TESTING FOR BREAKS IN REGRESSION MODELS WITH DEPENDENT DATA

J. HIDALGO AND V. DALLA

Abstract. The paper examines a test for smoothness/breaks in a nonparametric regression model with dependent data. The test is based on the supremum of the difference between the one-sided kernel regression estimates. When the errors of the model exhibit strong dependence, we have that the normalization constants to obtain the asymptotic Gumbel distribution are data dependent and the critical values are difficult to obtain, if possible. This motivates, together with the fact that the rate of convergence to the Gumbel distribution is only logarithmic, the use of a bootstrap analogue of the test. We describe a valid bootstrap algorithm and show its asymptotic validity. It is interesting to remark that neither subsampling nor the sieve bootstrap will lead to asymptotic valid inferences in our scenario. Finally, we indicate how to perform a test for k breaks against the alternative of $k+k_0$ breaks for some k_0.

JEL classifications. C14, C22.

Key words and phrases: Nonparametric regression. Breaks/smoothness. Strong dependence. Extreme-values distribution. Frequency domain bootstrap algorithms.

1. INTRODUCTION

The literature on breaks/continuity on parametric regression models is both extensive and exhaustive in both econometric and statistical literature, see Perron (2006) for a survey. Because as in many other situations an incorrect specification of the model can lead to misleading conclusions, see for instance Hidalgo (1995), it is of interest to develop tests which do not rely on any functional specification of the regression model. Although some work has been done in the nonparametric setup, the literature appears to focus mostly on the estimation of the break point, see for instance Müller (1992), Chu and Wu (1992) and Delgado and Hidalgo (2000), rather than on the testing of its existence. With this view, the purpose of this paper is to fill this gap by looking at testing for the hypothesis of continuity against the alternative of the existence of (at least) one discontinuity point in a nonparametric regression model, although we shall indicate how to perform a test for k breaks against the alternative of $k+k_0$ breaks for some k_0.

More specifically, we consider the regression model

$$y_t = r\left(x_t\right) + u_t; \quad t = 1, ..., n, \tag{1.1}$$

where we assume that the homoscedastic errors $\{u_t\}_{t\in\mathbb{Z}}$ follow a covariance stationary linear process, to be more precise in Condition $C1$ below. We shall assume that x_t is deterministic, say a time trend. A classical example of interest in time series is a polynomial trend, that is $x_t = \left(t, t^2, ..., t^p\right)$, and/or when regressors are of the type "$\cos t\lambda_0$" and/or "$\sin t\lambda_0$", where $\lambda_0 \neq 0$. The latter type of regressors can be convenient when the practitioner suspects that the data may exhibit some cyclical behaviour. Hence, one possible hypothesis of interest is to know if such a deterministic trend and/or cyclical behaviour has breaks. Of course, we can allow for stochastic covariates x,

Date: **20 March 2015**.

We like to thank Marie Huskova for their comments on a previous version of the paper. Of course, any remaining errors are our sole responsibility.

however this is beyond the scope of this paper as the technical aspects are quite different than those with deterministic regressors.

Our main goal is to test the null hypothesis $r(x) =: E(y \mid x)$ is continuous being the alternative hypothesis that there exists a point in $\mathcal{X}$ such that $r(x)$ is not continuous, and where herewith $\mathcal{X}$ denotes the domain of the variable x. We are also very much interested into the possible consequence of assuming that the errors u_t exhibit strong dependence, as opposed to weak dependence, and in particular the consequence on the asymptotic distribution of the test.

In this paper the methodology that we shall follow is based on a direct comparison between two "alternative" estimates of $r(x)$. More specifically, based on a sample $\{y_t, x_t\}_{t=1}^{n}$, the test is based on global measures of discrepancy between nonparametric estimates of $E(y \mid x)$ when we take only observations at the right and left of the point $x \in \mathcal{X}$. For that purpose, we have chosen the supremum norm, e.g. a Kolmogorov-Smirnov type of test. Alternatively we could have employed the $L_2 - norm$, see among others Bickel and Rosenblatt (1973).

One of our main findings of the paper is that the constant ζ_n used to normalize the statistic (see Theorem 1 below) depends on the so-called strong dependent parameter of the error term. However, due to the slow rate of convergence to the Gumbel distribution and that the implementation of the test can be quite difficult for a given data set, we propose and describe a bootstrap algorithm. So in our setup bootstrap algorithms are not only necessary because they provide more reliable inferences, but due to our previous comment regarding its implementation. The need to use resampling/subsampling algorithm leads to a rather surprising result. In our context, subsampling is not a valid method to estimate the critical values of the test. The reason being, as Theorem 1 below illustrates, see also the comments after Theorem 2 in Section 4, the implementation of the test requires the estimation of some normalization constants which subsampling is not able to compute consistently. Because the well known possible problems of the moving block bootstrap with strong dependence data, and that the sieve bootstrap is neither consistent when we allow for strong dependence, we will propose an algorithm in the frequency domain which overcomes the problem.

The paper is organized as follows. In the next section, we describe the model and test. Also, we present the regularity conditions and the one-sided kernel estimators of the regression function. Section 3 presents the main results of the paper. Due to the nonstandard results obtained in Section 3, Section 4 describes and examines a bootstrap algorithm, showing the validity in our context. The bootstrap is performed in the frequency domain and it extends results to the case when the errors are not necessarily weakly dependent. Section 5 gives the proofs of the results which rely on a series of lemmas in Section 6.

2. THE MODEL AND TEST. REGULARITY CONDITIONS

As we mentioned in the introduction, our main concern is to test the null hypothesis that $r(x)$ is continuous being the alternative hypothesis that there exist a point in $\mathcal{X}$ such that the function $r(x)$ is not continuous. So, noting that continuity of $r(x)$ means that $\forall x \in \mathcal{X}$, $r_+(x) = r_-(x)$, where $r_\pm(x) = \lim_{z \to x\pm} r(z)$, we can set our null hypothesis H_0 as

$$H_0 : r_+(x) = r_-(x), \quad \forall x \in \mathcal{X}, \tag{2.1}$$

being the alternative hypothesis the negation of the null.

The null hypothesis in (2.1) and the nonparametric nature of $r(x)$ suggests that we could base the test for the null hypothesis H_0 in (2.1) on the difference between the kernel regression estimates of $r_+(x)$ and $r_-(x)$. To that end, we shall employ one-sided kernels as proposed by Rice (1984) since in our context they appear necessary since the implementation of the test requires the estimation of $r_+(\cdot)$ and $r_-(\cdot)$, that is estimates of $r(z)$ at $z+$ and $z-$, respectively. Denoting by $K_+(x)$ and $K_-(x)$ one-sided kernels, that is kernel functions taking values for $x>0$ and $x<0$, respectively, we estimate $r_+(x)$ and $r_-(x)$ at points $x_q = q/n$, $q \in Q_n$, where $Q_n = \{q : \check{n} < q \le n - \check{n}\}$, by

$$\widehat{r}_{a,+}(q) := \widehat{r}_{a,+}(x_q) = \frac{1}{\check{n}} \sum_{t=q}^{n} y_t K_{+,t-q}, \quad \widehat{r}_{a,-}(q) := \widehat{r}_{a,-}(x_q) = \frac{1}{\check{n}} \sum_{t=1}^{q} y_t K_{-,t-q}, \tag{2.2}$$

where henceforth we abbreviate $K_\pm\left(\frac{t}{\check{n}}\right)$ by $K_{\pm,t}$, $\check{n} = [na]$ and $a = a(n)$ is a bandwidth parameter such that $a \to 0$ as n increases to infinity, and where for notational simplicity we shall take $x_t = t/n$ henceforth. Thus, the test for H_0 in (2.1) becomes

$$\mathcal{T}_d = \sup_{q \in Q_n} \left| \widehat{r}_{a,+}(q) - \widehat{r}_{a,-}(q) \right|. \tag{2.3}$$

Remark 1. *It is worth mentioning that to take the supremum on* $[0,1]$ *or at point* j/n*, for integer* j*, is the same as* $\widehat{r}_{a,+}(x_q) = \widehat{r}_{a,+}(x)$ *for all* $x \in (x_{q-1}, x_q]$.

Next, let's introduce the following regularity conditions:

C1: $\{u_t\}_{t\in\mathbb{Z}}$ is a covariance stationary linear process defined as

$$u_t = \sum_{j=0}^{\infty} \vartheta_j \varepsilon_{t-j}; \ \sum_{j=0}^{\infty} \vartheta_j^2 < \infty, \text{ with } \vartheta_0 = 1,$$

where $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ is an iid sequence with $E(\varepsilon_t) = 0$, $E\left(\varepsilon_t^2\right) = \sigma_\varepsilon^2$, $E\left(|\varepsilon_t|^\ell\right) = \mu_\ell < \infty$ for some $\ell > 4$. Also, the spectral density function of $\{u_t\}_{t\in\mathbb{Z}}$, denoted $f(\lambda)$, can be factorized as

$$f(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} g(\lambda) h(\lambda), \tag{2.4}$$

where $g(\lambda) = \left|1 - e^{ij\lambda}\right|^{-2d}$, $h(\lambda) = |B(\lambda)|^2$, $B(\lambda) = \sum_{j=0}^{\infty} b_j e^{-ij\lambda}$; and $\sum_{k=0}^{\infty} k^2 |b_k| < \infty$.

The case $d = 0$ refers to weak dependence, whereas the case $0 < d < \frac{1}{2}$ refers to strong dependence. One model satisfying (2.4) is the $FARIMA(p,d,q)$ process $(1-L)^d \Phi_p(L) u_t = \Theta_q(L) \varepsilon_t$, where $(1-L)^{-d} = \sum_{k=0}^{\infty} \overline{\vartheta}_k L^k$ with $\overline{\vartheta}_k = \Gamma(k+d) / (\Gamma(d)\Gamma(k+1))$, where $\Gamma(\cdot)$ denotes the gamma function such that $\Gamma(c) = \infty$ for $c = 0$ and $\Gamma(0)/\Gamma(0) = 1$, and $\Phi_p(L)$ and $\Theta_q(L)$ are the autoregressive and moving average polynomials with no common roots and outside the unit circle. The latter implies that $\Phi_p^{-1}(L)\Theta_q(L) = \sum_{j=0}^{\infty} b_j L^j$ with $b_j = O(j^{-c})$ for any $c > 0$. The condition $\sum_{k=0}^{\infty} k^2 |b_k| < \infty$ implies that $h(\lambda)$ is twice continuously differentiable for all $\lambda \in [0,\pi]$. We finish pointing out that the sole motivation to assume homoscedastic errors is only for notational simplicity as well as to shorten the arguments of the already technical proofs and eases some of the arguments for the proof of the validity of the bootstrap described in Section 4 below.

C2: For all $x \in [0,1]$, $r(x)$ satisfies

$$\lim_{y \to x} \left| \frac{r(y) - r(x) - R(x)}{|x-y|^\tau} \right| = o(1),$$

where $0 < \tau \le 2$ and $R(x)$ is a polynomial of degree $[\tau - 1]$ with $[z]$ denoting the integer part of z.

Condition $C2$ is only slightly stronger than functions $r(x)$ which are Lipschitz continuous of order τ if $0 < \tau \le 1$, or $r(x)$ is differentiable with derivative satisfying a Lipschitz condition of degree $\tau - 1$, if $1 < \tau \le 2$. For instance, when $\tau = 2$, $C2$ means that $r(x)$ is twice continuously differentiable.

C3: $K_+ : [0,1] \to \mathbb{R}$ and $K_- : [-1,0] \to \mathbb{R}$, where $K_+(x) = K_-(-x)$, $\int_0^1 K_+(x)\,dx = 1$ and $\int_0^1 xK_+(x)\,dx = 0$.

Kernels $K_+(x)$, and therefore $K_-(x)$, satisfying $C3$ can be obtained from any function $\upsilon(x)$ with domain in $[0,1]$ as $K_+(x) = \upsilon(x)(c_1 + c_2 x)$, where c_1 and c_2 are the solutions to $\int_0^1 K_+(x)\,dx = 1$ and $\int_0^1 xK_+(x)\,dx = 0$. As an example let $\upsilon(x) = x(x+1)$, then $K_+(x) = 12x(1-x)(3-5x)$, see Delgado and Hidalgo (2000).

Our next condition deals with the bandwidth parameter a.

C4: As $n \to \infty$, (i) $(na)^{-1} \to 0$ and (ii) $(na)^{\frac{1}{2}-d} a^{\tau} \le D < \infty$, with τ as in $C2$.

Part (i) is standard in kernel regression estimation, whereas part (ii) needs more explanation. The latter differs from the analogue assumed by Robinson (1997). Contrary to the latter work, we do not need to assume that $\check{n}^{\frac{1}{2}-d} a^{\tau} \to 0$ as $n \to \infty$. This allows us to choose the optimal bandwidth parameter a, in the sense of being the value a which minimizes the MSE of the nonparametric regression estimator. More precisely, suppose that $d = 0$ and $\tau = 2$. Then, it is known that the optimal choice of a satisfies $a = Dn^{-1/5}$ for some finite positive constant D, which corresponds to the choice of the bandwidth parameter by, say, cross-validation. Also, note that for a given degree of smoothness on $r(x)$, that is τ in $C2$, the bandwidth parameter converges to zero slower as d increases. That is, given a particular bandwidth it requires less smoothness in $r(x)$.

We finish indicating how we can extend our testing procedure to the case where we know that there exist k breaks and we want to test the existence of k_0 additional ones. That is, our null hypothesis is that

$$r(x) = \begin{cases} r_1(x) & x < x^1 \\ r_2(x) & x^1 \le x < x^2 \\ & \cdots \\ r_{k+1}(x) & x^k \le x, \end{cases}$$

where the functions $r_i(x)$ are continuous being the alternative hypothesis that there exist k_0 points in $\mathcal{X}$ for which $r_i(x)$ are not continuous, for some $i = 1, ..., k+1$. We now describe or envisage how we can modify our test in (2.3). To that end, let

$$\widehat{r}_{a,q} =: \left|\widehat{r}_{a,+}(x_q) - \widehat{r}_{a,-}(x_q)\right|, \quad q \in \widetilde{Q}_n,$$

where

$$\widetilde{Q}_n = \left\{ q : q \in Q_n \backslash \cup_{p=1}^{k} \widetilde{Q}_n^{(p)} \right\}$$

with $\widetilde{Q}_n^{(p)} = \{q : x^p - \check{n} < q \le x^p + \check{n}\}$, $p = 1, ..., k$. That is $\widetilde{Q}_n$ is the set of points $q \in Q_n$ which do not belong to the set $\cup_{p=1}^{k} \widetilde{Q}_n^{(p)}$. Next, denote $\widehat{r}_{(q)}$ the qth order statistic of $\{\widehat{r}_{a,q}\}_{q=1}^{\mathring{Q}_n}$, so that $\widehat{r}_{(1)} = \min_{q \in \widetilde{Q}_n} \widehat{r}_{a,q}$ and $\widehat{r}_{(\mathring{Q}_n)} = \max_{q \in \widetilde{Q}_n} \widehat{r}_{a,q}$, where $\mathring{Q}_n = \#\left\{\widetilde{Q}_n\right\}$. Then if k_0 is a known a priory positive integer, the test can be based on

$$\mathcal{T}_d^{k_0} =: \widehat{r}_{(\mathring{Q}_n - (k_0 - 1))}.$$

To examine the asymptotic behaviour of the test is beyond the scope of this paper and it will be discussed in a different manuscript.

3. RESULTS

Before we examine the properties of $\mathcal{T}_d$ in (2.3), we shall first examine the covariance of $\widetilde{r}_a(q)$ at two points $q_1 \le q_2 \in Q_n$, where in what follows $\widetilde{r}_a(q) =: \widehat{r}_{a,+}(q) - \widehat{r}_{a,-}(q)$. Also define $b(q_1, q_2) =: (q_2 - q_1)/\check{n}$ and $\vartheta(d) = 2\Gamma(1-2d)\cos\left(\pi\left(\frac{1}{2} - d\right)\right)$.

Proposition 1. *Assuming C1 − C4, under H_0, for any $\check{n} < q_1 \leq q_2 \leq n - \check{n}$, as $n \to \infty$,*

$$\check{n}^{1-2d} Cov\left(\widetilde{r}_a\left(q_1\right), \widetilde{r}_a\left(q_2\right)\right) \to \rho\left(b; d\right) =: \rho_+\left(b; d\right) + \rho_-\left(b; d\right) - \rho_\pm\left(b; d\right) - \rho_\mp\left(b; d\right),$$

where $b := \lim_{n\to\infty} b\left(q_1, q_2\right)$ is finite and

(*a*) *if* $0 < d < \frac{1}{2}$,

$$\begin{aligned}
\rho_+\left(b; d\right) &= h\left(0\right) \vartheta\left(d\right) \int_0^1 \int_b^{1+b} \left|v - w\right|^{2d-1} K_+\left(v\right) K_+\left(w - b\right) dvdw, \\
\rho_-\left(b; d\right) &= h\left(0\right) \vartheta\left(d\right) \int_{-1}^0 \int_{b-1}^{b} \left|v - w\right|^{2d-1} K_-\left(v\right) K_-\left(w - b\right) dvdw, \\
\rho_\pm\left(b; d\right) &= h\left(0\right) \vartheta\left(d\right) \int_0^1 \int_{b-1}^{b} \left|v - w\right|^{2d-1} K_+\left(v\right) K_-\left(w - b\right) dvdw, \\
\rho_\mp\left(b; d\right) &= h\left(0\right) \vartheta\left(d\right) \int_{-1}^0 \int_b^{1+b} \left|v - w\right|^{2d-1} K_-\left(v\right) K_+\left(w - b\right) dvdw.
\end{aligned}$$

(*b*) *if* $d = 0$,

$$\rho_+\left(b; d\right) = \rho_-\left(b; d\right) = 4^{-1} \sigma_u^2 \int_0^1 K_+\left(v\right) K_+\left(v - b\right) dv, \quad \rho_\pm\left(b; d\right) = \rho_\mp\left(b; d\right) = 0.$$

Proof. The proof is omitted since it proceeds as that in Robinson (1997). □

Proposition 1 indicates that the covariance structure is independent of the points at which $r_\pm\left(x\right)$ is estimated and only depends on the distance among the points where we estimate $r_\pm\left(x\right)$.

The next proposition deals with the correlation structure of $\widetilde{r}_a\left(q\right)$ as $b\left(q_1, q_2\right) \to 0$ and when $b\left(q_1, q_2\right) \to \infty$ as $n \to \infty$. In what follows, D will denote a positive finite constant.

Proposition 2. *Under C1 − C4, for some $\alpha \in (0, 2]$, as $n \to \infty$,*

$$(a) \quad \frac{\rho\left(b\left(q_1, q_2\right); d\right)}{\rho\left(b\left(q_1, q_1\right); d\right)} - 1 = -D\left|b\left(q_1, q_2\right)\right|^\alpha + o\left(\left|b\left(q_1, q_2\right)\right|^\alpha\right) \quad as \quad b\left(q_1, q_2\right) \to 0,$$

$$(b) \quad \rho\left(b\left(q_1, q_2\right); d\right) \log\left(b\left(q_1, q_2\right)\right) = o\left(1\right) \quad as \quad b\left(q_1, q_2\right) \to \infty.$$

Proof. The proof of this proposition or any other result is confined to Section 5 below. □

Proposition 3. *Assuming C1 − C4, for any finite collection q_j, $j = 1, ..., p$, such that $q_j \in Q_n$ and for any z such that $\left|q_{j_1} - q_{j_2}\right| \geq nz > 0$, as $n \to \infty$,*

$$\check{n}^{\frac{1}{2}-d} \rho^{-\frac{1}{2}}\left(0; d\right) \left(\widetilde{r}_a\left(q_j\right)\right)_{j=1}^p \xrightarrow{d} \mathcal{N}\left(0, diag\left(1, ..., 1\right)\right).$$

First of all, we observe that the lack of asymptotic bias when the bandwidth parameter a is chosen optimally. This is in clear contrast to standard kernel regression estimation results, for which a bias term appears in the asymptotic distribution, when a is chosen to minimize the MSE, e.g. when a is chosen as in $C4$. Moreover, the latter result together with Proposition 1 implies that $\widetilde{r}_a\left(q\right)$ has asymptotically stationary increments, which are key to obtain the asymptotic distribution of $\mathcal{T}_d$.

Before we present our main result, we shall give a proposition which may be of independent interest.

Proposition 4. *Let* $\overline{u}_t = \sum_{j=0}^{\infty} \vartheta_j \overline{\varepsilon}_{t-j}$ *and* $\{\overline{\varepsilon}_t\}_{t\in\mathbb{Z}}$ *is a zero mean iid sequence of standard normal random variables. Then under* $C1$ *and* $C3$*, we have that*

$$\sup_{1\le s\le n}\left|\sum_{t=1}^{s} K_{\pm,t}(0)\, u_t - \sum_{t=1}^{s} K_{\pm,t}(0)\, \overline{u}_t\right| = o_p\left(n^{d+1/4}\right). \tag{3.1}$$

We now give the main result of this section. Let $\upsilon_n = (-2\log a)^{1/2}$.

Theorem 1. *Assuming* $C1 - C4$*, under* H_0*,*

$$Prob\left\{\upsilon_n\left(\breve{n}^{\frac{1}{2}-d}\rho^{-\frac{1}{2}}(0;d)\,\mathcal{T}_d - \zeta_n\right)\le x\right\}\underset{n\uparrow\infty}{\to}\exp\left(-2e^{-x}\right),\quad for\ x>0,$$

where (*a*) *If* $0<d<1/2$*, then*

$$\zeta_n = \upsilon_n + \upsilon_n^{-1}\left\{\left(\frac{1}{2}-\frac{1}{\alpha}\right)\log\log a^{-1} + \log\left((2\pi)^{-\frac{1}{2}}\, 2^{\frac{2-\alpha}{2\alpha}} E^{\frac{1}{\alpha}}\mathcal{J}_\alpha\right)\right\}$$

for some $0<E<\infty$*, where* α *is as given in Proposition 2,*

$$0<\mathcal{J}_\alpha \equiv \lim_{a\to 0}\int_0^\infty e^s \Pr\left\{\sup_{0\le t\le [a]^{-1}} \mathcal{Y}(t) > s\right\} ds < \infty$$

and $\mathcal{Y}(t)$ *is a stationary mean zero Gaussian process with covariance structure*

$$Cov\left(\mathcal{Y}(t_1), \mathcal{Y}(t_2)\right) = |t_1|^\alpha + |t_2|^\alpha - |t_2 - t_1|^\alpha.$$

(*b*) *If* $d=0$*, then* $\zeta_n = \upsilon_n + \upsilon_n^{-1}\log\left((2\pi)^{-1}\left(\int_0^1 \left(\partial K_+(x)/\partial x\right)^2 dx\right)^{1/2}\right)$.

3.1. **POWER OF THE TEST.**

A desirable and important characteristic of any test is its consistency, that is under the alternative hypothesis the probability of rejection converges to 1 as $n\to\infty$. In addition to examine the limiting behaviour under local alternatives enables to make comparisons between different consistent tests. We begin with the latter. To that end, we consider the following sequence of local alternatives

$$H_a:\quad \exists x^0\in[0,1] \text{ such that } r_+\left(x^0\right) = r_-\left(x^0\right) + r_n\left(x^0\right), \tag{3.2}$$

where $r_n\left(x^0\right) = \breve{n}^{d-1/2}\,|2\log a|^{-1/2}\, r$ with $r\neq 0$ and $r(x)$ satisfies $C2$ for $x\neq x^0$. Then, we have the following:

Corollary 1. *Assuming* $C1 - C4$*, under* H_a *in* (*3.2*)

$$Prob\left\{\upsilon_n\left(\breve{n}^{\frac{1}{2}-d}\rho^{-\frac{1}{2}}(0;d)\,\mathcal{T}_d - \zeta_n\right)\le x\right\}\underset{n\uparrow\infty}{\to}\exp\left(-2e^{-\left(x-\frac{|r|\varrho(K_+)}{\rho^{1/2}(0;d)}\right)}\right),\ x>0,$$

where ζ_n *was given in Theorem 1 and* $\varrho(K_+) = \max_{\ell=1,\dots,\breve{n}}\int_{\ell/\breve{n}}^1 K_+(v)\,dv$.

Note that $\varrho(K_+)$ is not necessarily equal to 1 as would be the case if $K_+(\cdot)$ were nonnegative. This is because the condition $\int_0^1 xK_+(x)\,dx = 0$ implies that $K_+(\cdot)$ takes negative values in some subset of $[0,1]$.

From Corollary 1, one would expect that for fixed alternatives

$$H_1:\quad \exists x^0\in[0,1] \text{ such that } r_+\left(x^0\right) = r_-\left(x^0\right) + r;\quad |r|>0,$$

and $r(x)$ satisfies $C2$ for $x \neq x^0$, we should have

$$\lim_{n\to\infty} \text{Prob}\left\{\upsilon_n\left(\check{n}^{\frac{1}{2}-d}\rho^{-\frac{1}{2}}(0;d)\,\mathcal{T}_d - \zeta_n\right) \leq x\right\} = 0$$

that is, the test is consistent. This is confirmed in the next corollary.

Corollary 2. *Assuming $C1 - C4$, $\mathcal{T}_d$ is consistent.*

Although Theorem 1 gives asymptotic justification for our test $\mathcal{T}_d$ under H_0, we observe that the normalization constant ζ_n depends not only on d but more importantly on $\mathcal{J}_\alpha$. The latter quantity is very difficult to compute except for the special cases $\alpha = 1$ or 2, see Pickands (1969), where $\mathcal{J}_2 = \upsilon_n + \upsilon_n^{-1}\log\left(\pi^{-1}(E/2)^{1/2}\right)$ and $\mathcal{J}_1 = \upsilon_n + \upsilon_n^{-1}\log\left\{(E/\pi)^{1/2} + 2^{-1}\log\log a^{-1}\right\}$, where E is a constant which depends on K_+ although easy to obtain. More specifically, in our context, although d can be estimated, we face one potential difficulty when implementing the test. As we observe from (the proof of) Proposition 2, α depends on K_+ and d, so that to obtain $\mathcal{J}_\alpha$ does not seem an easy task. Under these circumstances, a bootstrap algorithm appears to be a sensible way to proceed.

4. THE BOOTSTRAP APPROACH

The comments made at the end of Section 3 and in the introduction suggest that to perform the test we need the help of bootstrap algorithms. In a context of time series, several approaches have been described in the literature. However, as we indicated in the introduction and after Corollary 2, the subsampling is not an appropriate method, neither the sieve bootstrap of Bühlmann (1997) as the latter is not consistent for the sample mean of the error term with strong dependent data. Recall that in our context the statistical properties of the sample mean plays an important role into the asymptotic distribution of the test.

Due to this, in this section we describe and examine a bootstrap algorithm in the frequency domain similar to that proposed by Hurvich and Zeger (1987), although they did not provide its justification and our conditions are significantly weaker than theirs. Two differences of our bootstrap procedure with moving block bootstrap (MBB) described in Künsch (1989), say, are that (a) it is not a subset of the original data, and (b) the bootstrap data, say $\{u_t^*\}_{t=1}^n$, is covariance stationary as we have that $Cov^*(u_t^*, u_s^*)$ is a function of $|t-s|$. Herewith, by $Cov^*(z_1, z_2)$ or, say $E^*(z)$, we mean the covariance or expectation conditional on the data.

We now describe our main ingredients of the bootstrap and its justification. Suppose that in $C1$, $d = 0$, that is $u_t = \sum_{k=0}^{\infty} b_k \varepsilon_{t-k}$. Then, using the identity

$$u_t = \frac{1}{n^{1/2}}\sum_{j=1}^{n} e^{it\lambda_j} w_u(\lambda_j)\,, \tag{4.1}$$

which can be considered as a "discrete" Cràmer representation of $\{u_t\}_{t=1}^n$, and Bartlett's approximation of $w_u(\lambda_j)$, see Brockwell and Davis's (1991) Theorem 10.3.2, we obtain that

$$u_t \approx \frac{1}{n^{1/2}}\sum_{j=1}^{n} e^{it\lambda_j} B(-\lambda_j)\, w_\varepsilon(\lambda_j)\,,$$

where "$\approx$" should be read as "approximately". Because $C1$ allows for strong dependence, the previous arguments suggests the approximation

$$w_u(\lambda_j) \approx \left(1 - e^{-i\lambda_j}\right)^{-d} B(-\lambda_j)\, w_\varepsilon(\lambda_j)\,. \tag{4.2}$$

However the lack of smoothness of $\left(1-e^{-i\lambda_j}\right)^{-d}$ around $\lambda_j=0$ and results given in Robinson's (1995a) Theorem 1 at frequencies λ_j for fixed j indicate that for those frequencies the approximation in (4.2) seems to be invalid. Observe that these frequencies are precisely the more relevant ones when examining the asymptotic behaviour of $\widehat{r}_{a,\pm}(q)$ in (2.2). So we consider

$$u_t \approx \widetilde{u}_t =: \frac{1}{n^{1/2}} \sum_{j=1}^{n} e^{it\lambda_j} \widetilde{g}^{1/2}(-\lambda_j; d) B(-\lambda_j) w_\varepsilon(\lambda_j), \tag{4.3}$$

where

$$\widetilde{g}^{1/2}(-\lambda_j; d) = \left| \sum_{\ell=-n+1}^{n-1} \gamma_\ell(d) e^{-i\ell\lambda_j} \right|^{1/2}, \tag{4.4}$$

with $\gamma_\ell(d) = \frac{(-1)^\ell \Gamma(1-2d)}{\Gamma(\ell-d+1)\Gamma(1-\ell-d)}$. It is easy to show that the right side of (4.3) preserves (asymptotically) the covariance structure of $\{u_t\}_{t\in\mathbb{Z}}$.

We now describe the bootstrap in the following *6 STEPS*.

***STEP 1*:** Let $\widehat{t} = \arg\max_{t\in Q_n} |\widehat{r}_{a,+}(t) - \widehat{r}_{a,-}(t)|$, and obtain the centered residuals $\widehat{u}_t = \widetilde{u}_t - n^{-1}\sum_{t=1}^{n} \widetilde{u}_t$, $t=1,...,n$, where $\widetilde{u}_t = y_t - \widehat{r}_a(t)$ with

$$\widehat{r}_a(t) = \begin{cases} \widehat{r}_{a,+}(t), & t \le \check{n} \\ \widehat{r}_{a,-}(t), & \check{n} < t < \widehat{t} \\ \widehat{r}_{a,+}(t), & \widehat{t} \le t \le n-\check{n} \\ \widehat{r}_{a,-}(t), & t > n-\check{n}, \end{cases} \tag{4.5}$$

and $\widehat{r}_{a,+}(t)$ and $\widehat{r}_{a,-}(t)$ given in (2.2).

It is worth indicating that we could have computed the residuals using an estimate of the regression model under the null hypothesis of continuity, i.e.

$$\widehat{r}_a(t) = \begin{cases} \widehat{r}_{a,+}(t), & t \le \check{n} \\ \frac{1}{2}\left(\widehat{r}_{a,+}(t) + \widehat{r}_{a,-}(t)\right), & \check{n} < t \le n-\check{n} \\ \widehat{r}_{a,-}(t), & t > n-\check{n}. \end{cases} \tag{4.6}$$

However, as it is well known, it is always preferable to obtain the residuals under the alternative hypothesis, as in (4.5), than under the null hypothesis.

***STEP 2*:** We estimate d by Robinson's (1995b) *GSE*,

$$\widehat{d} = \arg\min_{d\in[0,\Delta]} \widetilde{R}(d), \tag{4.7}$$

where $0<\Delta<1/2$, and

$$\widetilde{R}(d) = \log\left(\frac{1}{m}\sum_{j=1}^{m} \lambda_j^{2d} I_{\widehat{u}\widehat{u}}(\lambda_j)\right) - 2d\sum_{j=1}^{m} \log\lambda_j$$

for integer $m\in[1,\tilde{n})$, with $\tilde{n}=[n/2]$, and where $I_{\widehat{u}\widehat{u}}(\lambda) = |w_{\widehat{u}}(\lambda)|^2/(2\pi)$ is the periodogram of $\{\widehat{u}_t\}_{t=1}^{n}$, with $m^{-1}+mn^{-1}\to 0$.

We define our estimator of $2\pi h(\lambda) = |B(\lambda)|^2$ by

$$\widehat{h}(\lambda) = \frac{1}{2m+1}\sum_{j=-m}^{m} \left|1-e^{-i(\lambda+\lambda_j)}\right|^{2\widehat{d}} I_{\widehat{u}\widehat{u}}(\lambda+\lambda_j).$$

Our third step describes how to obtain $w_\varepsilon^*(\lambda_j)$, $j=1,...,\tilde{n}$.

***STEP 3*:** Let $\{\varepsilon_t^*\}_{t=1}^n$ be a random sample from standard normal and obtain its *discrete Fourier transform*,

$$\eta_j^* := w_\varepsilon^* (\lambda_j) = \frac{1}{n^{1/2}} \sum_{t=1}^n \varepsilon_t^* e^{-it\lambda_j}, \quad j = 1, ..., \tilde{n},$$

with $\eta_{n-j}^* = \overline{\eta}_j^*$, $j = 1, ..., \tilde{n}$, and $\overline{z}$ denoting the conjugate of z.

***STEP 4*:** Compute

$$u_t^* = \frac{1}{n^{1/2}} \sum_{j=1}^n e^{it\lambda_j} \widetilde{g}^{1/2} \left(\lambda_j; \widehat{d}\right) \widehat{A} (\lambda_j) \eta_j^*, \quad t = 1, ..., n,$$

where $\widetilde{g}^{1/2} \left(\lambda_j; \widehat{d}\right)$ is given in (4.4),

$$\widehat{A} (\lambda_j) = \exp \left\{ \frac{1}{2} \widehat{c}_0 + \sum_{r=1}^M \widehat{c}_r e^{-ir\lambda_j} \right\}, \quad j = 1, ..., \tilde{n} \tag{4.8}$$

with $\widehat{A} (\lambda_{n-j}) = \overline{\widehat{A} (\lambda_j)}$, and for $r = 0, ..., M = [n/4m]$, $\widehat{c}_r = \tilde{n}^{-1} \sum_{\ell=1}^{\tilde{n}} \log \left(\widehat{h} (\lambda_\ell)\right) \cos (r\lambda_\ell)$.

Remark 2. (**i**) *It is worth mentioning that the way to obtain the bootstrap observations u_t^* in* STEP 4*, together with the definition of $\widehat{h} (\lambda)$ in* (2.4)*, has some similarities with the autoregressive-aided bootstrap in Kreiss and Paparoditis* (2003)*.*

(**ii**) *There are doubts that* STEP 3 *can be modified to allow $\{\varepsilon_t^*\}_{t=1}^n$ to be a random sample from the empirical distribution of $\{\widehat{u}_t\}_{t=1}^n$, following arguments in Huskova et al.* (2008)*. However the latter will lengthen the arguments and the proof considerably and it is then beyond the scope of this paper.*

The modulus square of $\widehat{A} (\lambda)$ in (4.8) is an estimator of $h (\lambda)$ in (2.4) and it comes from the so-called canonical spectral decomposition of $h (\lambda)$, see for instance Brillinger (1981, *p.* 78 − 79).

***STEP 5*:** Compute $\widehat{r}_a (t)$ as in (4.6) and then

$$y_t^* = \widehat{r}_a (t) + u_t^*; \quad t = 1, ..., n.$$

STEP 5 employs the same bandwidth as that in *STEP 1*, so that the standard requirement of an additional bandwidth e, such that $a = o (e)$, in computing the bootstrap analogue of (1.1), see for instance Härdle and Marron (1991), is not needed. The reason comes from the observation that the bias of the nonparametric estimator of $r_+ (\cdot) - r_- (\cdot)$ is $o \left(a^2\right)$ instead of the usual $O \left(a^2\right)$. Our final step is:

***STEP 6*:** Compute $\widehat{r}_{a,+}^* (q)$ and $\widehat{r}_{a,-}^* (q)$, $q \in Q_n$, as in (2.2) but with y_t replaced by y_t^* and the same bandwidth parameter a employed in *STEP 1*. Then we compute the bootstrap version of $\mathcal{T}_d$ as

$$\mathcal{T}_d^* = \sup_{q \in Q_n} \left| \widehat{r}_{a,+}^* (q) - \widehat{r}_{a,-}^* (q) \right|.$$

The next proposition examines the behaviour of $\widehat{d}$ given in (4.7).

Proposition 5. *Under $C1$ and $C3$, $\left|\widehat{d} - d\right| = O_p \left((ma)^{-1} + a^4 n (m/n)^{2d}\right)$.*

Proof. The proof is omitted as it follows step by step that of Robinson's (1997) Theorem 3, after noting that in our case we do not have his terms $I_{\xi\xi}$ and $I_{\chi\chi}$. □

Let us introduce the following condition on the smoothing parameter m and the bandwidth parameter a.

C5: As $n \to \infty$, (i) $D^{-1}n^{-1/3} < a < Dn^{-1/4}$ and (ii) $D^{-1}n^{3/5} < m < Dn^{3/4}$.

The next proposition discusses the bias of the nonparametric estimator $\widetilde{r}^{*}_{a}(q) = \widehat{r}^{*}_{a,+}(q) - \widehat{r}^{*}_{a,-}(q)$.

Proposition 6. *Assuming C1 − C3, with $\tau = 2$ there, and C5, under $H_0 \cup H_1$, as $n \to \infty$, $E^{*}\widetilde{r}^{*}_{a}(q) = o_p\left(\check{n}^{d-1/2}\right)$.*

Proposition 7. *Denote $E(u_{\ell}u_0) = \delta_{|\ell|}$. Assuming C1 − C3 with $\tau = 2$ there, and C5, we have that for $q \in Q_n$,*

$$\frac{1}{\check{n}^{2d}} \sum_{t=q+1}^{\check{n}+q} \left| E^{*}\left(u^{*}_{t}u^{*}_{t+\ell}\right) - \delta_{|\ell|} \right| = o_p(1). \tag{4.9}$$

Proof. The proof proceeds as that of Hidalgo's (2007) Proposition 4.2 and thus it is omitted. □

Theorem 2. *Under the same conditions of Proposition 7 and $H_0 \cup H_1$, as $n \to \infty$*

$$\Pr\left\{ \upsilon_n \left(\check{n}^{\frac{1}{2}-\widehat{d}} \rho^{-\frac{1}{2}}\left(0;\widehat{d}\right) \mathcal{T}^{*}_{d} - \zeta_n \right) \leq x \middle| \mathcal{Y} \right\} \to \exp\left(-2e^{-x}\right), \quad \text{for } x > 0,$$

where υ_n and ζ_n were defined in Theorem 1.

We now comment on Theorem 2 and also the important reason why subsampling is not a valid procedure as we now argue. To obtain a critical value, say $x(\beta)$, for which $\exp\left(-2e^{-x(\beta)}\right) = 1 - \beta$, is the same as to find the value, say $z_n(\beta)$, which obeys the equality

$$\lim_{n\to\infty} \Pr\left\{ \mathcal{T}_d \leq z_n(\beta) \right\} = 1 - \beta.$$

However the value $z_n(\beta)$ depends on both ζ_n and υ_n and thus, indirectly, on the choice of the bandwidth parameter a. In fact, since the constants ζ_n and υ_n are not possible to be computed, in practice we would only hope to obtain the critical values via $z_n(\beta)$. So, when employing the bootstrap sample, we are bound to obtain $z^{*}_{n}(\beta)$, for which

$$\lim_{n\to\infty} \Pr\left\{ \mathcal{T}^{*}_{d} \leq z^{*}_{n}(\beta) \middle| \mathcal{Y} \right\} = 1 - \beta.$$

As with the original data, the value $z^{*}_{n}(\beta)$ depends on both ζ_n and υ_n and thus on the choice of a. The latter has thus to be kept in mind when computing $z^{*}_{n}(\beta)$. But recall that one requirement for the bootstrap to be (asymptotically) valid is that $z^{*}_{n}(\beta)$ needs to satisfy $\left| z^{*}_{n}(\beta)/z_n(\beta) - 1 \right| \overset{p}{\to} 0$. It is obvious that, for the latter expression to hold true, we need the constants ζ_n and υ_n to be the same for both $\mathcal{T}_d$ and $\mathcal{T}^{*}_{d}$, or that their ratio converges to one in probability. This is obviously possible only if the bandwidth parameter a is the same when estimating the regression function with both the original $\{y_t\}_{t=1}^{n}$ and bootstrap $\{y^{*}_{t}\}_{t=1}^{n}$ data.

5. PROOFS OF THE MAIN RESULTS

Proof of Proposition 2. We shall begin with the case $d>0$. Abbreviating $b(q_1, q_2)$ by b, Proposition 1 implies that

$$\rho(b;d) = \rho_{+}(b;d) + \rho_{-}(b;d) - \rho_{\pm}(b;d) - \rho_{\mp}(b;d) + o(1). \tag{5.1}$$

Noting that for $\ell > 0$ and $d > 0$, $\ell^{2d-1} = \frac{2}{\pi}\Gamma(2d)\cos(d\pi)\int_0^{\infty}\lambda^{-2d}\cos(\ell\lambda)\,d\lambda$, proceeding as in Hidalgo (2007), we have that the first term on the right of (5.1) is

$$\rho_{+}(b;d) = \int_{-\infty}^{\infty} |\lambda|^{-2d} \cos(|\lambda| b) \left| \int_0^1 K_{+}(v) e^{i\lambda v} dv \right|^2 d\lambda \tag{5.2}$$

for finite b. Likewise the last three terms on the right of (5.1) are, respectively,

$$\int_{-\infty}^{\infty} |\lambda|^{-2d} \cos(|\lambda| b) \left| \int_0^1 K_+(v) e^{-iv\lambda} dv \right|^2 d\lambda$$

$$\frac{1}{2} \int_{-\infty}^{\infty} |\lambda|^{-2d} \int_0^1 \int_0^1 \left(\left(e^{-i(v+w)\lambda} e^{-ib\lambda} + e^{i(v+w)\lambda} e^{ib\lambda} \right) K_+(v) K_+(w) \, dv dw \right) d\lambda$$

$$\frac{1}{2} \int_{-\infty}^{\infty} |\lambda|^{-2d} \int_0^1 \int_0^1 \left(\left(e^{i(v+w)\lambda} e^{-ib\lambda} + e^{-i(v+w)\lambda} e^{ib\lambda} \right) K_+(v) K_+(w) \, dv dw \right) d\lambda,$$

after noticing that by $C3$, $K_+(v) = K_-(-v)$. Hence, gathering (5.2) and the last displayed expressions, we conclude that

$$\rho(b; d) = \int_{-\infty}^{\infty} |\lambda|^{-2d} \cos(|\lambda| b) \left| \int_0^1 K_+(v) \left(e^{i\lambda v} - e^{-i\lambda v} \right) dv \right|^2 d\lambda.$$

The proof now proceeds as that of Hidalgo's (2007) Proposition 3.2.

Next, when $\mathbf{d} = \mathbf{0}$, the proof is omitted as it follows by standard arguments. See for instance Bickel and Rosenblatt's (1973) Theorem B.1. ■

Proof of Proposition 3. Proceeding as in Robinson's (1997) Theorem 1 and Lemma 1, it suffices to show that

$$E(\widetilde{r}_a(q)) = \begin{cases} o(a^\tau) & \text{if } 0 < \tau \leq 1 \\ o\left(a^\tau + n^{-1}\right) & \text{if } 1 < \tau \leq 2. \end{cases} \tag{5.3}$$

Observe that by uniform integrability of u_t^2 and that Propositions 1 and 2 imply that the covariance of $\widetilde{r}_a(q) = \widehat{r}_{a,+}(q) - \widehat{r}_{a,-}(q)$ at two points q_1 and q_2 converges to zero when $|q_1 - q_2| \geq nz > 0$, for any $z > 0$, we conclude that the covariance of the asymptotic distribution of the estimators is zero by Theorem A of Serfling (1980, p.14).

On the other hand, under H_0 and standard kernel manipulations, we obtain that

$$\begin{aligned} E(\widehat{r}_{a,+}(q) - r(q/n)) &= \int_0^1 K_+(v)(r(av + q/n) - r(q/n)) \, dv \\ &= \int_0^1 K_+(v) R(av + q/n) \, dv + \begin{cases} o(a^\tau) & \text{if } 0 < \tau \leq 1 \\ o\left(a^\tau + n^{-1}\right) & \text{if } 1 < \tau \leq 2 \end{cases} \end{aligned}$$

by $C2$. Similarly,

$$E(\widehat{r}_{a,-}(q) - r(q/n)) = \int_{-1}^0 K_-(v) R(av + q/n) \, dv + \begin{matrix} o(a^\tau) & \text{if } 0 < \tau \leq 1 \\ o\left(a^\tau + n^{-1}\right) & \text{if } 1 < \tau \leq 2. \end{matrix}$$

From here and after an obvious change of variables, (5.3) holds true because $Q(x)$ is twice continuously differentiable and $\int_0^1 x K_+(x) \, dx = 0$ by $C3$. ■

Proof of Proposition 4. We shall consider the case that $K_{+,t}(0) = 1$, $\overline{u}_{t,n} = \sum_{j=0}^{\infty} \vartheta_j \overline{\varepsilon}_{t-j}$ the general case follows after observing that by Abel summation by parts

$$\left| \sum_{t=1}^{s} K_{+,t}(0)(u_t - \overline{u}_t) \right| \leq \left| \sum_{t=1}^{s} (K_{+,t}(0) - K_{+,t+1}(0)) \sum_{l=1}^{t} (u_l - \overline{u}_l) \right| + \left| K_{+,s}(0) \sum_{l=1}^{s} (u_l - \overline{u}_l) \right|$$

and then that $\sum_{t=1}^{s} |K_{+,t}(0) - K_{+,t+1}(0)| < D$ by $C3$. First we observe that we can write u_t as follows

$$u_t = \sum_{j=0}^{t-1} \vartheta_j \varepsilon_{t-j} + \sum_{j=0}^{\infty} \vartheta_{j+t} \varepsilon_{-j} := u_{1,t} + u_{2,t}.$$

So, it suffices to show (3.1) when u_t is replaced by $u_{1,t}$ and $u_{2,t}$. That is,

$$\sup_{1\le s\le n} \left| \sum_{t=1}^{s} u_{j,t} - \sum_{t=1}^{s} \overline{u}_{j,t} \right| = o_p\left(n^{d+1/4}\right), \quad j = 1, 2 \tag{5.4}$$

and $\overline{u}_{1,t} = \sum_{j=0}^{t-1} \vartheta_j \overline{\varepsilon}_{t-j}$ and $\overline{u}_{2,t} = \sum_{j=0}^{\infty} \vartheta_{j+t} \overline{\varepsilon}_{-j}$.

We shall prove (5.4) for $j = 2$ first. After standard algebra and inequalities, for some $\chi \in (1, 2)$, we have that the left side is bounded by

$$\sup_{1\le s\le n} \left| \sum_{t=1}^{s} \sum_{j=s^{\chi}+1}^{\infty} \vartheta_{j+t} \left(\varepsilon_{-j} - \overline{\varepsilon}_{-j}\right) \right| + \sup_{1\le s\le n} \left| \sum_{t=1}^{s} \sum_{j=1}^{s^{2-\chi}} \vartheta_{j+t} \left(\varepsilon_{-j} - \overline{\varepsilon}_{-j}\right) \right| + \sup_{1\le s\le n} \left| \sum_{t=1}^{s} \sum_{j=s^{2-\chi}+1}^{s^{\chi}} \vartheta_{j+t} \left(\varepsilon_{-j} - \overline{\varepsilon}_{-j}\right) \right|. \tag{5.5}$$

The expression inside the absolute value of first term of (5.5) is

$$\sum_{j=s^{\chi}+1}^{\infty} \left(\sum_{t=1}^{s} \vartheta_{j+t} \right) \left(\varepsilon_{-j} - \overline{\varepsilon}_{-j}\right) = \sum_{j=s^{\chi}+1}^{\infty} \left\{ \sum_{t=1}^{s} \left(\vartheta_{j+t} - \vartheta_{j+t+1}\right) \right\} \sum_{p=s^{\chi}+1}^{j} \left(\varepsilon_{-p} - \overline{\varepsilon}_{-p}\right).$$

But by well known results due to Komlós, Major and Tusnady, $\sup_{s^{\chi}+1\le p\le j} \left| \sum_{p=s^{\chi}+1}^{j} \left(\varepsilon_{-p} - \overline{\varepsilon}_{-p}\right) \right| = o_p\left(j^{1/4}\right)$, so that the right side of the last displayed equality becomes

$$o_p(1) \sum_{j=s^{1+\chi}+1}^{\infty} j^{1/4} \sum_{t=1}^{s} (j+t)^{d-2} = o_p\left(s^{d+1/4}\right),$$

because $|\vartheta_j - \vartheta_{j+1}| \le D j^{d-2}$ by $C1$, and hence the first term of (5.5) is $o_p\left(n^{d+1/4}\right)$.

Next, proceeding similarly, we have that by Abel summation by parts, the expression inside the absolute value of the second term of (5.5) is

$$\sum_{j=1}^{s^{2-\chi}} \left\{ \sum_{t=1}^{s} \left(\vartheta_{j+t} - \vartheta_{j+t+1}\right) \right\} \sum_{p=1}^{j} \left(\varepsilon_{-p} - \overline{\varepsilon}_{-p}\right) + \sum_{t=1}^{s} \vartheta_{s^{2-\chi}+t} \sum_{p=1}^{s^{2-\chi}} \left(\varepsilon_{-p} - \overline{\varepsilon}_{-p}\right) = o_p\left(s^{d+1/4}\right),$$

because $\chi > 1$ implies that

$$\sum_{j=1}^{s^{2-\chi}} j^{1/4} \sum_{t=1}^{s} (j+t)^{d-2} \le K \sum_{j=1}^{s^{2-\chi}} j^{-1+d+1/q} \sum_{t=1}^{s} (j+t)^{-1} = o\left(s^{d+1/4}\right)$$

and $\sum_{t=1}^{s} \vartheta_{s^{2-\chi}+t} = O\left(s^{d}\right)$. So, it remains to examine the third term of (5.5). Proceeding as before we have, by Abel summation by parts, that it is

$$o_p(1) \sup_{1\le s\le n} \sum_{j=s^{2-\chi}+1}^{s^{\chi}} j^{1/4} \sum_{t=1}^{s} (j+t)^{d-2} = o_p(1) \sum_{t=1}^{n} \left(s^{2-\chi} + t\right)^{d-1+1/4} = o_p\left(n^{d+1/4}\right)$$

by standard manipulations. This completes the proof of (5.4) for $j = 2$.

We now show (5.4) for $j = 1$. First, $\sum_{t=1}^{s} u_{1,t} - \sum_{t=1}^{s} \overline{u}_{1,t}$ is

$$\sum_{t=1}^{s}\left(\sum_{j=0}^{t-1}\vartheta_j\left(\varepsilon_{t-j}-\overline{\varepsilon}_{t-j}\right)\right)=\sum_{\ell=1}^{s}\left(\sum_{j=0}^{s-\ell}\vartheta_j\right)\left(\varepsilon_{\ell}-\overline{\varepsilon}_{\ell}\right)=\sum_{\ell=1}^{s}\left(\sum_{j=0}^{s-\ell}\vartheta_j\right)\left\{\left(S_{\ell}-S_{\ell-1}\right)-\left(\overline{S}_{\ell}-\overline{S}_{\ell-1}\right)\right\},$$

where $S_\ell = \sum_{p=1}^{\ell} \varepsilon_p$ and $\overline{S}_\ell = \sum_{p=1}^{\ell} \overline{\varepsilon}_p$. From here the proof follows as in Lemma 5 of Marinucci and Robinson (2003) since $\Lambda_s = \sum_{j=0}^{s} \vartheta_j$ satisfies their Assumption 1. ■

Proof of Theorem 1. Because the asymptotic independence of the distributions of $\max_q$ and $\min_q$ and the asymptotic distributions of $\sup_i X_i$ and $\inf_i -X_i$ are the same, it suffices to show that, for $x > 0$,

$$\Pr\left\{\upsilon_n\left(\sup_{\check{n}<q\leq n-\check{n}}\check{n}^{\frac{1}{2}-d}\rho^{-\frac{1}{2}}\left(0;d\right)\left(\widetilde{r}_a\left(q\right)\right)-\zeta_n\right)\leq x\right\}\to\exp\left(2e^{-x}\right). \tag{5.6}$$

To that end, we will show that $\check{n}^{\frac{1}{2}-d}\widetilde{r}_a(q)$ converges to a Gaussian process $\mathcal{G}(u)$ in $\mathbb{D}[0,\infty)$, whose correlation structure satisfies conditions (v) and (vi) of Bickel and Rosenblatt's (1973) Theorem A1, for some $\alpha > 0$. See also Pickands's (1969) equations (1.2) and (2.1). From here and Proposition 4, the limiting distribution in (5.6) holds by Bickel and Rosenblatt's (1973) Theorem 1, for some $\alpha > 0$.

First by standard arguments, Propositions 3 and 4 implies that the finite dimensional distributions converge to those of a Gaussian process $\mathcal{G}(u)$, whereas, by Proposition 2, the correlation structure of $\mathcal{G}(u)$ satisfies the conditions in Bickel and Rosenblatt (1973) or Pickands (1969). So, to complete the proof it suffices to show the tightness condition for the process $\check{n}^{\frac{1}{2}-d}\left(\widetilde{r}_a(q)\right)$. To that end, we shall denote

$$X_{\pm,n}\left(\widetilde{q}\right)=\frac{1}{\check{n}^{\frac{1}{2}+d}}\sum_{t=1}^{n}u_tK_{\pm}\left(\frac{t}{\check{n}}-\widetilde{q}\right),\quad \widetilde{q}=\frac{1}{\check{n}},\frac{2}{\check{n}},...,[a]^{-1}.$$

So, we have that $X_{+,n}(\widetilde{q})$, say, is a process in $\mathbb{D}\left[0,[a]^{-1}\right]$ equipped with Skorohod's metric, where we extend $\mathbb{D}\left[0,[a]^{-1}\right]$ to $\mathbb{D}[0,\infty)$ by writing $X_{+,n}(\infty) = X_{+,n}\left([a]^{-1}\right)$. Then Pollard (1981, *Ch.V*) implies that we need only to show tightness in $\mathbb{D}[0,D]$ for any finite $D > 0$. To that end, let

$$\check{n}^{\frac{1}{2}-d}\left\{\left(\widehat{r}_{a,+}(q)-E\widehat{r}_{a,+}(q)\right)-\left(\widehat{r}_{a,-}(q)-E\widehat{r}_{a,-}(q)\right)\right\}:=X_{+,n}\left(\widetilde{q}\right)+X_{-,n}\left(\widetilde{q}\right).$$

Next Proposition 2 implies that the process $X_{+,n}(\widetilde{q})$ has independent and stationary increments, that is for $\widetilde{q} \in [c_1, d_1]$ and $\widetilde{q} \in [c_2, d_2]$ and $[c_1, d_1] \cap [c_2, d_2] = \emptyset$, $X_{+,n}(\widetilde{q})$ are (asymptotically) independent with the same finite dimensional distributions.

Because $\mathcal{G}(\bullet)$ has continuous paths, by Billingsley's (1968) Theorem 15.6, it suffices to show the Kolmogorov's moment condition

$$E\left(\left|X_{+,n}\left(\widetilde{q}_2\right)-X_{+,n}\left(\widetilde{q}\right)\right|^{\beta}\left|X_{+,n}\left(\widetilde{q}\right)-X_{+,n}\left(\widetilde{q}_1\right)\right|^{\beta}\right)\leq D\left|\widetilde{q}_2-\widetilde{q}\right|^{\frac{1+\delta}{2}}\left|\widetilde{q}-\widetilde{q}_1\right|^{\frac{1+\delta}{2}}$$

for some $\delta > 0$, $\beta > 0$ and where $0 \leq \widetilde{q}_1 < \widetilde{q} < \widetilde{q}_2 \leq D$. Observe that we can consider only the situation for which $\check{n}^{-1} < \widetilde{q}_2 - \widetilde{q}_1$, since otherwise the left side is trivially zero. Because for any $0 \leq a < b < c \leq D$, $|c-b|\,|b-a| \leq |c-a|^2$ by Cauchy-Schwarz inequality, the last displayed inequality holds true if

$$E\left|X_{+,n}\left(\widetilde{q}_2\right)-X_{+,n}\left(\widetilde{q}_1\right)\right|^{2\beta}\leq D\left|\widetilde{q}_2-\widetilde{q}_1\right|^{1+\delta}. \tag{5.7}$$

It suffices to consider $|\widetilde{q}_2 - \widetilde{q}_1| < 1$, the case $|\widetilde{q}_2 - \widetilde{q}_1| \geq 1$ is trivial since the left side of (5.7) is bounded provided that $\beta \leq 1$.

By definition, $X_{+,n}(\widetilde{q}_2) - X_{+,n}(\widetilde{q}_1)$ is

$$(5.8)\quad \frac{1}{\check{n}^{\frac{1}{2}+d}}\left\{\sum_{t=\check{n}-(q_2-q_1)+1}^{\check{n}} u_{t+q_2}K_{+,t} + \sum_{t=1}^{\check{n}-(q_2-q_1)} u_{t+q_2}\left(K_{+,t}-K_{+,t+q_1-q_2}\right) - \sum_{t=1}^{q_2-q_1} u_{t+q_1}K_{+,t}\right\}.$$

Choose $\beta = 1$ in (5.7). Because $\check{n}\widetilde{q} = q$, $C3$ implies that $K_{+,t} = D(t/\check{n})(1+o(1))$, and the second moment of the third term of (5.8) is bounded by

$$\frac{D}{\check{n}^{1+2d}}\sum_{t,s=1}^{q_2-q_1}|t-s|^{2d-1}\frac{t}{\check{n}}\frac{s}{\check{n}} \leq D\,|\widetilde{q}_2-\widetilde{q}_1|^{3+2d},$$

so that we have that the last term of (5.8) satisfies the inequality (5.7). Similarly, because $K_{+,t} = D(1-t/\check{n})(1+o(1))$ as $t \to \check{n}$ by $C3$, we obtain that the second term of (5.8) is bounded by $D|\widetilde{q}_2-\widetilde{q}_1|^2(1-(1-(\widetilde{q}_2-\widetilde{q}_1)))^{2d} \leq D|\widetilde{q}_2-\widetilde{q}_1|^{2+2d}$ because $0<\widetilde{q}_2-\widetilde{q}_1<1$. Finally, by continuous differentiability of $K_+(u)$ for $u \in (0,1)$, we obtain that the second moment of the middle term in (5.8) is bounded by $D(\widetilde{q}_2-\widetilde{q}_1)^2\frac{1}{\check{n}^{1+2d}}\sum_{t,s=1}^{\check{n}-(q_2-q_1)}|t-s|^{2d-1} \leq D|\widetilde{q}_2-\widetilde{q}_1|^2$ because $0 \leq d < 1/2$. So, (5.7) holds true choosing $\beta = 1$ and $\delta = 1$ and hence $X_{+,n}(\widetilde{q})$ is tight. By identical arguments, $X_{-,n}(\widetilde{q})$ is also tight, which implies that the process $\check{n}^{\frac{1}{2}-d}(\widehat{r}_{a,+}(q)-\widehat{r}_{a,-}(q))$ is tight. This concludes the proof of the theorem. ■

Proof of Corollary 1. From the proof of Theorem 1, we only need to show that

$$\sup_{\check{n}<q<n-\check{n}} \check{n}^{\frac{1}{2}-d}\upsilon_n\left|E\left(\widetilde{r}_a(q)\right)\right| \to r\varrho(K_+).$$

But this is the case because by standard kernel manipulations and that $C3$ implies that $K_+(x) = K_-(-x)$, we obtain that under H_a given in (3.2),

$$\check{n}^{\frac{1}{2}-d}E\left(\widetilde{r}_a(q_0)\right) = \frac{r_n\left(x^0\right)}{\check{n}^{\frac{1}{2}-d}}\frac{1}{\check{n}}\sum_{t=|q-q_0|}^{\check{n}}K_{+,t} = \frac{r}{\upsilon_n}\int_{|q-q_0|/\check{n}}^{1}K_+(x)\,dx\,(1+o(1)),$$

by Brillinger (1981, p.15) as $\int_0^1|\partial K_+(u)/\partial u|\,du < \infty$ and where q_0/n is the closest point to x^0. The conclusion is standard because $\sup_{q\in Q_n}\int_{|q-q_0|/\check{n}}^1 K_+(x)\,dx \to \varrho(K_+)$. But under H_a, $\upsilon_n\check{n}^{\frac{1}{2}-d}r_n\left(x^0\right) = r$, so following the arguments preceding (5.6), it suffices to show that

$$\Pr\left\{\upsilon_n\left(\sup_{q\in Q_n}\check{n}^{\frac{1}{2}-d}\rho^{-\frac{1}{2}}(0;d)\,\widetilde{r}_a(q) - \frac{r\varrho(K_+)}{\upsilon_n\rho^{1/2}(0;d)} - \zeta_n\right) \leq x\right\} \to \exp\left(-2e^{-x}\right),$$

which is the case as we now argue. Proceeding as with Theorem 1, the last expression holds true because: (a) the finite dimensional distributions of $\check{n}^{\frac{1}{2}-d}\widetilde{r}_a(q) - r\varrho(K_+)/\upsilon_n$ converge to those of a Gaussian process with correlation structure $Corr(b)$; (b) the process $\check{n}^{\frac{1}{2}-d}\widetilde{r}_a(q)$ is tight proceeding as in Theorem 1. ■

Proof of Corollary 2. Since for any sequence of random variables, $X_1, ..., X_n$, $\Pr\left\{\max_{i\leq n} X_i > x\right\} \geq \Pr\left\{\max_{i=k,...,n-\ell} X_i > x\right\}$, it suffices to show that there exists $q \in Q_n$, such that

$$\Pr\left\{\upsilon_n\left(\check{n}^{\frac{1}{2}-d}\rho^{-\frac{1}{2}}(0;d)\left(\widetilde{r}_a(q)\right)-\zeta_n\right) > x\right\} \to 1$$

for all $x > 0$. Choose $q = q_0$, with q_0 as in Corollary 1. Proceeding as in the proof of Proposition 3, we have that $\check{n}^{\frac{1}{2}-d}\left|\widetilde{r}_a(q_0) - r\right| = O_p(1)$, and hence $\left|\widetilde{r}_a(q_0)\right| = O_p\left(\check{n}^{d-\frac{1}{2}}\right) + |r|(1+o(1))$. So, we obtain that

$$\upsilon_n\left(\check{n}^{\frac{1}{2}-d}\left|\widetilde{r}_a(q_0)\right| - \zeta_n\right) \to \infty$$

because $C4$ and that $d < 1/2$ imply that $\check{n}^{\frac{1}{2}-d}\zeta_n^{-2} = D\check{n}^{\frac{1}{2}-d}\log^{-1} n \to \infty$. The conclusion now follows by standard arguments. ■

Proof of Proposition 6. By definition and because $E^* u_t^* = 0$,

$$E^*\widetilde{r}_a^*(q) = \frac{1}{\check{n}}\sum_{t=1}^{n}\widehat{r}_a(t)\left(K_{+,t-q} - K_{-,t-q}\right).$$

So, we need to show that the right side is $o_p\left(\check{n}^{d-1/2}\right)$. Because $K_+(t/\check{n}) = K_-(-t/\check{n})$ by $C3$, it suffices to show that

$$\begin{aligned}
(a)\ \ \frac{1}{\check{n}}\sum_{t=1}^{n}\left(E\left(\widehat{r}_a(t)\right) - r\left(\frac{q}{n}\right)\right)\left(K_{+,t-q} - K_{-,t-q}\right) &= o\left(\check{n}^{d-1/2}\right)\\
(b)\ \ \frac{1}{\check{n}}\sum_{t=1}^{n}\left(\widehat{r}_a(t) - E\left(\widehat{r}_a(t)\right)\right)\left(K_{+,t-q} - K_{-,t-q}\right) &= o_p\left(\check{n}^{d-1/2}\right).
\end{aligned}$$

We begin with (a). By Proposition 3, we have that $E\left(\widehat{r}_a(t)\right) - r(t/n) = o\left(a^2\right) = o\left(\check{n}^{d-1/2}\right)$ because by $C5$, $(na)^{1/2-d}a^2 \leq D$. On the other hand, because $d < 1/2$, and $\tau = 2$, the proof of Proposition 3 implies that

$$\check{n}^{-1}\sum_{t=1}^{n}\left(r(t/n) - r(q/n)\right)\left(K_{+,t-q} - K_{-,t-q}\right) = o\left(a^2\right) = o\left(\check{n}^{d-1/2}\right)$$

by $C5$ and that $d < 1/2$. Next we show part (b). By definition, it equals

$$\frac{1}{\check{n}}\sum_{t=1}^{n}\left\{\frac{1}{\check{n}}\sum_{s=1}^{n}u_s K_{t-s}\right\}\left\{K_{+,t-q} - K_{-,t-q}\right\}.$$

where $K_{t-s} = \frac{1}{2}\left(K_{+,t-s} + K_{-,t-s}\right)$. Let's examine the contribution due to $K_{+,t-q}$, that from $K_{-,t-q}$ being similarly handled. The second moment is

$$\frac{\sigma_u^2}{\check{n}^4}\sum_{s_1,s_2=1}^{n}\gamma_u\left(|s_1 - s_2|\right)\sum_{t_1,t_2=1}^{n}K_{t_1-s_1}K_{t_2-s_2}K_{+,t_1-q}K_{+,t_2-q} = o\left(\check{n}^{2d-1}\right).$$

From here we conclude that part (b) and the proof of the proposition. ■

Proof of Theorem 2. As we argue with (5.6), we only need to show that

$$\Pr\left\{\upsilon_n\left(\sup_{\check{n}<q<n-\check{n}}\check{n}^{\frac{1}{2}-\widehat{d}}\rho^{-\frac{1}{2}}\left(0;\widehat{d}\right)\widetilde{r}_a^*(q)-\zeta_n\right)\le x\,|\mathcal{Y}\right\}\to\exp\left(2e^{-x}\right), \tag{5.9}$$

for $x>0$. To that end, we will show that $\check{n}^{\frac{1}{2}-\widehat{d}}\rho^{-\frac{1}{2}}\left(0;\widehat{d}\right)\widetilde{r}_a^*(q)$ converges, in bootstrap sense, to the Gaussian process $\mathcal{G}(q)$ in $\mathbb{D}[0,\infty)$, whose correlation structure is that given in (**??**). Proceeding as with the proof of Theorem 1 but using Proposition 8 instead of Proposition 4 there, it suffices to show the tightness condition. To that end, denote

$$X_{+,n}^*(\widetilde{q})=\frac{1}{\check{n}^{\frac{1}{2}+\widehat{d}}}\sum_{t=1}^{n}u_t^*K_+\left(\frac{t}{\check{n}}-\widetilde{q}\right),\quad \widetilde{q}=\frac{1}{\check{n}},\frac{2}{\check{n}},...,[a]^{-1}.$$

Arguing as in the proof of Theorem 1, it suffices to show the moment condition

$$E^*\left|X_{+,n}^*(\widetilde{q}_2)-X_{+,n}^*(\widetilde{q}_1)\right|^{2\beta}\le DH_n(\widetilde{q}_2,\widetilde{q}_1)\left|\widetilde{q}_2-\widetilde{q}_1\right|^{1+\delta} \tag{5.10}$$

with $\check{n}^{-1}<\widetilde{q}_2-\widetilde{q}_1$ and $H_n(\widetilde{q}_2,\widetilde{q}_1)=O_p(1)$. It suffices to consider $|\widetilde{q}_2-\widetilde{q}_1|<1$, the case $|\widetilde{q}_2-\widetilde{q}_1|\ge 1$ is trivial since the left side of (5.10) is bounded in probability.

By definition, $X_{+,n}^*(\widetilde{q}_2)-X_{+,n}^*(\widetilde{q}_1)$ is

$$\frac{1}{\check{n}^{\frac{1}{2}+\widehat{d}}}\left\{\sum_{t=\check{n}-(q_2-q_1)+1}^{\check{n}}u_{t+q_2}^*K_{+,t}+\sum_{t=1}^{\check{n}-(q_2-q_1)}u_{t+q_2}^*\left(K_{+,t}-K_{+,t+q_1-q_2}\right)-\sum_{t=1}^{q_2-q_1}u_{t+q_1}^*K_{+,t}\right\}. \tag{5.11}$$

Choosing $\beta=1$ and because $\check{n}\widetilde{q}=q$, the contribution into the left of (5.10) due to the third term of (5.11) is

$$\left|\frac{1}{\check{n}^{1+2\widehat{d}}}\sum_{t,s=1}^{q_2-q_1}E^*\left(u_{t+q_1}^*u_{s+q_1}^*\right)K_{+,t}K_{+,s}\right|\le DH_n(\widetilde{q}_2,\widetilde{q}_1)\left|\widetilde{q}_2-\widetilde{q}_1\right|^{3+2d},$$

because by $C3$, $|K_{+,t}|\le Dt/\check{n}$ and Proposition 6 implies that

$$\sum_{t,s=1}^{q}\left|E^*\left(u_{t+q_1}^*u_{s+q_1}^*\right)\right|=D\sum_{t,s=1}^{q}\left|E(u_tu_s)\right|(1+o_p(1))=Dq^{1+2d}(1+o_p(1)).$$

Observe that $H_n(\widetilde{q}_2,\widetilde{q}_1)=D\check{n}^{2\left(d-\widehat{d}\right)}$, which is $O_p(1)$ because by Proposition 5 and $C5$, $\left|\widehat{d}-d\right|=o_p\left(\log^{-1}n\right)$. So, the last term of (5.11) satisfies the inequality (5.10). Similarly, we obtain that

$$\begin{aligned}E^*\left|\frac{1}{\check{n}^{\frac{1}{2}+\widehat{d}}}\sum_{t=\check{n}-(q_2-q_1)+1}^{\check{n}}u_{t+q_2}^*K_{+,t}\right|^2 &\le D\check{n}^{2\left(d-\widehat{d}\right)}\left|\widetilde{q}_2-\widetilde{q}_1\right|^2\left(1-\left(1-\left(\widetilde{q}_2-\widetilde{q}_1\right)\right)\right)^{2d}\\ &\le DH_n(\widetilde{q}_2,\widetilde{q}_1)\left|\widetilde{q}_2-\widetilde{q}_1\right|^{2+2d},\end{aligned}$$

because $0<\widetilde{q}_2-\widetilde{q}_1<1$ and choosing $H_n(\widetilde{q}_2,\widetilde{q}_1)=D\check{n}^{2\left(d-\widehat{d}\right)}$. Finally, the continuous differentiability of $K_+(u)$ for $u\in(0,1)$ implies that the bootstrap second moment of the middle term in (5.11) is bounded by

$$D\left(\widetilde{q}_2-\widetilde{q}_1\right)^2\frac{1}{\check{n}^{1+2\widehat{d}}}\sum_{t,s=1}^{\check{n}-(q_2-q_1)}|t-s|^{2d-1}\le DH_n(\widetilde{q}_2,\widetilde{q}_1)\left|\widetilde{q}_2-\widetilde{q}_1\right|^2,$$

because $d < 1/2$. So, (5.10) holds true choosing $\delta = 1$ and hence $X_{+,n}(\widetilde{q})$ is tight. On the other hand, proceeding similarly as with $X^*_{+,n}(\widetilde{q})$,

$$X^*_{-,n}(\widetilde{q}) = \frac{1}{\check{n}^{\frac{1}{2}+d}} \sum_{t=1}^{n} u_t^* K_- \left(\widetilde{q} - \frac{t}{\check{n}}\right), \quad \widetilde{q} = 1/\check{n}, 2/\check{n}, ..., [a]^{-1},$$

is also tight. So, $\check{n}^{\frac{1}{2}-\widehat{d}}\left(\widehat{r}^*_+(q) - \widehat{r}^*_-(q)\right)$ is tight, which concludes the proof of the theorem because from (**??**), we have that the correlation structure converges in probability to that given in Proposition 2. ■

6. AUXILIARY LEMMAS

In what follows $\varphi(\lambda_j)$ will be abbreviated as φ_j for a generic $\varphi(\lambda)$ function. Let us introduce the following notation. Let $\check{h}_\ell(d) = \frac{1}{2m+1}\sum_{j=-m}^{m} \psi_{\ell+j}^{2d} I_{uu,\ell+j}$, where $\psi_j = |2\sin(\lambda_j/2)|$. With this notation, we have that Taylor's expansion up to the βth term implies that

$$\begin{aligned} \check{h}_\ell\left(\widehat{d}\right) - \check{h}_\ell(d) \quad &= \quad \sum_{p=1}^{\beta-1}\left\{\frac{1}{2m+1}\sum_{j=-m}^{m}\psi_{\ell+j}^{2d}\phi_{\ell+j}(p)\, I_{uu,\ell+j}\right\} \\ &\quad + D\left|d-\widehat{d}\right|^\beta \frac{\log^\beta n}{2m+1}\sum_{j=-m}^{m}\psi_{\ell+j}^{2\widetilde{d}} I_{uu,\ell+j} \end{aligned} \tag{6.1}$$

$$\begin{aligned} \widehat{h}_\ell - \check{h}_\ell\left(\widehat{d}\right) &= \sum_{p=0}^{\beta-1}\left\{\frac{1}{2m+1}\sum_{j=-m}^{m}\psi_{\ell+j}^{2d}\phi_{\ell+j}(p)\left(I_{\widehat{u}\widehat{u},\ell+j} - I_{uu,\ell+j}\right)\right\} \\ &\quad + D\left|d-\widehat{d}\right|^\beta \frac{\log^\beta n}{2m+1}\sum_{j=-m}^{m}\psi_{\ell+j}^{2\widetilde{d}}\left(I_{\widehat{u}\widehat{u},\ell+j} - I_{uu,\ell+j}\right), \end{aligned} \tag{6.2}$$

where $\widetilde{d}$ is an intermediate point between d and $\widehat{d}$, and $\phi_j(p) = \frac{2^p\left(d-\widehat{d}\right)^p}{p!}\log^p \psi_j$. Denote $q_\ell(p) = (2m+1)^{-1}\sum_{j=-m}^{m}\phi_{\ell+j}(p)\frac{h_{\ell+j}}{h_\ell}\left(I_{\varepsilon\varepsilon,\ell+j} - \frac{\sigma_\varepsilon^2}{2\pi}\right)$.

Lemma 1. *Assuming C1′, C2 and C5, uniformly in $r \le M$,*

$$\frac{r}{\tilde{n}}\sum_{\ell=1}^{\tilde{n}}\frac{\check{h}_\ell\left(\widehat{d}\right) - \check{h}_\ell(d)}{h_\ell}\cos(r\lambda_\ell) = \sum_{p=1}^{\beta-1}\frac{r}{\tilde{n}}\sum_{\ell=1}^{\tilde{n}} q_\ell(p)\cos(r\lambda_\ell) + O_p\left(\left|\widehat{d}-d\right|\log r\right). \tag{6.3}$$

Proof. The proof proceeds as that of Hidalgo's (2007) Lemma 7.1, and so it is omitted. ■

Lemma 2. *Let υ_n be a sequence of random variables such that $E|\upsilon_n| = O\left(M^{-1}\right)$. Assuming C1′, C2, C3 and C5, for all $r \le M$ and uniformly in $\ell = 1, ..., \tilde{n}$,*

$$\textbf{(a)}\ \frac{1}{\tilde{n}}\sum_{\ell=1}^{\tilde{n}}\left(\widehat{h}_\ell - \check{h}_\ell\left(\widehat{d}\right)\right)\cos(r\lambda_\ell) = O_p\left(\left|\widehat{d}-d\right|\right)\upsilon_n, \quad \textbf{(b)}\ \sup_{\ell=1,...,\tilde{n}}\left|\widehat{h}_\ell - \check{h}_\ell\left(\widehat{d}\right)\right| = O_p\left(\left|\widehat{d}-d\right|\right).$$

Proof. We begin with (**a**). Writing $\varphi_j(p) = \psi_j^{2d} \log^p \psi_j$, the contribution of the first term on the right of (6.2) into the left of (6.3) has as typical term

$$\frac{2^p \left(\widehat{d} - d\right)^p}{p!} \frac{1}{\tilde{n}} \sum_{\ell=1}^{\tilde{n}} \left(\frac{1}{2m+1} \sum_{j=-m}^{m} \varphi_{\ell+j}(p) \left(I_{\widehat{u}\widehat{u},\ell+j} - I_{uu,\ell+j} \right) \right) \cos(r\lambda_\ell). \tag{6.4}$$

Now, by definition of $\widehat{u}_t$, we have that $I_{\widehat{u}\widehat{u},j} - I_{uu,j}$ is

$$\frac{1}{2\pi n} \sum_{t,s=1}^{n} \left\{ u_t \left(r(s) - \widehat{r}_a(s) \right) + u_s \left(r(t) - \widehat{r}_a(t) \right) + \left(r(t) - \widehat{r}_a(t) \right) \left(r(s) - \widehat{r}_a(s) \right) \right\} e^{i(t-s)\lambda_j}.$$

On the other hand, because $r(t) - \widehat{r}_a(t) = \xi_t - \theta_t$, where $\xi_t = \check{n}^{-1} \sum_{q=1}^{n} \left(r(t) - r(q) \right) \widetilde{K}_{t-q}$; $\theta_t = \check{n}^{-1} \sum_{q=1}^{n} u_q \widetilde{K}_{t-q}$, with an obvious notation for $\widetilde{K}_t$, we have that (6.4) is governed by $2^p \left(\widehat{d} - d\right)^p / p!$ times

$$\frac{1}{\tilde{n}} \sum_{\ell=1}^{\tilde{n}} \left(\frac{1}{2m+1} \sum_{j=-m}^{m} \varphi_{\ell+j} \frac{1}{2\pi n} \sum_{t,s=1}^{n} \xi_t \xi_s e^{i(t-s)\lambda_{\ell+j}} \right) \cos(r\lambda_\ell) \tag{6.5}$$

$$+ \frac{1}{\tilde{n}} \sum_{\ell=1}^{\tilde{n}} \left(\frac{1}{2m+1} \sum_{j=-m}^{m} \varphi_{\ell+j} \frac{1}{2\pi n} \sum_{t,s=1}^{n} \theta_t \theta_s e^{i(t-s)\lambda_{\ell+j}} \right) \cos(r\lambda_\ell) \tag{6.6}$$

$$+ \frac{1}{\tilde{n}} \sum_{\ell=1}^{\tilde{n}} \left(\frac{1}{2m+1} \sum_{j=-m}^{m} \varphi_{\ell+j} \frac{1}{2\pi n} \sum_{t,s=1}^{n} u_t \left(\xi_s - \theta_s \right) e^{i(t-s)\lambda_{\ell+j}} \right) \cos(r\lambda_\ell). \tag{6.7}$$

Because the expression inside the parenthesis in (6.5) is positive, $|\cos\lambda| \le 1$ and $n = 4mM$, the absolute value of (6.5) is bounded by

$$\frac{D}{n^2} \sum_{t,s=1}^{n} \xi_t \xi_s \sum_{j=1}^{\tilde{n}} \varphi_j e^{i(t-s)\lambda_j} = \frac{D}{n} \sum_{t,s=1}^{n} \frac{1}{|t-s|_+^{1+2d}} |\xi_t \xi_s| + \frac{D}{n^2} \sum_{t,s=1}^{n} |\xi_t \xi_s|$$

which is $O\left(a^4\right)$ because the integrability of $(\partial/\partial u)\varphi(u)$ and Brillinger (1981, *p*.15) implies that $\left| n^{-1} \sum_{j=1}^{n} \varphi_j - \int_0^1 \varphi(u)\,du \right| = O\left(n^{-1}\right)$, and using $\int_0^1 \varphi(u) e^{itu} du = O\left(t_+^{-1-2d}\right)$ and that $|\xi_t| = O\left(a^2\right)$ by Proposition 3. Next we handle (6.6), which following step by step the proof in Robinson (1997, pp. 2077-2078), the first absolute moment of (6.6) is bounded by

$$\frac{D}{nm} \sum_{\ell=1}^{\tilde{n}} \sum_{j=-m}^{m} \left\{ \min\left(1, \frac{1}{(\ell+j)^2 a^2}\right) + \frac{\log n}{\ell+j} \right\} = O\left(a^{-2} m^{-1} n^{-1} + \frac{\log^2 n}{n} \right).$$

So $(6.5) + (6.6) = O\left(M^{-1}\right) O_p\left(\left| \widehat{d} - d \right| \right)$ because Proposition 1 implies that $\left| \widehat{d} - d \right| = O_p\left(a^{-1} m^{-1}\right)$ and then $C5$ part (*ii*). On the other hand, $(6.7) = O\left(M^{-1}\right) O_p\left(\left| \widehat{d} - d \right| \right)$ by an obvious use of the Cauchy-Schwarz inequality and the previous arguments. Next, as in Lemma 7.1, the contribution of the second term on the right of (6.2) into the left of (6.3) is $O_p\left(n^{-1}\right)$ by choosing β large enough. This concludes the proof of part (**a**).

The proof of part (**b**) is obvious by part (**a**) and using the usual chaining rule after observing that $\sup_{\ell=p,\ldots,q} \left| \widehat{h}_\ell - \check{h}_\ell\left(\widehat{d}\right) \right| \le \sum_{\ell=p}^{q} \left| \widehat{h}_\ell - \check{h}_\ell\left(\widehat{d}\right) \right|$. ■

Let $\widetilde{h}_\ell = (2m+1)^{-1} \sum_{j=-m}^{m} h_{\ell+j}$ and define

$$c_{r,n} = \frac{1}{\tilde{n}} \sum_{\ell=1}^{\tilde{n}} \log(h_\ell)\cos(r\lambda_\ell); \quad \widetilde{c}_{r,n} = \frac{1}{\tilde{n}} \sum_{\ell=1}^{\tilde{n}} \log\left(\widetilde{h}_\ell\right)\cos(r\lambda_\ell).$$

Lemma 3. *Let v_n be as in Lemma 2. Assuming C1 – C3 and C5, uniformly in $r \le M$*

$$\textbf{(a)}\ \widehat{c}_r - \widetilde{c}_{r,n} = \frac{1}{\tilde{n}} \sum_{\ell=1}^{\tilde{n}} \frac{\check{h}_\ell(d) - \widetilde{h}_\ell}{h_\ell} \cos(r\lambda_\ell) + O_p\left(\frac{1}{m}\right) + O_p\left(\left|\widehat{d} - d\right|\right) v_n. \tag{6.8}$$

$$\textbf{(b)}\ \widetilde{c}_{r,n} - c_{r,n} = O\left(M^{-2}\right); \quad \textbf{(c)}\ c_{r,n} - c_r = O\left(n^{-1}\right).$$

Proof. The proof proceeds as that of Hidalgo's (2007) Lemma 7.3, and so it is omitted. ■

Lemma 4. *Assuming C1 – C3 and C5, $E\left|\sum_{\ell=1}^{\tilde{n}} \frac{\check{h}_\ell(d)-\widetilde{h}_\ell}{h_\ell}\cos(r\lambda_\ell)\right|^2 = O(1)$.*

Proof. The proof is a standard extension of Hidalgo and Yajima's (2002) Theorem 1, so it is omitted. ■

Let us define

$$\widetilde{A}_{\ell,n} = \exp\left\{\sum_{r=1}^{M-1} \widetilde{c}_{r,n} e^{-ir\lambda_\ell}\right\}; \quad A_{\ell,n} = \exp\left\{\sum_{r=1}^{M-1} c_{r,n} e^{-ir\lambda_\ell}\right\}; \quad A^*_\ell = \exp\left\{\sum_{r=1}^{M-1} c_r e^{-ir\lambda_\ell}\right\}.$$

Lemma 5. *Let v_n be such that $E|v_n| = O\left(M/n^{1/2}\right)$. Assuming C1 – C3 and C5, uniformly in ℓ,*

$$(a)\ \widehat{A}_\ell - \widetilde{A}_{\ell,n} = O_p\left(\left|\widehat{d} - d\right|\right) + v_n, \quad (b)\ \widetilde{A}_{\ell,n} - A_{\ell,n} = O\left(M^{-2}\right); \quad A_{\ell,n} - A^*_\ell = O\left(m^{-1}\right).$$

Proof. The proof follows as that of Lemma 7.5 of Hidalgo (2007) and thus it is omitted. ■

Lemma 6. *For any $q_1 \le q_2 \in Q_n$, as $n \to \infty$, assuming C1 – C3 and C5,*

$$\check{n}^{1-2d} Cov^*\left(\widehat{r}^*_{a,+}(q_1) - \widehat{r}^*_{a,-}(q_1), \widehat{r}^*_{a,+}(q_2) - \widehat{r}^*_{a,-}(q_2)\right) \xrightarrow{P} \rho(b; d),$$

where the right side is as defined in Proposition 1.

Proof. By definition, $\check{n}^{1-2d} Cov^*\left(\widehat{r}^*_{a,+}(q_1) - \widehat{r}^*_{a,-}(q_1), \widehat{r}^*_{a,+}(q_2) - \widehat{r}^*_{a,-}(q_2)\right)$ is

$$\begin{aligned} &\check{n}^{1-2d} Cov^*\left(\widehat{r}^*_{a,+}(q_1), \widehat{r}^*_{a,+}(q_2)\right) + \check{n}^{1-2d} Cov^*\left(\widehat{r}^*_{a,-}(q_1), \widehat{r}^*_{a,-}(q_2)\right) \\ &- \check{n}^{1-2d} Cov^*\left(\widehat{r}^*_{a,+}(q_1), \widehat{r}^*_{a,-}(q_2)\right) - \check{n}^{1-2d} Cov^*\left(\widehat{r}^*_{a,-}(q_1), \widehat{r}^*_{a,+}(q_2)\right). \end{aligned} \tag{6.9}$$

As was done in the proof of Proposition 1, we will only examine the first term of (6.9), the other three terms follow similarly. This term is

$$\frac{1}{\check{n}^{1+2d}} \left\{ \sum_{t=1}^{\check{n}} \sum_{s=q_2-q_1+1}^{\check{n}+q_2-q_1} \left\{E^*\left(u^*_{t+q_1} u^*_{s+q_1}\right) - \delta_{t-s}\right\} K_{+,t} K_{+,s+q_1-q_2} + \sum_{t=1}^{\check{n}} \sum_{s=q_2-q_1+1}^{\check{n}+q_2-q_1} \delta_{t-s} K_{+,t} K_{+,s+q_1-q_2} \right\}.$$

However, it suffices to show that the first term on the right of the last displayed expression converges to zero in probability because by Proposition 1, the second term converges to $\rho_+(b; d)$.

Because $|K_{+,t}| \le D$ by $C3$, we have that the first term of the last displayed expression is bounded in absolute value by

$$\frac{D}{\check{n}^{1+2d}} \sum_{t=1}^{\check{n}} \sum_{s=q_2-q_1+1}^{\check{n}+q_2-q_1} \left|E^*\left(u^*_{t+q_1} u^*_{s+q_1}\right) - \delta_{t-s}\right| = \frac{D}{\check{n}^{2d}} \sum_{t=q_2-q_1+1}^{\check{n}+q_2-q_1} \left|E^*\left(u^*_{t+1} u^*_1\right) - \delta_t\right|$$

is $o_p(1)$ by standard arguments and then by Proposition 7. ■

Lemma 7. *Assuming C1 − C3 and C5, we have that for all* $\phi > 0$,

$$\sum_{j=1}^{n} E^{*}\left(n^{-1}\check{n}^{-1-2\widehat{d}}\left|\varsigma_{\pm,q}\left(\lambda_j\right)\eta_j^{*}\right|^{2}\mathcal{I}\left(n^{-1}\check{n}^{-1-2\widehat{d}}\left|\varsigma_{\pm,q}\left(\lambda_j\right)\eta_j^{*}\right|^{2}>\phi\right)\right)\xrightarrow{P}0,$$

where $\varsigma_{\pm q}\left(\lambda_j\right)=\widehat{k}_{\pm,q}\left(\lambda_j\right)\widetilde{g}^{1/2}\left(\lambda_j;\widehat{d}\right)\widehat{B}\left(\lambda_j\right)$ *with* $\widehat{k}_{\pm,q}\left(\lambda_j\right)=\sum_{t=1}^{n}K_{\pm,t-q}e^{it\lambda_j}$.

Proof. The proof is identical to that of Hidalgo's (2007) Lemma 7.9 and thus it is omitted. ∎

References

[1] Bickel, P.J. and Rosenblatt, M. (1973): "On some global measures of the deviations of density function estimates," *Annals of Statistics*, **1**, 1071-1095.

[2] Billingsley, P. (1968): "Convergence of probability measures," Wiley and Sons, New York.

[3] Brillinger, D.R. (1981): Time series, data analysis and theory. Holden-Day, San Francisco.

[4] Brockwell, P.J. and Davis, R.A. (1991): "Time series: theory and methods," *Springer*, New York.

[5] Bühlmann, P. (1997): "Sieve-bootstrap for time series," *Bernoulli*, **3**, 123-148.

[6] Chu, J.S. and Wu, C.K. (1993): "Kernel-type estimators of jump points and values of a regression function," *Annals of Statistics*, **21**, 1545-1566.

[7] Delgado, M.A. and Hidalgo, J. (2000): "Nonparametric inference on structural breaks," *Journal of Econometrics*, **96**, 113-144.

[8] Fuller, W.A. (1996): "Introduction to statistical time series," *Wiley and Sons*, New York.

[9] Giné, E. and Zinn, J. (1989): "Necessary conditions for the bootstrap of the mean," *Annals of Statistics*, **17**, 684-691.

[10] Hall, P. (1991): "On convergence rates of suprema," *Probability Theory and Related Fields,* **89**, 447-455.

[11] Härdle, W. and Marron, J.S. (1991): "Bootstrapping simultaneous error bars for nonparametric regression," *Annals of Statistics*, **19**, 778-796.

[12] Hidalgo, J. (1995): "A nonparametric conditional moment test for structural stability," *Econometric Theory*, **11**, 671-698.

[13] Hidalgo, J. (2007): "Specification testing for regression models with dependent data," STICERD EM/2007/518. Published version *Journal of Econometrics*, **143**, 143-165.

[14] Hidalgo, J. and Yajima, Y. (2002): "Prediction in the frequency domain under long-range processes with application to the signal extraction problem," *Econometric Theory*, **18**, 584-624.

[15] Hurvich, C.M. and Zeger, S.L. (1987): "Frequency domain bootstrap methods for time series." *Technical report* **87-115**, Graduate School of Business Administration, New York University.

[16] Hušková, M., Kirch, C., Prášková, Z. and Steinebach, J. (2008): "On the detection of changes in autoregressive time series, II. Resampling procedures," *Journal of Statistical Planning and Inference*, **138**, 1697–1721.

[17] Kreiss, J.-P. and Paparoditis, E. (2003): "Autoregressive-aided periodogram bootstrap for time series," *Annals of Statistics,* **31**, 1923-1955.

[18] Künsch, H.A. (1989): "The jackknife and the bootstrap for general stationary observations," *Annals of Statistics,* **17**, 1217-1241.

[19] Marinucci, D. and Robinson, P.M. (2003): "Weak convergence of multivariate fractional processes," *Stochastic Processes and its Applications*, **80**, 103-120**.**

[20] Müller, H.-J. (1992): "Change-points in nonparametric regression analysis," *Annals of Statistics*, **20**, 737-761.

[21] Perron, P. (2006): "Dealing with structural breaks," in Palgrave Handbook of Econometrics, Vol. 1, Chapter 8. Edited by T.C. Mills and K. Patterson.

[22] Pickands, J. (1969): "Asymptotic properties of the maximum in a stationary gaussian process," *Transactions of the American Mathematical Society*, **145**, 75-86.

[23] Pollard, D. (1980): Convergence of stochastic processes. Springer, New York.

[24] Rice, J. (1984): "Boundary modification for kernel regression," *Communications in Statistics: Theory and Methods*, **13**, 893-900.

[25] Robinson, P.M. (1995a): "Log-Periodogram regression for time series with long range dependence," *Annals of Statistics*, **23**, 1048-1072.

[26] Robinson, P.M. (1995b): "Gaussian semiparametric estimation of long-range dependence," *Annals of Statistics*, **23**, 1630-1661.

[27] Robinson, P.M. (1997): "Large-sample inference for nonparametric regression with dependent errors," *Annals of Statistics*, **25**, 2054-2083.
[28] Serfling, R.J. (1980): "Approximation theorems of mathematical statistics," *Wiley and Sons*, New York.
[29] Woodroofe, M. (1967): "On the maximum deviation of the sample density," *Annals of Mathematical Statistics*, **38**, 475-481.

London School of Economics, Houghton Street, London WC2A 2AE, U.K.

National and Kapodistrian University of Athens, Sofokleous 1, Athens 10559, Greece