Spectral estimation for non-linear long range dependent discrete time trawl processes

Paul Doukhan, François Roueff, Joseph Rynkiewicz

Spectral estimation for non-linear long range dependent discrete time trawl processes

P. Doukhan[*], F. Roueff[†], and J. Rynkiewicz[‡]

January 6, 2020

Abstract

Discrete time trawl processes constitute a large class of time series parameterized by a trawl sequence $(a_j)_{j\in\mathbb{N}}$ and defined though a sequence of independent and identically distributed (i.i.d.) copies of a continuous time process $(\gamma(t))_{t\in\mathbb{R}}$ called the seed process. They provide a general framework for modeling linear or non-linear long range dependent time series. We investigate the spectral estimation, either pointwise or broadband, of long range dependent discrete-time trawl processes. The difficulty arising from the variety of seed processes and of trawl sequences is twofold. First, the spectral density may take different forms, often including smooth additive correction terms. Second, trawl processes with similar spectral densities may exhibit very different statistical behaviors. We prove the consistency of our estimators under very general conditions and we show that a wide class of trawl processes satisfy them. This is done in particular by introducing a weighted weak dependence index that can be of independent interest. The broadband spectral estimator includes an estimator of the long memory parameter. We complete this work with numerical experiments to evaluate the finite sample size performance of this estimator for various integer valued discrete time trawl processes.

Keywords: trawl processes; integer-valued time series; long memory parameter estimation
MSC: 62M10; 62F12; 60G51;

1 Introduction

A *discrete time trawl* process $X = \{X_k,\, k \in \mathbb{Z}\}$ is defined in [Doukhan et al., 2019] by

$$X_k \;=\; \sum_{j=0}^{\infty} \gamma_{k-j}(a_j), \qquad k \in \mathbb{Z}, \tag{1.1}$$

[*]University Cergy-Pontoise, AGM UMR8088 and CIMFAV, University of Valparaiso.
[†]LTCI, Telecom Paris, Institut Polytechnique de Paris
[‡]University Paris 1, la Sorbonne, SAMM EA4543.

where

(A-1) The sequence $(\gamma_k)_{k\in\mathbb{Z}}$ is a sequence of i.i.d. copies of a generic process $\gamma = \{\gamma(u), u \in \mathbb{R}\}$ and $a = \{a_j,\, j \geq 0\}$ is a sequence converging to zero.

Processes so defined can be interpreted as discrete time versions of the trawl processes introduced in [Barndorff-Nielsen et al., 2014]. The generic process γ is called the *seed process* and the sequence a is called the *trawl (height) sequence*.

Additional assumptions are required to have a converging sum in (1.1). The convergence in $\mathbb{L}^2$ is guaranteed if

$$\sum_{j=0}^{\infty} (|\mathbb{E}\,\gamma(a_j)| + \operatorname{Var}\gamma(a_j)) < \infty\,. \tag{1.2}$$

See [Doukhan et al., 2019, Proposition 1], where the covariance function is also given by the formula

$$r(k) = \operatorname{Cov}(X_0, X_k) = \sum_{j=0}^{\infty} \operatorname{Cov}(\gamma(a_j), \gamma(a_{j+k}))\,. \tag{1.3}$$

By [Doukhan et al., 2019, Proposition 3], we moreover know that if, in addition, the two following asymptotic behaviors hold:

$$\operatorname{Cov}(\gamma(u), \gamma(v)) = (|u| \wedge |v|)\ (1 + o(1)) \quad \text{as} \quad u, v \to 0\,, \tag{1.4}$$

$$a_j \ = \ c\, j^{-\alpha^*}(1 + o(1)) \quad \text{as} \quad j \to \infty\,, \tag{1.5}$$

with $c \neq 0$ and $\alpha^* > 1$, then, the covariance function behaves at large lags as

$$r(k) = c'\ k^{1-\alpha^*}\ (1 + o(1)) \quad \text{as} \quad k \to \infty\,. \tag{1.6}$$

In the following we will refer to α^* in (1.5) as the *trawl exponent*. In particular if

$$1 < \alpha^* < 2\,, \tag{1.7}$$

this behavior is often referred to as X being *long range dependent* with long memory parameter

$$d^* = 1 - \alpha^*/2\,. \tag{1.8}$$

Here (1.7) implies $d^* \in (0, 1/2)$ (sometimes referred to as *positive* long memory). We here use one of the several existing definitions of long range dependence, see for instance Condition II in [Pipiras and Taqqu, 2017, Section 2.1]. In fact in the cases considered here, the same long memory parameter d^* can also be defined through their condition IV, based on the spectral density. In the case where $\alpha^* \geq 2$, the two definitions may no longer coincide. The definition of *negative* long memory ($d^* < 0$) is generally relying on the behavior of the spectral density

at the origin (in particular imposing this spectral density to vanish there). Adopting this definition the formula (1.8) may not be valid anymore as the obtained process could have short memory ($d^* = 0$) even if $\alpha^* > 2$, or have negative long memory ($d^* < 0$). In the following, we will only consider the case where $d^* \geq 0$, avoiding the negative long memory case for convenience.

A very interesting feature of trawl processes is that under the fairly general assumption (1.4) on the seed process, the low frequency behavior of the spectral density is mainly driven by the trawl sequence. However, it is shown in [Doukhan et al., 2019] that, for a given trawl sequence, two different seed processes can yield different large scale behaviors, as can be seen by different types of limits in the invariance principle. In the case of a Lévy seed for instance, a Brownian seed process leads to an invariance principle with fractional Brownian motion limit, with Hurst parameter $(3 - \alpha^*)/2$, and a (centered) Poisson seed process leads to an invariance principle with Lévy α^*-stable limit, see [Doukhan et al., 2019, Theorems 1 and 2].

The goal of this paper is to investigate the spectral estimation of a long-range dependent process X from a sample $X_1, \ldots, X_n$. Deriving general results applying to a wide class of long range dependent trawl processes raise two major difficulties. First, as already noted about the asymptotic results derived in [Doukhan et al., 2019], the large scale behavior of such processes, can be very different from one trawl process to another, even with similar or even equal covariance structure. Second, the spectral density has a closed form only in particular cases for the seed process and the trawl sequence. The computation of the spectral density function depends both on the seed process γ and the sequence (a_j). For instance, in [Doukhan et al., 2019, Example 5], it is shown that for a large class of seed processes (that will be referred to as the Lévy seed process below), a specific sequence (a_j) leads to the same spectral density as an ARFIMA(0,d^*,0), namely,

$$f_{d^*}(\lambda) = \frac{1}{2\pi} \left|1 - \mathrm{e}^{-\mathrm{i}\lambda}\right|^{-2d^*} . \tag{1.9}$$

Here the spectral density is normalized in such a way to have the innovation process with unit variance. The general form that we will assume on the spectral density includes of course a multiplicative constant c^* but also an additive smooth function h^* belonging to the space $\mathcal{C}$ of continuous and (2π) periodic functions endowed with the sup norm. Namely, to encompass as many cases as possible, we assume that X has a spectral density function given by

$$f(\lambda) = c^* \left(f_{d^*}(\lambda) + h^*(\lambda)\right) , \qquad \lambda \in \mathbb{R} , \tag{1.10}$$

where $d^* \in [0, 1/2)$, $c^* > 0$ and $h^* \in \mathcal{C}$. The form (1.10) is the spectral behavior corresponding to that of the covariance in (1.6). Here d^* is again the long memory parameter, and it

characterizes the power law behavior of f at low frequencies while the function h^* encompasses the short-range behavior. As we will see, in many cases of interesting trawl processes, the function h^* is smooth in the Hölder sense, leading naturally to the additive parametric form (1.10) of the spectral density, which is different for the usual product parametric form usually encountered in linear models such as ARFIMA processes. Note however that such an additive form of the spectral density were already considered in [Hurvich et al., 2005] for completely different (non-linear) models.

We consider either pointwise or broadband estimation of the spectral density. In the first case, we estimate $f(\lambda)$ directly for a given λ, and, in the second case, we estimate the triplet (c^*, d^*, h^*) by assuming it belongs to a known parameter set. The first approach only makes sense for $\lambda \neq 0$ and will be investigated in Section 2.2 using a smoothed version of the periodogram. The second approach will be investigated in Section 2.3. The estimation of the long memory parameter is a widely studied problem in statistical inference, see the reference book [Doukhan et al., 2002], or, more recently, [Giraitis et al., 2012] and the references therein. Here, we propose to estimate the parameter (c^*, d^*, h^*) using a parametric Whittle approach. Define the periodogram

$$I_n(\lambda) = \frac{1}{2\pi n} \left| \sum_{k=1}^{n} (X_k - \bar{X}_n) \mathrm{e}^{-\mathrm{i}\lambda k} \right|^2 , \tag{1.11}$$

where $\bar{X}_n$ denotes the empirical mean of the sample $X_1, \ldots, X_n$, and denote the Whittle contrast by

$$\Lambda_n(d, h) = \ln \left(\int \frac{I_n}{f_d + h} \, \mathrm{dL} \right) + \int \ln (f_d + h) \, \mathrm{dL} , \tag{1.12}$$

where f_d is defined by (1.9) and L is the Lebesgue measure on $[-\pi, \pi]$ divided by 2π. Our estimator $(\hat{d}_n, \hat{h}_n, \hat{c}_n)$ is to find a near minimizer $(\hat{d}_n, \hat{h}_n)$ of $(d, h) \mapsto \Lambda_n(d, h)$ over a well chosen set of parameters for (d, h), and then set

$$\hat{c}_n = \int \frac{I_n}{f_{\hat{d}_n} + \hat{h}_n} \, \mathrm{dL} . \tag{1.13}$$

From which we can also define an estimator of the spectral density, namely,

$$\hat{f}_n = \hat{c}_n \left(f_{\hat{d}_n} + \hat{h}_n \right) .$$

Here we derive results that apply to a wide class of trawl processes, in particular to those of nature quite different from the well studied class of Gaussian or linear processes. For convenience, we focus on proving the consistency of our estimators under very general assumptions, that can be of interest beyond trawl processes:

(A-2) The process $X = (X_k)_{k \in \mathbb{Z}}$ is stationary, ergodic and L^2.

(A-3) There exist $C_0 > 0$ and $s_0 \in (0,1)$ such that, for all integers $t_1 \leq t_2 \leq t_3 \leq t_4$,

$$|\mathrm{Cov}(X_{t_1}\,,\, X_{t_2})| \leq C_0\,(1+t_2-t_1)^{-s_0}\,, \tag{1.14}$$

$$|\mathrm{Cov}(X_{t_1}X_{t_2}\,,\, X_{t_3}X_{t_4})| \leq C_1\,(1+t_3-t_2)^{-s_0}\,, \tag{1.15}$$

$$|\mathrm{Cov}(X_{t_1}X_{t_2}X_{t_3}\,,\, X_{t_4})| \leq C_1\,(1+t_4-t_3)^{-s_0}\,. \tag{1.16}$$

Assumption (A-2) is basically satisfied by all well defined discrete-time trawl processes. To show that a given trawl process satisfies (A-3) with a well chosen exponent s_0, we will rely on a weighted weak dependence property that is easy to prove for discrete-time trawl processes.

The paper is organized as follows. In Section 2, we present successively: 1) general conditions on the seed process and the trawl sequence so that the corresponding trawl process satisfies Condition (A-2) and (A-3) above, 2) general results on second order estimation under Assumption (A-3) and 3) a general consistency result on the parametric Whittle estimation of the parameters (d^*, h^*) of the unknown spectral density in (1.9). For this estimation result to hold, we only require on the observed process to satisfy (A-2). The assumption on the parameter set on which the Whittle contrast is maximized will be detailed in (A-4). We provide in Section 3 various examples of trawl processes. Although the usual causal linear models for long range dependence (such as ARFIMA processes) constitute specific examples of trawl processes, we here focus on the non-linear models introduced in [Doukhan et al., 2019], and specify simple sufficient conditions implying the assumptions used in the general results. The proofs of the results presented in Sections 2 and Section 3 are detailed in Section 5. Before that, we introduce in Section 4 some weighted weak dependence coefficients that can be of independent interest but which will mainly serve us here to check (A-3) for trawl processes. Finally in Section 6, we present numerical experiments focusing on the estimation of the long memory parameter d^* comparing our approach to the more classical local Whittle estimator, which is known to perform well for standard linear models. Concluding remarks including directions for future work are proposed in Section 7.

2 Main results

2.1 Results on trawl processes

As explained in the introduction, the L^2 convergence of (1.1) follows from (1.2). We provide hereafter a more precise statement, and a slight extension to a convergence in $\mathbb{L}^{2p}$ with $p \geq 1$. All the proofs of this section are postponed to Section 5.1.

Lemma 1. *Assume (A-1). Then (1.2) implies that the convergence (1.1) holds in* $\mathbb{L}^2$ *and*

the resulting process X is ergodic. A centered version of X can be obtained by setting

$$\tilde{\gamma}(u) = \gamma(u) - \mathbb{E}\gamma(u) \,, \tag{2.17}$$

$$\tilde{X}_k = X_k - \sum_{j=0}^{\infty} \mathbb{E}\,\gamma(a_j) = \sum_{j=0}^{\infty} \tilde{\gamma}_{k-j}(a_j) \,, \qquad k \in \mathbb{Z} \,. \tag{2.18}$$

If moreover, we have, for some $p > 1$,

$$\sum_{j=0}^{\infty} \|\tilde{\gamma}(a_j)\|_{2p}^{2p} < \infty \,, \tag{2.19}$$

then the convergence (1.1) also holds in $\mathbb{L}^{2p}$.

Having a condition for X to satisfy (A-2) and to be L^{2p}, we now provides conditions for obtaining (A-3), which requires $p \geq 2$ for the considered covariances to be well defined.

Theorem 1. *Assume (A-1) and (1.2). Then X is an $\mathbb{L}^2$ stationary process (by Lemma 1), and (A-3) follows from any of the two following assertions with the same $s_0 \in (0,1)$.*

(i) We have (2.19) with $p = 2$ and , as $r \to \infty$,

$$\sum_{j=r}^{\infty} \mathrm{Var}\,\gamma(a_j) = \mathcal{O}(r^{-2s_0}) \,, \tag{2.20}$$

$$\sum_{j=r}^{\infty} \|\tilde{\gamma}(a_j)\|_{2p}^{2p} = \mathcal{O}(r^{-2p\ s_0}) \,. \tag{2.21}$$

(ii) We have (2.19) with $p = 3$ and (2.20) holds as $r \to \infty$.

2.2 Second order estimation

In this section, we suppose that X is a weakly stationary process with auto-covariance r or spectral density f. All the proof of this section are postponed to Section 5.2 for convenience.

The main assumption that we will require on X is (A-3). It is interesting to note that, if (1.14) holds then assuming (1.15) and (1.16) is equivalent to assuming

$$|\mathrm{Cum}\,(X_{t_1}\ ,\ X_{t_2}\ ,\ X_{t_3}\ ,\ X_{t_4})| \leq C_2\ \left[(1 + t_4 - t_3)^{-s_0} \wedge (1 + t_3 - t_2)^{-s_0}\right] \,. \tag{2.22}$$

The precise statement is the following.

Lemma 2. *Let X be a weakly stationary process with zero mean such that (1.14) holds for all $t_1 \leq t_2$ in $\mathbb{Z}$. Then, for all $t_1 \leq t_2 \leq t_3 \leq t_4$ in $\mathbb{Z}$, (1.15) and (1.16) imply (2.22), with $C_2 = C_1 + 3C_0^2$ and (2.22) implies (1.15) and (1.16) with $C_1 = C_2 + 3C_0^2$.*

We denote the empirical covariance function by

$$\widehat{r}_n(m) = \frac{1}{n}\sum_{j=1}^{n-m}(X_j - \bar{X}_n)(X_{j+m} - \bar{X}_n) \tag{2.23}$$

where $\bar{X}_n$ denotes the empirical mean of the sample $X_1, \ldots, X_n$. The centering in the definitions of $\widehat{r}_n$ can be treated separately. Define non-centered covariance estimator

$$\widetilde{r}_n(k) = \frac{1}{n}\sum_{j=1}^{n-k} X_j X_{j+k} \, . \tag{2.24}$$

The empirical covariance function defined by (2.23) can then be written as

$$\widehat{r}_n(k) = \widetilde{r}_n(k) - R_n^r(k) \tag{2.25}$$

where $\widetilde{r}_n$ is the non-centered empirical covariance function defined in (2.24) and $R_n^n(k)$ is the reminder term defined by

$$R_n^r(k) = \frac{k}{n}\left(\bar{X}_n\right)^2 + \bar{X}_n \left(\frac{1}{n}\sum_{j=k+1}^{n-k} X_j\right) \, . \tag{2.26}$$

This term is "small" only if X is a centered process. Nevertheless, X can be assumed centered here, since the empirical covariance $\widehat{r}_n$ is unchanged when X is replaced by its centered version.

In the case where X has mean zero, we have the following result.

Proposition 1. *Let X be an L^4 process with zero mean and satisfying (A-3). Then there exists a constant C' only depending on C_0, C_1 and s_0 such that, for all $0 \le k \le \ell < n$,*

$$|\mathrm{Cov}(\widetilde{r}_n(k) \ , \ \widetilde{r}_n(\ell))| \le C' \, n^{-s_0} \, . \tag{2.27}$$

The following result follows.

Corollary 1. *Let X be a weakly stationary L^4 process satisfying (A-3) with covariance function r. Then there exists a constant C' only depending on C_0, C_1 and s_0 such that, for all $0 \le k \le \ell < n$,*

$$\max_{0 \le k < n} \mathbb{E}\,|\hat{r}_n(k) - r(k)| \le C' \, n^{-s_0/2} \, . \tag{2.28}$$

Another possible application of Proposition 1 is the pointwise Kernel estimation of the spectral density f wherever it is well defined and smooth. Let J denotes a two times continuously differentiable function with support $[-1/2, 1/2]$ and such that $\int J = 1$. For any $\beta > 0$ and $\lambda_0 \in [-\pi, \pi]$, let J_{β,λ_0} denotes its λ_0-shifted, β-scaled and (2π)-periodic version:

$$J_{\beta,\lambda_0}(\lambda) = \frac{1}{\beta}\sum_{k\in\mathbb{Z}} J\left(\frac{\lambda - \lambda_0 - 2\pi k}{\beta}\right) \, .$$

Define the Kernel estimator of $f(\lambda_0)$

$$\hat{f}_{n,\beta}(\lambda_0) = \int_{\mathbb{T}} I_n \, J_{\beta,\lambda_0} \ .$$

Let μ denote the spectral measure of X and suppose that it admits a density f in the neighborhood of λ_0, and that this density is continuous at λ_0. Then, it is easy to show that

$$\lim_{\beta \to 0} \int J_{\beta,\lambda_0} \, \mathrm{d}\mu = f(\lambda_0) \tag{2.29}$$

and the rate of convergence as $\beta \to 0$ can be obtained from the smoothness index of f at λ_0. This deterministic limit can be interpreted as a control on the bias of the estimator $\hat{f}_{n,\beta_n}(\lambda_0)$ of $f(\lambda_0)$. The deviation is bounded by the following result.

Corollary 2. *Let μ be the spectral density associated to the covariance function r. Define I_n and $\hat{r}_n$ by (1.11) and (2.23). Let J be a kernel function as above and define the kernel estimator $\hat{f}_{n,\beta}$ accordingly. Then (2.28) implies that there exists a constant C'' only depending on C', $r(0)$ and J such that, for any $\lambda_0 \in [-\pi, \pi]$,*

$$\mathbb{E} \left| \hat{f}_{n,\beta}(\lambda_0) - \int J_{h,\lambda_0} \, \mathrm{d}\mu \right| \leq C'' \, n^{-s_0/2} \, \beta^{-1} \ . \tag{2.30}$$

As usual, the deviation bound (2.30) (where $\beta = \beta_n$ should not converge to 0 at a rate faster than $n^{s_0/2}$) has to be balanced with the convergence (2.29) (where $\beta = \beta_n$ should converge to 0, the faster the better).

2.3 Parametric Whittle estimation

Although h^* is an unknown element in the infinite dimensional space $\mathcal{C}$, our approach is parametric in nature in the sense that we now assume that (d^*, h^*) belongs to a known compact subset K of $[0, 1/2] \times \mathcal{C}$. In practice, to get a good approximation of an element of $\mathcal{C}$, only a finite number of its Fourier coefficients needs to be estimated. More generally we denote by (K_n) a sequence of subsets of K in which we can always find (d^*, h_n^*) such that h_n^* approximates h^* well for n large. More precisely we consider the following assumption.

(A-4) Let K be a compact subset of $[0, 1/2] \times \mathcal{C}$ such that, for all $(d, h) \in K$, $f_d + h > 0$ on $\mathbb{R}$, and let (K_n) be a sequence of subsets of K such that for a well chosen sequence $(h_n^*) \in \mathcal{C}^{\mathbb{N}}$, we have $(d^*, h_n^*) \in K_n$ for all $n \in \mathbb{N}$ and h_n^* converges to h^* uniformly.

Remark 1. *If $h \in K$ is parameterized by finitely many parameters, one can take $K_n = K$ for all $n \geq 1$, in which case the last assertion of (A-4) is immediately satisfied for all $(d^*, h^*) \in K$ by taking $h_n^* = h^*$ for all n.*

An infinite dimensional setting can be set up as follows. For any $s, C > 0$, let $H(s, C)$ denote the ball of even, real and locally integrable (2π)-periodic functions $h : \mathbb{R} \to \mathbb{R}$ such that the Fourier coefficients

$$c_k(h) = \int h(\lambda)\, \mathrm{e}^{\mathrm{i}\lambda k}\, \mathrm{L}(\mathrm{d}\lambda) \quad \text{satisfy} \quad |c_k(h)| \leq C\,(1 + |k|)^{-1-s}\,, \qquad k \in \mathbb{Z}\,.$$

For any non-negative integer m, let moreover $\mathcal{P}_m$ denote the set of even real trigonometric polynomials of degree at most m. For any locally integrable (2π)-periodic function h, denote by $p_m[h]$ the projection of h onto $\mathcal{P}_m$, that is,

$$p_m[h](\lambda) = c_0(h) + \sum_{k=1}^{m} 2\, c_k(h)\, \cos(\lambda\, k)\,, \quad \lambda \in \mathbb{R}\,.$$

For any $s, C > 0$, it is easy to show that $\sup |h - P_m[h]| = \mathcal{O}(m^{-s})$ uniformly in $h \in H(s, C)$ as $m \to \infty$. The following result can be used to build a parameter space K and a sequence (K_n) satisfying (A-4) from a given set $A \subset [0, 1/2] \times H(s, C)$ of couples (d, h) containing the true parameters.

Lemma 3. *Let $s, C > 0$ and $A \subseteq [0, 1/2] \times H(s, C)$ such that $f_d + h > 0$ on $\mathbb{R}$ for all $(d, h) \in A$. Suppose that A is closed in $[0, 1/2] \times \mathcal{C}$ and let $(d^*, h^*) \in A$. Then there exists a positive integer m_0 such that $f_d + p_m[h] > 0$ on $\mathbb{R}$ for all $(d, h) \in A$ and $m \geq m_0$. Moreover, for any diverging sequence (m_n) of integers larger than or equal to m_0, Assumption (A-4) holds by setting $K_n = \{(d, p_{m_n}[h]) \ : \ (d, h) \in A\}$, for all $n \geq 1$ and $K = A \cup (\bigcup_n K_n)$.*

The proof of this lemma is postponed to Section 5.3. We can now state the consistency of our estimator which, in the same flavor as in [Giraitis et al., 2012, Theorem 8.2.1], only requires the observed process to be ergodic. Its proof is also postponed to Section 5.3.

Theorem 2. *Suppose that the process X satisfies (A-2) and admits a spectral density of the form (1.10) with parameter (c^*, d^*, h^*) satisfying (A-4) for some subsets K and (K_n) of $[0, 1/2] \times \mathcal{C}$.*

Let $(\hat{d}_n, \hat{h}_n) \in K_n$ such that, a.s., as $n \to \infty$,

$$\Lambda_n(\hat{d}_n, \hat{h}_n) \leq \inf_{(d,h)\in K_n} \Lambda_n(d, h) + o(1)\,, \tag{2.31}$$

where Λ_n is defined by (1.12), and define $\hat{c}_n$ by (1.13). Then, a.s., $\hat{d}_n$ and $\hat{c}_n$ converge to d^* and c^*, and $\hat{h}_n$ converges to h^* uniformly.

Assumption (A-4) provides a new framework of parametric models, different from the ones classically used in Whittle parameter estimation, and which seems to be well adapted for many examples of trawl processes, see Section 3. However it also includes many known

cases. Let us examine the celebrated ARFIMA model, in which the spectral density takes the form

$$f(\lambda) = \sigma_*^2 \, f_{d^*}(\lambda) \left| \frac{1 + \sum_{k=1}^{p} \theta_k^* \mathrm{e}^{-\mathrm{i}k\lambda}}{1 - \sum_{k=1}^{q} \phi_k^* \mathrm{e}^{-\mathrm{i}k\lambda}} \right|^2 , \tag{2.32}$$

where, for some positive integers p and q, the MA and AR coefficients $\theta^* = (\theta_1^*, \ldots, \theta_p^*)$ and $\phi^* = (\phi_1^*, \ldots, \phi_q^*)$ are assumed to make the corresponding ARMA process canonical. In the following this will be denoted by $(\phi, \theta) \in \Theta_{p,q}$, defined by

$$\begin{aligned} \Theta_{p,q} = \big\{ (\phi, \theta) \in \mathbb{R}^{p+q} \ : \ & \Phi \text{ and } \Theta \text{ have no common roots} \\ & \text{and for all } z \in \mathbb{C} \text{ such that } |z| \leq 1, \Phi(z) \neq 0 \text{ and } \Theta(z) \neq 0 \big\} , \end{aligned} \tag{2.33}$$

where Φ and Θ are the AR and MA polynomials defined by

$$\Phi(z) := 1 - \sum_{k=1}^{p} \phi_k z^k \text{ and } \Theta(z) := 1 + \sum_{k=1}^{q} \theta_k z^k .$$

The corresponding reduced Whittle contrast reads

$$\widetilde{\Lambda}_n(d, (\phi, \theta)) = \ln \left(\int \frac{I_n(\lambda)}{f_d(\lambda)} \, \frac{|\Phi(\mathrm{e}^{-\mathrm{i}\lambda})|^2}{|\Theta(\mathrm{e}^{-\mathrm{i}\lambda})|^2} \, \mathrm{L}(\mathrm{d}\lambda) \right) , \tag{2.34}$$

where f_d is defined by (1.9) and L is the Lebesgue measure on $[-\pi, \pi]$ divided by 2π. The form (2.32) is in fact a special case of (1.10) by setting

$$h^*(\lambda) = f_{d^*}(\lambda) \left(\left| \frac{\Theta(\mathrm{e}^{-\mathrm{i}\lambda}) \, \Phi(1)}{\Phi(\mathrm{e}^{-\mathrm{i}\lambda}) \, \Theta(1)} \right|^2 - 1 \right) \tag{2.35}$$

$$c^* = \sigma_*^2 \left| \frac{\Theta(1)}{\Phi(1)} \right|^2 . \tag{2.36}$$

Note that h^* is indeed continuous. The ARFIMA linear processes have been extensively studied. However the usual proof of the consistency relies on the Hannan's approach of [Hannan, 1973] but it does not hold if $d = 0$ is included in the set of parameters. Here, as a consequence of Theorem 2, we get the following, which provides an alternative proof.

Corollary 3. *Let p, q be two positive integers and $\tilde{K}$ be a compact subset of $[0, 1/2) \times \Theta_{p,q}$. Suppose that the process X satisfies (A-2) and admits a spectral density of the form (2.32), with $(d^*, (\phi^*, \theta^*)) \in \tilde{K}$ and $\sigma_* > 0$.*

Let $(\hat{d}_n, \hat{\vartheta}_n) \in \tilde{K}$ such that, a.s., as $n \to \infty$,

$$\widetilde{\Lambda}_n(\hat{d}_n, \hat{\vartheta}_n) \leq \inf_{(d,\vartheta) \in \tilde{K}} \widetilde{\Lambda}_n(d, \vartheta) + o(1) , \tag{2.37}$$

where $\widetilde{\Lambda}_n$ is defined by (2.34). Define moreover

$$\hat{\sigma}_n^2 = \exp\left(\widetilde{\Lambda}_n(\hat{d}_n, \hat{\vartheta}_n) \right) . \tag{2.38}$$

Then, a.s., $\hat{d}_n$, $\hat{\vartheta}_n$ and $\hat{\sigma}_n^2$ converge to d^*, $\vartheta^* = (\phi^*, \theta^*)$ and σ_*^2.

Proof. See Section 5.4. □

3 Examples of discrete time trawl processes

3.1 Random line seed

As explained in [Doukhan et al., 2019, Example 1], any causal linear process is a trawl process by setting the seed process to be the random line seed $\gamma(t) = t\epsilon$, where ϵ is a random variable with zero mean and finite variance.

The parametric estimation in the linear case is a well known topic, usually treated using ARFIMA parametrization, see e.g. [Giraitis et al., 2012, Section 8.3.2] for a complete statistical analysis of this model.

3.2 Lévy seed and non-increasing sequence

Consider the two following assumptions

(A-5) The process γ is a Lévy process with finite variance normalized so that $\operatorname{Var}\gamma(1) = 1$.

(A-6) The sequence a is non-increasing and there exist $c > 0$, and $\alpha^* > 1$ such that (1.5) holds.

They imply (1.2) since then we have, for all $t \geq 0$, $\mathbb{E}\,\gamma(t) = \delta\, t$ for some drift δ and $\operatorname{Var}\gamma(t) = t$. By Lemma 1 and Eq. (1.3), the trawl process X defined by (1.1) satisfies (A-2) and its autocovariance function r is given by

$$r(k) = \sum_{j \geq k} a_j \;, \qquad k \in \mathbb{N} \;. \tag{3.39}$$

If (A-5) and (A-6) hold and $\gamma(1)$ admits a finite q-th moment, we easily have that, for all $t_1 \leq \cdots \leq t_q$ in $\mathbb{Z}$,

$$\begin{aligned} \operatorname{Cum}\left(X_{t_1}\;,\;\ldots\;,\;X_{t_q}\right) &= \sum_{j \geq 0} \operatorname{Cum}\left(\gamma(a_{t_k - t_1 + j}),\; k = 1, \ldots, q\right) \\ &= \kappa_q \sum_{k \geq t_q - t_1} a_k \;, \end{aligned}$$

where κ_q is the q-th order cumulant of $\gamma(1)$. We then obtain

$$\operatorname{Cum}\left(X_{t_1}\;,\;X_{t_2}\;,\;X_{t_3}\;,\;X_{t_4}\right) = \mathcal{O}((t_4 - t_1)^{1-\alpha^*}) \;.$$

So, by Lemma 2, if $q = 4$, X satisfies (A-3) with $s_0 = \alpha^* - 1$. Theorem 1 shows that Condition (A-3) continues to hold for more general trawl processes, provided some adequate moment conditions, but with s_0 possibly higher than $\alpha^* - 1$ (see Section 3.3 for examples).

For such a process, we can specify (a_k) so that the spectral density is of the form (1.10) with (d^*, h^*) lying within a parameter space K satisfying Condition (A-4). A very special case, detailed in [Doukhan et al., 2019, Example 5], consists in setting

$$a_k = c^* \left(r_k^{(d^*)} - r_{k+1}^{(d^*)} \right) , \qquad k \in \mathbb{N} , \tag{3.40}$$

where, for all $d < 1/2$, $r^{(d)}$ is defined as the auto-covariance function of ARFIMA$(0, d, 0)$ with unit variance innovation, that is,

$$r^{(d)}(k) = \int \left| 1 - \mathrm{e}^{-\mathrm{i}\lambda} \right|^{-2d} \mathrm{e}^{\mathrm{i}\lambda k} \, \mathrm{L}(\mathrm{d}\lambda) , \qquad k \in \mathbb{Z} . \tag{3.41}$$

It is shown in [Doukhan et al., 2019] that, for any $d^* \in (0, 1/2)$ such a sequence (a_j) satisfies (A-6) with $\alpha^* = 2(1 - d^*) \in (1, 2)$, so that, under (A-5), following (3.39) and (3.41), the corresponding trawl process has a spectral density of the form (1.10) with $h^* = 0$.

We check in the following section that more general seed processes and trawl sequences can be used.

3.3 More general seeds and sequences

In this section, in contrast to (A-6), we consider trawl sequences (a_j) that may not be non-increasing but we specify (1.5) by assuming that, there exists $c > 0$ and $\alpha^* \in (1, 2)$ such that

$$0 \le a_j \;=\; c \, j^{-\alpha^*} (1 + O(j^{-1})) \quad \text{as} \quad j \to \infty . \tag{3.42}$$

We also consider the non Lévy seed processes introduced in [Doukhan et al., 2019], for which the covariance structure can still be derived precisely. Let us examine here the mixed Poisson seed and the Binomial seed processes of their Examples 3 and 4. The first case extends the (thus Lévy) Poisson seed by setting $\gamma(t) = N(\zeta \, t)$, where N is a homogeneous Poisson counting process with unit intensity and ζ is a positive random variable independent of N and with finite variance. Then we have, for all $u, v \ge 0$, $\mathbb{E}\, \gamma(u) = u \, \mathbb{E}\zeta$ and $\mathrm{Cov}(\gamma(u), \gamma(v)) = (u \wedge v)\mathbb{E}\zeta + uv\mathrm{Var}(\zeta)$. Thus, for any sequence (a_j) satisfying (3.42), Condition (1.2) holds and (A-2) follows from Lemma 1 and Eq. (1.3) yields the following auto-covariance function for X :

$$r(k) = \mathbb{E}\zeta \sum_{j=0}^{\infty} (a_j \wedge a_{j+k}) + \mathrm{Var}(\zeta) \sum_{j=0}^{\infty} a_j a_{j+k} , \qquad k \in \mathbb{N} .$$

If moreover $\mathbb{E}\zeta^6 < \infty$, then (2.19) holds with $p = 3$, and, by Theorem 1 (ii), we get (A-3) with $s_0 = (\alpha^* - 1)/2$. If we only assume that $\mathbb{E}\zeta^4 < \infty$, then (2.21) holds with $p = 2$ and $s_0 = (\alpha^* - 1)/4$, so that Theorem 1 (i) gives that (A-3) holds this time only with $s_0 = (\alpha^* - 1)/4$.

The Binomial seed process of [Doukhan et al., 2019, Example 4] is defined for some given $n \in \mathbb{N}^*$ by setting $\gamma(t) = \sum_{i=1}^{n} \mathbb{1}_{\{U_i \le t\}}$ with the U_i's i.i.d. and uniform on $[0,1]$. In this case, we have that, for all $u \ge 1$, $\gamma(u) = n$ and, for all $u, v \in [0,1]$, $\mathbb{E}\,\gamma(u) = n\,u$ and $\mathrm{Cov}(\gamma(u), \gamma(v)) = n(u \wedge v - uv)$. Thus, for any sequence (a_j) satisfying (3.42), similarly to the previous case, (A-2) holds and the trawl process X has auto-covariance function r given by

$$r(k) = n \sum_{j=0}^{\infty} (\tilde{a}_j \wedge \tilde{a}_{j+k}) - n \sum_{j=0}^{\infty} \tilde{a}_j \tilde{a}_{j+k} \,, \qquad k \in \mathbb{N} \,,$$

where, for all $j \in \mathbb{N}$, $\tilde{a}_j = a_j \mathbb{1}_{\{a_j < 1\}}$. Also, for the binomial seed and (a_j) satisfying (3.42), (2.19) holds for any integer p, and (A-3) holds with $s_0 = (\alpha^* - 1)/2$ by Theorem 1 (ii).

Having checked that the trawl process satisfies (A-2) and (A-3) for these seeds, we now turn to the form of its spectral density and show that it is indeed of the form (1.10) and can be used with Lemma 3 to form a parameter space K that satisfies (A-4).

Proposition 2. *Assume (A-1). Suppose that γ is Lévy seed process, a mixed Poisson seed process or a binomial seed process. Suppose moreover that $\gamma(1)$ has finite positive variance and (a_j) satisfies (3.42) with $\alpha^* \in (1,2)$. Then the trawl process defined by (1.1) has a spectral density of the form (1.10) with $d^* = 1 - \alpha^*/2 \in (0, 1/2)$ and $h^* \in H(\alpha^* - 1, C)$ for some $C > 0$.*

Proof. See Section 5.1. □

4 Weighted weak dependence indices

Here we introduce a somewhat general setting that will be used later to derive some important properties on the memory of Trawl processes. They can be, however, of independent interest.

We use the classical weak-dependence concept.

Definition 1 ([Dedecker et al., 2007])**.** *A random process $(X_t)_{t \in \mathbb{Z}}$ is said to be $\theta-$weakly dependent if*

$$|\mathrm{Cov}\left(f(X_{i_1}, \ldots, X_{i_u}), g(X_{j_1}, \ldots, X_{j_v})\right)| \le \theta_r \, v \,, \tag{4.43}$$

for $i_1 \le \cdots \le i_u \le j_1 - r \le j_1 \le \cdots \le j_v$ and functions $f : \mathbb{R}^u \to \mathbb{R}$ with $\|f\|_\infty \le 1$ and $g : \mathbb{R}^v \to \mathbb{R}$ with

$$|g(y_1, \ldots, y_v) - g(x_1, \ldots, x_v)| \le |y_1 - x_1| + \cdots + |y_v - x_v| \,.$$

Definition 2. *A time series (X_k) is said to be a causal Bernoulli shift process (CBS) if there exists an iid sequence $(\gamma_j)_{j \in \mathbb{Z}}$ valued in $(E, \mathcal{E})$ and a measurable function $\Phi : E^{\mathbb{N}} \to \mathbb{R}$ such*

that, for all $k \in \mathbb{Z}$, $X_k = \Phi((\gamma_{k-j})_{j\geq 0})$. *The* L^q *coefficients* $(\pi_r^{(q)})_{r\geq 1}$ *of* (X_k) *are then defined by*

$$\pi_r^{(q)} = \left\| \Phi((\gamma_j)_{j\geq 0}) - \Phi((\gamma_j)_{0\leq j<r}, (\gamma'_j)_{j\geq r}) \right\|_q \ , \tag{4.44}$$

where $(\gamma'_j)_{j\in\mathbb{Z}}$ *is an independent copy of* $(\gamma_j)_{j\in\mathbb{Z}}$.

Provided that a CBS process is well defined in $\mathbb{L}^2$, it is weakly dependent.

Lemma 4. *Let* X *be an* $\mathbb{L}^2$ *centered CBS process. Then it is* $\pi^{(2)}-$*weakly dependent.*

Proof. We write $X_k = X'_k + (X_k - X'_k)$ where (X'_k) is defined by $X'_k = \Phi((\gamma_{k-j})_{0\leq j<r}, (\gamma'_{k-j})_{j\geq r}))$. Observe now that X'_k is independent of $\sigma(\gamma_i,\, i \leq k-r)$, hence of $\sigma(X_i,\, i \leq k-r)$. On the other hand we have that

$$\mathbb{E}(X_k - X'_k)^2 = \pi_r^{(2)} \ .$$

Now take $f : \mathbb{R}^u \to \mathbb{R}$ with $\|f\|_\infty \leq 1$ and $g : \mathbb{R}^v \to \mathbb{R}$ Lipschitz and $i_1 \leq \cdots \leq i_u \leq j_1 - r \leq j_1 \leq \cdots \leq j_v$. Denoting

$$m = \mathbb{E}\, f(X_{i_1}, \ldots, X_{i_u}) \ ,$$

we get that $|\mathrm{Cov}\,(f(X_{i_1}, \ldots, X_{i_u}), g(X_{j_1}, \ldots, X_{j_v}))|$ is bounded from above by

$$\mathbb{E}\left[|f(X_{i_1}, \ldots, X_{i_u}) - m| \left(\sum_{k=1}^{v} |X_{j_k} - X'_{j_k}| \right) \right] \ .$$

And we conclude with the Cauchy-Schwartz inequality. □

Using the same proof we can include polynomial terms in the functions f and g.

Definition 3. *Let* $p_{(-)}, p_{(+)} \geq 1$. *A random process* $(X_t)_{t\in\mathbb{Z}}$ *is said to be* $(p_{(-)}, p_{(+)})$*-weighted* $\theta-$*weakly dependent if*

$$|\mathrm{Cov}\,(f(X_{i_1}, \ldots, X_{i_u}), g(X_{j_1}, \ldots, X_{j_v}))| \leq \theta_r\, v \ , \tag{4.45}$$

for all $i_1 \leq \cdots \leq i_u \leq j_1 - r \leq j_1 \leq \cdots \leq j_v$, *all functions* $f : \mathbb{R}^u \to \mathbb{R}$ *satisfying*

$$|f(x_{i_1}, \ldots, x_{i_u})| \leq \frac{1}{1+u} \left(1 + \sum_{k=1}^{u} |x_{i_k}| \right)^{p_{(-)}} \ ,$$

and all functions $g : \mathbb{R}^v \to \mathbb{R}$ *satisfying*

$$\frac{|g(y_1, \ldots, y_v) - g(x_1, \ldots, x_v)|\ (1+2v)}{(1 + |y_1| + \cdots + |y_v| + |x_1| + \cdots + |x_v|)^{p_{(+)}-1}} \leq (|y_1 - x_1| + \cdots + |y_v - x_v|) \ .$$

Remark 2. *Note that in Definition 3, the conditions on* f *and* g *are weaker as* $p_{(\pm)}$ *increases. Namely, if* $1 \leq p_{(\pm)} \leq p'_{(\pm)}$, *then a* $(p'_{(-)}, p'_{(+)})$*-weighted* $\theta-$*weakly dependent random process* $(X_t)_{t\in\mathbb{Z}}$ *is also* $(p_{(-)}, p_{(+)})$*-weighted* $\theta-$*weakly dependent.*

Using this new definition, we get the following result.

Lemma 5. *Let (X_k) be an $\mathbb{L}^{2p}$ centered CBS process. Then, for any $p_{(\pm)} \geq 1$ such that $p_{(-)} + p_{(+)} = 2p$, it is $(p_{(-)}, p_{(+)})$-weighted θ-weakly dependent with*

$$\theta_r \leq \left(1 \vee \|X_0\|_{2p}^{2p-1}\right) \pi_r^{(2p)} .$$

Proof. Let us now prove the bound of the p-weighted $\theta-$weak dependence coefficient $\theta_r^{(p)}$. We use the same notation as in the proof of Lemma 4 but this time with f and g as in Definition 3. We then obtain that

$$|\mathrm{Cov}\left(f(X_{i_1}, \ldots, X_{i_u}), g(X_{j_1}, \ldots, X_{j_v})\right)|$$

is bounded from above by

$$\frac{1}{1+2v}\,\mathbb{E}\left[|f(X_{i_1}, \ldots, X_{i_u}) - m| \left(\sum_{k=1}^{v} |X_{j_k} - X'_{j_k}|\right) \left(1 + \sum_{k=1}^{v} |X_{j_k}| + |X'_{j_k}|\right)^{p_{(+)}-1}\right] . \tag{4.46}$$

Using the Hölder inequality with $p_{(-)}/(2p) + 1/(2p) + (p_{(+)} - 1)/2p = 1$, we obtain

$$|\mathrm{Cov}\left(f(X_{i_1}, \ldots, X_{i_u}), g(X_{j_1}, \ldots, X_{j_v})\right)| \leq A\, B\, C\,, \tag{4.47}$$

with

$$\begin{aligned}
A &= \|f(X_{i_1}, \ldots, X_{i_u}) - m\|_2 \leq \|f(X_{i_1}, \ldots, X_{i_u})\|_{2p/p_{(-)}} \\
B &= \left\|\sum_{k=1}^{v} |X_{j_k} - X'_{j_k}|\right\|_{2p} \leq v\, \pi_r^{(2p)} \\
C &= \frac{1}{1+2v} \left\|\left(1 + \sum_{k=1}^{v} \left(|X_{j_k}| + |X'_{j_k}|\right)\right)^{p_{(+)}-1}\right\|_{2p/(p_{(+)}-1)} .
\end{aligned}$$

We immediately have that, by the assumption on f that

$$A \leq \frac{1}{1+u} \left(1 + \sum_{k=1}^{u} \|X_{i_k}\|_{2p}\right)^{p_{(-)}} \leq 1 \vee \|X_0\|_{2p}^{p_{(-)}} .$$

Finally, we note that $C \leq \left(1 \vee \|X_0\|_{2p}^{p_{(+)}-1}\right)$. The result follows from (4.47) and the above bounds of A, B and C. □

We also obtained this lemma with an improved weighted weakly dependent coefficient by conceding a bit of moment condition.

Lemma 6. *Let (X_k) be an $\mathbb{L}^{2p}$ centered CBS process. Then, for any $p_{(\pm)} \geq 1$*

$$p_{(-)} + p_{(+)} = p + 1 \; . \tag{4.48}$$

it is $(p_{(-)}, p_{(+)})$-weighted θ-weakly dependent with

$$\theta_r \leq 2 \left(1 \vee \|X_0\|_{2p}^p\right) \pi_r^{(2)} \; .$$

where C_p is a positive constant only depending on p and $S_0(2p)$ is defined in (5.51).

Proof. We use again the upper bound (4.46) of

$$|\mathrm{Cov}\left(f(X_{i_1}, \ldots, X_{i_u}), g(X_{j_1}, \ldots, X_{j_v})\right)| \; ,$$

but we apply the Hölder inequality with the weights $p_{(-)}/(2p) + 1/2 + (p_{(+)} - 1)/(2p) = 1$ (which holds by (4.48)) and obtain that

$$|\mathrm{Cov}\left(f(X_{i_1}, \ldots, X_{i_u}), g(X_{j_1}, \ldots, X_{j_v})\right)| \leq A' \; B' \; C' \; , \tag{4.49}$$

with

$$\begin{aligned}
A' &= \|f(X_{i_1}, \ldots, X_{i_u}) - m\|_{2p/p_{(-)}} \leq 2 \; \|f(X_{i_1}, \ldots, X_{i_u})\|_{2p/p_{(-)}} \\
B' &= \left\| \sum_{k=1}^{v} |X_{j_k} - X'_{j_k}| \right\|_2 \leq v \, \pi_r^{(2)} \\
C' &= \frac{1}{1+2v} \left\| \left(1 + \sum_{k=1}^{v} \left(|X_{j_k}| + |X'_{j_k}|\right)\right)^{p_{(+)}-1} \right\|_{2p/(p_{(+)}-1)} \; .
\end{aligned}$$

We immediately have that, by the assumption on f that

$$A' \leq \frac{2}{1+u} \left(1 + \sum_{k=1}^{u} \|X_{i_k}\|_{2p}\right)^{p_{(-)}} \leq 2 \left(1 \vee \|X_0\|_{2p}^{p_{(-)}}\right) \; .$$

Finally, we note that, similarly, $C' \leq \left(1 \vee \|X_0\|_{2p}^{p_{(+)}-1}\right)$. The result follows from (4.49) and the above bounds of A', B' and C'. □

5 Proofs

5.1 On trawl processes

Proof of Lemma 1. We prove the result under (1.2) and (2.19). The case where (2.19) is not assumed corresponds to setting $p = 1$ in the following. By the Rosenthal Inequality for sums

of independent random variables, see [Petrov, 1995, Theorem 2.9], we have, for any $1 \le i \le k$, for some constant C_p only depending on p,

$$\left\| \sum_{j=i}^{k} \gamma_j(a_j) \right\|_{2p} \le C_p \sum_{j=i}^{k} |\mathbb{E}\,\gamma(a_j)| + \left(\sum_{j=i}^{k} \|\tilde{\gamma}(a_j)\|_{2p}^{2p} \right)^{1/2p} + \left(\sum_{j=i}^{k} \operatorname{Var} \gamma(a_j) \right)^{1/2} \tag{5.50}$$

The convergence of (1.1) in $\mathbb{L}^{2p}$ follows.

It follows that we can write X_k as $X_k = F((\gamma_{k-j})_{j\ge 0})$ with F measurable from $\mathbb{R}^{\mathbb{R}}$ to $\mathbb{R}$, with $\mathbb{R}^{\mathbb{R}}$ endowed by the σ-field $\mathcal{B}(\mathbb{R})^{\otimes\mathbb{R}}$ (the smallest one that makes the $\mathbb{R}^{\mathbb{R}} \to \mathbb{R}$ mapping $x \mapsto x(t)$ measurable for all $t \in \mathbb{R}$). Since $(\gamma_j)_{j\in\mathbb{Z}}$ is i.i.d., it is ergodic, and so is $(X_k)_{k\in\mathbb{Z}}$.

All the other assertions of the lemma are obvious. □

Using the same idea and the results of Section 4, we now prove Theorem 1.

Proof of Theorem 1. Let us now prove Theorem 1. Note that

$$r(k) = \sum_{j\ge 0} \operatorname{Cov}\Big(\gamma(a_j), \gamma(a_{k+j})\Big) \le \left(\sum_{j\ge 0} \operatorname{Var}\ \gamma(a_j) \right)^{1/2} \left(\sum_{j\ge k} \operatorname{Var}\ \gamma(a_j) \right)^{1/2} .$$

Hence (2.20) implies (1.14).

It remains to show (1.15) and (1.16). We use that $(\tilde{X}_k)$ defined in (2.18) can be written the causal Bernoulli shift process

$$\tilde{X}_k = \Phi((\gamma_{k-j})_{j\ge 0}) \quad \text{with} \quad \Phi((\gamma_j)_{j\ge 0}) = \sum_{j=0}^{\infty} \tilde{\gamma}_j(a_j) .$$

where Φ is a measurable mapping on $E^{\mathbb{N}}$, with $E = \mathbb{R}^{\mathbb{R}}$ endowed with $\mathcal{B}(\mathbb{R})^{\otimes\mathbb{R}}$. Then the L^q coefficients defined in (4.44) with $(\gamma'_j)_{j\ge 0}$ denoting an independent copy of $(\gamma_j)_{j\ge 0}$, satisfy, for all $r \in \mathbb{N}$, and $q \ge 2$,

$$\begin{aligned} \pi_r^{(q)} &= \left\| \sum_{j=r}^{\infty} \big(\gamma_j(a_j) - \gamma'_j(a_j)\big) \right\|_q \\ &\le 2\, C_q S_r(q) , \end{aligned}$$

where the second inequality follows from (5.50) by setting

$$S_r(q) = \left(\sum_{j=r}^{\infty} \|\tilde{\gamma}(a_j)\|_q^q \right)^{1/q} + \left(\sum_{j=r}^{\infty} \operatorname{Var} \gamma(a_j) \right)^{1/2} . \tag{5.51}$$

We now separate the two cases.

In the case where $p = 2$ and both (2.20) and (2.21) hold, we apply Lemma 5 with $p_{(-)} = p_{(+)} = 2$ and $p_{(-)} = 3$, $p_{(+)} = 1$, successively. This gives (1.15) and (1.16), respectively.

In the case where $p = 3$, we only need (2.20) to hold, as we can apply Lemma 6 with $p_{(-)} = p_{(+)} = 2$ and $p_{(-)} = 3$, $p_{(+)} = 1$, successively. □

The following lemma is useful for proving Proposition 2.

Lemma 7. *Let $\alpha > 1$. Let (b_k) be a non-negative sequence such that $b_k = (1+k)^{-\alpha}(1 + O(k^{-1}))$ as $k \to \infty$. Then we have, as $k \to \infty$,*

$$\sum_{j\geq 0}(b_j \wedge b_{j+k}) = \sum_{j\geq k} b_j + O(k^{-\alpha}) \, .$$

Proof. First observe that, for all $k \in \mathbb{N}$,

$$\sum_{j\geq 0}(b_j \wedge b_{j+k}) \leq \sum_{j\geq 0} b_{j+k} \, . \tag{5.52}$$

Now, there exists $C > 0$ such that for all $j \in \mathbb{N}$,

$$(1+j)^{-\alpha}(1 - C(j+1)^{-1}) \leq b_j \leq (1+j)^{-\alpha}(1 + C(j+1)^{-1}) \, . \tag{5.53}$$

It follows from the first inequality that, for all $j \in \mathbb{N}$ and $k \geq (1+C)/\alpha$,

$$\begin{aligned} \frac{b_j}{(j+k+1)^{-\alpha}} &\geq (1 - C(j+1)^{-1})(1 + \alpha^{-1}(1+C)(j+1)^{-1})^{\alpha} \\ &= 1 + (j+1)^{-1} + O(j^{-2}) \quad \text{as } j \to \infty. \end{aligned}$$

In particular, the latter term is larger than or equal to 1 for j large enough and it follows that there exists j_0 only depending on α and C such that, for all $j \geq j_0$ and $k \geq (1+C)/\alpha$,

$$b_j \geq (j+k+1)^{-\alpha} \geq b_{j+k} - C(j+k+1)^{-1} \, ,$$

where we used the second inequality of (5.53). This now implies that, for all $k \geq (1+C)/\alpha$,

$$\sum_{j\geq 0}(b_j \wedge b_{j+k}) \geq \sum_{j\geq j_0} b_{j+k} - C\sum_{j\geq j_0}(j+k+1)^{-\alpha-1} = \sum_{j\geq 0} b_{j+k} - \sum_{j=0}^{j_0-1} b_{j+k} - C\sum_{j\geq j_0}(j+k+1)^{-\alpha-1} \, .$$

Since j_0 is fixed the two last term in the previous display are $O(k^{-\alpha})$ as $k \to \infty$ and we conclude from (5.52). □

We can now provide the proof of Proposition 2.

Proof of Proposition 2. From what precedes, we know that under these assumptions, the trawl process has an auto-covariance function of the form

$$r(k) = A\, S(k) + B\, P(k)\,, \qquad k \in \mathbb{N}\,, \tag{5.54}$$
$$\text{with} \quad S(k) = \sum_{j=0}^{\infty} (\tilde{a}_j \wedge \tilde{a}_{j+k}) \quad \text{and} \quad P(k) = \sum_{j=0}^{\infty} \tilde{a}_j \tilde{a}_{j+k}\,.$$

where $A > 0$, $B \in \mathbb{R}$ and $\tilde{a}_j = a_j \mathbb{1}_{\{a_j < \bar{a}\}}$ with $\bar{a}$ some positive constant. We treat the two terms in the right-hand side of (5.54) separately.

Term S: Since $\tilde{a}_k = a_k$ for k large enough, $(\tilde{a}_k)$ also satisfies Condition (3.42) and Lemma 7 gives that

$$S(k) = \sum_{j \geq k} \tilde{a}_j + O(k^{-\alpha^*}) = \sum_{j \geq k} a_j + O(k^{-\alpha^*})\,. \tag{5.55}$$

Recall the definition of $r^{(d)}$ in (3.41). Define, for all $k \geq 0$,

$$a_k^* = r^{(d^*)}(k) - r^{(d^*)}(k+1) = r^{(d^*)}(k) \frac{1 - 2d^*}{k + 1 - d^*}\,,$$

where the second equality is derived in [Doukhan et al., 2019, Example 5]. By [Giraitis et al., 2012, Theorem 72.1] and its proof, we have for any $d \in (-1/2, 1/2)$,

$$r_k^{(d)} = \frac{\Gamma(1-2d)}{\Gamma(1-d)\Gamma(d)}\, k^{-1+2d}\, (1 + O(k^{-1}))\,. \tag{5.56}$$

Hence the previous equation and the definition of d^* give that

$$a_k^* = \frac{\Gamma(2-2d^*)}{\Gamma(1-d^*)\Gamma(d^*)}\, k^{-\alpha^*}\, (1 + O(k^{-1}))\,. \tag{5.57}$$

And Condition (3.42) is equivalent to have

$$a_k = c^* \left(a_k^* + O(k^{-\alpha^* - 1}) \right)$$

with $c^* > 0$ only depending on c and α^*. Inserting this in (5.55) and using the definition of a^*, we obtain

$$S(k) = c^*\, r^{(d^*)}(k) + O(k^{-\alpha^*})\,.$$

This, with the definition (3.41) implies

$$S(k) = \int \left(c^* \left| 1 - \mathrm{e}^{-\mathrm{i}\lambda} \right|^{-2d^*} + h_S^* \right) \mathrm{e}^{\mathrm{i}\lambda k}\, \mathrm{L}(\mathrm{d}\lambda)\,, \qquad k \in \mathbb{N}\,,$$

where $h_S^* \in H(\alpha^* - 1, C_S)$ for some $C_S > 0$.

Term P: It only remains to prove that P defined in (5.54) satisfies $P(k) = O(k^{-\alpha^*})$ as $k \to \infty$ (so that the associated Fourier series belongs to $H(\alpha^* - 1, C_P)$ for some $C_P > 0$).

This follows immediately by observing that (3.42) with $\alpha^* > 1$ implies, for some constant $C > 0$ and all $k \in \mathbb{N}$,

$$|P(k)| \leq C \sum_{j=0}^{\infty} (j+1)^{-\alpha^*} (j+k+1)^{-\alpha^*} \leq C \left(\sum_{j=0}^{\infty} (j+1)^{-\alpha^*} \right) (k+1)^{-\alpha^*} .$$

This concludes the proof. □

5.2 Convergence of the empirical covariance function

We start with the proof of Lemma 2.

Proof of Lemma 2. Let r denote the autocovariance function of X. Let $t_1 \leq t_2 \leq t_3 \leq t_4$ in $\mathbb{Z}$. We use the identities

$$\begin{aligned} \mathrm{Cov}(X_{t_1} X_{t_2} \ , \ X_{t_3} X_{t_4}) = \mathrm{Cum}\,(X_{t_1} \ , \ X_{t_2} \ , \ X_{t_3} \ , \ X_{t_4}) &+ r(t_1 - t_3) r(t_2 - t_4) \\ &+ r(t_1 - t_4) r(t_2 - t_3) \ , \qquad (5.58) \\ \mathrm{Cov}(X_{t_1} X_{t_2} X_{t_3} \ , \ X_{t_4}) = \mathrm{Cum}\,(X_{t_1} \ , \ X_{t_2} \ , \ X_{t_3} \ , \ X_{t_4}) &+ r(t_4 - t_1) r(t_3 - t_2) \\ &+ r(t_4 - t_2) r(t_3 - t_1) \\ &+ r(t_4 - t_3) r(t_2 - t_1) \ . \qquad (5.59) \end{aligned}$$

This, with the bound (1.14), allows to go back and forth from (1.15) or (1.16) to (2.22). □

We can now prove Proposition 1.

Proof of Proposition 1. We have, using again the identity displayed in (2.22),

$$\begin{aligned} |\mathrm{Cov}(\widetilde{r}_n(k) \ , \ \widetilde{r}_n(\ell))| &\leq \frac{1}{n^2} \sum_{s=1}^{n-k} \sum_{s'=1}^{n-\ell} |\mathrm{Cov}(X_s X_{s+k} \ , \ X_{s'} X_{s'+\ell})| \\ &\leq \frac{1}{n^2} \sum_{s=1}^{n-k} \sum_{s'=1}^{n-\ell} |\mathrm{Cum}\,(X_s \ , \ X_{s+k} \ , \ X_{s'} \ , \ X_{s'+\ell})| \qquad (5.60) \\ &+ \frac{1}{n^2} \sum_{s=1}^{n-k} \sum_{s'=1}^{n-\ell} \left| r(s-s') r(s-s'+k-\ell) \right| \qquad (5.61) \\ &+ \frac{1}{n^2} \sum_{s=1}^{n-k} \sum_{s'=1}^{n-\ell} \left| r(s-s'+k) r(s-s'-\ell) \right| \ . \qquad (5.62) \end{aligned}$$

Using (1.14), we get that (5.61) and (5.62) are both less than or equal to

$$C_0^2 \frac{1}{n^2} \sum_{s=1}^{n} \sum_{s'=1}^{n} (1 + |s-s'|)^{-s_0} \leq C_0^2 \frac{1}{n} \sum_{\tau=-n+1}^{n-1} (1 + |\tau|)^{-s_0} \leq C \ n^{-s_0} \ ,$$

where $C > 0$ only depends on C_0 and s_0. To get (2.27), it thus only remains to show that a similar bound holds for the term appearing in (5.60). To this end we use the bound (2.22) that we have showed to hold under (A-3) in Lemma 2. More precisely we use the bound on left-hand side of the $\wedge$ sign in (2.22) in the first following case and the bound on right-hand side of the $\wedge$ sign for all the other cases:

1. For $s' \leq s \leq s+k \leq s'+\ell$,
$$|\mathrm{Cum}\,(X_s\ ,\ X_{s+k}\ ,\ X_{s'}\ ,\ X_{s'+\ell})| \leq C_2\ (1+|s-s'|)^{-s_0}\ .$$
2. For $s \leq s', s+k \leq s'+\ell$,
$$|\mathrm{Cum}\,(X_s\ ,\ X_{s+k}\ ,\ X_{s'}\ ,\ X_{s'+\ell})| \leq C_2\ (1+|s+k-s'|)^{-s_0}\ .$$
3. For $s' \leq s, s'+\ell \leq\leq s+k$,
$$|\mathrm{Cum}\,(X_s\ ,\ X_{s+k}\ ,\ X_{s'}\ ,\ X_{s'+\ell})| \leq C_2\ (1+|s'+\ell-s|)^{-s_0}\ .$$

(The case $s \leq s' \leq s'+\ell \leq s+k$ can only occur if $s = s'$ and $\ell = k$ since we assumed $0 \leq k \leq \ell$, so is included in the first case.) Hence we get that the term in (5.60) is bounded from above by

$$\frac{C_2}{n^2} \max_{j=0,k,\ell} \sum_{s=1}^{n}\sum_{s'=1}^{n}(1+|s-s'+j|)^{-s_0} \leq \frac{C_2}{n} \sum_{\tau=-2n}^{2n}(1+|\tau|)^{-s_0} \leq C\ n^{-s_0}\ ,$$

where C only depends on C_2 and s_0. □

Next, we prove Corollary 1.

Proof of Corollary 1. Since $\hat{r}_n(k)$, I_n, r and f are invariant by centering, we can assume in the following that X is centered without loss of generality.

Using Proposition 1 and $\mathbb{E}\,\tilde{r}_n(k) = (1-k/n)\ r(k)$, we have, for all $0 \leq k < n$,

$$\mathbb{E}\ (\tilde{r}_n(k)-r(k))^2 \leq C\,n^{-s_0} + \left(\frac{k}{n}\ r(k)\right)^2 \leq n^{-s_0}\left(C\ + C_0(1+k)^{2-2s_0}n^{s_0-2}\right) \leq n^{-s_0}\,(C+C_0)\ ,$$

where the second inequality follows from (1.14) in (A-3). The same bound gives that, for all $\ell \geq 1$,

$$\mathrm{Var}\ \sum_{j=1}^{\ell} X_j = \sum_{\tau=-\ell+1}^{\ell-1}(\ell-|\tau|)r(\tau) \leq \frac{2C_0}{1-s_0}\,(2+\ell)^{2-s_0}\ , \tag{5.63}$$

and we get, with (2.26), for all $0 \leq k < n$,

$$\mathbb{E}\,|R_n^r(k)| \leq n^{-2}\left(\mathrm{Var}\ \sum_{j=1}^{n} X_j + \left(\mathrm{Var}\ \sum_{j=1}^{n} X_j\right)^{1/2}\left(\mathrm{Var}\ \sum_{j=k+1}^{n-k} X_j\right)^{1/2}\right) \leq \frac{2C_0}{1-s_0}\,(2+n)^{-s_0}\ .$$

With (2.25), we conclude that (2.28) holds. □

Let $\mathbb{T}$ denote $\mathbb{R}/2\pi\mathbb{Z}$. Recall that the spectral measure μ of a weakly stationary process X is a finite measure on $\mathbb{T}$ such that the covariance function of X satisfies

$$r(k) = \int \mathrm{e}^{\mathrm{i}\lambda}\, \mu(\mathrm{d}\lambda)\;, \quad k \in \mathbb{Z}\;.$$

We derive the following useful lemma.

Lemma 8. *Let X be weakly stationary process with spectral measure μ and define the periodogram and the empirical covariance by (1.11) and (2.23). Let $h : \mathbb{R} \to \mathbb{R}$ be a (2π)-periodic bounded function. Then, for all $0 \le m < n$ and $(c_k)_{-m\le k\le m} \in \mathbb{C}^{2m+1}$,*

$$\left|\int_{\mathbb{T}} I_n\, h - \int h\, \mathrm{d}\mu\right| \le \sum_{k=-m}^{m} |\hat{r}_n(k) - r(k)|\; |c_k| + \sup_{\lambda\in\mathbb{R}} \left|h(\lambda) - \sum_{k=-m}^{m} c_k \mathrm{e}^{\mathrm{i}\lambda k}\right| \; (\hat{r}_n(0) + r(0))\;.$$

Proof. Let $0 \le m < n$ and $(c_k)_{-m\le k\le m} \in \mathbb{C}^{2m+1}$, and denote $h_m(\lambda) = \sum_{k=-m}^{m} c_k \mathrm{e}^{\mathrm{i}\lambda k}$ and $\epsilon_m = \sup|h - h_m|$. We write

$$\int_{\mathbb{T}} I_n\, h - \int h\, \mathrm{d}\mu = \int_{\mathbb{T}} I_n\, h_m - \int h_m\, \mathrm{d}\mu + \int_{\mathbb{T}} I_n\, (h - h_m) - \int (h - h_m)\, \mathrm{d}\mu\;.$$

Replacing h_m by its definition, we get

$$\int_{\mathbb{T}} I_n\, h_m - \int h_m\, \mathrm{d}\mu = \sum_{k=-m}^{m} c_k\, (\hat{r}_n(k) - r(k))\;.$$

Then, by definition of ϵ_m, we have

$$\left|\int_{\mathbb{T}} I_n\, (h - h_m)\right| \le \epsilon_m \int_{\mathbb{T}} I_n = \epsilon_m\, \hat{r}_n(0)\;,$$

and, similarly $\left|\int (h - h_m)\right| \mathrm{d}\mu \le \epsilon_m\, r(0)$. The result follows. □

We can no prove Corollary 2.

Proof of Corollary 2. For β small enough, since J is compactly supported, we have, for all $\lambda_0 \in \mathbb{R}$ and $k \in \mathbb{Z}$,

$$c_k(J_{\beta,\lambda_0}) = \frac{1}{2\pi\beta} \int J((\lambda - \lambda_0)/\beta)\, \mathrm{e}^{\mathrm{i}\lambda k}\, \mathrm{d}\lambda = \mathrm{e}^{-\mathrm{i}\lambda_0 k}\, J^*(\beta\, k)\;,$$

where $J^*(\xi) = \int J(x)\, \mathrm{e}^{\mathrm{i}x\xi}\, \mathrm{d}x$ is the Fourier transform of J. Since J is two times continuously differentiable and has compact support, we have $|J^*(\xi)| = \mathcal{O}(|\xi|^{-2})$ as $|\xi| \to \infty$. Hence J_{β,λ_0} has absloutely summable Fourier coefficients and the following identity holds

$$J_{\beta,\lambda_0}(\lambda) = \sum_{k\in\mathbb{Z}} c_k(J_{\beta,\lambda_0})\, \mathrm{e}^{\mathrm{i}\lambda k}\;, \quad \lambda \in \mathbb{R}\;.$$

Applying Lemma 8 with $m = n-1$ and $c_k = c_k(J_{\beta,\lambda_0})$ we get that

$$\left| \int_{\mathbb{T}} I_n \, J_{\beta,\lambda_0} - \int J_{\beta,\lambda_0} \, \mathrm{d}\mu \right| \leq \sum_{k=-n}^{n} |\hat{r}_n(k) - r(k)| \, |J^*(\beta\, k)| + \left(\sum_{|k|>n} |J^*(\beta\, k)| \right) \, (\hat{r}_n(0) + r(0)) \ .$$

Applying (2.28) and $|J^*(\xi)| = \mathcal{O}(|\xi|^{-2})$, we get

$$\left| \int_{\mathbb{T}} I_n \, J_{\beta,\lambda_0} - \int J_{\beta,\lambda_0} \, \mathrm{d}\mu \right| \leq C' \, n^{-s_0/2} \, \left(\sum_{k=-n}^{n} |J^*(\beta\, k)| \right) + C_1 \, (\beta\, n)^{-1} \, r(0)(1 + C') \ ,$$

where C_1 only depends on J. The result then follows from the fact that

$$\lim_{n\to\infty, \beta\to 0} \frac{1}{\beta} \sum_{k=-n}^{n} |J^*(\beta\, k)| = \int |J^*| < \infty \ .$$

□

5.3 Consistency of parametric Whittle estimation

We first introduce some notation valid throughout this section and derive useful lemmas. For any $d \in \mathbb{R}$ and $\epsilon > 0$, we define

$$\underline{f}_{(d,\epsilon)}(\lambda) = \frac{2^{-2(d+\epsilon)}}{2\pi} \, \left| \sin \frac{\lambda}{2} \right|^{-2(d-\epsilon)} \quad \text{and} \quad \overline{f}_{(d,\epsilon)}(\lambda) = \frac{2^{-2(d-\epsilon)}}{2\pi} \, \left| \sin \frac{\lambda}{2} \right|^{-2(d+\epsilon)} \ ,$$

so that, for all $d' \in [d-\epsilon, d+\epsilon]$ and $\lambda \in \mathbb{R}$,

$$\underline{f}_{(d,\epsilon)}(\lambda) \leq f_{d'}(\lambda) = \frac{2^{-2d'}}{2\pi} \, \left| \sin \frac{\lambda}{2} \right|^{-2d'} \leq \overline{f}_{(d,\epsilon)}(\lambda) \ . \tag{5.64}$$

Finally we denote

$$a_K := \inf_{(d,h)\in K} \inf_{\lambda\in\mathbb{R}} \left(f_d(\lambda) + h(\lambda) \right) \ . \tag{5.65}$$

We now introduce the useful lemmas.

Lemma 9. *Let $a > 0$ and $f : \mathbb{R} \to \mathbb{R}$ be (2π)-periodic. Let $g : \mathbb{R} \to \mathbb{R}_+$ be (2π)-periodic and such that $\int g \, \mathrm{dL} > 0$. Then $h \mapsto \ln \int \frac{g}{(f+h)\vee a} \, \mathrm{dL}$ is $(1/a)$-Lipschitz on $\mathcal{C}$. If moreover $\int \ln(f \vee a) \, \mathrm{dL} < \infty$, $h \mapsto \int \ln((f+h) \vee a) \, \mathrm{dL}$ is also $(1/a)$-Lipschitz on $\mathcal{C}$.*

Proof. We apply successively that, for all $0 < x \leq y$,

$$\left| \frac{1}{x} - \frac{1}{y} \right| = \frac{|y-x|}{xy} \quad \text{and} \quad |\ln y - \ln x| \leq \frac{|y-x|}{x} \ .$$

We obtain, for all $h, \tilde{h} \in \mathcal{C}$,

$$\left| \int \frac{g}{(f_d + h) \vee a} \, \mathrm{dL} - \int \frac{g}{(f_d + \tilde{h}) \vee a} \, \mathrm{dL} \right| \leq \int \frac{g \, |h - \tilde{h}|}{((f_d + h) \vee a) \, ((f_d + \tilde{h}) \vee a)} \, \mathrm{dL}$$
$$\leq \frac{1}{a} \sup |h - \tilde{h}| \int \frac{g}{((f_d + h) \vee a)} \, \mathrm{dL}$$
$$\text{or } \leq \frac{1}{a} \sup |h - \tilde{h}| \int \frac{g}{((f_d + \tilde{h}) \vee a)} \, \mathrm{dL} \, ,$$

hence is bounded from above by the min of the two last right-hand sides. Taking the difference of the log's then yields

$$\left| \ln \int \frac{g}{(f_d + h) \vee a} \, \mathrm{dL} - \ln \int \frac{g}{(f_d + \tilde{h}) \vee a} \, \mathrm{dL} \right| \leq \frac{1}{a} \sup |h - \tilde{h}| \, .$$

Hence we get the first assertion.

Similarly, we get that, for all $h, \tilde{h} \in \mathcal{C}$,

$$\left| \ln((f_d + h) \vee a) - \ln((f_d + \tilde{h}) \vee a) \right| \leq \frac{1}{a} \sup |h - \tilde{h}| \, .$$

And we get the second assertion. □

Lemma 10. *Let $d^* < 1/2$, $h^* \in \mathcal{C}$ and $a > 0$. For all $d \in \mathbb{R}$ and $h \in \mathcal{C}$, we have*

$$\lim_{\epsilon \to 0} \ln \int \frac{f_{d^*} + h^*}{(f_{(d,\epsilon)} + h) \vee a} \, \mathrm{dL} = \ln \int \frac{f_{d^*} + h^*}{(f_d + h) \vee a} \, \mathrm{dL} \, ,$$
$$\lim_{\epsilon \to 0} \int \ln((f_{(d,\epsilon)} + h) \vee a) \, \mathrm{dL} = \int \ln((f_d + h) \vee a) \, \mathrm{dL} \, ,$$

where $f_{(d,\epsilon)}$ denotes either $\underline{f}_{(d,\epsilon)}$ or $\overline{f}_{(d,\epsilon)}$.

Proof. We have, for all $\lambda \notin 2\pi\mathbb{Z}$, as $\epsilon \to 0$, $f_{(d,\epsilon)}(\lambda) \to f_d(\lambda)$. Moreover, for $\epsilon \in (0, 1)$,

$$\frac{f_{d^*} + h^*}{(f_{(d,\epsilon)} + h) \vee a} \leq \frac{1}{a} \, (f_{d^*} + h^*) \quad \text{and} \quad \ln a \leq \ln((f_{(d,\epsilon)} + h) \vee a) \leq \ln((\overline{f}_{(d,1)} + h) \vee a) \, .$$

We conclude by dominated convergence. □

Lemma 11. *Let $d^* < 1/2$, $h^* \in \mathcal{C}$ and $a > 0$. For all $d \in \mathbb{R}$ and $h \in \mathcal{C}$, we have*

$$\lim_{(d', \tilde{h}) \to (d,h)} \ln \int \frac{f_{d^*} + h^*}{(f_{d'} + \tilde{h}) \vee a} \, \mathrm{dL} = \ln \int \frac{f_{d^*} + h^*}{(f_d + h) \vee a} \, \mathrm{dL} \, ,$$
$$\lim_{(d', \tilde{h}) \to (d,h)} \int \ln((f_{d'} + \tilde{h}) \vee a) \, \mathrm{dL} = \int \ln((f_d + h) \vee a) \, \mathrm{dL} \, .$$

Proof. By Lemma 9, we have for all $d' \in \mathbb{R}$ and $h, \tilde{h} \in \mathcal{C}$,

$$\left| \ln \int \frac{f_{d^*} + h^*}{(f_{d'} + \tilde{h}) \vee a} \, \mathrm{dL} - \ln \int \frac{f_{d^*} + h^*}{(f_{d'} + h) \vee a} \, \mathrm{dL} \right| \leq \frac{1}{a} \sup |h - \tilde{h}| \, .$$

Then using (5.64) and Lemma 10, we get the first assertion. The second assertion is proved similarly. □

Lemma 12. *Let K be a compact subset of $\mathbb{R}\times\mathcal{C}$ such that, for all $(d,h)\in K$, $f_d+h>0$ on $\mathbb{R}$. Suppose moreover that*

$$(\{0\}\times\mathcal{C})\cap \mathrm{Cl}_K(\mathbb{R}^*_-\times\mathcal{C})=\emptyset\ , \tag{5.66}$$

where for any $A\subset\mathbb{R}\times\mathcal{C}$, $\mathrm{Cl}_K(A)$ denotes the closure of $A\cap K$ in K. Then we have $\inf_{(d,h)\in K}\inf(f_d+h)>0$.

Proof. Let

$$C_K=\sup_{(d,h)\in K,\lambda\in\mathbb{R}}|h(\lambda)|\ ,$$

which is finite since K is compact and $(h,\lambda)\mapsto h(\lambda)$ continuous.

We will need the following remark. Let $\epsilon\in(0,\pi)$. We have, for all $d>0$ and $\lambda\in[-\epsilon,\epsilon]$,

$$f_d(\lambda)=\frac{2^{-2d}}{2\pi}\ |\sin(\lambda/2)|^{-2d}\geq\frac{\epsilon^{-2d}}{2\pi}\ .$$

Then we get, if $\epsilon^{-2d}/(2\pi)\geq 2C_K+1$, which is equivalent to $d\geq\ln(2\pi(2C_K+1))/(-2\ln\epsilon)$, for all $(d,h)\in K$ and $\lambda\in[-\epsilon,\epsilon]$,

$$f_d(\lambda)+h(\lambda)\geq\epsilon^{-2d}/(2\pi)-C_K\geq C_K+1\ . \tag{5.67}$$

Let (d_n,h_n,λ_n) be a sequence valued in $K\times[-\pi,\pi]$ such that

$$a_K=\lim_{n\to\infty}\left(f_{d_n}(\lambda_n)+h_n(\lambda_n)\right)\ ,$$

where a_K is defined by (5.65). By compactness, there is an increasing sequence of integers (q_n) and $(d,h,\lambda)\in K\times[-\pi,\pi]$ such that $(d_{q_n},h_{q_n},\lambda_{q_n})$ converges to (d,h,λ). We now separate four cases, which cover all possible cases: 1) $\lambda\neq 0$, 2) $d<0$, 3) $\lambda=0$ and $d>0$ 4) $\lambda=d=0$.
Case 1) Suppose that $\lambda\neq 0$. Since the mapping $(d',\tilde{h},\lambda')\mapsto f_{d'}(\lambda')+\tilde{h}(\lambda')$ is continuous on $\mathbb{R}\times\mathcal{C}\times(\mathbb{R}\setminus(2\pi\mathbb{Z}))$, we get that $a_K=f_d(\lambda)+h(\lambda)>0$.
Case 2) Suppose that $d<0$. This case is similar to Case 1): it is sufficient to show that $(d',\tilde{h},\lambda')\mapsto f_{d'}(\lambda')+\tilde{h}(\lambda')$ is continuous on $(-\infty,0)\times\mathcal{C}\times\mathbb{R}$, which follows from the continuity of $(x,u)\mapsto u^x$ on $(0,\infty)\times\mathbb{R}_+$, which is easy to establish.
Case 3) Suppose that $\lambda=0$ and $d>0$. Then there exists an arbitrarily small $\epsilon\in(0,\pi)$ such that $d\geq\ln(2\pi(2C_K+1))/(-4\ln\epsilon)$, and (5.67) implies, for n large enough, $f_{d_{q_n}}(\lambda_{q_n})+h_{q_n}(\lambda_{q_n})\geq C_K+1$ hence $a_K>0$.
Case 4) Suppose that $\lambda=d=0$. Thanks to Condition (5.66), d_{q_n} must be non-negative for n large enough, in which case we have

$$f_{d_{q_n}}(\lambda_{q_n})=\frac{2^{-2d_{q_n}}}{2\pi}\ |\sin(\lambda_{q_n}/2)|^{-2d_{q_n}}\geq\frac{2^{-2d_{q_n}}}{2\pi}\ ,$$

which tends to $1/(2\pi)$ as $n \to \infty$, while $h_{q_n}(\lambda_{q_n})$ tends to $h(0)$. Hence

$$a_K = \lim_{n\to\infty} f_{d_{q_n}}(\lambda_{q_n}) + h_{q_n}(\lambda_{q_n}) \geq 1/(2\pi) + h(0) = f_d(\lambda) + h(\lambda) \ ,$$

since $d = \lambda = 0$. Again, we get $a_K > 0$. □

Remark 3. *Condition (5.66) means that any parameter $(0, h)$ in K is isolated from parameters $(d, \tilde{h})$ with $d < 0$. This assumption cannot be avoided in Lemma 12.(As a counterexample, take $K = \{(d, -d) \ : \ d \in [-1/2, 0]\}$ where here $-d$ is seen as the function in $\mathcal{C}$ that is constant equal to $-d$). It is of course trivially satisfied if $K \subset \mathbb{R}_+ \times \mathcal{C}$.*

We can now proceed with the proofs of Lemma 3 and Theorem 2.

Proof of Lemma 3. We first recall why $H(s, C)$ is a compact subset of $\mathcal{C}$. For all $u, v \in \mathbb{R}$, we have

$$|\mathrm{e}^{\mathrm{i}ku} - \mathrm{e}^{\mathrm{i}kv}| = \left| \int_0^{|u-v|} \mathrm{i}k \, \mathrm{e}^{\mathrm{i}kx} \, \mathrm{d}x \right| \leq k \, |u - v| \ .$$

We get that, for all $h \in H(s, C)$ and $u, v \in \mathbb{R}$,

$$|h(u) - h(v)| \leq C \left(|u - v| \sum_{|k| \leq |u-v|} (1 + |k|)^{-1-s} \, k + 2 \sum_{|k| > |u-v|} (1 + |k|)^{-1-s} \right) = \mathcal{O}(|u-v|^{s \wedge 1}) \ ,$$

where the $\mathcal{O}$ does not depend on h. By the Arzelà–Ascoli theorem, we get that $H(s, C)$ is a compact subset of $\mathcal{C}$. It follows that A, as a closed subset of $[0, 1/2] \times \mathcal{C}$ is also compact. Thus by Lemma 12, there exists $a_K > 0$ such that $f_d + h \geq a_K$ for all $(d, h) \in A$. Since $\sup(|p_m(h) - h|)$ tends to 0 uniformly in $h \in H(s, C)$ as $m \to \infty$, we get that there exists a positive integer m_0 such that $f_d + p_m[h] \geq a_K/2 > 0$ on $\mathbb{R}$ for all $(d, h) \in A$ and $m \geq m_0$. Let K_n and K be defined as in the lemma for some diverging sequence (m_n) of integers larger than or equal to m_0. It is straightforward to show that K is compact (because for any increasing or constant sequence $(\alpha_k)_{k\in\mathbb{N}}$ of integers and any sequence $((d_k, h_k))_{k\in\mathbb{N}}$ valued and converging in A, we have that $(d_k, p_{m_{\alpha_k}}[h_k])$ converges in K). Assumption (A-4) easily follows by setting $h_n^* = p_{m_n}[h^*]$. □

Proof of Theorem 2. Define

$$\Lambda^* := \ln c^* + \int \ln(f_{d^*} + h^*) \, \mathrm{dL} \ .$$

By (A-4), we have $(d^*, h_n^*) \in K_n$. Thus Equation (2.31) implies $(\hat{d}_n, \hat{h}_n) \to (d^*, h^*)$ in $[0, 1/2] \times \mathcal{C}$ a.s. provided that

$$\limsup_{n\to\infty} \Lambda_n(d^*, h_n^*) \leq \Lambda^* \quad \text{a.s.} \tag{5.68}$$

and that, for any $\epsilon_0 > 0$, we have

$$\liminf_{n\to\infty} \inf_{(d,h)\in K_0} \Lambda_n(d,h) > \Lambda^* \quad \text{a.s.} , \tag{5.69}$$

where K_0 is defined by

$$K_0 = \{(d,h) \in K \ : \ |d-d^*| + \sup|h-h^*| \geq \epsilon_0\} \ ,$$

We start with the proof of (5.68). By definition of a_K in (5.65), we have, for all $(d,h) \in K$,

$$\Lambda^* := \ln c^* + \int \ln((f_{d^*} + h^*) \vee a_K) \, \mathrm{dL} \ , \tag{5.70}$$

$$\Lambda_n(d,h) = \ln \int \frac{I_n}{(f_d + h) \vee a_K} \, \mathrm{dL} \ + \int \ln\left((f_d + h) \vee a_K\right) \, \mathrm{dL} \ . \tag{5.71}$$

And by Lemma 12, $a_K > 0$. Note that $\int I_n \, \mathrm{dL} = \hat{r}_n(0)/(2\pi) > 0$ for n large enough, a.s. Applying Lemma 9 with $a = a_K$ and since h_n^* converges to h^* uniformly by (A-4), we get

$$\lim_{n\to\infty} |\Lambda_n(d^*, h_n^*) - \Lambda_n(d^*, h^*)| = 0 \quad \text{a.s.}$$

Since $1/((f_{d^*}+h^*)\vee a_K)$ is continuous and X is ergodic with spectral density f given by (1.10), we have

$$\lim_{n\to\infty} \int \frac{I_n}{(f_{d^*} + h^*) \vee a_K} \, \mathrm{dL} = \int \frac{c^*(f_{d^*} + h^*)}{(f_{d^*} + h^*) \vee a_K} \, \mathrm{dL} \quad \text{a.s.}$$

(see e.g. [Giraitis et al., 2012, Theorem 8.2.1]). By definition of a_K, this limit is c^* and, with the three previous displayed equation, we get (5.68).

We conclude with the proof of (5.69), given some $\epsilon_0 > 0$. Equations (5.71) and (5.64) and Lemma 9 with $a = a_K$ yield, for all (d,h) and $(d',\tilde{h})$ in K such that $|d-d'| \leq \epsilon$ and $\sup|h-\tilde{h}| \leq \epsilon$,

$$\Lambda_n(d',\tilde{h}) \geq \ln \int \frac{I_n}{(\overline{f}_{(d,\epsilon)} + h) \vee a_K} \, \mathrm{dL} \ + \int \ln\left((\underline{f}_{(d,\epsilon)} + h) \vee a_K\right) \, \mathrm{dL} - \frac{2\epsilon}{a_K} \ .$$

Since $1/((\overline{f}_{(d,\epsilon)} + h) \vee a_K)$ is continuous and X is ergodic, we have

$$\lim_{n\to\infty} \int \frac{I_n}{(\overline{f}_{(d,\epsilon)} + h) \vee a_K} \, \mathrm{dL} = \int \frac{c^*(f_{d^*} + h^*)}{(\overline{f}_{(d,\epsilon)} + h) \vee a_K} \, \mathrm{dL} \quad \text{a.s.}$$

The last two displays give that, for all $(d,h) \in K$ and $\epsilon > 0$,

$$\begin{aligned} \liminf_{n\to\infty} \inf \Big\{ \Lambda_n(d',\tilde{h}) \ : \ (d',\tilde{h}) \in K, \, |d'-d| \leq \epsilon, \, \sup|h-\tilde{h}| \leq \epsilon \Big\} \\ \geq \ln \int \frac{c^*(f_{d^*} + h^*)}{(\overline{f}_{(d,\epsilon)} + h) \vee a_K} \, \mathrm{dL} \ + \int \ln\left((\underline{f}_{(d,\epsilon)} + h) \vee a_K\right) \, \mathrm{dL} - \frac{2\epsilon}{a_K} \quad \text{a.s.} \end{aligned} \tag{5.72}$$

For all $(d,h) \in K_0$, since $f_d + h$ and $f_{d^*} + h^*$ do not coincide almost everywhere, the Jensen inequality and the definition of Λ^* give that

$$\ln \int \frac{c^*(f_{d^*} + h^*)}{(f_d + h) \vee a_K} \, \mathrm{dL} + \int \ln((f_d + h) \vee a_K) \, \mathrm{dL} > \Lambda^* \, .$$

Since K_0 is compact, by Lemma 11, we can find $\eta > 0$ such that

$$\inf_{(d,h) \in K_0} \ln \int \frac{c^*(f_{d^*} + h^*)}{(f_d + h) \vee a_K} \, \mathrm{dL} + \int \ln((f_d + h) \vee a_K) \, \mathrm{dL} \geq \Lambda^* + 4\eta \, .$$

Applying Lemma 10 with $a = a_K$, we get that, for all $(d,h) \in K_0$, there exists $\epsilon > 0$ such that

$$\begin{aligned} \ln \int \frac{c^*(f_{d^*} + h^*)}{(\overline{f}_{(d,\epsilon)} + h) \vee a_K} \, \mathrm{dL} + \int \ln((\underline{f}_{(d,\epsilon)} + h) \vee a_K) \, \mathrm{dL} &\geq \ln \int \frac{c^*(f_{d^*} + h^*)}{(f_d + h) \vee a_K} \, \mathrm{dL} \\ &\quad + \int \ln((f_d + h) \vee a_K) \, \mathrm{dL} - 2\eta \\ &\geq \Lambda^* + 2\eta \, . \end{aligned}$$

Since K_0 is compact, we can thus cover K_0 with a finite collection $(B_i)_{i=1,\dots,N}$, for which, for any $i = 1, \dots, N$, there exists $\epsilon_i \in (0, \eta \, a_K/2)$ and $(d_i, h_i) \in K_0$ such that

$$\ln \int \frac{c^*(f_{d^*} + h^*)}{(\overline{f}_{(d_i,\epsilon_i)} + h_i) \vee a_K} \, \mathrm{dL} + \int \ln((\underline{f}_{(d_i,\epsilon_i)} + h_i) \vee a_K) \, \mathrm{dL} \geq \Lambda^* + 2\eta \, ,$$

and $(d,h) \in B_i$ implies $|d - d_i| \leq \epsilon_i$ and $\sup |h - h_i| \leq \epsilon_i$. Let $i = 1, \dots, N$. Applying (5.72) with $d = d_i$, $h = h_i$ and $\epsilon = \epsilon_i$, we get that

$$\liminf_{n \to \infty} \inf_{(d',\tilde{h}) \in B_i} \Lambda_n(d', \tilde{h}) \geq \Lambda^* + \eta \quad \text{a.s.}$$

Since $(B_i)_{i=1,\dots,N}$ covers K_0, we obtain (5.69).

We thus have proved that $(\hat{d}_n, \hat{h}_n) \to (d^*, h^*)$ in $[0, 1/2] \times \mathcal{C}$ a.s. and it only remains to show that $\hat{c}_n \to c^*$ a.s., where, by (1.13) and the definition of a_K,

$$\hat{c}_n = \int \frac{I_n}{(f_{\hat{d}_n} + \hat{h}_n) \vee a_K} \, \mathrm{dL} \, .$$

Let $\epsilon > 0$ and suppose that $|\hat{d}_n - d^*| \leq \epsilon$ and $\sup |\hat{h}_n - h^*| \leq \epsilon$, which happens for n large enough, a.s. By (5.64) and Lemma 9 successively, we get that

$$\ln \int \frac{I_n}{(\overline{f}_{(d^*,\epsilon)} + h^*) \vee a_K} \, \mathrm{dL} - \frac{\epsilon}{a_K} \leq \ln \hat{c}_n \leq \ln \int \frac{I_n}{(\underline{f}_{(d^*,\epsilon)} + h^*) \vee a_K} \, \mathrm{dL} + \frac{\epsilon}{a_K}$$

Letting n tend to ∞ and then ϵ to zero (using Lemma 10), we get that

$$\lim_{n \to \infty} \ln \hat{c}_n = \ln \int \frac{c^*(f_{d^*} + h^*)}{(f_{d^*} + h^*) \vee a_K} \, \mathrm{dL} \quad \text{a.s.}$$

By definition of a_K, the latter integral is c^* and the proof is concluded. □

5.4 The arfima case: Proof of Corollary 3

Denote, for any $d \in [0, 1/2)$ and $\vartheta = (\phi, \theta) \in \Theta_{p,q}$,

$$f_{d,\vartheta,\sigma^2}(\lambda) = \sigma^2 \, f_d(\lambda) \left| \frac{\Theta(\mathrm{e}^{-\mathrm{i}k\lambda})}{\Phi(\mathrm{e}^{-\mathrm{i}k\lambda})} \right|^2 .$$

From the discussion preceding Corollary 3 and leading to (2.35), we write

$$f_{d,\vartheta,\sigma^2} = \sigma^2 \frac{|\Theta(1)|^2}{|\Phi(1)|^2} \left(f_d + f_d \mathrm{R}(\vartheta) \right) ,$$

with $\mathrm{R}(\vartheta)$ defined for all $\vartheta \in \Theta_{p,q}$ as the $\mathbb{R} \to \mathbb{R}$ (2π)-periodic function

$$[\mathrm{R}(\vartheta)](\lambda) = \left| \frac{\Theta(\mathrm{e}^{-\mathrm{i}\lambda})\Phi(1)}{\Phi(\mathrm{e}^{-\mathrm{i}\lambda})\Theta(1)} \right|^2 - 1 .$$

Define the mapping

$$\begin{aligned} \Psi : [0, 1/2) \times \Theta_{p,q} &\to [0, 1/2) \times \mathcal{C} \\ (d, \vartheta) &\mapsto (d, f_d \mathrm{R}(\vartheta)) \end{aligned}$$

and denote by $K = \Psi(\tilde{K})$ its range over $\tilde{K}$. The following facts are established at the end of this proof section.

(i) We have $\widetilde{\Lambda}_n = \Lambda_n \circ \Psi$ over $[0, 1/2) \times \Theta_{p,q}$.

(ii) We have, for all $d \in [0, 1/2)$, $\vartheta = (\phi, \theta) \in \Theta_{p,q}$ and $\sigma^2 > 0$,

$$\exp \circ \widetilde{\Lambda}_n(d, \vartheta) = \left| \frac{\Phi(1)}{\Theta(1)} \right|^2 \int \frac{I_n}{f_d + f_d \mathrm{R}(\vartheta)} \, \mathrm{dL} .$$

(iii) The mapping R is continuous and one-to-one on $\Theta_{p,q}$. We denote by R^{-1} its inverse, which is continuous on $\mathrm{R}(K_0)$ for all compact subset $K_0 \subset \Theta_{p,q}$.

Then the condition (2.37) defining $(\hat{d}_n, \hat{\vartheta}_n)$ is equivalent to have (2.31) with $K_n = K$ and $\hat{h}_n = f_{\hat{d}_n} \mathrm{R}(\hat{\vartheta}_n)$ (and the same $\hat{d}_n$). To apply Theorem 2 on this sequence $(\hat{d}_n, \hat{h}_n)$, we need to check that (A-4) holds with $K_n = K$ and $h_n^* = h^* = f_{d^*} R_{\vartheta^*}$ (as in (2.35)) for all n. In this case, only the compactness of K is non-trivial and since $\tilde{K}$ is compact, this compactness follows from the following assertion:

(a) The mapping Ψ is continuous.

Assuming this fact proven, we can apply Theorem 2 and get that $(\hat{d}_n, \hat{h}_n)$ converges a.s. to $(d^*, h^*) = \Psi(d^*, \vartheta^*)$ (that is, h^* as in (2.35)). Also by Fact (ii) above, with (1.13), (2.36)

and (2.38), if we can apply Theorem 2, then we also get that $\hat{\sigma}_n^2$ is a consistent estimator of σ_*^2. Finally, it only remains to explain how to get that $\hat{\vartheta}_n$ converges to ϑ^* a.s. This follows from the assertion that $(\hat{d}_n, \hat{h}_n) = \Psi(\hat{d}_n, \hat{\vartheta}_n)$ converges a.s. to $(d^*, h^*) = \Psi(d^*, \vartheta^*)$, provided that Ψ can be continuously inversed on K. To summarize, to conclude the proof, we only need to prove the following assertion.

(b) The mapping Ψ is bijective and bi-continuous from $\tilde{K}$ to K (its range).

Define the mapping

$$\begin{aligned} \mathrm{A} : (-1/2, 1/2) \times \mathcal{A} &\to \mathcal{C} \\ (d, h) &\mapsto f_d h \ , \end{aligned}$$

where, for any $C > 0$,

$$\mathcal{A}(C) = \left\{ h \in \mathcal{C} \ : \ \sup_{t \in \mathbb{R}^*} |h(t)/t| \leq C \right\} \text{ and } \mathcal{A} = \left(\bigcup_{C>0} \mathcal{A}(C) \right) \ .$$

Note that we have, for all $d \in [0, 1/2)$, $h \in \mathcal{A}$ and $\vartheta \in \Theta_{p,q}$,

$$\begin{aligned} \Psi(d, \vartheta) &= (d, \mathrm{A}(d, \mathrm{R}(\vartheta))) \ , \\ \Psi\left(d, \mathrm{R}^{-1}\left(\mathrm{A}(-d, h)\right)\right)\big) &= (d, h) \ . \end{aligned}$$

Hence Assertion (b) follows from Assertion (iii) among with the following facts.

(iv) For all compact subset $K_0 \subset \Theta_{p,q}$ there exists $C > 0$ such that the range $\mathrm{R}(K_0) \subset \mathcal{A}(C)$.

(v) For any $C > 0$, A is continuous on $(-1/2, 1/2) \times \mathcal{A}(C)$.

Proof of Assertion (i): This follows directly from the definitions of Λ_n and $\tilde{\Lambda}_n$ in (1.12) and (2.34) and the well known fact that, for all $d \in (-1/2, 1/2)$ and $\vartheta \in \Theta_{p,q}$,

$$\int \ln(f_d + f_d \mathrm{R}(\vartheta)) \, \mathrm{dL} = \ln \left| \frac{\Phi(1)}{\Theta(1)} \right|^2 \ .$$

Proof of Assertion (ii): This is simple algebra using the above definitions.

Proof of Assertions (iii) and (iv): Using standard properties of canonical ARMA processes, we have, for all $\vartheta = (\phi, \theta) \in \Theta_{p,q}$,

$$[\mathrm{R}(\vartheta)](\lambda) = \left| \frac{\Phi(1)}{\Theta(1)} \right|^2 \sum_{k \geq 1} \alpha_k(\vartheta) \ (\cos(k\lambda) - 1) \ ,$$

where for any $k \geq 1$, the mapping $\vartheta \mapsto \alpha_k(\vartheta)$ is polynomial and for any compact subset $K_0 \subset \Theta_{p,q}$, there exists $C_0 > 0$ and $\rho_0 \in (0, 1)$ such that, for all $k \geq 1$,

$$|\alpha_k(\vartheta)| \leq C_0 \, \rho^k \ .$$

Assertion (iv) easily follows as well as the continuity of R over $\Theta_{p,q}$. Also since $\mathrm{R}(\vartheta)+1$ is the spectral density of the ARMA(p,q) process with ARMA polynomials Φ and Θ, it is obvious that R is one-to-one on $\Theta_{p,q}$. Let K_0 be a compact subset of $\Theta_{p,q}$. Then for all $h \in \mathrm{R}(K_0)$, using standard arguments, we can express the reciprocal $\mathrm{R}^{-1}(h)$ by

$$\mathrm{R}^{-1}(h) = \arg\min_{\vartheta \in K_0} \mathcal{K}(h;\vartheta) \ ,$$

where, for all $\vartheta = (\phi,\theta) \in \Theta_{p,q}$ and $\tilde{h} \in \mathcal{C}$,

$$\mathcal{K}(\tilde{h};\vartheta) = \int (\tilde{h}(\lambda)+1) \left| \frac{\Phi(\mathrm{e}^{-\mathrm{i}\lambda})}{\Theta(\mathrm{e}^{-\mathrm{i}\lambda})} \right|^2 \mathrm{L}(\mathrm{d}\lambda) \ .$$

Since $(\tilde{h},\vartheta) \mapsto \mathcal{K}(\tilde{h};\vartheta)$ is continuous on $\mathcal{C} \times \Theta_{p,q}$, we get that R^{-1} is obviously continuous on $R^{-1}(K_0)$ and Assertion (iii) is proved.

Proof of Assertion (v): Let $d \in (-1/2, 1/2)$ and $h \in \mathcal{A}(C)$ for some positive constant C. Let $\epsilon > 0$ such that $[d-\epsilon, d+\epsilon] \subset (-1/2, 1/2)$. Let $\eta \in (0, \pi/2)$. Then, for all $d' \in [d-\epsilon, d+\epsilon]$ and $\tilde{h} \in \mathcal{A}(C)$, we have, using (5.64),

$$\sup_{|\lambda| \le \eta} \left| f_d(\lambda)\, h(\lambda) - f_{d'}(\lambda)\, \tilde{h}(\lambda) \right| \le 2\, C \sup_{|\lambda| \le \eta} \left(\overline{f}_{(d,\epsilon)}(\lambda)\, |\lambda| \right) \ ,$$

which tends to 0 as $\eta \to 0$. On the other hand, we clearly have that

$$\lim_{(d',\tilde{h}) \to (d,h)} \sup_{\eta \le |\lambda| \le \pi} \left| f_d(\lambda)\, h(\lambda) - f_{d'}(\lambda)\, \tilde{h}(\lambda) \right| = 0 \ .$$

The last two displays yield (v), which concludes the proof.

6 Numerical experiments

6.1 Simulated trawl processes

We take a sample of exponents $\alpha^* \in (1,2) : \alpha^* \in \{1.1, 1.3, 1.5, 1.7, 1.9\}$, and for each of them, we generate two trawl processes obtained from two different sequences (a_k) and two different seed processes:

1. Poisson seed: $\gamma(t)$ is a homogeneous Poisson counting process with unit intensity and power sequence : $a_k = 10.\,(k+1)^{-\alpha^*}$.
2. Binomial seed with $n = 10$ (see Section 3.3) and power sequence : $a_k = (k+1)^{-\alpha^*}$.

We showed in Section 3.3 that the spectral densities of the two resulting trawl processes are of the form (1.10) with $(d^*, h^*) \in [0, 1/2] \times H(\alpha^* - 1, C)$, for the Poisson (thus Lévy) seed and the Binomial seed.

6.2 Estimation of the trawl exponent

To test our new estimator, we will compare it with local Whittle estimator for long range dependent sequences ([Robinson, 1995]). First, let us recall the definition of the local Whittle estimator.

Here the Hurst exponent is $H = (3-\alpha)/2$ and the spectral density writes

$$f(\lambda) = \frac{c_0}{\Gamma(\alpha)\cos\left(\frac{\pi(3-\alpha)}{2}\right)}\lambda^{\alpha-2}(1+o(\lambda)),$$

Let $\lambda_j = 2j\pi/n$ denote the canonical frequencies for $1 \leq j \leq n/2$, where n is the sample size. The local Whittle contrast is defined for a given bandwidth parameter $m \leq n/2$ by

$$R(\alpha) = \ln \widehat{G}(\alpha) + \frac{\alpha-2}{m}\sum_{j=1}^{m}\ln\lambda_j,\ \widehat{G}(\alpha) = \frac{1}{m}\sum_{j=1}^{m}\lambda_j^{\alpha-2}I_n(\lambda_j)\ ,$$

where I_n is the usual periodogram, see (1.11). Then the local Whittle estimator $\hat{\alpha}_{\mathrm{LW}}$ is computed through numerical minimization of $R(\alpha)$ over $\alpha \in [1,2]$. In the non-linear case, such as trawl processes with Poisson or binomial seed, the use of such an estimator is theoretically justified in[Dalla et al., 2005] under the assumption $\lim_{n\to\infty}\left(\frac{m}{n}+\frac{1}{m}\right)=0$.

The parametric Whittle estimator that we use is based on the parameterization (1.10). Thus we set $\hat{\alpha}_{\mathrm{PW}} = 2(1-\hat{d}_n)$ with $\hat{d}_n$ the estimator obtained through numerical minimization of $\Lambda_n(d,h)$ defined by (1.12) over $d \in [0,1/2]$ and $h \in \mathcal{P}_N$ (see Section 2.3), for a given N.

In our setting, both the local Whittle estimator $\hat{\alpha}_{\mathrm{LW}}$ and the parametric Whittle estimator $\hat{\alpha}_{\mathrm{PW}}$ rely on tuning parameters, respectively denoted by m and N. Observe that N and m have very different interpretations. As the bandwidth parameter m increases, a larger range of frequencies is used in the estimation, thus reducing the variance, and the estimator relies on the approximation $f(\lambda) \approx c\lambda^{\alpha-2}$ also over a larger range of frequencies, thus worsening the bias. In contrast, as N increases, we expect the variance to increase, since the number of parameters to estimate for h is larger, and the bias to decrease, since the approximation of h by a trigonometric polynomial is more accurate.

6.3 Results

We show here the comparison of the two estimators. We have to guess the hyperparameter of the two estimators: The "m" for the local Whittle and the number "N" of Fejér kernels for the parametric estimator. We give our results in function of the choice of these hyperparameters. For each experiment, we write in bold the choice of hyperparameters minimizing the sum of the square of the bias and the variance (the mean square error). In all cases, but especially for the Binomial seed, we can see in the following tables that our estimator outperforms the

local Whittle estimator. A right choice for the number of kernels seems to be around between 3 and 5, even if, best results may be obtained for higher number of kernels, but this may be due to local minima reached by numerical optimization.

Table 1: Estimation results for a local Whittle estimator, when $\alpha \in \{1.1, 1.3, 1.5, 1.7, 1.9\}$, 5000 observations, $m \in \{20, 50, 100, 200\}$, 100 replications.

α	Statistic	Poisson seed				Binomial seed			
	m	20	50	100	200	20	50	100	200
1.1	bias($\widehat{\alpha}_{\mathrm{LW}}$)	0.0679	0.0179	-0.0193	**-0.0563**	0.0441	-0.0332	**-0.0776**	-0.0953
	sd($\widehat{\alpha}_{\mathrm{LW}}$)	0.1973	0.1296	0.0854	**0.0513**	0.1943	0.1008	**0.0467**	0.0174
1.3	bias($\widehat{\alpha}_{\mathrm{LW}}$)	0.0305	-0.0072	**-0.0495**	-0.0941	-0.0887	**-0.1294**	-0.1795	-0.2352
	sd($\widehat{\alpha}_{\mathrm{LW}}$)	0.2697	0.168	**0.1078**	0.0765	0.2383	**0.144**	0.0969	0.0635
1.5	bias($\widehat{\alpha}_{\mathrm{LW}}$)	-0.0374	-0.0726	**-0.1053**	-0.1402	-0.0513	**-0.1595**	-0.224	-0.2932
	sd($\widehat{\alpha}_{\mathrm{LW}}$)	0.2939	0.1837	**0.111**	0.075	0.2861	**0.166**	0.1051	0.0872
1.7	bias($\widehat{\alpha}_{\mathrm{LW}}$)	-0.1025	-0.1447	**-0.17**	-0.2074	-0.1118	**-0.1998**	-0.2529	-0.336
	sd($\widehat{\alpha}_{\mathrm{LW}}$)	0.2594	0.1656	**0.1125**	0.0822	0.2509	**0.1841**	0.1159	0.0786
1.9	bias($\widehat{\alpha}_{\mathrm{LW}}$)	-0.1658	-0.1954	**-0.2148**	-0.2644	-0.1955	**-0.2588**	-0.3238	-0.4069
	sd($\widehat{\alpha}_{\mathrm{LW}}$)	0.2238	0.1465	**0.1125**	0.0804	0.2614	**0.1802**	0.1265	0.0859

Table 2: Estimation results for the parametric Whittle estimator and Poisson seed, when $\alpha \in \{1.1, 1.3, 1.5, 1.7, 1.9\}$, 5000 observations, $N \in \{2, 3, 4, 5, 6, 7, 8, 9\}$, 100 replications.

α	Statistic	Poisson seed							
	N	2	3	4	5	6	7	8	9
1.1	bias($\widehat{\alpha}_{\mathrm{PW}}$)	**-0.0037**	0.0524	0.1075	0.157	0.206	0.2595	0.2993	0.2927
	sd($\widehat{\alpha}_{\mathrm{PW}}$)	**0.058**	0.0754	0.0925	0.1114	0.1309	0.1376	0.1585	0.1552
1.3	bias($\widehat{\alpha}_{\mathrm{PW}}$)	**-0.0072**	0.0803	0.1483	0.2282	0.3085	0.3639	0.362	0.3233
	sd($\widehat{\alpha}_{\mathrm{PW}}$)	**0.0694**	0.0922	0.1117	0.1361	0.1569	0.1574	0.1308	0.1288
1.5	bias($\widehat{\alpha}_{\mathrm{PW}}$)	**0.0159**	0.1045	0.1974	0.2766	0.3353	0.356	0.2743	0.1949
	sd($\widehat{\alpha}_{\mathrm{PW}}$)	**0.0817**	0.1162	0.1381	0.1439	0.1312	0.1073	0.0918	0.0788
1.7	bias($\widehat{\alpha}_{\mathrm{PW}}$)	**0.0199**	0.1249	0.1959	0.2305	0.2544	0.2383	0.125	0.0781
	sd($\widehat{\alpha}_{\mathrm{PW}}$)	**0.102**	0.1144	0.0933	0.0774	0.0558	0.0441	0.0788	0.0836
1.9	bias($\widehat{\alpha}_{\mathrm{PW}}$)	0.0199	0.0653	0.0796	0.0841	0.0814	**0.0469**	-0.0201	-0.0696
	sd($\widehat{\alpha}_{\mathrm{PW}}$)	0.0696	0.0434	0.0216	0.0158	0.0151	**0.0288**	0.0578	0.0602

7 Conclusion

In this paper the consistency of pointwise and broadband spectral estimators have been proved under general conditions. We show in particular that a wide class of trawl pro-

Table 3: Estimation results for the parametric Whittle estimator and Binomial seed, when $\alpha \in \{1.1, 1.3, 1.5, 1.7, 1.9\}$, 5000 observations, $N \in \{2, 3, 4, 5, 6, 7, 8, 9\}$, 100 replications.

α	Statistic	Binomial seed							
	N	2	3	4	5	6	7	8	9
1.1	bias($\widehat{\alpha}_{\mathrm{PW}}$)	-0.0892	-0.0762	**-0.0458**	-0.0048	0.0443	0.0829	0.1193	0.1115
	sd($\widehat{\alpha}_{\mathrm{PW}}$)	0.0046	0.0316	**0.0602**	0.0804	0.1084	0.1232	0.141	0.1269
1.3	bias($\widehat{\alpha}_{\mathrm{PW}}$)	-0.2235	-0.1334	-0.0651	**-0.0025**	0.0579	0.1127	0.1251	0.0826
	sd($\widehat{\alpha}_{\mathrm{PW}}$)	0.0517	0.0817	0.1032	**0.1213**	0.1365	0.1498	0.1499	0.1322
1.5	bias($\widehat{\alpha}_{\mathrm{PW}}$)	-0.2472	-0.1348	**-0.0467**	0.042	0.1227	0.1925	0.1365	0.0578
	sd($\widehat{\alpha}_{\mathrm{PW}}$)	0.0634	0.0908	**0.1102**	0.1342	0.1533	0.1527	0.1296	0.132
1.7	bias($\widehat{\alpha}_{\mathrm{PW}}$)	-0.2484	-0.1068	0.0152	0.1083	0.1703	0.1808	**0.0263**	-0.0468
	sd($\widehat{\alpha}_{\mathrm{PW}}$)	0.0781	0.0941	0.118	0.1209	0.0998	0.0759	**0.1009**	0.0956
1.9	bias($\widehat{\alpha}_{\mathrm{PW}}$)	-0.268	-0.1074	-0.0025	0.0439	0.0631	**0.0278**	-0.1148	-0.2035
	sd($\widehat{\alpha}_{\mathrm{PW}}$)	0.0888	0.1118	0.0922	0.0693	0.0483	**0.0498**	0.0855	0.0903

cesses satisfy these conditions. However, in view of the sample mean behaviors exhibited in [Doukhan et al., 2019], finer results on the asymptotic behavior of these estimators should be treated under more specific assumptions. Up to our best knowledge, very few results are available for non-linear long-range dependent trawl processes. The rate of a wavelet based semi-parametric estimator of the long-range dependence parameter is studied in [Fay et al., 2007] for so called Infinite source Poisson, which can be seen as a specific trawl process with Poisson seed. A first step for future work could be to study the asymptotic behavior of such an estimator.

Acknowledgements. This work has been developed within the MME-DII center of excellence (ANR-11-LABEX-0023-01) and with the help of PAI-CONICYT MEC 80170072. We would like to thank Donatas Surgailis for fruitful discussions about this work.

References

[Barndorff-Nielsen et al., 2014] Barndorff-Nielsen, O. E., Lunde, A., Shepard, N., and Veraart, A. E. D. (2014). Integer-valued trawl processes: a class of stationary infinitely divisible processes. *Scand. J. Statist.*, 41:693–724.

[Dalla et al., 2005] Dalla, V., Giraitis, L., and Hidalgo, J. (2005). Consistent estimation of the memory parameter for nonlinear time series. *J. of Time series analysis*, 211-251:87–104.

[Dedecker et al., 2007] Dedecker, J., Doukhan, P., Lang, G., León, J. R., Louhichi, S., and Prieur, C. (2007). Weak dependence: With Examples and Applications, volume 190 of *Lecture Notes in Statistics*. Springer.

[Doukhan et al., 2019] Doukhan, P., Jakubowski, A., Lopes, S. R. C., and Surgailis, D. (2019). Discrete-time trawl processes. *Stochastic Process. Appl.*, 129(4):1326–1348.

[Doukhan et al., 2002] Doukhan, P., Oppenheim, G., and Taqqu, M. S. (2002). Theory and Applications of Long-Range Dependence. Birkhäuser, Boston.

[Fay et al., 2007] Fay, G., Roueff, F., and Soulier, P. (2007). Estimation of the memory parameter of the infinite-source Poisson process. *Bernoulli*, 13(2):473–491.

[Giraitis et al., 2012] Giraitis, L., Koul, H. L., and Surgailis, D. (2012). Large Sample Inference for Long Memory Processes. Imperial College Press, London.

[Hannan, 1973] Hannan, E. J. (1973). The asymptotic theory of linear time-series models. *J. Appl. Probability*, 10:130–145, corrections, ibid. 10 (1973), 913.

[Hurvich et al., 2005] Hurvich, C. M., Moulines, E., and Soulier, P. (2005). Estimating long memory in volatility. *Econometrica*, 73(4):1283–1328.

[Petrov, 1995] Petrov, V. V. (1995). *Limit theorems of probability theory*, volume 4 of *Oxford Studies in Probability*. The Clarendon Press, Oxford University Press, New York. Sequences of independent random variables, Oxford Science Publications.

[Pipiras and Taqqu, 2017] Pipiras, V. and Taqqu, M. S. (2017). *Long-Range Dependence and Self-Similarity*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

[Robinson, 1995] Robinson, P. (1995). Gaussian Semiparametric Estimation of Long Range Dependence. *Ann. Statist.*, 23-5:1630–1661.