ESAIM: Probability and Statistics November 2002, Vol. 6, 239–258
URL: http://www.emath.fr/ps/ DOI: 10.1051/ps:2002013

GOODNESS-OF-FIT TEST FOR LONG RANGE DEPENDENT PROCESSES

Gilles Fay[1] and Anne Philippe[1]

Abstract. In this paper, we make use of the information measure introduced by Mokkadem (1997) for building a goodness-of-fit test for long-range dependent processes. Our test statistic is performed in the frequency domain and writes as a non linear functional of the normalized periodogram. We establish the asymptotic distribution of our statistic under the null hypothesis. Under specific alternative hypotheses, we prove that the power converges to one. The performance of our test procedure is illustrated from different simulated series. In particular, we compare its size and its power with test of Chen and Deo.

Mathematics Subject Classification. 60F05, 62F03.

1. Introduction and preliminaries

This paper investigates a goodness-of-fit (GOF) test for possibly long range dependent (LRD) time series. This test is performed in the frequency domain, whereas many well known goodness-of-fit tests more likely check for the whiteness of the residuals for a fitted model (see Box and Pierre [4]). They include the so-called Portmanteau tests, which have recently been generalized in the spectral domain by (Chen and Deo [6]). The authors show the validity of a GOF procedure in short-range dependence (SRD) and also in long-range dependence, but under Gaussian assumption. However, the power of their tests against interesting alternatives (neither fixed nor local ones) is not derived.

The empirical spectral measure (or integrated periodogram) $\int_{-\pi}^{x} I_n(\lambda)\mathrm{d}\lambda$ became a very popular alternative approach to goodness of fit test for spectral densities (see Barlett [3]; Grenander and Rosenblatt [12]) since this quantity inherits of many of the asymptotic properties of the empirical distribution function of i.i.d. observations. Consequently, usual statistics (Kolmogorov–Smirnov, Cramér–von Mises) may apply (see Anderson [1]). Moreover, some of those results still hold true for linear infinite variance (stable) processes (Klueppelber and Mikosch [17]), where the shape of the linear filter is tested. (Kokoszka and Mikosch [18]) achieved convergence in long range dependence under finite or infinite variance hypotheses, using normalized and randomly centered integrated periodogram. Thanks to this device, the limit process is completely free from the distribution of the driving i.i.d sequence. Using the functional central limit theorem proved in [20], one can derive goodness-of fit test procedure for estimated spectral measure. Unfortunately, this last result is not proved to hold true in the long memory case. For a practical point of view, it is a major drawback since one would prefer to test for a composite hypothesis rather than for the exact specification of a model with precise numerical values.

Keywords and phrases: Goodness-of-fit test for spectral density, periodogram, long range dependence.

[1] Laboratoire de Mathématiques Appliquées, FRE 2222 du CNRS, UFR de Mathématiques, bâtiment M2, Université des Sciences et Technologies de Lille, 59655 Villeneuve-d'Ascq Cedex, France; e-mail: anne.philippe@univ-lille1.fr

Our procedure is based on the information measure (also known as "logarithmic contrast")

$$S(f,g) = \log \int_0^{2\pi} \frac{f(\lambda)}{g(\lambda)} \frac{\mathrm{d}\lambda}{2\pi} - \int_0^{2\pi} \log \frac{f(\lambda)}{g(\lambda)} \frac{\mathrm{d}\lambda}{2\pi}$$

introduced by (Mokkadem [21, 22]) to compare two spectral densities f and g. By Jensen inequality, $S(f,g)$ vanishes if and only if f and g are proportional on a set of Lebesgue measure 2π. The author uses a consistent estimation of f to estimate this quantity on turn and to test models for ARMA processes. This approach was furtherly extended to a linear short memory context by (Fay [7], Annex B). We propose here an adaptation to long range dependent processes that admit a parametric representation and that include ARFIMA processes and fractional Brownian motion increments (or fractional Gaussian noise). The procedure allows for testing composite hypotheses, which is an important practical issue.

To be more specific, the issue is the following: observing $X_1, \ldots, X_n$ and assuming that the underlying process $X = (X_t)_{t \in \mathbb{Z}}$ is stationary, we want to test for the hypothesis that the spectral density of X is of the following form

$$f(\lambda) = \sigma^2 g(\lambda; d_0, \theta_0) = \sigma^2 |1 - \mathrm{e}^{i\lambda}|^{-2d_0} g^*(\lambda; \theta_0), \qquad \sigma^2 > 0 \tag{1.1}$$

where $g^*(\cdot; \theta_0)$ is an even, positive continuous function completely defined up to the knowledge of a parameter $\theta_0 \in \mathbb{R}^l$. A particular case is the test for whiteness, *i.e.* $f \equiv \sigma^2$. As it is shown in (Fay [7], Annex B), one can test for the flatness of the spectral density f by using the following estimate of $S(f, 1)$:

$$S_n(\bar{I}, 1) = \log \left(\frac{1}{K_n} \sum_{k=1}^{K_n} \bar{I}_{n,k} \right) - \frac{1}{K_n} \sum_{k=1}^{K_n} \log\left(\bar{I}_{n,k}\right) + \gamma_{m,1} \tag{1.2}$$

with $\gamma_{m,1}$ a centering constant, K_n an increasing sequence, and $(\bar{I}_{n,k})_{k=1,\ldots,K_n}$ a modified version of the classical periodogram $(I_{n,k})$ of $X_1, \ldots, X_n$ at Fourier frequencies (λ_k), *i.e.*

$$I_{n,k} = \frac{1}{2\pi n} \left| \sum_{t=1}^{n} X_t \mathrm{e}^{-it\lambda_k} \right|^2 , \ \lambda_k = \frac{2k\pi}{n} , \ k = 1, \ldots, n. \tag{1.3}$$

Those quantities will be precisely defined below. The statistic $S_n(\bar{I}, 1)$ is a non-linear functional of the periodogram (see Taniguchi [23]; Janas and von Sachs [16]; Fay *et al.* [8], for the issues raised by those objects). Its weak convergence is derived under both whiteness and short range dependent linear hypotheses.

We use the heuristic idea that the normalized periodogram $(I_{n,k}/f(\lambda_k))$ is "close" to the periodogram of an i.i.d. sequence if f is the true spectral density. Loosely speaking, normalizing the periodogram by the true spectral density is a whitening operation in the spectral domain. More precisely, it is established that, for any fixed and distinct Fourier frequencies $\lambda_{k(1)}, \ldots, \lambda_{k(N)}$, the random vector $\left(\frac{I_{n,k(1)}}{f(\lambda_{k(1)})}, \ldots, \frac{I_{n,k(N)}}{f(\lambda_{k(N)})}\right)$ converges weakly to a vector of i.i.d. exponential random variables as soon as X is stationary and short range dependent with finite variance (see Brockwell and Davis [5], Th. 10.3.2). The limit distribution is exactly the distribution of the periodogram of a Gaussian white noise. This fact translates with slight modifications to the tapered and pooled periodogram, but fails in the long memory context (see *e.g.* Künsch [19]). Still, many statistical procedures translate into this framework. Thus, to test the hypothesis H_0: $f(\lambda) \equiv \sigma^2 g(\lambda; d, \theta)$ for some σ^2, d, θ, we shall consider the following periodogram estimate of $S(f, \sigma^2 g(\cdot; d, \theta))$.

$$S_n(\bar{I}, \sigma^2 g(\cdot; d, \theta)) = \log \left(\frac{1}{K_n} \sum_{k=1}^{K_n} \frac{\bar{I}_{n,k}}{\sigma^2 g(\lambda_k; d, \theta)} \right) - \frac{1}{K_n} \sum_{k=1}^{K_n} \log \left(\frac{\bar{I}_{n,k}}{\sigma^2 g(\lambda_k; d, \theta)} \right) + \gamma_{m,p}$$

which is expected to have the same weak limit under the null hypothesis as $S_n(\bar{I},1)$ for a white noise, by the above heuristic argument. Note that $S(f,g)$ being scale invariant, we have $S_n(\bar{I},\sigma^2 g(\cdot;d,\theta)) = S_n(\bar{I},g(\cdot;d,\theta))$. Only the "shape" of the spectral density is considered, not its "scale" parameter σ^2. In practice, one tests for composite hypothesis and d and θ have to be estimated from the data. Hence we assume that $(\hat{d},\hat{\theta})$ is a parametric estimate of (d,θ) and consider

$$S_n(\bar{I},g(\cdot;\hat{d},\hat{\theta})) = \log\left(\frac{1}{K_n}\sum_{k=1}^{K_n}\frac{\bar{I}_{n,k}}{g(\lambda_k;\hat{d},\hat{\theta})}\right) - \frac{1}{K_n}\sum_{k=1}^{K_n}\log\left(\frac{\bar{I}_{n,k}}{g(\lambda_k;\hat{d},\hat{\theta})}\right) + \gamma_{m,p}. \tag{1.4}$$

Definitions

The "raw" definition (1.3) of the periodogram can be modified by tapering and/or pooling.

Tapering

Before computing the periodogram, it may be useful and/or necessary to apply a taper (or windowing, or weighting function) to the observations. This operation is said to "reduce the leakage effect in the frequency domain". In this paper, we shall use a sligthly modified version of the data-taper introduced by (Hurvich and Chen [15]) which is very convenient to handle. For a given integer order p, it is defined by $w_{n,t}^{(p)} = \tilde{w}n,t/a_n^{(p)}$ with $\tilde{w}_{n,t} = (1-\mathrm{e}^{i2\pi t/n})^p$, $t=1,\dots,n$ and $a_n^{(p)} = (n^{-1}\sum_{t=1}^{n}|\tilde{w}_{n,t}|^2)^{1/2} = \binom{2p}{p}^{1/2}$. Now-on, the integer p is referred to as the "order of tapering" and $p=0$ means that no taper is applied. An effect of the Hurvich and Chen taper is to correlate each Fourier transform with its p right-neighbors. To be more specific, define

$$d_{n,k} = (2\pi n)^{-1/2}\sum_{t=1}^{n} X_t \mathrm{e}^{i\lambda_k t} \qquad \text{and} \qquad d_{n,k}^{(m,p)} = (2\pi n)^{-1/2}\sum_{t=1}^{n} w_{n,t}^{(p)} X_t \mathrm{e}^{i\lambda_k t}, \quad k=1,\dots,n$$

the discrete Fourier transform of the non-tapered and tapered observations $X_1,\dots,X_n$. Then

$$d_{n,k}^{(m,p)} = a_p^{-1}\sum_{j=0}^{p}\binom{p}{j}(-1)^j d_{n,k+j}. \tag{1.5}$$

Tapering is a major benefit when dealing with possibly over-differentiated (non-invertible) or non-stationary series (see Hurvich and Chen [15], see also Velasco [24]). Our GOF procedure may apply in those cases but we shall restrict our attention to invertible and stationary processes, which imply in particular $0 \le d < 1/2$.

Pooling

After computing the periodogram of the possibly tapered observations, divide the frequency axis in blocks of size $m \ge 1$ and sum the periodogram ordinates on each block. As m is fixed, this does not amount to smoothing the periodogram.

For notational convenience, we make the values of m and p implicit in the notation of the (possibly) tapered and (possibly) pooled periodogram

$$\bar{I}_{n,k} = (2\pi n m)^{-1}\sum_{j=(m+p)(k-1)+1}^{(m+p)k-p}\left|\sum_{t=1}^{n} h_{n,t}^{(p)} X_t \mathrm{e}^{it\lambda_k}\right|^2, \ k=1,\dots,K_n$$

with $K_n = (m+p)\lfloor\frac{n-1}{2(m+p)}\rfloor$. Define the frequencies $x_k = (m+p)(k+1/2)\times 2\pi/n, k=1,\dots,K_n$ so as to be central to pooling-tapering segment. Note that dropping out p discrete tapered Fourier transform in the definition of the tapered periodogram over each block ensures that the ordinates $\bar{I}_{n,k}$ are still i.i.d. when X is an i.i.d. Gaussian series. However, they are no more exponentials but chi-square. Also, this procedure may yield a loss of efficiency of order $m/(m+p)$ in further estimations.

Define now both constants $\gamma_{m,p}$ and $\sigma_{m,p}$,

$$\gamma_{m,p} = \mathbb{E}_0[\log 2\pi \bar{I}_{n,k}], \quad \sigma^2_{m,p} = \mathrm{var}_0[2\pi \bar{I}_{n,k} - \log 2\pi \bar{I}_{n,k}] \tag{1.6}$$

where $\mathbb{E}_0$ and var_0 stand for the mathematical expectation and the variance under the assumption that the process X is i.i.d. standard normal. When no taper is applied and X is a unit variance i.i.d. Gaussian noise, the $2\pi\bar{I}_{n,k}$'s are distributed as independent $(2m)^{-1}\chi^2_{2m}$, and then $\gamma_{m,0} = \Psi(m)$ and $\sigma^2_{m,0} = \Psi'(m) - 1/m$ where Ψ refers to the digamma function, *i.e.* $\Psi = \Gamma'/\Gamma$ where Γ denotes the gamma function.

The layout of the rest of the paper is as follows. Section 2 makes precise the assumptions on the process X. Section 3 is devoted to the statement of the weak convergence results on the statistic (1.4) for well-specified models and on the asymptotic power of the test. Those results are proved in Section 4. Section 5 describes a numerical study which compares our test to Chen and Deo [6]'s one.

2. Assumptions

In the following, it will be assumed that

(A1) The process X has spectral density f of the form

$$f(\lambda) = \sigma^2 |1 - \mathrm{e}^{i\lambda}|^{-2d} f^*(\lambda), \qquad \sigma^2 > 0$$

where $d \in [0, 1/2)$ and f^* is twice continuously differentiable w.r.t. $\lambda \in [-\pi, \pi]$ and bounded away from zero.

Assume also that X admits the linear representation

$$X_t = \sigma \sum_{j\in\mathbb{Z}} a_j Z_{t-j}\ ,\ Z = (Z_t)_{t\in\mathbb{Z}} \text{ i.i.d.}, \mathbb{E}Z_0 = 0,\ \mathbb{E}|Z_0|^2 = 1 \tag{2.1}$$

where $(a_j)_{j\in\mathbb{N}}$ is a real square summable sequence.

Define a parametric class of spectral densities by

$$\mathcal{F}_0 = \left\{ \sigma^2 g(\cdot; d, \theta), (d,\theta) \in D \times \Theta, \sigma^2 > 0 | \int_{-\pi}^{\pi} \log g(x; d, \theta) \mathrm{d}x = 0 \right\}$$

with D a compact subset of $[0, 1/2)$ and Θ a compact subset of $\mathbb{R}^l$. The following set of assumptions controls the regularity of this parameterization.

(A2) $\forall (d,\theta) \in D \times \Theta$, $\int_{-\pi}^{\pi} \log g(\lambda; d, \theta)\mathrm{d}\lambda$ is twice differentiable in (d, θ) under the integral sign.

(A3) $g^*(\lambda;\theta)$ and ${g^*}^{-1}(\lambda;\theta)$ are continuous at all (λ, θ).

(A4) $\frac{\partial^2}{\partial\theta\partial\lambda} g^*(\lambda;\theta)$ is continuous at all (λ,θ).

Remark 1. A straightforward consequence of assumption **(A3)** is that there exist positive and continuous functions $c_1(\theta)$ and $c_2(\theta)$ such that

$$\forall (\lambda, d, \theta) \in [-\pi, \pi] \times [0, 1/2) \times \Theta,\ c_1(\theta)\lambda^{-2d} \le g(\lambda; d, \theta) \le c_2(\theta)\lambda^{-2d}. \tag{2.2}$$

As we will use singular function of the periodogram, such as $\log x$ or $1/x$, we need the Fourier transform of $X_1, \ldots, X_n$ (hence its periodogram) to have no atom at zero eventually. This is ensured by the following hypothesis:

(A5) $\int_{\mathbb{R}} |\mathbb{E}\exp(itZ_0)|^q \mathrm{d}t < \infty$ for some $q \ge 1$.

Our goodness-of-fit test is designed for testing for the null hypothesis $(\mathbf{H}_0)$. Its power may be derived for fixed alternative $(\mathbf{H}_1)$. A particular case of sequence of alternatives $(\mathbf{H}_1^{(n)})$ is investigated through numerical simulations in Section 5.

$(\mathbf{H}_0)$ X is a process admitting the linear representation (2.1) with Z satisfying **(A5)** and with spectral density $f = \sigma^2 g(\cdot; d_0, \theta_0) \in \mathcal{F}_0$ where **(A2–A4)** hold.

$(\mathbf{H}_1)$ X is a process admitting the linear representation (2.1) with Z satisfying **(A5)** and with spectral density $f(x) = \sigma_1^2 |1 - \mathrm{e}^{ix}|^{-2d_1} f^*(x)$ with $d_1 \in [0, 1/2)$ and f^* satisfying **(A1)**. Moreover, there exist a unique (d_0, θ_0) in $D \times \Theta$ such that $S(f,h) = \inf_{g \in \mathcal{F}_0} S(f,g)$ where $h := h(.; d_0, \theta_0)$ and $S(f,h) > 0$.

Remark 2. Note that the last condition is equivalent to the statement the Kullback–Liebler divergence between f and any element of $\mathcal{F}_0$, $\inf_{g \in \mathcal{F}_0} \int_{-\pi}^{\pi} \left(\frac{f(x)}{h(x)} - 1 - \log \frac{f(x)}{h(x)} \right) \frac{\mathrm{d}x}{2\pi}$, is also bounded away from zero.

3. Main results

The following theorem provides the asymptotic behavior of the statistic S_n under the null hypothesis $(\mathbf{H}_0)$.

Theorem 3.1. *Under the null hypothesis* $(\mathbf{H}_0)$, *assume that* $(\hat{d}, \hat{\theta}) = (\hat{d}_n, \hat{\theta}_n)$ *is a* $\sqrt{n}$*-convergent estimation of the parameter* (d_0, θ_0), *i.e.*

$$\|(\hat{d}, \hat{\theta}) - (d_0, \theta_0)\| = O_P(n^{-1/2}). \tag{3.1}$$

Assume either

1. $d_0 = 0$ *and* $p = 0$ *or* 1;
2. $d_0 > 0$ *and* $p = 1$.

Let $m \geq 5$, *and* $\mathbb{E}|Z_0|^{4(m+p)+1} < \infty$. *Then*

$$\sqrt{K_n} S_n(\bar{I}, g(\cdot, \hat{d}, \hat{\theta})) \longrightarrow_d \mathcal{N}(0, \sigma_{m,p}^2 + \kappa_4 \alpha_{m,p}) \tag{3.2}$$

where $\sigma_{m,p}^2$ *is defined in (1.6),* κ_4 *is the fourth cumulant of* Z_1, *and* $\alpha_{m,p}$ *is* $2(m+p)$ *times the value of the integral (4.8), which vanishes if either* $p = 0$ *or* $p = 1$ *and* $m = 1$.

Remark 3. The limit distribution of $S_n(\bar{I}, g(\cdot, \hat{d}, \hat{\theta}))$ is free from the higher order of Z, especially from its fourth-order cumulant, for $p = 0$ and any m, or if $(p, m) = (1, 1)$. Although Monte-Carlo simulations strongly suggest that the integral (4.8) still vanishes for $p = 1, m > 1$, we were not able to establish this result.

Remark 4. Note that a $\sqrt{n}$-parametric estimate (3.1) exists in both SRD and LRD contexts. Our assumptions are compatible with the assumption set (B1–B6) of (Giraitis and Surgailis [11]) so that the Whittle contrast minimizer is proved to be $\sqrt{n}$-convergent for linear long-range dependent time series under some additional regularity conditions on the parametric set $\{g^*(\cdot; \theta), \theta \in \Theta\}$.

Let us now consider the behavior of the statistic under given fixed alternatives. We prove that the power of the test procedure converges to 1 *i.e.* under the hypothesis $(\mathbf{H}_1)$ the probability $\mathbb{P}(\sqrt{K_n} S_n(\bar{I}, g(\cdot, \hat{d}, \hat{\theta}) > C)$ converges to 1 when n tends to infinity.

Theorem 3.2. *Under* $(\mathbf{H}_1)$, *suppose that the estimator sequence* $(\hat{d}, \hat{\theta}) = (\hat{d}_n, \hat{\theta}_n)$ *is such that*

$$\|(\hat{d}, \hat{\theta}) - (d_0, \theta_0)\| = o_P((\log n)^{-1}) \tag{3.3}$$

and assume either

1. $d_1 = 0$ *and* $p = 0$ *or* 1;
2. $d_1 > 0$ *and* $p = 1$.

Let $m \geq 5$, *and* $\mathbb{E}|Z_0|^{4(m+p)+1} < \infty$. *Then, for any* C,

$$\lim_{n\to\infty} \mathbb{P}(\sqrt{K_n} S_n(\bar{I}, g(\cdot, \hat{d}, \hat{\theta}) > C) = 1. \tag{3.4}$$

Examples

The theory of parametric estimation and tests under misspecified models is not classical in long range dependence (see Hosoya [13]). We give below two practical examples for which the uniqueness of the minimizer $h \in \mathcal{F}_0$ of $S(f, \cdot)$ is established. Under the uniqueness hypothesis, the consistency (3.3) may be proved along the following lines. Take for $(\hat{d}, \hat{\theta})$ the Whittle estimate, *i.e.*

$$(\hat{\sigma}^2, \hat{d}, \hat{\theta}) := \arg \min_{(\sigma,d,\theta)} \sum_{k=1}^{K_n} \frac{\bar{I}_{n,k}}{\sigma^2 g(x_k; d, \theta)} + \log(\sigma^2 g(x_k; d, \theta)) \tag{3.5}$$

where the minimum is taken over the set $\mathbb{R}^+ \times D \times \Theta$. By standard manipulations, we get

$$\begin{aligned} \hat{\sigma}^2 &= K_n^{-1} \sum_{k=1}^{K_n} \frac{\bar{I}_{n,k}}{g(x_k; \hat{d}, \hat{\theta})} \\ (\hat{d}, \hat{\theta}) &= \arg \min_{(d,\theta)\in D\times\Theta} \log\left(K_n^{-1} \sum_{k=1}^{K_n} \frac{\bar{I}_{n,k}}{g(x_k; d, \theta)}\right) - K_n^{-1} \sum_{k=1}^{K_n} \log\left(\frac{\bar{I}_{n,k}}{g(x_k; d, \theta)}\right) \\ &= \arg \min_{(d,\theta)\in D\times\Theta} S_n(\bar{I}, g(., d, \theta)). \end{aligned}$$

Denote for short $\hat{f}(\cdot) = g(\cdot; \hat{d}, \hat{\theta})$. By hypothesis $(\hat{d}, \hat{\theta})$ and (d_0, θ_0) respectively minimize the functions $(d, \theta) \mapsto S_n(\bar{I}, g(\cdot, d, \theta))$ and $(d, \theta) \mapsto S(f, g(\cdot, d, \theta))$. Then,

$$\begin{aligned} 0 &\leq S(f, \hat{f}) - S(f, h) \\ &= S(f, \hat{f}) - S_n(\bar{I}, \hat{f}) + S_n(\bar{I}, \hat{f}) - S_n(\bar{I}, h) + S_n(\bar{I}, h) - S(f, h) \\ &\leq S(f, \hat{f}) - S_n(\bar{I}, \hat{f}) + S_n(\bar{I}, h) - S(f, h) \\ &\leq 2 \sup_{(d,\theta)\in D\times\Theta} |S(f, g(\cdot, d, \theta)) - S_n(\bar{I}, g(\cdot, d, \theta))|. \end{aligned}$$

Suppose now $D \subset [0, 1/2 - \eta]$. Then

$$\sup_{(d,\theta)\in D\times\Theta} |S(f, g(\cdot, d, \theta)) - S_n(\bar{I}, g(\cdot, d, \theta))| = O_P(n^{-\eta}).$$

To see this, define

$$J_{n,k} := \frac{\bar{I}_{n,k}}{f(x_k)} \quad \text{and} \quad \beta_{n,k} := \left(\frac{f_{n,k}/g_{n,k}}{\sum_{j=1}^{K_n} f_{n,j}/g_{n,j}}\right).$$

After algebraic calculation,

$$\begin{aligned} S(f, g(\cdot, d, \theta)) - S_n(\bar{I}, g(\cdot, d, \theta)) \quad = \quad & \log\left(1 + \sum_{k=1}^{K_n} \beta_{n,k}(J_{n,k} - 1)\right) - \frac{1}{K_n} \sum_{k=1}^{K_n} \log J_{n,k} - \gamma_{m,p} \\ & + S(f, g) - \log\left(K_n^{-1} \sum_{k=1}^{K_n} \frac{f(x_k)}{g(x_k)}\right) + K_n^{-1} \sum_{k=1}^{K_n} \log \frac{f(x_k)}{g(x_k)}. \end{aligned}$$

The last line is the error in a Riemann approximation, and it may be seen that it is bounded by $Cn^{-1+2(d_1-d)_+}$. Under mild conditions, it may be shown that

$$K_n^{-1/2}\sum_{k=1}^{K_n}(\log J_{n,k}-\gamma_{m,p}) \longrightarrow_d \mathcal{N}(0,s^2)$$

for some constant s^2. Finally, it may be shown (see the proof of Th. 3.2) that $\sum_{k=1}^{K_n}\beta_{n,k}(J_{n,k}-1) = O_P(n^{-(1\wedge(2-4(d_1-d_0)))})$. Then we have $|S(f,\hat{f})-S(f,h)| = O_P(n^{-\eta})$. Recall that $h=g(\cdot,d_0,\theta_0)$ minimizes over $\mathcal{F}_0$ the mapping $g\mapsto S(f,g)$. If (d_0,θ_0) belongs to the boundary of $D\times\Theta$, it may happen (this is the case of the particular case of Ex. 1 below) that the gradient $\nabla_{d,\theta}S(f,g)$ is non-zero at (d_0,θ_0). In which case it follows that $(\hat{d},\hat{\theta})-(d_0,\theta_0)=O_P(n^{-\eta})$. In any other case (as in Ex. 2 below), $\nabla_{d,\theta}S(f,g)$ is zero at (d_0,θ_0) and we make the further assumption that the second differential is positive definite here. Then, $(\hat{d},\hat{\theta})-(d_0,\theta_0)=O_P(n^{-\eta/2})$.

Example 1. Suppose that $\mathcal{F}_0$ is the set of ARFIMA$(0,d,0)$ spectral densities such that $d\in[d_{\min},d_{\max}]$ and that X satisfies **(A1)** with $d_1\notin[d_{\min},d_{\max}]$. Then X may be an ARFIMA(p,d_1,q) or a FEXP process. By the strict convexity of $\alpha\mapsto\log\int|1-\mathrm{e}^{ix}|^{-2\alpha}f^*(x)\mathrm{d}x/2\pi-\int\log|1-\mathrm{e}^{ix}|^{-2\alpha}\mathrm{d}x/2\pi$ on $(-1/2,1/2)$, a minimizer d_0 of $S(f,g), g\in\mathcal{F}_0$ exists and is unique. A particular case is fitting an ARFIMA(0,d,0), $d\in[d_{\min},d_{\max}]$ to an ARFIMA(0,d_1,0), $d_1\notin[d_{\min},d_{\max}]$. As one could expect it, the minimizer d_0 is $d_{\min}$ if $d_1<d_{\min}$ and $d_{\max}$ if $d_1>d_{\max}$).

Example 2. Suppose that $\mathcal{F}_0$ is the following collection of AR(P) spectral densities

$$\mathcal{F}_0=\left\{\frac{\sigma^2}{2\pi}g_\theta(\cdot),g_\theta(\cdot)=\left|1-\sum_{j=1}^{P}\theta_j\mathrm{e}^{ij\cdot}\right|^{-2},(\theta_j)\in\Theta,\sigma^2>0\right\}$$

where Θ is such that $\forall(\theta_j)\in\Theta, 1-\sum_{j=1}^{P}\theta_j z^j$ has no root in the closed unit circle. Note that $\int_{-\pi}^{\pi}|1-\sum_{j=1}^{p}\theta_j\mathrm{e}^{ij\lambda}|^{-2}\mathrm{d}\lambda=0$ (see Brockwell and Davis [5], p. 191). Put $\gamma_h=\mathrm{cov}(X_0,X+h)$, $\Gamma=(\gamma_{i-j})_{i\le i,j\le n}$ and $\gamma(P)=(\gamma_1,\dots,\gamma_P)'$. The minimizer θ_0 is the unique solution of the equation $\Gamma(P)\theta=\gamma(P)$. It is shown by (Yajima [25], Th. 2.1 and Ex. 1) that the parameter estimate of θ converges in probability to θ_0 at the rate $n^{-(1/2\wedge(1-2d))}$ (for $d\ne1/4$). The hypotheses of Theorem 3.2 are then satisfied for this example. Note that the weak limit of $\hat{\theta}$ suitably normalized is also available (Gaussian if $d<1/4$, Rosenblatt if $d>1/4$; for details, see Yajima 1993).

4. Proofs

4.1. **Proof of Theorem 3.1**

To establish the weak convergence of Theorem 3.1, it suffices to prove that

i) $\sqrt{K_n}S_n(\bar{I},f)$ has the asymptotic normal distribution of (3.2);

ii) $S_n(\bar{I},g(\cdot;\hat{d},\hat{\theta}))-S_n(\bar{I},f)=o_P(n^{-1/2})$ as $n\to\infty$.

The step **ii)** is established in Lemma 4.3. Together with **i)**, it says that the statistic with estimated spectral density has the same weak limit as the statistic with the true spectral density. Let now focus on the first step, which is taken by the use of the Bartlett decomposition which relates the normalized periodogram of X to the normalized periodogram of the unobserved sequence $Z_1,\dots,Z_n$, denoted $I_{n,k}^Z$ (note that the orders of pooling and tapering used to define the periodogram of Z are the same as for $\bar{I}_{n,k}$). Define

$$R_{n,k}=J_{n,k}-\tilde{J}_{n,k} \tag{4.1}$$

with $J_{n,k} := \frac{\bar I_{n,k}}{f(x_k)}$ and $\tilde J_{n,k} := 2\pi \bar I^Z_{n,k}$. Those $R_{n,k}$'s are stochastically small in a sense to be made precise. This formal decomposition suggests to write $S_n(\bar I, f) = T_n^{(1)} + T_n^{(2)}$ with

$$T_n^{(1)} = \log\left(\frac{1}{K_n}\sum_{k=1}^{K_n} \tilde J_{n,k}\right) - \frac{1}{K_n}\sum_{k=1}^{K_n} \log\Big(\tilde J_{n,k}\Big) + \gamma_{m,p}$$
$$T_n^{(2)} = \log\left(\frac{\sum_{k=1}^{K_n} J_{n,k}}{\sum_{k=1}^{K_n} \tilde J_{n,k}}\right) - \frac{1}{K_n}\sum_{k=1}^{K_n} \log\left(\frac{J_{n,k}}{\tilde J_{n,k}}\right).$$

Lemma 4.1 shows that $\sqrt{K_n}T_n^{(1)}$ is weakly convergent and Lemma 4.2 ensures that $\sqrt{K_n}T_n^{(2)}$ is asymptotically negligible in probability.

Lemma 4.1. *For $m \geq 5$, let $Z = (Z_t)_{t\in\mathbb{Z}}$ be a sequence of i.i.d random variables satisfying* **(A5)** *and $\mathbb{E}|Z_0|^{4(m+p)+1} < \infty$. Then the following weak limit holds*

$$\sqrt{K_n}\left[\log\left(\frac{1}{K_n}\sum_{k=1}^{K_n} \tilde J_{n,k}\right) - \frac{1}{K_n}\sum_{k=1}^{K_n} \log\Big(\tilde J_{n,k}\Big) + \gamma_{m,p}\right] \longrightarrow_d \mathcal{N}(0, \sigma^2_{m,p} + \kappa_4 \alpha_{m,p}). \tag{4.2}$$

Proof. Using the relation (1.5), we can write

$$\begin{aligned}\bar I^Z_{n,k} =& m^{-1} \sum_{j=(m+p)(k-1)+1}^{(m+p)k-p} |d^{(m,p)}_{n,j}|^2 \\ =& m^{-1} a_p^{-2} \sum_{j=(m+p)(k-1)+1}^{(m+p)k-p} \sum_{l=0}^{p} \tbinom{p}{l}^2 I_{n,j+l} + 2m^{-1}a_p^{-2} \sum_{j=(m+p)(k-1)+1}^{(m+p)k-p} \sum_{0\le l<l'\le p} \tbinom{p}{l}\tbinom{p}{l'} \Re(d_{n,j+l} d^{\dagger}_{n,j+l'})\end{aligned}$$

where $z^\dagger$ stands for the conjugate of the complex variable z. From well known results holding for the mean and the covariances of the periodogram of an i.i.d. sequence (see *e.g.* Brockwell and Davis [5], Prop. 10.3.2), it follows from the last expansion that

$$\mathbb{E}(\tilde J_{n,k}) = \mathbb{E}(2\pi \bar I^Z_{n,k}) = 1\ ,\ k = 1, \ldots, K_n \tag{4.3}$$
$$\mathrm{var}(\tilde J_{n,k}) = C + O(n^{-1})\ ,\ k = 1, \ldots, K_n \tag{4.4}$$
$$\mathrm{cov}(\tilde J_{n,k}, \tilde J_{n,j}) = O(n^{-1})\ ,\ 1 \le k < j \le K_n \tag{4.5}$$

where the $O(n^{-1})$ are uniform in j and k and C is a constant (equal to 1 if $m = 1$ and $p = 0$). We easily derive $\frac{1}{K_n}\sum_k \tilde J_{n,k} = 1 + O_P(n^{-1/2})$. With probability one, this last sum is also positive, since it follows from assumption **(A5)** that all the $I_{n,k}$'s have densities as soon as $n \ge q$. Thus, by Taylor formula, we obtain

$$\log\left(\frac{1}{K_n}\sum_{k=1}^{K_n} \tilde J_{n,k}\right) = \frac{1}{K_n}\sum_{k=1}^{K_n}\big(\tilde J_{n,k} - 1\big) + O_P(n^{-1}).$$

To establish (4.2), it remains to prove that

$$\frac{1}{\sqrt{K_n}}\sum_{k=1}^{K_n}\Big[\tilde J_{n,k} - 1 - \log\Big(\tilde J_{n,k}\Big) + \gamma_{m,p}\Big] \longrightarrow_d \mathcal{N}(0, \sigma^2_{m,p}). \tag{4.6}$$

If Z is Gaussian, the $\tilde{J}_{n,k}$ are i.i.d. so that (4.6) follows from classical result on sum of i.i.d. random variables with finite second order moment. In the non-Gaussian case, convergence of (4.6) follows from Theorem 1 in Fay and Soulier [9]. For $k = 1, \dots, K_n$ and $t = 1, \dots, n$, define the vectors

$$\mathbf{U}_{t,k} = \left(\cos t\lambda_{k_1}, \sin t\lambda_{k_1}, \dots, \cos t\lambda_{k_{m+p}}, \sin t\lambda_{k_{m+p}}\right)'$$

and the random vectors $\mathbf{W}_{n,k} = (2/n)^{1/2} \sum_{t=1}^{n} \mathbf{U}_{t,k} Z_t$ which components are sine and cosine transform of $Z_1, \dots, Z_n$ at the $(m+p)$ frequencies of the block k up to a factor $\sqrt{4\pi}$. Therefore the $\mathbf{W}_{n,k}$ are asymptotically normal with mean zero and covariance matrix $\mathrm{I}_{2(m+p)}$, the $2(m+p)$ rowed unit matrix. For $\boldsymbol{x} = (x_1, \dots, x_{2(m+p)}) \in \mathbb{R}^{2(m+p)}$, define

$$\psi_{m,p}(\boldsymbol{x}) := a_p^{-2} \sum_{k=1}^{m} \left| \sum_{j=0}^{p} \binom{p}{j} (-1)^j (x_{2(k+j)-1} + i x_{2(k+j)}) \right|^2 .$$

It is easy to check that, for $k = 1, \dots, K_n$,

$$\tilde{J}_{n,k} = \tfrac{1}{2m} \psi_{m,p}(\mathbf{W}_{n,k}).$$

The function $\psi_{m,p}$ is a quadratic form and let $\mathrm{A}_{m,p}$ be the symmetric matrix such that, for all $\boldsymbol{x}$, $\psi_{m,p}(\boldsymbol{x}) = \boldsymbol{x}' \mathrm{A}_{m,p} \boldsymbol{x}$. We have, for instance, $\mathrm{A}_{m,0} = \frac{1}{2} \mathrm{I}_{2m}$ and

$$\mathrm{A}_{1,1} = \frac{1}{2} \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}; \quad \mathrm{A}_{2,1} = \frac{1}{2} \begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ -1 & 0 & 2 & 0 & -1 & 0 \\ 0 & -1 & 0 & 2 & 0 & -1 \\ 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \end{bmatrix} \quad \text{etc.} \tag{4.7}$$

We set $\Phi_{m,p}(\boldsymbol{x}) := \psi_{m,p}(\boldsymbol{x})/m - 1 - \log(\psi_{m,p}(\boldsymbol{x})/m) + \gamma_{m,p}$. Then, the triangular array of (4.6) writes $K_n^{-1/2} \sum_{k=1}^{K_n} \Phi_{m,p}(\mathbf{W}_{n,k})$. Take $\boldsymbol{\xi} = (\xi_1, \dots, \xi_{2(m+p)})$ a standard Gaussian random vector. According to the definition of $\gamma_{m,p}$, we obtain

$$\mathbb{E}_0(\Phi_{m,p}(\mathbf{W}_{n,k})) = \mathbb{E}(\Phi_{m,p}(\boldsymbol{\xi})) = 0.$$

Lemma 4.1 follows by application of Theorem 1 in Fay and Soulier [9]: the array $K_n^{-1/2} \sum \Phi_{m,p}(\mathbf{W}_{nk})$ is asymptotically normal with mean zero and limit variance

$$\mathbb{E}\Phi_{m,p}^2(\boldsymbol{\xi}) + 2(m+p)\kappa_4 \left(\sum_{l=1}^{2(m+p)} \mathbb{E}[(\xi_l^2 - 1)\Phi_{m,p}(\boldsymbol{\xi})] \right)^2 = \sigma_{m,p}^2 + 2(m+p)\kappa_4 \mathbb{E}(\|\boldsymbol{\xi}\|^2 \Phi_{m,p}(\boldsymbol{\xi})) = \sigma_{m,p}^2 + \kappa_4 \alpha_{m,p}.$$

The moment writes

$$(2\pi)^{-(m+p)} \int_{\boldsymbol{x} \in \mathbb{R}^{2m+2p}} \|\boldsymbol{x}\|^2 \left[\frac{\boldsymbol{x}' \mathrm{A}_{m,p} \boldsymbol{x}}{m} - 1 - \log\left(\frac{\boldsymbol{x}' \mathrm{A}_{m,p} \boldsymbol{x}}{m} \right) + \gamma_{m,p} \right] \mathrm{e}^{-\|\boldsymbol{x}\|^2/2} \mathrm{d}\boldsymbol{x}. \tag{4.8}$$

□

Lemma 4.2. *Let X be a process satisfying (2.1)-***(A5)*** with $\mathbb{E}|Z_0|^{4(m+p)+1} < \infty$. Assume that the spectral density of X is of the form (1.1), with f^* twice continuously differentiable. Then, if either $d = 0$ or $d > 0$ and $p = 1$, it holds that $\sqrt{K_n} T_n^{(2)} = o_P(1)$.*

Proof. Under our assumptions, Lemmas 2 and 11 of Hurvich *et al.* [14] are valid and respectively yield

$$K_n^{-1/2} \sum_{k=1}^{K_n} \log\left(\frac{J_{n,k}}{\tilde{J}_{n,k}}\right) = o_P(1), \tag{4.9}$$

$$\mathbb{E}\left(J_{n,k} - \tilde{J}_{n,k}\right)^2 \leq Ck^{-2}. \tag{4.10}$$

The constant C is uniform with respect to f satisfying our hypotheses. Given (4.9), it remains to show that

$$\log\left(\frac{\sum_{k=1}^{K_n} J_{n,k}}{\sum_{k=1}^{K_n} \tilde{J}_{n,k}}\right) = \log\left(1 + \frac{\sum_{k=1}^{K_n}(J_{n,k} - \tilde{J}_{n,k})}{\sum_{k=1}^{K_n} \tilde{J}_{n,k}}\right) = o_P(n^{-1/2}).$$

By hypothesis **(A5)**, $\sum_{k=1}^{K_n} \tilde{J}_{n,k} > 0$ a.s. and $\sum_{k=1}^{K_n} J_{n,k} > 0$ a.s. Also, it is easily shown that $(\sum_{k=1}^{K_n} \tilde{J}_{n,k})^{-1} = O_P(n^{-1})$. Now, using (4.10), we get

$$\mathbb{E}\left|\sum_{k=1}^{K_n}(J_{n,k} - \tilde{J}_{n,k})\right| \leq C\sum_{k=1}^{K_n} k^{-1} = O(\log n)$$

so that $T_n^{(2)} = O_P(n^{-1}\log n) = o_P(n^{-1/2})$ which concludes the proof. □

Lemma 4.3. *Let f be a spectral density such that $f = \sigma^2 g(\cdot; d_0, \theta_0) \in \mathcal{F}_0$, $(d_0, \theta_0) \in D \times \Theta$ and suppose that $\|(\hat{d}_n, \hat{\theta}_n) - (d_0, \theta_0)\| = O_P(n^{-1/2})$. Then*

$$S_n(\bar{I}, g(\cdot, \hat{d}, \hat{\theta})) - S_n(\bar{I}, f(\cdot)) = o_P(n^{-1/2}).$$

Proof. Denote for short $f_{n,k} = f(x_k)/\sigma^2 = g(x_k; d_0, \theta_0)$, $\hat{f}_{n,k} = g(x_k, \hat{d}, \hat{\theta})$ for $k = 1, \ldots, K_n$. Write

$$\begin{aligned}
&S_n(\bar{I}, g(\cdot, \hat{d}, \hat{\theta})) - S_n(\bar{I}, \sigma^2 g(\cdot, d_0, \theta_0)) \\
&\quad = \log\left(\frac{1}{K_n}\sum_{k=1}^{K_n}\frac{\bar{I}_{n,k}}{\hat{f}_{n,k}}\right) - \frac{1}{K_n}\sum_{k=1}^{K_n}\log\left(\frac{\bar{I}_{n,k}}{\hat{f}_{n,k}}\right) - \log\left(\frac{1}{K_n}\sum_{k=1}^{K_n}\frac{\bar{I}_{n,k}}{f_{n,k}}\right) + \frac{1}{K_n}\sum_{k=1}^{K_n}\log\left(\frac{\bar{I}_{n,k}}{f_{n,k}}\right) \\
&\quad = \log\left(1 + \frac{\sum_{k=1}^{K_n}\frac{\bar{I}_{n,k}}{\hat{f}_{n,k}} - \frac{\bar{I}_{n,k}}{f_{n,k}}}{\sum_{k=1}^{K_n}\frac{\bar{I}_{n,k}}{f_{n,k}}}\right) - \frac{1}{K_n}\sum_{k=1}^{K_n}\log\left(1 + \left(\frac{f_{n,k}}{\hat{f}_{n,k}} - 1\right)\right).
\end{aligned}$$

A first order Taylor expansion of both terms in the last equation yields $S_n(\bar{I}, \hat{f}) - S_n(\bar{I}, f) = A_n + B_n$ with

$$\begin{aligned}
A_n &:= \frac{\sum_{k=1}^{K_n} J_{n,k}\left(\frac{f_{n,k}}{\hat{f}_{n,k}} - 1\right)}{\sum_{k=1}^{K_n} J_{n,k}} - \frac{1}{K_n}\sum_{k=1}^{K_n}\left(\frac{f_{n,k}}{\hat{f}_{n,k}} - 1\right) \\
B_n &:= -\frac{1}{2}\left(\frac{\sum_{k=1}^{K_n} J_{n,k}(\frac{f_{n,k}}{\hat{f}_{n,k}} - 1)}{\sum_{k=1}^{K_n} J_{n,k}}\right)^2\left(1 + r_n\right)^{-2} - \frac{1}{K_n}\sum_{k=1}^{K_n}\left(\frac{f_{n,k}}{\hat{f}_{n,k}} - 1\right)^2 (1 + u_{n,k})^{-2}
\end{aligned}$$

where r_n lies between 0 and $\left(\frac{\sum_{k=1}^{K_n} J_{n,k}(\frac{f_{n,k}}{\hat{f}_{n,k}}-1)}{\sum_{k=1}^{K_n} J_{n,k}}\right)$ and the $u_{n,k}$'s between 0 and $\frac{f_{n,k}}{\hat{f}_{n,k}} - 1$, $k = 1, \cdots, K_n$. Note that $r_n > -1$ a.s. and $\min_k u_{n,k} > -1$ a.s. Prove now that both A_n and B_n are $o_P(v_n)$. Now

$$\begin{aligned} A_n &= \left(\sum_{k=1}^{K_n} J_{n,k}\right)^{-1} \left(\sum_{k=1}^{K_n} J_{n,k}\left(\frac{f_{n,k}}{\hat{f}_{n,k}} - 1\right) - \frac{1}{K_n}\sum_{k=1}^{K_n} J_{n,k}\sum_{j=1}^{K_n}\left(\frac{f_{n,j}}{\hat{f}_{n,j}} - 1\right)\right) \\ &= \left(\sum_{k=1}^{K_n} J_{n,k}\right)^{-1} \sum_{k=1}^{K_n} J_{n,k}\left(\frac{f_{n,k}}{\hat{f}_{n,k}} - \frac{1}{K_n}\sum_{j=1}^{K_n}\frac{f_{n,j}}{\hat{f}_{n,j}}\right). \end{aligned}$$

Defining

$$W_{n,k} = \frac{f_{n,k}}{\hat{f}_{n,k}} - \frac{1}{K_n}\sum_{j=1}^{K_n}\frac{f_{n,j}}{\hat{f}_{n,j}} \tag{4.11}$$

we have $\sum_k W_{n,k} = 0$ and

$$A_n = \left(\sum_{k=1}^{K_n} J_{n,k}\right)^{-1} \sum_{k=1}^{K_n} J_{n,k} W_{n,k} = \left(\sum_{k=1}^{K_n} J_{n,k}\right)^{-1} \left(\sum_{k=1}^{K_n} (J_{n,k} - 1) W_{n,k}\right). \tag{4.12}$$

Put $T_{n,k} := \sum_{j=1}^{k} \alpha_{n,j}(J_{n,j} - 1)$ for $k \in \{1, \ldots, K_n\}$. Summing by parts in (4.12) yields

$$A_n = (T_{n,K_n} + K_n)^{-1} \left(\sum_{k=1}^{K_n - 1} T_{n,k}(W_{n,k} - W_{n,k+1}) + W_{n,K_n} T_{n,K_n}\right). \tag{4.13}$$

The following lemmas are proved at the end of the section:

Lemma 4.4. *For some constant C,*

$$\forall k \in \{1, \ldots, K_n\},\ \mathbb{E}|T_{n,k}| \le C\sqrt{k}. \tag{4.14}$$

Lemma 4.5. *Let $(d, d', \theta, \theta') \in D^2 \times \Theta^2$. Then there exists some positive constant C such that*

$$\left|\frac{g(x; d, \theta)}{g(x; d', \theta')} - 1\right| \le C|1 - \mathrm{e}^{ix_k}|^{-2(d-d')|}|(\|\theta - \theta'\| + \log n|d - d'|). \tag{4.15}$$

Lemma 4.6. *There exists a constant C such that the quantities $W_{n,k}$ defined in (4.11) satisfy*

$$|W_{n,K_n}| \le C(|d_0 - \hat{d}| \log n + \|\theta_0 - \hat{\theta}\|) \tag{4.16}$$

and for all $k = 1, \ldots, K_n - 1$

$$|W_{n,k} - W_{n,k+1}| \le C\big(k^{-1}|d_0 - \hat{d}| + n^{-1}\|\theta_0 - \hat{\theta}\|\big) n^{2|\hat{d} - d|}. \tag{4.17}$$

From (4.13) and (4.17), we may bound $|A_N|$ by

$$C|T_{n,K_n}+K_n|^{-1}\left[n^{2|\hat{d}-d_0|}\big(|d_0-\hat{d}|+\|\theta_0-\hat{\theta}\|\big)\sum_{k=1}^{K_n-1}|T_k|k^{-1}+|W_{n,K_n}||T_{n,K_n}|\right].$$

Using (4.14), we get

$$|T_{n,K_n}+K_n|^{-1}=O_P(n^{-1}). \tag{4.18}$$

Also,

$$\mathbb{E}\left[\sum_{k=1}^{K_n-1}|T_k|k^{-1}\right]\le Cn^{1/2}. \tag{4.19}$$

Applying (4.16), we obtain

$$|T_{n,K_n}W_{n,K_n}|\le C(|\hat{d}-d_0|\log n+\|\theta_0-\hat{\theta}\|)n^{1/2}. \tag{4.20}$$

Moreover, it is easy to see that $n^{2|d_0-\hat{d}|}=O_P(1)$, then

$$A_n\le\big(|d_0-\hat{d}|+\|\theta_0-\hat{\theta}\|\big)O_P(n^{-1/2}\log n)=O_P(n^{-1}\log n)=o_P(n^{-1/2}). \tag{4.21}$$

Let now consider B_n. By simple algebra and using the same summation by part as above,

$$\frac{\sum_{k=1}^{K_n}J_{n,k}\left(\frac{f_{n,k}}{\hat{f}_{n,k}}-1\right)}{\sum_{k=1}^{K_n}J_{n,k}}=A_n+(T_{n,K_n}+K_n)^{-1}\sum_{k=1}^{K_n}\left(\frac{f_{n,k}}{\hat{f}_{n,k}}-1\right). \tag{4.22}$$

Using Lemma 4.5,

$$\sum_{k=1}^{K_n}\left(\frac{f_{n,k}}{\hat{f}_{n,k}}-1\right)\le(\|\hat{\theta}-\theta_0\|+\log n|\hat{d}-d|)\sum_{k=1}^{K_n}|1-\mathrm{e}^{ix_k}|^{-2(d_0-\hat{d})}. \tag{4.23}$$

Note now that

$$\begin{aligned}\sum_{k=1}^{K_n}|1-\mathrm{e}^{ix_k}|^{-2(d_0-\hat{d})}&\le\sum_{k=1}^{K_n}|x_k|^{-2(d_0-\hat{d})}\\&\le(2\pi(m+p))^{2(d_0-\hat{d})}\sum_{k=1}^{K_n}\left(\frac{k}{n}\right)^{-2(d_0-\hat{d})}\le(2\pi(m+p))^{2(d_0-\hat{d})}n^{2|d_0-\hat{d}|+1}=O_P(n).\end{aligned}$$

It yields

$$|T_{n,K_n}+K_n|^{-1}\sum_{k=1}^{K_n}\left(\frac{f_{n,k}}{\hat{f}_{n,k}}-1\right)=O_P(n^{-1/2}\log n). \tag{4.24}$$

By (4.22, 4.24) and (4.21),

$$\left[\frac{\sum_{k=1}^{K_n} J_{n,k}\left(\frac{f_{n,k}-\hat{f}_{n,k}}{\hat{f}_{n,k}}\right)}{\sum_{k=1}^{K_n} J_{n,k}}\right]^2 = O_P(n^{-1}\log^2 n) = o_P(n^{-1/2}).$$

This result also implies that $(1+r_n)^{-2} = O_P(1)$. Consider now the sum $\frac{1}{K_n}\sum_{k=1}^{K_n}\left(\frac{f_{n,k}}{\hat{f}_{n,k}}-1\right)^2(1+u_{n,k})^{-2}$. Notice that for $k=1,\dots,K_n$,

$$(1+u_{n,k})^{-2} \le (\hat{f}_{n,k}/f_{n,k})^2 \vee 1 \le C x_k^{-4|d-\hat{d}|}$$

for sufficiently large n, as a consequence of (2.2). Using (4.23), it follows that

$$\begin{aligned}\frac{1}{K_n}\sum_{k=1}^{K_n}\left(\frac{f_{n,k}}{\hat{f}_{n,k}}-1\right)^2(1+u_{n,k})^{-2} &\le \frac{1}{K_n}C(\log n|d_0-\hat{d}|+\|\theta_0-\hat{\theta}\|)^2\sum_{k=1}^{K_n} x_k^{-8|d_0-\hat{d}|}\\ &= O_P(n^{-1}\log^2 n) = o_P(n^{-1/2}).\end{aligned}$$

This concludes the proof of Lemma 4.3. □

Proof of Lemma 4.4. Using the Bartlett decomposition,

$$\mathbb{E}|T_{n,k}| \le \mathbb{E}\left|\sum_{j=1}^{k}(\tilde{J}_{n,k}-1)\right| + \mathbb{E}\left|\sum_{j=1}^{k}(J_{n,k}-\tilde{J}_{n,k})\right|. \tag{4.25}$$

From (4.3, 4.4) and (4.5), the first expectation is bounded by $C\sqrt{k}$. The second expectation is bounded by $\log k$ as a consequence of the inequality (4.10).

Proof of Lemma 4.5. As $(g,g') \in \mathcal{F}_0^2$,

$$\left|\frac{g(x_k;d,\theta)}{g(x_k;d',\theta')}-1\right| \le |1-\mathrm{e}^{ix_k}|^{-2(d-d')}\left(\frac{g^*(x_k;\theta)}{g^*(x_k;\theta')}-1\right) - \left(1-|1-\mathrm{e}^{ix_k}|^{-2(d-d')}\right).$$

Under hypothesis **(A3)**, $|g^*(x_k;\theta)/g^*(x_k;\theta')|$ is bounded. Moreover, using **(A4)** $|(g^*(\cdot;\theta)/g^*(\cdot;\theta'))'| \le C\|\theta'-\theta\|$. Therefore, we get

$$\left|\frac{g^*(x_k;\theta)}{g^*(x_k;\theta')}-1\right| \le C\|\hat{\theta}-\theta_0\|.$$

Using the inequality $|1-\mathrm{e}^x| \le |x|\mathrm{e}^{|x|}$, we get

$$\left|1-|1-\mathrm{e}^{ix_k}|^{-2(d-d')}\right| \le 2|d-d'|(\log|1-\mathrm{e}^{ix_k}|)|1-\mathrm{e}^{ix_k}|^{-2(d-d')} \le C|1-\mathrm{e}^{ix_k}|^{-2(d-d')}\log n$$

and (4.15) follows.

Proof of Lemma 4.6. Write, for $k \in \{1,\dots,K_n-1\}$,

$$|W_{n,k}-W_{n,k+1}| = \left|\frac{f_{n,k}}{\hat{f}_{n,k}}-\frac{f_{n,k+1}}{\hat{f}_{n,k+1}}\right| \le \frac{2\pi}{n}\left|\sup_{\lambda\in[x_k,x_{k+1}]}\left(\frac{f}{\hat{f}}\right)'(\lambda)\right|.$$

Note that

$$\left(\frac{f}{\hat{f}}\right)'(\lambda) = \left(|1-\mathrm{e}^{i\lambda}|^{-2(\hat{d}-d_0)}\frac{f^*}{f^*}(\lambda)\right)'$$
$$= -2(\hat{d}-d_0)|1-\mathrm{e}^{i\lambda}|^{-2(\hat{d}-d_0)-1}\frac{f^*}{\hat{f}^*}(\lambda) + |1-\mathrm{e}^{i\lambda}|^{-2(\hat{d}-d_0)}\left(\frac{f^*}{\hat{f}^*}\right)'(\lambda).$$

As noticed in the proof of Lemma 4.6, under hypotheses **(A3)** and **(A4)**, we have $|f^*/\hat{f}^*| \leq C$ and $|(f^*/\hat{f}^*)'| \leq C\|\hat{\theta}-\theta_0\|$ for sufficiently large n. It follows that

$$\left|\sup_{\lambda\in[x_k,x_{k+1}]}\left(\frac{f}{\hat{f}}\right)'(\lambda)\right| \leq C|\hat{d}-d_0|(k/n)^{-2(\hat{d}-d_0)-1} + C\|\hat{\theta}-\theta_0\|(k/n)^{-2(\hat{d}-d_0)} \tag{4.26}$$

and

$$\begin{aligned}|W_{n,k}-W_{n,k+1}| &\leq Ck^{-1}|\hat{d}-d_0|(k/n)^{-2(\hat{d}-d_0)} + Cn^{-1}\|\hat{\theta}-\theta_0\|(k/n)^{-2(\hat{d}-d_0)}\\ &\leq C\big(k^{-1}|\hat{d}-d_0| + n^{-1}\|\hat{\theta}-\theta_0\|\big)n^{2|\hat{d}-d_0|}.\end{aligned}$$

Equation (4.16) is a straightforward consequence of Lemma 4.5.

4.2. Proof of Theorem 3.2

Proof. Under the hypothesis $(\mathbf{H}_1)$ we have $S(f,h) > 0$ thus for each constant C, it exists $n_0 \in \mathbb{N}$ and $\epsilon > 0$ such that $S(f,h) - C/\sqrt{K_n} > \epsilon > 0$ for $n \geq n_0$. Then

$$\begin{aligned}\forall n \geq n_0, \qquad \mathbb{P}(\sqrt{K_n}S_n(\bar{I}, g(\cdot,\hat{d},\hat{\theta}) > C) &\geq \mathbb{P}(|S_n(\bar{I}, g(\cdot,\hat{d},\hat{\theta}) - S(f,h)| < S(f,h) - C/\sqrt{K_n})\\ &= 1 - \mathbb{P}(|S_n(\bar{I}, g(\cdot,\hat{d},\hat{\theta}) - S(f,h)| > S(f,h) - C/\sqrt{K_n})\\ &\geq 1 - \mathbb{P}(|S_n(\bar{I}, g(\cdot,\hat{d},\hat{\theta}) - S(f,h)| > \epsilon).\end{aligned}$$

Therefore, it suffices to prove that

$$|S_n(\bar{I}, g(\cdot,\hat{d},\hat{\theta}) - S(f,g)| = o_P(1) \tag{4.27}$$

to obtain the convergence of the power to 1.

Denote for short $f_{n,k} = f(x_k)$, $h_{n,k} = h(x_k) = g(x_k; d_0, \theta_0)$ and $\hat{g}_{n,k} = g(x_k; \hat{d}, \hat{\theta})$. Write

$$S_n(\bar{I}, g(\cdot,\hat{d},\hat{\theta})) = \log\left(\frac{1}{K_n}\sum_{k=1}^{K_n}\frac{\bar{I}_{n,k}}{h_{n,k}}\right) - \frac{1}{K_n}\sum_{k=1}^{K_n}\log\left(\frac{\bar{I}_{n,k}}{h_{n,k}}\right) + \gamma_{m,p} \tag{4.28}$$
$$+S_n(\bar{I}, g(\cdot,\hat{d},\hat{\theta})) - S_n(\bar{I}, h) =: S_n^{(1)} + S_n^{(2)}. \tag{4.29}$$

Behavior of $S_n^{(2)}$: we prove hereafter that $S_n^{(2)} = o_P(1)$. As in the proof of Lemma 4.3, we have

$$S_n^{(2)} = \log\left(1 + \frac{\sum_{k=1}^{K_n}\frac{\bar{I}_{n,k}}{\hat{g}_{n,k}} - \frac{\bar{I}_{n,k}}{h_{n,k}}}{\sum_{k=1}^{K_n}\frac{\bar{I}_{n,k}}{h_{n,k}}}\right) - \frac{1}{K_n}\sum_{k=1}^{K_n}\log\left(1 + \frac{h_{n,k}-\hat{g}_{n,k}}{\hat{g}_{n,k}}\right)$$

and, by Lemma 4.5,

$$\forall k \in \{1,\dots,K_n\}, \left|\frac{h_{n,k}}{\hat{g}_{n,k}} - 1\right| \le C x_k^{-2|d_0-\hat{d}|}\|\theta_0 - \hat{\theta}\| + C\log n|d_0 - \hat{d}|x_k^{-2|d_0-\hat{d}|}.$$

Moreover, we have $x_k^{-2|d_0-\hat{d}|} \le n^{2|d_0-\hat{d}|} = O_p(1)$ uniformly in k, thus we get (by Eq. (3.3))

$$\left|\frac{h_{n,k}}{\hat{g}_{n,k}} - 1\right| \le C\|\theta_0 - \hat{\theta}\| + C\log n|d_0 - \hat{d}| = o_P(1) \text{ uniformly in } k$$

which implies that $\frac{1}{K_n}\sum_{k=1}^{K_n}\log\Big(1 + \frac{h_{n,k}-\hat{g}_{n,k}}{\hat{g}_{n,k}}\Big) = o_P(1)$. Similarly

$$\left|\frac{\sum_{k=1}^{K_n}\frac{\bar{I}_{n,k}}{\hat{g}_{n,k}} - \frac{\bar{I}_{n,k}}{h_{n,k}}}{\sum_{k=1}^{K_n}\frac{\bar{I}_{n,k}}{h_{n,k}}}\right| \le \frac{\sum_{k=1}^{K_n}\frac{\bar{I}_{n,k}}{h_{n,k}}\left|\frac{h_{n,k}}{\hat{g}_{n,k}} - 1\right|}{\sum_{k=1}^{K_n}\frac{\bar{I}_{n,k}}{h_{n,k}}} \le C\|\theta_0 - \hat{\theta}\| + C\log n|d_0 - \hat{d}| = o_P(1).$$

Behavior of $S_n^{(1)}$: prove now that $S_n^{(1)} - S(f,h) = o_P(1)$. After some algebra, we have

$$S_n^{(1)} - S(f,h) = \log\left(1 + \sum_{k=1}^{K_n}\beta_{n,k}(J_{n,k} - 1)\right) - \frac{1}{K_n}\sum_{k=1}^{K_n}\log(J_{n,k}) + \gamma_{m,p} \tag{4.30}$$

$$- \log\left(\frac{1}{K_n}\sum_{k=1}^{K_n}\frac{f_{n,k}}{h_{n,k}}\right) + \frac{1}{K_n}\sum_{k=1}^{K_n}\log\frac{f_{n,k}}{h_{n,k}} - S(f,h) \tag{4.31}$$

with

$$\beta_{n,k} := \left(\frac{f_{n,k}/h_{n,k}}{\sum_{j=1}^{K_n} f_{n,j}/h_{n,j}}\right).$$

It may be noted here that

$$\sum_{k=1}^{K_n}\beta_{n,k}^2 \le C\begin{cases} n^{-1} & \text{if } d_1 - d_0 < 1/4; \\ n^{-1}\log n & \text{if } d_1 - d_0 = 1/4; \\ n^{-2+4(d_1-d_0)} & \text{if } d_1 - d_0 > 1/4. \end{cases} \tag{4.32}$$

The second line of (4.30) is the difference between $S(f,h)$ and its estimation using approximation of integrals by Riemann sums. By properties of f and h, it goes to zero at the rate than $n^{-1+2(d_1-d_0)_+} = o(1)$. Using a Bartlett decomposition technique as in the proof of Theorem 3.1,

$$K_n^{-1/2}\sum_{k=1}^{K_n}\log(J_{n,k}) + \gamma_{m,p}$$

converges weakly to a mean zero normal random variable. It follows that

$$K_n^{-1}\sum_{k=1}^{K_n}\log(J_{n,k}) + \gamma_{m,p} = O_P(n^{-1/2}) = o_P(1).$$

It remains to show that the same conclusion holds for $\log(1+\sum_{k=1}^{K_n}\beta_{n,k}(J_{n,k}-1))$. We could proceed to obtain the weak convergence of this quantity, but only in the domain of application of Theorem 1 of Fay and Soulier [9] ($d_1 - d_0 < 1/4$ here). As we are only concerned here with convergence in probability to zero, we shall prove that $\sum_{k=1}^{K_n}\beta_{n,k}(J_{n,k}-1) = o_P(1)$ and greater that -1 a.s. which follows from Assumption **(A5)** as already noticed. Write now

$$\mathbb{E}\left(\sum_{k=1}^{K_n}\beta_{n,k}(J_{n,k}-1)\right)^2 = \sum_{k=1}^{K_n}\beta_{n,k}^2\mathbb{E}(J_{n,k}-1)^2 + \sum_{k\neq j}\beta_{n,k}\beta_{n,j}\mathbb{E}[(J_{n,k}-1)(J_{n,j}-1)].$$

One can bound the diagonal sum by

$$\sum_{k=1}^{K_n}\beta_{n,k}^2\mathbb{E}(J_{n,k}-1)^2 \le 2\sum_{k=1}^{K_n}\beta_{n,k}^2\mathbb{E}(\tilde{J}_{n,k}-1)^2 + 2\sum_{k=1}^{K_n}\beta_{n,k}^2\mathbb{E}(J_{n,k}-\tilde{J}_{n,k})^2.$$

The first sum is bounded by $2(1+O(1/N))\sum\beta_{n,k}^2 = o(1)$ using (4.3, 4.4) and (4.32). The second is bounded by $C\sum k^{-2}\beta_{n,k}^2 = o(1)$ using (4.10) and (4.32). The non-diagonal sum is equal to

$$\sum_{k\neq j}\beta_{n,k}\beta_{n,j}\mathbb{E}[(\tilde{J}_{n,k}-1)(\tilde{J}_{n,j}-1)] + \sum_{k\neq j}\beta_{n,k}\beta_{n,j}\mathbb{E}[(J_{n,k}-\tilde{J}_{n,k})(J_{n,j}-\tilde{J}_{n,j})]$$
$$+ 2\sum_{k\neq j}\beta_{n,k}\beta_{n,j}\mathbb{E}[(J_{n,k}-\tilde{J}_{n,k})(\tilde{J}_{n,j}-1)].$$

The first sum is bounded by $Cn^{-1}\sum_{k\neq j}\beta_{n,k}\beta_{n,j} \le Cn^{-1}(\sum_{k=1}^{K_n}\beta_{n,k})^2 = o(1)$ using (4.4). The last one is bounded by $C\sum_{k\neq j}\beta_{n,k}\beta_{n,j}\mathbb{E}^{1/2}(J_{n,k}-\tilde{J}_{n,k})^2\mathbb{E}^{1/2}(\tilde{J}_{n,j}-1)^2 \le C\sum_{k\neq j}\beta_{n,k}\beta_{n,j}k^{-1} \le C\sum_{k=1}^{K_n}k^{-1}\beta_{n,k} = o(1)$ by (4.5, 4.10, 4.32) and Cauchy–Schwartz inequality. The middle sum may be treated alike. It concludes the proof that $S_n^{(1)} - S(f,h) = o_P(1)$. □

5. Simulations

In this section, we illustrate the performance of the test procedure based on the statistic S_n (called S_n-test) from a variety of simulated processes. The choice of the parameters associated to the statistics S_n are $m = 5$ and $p = 0$ (resp. $p = 1$) for SRD (resp. LRD) processes.

We provide a comparison with the GOF procedure introduced in Chen and Deo [6] (called hereafter Chen and Deo's test). We restrict our comparison to their statistic associated to the Turkey kernel and the bandwidth $p_n = \lfloor 3n^{0.2}\rfloor$ (see Chen and Deo [6], for details). For each examples, we give the size-power curves based on 1000 independent replications. The curve is obtained as follows: for each replication, the test statistic is calculated and corresponding P-value is obtained. The size-power curve is the empirical distribution function of the P-values. In fact, this curve represents the power against nominal test size. Therefore, the optimal curve under the null hypothesis is the 45° line. Under the alternative hypotheses the optimal curve is the horizontal line with an intercept coefficient equal to 1.

First, we evaluate the sizes of both tests under different null hypotheses. We consider the three following null hypotheses:

(a) (X_n) is a white noise;
(b) (X_n) is an ARMA(P,Q) process (SRD);
(c) (X_n) is a FARIMA$(0,d,0)$ process (LRD).

The FARIMA$(0,d,0)$ processes are simulated using the circulant matrix embedding method (see Bardet *et al.* [2], for a review on the simulation of such processes).

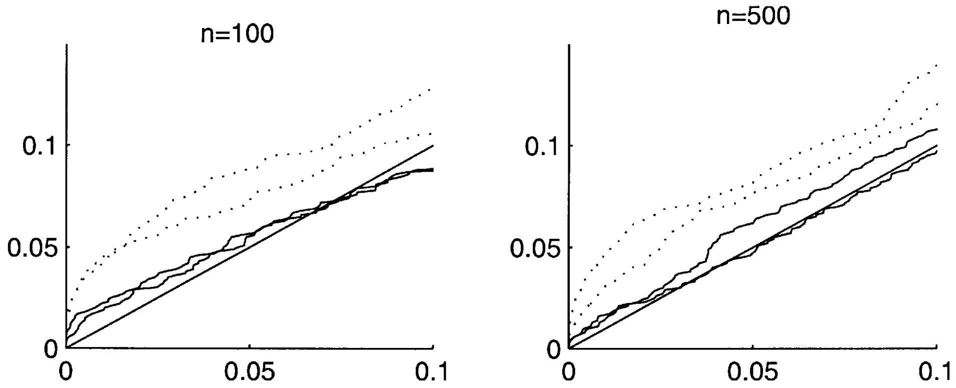

FIGURE 1. Size-power curve under white noise models for S_n-test (plain) and Chen and Deo's test (dots). Model: uniform and Gaussian white noise.

Figures 1, 2 and 3 give the size-power curves for these three null hypotheses. Whittle's estimator is used to estimate the different parameters of each models. The curves associated to the S_n-test are very close to the 45° line which is the optimal curve. For this class of examples, we can evaluate the improvement brought by the S_n-test upon Chen and Deo test. In particular, the size-power curve of the Chen and Deo test is above the 45° line, therefore the rejection rate under the null hypotheses is greater than the nominal level.

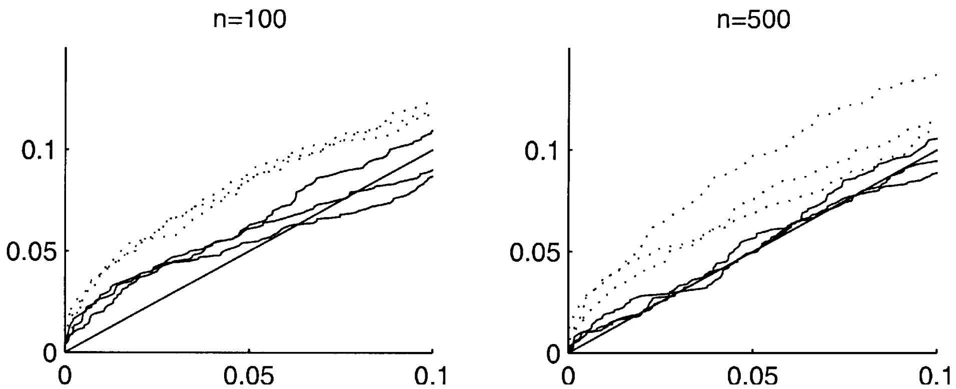

FIGURE 2. Size-power curve under ARMA processes for S_n-test (plain) and Chen and Deo's test (dots). Models: AR(1) $X_n{+}.5X_{n-1} = \epsilon_n$, MA(1) $X_n = .3\epsilon_{n-1}{+}\epsilon_n$, ARMA(1,1) $X_n{+}.5X_{n-1} = .3\epsilon_{n-1} + \epsilon_n$.

To evaluate the power of the test, we consider first a fixed alternative that satisfies the assumptions of Corollary 3.2.

(d) Generate non Gaussian processes of the form $(X_n + U_n)$ where (X_n) and (U_n) are independent processes, (X_n) is a FARIMA$(0, d, 0)$ process and (U_n) is an exponential white noise. We fit the FARIMA$(0, \hat{d}, 0)$ process where $\hat{d}$ denotes the Whittle estimate. Note that the process $(X_n + U_n)$ is outside the class of the FARIMA$(0, d, 0)$model.

Figure 4 shows that when we fit a FARIMA$(0, \hat{d}, 0)$ model on non Gaussian processes of the form $X_n + U_n$, the power of the S_n-test goes to 1 when the sample size n goes to infinity. For this model the power of Chen and Deo test is very poor. The size-power curves keep close to the 45° line.

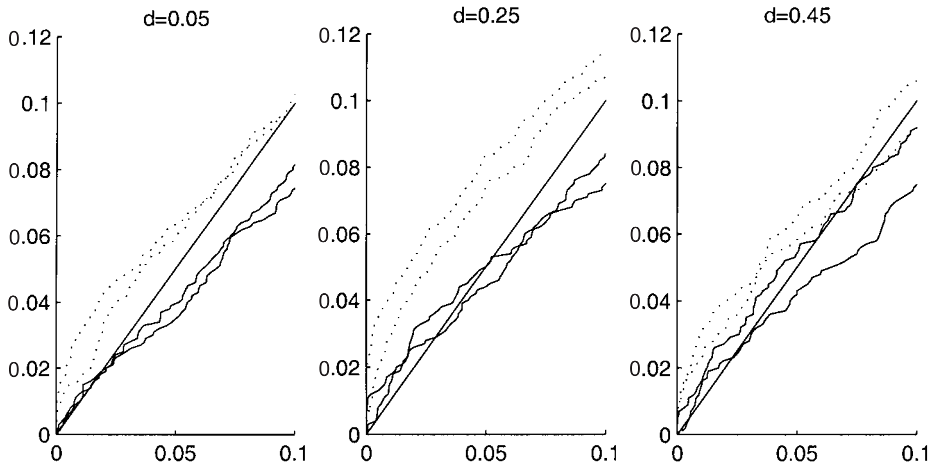

FIGURE 3. Size-power curve under FARIMA$(0, d, 0)$ model for S_n-test (plain) and Chen and Deo's test (dots). Model: $d = 0.05,\ 0.25,\ 0.45$.

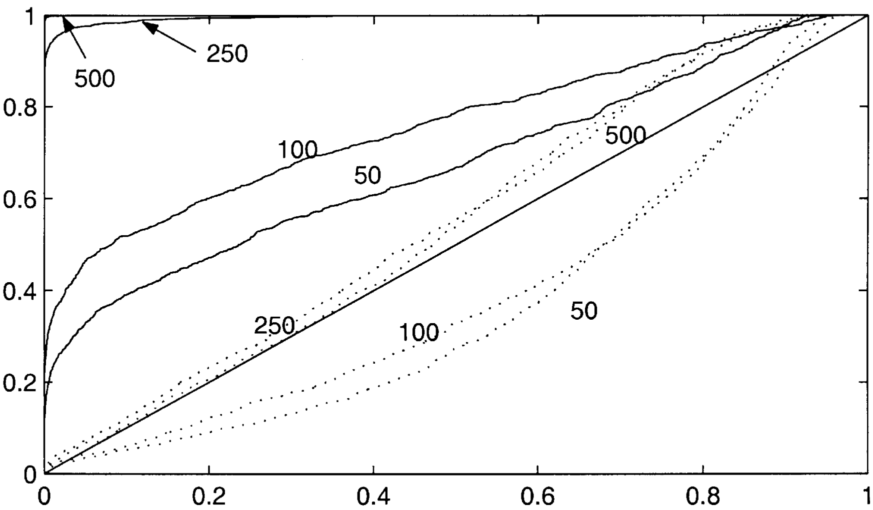

FIGURE 4. Size-power curves under non Gaussian model $X_n + U_n$ where $X_n \sim$ FARIMA$(0, d, 0)$ and U_n is an exponential white noise and fitting FARIMA$(0, \hat{d}, 0)$ model for S_n-test (plain) and Chen and Deo's test (dots). Model $d = 0.25$.

Finally we consider the following local alternative

(e) Generate FARIMA$(0, d_n, 0)$ with $d_n = n^{-\gamma}$ where n is the sample size. We fit a white noise.

This last case suggests the investigation of the power of the test against a sequence of alternatives tending to $(\mathbf{H}_0)$.

$(\mathbf{H}_1^{(n)})$ Let $(X_t^{(n)})_{t\in\mathbb{Z}}, n \in \mathbb{N}$ be a triangular array of processes admitting linear representations of the form (2.1) with innovation satisfying **(A5)** and with spectral densities $f_1^{(n)}(x) = \sigma_n^2 |1 - \mathrm{e}^{ix}|^{-2d_1^{(n)}} f^{*(n)}(x)$ with $(d_n) \in (0, 1/2)^{\mathbb{N}}$ and $f^{*(n)}$ is a sequence of functions twice continuously differentiable and bounded away from zero on $[-\pi, \pi]$, and such that the sequence $\epsilon_n = \inf_{f\in\mathcal{F}_0} S(f, f_1^{(n)})$ is positive and tends to zero as n goes to infinity.

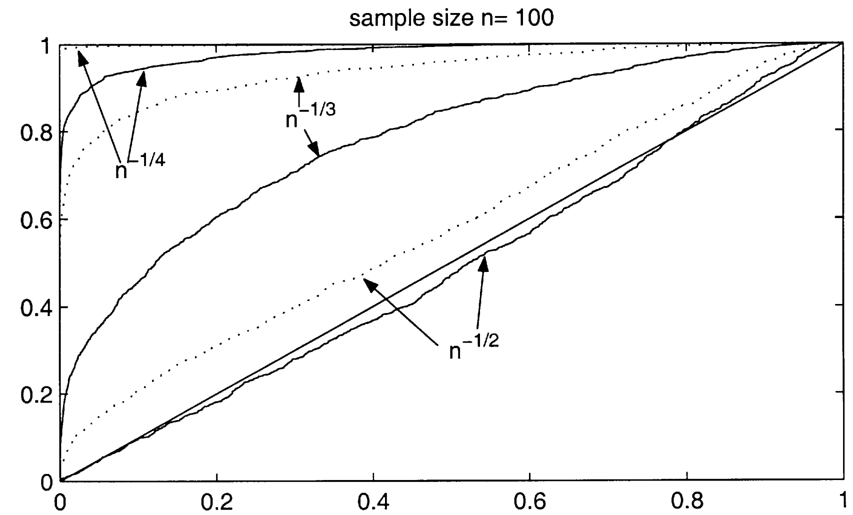

FIGURE 5. Size-power curve under an FARIMA$(0, d_n, 0)$ model and fitting white noise for S_n-test (plain) and Chen and Deo's test (dots).

Figure 5 gives the size-power curves when we fit a white noise on a simulated FARIMA$(0, d_n, 0)$ model. The last simulation seems to reveal that the S_n-test is not able to discriminate the alternative from the null if $\epsilon_n = O(n^{-1/2})$. It corresponds to a rate of $n^{-1/4}$ in the Kullback–Leibler divergence between f and $f_1^{(n)}$ or in the parameters for ARFIMA processes. Note that most of the goodness-of-fit procedures wish to discriminate contiguous alternatives at the optimal $\sqrt{n}$ rate on the parameters. This drawback of the S_n procedure may be explained by the fact that, if the Kullback–Leibler divergence $KL(f_n^{(1)}, f)$ between the sequence $f_n^{(1)}$ and f goes to zero, the divergence $S(f_n^{(1)}, f)$ goes to zero at a squared rate.

This remark suggests the investigation of others functionals of the periodogram with tractable asymptotic properties, in order to obtain more powerful procedures.

Figure 5 illustrates the preceding remark on the power when the parameter d_n satisfies the condition $d_n \gg 1/n^{1/4}$. It appears that for this example the Chen and Deo test is more powerful than S_n-test. However, no theoretical results are available to justify this phenomenon.

The distribution under $\mathbf{H}_1$ seem to be Gaussian in the case $d_1 - d_0 < 1/4$ and non-Gaussian (Rosenblatt?) in the case $d_1 - d_0 > 1/4$. This interpretation is motivated by Figure 6 and a Kolmogorv–Smirnoff test on the the empirical distribution of S_n.

REFERENCES

[1] T. Anderson, Goodness of fit tests for spectral distributions. *Ann. Statist.* **21** (1993) 830-847.

[2] J.-M. Bardet, G. Lang, G. Oppenheim, A. Philippe and M. Taqqu, *Generators of long-range dependent processes: A survey.* Birkhäuser (2002).

[3] M. Bartlett, *An introduction to stochastic processes.* Cambridge University Press (1955).

[4] G. Box and D.A. Pierce, Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *J. Am. Stat. Assoc.* **65** (1970) 1509-1526.

[5] P. Brockwell and R. Davis, *Time Series: Theory and Methods.* Springer-Verlag, *Springer Ser. in Statistics* (1991).

[6] W. Chen and R. Deo, *A generalized portmanteau goodness-of-fit test for time series models.* Preprint (2000).

[7] G. Fay, *Théorèmes limites pour les fonctionnelles du périodogramme*, Ph.D. Thesis. École Nationale Supérieure des Télécommunications (2000).

[8] G. Fay, E. Moulines and P. Soulier, *Non linear functionals of the periodogram* (submitted).

[9] G. Fay and P. Soulier, The periodogram of an i.i.d. sequence. *Stochastic Process. Appl.* **92** (2001) 315-343.

[10] R. Fox and M. Taqqu, Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Statist.* **14** (1986) 517-532.

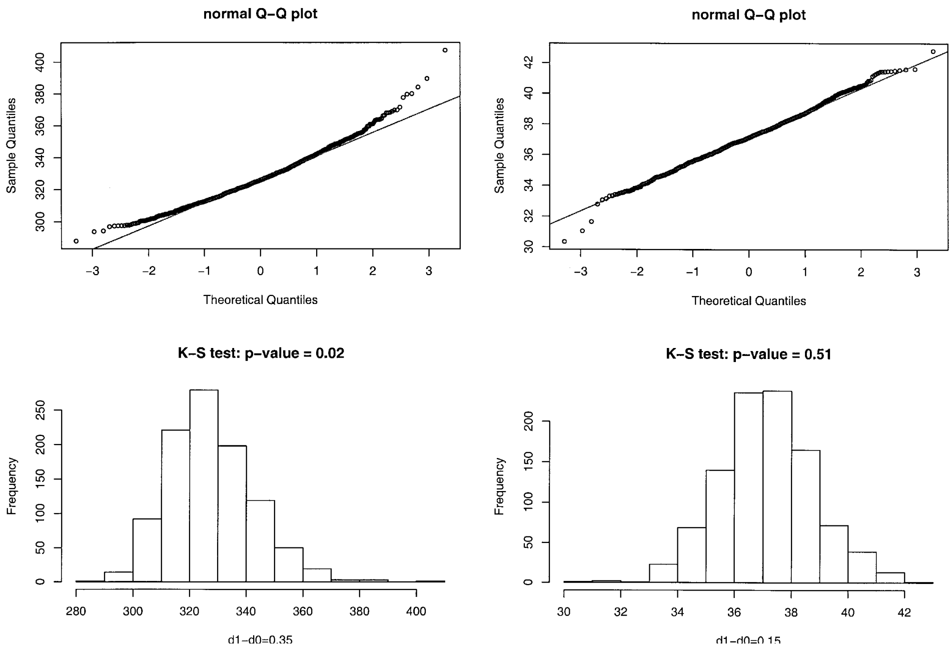

FIGURE 6. Empirical distributions under $\mathbf{H}_1$. Two regimes.

[11] L. Giraitis and D. Surgailis, A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotic normality of Whittles's estimate. *Probab. Theory Related Fields* **86** (1990) 87-104.

[12] U. Grenander and M. Rosenblatt, *Statistical analysis of stationary time series.* Wiley, New York (1957).

[13] Y. Hosoya, A limit theory for long-range dependence and statistical inference on related models. *Ann. Statist.* **25** (1997) 105-137.

[14] C. Hurvich, E. Moulines and P. Soulier, The FEXP estimator for potentially non-stationary linear time series. *Stochastic Process. Appl.* **97** (2002) 307-340.

[15] C.W. Hurvich and W. Chen, An efficient taper for potentially overdifferenced long-memory time series. *J. Time Ser. Anal.* **21** (2000) 155-180.

[16] D. Janas and R. von Sachs, Consistency for non-linear functions of the periodogram of tapered data. *J. Time Ser. Anal.* **16** (1995) 585-606.

[17] C. Klueppelberg and T. Mikosch, The integrated periodogram for stable processes. *Ann. Statist.* **24** (1996) 1855-1879.

[18] P. Kokoszka and T. Mikosch, The integrated periodogram for long-memory processes with finite or infinite variance. *Stochastic Process. Appl.* **66** (1997) 55-78.

[19] H. Künsch, Discrimination between monotonic trends and long-range dependence. *J. Appl. Probab.* **23** (1986) 1025-1030.

[20] T. Mikosch and R. Norvaisa, Uniform convergence of the empirical spectral distribution function. *Stochastic Process. Appl.* **70** (1997) 85-114.

[21] A. Mokkadem, Une mesure d'information et son application à des tests pour les processus arma. *C. R. Acad. Sci. Paris* **319** (1994) 197-200.

[22] A. Mokkadem, A measure of information and its applications to test for randomness against ARMA alternatives and to goodness-of-fit test. *Stochastic Process. Appl.* **72** (1997) 145-159.

[23] M. Taniguchi, On estimation of the integrals of certain functions of spectral density. *J. Appl. Probab.* **17** (1980) 73-83.

[24] C. Velasco, Non-stationary log-periodogram regression. *J. Econom.* **91** (1999) 325-371.

[25] Y. Yajima, Asymptotic properties of estimates in incorrect ARMA models for long-memory time series, in *New directions in time series analysis. Part II. Proc. Workshop, Minneapolis/MN (USA) 1990.* Springer, New York, *IMA Vol. Math. Appl.* **46** (1993) 375-382.