Annals of the Institute of Statistical Mathematics (2020) 72:1419–1447
https://doi.org/10.1007/s10463-019-00730-6

Comparing the marginal densities of two strictly stationary linear processes

Paul Doukhan[1] · Ieva Grublytė[2] · Denys Pommeret[3] · Laurence Reboul[4]

Received: 3 July 2018 / Revised: 5 July 2019 / Published online: 22 August 2019

Abstract
In this paper, we adapt a data-driven smooth test to the comparison of the marginal distributions of two independent, short or long memory, strictly stationary linear sequences. Some illustrations are shown to evaluate the performances of our test.

Keywords Linear processes · Local Whittle estimator · Long memory · Schwarz's rule · Smooth test · Strictly stationary process

1 Introduction

Long-range dependent (LRD) strictly stationary processes, empirically observed by a slowly decaying autocovariance function, is a topic of active research in probability theory (see e.g., Robinson 2003; Surgailis 2000; Beran et al. 2013) but also in applications (Hsing 2000; Taqqu 1975). The importance of these processes in many fields, such as Econometrics, Finance, Hydrology and other physical sciences, is abundantly demonstrated [see Doukhan et al. (2003), Baillie (1996) and the references therein]. Long-memory linear processes are an important class of such processes. Some of their theoretical properties have been studied in e.g., Giraitis et al. (2012), Surgailis

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10463-019-00730-6) contains supplementary material, which is available to authorized users.

✉ Denys Pommeret
denys.pommeret@univ-amu.fr

1 Department of Mathematics, University Cergy-Pontoise, AGM-UMR 8080, 2 Bd. Adolphe Chauvin, 95000 Cergy-Pontoise, France

2 Institute of Mathematics and Informatics of Vilnius University, 4 Akademijos, 08663 Vilnius, Lithuania

3 ISFA, LSAF, and Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Univ. Lyon 1, 163 Avenue de Luminy, 13009 Marseille, France

4 CNRS, Centrale Marseille, I2M, Aix Marseille Univ, 163 Avenue de Luminy, 13009 Marseille, France

(2000) and Ho and Hsing (1999). Classical examples are fractional autoregressive-integrated moving average (ARFIMA) time series (see Dittmann and Granger 2002; Hosking 1981), extensively used in Econometrics. Nonlinear transformations of such processes allow to construct ordinary statistics of linear processes, such as the empirical variance or the empirical process. These models received much attention in recent years (see Abadir et al. (2014); Giraitis and Surgailis (1999); Dittmann and Granger (2002); Ho and Hsing (1997); Wu (2006, 2002); Sang and Sang (2016); Giraitis et al. (2012); Giraitis (1985); Ho (2000)). One of the interesting problems highlighted by these previous works is to make nonparametric inference on the marginal distribution of such processes.

In this paper, we propose a nonparametric test of comparison for the marginal densities of strictly stationary linear processes. More precisely, consider the linear processes $X = (X_t)_{t\in\mathbb{Z}}$ and $Y = (Y_t)_{t\in\mathbb{Z}}$ such that

$$X_t = \sum_{j<t} \alpha_{t-j}\epsilon_j \quad \text{and} \quad Y_t = \sum_{j<t} \beta_{t-j} e_j, \tag{1}$$

where the innovations ϵ_i's and e_i's are i.i.d. centered standard variables, $\alpha_j \sim c_\alpha |j|^{\delta_X - 1}$, $\beta_j \sim c_\beta |j|^{\delta_Y - 1}$ and $\delta_Y, \delta_X < 1/2$, c_α and c_β are real constants such that $c_\alpha \neq 0$, $c_\beta \neq 0$. Thus, long-range-dependent (LRD) and short-range-dependent (SRD) models are considered here, according to the values of δ_X and δ_Y (see Sect. 2), the non-summability of the covariances following from $\delta_Y, \delta_X > 0$. Denoting by f_X and f_Y the unknown marginal densities at any given time of X and Y, we want to test the nonparametric hypothesis

$$H_0 : f_X = f_Y, \tag{2}$$

based on the observation of a sample of the series.

More generally, if L and K are two measurable functions, we want to test the nonparametric hypothesis

$$H_0 : f_{L(X)} = f_{K(Y)}, \tag{3}$$

based on the observation of a sample of the series from $L(X)$ and $K(Y)$.

Such a nonparametric hypothesis testing is well-addressed for i.i.d. observations but very few works exist when the series exhibits some kind of dependence [see Doukhan et al. (2015) and references therein for a review]. In the short-memory context, Munk et al. (2009) proposed a goodness-of-fit test for α-mixing sequences; Koul and Surgailis (2010), Koul et al. (2013) and Koul et al. (2016) proposed goodness-of-fit tests for the density of long-memory linear processes. In the two-sample case, Doukhan et al. (2015) recently proposed a data-driven test for comparing the marginal densities of a bivariate strictly stationary-dependent process Z under different assumptions on the dependence structure of Z, including m-dependence, α-mixing, θ-weakly dependent and Gaussian long or short-memory processes. Extending Janic-Wróblewska and Ledwina (2000), Ghattas et al. (2011), Ignaccolo (2004) and Munk et al. (2009), the test proposed by Doukhan et al. (2015) is based on a coefficient's comparison of densities expansions along an appropriated orthogonal basis. The number of coefficients to compare is chosen by a data-driven method (see Kallenberg and Ledwina 1995). In this paper, we extend this test strategy to LRD case, combining the approach proposed in Doukhan et al. (2015) with some properties of transformations of linear processes obtained by Ho and Hsing (1997).

The rest of the paper is organized as follows. In Sect. 2, we recall some properties of linear processes and their transformations. In Sect. 3, we briefly recall the methodology of Doukhan et al. (2015)'s test. Section 4 contains our main theoretical results. In Sect. 5, we discuss the practical implementation of the test. Section 6 is devoted to numerical results when both processes X and Y are ARFIMA processes. In Sect. 7, a real dataset is studied, consisting of measurements of the widths of the annual rings of Bristlecone Pine in Nevada hills. Section 8 contains a discussion and possible extensions. The simulation study is completed in the supplementary material file with quadratic and exponential transformations of ARFIMA processes.

2 Main tool on transformations of linear processes

Let $\gamma(s) = \mathrm{cov}(X_t, X_{t+s})$, for all $t, s \in \mathbb{Z}$, be the autocovariance function of a discrete time stationary process $X = (X_t)_{t\in\mathbb{Z}}$. The decay rate of $\gamma(s)$ as s goes to infinity characterizes the type of dependence of the process X. More precisely,

$$X \text{ is SRD if } \Sigma_X = \sum_{s=-\infty}^{\infty} \gamma(s) \neq 0, \sum_{s=-\infty}^{\infty} |\gamma(s)| < \infty,$$
$$X \text{ is LRD if } \sum_{s=-\infty}^{\infty} |\gamma(s)| = \infty.$$

Another way to characterize dependence is to fix a decay rate of the autocovariance function:

$$\gamma(s) \sim c|s|^{2\delta-1}, \quad s \to \infty, \text{ for some } \delta < 1/2, \ c \in \mathbb{R}.$$

Then, this is clear that $\delta < 0$ and $\Sigma_X \neq 0$ together imply SRD. If $0 < \delta < 1/2$ and $c > 0$, then X is LRD. The case $\delta = 0$ is another SRD special case, and may lead to normal behaviors with $\sqrt{n \log n}-$normalization under additional tail regularity conditions (see Szewcsak (2012)). Restraining to $|\delta| < 1/2$ concerns the memory models $I(d)$ defined from the properties of the spectral density, and the memory parameter is often denoted d instead of δ, see definition 3.1.3. page 36 in Giraitis et al. (2012);

Finally, in the case of linear processes

$$X_t = \sum_{j<t} \alpha_{t-j}\epsilon_j,$$

where the ϵ_i's are i.i.d. and centered standard variables, another way to characterize dependence is to assume that $\alpha_j \sim c|j|^{\delta-1}$ for $\delta < 1/2$, c is a real nonzero constant. Thus, X is SRD if $\delta < 0$ and $\sum_{s=-\infty}^{\infty} \gamma(s) \neq 0$; it is LRD if $0 < \delta < 1/2$.

These two classes of processes (SRD and LRD) especially differ with asymptotic behavior of their partial sum process: a classical $\sqrt{n}-$CLT holds under mild conditions for SRD processes. This is no more true for LRD for which partial sums may converge (or not) toward a possibly non-Gaussian process at a slower rate than $\sqrt{n}$.

Remark 1 The process X admits a negative memory (NRD) if

$$\Sigma_X = \sum_{s=-\infty}^{\infty} \gamma(s) = 0, \qquad \sum_{s=-\infty}^{\infty} |\gamma(s)| < \infty.$$

Due to simple heredity properties of dependence conditions, transformations of SRD processes are still short memory, under mild conditions [see Rosenblatt (1956) or Dedecker et al. (2007)]. Heredity properties of LRD are not so simple as quoted in Doukhan et al. (2003) and Giraitis et al. (2012). Two particular classes of LRD processes have been intensively studied in the literature: Gaussian processes and linear processes. Hereafter, we briefly recall main results on transformations of linear processes; the Gaussian case is detailed in Dobrushin and Major (1979), and the linear case is considered in Ho and Hsing (1997). For that, we need to introduce the notion of power rank which is a main tool [we refer the reader to Giraitis et al. (2012) for more details].

Let L be a measurable function such that $\mathbb{E}(L(X)) = 0$. Following Ho and Hsing (1997), we say that L has "power rank" $m \geq 1$ with respect to the linear process X if

$$L_\infty^{(m)}(0) \neq 0 \quad \text{and} \quad L_\infty^{(r)}(0) = 0, \quad \text{for all} \quad 1 \leq r < m,$$

$$L_\infty^{(r)}(w) = \frac{\partial^r}{\partial w^r} \mathbb{E}(L(w + X_0)). \tag{4}$$

Let $0 < \delta < 1/2$, that is, the process is LRD. Under mild conditions on L and the innovations of X, it can be shown (see Sang and Sang (2016), Theorem 2.1) that

(C_1) If $m(1-2\delta) < 1$ then $L(X)$ is LRD. The convergence of the partial sums of $L(X)$ holds at a $n^{1-m(1/2-\delta)}$−rate to a Rosenblatt process of order m at point 1. In case $m = 1$, this process is just the fractional Brownian motion with Hurst parameter $H = \delta + 1/2$.

(C_2) If $m(1-2\delta) > 1$ then $L(X)$ is SRD (if moreover $\Sigma_{L(X)} \neq 0$). The convergence of the partial sums of $L(X)$ holds at a $\sqrt{n}$−rate to a standard Gaussian.

Then, let us consider two transformations $K(X)$ and $L(Y)$ of the linear processes X and Y with known in advance power ranks. We will see that properties (C_1) and (C_2) will allow to restrict our study to $K = L = \mathrm{Id}$ (the identity function). This argument is developed in Sect. 4.

3 Methodology

3.1 Construction of the test statistic

Let $Z = (X, Y)$, with X and Y independent strictly stationary processes satisfying

$$X_t = \sum_{j<t} \alpha_{t-j}\epsilon_j, \quad \text{and} \quad Y_t = \sum_{j<t} \beta_{t-j} e_j, \tag{5}$$

where the innovations ϵ_i and e_i are i.i.d. and centered standard variables, $\alpha_j \sim c_\alpha |j|^{\delta_X - 1}$, $\beta_j \sim c_\beta |j|^{\delta_Y - 1}$ with $\delta_X, \delta_Y < 1/2$, c_α and c_β are real positive constants. For simplicity, we shall set $\delta_Y \leq \delta_X$.

Let ν be a given probability measure with density g with respect to a reference measure λ (Lebesgue's measure for instance). We denote by f_X and f_Y the respective common unknown marginal densities of the X_t's and Y_t's with respect to λ and assume that they belong to $\mathbb{L}_2(\nu)$.

We want to test

$$H_0 : f_X = f_Y, \quad \text{against} \quad H_1 : f_X \neq f_Y, \tag{6}$$

based on the observations $\{Z_1, \ldots, Z_s, \ldots, Z_n\}$ of the process Z. For that task, we consider the expansions of f_X and f_Y along a family $(Q_j)_{j\in\mathbb{N}}$ of orthonormal functions in $\mathbb{L}_2(\nu)$:

$$f_X = \sum_{j\geq 0} a_j\, Q_j \quad \text{and} \quad f_Y = \sum_{j\geq 0} b_j\, Q_j,$$

with

$$a_j = \mathbb{E}(\widetilde{Q}_j(X_1)) = \int_{\mathbb{R}} Q_j(t) f_X(t) d\nu(t), \; b_j = \mathbb{E}(\widetilde{Q}_j(Y_1)) = \int_{\mathbb{R}} Q_j(t) f_Y(t) d\nu(t),$$

and $\widetilde{Q}_j = g Q_j$ for all $j \in \mathbb{N}$. It is clear that H_0 can be rewritten as

$$H_0 : a_i = b_i, \; \text{for all } i = 1, 2, \ldots \tag{7}$$

In view to define a test strategy for (7), let us set, for $k \geq 1$,

$$V_s^{(j)} = \widetilde{Q}_j(X_s) - \widetilde{Q}_j(Y_s), \qquad V_s(k) = \left(V_s^{(j)}\right)_{1\leq j\leq k}. \tag{8}$$

Next, define the k-dimensional random vector of the renormalized partial sums of the process $(V_s(k))_{s\in\mathbb{Z}}$

$$U_n(k) = (U_n^{(j)})_{1\leq j\leq k} = \frac{1}{u_n}\sum_{s=1}^{n} V_s(k),$$

where $(u_n)_{n>0}$ is an appropriately chosen positive norming sequence such that $\lim_{n\to\infty} u_n = \infty$. We will consider the sequence of Neyman's-type test statistics defined as follows:

$$N_n(k) = \|U_n(k)\|^2 = \sum_{j=1}^{k} \left(U_n^{(j)}\right)^2, \qquad k = 1, \ldots, d(n), \tag{9}$$

where $d(n) \to \infty$ as $n \to \infty$. For any fixed $k > 0$, the statistic $N_n(k)$ is the renormalized sum of the squared differences between the empirical estimates of the k first coefficients.

3.2 Data-driven selection criterion

In view to select the number of components k, we use in the sequel a data-driven selection rule inspired from Schwarz's Bayesian information criterion. At first step, select among every $k = 1, \ldots, d(n)$ the value K_n that minimizes an information criterion:

$$K_n = \min\Big\{ \underset{1\le k\le d(n)}{\arg\max} \quad (N_n(k) - k\log(n))\Big\}. \tag{10}$$

Once K_n is determined, we use the test statistic

$$\widetilde{N}_n = N_n(K_n). \tag{11}$$

This criterion has been used in Kallenberg and Ledwina (1995) for independent observations and consists of a modified version of Schwarz (1978)'s Bayesian information rule, based on an expansion of the maximum likelihood function. The extension of this rule to the paired context and to the strong-mixing context has been heuristically justified by Ghattas et al. (2011) and Munk et al. (2009), respectively (see Remarks 3 of both papers).

The test statistic $\widetilde{N}_n$ has been proposed in Doukhan et al. (2015) to test (3) for the margins X and Y of a strictly stationary process Z under different assumptions on the dependence structure of Z.

Before setting our main results on the asymptotic behavior of the statistic, let us heuristically describe the output of the rule (10) and the information that it provides on the distribution. Assume that H_0 holds. Thus, all the statistics $\{N_n(k)\}_{k>0}$ are bounded in probability so that the rule will choose $K_n = 1$. Now, assume that H_1 holds. Let k_0 be the first index leading to an unbounded statistic $N_n(k_0)$; this means that the first difference between f_X and f_Y is detected on the order k_0 coefficients. Since $\delta_Y \le \delta_X$, it will be seen that the order of magnitude at which $N_n(k_0)$ explodes is $u_n = \max(\sqrt{n}, n^{1/2+\delta_X})$ which is more than $\log n$ thus the rule will choose $K_n = k_0$.

4 Main results

From now on, we assume that

$$d(n) = o(\log(n)) \tag{12}$$

and we set

$$\widetilde{Q}_k^* = \widetilde{Q}_k - \mathbb{E}_0(\widetilde{Q}_k(X_0)). \tag{13}$$

We will denote by m_k the power rank of $\widetilde{Q}_k^*$. Following Doukhan et al. (2015)'s Theorems 1 and 2, we give below the asymptotic distribution of $\widetilde{N}_n$ under the null, as well as consistency results under contiguous alternatives in the particular case where the power rank of $\widetilde{Q}_1^*$ is equal to 1, that is

$$m_1 = 1. \tag{14}$$

We will discuss this condition in Sect. 8.2. In particular, this condition is easily satisfied when the support of X and Y is $\mathbb{R}$. In that case, the choice of a Gaussian reference measure ν with associated Hermite polynomials leads to $m_1 = 1$. As it will be detailed in Remark 2 below, our results rely on suitable assumptions ensuring that a central limit theorem holds for the partial sums $U_n(1)$ of the process $V^{(1)}$ and allowing to control the asymptotic behavior of the partial sums of the series of absolute autocovariances $\gamma_k(t)$ of the stationary processes $V^{(k)}$ under H_0:

$$r_n(k) = \sum_{t=0}^{n-1} |\gamma_k(t)| = \sum_{t=0}^{n-1} \left|\mathbb{E}_0(\widetilde{Q}_k^*(X_0)\widetilde{Q}_k^*(X_t) + \widetilde{Q}_k^*(Y_0)\widetilde{Q}_k^*(Y_t))\right|. \tag{15}$$

4.1 Convergence under the null hypothesis

In the rest of the paper, we denote by Z a Chi-squared random variable with one degree of freedom.

We set

$$U_n(k) = \begin{cases} n^{-1/2} \sum_{s=1}^{n} V_s(k), & \text{if } \delta_X < 0, \\ n^{-1/2-\delta_X} \sum_{s=1}^{n} V_s(k), & \text{if } \delta_X > 0. \end{cases} \tag{16}$$

Recall here that we assumed $\delta_Y \leq \delta_X$, and the statistic will then depend on the greatest value δ_X.

Moreover, consider the following assumptions:

(i) $\widetilde{Q}_k^*$ is q_k-Lipschitz.

(ii) There exists some B such that for all $i > 0$, $\frac{1}{i}\sum_{k=1}^{i} q_k^2 < B$.

Then, we prove

Proposition 1 *Let $\delta_X < 0$. Assume that (i) and (ii) hold. In addition, assume*

(iii) $\mathbb{E}_0(\epsilon_0^4) < \infty$ *and* $\mathbb{E}_0(e_0^4) < \infty$.
(iv) $\widetilde{Q}_1^*$ *is differentiable with continuous bounded derivative.*
(v) $\mathbb{E}_0((\widetilde{Q}_1^*(X_0))^2) < \infty$.

Then, under H_0, $\widetilde{N}_n \xrightarrow{\mathcal{L}} \sigma^2 Z$, with

$$\sigma^2 = \sum_{t \in \mathbb{Z}} \gamma_1(t).$$

Proposition 2 *Let $\delta_Y < 0 < \delta_X$. Assume that the above conditions (i), (ii) and (v) hold. In addition, assume*

(vi) $\mathbb{E}_0(\epsilon_0^8) < \infty$ *and* $\mathbb{E}_0(\nu_0^4) < \infty$.
(vii) $\widetilde{Q}_1^*$ *has a continuous bounded third-order derivative.*

Then under H_0, $\widetilde{N}_n \xrightarrow{\mathcal{L}} \sigma^2 Z$, where

$$\sigma^2 = c_\alpha^2 \left((\widetilde{Q}_1^*)_\infty^{(1)}(0) \right)^2 C(\delta_X)^2, \tag{17}$$

$$C(\delta_X) = \sqrt{\frac{\mathcal{B}(1 - 2\delta_X, \delta_X)}{\delta_X(1 + 2\delta_X)}}, \tag{18}$$

$a, b > 0$, $\mathcal{B}(a, b)$ is the Beta function and $(\widetilde{Q}_1^*)_\infty^{(1)}(0)$ is defined as in (4).

Proposition 3 *Let $0 < \delta_Y \leq \delta_X$. Assume that (i), (ii), (v) and (vii) hold. In addition, assume*

(ix) $\mathbb{E}_0(\epsilon_0^8) < \infty$ *and* $\mathbb{E}_0(\nu_0^8) < \infty$.

Then under H_0

(a) if $\delta_Y < \delta_X$, $\widetilde{N}_n \xrightarrow{\mathcal{L}} \sigma^2 Z$, where σ^2 is defined by (17),
(b) if $\delta_Y = \delta_X$, $\widetilde{N}_n \xrightarrow{\mathcal{L}} \sigma^2 Z$, where $\sigma^2 = (c_\alpha^2 + c_\beta^2) \left((\widetilde{Q}_1^*)_\infty^{(1)}(0) \right)^2 C(\delta_X)^2$ and $C(\delta_X)$ is defined by (18).

Remark 2 Notice that assumptions (i) and (ii) imply that

$$\frac{1}{d(n)} \sum_{k=1}^{d(n)} r_n(k) = O\left(\frac{u_n^2}{n} \right). \tag{19}$$

This assumption [denoted (A) in Doukhan et al. (2015)] drives the asymptotic behavior of $r_n(k)$, the absolute series of autocovariances of the process $V^{(k)}$. When the $V^{(k)}$'s are SRD, this assumption is satisfied by definition. In the LRD case, it controls the rate at which $r_n(k)$ goes to infinity. Moreover, this assumption implies, see Doukhan et al. (2015), that under H_0:

$$K_n \xrightarrow{\mathcal{P}} 1. \tag{20}$$

This means that rule (10) selects the first statistic $N_n(1)$ to test (3) under the null. That is why we only need to get the asymptotic distribution of the partial sums of $V^{(1)}$ to calibrate the test.

Other assumptions involved in Propositions 1–3 are those of Ho and Hsing (1997)'s Theorem 4.1 and Corollary 3.3. They allow to get the limit distributions of the partial sums of the processes $\widetilde{Q}_1^*(X)$ and $\widetilde{Q}_1^*(Y)$ in the SRD and LRD cases, respectively. According to (C_1) and (C_2), these limits and their convergence rates $u_{n,X}$ and $u_{n,Y}$ depend on the power rank m_1, δ_X and δ_Y. Here, $m_1 = 1$ which implies Gaussian asymptotic distributions. Since

$$U_n^{(1)} = \frac{1}{u_{n,X}} \sum_{t=1}^{n} V_t^{(1)} = \frac{1}{u_{n,X}} \sum_{t=1}^{n} \widetilde{Q}_1^*(X_t) - \frac{u_{n,Y}}{u_{n,X}} \left(\frac{1}{u_{n,Y}} \sum_{t=1}^{n} \widetilde{Q}_1^*(Y_t) \right),$$

then it ensures Doukhan et al. (2015)'s condition (B):

$$U_n(1) \xrightarrow{\mathcal{L}} U. \tag{21}$$

The asymptotic behavior of $U_n(1)$ is overseen by the process $\widetilde{Q}^*(X)$ or $\widetilde{Q}^*(Y)$ having the longest memory. When $\delta_X > \delta_Y$ $u_{n,Y}/u_{n,X} \to \infty$ and the limit distribution is that of the partial sums of $\widetilde{Q}^*(X)$. When $\delta_X = \delta_Y$, $u_{n,X}/u_{n,Y} = 1$ and the independence between X and Y allows to obtain the limit distribution.

Remark 3 Conditions (i), (ii), (iv), (v), (vii) on $\widetilde{Q}_1$ in Propositions 1–3 would be satisfied as soon as the support of the processes is $\mathbb{R}$, choosing the Gaussian distribution for ν with associated Hermite polynomials (see Sect. 6.2 for detail). Moreover, Conditions (iv), (v), (vii) are satisfied as soon as $\widetilde{Q}_1(x) = x$.

To conclude this section, we extend Propositions 1–3 to any transformations of X and Y. We consider two measurable functions L and K such that both power ranks of $\widetilde{Q}_1^* \circ L$ and $\widetilde{Q}_1^* \circ K$ are equal to 1. We want to test

$$H_0 : f_{L(X)} = f_{K(Y)} \text{ against } H_1 : f_{L(X)} \neq f_{K(Y)}.$$

We adapt the test statistics $\tilde{N}_n$ and $N_n(k)$, replacing (8) by

$$V_s^{(j)} = \widetilde{Q}_j \circ L(X_s) - \widetilde{Q}_j \circ K(Y_s), \qquad V_s(k) = \left(V_s^{(j)}\right)_{1 \le j \le k},$$

and changing (i), (iv), (v) and (vii) by

(i) $(\widetilde{Q}_k \circ L)^*$ and $(\widetilde{Q}_k \circ K)^*$ are q_k-Lipschitz.
(iv) $(\widetilde{Q}_1 \circ L)^*$ and $(\widetilde{Q}_1 \circ K)^*$ are differentiable with continuous bounded derivatives.

(v) $\mathbb{E}_0(((\widetilde{Q}_1 \circ L)^*(X_0))^2) < \infty$ and $\mathbb{E}_0(((\widetilde{Q}_1 \circ K)^*(Y_0))^2) < \infty$.
(vii) $(\widetilde{Q}_1 \circ L)^*$ and $(\widetilde{Q}_1 \circ K)^*$ have a continuous bounded third-order derivatives.

Corollary 1 – *Let $\delta_X < 0$. Assume that (i), (ii), (iii), (iv) and (v) hold. Then, under H_0, $\widetilde{N}_n \xrightarrow{\mathcal{L}} \sigma^2 Z$, with*

$$\sigma^2 = \sum_{t \in \mathbb{Z}} \mathbb{E}_0((\widetilde{Q}_1 \circ L)^*(X_0)(\widetilde{Q}_1 \circ L)^*(X_t) + (\widetilde{Q}_1 \circ K)^*(Y_0)(\widetilde{Q}_1 \circ K)^*(Y_t)).$$

– *Let $\delta_Y < 0 < \delta_X$. Assume that (i), (ii), (v), (vi) and (vii) hold. Then under H_0, $\widetilde{N}_n \xrightarrow{\mathcal{L}} \sigma^2 Z$, where*

$$\sigma^2 = c_\alpha^2 \left(((\widetilde{Q}_1 \circ L)^*)^{(1)}_\infty (0) \right)^2 C(\delta_X)^2, \tag{22}$$

where $C(\delta_X)$ is given by (18).
– *$0 < \delta_Y \le \delta_X$. Assume that (i), (ii), (v), (vii) and (ix) hold. Then under H_0,*

(a) if $\delta_Y < \delta_X$, $\widetilde{N}_n \xrightarrow{\mathcal{L}} \sigma^2 Z$, where σ^2 is defined by (22),
(b) if $\delta_Y = \delta_X$, $\widetilde{N}_n \xrightarrow{\mathcal{L}} \sigma^2 Z$, where

$$\sigma^2 = \left(c_\alpha^2 \left(((\widetilde{Q}_1 \circ L)^*)^{(1)}_\infty (0) \right)^2 + c_\beta^2 \left(((\widetilde{Q}_1 \circ K)^*)^{(1)}_\infty (0) \right)^2 \right) C(\delta_X)^2.$$

4.2 Convergence under contiguous alternatives

In the sequel, we show that for suitable alternatives, the test based on the limiting quantile is consistent. Namely, let us set $\gamma_k = a_k - b_k$ and consider the following two alternatives:

$$\begin{aligned} H_1^* &: \max_{k \le d(n)} \gamma_k = O(u_n/n), \\ H_1^{**} &: \exists\, K > 1 \text{ such that } \max_{k<K} \gamma_k = O(u_n/n) \\ &\quad \text{and } \gamma_K \, n/(\log(n) u_n) \to \infty \text{ as } n \to \infty. \end{aligned}$$

We will denote by $\mathbb{E}_1$ and $\mathbb{P}_1$ the expectation and probability under the corresponding alternative.

Proposition 4 – *Let $0 > \delta_X$. Assume that (i)–(iv) hold. Then, under H_1^*, $K_n \xrightarrow{\mathcal{P}} 1$ and $\tilde{N}_n$ has the limiting distribution given in Proposition 1.*
– *Let $\delta_Y < 0 < \delta_X$. Assume that (i) and (v)–(vii) hold. Then, under H_1^*, $K_n \xrightarrow{\mathcal{P}} 1$ and $\tilde{N}_n$ has the limiting distribution given in Proposition 2.*
– *Let $0 < \delta_Y < \delta_X$. Assume that (i) and (vi)–(viii) hold. Then, under H_1^*, $K_n \xrightarrow{\mathcal{P}} 1$ and $\tilde{N}_n$ has the limiting distribution given in Proposition 3(a).*

– *Let* $\delta_X = \delta_Y > 0$. *Assume that (i) and (vi)–(viii) hold. Then, under* H_1^*, $K_n \xrightarrow{\mathcal{P}} 1$ *and* $\tilde{N}_n$ *has the limiting distribution given in Proposition* 3(b).

Proposition 5 *For every item of Proposition* 4 *and under the same assumptions, one has under* H_1^{**}, $K_n \xrightarrow{\mathcal{P}} K$ *and* $\tilde{N}_n \xrightarrow{\mathcal{P}} +\infty$, *that is,* $\forall \epsilon > 0$, $\mathbb{P}(\tilde{N}_n < \epsilon) \to 0$.

Therefore, under H_1^*, the perturbation will not be detected by the test procedure, while it will be detected under H_1^{**}.

The proofs of Propositions 4–5 are relegated in the supplementary material file.

5 Practical implementation of the test

Choice of $d(n)$ The computation of the test requires a numerical choice for $d(n)$. Previous studies (see Doukhan et al. 2015) have shown that the empirical levels and powers obtained do not depend on $d(n)$ for sufficiently large values of this parameter. In practice, $d(n)$ will be set at 10.

Choice of the basis $(Q_j)_{j\geq 0}$ *and of the reference measure* A family $(Q_j)_{j\geq 0}$ and a reference measure have to be selected. In practice, this choice depends on the support of the distributions considered. When the support is $\mathbb{R}$, we can use the standard Gaussian distribution and its associated Hermite polynomials or the classical orthogonal Hermite functions. For distributions on $\mathbb{R}^+$, we can use the exponential measure and its associated Laguerre polynomials. For bounded densities, we can use the uniform measure and its associated Legendre polynomials [see Doukhan et al. (2015) for more details]. We detail the Gaussian case with Hermite polynomials in Sect.6.2.

Estimation of δ_X *and* δ_Y The memory parameters are estimated here using the semi-parametric local Whittle estimator [see Giraitis et al. (2012) for other choices of estimators]. Namely, δ_X may be estimated as

$$\widehat{\delta}_X = \arg \min_{-1/2\leq\delta\leq 1/2} U_n(\delta, X), \tag{23}$$

with

$$U_n(\delta, X) = \log\left(\frac{1}{r}\sum_{j=1}^{r} j^{2\delta} I_X\big(2\pi\frac{j}{n}\big)\right) - \frac{2\delta}{r}\sum_{j=1}^{r}\log j,$$

r is a positive integer such that $r < n/2$ and I_X denotes the periodogram based on $(X_1, \ldots, X_n)$:

$$I_X(u) = \frac{1}{2\pi n}\left|\sum_{j=1}^{n} X_j e^{iju}\right|^2.$$

Estimation of σ^2 The limiting distribution of the test statistic requires the estimation of the unknown variance parameter.

– SRD case: For σ^2 in the SRD case of Proposition 1, we propose to use a class of estimators based on kernel estimate of the spectral density which has been introduced in econometrics literature by Newey and West (1987) and Andrews (1991) [see Newey and West (1994) and Andrews and Monahan (1992) for refinements]. They take the form

$$\hat{\sigma}^2 = \sum_{|s|<n} K\left(\frac{s}{\ell_n}\right) \widetilde{\sigma}_s^2, \tag{24}$$

with

$$\widetilde{\sigma}_s^2 = \begin{cases} \dfrac{1}{n}\displaystyle\sum_{h=1}^{n-s} V_h^{(1)} V_{h+s}^{(1)} & \text{if} \quad 0 \le s \le n, \\ \widetilde{\sigma}_{-s}^2 & \text{if} \quad -n \le s < 0. \end{cases}$$

In Formula (24), ℓ_n is a bandwidth satisfying $\ell_n \to +\infty$ and $\ell_n/n \to 0$; K is a symmetric kernel function satisfying $K(0) = 1$, $|K(x)| \le 1$ for all $x \in \mathbb{R}$, K is continuous at $x = 0$ and at almost all other points of $\mathbb{R}$. Consistency of $\hat{\sigma}^2$ has been obtained by several authors under more or less strong conditions on $V^{(1)}$ and ℓ_n [see e.g., Andrews (1991)'s Theorem 1]. We refer the reader to Doukhan et al. (2015) for a practical implementation.

– LRD case: The estimation of the variance in the LRD cases of Propositions 2 and 3 requires to estimate δ_X, δ_Y, c_α, c_β and $(\widetilde{Q}_1^*)_\infty^{(1)}(0)$. For the last quantity, since $\widetilde{Q}_1^*$ is known, we shall empirically estimate the order one moment of its derivative $\widetilde{Q}_1^{(1)} = \dfrac{\partial}{\partial w}\widetilde{Q}_1^*$ by

$$\widehat{(\widetilde{Q}_1^*)_\infty^{(1)}}(0) = \frac{1}{n}\sum_{s=0}^{n-1} \widetilde{Q}_1^{(1)}(X_s).$$

The former parameters are estimated with the Whittle estimator, for instance:

$$\widehat{c_\alpha} = \widehat{k_\alpha}\Gamma(2 - 2\widehat{\delta}_X)\cos\big(\pi(1 + 2\widehat{\delta}_X)\big),$$

with

$$\widehat{k_\alpha} = \frac{1}{r}\sum_{j=1}^{r} u_j^{2\widehat{\delta}_X} I_X\big(2\pi\frac{j}{n}\big),$$

where $\widehat{\delta}_X$ is defined as previously. We obtain $\widehat{c_\beta}$ similarly. Finally, we get variance estimator of the form

$$\widehat{\sigma}^2 = \widehat{c}_\alpha^2\left(\widehat{(\widetilde{Q}_1^*)_\infty^{(1)}}(0)\right)^2 C(\widehat{\delta}_X)^2. \tag{25}$$

6 Simulation study under ARFIMA models

In this section, we study the finite-sample performances of our test on Monte Carlo simulations, performed on several sample sizes and models. The models consist of marginal distributions of different autoregressive fractionally integrated moving average (ARFIMA) models, with different long-memory parameters. We will not develop the short-memory case here since it has been largely studied in Doukhan et al. (2015). The nominal level is fixed at $\alpha = 5\%$.

6.1 Models

In order to study a wide range of LRD situations, the simulated examples consist of observation of sequences with respective sizes $n \in \{100, 200, 500, 1000, 2000\}$ of independent ARFIMA processes X and Y that we briefly describe here [see Granger and Joyeux (1980) and Hosking (1981) for more details]. Setting $B^k x_t = x_{t-k}$, recall that X is an ARFIMA(p, δ, q) with mean μ if

$$\Phi(B)(1-B)^{\delta}(X_t - \mu) = \Theta(B)\epsilon_t, \tag{26}$$

with

$$\Phi(B) = 1 - \Phi_1 B - \cdots - \Phi_p B^p, \qquad \Theta(B) = 1 + \Theta_1 B + \cdots + \Theta_q B^q,$$

$$(1-B)^{\delta} = \sum_{k=0}^{\infty} \frac{\Gamma(k-\delta)}{\Gamma(-d)\Gamma(k+1)} B^k,$$

and the ϵ_i's is a white noise sequence with mean zero and variance σ_ϵ^2. We fix $\Phi = 0$ (resp. $\Theta = 0$) if $p = 0$ (resp. $q = 0$). ARFIMA are particular cases of $I(\delta)$ models (see Sect. 3). Parameters p and q model short-term dependency effects. The memory parameter δ is allowed to assume every real value. The restriction of δ to integer values gives rise to classical ARIMA processes. For $|\delta| \geq 1/2$, this is a non-stationary process. For $\delta \in (0, 1/2)$, the process is long memory, and for $\delta \in (-1/2, 0)$, this is a negative memory process. For $\delta = 0$, this is a particular case of short-memory process, which corresponds to a stationary ARMA model. When $|\Phi(z)| > 0$ and $|\Theta(z)| > 0$ for all $|z| \leq 1$ and $|d| < 0.5$, X has a moving average representation given on page 181 of Giraitis et al. (2012).

In our design, X is a pure fractional model ARFIMA$(0, \delta_X, 0)$ with $\delta_X \in (0, 1/2)$. This model has the moving average representation:

$$X_t = \mu + \sum_{k=0}^{\infty} \alpha_k B^k \epsilon_k, \text{ with } \alpha_k = \frac{\Gamma(k+\delta_X)}{\Gamma(\delta_X)\Gamma(k+1)}.$$

Notice (see Giraitis et al. 2012, pp. 176–177) that $\alpha_k \sim (\Gamma(\delta_X))^{-1} k^{\delta_X - 1}$ as $k \to \infty$, which sticks to our hypotheses. The marginal distribution of a mean μ ARFIMA with standard Gaussian innovations is Gaussian with mean μ and variance

$$\sigma_\epsilon^2 \cdot \frac{\Gamma(1-2\delta)}{(\Gamma(1-\delta))^2}.$$

Hereafter, we denote by $\mathcal{M}(p, \delta, \mu, \sigma_\epsilon^2)$ the ARFIMA$(p, \delta, 0)$ with mean μ and variance innovation σ_ϵ^2. With this notation X is $\mathcal{M}(0, \delta_X, 0, 1)$. Varying the parameters

δ, p, μ and σ_ϵ^2, we investigate several null and alternative distributions of the process Y.

- Null hypothesis: the process Y is $\mathcal{M}(0, \delta_Y, 0, 1)$, with $\delta_Y = \delta_X$, that is both processes X and Y are the same.
- Mean deviation alternative hypotheses: Y is $\mathcal{M}(0, \delta_Y, \mu, 1)$, with $\delta_Y = \delta_X$ and $\mu \neq 0$.
- Variance deviation alternative hypotheses: Y is $\mathcal{M}(0, \delta_Y, 0, \sigma_\epsilon^2)$, with $\delta_Y = \delta_X$ and $\epsilon \neq 1$.
- Long-memory deviation alternative hypotheses: Y is $\mathcal{M}(0, \delta_Y, 0, 1)$, with $\delta_Y \neq \delta_X$.

Remark 4 Another second null hypothesis should be the case where the process Y is $\mathcal{M}(0, \delta_Y, 0, \sigma_\epsilon'^2)$, such that

$$\frac{\Gamma(1-2\delta_X)}{(\Gamma(1-\delta_X))^2} = \sigma_\epsilon'^2 \cdot \frac{\Gamma(1-2\delta_Y)}{(\Gamma(1-\delta_Y))^2}, \tag{27}$$

in such a way X and Y have the same marginal distribution.

However, it can be seen that for all alternatives, there is no combination of parameters such that the marginal distributions of X and Y are the same.

The ARFIMA sequences were computed using the fracdiff package R. We used a burn-in period of 10000.

6.2 Test statistics

The limiting distribution under the null of the test statistics depends on the unknown values δ_X and δ_Y. We used the semi-parametric local Whittle estimators defined by (23) to estimate these parameters. The form of the test statistic depends of the sign of δ_X and δ_Y, and whether $\delta_X = \delta_Y$ or not. We suggest in Sect. 8 a possible method in order to test these values.

The computation of the test statistics requires also the estimation of the asymptotic variance. We used the variance estimators given in (24) and (25).

Finally, according to the support $\mathbb{R}$ of the processes considered in our simulation study, we used here the standard Gaussian distribution and its associated Hermite polynomials. The proof is given in the supplementary material file. We have

$$\tilde{Q}_k(x) = H_k(x)\exp(-x^2/2),$$

where H_k is the kth-order Hermite polynomial.

Lemma 1 *Conditions (i), (ii), (iv), (v) and (vii) are satisfied. Moreover, the power rank of Q_1^* is equal to* 1.

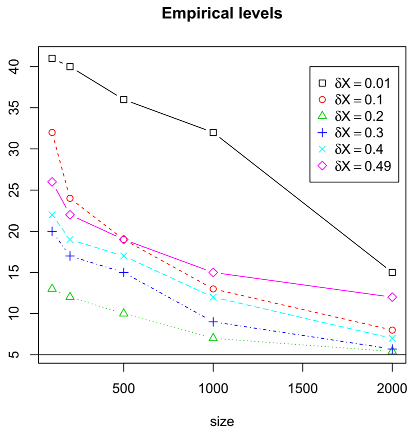


Fig. 1 Empirical levels for $\delta_X = 0.01, 0.1, 0.2, 0.3, 0.40.49$. X is $\mathcal{M}(0, \delta_X, 0, 1)$, Y is $\mathcal{M}(0, \delta_Y, 0, 1)$, and $\delta_X = \delta_Y$

6.3 Empirical levels

The empirical levels were defined as the percentage of rejection of the null hypothesis over 10,000 replications of the test statistics. We investigated their values for samples of size $n \in \{100, 200, 500, 1000, 2000\}$. We consider as null model the case where processes X and Y are the same and are governed by the model $\mathcal{M}(0, \delta_X, 0, 1)$. Different values of the long-memory parameter are considered, namely $\delta_X \in \{0.01, 0.1, 0.2, 0.3, 0.4, 0.45, 0.49\}$. The larger the value of δ_X is, the higher is the memory of the process. The empirical levels are represented in Fig. 1. Here, the values of the test statistic T are compared to the asymptotic distribution given in (b) of Proposition 3. We are confronted at two numerical difficulties when we browse the range $(0, 1/2)$ of δ_X. First, for small values of δ_X, the model is close to a short-memory process for which the asymptotic distribution is different. As seen in Proposition 1–3, there is no continuity between these asymptotic distributions and a larger size is necessary to distinguish LRD and SRD models. This phenomena is clearly represented in Fig. 1 where the empirical level associated with $\delta_X = 0.001$ is far from the asymptotic theoretical one of 5%. This is an instability due to the discontinuity of the type of range memory, that is, LRD or SRD. Second, for large values of δ_X, the convergence rate of the test statistic becomes very slow since $u_n = n^{1/2+\delta_X}$. Moreover, there is no more stationarity when $\delta_X = 0.5$. This phenomena is represented by the empirical level associated with $\delta_X = 0.49$ in Fig. 1 which is relatively far from the asymptotic level. This instability is due to the discontinuity of the stationarity. Globally, the convergence seems to be slow. It can be explained by the rate $1/u_n = n^{-1/2-\delta_X}$.

Figure 2 shows the empirical levels under the null, for $\sigma_\epsilon^2 \in \{1/2, 2\}$. These results show that the empirical level is smaller when the innovation variance σ_ϵ^2 is small. Then for $\sigma_\epsilon^2 = 1/2$, it can be observed that the empirical level is smaller than the theoretical one. Conversely, when $\sigma_\epsilon^2 = 2$, the empirical level is larger than 5%. This phenomena can be explained by the expression of the variance of the test statistic given in (17). Both quantities c_α and $C(\delta)$ are stable with respect to σ_ϵ^2. But the derivative

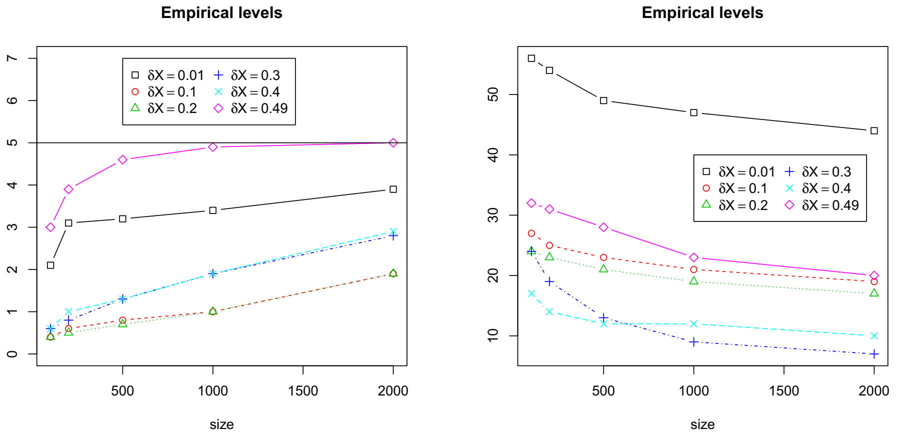


Fig. 2 Empirical levels for $\delta_X = 0.01, 0.1, 0.2, 0.3, 0.40.49$. X is $\mathcal{M}(0, \delta_X, 0, \sigma_\epsilon^2)$, Y is $\mathcal{M}(0, \delta_Y, 0, \sigma_\epsilon^2)$, and $\delta_X = \delta_Y$. On left: $\sigma_\epsilon^2 = 1/2$; on right: $\sigma_\epsilon^2 = 2$

$(\widetilde{Q}_1^*)_\infty^{(1)}(0)$ is very sensitive to the value of σ_ϵ^2. As an illustration, in the supplementary material file, Figure 10 represents the value of $(\widetilde{Q}_1^*)_\infty^{(1)}(0)$ with respect to the quantity σ_ϵ^2, when $n = 500$. Clearly, due to the factor $\left((\widetilde{Q}_1^*)_\infty^{(1)}(0)\right)^2$, the variance σ^2 of the limit null distribution decreases to zero when the innovation variance increases, and then the test statistic increases. To avoid this phenomena, one solution is to normalize both processes X and Y before testing their margins.

6.4 Empirical powers

The empirical powers were defined as the percentage of rejection of the null hypothesis over 1000 replications of the test statistics. We investigated their values for samples of size $n \in \{100, 200, 500, 1000, 2000\}$ and long-memory values $\delta_X \in \{0.01, 0.1, 0.2, 0.3, 0.4, 0.45, 0.49\}$. For the set of mean deviation alternative hypotheses, we used $\mu \in \{0.1, 1, 10\}$. For the set of variance deviation alternative hypotheses, we used $\sigma_\epsilon^2 \in \{0.5, 2\}$. For the set of long-memory deviation alternative hypotheses, we used (δ_Y, δ_X) varying in $\{(0.01, 0.49), (0.1, 0.4), (0.2, 0.3), (0.01, 0.1), (0.4, 0.49)\}$.

Long-memory deviations First alternatives considered are memory deviations where the process X is governed by the model $\mathcal{M}(0, \delta_X, 0, 1)$ and the process Y is governed by $\mathcal{M}(0, \delta_Y, 0, 1)$, with $\delta_X \neq \delta_Y$. For such alternatives, the adapted statistic is from Proposition 3(a). Figure 3 shows the empirical powers obtained with small or large difference between δ_X and δ_Y. Globally, alternatives seem to be well-detected, especially for large values of both δ_X and δ_Y, or for large differences between δ_X and δ_Y. Conversely, when δ_X and δ_Y are close together, or when both have small values, it can be observed that the power increases more slowly.

Mean deviation Second alternatives considered are mean deviations where the process X is governed by the model $\mathcal{M}(0, \delta_X, 0, 1)$ and the process Y is governed by

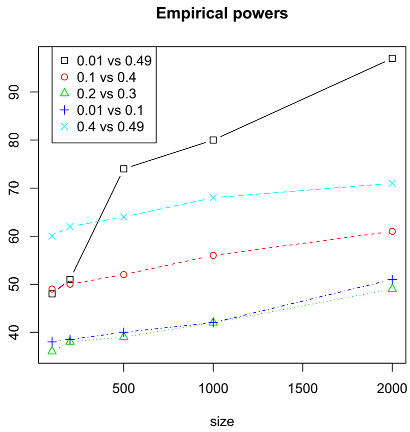


Fig. 3 Empirical powers. X id $\mathcal{M}(0, \delta_X, 0, 1)$, Y is $\mathcal{M}(0, \delta_Y, 0, 1)$, and $(\delta_Y, \delta_X) \in \{(0.01, 0.49), (0.1, 0.4), (0.2, 0.3), (0.01, 0.1), (0.4, 0.49)\}$

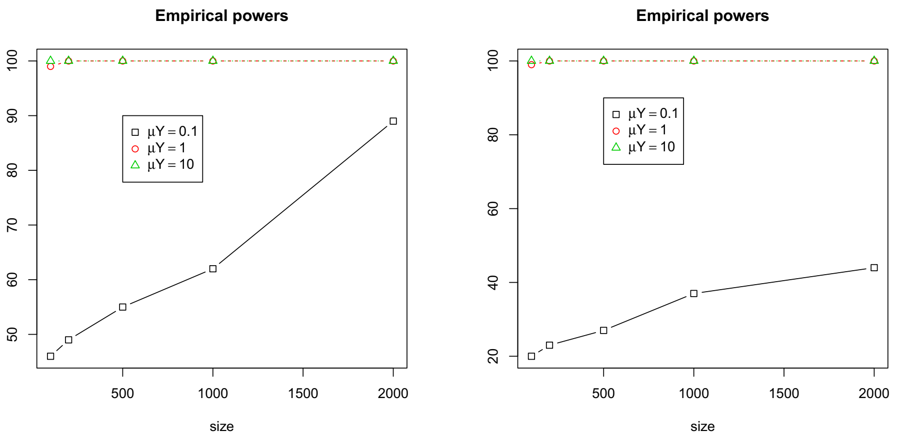


Fig. 4 Empirical powers. X is $\mathcal{M}(0, \delta_X, 0, 1)$ and Y is $\mathcal{M}(0, \delta_Y, \mu_Y, 1)$, with $\mu_Y = 0.1, 1, 10$. $\delta_Y = \delta_X = 0.01$ (left). $\delta_Y = \delta_X = 0.1$ (right)

$\mathcal{M}(0, \delta_X, \mu_Y, 1)$, with $\mu_Y \neq 0$. For such alternatives, the adapted statistic is from Proposition 3(b). Figures 4, 5 and 6 show the empirical powers obtained. As expected, the power is greater for larger values of μ_Y. The alternative with $\mu_Y = 0.1$ is not well-detected, except for $\delta_X = 0.01$, but this case corresponds to an instability of the test statistic. For large values of δ_X, the power decreases, as shown in Fig. 6, with $\delta_X = 0.4$ and $\delta_X = 0.49$.

Variance deviation The last alternatives are variance deviations where the process X is governed by the model $\mathcal{M}(0, \delta_X, 0, \sigma_\epsilon^2)$ and the process Y is governed by $\mathcal{M}(0, \delta_X, 0, \sigma_{\epsilon'}^2)$, with $\sigma_\epsilon^2 \neq \sigma_{\epsilon'}^2$. Figure 7 presents the empirical powers obtained with $\sigma_\epsilon^2 = 1$ and $\sigma_{\epsilon'}^2 \in \{1/2, 2\}$. The powers are relatively similar in both cases. Empirical powers are greater for small values of d, that is, close to the short-memory case ($d = 0$).

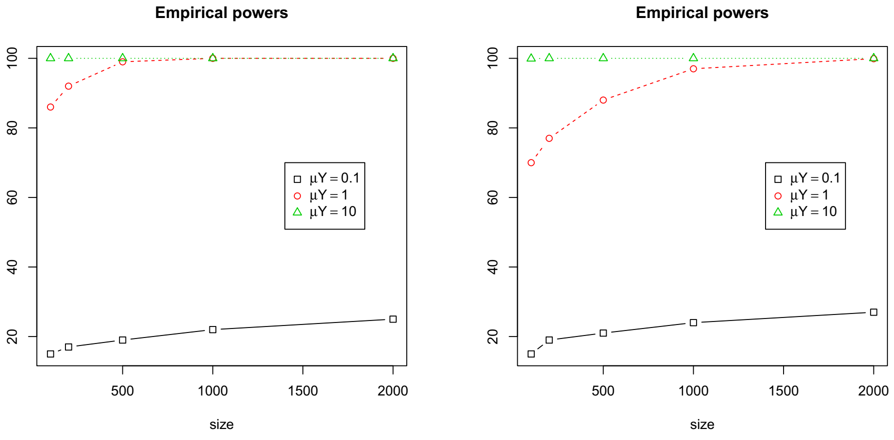


Fig. 5 Empirical powers. X is $\mathcal{M}(0, \delta_X, 0, 1)$ and Y is $\mathcal{M}(0, \delta_Y, \mu_Y, 1)$, with $\mu_Y = 0.1, 1, 10$. $\delta_Y = \delta_X = 0.02$ (left). $\delta_Y = \delta_X = 0.3$ (right)

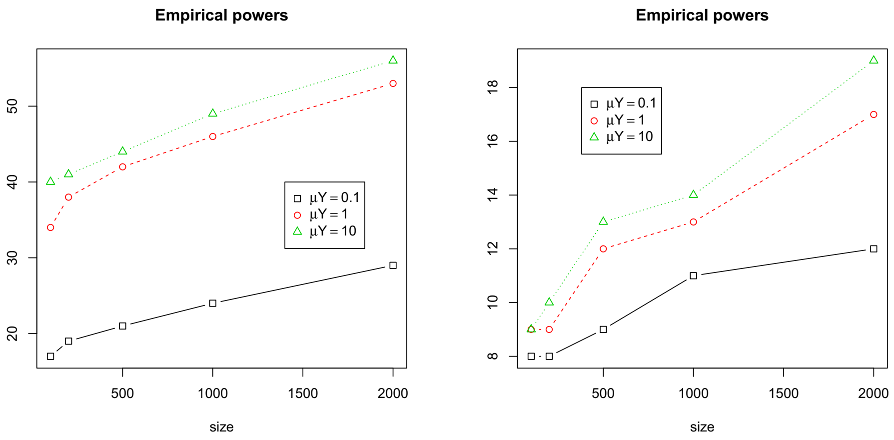


Fig. 6 Empirical powers. X is $\mathcal{M}(0, \delta_X, 0, 1)$ and Y is $\mathcal{M}(0, \delta_Y, \mu_Y, 1)$, with $\mu_Y = 0.1, 1, 10$. $\delta_Y = \delta_X = 0.04$ (left). $\delta_Y = \delta_X = 0.49$ (right)

7 Real data

One application of the two-sample test should detect a possible rupture in the stationarity of a LRD process. This topic exceeds the scope of this paper but we can explore briefly this idea as follows: we consider the Bristlecone Pine data which consist of width measurements of the annual rings of Bristlecone Pine in Nevada hills. The series contains yearly measurements on rings formed in the tree from year 1 to 1967. Such data are usually considered as LRD [see for instance the ARFIMA approach for the measurements of the widths of the annual rings of a Mount Campito Bristlecone pine in Hipel and McLeod (1994)]. The data are represented in Fig. 8.

The graph of this series could suggest a change around year 900. More precisely, minimum and maximum values are close, observed in year 810 and in year 896,

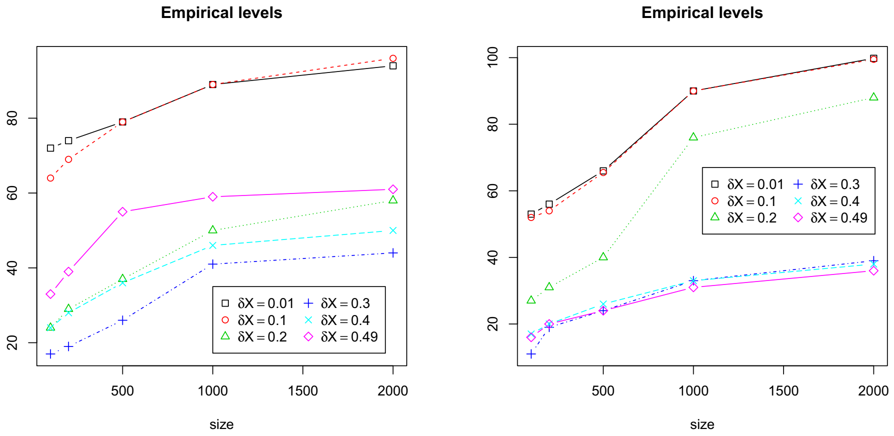
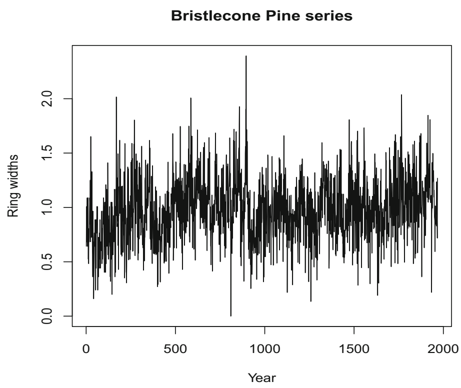


Fig. 7 Empirical powers for different values of d: under models $\mathcal{M}(0, \delta_X, 0, 1)$, and $\mathcal{M}(0, \delta_X, 0, 1/2)$ (left); under models $\mathcal{M}(0, \delta_X, 0, 1)$, and $\mathcal{M}(0, \delta_X, 0, 2)$ (right)

Fig. 8 Representation of the widths rings from year 1 to 1967



respectively. We then decided to cut the series into two new ones, one before 800, denoted by X, and the other after 1167, denoted by Y. We let enough delay time between the two series to assume independence of both processes. For simplicity, we keep the same length for both series, but this constraint could be easily relaxed. Figure 9 represents these series. We want to compare the stationary distributions on these two periods, that is, we want to test $H_0 : f_X = f_Y$. To use the appropriated statistics, we first estimated the δ parameters, getting $\delta_X = 0.27$ and $\delta_Y = 0.22$. At this step, we can decide to consider the test statistic given in Proposition 3(b), that is, when $\delta_X = \delta_Y$, or the statistic given in (a), that is, when $\delta_X \neq \delta_Y$. Again, this approach should be more detailed, with a more cautious cutting and with a test procedure on these parameters δ_X and δ_Y (see the Discussion in Sect. 8). But our purpose here is to indicate one possible way of application of our two-sample test. We applied the test statistics given in Proposition 3, in case (a) or (b). We obtained the following estimations and associated p-values:

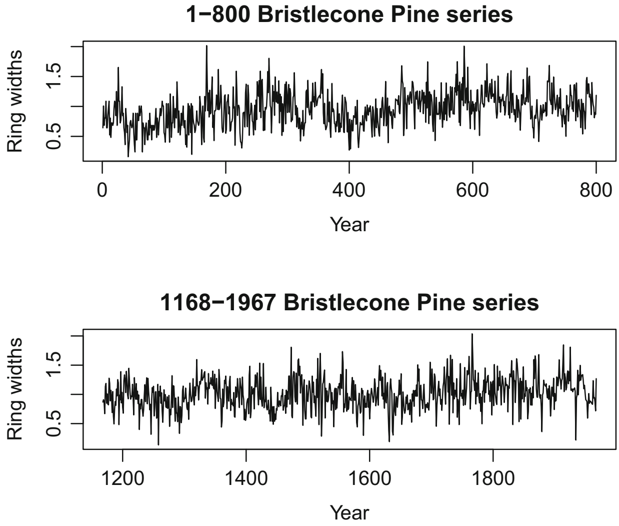


Fig. 9 Representation of the widths rings from year 1 to 800 (left) and from year 1168 to 1967 (right)

Case (a): $\tilde{N}_n/\sigma^2 = 0.17$ (p-value = 0.670).

Case (b): $\tilde{N}_n/\sigma^2 = 7.08$ (p-value = 0.008).

In conclusion, if we accept the equality of the dependence parameters δ_X and δ_Y (the b) case), then our test procedure leads to the rejection of the margins equality, that is, $f_X \neq f_Y$. Conversely, if we consider that $\delta_X \neq \delta_Y$ (the a) case), then we could accept the equality of the margins, that is, $f_X = f_Y$. In both cases, the conclusion is that the two series differ. There is a change even in the range of dependence, or in the stationary distribution.

8 Discussion

In this paper, we proposed a test for comparing two margins of SRD or LRD processes $X = (X_t)_{t\in\mathbb{Z}}$ and $Y = (Y_t)_{t\in\mathbb{Z}}$ having the form

$$X_t = \sum_{j<t} \alpha_{t-j}\epsilon_j \text{ and } \quad Y_t = \sum_{j<t} \beta_{t-j}e_j.$$

As shown in Corollary 1, we can extend our results to the case of transformations $K(X) = (K(X_t))_{t\in\mathbb{Z}}$ and $L(Y) = (L(Y_t))_{t\in\mathbb{Z}}$. A simulation study on LRD processes was considered, and it appeared that the test procedure yields good results for dependence parameters not too close to 0 or 0.5. The first case will lead to an instability due to the loss of stationarity, while the second case will lead to an instability due to the loss of LRD. In an ARFIMA context, it was noted that the test gave good results, with a possible normalization of both processes to conduct a stable test procedure. The results concerning simulation of transformations of ARFIMA processes are reported in the supplementary file. In such a situation, if we assume the knowledge of the power rank of the transformations, then from Theorem 2.2 of Sang and Sang (2016), we can deduce an estimator of the memory parameters of the original processes, and then we can construct our test statistic. In a more general context, it is necessary to know or to estimate the memory parameters of both original processes, possibly having observed them before transformation. The test was also used on the widths of the annual rings of Bristlecone Pine in Nevada hills to detect a possible rupture in the LRD process. The conclusion was that there exists a change around year 900. This change could be due to a difference of dependence parameters or a difference of margins. Eventually, such conditions on memory parameters or on power ranks are necessary to apply all our results. To conclude this work, we suggest possible ways for testing them.

8.1 Testing the value of the memory parameter

Propositions 1–3 of Sect. 4.1 highlight the dependence of limiting distribution of the test statistics under the null on the values of the parameters δ_X and δ_Y. However, the knowledge of these values may be not possible in some cases. Notice that several tests on the values of (δ_X, δ_Y) may be easily built from our results or those of Giraitis and Surgailis (1990).

Testing SRD versus LRD For I(d) processes, this may be done for instance using the V/S test developed in Giraitis et al. (2012). Based on our results, another test for

$$H_0 : \delta < 0 \text{ against } H_1 : \delta > 0,$$

consists of restricting the statistic given in (16) to the one sample case, with a sample $X_1, \cdots, X_n$ with memory parameter δ. We can modify our test statistic as follows: first we change (8) by

$$V_s^{(j)} = \widetilde{Q}_j(X_s) - \mathbb{E}(\widetilde{Q}_j(Y_s)), \qquad V_s(k) = \left(V_s^{(j)}\right)_{1\le j\le k}.$$

Then, in proof of **(A)**, (28) becomes

$$\mathbb{E}\big(U_n^{(k)}\big)^2 = \mathbb{E}\left(\frac{1}{u_n}\sum_{s=1}^{n}\widetilde{Q}_k^*(X_s)\right)^2 A_n(k).$$

The rest of the proof is mimicked to get (30), and we deduce the limit distribution of the statistic $\widetilde{N}_n$ under the null.

More precisely, under H_0, $\widetilde{N}_n \xrightarrow{\mathcal{L}} \sigma^2 Z$, where Z has a $\chi^2(1)$ distribution and

$$\sigma^2 = \sum_{t\in\mathbb{Z}} \left|\mathbb{E}_0(\widetilde{Q}_k^*(X_0)\widetilde{Q}_k^*(X_t))\right|,$$

while under H_1, $\widetilde{N}_n \xrightarrow{\mathcal{P}} +\infty$, which permits to detect the alternative.

Testing for the equality of the memory parameters To test

$$H_0 : \delta_Y = \delta_X \text{ against } H_1 : \delta_Y \neq \delta_X,$$

we can construct a test statistic based on the asymptotic normality of the Whittle estimator given by (23) under smoothness assumptions on the spectral density (see Giraitis and Surgailis (1990)). So, by independence of the samples from X and Y, $\sqrt{n}(\widehat{\delta}_X - \widehat{\delta}_Y)$ may be approached under H_0 by a centered Gaussian distribution with variance equal to the sum of the asymptotic variances of $\widehat{\delta}_X$ and $\widehat{\delta}_Y$ described in Giraitis and Surgailis (1990).

8.2 Testing the power rank

We have assumed throughout that the power rank m of $\tilde{Q}_1^*$ with respect to X (resp. Y) is equal to 1. This assumption leads to a simpler limit distribution of the test statistic in the case where $0 < \delta_X < 1/2$. First of all, this situation is the most encountered in the literature and there exists numerous cases of functions satisfying $m = 1$, see e.g., Bai and Taqqu (2018). In addition, the authors of this paper discuss the instability issue of power rank appearing in limit theorems under long memory and argue that a rank greater than 1 can be disturbed by a transformation and only a rank equal to 1 is stable. Indeed, assuming a higher-order rank when it is really not may result in underestimating the order of fluctuation of the statistic of interest. In order to perform valid inference, they suggest to adopt the assumption that the rank is always 1, regardless of any nonlinear transformation resulting from the statistical procedure.

However, we can construct a test to verify such a condition for linear processes. This test is related to the construction of our statistic but it could be extended to various situations. Considering the power rank of $\tilde{Q}_1^*$ associated with the process X, it consists of testing $H_0 : m = 2$ versus $H_1 : m = 1$, based on the observations of X. A more general test is in preparation in view to automatically detect the exact value of the power rank. From Ho and Hsing (1997), when $0 < \delta_X < 1/2$, we have

$$n^{m(1-2\delta_X)-2} \sum_{s=1}^{n} \widetilde{Q}_1^*(X_s) \to T_m = \gamma^{2m} \widetilde{Q}_\infty^{*(m)}(0) Z_m$$

for any value of the power rank m, where Z_m is a random variable with multiple Wiener-Itô integral representation given in Corollary 3.3 of Ho and Hsing (1997). We then consider the following test statistic

$$T = n^{m(1-2\delta_X)-2} \sum_{s=1}^{n} \tilde{Q}_1^*(X_s).$$

Under H_0, we have $T \to T_2 = \gamma^4 \widetilde{Q}_\infty^{*(2)}(0) Z_2$, and under H_1, we have the divergence of the test statistic, that is, $T \xrightarrow{\mathcal{P}} +\infty$. But this test procedure requires to estimate the memory parameter of X, so it does not apply to the transformation case where the original process X is not observed.

9 Proofs of Propositions 1 to 3

According to Doukhan et al. (2015)'s Theorem 1 and Remark 2, we have to prove

(A) : There exists some $C > 0$ and $n_0 > 0$ such that for all $n > n_0$,

$$\frac{1}{d(n)} \sum_{k=1}^{d(n)} \mathbb{E}_0 |U_n^{(k)}|^2 < C.$$

(B) : $U_n^{(1)} \xrightarrow{\mathcal{L}} U$ under H_0, where U is a random variable whose distribution possibly depends on a nuisance parameter.

9.1 Proof of (A)

Since X and Y are independent processes

$$\begin{aligned}
\mathbb{E}_0\big(U_n^{(k)}\big)^2 &= \mathbb{E}_0\left(\frac{1}{u_n}\sum_{s=1}^{n}\widetilde{Q}_k^*(X_s) - \frac{1}{u_n}\sum_{s=1}^{n}\widetilde{Q}_k^*(Y_s)\right)^2 \\
&= \mathbb{E}_0\left(\frac{1}{u_n}\sum_{s=1}^{n}\widetilde{Q}_k^*(X_s)\right)^2 + \mathbb{E}_0\left(\frac{1}{u_n}\sum_{s=1}^{n}\widetilde{Q}_k^*(Y_s)\right)^2 \\
&= A_n(k) + B_n(k). \qquad (28)
\end{aligned}$$

In order to control the terms at the right-hand side of (28), we will need the following Lemma

Lemma 2 *Let* $\delta < 1/2$ *and* $\{X_t = \sum_{j=0}^{\infty} \alpha_j \epsilon_{t-j}, t \in \mathbb{Z}\}$ *be a moving average process, where* $(\epsilon_t)_{t\in\mathbb{Z}}$ *is an i.i.d. standardized sequence and* $\sum \alpha_j^2 < \infty$. *Let us denote by* $(\mathcal{F}_t^\epsilon)_{t\in\mathbb{Z}}$ *the natural filtration of the process* (ϵ_t) *and set for all* $t > s$,

$$P_s(X_t) = \mathbb{E}\left(X_t|\mathcal{F}_s^\epsilon\right) - \mathbb{E}\left(X_t|\mathcal{F}_{s-1}^\epsilon\right).$$

Let Q *be a* q*-Lipschitz measurable function. Then,*

$$|P_s(Q(X_t))| \le q|\alpha_{t-s}|\left(|\epsilon_s| + \mathbb{E}(|\epsilon_0|)\right).$$

Proof

$$X_t = \sum_{j>0} \alpha_j \epsilon_{t-j} = \sum_{j \le t-s-1} \alpha_j \epsilon_{t-j} + \alpha_{t-s}\epsilon_s + \sum_{j\ge t-s+1} \alpha_j \epsilon_{t-j} = F + \alpha\epsilon + P$$

where $F \in \sigma(\epsilon_{s+1}, \ldots, \epsilon_t, \ldots)$, $\epsilon \equiv \epsilon_s \in \mathcal{F}_s^\epsilon$, and $P \in \mathcal{F}_{s-1}^\epsilon$.

Let $\mathbb{P}_F$ be the distribution of F and $\mathbb{P}_{(F,\epsilon)} = \mathbb{P}_F \otimes \mathbb{P}_\epsilon$ be the joint distribution of (ϵ, F). Then,

$$|P_s(Q(X_t))| = |\int Q(P + \alpha\epsilon + v)d\mathbb{P}_F(v) - \int Q(P + \alpha u + v)d\mathbb{P}_{(\epsilon,F)}(u, v)|.$$

By independence of present and future, one has $\mathbb{P}_{(\epsilon,F)} = \mathbb{P}_F\mathbb{P}_\epsilon$, so that

$$\begin{aligned}
|P_s(Q(X_t))| &= \left|\int Q(P + \alpha\epsilon + v)d\mathbb{P}_{(\epsilon,F)}(u, v) - \int Q(P + \alpha u + v)d\mathbb{P}_{(\epsilon,F)}(u, v)\right| \\
&= \left|\int (Q(P + \alpha\epsilon + v) - Q(P + \alpha u + v))d\mathbb{P}_{(\epsilon,F)}(u, v)\right| \\
&\le \int |Q(P + \alpha\epsilon + v) - Q(P + \alpha u + v)|d\mathbb{P}_{(\epsilon,F)}(u, v) \\
&\le q|\alpha| \int |\epsilon - u|d\mathbb{P}_{(\epsilon,F)}(u, v) \\
&\le q|\alpha_{t-s}| \left(|\epsilon_s| + \mathbb{E}_0|\epsilon_0|\right).
\end{aligned}$$

□

Noticing that for each measurable function Q one has $Q(X_t) = \sum_{s<t} P_s(Q(X_t))$ and applying Lemma 2 to $Q = \widetilde{Q}_k^*$ (it is a Lipschitz function by (i)),

$$\begin{aligned}
\mathbb{E}_0\Big(\sum_{s=1}^{n} \widetilde{Q}_k^*(X_s)\Big)^2 &= \sum_{r=1}^{n}\sum_{l=1}^{n} \mathbb{E}_0\left(\widetilde{Q}_k^*(X_r)\widetilde{Q}_k^*(X_l)\right) \\
&\le \sum_{r=1}^{n}\sum_{l=1}^{n}\sum_{s<\min(l,r)} \mathbb{E}_0\left(P_s(Q_k^*(X_l))P_s(Q_k^*(X_r))\right) \\
&\le \sum_{r=1}^{n}\sum_{l=1}^{n}\sum_{s<\min(l,r)} \mathbb{E}_0\left(|P_s(Q_k^*(X_l))||P_s(Q_k^*(X_r))|\right) \\
&\le 4\left(\mathbb{E}_0|\epsilon_0|\right)^2 q_k^2 \sum_{r=1}^{n}\sum_{l=1}^{n}\sum_{s<\min(l,r)} |\alpha_{l-s}||\alpha_{r-s}| \\
&\le Cq_k^2 L_n^{(X)}, \qquad (29)
\end{aligned}$$

where $C = 8\left(\mathbb{E}_0|\epsilon_0|\right)^2$ and

$$L_n^{(X)} = \sum_{r=1}^{n}\sum_{k=1}^{r}\sum_{j=0}^{n} |\alpha_j||\alpha_{j+k}|.$$

The way to obtain a upper bound for the right-hand side of (29) depends on δ_X.

– When $\delta_X < 0$, the α_js are absolutely summable and

$$\mathbb{E}_0\big(\sum_{s=1}^{n} \widetilde{Q}_k^*(X_s)\big)^2 \le C_X n \cdot q_k^2, \tag{30}$$

with

$$C_X = C\left(\sum_{j\ge 0} |\alpha_j|\right)^2.$$

– When $0 < \delta_X < 1/2$, then as in proof of (ii) in Proposition 3.2.1, p. 39 of Giraitis et al. (2012), as $k \to \infty$, one has

$$\sum_{j=0}^{\infty} |\alpha_j||\alpha_{j+k}| \sim \gamma^2 k^{2\delta_X - 1}\mathcal{B}(\delta_X, 1-2\delta_X),$$

where

$$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx, \;\; a > 0, b > 0$$

so that as $r \to \infty$

$$\sum_{k=1}^{r}\sum_{j=0}^{\infty} |\alpha_j||\alpha_{j+k}| \sim \gamma^2\,\mathcal{B}(\delta_X, 1-2\delta_X)\sum_{k=1}^{r} k^{2\delta-1} \sim \gamma^2 \cdot \frac{\mathcal{B}(\delta_X, 1-2\delta_X)}{2\delta_X} r^{2\delta_X}$$

and as $n \to \infty$

$$L_n^{(X)} \sim \gamma^2 \cdot \frac{\mathcal{B}(\delta_X, 1-2\delta_X)}{2\delta_X(2\delta_X+1)} \cdot n^{2\delta_X+1}$$

so that

$$\mathbb{E}_0\left(\sum_{s=1}^{n} \widetilde{Q}_k^*(X_s)\right)^2 \le C_{\delta_X} n^{1+2\delta_X} q_k^2, \tag{31}$$

with

$$C_{\delta_X} = C_X^* \cdot \frac{\mathcal{B}(\delta_X, 1-2\delta_X)}{\delta_X(2\delta_X+1)} \cdot n^{2\delta_X+1},$$

with $C_X^* = C_X/2$. Similar results occur for $\mathbb{E}_0\big(\sum_{s=1}^{n} \widetilde{Q}_k^*(Y_s)\big)^2$ replacing δ_X by δ_Y and the α_j's by the β_j's.

Therefore,

- When $\delta_Y \le \delta_X < 0$, $u_n = \sqrt{n}$ and

$$A_n(k) = \frac{1}{n}\mathbb{E}_0\big(\sum_{s=1}^{n} \widetilde{Q}_k^*(X_s)\big)^2 \le C_X q_k^2$$

by (30). Similarly,

$$B_n(k) \le C_Y q_k^2, \text{ with } C_Y = C\left(\sum_{j\ge 0}|\beta_j|\right)^2,$$

so that (A) obtains by (ii).

- When $\delta_Y < 0 < \delta_X$, then $u_n = n^{1/2+\delta_X}$,

$$A_n(k) = n^{-1/2-\delta_X}\mathbb{E}_0\big(\sum_{s=1}^n \widetilde{Q}_k^*(X_s)\big)^2 \le C_\delta q_k^2$$

and

$$B_n(k) = n^{-1/2-\delta_X}\mathbb{E}_0\big(\sum_{s=1}^n \widetilde{Q}_k^*(Y_s)\big)^2 \le C_Y n^{-2\delta_X} q_k^2 \to 0$$

so that (A) obtains by (ii).

- When $0 < \delta_Y \le \delta_X < 1$, then $u_n = n^{1/2+\delta_X}$,

$$A_n(k) \le C_{\delta_X} q_k^2 \quad \text{and} \quad B_n(k) \le C_{\delta_Y} n^{2(\delta_Y-\delta_X)} q_k^2,$$

with

$$C_{\delta_Y} = C''\gamma'^2 \frac{\mathcal{B}(\delta_Y, 1-2\delta_Y)}{2\delta_Y(1+2\delta_Y)}$$

which tends to 0 if $\delta_Y < \delta_X$ or equals $C_{\delta_Y} q_k^2$ if $\delta_Y = \delta_X$ so that (A) obtains by (ii). □

9.2 Proof of (B)

Let us set

$$U_n^{(1)} = \frac{1}{u_n}\sum_{s=1}^n \widetilde{Q}_1^*(X_s) - \frac{1}{u_n}\sum_{s=1}^n \widetilde{Q}_1^*(Y_s) = U_{X,n} - U_{Y,n}.$$

In order to prove **(B)**, we find in each case the limits U_X and U_Y of $U_{X,n}$ and $U_{Y,n}$ using Ho and Hsing (1997)'s results. Since $U_{X,n}$ and $U_{Y,n}$ are independent, the limiting distribution of $U_{X,n} - U_{Y,n}$ is that of $U_X - U_Y$.

– Assume $\delta \le \delta_Y < 0$, we can apply Ho and Hsing (1997)'s Theorem 4.1 to $U_{X,n}$ and $U_{Y,n}$ under (iii)–(v) and (i). Indeed, (iv) implies that the condition $C(1, 0, \lambda)$ of Theorem 4.1 is satisfied. Moreover (i) implies condition (3.3) of Theorem 4.1. Indeed, setting

$$X_{0,l} = \sum_{j=1}^l \alpha_j \epsilon_{-j},$$

one has $\mathbb{E}_0((\widetilde{Q}_1^*(X_0) - \widetilde{Q}_1^*(X_{0,l}))^2 \leq q_1^2 \sum_{j>l} \alpha_j^2$ which converges to zero as l goes to infinity. Therefore,

$$U_{X,n} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_X^2), \;\; U_{Y,n} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_Y^2)$$

$$\sigma_X^2 = \sum_{t\in\mathbb{Z}} \mathbb{E}_0(\widetilde{Q}_1^*(X_0)\widetilde{Q}_1^*(X_t)), \;\; \sigma_Y^2 = \sum_{t\in\mathbb{Z}} \mathbb{E}_0(\widetilde{Q}_1^*(Y_0)\widetilde{Q}_1^*(Y_t)),$$

so that

$$U_n^{(1)} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2),$$

where

$$\sigma^2 = \sigma_X^2 + \sigma_Y^2.$$

– When $0 < \delta_Y < \delta_X$, we have $u_n = n^{1/2+\delta}$ and we can apply on the one hand Ho and Hsing (1997)'s Corollary 3.3 with $k = m_1 = 1$ to $U_{X,n}$ under (vi), (vii) and (iii). Indeed, (vii) implies that $C(t, 0, \lambda)$ of the corollary is satisfied for $t = 0, 1, 2, 3$. Thus,

$$U_{X,n} \xrightarrow{\mathcal{L}} (\widetilde{Q}_1^*)_\infty^{(1)}(0)\sqrt{\gamma}C(\delta_X)Z_1,$$

where Z_1 is a standard Gaussian variable and $C(\delta_X)$ is defined by (18). On the other hand, under (v), (vi) and (i), one has

$$\frac{n^{1/2+\delta_X}}{n^{1/2}} U_{Y,n} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_Y^2),$$

where σ_Y^2 is defined as above. Then, $U_{Y,n} \xrightarrow{\mathcal{P}} 0$ so that

$$U_n^{(1)} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2),$$

with

$$\sigma^2 = \left((\widetilde{Q}_1^*)_\infty^{(1)}(0)\right)^2 c_\alpha C(\delta_X)^2.$$

– When $\delta_X = \delta_Y < 1$ Proposition 3(a), **(B)** follows from Ho and Hsing (1997)'s Corollary 3.3 under (i) and (vii)–(ix). Namely, $U_n^{(1)}$ has the same distribution as $(\widetilde{Q}_k^*)_\infty^{(1)}(0)C(\delta)\left(\sqrt{c_\alpha}Z_1 - \sqrt{c_\beta}Z_2\right)$, where Z_1 and Z_2 are two independent standard Gaussian variables. Therefore,

$$U_n^{(1)} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2),$$

with

$$\sigma^2 = \left((\widetilde{Q}_k^*)_\infty^{(1)}(0)\right)^2 c_\alpha C(\delta_X)^2(c_\alpha + c_\beta).$$

– For $0 < \delta_Y < \delta_X$, **(B)** follows from Ho and Hsing (1997)'s Corollary 3.3 under (i) and (vii)–(ix). Namely, as in Proposition 2,

$$U_{X,n} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2),$$

with

$$\sigma^2 = \left((\widetilde{Q}_1^*)^{(1)}_\infty(0)\right)^2 c_\alpha C(\delta_X)^2$$

and

$$\frac{n^{1/2+\delta_X}}{n^{1/2+\delta_Y}} U_{Y,n} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2),$$

so that $U_{Y,n} \xrightarrow{\mathcal{P}} 0$ and $U_n^{(1)}$ has the same distribution as in Proposition 2. $U_n^{(1)}$ has the same distribution as $(\widetilde{Q}_1^*)^{(1)}_\infty(0) K_{1,1-2\delta_X} Z_1$, where $K_{1,1-2\delta_X}$ is defined as in Equation (32) of Doukhan et al. (2015) and Z_1 is a standard Gaussian variable.

□

Acknowledgements The authors thank the referees and the associate editor for their careful reading and their suggestions which lead to improve the manuscript. The authors thank Donatas Surgailis for his kind help. This work has been developed within the MME-DII center of excellence (ANR-11-LABEX-0023-01) and PAI-CONICYT MEC Number 80170072.

References

Abadir, K. M., Distaso, W., Giraitis, L., Koul, H. L. (2014). Asymptotic normality for weighted sums of linear processes. *Econometric Theory*, *30*, 252–284.

Andrews, D. M. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, *59*, 817–856.

Andrews, D. W. K., Monahan, J. C. (1992). An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica*, *60*, 953–966.

Bai, S., Taqqu, M. S. (2018). How the instability of ranks under long memory affects large-sample inference. *Statistical Science*, *33*, 96–116.

Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, *73*, 5–59.

Beran, J., Feng, Y., Ghosh, S., Kulik, R. (2013). *Long-memory processes*: *Probabilistic properties and statistical methods*. Berlin: Springer.

Dedecker, J., Doukhan, P., Lang, G., Leon, J.R., Louhichi, S., Prieur, C. (2007). Weak dependence: With examples and applications. In*Lecture notes in statistics* (Vol. 190). New York: Springer.

Dittmann, I., Granger, C. (2002). Properties of nonlinear transformations of fractionally integrated processes. *Journal of Econometrics*, *110*, 113–133.

Dobrushin, R. L., Major, P. (1979). Non-central limit theorems for non-linear functionals of Gaussian fields. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, *50*, 27–52.

Doukhan, P., Oppenheim, G., Taqqu, M. S. (2003). *Theory and applications of long-range dependence*. Boston: Birkhauser.

Doukhan, P., Pommeret, D., Reboul, L. (2015). Data driven smooth test of comparison for dependent sequences. *Journal of Multivariate Analysis*, *139*, 147–165.

Ghattas, B., Pommeret, D., Reboul, L., Yao, A. F. (2011). Data driven smooth test for paired populations. *Journal of Statistical Planning and Inference*, *141*(1), 262–275.

Giraitis, L. (1985). Central limit theorem for functionals of a linear process. *Lithuanian Mathematical Journal*, *25*, 25–35.

Giraitis, L., Koul, H. L., Surgailis, D. (2012). *Large sample inference for long memory processes*. London: Imperial College Press.
Giraitis, L., Surgailis, D. (1999). Central limit theorem for the empirical process of a linear sequence with long memory. *Journal of Statistical Planning Inference*, *80*, 81–93.
Giraitis, L., Surgailis, D. (1990). A central limit theorem for quadratic forms in strongly dependent linear variables and its applications to the asymptotic normality of Whittle estimate. *Probabability Theory and Related Fields*, *86*, 87–104.
Granger, C. W. J., Joyeux, R. (1980). An introduction to long-memory series models and fractional differencing. *Journal of Time Series Analysis*, *1*, 15–29.
Hipel, K. W., McLeod, A. I. (1994). *Time series modelling of water resources and environmental system*. Amsterdam: Elsevier.
Ho, H. C. (2000). On functionals of linear processes with estimated parameters. *Statistica Sinica*, *12*, 1171–1190.
Ho, H. C., Hsing, T. (1997). Limit theorems for functionals of moving averages. *The Annals of Probability*, *25*, 1636–1669.
Ho, H. C., Hsing, T. (1999). On the asymptotic expansion of the empirical process of long memory moving averages. *The Annals Statistics*, *24*, 992–1024.
Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, *68*, 165–176.
Hsing, T. (2000). Linear processes, long-range dependence and asymptotic expansions. *Statistical Inference for Stochastic Processes*, *3*, 19–29.
Ignaccolo, R. (2004). Goodness-of-fit tests for dependent data. *Nonparametric Statistics*, *16*, 19–38.
Janic-Wróblewska, J. A., Ledwina, T. (2000). Data driven rank test for twosample problem. *The Scandinavian Journal of Statistics*, *27*, 281–297.
Kallenberg, W. C. M., Ledwina, T. (1995). Consistency and Monte Carlo simulation of a data driven version of smooth goodness-of-fit tests. *The Annals of Statistics*, *23*, 1594–1608.
Koul, H. L., Surgailis, D. (2010). Goodness-of-fit testing under long memory. *Journal of Statistical Planning and Inference*, *140*, 3742–3753.
Koul, H. L., Mimoto, N., Surgailis, D. (2013). Goodness-of-fit tests for long memory moving average marginal density. *Metrika*, *76*(205–224), 2013.
Koul, H. L., Mimoto, N., Surgailis, D. (2016). A goodness-of-fit test for marginal distribution of linear random fields with long memory. *Metrika*, *79*, 165–193.
Munk, A., Stockis, J. P., Valeinis, J., Giese, G. (2009). Neyman smooth goodness-of-fit tests for the marginal distribution of dependent data. *Annals of the Institute of Statistical Mathematics*, *63*, 939–959.
Newey, W. K., West, K. D. (1987). A simple positive semidefinite heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, *55*, 701–708.
Newey, W. K., West, K. D. (1994). Automatic lag selection in covariance matrix estimation. *Review of Economic Studies*, *61*, 631–653.
Robinson, P. M. (2003). *Time series with long memory*. Oxford: Oxford University Press.
Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proceedings of the National Academy of Sciences of the United States of America*, *42*, 43–47.
Sang, H., Sang, Y. (2016). Memory properties of transformations of linear processes. *Statistical Inference for Stochastic Processes*, *20*, 79–103.
Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464.
Surgailis, D. (2000). Long-range dependence and Appell rank. *The Annals of Probability*, *28*, 478–497.
Szewcsak, Z. (2012). Relative stability in strictly stationary random sequences. *Stochastic Processes and their Applications*, *122*(8), 2811–2829.
Taqqu, M. S. (1975). Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Zeitschrift fÃijr Wahrscheinlichkeitstheorie und Verwandte Gebiete*, *31*, 287–302.
Wu, W. B. (2002). Central limit theorems for functionals of linear processes and their applications. *Statistica Sinica*, *12*, 635–649.
Wu, W. B. (2006). Unit root testing for functionals of linear processes. *Econometric Theory*, *22*, 1–14.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.