REFERENCES

[1] W. O. Alltop, private communication.
[2] I. F. Blake and J. W. Mark, "A note on complex sequences with low correlations," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 814–816, Sept. 1982.
[3] F. D. Garber and M. B. Pursley, "Performance of offset quadriphase spread-spectrum multiple-access communications," *IEEE Trans. Commun.*, vol. COM-29, pp. 305–314, Mar. 1981.
[4] R. Gold, "Study of multistate PN sequences and their application to communication systems," *Rockwell Int. Corp. Rep.*, AD A025137, 1976.
[5] S. W. Golomb, *Shift Register Sequences*. San Francisco: Holden-Day, 1967.
[6] S. W. Golomb and R. A. Scholtz, "Generalized Barker sequences," *IEEE Trans. Inform. Theory*, vol. IT-11, pp. 533–537, Oct. 1965.
[7] S. M. Krone, "Quadriphase sequences for spread-spectrum communication," Coordinated Science Laboratory Tech. Rep. T-118, Univ. of Illinois, Urbana, Oct. 1982.
[8] R. M. Lerner, "Signals having good correlation functions," *WESCON Convention Record*, 1961.
[9] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam, The Netherlands: North-Holland, 1977.
[10] M. B. Pursley and F. D. Garber, "Quadriphase spread-spectrum multiple-access communications," *IEEE International Conference on Communications, Conference Record*, vol. 1, pp. 7.3.1–7.3.5, 1978.
[11] M. B. Pursley, F. D. Garber, and J. S. Lehnert, "Analysis of generalized quadriphase spread-spectrum communications," (Invited paper), *IEEE International Conference on Communications, Conference Record*, vol. 1, pp. 15.3.1–15.3.6, 1980.
[12] M. B. Pursley, "Spread-spectrum multiple-access communications," in *Multi-User Communication Systems*, G. Longo, Ed. Vienna and New York: Springer-Verlag pp. 139–199, 1981.
[13] D. V. Sarwate and M. B. Pursley, "Crosscorrelation properties of pseudorandom and related sequences," (Invited paper), *Proc. IEEE*, vol. 68, pp. 593–619, May 1980.
[14] W. M. Schmidt, *Equations Over Finite Fields* (Lecture Notes in Mathematics). New York: Springer-Verlag, 1976.
[15] R. A. Scholtz and L. R. Welch, "Group characters: Sequences with good correlation properties," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 537–545, Sept. 1978.
[16] V. M. Sidelnikov, "Some k-valued pseudo-random sequences and nearly equidistant codes," *Probl. Peredach. Informatsii*, vol. 5, pp. 16–22, Jan. 1969.
[17] V. M. Sidelnikov, "On mutual correlation of sequences," *Soviet Math. Dokl.*, vol. 12, pp. 197–201, 1971.
[18] R. Turyn, "Four-phase Barker codes," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 366–371, May 1974.
[19] L. R. Welch, "Lower bounds on the maximum cross correlation of signals," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 397–399, May 1974.
[20] G. R. Welti, "Quaternary codes for pulsed radar," *IRE Trans. Inform. Theory*, vol. IT-6, pp. 400–408, June 1960.

Min–Max Detection of Weak Signals in φ-Mixing Noise

GEORGE V. MOUSTAKIDES AND JOHN B. THOMAS, FELLOW, IEEE

Abstract—**Detection of weak signals in a special φ-mixing noise class is considered. The detector structure is restricted to sums of memoryless nonlinear transformations of the observations, correlated with the data sequence and compared to a fixed threshold. Using the efficacy to measure performance, the nonlinearity that has min–max performance is derived.**

I. INTRODUCTION

FOR the detection of signals in additive noise, a very commonly used detector structure consists of a sum of memoryless nonlinear transformations of the observations, correlated with the signals and compared to a fixed threshold. When the number of observations is large and the signals are weak, asymptotic performance is of interest. A common measure of asymptotic performance is the efficacy. It is well-known that for independent identically distributed (i.i.d.) observations the efficacy is maximized when the nonlinear transformation is given by the locally optimum nonlinearity defined by the marginal density. When this density is not known exactly, optimality is often defined in a min–max way. Following the ideas of Huber on robust estimation and hypothesis testing [1], [2] the min–max nonlinearities for detection are derived in [3]–[5] for the i.i.d. case and for densities belonging to an ε-contamination class. In [3], [4] the noise densities are assumed to be symmetric. In [5] noise symmetry is assumed inside an interval around the origin.

All of these approaches assume independent observations. For the dependent case, in order to calculate the efficacy we must know the bivariate densities of all pairs of observations. Finding the nonlinearity that maximizes the efficacy is not easy as in the i.i.d. case. This problem for stationary sequences is considered in [6] for the m-dependent case and in [7] for the φ-mixing case. It is shown that,

Manuscript received January 24, 1983; revised June 27, 1983. This work was supported in part by the U.S. Army under Grant DAAG29-82-K0095 and in part by the National Science Foundation under Grant ECS-79-18915.

G. V. Monstakides was with the Department of Electrical Engineering and Computer Science, Princeton University, Princeton, NJ 08544; he is now with the Institut de Recherche en Informatique et Systemes Aleatoires, Rennes, France.

J. B. Thomas is with the Department of Electrical Engineering and Computer Science, Princeton University, Princeton, NJ 08544.

0018-9448/84/0500-0529$01.00

for the m-dependent case, the optimum nonlinearity satisfies a Fredholm integral equation of the second kind.

Min–max detection with dependent observations is considered in [8]. Following similar ideas from [9], [10] the min–max nonlinearity is derived under the assumption that the observations are generated by a moving average process and are weakly dependent. In [11] the problem of min–max detection of a constant signal in stationary Markov noise is considered. It is shown that, for a special class of Markov noise processes, the min–max nonlinearity is very closely related to the one for the i.i.d. case. Here we consider an extension of [11]. We consider detection of nonconstant signals in a class of φ-mixing noise processes. The class defined in [11] is a special case of this φ-mixing class. We optimize over structures that consist of sums of memoryless nonlinear transformations. It is important to point out that, even though this structure is optimum for the i.i.d. case, this optimality does not hold under dependency. But, in any case, we would like to see how much the independence-assumption structure changes under dependency and also if the performance changes drastically.

II. Preliminaries

Let $\{N_i\}$ be a strictly stationary noise sequence. Denote by M_a^b the σ-algebra generated by the random variables $\{N_a, N_{a+1}, \cdots, N_b\}$. Let $f(x)$ be the common marginal density for the random variables N_i. We assume that this density is symmetric, that it has a continuous derivative different from zero almost everywhere with respect to $f(x)$ and that it has finite Fisher's information for location. For simplicity we will denote random variables with capital letters and sequences of random variables with boldface capital letters. We call the stationary sequence $\boldsymbol{N}$ a *φ-mixing* sequence if there exists a sequence $\{\varphi_n\}$ of real numbers satisfying

$$1 \geq \varphi_1 \geq \varphi_2 \geq \cdots \geq 0 \tag{1}$$

such that, for each positive integer n, if A an event from M_1^k and B from M_{k+n}^{∞} then

$$|P(A \cap B) - P(A)P(B)| \leq \varphi_n P(A). \tag{2}$$

This is the φ-mixing class defined in [13, p. 174]. We call a φ-mixing sequence *acceptable* if in addition to (1) and (2) it satisfies

$$\sum_{n=1}^{\infty} \varphi_n^{1/2} < \infty. \tag{3}$$

Here we consider a subclass of the acceptable φ-mixing sequences. We say that a sequence $\boldsymbol{N}$ belongs to the class $\mathscr{N}$ if it is an acceptable φ-mixing sequence and also satisfies the following conditions concerning the bivariate and univariate densities of two components N_k and N_{k+n}. If A is an event for N_k and B is an event for N_{k+n} then, for every k and n, we have

$$|P(A \cap B) - P(A)P(B)| \leq \gamma_n P(A)P(B) \tag{4}$$

with

$$\sum_{n=1}^{\infty} \gamma_n < \infty \tag{5}$$

and also

$$f(x) = (1-\epsilon)g(x) + \epsilon h(x). \tag{6}$$

Notice that (4) is different from (2) since it is defined only for two random variables. Also the right side of (4) involves the product of the two marginal probabilities rather than one marginal as in (2). Even though every bivariate density satisfies (2) for some φ_n (for example, $\varphi_n = 1$), such is not the case for (4). Finally (6) defines an ϵ-contamination model for the marginal density $f(x)$. We assume that $g(x)$ is a known symmetric strongly unimodal density, with continuous derivative different from zero almost everywhere with respect to $g(x)$ and with finite Fisher's information. For $h(x)$ we assume that it is a symmetric density; and ε a known constant in $[0,1)$.

Let us now consider the detection of a known signal sequence $\{s_i\}$. In particular, we would like to decide between the two hypotheses

$$\begin{aligned} H_0&: \quad X_i = N_i \\ H_1&: \quad X_i = N_i + \delta s_i, \qquad i = 1, 2, \cdots, \end{aligned} \tag{7}$$

where $\boldsymbol{X}$ is the observation sequence, $\boldsymbol{N}$ is the noise sequence, and δ tends to zero. We assume that the signal sequence is bounded and that the following limits exist

$$\nu_j = \lim_{n \to \infty} \frac{s_1 s_{j+1} + s_2 s_{j+2} + \cdots + s_{n-j} s_n}{n}, \qquad j = 0, 1, 2, \cdots. \tag{8}$$

Without loss of generality, assume $\nu_0 = 1$. For detection we use the nonlinear correlator (NC) detector, which is of the form

$$T_n(\boldsymbol{X}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} s_i \psi(X_i), \tag{9}$$

and the decision is made as follows

$$u(\psi, \boldsymbol{X}) = \begin{cases} 1, & \text{if } T_n(\boldsymbol{X}) > \tau, \\ p, & \text{if } T_n(\boldsymbol{X}) = \tau, \\ 0, & \text{if } T_n(\boldsymbol{X}) < \tau, \end{cases} \tag{10}$$

where $u(\psi, \boldsymbol{X})$ is the probability of deciding H_1. The threshold τ and the probability p are selected to control the false alarm probability.

As we mentioned before, the performance measure we consider here is the efficacy. In order for the efficacy to exist and to be a valid measure we must impose restrictions on $\psi(x)$ that will determine the class Ψ of allowable nonlinearities. Let E_δ denote expectation under H_1 and E_0 under H_0. Also define $\delta_n = k/\sqrt{n}$, where k is any nonnegative constant. We assume that $\psi(x)$ is a measurable function, with $E\{[\psi(N_1)]^2\} < \infty$, that satisfies the following:

$$\psi(x) = -\psi(-x) \tag{11}$$

$$\frac{\partial}{\partial\delta}\int_{-\infty}^{\infty}\psi(x)f(x-\delta)\,dx\bigg|_{\delta=0} = \int_{-\infty}^{\infty}\frac{\partial}{\partial\delta}\psi(x)f(x-\delta)\bigg|_{\delta=0}dx \tag{12}$$

$$\int_{-\infty}^{\infty}\psi(x)f'(x)\,dx < 0 \tag{13}$$

$$\lim_{n\to\infty}\int_{-\infty}^{\infty}\psi(x)f'(x-\delta_n)\,dx = \int_{-\infty}^{\infty}\psi(x)f'(x)\,dx \tag{14}$$

$$\lim_{n\to\infty}\frac{\left[\frac{\partial}{\partial\delta}E_\delta\{T_n(X)\}\Big|_{\delta=0}\right]^2}{nE_0\{[T_n(X)]^2\}} > 0 \tag{15}$$

$$\lim_{t\to 0}E\{[\psi(N_1+t)-\psi(N_1)]^2\} = 0 \tag{16}$$

$$\sigma_0^2(\psi) = E\{\psi(N_1)^2\} + 2\sum_{j=1}^{\infty}\nu_j E\{\psi(N_1)\psi(N_{j+1})\} > 0. \tag{17}$$

With condition (11) we restrict the nonlinearities to be odd symmetric, a reasonable restriction since all locally optimum nonlinearities are odd symmetric under our assumptions. We now present two lemmas that will give us useful properties of the φ-mixing and the acceptable φ-mixing sequences.

Lemma 1: Let $\boldsymbol{N}$ be a φ-mixing sequence and Ξ and Θ, two random variables defined on M_1^k and M_{k+n}^{∞}, respectively. If $E\{|\Xi|^r\} < \infty$ and $E\{|\Theta|^q\} < \infty$ with $(1/r) + (1/q) = 1$ and $r, q > 1$, then

$$|E\{\Xi\Theta\} - E\{\Xi\}E\{\Theta\}| \le 2\varphi_n^{1/r}E^{1/r}\{|\Xi|^r\}E^{1/q}\{|\Theta|^q\}. \tag{18}$$

(For a proof see [13, p. 170].)

Lemma 2: Let $\psi(x)$ be a measurable function with $E\{\psi(N_1)\} = 0$ and $E\{[\psi(N_1)]^2\} < \infty$. Let also $\boldsymbol{N}$ be an acceptable φ-mixing sequence. Then $\sigma_0^2(\psi)$ defined in (17) is absolutely summable.

Proof: Using the inequality $|ab| \le \frac{1}{2}(a^2 + b^2)$, it is easy to see that $|\nu_j| \le \nu_0 = 1$. Thus it is enough to show that

$$\sum_{j=1}^{\infty}|E\{\psi(N_1)\psi(N_{j+1})\}| < \infty. \tag{19}$$

Applying (18) for $r = q = 2$ and $\Xi = \psi(N_1)$ and $\Theta = \psi(N_{j+1})$ and remembering that $\boldsymbol{N}$ is stationary, we obtain

$$\sum_{j=1}^{\infty}|E\{\psi(N_1)\psi(N_{j+1})\}| \le 2\left(\sum_{j=1}^{\infty}\varphi_j^{1/2}\right)E\{[\psi(N_1)]^2\} < \infty. \tag{20}$$

We now prove a proposition that gives the limiting form of (9) under both hypotheses. □

Proposition 1: Let $\boldsymbol{N}$ be an acceptable φ-mixing sequence, let $\{s_i\}$ be a bounded sequence of real numbers satisfying (8), and let $\psi(x)$ be a measurable function with $E\{\psi(N_1)\} = 0$ and $E\{[\psi(N_1)]^2\} < \infty$ that satisfies (16) and (17). Define $\lambda(t) = E\{\psi(N_1 + t)\}$, then

$$T_n^0 = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}s_i\psi(N_i) \xrightarrow{D} \mathbb{N}(0, \sigma_0(\psi)) \tag{21}$$

$$T_n^1 = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}s_i\psi(N_i + \delta_n s_i) - \frac{1}{\sqrt{n}}\sum_{i=1}^{n}s_i\lambda(\delta_n s_i) \xrightarrow{D} \mathbb{N}(0, \sigma_0(\psi)), \tag{22}$$

where $\xrightarrow{D}$ means-convergence in distribution. The proof of this proposition is given in the Appendix. Notice that (21) or (22) is not a direct consequence of [13, th. 20.1, p. 174] since the sequence $\{s_i\psi(N_i)\}$ is not a stationary sequence. For the proof we apply a more general theorem [13, th. 19.2, p. 157] and we show that our case satisfies all the hypotheses of this theorem.

Under our assumptions the Pitman–Noether theorem may be invoked, and the efficacy takes the following form:

$$\text{eff}(\psi(x), \boldsymbol{N}) = \frac{\left[\sum_{-\infty}^{\infty}\psi(x)f'(x)\,dx\right]^2}{\sigma_0^2(\psi)}, \tag{23}$$

where $\sigma_0^2(\psi)$ is defined in (17). Now we prove a lemma that gives us a property that characterizes the class $\mathscr{N}$.

Lemma 3: Let $\psi(x)$ be a measurable function with $E\{[\psi(N_1)]^2\} < \infty$. Also let $\boldsymbol{N} \in \mathscr{N}$. Then

$$|E\{\psi(N_1)\psi(N_{j+1})\} - [E\{\psi(N_1)\}]^2| \le \gamma_j[E\{|\psi(N_1)|\}]^2. \tag{24}$$

Proof: It is enough to show (24) for simple functions. Thus let

$$\psi(x) = \sum_{i=1}^{K}\psi_i I_{A_i}. \tag{25}$$

Let B_i be the event $\{N_1 \in A_i\}$ and C_i, the event $\{N_{j+1} \in A_i\}$. Then using (4), we obtain

$$\begin{aligned}|E\{\psi(N_1)\psi(N_{j+1})\} &- [E\{\psi(N_1)\}]^2| \\ &\le \sum_{i=1}^{K}\sum_{l=1}^{K}|\psi_i\psi_l||P(B_i \cap C_l) - P(B_i)P(C_l)| \\ &\le \gamma_j\sum_{i=1}^{K}\sum_{l=1}^{K}|\psi_i\psi_l|P(B_i)P(C_l) = \gamma_j[E\{|\psi(N_1)|\}]^2.\end{aligned} \tag{26}$$

The difference again between (24) and (18) is that (24) involves only bivariate densities.

III. Min–Max Detection

The problem we would like to solve is the following. Find a nonlinearity $\psi_r(x) \in \Psi$ and a sequence $N_r \in \mathcal{N}$ such that

$$\sup_{\psi(x)\in\Psi} \inf_{N\in\mathcal{N}} \operatorname{eff}(\psi(x), N) = \operatorname{eff}(\psi_r(x), N_r), \quad (27)$$

subject to the constraint

$$\sup_{N\in\mathcal{N}} P_{FA}(\psi_r(x), N) \le \alpha, \quad (28)$$

where P_{FA} is the asymptotic false alarm (FA) probability of deciding H_1 given H_0. We now proceed as follows: for a given $\psi(x)$ we find the $N \in \mathcal{N}$ that minimizes the efficacy. Then the resulting expression is maximized over the nonlinearity $\psi(x)$. The minimization is done in two steps. First, we keep the marginal density fixed and minimize over all sequences that have the same marginal, and then we minimize over the marginal. If the marginal is fixed, we can see from (23) that, in order to minimize the efficacy, we need to maximize $\sigma_0^2(\psi)$. Using (24) and remembering that $\psi(x)$ is odd symmetric (zero-mean), we have

$$\sigma_0^2(\psi) \le \int_{-\infty}^{\infty} \psi^2(x) f(x)\,dx + 2\left(\sum_{j=1}^{\infty} |\nu_j|\gamma_j\right)\left[\int_{-\infty}^{\infty} |\psi(x)| f(x)\,dx\right]^2. \quad (29)$$

The series in (29) is summable because

$$\sum_{j=1}^{\infty} |\nu_j|\gamma_j \le \sum_{j=1}^{\infty} \gamma_j < \infty. \quad (30)$$

We have equality in (29) when the bivariate densities $f_j(x, y)$ of N_1 and N_{j+1} are given by

$$f_j(x, y) = f(x)f(y)\{1 + \gamma_j \operatorname{sgn}(\nu_j) sn_\psi(x) sn_\psi(y)\}. \quad (31)$$

The function $sn_\psi(x)$ is defined to be odd symmetric and for $x > 0$ is equal to the sign of $\psi(x)$ when $\psi(x) \neq 0$, and may take any value in $[-1, 1]$ when $\psi(x) = 0$. Also $\operatorname{sgn}(\nu_j)$ is equal to the sign of ν_j when $\nu_j \neq 0$. When $\nu_j = 0$ the bivariate density can be anything. The odd symmetry of $sn_\psi(x)$ is important because it makes $f_j(x, y)$ a legitimate bivariate density with marginal $f(x)$. Even though these densities are of the right form, it is possible that there is no sequence in $\mathcal{N}$ that will have them as bivariate densities. Here we will assume that such a sequence always exists and, in the examples we present, we show a way to construct its multivariate density. We must point out that if we cannot show the existence of a sequence in $\mathcal{N}$, then this approach does not necessarily lead to the min–max solution. Let us now substitute (29) into the expression for the efficacy and call the resulting expression eff*. We thus obtain

$$\operatorname{eff}^*(\psi(x), f(x)) = \frac{\left[\int_{-\infty}^{\infty} \psi(x) f'(x)\,dx\right]^2}{\int_{-\infty}^{\infty} \psi^2(x) f(x)\,dx + \gamma\left[\int_{-\infty}^{\infty} |\psi(x)| f(x)\,dx\right]^2}, \quad (32)$$

where $\gamma = 2\sum_{j=1}^{\infty} |\nu_j|\gamma_j$. The case $\gamma = 0$ is of little interest here since it is no different from the i.i.d. case. Thus we assume $\gamma > 0$. Next, we have to find a pair $\psi_r(x)$ and $f_r(x)$ such that

$$\sup_{\psi(x)\in\Psi} \inf_{f(x)} \operatorname{eff}^*(\psi(x), f(x)) = \operatorname{eff}^*(\psi_r(x), f_r(x)). \quad (33)$$

It turns out that this new min–max problem defined by (33) has a saddle point; in other words, the pair $\psi_r(x)$ and $f_r(x)$ satisfies the following double inequality

$$\operatorname{eff}^*(\psi(x), f_r(x)) \le \operatorname{eff}^*(\psi_r(x), f_r(x)) \le \operatorname{eff}^*(\psi_r(x), f(x)) \quad (34)$$

for any $\psi(x) \in \Psi$ and any $f(x)$ satisfying (6). Any pair that satisfies (34) is known to satisfy (33). Thus we will solve (34) instead of (33). The left inequality in (34) indicates that $\psi_r(x)$ is the optimum nonlinearity for $f_r(x)$ when the criterion function is the eff*. The following theorem gives the form of this optimum nonlinearity in terms of the marginal density.

Theorem 1: Let $f(x)$ be a symmetric density with finite Fisher's information and continuous derivative different from zero almost everywhere with respect to $f(x)$. Then the optimum nonlinearity $\psi_0(x)$ that maximizes the eff* is given by

$$\psi_0(x) = -\frac{f'(x)}{f(x)} - \mu\pi_0(x), \quad (35)$$

where $\pi_0(x)$ is defined as follows

$$\pi_0(x) = \begin{cases} -\dfrac{1}{\mu}\dfrac{f'(x)}{f(x)}, & \text{for } -1 \le -\dfrac{1}{\mu}\dfrac{f'(x)}{f(x)} \le 1, \\ 1, & \text{for } 1 \le -\dfrac{1}{\mu}\dfrac{f'(x)}{f(x)}, \\ -1, & \text{for } -1 \ge -\dfrac{1}{\mu}\dfrac{f'(x)}{f(x)}, \end{cases} \quad (36)$$

and μ is a positive constant that satisfies

$$S(\mu) = \mu + \frac{\int_{-\infty}^{\infty} f'(x)\pi_0(x)\,dx}{\dfrac{1}{\gamma} + \int_{-\infty}^{\infty} \pi_0^2(x) f(x)\,dx} = 0. \quad (37)$$

The proof of this theorem is given in the Appendix. From (35) and (36) we see that $\psi_0(x)$ is odd symmetric and closely related to the locally optimum nonlinearity. The function $\pi_0(x)$ is defined in a way such that $\psi_0(x)$ becomes zero whenever $-f'(x)/f(x)$ takes on values between $-\mu$ and μ. Now we are ready to define the pair that satisfies the saddle-point relation (34). Since $\psi_r(x)$ is optimum for $f_r(x)$, we need to define only $f_r(x)$ and this is done in the following theorem.

Theorem 2: The density $f_r(x)$ that gives the solution to the saddle-point problem defined by (34) is the following

$$f_r(x) = \begin{cases} (1-\epsilon)g(x_1)e^{x_1(x+x_1)}, & \text{for } x \le -x_1, \\ (1-\epsilon)g(x), & \text{for } |x| \le x_1, \\ (1-\epsilon)g(x_1)e^{-x_1(x-x_1)}, & \text{for } x \ge x_1, \end{cases} \tag{38}$$

where $x_1 \ge 0$ and such that $f_r(x)$ has total mass equal to unity.

Proof: This density is exactly the one defined by Huber in [1], [2] for the i.i.d. case. It belongs to the ϵ-contamination class with a density $h_r(x)$ that places all the mass outside the interval $[-x_1, x_1]$

$$\epsilon h_r(x) = \begin{cases} (1-\epsilon)\left[g(x_1)e^{x_1(x+x_1)} - g(x)\right], & \text{for } x \le -x_1, \\ 0, & \text{for } |x| \le x_1, \\ (1-\epsilon)\left[g(x_1)e^{-x_1(x-x_1)} - g(x)\right], & \text{for } x \ge x_1. \end{cases} \tag{39}$$

The nonnegativity of $h_r(x)$ can be proven using the strong unimodality of $g(x)$ (see [1], [2]). To find the nonlinearity $\psi_r(x)$ we use Theorem 1 and obtain

$$\psi_r(x) = \begin{cases} 0, & \text{for } 0 \le x \le x_2, \\ -\dfrac{g'(x)}{g(x)} + \dfrac{g'(x_2)}{g(x_2)}, & \text{for } x_2 \le x \le x_1, \\ -\dfrac{g'(x_1)}{g(x_1)} + \dfrac{g'(x_2)}{g(x_2)}, & \text{for } x_1 \le x. \end{cases} \tag{40}$$

For $x \le 0$ we recall that $\psi_r(x)$ is odd symmetric. We define x_2 as

$$-\frac{g'(x_2)}{g(x_2)} = \mu, \tag{41}$$

where μ is a solution to the equation defined by (37). In order for (40) to be valid, the x_2 defined by (41) must satisfy $0 \le x_2 \le x_1$. In the Appendix we show that such an x_2 always exists. Up to this point, because of Theorem 1, we have that $\psi_r(x)$ and $f_r(x)$ satisfy the left inequality in (34). To prove that they also satisfy the right inequality, notice that, since $g(x)$ is strongly unimodal, we have that $\psi_r(x)$ is a nondecreasing function. If we define

$$M = -\frac{g'(x_1)}{g(x_1)} + \frac{g'(x_2)}{g(x_2)}, \tag{42}$$

then, since $\psi_r(x)$ is odd and nondecreasing, we have $|\psi_r(x)| \le M$. Call $n(f)$ and $d(f)$ the numerator and the denominator of the eff*$(\psi_r(x), f(x))$; then

$$\begin{aligned} n(f) &= \left[\int_{-\infty}^{\infty} \psi_r(x) f'(x)\, d(x)\right]^2 \\ &= \left[(1-\epsilon)\int_{-\infty}^{\infty} \psi_r(x) g'(x)\, dx \right. \\ &\quad \left. + \epsilon \int_{-\infty}^{\infty} \psi_r(x) h'(x)\, dx\right]^2. \end{aligned} \tag{43}$$

Because $\psi_r(x)$ is nondecreasing, the two terms in the last expression are nonpositive; thus

$$n(f) \ge \left[(1-\epsilon)\int_{-\infty}^{\infty} \psi_r(x) g'(x)\, dx\right]^2 = n(f_r), \tag{44}$$

also

$$\begin{aligned} d(f) &= \int_{-\infty}^{\infty} \psi_r^2(x) f(x)\, dx + \gamma\left[\int_{-\infty}^{\infty} |\psi_r(x)| f(x)\, dx\right]^2 \\ &\le (1-\epsilon)\int_{-\infty}^{\infty} \psi_r^2(x) g(x)\, dx + \epsilon M^2 \\ &\quad + \gamma\left[(1-\epsilon)\int_{-\infty}^{\infty} |\psi_r(x)| g(x)\, dx + \epsilon M\right]^2 \\ &= d(f_r). \end{aligned} \tag{45}$$

Thus $f_r(x)$ simultaneously minimizes the numerator and maximizes the denominator of the eff*$(\psi_r(x), f(x))$, which means that the right-hand inequality in (34) is also satisfied. This concludes the proof. □

Returning now to our original min–max problem defined in (27), we have that $\psi_r(x)$ is the nonlinearity defined in (40) and N_r is any sequence from $\mathscr{N}$ that has bivariate densities given by

$$f_j^r(x, y) = f_r(x) f_r(y)\left\{1 + \gamma_j \operatorname{sgn}(\nu_j) sn_{\psi_r}(x) sn_{\psi_r}(y)\right\}, \tag{46}$$

where $f_r(x)$ is defined in (38).

In order now to satisfy (28) we must set the threshold for the detection structure

$$T_n^r(X) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi_r(X_i). \tag{47}$$

Notice that $T_n^r(X)$ under H_0 is Gaussian in the limit. Hence if (28) is satisfied for the sequence that has the maximum asymptotic variance it will be satisfied for any sequence. But the asymptotic variance of (47) is the square root of the denominator of the eff$(\psi_r(x), N)$. This denominator is maximized when $N = N_r$ and the maximum value is equal to $d(f_r)$ defined in (45). Thus the threshold τ is given by

$$1 - \Phi\left(\frac{\tau}{[d(f_r)]^{1/2}}\right) = \alpha, \tag{48}$$

where $\Phi(x)$ is the $\mathbb{N}(0,1)$ Gaussian cumulative distribution. Note that the sequence N_r achieves simultaneously the worst performance for the efficacy and for the false alarm probability. Before going to Section IV we must point out that the φ-mixing assumption and the boundedness of the sequence $\{s_i\}$ were used only in the proof of Proposition 1. Thus, as long as the central limit theorem holds, the optimum nonlinearity will be given by Theorem 2 for cases where the bivariate densities are given by (4)–(6). For the case where we cannot prove existence of a sequence in $\mathcal{N}$ that has bivariate densities defined by (31), we still satisfy a min–max relation. Only, instead of the class $\mathcal{N}$ of sequences, we will consider the class of bivariate densities that satisfy (4)–(6). In other words, the min–max problem will be defined over a larger class of bivariate densities than the one that $\mathcal{N}$ defines. Thus the lower performance bound on the min–max problem will not necessarily be the best for the class $\mathcal{N}$.

IV. Examples

As we can see from Theorem 2 the min–max nonlinearity and its worst performance depend on the density $g(x)$ and the constants ϵ and γ and not on the actual sequences $\{\gamma_i\}$ and $\{s_i\}$. In the following we give tables for the point x_2 and the performance of the min–max nonlinearity, for the case in which $g(x)$ is $\mathbb{N}(0,1)$. In Table I there are the values of x_2 for different γ and ϵ. The parameter x_1 depends only on ϵ, and it turns out that as $\gamma \to \infty$ then $x_2 \to x_1$. Thus the last column ($\gamma = \infty$ in the table) gives also the values for x_1. Table II gives the values of the ARE of $\psi_r(x)$ versus the locally optimum nonlinearity $-f_r'(x)/f_r(x)$ when the underlying sequence is the N_r. Notice that this locally optimum nonlinearity would have been the one to use if we had falsely assumed that the observations were i.i.d. Now we present two cases where the theory in Section III can be applied.

TABLE I
Values for x_2 (Last Column Values for x_1)

ϵ \ γ	1.0	2.0	3.0	4.0	5.0	10.0	20.0	∞
0.001	0.436	0.635	0.764	0.860	0.935	1.173	1.412	2.630
0.01	0.431	0.626	0.753	0.845	0.917	1.140	1.353	1.945
0.05	0.410	0.519	0.704	0.784	0.845	1.019	1.161	1.399
0.1	0.386	0.549	0.648	0.716	0.766	0.902	1.000	1.140
0.15	0.361	0.510	0.597	0.655	0.697	0.807	0.883	0.980
0.2	0.337	0.472	0.549	0.600	0.636	0.730	0.788	0.862
0.3	0.291	0.402	0.436	0.502	0.530	0.595	0.637	0.685
0.4	0.247	0.337	0.385	0.416	0.436	0.485	0.515	0.550
0.5	0.204	0.276	0.314	0.337	0.353	0.390	0.412	0.436
0.8	0.080	0.107	0.121	0.130	0.135	0.147	0.154	0.162

TABLE II
ARE of $\psi_r(x)$ Versus the Locally Optimum Nonlinearity

ϵ \ γ	1.0	2.0	3.0	4.0	5.0	10.0	20.0	∞
0.001	1.08	1.19	1.29	1.37	1.45	1.74	2.11	4.48
0.01	1.08	1.18	1.26	1.34	1.40	1.63	1.89	2.68
0.05	1.07	1.14	1.21	1.26	1.30	1.46	1.56	1.77
0.1	1.06	1.12	1.17	1.20	1.23	1.32	1.39	1.49
0.15	1.05	1.10	1.13	1.16	1.18	1.24	1.29	1.35
0.2	1.04	1.08	1.11	1.13	1.15	1.19	1.22	1.27
0.3	1.03	1.06	1.08	1.09	1.10	1.12	1.14	1.16
0.4	1.02	1.04	1.05	1.06	1.07	1.08	1.09	1.10

One-Dependent Case: Assume $\gamma_j = 0$ for $j \geq 2$. Here $\gamma = 2\gamma_1|\nu_1|$, and the bivariate densities defined in (31) take the form

$$f_1(x,y) = f(x)f(y)\{1 + \gamma_1 \operatorname{sgn}(\nu_1) sn_\psi(x) sn_\psi(y)\}$$
$$f_j(x,y) = f(x)f(y), \qquad j = 2,3,\cdots. \tag{49}$$

In order to be able to apply the results in Section III, we will show that there exists a multivariate density of a stationary sequence in $\mathcal{N}$ that has bivariate densities given by (49). To define such a density, let U be a stationary one-dependent sequence of random variables with U_i supported on $[-1,1]$ and

$$E\{U_1U_2\} = \gamma_1 \operatorname{sgn}(\nu_1). \tag{50}$$

Then the following expression is a multivariate density that possesses the properties we need

$$f(x_1,x_2,\cdots,x_n) = E_U\left\{\prod_{i=1}^{n} f(x_i)\left[1 + U_i sn_\psi(x_i)\right]\right\}, \tag{51}$$

where E_ν means expectation with respect to the sequence U. It is easy to see that (51) defines a density of a one-dependent sequence. To show this, notice that the sets $\{U_1,\cdots,U_{k-1}\}$ and $\{U_{k+1},\cdots,U_n\}$ are independent. Thus by interchanging integrals and expectations, we have

$$\begin{aligned} f(x_1,\cdots,x_{k-1},x_{k+1},\cdots,x_n) &= \int_{-\infty}^{\infty} f(x_1,\cdots,x_n)\,dx_k \\ &= E_U\left\{\left(\prod_{i=1}^{k-1} f(x_i)\left[1 + U_i sn_\psi(x_i)\right]\right)\cdot\left(\prod_{i=k+1}^{n} f(x_i)\left[1 + U_i sn_\psi(x_i)\right]\right)\right\} \\ &= E_U\left\{\prod_{i=1}^{k-1} f(x_i)\left[1 + U_i sn_\psi(x_i)\right]\right\}\cdot E_U\left\{\prod_{i=k+1}^{n} f(x_i)\left[1 + U_i sn_\psi(x_i)\right]\right\} \\ &= f(x_1,\cdots,x_{k-1}) f(x_{k+1},\cdots,x_n). \end{aligned} \tag{52}$$

Also this sequence is an acceptable φ-mixing sequence because it is an m-dependent sequence. The only problem is that (50) cannot be true for an arbitrary value of γ_1. For example, if we generate the U_i from the following model

$$U_i = \frac{\Lambda_i + r\Lambda_{i+1}}{1 + |r|}, \tag{53}$$

where the Λ_i are i.i.d. with support on $[-1, 1]$ and r a real number, we can realize only $\gamma_1 \le \frac{1}{4}$. The one-dependent case can be extended to the m-dependent by taking the sequence $\boldsymbol{U}$ to be m-dependent.

Markov Case: Assume $\gamma_j = m^j$ with $0 \le m < 1$ and $\nu_j \ge 0$. This is the case treated in [11]. Here $\gamma = 2\sum_{j=1}^{\infty} m^j \nu_j$, and

$$f_j(x, y) = f(x)f(y)\{1 + m^j sn_\psi(x) sn_\psi(y)\}. \tag{54}$$

Now $f_1(x, y)$, the density between consecutive points, defines a strictly stationary Markov sequence with multivariate density

$$f(x_1, x_2, \cdots, x_n) = \prod_{j=1}^{n} f(x_j) \prod_{j=1}^{n-1} \{1 + m sn_\psi(x_j) sn_\psi(x_{j+1})\}. \tag{55}$$

If the function $sn_\psi(x)$ takes values of only $+1$ or -1 (always possible), using induction we can show that the resulting Markov sequence has bivariate densities given by (54). To show that it is an acceptable φ-mixing sequence consider $A \in M_1^k$ and $B \in M_{k+n}^q$ with q any integer such that $q \ge k + n$. Let A_i be the range of values of N_i under the event A, for $i = 1, \cdots, k$. Similarly let B_j be the range of values of N_j under B, for $j = k + n, \cdots, q$. Define $A' = A_1 \times \cdots \times A_k$ and $B' = B_{k+n} \times \cdots \times B_q$. The two sets A' and B' are Borel sets. Then, because we have a Markov sequence,

$$P(A \cap B) = \int_{A' \times B'} f(x_1, \cdots, x_k) \frac{f_n(x_k, x_{k+n})}{f(x_k) f(x_{k+n})} \cdot f(x_{k+n}, \cdots, x_q)\, d\boldsymbol{x}, \tag{56}$$

where $\boldsymbol{x} = (x_1, \cdots, x_k, x_{k+n}, \cdots, x_q)$, and because of (54),

$$\begin{aligned} |P(A \cap B) - P(A)P(B)| &\le m^n \left| \int_{A' \times B'} f(x_1, \cdots, x_k) f(x_{k+n}, \cdots, x_q) \cdot \{sn_\psi(x_k) sn_\psi(x_{k+n})\}\, d\boldsymbol{x} \right| \\ &\le m^n P(A)P(B) \le m^n P(A). \end{aligned} \tag{57}$$

From (57) we conclude that the sequence is acceptable φ-mixing with $\varphi_j = m^j$.

V. Conclusion

In this paper we have found the min–max nonlinearity for detection of signals in a class of φ-mixing noise processes. This nonlinearity was shown to be closely related to the min–max nonlinearity for the i.i.d. case. However the performances of the two nonlinearities, in the example that was presented, were drastically different.

Appendix

Proof of Theorem 1: Notice that in (32) the value of the eff* does not change if we multiply the nonlinearity by a constant. Thus we maximize the numerator assuming that the denominator has some fixed value. Using (13) this is equivalent to maximizing the following expression

$$H(\psi) = -\int_{-\infty}^{\infty} \psi(x) f'(x)\, dx - \rho\left(\int_{-\infty}^{\infty} \psi^2(x) f(x)\, dx + \gamma\left[\int_{-\infty}^{\infty} |\psi(x)| f(x)\, dx\right]^2\right), \tag{A1}$$

where ρ is a Lagrange multiplier. We will show that (A1) is maximized by

$$\psi_0(x) = \frac{1}{2\rho}\left[-\frac{f'(x)}{f(x)} - \mu\pi_0(x)\right], \tag{A2}$$

where μ and $\pi_0(x)$ were defined in (36), (37). Let $\psi_1(x)$ be some other nonlinearity from the class Ψ. Define the following variation

$$\begin{aligned} J(\xi) = &-\int_{-\infty}^{\infty} [(1 - \xi)\psi_0(x) + \xi\psi_1(x)] f'(x)\, dx \\ &- \rho\left(\int_{-\infty}^{\infty} [(1 - \xi)\psi_0(x) + \xi\psi_1(x)]^2 f(x)\, dx \right. \\ &\left. + \gamma\left[\int_{-\infty}^{\infty} \{(1 - \xi)|\psi_0(x)| + \xi|\psi_1(x)|\} f(x)\, dx\right]^2\right), \end{aligned} \tag{A3}$$

where $\xi \in [0, 1]$. Notice that $J(0) = H(\psi_0)$ and $J(1) = H(\psi_1)$. By manipulating (A3) we can write it as

$$J(\xi) - J(0) = I_1 + I_2 + I_3, \tag{A4}$$

where

$$\begin{aligned} I_1 = \xi\int_{\infty}^{\infty} \Big\{ &-f'(x) - 2\rho\psi_0(x) f(x) \\ &- 2\gamma\rho\left[\int_{-\infty}^{\infty} |\psi_0(z)| f(z)\, dz\right] \pi_0(x) f(x)\Big\} [\psi_1(x) - \psi_0(x)]\, dx \end{aligned} \tag{A5}$$

$$\begin{aligned} I_2 = -2\xi\gamma\rho\left[\int_{-\infty}^{\infty} |\psi_0(z)| f(z)\, dz\right] &\cdot \left(\int_{-\infty}^{\infty} \{|\psi_1(x)| - |\psi_0(x)|\} f(x)\, dx \right. \\ &\left. - \int_{-\infty}^{\infty} \pi_0(x)[\psi_1(x) - \psi_0(x)] f(x)\, dx\right) \end{aligned} \tag{A6}$$

$$I_3 = -\xi^2\rho\left(\int_{-\infty}^{\infty} [\psi_1(x) - \psi_0(x)]^2 f(x)\, dx + \gamma\left[\int_{-\infty}^{\infty} \{|\psi_1(x)| - |\psi_0(x)|\} f(x)\, dx\right]^2\right). \tag{A7}$$

To prove that $J(0)$ is the maximum, it is enough to show that $J(\xi) - J(0) \le 0$ or that $I_i \le 0$ for $i = 1, 2, 3$. From the definition of $\pi_0(x)$ in (36) notice the following:

$$|\psi_0(x)| = \psi_0(x)\pi_0(x) \tag{A8}$$

$$|\pi_0(x)| \le 1. \tag{A9}$$

If we multiply (A2) by $\pi_0(x)f(x)$ and integrate and also use (37), we obtain

$$\int_{-\infty}^{\infty} |\psi_0(x)| f(x)\,dx = \frac{1}{2\rho}\left[-\int_{-\infty}^{\infty} f'(x)\pi_0(x)\,dx - \mu\int_{-\infty}^{\infty}\pi_0^2(x)\,dx\right] = \frac{\mu}{2\rho\gamma}. \tag{A10}$$

Substituting (A10) in the expression for I_1 and using (A2), we get zero. On using (A8), the term I_2 becomes

$$I_2 = -2\xi\rho\gamma\left[\int_{-\infty}^{\infty}|\psi_0(z)|f(z)\,dz\right] \cdot\left(\int_{-\infty}^{\infty}\{|\psi_1(x)| - \pi_0(x)\psi_1(x)\}f(x)\,dx\right). \tag{A11}$$

Because of (A9) we have $|\psi_1(x)| \geq \pi_0(x)\psi_1(x)$, and thus for $\rho > 0$, the I_2 becomes nonpositive. Finally, for $\rho > 0$, the I_S is clearly nonpositive too. If we define $\rho = 1/2$ then (A2) becomes the same as (35).

In order to complete the proof of Theorem 1 we must show that the equation defined in (37) has always a solution. Using continuity arguments, it is enough to show existence of two points μ_1 and μ_2 such that $S(\mu_1)S(\mu_2) \leq 0$. Notice that as $\mu \to 0$ then $-(1/\mu)(f'(x)/f(x)) \to \pm\infty$ except on sets of f-measure zero. Thus, $\pi_0(x) \to \operatorname{sgn}(-f'(x)/f(x)) = -\operatorname{sgn}(f'(x))$. Substituting in (37), we find

$$S(0) = \frac{\int_{-\infty}^{\infty}|f'(x)|\,dx}{\frac{1}{\gamma}+1} < 0. \tag{A12}$$

Now using (A9) and the Schwarz inequality, we have

$$\left|\int_{-\infty}^{\infty} f'(x)\pi_0(x)\,dx\right| \leq \int_{-\infty}^{\infty}|f'(x)|\,dx \leq \left[\int_{-\infty}^{\infty}\left(\frac{f'(x)}{f(x)}\right)^2 f(x)\,dx\right]^{1/2} = [I(f)]^{1/2} < \infty, \tag{A13}$$

where $I(f)$ is Fisher's information. Thus the second term in (37) is bounded by $\gamma[I(f)]^{1/2}$, and as $\mu \to +\infty$, we have that $S(\mu) \to +\infty$ or it becomes positive. And this concludes the proof of Theorem 1. □

Existence of x_2: In Theorem 2 we assumed that there exists an x_2, with $0 \leq x_2 \leq x_1$ that satisfies (41), where μ satisfies (37). Now, because $S(0) < 0$, if we show that $S(-g'(x_1)/g(x_1)) \geq 0$, then there exists a solution to (37) that will satisfy $0 < \mu \leq -g'(x_1)/g(x_1)$, and because of the monotonicity of $-g'(x)/g(x)$, we will have $0 \leq x_2 \leq x_1$. To prove this, notice that the locally optimum nonlinearity for the $f_r(x)$ defined in (38) is

$$-\frac{f_r'(x)}{f_r(x)} = \begin{cases} -\dfrac{g'(x)}{g(x)}, & \text{for } |x| \leq x_1, \\ -\dfrac{g'(x_1)}{g(x_1)}\operatorname{sgn}(x), & \text{for } |x| \geq x_1. \end{cases} \tag{A14}$$

Thus for $\mu = -g'(x_1)/g(x_1)$ we are always in the first case of (36), and we have

$$\pi_0(x) = -\frac{1}{\mu}\frac{f_r'(x)}{f_r(x)}. \tag{A15}$$

Substituting into (37) we obtain

$$S\left(-\frac{g'(x_1)}{g(x_1)}\right) = -\frac{g'(x_1)}{g(x_1)}\left(1 - \frac{I(f_r)}{\frac{1}{\gamma}\left[\frac{g'(x_1)}{g(x_1)}\right]^2 + I(f_r)}\right) \geq 0. \tag{A16}$$

And this proves the existence of x_2.

Proof of Proposition 1: The proof is based on the following theorem that gives sufficient conditions for convergence. □

Theorem 3: Let U be a sequence of random variables and define

$$R_n = \sum_{i=1}^{n} U_i \tag{A17}$$

$$X_n(t) = \frac{1}{\sqrt{n}}\sum_{i=1}^{[nt]} U_i, \qquad t \in [0,1], \tag{A18}$$

where by $[\cdot]$ we mean the integer part. If

a) $X_n(t)$ has asymptotically independent increments;
b) $E\{X_n(t)\} \to 0$ and $E\{X_n^2(t)\} \to \sigma^2 t$;
c) $X_n^2(t)$ is uniformly integrable; and
d) for every ϵ there exist a $\beta > 1$ and an integer n_0 such that

$$P\left\{\max_{i\leq n}|R_{k+i} - R_k| \geq \beta\sigma\sqrt{n}\right\} \leq \frac{\epsilon}{\beta^2}, \quad \text{for every } k \text{ and } n \geq n_0,$$

then $X_n(t)$ tends in distribution to a Brownian motion. For the proof see [13, ths. 8.4 and 19.2].

We will now show that our case satisfies the conditions given in Theorem 3. We will base our proof on the fact that an acceptable φ-mixing sequence satisfies all the above conditions (see [13, th. 20.1]). We prove first (21). Let us define

$$R_n = \sum_{i=1}^{n} s_i\psi(N_i) \tag{A19}$$

$$X_n(t) = \frac{1}{\sqrt{n}}\sum_{i=1}^{[nt]} s_i\psi(N_i) \tag{A20}$$

$$S_n = \sum_{i=1}^{n}\psi(N_i) \tag{A21}$$

$$Y_n(t) = \frac{1}{\sqrt{n}}\sum_{i=1}^{[nt]}\psi(N_i). \tag{A22}$$

Notice that $X_n(t)$ is the process which we want to show satisfies the conditions of Theorem 3 and $Y_n(t)$ is the stationary case we know satisfies the conditions.

To show condition a), let $0 \leq u_1 \leq v_1 < u_2 \leq v_2 < \cdots < u_r \leq v_r \leq 1$ and $b = \min_i(u_i - v_{i-1})$. Also let $A_i \in M_{[nu_i]}^{[nv_i]}$. Then from (2), using induction, we can show

$$\left|P\left(\bigcap_{i=1}^{r} A_i\right) - \prod_{i=1}^{r} P(A_i)\right| \leq r\varphi_{[nb]} \to 0. \tag{A23}$$

For condition b), we have $E\{X_n(t)\} = 0$. Also for $t > 0$,

$$E\{X_n^2(t)\} = \frac{[nt]}{n}E\left\{\left[\frac{1}{\sqrt{[nt]}}\sum_{i=1}^{[nt]} s_i\psi(N_i)\right]^2\right\} \to t\sigma_0^2(\psi). \tag{A24}$$

To show condition c), we have to show that

$$\lim_{\alpha\to\infty} \sup_n \int_{|X_n|>\alpha} X_n^2(t)\, dP = 0, \qquad \text{for every } t. \tag{A25}$$

Let M be a bound for the sequence $\{|s_i|\}$. Then, since $|s_i/M| \le 1$, we have for the two events

$$\left\{\frac{1}{\sqrt{n}}\left|\sum_{i=1}^{[nt]} \frac{s_i}{M}\psi(N_i)\right| > \alpha\right\} \subseteq \left\{\frac{1}{\sqrt{n}}\left|\sum_{i=1}^{[nt]} \psi(N_i)\right| > \alpha\right\}. \tag{A26}$$

Thus if we define $\lambda_1(\omega) = P\{X_n^2(t) > \omega\}$ and $\lambda_2(\omega) = P\{Y_n^2(t) > \omega\}$, we have that

$$\lambda_1(M^2\omega) \le \lambda_2(\omega). \tag{A27}$$

Because both $X_n(x)$ and $Y_n(t)$ have finite variance, we can write

$$\begin{aligned}\int_{X_n^2(t)>\alpha^2M^2} X_n^2(t)\, dP &= \alpha^2M^2\lambda_1(\alpha^2M^2) + \int_{\alpha^2M^2}^{\infty}\lambda_1(\omega)\, d\omega\\ &= M^2\left(\alpha^2\lambda_1(\alpha^2M^2) + \int_{\alpha^2}^{\infty}\lambda_1(\omega M^2)\, d\omega\right)\\ &\le M^2\left(\alpha^2\lambda_2(\alpha^2) + \int_{\alpha^2}^{\infty}\lambda_2(\omega)\, d\omega\right) = M^2\int_{|Y_n^2(t)|>\alpha^2} Y_n^2(t)\, dP\end{aligned} \tag{A28}$$

and, because $Y_n(t)$ is the stationary case, we have

$$\lim_{\alpha\to\infty}\sup_n \int_{|X_n(t)|>\alpha M} X_n^2(t)\, dP \le \lim_{\alpha\to\infty}\sup_n M^2\int_{|Y_n(t)|>\alpha} Y_n^2(t)\, dP = 0. \tag{A29}$$

To prove condition d), notice that

$$P\left\{\max_{i\le n}|R_{k+i} - R_k| \ge \beta\sigma\sqrt{n}\right\} = P\left\{\bigcup_{i=1}^{n}\left\{\left|\sum_{j=1}^{i} s_{j+k}\psi(N_j)\right| \ge \beta\sigma\sqrt{n}\right\}\right\}, \tag{A30}$$

and thus using (A26),

$$\begin{aligned}P\left\{\max_{i\le n}|R_{k+i} - R_k| > M\beta\sigma\sqrt{n}\right\} &\le P\left\{\bigcup_{i=1}^{n}\left\{|S_i| > \beta\sigma\sqrt{n}\right\}\right\}\\ &= P\left\{\bigcup_{i=1}^{n}\left\{|S_{i+k} - S_k| > \beta\sigma\sqrt{n}\right\}\right\}\\ &= P\left\{\max_{i\le n}|S_{i+k} - S_k| > \beta\sigma\sqrt{n}\right\} < \frac{\epsilon}{\beta^2}\end{aligned} \tag{A31}$$

for sufficiently large β and n. The last inequality is true because it comes from the stationary case. And if we define $\epsilon' = \epsilon M^2$ and $\beta' = \beta M$, we prove d). And this concludes the proof of (21).

To show (22) it is enough to prove that (see [13, p. 25])

$$\lim_{n\to\infty} E\left\{\left[T_n^0 - T_n^1\right]^2\right\} = 0. \tag{A32}$$

Notice that

$$E\left\{\left[T_n^0 - T_n^1\right]^2\right\} = \frac{1}{n}E\left\{\left[\sum_{i=1}^{n} s_iW_i\right]^2\right\}, \tag{A33}$$

where we define

$$W_i = \psi(N_i + \delta_n s_i) - \psi(N_i) - \lambda(\delta_n s_i). \tag{A34}$$

Because $\{s_i\}$ is bounded by M, if we use (16), for large enough n we can have $E\{W_i^2\} < \epsilon$ for every i, where ϵ an arbitrary positive real. Thus, using Lemma 1 for $r = q = 2$ and making n large enough, we have

$$\begin{aligned}&E\left\{\left[T_n^0 - T_n^1\right]^2\right\}\\ &\le \frac{1}{n}\left\{\sum_{i=1}^{n} s_i^2E\{W_i^2\} + 4\sum_{k<j}^{n}\sum^{n} s_ks_j\varphi_{j-k}^{1/2}\left[E\{W_k^2\}E\{W_j^2\}\right]^{1/2}\right\}\\ &\le M^2\left\{1 + 4\sum_{i=1}^{\infty}\varphi_i^{1/2}\right\}\epsilon,\end{aligned} \tag{A35}$$

and this concludes the proof of the Proposition 1. □

References

[1] P. J. Huber, "Robust estimation of a location parameter," *Ann. Math. Statist.*, vol. 35, pp. 73–101, 1964.
[2] ——, "A robust version of the probability ratio test," *Ann. Math. Statist.*, vol. 36, pp. 1753–1758, 1965.
[3] R. D. Martin and S. C. Schwartz, "Robust detection of a known signal in nearly Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 50–56, Jan. 1971.
[4] S. A. Kassam and J. B. Thomas, "Asymptotically robust detection of a known signal in contaminated non-Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 22–26, Jan. 1976.
[5] S. A. Kassam, G. V. Moustakides, and J. G. Shin, "Robust detection of known signals in asymmetric noise," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 84–91, Jan. 1982.
[6] H. V. Poor and J. B. Thomas, "Memoryless discrete time detection of a constant signal in *m*-dependent noise," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 54–61, Jan. 1979.
[7] D. R. Halverson and G. L. Wise, "Discrete time detection in φ-mixing noise," *IEEE Trans. Inform. Theory*, vol. IT-26, pp. 189–198, Mar. 1980.
[8] H. V. Poor, "Signal detection in the presence of weakly dependent noise—Part II: Robust detection," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 744–752, Sept. 1982.
[9] S. L. Portnoy, "Robust estimation in dependent situations," *Ann. Statist.*, vol. 5, pp. 22–43, 1977.
[10] ——, "Further remarks on robust estimation in dependent situations," *Ann. Statist.*, vol. 7, pp. 224–231, 1979.
[11] G. V. Moustakides and J. B. Thomas, "Min–max detection of a weak signal in stationary Markov noise," in *Proc. 20th Ann. Allerton Conf. on Commun., Control, and Computing*, Oct. 1982, pp. 713–722.
[12] N. N. Vorob'ev, "Consistent families of measures and their extensions," *Theory Prob. Appl.* vol. VII, no. 2, pp. 147–163, 1962.
[13] J. Billingsley, *Convergence of Probability Measures.* New York: Wiley, 1968.