Minimum Distance Lack-of-Fit Tests under Long Memory Errors[1]

Hira L. Koul[2] and Donatas Surgailis[3]

Michigan State University and Vilnius University

RM 700

Abstract

This paper discusses some tests of lack-of-fit of a parametric regression model when errors form a long memory moving average process with the long memory parameter $0 < d < 1/2$, and when design is non-random and uniform on $[0, 1]$. These tests are based on certain minimized distances between a nonparametric regression function estimator and the parametric model being fitted. The paper investigates the asymptotic null distribution of the proposed test statistics and of the corresponding minimum distance estimators under minimal conditions on the model being fitted. The limiting distribution of these statistics are Gaussian for $0 < d < 1/4$ and non-Gaussian for $1/4 < d < 1/2$. We also discuss the consistency of these tests against a fixed alternative.

1 Introduction

A stochastic process is said to have long memory if its lag k auto-covariances decay to zero like $k^{-\theta}$, for some $0 < \theta < 1$. Long memory processes have been found to arise in a variety of physical and social sciences, see, e.g. Beran (1994), Dehling, Mikosch, and Sørensen (2002), Doukhan, Oppenheim and Taqqu (2003), Robinson (2003), Giraitis, Koul and Surgailis (2012), and references therein.

Suppose (X_i, Y_i), $i = 1, \cdots, n$, are observed from the regression model

$$Y_i = \mu(X_i) + \varepsilon_i, \qquad i = 1, 2, \cdots, n, \tag{1.1}$$

where the design process X_i is a $p \geq 1$ dimensional random vector, μ is a real valued function and where the errors ε_i form a long memory moving average process, i.e., for some $0 < d < 1/2$ and $c_0 \neq 0$,

$$\varepsilon_i \;:=\; \sum_{j=0}^{\infty} \alpha_j \zeta_{i-j}, \qquad \alpha_k \sim c_0\, k^{-(1-d)}, \;\text{ as }\; k \to \infty. \tag{1.2}$$

---

[1] *AMS 1991 subject classification*: Primary 62G07; secondary 62M10.

Keywords and phrases: Moving average errors, local Whittle. Nonparametric residuals. **April 30, 2013**

[2]Research supported in part by NSF DMS grant 1205271.

[3]This author thanks the Michigan State University for hosting his visit from April 7 to May 10, 2013, during which this work was completed. Research of this author was also supported in part by grant MIP-063/2013 from the Research Council of Lithuania.

The innovations ζ_i are assumed to be i.i.d standardized random variables. These assumptions imply long-memory decay of the covariance of ε_i:

$$\gamma_j := E\varepsilon_0\varepsilon_j \sim c_1 j^{-(1-2d)}, \qquad j \to \infty, \quad c_1 := c_0^2 B(d, 1-2d), \tag{1.3}$$

where $B(a,b) := \int_0^1 u^{a-1}(1-u)^{b-1}du$, $a > 0, b > 0$.

In this paper we are interested in the classical problem of lack-of-fit testing of a parametric regression model when errors have long memory. Under the assumption of independent errors this problem has been well studied, cf., Hart (1997), Koul and Ni (2002), and Koul (2011). Not much is available in the literature under long memory errors. Kolmogorov-Smirnov type test based on a marked empirical process of residuals was analyzed in Koul, Baillie, and Surgailis (2004) and Guo and Koul (2007). One of the difficulty this process faces is that to fit a linear regression model with a non-zero intercept one must analyze the second order approximation to this process, and the weak limit thus obtained has non-trackable distribution. It is thus highly desirable to investigate tests that overcome this problem. In this paper we propose a class of such tests based on certain minimized distances in a regression model with non-random design.

We shall confine our attention to the case where $p = 1$, $X_i = i/n$, $i = 1, \cdots, n$, i.e., now our model is

$$Y_i = \mu(i/n) + \varepsilon_i, \tag{1.4}$$

where ε_i are as in (1.2). Let h_j, $j = 1, \cdots, q$, be continuous functions on $[0, 1]$, and let $H := (h_1, \cdots, h_q)'$. Consider the problem of testing

$$\mathcal{H}_0 : \mu(x) = \theta_0' H(x), \text{ for some } \theta_0 \in \mathbb{R}^q, \text{ and for all } x \in [0,1], \text{ vs.}$$
$$\mathcal{H}_1 : \mathcal{H}_0 \text{ is not true.}$$

The main reason for focusing on this relatively simpler problem, compared to fitting a more general nonlinear parametric model and with possibly non-random design, is to keep the exposition from becoming obscure and at the same time for illustrating the new challenges presented by having long memory errors.

To describe the class of tests for this problem, let K be a probability density kernel function on $[-1, 1]$, vanishing off $(-1, 1)$, and $b \equiv b_n$ be a deterministic bandwidth sequence. Let g be a probability densities on $[0, 1]$ and define,

$$\begin{aligned} M_n(\theta) &:= \int_0^1 \Big[\frac{1}{nb}\sum_{j=1}^n K\Big(\frac{nx-j}{nb}\Big)\Big(Y_j - \theta' H\big(\frac{j}{n}\big)\Big)\Big]^2 g(x)dx, \\ \hat{\theta}_n &:= \operatorname{argmin}_{\theta \in \mathbb{R}^q} M_n(\theta). \end{aligned} \tag{1.5}$$

In the next section we shall discuss asymptotic distribution of $\hat{\theta}_n$ and $M_n(\hat{\theta}_n)$. Under suitable conditions it turns out the limiting distribution of $n^{1/2-d}(\hat{\theta}_n - \theta_0)$ under $\mathcal{H}_0$ is Gaussian for all values of $0 < d < 1/2$. The description of the limiting distribution of $M_n(\hat{\theta}_n)$ is more complicated. For some sequence ℓ_n of real numbers, see (2.8) below, the limiting null distribution of $n^{1/2}(nb)^{1/2-2d}M_n(\hat{\theta}_n) - b^{-1/2}\ell_n$ is standard Gaussian, while for $1/4 < d < 1/2$, the limiting null distribution of $n^{1-2d}M_n(\hat{\theta}_n) - b^{2d-1}\ell_n$ is non-Gaussian.

Koul (2011) investigated the above problem in the case of i.i.d. errors. It was shown that for some sequences $\tau_n > 0$ and $\nu_n > 0$, the asymptotic null distribution of $\tau_n(M_n(\hat{\theta}_n) - \nu_n)$ is Gaussian. The entities τ_n, ν_n depend on some unknown parameters. It is further shown in the same paper that for some suitable estimators, $\hat{\tau}_n, \hat{\nu}_n$, the limiting null distribution of $\mathcal{D}_n := \hat{\tau}_n^{-1}(M_n(\hat{\theta}_n) - \hat{\nu}_n)$ is also standard normal. The test is then based on $\mathcal{D}_n$. One could proceed similarly here, i.e., we could plug in an estimate of ℓ_n and d and hope to obtain similar results. Because of the long memory set up things get complicated quickly. Instead we use the idea of symmetrizing the statistic M_n as follows. Let g_1, g_2 be two distinct probability densities on $[0, 1]$ and define M_{ni} and $\hat{\theta}_{ni}$ as M_n and $\hat{\theta}_n$, with g replaced by $g_i, i = 1, 2$. Then the proposed tests are based on the difference $\Delta M_n := M_{n1}(\hat{\theta}_{n1}) - M_{n2}(\hat{\theta}_{n2})$.

Limiting null distributions of $M_n(\hat{\theta}_n)$, $\hat{\theta}_n$ and ΔM_n are described in Theorem 2.1, Corollary 2.1, and Theorem 2.2, respectively. Theorem 2.3 discusses the asymptotic behavior of ΔM_n under some alternatives, including fixed alternatives, which in turn helps to prove consistency of these tests. All proofs are deferred to the last section.

2 Main results

In this section we shall discuss asymptotic distributions of $\hat{\theta}_n$, $M_n(\hat{\theta}_n)$, and ΔM_n. To proceed further, let

$$\begin{aligned} H_n(x) &:= \frac{1}{nb}\sum_{j=1}^{n} K\Big(\frac{nx-j}{nb}\Big)H\Big(\frac{j}{n}\Big), \quad 0 \le x \le 1, \\ \Sigma_n &:= \int H_n(x)H_n(x)'g(x)dx, \quad \Sigma := \int H(x)H(x)'g(x)dx. \end{aligned}$$

Here, and in the sequel, all limits are taken as $n \to \infty$, unless specified otherwise; for any two sequence of real numbers a_n, b_n tending to infinity, $a_n \sim b_n$, means that $\lim_{n\to\infty} a_n/b_n = 1$; $\mathcal{N}(\mu, \sigma^2)$ denote the normal distribution with mean μ and variance σ^2; and $\to_D$ denotes the convergence in distribution.

We need to assume the following about g, K, H and the window width sequence b.

(2.1) g are Lipschitz continuous probability densities on $[0, 1]$.

(2.2) Σ_n and Σ are positive definite for all $n \ge q$.

(2.3) K is an even Lipschitz continuous probability density with support $[-1,1]$.

(2.4) $b \to 0, \qquad nb \to \infty.$

Note that (2.1) - (2.3), $b \to 0$ and continuity of H imply

$$\int \|H_n(x) - H(x)\|^2 g(x)dx \to 0, \qquad \Sigma_n \to \Sigma. \tag{2.5}$$

Let θ_0 be as in $\mathcal{H}_0$, and let

$$U_n(x) \;:=\; \frac{1}{nb}\sum_{j=1}^{n} K(\frac{nx-j}{nb})\varepsilon_j, \quad S_n := \int U_n(x)H_n(x)g(x)dx. \tag{2.6}$$

Note that $M_n(\theta_0) = \int U_n^2(x)\, g(x)\, dx$ and

$$\begin{aligned} M_n(\theta) &= \int [U_n(x) - (\theta-\theta_0)'H_n(x)]^2 g(x)dx \\ &= \int U_n^2(x)\, g(x)\, dx - 2(\theta-\theta_0)'S_n + (\theta-\theta_0)'\Sigma_n(\theta-\theta_0), \quad \theta \in \mathbb{R}^q. \end{aligned}$$

Hence, in view of (2.2),

$$\hat\theta_n - \theta_0 \;=\; \Sigma_n^{-1}S_n, \quad M_n(\hat\theta_n) \;=\; M_n(\theta_0) - S_n'\Sigma_n^{-1}S_n, \quad \forall\, n \ge q. \tag{2.7}$$

Let

$$K_*(z) \;:=\; \int K(x)K(z+x)dx, \qquad K_{**}(z) \;:=\; \int K_*(x)K_*(z+x)dx.$$

Note that K_* and K_{**} are nonnegative even probability densities with supports $[-2,2]$ and $[-4,4]$, respectively. Also denote $\|g\|^2 = \int_0^1 g^2 dx$ and $\tau_+^\alpha := \tau^\alpha I(\tau > 0)$ $(\tau, \alpha \in \mathbb{R})$. Introduce

$$\ell_n := \frac{1}{nb}\sum_{j=-\infty}^{\infty} (nb)^{1-2d}\gamma_j\, K_*(\frac{j}{nb}). \tag{2.8}$$

Note that (1.3) and $nb \to \infty$ imply that

$$\lim_{n\to\infty} \ell_n \;=\; \ell(c_0, d) := c_1 \int |z|^{-(1-2d)} K_*(z)dz, \tag{2.9}$$

where $c_1 = c_0^2 B(d, 1-2d)$ is the asymptotic constant in (1.2).

Theorem 2.1 *Let ε_i be a long memory moving average process as in (1.2) with standardized i.i.d. innovations having finite fourth moment, and suppose assumptions (2.1) to (2.4) hold. Then the following holds.*

(i) For $0 < d < 1/4$,

$$n^{1/2}(nb)^{1/2-2d} M_n(\hat\theta_n) - b^{-1/2}\ell_n \ \ \to_D \ \ \mathcal{N}(0, \sigma^2(c_0, d)),$$

where

$$\sigma^2(c_0, d) \ := \ 2c_1^2(2B(2d, 1-4d) + B(2d, 2d))\|g\|^2 \int |z|^{4d-1} K_{**}(z)dz. \tag{2.10}$$

(ii) For $1/4 < d < 1/2$,

$$n^{1-2d} M_n(\hat\theta_n) - b^{2d-1}\ell_n \ \ \to_D \ \ \mathcal{W}(c_0, d) \ := \ \mathcal{W}^{(2)} - Z'\Sigma^{-1}Z,$$

where

$$\begin{aligned} \mathcal{W}^{(2)} \ &:= \ c_0^2 \int_{\mathbb{R}^2} W(dx_1)W(dx_2) \int_0^1 g(\tau)(\tau - x_1)_+^{d-1}(\tau - x_2)_+^{d-1} d\tau, \\ Z \ &:= \ c_0 \int_{\mathbb{R}} W(dx) \int_0^1 g(\tau)H(\tau)(\tau - x)_+^{d-1} d\tau, \end{aligned} \tag{2.11}$$

are stochastic Itô-Wiener integrals with respect to Gaussian white noise $W(dx)$ with zero mean and variance dx. In particularly, random vector Z has a Gaussian distribution with mean 0 and covariance

$$EZZ' \ = \ c_0^2 B(d, 1-2d) \int_0^1 \int_0^1 g(\tau_1)g(\tau_2)H(\tau_1)H(\tau_2)'|\tau_1 - \tau_2|^{2d-1} d\tau_1 d\tau_2. \tag{2.12}$$

From the proof of Theorem 2.1 (see (3.4)) and (2.7), the following statement is immediate.

Corollary 2.1 *Under the conditions of Theorem 2.1, $n^{1/2-d}(\hat\theta_n - \theta_0) \to_D \Sigma^{-1}Z$, for all $0 < d < 1/2$, where Z is a Gaussian vector with zero mean and covariance matrix as in (2.12).*

Remark 2.1 (i) The case $d = 1/4$ is open. We expect that in this case, the asymptotic distribution of $M_n(\hat\theta_n)$ is Gaussian under a normalization that includes an additional logarithmic factor.
(ii) The limit r.v. $\mathcal{W}(c_0, d)$ in Theorem 2.1(ii) is non-Gaussian, the r.v.'s $\mathcal{W}^{(2)}$ and $Z'\Sigma^{-1}Z$ being correlated. In particularly, $E\mathcal{W}^{(2)} = 0$ and

$$\mathrm{Var}(\mathcal{W}^{(2)}) \ = \ 2c_0^4 B(d, 1-2d)^2 \int_0^1 \int_0^1 g(\tau_1)g(\tau_2)|\tau_1 - \tau_2|^{2(2d-1)} d\tau_1 d\tau_2.$$

A natural idea for implementing the lack-of-fit tests based on Theorem 2.1 is to replace ℓ_n by $\ell(c_0, d)$ of (2.9). Then, the corresponding limit distributions of $M_n(\hat\theta_n)$ would be completely determined by parameters c_0 and d which can be estimated in principle from the residuals $\hat{\varepsilon}_i = Y_i - \hat{\theta}_n' H(i/n), 1 \le i \le n$. However, because of the presence of asymptotically divergent factors $b^{-1/2}$ and b^{2d-1} in front of ℓ_n in both cases (i) and (ii), this procedure requires additional assumptions on the convergence rate in (2.9), which in turn leads to severe restrictions on the bandwidth especially when d is close to 0. Moreover, the critical regions obtained in such a way might be very sensitive to the estimation error of these parameters, resulting in a poor empirical size of these tests.

In order to overcome the above difficulties, we shall now propose a test based on the difference of two minimized dispersions:

$$\Delta M_n \;:=\; M_{n1}(\hat{\theta}_{n1}) - M_{n2}(\hat{\theta}_{n2}). \tag{2.13}$$

Here, $M_{ni}(\theta)$ is the dispersion (1.5) with g replace by g_i, $\hat{\theta}_{ni} := \mathrm{argmin}_{\theta \in \mathbb{R}^q} M_{ni}(\theta)$ is the corresponding estimator of $\theta, i = 1, 2$, and $(g_1, g_2), g_1 \not\equiv g_2$ is a given pair of probability densities on $[0, 1]$. Let

$$\Sigma_{ni} \;:=\; \int H_n(x) H_n(x)' g_i(x) dx, \quad \Sigma_i := \int H(x) H(x)' g_i(x) dx$$

and assume for $i = 1, 2$ that

(2.14) $\quad g_i$ are Lipschitz continuous probability densities on $[0, 1]$.

(2.15) $\quad \Sigma_{ni}$ and Σ_i are positive definite for all $n \ge q$.

The following Theorem 2.2 shows that for the difference of two minimized dispersions in (2.13), the asymptotic bias terms containing ℓ_n cancel out, which is what can be expected from Theorem 2.1 since ℓ_n (2.8) does not depend on g. The proof of Theorem 2.2 largely repeats that of Theorem 2.1, with few changes.

Theorem 2.2 *Suppose assumptions (2.3), (2.4), (2.14), and (2.15) hold, and that ε_i are as in Theorem 2.1. Then the following results hold.*
(i) For $0 < d < 1/4$,

$$n^{1/2} (nb)^{1/2-2d} \Delta M_n \;\to_D\; \mathcal{N}(0, \sigma_\Delta^2(c_0, d)),$$

where

$$\sigma_\Delta^2(c_0, d) \;:=\; 2c_1^2 \{2B(2d, 1-4d) + B(2d, 2d)\} \|g_1 - g_2\|^2 \int |z|^{4d-1} K_{**}(z) dz. \tag{2.16}$$

(ii) For $1/4 < d < 1/2$,

$$n^{1-2d} \Delta M_n \;\to_D\; \mathcal{W}_\Delta(c_0, d) \;:=\; (\mathcal{W}_1^{(2)} - \mathcal{W}_2^{(2)}) - \big(Z_1' \Sigma_1^{-1} Z_1 - Z_2' \Sigma_2^{-1} Z_2\big),$$

where $\mathcal{W}_i^{(2)}, Z_i$ are defined as in (2.11) with g replaced by $g_i, i = 1, 2$.

From part (i) of the above theorem, in the case of $0 < d < 1/4$, we have an asymptotically distribution free test as follows. Let $\bar\theta_n := (\hat\theta_{n1}+\hat\theta_{n2})/2$. Let $\hat d$ be a $\log(n)$-consistent estimator of d and $\hat c_0$ be a consistent estimator of c_0 under $\mathcal{H}_0$, based on $Y_i - \bar\theta_n' H(i/n), 1 \le i \le n$. Let $\hat\sigma_n^2 := \sigma_\Delta^2(\hat c_0, \hat d)$. Then, for an $0 < \alpha < 1$, the test that rejects $\mathcal{H}_0$ whenever

$$n^{1/2}(nb)^{1/2-2\hat d}\big|\Delta M_n\big|/\hat\sigma_n > z_{\alpha/2},$$

is of the asymptotic size α, where z_α is the upper $(1-\alpha)100\%$ percentile of the $\mathcal{N}(0,1)$ distribution. The situation is far from standard in the case $1/4 < d < 1/2$.

Finally, we discuss the consistency of the ΔM_n test against a class of alternatives. To this end, let

$$\kappa_n := \begin{cases} n^{-1/2}(nb)^{2d-1/2}, & 0 < d < 1/4, \\ n^{2d-1}, & 1/4 < d < 1/2, \end{cases} \tag{2.17}$$

so that $\Delta M_n = O_p(\kappa_n)$ is the convergence rate of ΔM_n under the null hypothesis.

Theorem 2.3 *Let $\mu_n(x), x \in [0,1], n \ge 1$ be a sequence of continuous functions in $L_2[0,1]$ such that*

$$\Delta_{ni} := \inf_{\theta\in\mathbb{R}^q} \int_0^1 (\mu_n(x) - \theta' H(x))^2 g_i(x)dx > 0, \quad \forall\, n \ge 1, \quad i = 1,2. \tag{2.18}$$

Let $\Delta_n := \Delta_{n1} - \Delta_{n2}$, and $\theta_{0ni} := \inf_{\theta\in\mathbb{R}^q} \int_0^1 (\mu_n(x) - \theta' H(x))^2 g_i(x)dx$, $i = 1,2$.

Suppose the assumptions of Theorem 2.2 and the regression model (1.4) hold with $\mu \equiv \mu_n$. Moreover, assume that there exist real sequences δ_{ni} and continuous functions ψ_i such that

$$\mu_n(x) - \theta_{0ni}' H(x) = \delta_{ni}\psi_i(x), \quad x \in [0,1] \tag{2.19}$$

and

$$\delta_{ni}^2 = O(\Delta_n), \qquad i = 1,2. \tag{2.20}$$

Then

$$\Delta M_n \;=\; \Delta_n + O_p(\kappa_n), \tag{2.21}$$

where κ_n is defined at (2.17).

Note that in the case of fixed alternative $\mu(x)$ (independent of n) and such that with $\theta_{0i} := \mathrm{argmin}_{\theta\in\mathbb{R}^q} \int_0^1 (\mu(x) - \theta' H(x))^2 g_i(x)dx$, $i = 1,2$,

$$\Delta \;=\; \int_0^1 (\mu(x) - \theta_{01}' H(x))^2 g_1 dx - \int_0^1 (\mu(x) - \theta_{02}' H(x))^2 g_2 dx \;\neq\; 0,$$

conditions (2.19) and (2.20) are automatically satisfied with $\delta_{ni} \equiv 1$, implying

$$\Delta M_n \;=\; \Delta + O_p(\kappa_n) \;=\; \Delta + o_p(1).$$

This together with the fact $\kappa_n \to 0$ in turn implies the consistency of the ΔM_n test against the fixed alternatives $\mu(x)$ with $\Delta \neq 0$, for all $0 < d < 1/2$.

Example. $q = 1, H(x) = x, \mu(x) = x^2$. We get for $i = 1, 2$

$$\begin{aligned}
\hat\theta_{ni} &= \frac{\int_0^1 \big[\frac{1}{nb}\sum_{j=1}^n K(\frac{nx-j}{n})((\frac{j}{n})^2+\varepsilon_j)\big]\big[\frac{1}{nb}\sum_{j=1}^n K(\frac{nx-j}{n})(\frac{j}{n})\big] g_i(x)dx}{\int_0^1 \big[\frac{1}{nb}\sum_{j=1}^n K(\frac{nx-j}{n})(\frac{j}{n})\big]^2 g_i(x)dx}, \\
M_{ni}(\hat\theta_{ni}) &= \int_0^1 \big[\frac{1}{nb}\sum_{j=1}^n K(\frac{nx-j}{n})((\frac{j}{n})^2+\varepsilon_j)\big]^2 g_i(x)dx \\
&- \frac{\big\{\int_0^1 \big[\frac{1}{nb}\sum_{j=1}^n K(\frac{nx-j}{n})((\frac{j}{n})^2+\varepsilon_j)\big]\big[\frac{1}{nb}\sum_{j=1}^n K(\frac{nx-j}{n})(\frac{j}{n})\big] g_i(x)dx\big\}^2}{\int_0^1 \big[\frac{1}{nb}\sum_{j=1}^n K(\frac{nx-j}{n})(\frac{j}{n})\big]^2 g_i(x)dx}.
\end{aligned}$$

Clearly, as $n \to \infty$ for each $i = 1, 2$

$$M_{ni}(\hat\theta_{ni}) \;\to\; \int_0^1 x^4 g_i(x)dx - \frac{\big\{\int_0^1 x^3\, g_i(x)dx\big\}^2}{\int_0^1 x^2 g_i(x)dx}. \tag{2.22}$$

It is also clear that for any $i = 1, 2$ $\mu(x)$ can be written as

$$\mu(x) = x^2 = \theta_i x + \psi_i(x),$$

where

$$\theta_i = \frac{\int_0^1 x^3 g_i dx}{\int_0^1 x^2 g_i dx}, \qquad \psi_i(x) = x^2 - \theta_i x = x^2 - x\frac{\int_0^1 x^3 g_i dx}{\int_0^1 x^2 g_i dx} \tag{2.23}$$

satisfy $\int_0^1 x\psi_i(x)g_i(x)dx = 0$. We see that (2.22) can be rewritten as

$$M_{ni}(\hat\theta_{ni}) \;\to\; \int_0^1 \psi_i^2(x)g_i(x)dx. \tag{2.24}$$

and hence

$$\Delta M_n = M_{n1}(\hat\theta_{n1}) - M_{n2}(\hat\theta_{n2}) \;\to\; \int_0^1 (\psi_1^2(x)g_1(x) - \psi_2^2(x)g_2(x))dx \tag{2.25}$$

completely in agreement with (2.21). Clearly, θ_i, ψ_i in (2.23) depend on g_i although $\mu(x) = x^2$ is fixed alternative and does not depend on i.

3 Proof of Theorems 2.1 - 2.3

Proof of Theorem 2.1. Recall the decomposition of $M_n(\hat{\theta}_n)$ in (2.7). Theorem 2.1 follows from (2.7) and relations below:

$$
\begin{aligned}
&(3.1) & (nb)^{1-2d} EM_n(\theta_0) &= \ell_n + O(b), \\
&(3.2) & n^{1/2}(nb)^{1/2-2d}(M_n(\theta_0) - EM_n(\theta_0)) &\to_D \mathcal{N}(0, \sigma^2(c_0, d)), && 0 < d < 1/4, \\
&(3.3) & n^{1-2d}(M_n(\theta_0) - EM_n(\theta_0)) &\to_D \mathcal{W}^{(2)}, && 1/4 < d < 1/2, \\
&(3.4) & n^{1/2-d} S_n &\to_D Z, && 0 < d < 1/2,
\end{aligned}
$$

and the joint convergence

$$
(3.5) \qquad \big(n^{1-2d}(M_n(\theta_0) - EM_n(\theta_0)), n^{1/2-d} S_n\big) \ \to_D \ (\mathcal{W}^{(2)}, Z), \quad 1/4 < d < 1/2.
$$

Indeed, if $0 < d < 1/4$, then $n^{1/2}(nb)^{1/2-2d} S_n' \Sigma_n^{-1} S_n = O_p(b^{1/2-2d}) = o_p(1)$ according to (3.4), implying $n^{1/2}(nb)^{1/2-2d} M_n(\hat{\theta}_n) = n^{1/2}(nb)^{1/2-2d} EM_n(\theta_0) + n^{1/2}(nb)^{1/2-2d}(M_n(\theta_0) - EM_n(\theta_0)) + o_p(1)$ and the statement of the theorem follows from (3.1) and (3.2). In the case $1/4 < d < 1/2$, we have $n^{1-2d} EM_n(\theta_0) = b^{2d-1}\ell_n + o(1)$ by (3.1) and the statement of Theorem 2.1 obviously follows from (2.7) and (3.5).

Proof of Theorem 2.2. Write

$$
(3.6) \qquad \Delta M_n \ = \ Q_n - \big(S_{n1}' \Sigma_{n1}^{-1} S_{n1} - S_{n2}' \Sigma_{n2}^{-1} S_{n2}\big),
$$

where

$$
(3.7) \qquad Q_n := \big(M_{n1}(\theta_0) - M_{n2}(\theta_0)\big) = \int_0^1 U_n^2(x)(g_1(x) - g_2(x))dx,
$$

U_n is defined in (2.6) and S_{ni} are defined as in (2.6) with g replaced by g_i. Similarly as above, Theorem 2.2 follows from (3.6) and relations below:

$$
\begin{aligned}
&(3.8) & (nb)^{1-2d} EQ_n &= O(b), \\
&(3.9) & n^{1/2}(nb)^{1/2-2d}(Q_n - EQ_n) &\to_D \mathcal{N}(0, \sigma_\Delta^2(c_0, d)), && 0 < d < 1/4, \\
&(3.10) & n^{1-2d}(Q_n - EQ_n) &\to_D \mathcal{W}_\Delta^{(2)}, && 1/4 < d < 1/2, \\
&(3.11) & n^{1/2-d} S_{ni} &\to_D Z_i, \quad i = 1, 2, && 0 < d < 1/2,
\end{aligned}
$$

and the joint convergence

$$
(3.12) \quad \big(n^{1-2d}(Q_n - EQ_n), n^{1/2-d} S_{n1}, n^{1/2-d} S_{n2}\big) \ \to_D \ (\mathcal{W}_\Delta^{(2)}, Z_1, Z_2), \quad 1/4 < d < 1/2.
$$

Here, (3.8) is immediate from (3.1), while (3.9)-(3.12) are analogous to (3.2)-(3.5) and we omit the details. This concludes the proof of Theorem 2.2.

Proof of Theorem 2.3. Define

$$\Psi_{ni}(x) \;:=\; \frac{1}{nb}\sum_{j=1}^{n} K\big(\frac{nx-j}{nb}\big)\psi_i\big(\frac{j}{n}\big),$$

$$\tilde{U}_{ni}(x) \;:=\; U_n(x)+\delta_{ni}\Psi_{ni}(x), \quad \tilde{S}_{ni} := \int_0^1 \tilde{U}_{ni}(x)H_n(x)g_i(x)dx, \quad i=1,2.$$

Following the decomposition in (3.6), we have $\Delta M_n \;=\; \tilde{Q}_n - \tilde{R}_n$, where

$$\tilde{Q}_n := \int_0^1 (\tilde{U}^2_{n1}(x)g_1(x) - \tilde{U}^2_{n2}(x)g_2(x))dx, \qquad \tilde{R}_n \;:=\; \tilde{S}'_{n1}\Sigma^{-1}_{n1}\tilde{S}_{n1} - \tilde{S}'_{n2}\Sigma^{-1}_{n2}\tilde{S}_{n2}. \tag{3.13}$$

From (3.11) and $\int_0^1 \psi_i H g_i dx \;=\; 0$ we have that $\tilde{S}_{ni} \;=\; S_{ni} + \delta_{ni}\int_0^1 \Psi_{ni}(x)H_n(x)g_i(x)dx \;=\; O_p(n^{d-1/2}) + o(\delta_{ni})$ and therefore

$$\tilde{R}_n \;=\; O_p(n^{2d-1}) + o(\delta^2_{n1} + \delta^2_{n2}). \tag{3.14}$$

Next, $\tilde{Q}_n = Q_n + J_n + 2L_n$, where Q_n is defined in (3.7) and

$$J_n := \int_0^1 (\delta^2_{n1}\Psi^2_{n1}(x)g_1(x) - \delta^2_{n2}\Psi^2_{n2}(x)g_2(x))dx, \quad L_n := \int_0^1 U_n(x)(\delta_{n1}\Psi_{n1}(x) - \delta_{n2}\Psi_{n2}(x))dx.$$

We have with $\delta_{n\psi}$ as in (2.21) that

$$J_n \;=\; \delta_{n\psi} + o(\delta^2_{n1} + \delta^2_{n2}), \qquad L_n \;=\; O_p((|\delta_{n1}| + |\delta_{n2}|)n^{d-1/2}), \tag{3.15}$$

where the last relation follows similarly to (3.11). The statement (2.21) of Theorem 2.3 now follows by combining (3.14), (3.14), (2.20) and using the facts that $Q_n = O_p(\kappa_n)$ and $n^{2d-1} = O(\kappa_n)$.

The proof of the technical facts (3.1)-(3.5) used in the above proofs is presented in subsections 3.1-3.3. Before it, let us give a heuristic argument explaining the normalization and the limit distribution in (3.2) and (3.3). Let $m = nb$ be an integer, $m \to \infty, m = o(n)$. By 'discretizing' $M_n(\theta_0)$ we can write

$$\begin{aligned} M_n(\theta_0) \;&\approx\; \frac{1}{n}\sum_{t=1}^{n} g\big(\frac{t}{n}\big)\Big(\frac{1}{m}\sum_{j=1}^{n} K\big(\frac{t-j}{m}\big)\varepsilon_j\Big)^2 \\ &=\; m^{2d-1}\Big(\frac{m}{n}\Big)\sum_{\tau=\frac{1}{m},\frac{2}{m},\cdots,\frac{n}{m}} g\big(\tau\frac{m}{n}\big)\big(\mathcal{Y}_m(\tau)\big)^2\frac{1}{m} \\ &\approx\; m^{2d-1}\Big(\frac{m}{n}\Big)\int_0^{\frac{n}{m}} g\big(\tau\frac{m}{n}\big)\big(\mathcal{Y}_m(\tau)\big)^2 d\tau, \end{aligned}$$

where

$$\mathcal{Y}_m(\tau) \;:=\; \frac{1}{m^{1/2+d}}\sum_{j=1}^{n} K\big(\tau - \frac{j}{m}\big)\varepsilon_j$$

is a linear process in *continuous* time $\tau \in \mathbb{R}$. Using the moving average representation of ε_j in (1.2) and the asymptotics of its coefficients, it is easy to show that

$$\mathcal{Y}_m(\tau) \quad \to_{FDD} \quad \mathcal{Y}(\tau),$$

where

$$\mathcal{Y}(\tau) \quad := \quad c_0 \int K(\tau - u) dB_d(u) \quad = \quad \int \Big(\int K(\tau - u)(u - x)_+^{-(1-d)} du \Big) W(dx)$$

is a stationary Gaussian process in $\tau \in \mathbb{R}$, with zero mean, given by stochastic integral with respect to Gaussian white noise $W(dx)$ as in Theorem 2.1. Let us note that the process

$$B_d(\tau) := \int \Big(\int_0^\tau (u - x)_+^{-(1-d)} du \Big) W(dx), \qquad \tau \in \mathbb{R}$$

is a fractional Brownian motion with variance $EB_d^2(\tau) = B(d, 1-2d)\tau^{1+2d}$ (see, e.g., Taqqu (2003)). The covariance of $\mathcal{Y}$ equals

$$\begin{aligned} E\mathcal{Y}(0)\mathcal{Y}(\tau) \quad &= \quad c_0^2 \int \int K(-u)K(\tau - v) du dv \int (u - x)_+^{-(1-d)} (v - x)_+^{-(1-d)} dx \\ &= \quad c_0^2 B(d, 1-2d) \int \int K(-u) K(\tau - v) |u - v|^{-(1-2d)} du dv \\ &= \quad c_1 \int K_*(y) |y - \tau|^{-(1-2d)} dy, \end{aligned}$$

with c_1 given in (1.3), and hence it decays as $\tau^{-(1-2d)}$ meaning that the covariance of the squared Gaussian process, $\big(\mathcal{Y}\big)^2$, decays as $\tau^{-2(1-2d)}$, and hence it is integrable on the real line for $0 < d < 1/4$ and nonintegrable for $1/4 < d < 1/2$. Since

$$M_n(\theta_0) - EM_n(\theta_0) \quad \approx \quad m^{2d-1} \Big(\frac{m}{n} \Big) \int_0^{\frac{n}{m}} g\big(\tau \frac{m}{n}\big) \Big[\big(\mathcal{Y}(\tau)\big)^2 - E\big(\mathcal{Y}(\tau)\big)^2 \Big] d\tau \tag{3.16}$$

according to approximations above, one can expect that the limit distribution of $M_n(\theta_0) - EM_n(\theta_0)$ coincides with that of the integral on the right-hand side of (3.16). This is true indeed and the limit distribution of the right-hand side of (3.16) can be obtained by using usual techniques for subordinated Gaussian processes (see, e.g., Giraitis and Surgailis (1985), Giraitis *et al.* (2012)), and coincides with the limits in (3.2)-(3.3). However, a rigorous justification of the approximation in (3.16) in the case $0 < d < 1/4$ is difficult and in the proof of (3.2) we use a different approach based on finite-memory approximations and a central limit theorem for $m-$dependent r.v.'s.

3.1 Proof of (3.1)

We have

$$M_n(\theta_0) \quad = \quad \sum_{k,j=1}^{n} w_{jk} \varepsilon_j \varepsilon_k, \tag{3.17}$$

where

$$(3.18) \qquad w_{jk} := \frac{1}{(nb)^2}\int K\big(\frac{nx-j}{nb}\big)K\big(\frac{nx-k}{nb}\big)g(x)dx = \frac{1}{n^2b}\int K(z)K\big(z+\frac{k-j}{nb}\big)g\big(bz+\frac{k}{n}\big)dz.$$

Introduce

$$(3.19) \qquad \tilde{w}_{jk} := \frac{1}{n^2b}K_*\big(\frac{k-j}{nb}\big)g\big(\frac{k}{n}\big), \qquad \tilde{M}_n(\theta_0) := \sum_{k,j=1}^{n}\tilde{w}_{jk}\varepsilon_j\varepsilon_k.$$

Then, using Lipschitz condition for g and the fact that $|\gamma_u| = |E\varepsilon_0\varepsilon_u| \le C|u|^{2d-1}$, we obtain

$$\begin{aligned}
&(nb)^{1-2d}|EM_n(\theta_0) - E\tilde{M}_n(\theta_0)| \\
&= \frac{1}{n}\Big|\sum_{k=1}^{n}\frac{1}{nb}\sum_{u=1-k}^{n-k}[(nb)^{1-2d}\gamma_u]\int K(z)K\big(z+\frac{u}{nb}\big)\Big[g\big(bz+\frac{k}{n}\big)-g\big(\frac{k}{n}\big)\Big]dz\Big| \\
(3.20) \qquad &\le (Cb)\frac{1}{nb}\sum_{|u|\le 2(nb)}[(nb)^{1-2d}\gamma_u] = O(b).
\end{aligned}$$

Next, $(nb)^{1-2d}E\tilde{M}_n(\theta_0) = J_{n1} + J_{n2}$, where

$$\begin{aligned}
(3.21) \qquad J_{n1} &:= \Big(\frac{1}{n}\sum_{k=1}^{n}g\big(\frac{k}{n}\big)\Big)\ell_n = \ell_n + O(1/n), \\
|J_{n2}| &\le \frac{1}{n}\sum_{k=1}^{n}g\big(\frac{k}{n}\big)\frac{1}{nb}\sum_{u\notin[1-k,n-k]}(nb)^{1-2d}|\gamma_u|K_*\big(\frac{u}{nb}\big) \\
(3.22) \qquad &\le \frac{1}{n}\sum_{k:|k-1|\le 2nb \text{ or } |k-n|\le 2nb}\frac{1}{nb}\sum_{|u|\le 2(nb)}(nb)^{1-2d}|\gamma_u| = O(b)
\end{aligned}$$

similarly as above. Now, (3.1) follows from (3.20), (3.21) and (3.22).

3.2 Proof of (3.2)

Let us first prove the asymptotics of the variance:

$$(3.23) \qquad \mathrm{Var}(M_n(\theta_0)) \sim (nb)^{-(1-4d)}n^{-1}\sigma^2(c_0,d),$$

with $\sigma^2(c_0,d)$ given in (2.10). Let $W_{ts} := \sum_{k,j=1}^{n} w_{jk}\alpha_{j-s}\alpha_{k-t}$, with w_{jk} as in (3.18). Note $w_{jk} = w_{kj}$ and $W_{ts} = W_{st}$. Since $M_n(\theta_0) = \sum_{t,s}W_{ts}\zeta_t\zeta_s$, see (3.17), is a quadratic form in standardized i.i.d. r.v.'s ζ_t's, so

$$(3.24) \qquad \mathrm{Var}(M_n(\theta_0)) = 2\sum_{t,s}W_{ts}^2 + (\kappa_4+1)\sum_t W_{tt}^2,$$

where $\kappa_4 = \mathrm{Cum}_4(\zeta_0)$ is the fourth cumulant. Relation (3.23) for $0 < d < 1/4$ follows from

$$\sum_{t,s} W_{ts}^2 \sim (nb)^{-(1-4d)} n^{-1} c(d) \int |z|^{4d-1} K_{**}(z) dz, \tag{3.25}$$

$$\sum_{t} W_{tt}^2 = o\big((nb)^{-(1-4d)} n^{-1}\big), \tag{3.26}$$

where $c(d) := c_1^2(2B(2d, 1-4d) + B(2d, 2d))\|g\|^2$.

Let us prove (3.26). Using $|w_{jk}| \le C(n^2 b)^{-1} 1(|k-j| \le 2nb)$ and $\sum_s |\alpha_s \alpha_{t+s}| \le C|t|_+^{2d-1}$, we obtain

$$\begin{aligned}
\sum_t W_{tt}^2 &= \sum_t \Big(\sum_{k,j=1}^n w_{jk}\alpha_{j-t}\alpha_{k-t}\Big)^2 \\
&\le \frac{C}{n^2(nb)^2} \sum_{k=1}^n \sum_t \sum_{k'} \sum_{|u|\le 2nb,\ |u'|\le 2nb} |\alpha_{k-t}\alpha_{k+u-t}\alpha_{k'-t}\alpha_{k'+u'-t}| \\
&\le \frac{C}{n(nb)^2} \sum_{|u|\le 2nb,\ |u'|\le 2nb} |u|_+^{2d-1}|u'|_+^{2d-1} \\
&\le \frac{C(nb)^{4d-2}}{n},
\end{aligned}$$

thereby proving (3.26).

Consider (3.25). Write

$$\begin{aligned}
n(nb)^{1-4d}\sum_{t,s} W_{ts}^2 &= (nb)^{1-4d} n \sum_{k,k'=1}^n \gamma_{k-k'} \sum_{j,j'=1}^n w_{j,k} w_{j',k'} \gamma_{j-j'} \\
&= \frac{1}{nb}\sum_{|t|<n} (nb)^{1-2d}\gamma_t \sum_{|t+z|<n} \frac{1}{nb}(nb)^{1-2d}\gamma_{t+z} \\
&\quad \times \frac{1}{nb}\sum_u \int\int K(x)K\Big(x + \frac{u}{nb}\Big)K(y)K\Big(y + \frac{u+z}{nb}\Big)dxdy \\
&\quad \times \frac{1}{n}\sum_k^{\dagger} g\Big(bx + \frac{k}{n}\Big) g\Big(by + \frac{k-t}{n}\Big),
\end{aligned} \tag{3.27}$$

where the sum $\sum_k^{\dagger}$ is taken over all $k \in \{1, 2, \cdots, n\}$ such that $k - t \in \{1, 2, \cdots, n\}$. Consider the limit of the above sum as

$$t/nb \to \tau, \qquad z/nb \to \nu, \qquad u/nb \to \eta, \qquad k/n \to s \tag{3.28}$$

and $n \to \infty, b \to 0, nb \to \infty$. We claim that the last limit is

$$
\begin{aligned}
& c_1^2 \int |\tau|^{2d-1} d\tau \int |\tau+\nu|^{2d-1} d\nu \\
& \quad \times \int d\eta \int\int K(x)K(x+\eta)K(y)K(y+\eta+\nu)dxdy \int_0^1 g^2(s)ds \\
& = c_1^2 \|g\|^2 \int K_{**}(\nu)d\nu \int |\tau|^{2d-1}|\tau+\nu|^{2d-1}d\tau \\
& = c(d) \int |\nu|^{4d-1} K_{**}(\nu)d\nu, \qquad (3.29)
\end{aligned}
$$

with $c(d)$ as in (3.25). To rigorously show (3.29), we use the dominated convergence theorem and rewrite the right-hand side of (3.27) as

$$
(nb)^{1-4d} n \sum_{t,s} W_{ts}^2 = \frac{1}{(nb)^3 n} \sum_{|t/nb|\le 1/b, |(t+z)/nb|\le 1/b, |u/nb|\le 2, |k/n|\le 1} G_n(t/nb, z/nb, u/nb, k/n)
$$

with

$$
\begin{aligned}
G_n(t/nb, z/nb, u/nb, k/n) \;\to\; & c_1^2 |\tau|^{2d-1}|\tau+\nu|^{2d-1} g^2(s) \\
& \times \int\int K(x)K(x+\eta)K(y)K(y+\eta+\nu)dxdy
\end{aligned}
$$

in the limit (3.28), and then check the dominated bound

$$
|G_n(t/nb, z/nb, u/nb, k/n)| \le \bar{G}(t/nb, z/nb, u/nb, k/n), \qquad (3.30)
$$

with

$$
\bar{G}(\tau, \nu, \eta, s) := \frac{C}{|\tau|^{1-2d}|\tau+\nu|^{1-2d}} \qquad (3.31)
$$

an integrable function: $\int_{\mathbb{R}} d\tau \int_{-2}^2 d\nu \int_{-2}^2 d\eta \int_0^1 ds \bar{G}(\tau, \nu, \eta, s) < \infty$, due to $0 < d < 1/4$. Verification of the last equality in (3.29), viz.

$$
\int |\tau|^{2d-1}|\tau+\nu|^{2d-1} d\tau = (2B(2d, 1-4d) + B(2d, 2d))|\nu|^{4d-1},
$$

reduces to the case $\nu = 1$, by writing

$$
\begin{aligned}
I &:= \int_{-\infty}^{\infty} |\tau|^{2d-1}|\tau+1|^{2d-1} d\tau = 2I_1 + I_2, \\
I_1 &:= \int_0^{\infty} \tau^{2d-1}(\tau+1)^{2d-1} d\tau, \quad I_2 := \int_0^1 \tau^{2d-1}(1-\tau)^{2d-1} d\tau,
\end{aligned}
$$

where $I_1 = B(2d, 1-4d)$ and $I_2 = B(2d, 2d)$ (see e.g. Gradstein and Ryzhik (1962)).

Next, we turn to the central limit theorem in (3.2). It is possible that this result follows also from the central limit theorem for quadratic forms in Bhansali, Kokoszka and Giraitis

(2006), however, we were not able to verify the conditions on the kernel in the last paper. To prove (3.2), we approximate $Q_n := M_n(\theta_0)$ by a quadratic form $Q_{n,L}$ in $Lnb-$dependent r.v.'s, with $L < \infty$ large enough, for which the result follows by a central limit theorem for finitely dependent triangular arrays. To this end, let

$$\alpha_{j,L} := \alpha_j I(0 \le j \le Lnb), \quad \varepsilon_{t,L} := \sum_{j=0}^{\infty} \alpha_{j,L}\zeta_{t-j}, \quad Q_{n,L} := \sum_{k,j=1}^{n} w_{jk}\varepsilon_{j,L}\varepsilon_{k,L}. \tag{3.32}$$

We will prove that

$$\lim_{L\to\infty} \limsup_{n\to\infty} n(nb)^{1-4d}\mathrm{Var}(Q_n - Q_{n,L}) \;=\; 0 \tag{3.33}$$

and that, for any $L < \infty$, there exists a $C_L(d) < \infty$ such that, as $n \to \infty, nb \to \infty, b \to 0$,

$$\mathrm{Var}(Q_{n,L}) \;\sim\; (nb)^{-(1-4d)} n^{-1} C_L(d), \tag{3.34}$$

$$n^{1/2}(nb)^{1/2-2d}(Q_{n,L} - EQ_{n,L}) \;\to_D\; \mathcal{N}(0, C_L(d)). \tag{3.35}$$

These facts entail (3.2). Indeed, let

$$U_n := n^{1/2}(nb)^{1/2-2d}(Q_n - EQ_n) \quad \text{and} \quad U_{n,L} := n^{1/2}(nb)^{1/2-2d}(Q_{n,L} - EQ_{n,L}).$$

Then the difference of characteristic functions can be estimated as $|Ee^{\mathbf{i}aU_n} - e^{-a^2C^2(d)/2}| \le |Ee^{\mathbf{i}aU_{n,L}} - e^{-a^2C_L^2(d)/2}| + |Ee^{\mathbf{i}aU_n} - Ee^{\mathbf{i}aU_{n,L}}| + |e^{-a^2C^2(d)/2} - e^{-a^2C_L^2(d)/2}| =: J_1 + J_2 + J_3$, where $J_1 = o(1)$ as $n \to \infty$ for any L fixed, by (3.35), and $J_2 \le |\theta|\sqrt{\mathrm{Var}(U_n - U_{n,L})}$ can be made arbitrarily small by (3.33), by letting first $n \to \infty$ and then $L \to \infty$. Relation $J_3 \to 0$ $(L \to \infty)$, or $\lim_{L\to\infty} C_L^2(d) = C^2(d)$, is a consequence of (3.33) and the Cauchy-Schwarz inequality.

Let us prove (3.33). Let $W_{ts,L} := \sum_{j,k=1}^n w_{jk}\alpha_{j-s,L}\alpha_{k-t,L}$. Similarly to (3.24), $\mathrm{Var}(Q_n - Q_{n,L}) \le C\sum_{t,s}(W_{ts} - W_{ts,L})^2$. Also introduce

$$\tilde{\varepsilon}_{t,L} := \varepsilon_t - \varepsilon_{t,L}, \quad \tilde{\gamma}_{t,L} := E\tilde{\varepsilon}_{0,L}\tilde{\varepsilon}_{t,L}, \quad \gamma_{t,L} := E\varepsilon_{0,L}\varepsilon_{t,L}. \tag{3.36}$$

Then

$$\begin{aligned}\sum_{t,s}(W_{ts} - W_{ts,L})^2 \;&\le\; 2\sum_{j,k,j',k'=1}^{n} w_{jk}w_{j'k'}|\gamma_{j-j'}\tilde{\gamma}_{k-k',L}| + 2\sum_{j,k,j',k'=1}^{n} w_{jk}w_{j'k'}|\gamma_{j-j',L}\tilde{\gamma}_{k-k',L}| \\ &=:\; 2R_{n1} + 2R_{n2}.\end{aligned}$$

Note, by definitions (3.32) and (3.36), that $|\gamma_{t,L}| \le Ct^{2d-1} (t \ge 1)$ and $|\tilde{\gamma}_{t,L}| \le \tilde{\delta}(L)t^{2d-1}$ $(t \ge nb)$, with $\tilde{\delta}(L) \to 0$ $(L \to \infty$. Using these facts and similarly to the proof of (3.29) above, we can show that

$$\lim_{L\to\infty} \limsup_{n\to\infty} n(nb)^{1-4d}\Big(|R_{n1}| + |R_{n2}|\Big) \;=\; 0.$$

This proves (3.33).

Let us prove (3.35). For simplicity, assume that $m := nb$ and $\ell := 1/b$ are integers. Then, $m,\ \ell \to \infty,\ m = o(n),\ \ell = o(n)$. Let

$$\tilde U_{n,L} \ := \ \ell^{-1/2} \sum_{i=1}^{\ell} g\big(i\frac{m}{n}\big) \xi_{mi,L}, \tag{3.37}$$

where

$$\xi_{mi,L} \ := \ m^{-1-2d} \sum_{k=(i-1)m+1}^{im} \ \sum_{j:|j-k|\le 2m} K_*\big(\frac{k-j}{m}\big)(\varepsilon_{j,L}\varepsilon_{k,L} - E\varepsilon_{j,L}\varepsilon_{k,L}),$$

$i = 1, \cdots, \ell$. It is easy to check that $E(U_{n,L} - \tilde U_{n,L})^2 = O(1/\ell) = o(1)$ and therefore it suffices to show (3.35) for normalized sum $\tilde U_{n,L}$ in (3.37). Note that, for fixed n, m and L, the sequence $g\big(i\frac{m}{n}\big)\xi_{mi,L}, i \in \mathbb{Z}$ is $(L+2)-$dependent (and has zero mean). Orey (1958) proved that for such sequences, asymptotic normality follows from a Lindeberg-type condition, viz.

$$\max_{1\le i\le \ell} g\big(i\frac{m}{n}\big)^2 E\xi^2_{mi,L} I\big(\big|g\big(i\frac{m}{n}\big)\xi_{mi,L}\big| > u\ell^{1/2}\big) \ = \ o(1) \qquad \text{for any } u > 0.$$

Since g is bounded, and r.v.'s $\xi_{mi,L}, i \in \mathbb{Z}$ are identically distributed, the above fact follows from

$$E\xi^2_{m0,L} I(|\xi_{m0,L}| > u\ell^{1/2}) \ = \ o(1) \qquad \text{for any } u > 0. \tag{3.38}$$

To show (3.38), we shall verify the existence of the limit distribution:

$$\xi_{m0,L} \to_D \xi_L \qquad \text{and} \qquad E\xi^2_{m0,L} \to E\xi^2_L, \qquad m \to \infty. \tag{3.39}$$

Using Skorohod's theorem, the r.v.'s in (3.39) can be defined on the same probability space so that they converge in probability:

$$\xi_{m0,L} \to_p \xi_L \qquad \text{and} \qquad E\xi^2_{m0,L} \to E\xi^2_L, \qquad m \to \infty. \tag{3.40}$$

Since $\xi^2_{m0,L} I(|\xi_{m0,L}| > u\ell^{1/2}) \ = \ o_p(1)$, relation (3.38) follows from (3.40) and Pratt's lemma (Pratt (1960)).

Let us prove (3.39). To this end, rewrite $\xi_{m0,L}$ as a 'discrete stochastic integral':

$$\begin{aligned}\xi_{m0,L} \ &= \ \sum_{t,s\in\mathbb{Z}} \Big(\frac{\zeta_t}{m^{1/2}}\frac{\zeta_s}{m^{1/2}} - E\frac{\zeta_t}{m^{1/2}}\frac{\zeta_s}{m^{1/2}}\Big) \\ &\qquad \times \frac{1}{m}\sum_{k=1}^{m} \frac{1}{m}\sum_{j=1}^{n} K_*\big(\frac{k-j}{m}\big) \Big(\frac{\alpha_{k-t,L}}{m^{d-1}}\Big)\Big(\frac{\alpha_{k-s,L}}{m^{d-1}}\Big);\end{aligned}$$

see Giraitis et al. (2012, sec. 14.3), Surgailis (2003). Using the definition and asymptotics of the truncated moving average coefficients $\alpha_{j,L}$ and Surgailis (2003, Prop. 3.2), it can be easily shown that the above sum converges in distribution to a r.v. ξ_L given by the double Itô-Wiener integral:

$$\xi_L \;=\; c_0^2 \int\int W(dx)W(dy) \int_0^1 \int K_*(u-v)(u-x)_L^{d-1}(v-y)_L^{d-1} du dv,$$

where $y_L^{d-1} := y^{d-1} I(0 < y < L)$. This proves (3.23) and (3.2).

3.3 Proofs of (3.3), (3.4), and (3.5)

The proofs of (3.3)-(3.5) follow a similar 'scheme of discrete stochastic integral' as in (3.39). Let us start with the simplest relation, (3.4). Let $\tilde S_n := \sum_{j,k=1}^n \tilde w_{jk} \varepsilon_k H(j/n)$, where $\tilde w_{jk}$ are defined in (3.19). Similarly as in the proof of (3.1) we can show that it suffices to prove (3.4) for $\tilde S_n$ instead of $S_n = \sum_{j,k=1}^n w_{jk}\varepsilon_k H(j/n)$. Next, rewrite $\tilde S_n$ as a 'discrete stochastic integral':

$$\begin{aligned} n^{1/2-d}\tilde S_n &= \sum_{s\in\mathbb{Z}} (\zeta_s/n^{1/2}) h_{n1}(s), \qquad \text{with} \\ h_{n1}(x) &:= \frac{1}{n}\sum_{k=1}^n g\big(\frac{k}{n}\big)\big(\frac{\alpha_{k-[nx]}}{n^{d-1}}\big)\frac{1}{nb}\sum_{j=1}^n K_*\big(\frac{k-j}{nb}\big) H\big(\frac{j}{nb}\big), \quad x\in\mathbb{R}. \end{aligned}$$

The convergence $n^{1/2-d}\tilde S_n \to_D Z$, with Z in (2.11), follows from Giraitis et al. (2012, Prop. 14.3.2) and the fact that the integrand

$$h_{n1}(x) \to h_1(x) := c_0 \int_0^1 g(u)H(u)(u-x)_+^{d-1} du, \qquad \text{in} \quad L^2(\mathbb{R}), \tag{3.41}$$

where h_1 is the integrand of the limit stochastic integral Z, see (2.11). The proof of (3.41) is elementary by the dominated convergence theorem. This proves the convergence in (3.4), for any $0 < d < 1/2$.

Let us prove (3.3). Again, it is convenient to first approximate Q_n by $\tilde Q_n := \sum_{j,k=1}^n \tilde w_{jk}\varepsilon_k\varepsilon_j$, with $\tilde w_{jk}$ as in (3.19). Let us show (3.3) for $\tilde Q_n$ instead of Q_n and the convergence

$$n^{2-4d}\mathrm{Var}(\tilde Q_n) \;\to\; \mathrm{Var}(\mathcal{W}^{(2)}), \tag{3.42}$$

with the limit variance given in (2.13). To this end, rewrite the normalized quadratic form, $\tilde Q_n$, as a 'double discrete stochastic integral':

$$\begin{aligned} n^{1-2d}(\tilde Q_n - E\tilde Q_n) &= \sum_{t,s\in\mathbb{Z}} \Big(\frac{\zeta_t}{n^{1/2}}\frac{\zeta_s}{n^{1/2}} - E\frac{\zeta_t}{n^{1/2}}\frac{\zeta_s}{n^{1/2}}\Big) h_{n2}(t,s), \qquad \text{with} \\ h_{n2}(x,y) &:= \frac{1}{n}\sum_{k=1}^n g\big(\frac{k}{n}\big)\frac{1}{nb}\sum_{j=1}^n K_*\big(\frac{k-j}{nb}\big)\,\big(\frac{\alpha_{k-[nx]}}{n^{d-1}}\big)\big(\frac{\alpha_{j-[ny]}}{n^{d-1}}\big), \qquad (x,y)\in\mathbb{R}^2. \end{aligned}$$

Since

$$h_{n2}(x,y) \to h_2(x,y) := c_0^2 \int_0^1 g(u)(u-x)_+^{d-1} g(u)(u-y)_+^{d-1} du, \qquad \text{in} \quad L^2(\mathbb{R}^2), \tag{3.43}$$

where h_2 is the integrand of the double integral $\mathcal{W}^{(2)}$ (see (2.11)), from Giraitis et al. (2012, Prop. 14.3.2) we obtain

$$n^{1-2d}(\tilde{Q}_n - E\tilde{Q}_n) \to_D \mathcal{W}^{(2)} \quad \text{and} \quad n^{2-4d}\mathrm{Var}(\tilde{Q}_n) \ \to \ \mathrm{Var}(\mathcal{W}^{(2)}).$$

The proof of (3.43) uses the dominated convergence theorem and the asymptotics of α_k in (1.2). The above relations extend from $\tilde{Q}_n$ to Q_n using Lipschitz continuity of g as in the proof of (3.1). This proves (3.3). Finally, the joint convergence in (3.5) follows from (3.41), (3.43) and Giraitis et al. (2012, Prop. 14.3.3). This ends the proof of Theorem 2.1.

References

Beran, J. (1994). *Statistics for Long Memory Processes.* Monographs on Statistics and Applied Probability, vol. 61. Chapman and Hall, New York.

Bhansali, R.J., Giraitis, L., Kokoszka, P. (2006). Approximations and limit theory for quadratic forms of linear processes. *Stochastic Process. Appl.*

Dehling, H., Mikosch, T., Sørensen, M. (Eds.) (2002). *Empirical Process Techniques for Dependent Data.* Birkhäuser, Boston.

Doukhan, P., Oppenheim, G., Taqqu, M.S. (Eds.) (2003). *Theory and Applications of Long-Range Dependence.* Birkhäuser, Boston.

Giraitis, L., Surgailis, D. (1985). CLT and other limit theorems for functionals of Gaussian processes, *Zeitschrift f. Wahrscheinlichkeitstheorie verw. Geb.* **70**, 191–212.

Giraitis, L., Koul, H.L., Surgailis, D. (2012). *Large Sample Inference for Long Memory Processes.* Imperial College Press, London.

Gradstein, I.S., Ryzhik, I.M. (1962). *Tables of Integrals, Sums, Series and Products* (in Russian). Fizmatgiz, Moscow.

Guo, H., Koul, H.L. (2007). Nonparametric regression with heteroscedastic long memory errors. *J. Statist. Plan. Inference* **137**, 379–404.

Hart, J.D. (1997). *Nonparametric smoothing and lack-of-fit tests.* Springer Series in Statistics. Springer-Verlag, New York.

Koul, Hira L. (2011). Minimum distance lack-of-fit tests for fixed design. *J. Statist. Plann. Inference*, **141**, 65-79.

Koul, Hira L., Ni, Pingping. (2004). Minimum distance regression model checking. *J. Statist. Plann. Inference*, **119**, 109-141.

Koul, H.L., Baillie, R.T., Surgailis, D. (2004) Regression model fitting with a long memory covariate process. *Econometric Theory* **20**, 485–512

Orey, S. (1958). A central limit theorem for $m-$dependent random variables. *Duke Math. J.* **25**, 543–546.

Pratt, J. (1960). On interchanging limits and integrals. *Ann. Math. Statist.* **31**, 74-77.

Robinson, P.M. (1997). Large-sample inference for nonparametric regression with dependent errors, *Ann. Statist.* **25**, 2054-2083.

Robinson, P.M. (2003). *Long-memory Time Series.* Advanced Texts in Econometrics. Oxford University Press, Oxford.

Surgailis, D. (2003). Non CLTs: U-statistics, multinomial formula and approximations of multiple Ito-Wiener integrals. In: Doukhan, P.; Oppenheim, G.; Taqqu, M.S. (eds.), *Long Range Dependence: Theory and Applications*, pp. 129–142. Birkhäuser: Boston.

Taqqu, M.S. (2003). Fractional Brownian motion and long-range dependence. In: Doukhan, P.; Oppenheim, G.; Taqqu, M.S. (eds.), *Long Range Dependence: Theory and Applications*, pp. 5–38. Birkhäuser: Boston.