SOME RESULTS ON ASYMPTOTIC MEMORYLESS DETECTION IN STRONG MIXIN-ETC(U)

1980  D R HALVERSON, G L WISE

F49620-77-C-0101

AFOSR-TR-80-1303
In this work we consider the discrete time detection of strong mixing signals in strong mixing noise, and we allow a large degree of dependency to exist between the signal and the noise. We investigate the memoryless detector which is optimum in the sense of the asymptotic relative efficiency. It is shown that the design of this detector reduces to the solution of an integral equation in which knowledge of only the second-order statistics of the random processes involved is required.
SOME RESULTS ON ASYMPTOTIC MEMORYLESS DETECTION IN STRONG MIXING NOISE

D. R. Halverson
Department of Electrical Engineering
Texas A&M University
College Station, Texas 77843

G. L. Wise
Department of Electrical Engineering
University of Texas at Austin
Austin, Texas 78712

Abstract

In this work we consider the discrete time detection of strong mixing signals in strong mixing noise, and we allow a large degree of dependency to exist between the signal and the noise. We investigate the memoryless detector which is optimum in the sense of the asymptotic relative efficiency. It is shown that the design of this detector reduces to the solution of an integral equation in which knowledge of only the second-order statistics of the random processes involved is required.

I. INTRODUCTION

The detection of signals in corrupting noise has been an area of interest for some time. Because of modern high speed sampling, it is expected that the underlying random processes involved will not be "white", but will instead possess dependency to a certain degree. Neyman-Pearson optimal techniques [1] are tractable only in cases where the appropriate multivariate distributions are known. In many non-Gaussian situations these distributions are not known, which has thus led to the choice of an alternate fidelity criterion, commonly the asymptotic relative efficiency (ARE) criterion, which is especially appropriate in the weak signal and large sample situation. Because continuous time detection is often intractable in the non-Gaussian case, current efforts are directed toward discrete time detection. Results in this area have been obtained recently by Poor and Thomas [2,3] for the case of memoryless strong mixing model of [4-6] is easier to check. We therefore will consider the general situation where we are detecting the presence of a strong mixing signal in strong mixing noise.

II. PRELIMINARIES

Let \( (X_i ; i=1,2,...) \) be a strictly stationary sequence of random variables. For \( a < b \), define \( M(a,b) = \sigma(X_a, X_{a+1}, X_{a+2}, ... , X_{b-1}, X_b) \), the \( \sigma \)-algebra generated by the indicated random variables, where \( a \) and \( b \) may take on extended real values. Then \((X_i ; i=1,2,...)\) is symmetrically \( \varphi \)-mixing if there exists a nonnegative sequence \( \{\phi_i; i=1,2,...\} \) with \( \phi_i \rightarrow 0 \) such that for each \( k, 1 \leq k < \infty \), and for each \( i \geq 1 \), \( E_1 \cdot M(1,k) \) and \( E_2 \cdot M(k+1,\infty) \) together imply

\[
|P(E_1 \cap E_2) - P(E_1)P(E_2)| \leq \phi_i \min (P(E_1),P(E_2)).
\]

In [4-6] the above type of process is employed. Note that the left side of the above inequality provides a measure of dependence between events \( E_1 \) and \( E_2 \), and the right side bounds this quantity with a term involving \( E_1 \) and \( E_2 \). Such a definition has computational advantages; for example, it results in the very powerful Lemma 1 of [7, p.170]. However, it is a stronger requirement than our intuition might demand. Since we really wish to simply require a "decrease" in dependency as

Presented at the Twenty-Third Midwest Symposium on Circuits and Systems, August 4-6, 1980; to be published in the Proceedings of the Symposium.
and $E_2$ are more widely separated in time, it is thus more natural to employ the weaker requirement that there exists a nonnegative sequence $(a_i; i=1,2,...)$ with $a_i > 0$ such that for all $E_1$ and $E_2$ as above we have

$$|P(E_1 \cap E_2) - P(E_1)P(E_2)| \leq a_i.$$  

A process satisfying this condition is called strong mixing. We will consider the detection of a strong mixing signal $(S_i; i=1,2,...)$, where $0 < E(S_i^2) < \infty$, in additive strong mixing noise $(N_i; i=1,2,...)$, where we observe realizations $(y_i; i=1,2,...,n)$ of the process $(Y_i; Y_i=i=1,2,...,n)$. 

In order to apply the ARE fidelity criterion, this will amount to a choice between the two hypotheses

$$H_0: Y_i = N_i; i=1,2,...,n$$

$$H_1: Y_i = N_i + OS_i; i=1,2,...,n$$

where $0$ is a parameter which will be allowed to approach zero at the proper rate, thus yielding the asymptotic limit. Throughout the discussion we will assume that both the noise and signal processes possess (possibly different) $\alpha$-representations which satisfy

$$\sum_{i=1}^\infty \alpha_i(2+\delta) < \infty$$

for some appropriate $\delta > 0$. Such a strong mixing process will be called $\delta$-acceptable. For convenience we assume the existence of densities $f_j(\cdot,\cdot)$ of $N_k$ and $f_{k+j}(\cdot,\cdot)$ of $N_k$ and $S_k$, where the latter is assumed to be independent of $k$. We also assume

$$K_n(f,x,y) = \frac{1}{n} \sum_{j=1}^n [f_j(x,y) + f_j(y,x)]/\sqrt{f(x)f(y)}$$

is square integrable for all $n$, and that $f(\cdot)$ is strictly positive on the real line. We assume in addition that

$$\int y^{2+\delta} f(x,y)dy/\sqrt{|f(x)|}$$

and

$$\int y^{2+\delta} f(x,y)dy/\sqrt{|f(x)|}$$

are square integrable. Note that if the signal and noise are independent, the latter condition is equivalent to the assumption of finite Fisher's information number contained in [2-6] and [4,5], and [4,5].

We also assume that

$$\lim_{\theta \to 0} \int f(x-\theta y,y)dy = f(x).$$

As in [2-6], we will optimize over the class of optimal memoryless detectors designed under a "white noise" assumption, i.e. where a test statistic $T_g(y) = \sum_{i=1}^n g(y_i)$ is compared to a threshold. Specifying $g$ will therefore be of prime concern.

We will restrict the class $\mathcal{G}$ of nonlinearities $g$ to include those measurable realvalued functions for which we can find $\beta_1 > 0$ and $\delta_1 > \delta$ such that the random variable $g(N_i + OS_i)$ satisfies

$$E|g(N_i + OS_i)|^{2+\delta_1} < \infty$$

for all $\beta \in [0,\beta_1]$, and such that the following mild regularity conditions hold, where $E_0(\cdot)$ denotes expectation computed under $H_0$ with parameter $\beta$ (by proper choice of the threshold, we assume without loss of generality that the random variables $g(N_i)$ are zero mean):

(a) $\int g(x) f'(x)dx / \sqrt{|f(x)|} \neq 0$

if the signal and noise processes are independent;

(b) $\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^n g(T_i(y)) = 0$

if $\int g(x) dx > 0$, or

(b') $\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^n g(T_i(y)) = 0$

if $\int g(x) dx = 0$;

(c) $-\infty < \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^n E[g(N_i + OS_i)] = k_1 / \sqrt{n}$

for some constant $k_1 > 0$

if $\int g(x) dx = 0$;

(c') $-\infty < \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^n E[g(N_i + OS_i)] = k_2 / \sqrt{n}$

for some constant $k_2 > 0$

if $\int g(x) dx = 0$;

(d) $\lim_{\theta \to 0} E[g(N_i + OS_i)] = E[g(N_i)]$;

(e) $\int g(x) f(x-\theta y,y)dy = \int g(x) f(x-\theta y,y)dy / \sqrt{|f(x)|}$

if $\int g(x) dx = 0$, or
where we employ the notation of Lemma 1. Proof: Letting $T_{n,0}$ under $H_0$ it follows from condition (9) that \( \lim_{n \to \infty} \sigma_{n,0}^2 = 0 \), and hence, \( \lim_{n \to \infty} n \sigma_{n,0}^2 = 0 \). We thus obtain from Theorem 1.4 of [9] and Theorem 1 that

\[
\sqrt{n} T_{n,0} \xrightarrow{\mathcal{D}} N(0,1), \quad \text{and therefore}
\]

\[
T_{n,0} / \sigma_{n,0} \xrightarrow{\mathcal{D}} N(0,1). \quad \text{We also have from Lemma 1.3 of [9] and Theorem 1 that}
\]

\[
E(T_{n,0} - E(T_{n,0}^2)^{1/2}) \leq E(g(N_1)^2)^{1/2} \leq \sum_{i=1}^{2} \gamma_i \delta_{i+2} + 2(E(g(N_1)^2) \sum_{i=1}^{2} \gamma_i \delta_{i+2})^{1/2}
\]

and \( \sum_{i=1}^{2} \gamma_i < \infty \) and \( \gamma_i, i = 1, 2, \ldots \) is an $\alpha$-representation for the $\delta$-acceptable process

Lemma 2: Suppose $\{X_i, i = 1, 2, \ldots \}$ is $\delta$-acceptable, and for a fixed nonnegative integer $m$, $\mathcal{N}_i = \mathcal{N}(X_i, Z_i)$ for $i = 1, 2, \ldots$, where $X_i$ is a $\delta$-acceptable process, $\mathcal{N}_i$ is measurable. $G : \mathbb{R}^2 \to \mathbb{R}$ is measurable, and $(Z_i, i = 1, 2, \ldots)$ is $\delta$-acceptable and independent of $(X_i, i = 1, 2, \ldots)$. Then $\mathcal{N}_1, \mathcal{N}_2, \mathcal{N}_3, \ldots$ is $\delta$-acceptable.

Proof: This follows through an argument identical to the proof of Lemma 2 of [6]. Q.E.D.

We can thus obtain a useful result:

Theorem 1: Suppose \( \{ \mathcal{N}_i : \mathbb{R}^{2p+2} \to \mathbb{R}, i = 1, 2, \ldots \} \) is a family of measurable functions where $p$ is a fixed nonnegative integer. Then under the hypothesis of Lemma 1 or Lemma 2, we have that \( \{ \mathcal{N}_i(N_1, S_1, \ldots, N_{p+1}, S_{p+1}, i = 1, 2, \ldots) \) is $\delta$-acceptable.

Proof: This follows as a consequence of Lemma 1, Lemma 2, and a straightforward modification of Proposition 7 of [5]. Q.E.D.

We can now obtain the result which will allow employment of the Pitman-Noether Theorem [8].

Theorem 2: Suppose $\theta_0 \in \mathcal{R}$ with $\theta_0 = 0$, and $g \in \mathcal{W}$. Let $T_{n,0} = g / \sqrt{n}$ under $H_0$ with parameter $\theta_0$, where the noise and signal processes satisfy the hypothesis of Lemma 1 or Lemma 2. If

\[
\sigma_{n,0}^2 \geq E((T_{n,0} - E(T_{n,0}))^2),
\]

Then $\sqrt{n} T_{n,0} / \sigma_{n,0} \xrightarrow{\mathcal{D}} N(0,1)$, and therefore

\[
T_{n,0} / \sigma_{n,0} \xrightarrow{\mathcal{D}} N(0,1). \quad \text{We also have from Lemma 1.3 of [9] and Theorem 1 that}
\]

\[
E(T_{n,0} - E(T_{n,0}^2)^{1/2}) \leq E(g(N_1)^2)^{1/2} \leq \sum_{i=1}^{2} \gamma_i \delta_{i+2} + 2(E(g(N_1)^2) \sum_{i=1}^{2} \gamma_i \delta_{i+2})^{1/2}
\]

and \( \sum_{i=1}^{2} \gamma_i < \infty \) and \( \gamma_i, i = 1, 2, \ldots \) is an $\alpha$-representation for the $\delta$-acceptable process
(\sum_{j=1}^{m} \int [f_j(x,y)+f_j(y,x)]g(y)dy + h(x) = -f(x)g(x),
\sum_{j=1}^{m} \int [f_j(x,y)+f_j(y,x)]g(y)dy + h(x) = -f(x)g_m(x).

In this case we would then hope that the convergence in some appropriate sense of the \( g_m \) would lead to an optimal nonlinearity. This question is answered in the following:

Theorem 4: Under the hypothesis of Lemma 1 or Lemma 2, if there exists a \( g \in \mathcal{G} \) such that a subsequence \( \{g_{m_k}\}_{k=1}^{\infty} \) of \( \{g_m\}_{m=1}^{\infty} \) satisfies

\[ g_m(N_1) - g(N_1) \to 0 \quad \text{in } L_{2+\delta_1}, \]

then \( g \) is optimal (in the sense of the ARE) if and only if \( g \) satisfies (up to a scale factor) In this case we would then hope that the optimal nonlinearity. This question is answered in the following:

Theorem 4: Under the hypothesis of Lemma 1 or

\[ \sum_{j=1}^{m} \int [f_j(x,y)+f_j(y,x)]g_m(y)dy + f''(x) \]

as \( k \to \infty \), where \( \delta g \) is an arbitrary zero mean variation satisfying

\[ E(\delta g(N_1)) = \frac{\alpha_j}{(2+\delta_1)(1+c)} \]

where \( C_1 = E(\frac{\delta g(N_1)^2}{2+\delta_1}) < \infty \) and

\[ C_2 = E(\frac{\delta g(N_1)^2}{2+\delta_1}) < \infty. \]

Moreover, a similar application together with the Schwarz inequality shows that the first summand can be upper bounded by

\[ 4+2(C_1+C_2+\sqrt{C_1C_2}) \sum_{j=1}^{m} \frac{\alpha_j}{(2+\delta_1)+1+c} \]

for any \( c > 0 \), where

\[ a_j = E(\delta g(N_1))^{2+\delta_1} \]

Choosing \( c \) small enough so that

\[ \delta_1/[(2+\delta_1)(1+c)] \geq \delta/(2+\delta). \]
we obtain the desired result. Q.E.D.

Note the conditions on the optimal nonlinearity $g$ and the $\alpha$-representation are exceedingly mild.

For example, these results hold if

$$E[(g(N_1, \ldots, N_n))^4] < \infty$$

for all $\theta \in [0, 1]$, $\alpha_i < \infty$, and the subsequence of Theorem 4 converges in $L_4$. We remark finally that the nonlinearities $g_\alpha$ are obtainable through standard Hilbert-Schmidt techniques as solutions of Fredholm integral equations of the second kind.

IV. CONCLUSION

We have considered the design of the optimal detector for signal detection in corrupting noise, where both the signal and noise may be chosen from a large class of strong mixing processes and may be dependent on each other. We have seen that this design reduces to the solution of an integral equation in which knowledge of only the second-order statistics of the random processes involved is required. In particular, if the signal is independent of the noise and has nonzero mean, the optimal detector is the same as in the constant known signal case.

Acknowledgement

This research was supported by the Department of Defense Joint Services Electronics Program under Contract F49620-77-C-0101, and by the Air Force Office of Scientific Research under Grant AFOSR-76-3062.

REFERENCES


BIOGRAPHIES

D. R. Halverson was born in Menomonie, Wisconsin on March 1, 1947. He received the B.A. degree in mathematical sciences summa cum laude from the University of Iowa in 1968, the A.M. degree in mathematics from the University of Illinois in 1972, and the Ph.D. degree in electrical engineering from the University of Texas at Austin in 1979. He was employed as a Mathematical Statistician with the U.S. Air Force at Kelly Air Force Base from 1972-75 and as a Mathematics Instructor at the U.S. Air Force Academy Preparatory School from 1975-77. He is presently with the Department of Electrical Engineering at Texas A&M University. His research interests include statistical communication theory, detection theory, and estimation theory. He is a member of Phi Beta Kappa and IEEE.

Gary L. Wise was born in Texas City, Texas on July 29, 1945. He received the B.A. degree in mathematical sciences summa cum laude from Rice University in 1971 with a double major in electrical engineering and mathematics. He received the M.S.E., M.A., and Ph.D. degrees in electrical engineering from Princeton University in 1973, 1973, and 1974, respectively. He is presently with the Department of Electrical Engineering at the University of Texas at Austin. His research interests include statistical communication theory, random processes, and signal processing. He is a member of Phi Beta Kappa, Tau Beta Pi, Eta Kappa Nu, IEEE, SIAM, and AMS.