SFB 823

Goodness-of-fit tests in long-range dependent processes under fixed alternatives

Holger Dette, Kemal Sen

Discussion Paper

Nr. 48/2010

Goodness-of-fit tests in long-range dependent processes under fixed alternatives

Holger Dette, Kemal Sen

Ruhr-Universität Bochum

Fakultät für Mathematik

44780 Bochum

Germany

email: holger.dette@ruhr-uni-bochum.de, kemal.sen@ruhr-uni-bochum.de

FAX: +49 234 32 14559

December 3, 2010

Abstract

In a recent paper Fay and Philippe (2002) 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.

AMS Subject Classification: 60F05, 62F03

Keywords: long-range dependence, goodness-of-fit test, asymptotic power, periodogram

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 Koutsoyiannis et al. (2009), Stroe-Kunold et al. (2009), Park and Willinger (2000), Granger

(1980), Greene and Fielitz (1977) among many others]. Numerous parametric models have been proposed for the analysis with long-range dependent processes. The most important among them are fractional ARIMA processes which were independently introduced by Granger and Joyeux (1980) and Hosking (1981) and fractional Gaussian noise processes [see Beran (1994)]. Many of the methods assume that the specific form of the spectrum is known except for a finite dimensional parameter. The results of the statistical analysis depend sensitively on pre-specified model 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 (1992) 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 (2000) who investigated an integral of the squared periodogram. Chen and Deo (2004) 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 (2002) used a logarithmic contrast for the construction of a test for a specific parametric form of the spectral density [see also Mokkadem (1997) or Dette and Spreckelsen (2003) for an application of this measure in the context of ARMA processes]. These authors established the asymptotic normality of a corresponding test statistic under the null hypothesis and local alternatives.

As pointed out by Chen and Deo (2004), most theoretical results in the context of goodness-of-fit testing address the asymptotic behaviour of a test statistic when the null hypothesis is correctly specified, and an additional question of interest is the power property of the corresponding test when the null hypothesis is actually misspecified. This problem requires asymptotic inference under the alternative and has found considerable interest in the context of classical regression analysis [see Dette (1999) or Dette (2002) among others]. Dette and Spreckelsen (2003) investigated the asymptotic properties of an L^2-test proposed by Paparoditis (2000) 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 (2002) under fixed alternatives. In Section 2 we introduce the necessary notations and assumptions and review the results of Fay and Philippe (2002). Section 3 presents our main results which show that the test statistic proposed by Fay and Philippe (2002) 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

Let $X = (X_t)_{t\in\mathbb{Z}}$ denote a stochastic process which admits the linear representation

$$X_t = \sigma \sum_{j\in\mathbb{Z}} a_j Z_{t-j} \tag{2.1}$$

where $\sum_{j\in\mathbb{Z}} a_j^2 < \infty$ and $(Z_t)_{t\in\mathbb{Z}}$ denotes a Gaussian white noise process. Following Fay and Philippe (2002) we represent the spectral density f of the process $(X_t)_{t\in\mathbb{Z}}$ as

$$f(\lambda) = \sigma^2 |1 - e^{i\lambda}|^{-2d_1} f^*(\lambda);\ \ \lambda \in [-\pi, \pi] \tag{2.2}$$

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

$$H_0 :\ f \in \mathcal{F}_0\ . \tag{2.3}$$

Here $\mathcal{F}_0$ denotes a parametric class of spectral densities defined by

$$\mathcal{F}_0 = \left\{ f(\lambda) = \sigma^2 |1 - e^{i\lambda}|^{-2d} g^*(\lambda;\theta)\ \middle|\ (d,\theta) \in D \times \Theta, \sigma^2 > 0, g^* \in \mathcal{G} \right\}, \tag{2.4}$$

where 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

$$\int_{-\pi}^{\pi} \log g^*(x;\theta) dx = 0\ .$$

For a given $g^* \in \mathcal{G}$ we define

$$g(\lambda; d, \theta) = |1 - e^{i\lambda}|^{-2d} g^*(\lambda;\theta).$$

For the testing problem (2.3) Fay and Philippe (2002) proposed to measure deviations from the null hypothesis by

$$\inf_{d\in D, \theta\in\Theta} S(f, f(\cdot, d, \theta)) \tag{2.5}$$

where

$$S(f, f(\cdot, d, \theta)) = \log \int_{-\pi}^{\pi} \frac{f(\lambda)}{f(\lambda, d, \theta)} \frac{d\lambda}{2\pi} - \int_{-\pi}^{\pi} \log \frac{f(\lambda)}{f(\lambda, d, \theta)} \frac{d\lambda}{2\pi} \tag{2.6}$$

denotes a logarithmic contrast between the spectral density f and an element of the class $\mathcal{F}_0$. Note that the information measure in (2.6) is always nonnegative and that the null hypothesis is satisfied if and only if the expression in (2.5) vanishes. The logarithmic contrast has been used before by Mokkadem

(1997) and Dette and Spreckelsen (2003) for testing hypotheses in ARMA processes. In order to estimate the minimal distance Fay and Philippe (2002) proposed to consider a tapered Fourier transform of the series $\{X_1, \dots, X_n\}$ that is

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

where $\lambda_k = 2\pi k/n$ are the Fourier frequencies,

$$w_{n,t}^{(p)} = \binom{2p}{p}^{-\frac{1}{2}} \left(1 - e^{i\frac{2\pi t}{n}}\right)^p; \quad t = 1, \dots, n$$

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 to define a pooled periodogram by

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

Throughout this paper $\overline{I}_{n,k}^Z$ denotes the pooled and tapered periodogram of the Gaussian white noise $Z_1, \dots, Z_n$. Note that the interval $[0, \pi]$ is decomposed in $K_n = \lfloor \frac{n-1}{2(m+p)} \rfloor$ intervals ($m \in \mathbb{N}$) of the form

$$[\lambda_{(k-1)(m+p)}, , \lambda_{k(m+p)}]$$

and that the center of the kth interval is given by

$$x_k := (m+p)\frac{2\pi}{n}\left(k - \frac{1}{2}\right). \tag{2.7}$$

Fay and Philippe (2002) introduced the discretized version of (2.6), 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

$$\gamma_{m,p} = \mathbb{E}\left[\log 2\pi \overline{I}_{n,k}^Z\right] \tag{2.8}$$

which is a centering constant, such that the expectation under the null hypothesis vanishes asymptotically. For the cases

1.) $d_0 = 0$, $m \geq 5$ and $p = 0$ or $p = 1$

2.) $d_0 > 0$, $m \geq 5$ and $p = 1$

Fay and Philippe (2002) 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 Section 3 for details], the statistic $\sqrt{K_n} S_n(\overline{I}_n^X, g(\cdot; \hat{d}_n, \hat{\theta}_n))$ converges weakly, that is

$$T_n = \sqrt{K_n} S_n\Big(\overline{I}_n^X, g(\cdot; \hat{d}_n, \hat{\theta}_n)\Big) \xrightarrow{d} \mathcal{N}\Big(0, \tau_0^2\Big), \tag{2.9}$$

where $(\hat{d}_n, \hat{\theta}_n)$ is any estimator of the "true" parameter (d_0, θ_0) satisfying

$$\big\|(\hat{d}_n, \hat{\theta}_n) - (d_0, \theta_0)\big\| = O_p\Big(\frac{1}{\sqrt{n}}\Big)$$

and the asymptotic variance in (2.9) is given by

$$\tau_0^2 := \operatorname{Var}\Big(2\pi \overline{I}_{n,k}^Z - \log\Big(2\pi \overline{I}_{n,k}^Z\Big)\Big).$$

For a discussion of the quantities $\gamma_{m,p}$ and τ_0^2 we refer to Hurvich et al. (2002). Note that these authors did not assume a Gaussian white noise process, but considered a general white noise process $(Z_t)_{t\in\mathbb{Z}}$ with 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 we will study the asymptotic properties of the statistic T_n if the null hypothesis is not satisfied. For the sake of simplicity, we restrict ourselves to the Gaussian case. The general case is briefly discussed in Remark 3.2.

3 Weak convergence under fixed alternatives

If the null hypothesis is not satisfied, then the minimum distance in (2.5) 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 (2.4) 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 Chen and Deo (2006)]. 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 [Whittle (1953)] 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 (2.7). In the case where the model is correctly specified, the asymptotic behaviour of the maximum likelihood estimator was investigated by Dahlhaus (1989). The Whittle estimator was investigated by Fox and Taqqu (1986) and Giraitis and Surgailis (1990) for Gaussian and linear processes, respectively. Recently, Chen and Deo (2006) derived the asymptotic properties of an estimator minimizing an approximation to the negative of the exact Gaussian likelihood [Whittle (1953)] 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 (2006) 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 (2006), 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 (2002) in the case of a misspecified model. For this purpose we define by

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

as the minimal distance between the true spectral density f and the parametric class defined in (2.4) with respect to the logarithmic contrast introduced in (2.6). Note that the null hypothesis (2.3) is satisfied if and only if $D(d_0, \theta_0) = 0$. We assume that $(X_t)_{t\in\mathbb{Z}}$ is a stationary process with linear representation (2.1) where the innovations $(Z_t)_{t\in\mathbb{Z}}$ define a Gaussian white noise process and the spectral density of $(X_t)_{t\in\mathbb{Z}}$ is given by (2.2).

Theorem. 3.1. *Let $(X_t)_{t\in\mathbb{Z}}$ be a stationary process with linear representation (2.1) and Gaussian white noise $(Z_t)_{t\in\mathbb{Z}}$, $d_1 \in (0, 1/2)$, $p = 1$, $m \geq 5$, $d_1 - d_0 < 1/4$, 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) $\int_{-\pi}^{\pi} \log g^*(\lambda; \theta)\, d\lambda = 0$ *for all* $\theta \in \Theta$.

If $n \to \infty$, then

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

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 (2.2) and the asymptotic variance is given by

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

with

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

Proof. Recalling the definition of the statistic T_n in (2.9) we introduce the decomposition

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

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

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

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

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

respectively. In the Appendix we will show that

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

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

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

where $\overline{I}_{n,k}^Z$ 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

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

Observing the approximation

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

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

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

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

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

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

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

Observing the approximation

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

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

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

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. 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 (2002). These authors did not assume a Gaussian white noise in the linear representation (2.1). This assumption was made here for the sake of transparent presentation and Theorem 3.1 remains valid in the general case, 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

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

and ζ is a $2(m+p)-$dimensional standard Gaussian vector. Note that $\alpha_{m,p}$ is the same as in the asymptotic variance under the null hypothesis in Fay and Philippe (2002).

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 (2003).

(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

$$\left[0, \hat{S}_n\left(\overline{I}_n^X, g(\cdot; \hat{d}_n, \hat{\theta}_n)\right) + \frac{\hat{\tau}_\Delta}{\sqrt{K_n}}\, u_{1-\alpha}\right]$$

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 $u_{1-\alpha}$ denotes the $(1-\alpha)$ 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 (2002) 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

$$P_{H_1}(\text{" } H_0 \text{ is rejected" }) \approx \Phi\left(\sqrt{K_n}\, \frac{D(d_0, \theta_0)}{\tau_\Delta} - \frac{\tau_0}{\tau_\Delta}\, u_{1-\alpha}\right),$$

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.

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

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

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

$$f(\lambda) = \frac{1}{2\pi}|1 - 0.1e^{i\lambda}|^2|1 - e^{i\lambda}|^{-2\cdot 0.4}.$$

Using the formula 3.631(8) in Gradshteyn and Ryzhik (1980), we approximately calculate d_0 and $D(d_0)$ as 0.3400325 and 0.003725739, respectively. We generated 5000 replications of the process for sample

	$n = 100$	$n = 200$	$n = 500$	$n = 1000$
0.8	0.6822	0.7244	0.7846	0.7966
0.9	0.9076	0.896	0.9164	0.9122
0.95	0.9876	0.975	0.9682	0.9698

Table 1: *Simulated coverage probabilities of the asymptotic confidence intervals defined in Remark 3.3(1)*

sizes $n = 100, 200, 500$ and 1000 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

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

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 Table 1. We observe reasonable coverage probabilities in most cases. While the 90% confidence interval is already approximated accurately for the samples size $n = 100$, larger sample sizes are required for the 80% and 95% confidence interval.

4 Appendix: Technical details

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

4.1 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 (2002) to show that the difference is stochastically small, i.e.

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

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

4.2 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

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

where the second estimate follows from Lemma 2 in Hurvich et al. (2002). We note that by the central limit theorem

$$
\frac{1}{K_n}\sum_{k=1}^{K_n} \Big(2\pi \overline{I}_{n,k}^Z - 1\Big) = \frac{1}{K_n}\sum_{k=1}^{K_n} 2\pi \overline{I}_{n,k}^Z - 1 = O_p\Big(\frac{1}{\sqrt{n}}\Big). \tag{4.1}
$$

We will show at the end of this section that

$$
\sum_{k=1}^{K_n} \beta_{n,k}\Big(\frac{\overline{I}_{n,k}^X}{f(x_k)} - 1\Big) = O_p\Big(\frac{1}{\sqrt{n}}\Big), \tag{4.2}
$$

then the expansion of the function $\log(1+z) = z + o(z^2)$ yields with the estimates (4.2) and (4.1) (note that $\sum_{k=1}^{K_n} \beta_{n,k} = 1$)

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

Observing Lemma 11 in Hurvich et al. (2002) we have

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

which yields

$$\text{(4.3)} \qquad \sum_{k=1}^{K_n} \beta_{n,k} \Big(\frac{\overline{I}_{n,k}^X}{f(x_k)} - 2\pi \overline{I}_{n,k}^Z \Big) = o_p\Big(\frac{1}{\sqrt{n}}\Big).$$

(note that $d_1 - d_0 < \frac{1}{4}$ by assumption). Therefore the assertion in (3.8) follows from (4.2) and (4.3). We conclude this section with a proof of the statement (4.2) which is obtained observing the decomposition

$$\sum_{k=1}^{K_n} \beta_{n,k} \Big(\frac{\overline{I}_{n,k}^X}{f(x_k)} - 1 \Big) = \sum_{k=1}^{K_n} \beta_{n,k} \Big(\frac{\overline{I}_{n,k}^X}{f(x_k)} - 2\pi \overline{I}_{n,k}^Z \Big) + \sum_{k=1}^{K_n} \beta_{n,k} \Big(2\pi \overline{I}_{n,k}^Z - 1 \Big) = O_p\Big(\frac{1}{\sqrt{K_n}}\Big)$$

where the last estimate follows again from (4.3) and a straightforward application of Chebyshev's inequality.

4.3 Proof of (3.9)

Observing the definition (3.6) we decompose C_n as follows

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

where

$$\text{(4.4)} \qquad C_n^{(1)} = \log\Big(\frac{1}{K_n} \sum_{k=1}^{K_n} \frac{\overline{I}_{n,k}^X}{g(x_k; \hat{\Gamma}_n)} \Big) - \log\Big(\frac{1}{K_n} \sum_{k=1}^{K_n} \frac{\overline{I}_{n,k}^X}{g(x_k; \Gamma_0)} \Big),$$

$$\text{(4.5)} \qquad C_n^{(2)} = \frac{1}{K_n} \sum_{k=1}^{K_n} \log \frac{g(x_k; \hat{\Gamma}_n)}{g(x_k; \Gamma_0)},$$

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

$$\text{(4.6)} \qquad C_n^{(j)} = o_p\Big(\frac{1}{\sqrt{n}}\Big), \ j = 1, 2.$$

For a proof of (4.6) 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

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

An extension of Theorem 2, Lemma 2 and 3 in Chen and Deo (2006) to the objective function (3.1) yields

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

and assertion (4.6) follows in the case $j = 1$.
In order to prove the statement in the case $j = 2$ we recall the definition in (4.5) and obtain by a Taylor expansion

$$\begin{aligned} C_n^{(2)} &= \frac{1}{K_n}\sum_{k=1}^{K_n}\Big\{\frac{1}{g(x_k;\Gamma_0)}\frac{\partial g(x_k;\Gamma_0)}{\partial\Gamma}\Big(\hat{\Gamma}_n - \Gamma_0\Big)\Big\} + O_p\Big(\frac{1}{n}\Big) \\ &= O_p\Big(\frac{1}{\sqrt{n}}\Big)\frac{1}{K_n}\sum_{k=1}^{K_n}\Big\{\frac{1}{g(x_k;\Gamma_0)}\frac{\partial g(x_k;\Gamma_0)}{\partial\Gamma}\Big\} + o_p\Big(\frac{1}{\sqrt{n}}\Big) \end{aligned} \tag{4.7}$$

where we have again used an extension of Theorem 2 in Chen and Deo (2006) to the loss function (3.1). From the assumption $g(\lambda;\Gamma) \in \mathcal{F}_0$ we have

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

for all $\Gamma \in D \times \Theta$, which implies (observing the symmetry of the function g) that the sum in (4.7) converges to 0 (a.s.). This proves the statement (4.6) 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. 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

Beran, J. (1992). A goodness-of-fit test for time series with long-range dependence. *Journal of the Royal Statistical Society, Ser. B*, 54(3):749–760.

Beran, J. (1994). *Statistics for Long-Memory Processes.* Chapman and Hall, New York.

Chen, W. W. and Deo, R. S. (2004). A generalized Portmanteau goodness-of-fit test for time series models. *Econometric Theory*, 20:382–416.

Chen, W. W. and Deo, R. S. (2006). Estimation of misspecified long-memory models. *Journal of Econometrics*, 134(1):257–281.

Dahlhaus, R. (1989). Efficient parameter estimation for self-similar processes. *The Annals of Statistics*, 17(4):1749–1766.

Deo, R. S. and Chen, W. W. (2000). On the integral of the squared periodogram. *Stochastic Processes and their Applications*, 85:159–176.

Dette, H. (1999). A consistent test for the functional form of a regression based on a difference of variance estimators. *Annals of Statistics*, 27:1012–1040.

Dette, H. (2002). A consistent test for heteroscedasticity in nonparametric regression based on the kernel method. *Journal of Statistical Planning and Inference*, 103:311–329.

Dette, H. and Munk, A. (2003). Some methodological aspects of validation of models in nonparametric regression. *Statistica Neerlandica*, 57:207–244.

Dette, H. and Spreckelsen, I. (2003). A note on a specification test for time series models based on spectral density estimation. *Scandinavian Journal of Statistics*, 30:481–491.

Fay, G. and Philippe, A. (2002). Goodness-of-fit test for long range dependent processes. *ESAIM: Probability and Statistics*, 6:239–258.

Fox, R. and Taqqu, M. S. (1986). Large-sample properties of parameter estimates for strongly dependent stationary gaussian time series. *Annals of Statistics*, 14:517–532.

Giraitis, L. and Surgailis, D. (1990). A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotical normality of Whittle's estimate. *Probability Theory and Related Fields*, 86(1):87–104.

Gradshteyn, I. and Ryzhik, I. (1980). *Table of Integrals, Series, and Products.* Academic Press, New York.

Granger, C. (1980). Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics*, 14(2):227–238.

Granger, R. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1(1):15–29.

Greene, M. and Fielitz, B. (1977). Long-term dependence in common stock returns. *Journal of Financial Economics*, 4(3):339–349.

Hosking, J. (1981). Fractional differencing. *Biometrika*, 68(1):165–176.

Hurvich, C., Moulines, E., and Soulier, P. (2002). The FEXP estimator for potentially non-stationary linear time series. *Stochastic Processes and their Applications*, 97(2):307–340.

Koutsoyiannis, D., Makropoulos, C., Langousis, A., Baki, S., Efstratiadis, A., Christofides, A., Karavokiros, G., and Mamassis, N. (2009). *HESS* opinions: Climate, hydrology, energy, water: recognizing uncertainty and seeking sustainability. *Hydrology and Earth System Sciences*, 13:247–257.

Mokkadem, A. (1997). A measure of information and its applications to test for randomness against ARMA alternatives and to goodness-of-fit test. *Stochastic Processes and their Applications*, 72(2):145–159.

Paparoditis, E. (2000). Spectral density based goodness-of-fit tests for time series models. *Scandinavian Journal of Statistics*, 27:143–176.

Park, K. and Willinger, W. (2000). Self-similar network traffic: An overview. In Park, K. and Willinger, W., editors, *Self-Similar Network Traffic and Performance Evaluation*, pages 1–39. Wiley Interscience, New York.

Stroe-Kunold, E., Stadnytska, T., Werner, J., and Braun, S. (2009). Estimating long-range dependence in time series: An evaluation of estimators implemented in R. *Behavior Research Methods*, 41:909–923.

Whittle, P. (1953). Estimation and information in stationary time series. *Arkiv for Matematik*, 1:423–434.