Sequential Partition Detectors
With Dependent Sampling

A Paper Presented at the
1977 IEEE International Symposium
on Information Theory.

Roger F. Dwyer
Surface Ship
Sonar Systems Department
Ludwik Kurz
Polytechnic Institute of New York

17 Nov 1977

NUSC
NAVAL UNDERWATER SYSTEMS CENTER
Newport, Rhode Island & New London, Connecticut

Approved for public release; distribution unlimited.
PREFACE

This document was prepared under Project No. A75028, IR/IED investigation on "Adaptive Nonparametric Sequential Detectors." The Sponsoring Activity is Naval Material Command, Program Manager, J. H. Probus (MAT 035). This document was also prepared under NUSC Project No. A67840, "Escort Towed Array Sonar Systems," Principal Investigator, G. T. Finn (Code 3324); and Project No. A20005, "Adaptive Detection and Estimation Processor (ADEC)," Principal Investigator, A. A. Filippini (Code 333). The Sponsoring Activity is Naval Sea Systems Command, Program Manager, E. Landers (SEA 06H2).

REVIEWED AND APPROVED: 17 November 1977

W. A. VonWinkle
Associate Technical Director
for Technology

The authors of this document are located at the New London Laboratory, Naval Underwater Systems Center, New London, Connecticut, and the Polytechnic Institute of New York, Brooklyn, New York, respectively.
This document presents the oral version of a presentation given at the 1977 IEEE International Symposium on Information Theory, held at Cornell University, Ithaca, New York, 10-14 October. The paper is concerned with the effects of dependent sampling on sequential partition detectors. As stated in this document, sequential partition detectors show improved efficiency under dependent sampling. The text of this presentation will be available in the literature shortly.
INTRODUCTION

The purpose of this talk is to discuss the effects of dependent sampling on sequential partition detectors. A technique will be presented for adjusting the thresholds under Q-dependent sampling in order to maintain the same error probabilities as in the independent sampling case. The results will be consistent with fixed sample detectors in that the effects of dependent sampling depend upon the normalized correlation function of noise, and the cost of partitioning can be recovered by rapid sampling.

-Vu-graph 1 please-
q -DEPENDENT PROCESS

\[ \bar{X} \rightarrow \text{PARTITION SPACE} \rightarrow \sum_{i=1}^{N_0} T_i \rightarrow T_{jN_0} \]

ESTIMATE QUANTILES
I.I.D. SAMPLES

WHERE

\[ T_{jN_0} = \sum_{i=1}^{N_0} \sum_{k=1}^{M} b_{ik} N_{ik} \quad j=1,2,\ldots, N \]

VU-GRAPH 1
VU-GRAF 1

The basic structure of the sequential partition detector, which will be referred to as SPD, henceforth, is composed of a partitioning device which reduces the data samples into $M$-intervals based on knowledge of $M-1$ quantiles of the unknown noise distribution. For the purposes of this discussion, we will assume that the quantiles are known; however, Kersten and Kurz have shown that the quantiles can be estimated in real time using stochastic approximation. Therefore, the SPD is adaptive in its general form. Implicit in the discussion will be the assumption that the signal is constant at the input to the SPD and the noise is additive. Generalizations are possible to random and nonrandom time varying signals.

The problem is to decide if signal plus noise or noise only is present for fixed error probabilities under dependent sampling. The output of the partition is transformed from a $q$-dimensional space to a univariate space by summing over $N$ sub $0$ samples. The output of interest is represented by $T_j, N$ sub $0$ and will be referred to as the subsample where $j$ indexes the output from $1$ to $N$, $N$ being a random variable.

$B_k, k$ ranging over the $M$ intervals, are called scores or weighting constants. In general, they are nonlinear functions of time. For this presentation, the $B_k$'s will represent a log likelihood ratio at the quantile locations for a designed signal-to-noise ratio, under a selected alternative for independent sampling.

$N$ sub $I_k$ represents the interval the $I$th data sample falls in. Related work is given by Woinisky and Kurz, Kassam and Thomas, and others (for fixed sample detectors).

-Next VU-graph please-
**WALD'S FUNDAMENTAL IDENTITY**

\[
E \left[ e^{T_N t} \Phi(t) \right] = 1
\]

**WHERE**

\[
T_N = \sum_{j=1}^{N} T_{jN_0} \quad (a', b')
\]

\[
\Phi(t) = E \left[ e^{T_{jN_0}} \right] = 1
\]

\[
t = t_D(\theta) = -2 \left\{ \frac{E \left[ T_{jN_0} \right]}{\sigma^2} \right\}
\]

\(\theta \) small
Wald’s fundamental identity plays a central role in sequential detection. Both the operating characteristic function, or power function, and the average sample number are derived from this equation. For this relationship to hold in dependent sampling, the subsamples must be independent.

\( T_{\text{SUB N}} \) represents the sum of independent subsamples; it is the essential test statistic. \( A' \) and \( B' \) are the upper and lower thresholds, respectively.

If \( T_{\text{SUB N}} \) is between \( A' \) and \( B' \), the test continues. The signal plus noise condition is chosen if the upper threshold is reached. However, if the lower boundary is crossed, then the noise-only case is the choice.

\( \Phi(t) \) is the moment generating function of the subsample and \( t \) is a function of the signal-to-noise ratio. The problem is to find a solution for \( t \) not equal to zero which sets the moment generating function equal to one. This is a difficult problem in general. However, for small signals, higher order cumulants of the moment generating function are negligible, and a solution for \( t \) is approximately equal to minus 2 times the ratio of the mean over the variance of the subsample.

-Vu-graph 3 please-
MEAN

\[ E[T_{jN_0}] = N_0 E[T_i] \]

VARIANCE

\[ \sigma^2_{T_{jN_0}} = \sigma^2_{T_i} N_0 \left\{ 1 + 2 \sum_{Q=1}^{N_0-1} \frac{1 - Q/N_0}{Q} R(Q) \right\} \]

WHERE

\[ R(Q) = \sum_{K=1}^{M} \sum_{L=1}^{M} b_k b_L E_Q(N_k N_L) - \left[ E(T_i) \right]^2 \]

\[ \sigma^2_{T_i} \]

VU-GRAPH 3
VU-GRAH 3

The mean of the subsample is just equal to \( N \, \text{sub} \, 0 \) times the mean of the output of the partition for independent samples. However, the variance, as shown in the text, is composed of the variance for independent samples times \( N \, \text{sub} \, 0 \) times a term which depends upon the partition structure. The term \( R \) of \( \Omega \), which is defined as the partition correlation function, expresses this relationship. It is essentially a function of the scores, which are known, and the joint expectation of the partitioning intervals over \( M \), which are not known, where \( \Omega \) indexes the dependence upon the sampling rate.

The evaluation of \( R \) of \( \Omega \), therefore, usually requires a two-dimensional integration with knowledge of the bivariate distribution. For the Gaussian distribution with normalized correlation function \( \rho \) of \( \Omega \), it is shown in the text that only a one-dimensional integration is required.

-VU-GRAH 4 please-
\[ R(Q) = \rho(Q) \]

\[ R(Q) = \frac{2}{\pi} \sin^{-1} \rho(Q) \]

\[ M = 6 \]

\[ M = 4 \]

\[ M = 2 \]
By expanding the bivariate Gaussian distribution and integrating over Rho of Q as mentioned previously, the partition correlation function was evaluated using numerical analysis. The results are given for M equal to 2, 4, and 6 in the figure. The abscissa represents the normalized correlation function of the noise at the input of the partition, and the ordinate depicts its output.

For M equal to 2 and assuming a zero mean Gaussian distribution, the quantile is located at zero; and for the noise-only case, the partition correlation function equals 2 over Pi times arc sin of Rho of Q, which is shown in the curve to the right. The same result has been obtained for the hard clipper, or sign test, for fixed sample detectors.

As M increases under the same conditions, the partition correlation function approaches Rho of Q (the curve on the left) very rapidly. This result also holds under the signal plus noise condition.

By substituting Rho of Q into the equation for the variance of the subsample given previously, it can be shown that an upper bound for the variance is obtained. This has not been proven for all distributions; however, under somewhat restricted conditions as noted in the text, the normalized correlation function for a Rayleigh distribution under a Lehmann alternative also gives an upper bound on the variance. Therefore, for subsequent results, we will only consider distributions where this result holds.

In general, however, the partition correlation function can always be calculated.

-VU-graph 5 please-
SOLUTION FOR $t$

$$t_D(\theta) = \frac{t(\theta)}{\left[ 1 + 2 \sum_{Q=1}^{N_0-1} (1 - q/N_0) p(q) \right]}$$

WHERE

$$t(\theta) = -2 \left\{ E \left[ T_i \right] \right\} \text{ FOR INDEPENDENT SAMPLES}$$

$\theta$ small

$$a' = \left[ 1 + 2 \sum_{Q=1}^{N_0-1} (1 - q/N_0) p(q) \right] a$$

$$b' = \left[ 1 + 2 \sum_{Q=1}^{N_0-1} (1 - q/N_0) p(q) \right] b$$

VU-GRAPH 5
In order to use the results of Wald, a solution for \( t \) under dependent sampling was needed. For the small signal case, the solution \( \ddt D \) of Theta is characterized by the equation shown here. In the numerator, \( \ddt \) of Theta represents the independent sample solution as a function of signal-to-noise ratio. The denominator, which reflects the amount of dependence, depends upon the normalized correlation function, or the noise spectrum, and the number of samples summed. From the fundamental identity, both the operating characteristic function and the average sample number will be functions of the amount of dependence. However, if the thresholds A and B derived for independent sampling are adjusted by precisely the amount of dependence as given in the denominator, the new thresholds \( \A' \) and \( \B' \) will give lower bounds on the error probabilities, and as \( M \) approaches infinity they will approach the same error probabilities as in the independent sampling case. If \( R \) of \( Q \) is used, the error probabilities will be identical for all \( M \).

For example, if Rho of \( Q \) equals zero, then \( \A' \) equals \( A \) and \( \B' \) equals \( B \); and, if Rho of \( Q \) equals 1, then \( \A' \) equals \( N \) sub 0 times \( A, B \), respectively, which is what would be expected for completely correlated samples.

After the thresholds have been adjusted, the operating characteristic function for dependent sampling will reduce to the independent sampling relationship. However, the average sample number will still be a function of the amount of dependence.
AVERAGE SAMPLE NUMBER

\[
\text{ASND} = \frac{1+2 \sum_{Q=1}^{N_0-1} \left(1 - \frac{Q}{N_0}\right) \rho(Q)}{N_0} \cdot \frac{b(\epsilon^{t(\theta)a_1} + a_1(1-\epsilon^{t(\theta)b}))}{E[T_i]} \cdot \frac{\epsilon^{t(\theta)a_1 - \epsilon^{t(\theta)b}}}{E[T_i] \neq 0}
\]

\[
\text{ASND} = \frac{1+2 \sum_{Q=1}^{N_0-1} \left(1 - \frac{Q}{N_0}\right) \rho(Q)}{N_0} \cdot \frac{-a_1 b}{\sigma^2 \cdot \frac{2}{T_i}} \cdot \frac{E[T_i] = 0}{E[T_i] \neq 0}
\]

VU-GRAPH 6
VU-GRAPH 6

After the thresholds have been adjusted, the average sample number for dependent sampling, shown here as ASND, has been factored into a dependent and an independent term. The boxed expression represents the effects of dependence normalized by \(N_{\text{sub } 0}\). The term to the right of the boxed expression represents the average sample number for independent sampling, which is a function of \(t\) of \(\theta\) (\(\theta\) being the signal-to-noise ratio parameter, thresholds \(A\) and \(B\), and the output of the partitioning device \(T_{\text{sub } I}\)) when the average value does not equal zero. For the case when the average value equals zero, the average sample number reduces to a relationship which depends upon the thresholds \(A\) and \(B\) over the variance of \(T_{\text{sub } I}\). In this case, the average sample number takes on its maximum value. Both the average value and variance of the partition output depend upon the quantiles and the signal-to-noise ratio in a complicated way. For a small signal-to-noise ratio, Dwyer and Kurz have shown that the terms involving the quantiles and signal-to-noise ratio can be decoupled for both the mean and variance. This allows the average sample number to be minimized as a function of the quantiles alone. This result also carries over to the dependent sample case.

-VU-GRAPH 7 PLEASE-
EFFICIENCY

SHIFT OF THE MEAN ALTERNATIVE

\[ \mathcal{E} = \frac{N_0 r}{N_0 + r} \left[ 1 + 2 \sum_{q=1}^{N_0-1} (1-q/N_0) p(q) \right] \]

\[ \sum_{k=1}^{N_0-1} \frac{[f(a_k)-f(a_{k-1})]^2}{F(a_k)-F(a_{k-1})} \]

\[ \frac{\int_{-\infty}^{\infty} \left[ \frac{f'(x)}{f(x)} \right]^2 f(x) \, dx}{\sum_{k=1}^{N_0-1} \sum_{k=1}^{N_0-1} (1-q/N_0) p(q)} \]

LEHMANN ALTERNATIVE

\[ \mathcal{E} = \frac{N_0 r}{N_0 + r} \left[ 1 + 2 \sum_{q=1}^{N_0-1} (1-q/N_0) p(q) \right] \]

\[ \sum_{k=1}^{N_0-1} \frac{[F(a_k) \ln F(a_k)-F(a_{k-1}) \ln F(a_{k-1})]^2}{F(a_k)-F(a_{k-1})} \]

VÜ-GRAF 7

THE NUMERATOR IN THE BOXED EXPRESSION IS A FUNCTION OF THE INCREASED SAMPLING RATE P SUB S OVER INDEPENDENT SAMPLING, AND REFLECTS SKIPPING P SUB S MINUS 1 SAMPLES BETWEEN EACH SUBSAMPLE TO ASSURE THEIR INDEPENDENCE. THE DENOMINATOR REPRESENTS THE EFFECTS OF DEPENDENT SAMPLING AS BEFORE.


-VU-GRAH 8 PLEASE-
EFFICIENCY RELATIVE TO SEQUENTIAL SIGN TEST

\[ N_0 \to \infty \]

\[ \varepsilon \]

INDEPENDENT SAMPLES

\[ M = 8 \]

\[ M = 2 \]

\[ N_0 \]

VU-GRAPH 8
The efficiency under dependent sampling, relative to the sequential sign test for independent samples, depends upon the normalized correlation function, the increased sampling rate, and the number of samples summed. Assuming a Gauss-Markov process under a Lehmann alternative and a sampling rate of twice that needed for independent samples, the figure shows the effect of increasing the number of samples summed for \( M \), the number of intervals, equal to 2 and 8. The abscissa represents the number of samples summed, \( n \), and the ordinate gives the efficiency for each case. As \( n \) approaches infinity, for this example, the efficiency approaches 1.27 for \( M \) equal to 2 (indicated by the broken line in the lower half of the figure). For \( M \) equal to 8, the efficiency approaches 1.89 as seen in the upper half of the figure. The solid lines represent the independent sampling efficiency for each case.

As the sampling rate and the number of samples summed approach infinity, assuming the number of samples summed is much larger than the increased sampling rate, the efficiency for \( M \) equal to 2 approaches but does not equal the efficiency for the optimum SPD under independent sampling. Thus, the loss in efficiency due to partitioning can be recovered in the SPD by rapid sampling. A similar result is well known for the hard clipper in fixed sample detectors.

To summarize, the effects of dependent sampling on sequential partition detectors were shown to depend upon the normalized correlation function or spectrum shape of the input noise. The efficiency after adjusting the thresholds to maintain the same or lower error probabilities as in the independent sampling case, was improved for rapid sampling.

-VU-graph OFF PLEASE. ARE THERE ANY QUESTIONS?-
## INITIAL DISTRIBUTION LIST

<table>
<thead>
<tr>
<th>Addressee</th>
<th>No. of Copies</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAVSEASYSCOM, SEA 06H2 (CDR D. Bolka, E. Landers,</td>
<td>6</td>
</tr>
<tr>
<td>C. Tolsdorf, K. Wong), SEA 661C,</td>
<td></td>
</tr>
<tr>
<td>SEA 06H1 (CDR D. Griffiths)</td>
<td></td>
</tr>
<tr>
<td>Engineering Societies Library, United Engineering Center</td>
<td>1</td>
</tr>
<tr>
<td>CNM, Code 035</td>
<td>2</td>
</tr>
<tr>
<td>SACLANT ASW Research Centre</td>
<td>1</td>
</tr>
</tbody>
</table>