Goodness-of-fit tests for long memory moving average marginal density

Hira L. Koul, Nao Mimoto & Donatas Surgailis

Metrika
International Journal for Theoretical and Applied Statistics

ISSN 0026-1335
Volume 76
Number 2

Metrika (2013) 76:205-224
DOI 10.1007/s00184-012-0383-y

Springer

Metrika (2013) 76:205–224
DOI 10.1007/s00184-012-0383-y

Goodness-of-fit tests for long memory moving average marginal density

Hira L. Koul · Nao Mimoto · Donatas Surgailis

Received: 20 May 2011 / Published online: 18 February 2012

Abstract This paper addresses the problem of fitting a known density to the marginal error density of a stationary long memory moving average process when its mean is known and unknown. In the case of unknown mean, when mean is estimated by the sample mean, the first order difference between the residual empirical and null distribution functions is known to be asymptotically degenerate at zero, and hence can not be used to fit a distribution up to an unknown mean. In this paper we show that by using a suitable class of estimators of the mean, this first order degeneracy does not occur. We also investigate the large sample behavior of tests based on an integrated square difference between kernel type error density estimators and the expected value of the error density estimator based on errors. The asymptotic null distributions of suitably standardized test statistics are shown to be chi-square with one degree of freedom in both cases of the known and unknown mean. In addition, we discuss the consistency and asymptotic power against local alternatives of the density estimator based test in the case of known mean. A finite sample simulation study of the test based on residual empirical process is also included.

Keywords Kernel density estimator · Chi square distribution · Residual empirical process

Research was supported in part by the NSF DMS Grant 07-04130 and the Grant MIP-11155 from the Research Council of Lithuania.

H. L. Koul (✉) · N. Mimoto
Michigan State University, East Lansing, MI, USA
e-mail: koul@stt.msu.edu

D. Surgailis
Vilnius University, Vilnius, Lithuania

1 Introduction

The problem of fitting a parametric family of distributions to a probability distribution, known as the goodness-of-fit testing problem, is classical in statistics, and well studied when the underlying observations are i.i.d. See, for example, Durbin (1973, 1975), Khmaladze (1979, 1981), D'Agostino and Stephens (1986), among others.

A discrete time stationary stochastic process with finite variance is said to have long memory if its autocorrelations tend to zero hyperbolically in the lag parameter, as the lag tends to infinity, but their sum diverges. The importance of these processes in econometrics, hydrology and other physical sciences is abundantly demonstrated in the works of Beran (1992, 1994), Baillie (1996), Dehling et al. (2002) and Doukhan et al. (2003), and the references therein.

Consider the moving average time series

$$X_j = \sum_{i=0}^{\infty} b_i \zeta_{j-i}, \quad j \in \mathbb{Z} := \{0, \pm 1, \pm 2, \ldots\}, \tag{1.1}$$

where $\zeta_s, s \in \mathbb{Z}$ are i.i.d. with zero mean and unit variance. The constants $\{b_j, j \in \mathbb{Z}\}$ satisfy $b_k = 0, k < 0, \ b_0 = 1$ and

$$b_j \sim c j^{-(1-d)} \qquad \text{as } j \to \infty, \text{ for some } 0 < c < \infty \quad \text{and} \quad 0 < d < 1/2. \tag{1.2}$$

One can verify $\{X_j\}$ is a stationary process, $EX_0 = 0$ and $\mathrm{Cov}(X_0, X_j) \sim c^2 B(d, 1-2d) j^{-(1-2d)}$, as $j \to \infty$, where $B(a, b) := \int_0^1 x^{a-1}(1-x)^{b-1} dx, a > 0, b > 0$. Consequently the process $\{X_j, j \in \mathbb{Z}\}$, has long memory.

Now, let F and f denote the marginal distribution and density functions of X_0 and F_0 be a known distribution function (d.f.) with density f_0. The problem of interest is to test the hypothesis

$$\mathcal{H}_0 : f = f_0 \qquad \text{vs.} \qquad \mathcal{H}_1 : f \neq f_0.$$

A motivation for this problem is that often in practice one uses inference procedures that are valid under the assumption of $\{X_j\}$ being Gaussian. If one were to reject the hypothesis that the marginal error distribution is Gaussian, then the validity of the use of such inference procedures would be questionable.

Now, let throughout the paper, Z denote a $\mathcal{N}(0, 1)$ r.v., and define

$$\widehat{F}_n(x) := n^{-1} \sum_{j=1}^{n} I(X_j \le x), \quad x \in \mathbb{R}, \quad \kappa^2(\theta) := c^2 B(d, 1-2d)/d(1+2d),$$
$$\theta := (c, d)', \ \|f_0\|_\infty := \sup_{x \in \mathbb{R}} f_0(x).$$

A test of $\mathcal{H}_0$ is the Kolmogorov–Smirnov test based on $D_n := \sup_{x\in\mathbb{R}} |\widehat{F}_n(x) - F_0(x)|$. Giraitis et al. (1996) observed, under some conditions, that

$$\mathcal{D}_n(\theta) := \frac{n^{1/2-d} D_n}{\kappa(\theta)\|f_0\|_\infty} \to_D |Z|. \tag{1.3}$$

Here, and in the sequel, $\to_D$ stands for the convergence in distribution.

Let $\widehat{c}$, $\widehat{d}$ be consistent and $\log(n)$ consistent estimators of c and d, under $\mathcal{H}_0$, respectively, and set $\widehat{\theta} := (\widehat{c}, \widehat{d}\,)$. Let z_α be $100(1-\alpha)$th percentile of $\mathcal{N}(0, 1)$ distribution. From (1.3), we readily obtain that the test that rejects $\mathcal{H}_0$ whenever $\mathcal{D}_n(\widehat{\theta}\,) \geq z_{\alpha/2}$, is of asymptotic size α. Thus this test is relatively easy to implement, relative to the corresponding test in the i.i.d. case where one must use the distribution of the supremum of Brownian bridge to obtain critical values.

Before proceeding further, recall that in the case of dependent short memory stationary observations (satisfying some mixing conditions; see, e.g., Dedecker et al. 2007), the empirical process $n^{1/2}(\widehat{F}_n(x) - F_0(x))$ weakly converges to a centered Gaussian process $\{W(x), x \in \mathbb{R}\}$ with covariance $\mathrm{Cov}(W(x), W(y)) = \sum_{j \in \mathbb{Z}} \mathrm{Cov}(I(X_0 \le x), I(X_j \le y))$. For linear processes in (1.1) with summable weights $\sum_{j=0}^{\infty} |b_i| < \infty$, the last result holds both under short memory ($\sum_{i=0}^{\infty} b_i \neq 0$) and negative memory ($b_j \sim c j^{-(1-d)}$, $-1/2 < d < 0$, $\sum_{i=0}^{\infty} b_i = 0$) assumptions. See Doukhan and Surgailis (1998). In particular, under short ($d = 0$) or negative ($-1/2 < d < 0$) memory, the test based on $\mathcal{D}_n(\theta)$ in (1.3), as well as other tests discussed in this paper, is generally inconsistent and the limit distribution of $n^{1/2} D_n$ depends on the probability structure of $\{X_j\}$ (via $\sup_x |W(x)|$) in a complicated fashion. On the other hand, if (estimated) d is suspected to be close to zero, a visual inspection of $\widehat{F}_n(x) - F_0(x)$, $x \in \mathbb{R}$ might help the practitioner to decide between the two possibilities $d = 0$ and $d > 0$: in the former case, the empirical process behaves as a Gaussian process fluctuating around zero, while in the latter case, it resembles a signed probability density staying away from zero.

Now consider the problem of fitting f_0 to f up to an unknown location parameter, i.e. the problem of interest is to test

$$\mathcal{H}_{0\ell oc} : f(x) = f_0(x - \mu), \quad \forall\, x \in \mathbb{R}, \text{ for some } \quad \mu \in \mathbb{R}, \text{ vs.}$$
$$\mathcal{H}_{1,\ell oc} : \mathcal{H}_{0\ell oc} \quad \text{is not true.}$$

This is equivalent to stipulating that we observe Y_i's from the model $Y_i = \mu + X_i$, for some $\mu \neq 0$, and wish to test $\mathcal{H}_0$ based on Y_i, $1 \le i \le n$. Let $\bar{F}_n$ be the empirical d.f. based on $Y_i - \bar{Y}$, $1 \le i \le n$, and $\bar{D}_n := \sup_x |\bar{F}_n(x) - F_0(x)|$, where $\bar{Y}$ is the empirical mean. An interesting observation made in Koul and Surgailis (2002, 2010) is that in this case the null weak limit of the first order difference between the residual empirical process and the null model is degenerate at zero, i.e., $n^{1/2-d} \bar{D}_n \to_p 0$, and hence it can not be used asymptotically to test for $\mathcal{H}_{0\ell oc}$.

This is partly due to the uniform reduction principle that says that the weak limit of empirical process $n^{1/2-d}(\widehat{F}_n - F)$ is a degenerate process, and partly due to the choice of the estimator $\bar{Y}$ of μ. In this paper we first provide a class of estimators of μ for which the weak limit of the process $n^{1/2-d}(\bar{F}_n - F_0)$ under $\mathcal{H}_{0\ell oc}$ is a non-degenerate Gaussian distribution. In addition, we investigate tests based on kernel density estimators testing for both hypotheses $\mathcal{H}_0$ and $\mathcal{H}_{0\ell oc}$, using both the proposed class of estimators of μ and $\bar{Y}$.

To be precise, let φ be a piece-wise continuously differentiable function on [0, 1] and define

$$\widetilde{Y} := \frac{1}{n}\sum_{i=1}^{n} Y_i\left[1+\varphi\left(\frac{i}{n}\right)\right] = \mu(1+\bar{\varphi}_n)+\bar{X}+\bar{W}, \qquad \widetilde{\delta} := \widetilde{Y}-\mu, \tag{1.4}$$

$$\bar{X} := \frac{1}{n}\sum_{i=1}^{n} X_i, \quad \bar{W} := \frac{1}{n}\sum_{i=1}^{n} X_i\varphi\left(\frac{i}{n}\right), \quad \bar{\varphi}_n := \frac{1}{n}\sum_{i=1}^{n}\varphi\left(\frac{i}{n}\right), \quad \bar{\varphi} := \int_0^1 \varphi(u)du.$$

Under the assumed conditions of φ, $\bar{\varphi}_n \to \bar{\varphi}$, and $\bar{\varphi}_n - \bar{\varphi} = O(n^{-1})$. By the Ergodic Theorem, $\bar{X} \to EX_0 = 0$, a.s. Also, by Lemma 2.1, $\bar{W} = o_p(1)$. Hence, if $\bar{\varphi} = 0$, $\widetilde{Y} \to_p \mu$.

Let

$$\widetilde{F}_n(x) := n^{-1}\sum_{i=1}^{n} I(Y_i - \widetilde{Y} \le x) = \widehat{F}_n(x+\widetilde{\delta}), \quad \widetilde{D}_n(x) := \widetilde{F}_n(x) - F_0(x), \quad x \in \mathbb{R},$$
$$\widetilde{D}_n := \sup_{x\in\mathbb{R}} |\widetilde{D}_n(x)|.$$

We show that under $\mathcal{H}_{0\ell oc}$ and $\bar{\varphi} = 0$, $n^{1/2-d}\widetilde{D}_n \to_p \nu(\theta)\|f_0\|_\infty|Z|$, where $\nu(\theta)$ is as in Lemma 2.1; see Theorem 2.1. Consequently, the test that rejects $\mathcal{H}_{0\ell oc}$ whenever $\widetilde{\mathcal{D}}_n := \{\nu(\widetilde{\theta})\|f_0\|_\infty\}^{-1} n^{1/2-\widetilde{d}}\,\widetilde{D}_n > z_{\alpha/2}$ is asymptotically distribution free and of the asymptotic level α, where now $\widetilde{c}, \widetilde{d}$ are consistent and $\log(n)$-consistent estimators of c, d under $\mathcal{H}_{0\ell oc}$, and $\widetilde{\theta} := (\widetilde{c}, \widetilde{d}\,)'$.

Next, let K be a density kernel on $[-1, 1]$, $h \equiv h_n$ be bandwidth sequence, E_0 denote the expectation under $\mathcal{H}_0$, and define

$$\widehat{f}_n(x) := \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h}\right), \qquad \widetilde{f}_n(x) := \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-(Y_i-\widetilde{Y})}{h}\right), \quad x \in \mathbb{R},$$
$$T_n := \int \left(\widehat{f}_n(x) - E_0\widehat{f}_n(x)\right)^2 dx, \qquad \widetilde{T}_n := \int (\widetilde{f}_n(x) - E_0\widehat{f}_n(x))^2 dx.$$

Note that the choice $\varphi \equiv -1$ yields $\widetilde{Y} = 0$, and if one also has $\mu = 0$, then $\widetilde{f}_n(x) = \widehat{f}_n(x)$, $\widetilde{T}_n = T_n$. Statistics T_n and $\widetilde{T}_n$ are useful in testing for $\mathcal{H}_0$ and $\mathcal{H}_{0\ell oc}$, respectively.

It is shown in Theorem 2.2 below that the asymptotic null distribution of $n^{1-2d}T_n/\kappa_1$ is that of Z^2, a chi-square r.v. with one degree of freedom, where κ_1 is defined at (2.10) below. Surprisingly, a similar result holds for $\widetilde{T}_n$ provided $\bar{\varphi} = 0$ and φ is not identically zero.

The case of $\varphi \equiv 0$ makes $\widetilde{Y} = \bar{Y}$. The corresponding test statistic $\widetilde{T}_n$ then also has the first order degenerate behavior analogous to the residual empirical process. In this case we show that for $1/4 < d < 1/2$, $n^{2(1-2d)}\widetilde{T}_n$ has a nonstandard asymptotic null distribution; see Theorem 2.3 below.

We also discuss consistency and asymptotic power against certain local alternatives of the T_n-test. Let f be an alternate density that is differentiable and for which

$$m := \|f - f_0\|^2 > 0. \tag{1.5}$$

Let f' denote the first derivative of f and define

$$\Delta := \int f'(x)(f(x) - f_0(x))dx, \tag{1.6}$$

$$\begin{aligned} m(h) &:= \int \left\{ \int (f(x-uh) - f_0(x-uh))K(u)du \right\}^2 dx \\ &= m + o(1), \quad h \to 0. \end{aligned} \tag{1.7}$$

We show that under some conditions, $n^{1/2-d}(T_n - m(h)) \to_D \mathcal{N}(0, 4\kappa^2(\theta)\Delta^2)$, where $\kappa^2(\theta)$ is as in (1.3). Consequently, the T_n-test is consistent against all differentiable fixed densities f satisfying (1.5). We also investigate asymptotic distribution of T_n under a sequence of local alternatives; see Theorem 2.5.

2 Main results

In this section we shall give precise conditions under which the previously stated results are proved. We do this in the two subsections. The first one deals with $\mathcal{H}_0$ and $\mathcal{H}_{0\ell oc}$ while the second subsection deals with the asymptotic power analysis of the T_n-test under fixed and local alternatives.

2.1 Asymptotic null distribution of $\widetilde{D}_n$, T_n and $\widetilde{T}_n$

Our proofs here are based on some results derived in Koul and Surgailis (2002) (KS). Accordingly, we first specify the needed assumptions. Let ζ be a copy of ζ_0. Following KS, assume that the innovation distribution satisfies

$$E|\zeta|^3 < C, \tag{2.1}$$

$$|Ee^{\mathbf{i}u\zeta}| \le C(1+|u|)^{-\delta}, \qquad \text{for some } 0 < C < \infty,\ \delta > 0,\ \forall\, u \in \mathbb{R}. \tag{2.2}$$

Under (2.2), it is shown in KS that the d.f. F of X_0 is infinitely differentiable and for some universal positive constant C,

$$(f(x), |f'(x)|, |f''(x)|, |f'''(x)|) \le C(1+|x|)^{-2}, \quad \forall\, x \in \mathbb{R}, \tag{2.3}$$

where f'', f''' are the second and third derivatives of f, respectively. This fact in turn clearly implies f and these derivatives are square integrable.

About the kernel K, we assume that it is a symmetric density on $[-1, 1]$, with unit variance, vanishing off $(-1, 1)$, differentiable with derivative K' satisfying

$$\int |vK'(v)|dv < \infty. \tag{2.4}$$

For the bandwidth h, we assume

$$h \to 0, \quad \min(n^{2d}h,\ n^{1-2d}h) \to \infty, \quad \text{as } n \to \infty. \tag{2.5}$$

In the sequel, all limits are taken as $n \to \infty$, unless specified otherwise.

First, we recall the following results about $\bar{X}$ and $\bar{W}$ from Davydov (1970) and Koul and Surgailis (2000, Lemma 2.4 (iii)), respectively. Let, as before, $\theta = (c, d)'$, and

$$\nu^2(\theta) := c^2 B(d, 1-2d) \int_0^1 \int_0^1 \varphi(u)\varphi(v)|u-v|^{2d-1} du dv.$$

Lemma 2.1 *Let $\varphi(x), x \in [0, 1]$ be a piecewise continuously differentiable function and suppose $\{X_j\}$ satisfy (1.1) and (1.2). Then,*

$$\kappa^{-1}(\theta) n^{1/2-d} \bar{X} \to_D Z, \quad n^{1/2-d} \bar{W} \to_D \nu(\theta) Z. \tag{2.6}$$

We shall now describe the asymptotic null distribution of the statistic $\widetilde{D}_n$.

Theorem 2.1 *Suppose (1.1), (1.2), (2.1), (2.2) hold. Let $\varphi(x), x \in [0, 1]$ be a piecewise continuously differentiable function satisfying*

$$\bar{\varphi} = 0. \tag{2.7}$$

Then, under $\mathcal{H}_{0\ell oc}$,

$$n^{1/2-d} \sup_{x \in \mathbb{R}} \left| \widetilde{F}_n(x) - F_0(x) - \bar{W} f_0(x) \right| = o_p(1).$$

Consequently, $n^{1/2-d} \widetilde{D}_n \to_D \nu(\theta) \|f_0\|_\infty |Z|$.

Proof By Lemma 2.1, $n^{1/2-d}|\tilde{\delta}| = O_p(1)$. This fact together with results from KS imply that under the assumed conditions, and under $\mathcal{H}_{0\ell oc}$,

$$n^{1/2-d} \sup_{x \in \mathbb{R}} |\widetilde{F}_n(x) - \widehat{F}_n(x) - f_0(x)\,\tilde{\delta}| = o_p(1),$$
$$n^{1/2-d} \sup_{x \in \mathbb{R}} \left| \widehat{F}_n(x) - F_0(x) + f_0(x)\,\bar{X} \right| = o_p(1).$$

Hence, the decomposition $\widetilde{F}_n(x) - F_0(x) = \widetilde{F}_n(x) - \widehat{F}_n(x) + \widehat{F}_n(x) - F_0(x)$ yields

$$n^{1/2-d} \sup_{x\in\mathbb{R}} \left| \widetilde{F}_n(x) - F_0(x) - \left(\widetilde{\delta} - \bar{X}\right) f_0(x) \right| = o_p(1).$$

By assumption on φ and by (2.7), $\bar{\varphi}_n = O(n^{-1}) = o(n^{d-1/2})$. This fact and the identity $\widetilde{\delta} - \bar{X} = \bar{W} + \mu\bar{\varphi}_n$ now readily yields the theorem.

Remark 2.1 Let $\widetilde{c}$, $\widetilde{d}$ be consistent and $\log(n)$-consistent estimators of c, d, respectively, under $\mathcal{H}_{0\ell oc}$ and let $\widetilde{\theta} := (\widetilde{c}, \widetilde{d})'$. A consequence of the above theorem is that the test that rejects $\mathcal{H}_{0\ell oc}$ whenever $\widetilde{\mathcal{D}}_n := \left\{\nu(\widetilde{\theta})\|f_0\|_\infty\right\}^{-1} n^{1/2-\widetilde{d}}\widetilde{D}_n > z_{\alpha/2}$ is of the asymptotic level α.

Next, we turn to obtaining the asymptotic distributions of T_n and $\widetilde{T}_n$. Let E denote the expectation when density of X_0 is f, and define

$$\widetilde{S}_n(a) := \int (\widetilde{f}_n(x) - E\widehat{f}_n(x + a\bar{\varphi}))^2 dx, \quad a \in \mathbb{R}. \tag{2.8}$$

Note that when $f = f_0$, $\widetilde{S}_n(0) = \widetilde{T}_n$. Moreover, if in addition, $\bar{\varphi} = 0$, then $\widetilde{S}_n(a) = \widetilde{T}_n$, for all $a \in \mathbb{R}$.

First, we state a general result about the asymptotic distribution of $\widetilde{S}_n(\mu)$. Throughout, for any square integrable function g, $\|g\|^2 := \int_{\mathbb{R}} g^2(x)dx$.

Theorem 2.2 *Suppose* (1.1), (1.2), (2.1)–(2.5) *hold. Let* $\varphi(x), x \in [0, 1]$ *be a piece wise continuously differentiable function. Then,*

$$n^{1-2d}\widetilde{S}_n(\mu) \to_D \kappa_\varphi^2(\theta)Z^2, \qquad \kappa_\varphi^2(\theta) := \nu^2(\theta)\int \left(f'(x + \mu\bar{\varphi})\right)^2 dx. \tag{2.9}$$

Theorem 2.2 implies the following corollary about the limit distributions of the test statistics T_n and $\widetilde{T}_n$ for testing $\mathcal{H}_0$ and $\mathcal{H}_{0\ell oc}$.

Corollary 2.1 *Under the conditions of Theorem* 2.2 *the following hold.*
(i) *Suppose* $\mu = 0$, $\varphi(x) \equiv -1$, *and* $\mathcal{H}_0$ *holds. Then,* $\widetilde{S}_n(0) = T_n$, *and*

$$n^{1-2d}T_n \to_D \sigma^2(\theta)Z^2, \qquad \sigma^2(\theta) := \kappa^2(\theta)\|f_0'\|^2. \tag{2.10}$$

(ii) *Suppose* $\bar{\varphi} = 0$ *and* $\mathcal{H}_{0loc}$ *holds. Then,* $\widetilde{S}_n(\mu) = \widetilde{T}_n$, *and*

$$n^{1-2d}\widetilde{T}_n \to_D \kappa_2^2(\theta)Z^2, \qquad \kappa_2^2(\theta) := \nu^2(\theta)\|f_0'\|^2. \tag{2.11}$$

It thus follows that the test that rejects $\mathcal{H}_0$ whenever $n^{1-2\widehat{d}}T_n/\sigma^2(\widehat{\theta}) > k_\alpha$ has the asymptotic size α, $0 < \alpha < 1$, where k_α is the $(1-\alpha)100$th percentile of the χ_1^2 distribution. Similarly, the test that rejects $\mathcal{H}_{0\ell oc}$, whenever $n^{1-2\widetilde{d}}(\widetilde{T}_n/\nu^2(\widetilde{\theta})\|f_0'\|^2) > k_\alpha$, has the asymptotic size α, $0 < \alpha < 1$. Here, $\widehat{c}$, $\widehat{d}$, $\widehat{\theta}$, $\widetilde{c}$, $\widetilde{d}$ and $\widetilde{\theta}$ are as before.

Remark 2.2 A particularly simple choice of function φ in Corollary 2.1 is $\varphi(x) = I(0 < x \leq 1/2) - I(1/2 < x \leq 1)$. It satisfies $\bar{\varphi} = 0$ and leads to

$$\widetilde{Y} := \bar{Y} + \frac{1}{n}\sum_{i=1}^{[n/2]} Y_i - \frac{1}{n}\sum_{i=[n/2]+1}^{n} Y_i = \frac{2}{n}\sum_{i=1}^{[n/2]} Y_i.$$

Moreover, in this case

$$\nu^2(\theta) = c^2 \frac{B(d, 1-2d)}{d(1+2d)}\left(2^{1-2d} - 1\right). \tag{2.12}$$

In other words the corresponding $\widetilde{T}_n$-test uses the first half sample to estimate μ and then the entire set of residuals to perform the test.

The above example can help to understand why a partial centring by $\widetilde{Y}$ can be better than the natural or naive centring by the empirical mean, for testing $\mathcal{H}_{0loc}$. From the uniform reduction principle, we have $\widehat{F}_n(x) - F_0(x) = -f_0(x)\bar{X} + o_p(\bar{X})$, uniformly in x, and therefore,

$$\begin{aligned}\widetilde{F}_n(x) - F_0(x) &= [\widehat{F}_n(x+(\widetilde{Y}-\mu)) - F_0(x+(\widetilde{Y}-\mu))] + [F_0(x+(\widetilde{Y}-\mu)) - F_0(x)] \\ &= -f_0(x + (\widetilde{Y} - \mu))\bar{X} + f_0(x)(\widetilde{Y} - \mu) + o_p(\bar{X}) \\ &= f_0(x)(\widetilde{Y} - \mu - \bar{X}) + o_p(\bar{X}), \qquad \text{uniformly in } x.\end{aligned}$$

In other words, in the case of centering by the empirical mean $\bar{Y}$, the term $f_0(x)(\bar{Y} - \mu) = f_0(x)\bar{X}$ (coming from Taylor's expansion of F_0) completely cancels with the main expansion term $-f_0(x)\bar{X}$ of the empirical process $\widehat{F}_n - F_0$. On the other hand, in the case of the partial centring, such a cancelation need not occur, leading to the main term $f_0(x)(\widetilde{Y} - \mu - \bar{X}) = f_0(x)W$, which has a non-degenerate Gaussian limit. Indeed, since partial sums of $\{X_j\}$, normalized by $n^{d-1/2}$, approach fractional Brownian motion $\kappa(\theta)B_{d+1/2}$; see Theorem 2 in Davydov (1970), the distribution of $n^{1/2-d}(\widetilde{Y} - \mu - \bar{X})$ in the above example tends to Gaussian limit $\kappa(\theta)(2B_{d+1/2}(1/2) - B_{d+1/2}(1)) \sim \mathcal{N}(0, \nu^2(\theta))$, with $\nu^2(\theta)$ given in (2.12).

Remark 2.3 It is of interest to contrast the result of Corollary 2.1 with what is available under independence. When observations are i.i.d., Bickel and Rosenblatt (1973) proved, under $\mathcal{H}_0$ and the second order differentiability of f_0, that

$$\begin{aligned} n\sqrt{h}\left(T_n - \frac{1}{nh}\int K^2(t)dt\right) &\to_D \mathcal{N}(0, \tau^2), \\ \tau^2 := 2\int f_0^2(x)\int\left(\int K(x-u)K(u)du\right)^2 dx. \end{aligned}$$

Bachmann and Dette (2005) proved this result requiring f_0 to be only continuous and square integrable.

The first thing one notices missing under long memory is the centering of T_n. Under long memory there is no asymptotic bias in T_n. Secondly, the normalization n^{1-2d} does not depend on the window width. Under long memory the role of $n^{1/2}$ is played by $n^{1/2-d}$, so one sees the normalization is simply square of this. Finally, unlike in the i.i.d. case, asymptotic variance of standardized T_n does not depend on K in the present set up. Of course to implement the test under long memory, one has to have consistent and $\log(n)$-consistent estimators of c and d, respectively. But under fairly general conditions local Whittle estimators of c and d are known to satisfy these constraints. See, Dalla et al. (2006).

Typically if one uses $h \propto n^{-a}$, then (2.5) will be satisfied as long $a/2 < d < (1-a)/2$. For example, if $a = 1/5$, then (2.9) holds for all $0.1 < d < 0.4$.

The case $\varphi \equiv 0$ leads to $\widetilde{Y} = \bar{Y}$ and the trivial limit in (2.11). In this case the limit distribution of $\widetilde{T}_n$ is obtained from the second order expansion of the empirical distribution function in KS as described in the following theorem.

Theorem 2.3 *Assume the same conditions as in Theorem 2.2, with exception of (2.5), and let $\varphi \equiv 0$. Moreover, assume $1/4 < d < 1/2$ and*

$$h \to 0, \qquad \min(n^{4d-1}h,\, n^{1-2d}h) \to \infty, \quad \text{as} \quad n \to \infty. \tag{2.13}$$

Then, $n^{2(1-2d)}\widetilde{T}_n \to_D \|f_0''\|^2\{Z^{(2)} - 2^{-1}(Z^{(1)})^2\}^2$, where the r.v.'s $Z^{(2)}$ and $Z^{(1)}$ as defined in (4.11) below have a Rosenblatt and a Gaussian distribution, respectively.

2.2 Consistency and asymptotic power of T_n-test

Here we shall study the consistency and asymptotic power of the test based on T_n against some alternatives. Let f be a density of X_0 satisfying (1.5). Assume that the corresponding innovations in (1.1) satisfy (2.2) so that f also satisfies (2.3). Recall the definition of Δ from (1.6) and $m(h)$ from (1.7). Also, define

$$\tau := 2\kappa(\theta)\Delta.$$

Then, we have the following

Theorem 2.4 *Suppose (1.1), (1.2), (2.1)–(2.5) hold with density f of X_0 satisfying (1.5). Then, $n^{1/2-d}(T_n - m(h)) \to_D \mathcal{N}(0, \tau^2)$.*

This theorem is useful for discussing the consistency of the T_n-test. Consider the case $\Delta \neq 0$. Let P_A denoting the probability measure under the alternative (1.5). Then,

$$\begin{aligned} P_A\Big(n^{1-2\widehat{d}}T_n > \sigma^2(\widehat{\theta})k_\alpha\Big) &= P_A\Big(\frac{n^{1/2-\widehat{d}}T_n}{|\tau|} > \frac{\sigma^2(\widehat{\theta})k_\alpha}{n^{1/2-\widehat{d}}|\tau|}\Big) \\ &= P_A\Big(\frac{n^{1/2-\widehat{d}}(T_n - m(h))}{|\tau|} > \frac{\sigma^2(\widehat{\theta})k_\alpha}{n^{1/2-\widehat{d}}|\tau|} - \frac{n^{1/2-\widehat{d}}m(h)}{|\tau|}\Big). \end{aligned} \tag{2.14}$$

Clearly, the $\log(n)$-consistency of $\widehat{d}$ for d and consistency of $\widehat{c}$ for c implies $n^{2(d-\widehat{d})} \to_p 1$, $\sigma(\widehat{\theta})/\sigma(\theta) \to_p 1$, and by Theorem 2.4, $n^{1/2-\widehat{d}}(T_n - m(h))/|\tau| \to_D Z$. From these facts and the above relation, it readily follows that for every fixed f satisfying (1.5), $n^{1/2-\widehat{d}}|\tau| \to_p \infty$, $n^{1/2-\widehat{d}}m(h)/|\tau| \to_p \infty$, and hence, the power of the asymptotic level α T_n-test tends to 1.

Next, consider the case $\Delta = 0$. Then, by Theorem 2.4, and arguing as above,

$$P_A\Big(n^{1-2\widehat{d}}T_n > \sigma^2(\widehat{\theta})k_\alpha\Big) = P_A\Big(n^{1/2-\widehat{d}}(T_n - m(h)) > n^{\widehat{d}-1/2}\sigma^2(\widehat{\theta})k_\alpha - n^{1/2-\widehat{d}}m(h)\Big)$$
$$\approx P_A\Big(0 > -n^{1/2-\widehat{d}}m(h)\Big) \to 1$$

since $n^{1/2-\widehat{d}}m(h) \to_p \infty$, for all those f satisfying (1.5).

Summarizing, under the conditions of Theorem 2.4, the T_n-test is consistent against all those differentiable f's for which $m > 0$.

Next, we give a theorem that describes asymptotic distribution of T_n under certain sequences of local alternatives.

Theorem 2.5 *Let $f_n, n \geq 1$, be a sequence of densities on $\mathbb{R}$. Suppose $X_j \equiv X_{j,n} = \sum_{i=0}^{\infty} b_i \zeta_{j-i,n}$, $j \in \mathbb{Z}$, is a sequence of stationary MA processes with the coefficients b_j as in (1.2) and having marginal density f_n. Assume also that $\{\zeta_{s,n}, s \in \mathbb{Z}\}$ are standardized i.i.d. innovations satisfying (2.2) and (2.1) for each n with C, δ independent of n. Let ψ be a real valued square integrable and differentiable function on $\mathbb{R}$ satisfying*

$$\int |f_0'(x)\psi(x)|dx < \infty, \qquad n^{1-2d}m_n(h) \to \|\psi\|^2, \tag{2.15}$$
$$n^{1/2-d}\Delta_n \to \Delta_0 := \int f_0'(x)\psi(x)dx, \qquad \|f_n'\|^2 \to \|f_0'\|^2.$$

where $m_n(h)$, Δ_n are defined as in (1.6), (1.7), respectively, with f replaced by f_n.

Assume also (2.4) and (2.5) hold. Then,

$$n^{1-2d}(T_n - m_n(h)) \to_D -2Z\kappa^2(\theta)\Delta_0 + Z^2\sigma^2(\theta).$$

Recall $\sigma^2(\theta) = \kappa^2(\theta)\|f_0'\|^2$. Suppose $\widehat{c}, \widehat{d}$ continue to be consistent and $\log(n)$-consistent estimators of c, d under f_n. Then, from the above theorem asymptotic power of the asymptotic α-level T_n-test against the above sequence of alternative densities f_n is

$$P\left(-2Z\frac{\Delta_0}{\|f_0'\|^2} + Z^2 > k_\alpha - \frac{\|\psi\|^2}{\sigma^2(\theta)}\right).$$

Clearly, if $\Delta_0 = 0$, then this power is equal to $1 - G(k_\alpha - \|\psi\|^2/\sigma^2(\theta))$, where G is the d.f. of a χ_1^2 r.v.

Remark 2.4 **Another test** Suppose one bases test of $\mathcal{H}_0$ on the integrated square difference

$$\widehat{T}_n := \int \big(\widehat{f}_n(x) - f_0(x)\big)^2 dx.$$

Here we give sufficient conditions under which this test is equivalent to the test based on T_n under $\mathcal{H}_0$. To begin with note that

$$\widehat{T}_n = T_n + \int \big(E_0\widehat{f}_n(x) - f_0(x)\big)^2 dx + 2\int \big(\widehat{f}_n(x) - E_0\widehat{f}_n(x)\big)\big(E_0\widehat{f}_n(x) - f_0(x)\big)dx,$$

is a sum of T_n, bias term, and the cross product term.

Consider the bias term. Taylor expansion and symmetry of K with mean 0 imply

$$\begin{aligned} E_0\widehat{f}_n(x) - f_0(x) &= \int K(z) f_0(x - zh)dz - f_0(x) \\ &= \int K(z)\Big[f_0(x) + f_0'(x)(-zh) + 2^{-1} f_0''(x - \xi zh)(zh)^2\Big]dz - f_0(x) \\ &= (1/2)h^2 \int K(z) f_0''(x - \xi zh) z^2 dz, \end{aligned}$$

where $0 < \xi < 1$. Therefore, by the continuity and square integrability of f'',

$$\begin{aligned} & n^{1-2d} \int \big(E_0\widehat{f}_n(x) - f_0(x)\big)^2 dx \\ &= (1/4)\, n^{1-2d} h^4 \int \Big(\int f_0''(x - \xi zh) z^2 K(z) dz\Big)^2 dx \\ &= (1/4)\, n^{1-2d} h^4 \int\int\int f_0''(x - \xi zh) f_0''(x - \xi sh) s^2 z^2 K(s) K(z) ds\, dz\, dx \\ &\sim (1/4)\, n^{1-2d} h^4 \int \big(f''(x)\big)^2 dx \to 0, \end{aligned}$$

provided $n^{1-2d}h^4 \to 0$ holds. By the Cauchy-Schwarz inequality, this implies that n^{1-2d} times the cross term also tends to zero in probability. We summarize all this in the following

Corollary 2.2 *Suppose assumptions of Corollary* 2.1(i) *hold and* $n^{1-2d}h^4 \to 0$. *Then, under* $\mathcal{H}_0$, $n^{1-2d}|\widehat{T}_n - T_n| = o_p(1)$.

Consequently, $\widehat{T}_n$ has the same asymptotic null distribution as T_n.

3 Simulation

This section shows results of a simulation study that demonstrate the finite sample behavior of the $\widetilde{\mathcal{D}}_n$-test, using autoregressive fractionally integrated moving average (ARFIMA) process.

Table 1 Empirical size of the test based on $\widetilde{\mathcal{D}}_n$ when asymptotic size of the test is 0.05 and 0.1

α	n	d						
		0.1	0.15	0.2	0.25	0.3	0.35	0.4
0.05	100	0.387	0.327	0.290	0.230	0.210	0.209	0.248
	250	0.304	0.231	0.171	0.138	0.116	0.115	0.170
	500	0.259	0.180	0.127	0.103	0.081	0.082	0.123
	1,000	0.209	0.142	0.099	0.079	0.066	0.066	0.083
	2,000	0.185	0.112	0.081	0.070	0.055	0.054	0.066
	5,000	0.155	0.101	0.065	0.053	0.049	0.049	0.054
0.1	100	0.587	0.504	0.453	0.393	0.367	0.362	0.404
	250	0.481	0.381	0.298	0.262	0.229	0.239	0.297
	500	0.414	0.312	0.235	0.206	0.169	0.179	0.234
	1,000	0.352	0.257	0.198	0.164	0.141	0.145	0.182
	2,000	0.313	0.219	0.162	0.144	0.122	0.117	0.155
	5,000	0.276	0.191	0.138	0.116	0.099	0.101	0.133

Using the fracdiff package of Revolution R software, the X_i's were taken to be ARFIMA $(0, d, 0)$ process with standard normal innovations and for $d = 0.1,\ 0.15,\ 0.2,\ 0.25,\ 0.3,\ 0.35$, and 0.4. The Y_i's were generated by using the location model equation $Y_i = \mu + X_i$, where the true mean $\mu = 10$. As in Remark 2.2, $\hat{\mu}$ was taken to be the sample mean of the first half of the process, and the test statistic $\widetilde{\mathcal{D}}_n$ was based on the entire set of the residuals using this estimator. Parameters c and d were estimated using MLE based on the residuals. Sample size n was varied as 100, 250, 500, 1,000, 2,000, and 5,000, with the burn-in period of 10,000. This very long burn-in period was necessary to ensure the accuracy of marginal distribution of the simulated process because the fracdiff package uses truncated moving averages. The procedure was repeated 10,000 times for each sample size.

Table 1 shows the empirical size of the test for the asymptotic levels 0.05 and 0.1 based on Theorem 2.1. These simulation results show that one needs a fairly large sample to implement the test at the chosen levels. Empirical distribution of $\widetilde{\mathcal{D}}_n^2$ for the chosen sample sizes together with the χ^2 distribution with one degree of freedom are shown in Fig. 1. The plots show the convergence of the finite sample distributions to the asymptotic distribution.

4 Proofs

This section contains the proofs of Theorems 2.2 to 2.5. Key point is the bound (4.2) due to Koul and Surgailis (2002) (KS). For the sake of completeness we reproduce this bound and another needed result from this reference. Following KS, let

$$R_{i,1}(x) := I(X_i \le x) - F(x) + f(x)X_i,$$
$$R_{i,2}(x) := I(X_i \le x) - F(x) + f(x)X_i + f'(x)X_i^{(2)}, \quad x \in \mathbb{R},\ i \in \mathbb{Z},$$

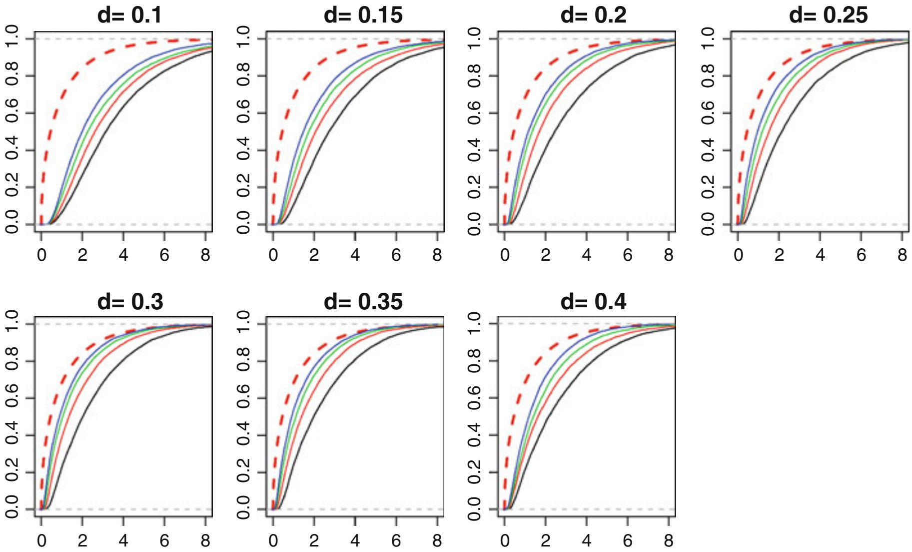

Fig. 1 Empirical distribution of $\widetilde{\mathcal{D}}_n^2$ for $n = 100, 250, 500$, and 1,000 (*bottom* to *top*) compared with the $\chi^2(1)$ (*dashed line*)

where

$$X_i^{(2)} := \sum_{j_2 > j_1 \geq 0} b_{j_1} b_{j_2} \zeta_{i-j_1} \zeta_{i-j_2},$$

and where the last sum converges in mean square; see KS. Let

$$a_1 := \begin{cases} (3-4d)/2, & 0 < d < 1/4, \\ 2(1-2d), & 1/4 < d < 1/2, \end{cases} \quad a_2 := \begin{cases} 3(1-2d), & 1/3 < d < 1/2, \\ 3(1-d)/2, & 1/4 < d < 1/3. \end{cases} \tag{4.1}$$

Note $1 - 2d < a_1$, $2(1-2d) < a_2$, for the values of d indicated in (4.1). Also, let $g(x) := (1 + |x|)^{-3/2}$, $x \in \mathbb{R}$. The first needed result of KS is given in the following

Lemma 4.1 (i) *For all* $i, j \in \mathbb{Z}$, $x \in \mathbb{R}$,

$$\begin{aligned} &\left| E\Big((R_{i,1}(x-uh) - R_{i,1}(x))(R_{j,1}(x-vh) - R_{j,1}(x)) \Big) \right| \\ &\quad \leq C \left(\int_x^{x-uh} g(t)dt \int_x^{x-vh} g(s)ds \right)^{1/2} (1 + |i-j|)^{-a_1} \\ &\quad \leq C|uv|^{1/2} h g(x)(1 + |i-j|)^{-a_1}, \qquad 0 < d < 1/2, \end{aligned} \tag{4.2}$$

$$\Big|E\Big((R_{i,2}(x-uh)-R_{i,2}(x))(R_{j,2}(x-vh)-R_{j,2}(x))\Big)\Big| \\ \le C|uv|^{1/2}hg(x)(1+|i-j|)^{-a_2}, \qquad 1/4<d<1/2. \tag{4.3}$$

(ii) *Let h be a real valued function on $\mathbb{R}$ such that for some $C>0$, $|h(x)|\le Cg(x)$. Then,*

$$\left|\int_0^y h(x+w)dw\right| \le Cg(x)\big(|y|\vee|y|^{3/2}\big), \quad \forall x,y\in\mathbb{R}. \tag{4.4}$$

Proof The inequality (4.4) is (5.12) in KS, p.225. The first inequality in (4.2) is proved in KS, p. 227. The second inequality in (4.2) follows from (4.4) and $|uh|^{3/2}\le|uh|$ for $|u|,|h|\le 1$. Inequality (4.3) follows similarly from KS, p. 227.

Now we return to the proof of Theorem 2.2. Recall (1.4). Let

$$U_n := \mu\bar{\varphi}_n+\bar{X}+\bar{W}, \qquad u_n := \mu(\bar{\varphi}_n-\bar{\varphi})+\bar{X}+\bar{W}.$$

Recall $\widehat{F}_n(y)=n^{-1}\sum_{i=1}^n I(X_i\le y)$. Integration by parts yields

$$\begin{aligned}\widetilde{f}_n(x)-E\widehat{f}_n(x+\mu\bar{\varphi}) &= \frac{1}{h}\int K\Big(\frac{x+U_n-y}{h}\Big)d\widehat{F}_n(y)\\ &\quad -\frac{1}{h}\int K\Big(\frac{x+\mu\bar{\varphi}-y}{h}\Big)dF(y)\\ &= \frac{1}{h}\int(\widehat{F}_n(x+U_n-uh)-F(x+\mu\bar{\varphi}-uh))K'(u)du\\ &= \sum_{i=1}^5\psi_{ni}(x),\end{aligned} \tag{4.5}$$

where

$$\begin{aligned}\psi_{n1}(x) &:= \frac{1}{h}\int\big(\widehat{F}_n(x+U_n-uh)-F(x+U_n-uh)+f(x+U_n-uh)\bar{X}\big)K'(u)du,\\ \psi_{n2}(x) &:= \frac{1}{h}\int\big(F(x+U_n-uh)-F(x+\mu\bar{\varphi}-uh)-f(x+U_n-uh)u_n\big)K'(u)du,\\ \psi_{n3}(x) &:= (\bar{W}+\mu(\bar{\varphi}_n-\bar{\varphi}))\frac{1}{h}\int[f(x+\mu\bar{\varphi}+u_n-uh)-f(x+\mu\bar{\varphi}-uh)]K'(u)du,\\ \psi_{n4}(x) &:= \mu(\bar{\varphi}_n-\bar{\varphi})\frac{1}{h}\int f(x+\mu\bar{\varphi}-uh)K'(u)du,\\ \psi_{n5}(x) &:= \bar{W}\frac{1}{h}\int f(x+\mu\bar{\varphi}-uh)K'(u)du.\end{aligned}$$

With $R_{i,1}(x)$ as in Lemma 4.1, $\psi_{n1}(x) = (nh)^{-1}\sum_{i=1}^{n}\int R_{i,1}(x + U_n - uh)K'(u)du$ and

$$\int \psi_{n1}^2(x)dx = \int \left\{\frac{1}{nh}\sum_{i=1}^{n}\int R_{i,1}(x+U_n-uh)K'(u)du\right\}^2 dx$$
$$= \int \left\{\frac{1}{nh}\sum_{i=1}^{n}\int R_{i,1}(x-uh)K'(u)du\right\}^2 dx$$
$$= \int \left\{\frac{1}{nh}\sum_{i=1}^{n}\int [R_{i,1}(x-uh)-R_{i,1}(x)]K'(u)du\right\}^2 dx,$$

where we used the fact that $\int K'(u)du = 0$. Therefore,

$$E\int \psi_{n1}^2(x)dx = \frac{1}{n^2h^2}\sum_{i,j=1}^{n}\int\int\int E[R_{i,1}(x-uh) - R_{i,1}(x)][R_{j,1}(x-vh) - R_{j,1}(x)]K'(u)K'(v)dudvdx.$$

Hence, using (4.2) we obtain

$$E\int \psi_{n1}^2(x)dx \le \frac{C_1 h}{n^2h^2}\sum_{i,j=1}^{n}(1+|i-j|)^{-a_1} \le \frac{C_1}{n^2 h}\begin{cases} n^{4d}, & 1/4 < d < 1/2, \\ n, & 0 < d < 1/4, \end{cases}, \tag{4.6}$$

with $C_1 = C\int g(x)dx\left(\int |u|^{1/2}|K'(u)|du\right)^2 < \infty$, not depending on n.

Next, consider

$$\begin{aligned} &F(x+U_n-uh) - F(x+\mu\bar{\varphi}-uh) - f(x+U_n-uh)u_n \\ &\quad= F(x+\mu\bar{\varphi}+u_n-uh) - F(x+\mu\bar{\varphi}-uh) - f(x+\mu\bar{\varphi}+u_n-uh)u_n \\ &\quad= -\int_0^{u_n} dz \int_z^{u_n} f'(x+\mu\bar{\varphi}-uh+w)dw. \end{aligned}$$

Then

$$\psi_{n2}(x) = h^{-1}\int_0^{u_n}\int_z^{u_n}\int [f'(x+\mu\bar{\varphi}+w) - f'(x+\mu\bar{\varphi}-uh+w)]K'(u)\,du\,dw\,dz.$$

By (2.3), $|f'(x+\mu\bar{\varphi}+w) - f'(x+\mu\bar{\varphi}-uh+w)| \le Cg(x+\mu\bar{\varphi}+w)|uh|$ and therefore

$$|\psi_{n2}(x)| \le C \int_{|z|\le|u_n|} dz \int_{|w|\le|u_n|} g(x+\mu\bar{\varphi}+w)dw \le C|u_n|(|u_n|+|u_n|^{3/2})g(x+\mu\bar{\varphi})$$

according to (4.4) in Lemma 4.1(ii). By Lemma 2.1, $|u_n| = O_p(n^{d-1/2}) + |\mu|O(n^{-1})$. We thus obtain

$$\int \psi_{n2}^2(x)dx \le Cu_n^4(1+|u_n|)\int g^2(x)dx = O_p(n^{4d-2}). \tag{4.7}$$

In a similar way,

$$\begin{aligned}
&|f(x+\mu\bar{\varphi}+u_n-uh) - f(x+\mu\bar{\varphi}-uh) - f(x+\mu\bar{\varphi}+u_n) + f(x+\mu\bar{\varphi})| \\
&= \Big|\int_0^{u_n} [f'(x+\mu\bar{\varphi}-uh+z) - f'(x+\mu\bar{\varphi}+z)]dz\Big| \\
&\le C|uh| \int_{|z|\le|u_n|} g(x+\mu\bar{\varphi}+z)dz,
\end{aligned}$$

yielding

$$\begin{aligned}
|\psi_{n3}(x)| &\le C(|\bar{W}| + |\mu|n^{-1}) \int_{|z|\le|u_n|} g(x+\mu\bar{\varphi}+z)dz \\
&\le C(|\bar{W}| + |\mu|n^{-1})(|u_n| + |u_n|^{3/2})g(x+\mu\bar{\varphi}),
\end{aligned}$$

and hence

$$\int \psi_{n3}^2(x)dx \le C\Big(|\bar{W}| + |\mu|n^{-1}\Big)^2\Big(|u_n| + |u_n|^{3/2}\Big)^2 = O_p\Big(n^{4d-2}\Big). \tag{4.8}$$

The term $\psi_{n4}(x) = \mu(\bar{\varphi}_n - \bar{\varphi})\frac{1}{h}\int[f(x+\mu\bar{\varphi}-uh) - f(x+\mu\bar{\varphi})]K'(u)du$ can be similarly bounded above by $|\psi_{n4}(x)| \le C|\mu|n^{-1}g(x+\mu\bar{\varphi})$, yielding

$$\int \psi_{n4}^2(x)dx \le C\mu^2 n^{-2}. \tag{4.9}$$

Finally, $h^{-1}\int f(x+\mu\bar{\varphi}-uh)K'(u)du = \int f'(x+\mu\bar{\varphi}-uh)K(u)du \to f'(x+\mu\bar{\varphi})$ and therefore, in view of Lemma 2.1,

$$n^{1-2d}\int \psi_{n5}^2(x)dx \to_D W^2\int (f'(x+\mu\bar{\varphi}))^2 dx, \qquad W \sim \mathcal{N}(0, v^2). \tag{4.10}$$

The claim of Theorem 2.2 follows from decomposition (4.5) and (4.6) to (4.10).

Proof of Theorem 2.3 Introduce r.v.'s

$$Z^{(k)} := \frac{c^k}{k!} \int_{\mathbb{R}^k} \left\{ \int_0^1 \prod_{j=1}^k (u - x_j)_+^{d-1} du \right\} W(dx_1) \cdots W(dx_k), \tag{4.11}$$

which is well-defined for any integer $1 \le k < 1/(1-2d)$, as a multiple Wiener–Itô integral w.r.t. a Gaussian white noise $W(dx)$, $EW(dx) = 0$, $E(W(dx))^2 = dx$.

Similarly to (4.5), write

$$\widetilde{f}_n(x) - E\widehat{f}_n(x) = \sum_{i=1}^3 \tilde{\psi}_{ni}(x),$$

where

$$\begin{aligned}
\tilde{\psi}_{n1}(x) &:= \frac{1}{h} \int \Big(\widehat{F}_n(x + \bar{X} - uh) - F(x + \bar{X} - uh) + f(x + \bar{X} - uh)\bar{X} \\
&\qquad - f'(x + \bar{X} - uh)\overline{X^{(2)}} \Big) K'(u) du, \\
\tilde{\psi}_{n2}(x) &:= \frac{1}{h} \int \Big(F(x + \bar{X} - uh) - F(x - uh) - f(x + \bar{X} - uh)\bar{X} \\
&\qquad + 2^{-1} f'(x + \bar{X} - uh)(\bar{X})^2 \Big) K'(u) du, \\
\tilde{\psi}_{n3}(x) &:= \Big(\overline{X^{(2)}} - 2^{-1}(\bar{X})^2 \Big) \frac{1}{h} \int f'(x + \bar{X} - uh) K'(u) du,
\end{aligned}$$

where $\overline{X^{(2)}} := n^{-1} \sum_{i=1}^n X_i^{(2)}$. With $R_{i,2}$ as in Lemma 4.1, write

$$\tilde{\psi}_{n1}(x) = (nh)^{-1} \sum_{i=1}^n \int [R_{i,2}(x + \bar{X} - uh) - R_{i,2}(x + \bar{X})] K'(u) du.$$

Then using Lemma 4.1, similarly as in the proof of Theorem 2.2, we obtain

$$\begin{aligned}
&E \int \tilde{\psi}_{n1}^2(x) dx \\
&= \frac{1}{n^2 h^2} \sum_{i,j=1}^n \int\int\int E[R_{i,2}(x - uh) - R_{i,2}(x)][R_{j,2}(x - vh) \\
&\quad - R_{j,2}(x)] K'(u) K'(v) du dv dx \\
&\le \frac{Ch}{n^2 h^2} \sum_{i,j=1}^n (1 + |i - j|)^{-a_2} \le \frac{C}{n^2 h} \begin{cases} n^{6d-1}, & 1/3 < d < 1/2, \\ n, & 1/4 < d < 1/3, \end{cases}.
\end{aligned} \tag{4.12}$$

Next as in the proof of Theorem 2.2,

$$\tilde{\psi}_{n2}(x) = h^{-1} \int K'(u)du \int_0^{\bar{X}} dz \int_z^{\bar{X}} dw \int_w^{\bar{X}} d\xi \int_0^{-uh} f'''(x+\xi-v)dv$$

implying $|\tilde{\psi}_{n2}(x)| \le C|\bar{X}|^2(|\bar{X}| + |\bar{X}|^{1/2})g(x) \int |uK'(u)|du$ and hence

$$\int \tilde{\psi}_{n2}^2(x)dx \le C(|\bar{X}|^6(1+|\bar{X}|) \int g^2(x)dx \;=\; O_p(n^{6d-3}). \tag{4.13}$$

Finally, $n^{1-2d}\big(\overline{X^{(2)}} - 2^{-1}(\bar{X})^2\big) \to_D \big(Z^{(2)} - 2^{-1}(Z^{(1)})^2\big)$, see Koul and Surgailis (2010, (5.3)) and $\int \big(\frac{1}{h}\int f'(x+\bar{X}-uh)K'(u)du\big)^2 dx \to_p \|f''\|^2$. These facts together with (4.12), (4.13) and (2.13) imply the statement of Theorem 2.3.

Proof of Theorem 2.4 We shall prove that

$$\begin{aligned} T_n &= m(h) - 2\bar{X}\Delta + (\bar{X})^2\|f'\|^2 \\ &\quad + O_p\big((nh)^{-1/2}\big) + O_p\big(n^{2d-1}h^{-1/2}\big) + O_p(n^{d-1/2}h^2). \end{aligned} \tag{4.14}$$

To this end, write $T_n = \int (t_{n1}(x) - \bar{X}t_{n2}(x) + t_{n3}(x))^2 dx$, where

$$\begin{aligned} t_{n1}(x) &:= \frac{1}{h}\int \big[\widehat{F}_n(x-uh) - F(x-uh) + f(x-uh)\bar{X}\big]K'(u)du, \\ t_{n2}(x) &:= \frac{1}{h}\int f(x-uh)K'(u)du \;=\; \int f'(x-uh)K(u)du, \\ t_{n3}(x) &:= \frac{1}{h}\int [F(x-uh) - F_0(x-uh)\big]K'(u)du \\ &= \int \big[f(x-uh) - f_0(x-uh)\big]K(u)du, \; \int t_{n3}^2(x)dx = m(h). \end{aligned}$$

Hence, with $B_{ij} := \int t_{ni}(x)t_{nj}(x)dx,\ i, j = 1, 2, 3,$

$$T_n = m(h) - 2\bar{X}B_{23} + 2B_{13} + B_{11} - 2\bar{X}B_{12} + (\bar{X})^2 B_{22}, \tag{4.15}$$

From the proof of Theorem 2.2,

$$\begin{aligned} &B_{11} = O_p((hn)^{-1} + h^{-1}n^{4d-2}), \qquad B_{22} \;=\; \|f'\|^2 + O(h^2), \\ &B_{23} = \Delta + O(h^2), \quad |B_{1i}| \le B_{11}^{1/2}B_{ii}^{1/2} \le CB_{11}^{1/2},\ i = 2, 3. \end{aligned} \tag{4.16}$$

Relations (4.15) and (4.16) imply (4.14).

By (2.5), $n^{1/2-d}(nh)^{-1/2} \to 0$, $n^{1/2-d}n^{-1+2d}h^{-1/2} \to 0$, and $n^{1/2-d}n^{d-1/2}h^2 = h^2 \to 0$. Hence, by (4.14),

$$n^{1/2-d}(T_n - m(h)) = -2n^{1/2-d}\bar{X}\Delta + n^{1/2-d}(\bar{X})^2\|f'\|^2 + o_p(1).$$

This in turn together with the first fact in (2.6) and the Ergodic Theorem that guarantees $\bar{X} \to EX_0 = 0$ imply $n^{1/2-d}(T_n - m(h)) \to_D \mathcal{N}(0, 4\kappa^2(\theta)\Delta^2)$. □

Proof of Theorem 2.5 Arguing as in the proof of the previous theorem, we have an analog of (4.15) for $f = f_n$, viz.,

$$T_n = m_n(h) - 2\bar{X}B_{23,n} + 2B_{13,n} + B_{11,n} - 2\bar{X}B_{12,n} + (\bar{X})^2 B_{22,n}, \quad (4.17)$$

with $B_{ij,n}$ defined similarly as above with f replaced by f_n. From the conditions of Theorem 2.5 we easily obtain that

$$\begin{aligned} B_{11,n} &= O_p((hn)^{-1} + h^{-1}n^{4d-2}), \qquad B_{22,n} = \|f_0'\|^2 + o(1), \qquad (4.18)\\ n^{1/2-d}B_{23,n} &\to \Delta_0, \quad n^{1/2-d}|B_{12,n}| \le Cn^{1/2-d}B_{11,n}^{1/2} = o(1),\\ n^{1/2-d}|B_{13,n}| &\le CB_{11,n}^{1/2}\left(n^{1-2d}m_n(h)\right)^{1/2} = o(1). \end{aligned}$$

From (4.17)–(4.18) we obtain, with $Z_n = n^{1/2-d}\bar{X}$,

$$n^{1-2d}(T_n - m_n(h)) = -2Z_n\Delta_0 + Z_n^2\|f_0'\|^2 + o_p(1).$$

Therefore, the claim of the theorem readily follows from (2.6).

Acknowledgments Authors are grateful to the two dedicated anonymous referees whose critical reading of the manuscript helped to improve the presentation.

References

Bachmann D, Dette H (2005) A note on the Bickel–Rosenblatt test in autoregressive time series. Stat Probab Lett 74:221–234
Baillie RT (1996) Long memory processes and fractional integration in econometrics. J Econom 73:5–59
Beran J (1992) Statistical methods for data with long-range dependence (with discussions). Stat Sci 7:404–427
Beran J (1994) Statistics for long-memory processes. Monogr Ser in Stat Appl Probab, vol 61. Chapman & Hall, New York
Bickel P, Rosenblatt M (1973) On some global measures of the deviations of density function estimates. Ann Stat 1:1071–1095
D'Agostino RB, Stephens MA (1986) Goodness-of-fit techniques. In: D'Agostino RB, Stephens MA (eds) Statistics: textbooks and monographs, vol 68. Marcel Dekker, New York
Dalla V, Giraitis L, Hidalgo J (2006) Consistent estimation of the memory parameter for nonlinear time series, preprint. J Time Ser Anal 27:211–251
Davydov YA (1970) The invariance principle for stationary processes. Theor Probab Appl 15:487–498
Dedecker J, Doukhan P, Lang G, León JR, Louichi S, Prieur C (2007) Weak dependence. Lecture notes in statistics, vol 190. Springer, New York
Dehling H, Mikosch T, Sørensen M (eds) (2002) Empirical process techniques for dependent data. Birkhäuser, Boston
Doukhan P, Surgailis D (1998) Functional central limit theorem for the empirical process of short memory linear sequence. C R Acad Sci Paris 326(Serié 1):87–92
Doukhan P, Oppenheim G, Taqqu MS (eds) (2003) Theory and applications of long-range dependence. Birkhäuser, Boston
Durbin J (1973) Weak convergence of the sample distribution function when parameters are estimated. Ann Stat 1:279–290

Durbin J (1975) Distribution theory for test based on the sample d.f. SIAM, Philadelphia
Giraitis L, Koul HL, Surgailis D (1996) Asymptotic normality of regression estimators with long memory errors. Stat Probab Lett 29:317–335
Khmaladze EV (1979) The use of ω^2 tests for testing parametric hypotheses. Theor Probab Appl **24**, 283–301
Khmaladze EV (1981) Martingale approach in the theory of goodness-of-fit tests. Theor Probab Appl **26**, 240–257
Koul HL, Surgailis D (2000) Second order behavior of M-estimators in linear regression models with long memory errors. J Stat Plan Inference 91:399–412
Koul HL, Surgailis D (2002) Asymptotic expansion of the empirical process of long memory moving averages. In: Dehling H, Mikosch T, Sørensen M (eds) Empirical process techniques for dependent data. Birkhäuser, Boston pp 213–239
Koul H, Surgailis D (2010) Goodness-of-fit testing under long memory. J Stat Plan Inference 140: 3742–3753