ESAIM: PS 17 (2013) 432–443
DOI: 10.1051/ps/2012006

ESAIM: Probability and Statistics
www.esaim-ps.org

GOODNESS-OF-FIT TESTS IN LONG-RANGE DEPENDENT PROCESSES UNDER FIXED ALTERNATIVES

Holger Dette[1] and Kemal Sen[1]

Abstract. In a recent paper Fay and Philippe [*ESAIM: PS* **6** (2002) 239–258] proposed a goodness-of-fit test for long-range dependent processes which uses the logarithmic contrast as information measure. These authors established asymptotic normality under the null hypothesis and local alternatives. In the present note we extend these results and show that the corresponding test statistic is also normally distributed under fixed alternatives.

Mathematics Subject Classification. 60F05, 62F03.

Received December 13, 2010. Revised October 17, 2011.

1. Introduction

Nowadays long-range dependent processes represent a well accepted class of stochastic processes for modelling real phenomena in such diverse areas as hydrology, behaviour research, network traffic or finance (see [15, 17, 20, 21] among many others). Many inference methods assume that the specific form of the spectrum is known except for a finite dimensional parameter. Therefore results of the statistical analysis depend sensitively on these assumptions, and the conclusions from the data may be misleading if these assumptions are violated. For this reason several authors have pointed out the importance of being able to check the goodness-of-fit of a specific model assumption in long-range dependent processes. Beran [2] proposed a method for testing how well a specified model, such as a fractional Gaussian noise, fits the data. His results were extended by Deo and Chen [7] who investigated an integral of the squared periodogram. Recently Chen and Deo [4] suggested a generalized Portmanteau test based on the discrete spectral average estimator and obtained the asymptotic null distribution for Gaussian long-memory time series. While most of the tests proposed by these authors are based on the estimation of the L^2 distance between the unknown spectral density and the best approximation by the parametric class, Fay and Philippe [12] proposed to measure the distance between two spectral densities f_1, f_2 by a modified Kullback–Leibler information

$$S(f_1, f_2) = \log \int_{-\pi}^{\pi} \frac{f_1(\lambda)}{f_2(\lambda)} \frac{\mathrm{d}\lambda}{2\pi} - \int_{-\pi}^{\pi} \log \frac{f_1(\lambda)}{f_2(\lambda)} \frac{\mathrm{d}\lambda}{2\pi}. \tag{1.1}$$

It was pointed out by Mokkadem [18] that the distance (1.1) is more adapted to derive useful tests than L_2-distances. Fay and Philippe [12] replaced f_1 by a pooled and tapered periodogram and f_2 by a parametric

Keywords and phrases. Long-range dependence, goodness-of-fit test, asymptotic power, periodogram.

[1] Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany. holger.dette@ruhr-uni-bochum.de; kemal.sen@ruhr-uni-bochum.de

Article published by EDP Sciences

© EDP Sciences, SMAI 2013

https://doi.org/10.1051/ps/2012006 Published online by Cambridge University Press

estimate of the spectral density and established the asymptotic normality of a corresponding test statistic under the null hypothesis and local alternatives.

As pointed out by Chen and Deo [4], most theoretical results in the context of goodness-of-fit testing address the asymptotic behaviour of a test statistic under the null hypothesis, and an additional question of interest is the power property of the corresponding test when the null hypothesis is actually not satisfied. This problem requires asymptotic inference under the alternative and has found considerable interest in the context of classical regression analysis (see [8] or [9] among others). Dette and Spreckelsen [11] investigated the asymptotic properties of an L^2-test proposed by Paparoditis [19] for the parametric form of the spectral density in stationary short-range dependent processes, but less results are available for goodness-of-fit tests in long range dependence processes.

The present paper is devoted to the asymptotic analysis of the test statistic proposed by Fay and Philippe [12] under fixed alternatives. In Section 2 we introduce the necessary notations and assumptions and review the results of Fay and Philippe [12]. Section 3 presents our main results which show that the test statistic proposed by Fay and Philippe [12] is also asymptotically normally distributed under fixed alternatives. We state a general result which contains the situation of a "true" null hypothesis as a special case and also discuss potential applications of our results. Finally, for the sake of a transparent presentation, some technical details are deferred to an appendix in Section 4.

2. Preliminaries

Throughout this paper let $X = (X_t)_{t\in\mathbb{Z}}$ denote a stationary Gaussian process. Following Fay and Philippe [12] we represent the spectral density f of the process $(X_t)_{t\in\mathbb{Z}}$ as

$$f(\lambda) = \sigma^2|1-\mathrm{e}^{\mathrm{i}\lambda}|^{-2d_1} f^*(\lambda); \quad \lambda\in[-\pi,\pi], \tag{2.1}$$

where $d_1 \in [0, 1/2)$ and f^* is a twice continuously differentiable function defined on the interval $[-\pi,\pi]$ and bounded away from zero. We are interested in the problem of testing for a specific parametric form of the spectral density of the process $(X_t)_{t\in\mathbb{Z}}$, that is

$$H_0:\ f\in\mathcal{F}_0 = \Big\{ f(\lambda) = \sigma^2|1-\mathrm{e}^{\mathrm{i}\lambda}|^{-2d} g^*(\lambda;\theta) \ \Big|\ (d,\theta)\in D\times\Theta, \sigma^2>0, g^*\in\mathcal{G}\Big\}. \tag{2.2}$$

Here D is a compact subset of the interval $[0, 1/2)$, $\Theta\subset\mathbb{R}^l$ denotes a compact set ($l\in\mathbb{N}$) and $\mathcal{G}$ is the set of positive and symmetric functions defined on the interval $[-\pi,\pi]$ satisfying $\int_{-\pi}^{\pi}\log g^*(x;\theta)\mathrm{d}x = 0$. For a given $g^*\in\mathcal{G}$ we define $g(\lambda;d,\theta) = |1-\mathrm{e}^{\mathrm{i}\lambda}|^{-2d} g^*(\lambda;\theta)$. Fay and Philippe [12] proposed to measure deviations from the null hypothesis (2.2) by

$$\inf_{d\in D,\theta\in\Theta} S(f, f(\cdot,d,\theta)) = \inf_{d\in D,\theta\in\Theta}\left\{\log\int_{-\pi}^{\pi}\frac{f(\lambda)}{f(\lambda,d,\theta)}\frac{\mathrm{d}\lambda}{2\pi} - \int_{-\pi}^{\pi}\log\frac{f(\lambda)}{f(\lambda,d,\theta)}\frac{\mathrm{d}\lambda}{2\pi}\right\}. \tag{2.3}$$

Note that the information measure in (2.3) is always nonnegative and that the null hypothesis is satisfied if and only if the expression in (2.3) vanishes. The logarithmic contrast has been used before by Mokkadem [18] and Dette and Spreckelsen [11] for testing hypotheses in ARMA processes. In order to estimate the minimal distance Fay and Philippe [12] proposed to consider a tapered Fourier transform of the series $\{X_1,\dots,X_n\}$ that is

$$d_{n,k}^{(p)} = \frac{1}{\sqrt{2\pi n}}\sum_{t=1}^{n} w_{n,t}^{(p)} X_t \mathrm{e}^{\mathrm{i}\lambda_k t}; \quad k=1,\dots,n$$

where $\lambda_k = 2\pi k/n$ are the Fourier frequencies, $w_{n,t}^{(p)} = \binom{2p}{p}^{-\frac{1}{2}}\big(1-\mathrm{e}^{\mathrm{i}\frac{2\pi t}{n}}\big)^p$; $t=1,\dots,n$ is the data taper and $p\in\mathbb{N}_0$ denotes the order of the taper (note that $p=0$ yields $w_{n,t}^{(0)}=1$, $t=1,\dots,n$). These quantities are used

https://doi.org/10.1051/ps/2012006 Published online by Cambridge University Press

to define a pooled periodogram by

$$\overline{I}_{n,k}^X := \frac{1}{m} \sum_{j=(m+p)(k-1)+1}^{(m+p)k-p} \left| d_{n,j}^{(p)} \right|^2; \quad k = 1, \ldots, K_n.$$

Note that the interval $[0, \pi]$ is decomposed in $K_n = \lfloor \frac{n-1}{2(m+p)} \rfloor$ intervals of the form $[\lambda_{(k-1)(m+p)}, \lambda_{k(m+p)}]$ $(m \in \mathbb{N})$ and that the center of the kth interval is given by $x_k := (m+p)\frac{2\pi}{n}\big(k - \frac{1}{2}\big)$. Fay and Philippe [12] introduced the discretized version of the term S_n in (2.3), *i.e.*

$$S_n\left(\overline{I}_n^X, g(\cdot; d, \theta)\right) = \log\left(\frac{1}{K_n}\sum_{k=1}^{K_n} \frac{\overline{I}_{n,k}^X}{g(x_k; d, \theta)}\right) - \frac{1}{K_n}\sum_{k=1}^{K_n} \log\left(\frac{\overline{I}_{n,k}^X}{g(x_k; d, \theta)}\right) + \gamma_{m,p}$$

where the constant $\gamma_{m,p}$ is defined by $\gamma_{m,p} = \mathbb{E}\big[\log 2\pi \overline{I}_{n,k}^Z\big]$ and $\overline{I}_{n,k}^Z$ denotes the pooled and tapered periodogram of a Gaussian white noise process $(Z_t)_{t\in\mathbb{Z}}$. Note that $\gamma_{m,p}$ is a centering constant, such that the expectation under the null hypothesis vanishes asymptotically. For the cases

$$(1)\ d_0 = 0,\ m \geq 5 \text{ and } p = 0 \text{ or } p = 1 \qquad (2)\ d_0 > 0,\ m \geq 5 \text{ and } p = 1.$$

Fay and Philippe [12] proved that under the null hypothesis, *i.e.* $f(\lambda) = g(\lambda, d_0, \theta_0)$ for some $(d_0, \theta_0) \in D \times \Theta$ and certain assumptions of regularity (see Sect. 3 for details), the statistic $\sqrt{K_n} S_n(\overline{I}_n^X, g(\cdot; \hat{d}_n, \hat{\theta}_n))$ converges weakly to a centered normal distribution with variance $\tau_0^2 := \mathrm{Var}(2\pi\overline{I}_{n,k}^Z - \log(2\pi\overline{I}_{n,k}^Z))$, where $(\hat{d}_n, \hat{\theta}_n)$ is any $\sqrt{n}$-consistent estimator of the "true" parameter (d_0, θ_0).

For a discussion of the quantities $\gamma_{m,p}$ and τ_0^2 we refer to Hurvich *et al.* [16]. Note that Fay and Philippe [12] did not assume a Gaussian stationary process, but considered a linear representation of the form $X_t = \sigma \sum_{j\in\mathbb{Z}} a_j Z_{t-j}$ with general white noise process $(Z_t)_{t\in\mathbb{Z}}$ satisfying several assumptions regarding the characteristic function $\mathbb{E}[\exp(iZ_t)]$. In this case there appears an additional constant in the asymptotic variance depending on the fourth cumulant of the white noise process. In the following section we will study the asymptotic properties of the statistic S_n under the assumption of a Gaussian stationary process if the null hypothesis is not satisfied.

3. Weak convergence under fixed alternatives

If the null hypothesis is not satisfied, then the minimum distance in (2.3) is positive. Throughout this paper we assume that there exists a unique pair $(d_0, \theta_0) \in (D \times \Theta)^0$ such that

$$\inf_{(d,\theta)\in D\times\Theta} S(f, g(\cdot; d, \theta)) = S(f, g(\cdot; d_0, \theta_0)),$$

where C^0 denotes the interior of the set $C \subset \mathbb{R}^{l+1}$ and D in (2.2) is defined by $D = [\delta, 1/2 - \delta]$ for some $0 < \delta < 1/4$. We further assume that the set Θ is additionally convex (see [5]). Note that (d_0, θ_0) is the parameter corresponding to the best approximation of the spectral density f by densities of the class $\mathcal{F}_0$. Throughout this paper let $(\hat{d}_n, \hat{\theta}_n)$ denote a Whittle type estimate [22] which is defined as the minimizer of the objective function

$$Q_n(d, \theta) = \frac{\pi}{K_n}\sum_{j=1}^{K_n} \frac{\overline{I}_{n,j}^x}{g(x_j; d, \theta)} \tag{3.1}$$

where x_j is defined in Section 2. In the case where the model is correctly specified, the asymptotic behaviour of the maximum likelihood estimator was investigated by Dahlhaus [6]. The Whittle estimator was investigated by Fox and Taqqu [13] and Giraitis and Surgailis [14] for Gaussian and linear processes, respectively. Recently, Chen

https://doi.org/10.1051/ps/2012006 Published online by Cambridge University Press

and Deo [5] derived the asymptotic properties of an estimator minimizing an approximation to the negative of the exact Gaussian likelihood [22] in the case of misspecified long-range dependent processes. Note that in contrast to these results, the objective function considered in (3.1) is based on the tapered and pooled periodogram in this definition, while Chen and Deo [5] considered the "classical" periodogram in the objective function (3.1). A careful inspection of the proofs in this reference shows that the main results, in particular Theorem 2 and Lemma 2 of Chen and Deo [5], remain valid in this case. It is also notable that the asymptotic properties – in particular the rate of convergence – depend sensitively on the distance $d_1 - d_0$. If $d_1 - d_0 \leq 1/4$ the estimator $\sqrt{n}((\hat{d}_n, \hat{\theta}_n) - (d_0, \theta_0))$ is asymptotically normal distributed, while in the case $d_1 - d_0 > 1/4$ the difference converges in distribution with a different rate to a non-Gaussian limit. In particular, the rate of convergence can be arbitrarily small in this case. In our main result we specify the asymptotic behaviour of the test statistic proposed by Fay and Philippe [12] in the case of a misspecified model. For this purpose we define by

$$D(d_0, \theta_0) := \log\left(\frac{1}{\pi}\int_0^\pi \frac{f(x)}{g(x; d_0, \theta_0)}\,\mathrm{d}x\right) - \frac{1}{\pi}\int_0^\pi \log\left(\frac{f(x)}{g(x; d_0, \theta_0)}\right)\mathrm{d}x \tag{3.2}$$

as the minimal distance between the true spectral density f and the parametric class defined in (2.2) with respect to the logarithmic contrast. Note that the null hypothesis (2.2) is satisfied if and only if $D(d_0, \theta_0) = 0$.

Theorem 3.1. *Let $(X_t)_{t\in\mathbb{Z}}$ be a stationary Gaussian process with spectral density given by (2.1), $d_1 \in (0, 1/2)$, $p = 1$, $m \geq 4$, $d_1 - d_0 < 1/4$, and assume that the following conditions are satisfied:*

(***A1***) *$g^*(\lambda; \theta)$ is three times continuously differentiable;*
(***A2***) $\inf_\theta \inf_\lambda g^*(\lambda;\theta) > 0$, $\sup_\theta \sup_\lambda g^*(\lambda;\theta) < \infty$;
(***A3***) $\sup_\lambda \sup_\theta \left|\frac{\partial g^*(\lambda;\theta)}{\partial \theta_i}\right| < \infty; \quad 1 \leq i \leq l;$
(***A4***) $\sup_\lambda \sup_\theta \left|\frac{\partial^2 g^*(\lambda;\theta)}{\partial \theta_i \partial \theta_j}\right| < \infty, \quad \sup_\lambda \sup_\theta \left|\frac{\partial^2 g^*(\lambda;\theta)}{\partial \theta_i \partial \lambda}\right| < \infty; \quad 1 \leq i, j \leq l;$
(***A5***) $\sup_\lambda \sup_\theta \left|\frac{\partial^3 g^*(\lambda;\theta)}{\partial \theta_i \partial \theta_j \partial \theta_k}\right| < \infty; \quad 1 \leq i, j, k \leq l;$
(***A6***) $\int_{-\pi}^{\pi} \log g^*(\lambda;\theta)\,\mathrm{d}\lambda = 0$ *for all $\theta \in \Theta$;*
(***A7***) *There exists a unique vector (d_0, θ_0), such that $(d_0, \theta_0) = \arg\min_{D\times\Theta} D(d, \theta)$.*

If $n \to \infty$, then

$$\sqrt{K_n}\left\{S_n\left(\overline{I}_n^X, g(\cdot; \hat{d}_n, \hat{\theta}_n)\right) - D(d_0, \theta_0)\right\} \xrightarrow{D} \mathcal{N}\left(0, \tau_\Delta^2\right)$$

where $D(d_0, \theta_0)$ denotes the minimal distance between the parametric class $\mathcal{F}_0$ and the unknown spectral density f defined in (2.1) and the asymptotic variance is given by

$$\tau_\Delta^2 := (\Delta - 1)\mathrm{Var}\left(2\pi \overline{I}_{n,k}^Z\right) + \mathrm{Var}\left(2\pi \overline{I}_{n,k}^Z - \log 2\pi \overline{I}_{n,k}^Z\right)$$

with

$$\Delta = \pi \int_0^\pi \left(\frac{f(x)}{g(x; d_0, \theta_0)}\right)^2 \mathrm{d}x \left(\int_0^\pi \frac{f(x)}{g(x; d_0, \theta_0)}\,\mathrm{d}x\right)^{-2}. \tag{3.3}$$

Proof. We define and decompose the statistic T_n

$$T_n = \sqrt{K_n}\left\{S_n\left(\overline{I}_n^X, g(\cdot; \hat{d}_n, \hat{\theta}_n)\right) - D(d_0, \theta_0)\right\} = \sqrt{K_n}\left\{A_n + B_n + C_n\right\},$$

where the random variables A_n, B_n and C_n are defined by

$$A_n := S_n\left(\overline{I}_n^X, f(\cdot)\right), \tag{3.4}$$

$$B_n := S_n\left(\overline{I}_n^X, g(\cdot; d_0, \theta_0)\right) - S_n\left(\overline{I}_n^X, f(\cdot)\right), \tag{3.5}$$

$$C_n := S_n\big(\overline{I}_n^X, g(\cdot; \hat{d}_n, \hat{\theta}_n)\big) - S_n\big(\overline{I}_n^X, g(\cdot; d_0, \theta_0)\big), \tag{3.6}$$

https://doi.org/10.1051/ps/2012006 Published online by Cambridge University Press

respectively. Since the process $(X_t)_{t\in\mathbb{Z}}$ is Gaussian, it has a linear representation of the form $X_t = \sigma \sum_{j\in\mathbb{Z}} a_j Z_j$, where $(a_j)_{j\in\mathbb{Z}}$ is square summable and $(Z_t)_{t\in\mathbb{Z}}$ is a Gaussian white noise process with unit variance. In the appendix we will show that

$$A_n = \frac{1}{K_n}\sum_{k=1}^{K_n}\left(2\pi\overline{I}^Z_{n,k} - 1 - \log 2\pi\overline{I}^Z_{n,k} + \gamma_{m,p}\right) + o_p\left(\frac{1}{\sqrt{K_n}}\right), \tag{3.7}$$

$$B_n = \sum_{k=1}^{K_n}\left(\beta_{n,k} - \frac{1}{K_n}\right)\left(2\pi\overline{I}^Z_{n,k} - 1\right) + \log\left(\frac{1}{K_n}\sum_{k=1}^{K_n}\frac{f(x_k)}{g(x_k;d_0,\theta_0)}\right) - \frac{1}{K_n}\sum_{k=1}^{K_n}\log\left(\frac{f(x_k)}{g(x_k;d_0,\theta_0)}\right) + o_p\left(\frac{1}{\sqrt{K_n}}\right), \tag{3.8}$$

$$C_n = o_p\left(\frac{1}{\sqrt{K_n}}\right) \tag{3.9}$$

where $\overline{I}^Z_{n,k}$ denotes the pooled and tapered periodogram of the Gaussian white noise process $(Z_t)_{t\in\mathbb{Z}}$ and the constants $\beta_{n,k}$ are defined by

$$\beta_{n,k} = \frac{\frac{f(x_k)}{g(x_k;d_0,\theta_0)}}{\sum_{j=1}^{K_n}\frac{f(x_j)}{g(x_j;d_0,\theta_0)}} = \frac{|1-\mathrm{e}^{\mathrm{i}x_k}|^{-2(d_1-d_0)}\frac{f^*(x_k)}{g^*(x_k;\theta_0)}}{\sum_{j=1}^{K_n}|1-\mathrm{e}^{\mathrm{i}x_j}|^{-2(d_1-d_0)}\frac{f^*(x_j)}{g^*(x_j;\theta_0)}}.$$

Observing the approximation

$$\begin{aligned}
&\log\left(\frac{1}{K_n}\sum_{k=1}^{K_n}\frac{f(x_k)}{g(x_k;d_0,\theta_0)}\right) - \frac{1}{K_n}\sum_{k=1}^{K_n}\log\left(\frac{f(x_k)}{g(x_k;d_0,\theta_0)}\right)\\
&= \log\left(\frac{1}{K_n}\sum_{k=1}^{K_n}|1-\mathrm{e}^{\mathrm{i}x_k}|^{-2(d_1-d_0)}\frac{f^*(x_k)}{g^*(x_k;\theta_0)}\right) - \frac{1}{K_n}\sum_{k=1}^{K_n}\log\left(|1-\mathrm{e}^{\mathrm{i}x_k}|^{-2(d_1-d_0)}\frac{f^*(x_k)}{g^*(x_k;\theta_0)}\right)\\
&= D(d_0,\theta_0) + O\left(n^{-1+2(d_1-d_0)^+}\right),
\end{aligned}$$

it follows that the weak convergence of the statistic T_n can be obtained from the asymptotic properties of the random variable

$$\tilde{T}_n = \sqrt{K_n}\sum_{k=1}^{K_n}\left\{\left(\beta_{n,k}2\pi\overline{I}^Z_{n,k} - \frac{1}{K_n}\log 2\pi\overline{I}^Z_{n,k}\right) - \left(\beta_{n,k} - \frac{1}{K_n}\gamma_{m,p}\right)\right\}.$$

For this purpose we use the central limit theorem of Ljapunov. To precise we note that the random variables $2\pi\overline{I}^Z_{n,k}$ are independent identically distributed with existing fourth moment satisfying

$$\mathbb{E}\left[2\pi\overline{I}^Z_{n,k}\right] = 1; \quad k = 1,\dots K_n. \tag{3.10}$$

Therefore we obtain for the variance of $\tilde{T}_n$ by a straightforward calculation

$$\mathrm{Var}[\tilde{T}_n] = \mathrm{Var}\left(2\pi\overline{I}^Z_{n,k}\right)K_n\sum_{k=1}^{K_n}\beta^2_{n,k} + \mathrm{Var}\left(\log 2\pi\overline{I}^Z_{n,k}\right) - 2\left\{\mathbb{E}\left[2\pi\overline{I}^Z_{n,k}\log 2\pi\overline{I}^Z_{n,k}\right] - \gamma_{m,p}\right\}.$$

Observing the approximation

$$\begin{aligned}
\frac{1}{K_n}\sum_{k=1}^{K_n}\left(\frac{f(x_k)}{g(x_k;d_0,\theta_0)}\right)^j &= \frac{1}{K_n}\sum_{k=1}^{K_n}\left(|1-\mathrm{e}^{\mathrm{i}x_k}|^{-2(d_1-d_0)}\frac{f^*(x_k)}{g^*(x_k,\theta_0)}\right)^j\\
&= \frac{1}{\pi}\int_0^\pi\left(\frac{f(x)}{g(x;d_0,\theta_0)}\right)^j\mathrm{d}x + O\left(n^{-1+2j(d_1-d_0)^+}\right); \quad j = 1,2
\end{aligned}$$

https://doi.org/10.1051/ps/2012006 Published online by Cambridge University Press

we obtain by a tedious calculation

$$\lim_{n\to\infty} K_n \sum_{k=1}^{K_n} \beta_{n,k}^2 = \Delta,$$

where Δ is defined in (3.3) and

$$\sum_{k=1}^{K_n} \beta_{n,k}^2 \le \frac{C}{n} \tag{3.11}$$

(note that $d_1 - d_0 < \frac{1}{4}$ by assumption). Combining these results gives for the asymptotic variance of $\tilde{T}_n$

$$\lim_{n\to\infty} \mathrm{Var}[\tilde{T}_n] = \tau_\Delta^2$$

where τ_Δ^2 is defined in Theorem 3.1. Note that $E[\log 2\pi \overline{I}_{n,k}^Z]^4$ is constant, then a similar calculation yields for the numerator in the Ljapunov condition

$$\begin{aligned}
&K_n^2 \sum_{k=1}^{K_n} \mathbb{E}\left[\beta_{n,k}\left(2\pi\overline{I}_{n,k}^Z - 1\right) - \frac{1}{K_n}\left(\log 2\pi\overline{I}_{n,k}^Z - \gamma_{m,p}\right)\right]^4 \\
&\le K_n^2 \sum_{k=1}^{K_n} \Bigg| \beta_{n,k}^4 \mathbb{E}\left[2\pi\overline{I}_{n,k}^Z - 1\right]^4 - 4\beta_{n,k}^3 \frac{1}{K_n}\mathbb{E}\left[\left(2\pi\overline{I}_{n,k}^Z - 1\right)^3\left(\log 2\pi\overline{I}_{n,k}^Z - \gamma_{m,p}\right)\right] \\
&\quad + 6\beta_{n,k}^2 \frac{1}{K_n^2}\mathbb{E}\left[\left(2\pi\overline{I}_{n,k}^Z - 1\right)^2\left(\log 2\pi\overline{I}_{n,k}^Z - \gamma_{m,p}\right)^2\right] \\
&\quad - 4\beta_{n,k}\frac{1}{K_n^3}\mathbb{E}\left[\left(2\pi\overline{I}_{n,k}^Z - 1\right)\left(\log 2\pi\overline{I}_{n,k}^Z - \gamma_{m,p}\right)^3\right] + \frac{1}{K_n^4}\mathbb{E}\left[\log 2\pi\overline{I}_{n,k}^Z - \gamma_{m,p}\right]^4 \Bigg| \\
&= O(1)\left\{K_n^2 \sum_{k=1}^{K_n} \beta_{n,k}^4 + K_n \sum_{k=1}^{K_n} \beta_{n,k}^3 + \sum_{k=1}^{K_n} \beta_{n,k}^2 + \frac{1}{K_n} + \frac{1}{K_n}\right\} = O\left(\frac{1}{n}\right),
\end{aligned}$$

where we have used (3.11) for the last estimate. This establishes the Lyapunov condition and the asymptotic normality of T_n follows observing that T_n and $\tilde{T}_n$ have the same asymptotic behavior. □

Remark 3.2.
(a) Note that Theorem 3.1 holds under the null hypothesis and under the alternative, in particular it reduces to Theorem 3.1 in Fay and Philippe [12]. These authors did not assume a Gaussian process and we conjecture that Theorem 3.1 remains valid under the assumptions made in this reference, where the asymptotic variance has to be replaced by

$$\tau_\Delta^2 + \frac{\kappa_4 \alpha_{m,p}}{8(m+p)}.$$

Here the constant $\alpha_{m,p}$ is defined by

$$\alpha_{m,p} = \mathbb{E}^2\left[\|\zeta\|^2 \Phi_{m,p}(\zeta)\right]$$

with

$$\begin{aligned}
\Phi_{m,p}(x) &= \frac{\psi_{m,p}(x)}{2m} - 1 - \ln\left(\frac{\psi_{m,p}(x)}{2m}\right) + \gamma_{m,p}, \\
\psi_{m,p}(x) &= \binom{2p}{p}^{-1} \sum_{j=1}^{m} \left| \sum_{l=0}^{p} \binom{p}{l} (-1)^l \left(x_{2(j+l)-1} + i x_{2(j+l)}\right) \right|^2
\end{aligned}$$

and ζ is a $2(m+p)$-dimensional standard Gaussian vector.

https://doi.org/10.1051/ps/2012006 Published online by Cambridge University Press

(b) Note that in the Gaussian case an alternative proof of Theorem 3.1, which avoids the Bartlett decomposition technique, could be obtained by applying general limit theorems for functionals of stationary Gaussian sequences (see *e.g.* [1]). In this case one could use the method of moments and the diagram formula.
(c) Note that the assumption $p = 1, m \geq 4$ is required for the proof of (3.8), which uses results in Hurvich *et al.* [16] (see Sect. 4.2).

Remark 3.3. In this remark we indicate two important applications of the Theorem 3.1. For a more detailed discussion we refer to Dette and Munk [10].

(1) If $D(d_0, \theta_0)$ is used as a measure for the deviation of the "true" spectral density from the parametric class $\mathcal{F}_0$, we obtain from Theorem 3.1 a consistent estimate of $D(d_0, \theta_0)$, and it follows that the interval

$$\left[S_n\left(\overline{I}_n^X, g(\cdot; \hat{d}_n, \hat{\theta}_n)\right) - \frac{\hat{\tau}_\Delta}{\sqrt{K_n}}\, u_{1-\alpha/2}, S_n\left(\overline{I}_n^X, g(\cdot; \hat{d}_n, \hat{\theta}_n)\right) + \frac{\hat{\tau}_\Delta}{\sqrt{K_n}}\, u_{1-\alpha/2} \right]$$

is an asymptotic $(1-\alpha)$ confidence interval for the logarithmic contrast $D(d_0, \theta_0)$, which measures the deviation from the parametric class $\mathcal{F}_0$. Here $u_{1-\alpha/2}$ denotes the $(1-\frac{\alpha}{2})$ quantile of the standard normal distribution and $\hat{\tau}_\Delta^2$ is a consistent estimate of the asymptotic variance τ_Δ^2.

(2) As pointed out by Fay and Philippe [12] an application of the asymptotic normality of the statistic $S_n(\overline{I}_n^X, g(\cdot; \hat{d}_n, \hat{\theta}_n))$ under the null hypothesis consists in the construction of an asymptotic level α test for the hypothesis of a parametric form of the spectral density of the long range dependence process. A consistent test is obtained by rejecting the null hypothesis whenever

$$S_n\left(\overline{I}_n^X, g(\cdot; \hat{d}_n, \hat{\theta}_n)\right) \geq \frac{\tau_0}{\sqrt{K_n}} u_{1-\alpha}$$

where τ_0^2 denotes the asymptotic variance under the null hypothesis (which has to be estimated in the case of a non Gaussian white noise). The asymptotic power of this test can now be approximated by Theorem 3.1, that is

$$\begin{aligned} P_{H_1}(\text{"}H_0 \text{ is rejected"}) &= P_{H_1}\left(S_n\left(\overline{I}_n^X, g(\cdot, \hat{d}_n, \hat{\theta}_n)\right) \geq \frac{\tau_0}{\sqrt{K_n}} u_{1-\alpha} \right) \\ &= P_{H_1}\left(\frac{\sqrt{K_n}}{\tau_\Delta} \left\{ S_n\left(\overline{I}_n^X, g(\cdot, \hat{d}_n, \hat{\theta}_n)\right) - D(d_0, \theta_0) \right\} \geq \frac{\tau_0}{\tau_\Delta} u_{1-\alpha} - \frac{\sqrt{K_n}}{\tau_\Delta} D(d_0, \theta_0) \right) \\ &\approx \Phi\left(\sqrt{K_n} \frac{D(d_0, \theta_0)}{\tau_\Delta} - \frac{\tau_0}{\tau_\Delta} u_{1-\alpha} \right), \end{aligned}$$

where τ_0 and τ_Δ denote the (asymptotic) standard deviation of $\sqrt{K_n} S_n\left(\overline{I}_n^X, g(\cdot; \hat{d}_n, \hat{\theta}_n)\right)$ under the null hypothesis and alternative, respectively, and Φ is the distribution function of the standard normal distribution.

(3) A further important application of Theorem 3.1 arises from the fact that in practice when testing the hypothesis of a parametric model the distance $D(d_0, \theta_0)$ will usually never be precisely identical 0. The more realistic question in this context is, if the parametric is approximately correct. Therefore we propose to investigate the so called *precise hypotheses* [3]

$$H_0 : D(d_0, \theta_0) > \varepsilon \quad \textit{versus} \quad H_1 : D(d_0, \theta_0) \leq \varepsilon\,, \tag{3.12}$$

where $D(d_0, \theta_0)$ is the measure defined by (3.2) and $\varepsilon > 0$ is a prespecified constant for which the statistician agrees to analyse the data under the additional assumption of the parametric model. An asymptotic level α test for the hypothesis (3.12) is obtained by rejecting the null hypothesis, whenever

$$\sqrt{\frac{K_n}{\tau_\Delta^2}} \left\{ S_n\left(\overline{I}_n^X, g\left(\cdot; \hat{d}_n, \hat{\theta}_n\right)\right) - \varepsilon \right\} \; < \; u_\alpha.$$

https://doi.org/10.1051/ps/2012006 Published online by Cambridge University Press

TABLE 1. Different short memory parameters in the FARIMA$(0, 0.4, 1)$-process with spectral density given by (3.14). The values d_0 and $D(d_0)$ correspond to the long range parameter and minimal distance of the best approximation of the process with density (3.14) by a process with density (3.13).

β	-0.05	-0.4	-0.65
d_0	0.3698	0.1731	0.0505
$d_1 - d_0$	0.0302	0.2269	0.3495
$D(d_0)$	0.0009555	0.0511878	0.1196544

TABLE 2. Simulated coverage probabilities of the asymptotic confidence intervals defined in Remark 3.3 (1). The parametric and the "true" density are given by (3.14) and (3.13), respectively, where $\beta = -0.05$.

α	$n = 100$	$n = 200$	$n = 500$
0.2	0.8308	0.8258	0.8142
0.1	0.9440	0.9326	0.9196
0.05	0.9644	0.9732	0.9664

Example 3.4. In this example we illustrate the accuracy of the confidence interval for the distance $D(d_0, \theta_0)$ in Remark 3.3 (1) by means of a small simulation study. We assume that the process $X = (X_t)_{t\in\mathbb{Z}}$ is a Gaussian FARIMA$(0, d, 0)$-process with spectral density

$$g(\lambda; d, \theta) = \frac{1}{2\pi}\left|1 - \mathrm{e}^{\mathrm{i}\lambda}\right|^{-2d} \tag{3.13}$$

but generated data from a Gaussian FARIMA(0,0.4,1)-process with spectral density given by

$$f(\lambda) = \frac{1}{2\pi}|1 + \beta\mathrm{e}^{\mathrm{i}\lambda}|^2|1 - \mathrm{e}^{\mathrm{i}\lambda}|^{-2\cdot 0.4}, \tag{3.14}$$

where $\beta = -0.65, -0.4, -0.1, -0.05$. The corresponding values for d_0 and $D(d_0)$ have been calculated as described in Chen and Deo [5] and are depicted in Table 1. It is notable that the changes in the short-memory component β also result in a change in the difference between the long memory parameters d_0 and d_1.

We have generated 5000 replications of the process for sample sizes $n = 100, 200$ and 500 using the farimaSim function in the fArma package in R. The parameter d_0 in the variance τ_Δ^2 was estimated by the Whittle estimator in (3.1). The other quantities in the asymptotic variance have been determined explicitly by numerical integration and are given by

$$\begin{aligned} \gamma_{m,p} &= -0.1400195, \\ \mathrm{Var}\left(2\pi\overline{I}^Z_{n,k}\right) &= 0.2795195, \\ \mathrm{Var}\left(2\pi\overline{I}^Z_{n,k} - \log 2\pi\overline{I}^Z_{n,k}\right) &= 0.03776237. \end{aligned}$$

For each series the 80%, 90% and 95% confidence intervals ($p = 1$, $m = 5$) were calculated and the proportion of the intervals containing the true value $D(0.34)$ are listed in Tables 2–4 corresponding to the cases $\beta = -0.05, -0.4$ and -0.65, respectively.

We observe reasonable coverage probabilities for the two scenarios $\beta = -0.05$ and $\beta = -0.4$ (see Tabs. 2 and 3). If $\beta = -0.05$ the coverage probability is already very good for the sample size $n = 100$ (see Tab. 2), while a larger sample (namely $n = 200$) is needed to receive satisfactory results in the case $\beta = -0.4$. Note that the results in Table 4 correspond to the situation $d_1 - d_0 = 0.3495$, where the assumptions of Theorem 3.1 are

https://doi.org/10.1051/ps/2012006 Published online by Cambridge University Press

TABLE 3. Simulated coverage probabilities of the asymptotic confidence intervals defined in Remark 3.3 (1). The parametric and the "true" density are given by (3.14) and (3.13), respectively, where $\beta = -0.4$.

α	$n = 100$	$n = 200$	$n = 500$
0.2	0.9118	0.8834	0.7886
0.1	0.9542	0.9456	0.8878
0.05	0.9750	0.9726	0.9406

TABLE 4. Simulated coverage probabilities of the asymptotic confidence intervals defined in Remark 3.3 (1). The parametric and the "true" density are given by (3.14) and (3.13), respectively, where $\beta = -0.65$.

α	$n = 100$	$n = 200$	$n = 500$
0.2	0.7674	0.6388	0.1458
0.1	0.8652	0.7832	0.2826
0.05	0.9264	0.8758	0.4294

TABLE 5. Simulated coverage probabilities of the asymptotic confidence intervals defined in Remark 3.3 (1). The "true" and the parametric density are given by (3.13) and a FARIMA$(0, 0.4, 1)$ process with an additive exponential white noise, respectively.

α	$n = 100$	$n = 200$	$n = 500$
0.2	0.6896	0.7874	0.8096
0.1	0.7874	0.8846	0.9142
0.05	0.8106	0.9250	0.9638

not satisfied. In this case the asymptotic properties of test statistic are different and this fact is clearly reflected by the empirical results, which show the coverage probability is strictly decreasing with increasing sample size.

We finally investigate the coverage probability of the confidence intervals in a model with additional noise. To be precise we consider a process of the form $X_n + U_n$ where X_n is a FARIMA$(0, 0.4, 1)$ process with $\beta = -0.1$ and U_n is an exponential white noise independent of the process X_n. The assumed model is again specified by spectral density (3.13). The corresponding results are depicted in Table 5 and we observe a reasonable approximation of the coverage probabilities for sample sizes larger than $n = 200$.

APPENDIX A: TECHNICAL DETAILS

In this appendix we provide the technical details for the stochastic expansions (3.7)–(3.9).

A.1. Proof of (3.7)

We use a Bartlett decomposition technique, *i.e.* we relate the periodogram of X to the periodogram of Z and then apply Lemma 4.2 in Fay and Philippe [12] to show that the difference is stochastically small, *i.e.*

$$\begin{aligned} A_n &= S_n\left(\overline{I}_n^X, f(\cdot)\right) = S_n\left(2\pi\overline{I}_n^Z, 1\right) + R_n \\ &= \log\left(\frac{1}{K_n}\sum_{k=1}^{K_n} 2\pi\overline{I}_{n,k}^Z\right) - \frac{1}{K_n}\sum_{k=1}^{K_n}\log\left(2\pi\overline{I}_{n,k}^Z\right) + \gamma_{m,p} + o_p\left(\frac{1}{\sqrt{K_n}}\right). \end{aligned}$$

Using (3.10) and the independence of the $\overline{I}_{n.k}^Z$ we can expand the first term into a Taylor series and obtain the stochastic expansion in (3.7).

https://doi.org/10.1051/ps/2012006 Published online by Cambridge University Press

A.2. Proof of (3.8)

Recall the definition of B_n in (3.5). For a proof of (3.8) we use the Bartlett decomposition twice, which yields

$$\begin{aligned}
B_n &= \log\left(\frac{\sum_{k=1}^{K_n} \overline{I}^X_{n,k}/g(x_k;d_0,\theta_0)}{\sum_{k=1}^{K_n} 2\pi \overline{I}^Z_{n,k}}\right) - \log\left(\frac{\sum_{k=1}^{K_n} \overline{I}^X_{n,k}/f(x_k)}{\sum_{k=1}^{K_n} 2\pi \overline{I}^Z_{n,k}}\right) - \frac{1}{K_n}\sum_{k=1}^{K_n} \log\left(\frac{f(x_k)}{g(x_k;d_0,\theta_0)}\right) \\
&= \log\left(\frac{1}{K_n}\sum_{k=1}^{K_n} \frac{\overline{I}^X_{n,k}}{g(x_k;d_0,\theta_0)}\right) - \log\left(\frac{1}{K_n}\sum_{k=1}^{K_n} 2\pi \overline{I}^Z_{n,k}\right) - \frac{1}{K_n}\sum_{k=1}^{K_n} \log\left(\frac{f(x_k)}{g(x_k;d_0,\theta_0)}\right) + o_p\left(\frac{1}{\sqrt{K_n}}\right) \\
&= \log\left(\sum_{k=1}^{K_n} \beta_{n,k}\frac{\overline{I}^X_{n,k}}{f(x_k)}\right) + \log\left(\frac{1}{K_n}\sum_{k=1}^{K_n} \frac{f(x_k)}{g(x_k;d_0,\theta_0)}\right) - \log\left(\frac{1}{K_n}\sum_{k=1}^{K_n} 2\pi \overline{I}^Z_{n,k}\right) \\
&\quad - \frac{1}{K_n}\sum_{k=1}^{K_n} \log\left(\frac{f(x_k)}{g(x_k;d_0,\theta_0)}\right) + o_p\left(\frac{1}{\sqrt{K_n}}\right),
\end{aligned}$$

where the second estimate follows from Lemma 2 in Hurvich *et al.* [16]. We note that by the central limit theorem

$$\frac{1}{K_n}\sum_{k=1}^{K_n}\left(2\pi \overline{I}^Z_{n,k} - 1\right) = \frac{1}{K_n}\sum_{k=1}^{K_n} 2\pi \overline{I}^Z_{n,k} - 1 = O_p\left(\frac{1}{\sqrt{n}}\right). \tag{A.1}$$

We will show at the end of this section that

$$\sum_{k=1}^{K_n} \beta_{n,k}\left(\frac{\overline{I}^X_{n,k}}{f(x_k)} - 1\right) = O_p\left(\frac{1}{\sqrt{n}}\right), \tag{A.2}$$

then the expansion of the function $\log(1+z) = z + o(z^2)$ yields with the estimates (A.2) and (A.1) (note that $\sum_{k=1}^{K_n} \beta_{n,k} = 1$)

$$\begin{aligned}
B_n &= \sum_{k=1}^{K_n} \beta_{n,k}\left(\frac{\overline{I}^X_{n,k}}{f(x_k)} - 1\right) - \frac{1}{K_n}\sum_{k=1}^{K_n}\left(2\pi \overline{I}^Z_{n,k} - 1\right) + o_p\left(\frac{1}{\sqrt{K_n}}\right) \\
&\quad + \log\left(\frac{1}{K_n}\sum_{k=1}^{K_n} \frac{f(x_k)}{g(x_k;d_0,\theta_0)}\right) - \frac{1}{K_n}\sum_{k=1}^{K_n} \log\left(\frac{f(x_k)}{g(x_k;d_0,\theta_0)}\right).
\end{aligned}$$

Observing Lemma 11 in Hurvich *et al.* [16] we have

$$\mathbb{E}\left|\sum_{k=1}^{K_n} \beta_{n,k}\left(\frac{\overline{I}^X_{n,k}}{f(x_k)} - 2\pi \overline{I}^Z_{n,k}\right)\right| \le \sum_{k=1}^{K_n} \beta_{n,k}\mathbb{E}\left|\frac{\overline{I}^X_{n,k}}{f(x_k)} - 2\pi \overline{I}^Z_{n,k}\right| \le \begin{cases} O\left(n^{-1+2(d_1-d_0)^+}\right) & \text{if } d_1 - d_0 \neq 0 \\ O\left(\frac{\log n}{n}\right) & \text{if } d_1 - d_0 = 0 \end{cases}$$

which yields

$$\sum_{k=1}^{K_n} \beta_{n,k}\left(\frac{\overline{I}^X_{n,k}}{f(x_k)} - 2\pi \overline{I}^Z_{n,k}\right) = o_p\left(\frac{1}{\sqrt{n}}\right). \tag{A.3}$$

(note that $d_1 - d_0 < \frac{1}{4}$ by assumption). Therefore the assertion in (3.8) follows from (A.2) and (A.3).

We conclude this section with a proof of the statement (A.2) which is obtained observing the decomposition

$$\sum_{k=1}^{K_n} \beta_{n,k}\left(\frac{\overline{I}^X_{n,k}}{f(x_k)} - 1\right) = \sum_{k=1}^{K_n} \beta_{n,k}\left(\frac{\overline{I}^X_{n,k}}{f(x_k)} - 2\pi \overline{I}^Z_{n,k}\right) + \sum_{k=1}^{K_n} \beta_{n,k}\left(2\pi \overline{I}^Z_{n,k} - 1\right) = O_p\left(\frac{1}{\sqrt{K_n}}\right)$$

where the last estimate follows again from (A.3) and a straightforward application of Chebyshev's inequality.

https://doi.org/10.1051/ps/2012006 Published online by Cambridge University Press

A.3. Proof of (3.9)

Observing the definition (3.6) we decompose C_n as follows

$$C_n = C_n^{(1)} + C_n^{(2)}$$

where

$$C_n^{(1)} = \log\left(\frac{1}{K_n}\sum_{k=1}^{K_n}\frac{\overline{I}_{n,k}^X}{g(x_k;\hat{\Gamma}_n)}\right) - \log\left(\frac{1}{K_n}\sum_{k=1}^{K_n}\frac{\overline{I}_{n,k}^X}{g(x_k;\Gamma_0)}\right), \tag{A.4}$$

$$C_n^{(2)} = \frac{1}{K_n}\sum_{k=1}^{K_n}\log\frac{g\left(x_k;\hat{\Gamma}_n\right)}{g(x_k;\Gamma_0)}, \tag{A.5}$$

and we have used the notation $\hat{\Gamma}_n = (\hat{d}_n, \hat{\theta}_n)$ and $\Gamma_0 = (d_0, \theta_0)$. The assertion in (3.9) is now obtained by treating these terms separately, that is

$$C_n^{(j)} = o_p\left(\frac{1}{\sqrt{n}}\right), \quad j = 1, 2. \tag{A.6}$$

For a proof of (A.6) in the case $j = 1$ we note that the estimate $\hat{\Gamma}_n = (\hat{d}_n, \hat{\theta}_n)$ is defined as a solution of the equation

$$\frac{\partial Q_n(\hat{\Gamma}_n)}{\partial \Gamma} = 0,$$

where the function Q_n is defined in (3.1). Therefore a Taylor expansion yields

$$\begin{aligned} C_n^{(1)} &= \log Q_n(\Gamma_0) - \log Q_n\left(\hat{\Gamma}_n\right) \\ &= \frac{1}{2}\left(\Gamma_0 - \hat{\Gamma}_n\right)^T \frac{1}{Q_n(\hat{\Gamma}_n)}\frac{\partial^2 Q_n(\hat{\Gamma}_n)}{\partial\Gamma\partial\Gamma^T}\left(\Gamma_0 - \hat{\Gamma}_n\right) + o\left(\left\|\Gamma_0 - \hat{\Gamma}_n\right\|^2\right). \end{aligned}$$

An extension of Theorem 2, Lemmas 2 and 3 in Chen and Deo [5] to the objective function (3.1) yields

$$\begin{aligned} \hat{\Gamma}_n - \Gamma_0 &= O_p\left(\frac{1}{\sqrt{n}}\right) \\ \frac{1}{Q_n\left(\hat{\Gamma}_n\right)} &\xrightarrow{P} \frac{1}{Q\left(\Gamma_0\right)} = \left(\int_0^\pi \frac{f(\lambda)}{g(\lambda;\Gamma_0)}\,\mathrm{d}\lambda\right)^{-1}, \\ \frac{\partial^2 Q_n\left(\hat{\Gamma}_n\right)}{\partial\Gamma\partial\Gamma^T} &\xrightarrow{P} \frac{\partial^2 Q(\Gamma_0)}{\partial\Gamma\partial\Gamma^T}, \end{aligned}$$

and assertion (A.6) follows in the case $j = 1$.

In order to prove the statement in the case $j = 2$ we recall the definition in (A.5) and obtain by a Taylor expansion

$$\begin{aligned} C_n^{(2)} &= \frac{1}{K_n}\sum_{k=1}^{K_n}\left\{\frac{1}{g(x_k;\Gamma_0)}\frac{\partial g(x_k;\Gamma_0)}{\partial\Gamma}\left(\hat{\Gamma}_n - \Gamma_0\right)\right\} + O_p\left(\frac{1}{n}\right) \\ &= O_p\left(\frac{1}{\sqrt{n}}\right)\frac{1}{K_n}\sum_{k=1}^{K_n}\left\{\frac{1}{g(x_k;\Gamma_0)}\frac{\partial g(x_k;\Gamma_0)}{\partial\Gamma}\right\} + o_p\left(\frac{1}{\sqrt{n}}\right) \end{aligned} \tag{A.7}$$

https://doi.org/10.1051/ps/2012006 Published online by Cambridge University Press

where we have again used an extension of Theorem 2 in Chen and Deo [5] to the loss function (3.1). From the assumption $g(\lambda;\Gamma) \in \mathcal{F}_0$ we have

$$\int_{-\pi}^{\pi} \log g(\lambda;\Gamma)\,\mathrm{d}\lambda = \int_{-\pi}^{\pi} \log g^*(\lambda,\theta)\mathrm{d}\lambda = 0$$

for all $\Gamma \in D \times \Theta$, which implies (observing the symmetry of the function g) that the sum in (A.7) converges to 0 (a.s.). This proves the statement (A.6) in the case $j = 2$.

Acknowledgements. The authors would like to thank Martina Stein, who typed parts of this manuscript with considerable technical expertise. We are also grateful to G. Fay for helpful discussion during the preparation of this manuscript and to two unknown referees for their constructive comments on an earlier version of this manuscript. The work of the authors has been supported in part by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823) of the German Research Foundation (DFG).

References

[1] M.A. Arcones, Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *Ann. Probab.* **22** (1994) 2242–2274.
[2] J. Beran, A goodness-of-fit test for time series with long-range dependence. *J. R. Stat. Soc., Ser. B Stat. Methodol.* **54** (1992) 749–760.
[3] J.O. Berger and M. Delampady, Testing precise hypotheses. *Stat. Sci.* **2** (1987) 317–335.
[4] W.W. Chen and R.S. Deo, A generalized Portmanteau goodness-of-fit test for time series models. *Econ. Theory* **20** (2004) 382–416.
[5] W.W. Chen and R.S. Deo, Estimation of misspecified long-memory models. *J. Econom.* **134** (2006) 257–281.
[6] R. Dahlhaus, Efficient parameter estimation for self-similar processes. *Ann. Stat.* **17** (1989) 1749–1766.
[7] R.S. Deo and W.W. Chen, On the integral of the squared periodogram. *Stoc. Proc. Appl.* **85** (2000) 159–176.
[8] H. Dette, A consistent test for the functional form of a regression based on a difference of variance estimators. *Ann. Statist.* **27** (1999) 1012–1040.
[9] H. Dette, A consistent test for heteroscedasticity in nonparametric regression based on the kernel method. *J. Statist. Plann. Inference* **103** (2002) 311–329.
[10] H. Dette and A. Munk, Some methodological aspects of validation of models in nonparametric regression. *Stat. Neerl.* **57** (2003) 207–244.
[11] H. Dette and I. Spreckelsen, A note on a specification test for time series models based on spectral density estimation. *Scand. J. Stat.* **30** (2003) 481–491.
[12] G. Fay and A. Philippe, Goodness-of-fit test for long range dependent processes. *ESAIM: PS* **6** (2002) 239–258.
[13] R. Fox and M.S. Taqqu, Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Stat.* **14** (1986) 517–532.
[14] L. Giraitis and D. Surgailis, A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotical normality of Whittle's estimate. *Probab. Theory Relat. Fields* **86** (1990) 87–104.
[15] M.T. Greene and B.D. Fielitz, Long-term dependence in common stock returns. *J. Financ. Econ.* **4** (1977) 339–349.
[16] C. Hurvich, E. Moulines and P. Soulier, The FEXP estimator for potentially non-stationary linear time series. *Stochastic Processes Appl.* **97** (2002) 307–340.
[17] D. Koutsoyiannis, C. Makropoulos, A. Langousis, S. Baki, A. Efstratiadis, A. Christofides, G. Karavokiros and N. Mamassis, *HESS* opinions: climate, hydrology, energy, water: recognizing uncertainty and seeking sustainability. *Hydrol. Earth Syst. Sci.* **13** (2009) 247–257.
[18] A. Mokkadem, A measure of information and its applications to test for randomness against ARMA alternatives and to goodness-of-fit test. *Stochastic Processes Appl.* **72** (1997) 145–159.
[19] E. Paparoditis, Spectral density based goodness-of-fit tests for time series models. *Scand. J. Statist.* **27** (2000) 143–176.
[20] K. Park and W. Willinger, Self-similar network traffic: an overview, in *Self-Similar Network Traffic and Performance Evaluation*, edited by K. Park and W. Willinger. Wiley Interscience, New York (2000) 1–39.
[21] E. Stroe-Kunold, T. Stadnytska, J. Werner and S. Braun, Estimating long-range dependence in time series: An evaluation of estimators implemented in R. *Behavior Res. Methods* **41** (2009) 909–923.
[22] P. Whittle, Estimation and information in stationary time series. *Ark. Mat.* **1** (1953) 423–434.

https://doi.org/10.1051/ps/2012006 Published online by Cambridge University Press