AD-A066 663 TRACOR INC AUSTIN TEX F/G 17/1

COMPUTER UTILIZATION OF SEQUENTIAL HYPOTHESIS TESTING FOR DETEC--ETC(U)

DEC 67 H A REEDER NOBSR-93352

UNCLASSIFIED TRACOR-67-717-U NL

| OF |

AD A066663

END
DATE
FILMED
5-79
DDC

Contract NObsr 93352
Project Serial SF001-03-16
Task 8523
TRACOR Project 002 016 01
Document Number 67-717-U

TECHNICAL MEMORANDUM

COMPUTER UTILIZATION OF SEQUENTIAL HYPOTHESIS TESTING FOR DETECTION AND CLASSIFICATION OF SONAR SIGNALS

by

H. A. Reeder

Submitted to:

Commander, Naval Ship Systems Command
Department of the Navy
Washington, D. C. 20360
Attn: Mr. Joe Manseau, Code OOV1

7 December 1967

TRACOR

6500 Tracor Lane, Austin, Texas 78721, AC 512/926-2800

TRACOR 6500 TRACOR LANE, AUSTIN, TEXAS 78721

(15) Contract NObsr-93352
Project Serial SF001-03-16
Task 8523
TRACOR Project 002 016 01
Document Number 67-717-U

(16) F00103

(17) SF0010316

(9) TECHNICAL MEMO~~RANDUM~~

(6)

COMPUTER UTILIZATION OF SEQUENTIAL HYPOTHESIS TESTING FOR DETECTION AND CLASSIFICATION OF SONAR SIGNALS.

(14) TRACOR-67-717-U

(10) Hugh ~~H~~. A. Reeder

Submitted to:

Commander, Naval Ship Systems Command
Department of the Navy
Washington, D. C. 20360
Attn: Mr. Joe Manseau, Code 00V1

(12) 44p.

(11) 7 Dec~~ember~~ 1967

Approved by:

E. A. Tucker

E. A. Tucker
Program Manager

Prepared by:

Hugh A. Reeder

Hugh A. Reeder
Project Director

352100

TRACOR 6500 TRACOR LANE, AUSTIN, TEXAS 78721

ABSTRACT

Active sonar systems furnish the operator large quantities of data to examine for possible targets. In order to detect small target returns, the display thresholds must be set such that clutter is significant. On a PPI display this clutter makes it difficult to visually integrate successive pings.

This report shows a technique utilizing a digital computer to examine the data out of a sonar system, reduce its volume, perform ping-to-ping integration, make simple decisions, and generate useful displays. The process takes advantage of a statistical decision procedure called sequential hypothesis testing.

Specifically, the technique is not dependent on a particular sonar system; it is well founded mathematically, and sufficiently simple to make real time implementation possible on a modest state-of-the-art computer.

There is no fundamental limit on the number of echo cycles to be considered for ping-to-ping integration, in the decision process.

TRACOR 6500 TRACOR LANE, AUSTIN, TEXAS 78721

TABLE OF CONTENTS

TRACOR 6500 TRACOR LANE, AUSTIN, TEXAS 78721

TABLE OF CONTENTS (Cont'd.)

LIST OF ILLUSTRATIONS

1. INTRODUCTION

1.1 HIGHLIGHTS

A technique for accomplishing computer-aided detection and classification has been developed. Highlights of the technique are:

1. The basic detection algorithm is not dependent upon a particular sonar system; rather, it is applicable to any one of a quite large class of active sonar systems.

2. The procedures used are well-founded mathematically, rather than being empirical, and can be shown to be optimal from the standpoint of statistical decision theory.

3. The necessary logical and arithmetic calculations are simple. This should allow real time implementation on a reasonably modest state-of-the-art digital computer.

4. In the basic decision process there is no fundamental limit on the number of echo cycles to be considered for ping-to-ping integration; as much of the available data is used as is necessary.

The remainder of the introduction gives a brief, general discussion of the technique and the parameters involved. A more complete discussion of the theory and information flow is given in Sections 2 and 3. Section 4 discusses how the technique was implemented for a particular problem and gives results for different situations.

1.2 PROBLEM

Active sonar systems furnish the operator large quantities of data to examine for possible targets. In order to detect small target returns, the display thresholds must be set such that clutter is significant. On a PPI display this clutter makes it difficult to visually integrate successive pings. By displaying several successive A-type scans, ping-to-ping integration may be accomplished and, consequently, the probability of detection is increased. However, the best results are obtained only when an alerted operator is handling this mass of data. On long searches with few targets, this alerted condition would be difficult to maintain, and a computer detection and classification system may be more responsive to low signal-to-noise ratio targets.

1.3 TECHNIQUE FOR A SOLUTION

This report shows a technique utilizing a digital computer to examine the data out of a sonar system, reduce its volume, perform ping-to-ping integration, make simple decisions, and generate useful displays. The process takes advantage of a statistical decision procedure called sequential hypothesis testing. In this procedure, the amplitude of a promising return peak is transformed by a function dependent on the output statistics of the sonar processor. If the resulting value is greater than a certain threshold value, the peak is accepted as a possible target and displayed. If it is below a lower threshold value, the peak is rejected as a non-target. If it is between the two thresholds, it is retained in the computer, along with other information about the peak, to form a track. On the next ping, the volume to which a target might move is searched for possible target peaks. Linkages are made with the new peaks, and a statistic based on both peaks is calculated. The decision process to test whether the track is target, non-target, or undetermined is repeated using the same thresholds. A track may be undetermined for as many pings as are necessary to make the appropriate final decision.

1.4 PARAMETERS

By choosing a display threshold for the track statistic that corresponds to the statistic obtained by a single peak return with an amplitude of 4 to 5 noise standard deviations above the noise mean, a significant reduction in clutter rate is achieved. Ping-to-ping integration of weak target tracks with individual peaks less than that value is left to the computer until a decision is reached.

Besides the display threshold there are other parameters involved in the procedure; however, these would be changed only occasionally. The lower rejection threshold should be set so that the probability of a true target track being rejected is small. The transformation of peak amplitude to the testing statistic requires an assumed target signal-to-noise ratio; the testing procedure is optimal for tracks of this signal-to-noise ratio, with only slight degradation for target tracks at higher signal-to-noise ratios.

For tracks below the assumed signal-to-noise ratio, degradation for the testing procedure increases slowly, at first, with decreasing track signal-to-noise, then rapidly until a sharp cutoff is reached about 6-8 dB below the assumed signal-to-noise ratio. Tracks with average signal-to-noise ratios below this cutoff value will not integrate up to the display threshold. It is desirable to have the assumed signal-to-noise ratio low, but if it is too low, the probability that a noise track or peak will be treated as a target track increases, and consequently the clutter rate increases.

Since a target only can move a certain distance each ping cycle, it is possible to limit the number of peaks in each cycle that can be linked with an established track. In general, the more stringent the limits, the faster the decision, target or non-target, will be made. On the other hand, too severe limits would not allow linkages on a track of a maneuvering target. Hence, these limits should be dependent on the characteristics of possible targets in a particular area.

1.5 SUMMARY

With the technique described above, a computer can be used not only to reduce the output rate of the sonar system to an acceptable level, but also to enhance the probability of detection of target tracks by ping-to-ping integration.

The selectivity of the process can be used to limit the number of possible targets to be subjected to further classification procedures, which take computer time and core memory to accomplish. The exact classification procedures to be implemented would depend on the sonar system and the pulse form transmitted. In addition to the two classification clues, target strength and track consistency, used in the process, some of the items that could be checked are Doppler consistency of echo, echo length, wake, highlights, and aspect. Further study is needed to apply these descriptors to the detection and classification process. The computer can perform other functions such as range prediction, mode selection, attack problems, and training, but these are outside the scope of the present report.

2. THE LIKELIHOOD RATIO AND ITS USEFULNESS IN STATISTICAL DECISION THEORY

2.1 HYPOTHESIS TESTING

The fundamental problem of statistical decision theory is that of choosing one of several hypotheses by making observations of some quantity. A great deal of generality can be included in defining the algorithm to optimally carry out this procedure. A simplified approach will be taken here. Assume that the number of hypotheses available is two, and refer to these hypotheses as H_0 and H_1. In the problem under consideration H_0 is the hypothesis "the track is non-target or noise to be rejected" and H_1 is "the track is a target to be displayed". It is also assumed that the quantity to be observed is a single numerical value. If the hypothesis H_0 is true, then observed values of the quantity x will be described by a known probability density function, $p_{0(x)}$, such as the one shown in Fig. 1. In a similar way, if H_1 is true, then there will be a different probability density function, $p_{1(x)}$, for the observed values of x. For the examples shown in Fig. 1, x might be thought of as the amplitude of a peak; p_0 as the distribution under the conditions of noise alone; and p_1 as the distribution when signal plus noise is present. The likelihood ratio, L(x), is defined as

$$L(x) \triangleq \frac{p_1(x)}{p_0(x)} .$$

In many cases the likelihood ratio is a monotonically increasing function of the observation quantity x. In this example, large values of the likelihood ratio tend to imply a choice of H_1, while small values of the likelihood ratio tend to imply a choice of H_0. Thresholds need to be established on L(x) for our decision process. There are several techniques for picking an optimal threshold which differ primarily in the amount of a priori

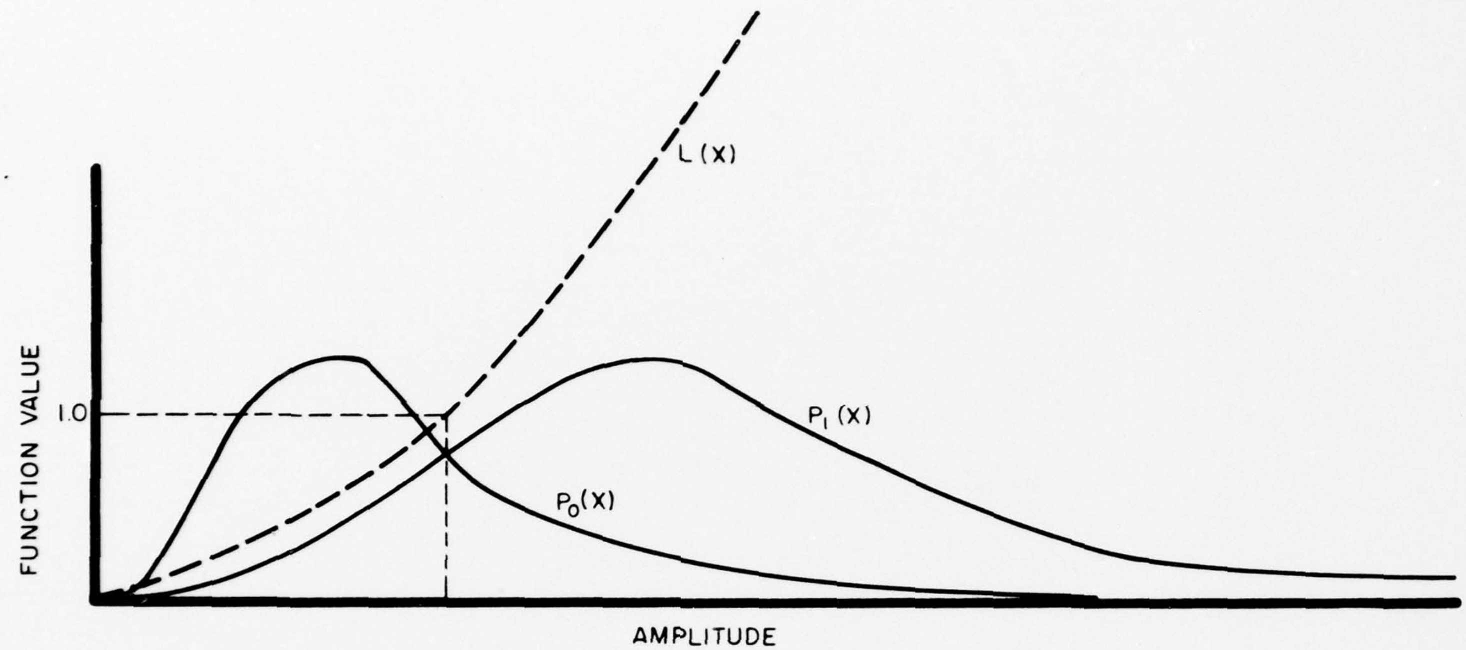

FIG. 1 – TYPICAL PROBABILITY DENSITY FUNCTION AND LIKELIHOOD RATIO

TRACOR DWG. A6-105-21
9/14/67 REEDER/SL

information that is known about the problem. For sonar applications most of the assumed quantities in the highly optimal techniques are simply not known. The threshold, in this case, will likely be based upon what false alarm rate or clutter rate can be tolerated rather than upon the classical approach from a standpoint of decision theory.

2.2 MULTIPLE OBSERVATIONS

If several observations of the quantity (x_1,x_2,x_3,etc.,) are to be made, then a joint likelihood ratio for the set of observations would be defined as the ratio of the appropriate multi-dimensional probability density functions,

$$L(x_1,x_2,x_3,\ldots) \triangleq \frac{p_1(x_1,x_2,x_3,\ldots)}{p_0(x_1,x_2,x_3,\ldots)} .$$

If the observations made are all statistically independent, then the appropriate multi-dimensional probability density function is simply the product of the individual one-dimensional probability density functions:

$$L(x_1,x_2,x_3,\ldots) = \frac{p_1(x_1)\cdot p_1(x_2)\cdot p_1(x_3)\ldots}{p_0(x_1)\cdot p_0(x_2)\cdot p_0(x_3)\ldots} .$$

This yields a significant simplification and also introduces the logarithm of the likelihood ratio as a very useful quantity. The multiplication of individual observation likelihood ratios to obtain a joint likelihood ratio can be avoided by using the logarithm of the likelihood ratio as the fundamental quantity. The logarithm of the joint likelihood ratio can be obtained as the sum of the logarithms of individual ratios:

$$\ell(x_i) \triangleq \mathrm{Log}(L(x_i)) = \mathrm{Log}\left(\frac{p_1(x_i)}{p_0(x_i)}\right) ,$$

and

$$\ell(x_1,x_2,x_3,\ldots) = \ell(x_1) + \ell(x_2) + \ell(x_3) + \ldots .$$

2.3 SEQUENTIAL TESTING

It is of interest to consider a system in which the number of observations to be taken is not a fixed quantity but, instead, a decision is to be made when specified confidence levels are reached. This technique of observation is known as sequential testing. In this approach, two thresholds are established. The first threshold t_1 is a threshold such that if the joint log of likelihood ratio falls below this level, Hypothesis 0, no signal is present, is chosen and the testing chain stops. On the other hand, if the joint log of likelihood ratio exceeds a second threshold t_2, then Hypothesis 1, signal is present, is chosen. The detection process is complete, though the testing does not stop but continues into an automatic track. If the joint log of likelihood ratio lies between the bounds of t_1 and t_2, no decision is made; another sample is taken and the decision process is repeated. The situation is very similar to that of the random walk problem, and it can be shown that eventually one of the two boundaries will be crossed and a decision will be reached.

The average number of samples needed to reach a decision for given probabilities of wrong decisions is less than the required number for a fixed sample size test with the same given probabilities. If the track is noise, the decision is reached relatively promptly, but, if the track is a target, a significantly greater number of samples may be needed for a decision[1].

1 Skolnik, M. I., "Introduction to Radar Systems", McGraw-Hill, New York (1962).

2.4 TRACKING

In sonar applications one difficulty arises which does not often occur in other statistical decision theory applications. The problem is that one does not really know uniquely how to take a single "next" observation. For example, receipt of a modestly high peak on one beam, at a given range, and perhaps with some Doppler, gives an indication of where to look in the next echo cycle, in terms of beam, range, and Doppler. This information, however, cannot give a precise specification of the location of the linking data in the next echo cycle. Rather, it shows a volume in the next echo cycle which must be scanned, the volume being defined by a range interval, a beam interval, and a Doppler interval. This volume must be searched to determine whether any linkages exist with the previous data.

Since multiple linkages are allowed in this process, the probability of having chosen the correct linkage must be used in calculating the likelihood ratio. Under the Hypothesis H_0, that the linkage is noise, the probability is 1.0 that a noise peak is chosen.

The probability of a correct linkage under the Hypothesis H_1 is the number of true target tracks divided by the number of possible linkages. It is unlikely that there will be more than one true target track in any given volume, and it will be assumed that the number of true targets is 1.0. Since it would be difficult to count the number of possible linkages each time, the average number of possible linkages will be used. Hence, the likelihood ratios must be multiplied by the reciprocal of the average number of possible linkages, or, equivalently, the logarithm of that number must be subtracted from the log likelihood ratio. This points up the advantages of having a highly localized target track. If it is possible by some means to confine the volume of legitimate linkages to a relatively modest volume, then the amount that must be subtracted from the log likelihood integration is decreased. On the other hand, if the volume of legitimate linkages can only be defined in a gross sense, then from each linkage that is established the logarithm of the number of independent samples in this total volume must be subtracted.

3. COMPUTER PROCESS DESCRIPTION

3.1 INTRODUCTION

The overall purpose of the system is to produce a display on which the returns most likely to have resulted from targets stand-out, and on which the tracks of these returns over a number of pings can readily be seen.

To accomplish this function, the system combines data derived from each ping, upon its return, with data in the status file, derived from previous pings. The information flow is summarized in Fig. 2.

3.2 PRELIMINARY DATA REDUCTION

The normalized output of the sonar signal processor from the current ping cycle is processed by the box labeled "Preliminary Data Reduction." This processor has three purposes. First, the data received is grouped into single ping event packages and converted from analog to digital format, if necessary. The event packages created contain information such as the time, amplitude, and Doppler of the event. Second, the box performs a thresholding function. The threshold can be set sufficiently low that any information of interest will pass, and still be sufficiently high to allow reduction of the volume passed by 90 to 99%. Third, the event amplitude is mapped to logarithm of the likelihood ratio. The specific transformation is dependent upon the particular sonar signal processing system being used, The techniques for the determination of mapping are discussed in Appendix B. Though the exact transformation equations are often complex, in every example considered to the transformation can be accurately approximated by a linear transformation equation of the form

$$l_i = A + B \cdot P_i \ .$$

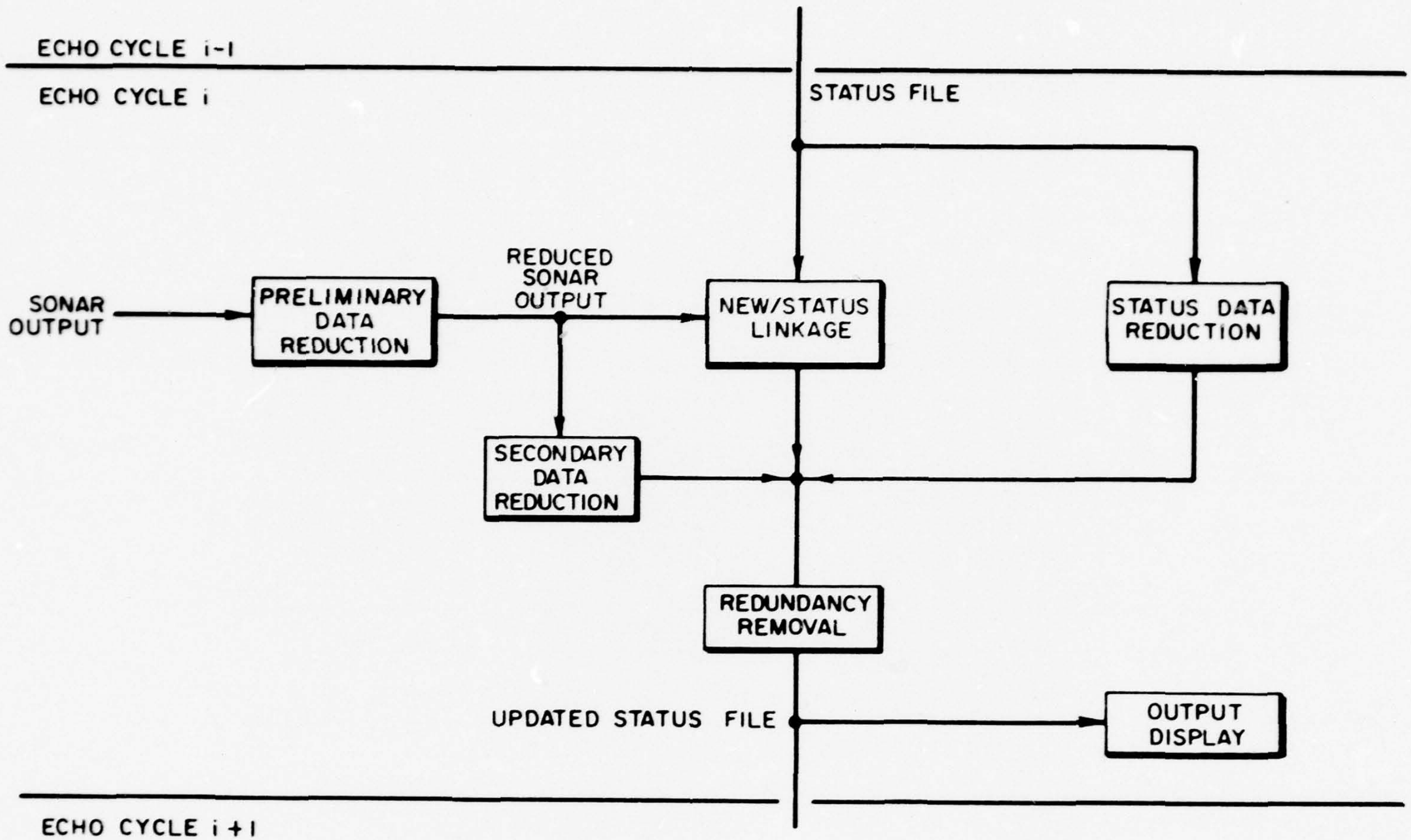

FIG. 2 - GENERAL ORGANIZATION OF LOGIC AND DATA FLOW

TRACOR DWG. A6-105-22
9/14/67 REEDER/SL

The output of the "Preliminary Data Reduction" box goes to two places, New/Status Linkage, and Secondary Data Reduction.

3.3 NEW/STATUS LINKAGE

The "New/Status Linkage" box receives two inputs; one is the reduced sonar output from Preliminary Data Reduction, and the other is the status file input. Where the reduced sonar output contains single ping event packages, the status file contains multi-ping event packages.

3.3.1 Status File

For each event logged in the status file, four functional quantities are maintained:

1. The position vector from the preceding echo cycle.
2. The expected position vector for the next echo cycle.
3. The variance vector for the next echo cycle.
4. The joint log likelihood ratio for the event.

For the three vectors, the number of dimensions required depends on the sonar system. For example, an omni-directional sonar system with no Doppler measuring capability would require only a one-dimensional vector indicating range. A more sophisticated sonar, capable of measuring range, bearing, and Doppler, however, would require a three-dimensional vector. The variance vector defines the volume, centered on the expected position vector, within which legitimate linkages can occur for the next echo cycle. In the one-dimensional situation, for example, the variance vector would define an interval of interest along the range scale.

3.3.2 Linkage Process

The "New/Status Linkage" process compares each status file entry with the single-ping event packages from the reduced sonar output. If the single ping event position vector lies within the volume of suspicion of a status file entry, the single-ping event is said to be linked with the status file entry. For linked events, information from the status file entry can be processed in conjunction with the single-ping event position vector to generate a new estimated position. This information can also be processed to obtain a new variance vector.

The joint log likelihood ratio of the new multi-ping event can be evaluated by adding the log likelihood ratios of the multi-ping event in the status file and the single-ping event and subtracting the logarithm of the number of independent samples in the volume of suspicion of the status file entry. If this new joint log likelihood ratio is less than the lower decision threshold, the Hypothesis H_0 "The track is noise" is accepted and the track is discarded. Otherwise, the multi-ping event is entered in the new status file.

A status file entry is allowed to link with all events which fall within its volume of suspicion. Similarly, a single-ping event can fall in the volumes of suspicion of several status file entries and hence be linked in several ways. The procedure allows many incorrect linkages, but, since all incorrect linkages will yield a track which will be integrating noise, the log likelihood ratio will tend to decrease, and the track will eventually be dropped.

3.4 SECONDARY DATA REDUCTION

The reduced sonar output is also processed by the "Secondary Data Reduction" box, which may make entries into the status file based upon amplitude of single-ping events. The criterion for making an entry in the status file is that the log likelihood ratio for the single event be greater than the lower decision threshold corresponding to Hypothesis H_0 "The track is noise".

The log likelihood ratio and present position vector are entered in the appropriate positions of the status file entry.

The estimated position vector and a rather gross variance factor are determined and entered in the status unit. These factors depend upon whether Doppler information is available.

During the first echo cycle, when there are no previously acquired status units, the "Secondary Data Reduction" box is the only one in operation. In each echo cycle it is here that new tracks are started. The entire process does not preclude a single large echo return being entered in the status file and, indeed, being placed on the output display immediately.

3.5 STATUS DATA REDUCTION

In addition to furnishing information to the "New/Status Linkage" box, the Status File is processed by the "Status Data Reduction" box. This procedure propagates a status file entry associated with a strong track even though the echo for the track may not occur during a particular echo cycle. This avoids losing a well-established track because of a single miss. The status data reduction procedure degrades the joint log likelihood ratio of each status file entry. If the degraded log likelihood ratio still exceeds the required lower threshold, then an entry is made propagating the track entry. The variance vector is also adjusted to accommodate the increased uncertainty of the target position.

3.6 REDUNDANCY REMOVAL

Because of the independent operation of the three status file entry-generating procedures, it is possible to develop several status file entries with the same predicted and present locations. These entries are redundant since they really describe different tracks that have merged into identical tracks. The redundancy removal procedure scans the entries to detect these redundancies and removes all except the single entry with the largest track log likelihood ratio.

3.7 OUTPUT DISPLAY

The updated Status File furnishes information for the output display section. The operator should have control over the display threshold on the log likelihood ratios stored in the status unit. If a log likelihood ratio exceeds this threshold, the Hypothesis H_1 "The track is a target" is accepted and the status unit is displayed. By varying this threshold an operator can look at the status file in more detail if something of interest turns up. A change in the output display threshold immediately changes what is on the display. There is no need to wait for events to start accumulating on the display, since the accumulation has already occurred, and the data are in storage.

3.8 SUMMARY

The output from the sonar processor for the most recent echo cycle is combined with information from previous echo cycles to form an updated status file that contains information for display and for the next echo cycle. The process does not include a fundamental specification of the number of pings over which ping integration will be carried. Indeed, a single status file entry could represent a track which has been propagated for an indefinite number of pings. A change in the lower decision threshold for rejecting status file entries affects only the amount of processing and storage, but does not necessarily affect the degree of clutter shown in the output display. This offers a significant improvement over some conventional approaches which integrate ping-to-ping on the display and which necessarily operate with a clutter rate sufficient to allow small echoes to paint the display at least a small amount so that the ping-to-ping integration process can get started.

4. IMPLEMENTATION OF A SIMPLIFIED PROBLEM

4.1 PROBLEM DESCRIPTION

In order to illustrate the effectiveness of this technique a simplified problem was considered, in which the output data were taken from a single sonar beam with no Doppler information.

With this simplification, the vectors described in the previous section become one-dimensional; for example, the position vector represents range only. The estimated position in the next ping cycle is determined by straight line extrapolation of the previous peak and present peak positions. The variance in expected peak position is assumed to take on two fixed values. The larger value is used when a track is initially started by the Secondary Data Reduction section and no range rate information is available. The small variance is used after a linkage has occurred and range rate is known. The volume of suspicion in the next echo cycle is given by adding and subtracting the variance from the predicted location. The output statistics of the linear correlator have been analyzed for signal plus noise and for noise alone. It is possible to form the likelihood ratio of a single event and, by a certain approximation, to reduce the log likelihood ratio to a linear equation that is quite accurate for the region of interest in these calculations.

4.2 OUTPUT DISPLAY

The output display can take many forms. The one chosen for this example is convenient to produce on the computer printer. The distance across the computer page is proportional to range, and successive ping cycles are displayed on succeeding lines of the page. The numbers printed correspond to the amount by which the log likelihood ratio exceeds the display threshold, truncated to an integer. This value is not directly related to units of standard deviation above the mean or any other commonly used measure of signal strength; however, the greater the value of the log likelihood value of a track, the more certainty that it is a target track. Since the display shows the amount by which the threshold is exceeded, an increase in threshold will cause the same data to appear to decrease in value, but the true value of the log likelihood ratio remains the same and is given by threshold plus the displayed number.

4.3 EXAMPLES

The first tracking problem was generated by adding three FM slides at appropriate times to simulate targets with range rates of 4 knots closing, 0 knots, and 5 knots opening. The output signal-to-noise ratios are 13.7 dB, 9.5 dB, and 10.9 dB, respectively. The tracks are designated in the order 1, 2, and 3. The larger fixed variance was selected to accommodate targets with range rates of ± 10 knots and the smaller variance to accommodate targets with a variation of ± 1 knot in the established range rate. In Fig. 3 the display threshold on the log likelihood ratio is 2.94, corresponding to an approximate probability 0.05 for accepting a noise track as a target. In Fig. 4 this threshold is 9.15 and the corresponding probability is 0.01. In each case the assumed output signal-to-noise ratio was 13.7 dB. Track 1, with a matched output signal-to-noise ratio, was tracked very well. Track 3, with a signal-to-noise ratio about 3 dB lower, performed reasonably well. It took considerably more ping samples to reach a decision on track 2, with a signal-to-noise ratio of 4 dB less than the assumed value.

In the next example actual sea data were processed. The data consist of the output of a single receiving beam of an AN/SQS-26 sonar system for 92 ping cycles. The target was closing, and had an average output signal-to-noise ratio of 16 dB. Two displays are presented; both have a display threshold of 9.15 and a larger variance of ± 10 knots. The first display, in Fig. 5, has a smaller variance of ± 2.5 knots. The second, in Fig. 6, has a smaller variance of ± 1 knots and a slightly larger assumed signal-to-noise ratio of 17 dB which is closer to the average output signal-to-noise ratio than 13.7 dB used in the preceding example.

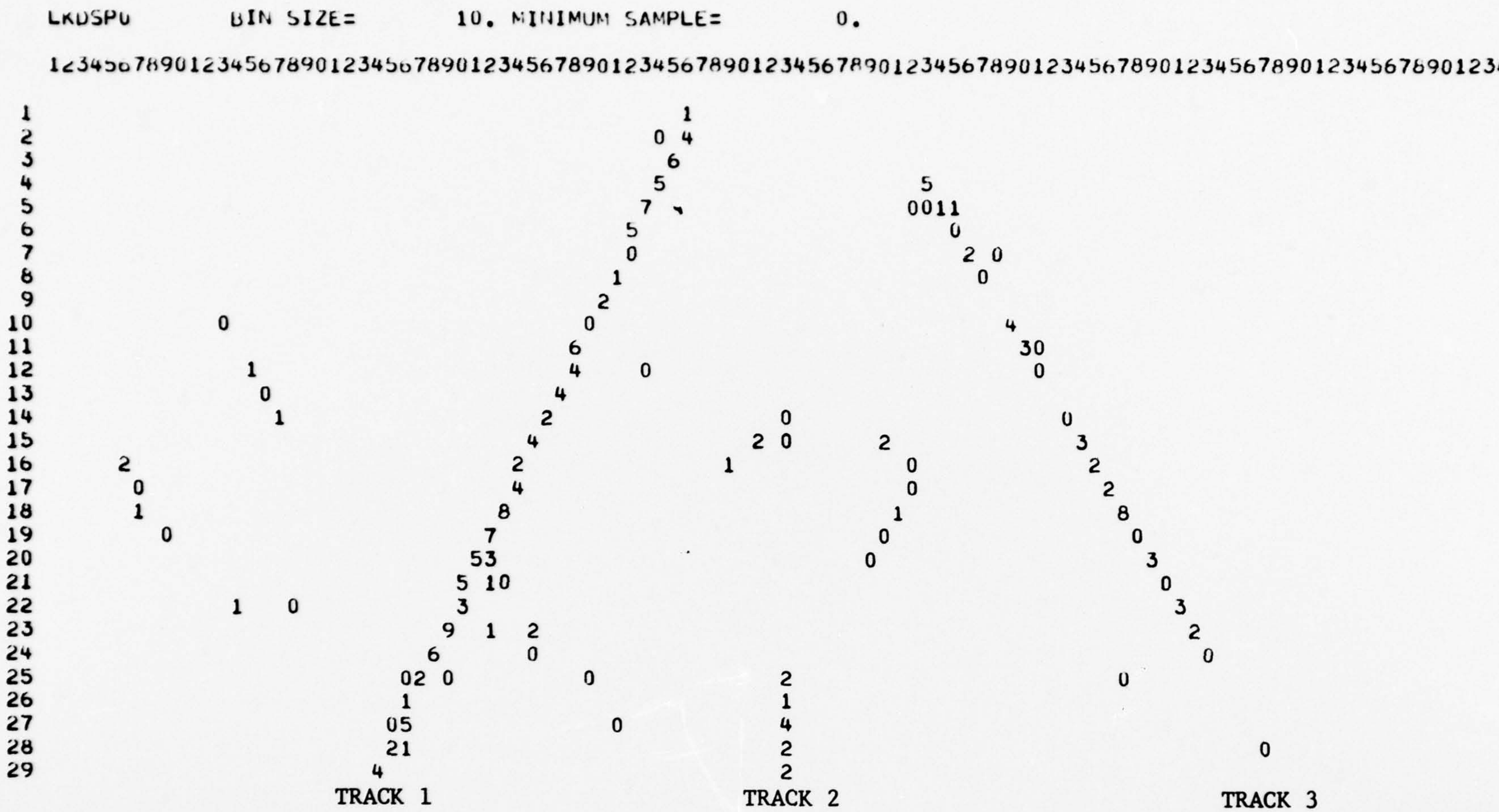

FIG. 3 THREE SIMULATED TARGET TRACKS. DISPLAY THRESHOLD ON LOG LIKELIHOOD RATIO IS 2.94

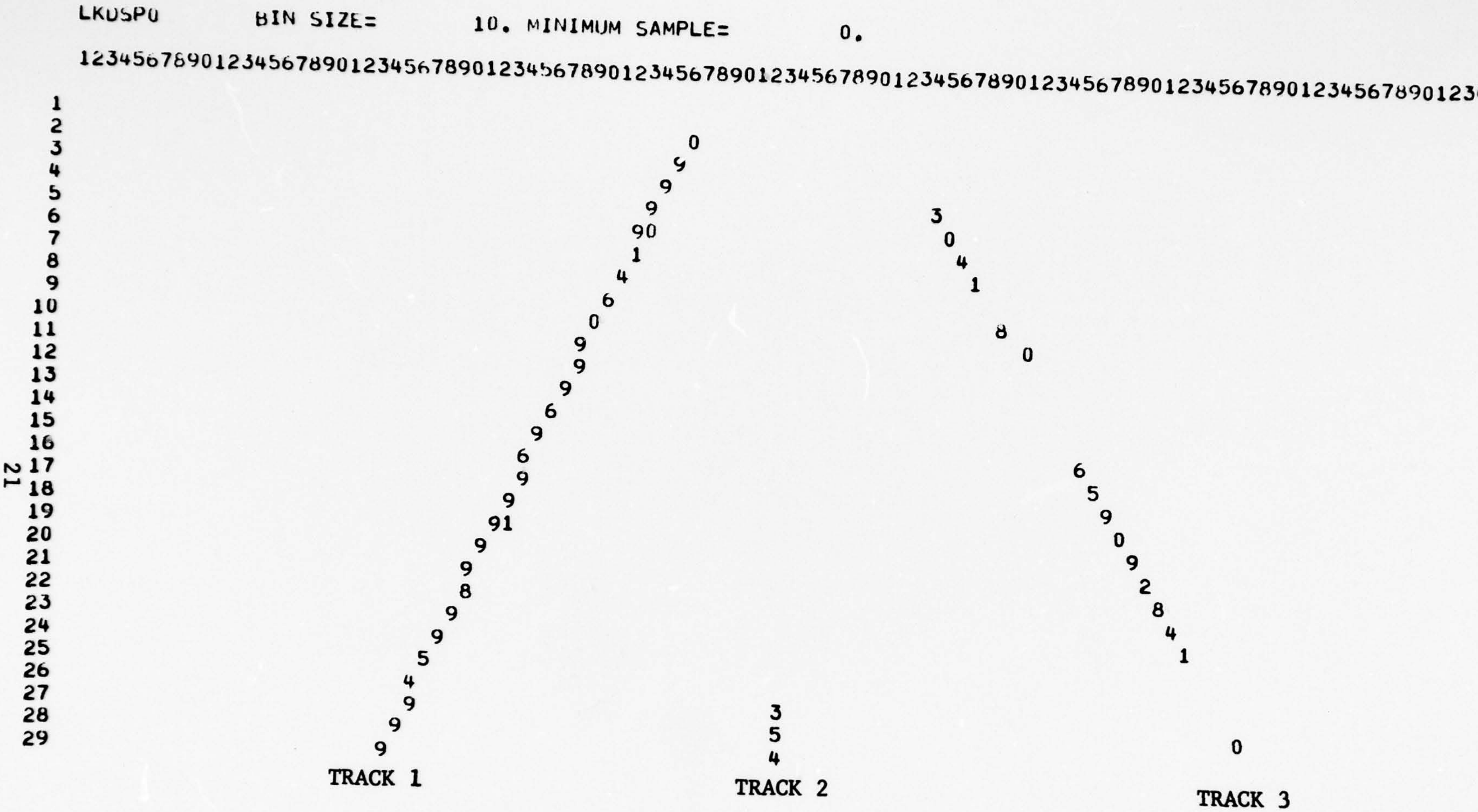

FIG. 4 THREE SIMULATED TARGET TRACKS. DISPLAY THRESHOLD ON LOG LIKELIHOOD RATIO IS 9.15

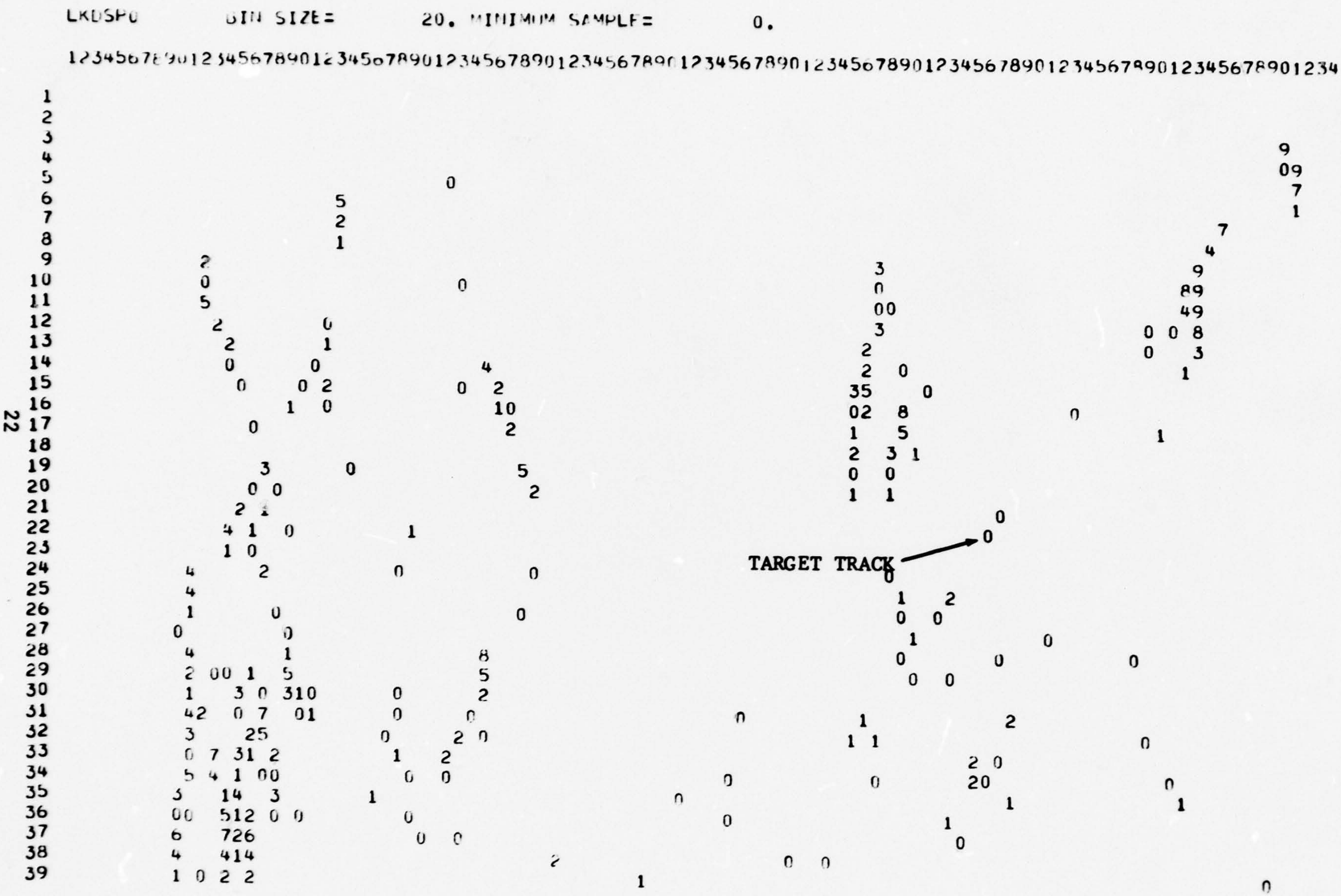

FIG. 5 SEA TEST DATA. ASSUMED OUTPUT S/N = 13.7 dB

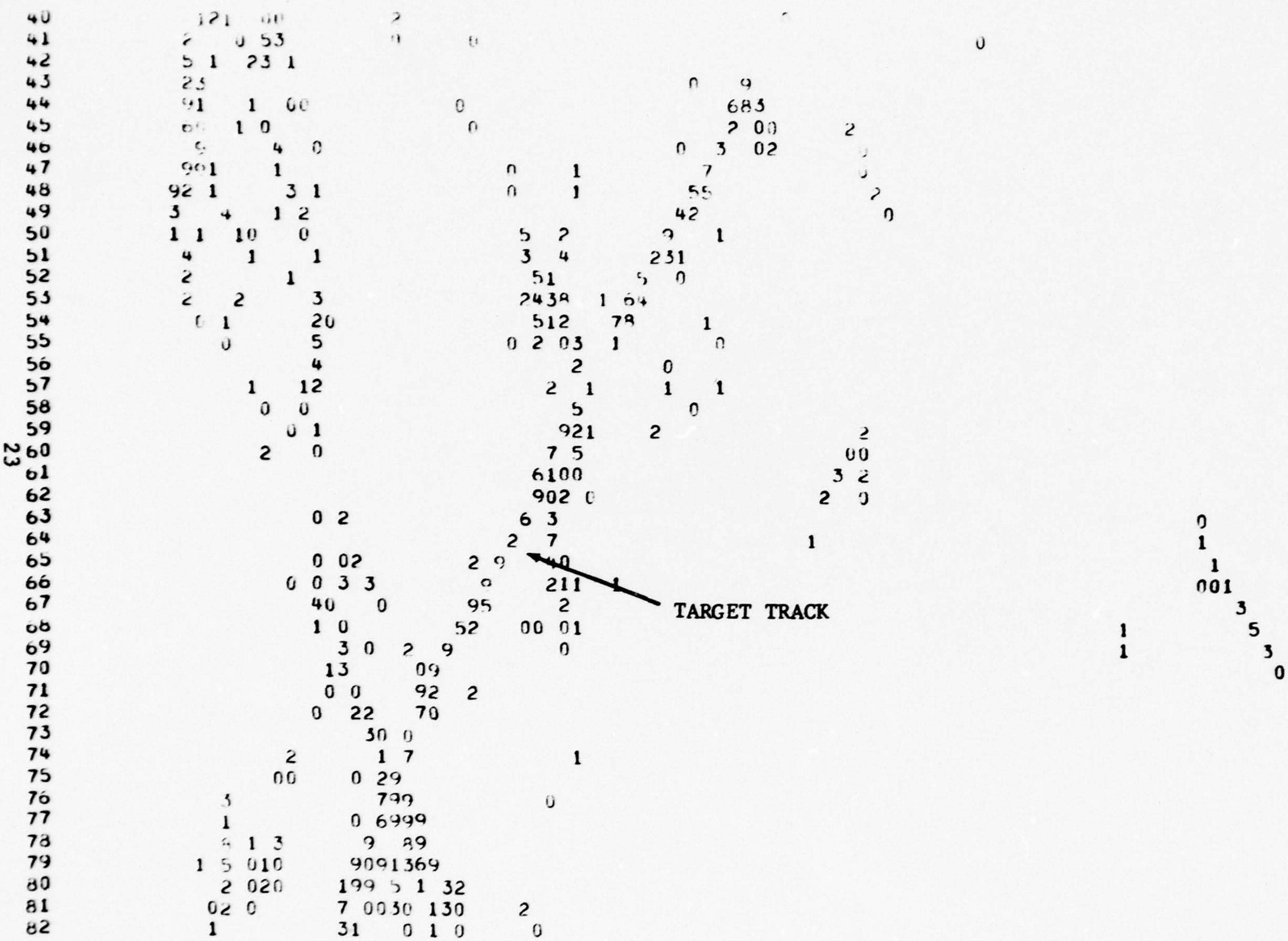

FIG. 5 (Cont'd.)

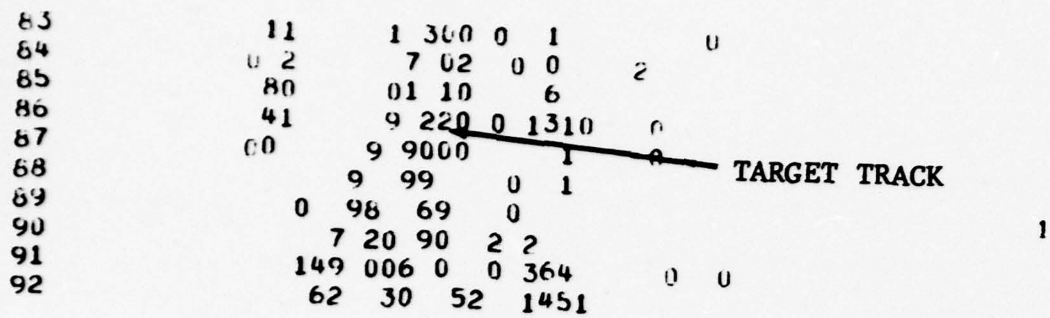
83 11 1 300 0 1 0
84 0 2 7 02 0 0 2
85 80 01 10 6
86 41 9 220 0 1310 0
87 00 9 9000 1 0
88 9 99 0 1
89 0 98 69 0
90 7 20 90 2 2 1
91 149 006 0 0 364 0 0
92 62 30 52 1451
TARGET TRACK

FIG. 5

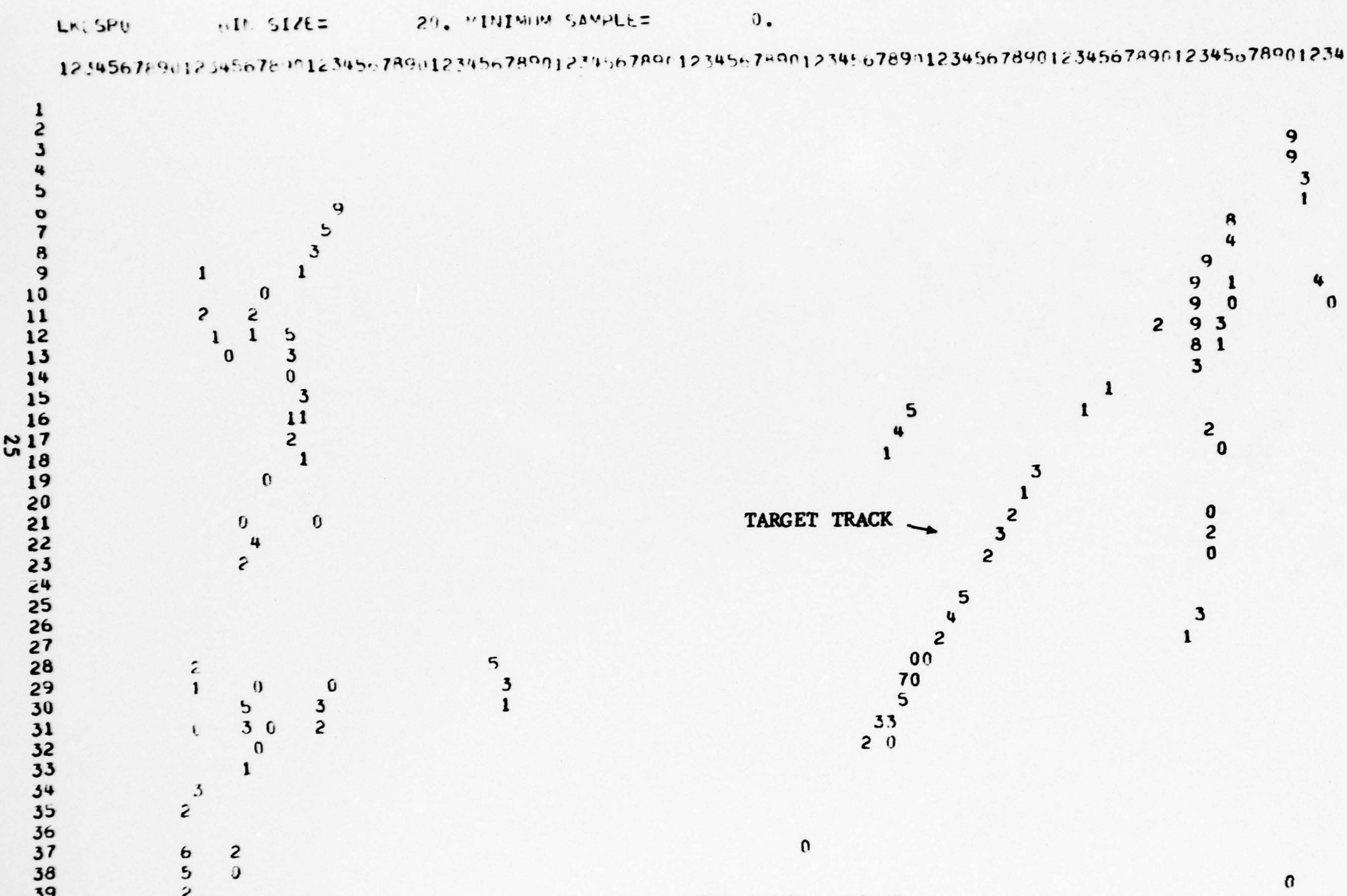

FIG. 6 SEA TEST DATA. ASSUMED OUTPUT S/N = 17 dB

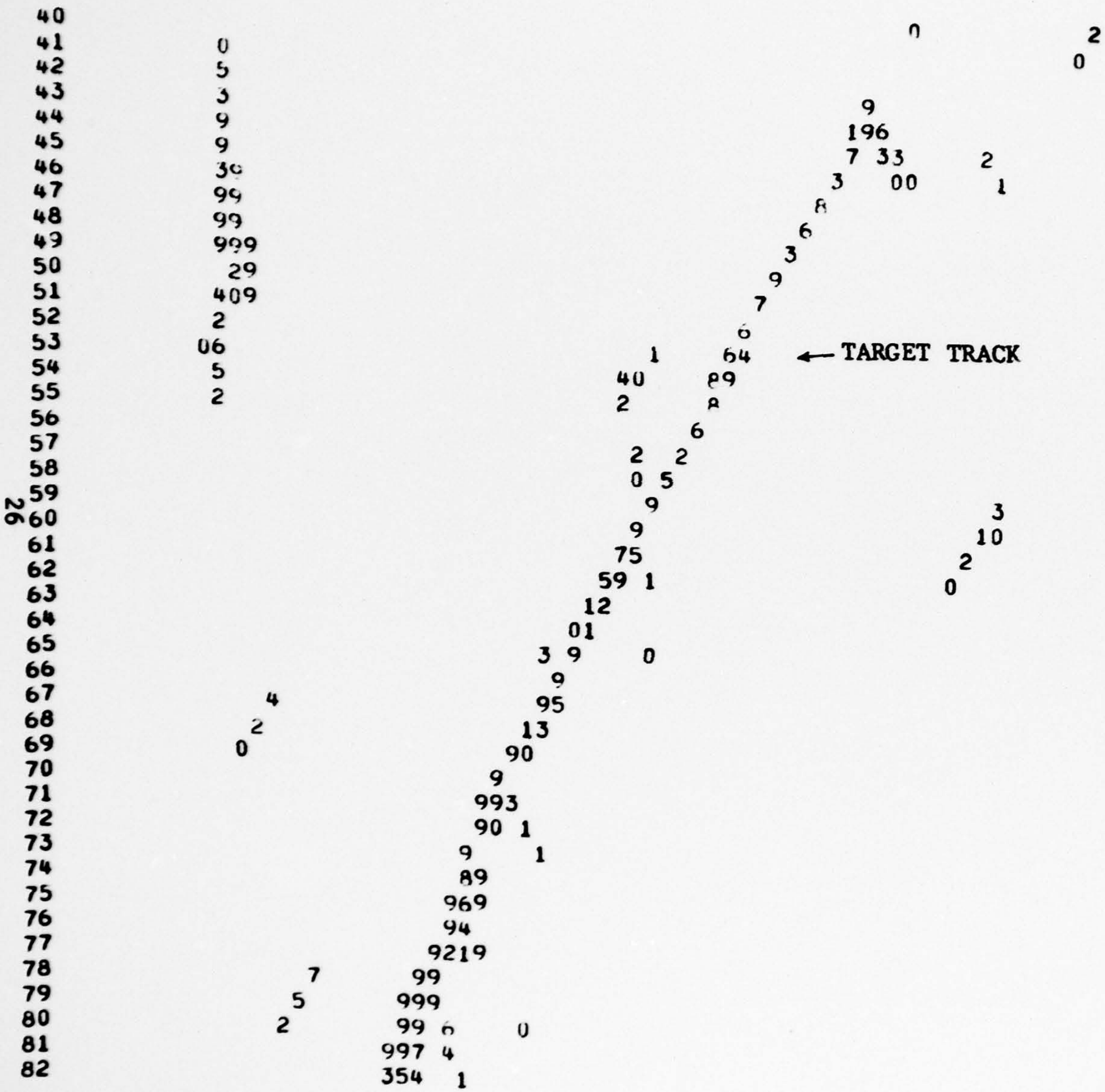

FIG. 6 (Cont'd.)

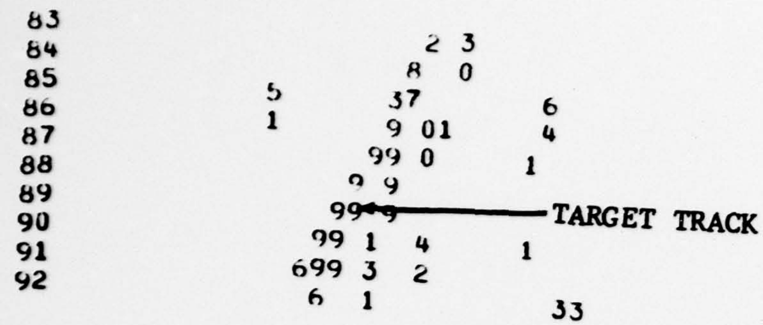

FIG. 6 (Cont'd.)

APPENDIX A

SEQUENTIAL TESTING OF HYPOTHESES

A.1 STRUCTURE OF SEQUENTIAL TEST

This appendix discusses some aspects of the theory of sequential testing of hypotheses. Suppose it is desired to make a choice between two hypotheses H_o and H_1. Each hypothesis has an associated probability density function, $f_o(x)$ and $f_1(x)$. To reach the decision, a sequence of observations x_1, x_2, x_3, . . ., are taken, the subscripts indicating their order, and two positive thresholds $A > 1$ and $B < 1$ are chosen. The likelihood ratio for m observations,

$$L_m = \prod_{i=1}^{m} \frac{f_1(x_i)}{f_o(x_i)} ,$$

will be used in the sequential test in the following manner. New observations are taken, and the ratios L_1, L_2, L_3, . . . are calculated as long as

$$B < L_m < A . \qquad \text{(A-1)}$$

If, for some m, L_m is less than or equal to B the test stops and the hypothesis H_o is accepted. If L_m is greater than or equal to A, the hypothesis H_1 is accepted.

A.2 TERMINATION OF THE TEST

While it may seem that the test could go on forever, the probability that it will end is 1. In order to show this, let

$$\ell = \log[f_1(x)/f_2(x)].$$

now, ℓ will have some probability density function, $g(\ell)$, determined by the density of x, which is not necessarily $f_0(x)$ or $f_1(x)$. The observations $x_1, x_2, x_3, \ldots$, determine a sequence $\ell_1, \ell_2, \ell_3, \ldots$, and the inequality (A-1) becomes

$$\log B < \sum_{i=1}^{m} \ell_i < \log A \,, \qquad \text{(A-2)}$$

where log B is negative and log A is positive. Let C = log A - log B and p be the area under $g(\ell)$ between -C and C. If any ℓ_i is outside of the interval -C to C, one of the inequalities in (A-2) will be violated either at that stage or at a previous stage. Assume that for step m - 1

$$\text{Log } B < \sum_{i=1}^{m-1} \ell_i < \log A \,, \qquad \text{(A-3)}$$

and $\ell_m > C = \log A - \log B$. Add the left hand inequality of (A-3) to the second inequality

$$\log B + (\log A - \log B) < \sum_{i-1}^{m-1} \ell_i + \ell_m$$

$$\log A < \sum_{i=1}^{m} \ell_i \;.$$

The inequality (A-2) is violated similarly if $\ell_m < -C$:

$$\sum_{i=1}^{m} \ell_i < \log B \;.$$

Hence, a necessary condition for (A-2) to hold for all m is that ℓ_i fall between -C and C. It should be noted that the inequalities can be violated even with this condition. The probability of a ℓ_i being between -C and C is the area under

$g(\ell)$ in that interval, that is, p. The probability that the first m observations will fall in that interval is p^m. Since p is less than 1 this probability approaches 0 as m increases, or the probability of termination goes to 1. If $g(\ell)$ is 0 outside -C and C, a new variable λ may be defined by letting λ_1 be the sum of the first r ℓ's, λ_2 the sum of the next r ℓ's, and so forth, choosing r large enough that the probability of a value of λ_i outside of -C to C is not zero.

A.3 THRESHOLDS

Let α be the desired probability that the hypothesis H_o will be rejected when, in fact, it is true. Since H_o will be rejected if L_m exceeds A before it is less than B, A and B must be chosen so that

$$\alpha = P(L_1 \geq A) + P(B < L_1 < A,\ L_2 \geq A)$$

$$+ P(B < L_1 < A,\ B < L_2 < A,\ L_3 \geq A) + \ldots \quad \text{(A-4)}$$

Similarly, if β is the desired probability that H_o will be accepted when H_1 is true, A and B must also satisfy

$$\beta = P(L_1 \leq B) + P(B < L_1 < A,\ L_2 \leq B)$$

$$+ P(B < L_1 < A,\ B < L_2 < A,\ L_3 \leq B) + \ldots \quad \text{(A-5)}$$

Using the density $f_o(x)$ in (A-4) the probabilities can be calculated. Similarly $f_1(x)$ can be used in (A-5). This gives α and β as functions of A and B. The inverse relation could be solved for A and B in terms of α and β.

As can be imagined, this could be a major computational project. Fortunately, there are simple and accurate approximations available:

$$A \cong \frac{1 - \beta}{\alpha} , \tag{A-6}$$

$$B \cong \frac{\beta}{1 - \alpha} . \tag{A-7}$$

Let Λ be the space of all observational sequences $x_1, x_2, x_3, \ldots$, and Λ_o be the subspace of sequences that result in the decision H_o. That is, for any sequence in Λ_o, for some m, and for every $n < m$,

$$B < \prod_{i=1}^{n} \frac{f_1(x_i)}{f_o(x_i)} < A ,$$

and

$$\prod_{i=1}^{m} \frac{f_1(x_i)}{f_o(x_i)} \leq B ,$$

or

$$\lambda_1 = \prod_{i=1}^{m} f_1(x_i) \leq B \prod_{i=1}^{m} f_o(x_i) = B\lambda_o .$$

Now, the integral of λ_1 over Λ_o is the probability that H_o will be accepted when H_1 is really true; hence, it is equal to β. The integral of λ_o over Λ_o is the probability that H_o will be accepted when H_o is true; hence, it is equal to $1 - \alpha$. Therefore,

$$\beta \leq B(1 - \alpha) . \tag{A-8}$$

Similarly, if the Λ_1, the subspace of all sequences that result in the decision H_1, is considered,

$$A\alpha \geq 1 - \beta. \tag{A-9}$$

The error in the use of (A-6) and (A-7) is a result of using equalities in (A-8) and (A-9). This error has been shown to be small if α and β are less than 1/2[1].

A.4 RULES FOR SEQUENTIAL TEST

The sequential test may be summarized in a few simple rules. If it is desired to make a decision between hypothesis H_o, with probability density function $f_o(x)$, and hypothesis H_1, with probability density function $f_1(x)$, and if α is the error probability of choosing H_1 when H_o is true and β is the error probability of choosing H_o when H_1 is true, then compute

1. A, where $A = (1 - \beta)/\alpha$,
2. B, where $B = \beta/(1 - \alpha)$,
3. Take an observation x_1 and calculate

$$L_1 = \frac{f_1(x_1)}{f_o(x_1)} ,$$

4. If $L_1 \leq B$ accept H_o,
5. If $L_1 \geq A$ accept H_1,
6. If $B < L_1 < A$, take another observation and calculate

$$L_2 = \frac{f_1(x_1)f_1(x_2)}{f_o(x_1)f_o(x_2)} ,$$

---

[1] Wald, A., Sequential Analysis, Wiley, New York (1947).

7. Repeat 4 and 5 with L_1 replaced by L_2,

8. Continue taking observations until 4 or 5 are satisfied.

In many cases it is easier to work with logarithms of A, B, and L_m.

APPENDIX B
DERIVATION OF A LOG LIKELIHOOD RATIO

The purpose of this appendix is to derive a linear approximation of the log likelihood ratio for the processor used in the examples of Section 4.

After the correlation process, the signal peaks are similar to a short CW pulse, and the linear rectifier and perfect averager act as an envelope detector for this pulse. The probability density function of the envelope of a sine wave added to Gaussian noise is given by S. O. Rice[1] as

$$p(x,S) = \frac{x}{N^2} \exp(-(x^2+S^2)/2N^2)\, I_0((xS)/N^2),$$

where x is the height of the envelope,

N is the standard deviation of the noise before envelope detection,

S is the amplitude of the sine wave, and

$I_0()$ is the zero order modified Bessel function.

The likelihood ratio L(x) is given by

$$L(x) = \frac{p(x,S)}{p(x,0)} = I_0(x\, S/N^2) \exp(-S^2/2N^2). \qquad \text{(B-1)}$$

A value of four noise standard deviations above the noise mean will be chosen as an average peak height. Since the envelope will be subjected to thresholding to eliminate small values of x and the assumed value for S will be about $4 \cdot N$, the smallest argument to be expected for I_0 is approximately 10.

---

1 Rice, S. O., "Mathematical Analysis of Random Noise." Bell System Technical Journal, Vol. 24, 1945, p. 100.

This large value allows the *asymptotic expansion of* I_0 *to be* used,

$$I_0(x\ S/N^2) \simeq \frac{\exp(x\ S)/N^2}{2\pi x\ S/N^2} \quad .$$

Substituting this expression in Equation (B-1) and taking the natural logarithm yields

$$\ell(x) = \log L(x) \simeq x\ S/N^2 - S^2/2N^2 - \tfrac{1}{2}\log(2\pi x\ S/N^2)\ . \quad \text{(B-2)}$$

A plot of the above function shows an almost straight line in the region of interest greater than one noise standard deviation above the noise mean. A linear approximation will be found so that its slope and value equals that of Equation (B-2) at the average peak height of four standard deviations above the mean.

As given by Rice[2] the output noise mean is $\sqrt{N}\frac{\pi}{2}$ and the output noise standard deviation is $N\sqrt{2-\frac{\pi}{2}}$. If r is the peak height in units of the standard deviation about the noise mean, then

$$x = r\sqrt{2-\frac{\pi}{2}} + \sqrt{\frac{\pi}{2}} \quad ,$$

or

$$x = (0.655\ r + 1.254)N.$$

Using the average peak height of four standard deviations,

$$x = 3.87\ N.$$

The slope of the log likelihood ratio is given by

$$a = \ell'(x) = S/N^2 - \frac{1}{2x} \quad ,$$

and, at x = 3.87 N,

---

[2] S. O. Rice, op.cit., p. 100.

$$a = (S/N - 1/6.74)/N;$$

also

$$b = \ell(3.87\ N) - a \cdot 3.87\ N$$

$$= 3.87 \cdot S/N - \tfrac{1}{2}(S/N)^2 - \tfrac{1}{2} \log (6.75\pi\ S/N).$$

Using these values, Equation (B-2) can be approximated with

$$\ell(x) \simeq ax + b\ .$$

UNCLASSIFIED

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) TRACOR, Inc. 6500 Tracor Lane Austin, Texas 78721	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

"COMPUTER UTILIZATION OF SEQUENTIAL HYPOTHESIS TESTING FOR DETECTION AND CLASSIFICATION OF SONAR SIGNALS"

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Memorandum

5. AUTHOR(S) (Last name, first name, initial)

Reeder, Hugh A.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
27 October 1967	40	3

8a. CONTRACT OR GRANT NO. NObsr-93352 b. PROJECT NO. SF001-03-16 c. Task 8523 d.	9a. ORIGINATOR'S REPORT NUMBER(S) 67-717-U 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. AVAILABILITY/LIMITATION NOTICES

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Naval Ship Systems Command Department of the Navy Washington, D. C. 20360

13. ABSTRACT

Active sonar systems furnish the operator large quantities of data to examine for possible targets. In order to detect small target returns, the display thresholds must be set such that clutter is significant. On a PPI display this clutter makes it difficult to visually integrate successive pings.

This report shows a technique utilizing a digital computer to examine the data out of a sonar system, reduce its volume, perform ping-to-ping integration, make simple decisions, and generate useful displays. The process takes advantage of a statistical decision procedure called sequential hypothesis testing.

DD FORM 1 JAN 64 1473

UNCLASSIFIED

Security Classification

UNCLASSIFIED
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Detection (Sonar)						
Classification (Sonar)						
Likelihood Ratio						
Sequential Hypothesis Testing						
Displays (Sonar)						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2*a*. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2*b*. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7*a*. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7*b*. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8*a*. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8*b*, 8*c*, & 8*d*. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9*a*. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9*b*. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through ______________________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through ______________________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through ______________________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.

UNCLASSIFIED
Security Classification