Goodness-of-fit tests for marginal distribution of long memory error fields in spatial nonparametric regression models

Hira L. Koul Lihong Wang Nao Mimoto
Michigan State University Nanjing University, China
University of Akron, Akron, OH, USA

Abstract

This paper presents a test for fitting the marginal error density of a stationary long memory error random field in spatial nonparametric regression models, where the one dimensional covariate process is also assumed to be a long memory random field, independent of the error random field. The proposed test is based on the integrated square distance between an error density estimator obtained from suitable nonparametric residuals and the expected value of the null error density estimator based on the true errors. Unlike the tests based on the residual empirical distribution functions, the asymptotic null distribution of this statistic is not affected by the regression function estimate and the window widths and the kernels chosen to estimate the regression and density functions. It is chi-square up to a scale factor with one degree of freedom. Moreover, the asymptotic distributions of the proposed test statistic under the null hypothesis and under a class of fixed alternatives are seen to be the same as in the case when the regression function is known. A simulation study shows a reasonable performance of the finite sample level of the proposed test, compared to the asymptotic level.

1 Introduction

One of the classical goodness-of-fit testing problems in statistics is to fit a distribution up to an unknown location parameter in the one sample location model. The classical tests based on certain distances between the residual empirical distribution function and the null model being fitted such as Kolmogorov-Smirnov and Cramer-von Mises tests are not asymptotically distribution free for testing this hypothesis. Their asymptotic null distribution depends on the null model and the parameter estimator in a complicated way which renders them infeasible. In contrast, the asymptotic null distribution of the test based on an integrated square distance between a nonparametric density estimate and its expected value under the null hypothesis with known location parameter for fitting an error density, first proposed by Bickel and Rosenblatt (1973) (B-R) in the case of known location parameter, is not affected by whether the location parameter is known or unknown. This property continues to hold in more complicated models including parametric autoregressive and generalized autoregressive conditionally heteroscedastic models with independent and identically distributed (i.i.d.) errors, see, e.g. Lee and Na (2002), Chebana (2004), Bachmann and Dette (2005), Cheng and Sun (2008), Na (2009), Koul and Mimoto (2012), among others. Moreover, Ghosh and

Huang demonstrated that the B-R test performs better than the Kolmogorov-Smirnov test in detecting sharp peak alternatives. It is thus desirable to expand the domain of applications of this type of the B-R test to long memory random fields.

Long memory time series and long memory random fields (RFs) have been receiving considerable attention in the recent years, due to the fact that more and more empirical studies ranging from astrophysics to agriculture and atmospheric sciences indicate that spatial data may exhibit non-summable correlations and strong dependence, as is evidenced for example in Kashyap and lapsa (1984), Gneiting (2000), Perival Rothrock, Thornfike and Gneiting (2008) and Carlos-Davila, Mejia-Lira and Moran-Lopez (1985). For theoretical advances and more applications of RFs see Dobrushin and Major (1979), Surgailis (1982), Ivanov and Leonenko (1989), Leonenko (1999), Doukhan, Lang and Surgailis (2002) and Lavancier (2006). The recent monographs of Giraitis, Koul and Surgailis (2012) and Beran, Feng, Ghosh and Kulik (2013) contain numerous additional references on long memory time series.

Koul and Surgailis (2010) and Koul, Mimoto and Surgailis (2013) discussed the problem of fitting a known density to the marginal error density of a stationary long memory moving average time series with unknown mean. The asymptotic distributional properties of the B-R and Kolmogorov-Smirnov tests were studied. Koul and Surgailis (2013)considered the same problem for linear random fields with long memory. In this paper we investigate the asymptotic behavior of an analog of the B-R test for fitting marginal error density in the spatial nonparametric regression models where the errors form long memory moving average random fields, and when the covariate random fields also have long memory.

Among the interesting and somewhat surprising findings are that the asymptotic distributions under H_0 and under a large class of alternatives of an analogue of the B-R test based on nonparametric residuals are the same as in the case of known regression function. Moreover, the asymptotic null distribution of the integrated square distance between a nonparametric density estimate and its expected value under the null hypothesis with known regression function has no bias, does not depend on the kernels used in the density and regression function estimates, and the normalization factor, which is the same as in parametric regression models, does not depend on the window widths used in the nonparametric estimation. These findings are unlike the results available under the i.i.d. one sample location set up where the above mentioned integrated squared difference needs to be corrected for the asymptotic bias and where the normalization depends on the window width and the kernel being used in the density estimation.

Proceeding a bit more precisely, we consider a random field defined as a collection of random vectors $(X_i, Y_i) \in \mathbb{R} \times \mathbb{R}$ indexed in the p-dimensional lattice $\mathbb{Z}^p$, $p \geq 1$, where $\mathbb{Z} := \{0, \pm 1, \pm 2, \cdots\}$. Here X_i denotes the ith covariate and Y_i the corresponding response. We assume that the random field $\{X_i, Y_i; i \in \mathbb{Z}^p\}$ is stationary and $E|Y_0| < \infty$. Let $I_n = [1, n]^p \subset \mathbb{Z}^p$. With $\mu(x) = E(Y_0|X_0 = x)$, the observed data consists of (X_i, Y_i), $i \in I_n$, obeying the model

$$Y_i = \mu(X_i) + \varepsilon_i, \quad i \in I_n, \tag{1.1}$$

$$\varepsilon_i = \sum_{j \in \mathbb{Z}^p} a_j \zeta_{i-j}, \quad i \in \mathbb{Z}^p, \tag{1.2}$$

where ζ_j, $j \in \mathbb{Z}^p$, are i.i.d. standardized r.v.'s, independent of $\{X_i,\ i \in \mathbb{Z}^p\}$. Moreover, a_j are non-random weights of the form

$$a_j = B(j/\|j\|)\|j\|^{-(p-d)}, \qquad j \in \mathbb{Z}^p,$$

for some $d \in (0, p/2)$. Here, $B(y)$ is a bounded piecewise continuous function on $S_{p-1} = \{y \in \mathbb{R}^p : \|y\| = 1\}$, and $\|\cdot\|$ denotes the Euclidean norm. Direct calculations yield

$$\gamma(i) = \mathrm{Cov}(\varepsilon_0, \varepsilon_i) \sim R(i/\|i\|)\|i\|^{-(p-2d)}, \quad \|i\| \to \infty, \tag{1.3}$$

with $R(\cdot)$ being some positive, bounded and continuous function on S_{p-1}. Since $0 < d < p/2$, the sum $\sum_{i \in \mathbb{Z}} \|i\|^{-(p-2d)}$ diverges, and hence the random error field $\{\varepsilon_i,\ i \in \mathbb{Z}^p\}$ has long memory in the covariance sense. The long memory property of a stationary random field can be characterized in various ways as is evidenced in Guégan (2005) and Lavancier (2006).

We shall additionally assume that the covariate random field $\{X_i,\ i \in \mathbb{Z}^p\}$ has long memory in the sense that for some $0 < d_X < p/2$ and some positive, bounded continuous function R_X on S_{p-1},

$$\mathrm{Cov}(X_0, X_i) \sim R_X(i/\|i\|)\|i\|^{-(p-2d_X)}, \quad \|i\| \to \infty. \tag{1.4}$$

Additional needed assumptions will be described in the next section.

Now, let f denote the common marginal density function of ε_0, and f_0 be a prescribed density. The problem of interest is to test the hypothesis

$$H_0 : f = f_0 \quad vs. \quad H_a : f \neq f_0,$$

based on the data $\{X_i, Y_i\}$, $i \in I_n$ from the model (1.1) and (1.2). A motivation for this problem is as follows. Numerous statistical inference procedures for long memory processes are developed under the assumption of Gaussianity. A rejection of the null hypothesis that the marginal distribution is Gaussian would cast some doubt about the validity of such inference procedures.

In this paper we investigate an analogue of the B-R type test for H_0. To define this analogue we need to construct appropriate residuals, which can be used to construct the error density estimate. Koul and Surgailis (2002, 2010) observed that in the one sample location model, and more generally in the parametric multiple linear regression models with non-zero intercept and long memory errors, the first order asymptotic behavior of the residual empirical process based on the least square residuals is degenerate. Hence it can not be used to implement any tests of goodness-of-fit of the error distribution in these models. Moreover, as shown in these two papers, the second or higher order distributional approximations of the least square residual empirical process yield some unknown limiting distributions, and hence generally these approximations do not yield any implementable testing procedure.

When testing for H_0 in the case of $p = 1$ in the one sample location model with long memory errors, Koul et al. (2013) proposed a modified estimator of the location parameter so that the corresponding residual empirical process is not the first order degenerate asymptotically. In this paper we introduce an analog of this location parameter estimator to define a nonparametric regression function estimator, which in turn is used to construct nonparametric residuals and the marginal error density estimate.

Accordingly, let $\tilde{K}$ be a kernel density functions on $\mathbb{R}$, h_n be a bandwidth sequence, tending to zero, and let $\tilde{K}_{h_n}(\cdot) = h_n^{-1}\tilde{K}(\cdot/h_n)$. Let ϕ be a piece-wise continuously differentiable function on $[0,1]^p$ and define

$$
\begin{aligned}
(1.5) \qquad \phi_{ni} &= n^p \int_{\prod_{j=1}^{p}((i_j-1)/n, i_j/n]} \phi(u)du, \quad i = (i_1, \cdots, i_p) \in I_n, \\
f_{n,X}(x) &= \frac{1}{n^p}\sum_{i \in I_n} \tilde{K}_{h_n}(x - X_i), \quad x \in \mathbb{R}, \\
\hat{\mu}_n(x) &= \frac{1}{n^p}\sum_{i \in I_n} Y_i(1+\phi_{ni})\tilde{K}_{h_n}(x - X_i)/f_{n,X}(x).
\end{aligned}
$$

Note that $f_{n,X}$ and $\hat{\mu}_n$ are the kernel estimators of the density f_X of X_0 and the regression function μ, respectively. Use $\hat{\mu}_n$ to define the residuals

$$\hat{\varepsilon}_{ni} := Y_i - \hat{\mu}_n(X_i), \quad i \in I_n.$$

Next, let K be another kernel density function and b_n be another bandwidth sequence tending to zero, and define

$$(1.6) \qquad \hat{f}_n(x) = \frac{1}{n^p}\sum_{i \in I_n} K_{b_n}(x - \hat{\varepsilon}_{ni}), \quad \hat{T}_n = \int [\hat{f}_n(x) - K_{b_n} * f_0(x)]^2 dx,$$

$$(1.7) \qquad f_n(x) = \frac{1}{n^p}\sum_{i \in I_n} K_{b_n}(x - \varepsilon_i), \quad T_n = \int [f_n(x) - K_{b_n} * f_0(x)]^2 dx,$$

where $K_{b_n} * f(x) := \int K(u)f(x - b_n u)du$.

In the case of $p = 1$ and i.i.d. observable ε_i's, Bickel and Rosenblatt (1973) proposed a test of H_0 based on T_n. Clearly, $\hat{T}_n$ is an analog of this test statistic in the present set up, with its large values being significant.

The asymptotic distributions of T_n under the null hypothesis $f = f_0$ and under the alternative $f \neq f_0$ with different rates of convergence in each case are described in section 2 below, along with the needed assumptions. Some of the unusual findings are that under long memory, the asymptotic null distribution of $n^{p-2d}T_n$ is a chi-square distribution up to a scale parameter while under certain alternatives, $n^{p/2-d}(T_n - B_n)$ converges in distribution to a

Gaussian r.v. with mean zero and some positive variance, where B_n is a centering sequence depending on K and b_n, see Theorem 2.1 below. These findings are similar to those reported in Corollary 2.1(i) and Theorem 2.4 of Koul et al. (2013) in the case $p = 1$. They are completely different from what is known in the i.i.d. or weakly dependent set up, where the asymptotic null distribution of a suitably centered and scaled T_n is Gaussian, and where the rate of standardization under H_0 depends on the bandwidth h_n. More importantly, we find that the asymptotic distributions of the test statistic $\hat{T}_n$ are quite similar to those of T_n, see Theorem 2.2 below.

The paper is concluded with a simulation study assessing the finite sample level performance of this test in section 4.

Throughout the paper, all limits are taken as $n \to \infty$, unless specified otherwise, $\to_D$ denotes convergence in distribution, $\to_P$ denotes convergence in probability, and Z stands for a standard normal random variable.

2 Asymptotic Distributions of T_n and $\hat{T}_n$

In this section we present the asymptotic distributions under H_0 and H_a of T_n and $\hat{T}_n$. To begin with, we state the needed assumptions for the kernels K, $\tilde{K}$, regression function μ, the bandwidths h_n and b_n, and the underlying model.

Assumption (K): K is a symmetric density, two times differentiable with derivatives satisfying $\int |u|^{1/2}|K'(u)|\, du < \infty$, $\int |K'(u)|^2 du < \infty$, $\int |K''(u)| du < \infty$ and K'' is bounded.

Assumption ($\tilde{\mathrm{K}}$): $\tilde{K}$ is a symmetric bounded density and $\int |u|^2 \tilde{K}(u) du < \infty$.

Assumption (A): The distribution of ζ_0 in (1.2) satisfies the following two conditions: there exists constants $C, \delta > 0$ such that $|Ee^{it\zeta_0}| \le C(1+|t|)^{-\delta}$, $t \in \mathbb{R}$ and $E|\zeta_0|^3 < \infty$.

Assumption (B):

(B1) The regression function μ is twice differentiable with bounded derivatives, and $E\mu^2(X_0) < \infty$.

(B2) The marginal density f_X is bounded, positive and differentiable with bounded derivative f'_X satisfying $\int \big(|f'_X(x)|^2 / f_X(x)\big) dx < \infty$.

(B3) The joint density function $f_{X,j,k}$ of (X_j, X_k) is positive, bounded and such that for some positive constant C and for all $x, y \in \mathbb{R}$,

$$|f_{X,j,k}(x,y) - f_X(x) f_X(y)| \le C\|j-k\|^{-(p-2d_X)} |f'_X(x) f'_X(y)| + o(\|j-k\|^{2d_X - p})$$

as $\|j-k\| \to \infty$, where $d_X \in (0, p/2)$ is as in (1.4).

Moreover, the joint density function $f_{X,i,j,k}$ of (X_i, X_j, X_k) is positive, bounded and such that for some positive constant C' and for all $x, y \in \mathbb{R}$,

$$
\begin{aligned}
&|f_{X,i,j,k}(x,y,z) - f_X(x) f_{X,j,k}(y,z)| \\
&\le C'\Big(\|i-j\|^{-(p-2d_X)} |f'_X(x) f'_X(y) f_X(z)| + \|i-k\|^{-(p-2d_X)} |f'_X(x) f'_X(z) f_X(y)| \Big) \\
&\quad + o(\|i-j\|^{2d_X - p}) + o(\|i-k\|^{2d_X-p}),
\end{aligned}
$$

as $\|j-k\| \to \infty$, $\|i-j\| \to \infty$ and $\|i-k\| \to \infty$.

(B4) $h_n b_n^{-3} \to 0$, $n^{-2d} h_n^{-1} b_n^{-3} \to 0$, $n^{p-2d} h_n^4 b_n^{-3} \to 0$, $n^{2d_X - 2d} b_n^{-3} \to 0$, $n^{2d_X - p} h_n^{-2} b_n^{-3} \to 0$. Moreover, $n^{2d-p} b_n^{-1} \to 0$ when $p/4 < d < p/2$.

Remark 2.1 Here we remark about the validity of the above assumptions.

Assumption (B4) actually implies the following assumption:

(B4$'$) $n^{-2d} h_n^{-1} \to 0$, $n^{p/2-d} h_n^2 \to 0$, $n^{d_X - p/2} h_n^{-1} \to 0$, and $0 < d_X < d < p/2$.

The Assumptions (K) and ($\tilde{\mathrm{K}}$) are the usual standard conditions for the kernel type estimation. Note that Assumption ($\tilde{\mathrm{K}}$) implies that $\int |u|^j \tilde{K}^2(u) du < \infty$, $j = 0, 1, 2$. Examples of the kernels satisfying Assumptions (K), ($\tilde{\mathrm{K}}$) and the Assumption (C2) below are the Gaussian kernel and the uniform kernel vanishing off $(-1, 1)$.

The Assumption (A) is imposed for the first time in Giraitis, Koul and Surgailis (1996) and is then used by Doukhan et al. (2002) to derive the following result (2.5) for the empirical processes of long memory random fields. Under Assumption (A), Koul and Surgailis (2002) (see also Lemma 1 of Giraitis et al. (1996)) showed that, for the case $p = 1$, the error density f is continuously differentiable, and f and its derivative f' are square integrable. Lemma 3.1 of Doukhan et al. (2002) extended this result to the stationary linear random fields with long memory.

Furthermore, from Lemma 2 of Giraitis et al. (1996), Assumption (A) implies that

$$
|f_{0,i}(x,y) - f(x)f(y)| \le C\|i\|^{-(p-2d)} |f'(x) f'(y)| + o(\|i\|^{2d-p}), \tag{2.1}
$$

for any $x, y \in \mathbb{R}$ as $\|i\| \to \infty$ for the $p = 1$ case, where $f_{0,i}$ is the joint density of $(\varepsilon_0, \varepsilon_i)$. Hence, using the same arguments as in Lemma 2 of Giraitis et al. (1996), (see also Lemma 5.1 of Doukhan et al. (2002)), we see that (2.1) holds for the case of $p > 1$. This property actually defines the long memory behavior in distribution for the error field.

The first part of Assumption (B3) imposes a similar long memory property on the random field of covariates. Section 3.5 of Guégan (2005) more details for this concept of long memory. Moreover, for the special case of Gaussian random fields, the second part of Assumption (B3) is similar to the moment bounds studied in Soulier (2001) and Guo and Koul (2008). The second part of Assumption (B3) is used in the proof of the following Lemma 3.2. Direct calculations yield that the long memory Gaussian random fields with the covariance function as in (1.4) satisfy Assumption (B3). Many commonly used densities including Gaussian and logistic can be shown to satisfy (B2).

Moreover, if one uses $h_n \sim n^{-a}$, $a > 0$, then assumption (B4$'$) will be satisfied as long as $p/4 - d/2 < a < \min(2d, p/2 - d_X)$ for $p/10 < d < p/2$ and $d_X < d$. In addition, if

$b_n \sim n^{-r}$, $r > 0$, then Assumption (B4) holds for $0 < r < \min(a/3, 2d/3 - a/3, 2d/3 + 4a/3 - p/3, 2d/3 - 2d_X/3, p/3 - 2d_X/3 - 2a/3, p - 2d)$. For example, if $p = 2$, $d = 0.5$, $d_X = 0.1$, then $0.25 < a < 0.9$. Let $a = 0.6$, then $0 < r < 2/15$.

We now state some needed preliminaries. Let $\bar{\varepsilon}_n \equiv n^{-p} \sum_{i \in I_n} \varepsilon_i$. From Surgailis (1982), we recall that for the random field (1.2),

$$n^{p/2-d} \bar{\varepsilon}_n \to_D cZ, \qquad n^{p-2d} E\bar{\varepsilon}_n^2 = (c^2 + o(1)), \tag{2.2}$$

where c is the positive square root of

$$c^2 = \int_{[0,1]^p} \int_{[0,1]^p} R((u-v)/\|u-v\|)\|u-v\|^{2d-p} du dv.$$

Theorem 2.1 below describes the asymptotic distributions of a suitably standardized T_n under the null hypothesis and under some fixed alternatives, respectively. Let $\kappa_1 := \int (f_0{}'(x))^2 dx$ and $\kappa_2 := \int f'(x)(f(x) - f_0(x))dx$. From Assumption (A) and Remark 2.1, $\kappa_1 < \infty$ and $|\kappa_2| < \infty$.

Theorem 2.1 *Suppose that the Assumptions (K) and (A) hold, and* $b_n^{-1} = o(n^\rho)$, *where* $\rho = \min(p/2 - d, (p-d)\delta)$, *with* δ *the same as in Assumption (A). Then the following hold.*
(i) *Under the null hypothesis* H_0,

$$n^{p-2d} T_n \to_D \kappa_1 c^2 Z^2.$$

(ii) *Under a fixed alternative* $f \neq f_0$ *of the form* $\int (f(x) - f_0(x))^2 dx > 0$,

$$n^{p/2-d}\Big(T_n - \int [K_{b_n} * (f - f_0)]^2(x) dx\Big) \to_D 2\kappa_2 cZ.$$

Proof. The proof here is based on some techniques used in Koul et al. (2013) and Bachmann and Dette (2005). Let $H(x, \varepsilon_i) = K_{b_n}(x - \varepsilon_i) - K_{b_n} * f(x)$, $h(\varepsilon_i) = K_{b_n} * (K_{b_n} * (f - f_0))(\varepsilon_i)$, and consider the following decomposition for T_n.

$$\begin{aligned} T_n &= \int [f_n(x) - K_{b_n} * f_0(x)]^2 dx \\ &= \int [f_n(x) - K_{b_n} * f(x)]^2 dx + \int [K_{b_n} * (f - f_0)(x)]^2 dx \\ &\quad + 2\int [f_n(x) - K_{b_n} * f(x)][K_{b_n} * (f - f_0)(x)] dx \\ &= \int \Big(\frac{1}{n^p} \sum_{i \in I_n} [K_{b_n}(x - \varepsilon_i) - K_{b_n} * f(x)]\Big)^2 dx + \int [K_{b_n} * (f - f_0)(x)]^2 dx \\ &\quad + \frac{2}{n^p} \sum_{i \in I_n} \{K_{b_n} * (K_{b_n} * (f - f_0))(\varepsilon_i) - E[K_{b_n} * (K_{b_n} * (f - f_0))(\varepsilon_i)]\} \end{aligned} \tag{2.3}$$

$$= \int \Big(\frac{1}{n^p}\sum_{i\in I_n} H(x,\varepsilon_i)\Big)^2 dx + \int [K_{b_n} * (f-f_0)(x)]^2 dx + \frac{2}{n^p}\sum_{i\in I_n}(h(\varepsilon_i) - Eh(\varepsilon_i)).$$

Let $F_n(x) = n^{-p}\sum_{i\in I_n} I(\varepsilon_i \le x)$, $x \in \mathbb{R}$. Following the arguments used in the proof of Theorem 2.2 of Koul et al. (2013) and integration by parts yield

$$\begin{aligned}
&\frac{1}{n^p}\sum_{i\in I_n} H(x,\varepsilon_i)\\
&= \frac{1}{b_n}\int [F_n(x-ub_n) - F(x-ub_n) + f(x-ub_n)\bar\varepsilon_n]K'(u)du\\
&\qquad\qquad -\bar\varepsilon_n \frac{1}{b_n}\int f(x-ub_n)K'(u)du\\
&= \frac{1}{b_n}\int [F_n(x-ub_n) - F(x-ub_n) + f(x-ub_n)\bar\varepsilon_n]K'(u)du\\
&\qquad\qquad -\bar\varepsilon_n \int f'(x-ub_n)K(u)du\\
&=: Q_{n1}(x) + Q_{n2}(x),
\end{aligned}$$

where F is the distribution function of ε_i. Hence

$$\int \Big(\frac{1}{n^p}\sum_{i\in I_n} H(x,\varepsilon_i)\Big)^2 dx = \int Q_{n1}^2(x)dx + \int Q_{n2}^2(x)dx + 2\int Q_{n1}(x)Q_{n2}(x)dx. \tag{2.4}$$

Now, let $S_n(x) = \sum_{i\in I_n}\{I(\varepsilon_i \le x) - F(x) + f(x)\varepsilon_i\}$, $S_n(x, x-ub_n) = S_n(x-ub_n) - S_n(x)$, and $V(x, x-ub_n) = F(x-ub_n) - F(x) + \int_x^{x-ub_n}(1+|t|)^{-2}dt$. Because $\int K'(u)du = 0$, we can write $Q_{n1}(x) = (n^p b_n)^{-1}\int S_n(x, x-ub_n)K'(u)du$. It follows from the proof of Lemma 1.4 of Doukhan et al. (2002) that for some constant $0 < C < \infty$,

$$E(S_n(x, x-ub_n))^2 \le CV(x, x-ub_n)n^{p+2d-\rho}. \tag{2.5}$$

This, together with Assumption (K) which guarantees $\int |u|^{1/2}|K'(u)|du < \infty$, yields

$$\begin{aligned}
(2.6)\quad & E\int Q_{n1}^2(x)dx\\
&= \frac{1}{n^{2p}b_n^2}\int E\{S_n(x,x-ub_n)S_n(x,x-vb_n)\}K'(u)K'(v)dudvdx\\
&\le \frac{1}{n^{2p}b_n^2}\int \{E(S_n(x,x-ub_n))^2 E(S_n(x,x-vb_n))^2\}^{1/2}|K'(u)K'(v)|dudvdx\\
&\le \frac{1}{n^{2p}b_n^2}\int C\{V(x,x-ub_n)V(x,x-vb_n)\}^{1/2}n^{p+2d-\rho}|K'(u)K'(v)|dudvdx\\
&\le Cn^{2d-p-\rho}b_n^{-1}\int |uv|^{1/2}|K'(u)||K'(v)|\{f(x) + (1+|x|)^{-2}\}dudvdx\\
&= O\Big(n^{2d-p-\rho}b_n^{-1}\Big).
\end{aligned}$$

For $Q_{n2}(x)$, under H_0, $\int f'(x - ub_n)K(u)du \to f_0'(x)$. This, together with (2.2), yields

$$n^{p-2d} \int Q_{n2}^2(x)dx \to_D \kappa_1 c^2 Z^2. \tag{2.7}$$

Upon combining this with (2.6) and the Cauchy-Shwarz inequality yields that

$$\int |Q_{n1}(x)Q_{n2}(x)|dx \to_P 0. \tag{2.8}$$

Moreover, under H_0, the second term in the right hand side of (2.3) is zero and $h(\varepsilon_i) \equiv 0$. Therefore, upon combining (2.3), (2.4), (2.6), (2.7) with (2.8), we arrive at the result (i) of Theorem 2.1.

On the other hand, under the alternative, (2.3), (2.6), (2.7) and the assumption $b_n^{-1} = o(n^\rho)$ yield

$$\begin{aligned} & n^{p/2-d}\Big(T_n - \int [K_{b_n} * (f - f_0)]^2(x)dx\Big) \\ &= O_P(n^{d-p/2}) + 2n^{-d-p/2} \sum_{i \in I_n} (h(\varepsilon_i) - Eh(\varepsilon_i)). \end{aligned} \tag{2.9}$$

Note that

$$\sum_{i \in I_n} (h(\varepsilon_i) - Eh(\varepsilon_i)) = n^p \int h(x)d[F_n(x) - F(x)],$$

and

$$\int h(x)f'(x)dx = \int f'(x)K(u)K(v)(f - f_0)(x - b_n u - b_n v)dudvdx \longrightarrow \kappa_2.$$

By Corollary 1.2 and Corollary 1.3 of Doukhan et al. (2002), we obtain

$$n^{-d-p/2} \sum_{i \in I_n} (h(\varepsilon_i) - Eh(\varepsilon_i)) \to_D \kappa_2 c Z. \tag{2.10}$$

The proof of part (ii) of Theorem 2.1 is completed upon combining (2.9) and (2.10).

Remark 2.2 It is of interest to compare part (ii) of the above theorem with Theorem 2.1 in Bachmann and Dette (2005), where they prove, when ε_i are i.i.d. f satisfying some other mild conditions, that $n^{1/2}\Big(T_n - \int [K_{b_n} * (f - f_0)]^2(x)dx\Big) \to_D 2\rho Z$, where $\rho^2 = \mathrm{Var}(f - f_0)(\varepsilon_1)$. Clearly, part (ii) of the above theorem is an appropriate extension of this result to the long memory moving average random fields.

Remark 2.3 Note that, if in the case $p = 1$ we assume that $a_j = c_0 j^{-(1-d)}$, where c_0 is a finite positive constant, then one can verify that $c^2 = c_0^2 B(d, 1 - 2d)/d(1 + 2d)$, where $B(a, b) = \int_0^1 x^{a-1}(1 - x)^{b-1}dx$, $a > 0$, $b > 0$. Then, comparing Theorem 2.1(i) and Theorem 2.1(ii) with Corollary 2.1(i) and Theorem 2.4 of Koul et al. (2013) , respectively, in the case $p = 1$, we can conclude that, even if one is in the nonparametric regression set up, one still obtains results analogous to those available in the parametric mean model. Theorem 2.1(i) is also analogous to Proposition 3.1(i) in Koul and Surgailis (2013) for linear random fields with zero mean.

Some additional conditions on the kernel $\tilde{K}$ and the explanatory random field $\{X_i, i \in \mathbb{Z}^p\}$ are needed for deriving the asymptotic distributions of the test statistic $\hat{T}_n$.

Assumption (C):

(C1) $X_i \in \Delta$, $i \in I_n$, where Δ is a compact set and $f_X(x) \geq M$ for $x \in \Delta$, where M is a finite positive constant.

(C2) The Fourier transform $\Phi(r) = 2\pi \int e^{irx}\tilde{K}(x)dx$ of $\tilde{K}$ is integrable: $\int |\Phi(r)|dr < \infty$.

The next theorem describes the asymptotic distributions of $\hat{T}_n$. Its proof is given in the next section. Let $c^2(\phi) = \int_{[0,1]^p}\int_{[0,1]^p} \phi(u)\phi(v)R((u-v)/\|u-v\|)\|u-v\|^{2d-p}dudv$ and $\bar{\phi} = n^{-p}\sum_{i\in I_n}\phi_{ni}$.

Theorem 2.2 *Suppose that the Assumptions (K), ($\tilde{K}$), (A), (B) and (C) hold. Let $\phi(u)$, $u \in [0,1]^p$ be a piecewise continuously differentiable function satisfying $\bar{\phi} = 0$. Then the following hold.*

(i) Under the null hypothesis H_0,

$$n^{p-2d}\hat{T}_n \to_D \kappa_1 c^2(\phi) Z^2.$$

(ii) Under a fixed alternative $f \neq f_0$ of the form $\int (f(x) - f_0(x))^2 dx > 0$,

$$n^{p/2-d}\Big(\hat{T}_n - \int [K_{b_n} * (f - f_0)]^2(x)dx\Big) \to_D 2\kappa_2 c(\phi) Z.$$

Remark 2.4 Note that, the asymptotic distribution in Theorem 2.2(i) is the same as that in Proposition 3.1(ii) of Koul and Surgailis (2013) for the location model of long memory linear random fields, while Theorem 2.2(ii) is new even in the case $p = 1$.

Estimation of d and $c(\phi)$. To implement the test of H_0 based on $\hat{T}_n$, we need to have consistent and $\log(n)$-consistent estimates of $c(\phi)$ and d, respectively. Several approaches for estimating the underlying parameters in long memory random fields have been suggested in the literature. Frias, Alonso, Ruiz-Medina and Angulo (2008) suggested an averaged periodogram estimator for long memory time series in the two-dimensional spatial case. Wang (2009) investigated the GPH-estimator for long memory random fields. Under some general conditions these estimators of the long memory parameter d are all $\log(n)$-consistent.

In the simulation studies below, we used the GPH-estimator $\hat{d}$ studied in Hurvich, Deo and Brodsky (1998) andWang (2009). To describe this estimator, let $\lambda^{(j)} = (\lambda_{j_1}, \cdots, \lambda_{j_p})' = 2\pi j/n$, $j = (j_1, \cdots, j_p)' \in I_n$, and m be a "bandwidths" sequence that tends to infinity slower than n. Let $U_j = \log|1 - \exp(-i\lambda^{(j)})| = \frac{1}{2}\log\Big(\sum_{k=1}^p |1-\exp(-i\lambda_{j_k})|^2\Big)$, $\bar{U} = m^{-p}\sum_{j\in I_m} U_j$, and define

$$(2.11)\ Q_j = \frac{1}{(2\pi n)^p}\Big|\sum_{t\in I_n}\hat{\varepsilon}_{nt}\exp(it'\lambda^{(j)})\Big|^2,\ j \in I_m, \quad \hat{d} = -\frac{\sum_{j\in I_m}(U_j - \bar{U})\log Q_j}{2\sum_{j\in I_m}(U_j - \bar{U})^2},$$

$$\tilde{Q}_j = \frac{1}{(2\pi n)^p}\Big|\sum_{t\in I_n}\varepsilon_t\exp(it'\lambda^{(j)})\Big|^2,\ j \in I_m, \quad \tilde{d} = -\frac{\sum_{j\in I_m}(U_j - \bar{U})\log \tilde{Q}_j}{2\sum_{j\in I_m}(U_j - \bar{U})^2}.$$

As stated in Lavancier (2006), if the covariance of a random field satisfies (1.3), then its spectral density satisfies $f(\lambda) \sim \|\lambda\|^{-2d} R_0(\lambda/\|\lambda\|)$, as $\|\lambda\| \to 0$, with $R_0(\cdot)$ being some continuous function on the unit sphere. Therefore, by Theorem 1 of Wang (2009), $\tilde{d} - d = O_P(m^{-1}(\log m)^7)$ for $p \geq 2$, where m satisfies $m \log m/n \to 0$. If we let $m = n^{\epsilon}$, where $0 < \epsilon < 1$, then we obtain $|\tilde{d} - d| = o_P(n^{-\epsilon_1})$, for any $0 < \epsilon_1 < \epsilon$. In addition,

$$\hat{d} - \tilde{d} = -\frac{\sum_{j \in I_m} (U_j - \bar{U})(\log Q_j - \log \tilde{Q}_j)}{2 \sum_{j \in I_m} (U_j - \bar{U})^2}.$$

From the following proofs of Theorem 2.2 and Lemma 3.3, we note that, for every $j \in I_m$,

$$Q_j = O_P(n^p), \quad \tilde{Q}_j = O_P(n^p), \quad |Q_j - \tilde{Q}_j| = O_P(n^{d+p/2}).$$

This, together with Lemma 1 of Wang (2009), implies that

$$|\hat{d} - \tilde{d}| = O_P(n^{d-p/2} \log m), \qquad \hat{d} - d = o_P((\log n)^{-1}).$$

To estimate $c(\phi)$, Koul and Surgailis (2013) used the fact that

$$c^2(\phi) = \lim_{n \to \infty} n^{-p-2d} \sum_{j,k \in I_n} \phi_{nj} \phi_{nk} \gamma(j-k)$$

to propose a consistent estimator of $c(\phi)$ based on the sample auto-covariance function of the residuals. Their estimate is

$$\hat{c}^2(\phi) = q^{-p-2\hat{d}} \sum_{j,k \in I_q} \phi_{qj} \phi_{qk} \hat{\gamma}(j-k),$$

where $q \to \infty$, $q = o(n)$ is a bandwidth sequence, $\hat{\gamma}(k)$'s are the sample auto-covariances based on the residuals $Y_i - \bar{Y}$, and $\hat{d}$ is a $\log(n)$-consistent estimator of d based on these residuals. Lemma 3.3 below shows that $\hat{c}^2(\phi)$ with $\hat{\gamma}(k) = n^{-p} \sum_{i \in I_n} \hat{\varepsilon}_{ni} \hat{\varepsilon}_{n,i+k}$ and $\hat{d}$ as in (2.11) continues to be consistent for $c^2(\phi)$ in the above nonparametric regression model.

It thus follows from Theorem 2.2(i) that the test that rejects H_0 whenever $\kappa_1^{-1} \hat{c}^{-2}(\phi)$ $n^{p-2\hat{d}}\ \hat{T}_n > \chi^2_\alpha(1)$ is asymptotically distribution free and of the asymptotic level α, where $0 < \alpha < 1$ and $\chi^2_\alpha(1)$ is the $(1-\alpha)100$th percentile of the $\chi^2(1)$ distribution.

Using the same discussions as in Koul et al. (2013) , we see that Theorem 2.2(ii) implies that the $\hat{T}_n$-test is consistent against all those density function f's for which $\int (f(x) - f_0(x))^2 dx > 0$.

3 Proof of Theorem 2.2

The proof of Theorem 2.2 is facilitated by the following lemmas.

Lemma 3.1 *Under the Assumptions ($\tilde{K}$), (B2), (B3) and (C2),*

$$E\Big(\sup_{x \in \mathbb{R}} |f_{n,X}(x) - f_X(x)|\Big)^2 = O(\max\{n^{2d_X - p} h_n^{-2}, h_n^2\}).$$

Proof. The proof is similar to that of Bierens (1983). Write

$$f_{n,X}(x) = \frac{1}{n^p}\sum_{j\in I_n}\tilde{K}_{h_n}(x-X_j) = \frac{1}{n^p h_n}\sum_{j\in I_n}\int e^{-ir(x-X_j)/h_n}\Phi(r)dr.$$

Then

$$
\begin{aligned}
(3.1)\qquad & \sup_{x\in\mathbb{R}}|f_{n,X}(x)-E(f_{n,X}(x))| \\
&= \sup_{x\in\mathbb{R}}\Big|\frac{1}{n^p h_n}\sum_{j\in I_n}\Big\{\int e^{-ir(x-X_j)/h_n}\Phi(r)dr - E\Big[\int e^{-ir(x-X_j)/h_n}\Phi(r)dr\Big]\Big\}\Big| \\
&\le \int\Big|\frac{1}{n^p h_n}\sum_{j\in I_n}\Big\{e^{irX_j/h_n}-E(e^{irX_j/h_n})\Big\}\Big|\sup_{x\in\mathbb{R}}|e^{-irx/h_n}||\Phi(r)|dr. \\
&\le \int\Big|\frac{1}{n^p}\sum_{j\in I_n}\{e^{irX_j}-E(e^{irX_j})\}\Big||\Phi(rh_n)|dr.
\end{aligned}
$$

By Assumption (B3), for some large enough N,

$$
\begin{aligned}
(3.2)\qquad & \mathrm{Var}\Big(\frac{1}{n^p}\sum_{j\in I_n}\cos(rX_j)\Big) \\
&= \frac{1}{n^{2p}}\sum_{j\in I_n}\mathrm{Var}(\cos(rX_j)) + \frac{1}{n^{2p}}\sum_{i\neq j}\mathrm{Cov}(\cos(rX_i),\cos(rX_j)) \\
&\le Cn^{-p} + \frac{1}{n^{2p}}\Big(\sum_{0<\|j-i\|\le N} + \sum_{\|j-i\|>N}\Big) \\
&\qquad\qquad \cdot\int\cos(rx)\cos(ry)\{f_{X,i,j}(x,y)-f_X(x)f_X(y)\}dxdy \\
&\le Cn^{-p} + Cn^{-2p}\sum_{\|j-i\|>N}\Big(\|j-i\|^{2d_X-p} + o(\|j-i\|^{2d_X-p})\Big) \\
&= O(n^{-p}) + O(n^{2d_X-p}).
\end{aligned}
$$

The same inequality holds with $\cos(\cdot)$ replaced by $\sin(\cdot)$. Hence,

$$\sup_{r\in\mathbb{R}} E\Big|\frac{1}{n^p}\sum_{i\in I_n}\{e^{irX_i}-E(e^{irX_i})\}\Big|^2 = O(n^{2d_X-p}).$$

This, together with the fact that $\int|\Phi(rh_n)|dr = O(h_n^{-1})$, implies

$$(3.3)\qquad E\Big(\sup_{x\in\mathbb{R}}|f_{n,X}(x)-E(f_{n,X}(x))|\Big)^2 = O(n^{2d_X-p}h_n^{-2}).$$

Furthermore, by Assumption (B2),

$$\Big(\sup_{x\in\mathbb{R}}|E(f_{n,X}(x))-f_X(x)|\Big)^2 = \Big(\sup_{x\in\mathbb{R}}\Big|\int\tilde{K}(u)(f_X(x-h_nu)-f_X(x))du\Big|\Big)^2 = O(h_n^2).$$

Combining this with (3.3) concludes the proof of Lemma 3.1.

To state the next result, recall the definition of $\bar{\varepsilon}_n$ from (2.2), and let

$$\hat{Z}_i = \varepsilon_i - \hat{\varepsilon}_{ni}, \ i \in I_n, \quad \bar{W}_n = \frac{1}{n^p} \sum_{j \in I_n} \varepsilon_j \phi_{nj}, \quad Z_n = \bar{\varepsilon}_n + \bar{W}_n,$$
$$\gamma_n = \max\big(h_n^2, n^{-p/2} h_n^{-1/2}, n^{d-p/2} h_n^{1/2}, n^{d_X - p/2}\big).$$

We are now ready to state and prove

Lemma 3.2 *Assume ($\tilde{K}$), (B1)-(B3), (B4′), (C1), and that ϕ is a piecewise continuously differentiable function satisfying $\bar{\phi} = 0$. Then the following holds.*

$$\sum_{i \in I_n} \Big(\frac{f_{n,X}(X_i)}{f_X(X_i)} \hat{Z}_i - Z_n \Big)^2 = O_P(n^p \gamma_n^2). \tag{3.4}$$

Proof. Let $\tilde{K}_j(X_i) = \tilde{K}((X_i - X_j)/h_n)$. From (1.1) and (1.5), we obtain

$$\begin{aligned}
\frac{f_{n,X}(X_i)}{f_X(X_i)} \hat{Z}_i - Z_n &= \frac{f_{n,X}(X_i)}{f_X(X_i)} (\hat{\mu}_n(X_i) - \mu(X_i)) - Z_n \\
&= \frac{1}{n^p h_n} \sum_{j \in I_n} \varepsilon_j (1 + \phi_{nj}) \frac{\tilde{K}_j(X_i)}{f_X(X_i)} - Z_n \\
&\quad + \frac{1}{f_X(X_i)} \frac{1}{n^p h_n} \sum_{j \in I_n} (\mu(X_j) - \mu(X_i)) \tilde{K}_j(X_i) \\
&\quad + \frac{1}{f_X(X_i)} \frac{1}{n^p h_n} \sum_{j \in I_n} \mu(X_j) \phi_{nj} \tilde{K}_j(X_i).
\end{aligned} \tag{3.5}$$

In all sums below, the summation index varies over I_n, unless mentioned otherwise. Let

$$\xi_{ni} = \frac{1}{n^p h_n} \sum_j (\mu(X_j) - \mu(X_i)) \tilde{K}_j(X_i) = \frac{1}{n^p h_n} \sum_{j \neq i} (\mu(X_j) - \mu(X_i)) \tilde{K}_j(X_i).$$

Then

$$\begin{aligned}
&E(\xi_{ni}^2) \\
&\leq \frac{1}{n^{2p} h_n} \sum_{j \neq i} \int [\mu(x - h_n u) - \mu(x)]^2 \tilde{K}^2(u) f_{X,i,j}(x, x - h_n u) du dx \\
&\quad + \frac{1}{n^{2p}} \sum_{j \neq k, j, k \neq i} \int (\mu(x - h_n u) - \mu(x))(\mu(x - h_n v) - \mu(x)) \tilde{K}(u) \tilde{K}(v) \\
&\qquad\qquad \times f_{X,i,j,k}(x, x - h_n u, x - h_n v) dx du dv \\
&=: I_{1i} + I_{2i}.
\end{aligned}$$

By (B1), the fact $d_X < p/2$, and arguing as in the proof of (3.2),

$$
\begin{aligned}
(3.6)\ \sum_i I_{1i} &\le Cn^{-2p}h_n \sum_i \sum_{j\neq i} \int |u|^2 \tilde{K}^2(u) f_{X,i,j}(x, x-h_n u)dudx \\
&= Cn^{-2p}h_n \sum_i \sum_{j\neq i} \int |u|^2 \tilde{K}^2(u) f_X(x) f_X(x-h_n u)dxdu \\
&\quad +Cn^{-2p}h_n \sum_i \sum_{j\neq i} \int |u|^2 \tilde{K}^2(u) \\
&\qquad \times \big[f_{X,i,j}(x, x-h_n u) - f_X(x) f_X(x-h_n u)\big]dxdu \\
&= O(h_n) + O(n^{-p}h_n) + Cn^{-2p}h_n \sum_{\|j-i\|>N} \Big(\|j-i\|^{2d_X-p} + o(\|j-i\|^{2d_X-p})\Big) \\
&= O(h_n) + O(n^{2d_X-p}h_n) = O(h_n).
\end{aligned}
$$

Moreover, by $(\tilde{K})$ and (B1), since $\int u\tilde{K}(u)du = 0$, we obtain

$$
\begin{aligned}
I_{2i} \;\le\; & \frac{Ch_n^4}{n^{2p}} \sum_{j\neq k, j,k\neq i} \int |uv|^2 \tilde{K}(u)\tilde{K}(v) f_X(x) f_X(x-h_n u) f_X(x-h_n v) dxdudv \\
&+\frac{Ch_n^4}{n^{2p}} \sum_{j\neq k, j,k\neq i} \int |uv|^2 \tilde{K}(u)\tilde{K}(v) \big| f_{X,i,j,k}(x, x-h_n u, x-h_n v) \\
&\qquad\qquad -f_X(x) f_{X,j,k}(x-h_n u, x-h_n v)\big| dxdudv \\
&+\frac{Ch_n^4}{n^{2p}} \sum_{j\neq k, j,k\neq i} \int |uv|^2 \tilde{K}(u)\tilde{K}(v) f_X(x) \\
&\qquad \times \big| f_{X,j,k}(x-h_n u, x-h_n v) - f_X(x-h_n u) f_X(x-h_n v)\big| dxdudv \\
=:\; & R_{1i} + R_{2i} + R_{3i}.
\end{aligned}
$$

First note that

$$\sum_i R_{1i} = O(n^p h_n^4).$$

By Assumption (B3) and arguing as in the proof of (3.2),

$$
\begin{aligned}
\sum_i R_{2i} &\le Ch_n^4 + C'n^{-p}h_n^4 \Big\{ \sum_{\|j-i\|>N} \Big(\|j-i\|^{2d_X-p} + o(\|j-i\|^{2d_X-p})\Big) \\
&\qquad + \sum_{\|k-i\|>N} \Big(\|k-i\|^{2d_X-p} + o(\|k-i\|^{2d_X-p})\Big)\Big\} \\
&= O(h_n^4) + O(n^{2d_X} h_n^4).
\end{aligned}
$$

Similarly,

$$\sum_i R_{3i} = O(h_n^4) + O(n^{2d_X} h_n^4).$$

Combine these bounds with the fact $n^{2d_X} h_n^4 / n^p h_n^4 = n^{2d_X - p} \to 0$, to obtain

$$\sum_i I_{2i} = O(n^p h_n^4).$$

This bound together with (3.6) yields

$$E\Big(\sum_i \xi_{ni}^2\Big) = O(\max\{h_n, n^p h_n^4\}).$$

Then it follows from Assumption (C1) that

$$\sum_i \Big(\frac{1}{f_X(X_i)} \frac{1}{n^p h_n} \sum_{j \in I_n} (\mu(X_j) - \mu(X_i)) \tilde{K}_j(X_i)\Big)^2 = O_P(\max\{h_n, n^p h_n^4\}). \tag{3.7}$$

Next, consider the third term in the right hand side of (3.5). Note that

$$\begin{aligned}
& E\Big(\frac{1}{n^p h_n} \sum_j \mu(X_j) \phi_{nj} \tilde{K}_j(X_i)\Big)^2 \\
= {} & \frac{\phi_{ni}^2 \tilde{K}^2(0)}{n^{2p} h_n^2} E\mu^2(X_i) + \frac{1}{n^{2p} h_n^2} \sum_{j \neq i} \phi_{nj}^2 E\Big(\mu(X_j) \tilde{K}_j(X_i)\Big)^2 \\
& + \frac{2\phi_{ni} \tilde{K}(0)}{n^{2p} h_n^2} \sum_{j \neq i} \phi_{nj} E\Big(\mu(X_i) \mu(X_j) \tilde{K}_j(X_i)\Big) \\
& + \frac{1}{n^{2p} h_n^2} \sum_{j \neq k, j, k \neq i} \phi_{nj} \phi_{nk} E\Big(\mu(X_j) \tilde{K}_j(X_i) \mu(X_k) \tilde{K}_k(X_i)\Big) \\
= {} & Q_{0i} + Q_{1i} + Q_{2i} + Q_{3i}, \qquad \text{say.}
\end{aligned}$$

Standard kernel arguments yield that,

$$\begin{aligned}
\sum_i Q_{0i} &= O(n^{-p} h_n^{-2}), \qquad \sum_i Q_{2i} = O(h_n^{-1}), \\
\sum_i Q_{1i} &= \frac{1}{n^{2p} h_n^2} \sum_i \sum_{j \neq i} \phi_{nj}^2 \int \mu^2(y) \tilde{K}^2\Big(\frac{x-y}{h_n}\Big) f_{X,i,j}(x,y)\, dx dy = O(h_n^{-1}).
\end{aligned}$$

For the sake of brevity, let $dK(u,v) := \tilde{K}(u)\tilde{K}(v) du dv$. Then, by Assumption (B3) and $\bar{\phi} = 0$, we obtain

$$\begin{aligned}
Q_{3i} &= \frac{1}{n^{2p}} \sum_{j \neq k, j, k \neq i} \phi_{nj} \phi_{nk} \int \mu(x - h_n u) \mu(x - h_n v) \\
&\qquad \times f_{X,i,j,k}(x, x - h_n u, x - h_n v)\, dx dK(u,v) \\
&= \frac{1}{n^{2p}} \sum_{j \neq k, j, k \neq i} \phi_{nj} \phi_{nk} \int \mu(x - h_n u) \mu(x - h_n v) \\
&\quad \times \{f_{X,i,j,k}(x, x - h_n u, x - h_n v) - f_X(x) f_{X,j,k}(x - h_n u, x - h_n v)\}\, dx dK(u,v) \\
&\quad + \frac{1}{n^{2p}} \sum_{j \neq k, j, k \neq i} \phi_{nj} \phi_{nk} \int \mu(x - h_n u) \mu(x - h_n v) f_X(x) \\
&\quad \times \{f_{X,j,k}(x - h_n u, x - h_n v) - f_X(x - h_n u) f_X(x - h_n v)\}\, dx dK(u,v)
\end{aligned} \tag{3.8}$$

$$
\begin{aligned}
& +\frac{1}{n^{2p}} \sum_{j\neq k, j,k\neq i} \phi_{nj}\phi_{nk} \int \mu(x-h_n u)\mu(x-h_n v) \\
& \qquad\qquad\qquad \times f_X(x) f_X(x-h_n u) f_X(x-h_n v)\, dx dK(u,v) \\
\leq\ & C n^{-2p}\Big[n^p + \sum_{j\in I_n; \|j-i\|>N} \Big(\|j-i\|^{2d_X-p} + o(\|j-i\|^{2d_X-p})\Big) \\
& \qquad + \sum_{k\in I_n; \|k-i\|>N} \Big(\|k-i\|^{2d_X-p} + o(\|k-i\|^{2d_X-p})\Big) \\
& \qquad + \sum_{\|j-k\|>N} \Big(\|j-k\|^{2d_X-p} + o(\|j-k\|^{2d_X-p})\Big)\Big] \\
& \qquad + \bar{\phi}^2 \int \mu^2(x) f_X^3(x) dx (1+O(h_n)) + O(n^{-p}) \\
=\ & O(n^{2d_X-p}), \ \text{ not depending on } i.
\end{aligned}
$$

This implies that

$$
\sum_i \Big(\frac{1}{f_X(X_i)} \frac{1}{n^p h_n} \sum_{j\in I_n} \mu(X_j)\phi_{nj}\tilde{K}_j(X_i)\Big)^2 = O_P(\max\{h_n^{-1}, n^{2d_X}\}). \tag{3.9}
$$

Now consider the first term in the right hand side of (3.5). Let

$$
D_{ij} = \frac{1}{h_n} E\Big(\frac{\tilde{K}_j(X_i)}{f_X(X_i)}\Big) - 1, \ \ j\in I_n.
$$

We have

$$
\begin{aligned}
& \frac{1}{n^p h_n}\sum_j \varepsilon_j(1+\phi_{nj})\frac{\tilde{K}_j(X_i)}{f_X(X_i)} - Z_n \\
=\ & \frac{1}{n^p}\frac{\tilde{K}(0)}{h_n f_X(X_i)}\varepsilon_i(1+\phi_{ni}) + \frac{1}{n^p h_n}\sum_{j\neq i}\varepsilon_j(1+\phi_{nj})\frac{\tilde{K}_j(X_i)}{f_X(X_i)} - Z_n \\
=\ & \frac{1}{n^p}\Big(\frac{\tilde{K}(0)}{h_n f_X(X_i)} - 1\Big)\varepsilon_i(1+\phi_{ni}) + \frac{1}{n^p}\sum_{j\neq i}\varepsilon_j(1+\phi_{nj})D_{ij} \\
& + \frac{1}{n^p h_n}\sum_{j\neq i}\varepsilon_j(1+\phi_{nj})\Big[\frac{\tilde{K}_j(X_i)}{f_X(X_i)} - E\Big(\frac{\tilde{K}_j(X_i)}{f_X(X_i)}\Big)\Big] \\
=\ & A_{0i} + A_{1i} + A_{2i}, \quad \text{say.}
\end{aligned} \tag{3.10}
$$

First, by Assumption (C1),

$$
\sum_i A_{0i}^2 = \frac{1}{n^{2p}}\sum_i \Big(\frac{\tilde{K}(0)}{h_n f_X(X_i)} - 1\Big)^2 \varepsilon_i^2(1+\phi_{ni})^2 = O_P(n^{-p}h_n^{-2}). \tag{3.11}
$$

Now consider the term A_{1i}. Note that, for $j\neq i$,

$$
\begin{aligned}
D_{ij} &= \int \frac{\tilde{K}(u)}{f_X(x)} f_{X,i,j}(x, x - h_n u) du dx - 1 \\
&= \int \frac{\tilde{K}(u)}{f_X(x)} f_X(x) f_X(x - h_n u) du dx - 1 \\
&\quad + \int \frac{\tilde{K}(u)}{f_X(x)} \big[f_{X,i,j}(x, x - h_n u) - f_X(x) f_X(x - h_n u) \big] du dx \\
&= O(h_n) + \int \frac{\tilde{K}(u)}{f_X(x)} \big[f_{X,i,j}(x, x - h_n u) - f_X(x) f_X(x - h_n u) \big] du dx.
\end{aligned}
$$

Then

$$
\begin{aligned}
E(A_{1i}^2) &= \frac{1}{n^{2p}} \sum_{j \neq i} (1 + \phi_{nj})^2 D_{ij}^2 E\varepsilon_j^2 + \frac{1}{n^{2p}} \sum_{j \neq k, j, k \neq i} (1 + \phi_{nj})(1 + \phi_{nk}) D_{ij} D_{ik} E(\varepsilon_j \varepsilon_k) \\
&= A_{11i} + A_{12i}.
\end{aligned}
$$

But,

$$
\begin{aligned}
A_{11i} &\leq \frac{C}{n^{2p}} \sum_{j \neq i} D_{ij}^2 = \frac{C}{n^{2p}} \Big(\sum_{j \in I_n; 0 < \|j-i\| \leq N} + \sum_{j \in I_n; \|j-i\| > N} \Big) D_{ij}^2 \\
&\leq O(n^{-2p}) + \frac{1}{n^{2p}} \sum_{j \in I_n; \|j-i\| > N} \Big(O(h_n^2) + C \Big(\|j-i\|^{2d_X - p} + o(\|j-i\|^{2d_X - p}) \Big)^2 \Big) \\
&= O(n^{-2p}) + O(n^{-p} h_n^2) + O(n^{4d_X - 3p}), \qquad \forall\, i \in I_n, \\
\sum_i A_{11i} &= O(n^{-p}) + O(h_n^2) + O(n^{4d_X - 2p}).
\end{aligned}
$$

For the sake of brevity, we omit the $o(\|j-i\|^{2d_X-p})$, $o(\|k-i\|^{2d_X-p})$ and $o(\|j-k\|^{2d_X-p})$ terms in the following derivations.

$$
\begin{aligned}
A_{12i} &\leq \frac{C}{n^{2p}} \sum_{j \neq k, j, k \neq i} |D_{ij} D_{ik}| |\gamma(j-k)| \\
&\leq \frac{C}{n^{2p}} \Big[n^p + \sum_{\|j-k\| > N, j, k \neq i} \Big(h_n + 1 + \|i-j\|^{2d_X - p} \Big) \\
&\qquad\qquad \times \Big(h_n + \|i-k\|^{2d_X - p} \Big) \|j-k\|^{2d-p} \Big] \\
&\leq C n^{-p} + \frac{C}{n^{2p}} \sum_{\|j-k\| > N, j, k \neq i} \Big\{ h_n \|j-k\|^{2d-p} + \|j-k\|^{2d-p} \|i-k\|^{2d_X - p} \\
&\qquad + h_n \|j-k\|^{2d-p} \|i-j\|^{2d_X - p} + \|j-k\|^{2d-p} \|i-j\|^{2d_X - p} \|i-k\|^{2d_X - p} \Big\} \\
&= O(n^{2d-p} h_n) + O(n^{2d + 2d_X - 2p}) + O(n^{2d + 4d_X - 3p}), \qquad \forall\, i \in I_n, \\
\sum_i A_{12i} &= O(n^{2d} h_n) + O(n^{2d + 2d_X - p}) + O(n^{2d + 4d_X - 2p}).
\end{aligned}
$$

Hence, by the fact that $n^{2d+2d_X-p}/n^{2d}h_n \to 0$,

$$\sum_i A_{1i}^2 \;=\; O_P(n^{2d}h_n). \tag{3.12}$$

Next, we shall show that

$$\sum_i A_{2i}^2 \;=\; O_P(\max\{h_n^{-1}, n^{2d}h_n\}). \tag{3.13}$$

Note that

$$\begin{aligned}
A_{2i}^2 \;&=\; \frac{1}{n^{2p}h_n^2}\sum_{j\neq i}\varepsilon_j^2(1+\phi_{nj})^2\Big(\frac{\tilde K_j(X_i)}{f_X(X_i)} - E\Big(\frac{\tilde K_j(X_i)}{f_X(X_i)}\Big)\Big)^2\\
&\quad + \frac{1}{n^{2p}h_n^2}\sum_{j\neq k, j,k\neq i}(1+\phi_{nj})(1+\phi_{nk})\varepsilon_j\varepsilon_k\\
&\qquad\quad \times\Big(\frac{\tilde K_j(X_i)}{f_X(X_i)} - E\Big(\frac{\tilde K_j(X_i)}{f_X(X_i)}\Big)\Big)\Big(\frac{\tilde K_k(X_i)}{f_X(X_i)} - E\Big(\frac{\tilde K_k(X_i)}{f_X(X_i)}\Big)\Big)\\
&=:\; A_{21i} + A_{22i}.
\end{aligned}$$

Fix an $i \in I_n$, by Assumption ($\tilde{\mathrm{K}}$), we obtain

$$\begin{aligned}
&E\Big(\frac{1}{n^{2p}h_n^2}\sum_{j\neq i}\varepsilon_j^2(1+\phi_{nj})^2\tilde K_j^2(X_i)\Big)\\
&\quad = n^{-2p}h_n^{-1}E(\varepsilon_1^2)\sum_{j\neq i}(1+\phi_{nj})^2\int \tilde K^2(u) f_{X,i,j}(x, x-h_nu)\,du\,dx = O(n^{-p}h_n^{-1}).
\end{aligned}$$

Then by Assumption (C1),

$$\sum_i A_{21i} = O_P(h_n^{-1}).$$

Moreover, by the independence of $\{\varepsilon_j\}$ and $\{X_j\}$,

$$E(A_{22i}) \;\le\; n^{-2p}h_n^{-2}\sum_{j\neq k, j,k\neq i} E(\varepsilon_j\varepsilon_k)\mathrm{Cov}\Big(\frac{\tilde K_j(X_i)}{f_X(X_i)}, \frac{\tilde K_k(X_i)}{f_X(X_i)}\Big).$$

Arguing as for (3.8), by Assumptions (B2) and (B3),

$$\begin{aligned}
&h_n^{-2}\sum_{j\neq k, j,k\neq i} E(\varepsilon_j\varepsilon_k)\mathrm{Cov}\Big(\frac{\tilde K_j(X_i)}{f_X(X_i)}, \frac{\tilde K_k(X_i)}{f_X(X_i)}\Big)\\
&\quad = \sum_{j\neq k, j,k\neq i}|\gamma(j-k)|\Big|\int \frac{\tilde K(u)\tilde K(v)}{f_X^2(x)} f_{X,i,j,k}(x, x-h_nu, x-h_nv)\,dx\,du\,dv\\
&\qquad\qquad - \int \frac{\tilde K(u)}{f_X(x)} f_{X,i,j}(x, x-h_nu)\,dx\,du \int \frac{\tilde K(v)}{f_X(x)} f_{X,i,k}(x, x-h_nv)\,dx\,dv\Big|
\end{aligned}$$

$$
\begin{aligned}
\leq\ & C\Big[n^p + \sum_{\|j-k\|>N, j,k\neq i} \|j-k\|^{2d-p}\Big\{C'\|i-j\|^{2d_X-p} + C'\|i-k\|^{2d_X-p} \\
&\qquad + \int \frac{\tilde K(u)\tilde K(v)}{f_X^2(x)} f_X(x) f_{X,j,k}(x-h_nu, x-h_nv)\,dxdudv \\
&\qquad - \Big(C\|i-j\|^{2d_X-p} + \int \frac{\tilde K(u)}{f_X(x)} f_X(x) f_X(x-h_nu)\,dxdu\Big) \\
&\qquad\qquad \times\Big(C\|i-k\|^{2d_X-p} + \int \frac{\tilde K(v)}{f_X(x)} f_X(x) f_X(x-h_nv)\,dxdv\Big)\Big\}\Big] \\
\leq\ & O(n^p) + C\sum_{\|j-k\|>N, j,k\neq i} \|j-k\|^{2d-p}\Big\{C'\|i-j\|^{2d_X-p} + C'\|i-k\|^{2d_X-p} \\
&\qquad + C\|j-k\|^{2d_X-p} + \int \frac{\tilde K(u)\tilde K(v)}{f_X(x)} f_X(x-h_nu) f_X(x-h_nv)\,dxdudv \\
&\qquad - \Big(C\|i-j\|^{2d_X-p} + \int \tilde K(u) f_X(x-h_nu)\,dxdu\Big) \\
&\qquad\qquad \times\Big(C\|i-k\|^{2d_X-p} + \int \tilde K(v) f_X(x-h_nv)\,dxdv\Big)\Big\} \\
\leq\ & O(n^p) + C\Big\{\sum_{\|j-k\|>N, j,k\neq i} \|j-k\|^{2d-p}\|i-j\|^{2d_X-p} + \sum_{\|j-k\|>N} \|j-k\|^{2(d+d_X)-2p} \\
&\quad + \sum_{\|j-k\|>N, j,k\neq i} \|j-k\|^{2d-p}\|i-k\|^{2d_X-p} + \sum_{\|j-k\|>N} \|j-k\|^{2d-p} O(h_n) \\
&\quad + \sum_{\|j-k\|>N, j,k\neq i} \|j-k\|^{2d-p}\|i-j\|^{2d_X-p}\|i-k\|^{2d_X-p}\Big\} \\
=\ & O(n^{2(d+d_X)}) + O(n^{2d+4d_X-p}) + O(n^{2d+p}h_n) \\
=\ & O(n^{2(d+d_X)}) + O(n^{2d+p}h_n).
\end{aligned}
$$

This implies the claim in (3.13).

Assumptions (B4′) combined with (3.10) to (3.13) yield

$$
\sum_i \Big(\frac{1}{n^p h_n}\sum_{j\in I_n} \varepsilon_j(1+\phi_{nj})\frac{\tilde K_j(X_i)}{f_X(X_i)} - Z_n\Big)^2 = O_P(\max\{h_n^{-1}, n^{2d}h_n\}), \tag{3.14}
$$

Claim (3.4) is now established upon combining (3.5), (3.7), (3.9) and (3.14), thereby completing the proof of Lemma 3.2.

Let $\xi_n = \max(h_n^2, n^{-p/2}h_n^{-1/2}, n^{d-p/2}h_n^{1/2}, n^{d_X-p/2}, n^{d+d_X-p}h_n^{-1})$.

Proof of Theorem 2.2. (i). The proof here uses the argument of the proof of Theorem 2.2 of Koul et al. (2013) . Recall that $\hat Z_i = \varepsilon_i - \hat\varepsilon_{ni}$, and $Z_n = \bar\varepsilon_n + \bar W_n$. Lemmas 3.1, 3.2, and Assumption (C1), imply

$$\sum_{i\in I_n}\Big(\frac{f_{n,X}(X_i)}{f_X(X_i)}\hat{Z}_i - Z_n\Big)^2 = O_P(n^p\gamma_n^2),$$
$$\max_{i\in I_n}\Big|\frac{f_X(X_i)}{f_{n,X}(X_i)}-1\Big| = O_P(\max\{n^{d_X-p/2}h_n^{-1}, h_n\}).$$

Moreover, by Koul and Surgailis (2013),

$$n^{p/2-d}\bar{W}_n \to_D W = c(\phi)Z. \tag{3.15}$$

It follows that $Z_n = O_P(n^{d-p/2})$ and,

$$\begin{aligned}\sum_{i\in I_n}(\hat{Z}_i - Z_n)^2 &= O_P(n^p\gamma_n^2) + O_P(n^{2d})O_P(\max\{n^{2d_X-p}h_n^{-2}, h_n^2\}) \\ &= O_P(n^p\xi_n^2).\end{aligned} \tag{3.16}$$

Let $\tilde{f}_n(x) = \frac{1}{n^p}\sum_{i\in I_n}K_{b_n}(x+Z_n-\varepsilon_i)$, $\psi_{n1}(x) = \hat{f}_n(x) - \tilde{f}_n(x)$, and let

$$\begin{aligned}\psi_{n2}(x) &= \frac{1}{b_n}\int\big[F_n(x+Z_n-ub_n) - F(x+Z_n-ub_n) + f(x+Z_n-ub_n)\bar{\varepsilon}_n\big] \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \times K'(u)du, \\ \psi_{n3}(x) &= \frac{1}{b_n}\int\big[F(x+Z_n-ub_n) - F(x-ub_n) - f(x+Z_n-ub_n)Z_n\big] \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \times K'(u)du, \\ \psi_{n4}(x) &= (Z_n-\bar{\varepsilon}_n)\frac{1}{b_n}\int\big[f(x+Z_n-ub_n) - f(x-ub_n)\big]K'(u)du, \\ \psi_{n5}(x) &= \bar{W}_n\frac{1}{b_n}\int f(x-ub_n)K'(u)du.\end{aligned}$$

Integration by parts yields

$$\begin{aligned}&\hat{f}_n(x) - K_{b_n}*f(x) \\ &\quad= \hat{f}_n(x) - \tilde{f}_n(x) + \tilde{f}_n(x) - K_{b_n}*f(x) \\ &\quad= \hat{f}_n(x) - \tilde{f}_n(x) + \int K_{b_n}(x+Z_n-y)dF_n(y) - \int K_{b_n}(x-y)dF(y) \\ &\quad= \hat{f}_n(x) - \tilde{f}_n(x) + \frac{1}{b_n}\int(F_n(x+Z_n-ub_n) - F(x-ub_n))K'(u)du \\ &\quad= \sum_{i=1}^5\psi_{ni}(x).\end{aligned}$$

Observe that, with $\Delta_i := \hat{Z}_i - Z_n$, $b := b_n$,

$$\begin{aligned}&\int\psi_{n1}^2(x)dx \\ &\quad= \frac{1}{n^{2p}b^2}\int\Big[\sum_{i\in I_n}\Big\{K(\frac{x+\hat{Z}_i-\varepsilon_i}{b}) - K(\frac{x+Z_n-\varepsilon_i}{b})\Big\}\Big]^2dx\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{n^{2p}b^2}\int\Big[\sum_{i\in I_n}\Big\{\frac{\Delta_i}{b}K'\Big(\frac{x+Z_n-\varepsilon_i}{b}\Big)+R_i(x)\Big\}\Big]^2dx\\
&\le \frac{2}{n^{2p}b^2}\Big\{\int\Big[\sum_{i\in I_n}\frac{\Delta_i}{b}K'\Big(\frac{x+Z_n-\varepsilon_i}{b}\Big)\Big]^2dx+\int\Big[\sum_{i\in I_n}R_i(x)\Big]^2dx\Big\}\\
&= \frac{2}{n^{2p}b^2}\{A+B\},\qquad \text{say.}
\end{aligned}
$$

Here

$$
\begin{aligned}
R_i(x) &:= K(\frac{x+\hat{Z}_i-\varepsilon_i}{b})-K(\frac{x+Z_n-\varepsilon_i}{b})-\frac{\hat{Z}_i-Z_n}{b}K'\Big(\frac{x+Z_n-\varepsilon_i}{b}\Big)\\
&= \int_0^{\Delta_i/b}(\frac{\Delta_i}{b}-s)K''(\frac{x+Z_n-\varepsilon_i}{b}+s)ds.
\end{aligned}
$$

Note that

$$
|R_i(x)| \;\le\; \frac{|\Delta_i|}{b}\int_{-|\Delta_i|/b}^{|\Delta_i|/b}\big|K''(\frac{x+Z_n-\varepsilon_i}{b}+s)\big|ds,\quad \forall\, i\in I_n, x\in\mathbb{R}.
$$

Hence, with $\|K''\|_\infty := \sup_{x\in\mathbb{R}}|K''(x)|$,

$$
\begin{aligned}
B &\le \sum_{i\in I_n,j\in I_n}\frac{|\Delta_i\Delta_j|}{b^2}\int_{-|\Delta_i|/b}^{|\Delta_i|/b}\int_{-|\Delta_j|/b}^{|\Delta_j|/b}\int\big|K''(\frac{x+Z_n-\varepsilon_i}{b}+s)\big|\\
&\qquad\qquad\times\big|K''(\frac{x+Z_n-\varepsilon_j}{b}+t)\big|dxdsdt\\
&\le 2\|K''\|_\infty\sum_{i\in I_n,j\in I_n}\frac{|\Delta_i|\Delta_j^2}{b^3}\int_{-|\Delta_i|/b}^{|\Delta_i|/b}\int\big|K''(\frac{x+Z_n-\varepsilon_i}{b}+s)\big|dxds\\
&\le 4\|K''\|_\infty\sum_{i\in I_n,j\in I_n}\frac{\Delta_i^2\Delta_j^2}{b^3}\int|K''(u)|du.
\end{aligned}
$$

Next, similarly,

$$
\begin{aligned}
A &= \sum_{i\in I_n}\frac{\Delta_i^2}{b^2}\int K'^2(\frac{x+Z_n-\varepsilon_i}{b})dx\\
&\qquad +\sum_{i\ne j}\frac{\Delta_i\Delta_j}{b^2}\int K'(\frac{x+Z_n-\varepsilon_i}{b})K'(\frac{x+Z_n-\varepsilon_j}{b})dx\\
&\le \sum_{i\in I_n}\frac{\Delta_i^2}{b}\int K'^2(u)du+\|K'\|_\infty\sum_{i\ne j}\frac{|\Delta_i||\Delta_j|}{b}\int|K'(u)|du\\
&\le C\,b^{-1}\Big(\sum_{i\in I_n}|\Delta_i|\Big)^2.
\end{aligned}
$$

Hence,

$$\begin{aligned}\int \psi_{n1}^2(x)dx &\le \frac{C}{n^{2p}b^2}\Big\{b^{-1}\Big(\sum_{i\in I_n}|\Delta_i|\Big)^2 + b^{-3}\Big(\sum_{i\in I_n}\Delta_i^2\Big)^2\Big\}\\ &\le \frac{C}{n^{2p}b^2}\Big\{n^p b^{-1}\Big(\sum_{i\in I_n}\Delta_i^2\Big) + b^{-3}\Big(\sum_{i\in I_n}\Delta_i^2\Big)^2\Big\}.\end{aligned}$$

It follows from (3.16) and Assumption (B4) that

$$\int \psi_{n1}^2(x)dx = O_P(\xi_n^2 b_n^{-3}) + O_P(\xi_n^4 b_n^{-5}) = o_P(n^{2d-p}).$$

Next, following the proof of Theorem 2.2 of Koul et al. (2013) and using Lemma 6.1 of Koul and Surgailis (2013), we obtain

$$\begin{aligned}\int \psi_{n2}^2(x)dx &= O_P(n^{4d-2p}b_n^{-1}), \qquad p/4 < d < p/2,\\ &= O_P(n^{-p}b_n^{-1}), \qquad 0 < d \le p/4,\\ \int \psi_{n3}^2(x)dx &= O_P(n^{4d-2p}), \qquad \int \psi_{n4}^2(x)dx = O_P(n^{4d-2p}),\end{aligned}$$

and, under the null hypothesis, by (3.15),

$$n^{p-2d}\int \psi_{n5}^2(x)dx \to_D \kappa_1 W^2.$$

This concludes the proof of (i).

(ii). From the proof of (i), we obtain that

$$\begin{aligned}&\hat{T}_n - \int [K_{b_n} * (f - f_0)(x)]^2 dx\\ &= \int (\hat{f}_n - K_{b_n} * f(x))^2 dx + 2\int (\hat{f}_n - K_{b_n} * f(x)) K_{b_n} * (f - f_0)(x)dx\\ &= O_P(n^{2d-p}) + 2\sum_{i=1}^{5}\int \psi_{ni}(x) K_{b_n} * (f - f_0)(x)dx\end{aligned}$$

Note that, for $i = 1, 2, 3, 4$,

$$\int \psi_{ni}(x)K_{b_n} * (f - f_0)(x)dx \le \Big(\int \psi_{ni}^2(x)dx \int [K_{b_n} * (f - f_0)(x)]^2 dx\Big)^{1/2} = o_P(n^{d-p/2}),$$

and

$$\begin{aligned}&n^{p/2-d}\int \psi_{n5}(x)K_{b_n} * (f - f_0)(x)dx\\ &= n^{p/2-d}\bar{W}_n \int f'(x - b_n u)K(u)K(v)(f - f_0)(x - b_n v)dudvdx \to_D \kappa_2 W.\end{aligned}$$

This completes the proof of Theorem 2.2. □

Lemma 3.3 *Suppose that ($\tilde{K}$), (B1)-(B3), (B4′), (C) hold,* $E|\zeta_0|^3 < \infty$*,* $q = o(n^{1/2})$*, and* ϕ *is a piecewise continuously differentiable function satisfying* $\bar{\phi} = 0$*. Then*

$$\hat{c}^2(\phi) \to_P c^2(\phi).$$

Proof. As discussed above, $\hat{d}$ is $\log(n)$-consistent estimator of d. Hence, to prove Lemma 3.3, it suffices to show that

$$\tilde{c}^2(\phi) \to_P c^2(\phi),$$

where $\tilde{c}^2(\phi) = q^{-p-2d} \sum_{j,k \in I_q} \phi_{qj}\phi_{qk}\hat{\gamma}(j-k)$.

We will follow the argument as in the proof of Lemma 2.2 of Koul and Surgailis (2013). Write $\tilde{c}^2(\phi) = \tilde{c}_1^2(\phi) + \tilde{c}_2^2(\phi)$, $\tilde{c}_i^2(\phi) = q^{-p-2d} \sum_{j,k \in I_q} \phi_{qj}\phi_{qk}\hat{\gamma}_i(j-k)$, $i = 1, 2$, where

$$\hat{\gamma}_1(k) = \frac{1}{n^p} \sum_{i \in I_n} \varepsilon_i \varepsilon_{i+k}, \qquad \hat{\gamma}_2(k) := \hat{\gamma}(k) - \hat{\gamma}_1(k).$$

By Lemma 2.2 of Koul and Surgailis (2013), $\tilde{c}_1^2(\phi) \to_P c^2(\phi)$. Hence, Lemma 3.3 will follow if we show that

$$n^{-p}q^{-p-2d} \sum_{j,k \in I_q} \sum_{i \in I_n} \phi_{qj}\phi_{q,j+k}\varepsilon_i \hat{Z}_{i+k} = o_P(1), \tag{3.17}$$

$$n^{-p}q^{-p-2d} \sum_{j,k \in I_q} \sum_{i \in I_n} \phi_{qj}\phi_{q,j+k}\varepsilon_{i+k} \hat{Z}_{i} = o_P(1), \tag{3.18}$$

$$n^{-p}q^{-p-2d} \sum_{j,k \in I_q} \sum_{i \in I_n} \phi_{qj}\phi_{q,j+k}\hat{Z}_i \hat{Z}_{i+k} = o_P(1). \tag{3.19}$$

(3.16) with Assumption (B4′) implies the bound

$$\sum_i \hat{Z}_i^2 = O_P(n^p \xi_n^2) + O_P(n^{2d}) = O_P(n^{2d}).$$

Then we obtain

$$\begin{aligned} & n^{-p}q^{-p-2d} \sum_{j,k \in I_q} \sum_{i \in I_n} \phi_{qj}\phi_{q,j+k}\varepsilon_i \hat{Z}_{i+k} \\ \le\ & Cn^{-p}q^{p-2d} \Big\{ \sum_{i \in I_n} \varepsilon_i^2 \sum_{i \in I_n} \hat{Z}_{i+k}^2 \Big\}^{1/2} \\ =\ & Cn^{-p}q^{p-2d} n^{p/2} n^d = O_P(q^{p-2d} n^{d-p/2}) = o_P(1). \end{aligned}$$

This implies (3.17), while (3.18) follows in a similar way. Finally, the left hand side of (3.19) is bounded from the above by

$$Cn^{-p}q^{p-2d} \sum_{i \in I_n} \hat{Z}_i^2 = Cn^{-p}q^{p-2d} n^{2d} = O_P((n/q)^{2d-p}) = o_P(1).$$

This completes the proof of Lemma 3.3.

4 A Simulation Study

This section contains results from a simulation study investigating empirical size of the test based on Theorem 2.2(i). In this simulation we chose $p = 2$, $n = 32, 64, 128$, the values of the long memory parameter $d = 0.2, 0.4, 0.6,\ 0.8$, and the asymptotic sizes $\alpha = 0.05,$ and 0.1. We wish to test $H_0 : f = f_0$, versus the alternative $H_a : f \neq f_0$, where f_0 is the normal density with zero mean and variance of 1.5. As mentioned at end of Section 2, the test that rejects H_0 when $n^{p-2\hat{d}}\hat{T}_n/\kappa_1\hat{c}^2(\phi) > \chi^2_\alpha(1)$, has the asymptotic size of α. The purpose of this simulations is to see how good is this approximation.

We shall first describe how we generated the two random fields ε_i and X_i, $i \in I_n$. The error random field ε_i of (1.2) are generated as follows. First, we chose $a_0 := 1$, and $a_j = c\,\|j\|^{d-2}$, for $j \in I_n, j \neq 0$, where the mechanism for choosing $c > 0$ will be described shortly. In other words $B(j/\|j\|) = c$ in the definition of a_j of (1.2). Now, let $A := \{a_j\}$, $j \in \mathbb{Z}^2$, $B := \{\zeta_{i-j}\}$, $i, j \in \mathbb{Z}^2$ denote the two infinite-size matrices. Then ε_i of (1.2) can be viewed as the sum of all elements after element-wise multiplication of matrices A and B. Let $A_{k_1;k_2}$ and $B_{k_1;k_2}$ denote the k_1th row and k_2th column element of matrices A and B, respectively. Then one can rewrite

$$\varepsilon_i = \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} A_{j_1;j_2} B_{i_1-j_1;i_2-j_2}.$$

Because we use computer, the matrices A and B need to be approximated by some finite dimensional matrices. Let $C := [-1000, 1000]^2 \subset \mathbb{Z}^2$ and $\tilde{A} := \{a_j\}$, $j \in C$. Note that all elements of A not in $\tilde{A}$ are of small magnitude, e.g., if $d = 0.2$ and $j = (0, 1000)'$, then $a_j = c(3.98)10^{-6}$. Hence we approximate ε_i by

$$\sum_{j_1=-1000}^{1000} \sum_{j_2=-1000}^{1000} \tilde{A}_{j_1;j_2} B_{i_1-j_1;i_2-j_2}.$$

This expression tells us how large a B we have to generate. For example, if $i = (0, 0)'$, then we only need B on $[-1000; 1000]^2$ lattice. On the other hand, if $i = (64, 64)'$, then generating B on the lattice $[-936, 1064]^2$ is required. Thus, in general, to generate ε_i for $i \in I_n$ using $\tilde{A}$, we need to generate B only on the lattice $[-1000, 1000 + n]^2$.

To summarize, to generate $\{\varepsilon_i,\ i \in I_n\}$ for this simulation study, we first generate i.i.d. standardized normal r.v.'s ζ_j, $j \in C$, and then calculate

$$\varepsilon_i = \sum_{j \in C} a_j \zeta_{i-j}, \quad i \in I_n.$$

The value of c was adjusted for each value of d to make variance of ε_i equal to 1.5. The covariate random field $X_i, i \in I_n$ is generated similarly. Since the asymptotic distribution of the test statistic does not depend on d_X, it is chosen to be 0.1, with the corresponding $c = 0.28$. Note that $d_X = 0.1$ and the above choices of d satisfy the constraint that $d_X < d$.

Once ε_i and X_i are generated, we then generate response field Y_i using the model

$$Y_i = 3\sin(X_i) + \varepsilon_i, \quad i \in I_n.$$

Consistent with the assumption (B4), the bandwidth h was chosen to be $h = 10n^{-0.85}$. Nonparametric regression was performed and residuals were calculated, using Normal kernel and $\phi_{ni} = -I\{i_1 \in [1, n/2]\} + I\{i_1 \in (n/2, n]\}$, $i = (i_1, i_2) \in I_n$.

Using these residuals, the GPH-estimator $\hat{d}$ was calculated with the bandwidth $m = n^{2/5}$. Then using the residuals and $\hat{d}$, $\hat{c}^2(\phi)$ were calculated with the bandwith $q = \sqrt{n}$. To calculate $\hat{T}_n$, Normal kernel and the data-dependent bandwith $b_n = (4/3)\hat{\sigma}_\varepsilon n^{-1/5}$ was used to obtain density estimate, where $\hat{\sigma}_\varepsilon$ is the sample standard deviation of the residuals.

Table 1 shows the empirical size of the test for various value of n and d, obtained from 300 iterations. The empirical size of the test shows reasonable convergence to the asymptotic value, except for the case $d = 0.8$.

n	d=0.2	0.4	0.6	0.8	d=0.2	0.4	0.6	0.8
32	0.027	0.027	0.030	0.023	0.040	0.047	0.050	0.030
64	0.053	0.053	0.050	0.030	0.077	0.080	0.060	0.046
128	0.060	0.057	0.047	0.027	0.107	0.087	0.087	0.050

Table 1: Empirical size of $\hat{T}_n$-test with asymptotic size $\alpha = 0.05$(left) and 0.1 for various values of d and n.

Acknowledgements

Research of the first author is supported in part by the NSF DMS grant 1205271. Work was completed while the second author was visiting Michigan State University in AY 2013-2014. Research of the second author is supported in part by National Natural Science Foundation of China (NSFC) Grant (No.11171147), the funds from Qing Lan Project, Jiangsu Province, and the Cultivation Fund of the Key Scientific and Technical Innovation Project, Ministry of Education of China (No.708044).

References

[1] Bachmann, D. and Dette, H. (2005). A note on the Bickel-Rosenblatt test in autoregressive time series. *Statist. Probab. Lett.*, **74**, 221–234.

[2] Beran, J., Feng, Y., Ghosh, S. and Kulik, R. (2013). *Long-memory processes. Probabilistic properties and statistical methods.* Springer, Hiedelberg, Germany.

[3] Bickel, P.J. and Rosenblatt, M. (1973). On some global measures of the deviations of density function estimators. *Ann. Statist.*, **6**, 1071–1095.

[4] Bierens, H.J. (1983). Uniform consistency of kernel estimators of a regression function under generalized conditions. *J. Amer. Statist. Assoc.*, **77**, 699–707.

[5] Carlos-Davila, E., Mejia-Lira, F. and Moran-Lopez, J.L. (1985). Ferromagnetism and spatial long range order in binary alloys: Systems with face-centered cubic structure. *J. Phys. C: Solid State Phys.*, **18**, 1217–1224.

[6] Chebana, F. (2004). On the optimization of the weighted Bickel- Rosenblatt test. *Statist. Probab. Lett.*, **68**, 333–345.

[7] Cheng, F. and Sun, S. (2008). A goodness-of-fit test of the errors in nonlinear autoregressive time series models. *Statist. Probab. Lett.*, **78**, 50–59.

[8] Dobrushin, R.L. and Major, P. (1979). Non-central limit theorems for non-liear functions of Gaussian fields. *Z. Wahrsch. Verw. Geb.*, **50**, 27–52.

[9] Doukhan, P., Lang, G. and Surgailis, D. (2002). Asymptotics of weighted empirical processes of linear fields with long-range dependence. *Ann. Inst. H. Poincaré Probab. Statist.*, **38**, 879–896.

[10] Frias, M.P., Alonso, F.J., Ruiz-Medina, M.D. and Angulo, J.M. (2008). Semiparametric estimation of spatial long-range dependence. *J. Statist. Plann. Inference*, **138**, 1479–1495.

[11] Ghosh, B.K. and Huang, W.M. (1991). The power and optimal kernel of the Bickel-Rosenblatt test for goodness of fit. *Ann. Statist.*, **19**, 999–1009.

[12] Giraitis, L., Koul, H.L. and Surgailis, D. (1996). Asymptotic normality of regression estimators with long memory errors. *Statist. Probab. Lett.*, **29**, 317–335.

[13] Giraitis, L., Koul, H.L. and Surgailis, D. (2012). *Large sample inference for long memory processes.* Imperial College Press, London, UK.

[14] Gneiting, T. (2000). Power-law correlations, related models for long range dependence, and their simulation. *J. Appl. Prob.*, **37**, 1104–1109.

[15] Guégan, D. (2005). How can we define the concept of long memory? An econometric survey. *Econometric Rev.*, **24**, 113–149.

[16] Guo, H. and Koul, H.L. (2008). Asymptotic inference in some heteroscedastic regression models with long memory design and errors. *Ann. Statist.*, **36**, 458–487.

[17] Hurvich, C.M., Deo, R. and Brodsky, J. (1998). The mean squared error of Geweke and Porter-Hudak's estimator of a long-memory time series. *J. Time Ser. Anal.*, **19**, 19–46.

[18] Ivanov, A.V. and Leonenko, N.N. (1989). *Statistical Analysis of Random Fields.* Kluwer, Dordrecht, Germany.

[19] Kashyap, R.L. and Lapsa, P.M. (1984). Synthesis and estimation of random fields using long correlations. *IEEE Trans. Pattern Anal. Mach. Intell.*, **6**, 800–809.

[20] Koul, H.L. and Mimoto, N. (2012). A goodness-of-fit test for GARCH innovation density. *Metrika*, **75**, 127–149.

[21] Koul, H.L., Mimoto, N. and Surgailis, D. (2013). Goodness-of-fit tests for long memory moving average marginal density. *Metrika*, **76**, 205–224.

[22] Koul, H.L and Surgailis, D. (2002). Asymptotic expansion of the empirical process of long memory moving averages. In: Dehling, H., Mikosch, T., Søensen, M. (Eds.), *Empirical Process Techniques for Dependent Data*, 213–239. Birkhäser, Boston, USA.

[23] Koul, H.L. and Surgailis, D. (2010). Goodness-of-fit testing under long memory. *J. Statist. Plann. Inference*, **140**, 3742–3753.

[24] Koul, H.L. and Surgailis, D. (2013). Goodness-of-fit tests for marginal distribution of linear random fields with long memory. RM #702, Department of Statistics & Probability, Michigan State University.

[25] Lavancier, F. (2006). Long memory random fields. In: *Dependence in Probability and Statistics.* In: **Lecture Notes in Statist.**, **187**, 195–220. Springer, New York, USA.

[26] Lee, S. and Na, S. (2002). On the Bickel-Rosenblatt test for first-order autoregressive models. *Statist. Probab. Lett.*, **56**, 23–35.

[27] Leonenko, N.N. (1999). *Random Fields with Singular Spectrum.* Kluwer, Dordrecht, Germany.

[28] Na, S. (2009). Goodness-of-fit test using residuals in infinite-order autoregressive models. *J. Korean Statist. Soc.*, **38**, 287–295.

[29] Percival, D.B., Rothrock, D.A., Thorndike, A.S. and Gneiting, T. (2008). The variance of mean sea-ice thickness: Effect of long range dependence. *J. Geophys. Research: Oceans*, **113**.

[30] Soulier, P. (2001). Moment bounds and central limit theorem for functions of Gaussian vectors. *Statist. Probab. Lett.*, **54**, 193–203.

[31] Surgailis, D. (1982). Domains of attraction of self-similar multiple integrals. *Lithuanian Math. J.*, **22**, 185–201.

[32] Wang, L. (2009). Gaussian semiparametric estimation for long range dependent random fields. *Statist. Probab. Lett.* **79**, 2297–2306.