The increment ratio statistic

DONATAS SURGAILIS[1*], GILLES TEYSSIÈRE[2†] AND MARIJUS VAIČIULIS[1*]

[1] *Institute of Mathematics and Informatics, Vilnius*

[2] *SAMOS–MATISSE, CES, University Paris 1 Panthéon–Sorbonne*

Centre for Finance, Göteborg University

December 2006

To appear in the *Journal of Mutivariate Analysis*

Abstract

We introduce a new statistic written as a sum of certain ratios of second order increments of partial sums process $S_n = \sum_{t=1}^n X_t$ of observations, which we call the Increment Ratio (IR) statistic. The IR statistic can be used for testing nonparametric hypotheses for $d-$integrated $(-1/2 < d < 3/2)$ behavior of time series X_t, including short memory $(d = 0)$, (stationary) long–memory $(0 < d < 1/2)$ and unit roots $(d = 1)$. If S_n behaves asymptotically as an (integrated) fractional Brownian motion with parameter $H = d+1/2$, the IR statistic converges to a monotone function $\Lambda(d)$ of $d \in (-1/2, 3/2)$ as both the sample size N and the window parameter m increase so that $N/m \to \infty$. For Gaussian observations X_t, we obtain a rate of decay of the bias $\mathrm{E}IR - \Lambda(d)$ and a central limit theorem $(N/m)^{1/2}(IR - \mathrm{E}IR) \to \mathcal{N}(0, \sigma^2(d))$, in the region $-1/2 < d < 5/4$. Graphs of the functions $\Lambda(d)$ and $\sigma(d)$ are included. A simulation study shows that the IR test for short memory $(d = 0)$ against stationary long–memory alternatives $(0 < d < 1/2)$ has good size and power properties and is robust against changes in mean, slowly varying trends and nonstationarities. We apply this statistic to sequences of squares of returns on financial assets and obtain a nuanced picture of the presence of long–memory in asset price volatility.

*Partially supported by the bilateral France-Lithuania scientific project Gilibert and the Lithuanian State Science and Studies Foundation, grant no.T-10/06.

†Address for correspondence: Gilles Teyssière, Centre for Finance, Göteborg University. Box 640, Vasagatan 1. SE 405 30 Göteborg, Sweden. e-mail: stats@gillesteyssiere.net

1 Introduction

The paper introduces a new statistic

$$IR := \frac{1}{N-3m}\sum_{k=0}^{N-3m-1}\frac{\left|\sum_{t=k+1}^{k+m}(X_{t+m}-X_t) \ + \ \sum_{t=k+m+1}^{k+2m}(X_{t+m}-X_t)\right|}{\left|\sum_{t=k+1}^{k+m}(X_{t+m}-X_t)\right| + \left|\sum_{t=k+m+1}^{k+2m}(X_{t+m}-X_t)\right|}, \tag{1.1}$$

with the convention $\frac{0}{0} := 1$. Here, $X_1,\ldots,X_N$ is a given sample of length N and $m = 1,2,\ldots$ is a bandwidth parameter. We call (1.1) the *Increment Ratio (IR)* statistic, since the sums in the numerator and denominator in (1.1) are second order increments, or differences, of partial sums $S_n := \sum_{t=1}^n X_t$. In fact, (1.1) can be rewritten as the integral:

$$IR \ = \ \frac{1}{(N/m)-3}\int_0^{(N/m)-3}\frac{\left|\Delta^2 S_{[m\tau]} \ + \ \Delta^2 S_{[m(\tau+1)]}\right|}{\left|\Delta^2 S_{[m\tau]}\right| + \left|\Delta^2 S_{[m(\tau+1)]}\right|}\,\mathrm{d}\tau, \tag{1.2}$$

where $\Delta f(\tau) := f(\tau+1) - f(\tau),\ \Delta^2 f(\tau) := \Delta(\Delta f(\tau))$ is the difference operator.

By definition, the IR statistic is always bounded by 0 and 1: $0 \le IR \le 1$ a.s. It is also location and scale free, i.e., does not change when X_t is replaced by an arbitrary linear combination $aX_t + b$, where $a \ne 0, b$ are arbitrary constants. Empirical simulations show that the IR statistic is quite insensitive to trends, local nonstationarities and heavy tails, see section 3 below. The limit of the IR statistic as $N, m, N/m \to \infty$ is related to the limit behavior of a (rescaled) partial sums process $S_{[m\tau]}, \tau \in [0,\infty)$, or the differenced process $\Delta^2 S_{[m\tau]}, \tau \in [0,\infty)$. In particular, if X_t is stationary and its partial sums process converges to a fractional Brownian motion (fBm) $B_{d+.5}(\tau), \tau \in [0,\infty)$ with (Hurst) parameter $d + .5 \in (0,1)$, in the way described in Assumption 1 (section 2), the IR statistic converges in probability to the expectation

$$\Lambda(d) := \mathrm{E}\left[\frac{|Z_d(0) + Z_d(1)|}{|Z_d(0)| + |Z_d(1)|}\right], \tag{1.3}$$

where $(Z_d(0), Z_d(1))$ have a jointly Gaussian distribution, with zero mean, unit variances and the covariance

$$\rho(d) := \mathrm{cov}(Z_d(0), Z_d(1)) \ = \ \frac{-9^{d+.5} + 4^{d+1.5} - 7}{2(4 - 4^{d+.5})}. \tag{1.4}$$

A similar convergence to the function $\Lambda(d)$ in (1.3) holds also in the case when X_t is nonstationary but the differenced process $U_t := X_t - X_{t-1}$ is stationary and the partial sums of U_t tends, in the way described in Assumption 2 (section 2), to a fBm $B_{d-.5}$ with Hurst parameter $d - .5 \in (0,1)$. The limit function $\Lambda(d)$ is defined in (1.3) for all $-.5 < d < 1.5, d \ne .5$, where

$$Z_d(\tau) \ := \ \frac{1}{\sqrt{|4 - 4^{d+.5}|}}\begin{cases}\Delta^2 B_{d+.5}(\tau), & -.5 < d < .5,\\ \sqrt{2d(2d+1)}\int_0^1 \Delta B_{d-.5}(\tau + s)\mathrm{d}s, & .5 < d < 1.5,\end{cases} \tag{1.5}$$

is a stationary Gaussian process with continuous time $\tau \in \mathbb{R}$, with zero mean unit variance and the covariance

$$\mathrm{E}Z_d(0)Z_d(\tau) \ = \ \frac{1}{2(4^{d+.5} - 4)}\Delta_s^2\Delta_t^2|t-s|^{2d+1}\Big|_{t-s=\tau}. \tag{1.6}$$

(For $d = .5$, (1.3) -(1.6) exist as the corresponding limits when d tends to .5). We call the process $Z_d(\tau)$ a *second increment fBm*. The function $\Lambda(d)$ is strictly monotone increasing on $(-.5, 1.5)$ (see the graph in Figure 1) and can be explicitly written as

$$\Lambda(d) = \Lambda_0(\rho(d)),$$

where

$$\Lambda_0(r) \ := \ \frac{2}{\pi}\arctan\left(\sqrt{\frac{1+r}{1-r}}\right) + \frac{1}{\pi}\sqrt{\frac{1+r}{1-r}}\log\left(\frac{2}{1+r}\right). \tag{1.7}$$

The above mentioned consistency property of the IR statistic is very general and essentially uses only a "fBm asymptotics" of the partial sums process $S_{[m\tau]}$, see section 2 for details. To obtain more detailed information concerning convergence rates and the asymptotic distribution of the IR statistic, we assume that X_t is a Gaussian process. Theorem 2.4 obtains the decay rate of the bias $\mathrm{E}IR - \Lambda(d)$, as the window parameter $m \to \infty$, under semiparametric assumptions on the spectral density of stationary processes X_t (case $-.5 < d < .5$) and $U_t = X_t - X_{t-1}$ (case $.5 < d < 1.5$). Under similar assumptions on X_t and U_t we obtain the central limit theorem:

$$(N/m)^{1/2}(IR - \mathrm{E}IR) \ \to_{\mathrm{D}} \ \mathcal{N}(0, \sigma^2(d)) \qquad (N, m, N/m \to \infty), \tag{1.8}$$

see Theorem 2.5 where

$$\sigma^2(d) \ := \ 2\int_0^\infty \mathrm{cov}\left(\frac{|Z_d(0)+Z_d(1)|}{|Z_d(0)|+|Z_d(1)|}, \frac{|Z_d(\tau)+Z_d(\tau+1)|}{|Z_d(\tau)|+|Z_d(\tau+1)|}\right)\,\mathrm{d}\tau, \tag{1.9}$$

and where $Z_d(\tau)$ is defined in (1.5). The CLT in (1.8) holds for $-.5 < d < 1.25$, $d \neq .5$. (For $d \in (1.25, 1.5)$ the integral in (1.9) diverges and the CLT in (1.8) most likely fails.) The graph of $\sigma(d)$ obtained with the help of *Mathematica 4.0* is shown in Figure 2.

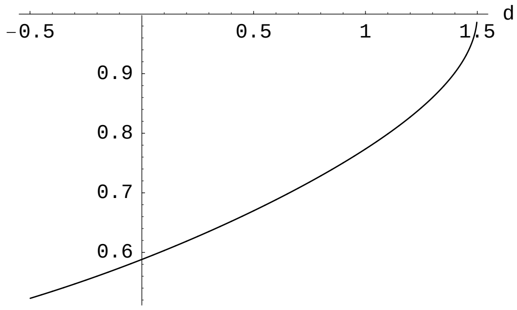

Figure 1: The graph of $\Lambda(d)$

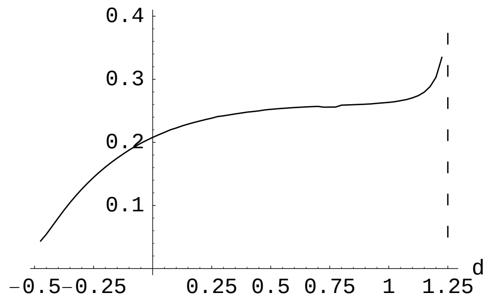

Figure 2: The graph of $\sigma(d)$

The above mentioned results suggest using the IR statistic for testing various nonparametric hypotheses, e.g., stationary short memory vs. stationary long memory, stationary long–memory vs. nonstationary unit root, etc. Several statistics and tests have been proposed in the literature for testing such hypotheses. Among them, we mention the score test (Robinson, 1994), the Lagrange multiplier test (Lobato and Robinson, 1998), the modified R/S statistic (Lo, 1991), the KPSS statistic (Kwiatkowski *et al.*, 1992), the V/S statistic (Giraitis *et al.*, 2003). The last three statistics are essentially based on fBm-type behavior of the partial sums process of X_t; however, their limit distributions are non Gaussian and normalizations depend on the (possibly unknown) memory parameter d. Section 3 provides a finite sample simulation study of the IR test of short memory ($d = 0$) vs. long–memory ($d > 0$), with the critical region

$$IR - \Lambda(0) > z_\alpha \sigma(0)\sqrt{\frac{m}{N-3m}}, \tag{1.10}$$

where $\Lambda(0) \approx .5881$, $\sigma(0) \approx .2080$, and z_α is the standard normal quantile. We study the empirical size of the test (1.10) under "AR+stochastic trend" and "AR+deterministic trend" models, the empirical power under "FARIMA with memory breaks" model, and the robustness of that test under nonstationary models and heavy–tailed α–stable distributions.

Long–range dependent processes can be confused with trended processes and change–point processes; see e.g., Bhattacharya *et al.* (1983). One can distinguish between these alternatives by resorting to estimators of the long–range dependent parameter that are robust to the presence of trends, change–points and nonstationarities. Abry and Veitch (1998) introduced a wavelet estimator of the memory parameter robust to deterministic linear and polynomial trends, which works for large samples, e.g., $N = 10000$; see also Abry *et al.* (2003), Teyssière and Abry (2005). However, the asymptotic variance of this estimator depends on the memory parameter and the corresponding confidence intervals with the sample size used in this paper ($N = 1000$) are inconclusive; see also Bardet *et al.* (2000).

Künsch (1986) and later Sibbertsen (2003) proposed procedures for discriminating between trends and long–range dependence based on the periodogram. Since tapering the periodogram allows to get rid of small trends and slowly varying trends, the discrepancy between the spectral estimates obtained with and without tapering the periodogram constitutes an evidence of spurious long–range dependence.

Dolado *et al.* (2005) proposed an extension of the fractional Dickey–Fuller test for long–range dependence against the alternative of short–range dependence, robust to the presence of a single break. Recently, Berkes *et al.* (2006) proposed a CUSUM test for discriminating between long–range dependence and change–points, including the case multiple change–points. This is of interest when dealing with large samples, as for large samples the occurrence of a single change–point is unlikely. We then compare the performance of our test with this one for the case of nonhomogeneous processes.

Comparisons with the V/S and Robinson's (1994) tests are provided, indicating that in the presence of stochastic trend, deterministic trends or change–points, the IR test clearly outperforms the other tests.

The robustness of the IR test with respect to change–points and other structural changes can be explained by the fact that the IR statistic uses "local data" or "moving" subsamples of length $3m$, while other above mentioned tests use "global" quantities such as the sample mean or periodogram estimates. In the case of a few change–points, only a small fraction of subsamples of length $3m$ (ratios in (1.1)) near the change points feel the changes. On the other hand, the sample mean can be severely affected by a single change in the mean.

The present study can be extended into several directions. From the theoretical point of view, it is desirable to relax the Gaussianity assumption, e.g., by extending Theorems 2.4 and 2.5 to moving averages X_t in general iid innovations. The cases of stationary weakly dependent X_t (corresponding to $d = 0$) and stationary weakly dependent $U_t = X_t - X_{t-1}$ (corresponding to $d = 1$) are of particular interest, where the distributional assumptions on X_t should be kept to minimum. The IR statistic in (1.1) allows for a number of modifications which in principle might have better asymptotic or finite sample properties. Further generalizations may involve observations in continuous and/or multidimensional time (random fields). We hope to study some of these issues in the future.

The paper is organised as follows: section 2 provides asymptotic results, section 3 studies the size, power and robustness of the IR statistic, and provides comparisons with other statistics. An application of this statistic to real data is given in section 4. The proofs of all statements in section 2 and the properties of the second increment of fractional Brownian motion are relegated in sections 5 and 6 respectively.

2 Asymptotic results

In this section, we introduce general Assumptions (A1) and (A2) which guarantee the convergence of the IR statistic to the function $\Lambda(d)$ in (1.3) (see Proposition 2.1). Neither Gaussianity nor stationarity of the observations is required by these assumptions. Write $\to_{\mathrm{D}}$ (respectively, $\to_{\mathrm{FDD}}$) for weak convergence of distributions (respectively, of finite dimensional distributions). Recall that a fractional Brownian motion (fBm) with Hurst parameter $0 < H < 1$ is a Gaussian process $B_H(\tau), \tau \in \mathbb{R}$, with zero mean and the covariance

$$\mathrm{E}B_H(\tau_1)B_H(\tau_2) = \frac{1}{2}\left(|\tau_1|^{2H} + |\tau_2|^{2H} - |\tau_1 - \tau_2|^{2H}\right). \tag{2.1}$$

<u>Assumption (A1)</u> For $-.5 < d < .5$, there exists a constant $G(d) \neq 0$ and normalizations $G_m = G_m(d) \to \infty, A_m = A_m(d)$ such that

$$G_m^{-1}\left(\sum_{t_1=1+[mT_1]}^{[m(T_1+\tau_1)]}(X_{t_1} - A_m), \sum_{t_2=1+[mT_2]}^{[m(T_2+\tau_2)]}(X_{t_2} - A_m)\right) \to_{\mathrm{FDD}} G(d)\left(B^1_{d+.5}(\tau_1), B^2_{d+.5}(\tau_2)\right) \tag{2.2}$$

as $m, T_1, T_2 - T_1 \to \infty$, where $B^1_{d+.5}, B^2_{d+.5}$ are independent copies of fBm $B_{d+.5}$ with Hurst parameter $H = d + .5 \in (0,1)$.

<u>Assumption (A2)</u> For $.5 < d < 1.5$, there exists a constant $G(d) \neq 0$ and a normalization $G_m = G_m(d) \to \infty$ such that

$$G_m^{-1}\left(X_{[m(T_1+\tau_1)]} - X_{[mT_1]}, X_{[m(T_2+\tau_2)]} - X_{[mT_2]}\right) \to_{\mathrm{FDD}} G(d)(B^1_{d-.5}(\tau_1), B^2_{d-.5}(\tau_2)) \tag{2.3}$$

as $m, T_1, T_2 - T_1 \to \infty$, where $B^1_{d-.5}, B^2_{d-.5}$ are independent copies of fBm $B_{d-.5}$ with Hurst parameter $H = d - .5 \in (0,1)$. Moreover, there exists a constant $C_2 < \infty$ such that for any $m, j \geq 1$

$$\mathrm{E}(X_{m+j} - X_j)^2 \leq C_2 G_m^2. \tag{2.4}$$

Proposition 2.1 (i) *Let Assumption (A1) be satisfied,* $-.5 < d < .5$. *Then, as* $N \to \infty$, $m \to \infty$, $m/N \to 0$

$$\mathrm{E}IR \to \Lambda(d), \tag{2.5}$$

where the function $\Lambda(d)$ *is defined in (1.3). Moreover,*

$$\mathrm{E}(IR - \Lambda(d))^2 \to 0. \tag{2.6}$$

(ii) *Let Assumption (A2) be satisfied,* $.5 < d < 1.5$. *Then, as* $N \to \infty$, $m \to \infty$, $m/N \to 0$, *relations (2.5) and (2.6) hold. The function* $\Lambda(d)$ *is defined in (1.3), with* $Z_d(0), Z_d(1)$ *as in (1.5).*

In the literature, convergence of partial sums towards a fBm has been proved for a number of linear and nonlinear (stationary and nonstationary) processes. See Davydov (1970), Taqqu (1977), Ho and Hsing (1997), Giraitis *et al.* (2000), Giraitis and Surgailis (2002), Philippe *et al.* (2006a, 2006b, 2007) and the references therein. A new feature of Assumptions (A1)/(A2) concerns the asymptotic independence of increments of partial sums separated by long interval $T = T_2 - T_1 \to \infty$ (i.e., the independence of the limiting fBm's). For Gaussian processes, Assumptions (A1)/A(2) can be easily verified; see Proposition 2.2 below. Csörgő and Mielniczuk (1995), Bružaitė and Vaičiulis (2005) discuss the validity of Assumption (A1) for Gaussian subordinated and linear processes.

Proposition 2.2 (i) *Let X_t be a stationary Gaussian process having spectral density $f(x)$ such that*

$$f(x) \ = \ L(1/|x|)|x|^{-2d}, \tag{2.7}$$

where $-.5 < d < .5$ and L is slowly varying at infinity. Then X_t satisfies Assumption (A1), with $G_m^2 = L(m)m^{2d+1}, A_m = \mathrm{E}X_0$ and $G^2(d) = K(d + .5)$, where

$$K(H) := \frac{\pi}{H\Gamma(2H)\sin(H\pi)}. \tag{2.8}$$

(ii) *Let $U_t = X_t - X_{t-1}$ be a stationary Gaussian process having spectral density $f(x)$ such that*

$$f(x) \ = \ L(1/|x|)|x|^{2-2d}, \tag{2.9}$$

where $.5 < d < 1.5$ and L is slowly varying at infinity. Then X_t satisfies Assumption (A2), with $G_m^2 = L(m)m^{2d-1}$, $G^2(d) = |K(d - .5)|$ and $K(H)$ as in (2.8).

Let us note that Assumption (A1)(respectively, (A2)) refers to "distant increments" of partial sums of the observations (respectively, of the observations themselves) on intervals of length $O(m)$ which are far away from each other and also from the origin, due to the fact that $T_1 \to \infty, T_2 - T_1 \to \infty$. Therefore (A1)/(A2) may apply also in the case when the limit of partial sums is a process with *asymptotically stationary increments* (see Philippe *et al.* (2007), Bružaitė *et al.* (2006) for the definition and examples of such processes). In particular, consider a $d-$integrated ($d > -.5$) process X_t defined as a solution of $(1 - L)^d X_t = \xi_t \mathbb{I}_{\{t\ge 1\}}$:

$$X_t \ = \ \sum_{s=1}^{t} \psi(t-s)\xi_s, \qquad \psi(j) := \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)} \quad (j \ge 0), \tag{2.10}$$

where $LX_t = X_{t-1}$ is the backward shift, $\mathbb{I}$ denotes the indicator function, $\psi(j)$ $(j \ge 0)$ are the coefficients of FARIMA$(0, d, 0)$ filter, and where $\xi_t, t \in \mathbb{Z}$ are standard iid random variables, with zero mean and variance 1. One can show (see Marinucci and Robinson (1999) and the references therein) that for any $d > -.5$

$$m^{-d-.5} \sum_{t=1}^{[m\tau]} X_t \ \to_{\mathrm{FDD}} \ \frac{1}{\Gamma(d)} \int_0^\tau (\tau - x)^d M(\mathrm{d}x), \tag{2.11}$$

where $M(\mathrm{d}x)$ is a standard Gaussian white noise (see Sec. 6). The limit process in (2.11) is called a type II fractional Brownian motion (Marinucci and Robinson, 1999) and has asymptotically stationary increments tending to increments of a (usual) fBm (Philippe *et al.*, 2007).

Proposition 2.3 *Let X_t be the moving average in (2.10), $d \in (-.5, 1.5), d \ne .5$. Then X_t satisfies (A1)/(A2).*

In the remainder of this section we assume that the time series $X_t, t = 1, \ldots, N$ is a Gaussian process. This assumption and the following assumptions on the covariance structure of X_t allows us to obtain a convergence rate of the bias $\mathrm{E}IR - \Lambda(d)$, as well as a central limit theorem for the IR statistic, when N and m increase in a suitable way. We separately discuss the cases (i) $-.5 < d < .5$ and (ii) $.5 < d < 1.5$. In the Case (i), we assume that X_t is a stationary Gaussian process, while in the Case (ii), we assume that X_t is an integrated process so that the process $U_t = X_t - X_{t-1}$ is stationary.

Theorem 2.4 (i) *Let X_t be a stationary Gaussian process having spectral density $f(x)$ such that there exist constants $c_0 > 0, \beta > 0, -.5 < d < .5$ such that*

$$f(x) \;=\; |x|^{-2d}\left(c_0 + O(|x|^{\beta})\right) \qquad (x \to 0). \tag{2.12}$$

Moreover, assume that $f(x)$ is bounded outside zero frequency, and $0 < \beta < 2d + 1$. Then

$$\mathrm{E}IR - \Lambda(d) \;=\; O(m^{-\beta}). \tag{2.13}$$

(ii) *Let $U_t = X_t - X_{t-1}$ be a zero mean stationary Gaussian process, with zero mean and spectral density $f(x)$. Assume that there exist constants $c_0 > 0, \beta > 0, .5 < d < 1.5$ such that*

$$f(x) \;=\; |x|^{2-2d}\left(c_0 + O(|x|^{\beta})\right) \qquad (x \to 0). \tag{2.14}$$

Moreover, assume that $f(x)$ is bounded outside zero frequency, and $0 < \beta < 2d - 1$. Then relation (2.13) holds.

Theorem 2.4 is proved in Section 5. Let us explain the main idea of its proof. Define

$$V_m^2 \;:=\; \mathrm{E}\Big(\sum_{t=1}^{m}(X_{t+m} - X_t)\Big)^2, \tag{2.15}$$

$$R_m \;:=\; \mathrm{E}\Big(\sum_{t,s=1}^{m}(X_{t+m} - X_t)(X_{s+2m} - X_{s+m})\Big). \tag{2.16}$$

By stationarity, in both cases (i) and (ii)

$$\mathrm{E}IR - \Lambda(d) \;=\; \mathrm{E}\left[\frac{|Y^0 + Y^1|}{|Y^0| + |Y^1|} - \frac{|Z^0 + Z^1|}{|Z^0| + |Z^1|}\right], \tag{2.17}$$

where

$$Y^0 := V_m^{-1}\sum_{t=1}^{m}(X_{t+m} - X_t), \quad Y^1 := V_m^{-1}\sum_{t=m+1}^{2m}(X_{t+m} - X_t), \qquad Z^0 := Z_d(0), \quad Z^1 := Z_d(1),$$

are Gaussian variables, with zero mean, unit variances $\mathrm{E}(Y^0)^2 = \mathrm{E}(Y^1)^2 = \mathrm{E}(Z^0)^2 = \mathrm{E}(Z^1)^2 = 1$ and the covariances

$$\mathrm{E}Y^0Y^1 \;=\; \frac{R_m}{V_m^2}, \qquad \mathrm{E}Z^0Z^1 = \rho(d), \tag{2.18}$$

respectively (the variables $Z_d(0), Z_d(1)$ and $\rho(d)$ were defined earlier in (1.5)-(1.4)). Using (2.17) and the Gaussianity, it is easy to show the bound

$$|\mathrm{E}IR - \Lambda(d)| \;\le\; C|\mathrm{E}Y^0Y^1 - \mathrm{E}Z^0Z^1| \tag{2.19}$$

where the constant C does not depend on m. As shown in the proof of Theorem 2.4, under the assumptions on the spectral density as in (2.12), (2.14), one has the following asymptotics

$$V_m^2 = c_0 m^{2d+1}\big(V(d) + O(m^{-\beta})\big), \tag{2.20}$$

$$R_m = c_0 m^{2d+1}\big(R(d) + O(m^{-\beta})\big), \tag{2.21}$$

where

$$V(d) := (4 - 4^{d+.5})K(d+.5), \tag{2.22}$$

$$R(d) := (1/2)\big(-9^{d+.5} + 4^{d+1.5} - 7\big)K(d+.5), \tag{2.23}$$

with $K(H)$ given in (2.8) above. (Note the relation $R(d)/V(d) = \rho(d)$ and the fact that (2.20)-(2.23) hold in both cases (i) and (ii).) Clearly, (2.18)-(2.23) imply (2.13).

We now turn to the central limit theorem for the IR statistic.

Theorem 2.5 (i) *Let X_t be a stationary Gaussian process whose spectral density $f(x)$ satisfies condition (2.12), for some $-.5 < d < .5, c_0 > 0, \beta > 0$. Moreover, assume that $f(x)$ is differentiable on $(0, \pi)$ and*

$$|f'(x)| \le C|x|^{-2d-1}, \tag{2.24}$$

where $C > 0$ is some constant. Then, as $N, m, N/m \to \infty$,

$$(N/m)\mathrm{var}(IR) \to \sigma^2(d), \tag{2.25}$$

and

$$(N/m)^{1/2}(IR - \mathrm{E}IR) \to_{\mathrm{D}} \mathcal{N}(0, \sigma^2(d)), \tag{2.26}$$

where $\sigma^2(d)$ is defined in (1.9).

(ii) *Let $X_t - X_{t-1} = U_t$ be a stationary Gaussian process whose spectral density $f(x)$ satisfies (2.14), for some $.5 < d < 1.25, c_0 > 0, \beta > 0$. Moreover, assume that $f(x)$ is differentiable on $(0, \pi)$ and*

$$|f'(x)| \le C|x|^{1-2d}, \tag{2.27}$$

where $C > 0$ is some constant. Then the relations (2.25) and (2.26) hold.

Let us explain the idea of the proof of the above theorem. Let

$$Y_m(j) := V_m^{-1} \sum_{t=j+1}^{j+m} (X_{t+m} - X_t), \tag{2.28}$$

where V_m is defined in (2.15). Note, for m fixed, $Y_m(j), j \in \mathbb{Z}$ is a stationary Gaussian process, with zero mean and unit variance, and

$$IR = \frac{1}{N-3m} \sum_{j=0}^{N-3m-1} \eta_m(j), \qquad \eta_m(j) := \frac{|Y_m(j) + Y_m(j+m)|}{|Y_m(j)| + |Y_m(j+m)|}. \tag{2.29}$$

The proof of (2.25) and (2.26) uses Hermite expansion of the nonlinear function $\eta_m(j)$ in Gaussian variables (2.28). It is easy to see from the definition in (2.29) that the linear terms of the Hermite expansion are zero and therefore the covariance of $\eta_m(j)$ behaves as the squared covariance of $Y_m(j)$'s, which turns to be summable for $. - 5 < d < 1.25$; see (5.35)-(5.36).

3 The power and robustness of the IR test for short memory: an empirical study

As noted in the Introduction, the IR statistic can be used to test hypotheses about unknown parameter d, e.g., the null hypothesis H_0: $d = d_0$, where $d_0 \in (-.5, 1.25), d_0 \neq .5$. A more precise meaning of the null hypothesis is that X_t satisfies Assumptions (A1)/(A2) with $d = d_0$, as well as the additional conditions guaranteeing the asymptotic behavior of the IR statistics as in Theorems 2.4 and 2.5. Obviously, the assumption of Gaussianity in these theorems is quite restrictive and the IR test needs to be further developped. Nevertheless, an empirical study of the IR statistic and its performance against other tests for testing similar hypotheses is clearly of interest. The choice of benchmark tests for IR is somewhat arbitrary and also limited by the length of the paper. In the present section, we compare the size, power and robustness of the the IR test (1.10) for short memory ($d = 0$) against the long–range dependent alternative ($d > 0$) to the V/S test, the Robinson (1994) test, the CUSUM test of Berkes *et al.* (2006) and the SB-FDF test. More complete comparison results can be found on the the following web site `http://samos.univ-paris1.fr/ppub2005.html#prepub2006`, as supplementary material of this paper.

The V/S statistic introduced in Giraitis *et al.* (2003) is defined as

$$\frac{V}{N\widehat{s}^2(q)} = \frac{N^{-1}\Big[\sum_{k=1}^{N}\Big(\sum_{j=1}^{k}(X_j - \bar{X})\Big)^2 - \frac{1}{N}\Big(\sum_{k=1}^{N}\sum_{j=1}^{k}(X_j - \bar{X})\Big)^2\Big]}{N\widehat{s}^2(q)}. \tag{3.1}$$

The numerator V is an estimator of the variance of the partial sums process, while

$$\widehat{s}^2(q) = \hat{\gamma}(0) + 2\sum_{j=1}^{q}(1 - \frac{j}{q+1})\hat{\gamma}(j), \qquad \hat{\gamma}(j) := N^{-1}\sum_{i=1}^{N-j}(X_i - \bar{X})(X_{i+j} - \bar{X}), \tag{3.2}$$

is a spectral estimator of $s^2 = \sum_{j\in\mathbb{Z}}\mathrm{cov}(X_0, X_j)$, and $q = q_N$ is the bandwidth parameter satisfying $q \to \infty, \quad q/N \to 0$. This estimator of s^2 has been used by Lo (1991) and Kwiatkowski *et al.* (1992) for respectively the R/S and the KPSS statistic. For all values of q, the V/S statistic has more power than the KPSS statistic and is less sensitive to q than the R/S statistic; see Giraitis *et al.* (2003a, 2003b) for further details. Thus, we do not consider the R/S and KPSS statistics in this comparative study.

Under general stationarity and "short memory" assumptions on X_t (see Giraitis *et al.* (2003a, Assumption S), the V/S statistic has a limit distribution $N^{-1}V/\widehat{s}^2(q) \to_{\mathrm{D}} W$, with

$$\mathrm{P}(W \le x) = 1 + 2\sum_{k=1}^{\infty}(-1)^k \mathrm{e}^{-2k^2\pi^2 x}.$$

A test for short-memory against LRD alternatives has a critical region of the form

$$\frac{V}{\widehat{s}^2(q)} > c_\alpha N, \tag{3.3}$$

c_α being the critical values of this distribution. The V/S statistic was also studied in Leipus and Viano (2003), Giraitis *et al.* (2003b), Giraitis *et al.* (2006), Aue *et al.* (2005). As it should be clear from equation (3.2), the V/S statistic strongly relies on the constancy of the mean $\bar{X}$. When working with financial data that are not homogeneous, e.g., volatility series, this assumption is too strong. Although the V/S statistic solves the issue of extreme sensitivity to q, the issue of sensitivity to changes in $\bar{X}$ remains.

The score $\hat{r}$ test developed in Robinson (1994) and Gil-Alaña and Robinson (1997) tests H_0 : $d = d_0$ against the fractional alternative $d > d_0$, for models of the form

$$\phi(L)X_t = \xi_t, \tag{3.4}$$

where $\phi(z) = (1-z)^d$ and ξ_t is a covariance stationary sequence with zero mean and parametric spectral density $f(\lambda) = (\sigma^2/2\pi)g(\lambda;\tau)$ depending on unknown parameters $\tau \in \mathbb{R}^k$ and σ^2. Let

$$\varphi(\lambda) = \mathrm{Re}\Big\{ \log\big(\phi(\mathrm{e}^{\mathrm{i}\lambda})\big)'_{d=d_0} \Big\} = \log|2\sin(\lambda/2)|\,, \quad \lambda \in [-\pi, \pi). \tag{3.5}$$

Define $\tilde{\xi}_t = (1-L)^{d_0}X_t$, $I_{\tilde{\xi}}(\lambda) = (1/2\pi N)\left|\sum_{t=1}^N \tilde{\xi}_t \mathrm{e}^{\mathrm{i}t\lambda}\right|^2$, $\lambda_j = 2\pi j/N$, $\widehat{\zeta}(\lambda) = (\partial/\partial\tau)\log g(\lambda;\widehat{\tau})$,

$$\begin{aligned} \sigma^2(\tau) &= \frac{2\pi}{N}\sum_j \frac{I_{\tilde{\xi}}(\lambda_j)}{g(\lambda_j;\tau)}, \qquad \widehat{\sigma}^2 = \sigma^2(\widehat{\tau}), \qquad \widehat{a} = -\frac{2\pi}{N}\sum_j{}' \varphi(\lambda_j)\frac{I_{\tilde{\xi}}(\lambda_j)}{g(\lambda_j;\widehat{\tau})}, \\ \widehat{A} &= \frac{2}{N}\left(\sum_j{}' |\varphi(\lambda_j)|^2 - \sum_j{}'\varphi(\lambda_j)\widehat{\zeta}(\lambda_j)'\Big\{\sum_j{}'\widehat{\zeta}(\lambda_j)\widehat{\zeta}(\lambda_j)'\Big\}^{-1}\sum_j{}'\widehat{\zeta}(\lambda_j)\varphi(\lambda_j)\right), \end{aligned}$$

where the sum $\sum_j$ (respectively, $\sum_j'$) is taken over all $\lambda_j \in (-\pi, \pi)$ (respectively, over all $\lambda_j \in (-\pi,\pi), \lambda_j \neq 0$), and $\widehat{\tau}$ is a consistent estimator of τ. Note $\tilde{\xi}_t = X_t$ for testing the short memory hypothesis $d = 0$. The score $\hat{r}$ statistic is defined as

$$\hat{r} = \frac{N^{1/2}}{\widehat{\sigma}^2}\widehat{A}^{-1/2}\widehat{a}. \tag{3.6}$$

Under H_0: $d = d_0$ and some additional assumptions on ξ_t in (3.4), see Robinson (1994), $\hat{r} \to_{\mathrm{D}} \mathcal{N}(0,1)$, and a critical region is given by

$$\hat{r} > z_\alpha, \tag{3.7}$$

where z_α is the standard normal quantile.

In our study, $d_0 = 0$ and ξ_t is a weakly dependent AR(k) process, i.e., $g(\lambda;\tau) = \left|1 - \sum_{j=1}^k \tau_j \mathrm{e}^{\mathrm{i}j\lambda}\right|^{-2}$, $\tau = (\tau_1, \dots, \tau_k)$, with $k = 1$ and $k = 3$. Results for AR(k) for other values of k and for the Bloomfield process can be found at `http://samos.univ-paris1.fr/ppub2005.html#prepub2006`, as supplementary material of this paper.

The M_N statistic of Berkes *et al.* (2006) is based on a change–point estimator and two CUSUM statistics applied to the sub-samples before and after the detected change-point. The last paper also extends the M_N test to the case of multiple change-points, using a binary segmentation procedure.

3.1 Stochastic and deterministic trends

The empirical sizes (probabilities of Type I error) of the tests (1.10), (3.3) and (3.7) are studied for short memory observations X_t of the form

$$\begin{aligned} X_t &= Y_t + f_{t,N}, \quad t = 1, \dots, N, && (3.8)\\ Y_t &= aY_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{iid } \mathcal{N}(0,1), && (3.9)\\ f_{t,N} &= \sum_{i=1}^t b_{i,N}c_i, \quad c_i \sim \text{iid } \mathcal{N}(0, b^2), \quad b_{i,N} \text{ iid Bernoulli}, && (3.10) \end{aligned}$$

i.e., X_t is the sum of an AR(1) process Y_t and a stochastic trend $f_{t,N}$ with $\mathrm{P}(b_{i,N} = 1) = \pi_N = 1 - \mathrm{P}(b_{i,N} = 0)$. The three processes $\{\varepsilon_i, 1 \le i \le N\}, \{b_{i,N}, 1 \le i \le N\}$ and $\{c_i, 1 \le i \le N\}$ are mutually independent. Model (3.8), called a mixture model in the literature, can generate the so-called "spurious long–memory" effect; see Diebold and Inoue (2001), Granger and Hyung (2004). The V/S test in the presence of stochastic trend (3.10) was studied in Leipus and Viano (2003), Aue *et al.* (2005). For $a = b = 0$ this is an iid process, while for $b = 0$, this is a weakly dependant process, that tends to a process with a unit root as a tends to one.

Table 1 illustrates empirical sizes of the IR, the V/S and the score tests at the level $\alpha = 5\%$ under the model (3.8) for $N = 1000$, and selected values of parameters a, b; the probability of "trend jump" is $\pi_N = 5/N = 0.005$ in all samples. The choice of q in the range $N^{1/3}$ to $N^{1/2}$, as a reasonable compromise between size and power distortions for the V/S test, was suggested in Giraitis *et al.* (2003a, 2003b). Our simulations suggest a similar choice of $m = O(N^{1/3})$ to $O(N^{1/2})$ for the IR statistic.

Table 1: Frequency of rejection of the null hypothesis of short memory for sequences of AR(1) + mixture trend processes, having on average 5 $\mathcal{N}(0, b^2)$-distributed jumps in a sample, ($\pi_N = 5/1000$). Test size 5%. $N = 1000$ (based on 10000 replications)

		V/S		IR		$\hat{r},\ \xi_t \sim \mathrm{AR}(k)$	
a	b	$q = 10$	$q = 30$	$m = 10$	$m = 30$	$k = 1$	$k = 3$
0.0	0.0	0.0444	0.0363	0.0515	0.0465	0.0795	0.0691
0.0	0.2	0.6709	0.6062	0.0566	0.0914	0.6838	0.6987
0.0	1.0	0.9581	0.9103	0.1833	0.4669	0.9516	0.6868
0.2	0.0	0.0531	0.0387	0.0845	0.0560	0.0777	0.0686
0.2	0.2	0.5947	0.5286	0.0874	0.0848	0.5974	0.6015
0.2	1.0	0.9484	0.8969	0.2026	0.4120	0.9288	0.7514
0.4	0.0	0.0648	0.0417	0.1351	0.0679	0.0719	0.0675
0.4	0.2	0.4885	0.4125	0.1438	0.0834	0.4706	0.4660
0.4	1.0	0.9292	0.8724	0.2410	0.3394	0.8945	0.7981
0.6	0.0	0.0867	0.0472	0.2802	0.0885	0.0685	0.0667
0.6	0.2	0.3636	0.2679	0.2854	0.0946	0.2873	0.2841
0.6	1.0	0.8893	0.8146	0.3634	0.2689	0.8076	0.8070
0.8	0.0	0.1836	0.0680	0.8023	0.1483	0.0940	0.0555
0.8	0.2	0.2918	0.1422	0.7960	0.1492	0.1104	0.1168
0.8	1.0	0.7929	0.6545	0.8163	0.2231	0.1916	0.5735
1.0	0.0	0.9990	0.9563	1.0000	0.9926	0.1307	0.1501

The results in Table 1 indicate that in the absence of a trend ($b = 0$), the V/S test has a better size than the IR test, mainly for the highest values of the parameter a and the smallest windows $m = q = 10$. The size of the V/S test rapidly deteriorates as b increases, while the IR test shows a much better robustness to trends for the largest values of the bandwidth parameters m and q. Note that the bandwidths m and q are not directly comparable.

Note that for the highest values of a, the $\hat{r}$ test with the AR(k) specifications has a better size than both the IR and V/S tests. However, for lower values of a ($a < 0.8$) and in the presence of a trend, the IR test has a better size than the two other tests. The last line in Table 1 shows that the score test with the given AR(k) specifications has a low power against the unit root, in contrast to the V/S and R/S tests.

We also consider the case of deterministic trends, with a possible break at time $t = [\delta N]$

$$X_t \;=\; X_0 + c_0 t + c_1 \mathbb{I}_{\{t>[\delta N]\}}(t - [\delta N]) + \varepsilon_t, \quad \delta \in (0,1), \quad \varepsilon_t \sim \text{iid}\ \mathcal{N}(0,1). \tag{3.11}$$

We set $\delta = 0.5$, i.e., the break in the trend occurs in the middle of the sample.

Table 2: Frequency of rejection of the null hypothesis of short memory for sequences of iid process with a deterministic linear trend and a possible break. Test size 5%. $N = 1000$ (based on 10000 replications)

		V/S		IR		$\hat{r},\ \xi_t \sim$ AR(k)	
c_0	c_1	$q = 10$	$q = 30$	$m = 10$	$m = 30$	$k = 1$	$k = 3$
0.001	0.0	0.0444	0.0363	0.0548	0.0654	1.0000	1.0000
0.001	0.002	1.0000	1.0000	0.0581	0.2043	1.0000	1.0000

From Table 2 we may conclude that the IR test is far more robust to deterministic trends than both the V/S and the score $\hat{r}$ tests.

Dolado *et al.* (2005) studied the power of their test only for a process similar to the one defined by equation (3.11), so that we study the performance of their test for that process. Note that the null hypothesis of their test is that the process is I(d), and the alternative hypothesis is that the process is I(0) with a single break, so that it is not directly comparable with the IR, V/S and $\hat{r}$ score tests.

Table 3: Frequency of rejection of the null hypothesis of I(d) for sequences of iid process with a deterministic linear trend and a break, i.e., $c_0 = 0.001$ $c_1 = 0.002$. Test size 5%. $N = 1000$ (based on 10000 replications)

d	SB-FDF Model B	SB-FDF Model C
0.40	1.0000	1.0000
0.30	1.0000	1.0000
0.20	1.0000	1.0000
0.10	0.8832	0.8749

Model B corresponds to the "changing growth" model,

$$X_t = \mu + \nu_1 t + (\nu_2 - \nu_1) DN_t^\star + \varepsilon_t, \quad DN_t^\star = \mathbb{I}_{(t>[\delta N])}(t - [\delta N]),$$

i.e., under the alternative hypothesis, the slope of the trend changes without change in the level, while Model C corresponds to "the changing growth with crash" model,

$$X_t = \mu_1 + \nu_1 t + (\nu_2 - \nu_1) DN_t + (\mu_2 - \mu_1) DU_t + \varepsilon_t, \quad DN_t = t\mathbb{I}_{(t>[\delta N])}, \quad DU_t = \mathbb{I}_{(t>[\delta N])},$$

i.e., under the alternative hypothesis there is a change in both the level and slope of the trend; see Perron (1989) for further details. This test always rejects the null hypothesis of I(d) process for $d = 0.40, 0.30, 0.20$, and nearly 90% of the times the null hypothesis $d = 0.10$.

Teyssière and Abry (2005) studied the performance of the wavelet estimator on a more general process: an additive combination of a fractionally integrated process and a broken polynomial trend. The wavelet estimator was not fooled by the overimposition of the broken polynomial trend, and estimation biases were of the same order as the ones for the process without trend and break, provided that the number of vanishing moments of the mother wavelet is large enough.

3.2 Robustness to memory breaks and heavy tails

Consider the so-called "FARIMA $(0, d, 0)$ with memory breaks" model, defined by

$$X_t = \varepsilon_t + \sum_{j=1}^{\infty} \varepsilon_{t-j}\psi(j) \prod_{i=1}^{j}(1 - b_{t-i,N}), \quad \varepsilon_t \sim \text{iid } \mathcal{N}(0, 1), \tag{3.12}$$

where $\psi(j)$ are the FARIMA$(0, d, 0)$ coefficients, see (2.10), and $b_{t,N}$ are iid Bernoulli as in (3.10). Conditionally on $b_{i,N}, i \in \mathbb{Z}$, the process in (3.12) is nonstationary and satisfies the FARIMA$(0, d, 0)$ equation $(1-L)^d X_t = \varepsilon_t$ on intervals $t_k \le t < t_{k+1}$ between consecutive moments t_k with $b_{t_k,N} = 1$, with zero "initial condition" $X_u = 0, u < t_k$; moreover, $X_t, t \ge t_k$ are conditionally independent of $\varepsilon_u, u < t_k$. The moments t_k can be thus identified with "memory breaks". If the probability $\pi_N = \mathrm{P}(b_{0,N} = 1) = c/N$ is small, there are few "memory breaks" in the interval $[1, N]$ and their number has approximate Poisson distribution with mean c. Note also that unconditionally the process X_t in (3.12) is (strictly) stationary and exists for any $d \in \mathbb{R}$, unless $\mathrm{P}(b_{0,N} = 0) = 1$. In the last case, (3.12) is nothing but the usual stationary FARIMA$(0, d, 0)$ process $(d < 0.5)$.

From Table 5 one may infer that the V/S test has a slightly better power than the IR test under the "pure FARIMA" model with Gaussian $(\alpha = 2)$ innovations. However, the advantage of the V/S test disappears with the presence of memory breaks, see Table 4, in which case the IR test seems to have somewhat better power against fractional alternatives.

The excellent power of the $\hat{r}$ test vs. the two other tests in Tables 4 and 5 can be largely explained by the fact that the null hypothesis of this test is very "narrow" and the test is simply unable to fit FARIMA into the given AR(1) or AR(3) specifications. The power of the $\hat{r}$ test consistently drops with increase of k for values of d; see `http://samos.univ-paris1.fr/ppub2005.html#prepub2006`.

Table 4: Frequency of rejection of the null hypothesis of short memory for sequences of FARIMA$(0, d, 0)$ with memory breaks processes, with the average distance 333.3 between breaks $(\pi_N = 15/5000)$. Test size 5%. $N = 5000$ (based on 10000 replications)

	V/S		IR		$\hat{r},\ \xi_t \sim$ AR(k)	
d	$q = 10$	$q = 30$	$m = 10$	$m = 30$	$k = 1$	$k = 3$
0.40	0.9775	0.8329	1.0000	0.9753	1.0000	0.9998
0.30	0.8946	0.6692	0.9973	0.8602	1.0000	1.0000
0.20	0.6564	0.4363	0.9177	0.5826	1.0000	1.0000
0.10	0.3017	0.2069	0.4678	0.2473	1.0000	0.9458

Table 5 is motivated by applications to financial econometrics, where it is argued that asset returns, or their squares, may follow a heavy-tailed (e.g., $\alpha-$stable) distribution. From this table we can see that for the largest values of m the IR statistic is more robust than the V/S statistic for α–stable

Table 5: Frequency of rejection of the null hypothesis of short memory for sequences of FARIMA$(0, d, 0)$ processes with Gaussian ($\alpha = 2$) and symmetric α–stable innovations. Test size 5%. $N = 1000$ (based on 10000 replications)

		V/S		IR		$\hat{r},\ \xi_t \sim$ AR(k)	
α	d	$q = 10$	$q = 30$	$m = 10$	$m = 30$	$k = 1$	$k = 3$
2.0	0.30	0.7182	0.4486	0.6752	0.3733	0.9999	0.9238
2.0	0.20	0.4816	0.2809	0.4170	0.2327	0.9956	0.8361
2.0	0.10	0.2209	0.1300	0.1864	0.1199	0.8208	0.4942
2.0	0.00	0.0432	0.0358	0.0514	0.0489	0.0814	0.0698
1.5	0.30	0.7538	0.4851	0.8906	0.5416	0.9979	0.9433
1.5	0.20	0.5228	0.2763	0.6441	0.3362	0.9935	0.8773
1.5	0.10	0.2025	0.1101	0.2773	0.1588	0.8620	0.5232
1.5	0.00	0.0303	0.0245	0.0648	0.0487	0.0508	0.0490
1.25	0.30	0.7920	0.5153	0.9656	0.6851	0.9929	0.9450
1.25	0.20	0.5660	0.2861	0.8093	0.4534	0.9910	0.9011
1.25	0.10	0.1984	0.1016	0.3966	0.2031	0.8971	0.5542
1.25	0.00	0.0224	0.0177	0.0762	0.0544	0.0387	0.0391

innovations: unlike the V/S statistic, the IR statistic has still the correct size and its power is not much affected. Abry *et al.* (2003) observed that the wavelet estimator of the memory parameter is also robust to heavy–tailed distributions.

The above mentioned robustness of the IR test can be explained by the fact that the limit of the IR statistic is quite insensitive to heavy tails and asymmetry of the DGP. In the case of iid X_t in the domain of attraction of a stable law with index $0 < \alpha < 2$ and skewness parameter $\beta \in [-1, 1]$, the IR statistic converges to the expectation $\Lambda(\alpha, \beta) = \mathrm{E}[|\Delta^2 Z_{\alpha,\beta}(0) + \Delta^2 Z_{\alpha,\beta}(1)|/(|\Delta^2 Z_{\alpha,\beta}(0)| + |\Delta^2 Z_{\alpha,\beta}(1)|)]$ where $Z_{\alpha,\beta}(\tau)$ is a corresponding Lévy process with independent and homogeneous increments. Monte-Carlo simulations with large $N = 10^7$ show that the "bias" $\Lambda(\alpha, \beta) - \Lambda(0)$ in the IR test (1.10) due to a change of the limiting value of the IR statistic is quite small: $\Lambda(1.5, 0) - \Lambda(0) \approx 0.5905 - 0.5881 = 0.0027$, $\Lambda(1.5, 1) - \Lambda(0) \approx 0.5914 - 0.5881 = 0.0033$, and does not change much the outcome of the test.

3.3 Robustness to single change–point in the mean of an iid process

We consider the following iid process

$$X_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, 1). \tag{3.13}$$

We consider two cases for μ_t:

- DGP A: $\mu_t = 0$ for $t = 1, \dots, N$,
- DGP B: $\mu_t = 0$ for $t = 1, \dots, [N/2]$, $\mu_t = 1/4$ for $t = [N/2] + 1, \dots, N$.

From Table 6 we infer that, unlike the V/S and $\widehat{r}$ statistics, the IR statistic is not much affected by changes in the mean.

Table 6: Frequency of rejection of the null hypothesis of short memory for sequences of iid $\mathcal{N}(0,1)$ processes. Test size 5%. $N = 1000$ (based on 10000 replications)

	V/S		IR		$\hat{r},\ \xi_t \sim \mathrm{AR}(k)$		M_N	
DGP	$q = 10$	$q = 30$	$m = 10$	$m = 30$	$k = 1$	$k = 3$	$q = 10$	$q = 45$
DGP A	0.0432	0.0358	0.0514	0.0489	0.0796	0.0705	0.0000	0.0000
DGP B	0.8780	0.8254	0.0562	0.0585	0.7253	0.7830	0.0022	0.0001

The M_N statistic of Berkes *et al.* (2006) strongly rejects the hypothesis $d > 0$ in both cases DGP A and DGP B, although with this small change in the mean (1/4 in the case of DGP B), it rarely detects the change itself.

3.4 Squares of nonhomogeneous GARCH(1,1) processes

We consider several GARCH(1,1) volatility processes defined as

$$X_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \theta X_{t-1}^2, \tag{3.14}$$

with two possible distributions for ε_t: $\varepsilon_t \sim \mathcal{N}(0,1)$ and $\varepsilon_t \sim t(7)$; the latter choice is motivated by empirical evidence for financial returns; see Bollerslev (1987) and Teräsvirta (1996).

For one of these processes, the parameters (ω, β, θ) are constant so that the unconditional variance of the process $\sigma^2 = \omega/(1-\theta-\beta)$ is constant as well. For the other processes, the parameters (ω, β, θ) change at time $t = [N/2]$ with different magnitudes for the change in the unconditional variance of the process. Mikosch and Stărică (1999, 2003) have shown that nonstationarity in GARCH processes generate spurious long–range dependence in the power transformation of level series, the intensity of this spurious long–range dependence is positively correlated with the magnitudes of the changes in the unconditional variance.

- DGP 0: GARCH(1,1):

$$\omega = 0.1, \quad \beta = 0.3, \quad \theta = 0.3.$$

- DGP 1: GARCH(1,1) process with abrupt change–point in the middle of the sample (large changes in the parameters, large change in the unconditional variance):

$$\omega = 0.1, \quad \beta = 0.3, \quad \theta = 0.3 \quad \text{for } t = 1, \ldots, [\tfrac{N}{2}] \quad (\sigma^2 = 0.25), \tag{3.15}$$

$$\omega = 0.15, \quad \beta = 0.65, \quad \theta = 0.25 \quad \text{for } t = [\tfrac{N}{2}] + 1, \ldots, N \quad (\sigma^2 = 1.5). \tag{3.16}$$

- DGP 2: GARCH(1,1) process with abrupt change–point in the middle of the sample (large changes in the parameters, small change in the unconditional variance):

$$\omega = 0.1, \quad \beta = 0.3, \quad \theta = 0.3 \quad \text{for } t = 1, \ldots, [\tfrac{N}{2}] \quad (\sigma^2 = 0.25), \tag{3.17}$$

$$\omega = 0.125, \quad \beta = 0.6, \quad \theta = 0.1 \quad \text{for } t = [\tfrac{N}{2}] + 1, \ldots, N \quad (\sigma^2 = 0.4667). \tag{3.18}$$

- DGP 3: GARCH(1,1) process with change–point in the middle of the sample, such that the unconditional variance $\omega/(1-\theta-\beta)$ remains unchanged ($\sigma^2 = 0.25$)

$$\omega = 0.1, \quad \beta = 0.3, \quad \theta = 0.3 \quad \text{for } t = 1, \ldots, [\tfrac{N}{2}], \tag{3.19}$$

$$\omega = 0.15, \quad \beta = 0.25, \quad \theta = 0.15 \quad \text{for } t = [\tfrac{N}{2}] + 1, \ldots, N.$$

- DGP 4: Smooth transition GARCH(1,1) process,

$$\sigma_t^2 = \omega + \omega^* F\left(t, [\tfrac{N}{2}]\right) + \left(\beta + \beta^* F\left(t, [\tfrac{N}{2}]\right)\right) \sigma_{t-1}^2 + \left(\theta + \theta^* F\left(t, [\tfrac{N}{2}]\right)\right) X_{t-1}^2, \tag{3.20}$$

  with

$$\begin{aligned} \omega &= 0.1, \quad \beta = 0.3, \quad \theta = 0.3, \\ \omega^* &= 0.05, \quad \beta^* = 0.35, \quad \theta^* = -0.05, \quad \gamma = 0.05, \end{aligned}$$

  where $F(t,k) = (1 + \exp(-\gamma(t-k)))^{-1}$, γ is a strictly positive parameter controlling the smoothness of the transition. If γ is large, DGP 4 reduces to DGP 1. We choose here a small value for γ, i.e., the transition between the two processes is smooth.

- DGP 5: The parameters of this DGP are similar to DGP 2. However, there are two change–points, at times $[\frac{N}{3}]$ and $[\frac{2N}{3}]$, i.e.,

$$\begin{aligned} \omega &= 0.1, \quad \beta = 0.3, \quad \theta = 0.3 \text{ for } t = 1, \ldots, [\tfrac{N}{3}] \text{ and } t = [\tfrac{2N}{3}] + 1, \ldots, N \quad (\sigma^2 = 0.25), \\ \omega &= 0.125, \quad \beta = 0.6, \quad \theta = 0.1 \text{ for } t = [\tfrac{N}{3}] + 1, \ldots, [\tfrac{2N}{3}] \quad (\sigma^2 = 0.4667). \end{aligned}$$

The behavior of the V/S statistic for the sequences of absolute values $|X_t|$ for the DGP 0, DGP 1 and DGP 4 has been studied in Teyssière (2003). For DGP 0, the sum of the parameters $\beta + \theta = 0.6$, which differs from what is observed with real data. We check whether this choice does not affect the results of the Monte Carlo experiment by choosing $\beta = 0.75$ and $\theta = 0.07$ from empirical estimation results on homogeneous samples of the S&P 500 index by Mikosch and Stărică (2004). The empirical size for the IR statistic is equal to 0.2715 and 0.0990 for $m = 10$ and $m = 30$ respectively, while the empirical size for the V/S statistic is equal to 0.0971 and 0.0494 for for $q = 10$ and $q = 30$ respectively, which are close to the results reported in Table 7.

The GARCH processes satisfy Assumption 2.1 by Berkes *et al.* (2006). Note that DGP 5 contains two change–points so that we use their testing procedure in the case of at most two change–points. The bandwidth parameter q in this statistic is analogous to the V/S case; the choice $q = [15 \log_{10} N] = 45$ is suggested in Berkes *et al.* (2006).

Table 7: Frequency of rejection of the null hypothesis of short memory for sequences of squares X_t^2 of GARCH(1,1) processes with $\mathcal{N}(0,1)$ innovations. Test size 5%. $N = 1000$ (based on 10000 replications)

	V/S		IR		$\hat{r},\ \xi_t \sim \text{AR}(k)$		M_N	
DGP	$q = 10$	$q = 30$	$m = 10$	$m = 30$	$k = 1$	$k = 3$	$q = 10$	$q = 45$
DGP 0	0.0648	0.0379	0.2394	0.0910	0.5621	0.0871	0.0006	0.0000
DGP 1	0.9958	0.9468	0.6153	0.2548	0.9950	0.8714	0.3247	0.0899
DGP 2	0.8507	0.7764	0.2239	0.1090	0.9762	0.8052	0.0119	0.0043
DGP 3	0.0690	0.0465	0.1716	0.0789	0.5046	0.0925	0.0010	0.0000
DGP 4	0.9962	0.9584	0.5753	0.2488	0.9951	0.8801	0.3904	0.1205
DGP 5	0.7844	0.6899	0.2458	0.1390	0.9566	0.7360	0.0009	0.0000

Table 8: Frequency of rejection of the null hypothesis of short memory for sequences of squares X_t^2 of GARCH(1,1) processes with $t(7)$ innovations. Test size 5%. $N = 1000$ (based on 10000 replications)

	V/S		IR		$\widehat{r},\ \xi_t \sim$ AR(k)		M_N	
DGP	$q = 10$	$q = 30$	$m = 10$	$m = 30$	$k = 1$	$k = 3$	$q = 10$	$q = 45$
DGP 0	0.0580	0.0339	0.2435	0.0951	0.5414	0.0906	0.0005	0.0002
DGP 1	0.9707	0.8583	0.5801	0.2369	0.9940	0.8816	0.1591	0.0375
DGP 2	0.6686	0.5730	0.2351	0.1099	0.8980	0.6170	0.0082	0.0019
DGP 3	0.0613	0.0426	0.1852	0.0827	0.4528	0.0922	0.0004	0.0000
DGP 4	0.9729	0.8758	0.5381	0.2251	0.9939	0.8859	0.1735	0.0433
DGP 5	0.5954	0.4884	0.2505	0.1322	0.8734	0.5508	0.0000	0.0000

From Tables 7 and 8 we see that, unlike the V/S statistic, the IR statistic is not much affected by nonstationarities of the GARCH processes. This is of real interest when analyzing the long–memory properties of the squares of asset prices returns, as the empirical finding of the presence of long–range dependence in the squares of financial returns might be the consequence of both nonstationarity in the data and the use of statistical tools not robust to these nonstationarities; see Mikosch and Stărică (2003). The test $\hat{r}$ rejects the null hypothesis of an $I(0)$ process when the unconditional variance of the process is not constant, i.e., for all DGP except DGP 0 and DGP 3. The statistic M_N, designed with the purpose to discriminate between change-points and long memory, performs remarkably well in this context.

Teyssière and Abry (2005) carried a wavelet analysis on the squares of DGP 0, DGP 1 and DGP 2, and multiple change–points GARCH processes, and observed that unlike the local Whittle and log periodogram spectral estimators, the wavelet estimator of the memory parameter is not fooled by the nonstationarities, and does not detect long–range dependence in the squared series.

4 Application to financial times series

The discussion below is similar to the so-called "R/S analysis", which consists in analyzing the long–memory properties of financial time series using the R/S statistic. As it has been shown in Giraitis *et al.* (2003a,b), the V/S statistic is more of interest as it is less sensitive to the choice of the bandwidth parameter q so that the conclusions on the presence of long–range dependence reached by the investigator do not depend too much on the choice of the bandwidth parameter. As for the simulation study presented above, we will compare the results of the V/S and IR analysis, by using their P–values, i.e., the observed size, instead of the standard α%–size tests.

We first consider three series of daily returns $X_{1,t}, X_{2,t}, X_{3,t}$, where $X_{i,t} = 100 \times \log(P_{i,t}/P_{i,t-1})$, where $P_{i,t}$ are shares on Bank of America (BoA), Oracle, and SAP, observed between April 1999 and April 2002, $N = 752$. For these series, see Table 9, while both the V/S statistic and the score statistic $\hat{r}$ detect long–range dependence in the series of squared returns, the results of the IR statistic lead us to the opposite conclusion: the null hypothesis $d = 0$ is accepted.

Table 9: V/S, IR and score $\hat{r}$ statistics for the series of squared returns

	V/S			IR			$\hat{r},\ \xi_t \sim$ AR(k)		
Series	q	V/S	P–values	m	IR	P–values	k	$\hat{r}$	P–values
BoA	10	0.4662	0.0002	10	0.6433	0.0121	0	6.8593	3.4585e-12
	20	0.3524	0.0019	20	0.6291	0.1229	1	6.5906	2.1905e-11
	30	0.2919	0.0063	30	0.6059	0.3441	3	4.5082	3.2685e-06
	$[N^{1/2}]$	0.3051	0.0048	$[N^{1/2}]$	0.6029	0.3612	5	5.5868	1.1564e-08
	$[N^{1/3}]$	0.4830	0.0001	$[N^{1/3}]$	0.6525	0.0027			
Oracle	10	0.2931	0.0061	10	0.6251	0.0652	0	4.7063	1.2614e-06
	20	0.2327	0.0202	20	0.6843	0.0033	1	5.5311	1.5909e-08
	30	0.1979	0.0402	30	0.6270	0.1895	3	3.9880	3.3320e-05
	$[N^{1/2}]$	0.2072	0.0335	$[N^{1/2}]$	0.6308	0.1527	5	3.8839	5.1391e-05
	$[N^{1/3}]$	0.3008	0.0053	$[N^{1/3}]$	0.6027	0.2639			
SAP	10	0.2842	0.0073	10	0.5957	0.3774	0	4.8614	5.269e-07
	20	0.2302	0.0212	20	0.6517	0.0350	1	6.7675	6.549e-12
	30	0.2007	0.0380	30	0.5863	0.5163	3	3.5581	0.0002
	$[N^{1/2}]$	0.2058	0.0344	$[N^{1/2}]$	0.5866	0.5142	5	2.7533	0.0029
	$[N^{1/3}]$	0.2927	0.0062	$[N^{1/3}]$	0.6228	0.0660			

For the BoA series, the test by Berkes *et al.* (2006) detects one change point for $q = 5, 10, 15$, and neither change–point nor long–range dependence for $q = [15 \log_{10} N] = 43$. For both the Oracle and SAP series, this test does not detect neither long–range dependence nor change–points for all values of q.

Consider now a series of financial returns at higher frequency, i.e., 30 minutes spaced returns on US dollar/British Pound Foreign Exchange (FX) rate, in ϑ–time (the daily seasonal components have been removed; see Dacorogna *et al.*, 1993, for the definition of ϑ–time) observed in 1996, i.e., $N = 17520$.

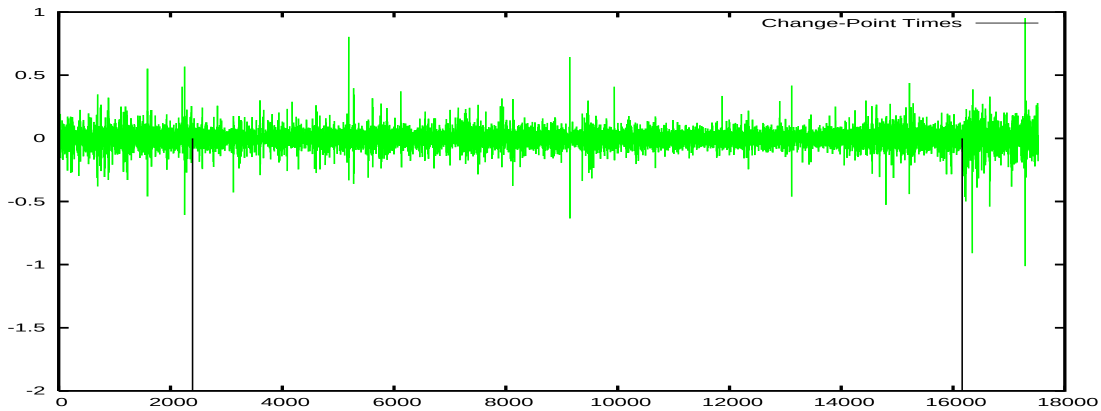

Figure 3 : The series of returns on US dollar/British pound FX rate with the two estimated change–points in variance (using the adaptive method) at times $t = 2394$ and $t = 16164$ represented by the two vertical dark lines

The plot of this series, see Figure 3, shows that this series displays intermittency, and two significant changes in variance: we use a Gaussian penalized contrast function, and estimate the number of intervals with an adaptive method; see Lavielle and Ludeña (2005), Lavielle and Teyssière (2005, 2006) for further details on this method.

We estimate both statistics on the whole sample, for a grid of bandwidths between $[N^{1/3}]$ and $[N^{1/2}]$, i.e., 26, 40, 60, 80, 100, 132, see Table 10. While the V/S statistic detects long–range dependence in the series of squared returns, with very low P–values, the IR statistic yields mixed results, as for $m = 40, 100, 132$ the null hypothesis of no long–range dependence is accepted. For all values of the bandwidth parameters, the P–values of the IR statistic are far greater than the ones of the V/S statistic. We obtain here a more nuanced view on the presence of long-range dependence in volatility: there might be long-memory in squared returns, but with a lower intensity than the one that can be inferred from the results of the V/S statistic. This result is consistent with the wavelet analysis of long–range dependence by Teyssière and Abry (2005), who observed that long–range dependence is present in this series of squared returns, but with a far lower intensity, i.e., $\hat{d}_W = 0.0491$, than the one obtained with the local Whittle and local log-periodogram spectral estimators.

Table 10: V/S, IR and score $\hat{r}$ statistics statistics for the series of squared returns on 30-minutes spaced GBP-USD FX rate

V/S			IR			$\hat{r},\ \xi_t \sim \mathrm{AR}(k)$		
q	V/S	P–values	m	IR	P–values	k	$\hat{r}$	P–values
$[N^{1/3}]$	1.2072	8.9425e-11	$[N^{1/3}]$	0.6100	0.0030	0	50.3717	0.0000
40	1.0868	9.6421e-10	40	0.5918	0.3540	1	6.9013	2.5765e-12
60	0.9798	7.9703e-09	60	0.6127	0.0222	3	9.6007	3.9711e-22
80	0.8898	4.7048e-08	80	0.6179	0.0175	5	8.7394	1.1716e-18
100	0.8148	2.0709e-07	100	0.6106	0.0778			
$[N^{1/2}]$	0.7232	1.2618e-06	$[N^{1/2}]$	0.5992	0.2721			

The score test $\hat{r}$ rejects always the null hypothesis of long-range dependence, which is not surprising since this test is not robust to the presence of changes in the unconditional variance of the process.

The CUSUM test by Berkes *et al.* (2006) detects a change point for $q = 5$, and does not reject the null hypothesis of weak dependence, for $q = 10, 15$, and $q = [15 \log_{10} N] = 63$.

5 Proofs

Proof of Proposition 2.1 (i) Let $S_n := \sum_{t=1}^{n}(X_t - A_m), n = 1, 2, \ldots$, then $\sum_{t=j+1}^{m+j}(X_{t+m} - X_t) = \Delta^2 S_{[mT]}$, where $T = j/m$ and Δ^2 is the 2nd difference operator defined in section 1. Relation (2.2) can be rewritten as

$$G_m^{-1}\left(S_{[m(T_1+\tau_1)]} - S_{[mT_1]}, S_{[m(T_2+\tau_2)]} - S_{[mT_2]}\right) \to_{\mathrm{FDD}} G(d)\left(B^1_{d+.5}(\tau_1), B^2_{d+.5}(\tau_2)\right) \tag{5.1}$$

as $m, T_1, T_2 - T_1 \to \infty$. In particular, (5.1) implies

$$\begin{aligned} G_m^{-1} \sum_{t=j+1}^{m+j} (X_{t+m} - X_t) &= G_m^{-1} \Delta^2 S_{[mT]} \qquad (T = j/m) \\ &= G_m^{-1}(S_{[m(T+2)]} - S_{[mT]}) - 2D_m^{-1}(S_{[m(T+1)]} - S_{[mT]}) \\ &\to_{\mathrm{D}} G(d)\,(B_{d+.5}(2) - 2B_{d+.5}(1)) \;=\; G(d)\Delta^2 B_{d+.5}(0) \end{aligned}$$

as $m \to \infty, j/m = T \to \infty$. In a similar way,

$$\begin{aligned} G_m^{-1}\Big(\sum_{t=j+1}^{m+j} (X_{t+m} - X_t), \sum_{t=j+m+1}^{2m+j} (X_{t+m} - X_t)\Big) &= G_m^{-1}\left(\Delta^2 S_{[mT]}, \Delta^2 S_{[m(T+1)]}\right) \\ &\to_{\mathrm{D}} G(d)\left(\Delta^2 B_{d+.5}(0), \Delta^2 B_{d+.5}(1)\right). \end{aligned} \tag{5.2}$$

Therefore, as $m \to \infty, j/m \to \infty$, so

$$\begin{aligned} \eta_m(j) &:= \frac{\left|G_m^{-1} \sum_{t=j+1}^{j+m}(X_{t+m} - X_t) \;+\; G_m^{-1} \sum_{t=j+m+1}^{j+2m}(X_{t+m} - X_t)\right|}{\left|G_m^{-1} \sum_{t=j+1}^{j+m}(X_{t+m} - X_t)\right| + \left|G_m^{-1} \sum_{t=j+m+1}^{j+2m}(X_{t+m} - X_t)\right|} \\ &\to_{\mathrm{D}} \frac{\left|\Delta^2 B_{d+.5}(0) + \Delta^2 B_{d+.5}(1)\right|}{\left|\Delta^2 B_{d+.5}(0)\right| + \left|\Delta^2 B_{d+.5}(1)\right|} \end{aligned} \tag{5.3}$$

and hence

$$\mathrm{E}\eta_m(j) \to \Lambda(d) \qquad (m \to \infty,\ j/m \to \infty), \tag{5.4}$$

by definition of $\Lambda(d)$ in (1.3). Relation (2.5) now easily follows by the dominated convergence theorem, as $0 \le \mathrm{E}\eta_m(j) \le 1$ and

$$IR \;=\; \frac{1}{N - 3m} \sum_{j=0}^{N-3m-1} \eta_m(j). \tag{5.5}$$

Consider (2.6). With (2.5) in mind, it suffices to show

$$\mathrm{var}(IR) \to 0. \tag{5.6}$$

We have by (5.5)

$$\mathrm{var}(IR) \;=\; \frac{1}{(N - 3m)^2} \sum_{j_1, j_2=0}^{N-3m-1} \mathrm{cov}\,(\eta_{j_1}, \eta_{j_2}). \tag{5.7}$$

It suffices to show that

$$\mathrm{E}\eta_m(j_1)\eta_m(j_2) \to \Lambda(d)^2, \qquad \mathrm{E}\eta_m(j_1)\mathrm{E}\eta_m(j_2) \to \Lambda(d)^2, \tag{5.8}$$

as $m, j_1/m, (j_2 - j_1)/m \to \infty$. Clearly, the second relation in (5.8) follows from (5.4). Next, by Assumption (A1),

$$\begin{aligned} &G_m^{-1}\Big(\sum_{t=j_1+1}^{m+j_1} (X_{t+m} - X_t), \sum_{t=j_1+m+1}^{2m+j_1} (X_{t+m} - X_t), \sum_{t=j_2+1}^{m+j_2} (X_{t+m} - X_t), \sum_{t=j_2+m+1}^{2m+j_2} (X_{t+m} - X_t)\Big) \\ &= G_m^{-1}\left(\Delta^2 S_{[mT_1]}, \Delta^2 S_{[m(T_1+1)]}, \Delta^2 S_{[mT_2]}, \Delta^2 S_{[m(T_2+1)]}\right) \\ &\to_{\mathrm{D}} G(d)\left(\Delta^2 B^1_{d+.5}(0), \Delta^2 B^1_{d+.5}(1), \Delta^2 B^2_{d+.5}(0), \Delta^2 B^2_{d+.5}(1)\right). \end{aligned} \tag{5.9}$$

as $m \to \infty, j_1/m = T_1 \to \infty, j_2/m = T_2 \to \infty, (j_2 - j_1)/m = T_2 - T_1 \to \infty$. Whence and from the definition of $\eta_m(j)$ the first relation in (5.8) easily follows. This proves (5.6) and part (i).

(ii) The proof is similar to that of part (i). Write $(mG_m)^{-1} \sum_{t=j+1}^{j+m} (X_{t+m} - X_t) = \int_0^1 h_{m,T}(\tau) \mathrm{d}\tau$, where $T := (j+1)/m$,

$$h_{m,T}(\tau) := G_m^{-1}(X_{[m(\tau+1+T)]} - X_{[mT]}) - G_m^{-1}(X_{[m(\tau+T)]} - X_{[mT]}).$$

By Assumption (A2),

$$h_{m,T}(\tau) \to_{\mathrm{FDD}} G(d)(B_{d-.5}(\tau+1) - B_{d-.5}(\tau)) \tag{5.10}$$

as $m, (j+1)/m = T \to \infty$. It is easy to check that the sequence of random processes $\{h_{m,T}(\tau), \tau \in [0,1]\}$ satisfies the weak convergence criterion in $L^1[0,1]$ due to Cremers and Kadelka (1986). Indeed, from (2.4), for any $\tau \in [0,1]$

$$\mathrm{E}|h_{m,T}(\tau)| \le (\mathrm{E}h_{m,T}^2(\tau))^{1/2} = G_m^{-1} \left(\mathrm{E}(X_{[m\tau]+m+j+1} - X_{[m\tau]+j+1})^2\right)^{1/2} \le C_2^{1/2}$$

which also implies $\mathrm{E}|h_{m,T}(\tau)| \to |G(d)|\mathrm{E}|B_{d-.5}(\tau+1) - B_{d-.5}(\tau)|$ and therefore the convergence in (5.10) extends to the weak convergence in $L^1[0,1]$ so that

$$(mG_m)^{-1} \sum_{t=j+1}^{j+m} (X_{t+m} - X_t) \ \to_{\mathrm{D}} \ G(d) \int_0^1 \Delta B_{d-.5}(\tau) \mathrm{d}\tau$$

as $m, (j+1)/m = T \to \infty$. In a similar way, from Assumption (A2) we obtain

$$\begin{aligned}
(mG_m)^{-1}\Big(\sum_{t=j_1+1}^{m+j_1} (X_{t+m} - X_t), &\sum_{t=j_1+m+1}^{2m+j_1} (X_{t+m} - X_t), \\
&\sum_{t=j_2+1}^{m+j_2} (X_{t+m} - X_t), \sum_{t=j_2+m+1}^{2m+j_2} (X_{t+m} - X_t)\Big) \\
\to_{\mathrm{D}} \ G(d)\Big(\int_0^1 \Delta B^1_{d-.5}(\tau)\mathrm{d}\tau, \int_0^1 \Delta B^1_{d-.5}(\tau+1)\mathrm{d}\tau, &\int_0^1 \Delta B^2_{d-.5}(\tau)\mathrm{d}\tau, \int_0^1 \Delta B^2_{d-.5}(\tau+1)\mathrm{d}\tau\Big),
\end{aligned}$$

as $m \to \infty, j_1/m = T_1 \to \infty, j_2/m = T_2 \to \infty, (j_2 - j_1)/m = T_2 - T_1 \to \infty$. The remaining details are similar as in the proof of (i). Proposition 2.1 is proved. □

PROOF OF PROPOSITION 2.2 follows by standard Fourier series argument and is omitted. □

PROOF OF PROPOSITION 2.3. Let first $-.5 < d < .5$. Write

$$\sum_{t=1+[mT]}^{[m(T+\tau)]} X_t = \sum_{t=1+[mT]}^{[m(T+\tau)]} X_t^0 + S_{m,T}(\tau),$$

where $X_t^0 := \sum_{s=-\infty}^{t} \psi(t-s)\xi_s$ is a stationary FARIMA$(0,d,0)$ process, and

$$S_{m,T}(\tau) := \sum_{s=-\infty}^{0} \sum_{t=1+[mT]}^{[m(T+\tau)]} \psi(t-s)\xi_s,$$

Therefore relation (2.2) for X_t follows from the fact that this relation holds for the stationary FARIMA$(0,d,0)$ process X_t^0, under the normalization $G_m = m^{-d-.5}$, see Bružaitė and Vaičiulis (2005), and

$$m^{-2d-1}\mathrm{E}S^2_{m,T}(\tau) \ \to \ 0 \qquad (m\to\infty,\ T\to\infty). \tag{5.11}$$

But

$$\begin{aligned}\mathrm{E}S^2_{m,T}(\tau) \ &= \ \sum_{s=0}^{\infty}\Bigg(\sum_{t=1+[mT]}^{[m(T+\tau)]}\psi(t+s)\Bigg)^2 \ \le \ C\int_0^\infty\Bigg(\int_{mT}^{m(T+\tau)}(t+s)^{d-1}\mathrm{d}t\Bigg)^2\mathrm{d}s\\ &= \ Cm^{2d+1}\int_0^\infty\Bigg(\int_T^{T+\tau}(t+s)^{d-1}\mathrm{d}t\Bigg)^2\mathrm{d}s \ = \ o(m^{2d+1}).\end{aligned}$$

Next, let $.5<d<1.5, d\ne 1$. Let

$$Y_t^0 := \sum_{s=-\infty}^{t}\Delta\psi(t-s)\xi_s, \qquad \Delta\psi(t):=\psi(t)-\psi(t-1)\quad (t\ge 1), \qquad \Delta\psi(0):=1. \tag{5.12}$$

Note $\Delta\psi(j)\sim j^{d-2}/\Gamma(d-1)$ $(j\to\infty)$ and so $\sum_{j=0}^\infty(\Delta\psi(j))^2<\infty$ $(1<d<1.5)$, $\sum_{j=0}^\infty|\Delta\psi(j)|<\infty$, $\sum_{j=0}^\infty\psi(j)=0$ $(.5<d<1)$. Therefore Y_t^0 in (5.12) is well-defined, as a stationary moving average process, and satisfies Assumption (A.1) with d replaced by $d-1\in(-.5,.5)$, see Bružaitė and Vaičiulis (2005). We have

$$X_{[m(T+\tau)]}-X_{[mT]} = \sum_{t=[mT]+1}^{[m(T+\tau)]}Y_t^0+U_{m,T}(\tau),$$

where

$$U_{m,T}(\tau) := \sum_{s=-\infty}^{0}\sum_{t=1+[mT]}^{[m(T+\tau)]}\Delta\psi(t-s)\xi_s$$

satisfies $m^{-2(d-1)-1}\mathrm{E}U^2_{m,T}(\tau) \ \to \ 0 \qquad (m\to\infty,\ T\to\infty)$ (the proof of the last relation is analogous to (5.11). We have proved that X_t satisfies Assumption (A.2). The statement of Proposition 2.3 in the case $d=1$ is obvious, as X_t reduces to a sum of iid rv's. Proposition 2.3 is proved. □

Proof of Theorem 2.4 As explained in Section 2, the theorem follows from the inequality (2.19) and the asymptotics (2.20) - (2.23).

Proof of (2.20) - (2.23). Without loss of generality, assume $c_0=1$. We shall separately consider the cases (i) $(-.5<d<.5)$ and (ii) $(.5<d<1.5)$.

Case (i). Let $r(t)=\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}tx}f(x)\mathrm{d}x$ be the covariance of X_t. Then

$$V_m^2 \ = \ 2\sum_{t,s=1}^m r(t-s)-2\sum_{t,s=1}^m r(t-s+m) \ = \ 4\int_{-\pi}^{\pi}f(x)\frac{\sin^4(mx/2)}{\sin^2(x/2)}\mathrm{d}x, \tag{5.13}$$

and, similarly,

$$R_m \ = \ 4\int_{-\pi}^{\pi}f(x)\cos(mx)\frac{\sin^4(mx/2)}{\sin^2(x/2)}\mathrm{d}x. \tag{5.14}$$

Consider the integral

$$J(a,m) := \int_0^\pi x^a\frac{\sin^4(mx/2)}{\sin^2(x/2)}\mathrm{d}x \ = \ 4m^{1-a}\int_0^\infty x^{a-2}\sin^4(x/2)g_m(x)\mathrm{d}x,$$

where $-1 < a < 1$ and

$$g_m(x) := \begin{cases} \frac{(x/2m)^2}{\sin^2(x/2m)}, & 0 < x < m\pi, \\ 0, & x > m\pi. \end{cases}$$

Note for each $x > 0$, $g_m(x) \to 1$ $(m \to \infty)$. Next, $\int_0^\infty x^{a-2} \sin^4(x/2)|g_m(x) - 1| \mathrm{d}x \le I_1 + I_2$, where

$$I_1 := \int_{m\pi}^\infty x^{a-2} \mathrm{d}x, \qquad I_2 := \int_0^{m\pi} x^{a-2} \sin^4(x/2) \frac{(x/2m)^2 - \sin^2(x/2m)}{\sin^2(x/2m)} \mathrm{d}x$$

Here, $I_1 \le Cm^{a-1}$. Using the bounds: $\sin(x) \ge c_1 x$ and $|x^2 - \sin^2(x)| \le c_2 x^4$ $(0 < x < \pi/2)$, where $c_1, c_2 > 0$ are some constants, we obtain

$$I_2 \le C \int_0^{m\pi} x^{a-2} \min(x^4, 1)(x/m)^2 \mathrm{d}x \ \le \ Cm^{a-1}.$$

Note also $I(a) := \int_0^\infty x^{a-2} \sin^4(x/2) \mathrm{d}x < \infty$. We thus obtain that for $a \in (-1, 1)$,

$$J(a, m) = 4I(a)m^{1-a} + O(1) \qquad (m \to \infty). \tag{5.15}$$

Applying (5.15) with $a = -2d \in (-1, 1)$ and $a = -2d + \beta \in (-1, 1)$ (as $0 < \beta < 2d + 1$) we obtain relation (2.20), with $V(d) = 4 \int_0^\infty x^{-2d-2} \sin^4(x/2) \mathrm{d}x$. In a similar way, (2.21) follows with $R(d) = 4 \int_0^\infty x^{-2d-2} \cos(x) \sin^4(x/2) \mathrm{d}x$. To explicitly obtain the above integrals, use (2.8) and the identities $\int_0^\infty x^{-2d-2} \sin^2(x/2) \mathrm{d}x = K(d + .5)/8$ (see Taqqu (2003, (9.8)) and $\sin^4(x/2) = -\frac{1}{4} \sin^2(x) + \sin^2(x/2)$, $\cos(x) \sin^4(x/2) = \frac{1}{8}\Big(- \sin^2(3x/2) + 4 \sin^2(x) - 7 \sin^2(x/2)\Big)$.

Case (ii) follows similarly to (i), by writing V_m, R_m in terms of the spectral density $f(x)$ of $U_t = X_t - X_{t-1}$:

$$V_m^2 = \int_{-\pi}^{\pi} \frac{\sin^4(mx/2)}{\sin^4(x/2)} f(x) \mathrm{d}x, \tag{5.16}$$

$$R_m = \int_{-\pi}^{\pi} \cos(mx) \frac{\sin^4(mx/2)}{\sin^4(x/2)} f(x) \mathrm{d}x, \tag{5.17}$$

c.f. (5.13), (5.14). The remaining details are similar to those of part (i) and are omitted.

Proof of (2.19). Let $\rho_m := \mathrm{E}Y^0 Y^1, \rho := \mathrm{E}Z^0 Z^1 = \rho(d)$. Let ξ_0, ξ_1 be mutually independent standard $\mathcal{N}(0, 1)$ random variables. Then

$$(Y^0, Y^1) =_{\mathrm{D}} (\xi_0, \rho_m \xi_0 + (1 - \rho_m^2)^{1/2} \xi_1), \qquad (Z^0, Z^1) =_{\mathrm{D}} (\xi_0, \rho \xi_0 + (1 - \rho^2)^{1/2} \xi_1), \tag{5.18}$$

in the sense of equality of distributions. By (2.17),

$$\begin{aligned} \mathrm{E}IR - \Lambda(d) &= \mathrm{E}\left[\frac{|\xi_0 + (\rho_m \xi_0 + (1 - \rho_m^2)^{1/2} \xi_1)|}{|\xi_0| + |\rho_m \xi_0 + (1 - \rho_m^2)^{1/2} \xi_1|} - \frac{|\xi_0 + (\rho \xi_0 + (1 - \rho^2)^{1/2} \xi_1)|}{|\xi_0| + |\rho \xi_0 + (1 - \rho^2)^{1/2} \xi_1|} \right] \\ &= \int_\rho^{\rho_m} \mathrm{E}\left[\frac{\partial \phi(r; \xi_0, \xi_1)}{\partial r} \right] \mathrm{d}r, \end{aligned} \tag{5.19}$$

where $\phi(r; x_0, x_1) := (|x_0 + (rx_0 + (1 - r^2)^{1/2} x_1))/(|x_0| + |rx_0 + (1 - r^2)^{1/2} x_1|)$. It is easy to check that $|\partial \phi(r; x_0, x_1)/\partial r| \le C/(1 - r^2)$ is bounded uniformly in x_0, x_1 provided r^2 is separated from 1: $1 - r^2 > C_1 > 0$. Then (2.19) is immediate from (5.19). Theorem 2.4 is proved. □

Proof of Theorem 2.5 (i) Recall the definitions of $Y_j(m)$ in (2.28) and $Z_d(\tau)$ in (1.5). We start with the relation

$$Y_m([m\tau]) \to_{\rm FDD} Z_d(\tau), \tag{5.20}$$

which holds in view of Gaussianity of X_t and the assumptions on the spectral density $f(x)$. Next,

$$(N/m)\mathrm{var}(IR) \;=\; \frac{N}{N-3m}\sum_{|j|<N-3m} m^{-1}\mathrm{cov}\left(\eta_m(0),\eta_m(j)\right)\left(1-\frac{|j|}{N-3m}\right), \tag{5.21}$$

Using the convergence (5.20), one can easily show that as $m\to\infty$ and $j/m\to\tau\in\mathbb{R}$, so

$$\mathrm{cov}\left(\eta_m(0),\eta_m(j)\right) \;\to\; \mathrm{cov}\left(\frac{|Z_d(0)+Z_d(1)|}{|Z_d(0)|+|Z_d(1)|},\frac{|Z_d(\tau)+Z_d(\tau+1)|}{|Z_d(\tau)|+|Z_d(\tau+1)|}\right). \tag{5.22}$$

From Lemma 5.1, (5.35) and Arcones (1994, Lemma 1) it follows that there exists a constant $C>0$ such that for any $m,j\ge 1$

$$|\mathrm{cov}\left(\eta_m(0),\eta_m(j)\right)| \;\le\; C(j/m)^{-2}. \tag{5.23}$$

Then (2.25) follows from (5.21), (5.22), the dominated convergence theorem and (5.23).

The proof of (2.26) follows the usual scheme of the proof of CLT's for sums of subordinated Gaussian functionals using Hermite expansion and the diagram formula; see e.g., Breuer and Major (1983), Giraitis and Surgailis (1985), Chambers and Slud (1989), Arcones (1994). (However, these results do not directly apply to our situation since $Y_m(j), 0\le j<N-3m$ form a triangular array.) Therefore, we present an outline of the proof of the CLT (Steps 1-3 below).

Step 1: Hermite expansion. Let

$$\xi_{0m}(j) := Y_m(j), \qquad \xi_{1m}(j) := (1-\rho_m^2)^{-1/2}(Y_m(j+m)-\rho_m Y_m(j)). \tag{5.24}$$

Then for each j,m, $\xi_{0m}(j),\xi_{1m}(j)$ are independent and have a standard Gaussian distribution; moreover, $Y_m(j)=\xi_{0m}(j)$, $Y_m(j+m)=\rho_m\xi_{0m}(j)+(1-\rho_m^2)^{1/2}\xi_{1m}(j)$; see (5.18). Let

$$g_m(x_0,x_1) \;:=\; \frac{|x_0+\rho_m x_0+(1-\rho_m^2)^{1/2}x_1|}{|x_0|+|\rho_m x_0+(1-\rho_m^2)^{1/2}x_1|}. \tag{5.25}$$

Then $\eta_m(j)=g_m\left(\xi_{0m}(j),\xi_{1m}(j)\right)$ is a nonlinear function (bounded by 1) in standardized Gaussian variables of (5.24). One can write the Hermite expansion:

$$\eta_m(j) \;=\; \mathrm{E}\eta_m(j)+\sum_{k_0,k_1\ge 0:k_0+k_1\ge 2}\frac{c^{(m)}_{k_0,k_1}}{k_0!\;k_1!}He_{k_0}\left(\xi_{0m}(j)\right)He_{k_1}\left(\xi_{1m}(j)\right), \tag{5.26}$$

convergent in mean square, where

$$c^{(m)}_{k_0,k_1} := \mathrm{E}\left[g_m(\xi_0,\xi_1)He_{k_0}(\xi_0)He_{k_1}(\xi_1)\right], \tag{5.27}$$

where $\xi_0,\xi_1\sim\mathcal{N}(0,1)$ are uncorrelated, and where $He_k(x)=(-1)^k\mathrm{e}^{x^2/2}(\mathrm{e}^{-x^2/2})^{(k)}$, $k=0,1,\dots$ are Hermite polynomials. Note $\mathrm{E}\eta_m(j) \;=\; \mathrm{E}g_m\left(\xi_{0m}(j),\xi_{1m}(j)\right) \;=\; c^{(m)}_{0,0}$ and $c^{(m)}_{1,0}=\mathrm{E}[\xi_0 g_m(\xi_0,\xi_1)]=0$, $c^{(m)}_{0,1}=\mathrm{E}[\xi_1 g_m(\xi_0,\xi_1)]=0$ which follows by symmetry of g_m in (5.25).

Step 2: approximation by finite sum of Hermite polynomials. Let $K\ge 1$ be a sufficiently large integer. From (5.26) we can write

$$IR-\mathrm{E}IR \;=\; S_K+\widetilde{S}_K, \tag{5.28}$$

where $S_K := (N-3m)^{-1}\sum_{j=0}^{N-3m-1}\eta_{m,K}(j)$, $\widetilde S_K := (N-3m)^{-1}\sum_{j=0}^{N-3m-1}\widetilde\eta_{m,K}(j)$, and where

$$\eta_{m,K}(j) \;:=\; \sum_{2\le k_0+k_1\le K}\frac{c^{(m)}_{k_0,k_1}}{k_0!k_1!}He_{k_0}(\xi_{0m}(j))He_{k_1}(\xi_{1m}(j)), \tag{5.29}$$

$$\widetilde\eta_{m,K}(j) \;:=\; \sum_{k_0+k_1>K}\frac{c^{(m)}_{k_0,k_1}}{k_0!k_1!}He_{k_0}(\xi_{0m}(j))He_{k_1}(\xi_{1m}(j)). \tag{5.30}$$

Similarly to (5.23) we obtain $|\mathrm{cov}\,(\widetilde\eta_{m,K}(0),\widetilde\eta_{m,K}(j))| \le \delta(K)(j/m)^{-2}$, where $\delta(K)$ does not depend on $m, j\ge 1$, and vanishes as $K\to\infty$. As a consequence, the second term on the r.h.s. in (5.28) is negligible, and it suffices to prove the CLT for the (truncated) term S_K only, namely,

$$(N/m)^{1/2}S_K \;\to_{\mathrm{D}}\; \mathcal{N}(0,\sigma^2_K), \tag{5.31}$$

as $N, m, N/m\to\infty$, where $\lim_{K\to\infty}\sigma^2_K = \sigma^2(d)$ (the proof of the last fact is similar to (2.25) above). *Step 3: proof of (5.31).* Similarly to (2.25), one can show $(N/m)\mathrm{E}S^2_K \to \sigma^2_K$ $(N, m, N/m\to\infty)$. Therefore the proof of (5.31) reduces to asymptotic normality of sums of (bivariate) Hermite polynomials in (5.29). In other words, it suffices to show that for any $p\ge 3$ and all sufficiently large $N, m\ge 1$

$$\mathrm{cum}(S(q_{01},q_{11}),\dots,S(q_{0p},q_{1p})) \;=\; o\left((N/m)^{-p/2}\right), \tag{5.32}$$

where $\mathrm{cum}(\cdot,\dots,\cdot)$ stands for joint cumulant,

$$S(k_0,k_1) := \frac{1}{N-3m}\sum_{j=0}^{N-3m-1} He_{k_0}(\xi_{0m}(j))He_{k_1}(\xi_{1m}(j)),$$

and where $q_{ij}\ge 0$ $(i=0,1, j=1,\dots,p)$ are arbitrary integers such that $q_{01}+q_{11}\ge 2,\dots,q_{0p}+q_{1p}\ge 2$. By the diagram formula, see e.g., Arcones (1994), Giraitis and Surgailis (1985), the cumulant in (5.32) can be written as a sum of contributions $J(\gamma)$ corresponding to all *connected diagrams* γ of the table

$$T \;=\; \begin{pmatrix} (1,1) & (1,2) & \dots & (1,q_{01}+q_{11}) \\ (2,1) & (2,2) & \dots & (2,q_{02}+q_{12}) \\ \dots & & & \\ (p,1) & (p,2) & \dots & (p,q_{0p}+q_{1p}) \end{pmatrix}, \tag{5.33}$$

and (5.32) follows from $J(\gamma) = o\left((N/m)^{-p/2}\right)$, for any given connected diagram γ. The last relation can be proved using the bound in Lemma 5.1 (i) and the (generalized) Hölder inequality in Giraitis and Surgailis (1985, (2.13)) (see also Surgailis (2003, Proposition 3.1)). This concludes the proof of (5.31) and part (i) of Theorem 2.5, too.

(ii) is very similar to that of (i). Consider the representation (2.29), with Y_j defined as in (2.28). From the assumptions on X_t and spectral density $f(x)$ of U_t, it easy to verify the relation (5.20). Then (2.25) follows from (5.22) as in part (i), and from the bound

$$|\mathrm{cov}\,(\eta_m(0),\eta_m(j))| \;\le\; C|j/m|^{-2\min(1,3-2d)}, \tag{5.34}$$

as $2\min(1,3-2d)>1$ for $.5<d<1.25$. The proof of (5.34) is exactly similar to that of (5.23), with the difference that Lemma 5.1 (i) must be replaced by Lemma 5.1 (ii). Theorem 2.5 is proved. □

Lemma 5.1 (i) *Let the assumptions of Theorem 2.5 (i) be satisfied. Then there exists a constant $C > 0$ such that for any integers $m, j \ge 1$*

$$|\mathrm{cov}\,(Y_m(0), Y_m(j))| \;\le\; C(j/m)^{-1}. \tag{5.35}$$

(ii) *Let the assumptions of Theorem 2.5 (ii) be satisfied. Then there exists a constant $C > 0$ such that for any integers $m, j \ge 1$*

$$|\mathrm{cov}\,(Y_m(0), Y_m(j))| \;\le\; C \begin{cases} (j/m)^{-1}, & .5 < d < 1, \\ (j/m)^{-1}\,(1 + \log(1 + (j/m))), & d = 1, \\ (j/m)^{3-2d}, & 1 < d < 1.5. \end{cases} \tag{5.36}$$

Proof. (i) Let $\rho_m(j) = \mathrm{cov}(Y_m(0), Y_m(j))$. Similarly as in (5.13)-(5.14),

$$\begin{aligned}
\rho_m(j) &= V_m^{-2} \sum_{t,s=1}^{m} (2r(j+t-s) - r(j+m+t-s) - r(j-m+t-s)) \\
&= 4V_m^{-2} \int_{-\pi}^{\pi} f(x)\cos(jx) \frac{\sin^4(mx/2)}{\sin^2(x/2)} \mathrm{d}x \\
&= -8V_m^{-2} j^{-1} \int_0^{\pi} \sin(jx) F_m'(x) \mathrm{d}x,
\end{aligned}$$

where $F_m(x) := f(x)\sin^4(mx/2)/\sin^2(x/2)$. Hence using $V_m^2 \sim \mathrm{const.}m^{2d+1}$ we obtain

$$|\rho_m(j)| \le C m^{-2d-1} j^{-1} \int_0^{\pi} |F_m'(x)| \mathrm{d}x. \tag{5.37}$$

Note $F_m'(x) \to 0$ $(x \to 0)$ by condition (2.12) and

$$\Big| \Big(\frac{\sin^4(mx/2)}{\sin^2(x/2)} \Big)' \Big| \;\le\; C \begin{cases} m^4 x, & 0 < x < 1/m, \\ m x^{-2}, & x > 1/m \end{cases}. \tag{5.38}$$

From (2.12) and (5.38) we obtain

$$\begin{aligned}
\int_0^{1/m} |F_m'(x)| \mathrm{d}x &\le C \int_0^{1/m} \Big(x^{-2d} m^4 x + x^{-2d} m^2 \Big) \mathrm{d}x \;=\; C m^{2d+2}, \\
\int_{1/m}^{\pi} |F_m'(x)| \mathrm{d}x &\le C \int_{1/m}^{\pi} \Big(x^{-2d} m x^{-2} + x^{-2d-1} x^{-2} \Big) \mathrm{d}x \;\le\; C m^{2d+2},
\end{aligned}$$

implying (5.35) by (5.37).

(ii) As $\mathrm{E}Y_m^2(j) = 1$, it suffices to prove the statement for $j \ge m$. Furthermore, for simplicity we shall assume that j is an even integer, $j \ge 2$. Similarly as in (5.13)-(5.14),

$$\begin{aligned}
\rho_m(j) &= V_m^{-2} \mathrm{E} \Big(\sum_{t=1}^{m} \sum_{t < s \le t+m} U_s \Big) \Big(\sum_{t'=j+1}^{j+m} \sum_{t' < s' \le t'+m} U_{s'} \Big) \\
&= 2V_m^{-2} \int_0^{\pi} f(x)\cos(jx) \frac{\sin^4(mx/2)}{\sin^4(x/2)} \mathrm{d}x \;=:\; 2V_m^{-2} I.
\end{aligned} \tag{5.39}$$

Write $I = \int_0^{2\pi/j} \ldots + \int_{2\pi/j}^{\pi} \ldots =: I_1 + I_2$. Here,

$$I_2 \;=\; j^{-1} \int_{2\pi}^{\pi j} \cos(y) f(y/j) \frac{\sin^4(my/2j)}{\sin^4(y/2j)} \mathrm{d}y \;=\; j^{-1} \sum_{q=1}^{j/2-1} I_2(q),$$

where

$$I_2(q) := \int_{2\pi q}^{2\pi q+\pi} \cos(y)\left(f\left(\frac{y}{j}\right)\frac{\sin^4(my/2j)}{\sin^4(y/2j)} - f\left(\frac{y+\pi}{j}\right)\frac{\sin^4(m(y+\pi)/2j)}{\sin^4((y+\pi)/2j)}\right)\mathrm{d}y$$
$$= \int_{2\pi q}^{2\pi q+\pi} \cos(y)\left(\tilde F_m\left(\frac{y}{j}\right) - \tilde F_m\left(\frac{y+\pi}{j}\right)\right)\mathrm{d}y,$$

where $\tilde F_m(x) := f(x)\sin^4(mx/2)/\sin^4(x/2)$. Using condition (2.27) and the bound

$$\left|\Big(\frac{\sin^4(mx/2)}{\sin^4(x/2)}\Big)'\right| \le C\begin{cases} m^6 x, & 0<x<1/m,\\ x^{-5}, & 1/m<x<\pi,\end{cases}$$

we obtain

$$|\tilde F_m'(x)| \le |f'(x)|\Big(\frac{\sin^4(mx/2)}{\sin^4(x/2)}\Big) + |f(x)|\left|\Big(\frac{\sin^4(mx/2)}{\sin^4(x/2)}\Big)'\right|$$
$$\le C\begin{cases} x^{1-2d}m^4 + x^{2-2d}m^6x = m^4x^{1-2d}, & \text{if } 0<x<1/m,\\ x^{1-2d}x^{-4} + x^{2-2d}x^{-5} = x^{-3-2d}, & \text{if } 1/m<x<\pi.\end{cases} \tag{5.40}$$

Then for $1 \le q \le j/m$, using the first bound in (5.40), we obtain

$$|I_2(q)| \le C|\tilde F_m'(2\pi q/j)|j^{-1} \le Cm^4 j^{2d-2}q^{1-2d},$$

and

$$j^{-1}\sum_{q=1}^{j/m}|I_2(q)| \le Cm^4j^{2d-3}\sum_{q=1}^{j/m}q^{1-2d} \le Cm^{2d+1}\begin{cases}(j/m)^{-1}, & .5<d<1,\\ (j/m)(1+\log(j/m)), & d=1,\\ (j/m)^{2d-3}, & 1<d<1.5.\end{cases}$$

On the other hand, for $j/m \le q \le j$, using the second bound in (5.40), we obtain

$$|I_2(q)| \le C|\tilde F_m'(2\pi q/j)|j^{-1} \le Cq^{-2d-3}j^{2+2d},$$

and

$$j^{-1}\sum_{q\ge j/m}|I_2(q)| \le Cj^{2d+1}\sum_{q\ge j/m}q^{-3-2d} \le Cm^{2d+1}(j/m)^{-1}.$$

Consequently,

$$|I_2| \le Cm^{2d+1}\begin{cases}(j/m)^{-1}, & .5<d<1,\\ (j/m)^{-1}\left(1+\log(1+(j/m))\right), & d=1,\\ (j/m)^{2d-3}, & 1<d<3/2.\end{cases} \tag{5.41}$$

Finally, using (2.14),

$$|I_1| \le Cm^4\int_0^{2\pi/j} f(x)\mathrm{d}x \le Cm^4j^{2d-3} = Cm^{2d+1}(j/m)^{2d-3}. \tag{5.42}$$

The statement of the lemma follows from (5.41), (5.42) and $V_m^2 \sim c_0V(d)m^{2d+1}$ (see (2.20)). □

6 Properties of the second increment of fBm

In this section, we discuss some properties of the process $Z_d(\tau), \tau \in \mathbb{R}$ (the second increment of fBm) defined in (1.5).

Proposition 6.1 *The processes $Z_d(\tau), \tau \in \mathbb{R}$ in (1.5) is well-defined and stationary Gaussian process, for any $-.5 < d < 1.5, d \neq .5$. It has zero mean, unit variance $\mathrm{E}Z_d^2(\tau) = 1$ and the covariance given in (1.6); more explicitly,*

$$\mathrm{E}Z_d(0)Z_d(\tau) \;=\; \frac{1}{2(4-4^{d+.5})}\Big(-|\tau+2|^{2d+1} + 4|\tau+1|^{2d+1} - 6|\tau|^{2d+1} + 4|\tau-1|^{2d+1} - |\tau-2|^{2d+1}\Big), \qquad (6.1)$$

The process $Z_d(\tau)$ admits the stochastic integral representation

$$Z_d(\tau) \;=\; C(d+.5)\int_{\mathbb{R}} \Delta_\tau^2 (\tau - x)_+^d M(\mathrm{d}x), \qquad (6.2)$$

where $M(\mathrm{d}x)$ is a standard Gaussian white noise with zero mean and variance $\mathrm{d}x$, and where $C(H) := \sqrt{\Gamma(2H+1)|\sin(\pi H)|/\Gamma(H+.5)^2|4-4^H|}$.

Proof. Equation (6.1) follows from (1.5), (2.1) and elementary integration; equation (6.2) is immediate from (1.5) and the stochastic integral representation of fBm given in Taqqu (2003). (One can easily check that the integrand in (6.2) belongs to $L^2(\mathbb{R})$ so that the stochastic integral is well-defined.) $\square$

Remark 6.1 (i) For $d = .5$, the process $Z_{.5}(\tau), \tau \in \mathbb{R}$ can be defined by continuity, as a stationary Gaussian process with zero mean and the covariance

$$\begin{aligned}\mathrm{E}Z_{.5}(0)Z_{.5}(\tau) &= \frac{1}{16\log 2}\Big((\tau+2)^2\log(\tau+2)^2 - 4(\tau+1)^2\log(\tau+1)^2 + 6\tau^2\log\tau^2 \\ &\qquad -4(\tau-1)^2\log(\tau-1)^2 + (\tau-2)^2\log(\tau-2)^2\Big) \\ &= -\frac{1}{16\log 2}\Delta_s^2\Delta_t^2 (t-s)^2\log(t-s)^2\Big|_{t-s=\tau}. \end{aligned} \qquad (6.3)$$

(ii) From Taylor expansion of (6.1),

$$\mathrm{E}Z_d(0)Z_d(\tau) \;\sim\; \frac{(2d+1)(2d)(2d-1)(2d-2)}{2(4^{d+.5}-4)}\tau^{2d-3} \qquad (\tau\to\infty). \qquad (6.4)$$

The asymptotic relation (6.4) holds for all $-.5 < d < 1.5, d \neq 0, .5, 1$. Note the asymptotic constant vanishes for $d = 0$ and $d = 1$. When $d = 0$ or $d = 1$, the autocovariance function is a piecewise polynomial, in particular,

$$\mathrm{E}Z_1(0)Z_1(\tau) \;=\; 2^{-3}\begin{cases} 6\tau^3 - 12\tau^2 + 8, & 0 \le \tau \le 1 \\ -2\tau^3 + 12\tau^2 - 24\tau + 16, & 1 \le \tau \le 2 \\ 0, & \tau \ge 2. \end{cases}$$

Proposition 6.2 *The function $\Lambda(d)$ in (1.3) satisfies $\Lambda(d) = \Lambda_0(\rho(d))$, where $\rho(d), \Lambda_0(r)$ are given in (1.4), (1.7), respectively. The function $\Lambda(d)$ is strictly increasing on the interval $(-.5, 1.5)$ and $\Lambda(1.5) = 1$. In particular,*

$$\begin{aligned}\Lambda(0) &= \Lambda_0(-.5) = \frac{2}{\pi}\arctan\sqrt{\frac{1}{3}} + \frac{1}{\pi}\sqrt{\frac{1}{3}}\log 4 = .588101\ldots,\\ \Lambda(1) &= \Lambda_0(.25) = \frac{2}{\pi}\arctan\sqrt{\frac{5}{3}} + \frac{1}{\pi}\sqrt{\frac{5}{3}}\log\Big(\frac{8}{5}\Big) = .773572\ldots.\end{aligned}$$

Proof. From the definition and the change of variables $x_1 = a\cos\phi, x_2 = a\sin\phi$,

$$\begin{aligned}\Lambda_0(r) &= \frac{1}{2\pi\sqrt{1-r^2}}\int_{\mathbb{R}^2}\frac{|x_1+x_2|}{|x_1|+|x_2|}\mathrm{e}^{-\frac{1}{2(1-r^2)}\left(x_1^2-2rx_1x_2+x_2^2\right)}\mathrm{d}x_1\mathrm{d}x_2\\ &= \frac{2\sqrt{1-r^2}}{\pi}\int_0^{\pi/4}\left(\frac{1}{1-r\sin(2\phi)}+\frac{\cos\phi-\sin\phi}{\cos\phi+\sin\phi}\,\frac{1}{1+r\sin(2\phi)}\right)\mathrm{d}\phi\\ &= \frac{2}{\pi}\arctan\left(\sqrt{\frac{1+r}{1-r}}\right)+\frac{\sqrt{1-r^2}}{\pi(1-r)}\big(\log 2-\log(1+r)\big),\end{aligned}$$

proving (1.7). The strict monotonicity of $\Lambda(d)$ follows from the monotonicity of $\rho(d), d \in (-.5, 1.5)$ and $\Lambda_0(r), r \in (-1, 1)$, which follows from

$$\Lambda_0'(r) = \frac{1}{\pi\sqrt{(1+r)(1-r)^3}}\log\left(\frac{2}{1+r}\right) > 0 \qquad (-1 < r < 1).$$

Proposition 6.2 is proved. □

Acknowledgments

We thank the referees and the editor for comments which helped to improve the presentation, Paul Doukhan for his invitations to present this paper to the LS-CREST ENSAE workshop “Données de grande dimension et processus dépendants”, April 2006, and the MAS–SMAI conference, Lille, September 2006.

References

Abadir, K., Distaso, W. and Giraitis, L. (2006) Estimation and inference for trending I(d) and related processes. *Preprint.*

Abry, P. and Veitch, D. (1998) Wavelet analysis of long-range dependent traffic. *IEEE Trans. Info. Theory* **44**, 2–15.

Abry, P., Flandrin, P., Taqqu, M. and Veitch, D. (2003) Self–similarity and long–range dependence through the wavelet lens. In: Doukhan, P., Oppenheim, G. and Taqqu, M.S. (Eds.), *Long–Range Dependence: Theory and Applications.* Birkhäuser, Boston, pp. 527–556.

Arcones, M.A. (1994) Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *Ann. Probab.* **22**, 2242–2274.

Aue, A., Horváth, L. and Steinebach, J. (2005) Long memory versus stochastic trend. *Preprint.*

Bardet, J.M., Lang, G., Moulines, E. and Soulier, P. (2000) Wavelet estimator of long–range dependent processes. *Statist. Inf. Stoch. Process.* **3**, 85–99.

Berkes, I., Horváth, L., Kokoszka, P.S. and Shao, Q-M. (2006) On discriminating between long–range dependence and changes in mean. *Ann. Statist.* **34**, 1140–1165.

Bhattacharya, R.N., Gupta, V.K. and Waymire, E. (1983) The Hurst effect under trends. *J. Appl. Probab.* **20**, 649–662.

Bollerslev, T. (1987) A conditional heteroskedastic time series model for speculative prices and rates of return. *Rev. Econ. Statist.*, **69**, 542–547.

Breuer, P. and Major, P. (1983) Central limit theorems for nonlinear functionals of Gaussian fields. *J. Multivariate Anal.* **13**, 425–441.

Bružaitė, K. and Vaičiulis, M. (2005) Asymptotic independence of distant partial sums of linear process. *Lithuanian Math. J.* **45**, 387–404.

Bružaitė, K., Surgailis, D. and Vaičiulis, D. (2006) Time-varying fractionally integrated processes with finite or infinite variance and nonstationary long memory. *Preprint.*

Csörgő, S. and Mielniczuk, J. (1995) Distant long-range dependent sums and regression estimation. *Stoch. Process. Appl.* **59**, 143–155.

Chambers, D. and Slud, E. (1989) Central limit theorems for nonlinear functional of stationary Gaussian processes. *Probab. Th. Rel. F.* **80**, 323–346.

Cox, D.R. (1984) Long-range dependence: A review. In: David, H.A. and David, H.T. (Eds.) *Statistics: An Appraisal.* Iowa State Univ. Press, Iowa, pp. 55–74.

Cremers, H. and Kadelka, D. (1986) On weak convergence of integral functions of stochastic processes with applications to processes taking paths in $L(p;E)$. *Stoch. Process. Appl.* **21**, 305–317.

Dacorogna, M.M., Müller, U.A., Nagler, R.J., Olsen, R.B. and Pictet, O.V. (1993) A geographical model for the daily and weekly seasonal volatility in the foregn exchange Market. *J. Int. Money Finance* **12,** 413–438.

Davydov, Yu. (1970) The invariance principle for stationary processes. *Theory Probab. Appl.* **15**, 487–498.

Dolado, J., Gonzalo, J. and Mayoral, L. (2005) What is what ? A simple time-domain test of long-memory vs. structural breaks. *Preprint.*

Diebold, F.X. and Inoue, A. (2001) Long memory and regime switching. *J. Econometrics* **105**, 131–159.

Gil–Alaña, L.A. and Robinson, P.M. (1997) Testing of unit root and other nonstationary hypotheses in macroeconomic time series. *J. Econometrics* **80**, 241–268.

Giraitis, L., Robinson, P.M. and Surgailis, D. (2000) A model for long memory conditional heteroscedasticity, *Ann. Appl. Probab.* **10**, 1002–1024.

Giraitis, L., Kokoszka, P. and Leipus, R. (2001) Testing for long memory in the presence of a general trend. *J. Appl. Probab.* **38**, 1033–1054.

Giraitis, L., Kokoszka, P., Leipus, R. and Teyssière, G. (2003a) Rescaled variance and related tests for long memory in volatility and levels. *J. Econometrics* **112**, 265–294.

Giraitis, L., Kokoszka, P., Leipus, R. and Teyssière, G. (2003b) On the power of R/S tests for stationarity against contiguous and semi-long memory alternatives. *Acta Appl. Math.* **78**, 285-299.

Giraitis, L., P., Leipus, R. and Philippe, A. (2006) A test for stationarity versus trends and unit roots for a wide class of dependent errors. *Econometric Th.* **22**, 989–1029.

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

Giraitis, L. and Surgailis, D. (2002) ARCH-type bilinear models with double long memory, *Stoch. Proc. Appl.* **100**, 275–300.

Gourieroux, C. and Jasiak, J. (2001) Memory and infrequent breaks. *Economics Lett.* **70**, 29–41.

Granger, C.W.J. and Hyung, N. (2004) Occasional structural breaks and long memory with an application to the S&P500 absolute stock returns. *J. Empirical Finance* **11**, 399–421.

Ho, H.-C. and Hsing, T. (1997) Limit theorems for functionals of moving averages. *Ann. Probab.* **25**, 1636–1669.

Ibragimov, I.A. and Linnik, Yu.V. (1971) *Independent and Stationary Sequences of Random Variables.* Wolters-Noordholf, Groningen.

Kokoszka, P. and Teyssière, G. (2002) Change–point detection in GARCH models: Asymptotic and bootstrap tests. *Preprint.*

Künsch, H. (1986) Discrimination between monotonic trends and long–range dependence. *J. Appl. Probab.* **23**, 1025–1030.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Shin, Y. (1992) Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root? *J. Econometrics* **54**, 159–178.

Lavielle, M. (1999) Detection of multiple changes in a sequence of dependent variables. *Stoch. Proc. Appl.* **83**, 79–102.

Lavielle, M. and Ludeña, C. (2005) Random thresholds for linear model selection. *Publication of the INRIA*, **5572**.

Lavielle, M. and Moulines, E. (2000) Least–squares estimation of an unknown number of shifts in a time series. *J. Time Series Anal.* **21**, 33–59.

Lavielle, M. and Teyssière, G. (2005) Adaptive detection of multiple change–points in asset price volatility. In: Teyssière, G. and Kirman, A. (Eds.), *Long Memory in Economics.* Springer Verlag, Berlin, pp. 129–156.

Lavielle, M. and Teyssière, G. (2006) Detection of multiple change–points in multivariate time series. *Lithuanian Math. J.* **46**, 287–306.

Leipus, R. and Viano, M.-C. (2003) Long memory and stochastic trend. *Statist. Probab. Lett.* **61**, 177–190.

Lo, A.W. (1991) Long-term memory in stock market prices. *Econometrica* **59**, 1279–1313.

Lobato, I. and Robinson, P.M. (1998) A nonparametric test for I(0). *Rev. Econ. Studies*, **65**, 475–495.

Marinucci, D. and Robinson, P.M. (1999) Alternative forms of fractional Brownian motion. *J. Statist. Plan. Inf.* **80**, 111–122.

Mikosch, T. and Stărică, C. (2004) Nonstationarities in financial time series, the long-range dependence and the IGARCH effects. *Rev. Econ. Statist.* **86**, 378–390.

Mikosch, T. and Stărică, C. (2003) Long–range dependence effects and ARCH modeling. In: Doukhan, P., Oppenheim, G. and Taqqu, M.S. (Eds.), *Long–Range Dependence: Theory and Applications.* Birkhäuser, Boston, pp. 439–459.

Mikosch, T. and Stărică, C. (2004) Change of structure in financial time series, long range dependence and the GARCH model. *REVSTAT Statistical Journal* **2**, 16–41.

Perron, P. (1989) The great crash, the oil shock and the unit root hypothesis. *Econometrica* **57**, 1361–1401.

Petrov, V.V. (1975) *Sums of Independent Random Variables.* Springer-Verlag, New York.

Philippe, A., Surgailis, D. and Viano, M.-C. (2006a) Invariance principle for a class of non stationary processes with long memory. *C. R. Acad. Sci. Paris Ser. 1* **342**, 269–274.

Philippe, A., Surgailis, D. and Viano, M.-C. (2006b) Almost periodically correlated processes with nonstationary long memory. In: Bertail, P., Doukhan, P. and Soulier, Ph. (Eds.), *Dependence in Probability and Statistics.* Lecture Notes in Statistics, Vol. **187**. Springer Verlag, Berlin, pp. 159–194.

Philippe, A., Surgailis, D. and Viano, M.-C. (2007) Time varying fractionally integrated processes with nonstationary long memory. *Th. Probab. Appl.* (forthcoming).

Robinson, P.M. (1994) Efficient tests of nonstationary hypotheses. *J. Amer. Statist. Assoc.* **89**, 1420–1437.

Sibbertsen, P. (2003) Log-periodogram estimation of the memory parameter of a long memory process under trend. *Statist. Probab. Lett.* **61**, 261–268.

Surgailis, D. (2003) CLTs for polynomials of linear processes: diagram formula with illustrations. In: Doukhan, P., Oppenheim, G. and Taqqu, M.S. (Eds.) *Long-Range Dependence: Theory and Applications.* Birkhäuser, Boston, pp. 111–127.

Taqqu, M.S. (1975) Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Zeit. Wahrsch. verw. Geb.* **31**, 287–302.

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

Teräsvirta, T. (1996) Two stylized facts and the GARCH(1,1) model. *Stockholm School of Economics, Working Paper Series in Economics and Finance* **96**.

Teyssière, G. and Abry, P. (2005) Wavelet analysis of nonlinear long–range dependent processes. Applications to financial time Series. In: Teyssière, G. and Kirman, A. (Eds.) *Long Memory in Economics.* Springer Verlag, Berlin, pp. 173–238.

Teyssière, G. (2003) Interaction models for common long–range dependence in asset price volatility. In: Rangarajan, G. and Ding, M. (Eds.) *Processes with Long Range Correlations: Theory and Applications.* Lecture Notes in Physics, Vol. **621**. Springer Verlag, Berlin, pp. 251–269.

Donatas Surgailis and Marijus Vaičiulis
Vilnius Institute of Mathematics and Informatics,
Akademijos 4, 2600 Vilnius. LITHUANIA.
email: sdonatas@ktl.mii.lt, marijus@ktl.mii.lt

Gilles Teyssière
SAMOS, CES, Université Paris 1 Panthéon–Sorbonne
Centre Pierre Mendès France
90, Rue de Tolbiac - 75634 Paris Cedex 13. FRANCE.
& Centre for Finance, Göteborg University.
Box 640, Vasagatan 1.
SE 405 30 Göteborg, SWEDEN.
e-mail: stats@gillesteyssiere.net