AD-A053 722 MOORE SCHOOL OF ELECTRICAL ENGINEERING PHILADELPHIA PA F/G 9/4
SAMPLING RATE AND THE J-DIVERGENCE IN DETECTION BASED ON CONTIN--ETC(U)
MAR 78 J G SHIN, S A KASSAM AFOSR-77-3154

UNCLASSIFIED AFOSR-TR-78-0726 NL

| OF |

AD
A053722

END
DATE
FILMED
6 -78
DDC

SAMPLING RATE AND THE J-DIVERGENCE IN DETECTION BASED ON CONTINUOUS-TIME OBSERVATIONS

Jung Gil Shin and Saleem A. Kassam
Department of Systems Engineering
The Moore School of Electrical Engineering
University of Pennsylvania
Philadelphia, PA. 19104

ABSTRACT

The J-divergence is considered as a measure of discrimination information in continuous-time observation for a binary hypothesis testing problem. Relations between the sampling rate, the observation time interval and the amount of information are discussed from a detection view point. General asymptotic relations are obtained as well as some finite observation-time results for stationary Gaussian processes.

I. INTRODUCTION

Studies on continuous-time waveform sampling problems have been focused mainly on the reconstruction of the original waveform according to the minimum mean-square-error criterion. For detection purposes, however, where we are interested in determining whether a signal is present in the observation or not, the goodness of the estimate of the waveform itself from samples may not be the most relevant aspect of the problem. In [1], for example, it has been shown that for detection of weak signals in additive noise the characteristic of the optimum quantizer approximates the nonlinearity p'/p, where p and p' are the noise density function and its derivative; $p'(x)/p(x)$ may not be equal to x. Therefore it is of interest to have results on the degree of degradation of detector performance due to sampling, results which would enable one to find a reasonable sampling rate affording an acceptable compromise between the reduction of volume of computation and data, and optimization of detection performance. In this paper we consider the relations between the sampling rate, the observation-time interval, and discrimination information in samples from a continuous-time observation.

Consider a simple binary detection problem of testing a null hypothesis H_0 and an alternative H_1, using uniformly sampled observations from a stationary mean-square continuous Gaussian process. A stationary Gaussian process is considered since it plays an important role in many types of problems, and allows analytical results to be obtained. As a measure of discrimination information in an observation vector of samples the J-divergence [2,3] will be employed. Denote by $J(1,0)$ the J-divergence between the distributions of the sampled-observation vector under the hypothesis H_0 and H_1. It is defined by

$$J(1,0) = \int\{f_1(\underline{x}) - f_0(\underline{x})\}\ln[f_1(\underline{x})/f_0(\underline{x})]d\underline{x} \quad (1)$$

where f_i denotes the probability density function of the observation $\underline{X}$ under hypothesis H_i. It is a measure of the degree of "closeness" of two distributions, and is closely related to the detection probability [4].

In the following section the problems of detection of a constant signal and a time-varying signal in additive noise are considered. For both cases the J-divergence in some finite number N of samples, denoted by J^N, is obtained and compared with the total J-divergence, denoted by J^T, for the entire continuous-time observation in the finite observation time interval T. This enables us, for instance, to pick a reasonable sampling rate which will yield samples containing a certain percentage of the total information available. In the asymptotic case where both T and N tend to infinity, with T/N remaining fixed at some finite value, exact analytic expressions can be obtained for J^N/J^T. Detection probabilities for some given probability of false alarm have also been computed to give an example of the relation between the J-divergence and the detection probability.

It is interesting to note that the results obtained indicate that when the problem is the detection of a constant signal, for instance, and the noise is bandlimited with bandwidth W, a number of samples which slightly exceeds WT is all that we need for detection purposes, rather than the 2WT samples taken at the Nyquist rate.

Specifically, the two hypotheses H_0 and H_1 that we will consider for our continuous-time observation X(t) are defined as follows:

$$\begin{aligned} H_1&: X(t) = s(t) + N(t) \\ H_0&: X(t) = N(t) \end{aligned} \quad , \quad -T/2 \le t \le T/2 \quad (2)$$

This research is supported by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant AFOSR 77-3154.

Presented at the Conference on Information Sciences and Systems, Johns Hopkins University, March 1978. To be published in the Proceedings of the Conference.

Here, s(t) denotes a known finite-energy signal and N(t) denotes a stationary Gaussian process with an autocorrelation function $K(\tau)$, τ being the time difference between two observations. Components X_i of the observation vector $\underline{X}$ are defined as

$$X_{i+1} = X(i\Delta - \frac{T}{2}), \quad i=0,1,\ldots,N-1 \tag{3}$$

where Δ denotes the time difference between any two consecutive samples, so that

$$\Delta = T/(N-1) \; . \tag{4}$$

We will use the term "J-Ratio" to denote the ratio of J-divergence, J^N/J^T. By total J-divergence, J^T, we mean

$$J^T = \lim_{N\to\infty} J^N \tag{5}$$

for fixed T.

II. J-DIVERGENCE AND SAMPLING RATE

We will first show that J^T can be given by the J-divergence of the coefficients of the Karhunen-Loeve expansion of the process X(t). Then the expression for asymptotic J-Ratio will be derived.

A. N-Sample J-divergence

We obtain first the J-divergence in N uniformly sampled observations in the observation time interval of length T. We will assume that the noise is a zero-mean process to simplify our discussion. Let R_N be the covariance matrix of the N-samples. Then we can write

$$R_N = \begin{bmatrix} K(0) & K(\Delta) & \cdots & K(T) \\ K(\Delta) & K(0) & \cdots & K(T-\Delta) \\ \vdots & \vdots & & \vdots \\ K(T) & K(T-\Delta) & \cdots & K(0) \end{bmatrix} \tag{6}$$

which is a symmetric Toeplitz matrix [5]. Let Φ_N denote the matrix whose columns are the normalized eigenvectors of R_N and let λ_k denote the eigenvalues of R_N, k=1,2,...,N. We note that

$$\lambda_i \underline{\phi}_i = R_N \underline{\phi}_i \tag{7}$$

where $\underline{\phi}_i$ is the i-th column of Φ_N. By transforming $\underline{X}$ to $\underline{X}'$ such that

$$\underline{X}' = \Phi_N^{-1} \underline{X} \; , \tag{8}$$

and because Φ_N^{-1} is an invertible mapping [6], we can easily see that J^N is given by

$$J^N = \sum_{i=1}^{N} \frac{\left\{ \sum_{j=1}^{N} s_j \, \phi_{ji} \right\}^2}{\lambda_i} \tag{9}$$

where s_j is the signal component of the j-th sample and ϕ_{ji} is the ji-th element of Φ_N, i.e. the j-th component of the eigenvector $\underline{\phi}_i$.

B. Total J-divergence

Let us denote the Karhunen-Loeve (KL) expansion of X(t) in the interval [-T/2,T/2] by

$$X(t) = \sum_{i=1}^{\infty} Y_i \, \psi_i(t) \tag{10}$$

where the coefficients Y_i are

$$Y_i = \int_{-T/2}^{T/2} X(t)\psi_i(t)dt. \tag{11}$$

The orthonormal basis functions ψ_i's satisfy the following integral equation:

$$\mu_i \psi_i(t) = \int_{-T/2}^{T/2} K(t-\tau)\psi_i(\tau)d\tau \tag{12}$$

We know that Y_i's are orthogonal and thus independent random variables, since they are normally distributed. Defining S_i to be

$$S_i = \int_{-T/2}^{T/2} s(t)\psi_i(t)dt, \tag{13}$$

we get the J-divergence, J^{KL}, in the coefficients Y_i, i=1,2,..., to be

$$J^{KL} = \sum_{i=1}^{\infty} \frac{S_i^2}{\mu_i} \tag{14}$$

The integral equation, (12), holds for all t, $-T/2 \leq t \leq T/2$. We can approximate the integral in (12) by a finite sum and consider the equations for t = $-T/2, \Delta-T/2, \ldots, T/2$. Then the set of equations can be written as an approximate matrix equation of finite dimensions,

$$(\mu_i/\Delta)\underline{\psi}_i\sqrt{\Delta} \doteq R_N \underline{\psi}_i \sqrt{\Delta} \; , \tag{15}$$

where $\underline{\psi}_i$ is a vector whose (j+1)-th component is $\psi_i(j\Delta - T/2)$ for j=0,1,...,N-1. This becomes the same equation as we had in (7) if we let $N\to\infty$ for fixed T. Therefore, with ψ_{ji} the j-th component of ψ_i,

$$J^T = \lim_{N\to\infty} J^N = \lim_{\substack{\Delta\to 0\\ T \text{ fixed}}} \sum_{i=1}^{\frac{T}{\Delta}+1} \frac{\left[\sum_{j=1}^{\frac{T}{\Delta}+1} s_j \psi_{ji} \sqrt{\Delta}\right]^2}{\mu_i/\Delta}$$

$$= \sum_{i=1}^{\infty} \frac{\left[\int_{-T/2}^{T/2} s(t)\ \psi_i(t)dt\right]^2}{\mu_i} \qquad (16)$$

which is the J-divergence in the KL expansion coefficients.

C. Asymptotic Ratio of J^N and J^T

In many cases where T is finite, finding a closed-form expression for J^N may not be simple. Numerical computations also have a limit with increasing number of samples since they involve, in general, a matrix inversion. However, in the asymptotic case where both T and N tend to infinity with Δ remaining fixed at some finite value, we can examine the ratio of J^N and J^T as a function of Δ.

For T>>1, the eigenvalues and the eigenfunctions satisfying the integral equation, (12), may be approximated by [7],

$$\mu_n \cong N_c(nf_0) \qquad (17)$$

and

$$, n=0,\pm 1,\pm 2...$$

$$\psi(t) \cong (1/\sqrt{T})\exp(-j2\pi f_0 nt) \qquad (18)$$

where $f_0=1/T$ and $N_c(f)$ is the power spectrum of the noise process given by

$$N_c(f) = \int_{-\infty}^{\infty} K(\tau)e^{-j2\pi f\tau}d\tau. \qquad (19)$$

(Note that, in (17) and (18), the eigendata are indexed over both positive and negative integers [7].) Also, define

$$S_c(f) = \int_{-\infty}^{\infty} s(t)e^{-j2\pi ft}dt, \qquad (20)$$

the Fourier transform of the energy signal s(t). For T>>1, using the approximation given by (17) and (18) in (16), we get

$$\lim_{T\to\infty} J^T = \int_{-\infty}^{\infty} \frac{|S_c(f)|^2}{N_c(f)} df \qquad (21)$$

A more rigorous derivation of (21) is possible along the lines in the Appendix.

Let us now define

$$N_d(f) = \sum_{n=-\infty}^{\infty} K(n\Delta)e^{-j2\pi fn\Delta} ; \qquad (22)$$

then

$$K(m\Delta) = \Delta \int_{-\frac{1}{2\Delta}}^{\frac{1}{2\Delta}} N_d(f)e^{j2\pi fm\Delta}df. \qquad (23)$$

Similarly, define $S_d(f)$ for signal s(t). After some manipulation we can show that

$$\lim_{\substack{N,T\to\infty\\ \Delta \text{ fixed}}} J^N = \Delta \int_{-\frac{1}{2\Delta}}^{\frac{1}{2\Delta}} \frac{|S_d(f)|^2}{N_d(f)} df . \qquad (24)$$

The derivation of (24) is outlined in the Appendix. From (21) and (24) we get the asymptotic ratio of J^N and J^T,

$$\text{Asymp. J-Ratio} = \frac{\Delta \int_{-\frac{1}{2\Delta}}^{\frac{1}{2\Delta}} \frac{|S_d(f)|^2}{N_d(f)} df}{\int_{-\infty}^{\infty} \frac{|S_c(f)|^2}{N_c(f)} df} \qquad (25)$$

In the constant signal case, assuming finite energy by letting $s(t) = s\to 0$ as $T\to\infty$, the asymptotic J-Ratio reduces to

$$\text{Asymp. J-Ratio} = \frac{N_c(0)}{\Delta\ N_d(0)} \qquad (26)$$

By the Poisson sum formula [8] we can write

$$N_d(f) = \frac{1}{\Delta} \sum_{n=-\infty}^{\infty} N_c(f + n/\Delta). \qquad (27)$$

Therefore, if the noise spectrum has a bandwidth W, the ratio given by (26) remains unity as long as $1/\Delta > W$. In other words a sampling rate higher than half of the Nyquist rate should yield samples which contain full discrimination information in terms of J-divergence. In time-varying signal cases if the signal is frequency limited with a bandwidth W_s and with $W_s < W$, then a sampling rate of $(W+W_s)$ should be sufficient for the detection problem.

From this asymptotic characteristic of J-divergences with respect to sampling rate we can see that the degradation of detector performance due to sampling at a rate below the Nyguist rate could be insignificant. If we plot the J-Ratio for a specific detection problem of interest, we can use it as a guide to picking a sampling rate. We would then have a reasonable theoretical justification for this choice in terms of the percentage of full discrimination information it preserves in the samples.

III. NUMERICAL EXAMPLES

In the following examples we assume until energy signals for convenience.

Example 1. Constant Signal and Ideal Low Pass Noise

To see if we have characteristics for the J-Ratio in a finite number of samples similar to what we noticed in the asymptotic case, we consider an example for constant signal detection where the additive noise has the power spectral density as shown in Fig. 1. In order to obtain the finite-T J-Ratio and the probability of detection, the observation time duration T is assumed to be 15. Thus the number of samples taken at the Nyquist rate is 31. The detection probability, P_D, is computed for $P_{FA} = 10^{-4}$ where P_{FA} denotes the false alarm probability. In Fig. 1, the asymptotic and the finite-T J-Ratio and P_D are plotted. We note that both J-Ratio and P_D show a jump to their maximum values as N exceeds 16 and remain at the value for N>17.

Example 2. Constant Signal and Gauss-Markov Noise

When the noise is Gauss-Markov where the covariance function is of the form

$$K(\tau) = K_0 e^{-a|\tau|}, \tag{28}$$

analytic derivation of J^N, J^T and the asymptotic J-Ratio is possible. We list the results below. For convenience, we let $K_0=1$.

$$J^N = [N(1-\rho) + 2\rho]/(1+\rho) \tag{29}$$

$$J^T = aT/2 + 1 \tag{30}$$

$$\text{Asymp. J-Ratio} = (1-e^{-a\Delta})/[a\Delta(1+e^{-a\Delta})] \tag{31}$$

where

$$\rho = e^{-aT/(N-1)}. \tag{32}$$

For a=2 and T=15, these are plotted in Fig. 2. It shows the close relationship between the J-divergence and the detection probability. Suppose we are to find a sampling rate such that the samples obtained at the rate preserve 90% of the total discrimination information, we can read it from the asymptotic J-Ratio curve which is, in this case, Δ=0.56.

Example 3. Time-varying Signal and Ideal Low Pass Noise

As an example for the time-varying signal case we consider a signal whose Fourier transform is of a triangular shape against the ideal low pass noise as in Ex. 1. The asymptotic J-Ratios for W_s = 0.3 and 0.4 are plotted in Fig. 3, where W_s denotes the bandwidth of the signal frequency spectrum. As we mentioned in Part C, Sec. II, the J-Ratio remains at one for $1/\Delta > W+W_s$.

APPENDIX. Derivation of (24)

From (6) and (23), R_N can be written as

$$R_N = \Delta \int_{-\frac{1}{2\Delta}}^{\frac{1}{2\Delta}} N_d(f)\, \Phi_f \, df \tag{A1}$$

where Φ_f is a N×N Hermitian symmetric matrix whose ij-th element is $\exp[-j2\pi\Delta f(i-j)]$.

For N>>1, we can approximate R_N^{-1}, the inverse of R_N, by

$$R_N^{-1} \cong \Delta \int_{-\frac{1}{2\Delta}}^{\frac{1}{2\Delta}} \frac{1}{N_d(f)} \Phi_f \, df \tag{A2}$$

It can be shown that $R_N R_N^{-1} \cong I$, the identity matrix, by using the Poisson sum formula in the intermediate stages.

Let $\underline{S}$ be the signal component vector, i.e. $\underline{S}^t = \{s(i\Delta)\}_{i=-N/2}^{N/2-1}$. Then J^N can be written as

$$J^N = \underline{S}^t R_N^{-1} \underline{S}. \tag{A3}$$

We can then see that, for N>>1,

$$\begin{aligned} \underline{S}^t \Phi_f \underline{S} &= |\underline{S}^t \underline{\phi}_f|^2 \\ &\cong |S_d(f)|^2 \end{aligned} \tag{A4}$$

where

$$\underline{\phi}_f = \left\{ e^{j\frac{1}{2}\pi f \Delta_i} \right\}_{i=-N/2}^{N/2-1}.$$

Thus we get (24).

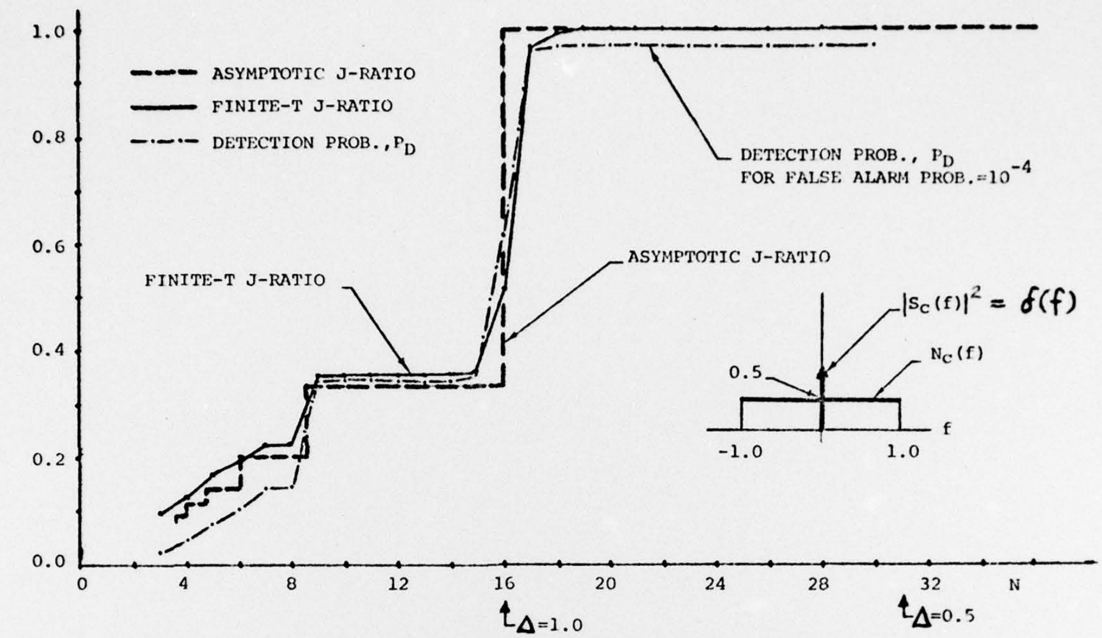

Fig. 1. Constant Signal and Ideal Low Pass Noise. T=15 is assumed for finite-T J-Ratio and P_D.

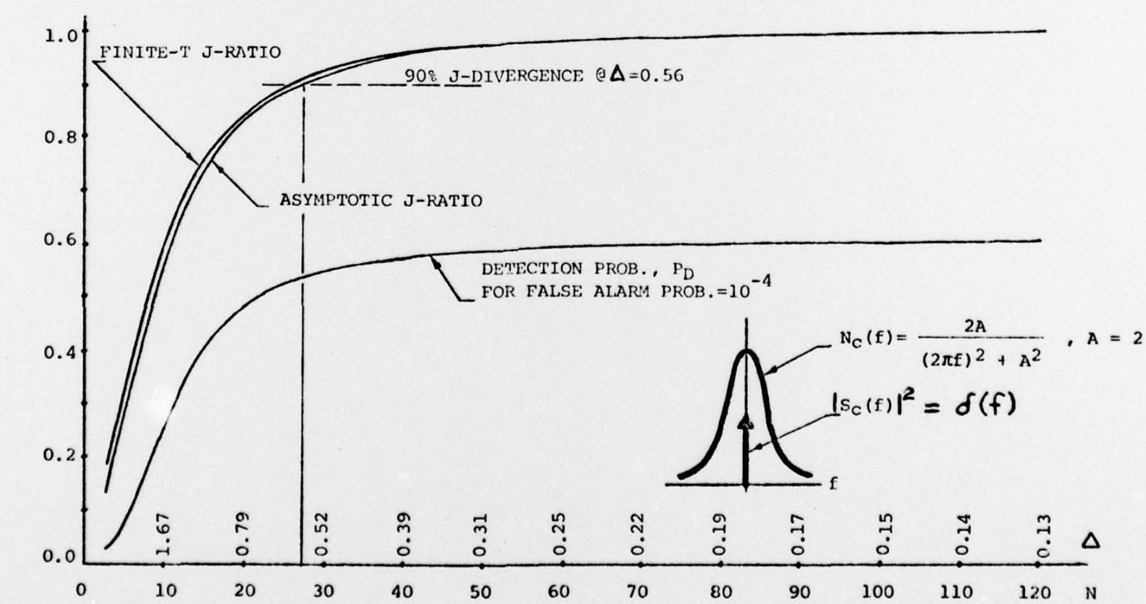

Fig. 2. Constant Signal and Gauss-Markov Noise.

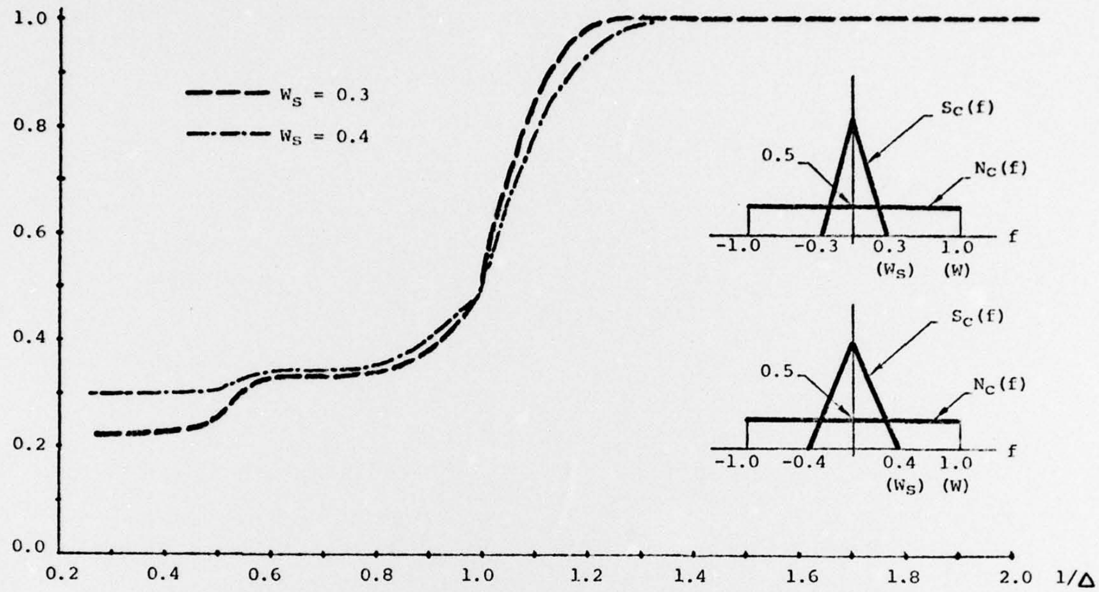

Fig. 3. Time-Varying Signal and Ideal Low Pass Noise. Asymptotic J-Ratio.

REFERENCES

1. S. A. Kassam, "Optimum Quantization for Signal Detection," IEEE Trans. Commun., vol. COM-25, pp. 479-484, May 1977.

2. S. Kullback, Information Theory and Statistics. New York, Wiley, 1959.

3. H. Jeffery, Theory of Probability. Oxford University Press, 1948.

4. T. Kailath, "The Divergence and Bhattacharyya Distance Measures in Signal Selection," IEEE Trans. Commun., vol. COM-15, pp. 52-60, February 1967.

5. U. Grenander and G. Szego, Toeplitz Forms and Their Application. University of California Press, 1958.

6. T. Berger, Rate Distortion Theory. Englewood Cliffs, Prentice-Hall, 1971.

7. H. L. VanTrees, Detection, Estimation, and Modulation Theory, Part I. New York, Wiley, 1968.

8. A. Papoulis, Signal Analysis. New York, McGraw-Hill, 1977.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR-TR- 78 - 0726	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): SAMPLING RATE ~~ON~~ and THE J-DIVERGENCE IN DETECTION BASED ON CONTINUOUS-TIME OBSERVATIONS	5. TYPE OF REPORT & PERIOD COVERED: Interim rept.
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Jung Gil Shin [redacted] Saleem A. Kassam	8. CONTRACT OR GRANT NUMBER(s): AFOSR-77-3154
9. PERFORMING ORGANIZATION NAME AND ADDRESS: University of Pennsylvania, The Moore School of Electrical Engineering, Philadelphia, PA 19174	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research/NM, Bolling AFB, DC 20332	12. REPORT DATE: Mar[redacted]78
	13. NUMBER OF PAGES: 6
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The J-divergence is considered as a measure of discrimination information in continuous-time observation for a binary hypothesis testing problem. Relations between the sampling rate, the observation time interval and the amount of information are discussed from a detection view point. General asymptotic relations are obtained as well as some finite observation-time results for stationary Gaussian processes.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)