



HAL
open science

Frequency approach for detecting nonstationarity in dependent data

Mohamedou Ould Haye, Anne Philippe

► **To cite this version:**

Mohamedou Ould Haye, Anne Philippe. Frequency approach for detecting nonstationarity in dependent data. 2019. <hal-02126749v3>

HAL Id: hal-02126749

<https://hal.science/hal-02126749v3>

Preprint submitted on 19 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

A Frequency approach for detecting nonstationarity in dependent data

M. Ould Haye¹ A. Philippe^{2*}

¹School of Mathematics and Statistics.

Carleton University, 1125 Colonel By Dr. Ottawa, ON, Canada, K1S 5B6

²Laboratoire de Mathématiques Jean Leray, UMR6629 CNRS

2 rue de la Houssiniere, Nantes Université, 44 322 Nantes France.

Abstract

Distinguishing long memory behaviour from nonstationarity can be very difficult as in both cases the sample autocovariance function decays very slowly. Available stationarity tests either do not include long memory or fare poorly in terms of empirical size, especially near the boundary between long memory and nonstationarity. We propose a testing procedure based on evaluating periodograms at different epochs. Limiting distributions established here are easily tractable as sum of weighted independent χ^2 random variables. Moreover, numerical studies are provided to show that the proposed approach seems to outperform existing methods.

Keywords : Long memory, dependence, stationarity, frequency

*corresponding author : anne.philippe@univ-nantes.fr

1 Introduction

Let us consider a stationary process (X_t) with a spectral density of some semi-parametric form:

$$f(\lambda) = |\lambda|^{-2d} f^*(\lambda) \quad (1)$$

with $-1/2 < d < 1/2$ and f^* is an even, positive, continuous function on $[-\pi, \pi]$. Hereafter, we will refer to processes satisfying condition (1) as satisfying the null hypothesis of stationarity H_0 . It is common to denote this class of processes $I(d)$, $-1/2 < d < 1/2$. It is worth noticing that spectral densities of the form (1) encompass a large variety of stationary processes such as the very popular fractionally autoregressive integrated moving average processes FARIMA(p, d, q)

$$\Phi(B)(1 - B)^d X_t = \Theta(B)\epsilon_t,$$

where ϵ_t is a white noise with variance σ^2 , B is the lag operator $BX_t = X_{t-1}$, and where $\Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$, $\Theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$ are the autoregressive and moving average polynomials. The FARIMA spectral density is then given by

$$f(\lambda) = \frac{\sigma^2}{2\pi} \left(\left| \frac{\Theta(e^{i\lambda})}{\Phi(e^{i\lambda})} \right|^2 \right) |1 - e^{i\lambda}|^{-2d} = |\lambda|^{-2d} \left[\frac{\sigma^2}{2\pi} \left(\left| \frac{\sin(\lambda/2)}{(\lambda/2)} \right|^{-2d} \right) \left(\left| \frac{\Theta(e^{i\lambda})}{\Phi(e^{i\lambda})} \right|^2 \right) \right].$$

We say that the process (X_t) is of short memory if $d = 0$ (i.e. the spectral density $f = f^*$ in (1) is continuous and positive over $[-\pi, \pi]$), and we say that the process is of long memory if $0 < d < 1/2$ (i.e. the spectral density is unbounded at zero). The case $-1/2 < d < 0$ corresponds to the so-called persistent memory, where $f(0) = 0$. In time-domain terms, short memory means that the covariance function is summable (i.e. covariance function $\gamma(h) = \text{Cov}(X_i, X_{i+h})$ goes to zero fast as the lag h increases) and long memory means that the covariance function is not summable (i.e. the covariance function goes to zero

slowly as the lag increases). Literature is full of real world examples of both types of processes. Over the past decade, several monographs have been written on statistical and probabilistic aspects of long memory processes. The reader is referred to the recent book by Beran et al. (2013) as well as the book of Giraitis et al. (2012) and references therein. Under mild conditions, the switch between time-domain and frequency-domain is easily made via the two representations

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-ih\lambda} \gamma(h) = \frac{1}{2\pi} \left[\gamma(0) + 2 \sum_{h=1}^{\infty} \cos(\lambda h) \gamma(h) \right], \quad |\lambda| \leq \pi,$$

and reciprocally

$$\gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda} f(\lambda) d\lambda = \int_{-\pi}^{\pi} \cos(\lambda h) f(\lambda) d\lambda, \quad h = 0, \pm 1, \pm 2, \dots$$

From an applied point of view, long memory behaviour seems to be often confused with a lack of stationarity, as time series exhibiting such a behaviour tend to have a sample autocorrelation function with large spikes at several lags which is well known to be the signature of non-stationarity for many practitioners. There are many stationarity tests such as Phillips and Perron (1988) tests and the frequency test of Bailey and Giraitis (2016), which can be seen as extensions of the augmented Dickey-Fuller tests since their stationarity assumption includes larger class than i.i.d. case. For instance the former allows for mixing and the latter includes short memory linear processes. However, the first statistical test for stationarity under a large umbrella of dependence, i.e. short and long memory dependence versus non stationarity is the so-called V/S test introduced by Giraitis, Leipus and Philippe (2006).

We develop a statistical test that uses an estimate of the memory parameter d . Contrary to V/S test, this one detects the stationarity for all d in $(-1/2, 1/2)$ unlike V/S which requires d to be in a compact $[a, b] \subset (-1/2, 1/2)$. This drawback in V/S results in it having high

empirical sizes starting from $d = .4$. The proposed test does not suffer from this issue and keeps empirical sizes close to the nominal levels even for d close to $1/2$.

The simple idea of using an estimate \hat{d} of d with its confidence interval to test the stationarity poses some fundamental problems even if we rewrite the null and alternative hypotheses with semi-parametric forms: $X_t \in I(d)$, $-1/2 < d < 1/2$ versus it exists integer $p \geq 1$ such $(1 - B)^p X_t \in I(d)$ for certain $-1/2 < d < 1/2$. In this context, the most common d estimators are local Whittle or fully extended local Whittle estimator as defined in Abadir et al. (2007) or Giraitis et al. (2012). Their asymptotic normality requires an assumption on its bandwidth parameter m_n that depends in particular on the second order expansion (determined by a coefficient $\beta \in (0, 2]$) of the spectral density near zero. An optimal choice of the bandwidth, needed to ensure a well-sized test, requires additional assumptions on the null hypothesis stronger than (1), such as knowing β . For instance taking $\beta = 2$ we obtain a class of linear models containing FARIMA models. Note that our approach only requires the use of a consistent estimator of d . Moreover, in the simulations section, we will also see that the sole use of \hat{d} will not suffice to detect slight structural breaks (see Figure 5). An additional difficulty of writing a stationarity test as one side test $H_0: d < 1/2$ is to obtain a testing procedure with asymptotic level $\alpha (= .05$ for example). Assuming that asymptotic normality conditions of local Whittle estimator \hat{d} are satisfied,

$$2\sqrt{m_n}(\hat{d} - d) \rightarrow \mathcal{N}(0, 1)$$

(see Giraitis et al. (2012) for details). Hence under H_0 we have

$$\sup_{d < 1/2} \lim_{n \rightarrow \infty} P_d \left(\hat{d} > 1/2 + \frac{1.65}{2\sqrt{m_n}} \right) = 0.$$

A solution to obtain a test of asymptotic level $\alpha = .05$ consists in considering a sub hypothesis $d \leq 1/2 - \epsilon$ ($\epsilon > 0$). In this case we have

$$\sup_{d \leq 1/2 - \epsilon} \lim_{n \rightarrow \infty} P_d \left(\hat{d} > 1/2 - \epsilon + \frac{1.65}{2\sqrt{m_n}} \right) = .05.$$

However, this test excludes the main important part of the stationarity domain $1/2 - \epsilon < d < 1/2$ where non stationarity is most suspected in practice.

From a sample X_1, \dots, X_n of the process (X_t) , we are interested in building a testing procedure to discriminate between dependence and non-stationarity.

The proposed statistic is built from periodograms taken at different epochs. More precisely, the procedure is as follows: we split our initial sample X_1, \dots, X_n into m blocks (or epochs), each of size ℓ , and we construct the periodogram $I_{n,i}$ on the i th block $X_{(i-1)\ell+1}, \dots, X_{i\ell}$.

Let

$$Q_{n,m}(s, d) = (m^{-2d}) \sum_{j=1}^s \frac{I_n(\lambda_j)}{\frac{1}{m} \sum_{i=1}^m I_{n,i}(\lambda'_j)}, \quad (2)$$

with

$$I_n(\lambda) = \frac{1}{2\pi n} \left| \sum_{t=1}^n X_t e^{it\lambda} \right|^2, \quad I_{n,h}(\lambda') = \frac{1}{2\pi \ell} \left| \sum_{t=(h-1)\ell+1}^{h\ell} X_t e^{it\lambda'} \right|^2, \quad (3)$$

where $\lambda_j = 2\pi j/n$, and $\lambda'_j = 2\pi j/\ell$, $j = 1, \dots, s$, for fixed s , are the Fourier frequencies.

The statistic $Q_{n,m}(s, d)$ can be viewed as sum of self normalized periodograms $I_n(\lambda_j)$, $j = 1, \dots, s$. A particular case of the statistic defined in (2) was studied in Gromykov et al. (2018), where the statistic $Q_{n,m}(s, 0)$ was considered to test short memory versus long memory, and we established its asymptotic distribution (i.e. under short memory).

It is worth noticing that establishing limiting distributions under long memory necessitates modifying the short memory statistic, including a normalizing coefficient m^{-2d} , as it appears in (2), as well as two sets of Fourier frequencies λ_j and λ'_j . Of course, this will

result in different limiting distributions due to the long memory that will strongly affect the asymptotic behaviour of the periodogram.

We assume that $m = m(n)$, and $\ell = \ell(n) \rightarrow \infty$ as $n \rightarrow \infty$. It is important to emphasize the fact that m and ℓ increase with n and are not constant, and that $m = n/\ell$. We are simply using notation m and ℓ rather than $m(n)$ and $\ell(n)$ only for the sake of simplicity. The rest of the paper is organized as follows. Section 2 contains main limiting theorems related to the statistic $Q_{n,m}(s, d)$, including its behaviour under a wide range of nonstationary alternatives. In Section 3 we present our test statistic to detect nonstationarity using asymptotic properties obtained in Section 2. Section 4 contains some Monte Carlo study to illustrate the performance of our proposed method. The paper ends with an appendix containing all the proofs.

2 Main results

Let $(X_t)_{t \geq 0}$ be a linear process of the form

$$X_t = \sum_{j=0}^{\infty} a_j \epsilon_{t-j}, \quad (4)$$

where a_j are the Fourier coefficients of an L^2 -Hermitian function and $a_0 = 1$, and where (ϵ_j) are i.i.d. random variables with zero mean and finite fourth moment. Denote $\sigma^2 = \mathbb{E}(\epsilon_1^2)$ and $\eta = \mathbb{E}(\epsilon_1^4)$.

For $-1/2 < d < 1/2$, we define the following memory condition for stationary processes.

Condition \mathcal{H}_d The process X_t is a linear as defined in (4), and its spectral density (which exists) f_d is of the form (1). In addition, for $d = 0$, we assume that

$$\sum_{j=0}^{\infty} |a_j| < \infty, \quad (5)$$

and when $-1/2 < d < 0$, we will also assume that

$$a_j = c(d)j^{-1+d}(1 + O(j^{-1})), \quad j \geq 1, \quad \text{for some } c(d) \neq 0 \quad \text{and} \quad \sum_{j=0}^{\infty} a_j = 0. \quad (6)$$

Condition (5), imposed when $d = 0$, implies that

$$\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty.$$

Note that when $0 < d < 1/2$, condition (1) on the spectral density implies that

$$\gamma(h) \sim C(d)h^{2d-1}, \quad \text{as } h \rightarrow \infty, \quad \text{for some constant } C(d) > 0, \quad (7)$$

where $a_n \sim b_n$ means that $a_n/b_n \rightarrow 1$ as $n \rightarrow \infty$. (see Proposition 3.1.1. (a) of Giraitis et al. (2012)). Also when $-1/2 < d < 0$, condition (6) implies that (7) also still holds. (See Proposition 3.2.1. (iii) of Giraitis et al. (2012)). In what follows, we will focus on the statistic $Q_{n,m}(s, d)$ defined in (2). In the next theorem we give the asymptotic distribution of $Q_{n,m}(s, d)$ under the null hypothesis "the process is stationary". As for its behaviour under the alternative of "the process has a stochastic or deterministic trend", it is given in the Theorem 2. We first introduce a covariance matrix that will appear in the limiting distribution of $Q_{n,m}(s, d)$ in (11). The block diagonal form of the matrix is due to the dependence, inherited from the long memory, among normalized real parts and imaginary parts of the discrete Fourier transforms (DFT). Note that DFT's real and imaginary parts will be asymptotically independent even if they are taken at different Fourier frequencies.

Let $\Sigma(d)$ be a $(2s) \times (2s)$ covariance matrix of the form

$$\Sigma(d) = \begin{pmatrix} \Sigma^{(1)}(d) & 0 \\ 0 & \Sigma^{(2)}(d) \end{pmatrix} \quad (8)$$

where $\Sigma^{(1)}(d)$ and $\Sigma^{(2)}(d)$ are defined by

$$\Sigma_{j,k}^{(1)}(d) = \frac{1}{2} \frac{L_{j,k}(d) + L_{j,-k}(d)}{\sqrt{L_{j,j}(d)L_{k,k}(d)}} \quad \text{and} \quad \Sigma_{j,k}^{(2)}(d) = \frac{1}{2} \frac{L_{j,k}(d) - L_{j,-k}(d)}{\sqrt{L_{j,j}(d)L_{k,k}(d)}}, \quad (9)$$

for $j, k = 1, \dots, s$, and where for all integers $j, k \neq 0$

$$L_{j,k}(d) = \frac{|jk|^d}{\pi} \int_{-\infty}^{\infty} \frac{\sin^2(\lambda/2)}{(2\pi k - \lambda)(2\pi j - \lambda)} \left| \frac{\lambda}{2\pi} \right|^{-2d} d\lambda. \quad (10)$$

The next theorem gives the central result that is used to build the critical regions of the statistical test proposed in Section 3.

Theorem 1. *Let X_t be a stationary process satisfying \mathcal{H}_d for certain $d \in (-1/2, 1/2)$, and let $s \geq 1$ be fixed. Assume that there exist $\delta, \delta' > 0$ such that $\ell = O(n^{1-\delta})$ and $m = O(n^{1-\delta'})$ (with $n = m\ell$ by construction). Then the statistic $Q_{n,m}(s, d)$ defined in (2) satisfies*

$$Q_{n,m}(s, d) \xrightarrow{\mathcal{D}} Q(s, d) = \sum_{i=1}^{2s} \zeta_i(d) Q_i, \quad (11)$$

where Q_1, \dots, Q_{2s} are i.i.d. χ_1^2 random variables and $\zeta_1(d), \dots, \zeta_{2s}(d)$ are the eigenvalues of $\Sigma(d)$ defined in (8) and (9). In (11), $\xrightarrow{\mathcal{D}}$ represents the convergence in distribution.

Remark 1. *The limiting distribution form (11) is an elaborated form of the limiting form (17) (see in the proof below). Actually (11) can be seen as a sum of independent non-identically Gamma random variables and therefore its quantiles can easily be computed with high precision (see for instance Hu et al. (2020)). Evaluating the limit in (17) requires Monte Carlo simulations which may yield unstable quantiles.*

Proof. The proof is relegated to Section 5.1 of the appendix. □

The following propositions provide some insight on the limiting distribution obtained in Theorem 1. Namely we show that this distribution can be continuously extended to the case

$d = 1/2$. Moreover, we show that the quantiles (of fixed order) of the limiting distribution are continuous with respect to the memory parameter d . These two propositions will be useful in the next section when it comes to implementing the stationarity test.

Proposition 1. *The limiting distribution $Q(s, d)$ converges in distribution, as $d \rightarrow 1/2$, to a non degenerated random variable*

$$Q(s, 1/2) := \sum_{i=1}^{2s} \zeta_i(1/2) Q_i$$

where Q_1, \dots, Q_{2s} are i.i.d. χ_1^2 random variables and $\zeta_1(1/2), \dots, \zeta_{2s}(1/2)$ are the eigenvalues of the covariance matrix $\Sigma(1/2) = \lim_{d \rightarrow 1/2} \Sigma(d)$ where $\Sigma(d)$ is defined in (8) and (9).

Proposition 2. *Let $q_\alpha(s, d)$ be $Q(s, d)$'s quantile of order $1 - \alpha$ where $0 < \alpha < 1$. For fixed α , the mapping $(-1/2, 1/2] \rightarrow \mathbb{R}$, $d \mapsto q_\alpha(s, d)$ is continuous.*

The next theorem gives the asymptotic behaviour of the statistic $Q_{n,m}(s, d)$ for a wide variety of non stationary processes, such as random walk (with dependent increments) as well stationary noise with additive deterministic signals. This will ensure the consistency of the testing procedure described in the next section.

Theorem 2. *Let Y_t be a stationary process satisfying \mathcal{H}_d for certain $d \in (-1/2, 1/2)$, and consider the following two classes of non stationary processes X_t*

- *Stochastic trend (unit root) : $X_t = X_{t-1} + Y_t$*
- *Deterministic trend (structural breaks): $X_t = g_n(t) + Y_t$ where $g_n(t) = n^\beta g(t/n)$, with $\beta \geq 0$, and where g is a piece-wise continuous¹ satisfying*

$$\int_0^1 g(x) e^{i2\pi jx} dx \neq 0, \quad \text{for some } j \in \{1, 2, \dots, s\}. \quad (12)$$

¹the interval $[0,1]$ can be divided into a finite number of sub-intervals such that the function is continuous on each open sub-interval, with finite limit at each endpoint.

Then, for all $\delta \leq 1/2$

$$Q_{n,m}(s, \delta) \xrightarrow{P} \infty, \quad (13)$$

where \xrightarrow{P} denotes the convergence in probability.

Proof. The proof is relegated to Section 5.4 of the appendix. \square

Remark 2. We note in passing that condition (12) is satisfied by most of nonconstant functions. For example, nonnegative step functions (allowing change in the mean), polynomials, etc.

3 Testing procedures.

We want to check the stationarity of a process. More precisely we define the following hypotheses: the null hypothesis H_0 : "the process X_t is stationary" in the sense that there exists $-1/2 < d < 1/2$ such that X_t satisfies the condition \mathcal{H}_d . As such the process X_t satisfies Theorem 1 conditions. The alternative H_1 considered here is "the process X_t has a stochastic or deterministic trend" as defined in Theorem 2.

Hereafter, and for simplicity, we will refer to H_0 and H_1 as respectively stationarity and nonstationarity hypotheses.

We now describe different decision rules derived from the statistic $Q_{n,m}(s, d)$.

A decision rule can be derived from the statistic $Q_{n,m}(s, d)$ after estimating d . Let \hat{d} be a consistent estimator of d , for a process X_t satisfying \mathcal{H}_d assumption. We restrict \hat{d} to $(-1/2, 1/2]$ by forcing it to be equal to $1/2$ if \hat{d} exceeds $1/2$. If $\hat{d} - d = o_P(1/\ln m)$ then the convergence results established in Theorems 1 and 2 are still valid for the $Q_{n,m}(s, \hat{d})$ since

$$Q_{n,m}(s, \hat{d}) = e^{-2(\hat{d}-d)\ln m} Q_{n,m}(s, d).$$

To carry out such test, for fixed $\alpha \in (0, 1)$ we consider the rejection region

$$R_n = \{(X_1, \dots, X_n) \in \mathbb{R}^n, \quad Q_{m,n}(s, \hat{d}) > q_\alpha(s, \hat{d})\} \quad (14)$$

where $q_\alpha(s, \hat{d})$ is the quantile of order $(1 - \alpha)$ of $Q(s, \hat{d})$. Note that the rejection region R_n is still defined even if $\hat{d} = 1/2$ since $Q(s, 1/2)$ is non degenerated (see Proposition 1). This test has asymptotic level α thanks to Proposition 2 and is also clearly consistent.

Remark 3. *The fact that our limiting distribution can be extended to $d = 1/2$ gives our test a considerable advantage over V/S . Indeed, as the limiting distribution of V/S is degenerated at $d = 1/2$, \hat{d} had to be restricted to a compact $[-a, a] \subset (-1/2, 1/2)$ in V/S and its implementation. This leads to a blind spot $(a, 1/2]$ in the stationarity null hypothesis. This explains V/S 's high empirical sizes observed when d is close to $1/2$ (see simulation results in Giraitis et al. (2006)).*

4 Numerical Studies

4.1 Simulated models

We investigate two kinds of null hypothesis:

- Long memory processes ($d \neq 0$).

We simulate FARIMA(0, d , 0) processes $X_t = (1 - B)^{-d} \epsilon_t$ where $d \in (-1/2, 1/2)$ and $(\epsilon_t)_t$ are standard Gaussian innovations.

- Short memory processes ($d = 0$).

We simulate AR(1) processes: $X_t = aX_{t-1} + \epsilon_t$ where $a \in (-1, 1)$ and ϵ_t are standard Gaussian innovations.

Two alternatives are simulated as follows: starting from a FARIMA(0, d ,0) processes Y_t with $d \in (-1/2, 1/2)$, we take

- unit-root process $X_t = \sum_{i=1}^t Y_i$.

Note that, as an integrated process, X_t can be written as a FARIMA(0, $d+1$,0) with $d \in (-1/2, 1/2)$ which is not stationary (See Taqqu (2003) for a precise definition)

- Structural breaks (a change point in the mean in the middle $n/2$): $X_t = \mu_t + Y_t$ where

$$\mu_t = \begin{cases} 0 & \text{if } t < n/2, \\ \Delta & \text{if } t \geq n/2. \end{cases} \quad (15)$$

	$Q_{n,m}(s, \hat{d})$						V/S
	$m = n^{1/2}$			$m = n^{1/3}$			
d	s = 1	s = 3	s = 5	s = 1	s = 3	s = 5	
.41	0.05	0.03	0.03	0.07	0.09	0.10	0.10
.49	0.09	0.06	0.05	0.09	0.10	0.11	0.19

Table 1: Empirical size of tests based on $Q_{n,m}(s, \hat{d})$ and V/S for FARIMA(0, d , 0) processes when d is close to 1/2, $n=500$ with 5000 replications. The nominal level is $\alpha = 5\%$.

4.2 Empirical size and power evaluation

Fundamentally, the statistic $Q_{n,m}(s, \hat{d})$ depends on the choice of the bandwidth parameter m and the number of Fourier frequencies s . Our strategy is to calibrate m, s in order to ensure an empirical size that remains reasonably close to the nominal level $\alpha (= 5\%)$ set in advance, in particular when d is very close to 1/2. In doing so, and contrary to V/S test which is known to suffer from high empirical size, we are trying to drastically reduce the probability of rejecting the stationarity of processes with strong long memory processes.

Table 1 provides empirical size of our test for different values of s and m taking d close to the stationarity boundary 1/2. We also compare them with V/S test. The results show

that $m = \sqrt{n}$ performs better than $m = n^{1/3}$ in recognizing the stationarity of strong long memory processes. Moreover, the statistic seems to be little sensitive to the choice of s . Finally, our procedure outperforms V/S in recognizing the stationarity of strong long memory processes. This also confirms that V/S does suffer from very large empirical sizes in the neighbourhood of $d = 1/2$. In the subsequent numerical studies, we will focus on the bandwidth choice $m = \sqrt{n}$, where n is taking a moderate value of 500.

Figure 1 shows the empirical size as function of d of the tests $Q_{n,\sqrt{n}}(s, \hat{d})$ and V/S test. The figure also shows that the empirical size performance observed near non stationarity of $Q_{n,\sqrt{n}}(s, \hat{d})$ extends to all possible stationarity values of d .

In Figure 2 we consider stationary AR(1) process with coefficient a ranging from .8 to .99. For this model with short memory, it is legitimate to include Ducky-Fuller test. The comparison shows that, with $s = 1$, the performance of our statistics is comparable to that of V/S for stationary AR(1) processes near unit root (a close to 1). We notice that, near unit root, the choice of the number of frequencies s is important and the test performs better for $s = 1$. For this value, the results are similar to those of V/S . Of course Dickey Fuller's test is better than both V/S and our statistics for testing stationary AR processes against unit root. However our null hypothesis H_0 is much wider, as it contains long memory which cannot be seen as stationary by all sorts of Dickey Fuller's test. Note that, when we stay away from unit root, the stationarity is accepted, with probability one by all the above tests.

Figures 3 and 4 illustrate the power function of $Q_{n,\sqrt{n}}(s, \hat{d})$ under the alternatives: stochastic and deterministic trends respectively, and confirm the consistency of our test. The test V/S is artificially more powerful than ours in detecting slightly non stationary processes (d near 1/2 from the right). Indeed, as shown in Figure 1, this is due to its high empirical size (about 20% instead of the nominal level of 5%). Despite such high empirical level, with the right choice of s , our test becomes rapidly more powerful than V/S as d increases.

When it comes to choosing an s (the number of lower frequencies to be used), there does not seem to be a need for a trade-off between a good empirical size and a good power. In fact, as seen before, the empirical size is not affected by the choice of s among the tested values and the empirical power increases with s . The choice $s = 5$ seems to be satisfactory for both size and power.

From Figure 4, we notice a sharp difference in terms of empirical power behaviour between the stochastic trend and structural breaks. Unlike the stochastic trend, the power curves decrease with s . In this case, we recommend the use of $s = 1$, as it detects small structural breaks as represented in Figure 5. Moreover, this choice guarantees comparable power as V/S . In the structural breaks case, the non-increasing behaviour of the power function (in d) is consistent with previous findings in time-domain context (see for instance Giraitis et al., 2006). This is due to the fact that the marginal variance $\text{Var}(Y_1) = \sigma^2\Gamma(1 - 2d)/\Gamma^2(1 - d)$ increases with d , so that any small to moderate jump would be more difficult to detect for d close $1/2$, as the stationary noise Y_t will dominate.

4.3 Conclusion

We provided a periodogram based statistic to test stationarity versus non stationarity. The alternatives being non stationary FARIMA(p, d, q) with $1/2 < d < 3/2$ or deterministic trend plus a stationary FARIMA(p, d, q) with $-1/2 < d < 1/2$. The limiting distribution of $Q_{n,m}(s, d)$ extends to $d = 1/2$ without being degenerated. This, in part, renders the proposed test well empirically sized (i.e. with an empirical level close to the nominal one α) near non stationarity boundary $d = 1/2$. Consequently, empirical size-wise, the proposed method outperforms V/S in testing long memory versus stochastic or deterministic trends.

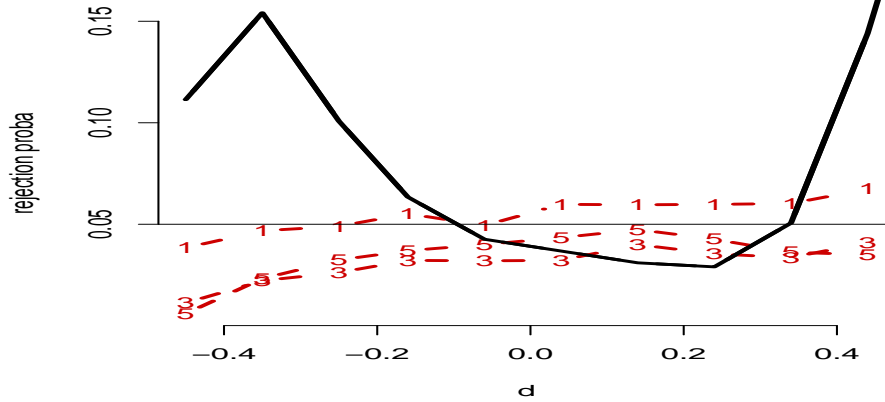


Figure 1: Empirical levels evaluated on FARIMA(0, d ,0) process as function of the memory parameter $d \in (-1/2, 1/2)$. We compare the tests based on the statistics V/S (solid line) and $Q_{n, \sqrt{n}}(s, \hat{d})$ (dashed red lines) with $n = 500$, $s = 1, 3, 5$ and with 5000 replications. The nominal level is $\alpha = 5\%$.

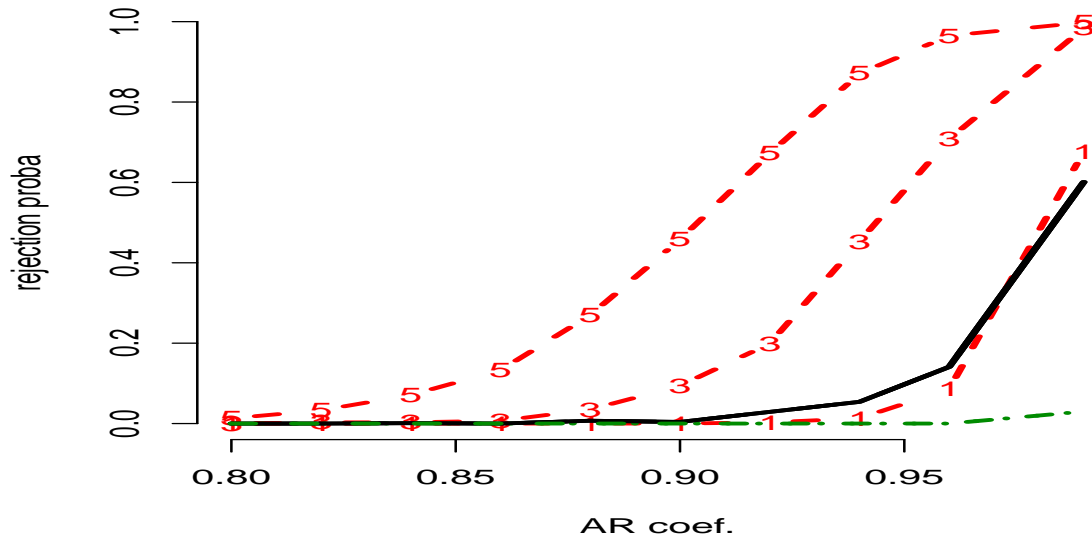


Figure 2: Empirical levels evaluated on AR(1) process $x_t = ax_{t-1} + \epsilon_t$ as function of its coefficient a . We compare the tests based on the statistics V/S (solid line) and $Q_{n, \sqrt{n}}(s, \hat{d})$ (dashed red line) and Dickey-Fuller (dashed green line) with $n = 500$, $s = 1, 3, 5$ and with 5000 replications. The nominal level is $\alpha = 5\%$.

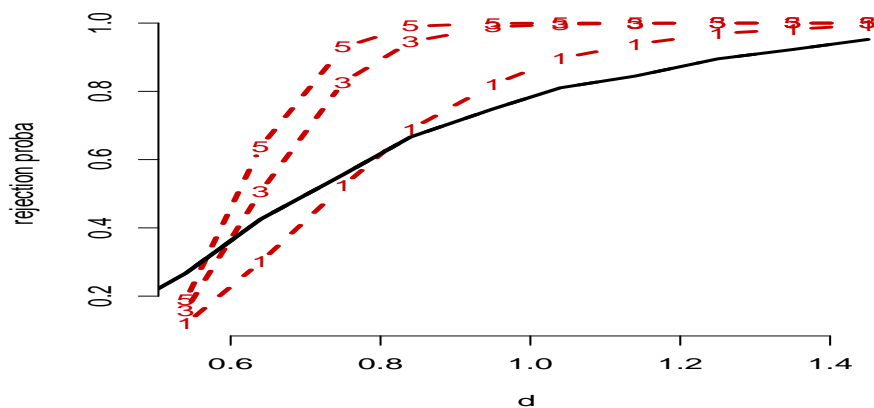


Figure 3: Empirical power as function of d , the parameter of nonstationary fractional processes with $d \in (1/2, 3/2)$. We compare the tests based on statistics V/S (solid line) and $Q_{n, \sqrt{n}}(s, \hat{d})$ (dashed red lines) with $n = 500$, $s = 1, 3, 5$ and with 5000 replications. The nominal level is $\alpha = 5\%$.

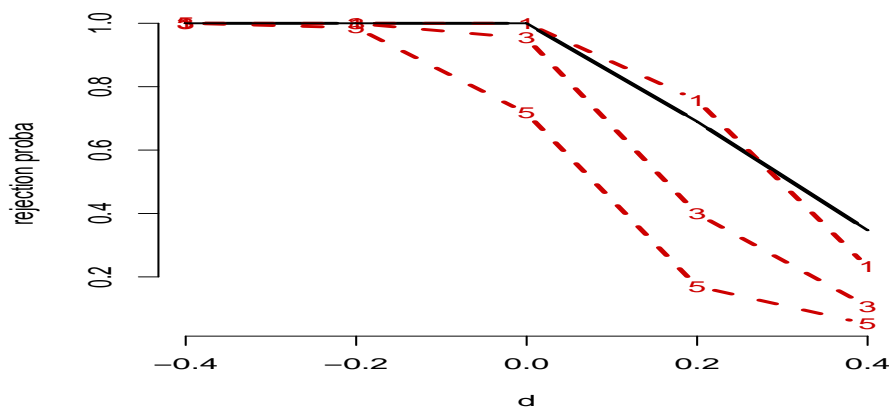


Figure 4: Empirical power as function of d , the memory parameter, for the change point model with $\Delta = 1$ defined in (15). We compare the tests based on statistics V/S (solid line) and $Q_{n, \sqrt{n}}(s, \hat{d})$ (dashed red lines) with $n = 500$, $s = 1, 3, 5$ and with 5000 replications. The nominal level is $\alpha = 5\%$.

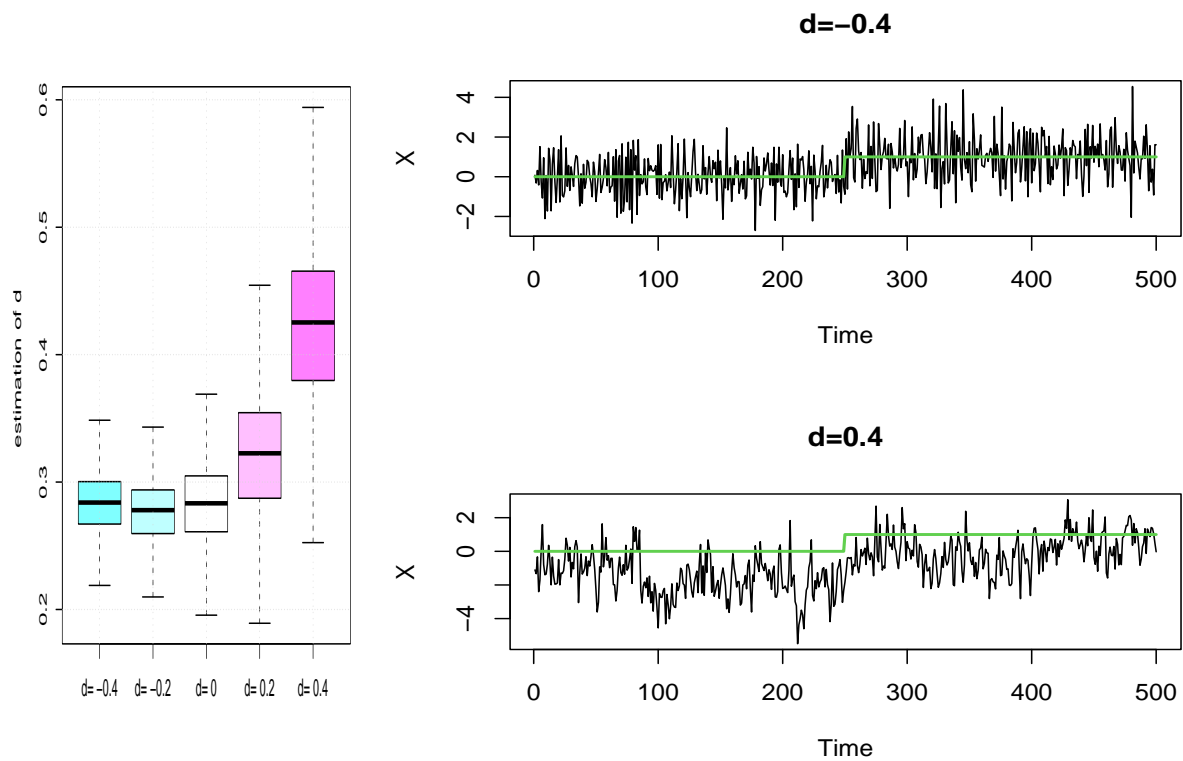


Figure 5: Left: Boxplots of \hat{d} for different values of d , calculated on FARIMA model with change point $\Delta = 1$, with sample size 500 and 5000 replications. Right: Two change point trajectories with $d = \pm 0.4$. The sample size is 500. The green line is the level μ_t defined in (15) with $\Delta = 1$.

5 Appendix

In what follows and for simplicity, c designates a generic positive constant that may be different from one expression to another.

5.1 Proof of Theorem 1

From (2) we get

$$\begin{aligned} Q_{n,m}(s, d) &= m^{-2d} \sum_{j=1}^s \left(\frac{f(\lambda_j)}{f(\lambda'_j)} \right) \frac{I_n(\lambda_j)/f(\lambda_j)}{\frac{1}{m} \sum_{i=1}^m I_{n,i}(\lambda'_j)/f(\lambda'_j)} \\ &\sim \sum_{j=1}^s \frac{I_n(\lambda_j)/f(\lambda_j)}{\frac{1}{m} \sum_{i=1}^m I_{n,i}(\lambda'_j)/f(\lambda'_j)}, \quad \text{as } n \rightarrow \infty. \end{aligned}$$

We know from Hurvich and Beltrão (1993)² and Terrin and Hurvich (1994) that

$$\left(\frac{I_n(\lambda_1)}{f(\lambda_1)}, \dots, \frac{I_n(\lambda_s)}{f(\lambda_s)} \right) \xrightarrow{\mathcal{D}} [L_1(d)(Z_1^2(1) + Z_2^2(1)), \dots, L_s(d)(Z_1^2(s) + Z_2^2(s))],$$

where $Z(s, d) = (Z_1(1), \dots, Z_1(s), Z_2(1), \dots, Z_2(s))$ is a zero-mean Gaussian random vector with covariance matrix $\Sigma(d)$ defined in (8) and (9), and $L_j(d) = 2L_{j,j}(d)$, defined in (10).

Actually, the covariances between $Z_i(j)$ and $Z_i(k)$, $i = 1, 2$ and $j, k = 1, \dots, s$ given in the references above can easily be worked out to obtain a simplified form under the covariance matrix $\Sigma(d)$ expressed only in terms of our coefficients $L_{j,k}(d)$ and $L_{j,-k}(d)$.

Using Slutsky Lemma, it will then be enough to show that, for each fixed j , as $n \rightarrow \infty$ (and therefore as $\ell, m \rightarrow \infty$),

$$\frac{1}{m} \sum_{i=1}^m \frac{I_{n,i}(\lambda'_j)}{f(\lambda'_j)} \xrightarrow{\text{P}} L_j(d), \quad (16)$$

²We note in passing that two errors have slipped into the proof of Theorem 5 of Hurvich and Beltrão (1993) on page 471 of the article leading to a coefficient of $(-1)^{j+k+1}$ in their formulae (7) and (8) while there should be no such a coefficient.

to obtain

$$Q_{n,m}(s, d) \stackrel{\mathcal{D}}{\rightarrow} Z_1^2(1) + Z_2^2(1) + \cdots + Z_1^2(s) + Z_2^2(s). \quad (17)$$

The limiting distribution above is that of the squared norm of $Z(s, d)$ whose covariance matrix can be decomposed as

$$\Sigma(d) = \Omega'(d)\Lambda(d)\Omega(d),$$

where $\Lambda(d)$ is the diagonal matrix of $\Sigma(d)$'s eigenvalues (denoted $\zeta_i(d)$, $i = 1, \dots, 2s$) and $\Omega(d)$ is an orthogonal matrix. Therefore, we can write

$$\begin{aligned} Z_1^2(1) + Z_2^2(1) + \cdots + Z_1^2(s) + Z_2^2(s) &= \|Z(s, d)\|^2 = \|\Omega(d)Z(s, d)\|^2 \\ &\stackrel{\mathcal{D}}{=} \|\Lambda(d)^{1/2}W\|^2 = \sum_{i=1}^{2s} \zeta_i(d)W_i^2, \end{aligned} \quad (18)$$

where $\stackrel{\mathcal{D}}{=}$ means equality in distribution and W is a $(2s \times 1)$ standard Gaussian vector. To complete the proof of the Theorem, we now establish (16). Using the the fact that $I_{n,i}(\lambda'_j)/f(\lambda'_j)$, $i = 1, \dots, m$ are identically distributed, and then Hurvich and Beltrão (1993), we have for fixed j ,

$$\mathbb{E} \left(\frac{1}{m} \sum_{i=1}^m \frac{I_{n,i}(\lambda'_j)}{f(\lambda'_j)} \right) = \mathbb{E} \left(\frac{I_{n,i}(\lambda'_j)}{f(\lambda'_j)} \right) \rightarrow L_j(d).$$

Since $f(\lambda'_j) \sim c\ell^{2d}$, it will then be enough to show that

$$\ell^{-4d} \text{Var} \left(\frac{1}{m} \sum_{i=1}^m I_{n,i}(\lambda'_j) \right) \rightarrow 0.$$

Let us first consider the case $0 \leq d < 1/2$. Similar computations as in Gromykov et al. (2018) show that the LHS of the expression above can be asymptotically bounded by

$$2 \frac{\ell^{-4d}}{m} \sum_{u=1}^m \frac{1}{\ell} \sum_{p=1}^{\ell} \sum_{q=1}^{\ell} \sum_{h=1+q-\ell}^{\ell-p+1} T_{\ell,u}(h, p, q), \quad (19)$$

where

$$T_{\ell,u}(h, p, q) = T_{\ell,u}^{(1)}(h, p, q) + T_{\ell,u}^{(2)}(h, p, q) + T_{\ell,u}^{(3)}(h, p, q),$$

with

$$T_{\ell,u}^{(1)}(h, p, q) = \gamma(\ell u - h) \gamma(\ell u - h + q - p),$$

$$T_{\ell,u}^{(2)}(h, p, q) = \gamma(\ell u - h + q) \gamma(\ell u - h - p),$$

and

$$T_{\ell,u}^{(3)}(h, p, q) = (\eta - 3) \sigma^4 \sum_{i=0}^{\infty} a_i a_{i+p} a_{i+h+\ell u} a_{i+h+\ell u+q}.$$

Let us handle $T_{\ell,u}^{(1)}(h, p, q)$ first: with $k = q - p$ and making u starts from 3,

$$\begin{aligned} & \frac{1}{m\ell^{4d+1}} \left| \sum_{u=3}^m \sum_{p=1}^{\ell} \sum_{q=1}^{\ell} \sum_{h=1-\ell+q}^{\ell-p} \gamma(\ell u - h) \gamma(\ell u - h - p + q) \right| \\ &= \frac{1}{m\ell^{4d+1}} \left| \sum_{u=3}^m \sum_{k=-\ell+1}^{\ell-1} \sum_{p=1}^{\ell-k} \sum_{h=1-\ell+k+p}^{\ell-p} \gamma(\ell u - h) \gamma(\ell u - h + k) \right| \\ &\leq \frac{1}{m\ell^{4d+1}} \sum_{u=3}^m \sum_{k=-\ell+1}^{\ell-1} \sum_{p=1}^{\ell-k} \sum_{h=-\ell}^{\ell} |\gamma(\ell u - h) \gamma(\ell u - h + k)| \\ &\leq \frac{1}{m\ell^{4d}} \sum_{u=3}^m \sum_{k=-\ell+1}^{\ell-1} \sum_{h=-\ell}^{\ell} |\gamma(\ell u - h) \gamma(\ell u - h + k)| \\ &= \frac{1}{m\ell^{4d}} \sum_{u=3}^m \sum_{k=-\ell+1}^{\ell-1} \sum_{s=\ell(u-1)}^{\ell u} |\gamma(s) \gamma(s + k)| \\ &= \frac{1}{m\ell^{4d}} \sum_{s=2\ell}^{m\ell} \sum_{k=-\ell+1}^{\ell-1} |\gamma(s) \gamma(s + k)|. \end{aligned} \quad (20)$$

When $d = 0$ then $\gamma(h)$ is summable and hence (20) is $O(1/m)$. When $0 < d < 1/2$ then $\gamma(h) = O(h^{2d-1})$ and hence (20) is bounded by

$$c \frac{1}{m} \int_2^m \left(\int_{-1}^1 (x+y)^{2d-1} dy \right) x^{2d-1} dx \leq \frac{2c}{m} \int_2^m x^{2d-1} dx = O(m^{2d-1}) \rightarrow 0.$$

$T_{\ell,u}^{(2)}(h,p,q)$ treats the same way. For the last term $T_{\ell,u}^{(3)}(h,p,q)$, we have

$$\begin{aligned} \left| \sum_{h=-\ell}^{\ell} T_{\ell,u}^{(3)}(h,p,q) \right| &= |(\eta-3)|\sigma^4 \left| \sum_{i=0}^{\infty} a_i a_{i+p} \right| \left| \sum_{s=i+(u-1)\ell}^{i+(u+1)\ell} a_s a_{s+q} \right| \\ &\leq |(\eta-3)|\sigma^4 |\gamma(p)| \sum_{s=\ell}^{\infty} |a_s a_{s+q}| \\ &\leq |(\eta-3)|\sigma^4 |\gamma(p)| \left(\sum_{s=\ell}^{\infty} a_s^2 \right)^{1/2} \left(\sum_{s=0}^{\infty} a_s^2 \right)^{1/2} = |\gamma(p)| |o(1)| \rightarrow 0, \text{ as } \ell \rightarrow \infty, \end{aligned} \quad (21)$$

where $o(1)$ is uniform in u, q and p . Therefore

$$\ell^{-4d} \frac{1}{m} \frac{1}{\ell} \sum_{u=1}^m \sum_{p=1}^{\ell} \sum_{q=1}^{\ell} \sum_{h=-\ell}^{\ell} T_{\ell,u}^{(3)}(h,p,q) = \ell^{-4d} \left(\sum_{p=1}^{\ell} |\gamma(p)| \right) \frac{o(\ell)}{\ell} = \ell^{-2d} \frac{o(\ell)}{\ell} \rightarrow 0, \text{ as } \ell \rightarrow \infty.$$

Let us consider the case $-1/2 < d < 0$. Noticing that quantity in (20) is bounded by

$$\frac{1}{m\ell^{4\ell}} \left(\sum_{h=\ell}^{\infty} |\gamma(h)| \right)^2 = O(1/m) \rightarrow 0$$

since under condition \mathcal{H}_d , we have $|\gamma(h)| \sim C(d)h^{-1+2d}$ as $h \rightarrow \infty$ and therefore, the result obtained above for $T_{\ell,u}^{(1)}(h,p,q)$ in case $0 < d < 1/2$ still holds when $-1/2 < d < 0$. Same thing remains true for $T_{\ell,u}^{(2)}(h,p,q)$. Concerning $T_{\ell,u}^{(3)}(h,p,q)$, using (21) and the summability

of $\gamma(p)$ as well as the expression of a_j given in \mathcal{H}_d , we obtain that

$$\begin{aligned} \ell^{-4d} \frac{1}{m} \frac{1}{\ell} \sum_{u=1}^m \sum_{p=1}^{\ell} \sum_{q=1}^{\ell} \sum_{h=-\ell}^{\ell} |T_{\ell,u}^{(3)}(h, p, q)| &\leq c \ell^{-4d-1} \sum_{s=\ell}^{\infty} |a_s| \sum_{q=1}^{\infty} |a_{s+q}| \\ &\leq c \ell^{-4d-1} \sum_{s=\ell}^{\infty} s^{-1+2d} = c \ell^{-1-2d} \rightarrow 0 \text{ as } \ell \rightarrow \infty. \end{aligned}$$

The proof of Theorem 1 is now complete.

5.2 Proof of Proposition 1

The first step consists in establishing the following properties for the entries of $\Sigma(d)$ (which are proven in the appendix)

1. For all fixed $j, k = 1, \dots, s$, $L_{j,k}(d)$ (defined in (10)) are well defined for $d = 1/2$.
Actually all the singularities of the integrands in the previous formula are removable and the integrands are asymptotically equivalent to $K(d)\lambda^{-2-2d}$ for some positive constant $K(d)$ (finite for all $d \leq 1/2$). Hence the integrals are finite when $d = 1/2$ (and even for all $d > -1/2$).
2. As functions of d , $L_{j,k}(d)$ are continuous at $d = 1/2$. Indeed,

$$L_{j,k}(d) = K(d) \int_{-\infty}^{\infty} h(\lambda, d) d\lambda$$

where

$$h(\lambda, d) = \frac{\sin^2(\lambda/2)}{(2\pi k - \lambda)(2\pi j - \lambda)} |\lambda|^{-2d}$$

is clearly continuous at $d = 1/2$ and

$$h(\lambda, d) \leq \begin{cases} h(\lambda, 1/2), & \text{if } |\lambda| \leq 1, \\ \frac{\sin^2(\lambda/2)}{(2\pi k - \lambda)(2\pi j - \lambda)}, & \text{if } |\lambda| \geq 1. \end{cases}$$

That is $h(\lambda, d)$ is uniformly bounded (in d) by an integrable function and hence $L_{j,k}(d) \rightarrow L_{j,k}(1/2)$.

Then these two points imply that $\Sigma(1/2)$ is well defined and a covariance matrix. Moreover $\Sigma(d)$ converges to $\Sigma(1/2)$ when $d \rightarrow 1/2$. For $-1/2 < d \leq 1/2$, define $Z(s, d)$ to be a zero-mean Gaussian vector with covariance matrix $\Sigma(d)$. Then the convergence of the covariance matrices ensures that $Z(s, d) \xrightarrow{\mathcal{D}} Z(s, 1/2)$ as $d \rightarrow 1/2$ and hence $Q(s, d) \xrightarrow{\mathcal{D}} Q(s, 1/2)$.

Writing $Z(s, 1/2) = (Z_1(1), Z_2(1), \dots, Z_1(s), Z_2(s))'$, we have $\text{Var}(Z_1(j)) + \text{Var}(Z_2(j)) = 1$, and therefore $Z(s, 1/2)$ has at least s non-degenerated components. This ensures that the covariance matrix $\Sigma(1/2)$ has at least s positive eigenvalues. This guarantees that $Q(s, 1/2)$ is a nondegenerated random variable.

5.3 Proof of Proposition 2

Let s be fixed and let d_n be an arbitrary sequence in $(-1/2, 1/2]$ converging to $d \in (-1/2, 1/2]$. Using the representation (18) and the continuity of the eigenvalues $\zeta_i(d)$, then $Q(s, d_n)$ converges in distribution to $Q(s, d)$ (the proof is identical to that of Proposition 1). Since $Q(s, d)$ has a continuous distribution function $F_{s,d}$, by Polya's theorem,

$$\|F_{s,d_n} - F_{s,d}\| \rightarrow 0.$$

Moreover, the invertibility of $F_{s,d}$ implies that (with $F = F_{s,d}$ and $F_n = F_{s,d_n}$)

$$F(F_n^{-1}(\alpha)) = |F(F_n^{-1}(\alpha)) - F_n(F_n^{-1}(\alpha)) + F_n(F_n^{-1}(\alpha))| \leq \|F_n - F\| + \alpha \rightarrow \alpha$$

which means that $F_n^{-1}(\alpha) \rightarrow F^{-1}(\alpha)$ since F^{-1} is continuous. This completes the proof.

5.4 Proof of Theorem 2

As $\delta \leq 1/2$ appears only in the normalization $m^{-2\delta}$ of the statistic $Q_{n,m}(s, \delta)$, it is enough to prove the convergence (13) for $\delta = 1/2$, i.e. $Q_{n,m}(s, 1/2) \xrightarrow{P} \infty$.

For a stochastic or deterministic trend, we want to show that

$$(m^{-1}) \sum_{j=1}^s \frac{I_n(\lambda_j)}{\frac{1}{m} \sum_{i=1}^m I_{n,i}(\lambda'_j)} \xrightarrow{P} \infty.$$

It will be enough to show that for fixed j ,

$$(m^{-1}) \frac{I_n(\lambda_j)}{\frac{1}{m} \sum_{i=1}^m I_{n,i}(\lambda'_j)} \xrightarrow{P} \infty. \quad (22)$$

First, consider the deterministic trend case: $X_t = g_n(t) + Y_t$. For fixed $j = 1, \dots, s$, and omitting the coefficient 2π in the denominator, we have from (3), with $I_{n,Y}$ denoting the periodogram built from Y_1, \dots, Y_n ,

$$I_n(\lambda_j) = I_{n,Y}(\lambda_j) + n^{2\beta+1} |D_n(g)|^2 + 2n^{\beta+1/2} \operatorname{Re} \left(D_n(g) \overline{D_n(Y)} \right) \quad (23)$$

where

$$D_n(g) = \frac{1}{\sqrt{2\pi}} \sum_{t=1}^n g\left(\frac{t}{n}\right) e^{i2\pi t j/n} \frac{1}{n},$$

and

$$D_n(Y) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^n Y_t e^{i2\pi t j/n}.$$

Since Y satisfies Theorem 1's assumptions, we have as $n \rightarrow \infty$,

$$\mathbb{E}(I_{n,Y}(\lambda_j)) \sim L_j(d)f(\lambda_j) \sim L_j(d)n^{2d}.$$

Under assumption (12), we get by Riemann sum approximation of an integral,

$$|D_n(g)|^2 \sim c > 0.$$

Finally, for the cross-term, we have

$$\mathbb{E}|D_n(g)D_n(Y)| = |D_n(g)| \mathbb{E}|D_n(Y)| \leq |D_n(g)| \mathbb{E}^{1/2}(I_{n,Y}(\lambda_j)) \leq cn^d. \quad (24)$$

Therefore

$$\frac{I_n(\lambda_j)}{n^{2\beta+1}} \xrightarrow{P} c. \quad (25)$$

We now study the behaviour of the term

$$\frac{1}{m} \sum_{i=1}^m I_{n,i}(\lambda'_j) = \frac{1}{m} \sum_{i=1}^m \frac{1}{2\pi\ell} \left| \sum_{t=(i-1)\ell+1}^{i\ell} (Y_t + g_n(t)) e^{it\lambda'_j} \right|^2.$$

Following the same decomposition as in (23), we will have a full stochastic term, a trend and a cross term to examine.

Denote $I_{n,i,Y}$ the periodogram built from the i th epoch of the stationary process Y_t . Using (16), we then obtain

$$\frac{1}{m} \sum_{i=1}^m I_{n,i,Y}(\lambda'_j) = O_P(\ell^{2d}).$$

For the trend part, we have

$$\begin{aligned}
& \frac{1}{m} \sum_{u=1}^m \frac{1}{\ell} \left| \sum_{t=1}^{\ell} g_n(t + (u-1)\ell) e^{it\lambda'_j} \right|^2 \\
&= \ell^{-1} n^{2\beta} \sum_{t=1}^{\ell} \sum_{s=1}^{\ell} \left(\sum_{u=1}^m g \left(\frac{t + (u-1)\ell}{n} \right) g \left(\frac{s + (u-1)\ell}{n} \right) \frac{1}{m} \right) e^{i(t-s)\lambda'_j} \\
&\sim \ell^{-1} n^{2\beta} \sum_{t=1}^{\ell} \sum_{s=1}^{\ell} \left[\int_0^1 g \left(\frac{t}{n} + x \right) g \left(\frac{s}{n} + x \right) dx \right] e^{i(t-s)\lambda'_j} \\
&= \ell^{-1} n^{2\beta} \int_0^1 \left(\left| \sum_{t=1}^{\ell} g \left(\frac{t}{n} + x \right) e^{it\lambda'_j} \right|^2 \right) dx. \tag{26}
\end{aligned}$$

(In the integrals above we extended g beyond $(0,1)$ with value 0). If g is continuous over $[0,1]$ then it is uniformly continuous and hence

$$g \left(\frac{t}{n} + x \right) = g(x) + \epsilon_n(t, x)$$

where $\epsilon_n(t, x) \rightarrow 0$ uniformly in $t = 1, \dots, \ell$ and x , and hence, since

$$\sum_{t=1}^{\ell} e^{it\lambda'_j} = 0,$$

the last expression in (26) is bounded by $n^{2\beta} o(\ell)$.

If g is piece-wise continuous on $[0,1]$ then there exists M sub-intervals (A_k, B_k) , $j = 1, \dots, M$, such that g is continuous on such intervals with finite limits at A_k and B_k .

The previous result (when g is continuous) remains true for each interval $[A_k, B_k]$ Hence we find that (26) is still bounded by $M n^{2\beta} o(\ell) = n^{2\beta} o(\ell)$.

For the cross-terms and similarly to (24), we easily obtain that

$$\begin{aligned}
& \mathbb{E} \left(\frac{1}{m} \sum_{u=1}^m \frac{1}{\ell} \sum_{t=1}^{\ell} \left| g_n(t + (u-1)\ell) e^{it\lambda'_j} \right| \left| \sum_{t=1}^{\ell} Y_{t+(u-1)\ell} e^{it\lambda'_j} \right| \right) \\
&= \frac{1}{n} \ell^{1/2} \sum_{u=1}^m \sum_{t=1}^{\ell} \left| g_n(t + (u-1)\ell) e^{it\lambda'_j} \right| \mathbb{E} \left[\frac{1}{\sqrt{\ell}} \left| \sum_{t=1}^{\ell} Y_{t+(u-1)\ell} e^{it\lambda'_j} \right| \right] \\
&= \ell^{1/2} \left(\sum_{u=1}^m \sum_{t=1}^{\ell} \left| g_n(t + (u-1)\ell) e^{it\lambda'_j} \right| \frac{1}{n} \right) \mathbb{E} \left[\frac{1}{\sqrt{\ell}} \left| \sum_{t=1}^{\ell} Y_t e^{it\lambda'_j} \right| \right] \\
&\leq \ell^{1/2} n^{\beta} \sum_{k=1}^n \left| g\left(\frac{k}{n}\right) \right| \frac{1}{n} (\mathbb{E}(I_{n,1,Y}(\lambda'_j)))^{1/2} \\
&\sim n^{\beta} \ell^{1/2} \int_0^1 |g(x)| dx [L_j(d)f(\lambda'_j)]^{1/2} \\
&= cn^{\beta} \ell^{1/2+d}.
\end{aligned}$$

Hence in total, the denominator is of order $\ell n^{2\beta} o_P(1)$. Therefore, using (25), for each fixed frequency λ_j , the ratio statistic

$$\left(\frac{n}{\ell}\right)^{-1} \frac{I_n(\lambda_j)}{\frac{1}{m} \sum_{i=1}^m I_{n,i}(\lambda'_j)} = \frac{\ell}{n} \frac{n^{2\beta+1}}{n^{2\beta} \ell o_P(1)} = \frac{1}{o_P(1)} \xrightarrow{P} \infty,$$

which finally implies that $Q_{n,m}(s, 1/2)$ goes to infinity as $n \rightarrow \infty$.

We now consider a stochastic trend $X_t = X_{t-1} + Y_t$ where Y_t has density of the form (1). Let λ_j be a fixed Fourier frequency. We begin with the numerator in (22). With $S_t = Y_1 + \dots + Y_t$, we write $X_t = X_0 + S_t$ and using the well known summation by parts formula

$$\sum_{t=1}^n U_t v_t = U_n V_n - \sum_{t=1}^n u_t V_t, \quad \text{where } U_k = u_1 + \dots + u_k, \quad V_k = v_1 + \dots + v_k,$$

and the fact that, $j = 1, 2, \dots, s$,

$$\sum_{t=1}^n e^{2\pi itj/n} = 0,$$

we obtain

$$\begin{aligned} I_n(\lambda_j) &= \frac{1}{2\pi n} \left| \sum_{t=1}^n (X_0 + S_t) e^{it\lambda_j} \right|^2 = \frac{1}{2\pi n} \left| \sum_{t=1}^n S_t e^{it\lambda_j} \right|^2 \\ &= \frac{1}{2\pi n} \left| \sum_{t=1}^n Y_t \left(\frac{1 - e^{it\lambda_j}}{1 - e^{i\lambda_j}} \right) \right|^2 \\ &\sim \frac{n^{2+2d}}{(2\pi)^3(j)^2} \left| \frac{1}{n^{1/2+d}} \sum_{t=1}^n (1 - e^{it\lambda_j}) Y_t \right|^2 \geq \frac{n^{2+2d}}{(2\pi)^3(j)^2} \left(\frac{1}{n^{1/2+d}} \sum_{t=1}^n \sin(t\lambda_j) Y_t \right)^2. \end{aligned}$$

According to formula (5) in Theorem 4 of Hurvich and Beltrão (1993), the last term can be evaluated as follows:

$$\begin{aligned} &\text{Var} \left(\frac{1}{n^{1/2+d}} \sum_{t=1}^n \sin(t\lambda_j) Y_t \right) \\ &= \frac{2\pi f(\lambda_j)}{n^{2d}} \text{Var} \left(\frac{1}{(2\pi n f(\lambda_j))^{1/2}} \sum_{t=1}^n \sin(t\lambda_j) Y_t \right) \sim [(2\pi)^{1-2d} j^{-2d} f^*(0)] \frac{1}{2} (L_{j,j}(d) - L_{j,-j}(d)) \end{aligned}$$

where $L_{j,\pm j}(d)$ is defined in (10).

Straightforward computations show that $L_{j,j}(d) - L_{j,-j}(d) > 0$ and therefore $I_n(\lambda_j)$ is at least of order n^{2+2d} . Now let us handle the denominator in (22). Similarly to the numerator, we can easily write, from the summation by parts formula,

$$\frac{1}{m} \sum_{u=1}^m |I_{n,u}(\lambda'_j)| = \frac{1}{m} \sum_{u=1}^m \frac{1}{2\pi\ell} \left| \sum_{t=1}^{\ell} S_{t+(u-1)\ell} e^{it\lambda'_j} \right|^2 = \frac{1}{m} \sum_{u=1}^m \frac{1}{2\pi\ell} |S_{\ell,u} - W_{\ell,u}(\lambda'_j)|^2 (|1 - e^{i\lambda'_j}|^{-2}),$$

where

$$S_{\ell,u} = \sum_{t=1}^{\ell} Y_{t+(u-1)\ell}, \quad \text{and } W_{\ell,u}(\lambda'_j) = \sum_{t=1}^{\ell} Y_{t+(u-1)\ell} e^{it\lambda'_j}.$$

Using the stationarity of $S_{\ell,u}$, $u = 1, \dots, m$ and the fact that $\mathbb{E}(S_{\ell,1}^2) \sim c\ell^{2d+1}$ (see Proposition 3.3.1 of Giraitis et al. (2012)) and that $\mathbb{E}\left(|W_{\ell,u}(\lambda'_j)|^2\right) \sim c\ell^{1+2d}$ (see Theorem 1 of Hurvich and Beltrão (1993)),

$$\frac{1}{m} \sum_{u=1}^m |I_{n,u}(\lambda'_j)| \leq \frac{2}{2\pi n} \sum_{u=1}^m (S_{\ell,u}^2 + |W_{\ell,u}(\lambda'_j)|^2) O(\ell^2) = O(\ell^2) \frac{m}{n} O_P(\ell^{2d+1}) = O_P(\ell^{2d+2}).$$

Finally, combining the results on the numerator and denominator, we obtain that asymptotically

$$\left(\frac{n}{\ell}\right)^{-1} \frac{I_n(\lambda_j)}{\frac{1}{m} \sum_{u=1}^m I_{n,u}(\lambda'_j)} = \frac{n^{2+2d-1}}{O_P(\ell^{2+2d-1})} = \frac{m^{1+2d}}{O_P(1)} \xrightarrow{P} \infty, \quad \text{as } n \rightarrow \infty. \quad \square$$

References

- Abadir, K. M., Distaso, W., and Giraitis, L. (2007). Nonstationarity-extended local whittle estimation. *Journal of Econometrics*, 141(2):1353–1384.
- Bailey, N. and Giraitis, L. (2016). Spectral approach to parameter-free unit root testing. *Comput. Statist. Data Anal.*, 100:4–16.
- Beran, J., Feng, Y., Ghosh, S., and Kulik, R. (2013). *Long-memory processes*. Springer, Heidelberg. Probabilistic properties and statistical methods.
- Giraitis, L., Koul, H. L., and Surgailis, D. (2012). *Large sample inference for long memory processes*. Imperial College Press, London.
- Giraitis, L., Leipus, R., and Philippe, A. (2006). A test for stationarity versus trends and unit roots for a wide class of dependent errors. *Econometric Theory*, 22(6):989–1029.
- Gromykov, G., Ould Haye, M., and Philippe, A. (2018). A frequency-domain test for long range dependence. *Statistical Inference for Stochastic Processes*, 21(3):513–526.

- Hu, C., Pozdnyakov, V., and Yan, J. (2020). Density and distribution evaluation for convolution of independent gamma variables. *Computational Statistics*, 35(1):327–342.
- Hurvich, C. M. and Beltrão, K. I. (1993). Asymptotics for the low-frequency ordinates of the periodogram of a long-memory time series. *J. Time Ser. Anal.*, 14(5):455–472.
- Phillips, P. C. B. and Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75:335–346.
- Taqqu, M. (2003). Fractional brownian motion and long-range dependence. *In Theory and Applications of Long-Range Dependence*, pages 5–38. Eds Doukhan, P. and Oppenheim, G. and Taqqu, M. Birkhauser, Boston.
- Terrin, N. and Hurvich, C. M. (1994). An asymptotic Wiener-Itô representation for the low frequency ordinates of the periodogram of a long memory time series. *Stochastic Process. Appl.*, 54(2):297–307.