farther the quantiles move out on the tails of the distribution. Second, the "normal" distributions that one encounters in practical situations are very often only approximately normal with deviations from normality greater out on the tails than toward the center of the distribution. Thus it is important on both counts to investigate the effect on the efficiency of the estimators when

optimum and suboptimum estimates of μ and σ are obtained when p_1 is restricted to be not less than some specified value. If the loss in efficiency is not excessive, it may well prove advantageous to adopt the cautious policy of restricting the value of p_1, and thus avoid or limit a bias in the estimates of μ and σ. This loss can be due either to a sufficiently large deviation from the assumed normality of the distribution of the extreme quantiles or to the erratic behavior out on the tail of an approximately normal parent distribution.

Optimum and suboptimum estimators of μ and σ were constructed when p_1 was restricted to be not less than 0.01 and again for p_1 not less than 0.025. The results were quite satisfactory in that losses in efficiency were small. Although the restrictions on p_1 affect the efficiencies of $\hat{\sigma}$ to a greater extent than those of $\hat{\mu}$, nevertheless efficiencies greater than 0.90 can be achieved for suboptimum estimators of σ for $k \geq 10$ when p_1 is restricted to not less than 0.025. In fact, for this case very little is gained by using more than 10 quantiles!

A type of statistical analysis closely related to the estimation of parameters is the testing of hypotheses concerning the values of the parameters. For example, one may wish to test the simple hypothesis that a normal population with known mean has a variance of σ_1^2 against the alternative hypothesis that the variance is σ_2^2. A number of tests of this nature have been devised using up to four quantiles. As in the case of estimators, one naturally seeks to maximize the efficiency of the test (relative reliability) as defined above. As one might expect, it turns out that for the tests which are concerned with the value of the mean, the optimum quantiles are those which maximize $Eff(\hat{\mu})$, and for the tests which are concerned with the value of the variance, the optimum quantiles are those which maximize $Eff(\hat{\sigma})$. Consequently, not only have optimum test statistics been devised for these tests, but suboptimum test statistics have also been devised using the suboptimum quantiles for estimating μ and σ. The mathematical and statistical details are given in a forthcoming Technical Report to be published by the Jet Propulsion Laboratory. Also included in the forthcoming report are quantile estimators of the degree of correlation between two normal populations.

Thus, it can be seen from the above discussion that by using quantiles to obtain certain types of statistical information, not only can a significant amount of data compression be achieved, but equally important, the reduction in uncertainty which accompanies a large sample size is also retained. The investigation into fur-

ther statistical uses of quantiles is being continued at the Jet Propulsion Laboratory. We shall describe, hereafter, a system which allows the quantiles to be found on board a spacecraft. The simplicity of this system, coupled with the large data compression ratio achievable, makes the use of the Quantiler appealing.

II. DATA COMPRESSION RATIOS

For quantile data to be a useful form of data compression, a Quantiler must be simple to construct. The design of the Quantiler introduced below is one of extreme simplicity; notably, no arithmetic operations are performed. As discussed in Section I, for quantile data to be useful, the loss of information that is involved in using quantiles instead of the entire sample must be considered in relation to the degree of data compression that can be achieved. The theory will be discussed from a practical point of view, as it applies to the design of the Quantiler.

A particle count experiment, the outcome of which has a *discrete distribution*, or *empirical distribution* or *histogram*, is used as an example. Before the quantiles are computed, the data are grouped into a familiar histogram. Quantiles of empirical distributions are called sample quantiles. Figure 1 shows a typical histogram with four such sample quantiles of orders 0.067, 0.291, 0.709, 0.933, for 1024 samples.

The sample quantiles are determined by first deciding how many quantiles are to be computed and transmitted. As few as four sample quantiles, as we have seen, convey a great deal of useful information. The answer to the statistical questions "Which quantiles should they be?" and "How good is the information which is obtained from the four quantiles relative to that obtained from the entire sample?" depends on what one wishes to compute.

It should be noted that transmitting quantiles is better than deriving the mean and standard deviation at the test site and then transmitting these moments, even ignoring the greater design simplicity of a quantile system over an arithmetic system. The knowledge of the quantiles affords additional knowledge of the distribution. For example, the distribution might be bimodal (i.e., have two local maxima) rather than being unimodal. Such a condition would be detected using the quantile method but not using the simple mean and standard deviation data. Furthermore, goodness-of-fit tests have been devised for use with quantiles. These are tests which tell an experimenter whether it is reasonable to assume that a given set of experimental data arises from a supposed probability law. This is an important type of statistical test in all phases of experimental work. Discriminating between two similar distributions using quantiles extends the applicability of quantiles to data compression still further. Further mathematical details will be found in Ref. 1.

The data compression ratios that can be achieved using four quantiles are demonstrated by the following example: Say the 1024 samples are particle counts from 1024 sampling periods. The spread between samples is assumed to be within 128; that is, no more than 127 particles arrive in any one second. A sample size of 1024 is chosen, since it is in the interest of the best estimates of the mean and standard deviation to minimize the variance of the estimates, the variance being inversely proportional to the total number of samples. In fact, the total number of samples should be as large as possible, subject to the incoming data remaining stationary. The sample spread should also be large in order to be able to handle maximum variability in the data rate. Both of these parameters must be limited, however, to hold
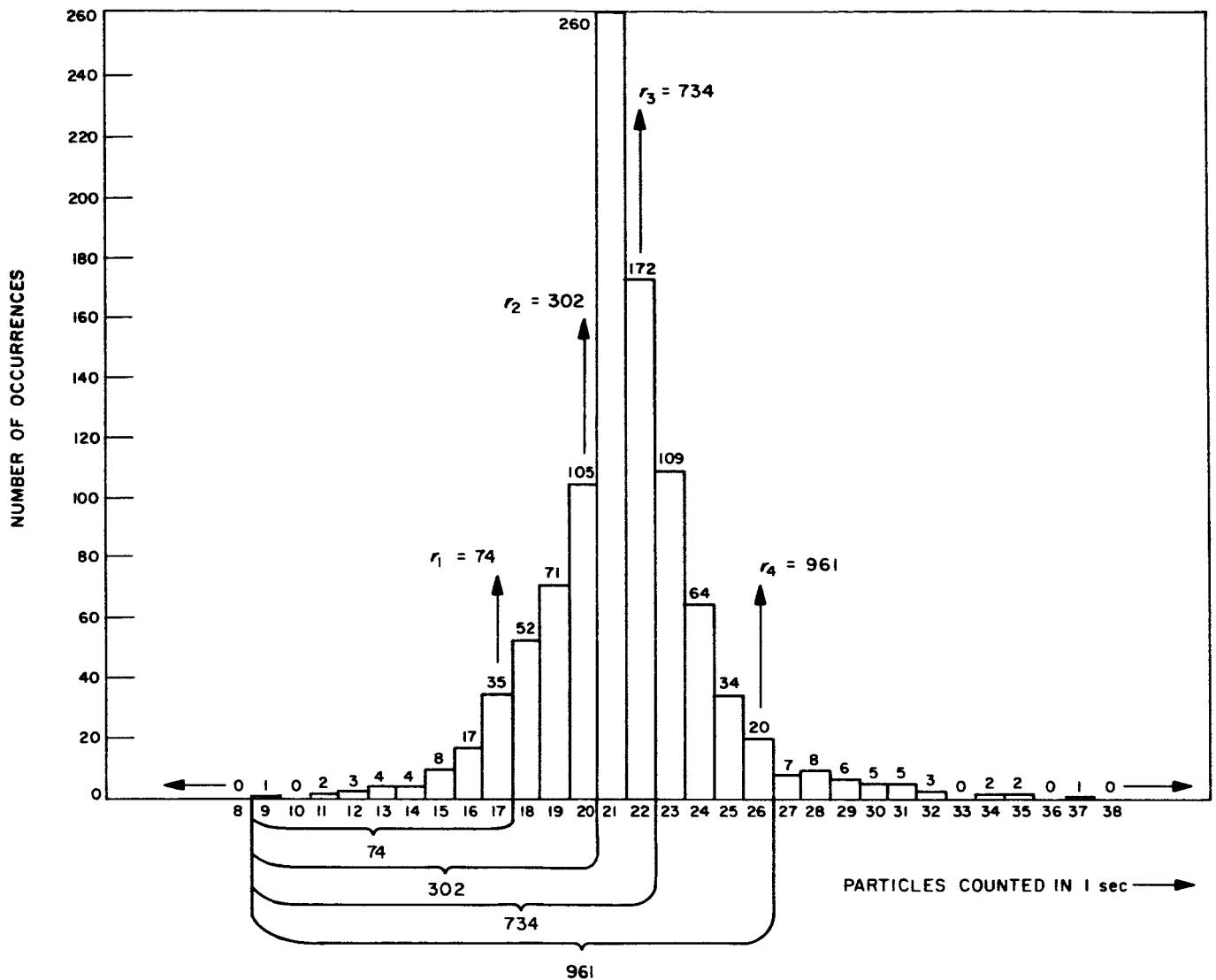
Fig. 1. A typical histogram

down the volume of hardware required. As for the sample spread, additional hardware can more efficiently be applied to vary the sampling time, instead of providing extra storage capacity. This self-adaptive feature will be discussed later.

In comparing the transmission of all samples (7 × 1024 bits) with the transmission of four quantiles (28 bits) per sampling cycle, a compression ratio of 250:1 results. However, consideration of the effect of bit errors for the two types of data gathering and transmission methods leads one to conclude that the signal-to-noise ratio for the quantile transmission should be

increased over the raw-data system. Doubling of the signal-to-noise ratio, for example, is equivalent to doubling the bit time with constant power; the compression ratio would thus be reduced to 125:1. A further reduction is demanded by the quantile estimation efficiency. With a typical efficiency of, for example, 80%, the final data compression ratio is found to be in the order of 100:1. By transmitting 1/100 as much data, estimates of the mean and standard deviation of the parent population can thus be obtained with variances no higher than if the full sample of uncompressed data were used. This means that, in effect, 100% efficiency is obtained with data compression ratios of 100:1.

III. DETAILED DESCRIPTION OF THE QUANTILER

One important novelty of the Quantiler design is to be found in the method by which the histogram is formed. Other operational functions will be described in reference to the Quantiler functional block diagram (Fig. 2), and also the detailed mechanization of the various control functions.

A. Histogram Storage Capacity with Relocation Capability

A Quantiler that generates the quantiles for a total of 1024 samples with a sample range of 256 is described in the remainder of this report. In order to cover all possible sample distributions, 8192 individual storage points would be required to store a histogram. With the histogram-forming method used by the Quantiler, however, only 1280 storage elements are needed: 1024 elements represent the actual samples, and 256 elements distinguish the 256 sample sizes.

In the prototype Quantiler, a 1280-bit recycling delay line is used as the memory (see Ref. 2). The line con-

tains at all times 256 *markers* and 1024 *spaces*. The spaces between two markers represent the number of samples of the size indicated by the order of the markers. This histogram-forming technique may better be understood perhaps from the following brief explanation of how the serial memory is loaded.

All markers are first arranged adjacent to each other at one end of the line; the line is then made to recycle. An address counter then counts the markers. As a sample has been taken, it is compared with the count in the address counter, and upon coincidence, a space is entered between the markers at that time. Thus, the number of spaces that follow a certain marker indicates the number of samples of the size equivalent to the address of that marker. When the first marker is adjacent to the last, all samples have been cumulated and a sampling cycle has been completed. At this point the quantiles are released and the line is reloaded by again placing the markers adjacent to one another at the beginning of the line. The detail mechanization of this reloading function will be discussed later.
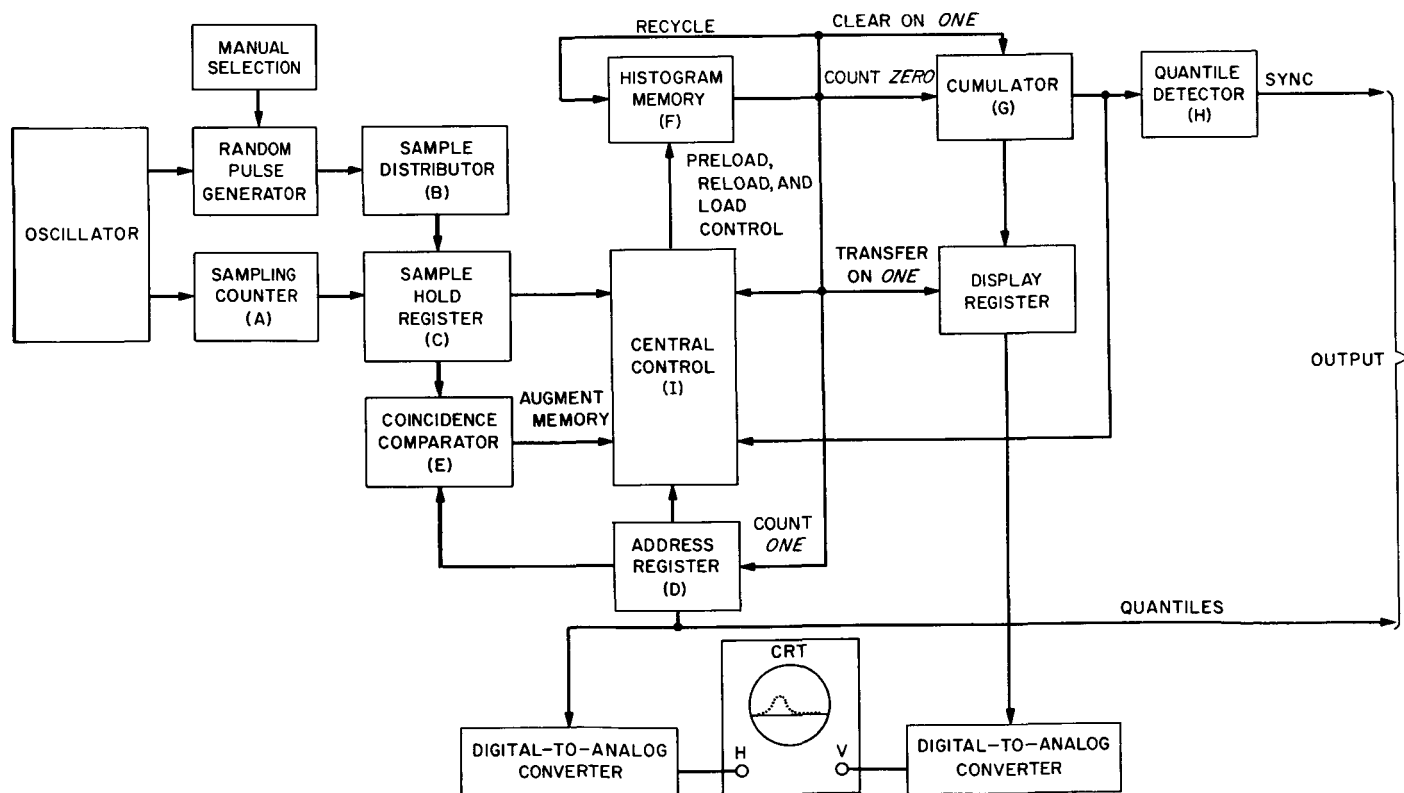

Fig. 2. The quantiler block diagram

Another feature of the Quantiler is the automatic control of the sampling rate. This automatic control of the sampling rate as a result of the incoming data rate can, in fact, be considered a secondary data compression function. With each change of the sampling rate, the memory line is synchronized. No extra data need be transmitted to indicate the sampling rate, since this information is inherent in the rate at which the quantiles are transmitted.

B. Block Diagram Analysis

Following the above brief outline of the most outstanding features of the Quantiler, an explanation of the joint operation of all the function blocks shown in Fig. 2 is needed.

Except for the sampling time controller which will be described separately (though part of the Quantiler) the Quantiler consists basically of three counters and a programmer or sequencer. This sequencer is, in effect, also a counter but of short length. These counters then are labeled:

1. The sample distributor.

2. The address register.

3. The cumulator.

4. The mode control.

With a sample spread of 256, 8 bits are required in the sample distributor. Eight bits are then required for the address register. With a sample size of 1024 samples, 10 bits are required for the cumulator. Three bits are used in the mode control to provide for six modes of operation (two positions are not used).

An intermediate hold or copy register is required for the sample distributor for the case in which all data must be sampled during a certain time. Because of the added display feature in the experimental model of the Quantiler, hold registers are also required for the address register and the cumulator. The value of the address register and the cumulator must be converted to an analogue voltage for display purposes. The two counters should then operate with a weighted code. Straight binary was selected. When all three counters require copy registers, the counters can readily be implemented as binary ripple counters. The ambiguous condition of the counter during the ripple period is then bypassed by the fact that both the input and the copy command are synchronized to the system clock. The three counters are utilized during the various modes of operation for many different purposes; that is, a great deal of time sharing of equipment is employed.

C. Modes of Operation

With the basic Quantiler consisting of three binary ripple counters and associated copy registers, the control of the entire operation (the mode control) simply amounts to control of the three signals pertinent to each of the counter-register sets. These signals are: (1) the input, (2) the copy command, and (3) the reset.

The modes of operation to be controlled by the mode control counter are:

1. M_1: preload of ones or markers.

2. M_2: preload of zeros or spaces.

3. M_3: load samples (the moving of a marker one space per sample).

4. M_4: histogram display.

5. M_5: cumulative function display.

6. M_6: quantile numerical display.

Depending on the particular mode at the time, the input to the mode control counter is either the address register full-scale signal or the cumulator full-scale signal. These full-scale conditions are not derived from detector gates. Since their operators are transitions rather than static conditions (the signals are used as clock for the mode control counter), the full-scale signals are derived as the most significant bits of the two counters change from ones to zeros.

The following paragraphs describe in detail the purpose of each individual mode of operation, as well as the utilization of each counter in that mode and also the control of the three pertinent inputs: (1) the clock, (2) the transfer command, and (3) the reset.

D. M_1, Preload of "Markers"

At the very beginning of a sample cycle and upon completion of a regular load operation and extraction of the quantile data or at the change of the sample rate, the serial memory must be prepared for the subsequent load cycle (all the markers must again be placed at one end of the memory line and all remaining spaces cleared).

During M_1, *ones* are entered into the serial memory. They are tallied by the address register by the fact that

the system clock is the input to the address register during M_1. System clock is also the copy command for the copy register at that time.

The address register is automatically cleared by the fact that the transition from M_6 to M_1 occurs on address full-scale; the mode control counter assumes its M_1 condition upon the transition from address register full-scale to the all-zero position.

Should a reload command occur (this could happen upon an automatic change of sample rate), the mode control would go from M_3, the only mode in which such a signal can occur, to M_1. The reload signal also constitutes a reset bus for the address register. No other units are involved in the preloading of the markers.

E. M_2, Preload of "Spaces"

During M_2, the input to the serial memory is inhibited; this is equivalent to the entering of zeros or clearing of the delay line. The number of spaces to be cleared or the length of time the input is to be inhibited is tallied by the cumulator. The input to the cumulator in mode 2 is then the system clock (or the frequency to which the delay line is tuned).

The cumulator starts from the all-zero condition. It is reset by a reset bus which is energized for one clock period (the m_1 address full-scale clock period). The m_1 ADDR FS signal is the signal used when power is first applied. This function is then part of the continuous resynchronization feature of the Quantiler.

During mode 2, transfer from the cumulator to the cumulator-hold register occurs upon system clock. In M_2 the cumulator full-scale is the clock for the mode control to advance to M_3. The output of the most significant bit of the cumulator-hold register is used as that clock. No other basic functions are involved in mode 2.

F. M_3, The Regular Load Operation

During M_3 the address register is counting the ones or markers in the recycling memory. When the count coincides with the count in the distributor-hold register, an extra space is inserted in the memory at that time. In the regular load operation, the markers are then spread apart throughout this line.

As stated earlier when introducing this histogram-forming technique, the distance (number of zeros) between markers on a line is then equal to the number of

samples of the size indicated by the order of the preceding *ones* loaded into the memory. Though many coincidences may occur as the memory recycles many times between transfers from the sample distributor to its hold register, only one coincidence per transfer is permitted to be operative. As shown in Fig. 3, this action is controlled by control bit c_8. The c_8 flip-flop sets on a transfer signal and resets on the next coincidence signal.[1]

Part of the forced sync arrangement used in this design is the detection of the "load complete" or "end-of-sample cycle" event. For this purpose, the cumulator in mode 3 is simply to tally the sample or transfer pulses, and the cumulator full-scale signal is the load-complete signal which will augment the mode control counter.

In the actual load operation another two control bits are involved. One is c_{11}, the extension of the memory line, and the other is c_{10}, which controls this extension. The detail mechanism is as follows: Upon coincidence between the distributor and the address register, the output of the memory line is stored in c_{10} and the regular input inhibited for one clock period. Upon the next clock pulse, c_{10} is set and the content of c_{11} is entered into the memory. The memory will then recycle through c_{10} until the last marker has been tallied.

G. M_4, Display of the Histogram

Mode 4 is one of three display modes. The quantile data present in the memory are displayed on a CRT in three different forms, the histogram form being the first. After the memory load operation has been completed, it will continuously recycle. In the three read-out modes, the address register and the cumulator are engaged to count markers and spaces on the memory line, respectively. The histogram will be displayed by simply transferring the data in the cumulator to the cumulator-hold register and by resetting the cumulator upon every marker, and then connecting the address-hold register and the cumulator-hold register to the X- and Y-deflection of the CRT. Both the X- and Y-deflection increments will, in this manner, be constant. However, the length of time a spot remains will vary between points. Each display cycle is equivalent to one complete cycle of the serial memory. The address full-scale signal augments the mode control counter. Figure 4 shows the block diagram for the CRT display arrangement.

H. M_5, Display of the Cumulative Function

The only difference between M_5 and M_4 is that the cumulator is *not* reset upon a marker. The transfer from

---

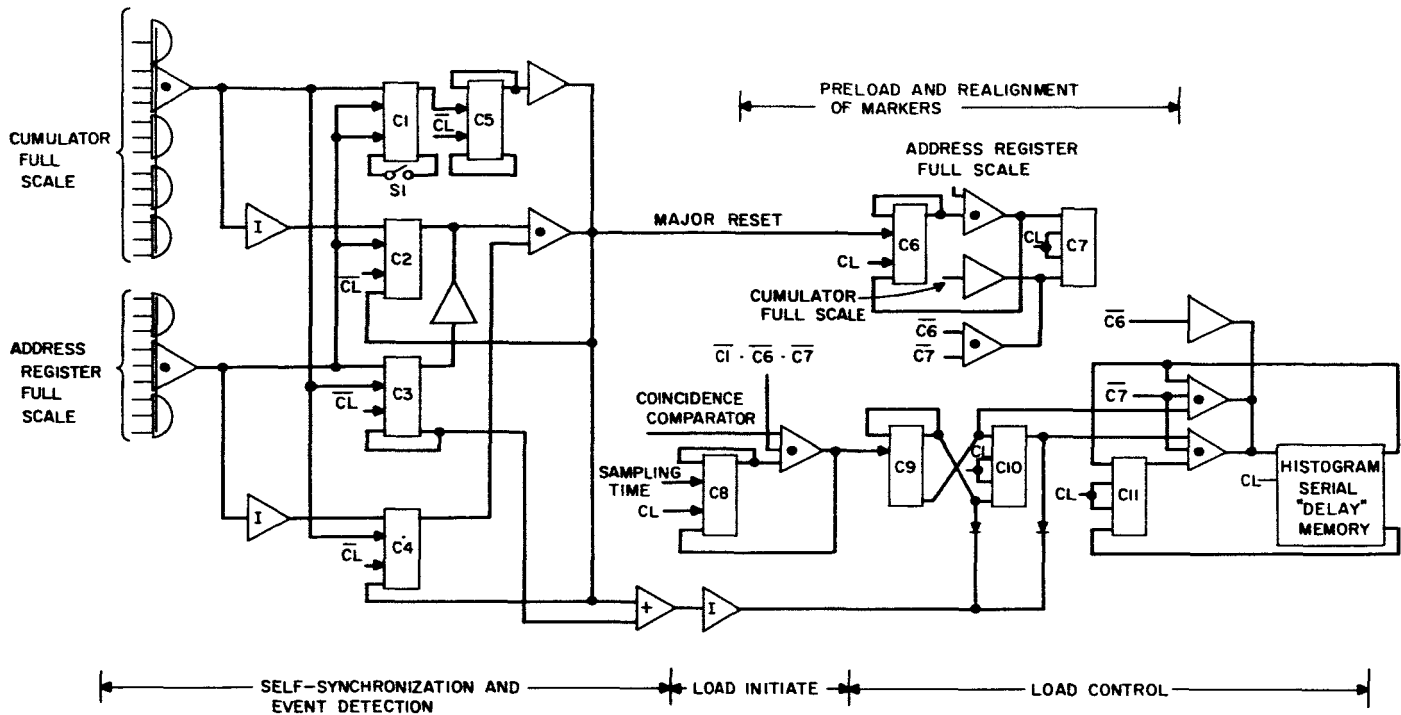[1] This section of the system has been updated and is described in Ref. 3.

Fig. 3. Logic diagram of the load control unit

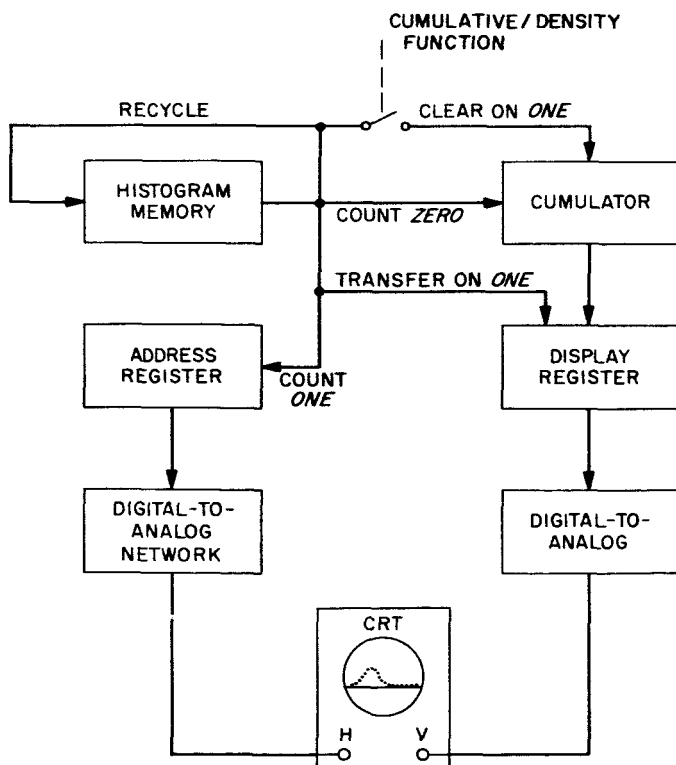Fig. 4. Block diagram of CRT display

the cumulator to the cumulator-hold register, however, does occur.

I. M_6, Display of the Quantile Numeral

In all other modes the transfer from the address register proper to the address-hold register occurs upon each clock pulse. In mode 6 the address-hold register is also the quantile register; however, the common output from the quantile detectors constitutes that transfer command. The quantile detector is simply a coincidence gate connected to the cumulator-hold register. In M_6 the transfer to the cumulator-hold register occurs upon the clock.

J. Mode Control

The operation of the mode control when power is first applied (or at power interrupt of any kind) needs to be discussed. In M_1, M_2, and M_3 the address register and the cumulator are counting clock pulses or sample pulses, respectively, and will thus generate full scale of these counters; this event will change the state of the mode counter. With either the address full-scale or the cumulator full-scale used to augment the mode control counter in M_4, M_5, and M_6, any hang-up, due to the

serial memory containing all ones or all zeros, will be prevented.

K. Automatic Sampling Control

To allow for variations of incoming data-rate an automatic sampling control has been provided. Shown in Fig. 5 is a block diagram of this arrangement. An oscillator frequency is divided down by a 6-bit binary ripple counter. One out of seven frequencies (the oscillator frequency or the output from any of the six counter stages) is automatically selected by a 3-bit selector counter. The selected output is synchronized with the system clock. The input to the selector counter is the carry from the primary control counter.

This primary control counter is a 5-bit up-down counter, the up-down control terms being the distributor empty- and full-scale conditions and the clock being the sampling signal (the transfer and reset signal for the distributor). The 3-bit selector counter is also an up-down counter and is augmented by a carry from the primary control counter. Also upon this carry, the primary counter itself

resets to mid-scale so as not to cause oscillations upon alternating distributor empty- and full-scale signals; such oscillation would waste too much of valuable data since every carry from the primary control counter also constitutes a reload (M_1 and reset of address register) command.

Both the primary control counter and the sample rate selector counter are implemented as PN-sequence shift register counters; in this case, such counters proved to be the most economical method. The count up-down mechanism then is implemented as a right-left shift with the feedback interchanged. For the primary control counter, the output is combined with the all-zero starter gate; for the selector counter, an all-zero gate is added to the middle stage.

The carry from the primary control counter simultaneously constitutes: (a) the clock for the selector counter, (b) the reset-to-mid-scale signal for the primary control counter, and (c) the reload command, which resets the address register and sets the mode control counter to the M_1 stage.
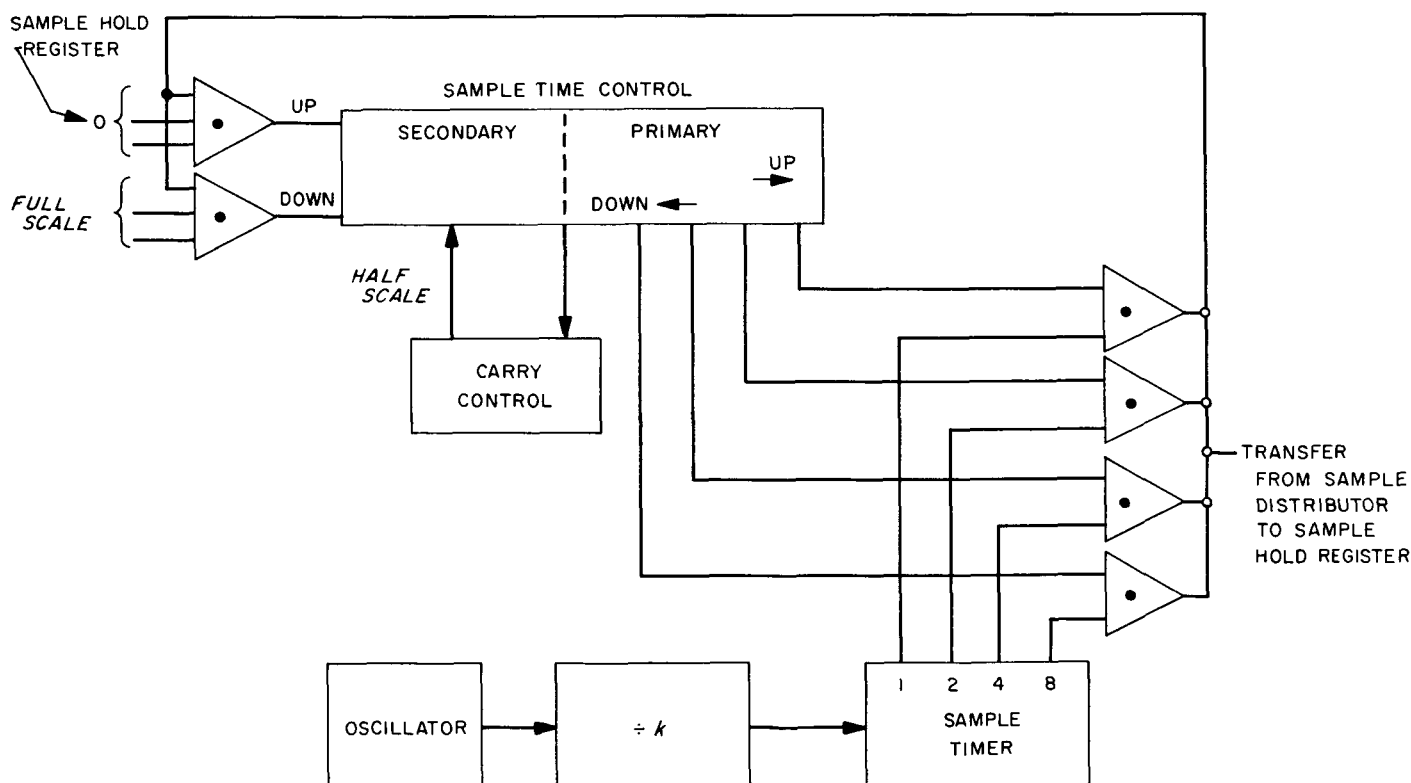

Fig. 5. Automatic sampling time control

L. X-Y Recorder Display

The type of display best suited for any one demonstration depends on the data rate and, thereby, also on the sampling rate. Functional units auxiliary to the Quantiler are needed to display the histogram and the sample cumulative distribution on an X-Y recorder. The operation of the various auxiliary units will now be described. Figure 6 shows the block diagram of the Quantiler with the associated display functions.

Auxiliary devices necessary for an X-Y recorder display are:

1. A display cumulator.

2. A display register.

3. A digital-to-analog converter.

4. An address register.

5. A coincidence comparator.

6. Associated control logic.

The display cumulator can be made to cumulate the spaces between any two selected markers only, or it can cumulate from the first marker to any selected marker. The cumulation of all spaces between two markers is accomplished with the aid of the display address register and the coincidence comparator. The coincidence comparator is connected between the Quantiler address register and the display register. Upon coincidence, a control bit C_d is set. This bit controls the input to the cumulator. Then C_d is reset upon the next marker on the line. For display of the cumulative distribution function, C_d is set at the end of the line and then reset upon coincidence between the two address registers.

In order to maintain equal time spacing for the horizontal deflection, transfer of the cumulator data to the hold register always occurs at the end of the line. Also, the display address register is augmented by one count, and the cumulator is reset at the end of the line. Connected to the display hold register is a digital-to-analog network. The analog output of this network is applied to the vertical deflection amplifier of the X-Y recorder. The input to the horizontal deflection amplifier is the continuous sweep, internal to the recorder. A proper signal to be used as a sync signal is the Quantiler address-zero signal. The length of this horizontal deflection shall correspond to as many recycling times as there are addresses. In this display scheme, the histogram will be retained for display for the same number of cycles as there are addresses. The addresses are tallied by the display address register which is augmented once per cycle.

M. Removing of Additive Bias

A given histogram may be located between A and B_1 as shown in Fig. 7. The spread of the samples is larger than the available histogram space and the absolute value of each sample is greater than the full scale of the sample distributor.

If a sample A is greater than B_1 with an amount equal to or greater than S, the sampling time should be decreased. The histogram is, thereby, relocated to be between B_2 and $B_2 + S$. By deducting the bias of B_2 from each sample, the histogram will be located between O and S, using all the available histogram memory capacity. In mechanizing the above mentioned function, points B_1 and B_2 must first be located.

The method by which to accomplish the compression and shift of the histogram simultaneously is relatively simple, although it can involve loss of data. In favor of the method is the fact that the data collected are the best possible using the quantile method. The histograms will fit between O and S with most of the slots occupied by at least one sample. The reduction process is described with reference to the block diagram shown in Fig. 8.

Except for the distributor A and its hold register B, a second hold register C, two comparators and an arithmetic transfer control from A to C are required for bias removal. The two comparators are: (1) $A < B$, whose output constitutes a "transfer from A to B" command, (2) $A > B + S$, whose output constitutes "divide sample time" command. Any one of these commands also constitutes a major reset command for the memory line and its control counters.

With the smallest of all samples in the sampling cycle in B and with a sample within the allotted sample spread in A, a transfer of $A - B$ into the sample hold register will provide a proper sample value to be loaded into the histogram memory line. The $A > B + S$ comparison and the $A - B$ transfer are implemented simply by the counting up and down, respectively, of the A and B registers. This arrangement for removing of additive bias is not included in the preset prototype model of the Quantiler. Another less involved method for handling samples of excessively large size and spread is suggested for the flight model, that of using a logarithmic sample distributor (suggested by David Schaeffer of NASA-Goddard Space Flight Center).

N. Flight Model

With the input sample counter counting in a logarithmic fashion (see Ref. 4), the automatic sample time con-
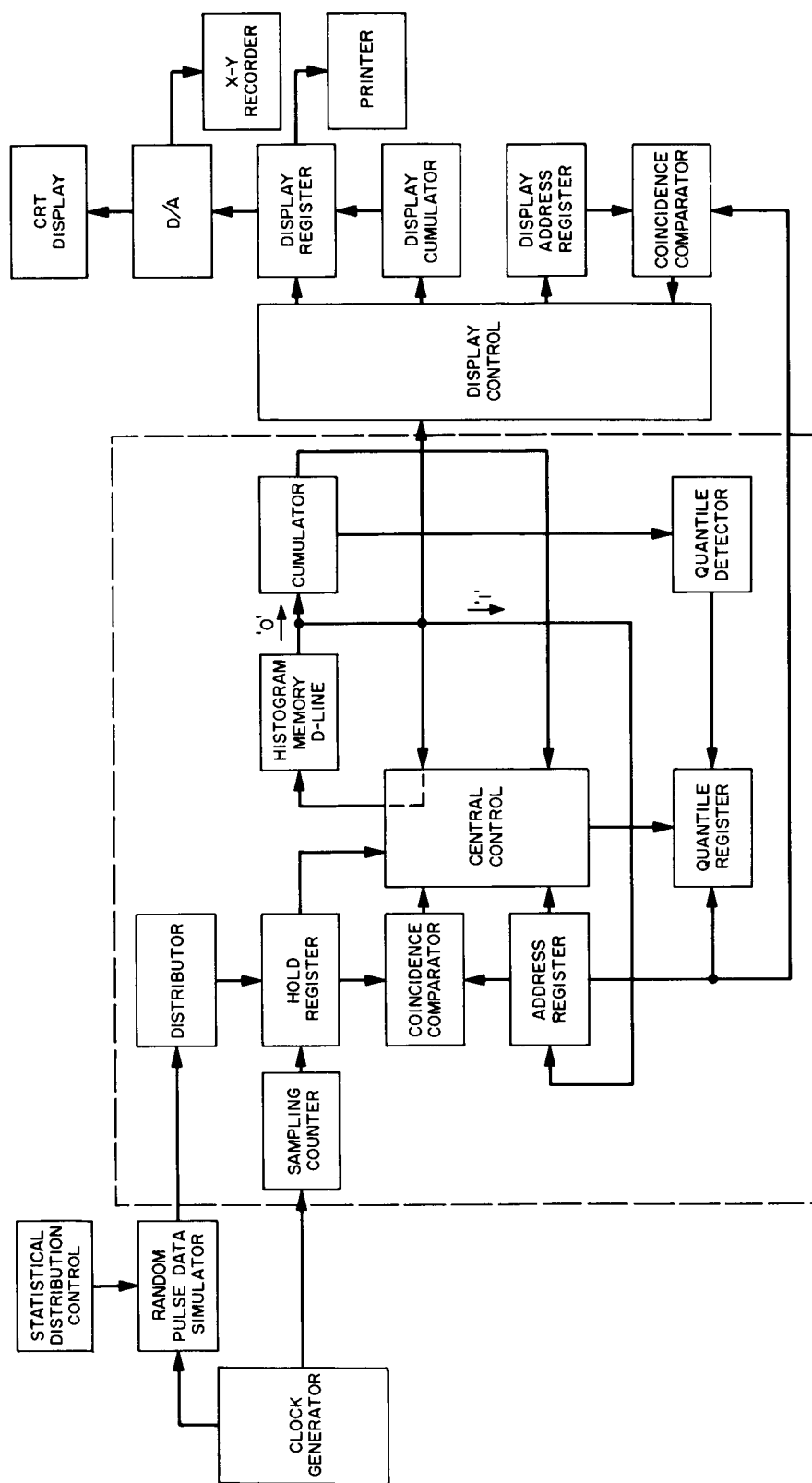
Fig. 6. Block diagram of the quantiler and input/output and display functions
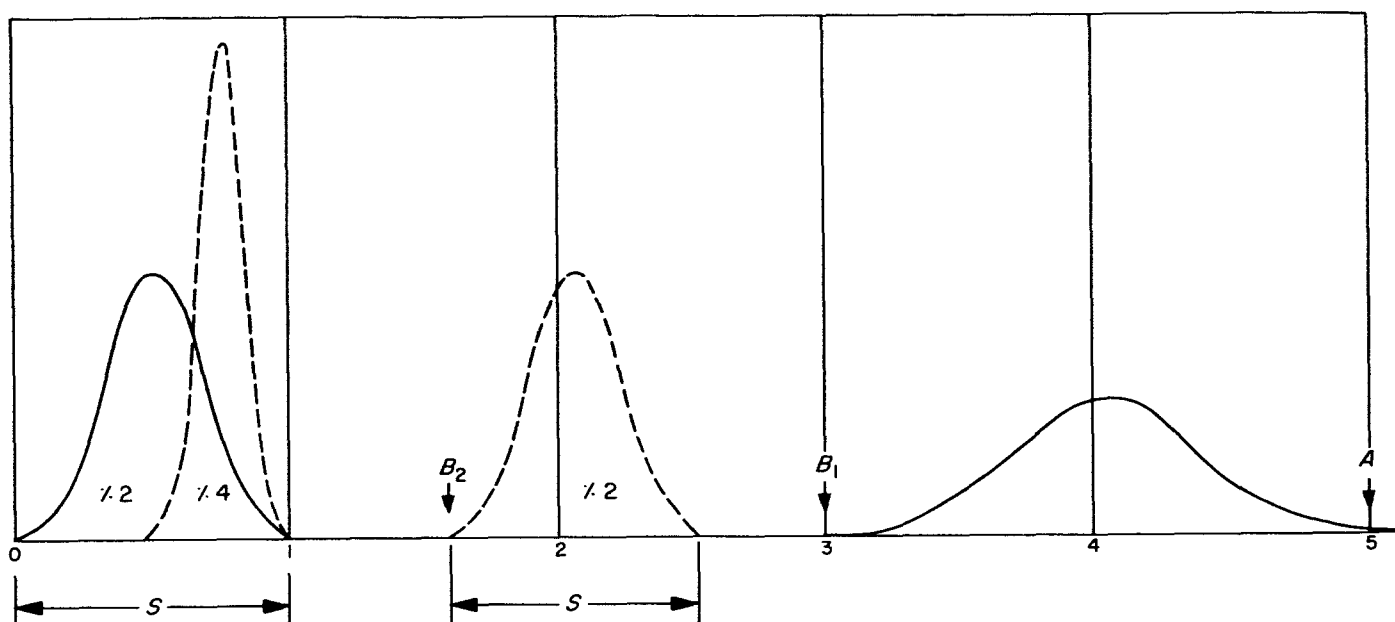
Fig. 7. Relocation of a histogram, graphic presentation

trol can be eliminated. This design includes the same functional units as the sampling time control. The logarithmic counter would replace the sample distributor with a saving in the total amount of hardware but at the penalty of introducing a 3% sample error in the worst case.

In a flight model, none of the three copy registers will be necessary. The distributor will then be controlled to stop when a sample has been taken, and the distribution will not be cleared (reset) and operative again until that particular sample has been loaded into the memory; or, for a constant sampling rate, until a stop time, sufficiently long to load any sample, can be used.

Fig. 8. Block diagram of the histogram relocation function

The ambiguities in the rippling address register and the cumulator—an ambiguity detrimental to any associated comparator—would be bypassed by using the input to these counters as a single term strobe in the respective comparators. The mode control in the flight model need only include four modes, M_1, M_2, M_3, and M_6.

A prototype flight model of the Quantiler can be built using micro-electronics and a glass line memory operating at 5 mc. As shown in the block diagram (Fig. 9), this model would consist basically of only three counters: (1) the sample counter, (2) the address register, and (3) the cumulator.

The sample counter will be of a wide dynamic-range-type counter requiring fewer bits. The address register will then also require fewer bits. The serial memory can then be made shorter and allow the sampling rate to be increased. Functions associated with the three counters will be: (1) a coincidence comparator and (2) a quantile detector (fixed-wired AND-gates connected to the cumulator).

A 2-bit program control counter could control the sync cycle, the load cycle, and the quantile readout cycle. Thus, an on-board Quantiler for data compression can be built with an extreme minimum of hardware, still yielding compression ratios of 100:1 or greater with 100% efficiency.
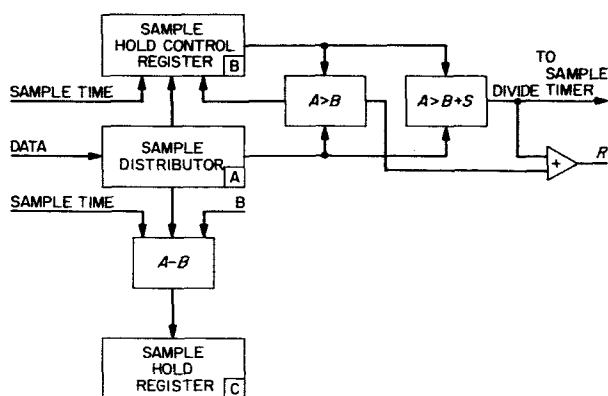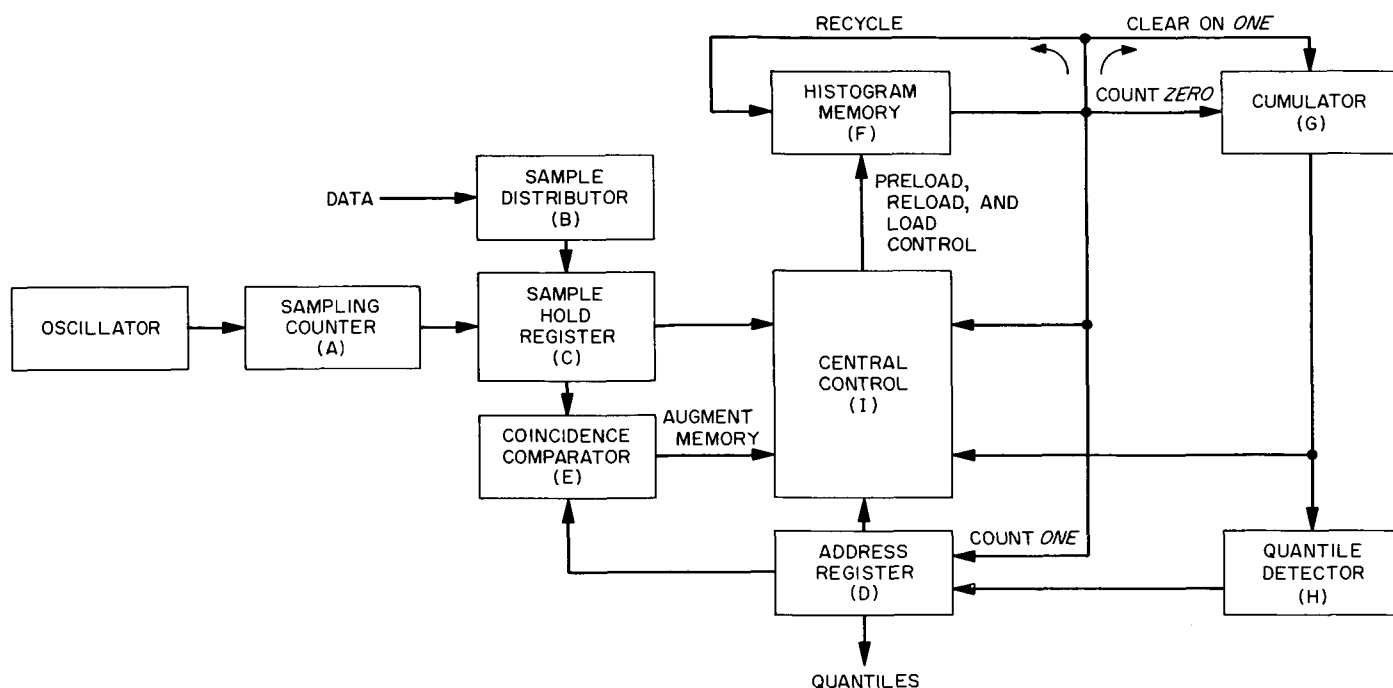
Fig. 9. Block diagram of a quantiler flight model

IV. THE CONSTRUCTED QUANTILER

Photographs of the Quantiler prototype are shown in Fig. 10. It is constructed from a set of standard digital circuit modules comprising three basic modules: (1) a decision element module, (2) a memory element module, and (3) a power amplifier module.

There are four identical circuit modules per systems module or card in a 28-pin-connector arrangement. DTL circuitry is used. The decision element is a three-input-plus-node NAND-circuit with separate collector resistor. Paralleling of the collectors of several circuits provides a second logic level with no increase in propagation time. The memory element is a triggered symmetrical flip-flop. A negative voltage transition will set or reset a flip-flop, provided the associated control level is at ground level. Individual trigger terms for set and reset are provided. DC set or reset is accomplished through collector set or reset by connecting the output of a decision element to the output of the flip-flop.

A power amplifier is provided for driving excessive logic loads as well as large capacitive loads. The output of this circuit is basically a complementary inverter. NPN silicon planar epitaxial transistors are used throughout this standardized set of logic modules. The voltage supply levels are ± 12 v, and the logic swing is nominally $+10$ v and ground—the transition being the trigger transition. These modules are intended for use in synchronous systems operating at clock rates up to 3 Mc (see Ref. 5). The serial memory is a magnetostrictive delay line with associated circuitry.
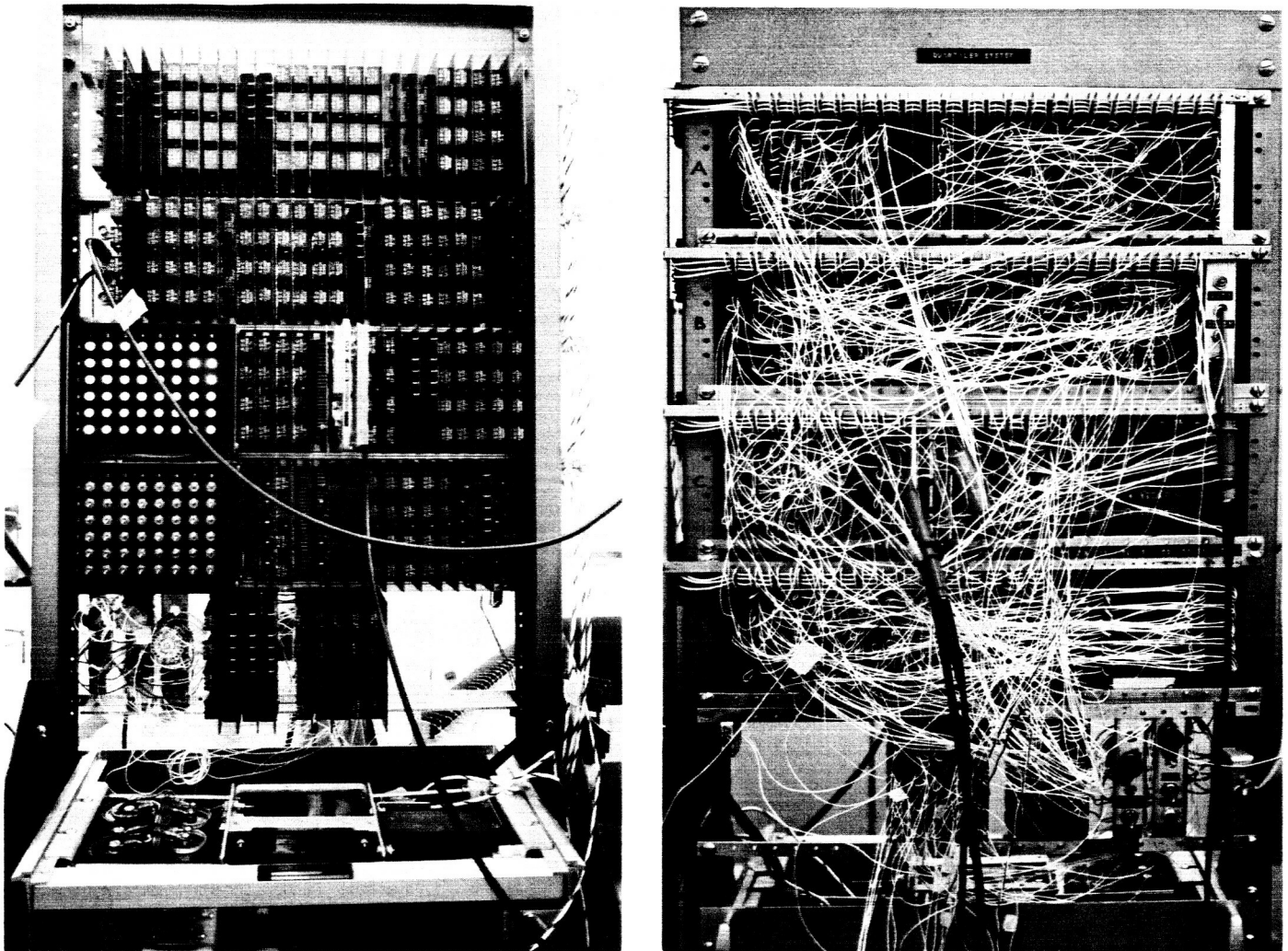
Fig. 10. The quantiler prototype, front and rear

V. EXPERIMENTAL RESULTS

Upon completion and checkout of the Quantiler, known distributions generated by the random pulse generator (Ref. 6) were sampled and displayed by the Quantiler display networks. Figures 11 through 13 show the three displays: the sample histogram, the sample cumulative function, and the four sample quantiles, displayed in octal, superimposed on one picture. Figures 11 and 12 show histograms of data taken from a population with a binomial distribution (close to being normal) but with different means. Figure 13 shows a *bimodal* distribution obtained by changing the sample population about half way through the sample. The effect on the
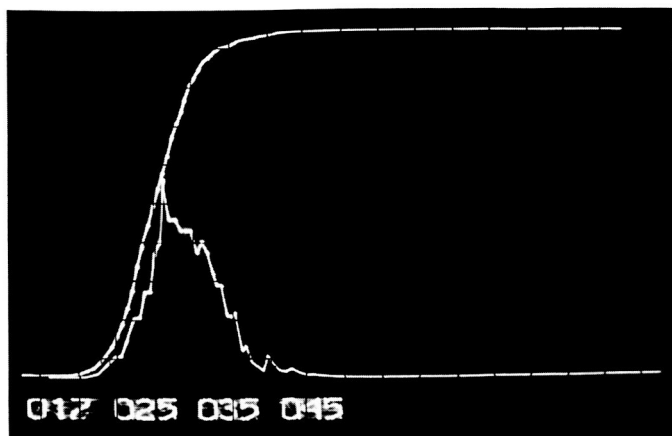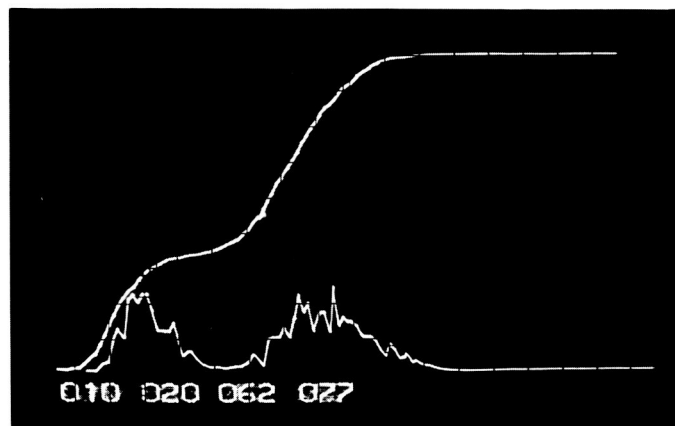
Fig. 11. First binomial distribution
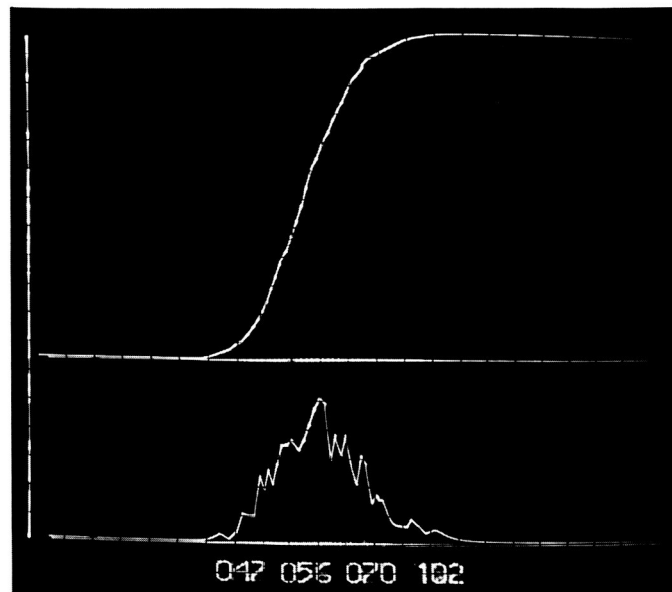

Fig. 13. A bimodal distribution

Fig. 12. Second binomial distribution

quantiles, in this case, is that there is a much larger gap between the second and third quantiles than in the unimodal case.

Tests were then performed to determine the mean and standard deviation of known distributions by sampling and computing these parameters from the observed quantiles. The estimates of these parameters were calculated by the following formulas from Section I:

$$\hat{\mu} = 0.141\ [Z(0.0668) + Z(0.9332)] + 0.359\ [Z(0.2912) + Z(0.7088)]$$

$$\hat{\sigma} = 0.258\ [Z(0.9332) - Z(0.0668)] + 0.205\ [Z(0.7088) - Z(0.2912)]$$

The results of these tests are given in Table 1. It is observed that in all cases the agreement between the observed and theoretical values is excellent. In fact, in the case of the "renewal" data, the variance calculated from the quantiles disclosed an error in the calculation of the theoretical variance!

Table 1. Comparison of experimental and theoretical results

Known distribution	Sample quantiles				$\hat{\mu}$	$\hat{\sigma}$	$\overline{\hat{\mu}}$	$\overline{\hat{\sigma}}$
Binomial $\mu = 50$	43	47	52	59	49.9	5.15		
$\sigma = 5.0$	43	47	52	58	49.8	4.89	49.8	4.90
	43	47	52	57	49.6	4.64		
Binomial $\mu = 125$	109	118	130	140	124.1	10.46		
$\sigma = 10.825$	109	118	131	142	124.8	11.18	124.5	10.71
	109	119	130	141	124.6	10.51		
Overlapping windows—111	104	116	133	149	125.1	15.10		
$\mu = 125$	103	116	133	149	124.9	15.35	124.99	15.29
$\sigma = 15.31$	104	116	132	151	125.0	15.41		
Overlapping windows—101	106	117	132	149	125.4	10.00		
$\mu = 125$	106	116	131	145	125.1	10.51	125.22	10.25
$\sigma = 10.45$	106	117	132	147	125.2	10.25		
Renewal process—111	61	67	75	82	71.9	7.06		
$\mu = 71.43$	61	67	75	83	71.3	7.32	71.47	7.23
$\sigma = 7.19$	61	67	75	83	71.3	7.32		

VI. APPLICATIONS OF THE QUANTILER

Although our motivation for developing the Quantiler was the need for data compression on deep space flights, we quickly became aware of the relevance of quantile technology to other telemetry situations including those in civilian technology. One situation in which the channel is overloaded is in the stock market. A Quantiler can be used here on the distribution of weighted price changes on each stock approximately every ten minutes. While this information will not tell the individual investor how his own stocks are doing, it will give a clear and continuous picture of how the market is doing. Very little extra channel capacity, above that needed to transmit individual transactions and prices, will be required.

Another use is the industrial quality control of a product manufactured automatically in an inaccessible location. Many quality control techniques in effect use the mean and variance of certain features of the manufac-tured product to detect departures from the so-called state of statistical quality control. A Quantiler at the data source would compute the quantiles for transmission to the quality control engineer. From these quantiles, the engineer would receive information not only of the mean and variance, but of other features which may be even more important in quality control. For example, if a bimodal distribution is detected, this could indicate that a machine malfunctioning occurs part of the time, chang-ing some statistical characteristics of a certain fraction of the output.

A still different application is the metering of utility usage to determine utility bills. Some utilities charge on a basis not only of average use but of fluctuations about average. If the use per half-hour is recorded in the sub-scribing facility and quantiles are formed locally once a week and transmitted to the utility, a charge can be

made based on mean and variance, as required by a rate table. The rate commissions can probably be convinced to go along with such schemes.

An application to automatic traffic control also exists. At remote sites throughout a city, the spectrum of vehicle velocities is taken. Quantiles of this velocity spectrum are transmitted to a central computer, which uses the information to monitor traffic. For example, a low average velocity with a large variance indicates an impending traffic jam.

Other areas for civilian use that occurred to us will be mentioned briefly. Workers in highly radioactive environments can use quantiles to monitor the radiation spectrum. The Weather Bureau can use quantiles to give usable information on quarter-hour temperatures at a given location without having to transmit the temperature every fifteen minutes.

Returning to space applications, there are other types of experiments (other than particle count) to which quantiles apply. Energy spectra of incoming particles are distributions and, as such, can be effectively transmitted by quantiles; the same applies to mass spectrograms in planetary surface chemical experiments.

The following application was suggested by Michall Dix of NASA Ames Research Center. Certain types of interplanetary probes rotate and an energy sensor records an energy distribution as a function of angle. This distribution is of course compressible by quantiles.

In fact, quantiles afford an extraordinary saving whenever any kind of curves, spectra, or distributions have to be transmitted. The potential applications (especially in a technology where remote sites are linked together by computers) are far-reaching indeed.

REFERENCES

1. Eisenberger, Isidore, and Edward C. Posner, *Systematic Statistics Used for Data Compression in Space Telemetry*, Technical Report No. 32-510, Jet Propulsion Laboratory, Pasadena, California, October 1963. (Also accepted for publication by the *Journal of the American Statistical Association*.)

2. Goldstein, Samuel, "A Bibliography on Magnetostrictive Delay Lines," *IRE Transactions on Ultrasonics Engineering*, Vol. UE-10, pp. 58–71, 1963.

3. *Space Programs Summary No. 37-27*, "Design of a Quantile System for Data Compression of Space Telemetry," Vol. III, pp. 103–112, Jet Propulsion Laboratory, May 31, 1964.

4. Schaefer, David H., "Logarithmic Compression of Binary Numbers," *Proceedings of the IRE*, Vol. 49, No. 7, July 1961.

5. Anderson, Tage O., *Logic Design Manual for a Set of Digital Circuit Modules*, Engineering Planning Document No. 196, Vol. I, Jet Propulsion Laboratory, Pasadena, California, August 21, 1964.

6. Anderson, Tage O., and Edward C. Posner, "Design and Construction of a Random Pulse Generator," *Space Programs Summary No. 37-29*, Vol. III, pp. 92–95, Jet Propulsion Laboratory, Pasadena, California, September 1964.