Received: 6 July 2018 | Revised: 8 May 2019 | Accepted: 20 July 2019

DOI: 10.1111/sjos.12404

ORIGINAL ARTICLE

Scandinavian Journal of Statistics

Estimation of cyclic long-memory parameters

Huda Mohammed Alomari[1] | **Antoine Ayache**[2] | **Myriam Fradon**[2] | **Andriy Olenko**[1]

[1]Department of Mathematics and Statistics, La Trobe University, Melbourne, VIC, Australia

[2]CNRS, UMR 8524 - Laboratoire Paul Painlevé, University of Lille, Lille, France

Correspondence
Andriy Olenko, Department of Mathematics and Statistics, La Trobe University, Melbourne, VIC 3086, Australia.
Email: a.olenko@latrobe.edu.au

Funding information
Laboratoire d'Excellence, Centre Européen pour les Mathématiques, la Physique et leurs interactions, Laboratoire de Mathématiques Paul Painlevé, France, Grant/Award Number: ANR-11-LABX-0007-01, CEMPI; Australian Research Council's Discovery Projects funding scheme, Grant/Award Number: DP160101366; La Trobe University DRP Grant in Mathematical and Computing Sciences

Abstract

This paper studies cyclic long-memory processes with Gegenbauer-type spectral densities. For a semiparametric statistical model, new simultaneous estimates for singularity location and long-memory parameters are proposed. This generalized filtered method-of-moments approach is based on general filter transforms that include wavelet transformations as a particular case. It is proved that the estimates are almost surely convergent to the true values of parameters. Solutions of the estimation equations are studied, and adjusted statistics are proposed. Monte-Carlo study results are presented to confirm the theoretical findings.

KEYWORDS

estimators of parameters, filter, Gegenbauer-type spectral densities, seasonal/cyclic long memory, stochastic process, wavelet transformation

1 | INTRODUCTION

The importance of long memory can be seen in various applications, for instance in finance, internet modeling, hydrology, linguistics, DNA sequencing, and other areas (see Cont, 2005; Leonenko & Olenko, 2014; Park et al., 2011; Pipiras & Taqqu, 2017; Samorodnitsky, 2007; Samorodnitsky, 2016; Willinger, Paxson, Riedi, & Taqqu, 2003, and the references therein).

Usually, for a stationary finite-variance random process $X(t), t \in \mathbb{R}$, long-memory or long-range dependence is defined as nonintegrability of its covariance function $\mathsf{B}(r) = \mathrm{cov}(X(t+r), X(t))$, that is, $\int_0^\infty |\mathsf{B}(r)|\, dr = +\infty$, or, more precisely, as a hyperbolic asymptotic behavior of $\mathsf{B}(\cdot)$.

It is known that the phenomenon of long-range dependence is related to singularities of spectral densities; see Leonenko and Olenko (2013). The majority of publications study the case when

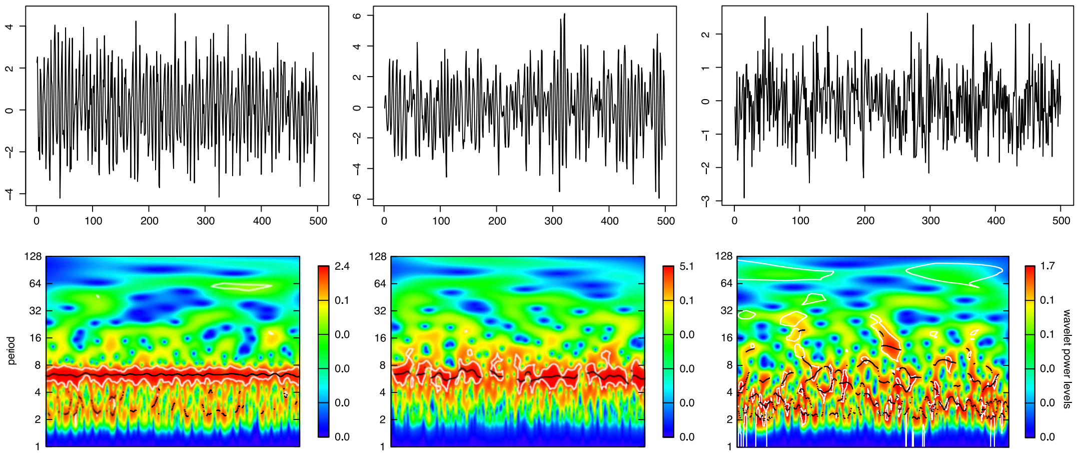

FIGURE 1 Realizations of three types of time series (with a periodic deterministic trend, ARMA, and cyclic long-range dependent) and their wavelet power spectra [Colour figure can be viewed at wileyonlinelibrary.com]

spectral densities are unbounded at the origin. However, singularities at nonzero frequencies play an important role in investigating cyclic or seasonal behavior of time series. Contrary to seasonal time series, nonseasonal cycles are often unknown in advance for various physical or financial data.

Two classical models in the literature to describe cyclic behaviors of time series are

(i) a sum of a periodic deterministic trend and a stationary random noise, and
(ii) an ARMA model with a spectral peak outside the origin.

A cyclic long-range dependent process, which will be referred to as (iii), is an intermediate case between (i) and (ii) as it has a pole in its spectral density; see Giraitis, Hidalgo, and Robinson (2001).

The first row of Figure 1 shows realizations of models (i), (ii), and (iii) from left to right. The stochastic processes $X(t) = 2\sin(t) + \varepsilon_t$ and $X(t) = 0.9X(t-1) - 0.8X(t-2) + \varepsilon_t$, where ε_t is a zero-mean white noise, were used as models (i) and (ii), respectively. The Gegenbauer random process from Section 6 was used as model (iii) in the simulations. For each simulation, the second row of Figure 1 gives the corresponding wavelet power spectra. It suggests that estimation of parameters might be more challenging problem for model (iii) than for cases (i) and (ii). Unexpectedly, relatively few publications on the matter are related to cyclic or seasonal long-memory processes. A survey of some recent asymptotic results for cyclic long-range dependent random processes and fields can be found in Arteche and Robinson (1999); Ivanov, Leonenko, Ruiz-Medina, and Savich (2013); Klykavka, Olenko, and Vicendese (2012); and Olenko (2013). It was demonstrated in Olenko (2013) that singularities at nonzero frequencies can play an important role in limit theorems even for the case of linear functionals.

Several parametric and semiparametric methods were proposed for the case when poles of spectral densities are unknown; see Arteche and Robinson (2000), Giraitis et al. (2001), Hidalgo (1996), Whitcher (2004), and the references therein. Various problems in statistical inference of random processes and fields characterized by certain singular properties of their spectral densities were investigated in Leonenko (1999); Tsai, Rachinger, and Lin (2015); and Dehling, Rooch, and Taqqu (2013). Some methods for estimating a singularity location were suggested by Arteche and Robinson (2000), Giraitis et al. (2001), and Ferrara and Guígan (2001).

The asymptotic theory for Gaussian maximum likelihood estimates (MLE) of seasonal long-memory models was developed in Giraitis et al. (2001). The quasilikelihood methods were studied in Hosoya (1997). The paper of Hidalgo (1996) studied limit theorems for spectral density estimators and functionals with spectral density singularities at the origin and possibly at other frequencies. Some results about consistency and asymptotic normality of the spectral density estimator were obtained. The paper of Yajima (1985) proposed the MLE and the least squares estimator for the long-memory time series models from Granger and Joyeux (1980) and the ARFIMA model in Hosking (1981). They examined the consistency, the limiting distribution, and the rate of convergence of these estimators. The least square estimator method was used in Beran, Ghosh, and Schell (2009) to estimate the long-range dependence parameter assuming that the singularity point is at the origin.

The minimum contrast estimator (MCE) methodology has been applied in a variety of statistical areas, in particular, for long-range dependent models. The article of Anh, Leonenko, and Sakhno (2007) discussed consistency and asymptotic normality of a class of MCEs for random processes with short- or long-range dependence based on the second- and third-order cumulant spectra. In Guo, Lim, and Meerschaert (2009), it was demonstrated that the Whittle maximum likelihood estimator is consistent and asymptotically normal for stationary seasonal autoregressive fractionally integrated moving-average processes. Consistency and asymptotic normality of MCEs for parameters of Gegenbauer random processes and fields were obtained in Espejo, Leonenko, Olenko, and Ruiz-Medina (2015). More details on the current state of the MCE theory for long-memory processes and additional references can be found in Alomari et al. (2017).

Unfortunately, the results in Alomari et al. (2017), Espejo et al. (2015), and Leonenko and Sakhno (2006) for Gegenbauer-type long-memory models use specific weight functions that approach zero at the known singularity point, which, for example, was estimated before or determined by particular applications. It is not obvious how to modify these methods if locations of singularities are unknown. New approaches are required for the situations where both the long-memory and seasonality parameters are unknown or have to be estimated simultaneously. It is known that, for nonlarge sample cases, simultaneous estimation methods yield substantial gains compared to multistage estimation methods.

Another disadvantage of fully parametric approaches for cyclic long-memory models is that incorrect specifications of models can result in inconsistent estimates of long-memory parameters; see Hidalgo (2005), and the references therein.

The article of Whitcher (2004) proposed to use wavelet transforms to estimate parameters of cyclic long-memory time series. Simulation studies were used to validate the approach and to compare it with other techniques. Unfortunately, there were no rigorous studies to justify the method and establish statistical properties of the estimators, except the case of the singularity at the origin; see Clausel, Roueff, Taqqu, and Tudor (2014), and the references therein.

This research addresses the above problem and gives first steps in developing simultaneous estimators for both the parameters. This paper deals with Gegenbauer-type cyclic long-memory semiparametric models. The Gegenbauer spectral density $f(\cdot)$ has the following form and asymptotic behavior around its poles $\pm\nu$:

$$f(\lambda) = C(2\,|\cos\lambda - \cos\nu|)^{-2\alpha} = C\left(4\left|\sin\left(\frac{\lambda+\nu}{2}\right)\sin\left(\frac{\lambda-\nu}{2}\right)\right|\right)^{-2\alpha} \sim C|\lambda^2-\nu^2|^{-2\alpha},$$

when $\lambda \to \pm\nu$.

The detailed review of the statistical inference theory for Gegenbauer random processes and fields can be found in Espejo et al. (2015).

We use the idea from Bardet and Bertrand (2010) to develop the first estimation equation. Namely, we study asymptotic properties of a filter transformation of cyclic long-memory processes. As a particular case, this transformation includes wavelet transformations. To get the second estimation equation, we propose a new approach that is based on asymptotic behavior of increments of the filter transformation. Finally, we investigate properties of the solutions to the estimation equations and propose adjusted statistics for both the cyclic and long-memory parameters. This generalized filtered method-of-moments methodology includes wavelet transformations as a particular case. Therefore, it is potentially very useful for real applications as it can employ the existing wavelet methods and software, which are more powerful and faster than programs for numerical integration and optimization required by the ordinary least squares, MLE, and MCE methods.

The simulation studies in Section 6 show that the proposed estimates rapidly converge to true solutions. One needs only a small number of different widths of a filter response function to reliably estimate the parameters. However, as feasible values of the parameters are nonnegative, it would be impossible to obtain exact nonasymptotic distributions for a fixed filter width. Small-sample simulation studies of the distributions of the estimators, limit distributions when widths of filters decrease, and other extensions (see Section 7) are omitted due to the page limit, but will be addressed in the following publications.

This article is organized as follows. In Section 2, we give basic definitions and notations. The first equation to estimate the parameters is derived in Section 3. Section 4 further studies properties of filter transforms and their increments. Then, these results are used to derive the second estimation equation. In Section 5, estimators of location and long-memory parameters are proposed and studied. Simulation studies that support the theoretical findings are presented in Section 6. The proofs of all results are in the Appendix.

All computations and simulations in this article were performed using the software R version 3.5.0 and Maple 17, Maplesoft, see Supporting information.

2 | DEFINITIONS AND AUXILIARY RESULTS

This section introduces classes of stochastic processes and their filter transforms that are studied in this paper.

We consider a measurable mean-square continuous stationary zero-mean Gaussian stochastic process $X(t), t \in \mathbb{R}$, defined on a probability space $(\Omega, \mathcal{F}, P)$, with the covariance function

$$B(r) := \mathrm{Cov}(X(t), X(t')) = \int_{\mathbb{R}} \mathrm{e}^{iu(t-t')} F(du), \quad t, t' \in \mathbb{R},$$

where $r = t - t'$ and $F(\cdot)$ is a nonnegative finite measure on $\mathbb{R}$.

Definition 1. The random process $X(t), t \in \mathbb{R}$, is said to possess an absolutely continuous spectrum if there exists a nonnegative function $f(\cdot) \in L_1(\mathbb{R})$ such that

$$F(u) = \int_{-\infty}^{u} f(\lambda) d\lambda, \quad u \in \mathbb{R}.$$

The function $f(\cdot)$ is called the spectral density of the process $X(t)$.

The process $X(t), t \in \mathbb{R}$, with absolutely continuous spectrum has the following isonormal spectral representation:

$$X(t) = \int_{\mathbb{R}} \mathrm{e}^{it\lambda} \sqrt{f(\lambda)} dW(\lambda),$$

where $W(\cdot)$ is a complex-valued Gaussian orthogonal random measure on $\mathbb{R}$.

For simplicity, in this paper, we consider the case of real-valued $X(t)$. Therefore, we assume that $f(\cdot)$ is an even function and the random measure is such that $W([\lambda_1, \lambda_2]) = W([-\lambda_2, -\lambda_1])$ for any $\lambda_2 > \lambda_1 > 0$; see Section 6 in Taqqu (1979). As all estimates in this paper use absolute values of integrands, the obtained results can also be rewritten for complex-valued processes.

Assumption 1. Let the spectral density $f(\cdot)$ of $X(t)$ admit the following semiparametric representation:

$$f(\lambda) = \frac{h(\lambda)}{|\lambda^2 - s_0^2|^{2\alpha}}, \quad \lambda \in \mathbb{R},$$

where $s_0 > 0$, $\alpha \in (0, \frac{1}{2})$ and $h(\cdot)$ is an even nonnegative bounded function that is four times boundedly differentiable on $[-\frac{1}{2}, \frac{1}{2}]$. Its derivatives of order i are denoted by $h^{(i)}(\cdot)$ and satisfy $h^{(i)}(0) = 0$, $i = 1, 2, 3, 4$. In addition, $h(0) = 1$, $h(\cdot) > 0$ in some neighborhood of $\lambda = \pm s_0$, and for all $\epsilon > 0$, it holds

$$\int_{\mathbb{R}} \frac{h(\lambda)}{(1 + |\lambda|)^{\epsilon}} d\lambda < \infty.$$

Remark 1. Stochastic processes satisfying Assumption 1 have seasonal/cyclic long memory, as their spectral densities have singularities at nonzero locations $\pm s_0$. The boundedness of $h(\cdot)$ guarantees that the singularities of $f(\cdot)$ are only in $\pm s_0$. The parameter α is a long-memory parameter. The parameter s_0 determines seasonal or cyclic behavior. Covariance functions of such processes exhibit hyperbolically decaying oscillations and $\int_{\mathbb{R}} |B(r)| dr = +\infty$ as $\alpha \in (0, 1/2)$; see Arteche and Robinson (1999). The Gegenbauer random processes have cyclic long-memory behavior determined by Assumption 1 as their spectral densities $f(\lambda) \sim c|\lambda^2 - s_0^2|^{-2\alpha}$, $\lambda \to \pm s_0$, around the Gegenbauer frequency s_0; see Chung (1966) and Espejo et al. (2015).

Remark 2. The conditions on $h(\cdot)$ guarantee that $f(\cdot)$ is a spectral density with only singularity locations at $\lambda = \pm s_0$. The differentiability conditions on $h(\cdot)$ and its derivatives can be relaxed and replaced by Hölder assumptions in some neighborhood of the origin.

The smoothness conditions guarantee the following technical inequalities required for the proof.

Lemma 1. *For* $0 \le \lambda \le \tilde{\lambda} \le \frac{1}{2}$, *it holds*

$$|h(\tilde{\lambda}) - h(\lambda)| \le c_1 |\tilde{\lambda}^2 - \lambda^2|,$$
$$h(\lambda) \le 1 + c_1,$$
$$|h(\lambda) - 1| \le c_1 \cdot \lambda^4,$$

where $c_1 := \max_{\lambda \in [0, \frac{1}{2}]} (h^{(2)}(\lambda), h^{(4)}(\lambda))$.

Example 1. The asymptotic behavior of the function $h(\lambda)$ when $\lambda \to \infty$ must guarantee that $f(\cdot) \in L_1(\mathbb{R})$.

For example, the function $f(\lambda) = |\lambda^2 - s_0^2|^{-2\alpha} I_{[-M,M]}(\lambda)$ satisfies Assumption 1, where $M > s_0$,

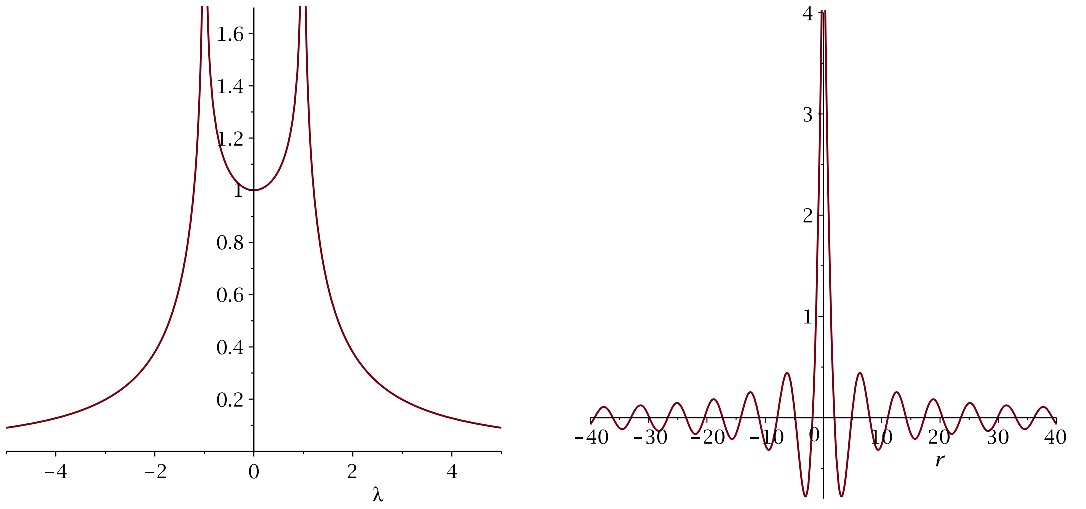

FIGURE 2 Plots of the spectral density $f(\lambda)$ with singularities at nonzero frequencies and the corresponding covariance function $B(r)$ [Colour figure can be viewed at wileyonlinelibrary.com]

$$I_{[-M,M]}(\lambda) = \begin{cases} 1, & \lambda \in [-M, M] \\ 0, & \lambda \notin [-M, M] \end{cases}$$ is the indicator function of the interval $[-M, M]$.

Another example satisfying Assumption 1 is the following spectral density and corresponding covariance function:

$$f(\lambda) = \begin{cases} \frac{1}{(1-\lambda^2)^{1/4}}, & |\lambda| \leq 1 \\ \frac{1}{\lambda(\lambda^2-1)^{1/4}}, & |\lambda| > 1, \end{cases}$$

$$B(r) = \sqrt{\pi}\left(\sqrt{2\pi} + \left(\frac{2}{r}\right)^{1/4}\Gamma\left(\frac{3}{4}\right)J_{\frac{1}{4}}(r) - 2\sqrt{2r}\,{}_1F_2\left(\frac{1}{4}; \frac{3}{4}, \frac{5}{4}; -\frac{r^2}{4}\right)\right),$$

where J_ν is the Bessel function of the first kind, ${}_1F_2$ is the hypergeometric function, and $s_0 = 1$ and $\alpha = 1/8$ were chosen. Plots of $f(\lambda)$ and $B(r)$ are shown in Figure 2.

Remark 3. As we study seasonal or cyclic long-memory models, in this paper, we consider the case of singularities at nonzero frequencies. The discussion about differences between the cases with spectral singularities at the origin and at other locations can be found in Arteche and Robinson (1999). As s_0 is separated from zero, without loss of generality, we assume that $s_0 > 1$. Indeed, if a time series has a periodic component with the period T, then the corresponding frequency $s_0 = 1/T$. Changing the time unit, the parameter s_0 can be made greater than 1.

Now, we introduce filter transforms of stochastic processes. To define filters, we use real-valued functions $\psi(t), t \in \mathbb{R}$, with the Fourier transforms $\hat{\psi}(\cdot)$.

Throughout this article, we use the convention that the Fourier transform of an arbitrary function ψ belonging to $L_1(\mathbb{R})$ is the function $\hat{\psi}$ defined, for every $\lambda \in \mathbb{R}$, as

$$\hat{\psi}(\lambda) = \int_{\mathbb{R}} e^{-i\lambda t}\psi(t)dt.$$

Assumption 2. Let $\psi \in L_2(\mathbb{R})$ be a real-valued function such that supp $\hat{\psi} \subset [-A, A], A > 0$, and $\hat{\psi}(\cdot)$ is continuous except at a finite number of points and of bounded variation on $[-A, A]$.

Remark 4. It follows from Assumption 2 that $\hat{\psi}$ is bounded and $\psi \in L_\infty(\mathbb{R})$ is an analytic function.

Let us define the constants $c_2 := \int_{\mathbb{R}} |\hat{\psi}(\lambda)|^2 d\lambda$ and $c_3 := 2\int_{\mathbb{R}} \lambda^2 |\hat{\psi}(\lambda)|^2 d\lambda$.

Some important for applications functions $\psi(\cdot)$ satisfying Assumption 2 are the wavelets; see Daubechies (1992) and Meyer (1992), given in the next examples. However, in general, $\psi(\cdot)$ is not required to be a wavelet.

Example 2. The function $\psi(\cdot)$ can be selected as the Shannon father or mother wavelets. Indeed, the Shannon father wavelet

$$\psi_f(t) = \operatorname{sinc}(\pi t) = \begin{cases} \frac{\sin(\pi t)}{\pi t}, & t \neq 0, \\ 1, & t = 0 \end{cases}$$

has the Fourier transform

$$\hat{\psi}_f(\lambda) = I_{[-\pi,\pi]}(\lambda).$$

The corresponding constants are $c_2 = 2\pi$ and $c_3 = \frac{4}{3}\pi^3$.

The Shannon mother wavelet

$$\psi_m(t) = \frac{\sin(2\pi t) - \cos(\pi t)}{\pi\left(t - \frac{1}{2}\right)}$$

has the Fourier transform

$$\hat{\psi}_m(\lambda) = -e^{-\frac{i\lambda}{2}} I_{[-2\pi,-\pi)\cup(\pi,2\pi]}(\lambda).$$

The corresponding constants are $c_2 = 2\pi$ and $c_3 = 9\frac{1}{3}\pi^3$.

Plots of $\psi_f(t)$ and $\psi_m(t)$ are shown in Figure 3.

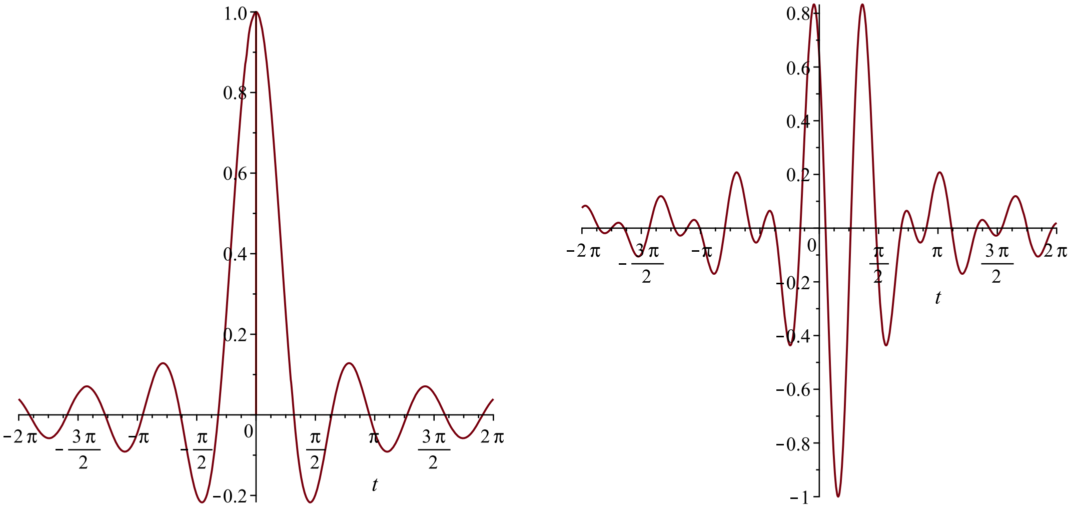

FIGURE 3 Plots of the Shannon father and mother wavelets $\psi_f(t)$ and $\psi_m(t)$ [Colour figure can be viewed at wileyonlinelibrary.com]

Example 3. The function $\psi(\cdot)$ can be selected as the Meyer father or mother wavelets (see, e.g., Daubechies, 1992; Meyer, 1992). Indeed, the Meyer wavelets have the Fourier transforms

$$\hat{\psi}_m(\lambda) := \begin{cases} \sin\left(\frac{\pi}{2}\nu\left(\frac{3|\lambda|}{2\pi}-1\right)\right)e^{\frac{i\lambda}{2}}, & \text{if } \frac{2\pi}{3} \le |\lambda| \le \frac{4\pi}{3}, \\ \cos\left(\frac{\pi}{2}\nu\left(\frac{3|\lambda|}{4\pi}-1\right)\right)e^{\frac{i\lambda}{2}}, & \text{if } \frac{4\pi}{3} \le |\lambda| \le \frac{8\pi}{3}, \\ 0, & \text{otherwise,} \end{cases}$$

$$\hat{\psi}_f(\omega) := \begin{cases} 1, & \text{if } |\lambda| \le \frac{2\pi}{3}, \\ \cos\left(\frac{\pi}{2}\nu\left(\frac{3|\lambda|}{4\pi}-1\right)\right), & \text{if } \frac{2\pi}{3} \le |\lambda| \le \frac{4\pi}{3}, \\ 0, & \text{otherwise,} \end{cases}$$

where $\nu(\cdot)$ is a function with values in [0,1] that satisfies $\nu(x)+\nu(1-x) = 1, x \in \mathbb{R}$. For example, one can use

$$\nu(x) = \begin{cases} 0, & \text{if } x < 0, \\ x, & \text{if } 0 \le x \le 1, \\ 1, & \text{if } x > 1. \end{cases}$$

The corresponding constants for $\hat{\psi}_m(\lambda)$ are $c_2 = 2\pi$ and $c_3 = 12\frac{4}{9}\pi(\pi^2+2)$ and $c_2 = 2\pi$ and $c_3 = 1\frac{7}{9}\pi(\pi^2-2)$ for $\hat{\psi}_f(\lambda)$.

Plots of $|\hat{\psi}_m(\lambda)|$ and $\hat{\psi}_f(\lambda)$ are shown in Figure 4.

Now, for any pair $(a,b) \in \mathbb{R}_+ \times \mathbb{R}$, where $\mathbb{R}_+ = (0,+\infty)$, we define the following filter transform of the process $X(t)$:

$$d_x(a,b) := \frac{1}{\sqrt{a}} \int_{\mathbb{R}} \psi\left(\frac{t-b}{a}\right) X(t)dt. \tag{1}$$

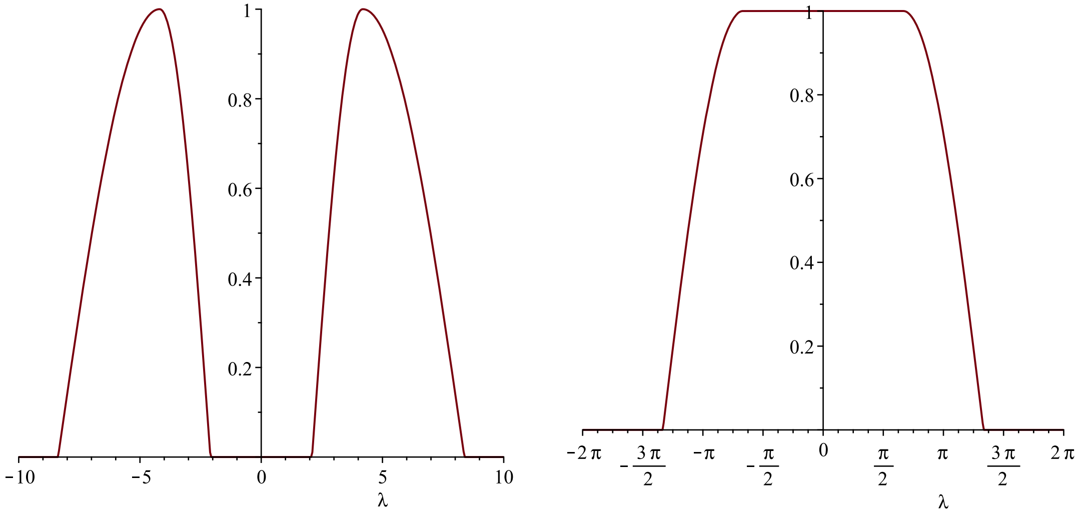

FIGURE 4 Plots of the Fourier transforms of the Meyer mother and father wavelets $|\hat{\psi}_m(\lambda)|$ and $\hat{\psi}_f(\lambda)$ [Colour figure can be viewed at wileyonlinelibrary.com]

Remark 5. If $\psi(\cdot)$ is a wavelet, then $d_x(a,b)$ given by (1) defines the wavelet transform of the process $X(t)$.

The general filtration theory of stochastic processes guarantees that (1) is correctly defined if the following assumption is satisfied; see Chapter V, Section 6, in Gikhman and Skorokhod (2004).

Assumption 3. Let the integral $\int_{\mathbb{R}^2}\psi(t)B(t-t')\overline{\psi(t')}dtdt'$ exist as an improper Cauchy integral on the plane.

Remark 6. Different assumptions on the process and the filter were used in Bardet and Bertrand (2010). Namely, they assumed that $\psi(\cdot)$ is a mother wavelet that has two vanishing moments and there are constants $c_\psi, c'_\psi > 0$ such that $(1+|t|)|\psi(t)| \le c_\psi$, $|\hat{\psi}(\lambda)|+|\hat{\psi}'(\lambda)| \le c'_\psi$, for all $t, \lambda \in \mathbb{R}$.

Using the above notations $d_x(a,b)$ can be rewritten in the frequency domain as

$$d_x(a,b) = \sqrt{a}\int_{\mathbb{R}} e^{ib\lambda}\overline{\hat{\psi}(a\lambda)}\sqrt{f(\lambda)}dW(\lambda).$$

This Gaussian random variable has a zero mean, that is, $Ed_x(a,b) = 0$. Its variance equals

$$E|d_x(a,b)|^2 = a\int_{\mathbb{R}}|\hat{\psi}(a\lambda)|^2 f(\lambda)d\lambda := J(a) \tag{2}$$

and, thus, does not depend on b.

In the following sections, we assume that Assumptions 1–3 are satisfied.

3 | FIRST STATISTICS

Spectral densities satisfying Assumption 1 have two parameters of interest (α and s_0). This section derives some properties of $d_x(a,b)$ and suggests a statistic based on $d_x(a,b)$ that can be used as an estimate of $s_0^{-4\alpha}$.

Let $\{a_j\}, \{\gamma_j\} \subset \mathbb{R}_+, \{m_j\} \subset \mathbb{N}, j \in \mathbb{N}$, be sequences of positive numbers and $\{b_{jk}\} \subset \mathbb{R}, j \in \mathbb{N}, k \in \mathbb{Z}$, be an infinite array. In the following proofs, we assume that $\{a_j\}$ is an unboundedly monotone increasing sequence and $b_{jk_1} \neq b_{jk_2}$ for all $j \in \mathbb{N}$ and $k_1 \neq k_2$.

We will use the following notation:

$$\delta_{jk} := d_x(a_j, b_{jk}) = \sqrt{a_j}\int_{\mathbb{R}} e^{ib_{jk}\lambda}\overline{\hat{\psi}(a_j\lambda)}\sqrt{f(\lambda)}dW(\lambda).$$

By Assumption 1,

$$\delta_{jk} = \sqrt{a_j}\int_{\mathbb{R}} e^{ib_{jk}\lambda}\frac{\overline{\hat{\psi}(a_j\lambda)}\sqrt{h(\lambda)}}{|\lambda^2 - s_0^2|^\alpha}dW(\lambda).$$

Therefore, for all $j \in \mathbb{N}$ and $k_1, k_2 \in \mathbb{Z}$, it holds $E\delta_{jk_1} = 0$ and

$$\begin{aligned} I(j,k_1,k_2) &:= \mathrm{Cov}\left(\delta_{jk_1},\delta_{jk_2}\right) = E\left(\delta_{jk_1}\overline{\delta}_{jk_2}\right) \\ &= a_j\int_{\mathbb{R}} \mathrm{e}^{i\left(b_{jk_1}-b_{jk_2}\right)\lambda}\frac{|\hat{\psi}(a_j\lambda)|^2 h(\lambda)}{\left|\lambda^2-s_0^2\right|^{2\alpha}}d\lambda \\ &= s_0^{-4\alpha}\int_{\mathbb{R}} \mathrm{e}^{i\frac{b_{jk_1}-b_{jk_2}}{a_j}\lambda}\frac{|\hat{\psi}(\lambda)|^2 h\left(\frac{\lambda}{a_j}\right)}{\left|\left(\frac{\lambda}{a_j s_0}\right)^2-1\right|^{2\alpha}}d\lambda. \end{aligned} \tag{3}$$

Lemma 2. *Let Assumption 2 hold true. Then, for all $k_1, k_2 \in \mathbb{Z}^2$ and $j \in \mathbb{N}$ such that $\frac{a_j}{2A} \geq 1$*

$$|I(j,k_1,k_2)| \leq c_4(s_0,\alpha)\begin{cases} 1, & \text{if } k_1 = k_2, \\ \frac{a_j}{|b_{jk_1}-b_{jk_2}|}, & \text{if } k_1 \neq k_2, \end{cases}$$

where

$$c_4(s_0,\alpha) := 2s_0^{-4\alpha}\max\left(\frac{2}{3}c_2(1+c_1), \max_{\lambda\in[-A,A]}|\hat{\psi}(\lambda)|^2 V_{-1/2}^{1/2}(h(\cdot)) + (1+c_1)\cdot V_{-A}^{A}\left(|\hat{\psi}(\cdot)|^2\right)\right),$$

$V_a^b(f)$ denotes the total variation of a function $f(\cdot)$ on an interval $[a,b]$.

Let us define

$$\bar{\delta}_{j\cdot}^{(2)} := \frac{1}{m_j}\sum_{k=1}^{m_j}\delta_{jk}^2.$$

It follows from (2) that $E\bar{\delta}_{j\cdot}^{(2)} = J(a_j)$.

Lemma 3. *Suppose that $\frac{a_j}{2A} \geq 1$, the sequences $\{b_{jk}\}$ and $\{\gamma_j\}$ are such that, for all $j \in \mathbb{N}, k_1, k_2 \in \mathbb{Z}$, it holds $|b_{jk_1} - b_{jk_2}| \geq |k_1 - k_2|\gamma_j$. Then,*

$$Var\left(\bar{\delta}_{j\cdot}^{(2)}\right) \leq \frac{c_{5j}(s_0,\alpha)}{m_j},$$

where $c_{5j}(s_0,\alpha) := 2c_4^2(s_0,\alpha)(1+\frac{\pi^2}{3}\frac{a_j^2}{\gamma_j^2})$.

Lemma 4. *Let $\{r_j\} \subset \mathbb{R}_+$ be a decreasing sequence such that $\lim_{j\to\infty} r_j = 0$. Let us choose such $\{m_j\}$ that $\sum_{j=1}^{\infty}\frac{1}{r_j^2 m_j} < +\infty$ and $\sum_{j=1}^{\infty}\frac{a_j^2}{r_j^2\gamma_j^2 m_j} < +\infty$, where $\{\gamma_j\}$ is from Lemma 3. Then, there exists an almost surely finite random variable c_6 such that, for all $j \in \mathbb{N}$,*

$$\left|\bar{\delta}_{j\cdot}^{(2)} - J(a_j)\right| \leq c_6 r_j.$$

The lemma below gives an upper bound on the deviation of $J(a_j)$ from $s_0^{-4\alpha}c_2$.

Lemma 5. *If $j \in \mathbb{N}$ is such that $\frac{a_j}{A} \geq 2$, then one has*

$$\left|J(a_j) - c_2 s_0^{-4\alpha}\right| \leq \frac{c_7(s_0,\alpha)}{a_j^2},$$

where $c_7(s_0,\alpha) := (\frac{4(1+c_1)}{s_0^2} + c_1)c_2\frac{A^2}{s_0^{4\alpha}}$.

Combining Lemmas 4 and 5, we obtain the following.

Proposition 1. *Under the conditions of Lemma 4, it holds* $\bar{\delta}^{(2)}_{j\cdot} \xrightarrow{a.s.} c_2 s_0^{-4\alpha}$. *Moreover, there exists an almost surely finite random variable* c_8 *such that, for all* $j \in \mathbb{N}$,

$$\left|\bar{\delta}^{(2)}_{j\cdot} - c_2 s_0^{-4\alpha}\right| \leq c_8 \max\left(r_j, a_j^{-2}\right).$$

4 | SECOND STATISTICS

In this section, we further study properties of $\bar{\delta}^{(2)}_{j\cdot}$ and $J(a_j)$. It allows us to suggest a new estimate of $\alpha s_0^{-4\alpha-2}$. The main idea is to find the asymptotic behavior of increments of $\bar{\delta}^{(2)}_{j\cdot}$. Therefore, we start by deriving some results about increments of $J(a_j) = \mathbb{E}\bar{\delta}^{(2)}_{j\cdot}$.

Lemma 6. *If* $\{a_j\}$ *is an unboundedly monotone increasing, then*

$$\lim_{j\to+\infty} \frac{J(a_j) - J(a_{j+1})}{a_j^{-2} - a_{j+1}^{-2}} = \alpha c_3 s_0^{-4\alpha-2}.$$

Now, we investigate the rate of convergence in Lemma 6.

Lemma 7. *There is* $j_0 \in \mathbb{N}$ *such that, for all* $j \geq j_0$, *it holds*

$$\left|\frac{J(a_j) - J(a_{j+1})}{a_j^{-2} - a_{j+1}^{-2}} - \alpha s_0^{-4\alpha-2} c_3\right| \leq \frac{c_3 2^6 A^2 s_0^{-4\alpha-4}\left(1 + 3^3 c_1 s_0^4/2^6\right)}{3^3 a_j^2 \left(1 - \left(\frac{a_j}{a_{j+1}}\right)^2\right)}.$$

Now, let us define $\Delta\bar{\delta}^{(2)}_{j\cdot} = \frac{\bar{\delta}^{(2)}_{j\cdot} - \bar{\delta}^{(2)}_{j+1\cdot}}{a_j^{-2} - a_{j+1}^{-2}}$. Then, the following result holds.

Proposition 2. *Let the assumptions of Lemma 4 hold true and there exist* $\varepsilon > 0$ *and* $j_0 \in \mathbb{N}$ *such that* $a_{j+1} \geq (1+\varepsilon)a_j$ *for all* $j \geq j_0$. *Then,*

$$\Delta\bar{\delta}^{(2)}_{j\cdot} \xrightarrow{a.s.} \alpha c_3 s_0^{-4\alpha-2}, \quad j \to +\infty.$$

Moreover, there exists an almost surely finite random variable c_9 *such that, for all* $j \in \mathbb{N}$, *it holds*

$$\left|\Delta\bar{\delta}^{(2)}_{j\cdot} - \alpha c_3 s_0^{-4\alpha-2}\right| \leq c_9 \max\left(a_j^2 r_j, a_j^{-2}\right). \tag{4}$$

Remark 7. As the rate of decay of $\{r_j\}$ can be arbitrary selected, the best upper bound given by (4) has order a_j^{-2}.

5 | ESTIMATION OF (s_0, α)

In the previous sections, we proved that if the true values of parameters are (s_0, α), then the vector statistics

$$\begin{pmatrix} \bar{\delta}^{(2)}_{j\cdot}/c_2 \\ \Delta\bar{\delta}^{(2)}_{j\cdot}/c_3 \end{pmatrix} \xrightarrow{\text{a.s.}} \begin{pmatrix} s_0^{-4\alpha} \\ \alpha s_0^{-4\alpha-2} \end{pmatrix}, \quad j \to +\infty.$$

In this section, we investigate properties of the pair $(\hat{s}_{0j}, \hat{\alpha}_j)$ that is a solution of the system

$$\begin{cases} \hat{s}_{0j}^{-4\hat{\alpha}_j} = \bar{\delta}_{j\cdot}^{(2)}/c_2, \\ \hat{\alpha}_j \hat{s}_{0j}^{-4\hat{\alpha}_j-2} = \Delta\bar{\delta}_{j\cdot}^{(2)}/c_3. \end{cases} \tag{5}$$

To handle the cases, where $(\frac{\bar{\delta}_{j\cdot}^{(2)}}{c_2}, \frac{\Delta\bar{\delta}_{j\cdot}^{(2)}}{c_3})$ may not be in the feasible region of $(s_0^{-4\alpha}, \alpha s_0^{-4\alpha-2})$, we propose adjusted estimates.

First, we discuss existence of solutions.

Lemma 8. *Let* $(y_1, y_2) \in R_{\mathbf{y}}$, *where* $R_{\mathbf{y}} := \{(y_1, y_2) \in (0,1) \times (0, y_1^2/2)\}$. *Then, the system*

$$\begin{cases} s_0^{-4\alpha} = y_1, \\ \alpha s_0^{-4\alpha-2} = y_2 \end{cases} \tag{6}$$

has a solution $(s_0, \alpha) \in (1, +\infty) \times (0, \frac{1}{2})$.

Thus, if $(\frac{\bar{\delta}_{j\cdot}^{(2)}}{c_2}, \frac{\Delta\bar{\delta}_{j\cdot}^{(2)}}{c_3}) \in R_{\mathbf{y}}$, then there is a pair $(\hat{s}_{0j}, \hat{\alpha}_j) \in (1, +\infty) \times (0, \frac{1}{2})$ that satisfies the system of equations (5). Now, we will investigate uniqueness of solutions.

Lemma 9. *Let* $(y_1, y_2) \in R_{\mathbf{y}}$. *Then, system (6) has a unique solution.*

Now, we provide solutions to system (6). These solutions are given in terms of the Lambert W function, which is defined as a solution of the equation $te^t = x$, $x \geq -\mathrm{e}^{-1}$, that is, $t = \mathrm{Lambert}\,W(x)$; see Corless, Gonnet, Hare, Jeffrey, and Knuth (1996).

Proposition 3. *Let* $(y_1, y_2) \in R_{\mathbf{y}}$. *Then, the solution to system (6) is*

$$\begin{aligned} s_0 &= exp\left(\frac{1}{2} LambertW\left(\frac{y_1}{y_2} \ln\left(y_1^{-\frac{1}{2}}\right)\right)\right), \\ \alpha &= \frac{y_2}{y_1} exp\left(LambertW\left(\frac{y_1}{y_2} \ln\left(y_1^{-\frac{1}{2}}\right)\right)\right). \end{aligned} \tag{7}$$

Remark 8. The Lambert W function has two real branches $\mathrm{Lambert}W_0$ and $\mathrm{Lambert}W_{-1}$. The branch $\mathrm{Lambert}W_0$ is defined on the interval $[-\frac{1}{\mathrm{e}}, +\infty)$, but the branch $\mathrm{Lambert}W_{-1}$ is defined only on the interval $[-\frac{1}{\mathrm{e}}, 0)$. The point $(-\frac{1}{\mathrm{e}}, -1)$ is a branch point for $\mathrm{Lambert}W_0$ and $\mathrm{Lambert}W_{-1}$. Hence, for $y_1 \in (0,1)$, it holds $(\frac{y_1}{y_2}) \ln(y_1^{-\frac{1}{2}}) > 0$ and (7) gives a unique solution to (6) with the branch $\mathrm{Lambert}W_0$.

Now, we see that, for $(y_1, y_2) \in R_{\mathbf{y}}$, there is a unique solution (s_0, α) to (6). If s_0 and α are the true value of parameters, then the corresponding $(y_1, y_2) \in R_{\mathbf{y}}$. As $R_{\mathbf{y}}$ is an open set, then $(y_1, y_2) \in \mathrm{int}(R_{\mathbf{y}}) = R_{\mathbf{y}}$, and there is some $j_0 \in \mathbb{N}$ such that $(\frac{\bar{\delta}_{j\cdot}^{(2)}}{c_2}, \frac{\Delta\bar{\delta}_{j\cdot}^{(2)}}{c_3}) \in R_{\mathbf{y}}$ for all $j \geq j_{0\cdot}$, where $\mathrm{int}(\cdot)$ denotes the interior of a set. Therefore, starting from j_0, system (5) has a unique solution.

However, it might happen that $(\frac{\bar{\delta}_{j\cdot}^{(2)}}{c_2}, \frac{\Delta\bar{\delta}_{j\cdot}^{(2)}}{c_3}) \notin R_{\mathbf{y}}$ for some $j < j_0$ even if $(y_1, y_2) \in R_{\mathbf{y}} = \text{int}(R_{\mathbf{y}})$ for the corresponding true value (s_0, α). For the cases $(\frac{\bar{\delta}_{j\cdot}^{(2)}}{c_2}, \frac{\Delta\bar{\delta}_{j\cdot}^{(2)}}{c_3}) \notin R_{\mathbf{y}}$ to define $(\hat{s}_{0j}, \hat{\alpha}_j)$, we introduce "adjusted" values $(\frac{\bar{\delta}_{j\cdot}^{(2,a)}}{c_2}, \frac{\Delta\bar{\delta}_{j\cdot}^{(2,a)}}{c_3}) \in R_{\mathbf{y}}$.

Definition 2. The adjusted statistics $\frac{\bar{\delta}_{j\cdot}^{(2,a)}}{c_2}$ and $\frac{\Delta\bar{\delta}_{j\cdot}^{(2,a)}}{c_3}$ are defined as follows:

- if $\frac{\bar{\delta}_{j\cdot}^{(2)}}{c_2} \in (0,1)$ and $\frac{\Delta\bar{\delta}_{j\cdot}^{(2)}}{c_3} \geq \frac{1}{2}(\frac{\bar{\delta}_{j\cdot}^{(2)}}{c_2})^2$, then

$$\bar{\delta}_{j\cdot}^{(2,a)} = \bar{\delta}_{j\cdot}^{(2)} \quad \text{and} \quad \Delta\bar{\delta}_{j\cdot}^{(2,a)} = c_3 \max\left(\left(\frac{\bar{\delta}_{j\cdot}^{(2)}}{c_2}\right)^2 - \frac{\Delta\bar{\delta}_{j\cdot}^{(2)}}{c_3}, \frac{1}{4}\left(\frac{\bar{\delta}_{j\cdot}^{(2)}}{c_2}\right)^2\right);$$

- if $\frac{\bar{\delta}_{j\cdot}^{(2)}}{c_2} \in (0,1)$ and $\frac{\Delta\bar{\delta}_{j\cdot}^{(2)}}{c_3} \leq 0$, then

$$\bar{\delta}_{j\cdot}^{(2,a)} = \bar{\delta}_{j\cdot}^{(2)} \quad \text{and} \quad \Delta\bar{\delta}_{j\cdot}^{(2,a)} = c_3 \min\left(-\frac{\Delta\bar{\delta}_{j\cdot}^{(2)}}{c_3}, \frac{1}{4}\left(\frac{\bar{\delta}_{j\cdot}^{(2)}}{c_2}\right)^2\right);$$

- if $\frac{\bar{\delta}_{j\cdot}^{(2)}}{c_2} \geq 1$ and $0 < \frac{\Delta\bar{\delta}_{j\cdot}^{(2)}}{c_3} < \frac{1}{2}$, then

$$\Delta\bar{\delta}_{j\cdot}^{(2,a)} = \Delta\bar{\delta}_{j\cdot}^{(2)} \quad \text{and} \quad \bar{\delta}_{j\cdot}^{(2,a)} = c_2 \max\left(2 - \frac{\bar{\delta}_{j\cdot}^{(2)}}{c_2}, \frac{1}{2}\left(1 + \left(2\frac{\Delta\bar{\delta}_{j\cdot}^{(2)}}{c_3}\right)^{\frac{1}{2}}\right)\right);$$

- if $\frac{\bar{\delta}_{j\cdot}^{(2)}}{c_2} \geq 1$ and $\frac{\Delta\bar{\delta}_{j\cdot}^{(2)}}{c_3} \geq \frac{1}{2}$, then

$$\Delta\bar{\delta}_{j\cdot}^{(2,a)} = c_3 \max\left(1 - \frac{\Delta\bar{\delta}_{j\cdot}^{(2)}}{c_3}, \frac{1}{4}\right) \quad \text{and} \quad \bar{\delta}_{j\cdot}^{(2,a)} = c_2 \max\left(2 - \frac{\bar{\delta}_{j\cdot}^{(2)}}{c_2}, \frac{1}{2}\left(1 + \left(2\frac{\Delta\bar{\delta}_{j\cdot}^{(2,a)}}{c_3}\right)^{\frac{1}{2}}\right)\right);$$

- if $\frac{\bar{\delta}_{j\cdot}^{(2)}}{c_2} \geq 1$ and $\frac{\Delta\bar{\delta}_{j\cdot}^{(2)}}{c_3} \leq 0$, then

$$\Delta\bar{\delta}_{j\cdot}^{(2,a)} = c_3 \min\left(-\frac{\Delta\bar{\delta}_{j\cdot}^{(2)}}{c_3}, \frac{1}{4}\right) \quad \text{and} \quad \bar{\delta}_{j\cdot}^{(2,a)} = c_2 \max\left(2 - \frac{\bar{\delta}_{j\cdot}^{(2)}}{c_2}, \frac{1}{2}\left(1 + \left(2\frac{\Delta\bar{\delta}_{j\cdot}^{(2,a)}}{c_3}\right)^{\frac{1}{2}}\right)\right);$$

- otherwise, $\bar{\delta}_{j\cdot}^{(2,a)} = \bar{\delta}_{j\cdot}^{(2)}$ and $\Delta\bar{\delta}_{j\cdot}^{(2,a)} = \Delta\bar{\delta}_{j\cdot}^{(2)}$.

In the fourth and fifth cases the value of $\Delta\bar{\delta}_{j\cdot}^{(2,a)}$ is computed first, and then, it is used to compute the adjusted value $\bar{\delta}_{j\cdot}^{(2,a)}$.

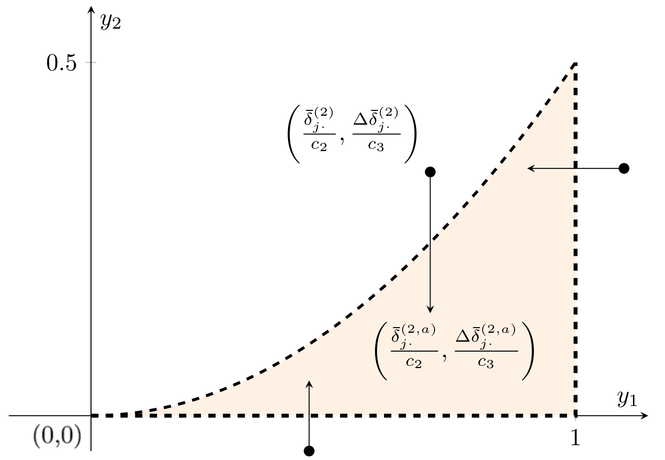

FIGURE 5 Plot of $R_{\mathbf{y}}$, $(\frac{\bar{\delta}^{(2)}_{j\cdot}}{c_2}, \frac{\Delta\bar{\delta}^{(2)}_{j\cdot}}{c_3})$ and the corresponding adjusted estimates $(\frac{\bar{\delta}^{(2,a)}_{j\cdot}}{c_2}, \frac{\Delta\bar{\delta}^{(2,a)}_{j\cdot}}{c_3})$ [Colour figure can be viewed at wileyonlinelibrary.com]

Figure 5 clarifies geometric reasons to introduce the adjusted values $(\frac{\bar{\delta}^{(2,a)}_{j\cdot}}{c_2}, \frac{\Delta\bar{\delta}^{(2,a)}_{j\cdot}}{c_3})$. Vertical or horizontal reflections over boundaries of $R_{\mathbf{y}}$ are used with an additional constraint that the reflected points do not go beyond the opposite boundaries of $R_{\mathbf{y}}$. In addition, the reflected points should not belong to the boundaries. For instance, this might happen, in the first case of Definition 2 if one has $\frac{\Delta\bar{\delta}^{(2)}_{j\cdot}}{c_3} = \frac{1}{2}(\frac{\bar{\delta}^{(2)}_{j\cdot}}{c_2})^2$.

Remark 9. The main practical advantage of using the adjusted statistics, as opposed to removing nonfeasible values, is the case when the true values of parameters correspond to a point that is close to the boundary of $R_{\mathbf{y}}$. Then, for a finite sample, it might happen that all $(\frac{\bar{\delta}^{(2)}_{j\cdot}}{c_2}, \frac{\Delta\bar{\delta}^{(2)}_{j\cdot}}{c_3}) \notin R_{\mathbf{y}}$. Nevertheless, even in this case, the adjusted statistics provide reliable estimates for the parameters.

Remark 10. By the construction in Definition 2, the adjusted pair $(\frac{\bar{\delta}^{(2,a)}_{j\cdot}}{c_2}, \frac{\Delta\bar{\delta}^{(2,a)}_{j\cdot}}{c_3}) \in R_{\mathbf{y}}$ and both the $(\frac{\bar{\delta}^{(2,a)}_{j\cdot}}{c_2}, \frac{\Delta\bar{\delta}^{(2,a)}_{j\cdot}}{c_3})$ and $(\frac{\bar{\delta}^{(2)}_{j\cdot}}{c_2}, \frac{\Delta\bar{\delta}^{(2)}_{j\cdot}}{c_3})$ converge to the same value $(s_0^{-4\alpha}, \alpha s_0^{-4\alpha-2})$ when $j \to +\infty$.

Remark 11. As only a finite number of $(\frac{\bar{\delta}^{(2)}_{j\cdot}}{c_2}, \frac{\Delta\bar{\delta}^{(2)}_{j\cdot}}{c_3})$ fall outside of $R_{\mathbf{y}}$, then there is $j_0 \in \mathbb{N}$ such that $(\frac{\bar{\delta}^{(2,a)}_{j\cdot}}{c_2}, \frac{\Delta\bar{\delta}^{(2,a)}_{j\cdot}}{c_3}) = (\frac{\bar{\delta}^{(2)}_{j\cdot}}{c_2}, \frac{\Delta\bar{\delta}^{(2)}_{j\cdot}}{c_3})$ for all $j \geq j_0$. Therefore, in this case, $(\frac{\bar{\delta}^{(2,a)}_{j\cdot}}{c_2}, \frac{\Delta\bar{\delta}^{(2,a)}_{j\cdot}}{c_3})$ and $(\frac{\bar{\delta}^{(2)}_{j\cdot}}{c_2}, \frac{\Delta\bar{\delta}^{(2)}_{j\cdot}}{c_3})$ have the same rate of convergence to $(s_0^{-4\alpha}, \alpha s_0^{-4\alpha-2})$ when $j \to +\infty$.

Now, we are ready to formulate the main result.

Theorem 1. *Let the process $X(t)$ and the filter $\psi(\cdot)$ satisfy Assumptions 1 to 3. Let $(\hat{s}_{0j}, \hat{\alpha}_j)$ be a solution of the system of equations*

$$\begin{cases} \hat{\alpha}_j(\hat{s}_{0j})^{-4\hat{\alpha}_j-2} = \Delta\bar{\delta}^{(2,a)}_{j\cdot}/c_3, \\ \hat{s}_{0j}^{-4\hat{\alpha}_j} = \bar{\delta}^{(2,a)}_{j\cdot}/c_2, \end{cases}$$

where $\Delta\bar{\delta}_{j\cdot}^{(2,a)}$ *and* $\bar{\delta}_{j\cdot}^{(2,a)}$ *are the adjusted statistics. Then,*

$$\begin{aligned} \hat{s}_{0j} &= exp\left(\frac{1}{2}LambertW\left(\frac{\ln\left(c_2/\bar{\delta}_{j\cdot}^{(2,a)}\right)}{2q_j}\right)\right), \\ \hat{\alpha}_j &= q_j\, exp\left(LambertW\left(\frac{\ln\left(c_2/\bar{\delta}_{j\cdot}^{(2,a)}\right)}{2q_j}\right)\right), \end{aligned} \tag{8}$$

where $q_j = \frac{c_2}{c_3}\frac{\Delta\bar{\delta}_{j\cdot}^{(2,a)}}{\bar{\delta}_{j\cdot}^{(2,a)}}$.

If s_0 *and* α *are the true values of parameters and the assumptions of Proposition 2 hold true, then* $\hat{s}_{0j} \xrightarrow{a.s.} s_0$ *and* $\hat{\alpha}_j \xrightarrow{a.s.} \alpha$, *when* $j \to +\infty$. *Moreover, there are almost surely finite random variables* c_{10} *and* c_{11} *such that, for all* $j \in \mathbb{N}$, *it holds*

$$|\hat{s}_{0j} - s_0| \le c_{10} \max\left(a_j^2 r_j, a_j^{-2}\right)$$

and

$$|\hat{\alpha}_j - \alpha| \le c_{11} \max\left(a_j^2 r_j, a_j^{-2}\right).$$

Remark 12. Note that, for any sequences $\{a_j\}$ and $\{r_j\}$, there exists a sequence $\{m_j\}$ such that the sums $\sum_{j=1}^{\infty} \frac{1}{r_j^2 m_j}$ and $\sum_{j=1}^{\infty} \frac{a_j^2}{r_j^2\gamma_j^2 m_j}$ are finite. Hence, one can get a specified convergence rate at levels j by choosing a sufficiently fast increasing range of averaging over k.

Example 4. Let us consider the case of $a_j \ge (1+\epsilon)^{j-1}$, $b_{jk} = k$, and $r_j = a_j^{-2.5}$, $j \in \mathbb{N}$. Then $|b_{jk_1} - b_{jk_2}| = |k_1 - k_2|$ and $\gamma_j = 1$. For $m_j = a_j^{\beta}$, the assumptions of Theorem 1 are satisfied if

$$\sum_{j=1}^{\infty}\frac{1}{r_j^2 m_j} = \sum_{j=1}^{\infty}\frac{1}{a_j^{\beta-5}} < +\infty \quad \text{and} \quad \sum_{j=1}^{\infty}\frac{a_j^2}{r_j^2\gamma_j^2 m_j} = \sum_{j=1}^{\infty}\frac{1}{a_j^{\beta-7}} < +\infty,$$

which is true for $\beta > 7$. Then, the order of the rate of convergence in Theorem 1 is $\max(a_j^2 r_j, a_j^{-2}) = a_j^{-0.5}$.

6 | SIMULATION STUDIES

This section presents some numerical studies to confirm the theoretical findings. It intends to give an example illustrating the application of the developed methodology, rather than conducting extensive numerical studies and providing practical recommendations for various possible scenarios, which will be explored in future publications. The results demonstrate that the approach can be extended to other processes and filters.

The generalized filtered method-of-moments approach in this paper was developed for functional time series. However, only discrete time can be used for computer simulations. In simulation studies and real-world applications, different discretization strategies are used for approximating functional data (see, e.g., §6.4.3 in Ramsay & Silverman, 2013). In the available literature, it is usually assumed as a matter of fact that the discretization error is negligible with respect to the estimation error. There are only few known results that rigorously prove it (see Alodat & Olenko, 2018; Ayache & Bertrand, 2011; Bardet & Bertrand, 2010; Leonenko & Taufer, 2006). We intend to devote another publication to investigating discretization errors. Meantime, to illustrate the generalized filtered method-of-moments approach, we use Riemann sums to approximate the integrals in δ_{jk}. To obtain high-precision approximations of the integrals, we use 100 points per time unit in the Riemann sums and the length of time intervals equal 100,000 time units.

Remark 13. Theorem 1 shows that the accuracy of the parameter estimates increases with increasing the number of levels j. However, in the case of real-world or simulated data, functional time series are observed only at a finite number N of time points. In the wavelet analysis of time series, it is recommended to select the maximum number of levels j satisfying $N \approx 2^j$. New information tends to zero if j substantially exceeds $\log_2 N$. The generalized filtered method-of-moments approach deals with the case of general filters that allow using m_j points at each level. Hence, the maximum number of levels j should be selected to satisfy $N \approx m_j$.

We consider the Gegenbauer random process $X(t), t \in \mathbb{Z}$; see Alomari et al. (2017), Espejo et al. (2015), and the references therein. This random process satisfies the following equation:

$$\Delta_u^d X(t) := (1 - 2uB + B^2)^d X(t) = \varepsilon_t,$$

where Δ_u^d is the fractional difference operator given by

$$\Delta_u^d = (1 - 2uB + B^2)^d = (1 - 2\cos(\nu)B + B^2)^d = [(1 - e^{i\nu}B)(1 - e^{-i\nu}B)]^d;$$

B denotes the backward-shift operator for the time coordinate t, that is, $BX_t = X_{t-1}$, $u = \cos\nu$ (i.e., $\nu = \arccos(u)$, $|u| \le 1$), $d \in (-\frac{1}{2}, \frac{1}{2})$; and ε_t is a zero-mean white noise with the common variance $E(\varepsilon_t^2) = \sigma_\varepsilon^2$.

There exists the following representation of a stationary Gegenbauer random process:

$$X(t) = \sum_{n=0}^{\infty} C_n^{(d)}(u)\varepsilon_{t-n}, \tag{9}$$

where $d \neq 0$ and the Gegenbauer polynomial $C_n^{(d)}(u)$ is given by

$$C_n^{(d)}(u) = \sum_{k=0}^{[n/2]} (-1)^k \frac{(2u)^{n-2k}\Gamma(d-k+n)}{k!(n-2k)!\Gamma(d)},$$

where $[n/2]$ is the integer part of $n/2$ and $\Gamma(\cdot)$ is the gamma function.

We generated the random process $X(t)$ using the parameter values $d = 0.1$ and $u = 0.3$. The chosen parameters d and u correspond to s_0 and α inside of the admissible region R_{y}. The realizations of $X(t)$ were approximated by truncated sums with 100 terms in (9). Note that, for $n \to \infty$, it holds

$$C_n^{(d)}(u) \sim \frac{\cos((n+d)\nu - d\pi/2)}{\Gamma(d)\sin^d(\nu)} \left(\frac{2}{n}\right)^{1-d},$$

where $\nu = \arccos(u)$. Thus, for simulating Gegenbauer processes with different parameters, especially when values of ν are close to 0, more terms may be required.

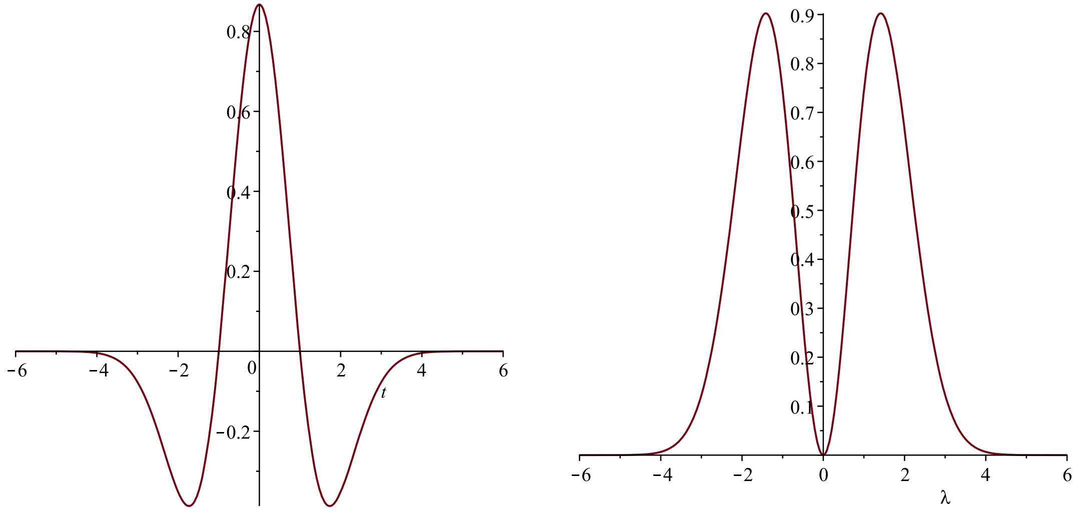

FIGURE 6 Plots of the Mexican hat wavelet $\psi(t)$ and its Fourier transform $\hat{\psi}(\lambda)$ [Colour figure can be viewed at wileyonlinelibrary.com]

The Mexican hat wavelet was used as a filter. It is defined by

$$\psi(t) = \frac{2}{\sqrt{3\sigma\pi^{\frac{1}{4}}}}\left(1 - \left(\frac{t}{\sigma}\right)^2\right)e^{-\frac{t^2}{2\sigma^2}}.$$

Its Fourier transform is (see Liu, 2010)

$$\hat{\psi}(\lambda) = \frac{\sqrt{8}\pi^{\frac{1}{4}}\sigma^{\frac{5}{2}}}{\sqrt{3}}\lambda^2 e^{-\frac{\sigma^2\lambda^2}{2}}.$$

Note that the Fourier transform $\hat{\psi}(\lambda)$ does not have a finite support but approaches zero very quickly when $\lambda \to +\infty$. The value $\sigma = 1$ was used in computations. Plots of $\psi(t)$ and $\hat{\psi}(\lambda)$ are shown in Figure 6. In this case, c_2 and c_3 are 2 and 10, respectively. The filter transform of $X(t)$ defined by (1) was computed using the R package WMTSA.

The random process $X(t)$ was generated 1,000 times over the time grid specified above, and the corresponding wavelet coefficients δ_{jk} were calculated for each generated trajectory. At first, to compute $\hat{s}_{0j}$ and $\hat{\alpha}_j$, the statistics $\bar{\delta}_{j\cdot}^{(2)}$ and $\Delta\bar{\delta}_{j\cdot}^{(2)}$ were found using the values $a_j = j$, $b_{jk} = k$, $\gamma_j = 1$, $r_j = a_j^{-2.5}$, and $m_j = a_j^9$, for $j = 1, \ldots, 7$. By Example 4, for $\varepsilon = 0.3$, these values satisfy the assumptions of Theorem 1. Figure 7 displays box plots of $\bar{\delta}_{j\cdot}^{(2)}$ and $\Delta\bar{\delta}_{j\cdot}^{(2)}$ for the simulated realizations. The values of a_j are shown along the horizontal axe. The horizontal dashed lines show the true values of the corresponding parameters. These plots confirm that $\bar{\delta}_{j\cdot}^{(2)}$ and $\Delta\bar{\delta}_{j\cdot}^{(2)}$ converge as j increases. As expected, consult the upper bound (4) in Proposition 2, the rate of convergence of $\Delta\bar{\delta}_{j\cdot}^{(2)}$ is slower than in the case of $\bar{\delta}_{j\cdot}^{(2)}$. Finally, the estimates $\hat{s}_{0j}$ and $\hat{\alpha}_j$ were calculated by (8) for each simulation. Figure 8 demonstrates that $\hat{s}_{0j}$ convergence to $s_0 = \arccos(u)$ and $\hat{\alpha}_j$ to $\alpha = d$ as j increases.

As the true values of parameters correspond to a point inside of $R_{\mathbf{y}}$, the majority of the parameter estimates are in the admissible region. Figures 7 and 8 suggest that the adjusted statistics should be applied mainly for the cases $j = 1$ and 2.

Table 1 below gives numerical values of root mean square errors (RMSEs) for each parameter estimated in Figures 7 and 8. These results numerically confirm the theoretical convergence.

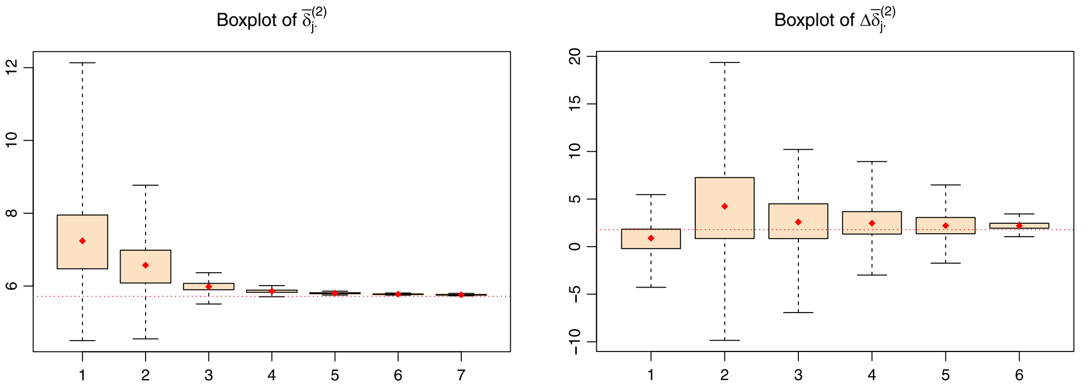

FIGURE 7 Boxplots of the first and second statistics $\bar{\delta}_{j\cdot}^{(2)}$ and $\Delta\bar{\delta}_{j\cdot}^{(2)}$ and their true values (horizontal dashed lines) [Colour figure can be viewed at wileyonlinelibrary.com]

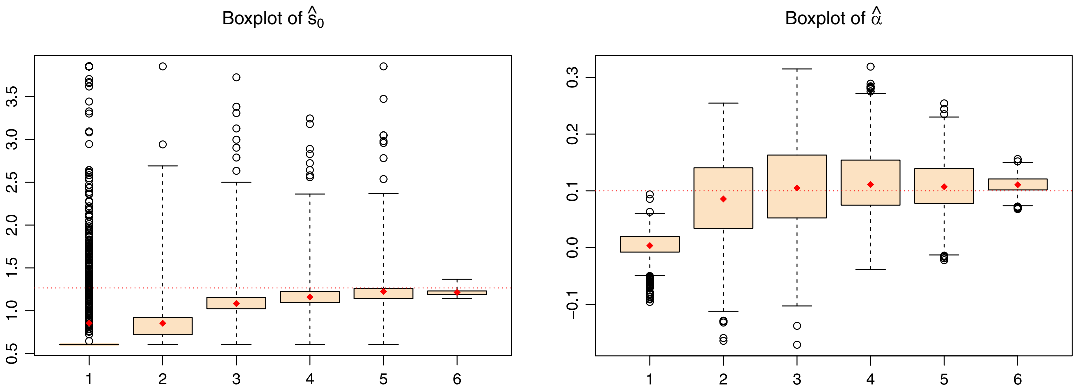

FIGURE 8 Boxplots of the estimates $\hat{s}_{0j}$ and $\hat{\alpha}_{j\cdot}$ and the corresponding true values of parameters (horizontal dashed lines) [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Root mean square error (RMSE) for each statistic and parameter

RMSEs	1	2	3	4	5	6
RMSE($\bar{\delta}_{j\cdot}^{(2)}$)	1.86690313	1.08704874	0.29951503	0.14773844	0.08741261	0.05930194
RMSE($\Delta\bar{\delta}_{j\cdot}^{(2)}$)	1.7436818	5.3719516	2.8432257	1.9604925	1.3885922	0.5747705
RMSE($\hat{s}_{0j}$)	0.72478156	0.48657429	0.37222642	0.31216483	0.27854370	0.06254635
RMSE($\hat{\alpha}_j$)	0.09968069	0.07616249	0.08119757	0.06502088	0.04983931	0.01845817

7 | DIRECTIONS FOR FUTURE RESEARCH

This paper has discussed statistical inference for parameters of cyclic long-memory processes with a spectral singularity at a nonzero frequency. The results were derived for wide classes of models with Gegenbauer-type spectral densities using very general filter transforms.

An important area for future explorations is obtaining similar results for the case of multiple singularities with the long-memory parameters varying across singularity locations; see the discussion on SCLM (seasonal/cyclical long memory) in Arteche and Robinson (1999). For the case of multiple unknown parameters, one can derive additional estimation equations similar to the ones in Section 4 using higher order differences of $\bar{\delta}_{j\cdot}^{(2)}$.

As this paper studied the case of Gegenbauer-type spectral densities given in Assumption 1, it would be interesting to apply the developed methodology to other cyclic long-memory models.

This paper develops statistical inference for parameters using functional data. There are numerous applications where $X(t)$ is observed only on a discrete grid or at random moments of a finite time interval. In addition, cyclic long-memory processes are often determined by discrete-time fractional autoregressive integrated moving average (FARIMA) models. In such cases, approximate formulas are used to compute filter transforms; see Section 3.2 in Bardet and Bertrand (2010). We plan to investigate statistical properties of the corresponding "approximate" estimates using approaches similar to Ayache and Bertrand (2011) and Bardet and Bertrand (2010).

Finally, it is important to extend the methodology to the multidimensional case of random fields; see Espejo et al. (2015).

ACKNOWLEDGEMENTS

We appreciate the editor and referees for their comments and suggestions that helped to improve this paper.

Andriy Olenko is grateful to Laboratoire d'Excellence, Centre Européen pour les Mathématiques, la Physique et leurs interactions (CEMPI, ANR-11-LABX-0007-01), Laboratoire de Mathématiques Paul Painlevé, France, for the support and for giving him the opportunity to pursue research at the Université Lille 1 for 2 months.

Antoine Ayache is grateful to the Sydney Mathematical Research Institute at the University of Sydney for having financially supported his 7-week visit to Australia in 2019, and for the position of visiting researcher at the University of Sydney. In addition, he is thankful to the School of Engineering and Mathematical Sciences at La Trobe University Australia for the honorary visiting professorship for 1 month at this university.

Andriy Olenko was partially supported under the Australian Research Council's Discovery Projects funding scheme (project number DP160101366).

SUPPORTING INFORMATION

Additional information for this article is available online in the folder "Research materials" from https://sites.goggle.com/site/olenkoandriy/ including the code used for simulations and examples.

ORCID

Andriy Olenko https://orcid.org/0000-0002-0917-7000

REFERENCES

Alodat, T., & Olenko, A. (2018). Weak convergence of weighted additive functionals of long-range dependent fields. *Theory of Probability and Mathematical Statistics*, *97*, 9–23.

Alomari, H. M., Frías, M. P., Leonenko, N. N., Ruiz-Medina, M. D., Sakhno, L., & Torres, A. (2017). Asymptotic properties of parameter estimates for random fields with tapered data. *Electronic Journal of Statistics*, *11*, 3332–3367.

Anh, V. V., Leonenko, N. N., & Sakhno, L. M. (2007). Statistical inference based on the information of the second and third order. *Journal of Statistical Planning and Inference*, *137*, 1302–1331.

Arteche, J., & Robinson, P. M. (1999). Seasonal and cyclical long memory. In S. Ghosh (Ed.), *Asymptotics, nonparametrics, and time series* (pp. 115–148). New York, NY: Marcel Dekker, Inc.

Arteche, J., & Robinson, P. M. (2000). Semiparametric inference in seasonal and cyclical long memory processes. *Journal of Time Series Analysis*, *21*, 1–25.

Ayache, A., & Bertrand, P. R. (2011). Discretization error of wavelet coefficient for fractal like processes. *Advances in Pure and Applied Mathematics*, *2*, 297–321.

Bardet, J.-M., & Bertrand, P. R. (2010). A non-parametric estimator of the spectral density of a continuous-time Gaussian process observed at random times. *Scandinavian Journal of Statistics*, *37*, 458–476.

Beran, J., Ghosh, S., & Schell, D. (2009). On least squares estimation for long-memory lattice processes. *Journal of Multivariate Analysis*, *100*, 2178–2194.

Chung, C.-F. (1966). A generalized fractional integrated autoregressive moving-average process. *Journal of Time Series Analysis*, *17*, 111–140.

Clausel, M., Roueff, F., Taqqu, M. S., & Tudor, C. (2014). Wavelet estimation of the long memory parameter for hermite polynomial of Gaussian processes. *ESAIM: Probability and Statistics*, *18*, 42–76.

Cont, R. (2005). Long range dependence in financial markets. In J. Lévy-Véhel & E. Lutton (Eds.), *Fractals in engineering: New trends in theory and applications* (pp. 159–179). London, UK: Springer.

Corless, R. M., Gonnet, G. H., Hare, D. E. G., Jeffrey, D. J., & Knuth, D. E. (1996). On the Lambert *W* function. *Advances in Computational Mathematics*, *5*, 329–359.

Daubechies, I. (1992). *Ten lectures on wavelets*. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Dehling, H., Rooch, A., & Taqqu, M. S. (2013). Non-parametric change-point tests for long-range dependent data. *Scandinavian Journal of Statistics*, *40*, 153–173.

Espejo, R. M., Leonenko, N. N., Olenko, A., & Ruiz-Medina, M. D. (2015). On a class of minimum contrast estimators for Gegenbauer random fields. *TEST*, *24*, 657–680.

Ferrara, L., & Guígan, D. (2001). Comparison of parameter estimation methods in cyclical long memory time series. In A. Timmermann, C. Dunis, & J. Moody (Eds.), *Developments in forecast combination and portfolio choice*. New York, NY: Wiley.

Gikhman, I. I., & Skorokhod, A. V. (2004). *The theory of stochastic processes II*. Heidelberg, Germany: Springer.

Giraitis, L., Hidalgo, J., & Robinson, P. M. (2001). Gaussian estimation of parametric spectral density with unknown pole. *Annals of Statistics*, *29*, 987–1023.

Granger, C. W. J., & Joyeux, R. (1980). An introduction to long-range time series models and fractional referencing. *Journal of Time Series Analysis*, *1*, 15–30.

Guo, H., Lim, C. Y., & Meerschaert, M. M. (2009). Local Whittle estimator for anisotropic random fields. *Journal of Multivariate Analysis*, *100*, 993–1028.

Hidalgo, J. (1996). Spectral analysis for bivariate time series with long memory. *Econometric Theory*, *12*, 773–792.

Hidalgo, J. (2005). Semiparametric estimation for stationary processes whose spectra have an unknown pole. *The Annals of Statistics*, *33*, 1843–1889.

Hosking, J. R. M. (1981). Fractional differencing. *Biomertrika*, *28*, 165–176.

Hosoya, Y. (1997). A limit theory for long-range dependence and statistical inference on related models. *Annals of Statistics*, *28*, 105–137.

Ivanov, A. V., Leonenko, N., Ruiz-Medina, M. D., & Savich, I. N. (2013). Limit theorems for weighted non-linear transformations of Gaussian stationary processes with singular spectra. *Annals of Probability*, *41*, 1088–1114.

Klykavka, B., Olenko, A., & Vicendese, M. (2012). Asymptotic behaviour of functionals of cyclical long-range dependent random fields. *Journal of Mathematical Sciences*, *187*, 35–48.

Leonenko, N. N. (1999). *Limit theorems for random fields with singular spectrum*. Dordrecht, The Netherlands: Kluwer.

Leonenko, N. N., & Olenko, A. (2013). Tauberian and Abelian theorems for long-range dependent random fields. *Methodology and Computing in Applied Probability*, *15*, 715–742.

Leonenko, N. N., & Olenko, A. (2014). Sojourn measures of Student and Fisher–Snedecor random fields. *Bernoulli*, *20*, 1454–1483.

Leonenko, N. N., & Sakhno, L. M. (2006). On the Whittle estimators for some classes of continuous-parameter random processes and fields. *Statistics and Probability Letters*, *76*, 781–795.

Leonenko, N. N., & Taufer, E. (2006). Weak convergence of functionals of stationary long memory processes to Rosenblatt-type distributions. *Journal of Statistical Planning and Inference*, *136*, 1220–1236.

Liu, C. L. (2010). A tutorial of the wavelet transform. Taipei, Taiwan: NTUEE. Retrieved from http://disp.ee.ntu.edu.tw/tutorial/WaveletTutorial.pdf

Meyer, Y. (1992). *Wavelets and operators*. Cambridge, UK: Cambridge University Press.

Olenko, A. (2013). Limit theorems for weighted functionals of cyclical long-range dependent random fields. *Stochastic Analysis and Applications*, *31*, 199–213.

Park, C., Hernández-Campos, F., Le, L., Marron, J. S., Park, J., Pipiras, V., … Zhu, Z. (2011). Long-range dependence analysis of internet traffic. *Journal of Applied Statistics*, *38*, 1407–1433.

Pipiras, V., & Taqqu, M. S. (2017). *Long-range dependence and self-similarity*. Cambridge, UK: Cambridge University Press.

Ramsay, J. O., & Silverman, B. W. (2013). *Functional data analysis*. New York, NY: Springer.

Samorodnitsky, G. (2007). Long range dependence. *Foundations and Trends in Stochastic Systems*, *1*, 163–257.

Samorodnitsky, G. (2016). *Stochastic processes and long range dependence*. Cambridge, UK: Cambridge University Press.

Taqqu, M. S. (1979). Convergence of integrated processes of arbitrary Hermite rank. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, *50*, 53–83.

Tsai, H., Rachinger, H., & Lin, E. M. H. (2015). Inference of seasonal long-memory time series with measurement error. *Scandinavian Journal of Statistics*, *42*, 137–154.

Ushakov, N. G. (1999). *Selected topics in characteristic functions*. Utrecht, The Netherlands: VSP.

Whitcher, B. (2004). Wavelet-based estimation for seasonal long-memory processes. *Technometrics*, *46*, 225–238.

Willinger, W., Paxson, V., Riedi, R. H., & Taqqu, M. S. (2003). Long-range dependence and data network traffic. In P. Doukhan, G. Oppenheim, & M. S. Taqqu (Eds.) *Theory and applications of long-range dependence*. Boston, UK: Birkhauser Press.

Yajima, Y. (1985). On estimation of long-memory time series models. *Australian and New Zealand Journal of Statistics*, *27*, 303–320.

How to cite this article: Alomari HM, Ayache A, Fradon M, Olenko A. Estimation of cyclic long-memory parameters. *Scand J Statist*. 2019;1–30. https://doi.org/10.1111/sjos.12404

APPENDIX A

Here, we give the proofs for all results.

Proof of Lemma 1. The first inequality follows from the estimate

$$
\begin{aligned}
|h(\tilde{\lambda}) - h(\lambda)| &\leq \max_{\lambda_0 \in [\lambda, \tilde{\lambda}]} |h'(\lambda_0)||\tilde{\lambda} - \lambda| = \max_{\lambda_0 \in [\lambda, \tilde{\lambda}]} |h'(\lambda_0) - h'(0)||\tilde{\lambda} - \lambda| \\
&\leq \max_{\lambda_0 \in [\lambda, \tilde{\lambda}]} \sup_{\tilde{\lambda}_0 \in [0, \lambda_0]} |h''(\tilde{\lambda}_0)| \cdot \lambda_0 \cdot |\tilde{\lambda} - \lambda| \\
&\leq \sup_{\tilde{\lambda}_0 \in \left[0, \frac{1}{2}\right]} |h''(\tilde{\lambda}_0)| \cdot |\tilde{\lambda} + \lambda| \cdot |\tilde{\lambda} - \lambda| \leq c_1 |\tilde{\lambda}^2 - \lambda^2|.
\end{aligned}
$$

Substituting $\lambda = 0$, we get the second inequality. Finally, the third upper bound is obtained using the mean value theorem four times. □

Proof of Lemma 2. If $k_1 = k_2$, then by (3),

$$I(j,k_1,k_1) = s_0^{-4\alpha}\int_{-A}^{A}\frac{|\hat{\psi}(\lambda)|^2 h\left(\frac{\lambda}{a_j}\right)}{\left|\left(\frac{\lambda}{a_j s_0}\right)^2 - 1\right|^{2\alpha}}d\lambda \le c_2 s_0^{-4\alpha}\sup_{u\in\left[0,\frac{A}{a_j}\right]} h(u)\left|1-\left(\frac{A}{a_j s_0}\right)^2\right|^{-2\alpha}.$$

By the conditions $\frac{a_j}{2A} \ge 1$ and $s_0 > 1$, we get

$$\frac{A}{a_j} \le \frac{1}{2} \quad \text{and} \quad \frac{A}{a_j s_0} \le \frac{1}{2}. \tag{A1}$$

Hence, by Lemma 1,

$$|I(j,k_1,k_1)| \le c_2 s_0^{-4\alpha}\frac{1+c_1}{\left(1-\frac{1}{4}\right)^{2\alpha}} \le \frac{4}{3}(1+c_1)c_2 s_0^{-4\alpha}. \tag{A2}$$

Let us denote $p(\lambda) := \frac{|\hat{\psi}(\lambda)|^2 h(\frac{\lambda}{a_j})}{|(\frac{\lambda}{a_j s_0})^2-1|^{2\alpha}}$. By Assumptions 1 and 2, the function $p(\cdot)$ is a nonnegative integrable function on $[-A,A]$. Therefore, $\tilde{p}(\lambda) := p(\lambda)/\int_{-A}^{A} p(\lambda)d\lambda$ is a probability density. Moreover, by Assumption 1, we get $V_{-\frac{1}{2}}^{\frac{1}{2}}(h(\cdot)) = \int_{-\frac{1}{2}}^{\frac{1}{2}} |h'(\lambda)|d\lambda \le 2\int_0^{\frac{1}{2}} c_1 \cdot \lambda d\lambda < +\infty$. Hence, it follows from Assumption 2 that $\tilde{p}(\lambda)$ is a function of bounded variation on $[-A,A]$.

Therefore, for $k_1 \ne k_2$ by Theorem 2.5.3 in Ushakov (1999),

$$\left|\int_{-A}^{A} e^{i\frac{b_{jk_1}-b_{jk_2}}{a_j}\lambda}\tilde{p}(\lambda)d\lambda\right| \le \frac{a_j}{|b_{jk_1}-b_{jk_2}|}V_{-A}^{A}(\tilde{p}).$$

Hence, if $k_1 \ne k_2$, we obtain

$$|I(j,k_1,k_2)| \le s_0^{-4\alpha}\int_{-A}^{A} p(\lambda)\,d\lambda \cdot V_{-A}^{A}(\tilde{p})\cdot\frac{a_j}{|b_{jk_1}-b_{jk_2}|} = s_0^{-4\alpha}V_{-A}^{A}(p)\cdot\frac{a_j}{|b_{jk_1}-b_{jk_2}|}.$$

It follows from (A1) and $\alpha \in (0,1/2)$ that

$$\begin{aligned}
V_{-A}^{A}(p) &\le 2V_{-A}^{A}\left(|\hat{\psi}(\cdot)|^2 h\left(\frac{\cdot}{a_j}\right)\right)\\
&\le 2\left[\max_{\lambda\in[-A,A]}|\hat{\psi}(\lambda)|^2 V_{-A}^{A}\left(h\left(\frac{\cdot}{a_j}\right)\right) + \max_{\lambda\in[-A,A]} h\left(\frac{\lambda}{a_j}\right)V_{-A}^{A}(|\hat{\psi}(\cdot)|^2)\right]\\
&\le 2\left[\max_{\lambda\in[-A,A]}|\hat{\psi}(\lambda)|^2 V_{-1/2}^{1/2}(h(\cdot)) + \max_{\lambda\in\left[-\frac{1}{2},\frac{1}{2}\right]} h(\lambda)\,V_{-A}^{A}(|\hat{\psi}(\cdot)|^2)\right].
\end{aligned}$$

Note that this upper bound does not depend on j, a_j, b_{jk}.

Hence,

$$|I(j,k_1,k_2)| \le 2s_0^{-4\alpha}\left[\max_{\lambda\in[-A,A]}|\hat{\psi}(\lambda)|^2 V_{-1/2}^{1/2}(h(\cdot)) + \max_{\lambda\in\left[-\frac{1}{2},\frac{1}{2}\right]} h(\lambda)\cdot V_{-A}^{A}(|\hat{\psi}(\cdot)|^2)\right]. \tag{A3}$$

Comparing (A2) and (A3), we obtain the value of $c_4(s_0,\alpha)$ at the statement of Lemma 2. □

Proof of Lemma 3. Notice that, by (1), random variables δ_{jk} are centered Gaussian. For any centered two-dimensional Gaussian vector (z_1, z_2), it holds $\mathrm{Cov}(z_1^2, z_2^2) = 2\mathrm{Cov}^2(z_1, z_2)$. Therefore, recalling (2) and the definition of $I(j, k_1, k_2)$ in (3), we obtain

$$\mathrm{Var}\left(\bar{\delta}_{j\cdot}^{(2)}\right) = \mathrm{Var}\left(\frac{1}{m_j}\sum_{k=1}^{m_j}\delta_{jk}^2\right) = \frac{1}{m_j^2}\sum_{1\le k_1,k_2\le m_j}\mathrm{Cov}\left(\delta_{jk_1}^2, \delta_{jk_2}^2\right)$$
$$= \frac{2}{m_j^2}\sum_{1\le k_1,k_2\le m_j} I^2(j,k_1,k_2) = \frac{2}{m_j^2}\left(\sum_{k_1=1}^{m_j} I^2(j,k_1,k_1) + \sum_{\substack{1\le k_1,k_2\le m_j\\ k_1\ne k_2}} I^2(j,k_1,k_2)\right).$$

By Lemma 2, it follows that

$$\mathrm{Var}\left(\bar{\delta}_{j\cdot}^{(2)}\right) \le \frac{2}{m_j^2}\left(c_4^2(s_0,\alpha)m_j + 2c_4^2(s_0,\alpha)\sum_{k_1=1}^{m_j}\sum_{k_2=k_1+1}^{m_j}\frac{a_j^2}{(b_{jk_1}-b_{jk_2})^2}\right)$$
$$\le \frac{2c_4^2(s_0,\alpha)}{m_j^2}\left(m_j + \frac{2a_j^2}{\gamma_j^2}\sum_{k_1=1}^{m_j}\sum_{k_2=k_1+1}^{m_j}\frac{1}{(k_2-k_1)^2}\right)$$
$$\le \frac{2c_4^2(s_0,\alpha)}{m_j^2}\left(m_j + \frac{2a_j^2}{\gamma_j^2}\sum_{k_1=1}^{m_j}\sum_{\tilde{k}_2=1}^{\infty}\frac{1}{\tilde{k}_2^2}\right) = \frac{2c_4^2(s_0,\alpha)}{m_j}\left(1+\frac{\pi^2}{3}\frac{a_j^2}{\gamma_j^2}\right).$$

□

Proof of Lemma 4. Using Chebyshev inequality and Lemma 3, we obtain

$$P\left(\left|\bar{\delta}_{j\cdot}^{(2)} - J(a_j)\right| > r_j\right) \le \frac{\mathrm{Var}\left(\bar{\delta}_{j\cdot}^{(2)}\right)}{r_j^2} \le \frac{c_{5j}(s_0,\alpha)}{r_j^2 m_j} = 2c_4^2(s_0,\alpha)\frac{\left(1+\frac{\pi^2}{3}\frac{a_j^2}{\gamma_j^2}\right)}{r_j^2 m_j}.$$

By the choice of $\{m_j\}$,

$$\sum_{j=1}^{\infty} P\left(\left|\bar{\delta}_{j\cdot}^{(2)} - J(a_j)\right| > r_j\right) < +\infty.$$

Therefore, applying the Borel–Cantelli lemma, we obtain the required statement. □

We will use the next technical result.

Lemma 10. *For all* $\alpha \in (0, \frac{1}{2})$ *and* $x \in [0, \frac{1}{2}]$, *it holds*

$$0 \le (1-x)^{-2\alpha} - 1 \le 4x.$$

Proof of Lemma 10. Applying the mean value theorem to the function $\gamma_\alpha(y) = (1-y)^{-2\alpha}$ on the interval $[0, x]$, we obtain

$$(1-x)^{-2\alpha} - 1 = \gamma_\alpha(x) - \gamma_\alpha(0) = \gamma_\alpha'(a_\alpha)x = 2\alpha(1-a_\alpha)^{-2\alpha-1}x \le (1-a_\alpha)^{-2\alpha-1}x,$$

where $a_\alpha \in [0, x] \subset [0, \frac{1}{2}]$.

Therefore, as $\alpha \in (0, \frac{1}{2})$, we get $(1-x)^{-2\alpha} - 1 \le (1-\frac{1}{2})^{-2}x = 4x$. □

Proof of Lemma 5. Noting that $J(a_j) = I(j, k_1, k_1)$ and using (3), we get

$$J(a_j) - c_2 s_0^{-4\alpha} = s_0^{-4\alpha} \int_{-A}^{A} |\hat{\psi}(\lambda)|^2 \left(\frac{h\left(\frac{\lambda}{a_j}\right)}{\left|1 - \left(\frac{\lambda}{a_j s_0}\right)^2\right|^{2\alpha}} - 1 \right) d\lambda.$$

Now, by Lemma 10, the conditions on $h(\cdot)$ in Assumption 1 and Lemma 1, it follows

$$\begin{aligned}
\left|J(a_j) - c_2 s_0^{-4\alpha}\right| &\le s_0^{-4\alpha} \int_{-A}^{A} |\hat{\psi}(\lambda)|^2 \left| h\left(\frac{\lambda}{a_j}\right) \left[\left(1 - \left(\frac{\lambda}{a_j s_0}\right)^2\right)^{-2\alpha} - 1 \right] + h\left(\frac{\lambda}{a_j}\right) - 1 \right| d\lambda \\
&\le s_0^{-4\alpha} \int_{-A}^{A} |\hat{\psi}(\lambda)|^2 \left(4h\left(\frac{\lambda}{a_j}\right) \left(\frac{\lambda}{a_j s_0}\right)^2 + c_1 \left(\frac{\lambda}{a_j}\right)^2 \right) d\lambda.
\end{aligned}$$

Moreover, it follows from Lemma 1 and the conditions of the lemma that, for $\lambda \in [-A, A]$,

$$4h\left(\frac{\lambda}{a_j}\right) \left(\frac{\lambda}{a_j s_0}\right)^2 + c_1 \left(\frac{\lambda}{a_j}\right)^2 \le \left(4(1 + c_1)\frac{A^2}{s_0^2} + c_1 A^2 \right) a_j^{-2},$$

which completes the proof. □

Proof of Lemma 6. By (2) and Assumption 1, we get

$$\begin{aligned}
\frac{J(a_{j+1}) - J(a_j)}{a_j^{-2} - a_{j+1}^{-2}} &= \int_{\mathbb{R}} \frac{|\hat{\psi}(\lambda)|^2}{a_j^{-2} - a_{j+1}^{-2}} \left(\frac{h\left(\frac{\lambda}{a_{j+1}}\right)}{\left|s_0^2 - \left(\frac{\lambda}{a_{j+1}}\right)^2\right|^{2\alpha}} - \frac{h\left(\frac{\lambda}{a_j}\right)}{\left|s_0^2 - \left(\frac{\lambda}{a_j}\right)^2\right|^{2\alpha}} \right) d\lambda \\
&= \int_{-A}^{A} |\hat{\psi}(\lambda)|^2 \frac{h\left(\frac{\lambda}{a_{j+1}}\right) \left|s_0^2 - \left(\frac{\lambda}{a_j}\right)^2\right|^{2\alpha} - h\left(\frac{\lambda}{a_j}\right) \left|s_0^2 - \left(\frac{\lambda}{a_{j+1}}\right)^2\right|^{2\alpha}}{\left(a_j^{-2} - a_{j+1}^{-2}\right) \left|s_0^2 - \left(\frac{\lambda}{a_{j+1}}\right)^2\right|^{2\alpha} \left|s_0^2 - \left(\frac{\lambda}{a_j}\right)^2\right|^{2\alpha}} d\lambda.
\end{aligned}$$

As $\lambda \in [-A, A]$ and $a_j \to +\infty$ when $j \to +\infty$, then there is $j_0 \in \mathbb{N}$, such that $\frac{|\lambda|}{a_j} \le \frac{1}{2}$, for all $\lambda \in [-A, A]$ and $j \ge j_0$. Hence, using the inequality $s_0 > 1$, for sufficiently large $j \ge j_0$, the integrand can be bounded as it is shown as follows:

$$\begin{aligned}
&|\hat{\psi}(\lambda)|^2 \frac{\left| h\left(\frac{\lambda}{a_{j+1}}\right) \left|s_0^2 - \left(\frac{\lambda}{a_j}\right)^2\right|^{2\alpha} - h\left(\frac{\lambda}{a_j}\right) \left|s_0^2 - \left(\frac{\lambda}{a_{j+1}}\right)^2\right|^{2\alpha} \right|}{\left(a_j^{-2} - a_{j+1}^{-2}\right) \left|s_0^2 - \left(\frac{\lambda}{a_{j+1}}\right)^2\right|^{2\alpha} \left|s_0^2 - \left(\frac{\lambda}{a_j}\right)^2\right|^{2\alpha}} \\
&\le \frac{|\hat{\psi}(\lambda)|^2 \left| h\left(\frac{\lambda}{a_{j+1}}\right) \left|s_0^2 - \left(\frac{\lambda}{a_j}\right)^2\right|^{2\alpha} - h\left(\frac{\lambda}{a_j}\right) \left|s_0^2 - \left(\frac{\lambda}{a_{j+1}}\right)^2\right|^{2\alpha} \right|}{\left(a_j^{-2} - a_{j+1}^{-2}\right) \left|s_0^2 - s_0^2 \max\limits_{\lambda \in [-A,A]} \left(\frac{\lambda}{a_{j+1}}\right)^2\right|^{2\alpha} \left|s_0^2 - s_0^2 \max\limits_{\lambda \in [-A,A]} \left(\frac{\lambda}{a_j}\right)^2\right|^{2\alpha}} \\
&\le |\hat{\psi}(\lambda)|^2 \frac{\left| h\left(\frac{\lambda}{a_{j+1}}\right) - h\left(\frac{\lambda}{a_j}\right) \right| \left|s_0^2 - \left(\frac{\lambda}{a_j}\right)^2\right|^{2\alpha} + h\left(\frac{\lambda}{a_j}\right) \left| \left|s_0^2 - \left(\frac{\lambda}{a_j}\right)^2\right|^{2\alpha} - \left|s_0^2 - \left(\frac{\lambda}{a_{j+1}}\right)^2\right|^{2\alpha} \right|}{\left(3s_0^2/4\right)^{4\alpha} \left(a_j^{-2} - a_{j+1}^{-2}\right)}. \qquad \text{(A4)}
\end{aligned}$$

Applying the first inequality in Lemma 1 to $h(\frac{\lambda}{a_{j+1}}) - h(\frac{\lambda}{a_j})$ and the mean value theorem to $(s_0^2 - (\frac{\lambda}{a_j})^2)^{2\alpha} - (s_0^2 - (\frac{\lambda}{a_{j+1}})^2)^{2\alpha}$, we obtain that the upper bound for the right part of (A4) is

$$\left|\widehat{\psi}(\lambda)\right|^2 \frac{c_1 s_0^{4\alpha}\lambda^2 + 2\alpha\,(1+c_1)\left(3s_0^2/4\right)^{2\alpha-1}\lambda^2}{\left(3s_0^2/4\right)^{2\alpha}}.$$

This upper bound is integrable and does not depend on j. Hence, one can use the dominated convergence theorem. For $\lambda \in [-A, A]$, it holds $\lim_{j\to+\infty} h(\frac{\lambda}{a_j}) = 1$ and $\lim_{j\to+\infty} |s_0^2 - (\frac{\lambda}{a_j})^2|^{2\alpha} = s_0^{4\alpha}$. Hence,

$$\lim_{j\to+\infty} \frac{h\left(\frac{\lambda}{a_{j+1}}\right)\left|s_0^2 - \left(\frac{\lambda}{a_j}\right)^2\right|^{2\alpha} - h\left(\frac{\lambda}{a_j}\right)\left|s_0^2 - \left(\frac{\lambda}{a_{j+1}}\right)^2\right|^{2\alpha}}{\left(a_j^{-2} - a_{j+1}^{-2}\right)\left|s_0^2 - \left(\frac{\lambda}{a_{j+1}}\right)^2\right|^{2\alpha}\left|s_0^2 - \left(\frac{\lambda}{a_j}\right)^2\right|^{2\alpha}}$$

$$= \lim_{j\to+\infty} \frac{h\left(\frac{\lambda}{a_{j+1}}\right)\left|s_0^2 - \left(\frac{\lambda}{a_j}\right)^2\right|^{2\alpha}}{\left(a_j^{-2} - a_{j+1}^{-2}\right)s_0^{8\alpha}}\left(1 - \frac{h\left(\frac{\lambda}{a_j}\right)}{h\left(\frac{\lambda}{a_{j+1}}\right)}\left|\frac{s_0^2 - \left(\frac{\lambda}{a_{j+1}}\right)^2}{s_0^2 - \left(\frac{\lambda}{a_j}\right)^2}\right|^{2\alpha}\right)$$

$$= \lim_{j\to+\infty} \frac{s_0^{-4\alpha}}{a_j^{-2} - a_{j+1}^{-2}}\left(1 - \frac{h\left(\frac{\lambda}{a_j}\right)}{h\left(\frac{\lambda}{a_{j+1}}\right)}\left|1 + \frac{\left(\frac{\lambda}{a_j}\right)^2 - \left(\frac{\lambda}{a_{j+1}}\right)^2}{s_0^2 - \left(\frac{\lambda}{a_j}\right)^2}\right|^{2\alpha}\right). \tag{A5}$$

Using L'Hôpital's rule, one can see that, for $\alpha \in (0, 1/2)$, it holds $\lim_{x\to 0} \frac{1-(1+x)^{2\alpha}}{x} = -2\alpha$.

Noting that, for $\lambda \in [-A, A]$, we get

$$\frac{\left|h\left(\frac{\lambda}{a_{j+1}}\right) - h\left(\frac{\lambda}{a_j}\right)\right|}{a_j^{-2} - a_{j+1}^{-2}} \le \sup_{\lambda_0\in[0,A]} \left|h''\left(\frac{\lambda_0}{a_j}\right)\right| \cdot \lambda^2 \to 0,$$

when $a_j \to +\infty$; we obtain that (A5) equals

$$\lim_{j\to+\infty} \frac{s_0^{-4\alpha}}{\left(a_j^{-2} - a_{j+1}^{-2}\right)} \frac{h\left(\frac{\lambda}{a_j}\right)}{h\left(\frac{\lambda}{a_{j+1}}\right)}\left(\frac{h\left(\frac{\lambda}{a_{j+1}}\right)}{h\left(\frac{\lambda}{a_j}\right)} - 1 + 1 - \left|1 + \frac{\left(\frac{\lambda}{a_j}\right)^2 - \left(\frac{\lambda}{a_{j+1}}\right)^2}{s_0^2 - \left(\frac{\lambda}{a_j}\right)^2}\right|^{2\alpha}\right)$$

$$= \lim_{j\to+\infty} \frac{s_0^{-4\alpha}}{\left(a_j^{-2} - a_{j+1}^{-2}\right)}\left(1 - \left|1 + \frac{\left(\frac{\lambda}{a_j}\right)^2 - \left(\frac{\lambda}{a_{j+1}}\right)^2}{s_0^2 - \left(\frac{\lambda}{a_j}\right)^2}\right|^{2\alpha}\right) = -\lim_{j\to+\infty} \frac{2\alpha s_0^{-4\alpha}\left(\left(\frac{\lambda}{a_j}\right)^2 - \left(\frac{\lambda}{a_{j+1}}\right)^2\right)}{\left(a_j^{-2} - a_{j+1}^{-2}\right)\left(s_0^2 - \left(\frac{\lambda}{a_j}\right)^2\right)}$$

$$= -2\alpha s_0^{-4\alpha-2}\lambda^2.$$

□

Proof of Lemma 7. Note that

$$\left|J(a_j) - J(a_{j+1}) - \alpha s_0^{-4\alpha-2} c_3 \cdot \left(a_j^{-2} - a_{j+1}^{-2}\right)\right|$$

$$= \left| \int_{-A}^{A} |\hat{\psi}(\lambda)|^2 \left(\frac{h\left(\frac{\lambda}{a_j}\right)}{\left|s_0^2 - \left(\frac{\lambda}{a_j}\right)^2\right|^{2\alpha}} - 2\alpha s_0^{-4\alpha-2} \left(\frac{\lambda}{a_j}\right)^2 - s_0^{-4\alpha} \right) d\lambda \right.$$

$$\left. - \int_{-A}^{A} |\hat{\psi}(\lambda)|^2 \left(\frac{h\left(\frac{\lambda}{a_{j+1}}\right)}{\left|s_0^2 - \left(\frac{\lambda}{a_{j+1}}\right)^2\right|^{2\alpha}} - 2\alpha s_0^{-4\alpha-2} \left(\frac{\lambda}{a_{j+1}}\right)^2 - s_0^{-4\alpha} \right) d\lambda \right|$$

$$\leq s_0^{-4\alpha} \int_{-A}^{A} |\hat{\psi}(\lambda)|^2 \left(\frac{\left|h\left(\frac{\lambda}{a_j}\right) - 1\right|}{\left|1 - \left(\frac{\lambda}{a_j s_0}\right)^2\right|^{2\alpha}} + \left| \frac{1}{\left|1 - \left(\frac{\lambda}{a_j s_0}\right)^2\right|^{2\alpha}} - 2\alpha \left(\frac{\lambda}{a_j s_0}\right)^2 - 1 \right| \right.$$

$$\left. + \frac{\left|h\left(\frac{\lambda}{a_{j+1}}\right) - 1\right|}{\left|1 - \left(\frac{\lambda}{a_{j+1} s_0}\right)^2\right|^{2\alpha}} + \left| \frac{1}{\left|1 - \left(\frac{\lambda}{a_{j+1} s_0}\right)^2\right|^{2\alpha}} - 2\alpha \left(\frac{\lambda}{a_{j+1} s_0}\right)^2 - 1 \right| \right) d\lambda. \tag{A6}$$

Let us consider the function $f(x) := (1-x)^{-2\alpha} - 2\alpha x - 1, x \in [0, 1/4], \alpha \in (0, 1/2)$. Notice that $f(\cdot) \in C^2[0, 1/4], f(0) = 0, f'(x) = 2\alpha(1-x)^{-2\alpha-1} - 2\alpha, f'(0) = 0, f''(x) = 2\alpha(2\alpha+1)(1-x)^{-2\alpha-2}$. Then, applying the mean value theorem twice, we get

$$|f(x) - f(0)| \leq \sup_{u \in [0,1/4]} f''(u) x^2, \quad x \in [0, 1/4].$$

Noting that $\sup_{u \in [0,1/4]} f''(u) \leq \frac{2^7}{3^3}$; it follows $|f(x) - f(0)| \leq \frac{2^7}{3^3} x^2, \quad x \in [0, 1/4]$.

Therefore, if $\frac{A}{a_j s_0} \leq \frac{1}{2}$, one can bound the second and the fourth terms of the integrand in (A6) by $\frac{2^7}{3^3}(\frac{\lambda}{a_j s_0})^4$ and $\frac{2^7}{3^3}(\frac{\lambda}{a_{j+1} s_0})^4$, respectively. By Lemma 1, if $\frac{A}{a_j} \leq \frac{1}{2}$, then the first and third terms in (A6) can be bounded by $2c_1(\frac{\lambda}{a_j})^4$ and $2c_1(\frac{\lambda}{a_{j+1}})^4$, respectively.

Combining the above bounds for sufficiently large j, we get

$$\left|J(a_j) - J(a_{j+1}) - \alpha c_3 s_0^{-4\alpha-2} \cdot \left(a_j^{-2} - a_{j+1}^{-2}\right)\right|$$

$$\leq s_0^{-4\alpha} \int_{-A}^{A} |\hat{\psi}(\lambda)|^2 \left[\frac{2^7}{3^3} \left(\frac{\lambda}{a_j s_0}\right)^4 + \frac{2^7}{3^3} \left(\frac{\lambda}{a_{j+1} s_0}\right)^4 + 2c_1 \left[\left(\frac{\lambda}{a_j}\right)^4 + \left(\frac{\lambda}{a_{j+1}}\right)^4 \right] \right] d\lambda$$

$$\leq \frac{2^8}{3^3} s_0^{-4\alpha-4} a_j^{-4} \int_{-A}^{A} |\hat{\psi}(\lambda)|^2 \lambda^4 \left(1 + 3^3 c_1 s_0^4 / 2^6\right) d\lambda.$$

Thus, for sufficiently large j, we obtain

$$\left|\frac{J(a_j)-J(a_{j+1})}{a_j^{-2}-a_{j+1}^{-2}}-\alpha c_3 s_0^{-4\alpha-2}\right|\le \frac{2^7A^2s_0^{-4\alpha-4}\left(1+3^3c_1s_0^4/2^6\right)}{3^3a_j^2\left(1-\frac{a_j^2}{a_{j+1}^2}\right)}c_3,$$

which completes the proof of Lemma 7. □

Proof of Proposition 2. Note that $\Delta\bar{\delta}_{j\cdot}^{(2)}$ can be rewritten as

$$\Delta\bar{\delta}_{j\cdot}^{(2)}=\frac{\left(\bar{\delta}_{j\cdot}^{(2)}-J(a_j)\right)-\left(\bar{\delta}_{j+1\cdot}^{(2)}-J(a_{j+1})\right)+\left(J(a_j)-J(a_{j+1})\right)}{a_j^{-2}-a_{j+1}^{-2}}.$$

Thus, by Lemmas 4 and 7,

$$\left|\Delta\bar{\delta}_{j\cdot}^{(2)}-\alpha c_3 s_0^{-4\alpha-2}\right|\le\frac{c_6(r_j+r_{j+1})}{a_j^{-2}-a_{j+1}^{-2}}+\frac{2^6A^2s_0^{-4\alpha-4}\left(1+3^3c_1s_0^4/2^6\right)}{3^3a_j^2\left(1-\left(\frac{a_j^2}{a_{j+1}^2}\right)^2\right)}c_3$$

$$\le\frac{2c_6}{1-\left(\frac{a_j}{a_{j+1}}\right)^2}a_j^2r_j+\frac{2^6A^2s_0^{-4\alpha-4}\left(1+3^3c_1s_0^4/2^6\right)c_3}{3^3\left(1-\left(\frac{a_j}{a_{j+1}}\right)^2\right)}a_j^{-2}$$

$$\le\frac{2c_6a_j^2r_j}{1-\frac{1}{(1+\varepsilon)^2}}+\frac{2^6A^2s_0^{-4\alpha-4}\left(1+3^3c_1s_0^4/2^6\right)c_3}{3^3\left(1-\frac{1}{(1+\varepsilon)^2}\right)}a_j^{-2}\le c_9\max\left(a_j^2r_j,a_j^{-2}\right),$$

where $c_9:=\frac{2c_6}{1-\frac{1}{(1+\varepsilon)^2}}+\frac{2^6A^2s_0^{-4\alpha-4}(1+3^3c_1s_0^4/2^6)c_3}{3^3(1-\frac{1}{(1+\varepsilon)^2})}$. □

Proof of Lemma 8. Let us find the range of the two-d–valued function

$$\mathbf{y}(s_0,\alpha)=\begin{pmatrix}y_1(s_0,\alpha)\\ y_2(s_0,\alpha)\end{pmatrix}:=\begin{pmatrix}s_0^{-4\alpha}\\ \alpha s_0^{-4\alpha-2}\end{pmatrix}$$

defined on the domain $(s_0,\alpha)\in(1,+\infty)\times(0,\frac{1}{2})$.

For simplicity, we use the notations y_1 and y_2 instead of $y_1(s_0,\alpha)$ and $y_2(s_0,\alpha)$ for the following computations.

As $s_0>1$, then for each $\alpha\in(0,\frac{1}{2})$, the range of possible values of y_1 is $(0,1)$. For each $\alpha\in(0,\frac{1}{2})$ and $y_1\in(0,1)$, there is such s_0 that $y_1=s_0^{-4\alpha}$. The variable y_2 can be expressed in terms of y_1 as $y_2=\alpha\cdot y_1^{1+\frac{1}{2\alpha}}$. Therefore, we can assume that y_1 is fixed and change only α to investigate the range of y_2.

Notice that

$$(y_2)'_\alpha=y_1^{1+\frac{1}{2\alpha}}+\alpha y_1^{1+\frac{1}{2\alpha}}\ln(y_1)\cdot\left(-\frac{1}{2\alpha^2}\right)=y_1^{1+\frac{1}{2\alpha}}\left(1-\frac{\ln(y_1)}{2\alpha}\right).$$

If $y_1\in(0,1)$, then $(y_2)'_\alpha>0$ and y_2 is an increasing function of α with the range $(0,y_1^2/2)$. Hence, the range of the function $\mathbf{y}(s_0,\alpha)$ on the domain $(1,+\infty)\times(0,\frac{1}{2})$ is $R_{\mathbf{y}}$, which completes the proof. □

Proof of Lemma 9. If (s_0, α) and (s_0', α') are two solutions of the system (6) for some $(y_1, y_2) \in R_y$, then

$$\begin{cases} s_0^{-4\alpha} = \left(s_0'\right)^{-4\alpha'}, \\ \alpha s_0^{-2} = \alpha'\left(s_0'\right)^{-2}, \end{cases}$$

and therefore,

$$\begin{cases} s_0^{\alpha} = \left(s_0'\right)^{\alpha'}, \\ \left(\frac{\alpha}{s_0^2}\right)^{\alpha} = \left(\frac{\alpha'}{(s_0')^2}\right)^{\alpha}. \end{cases}$$

Hence, $(\frac{\alpha'}{\alpha})^{\alpha} = (\frac{(s_0')^2}{s_0^2})^{\alpha} = (s_0')^{2(\alpha-\alpha')}$ and $\frac{\alpha'}{\alpha} = (s_0')^{2(1-\frac{\alpha'}{\alpha})}$.

Denoting $\frac{\alpha'}{\alpha} = t$ and $(s_0')^2 = a$, we obtain the equation

$$ta^{t-1} = 1, \quad t \in \mathbb{R}_+. \tag{A7}$$

As $s_0 > 1$, then s_0' must also be greater than 1 (otherwise, $\alpha' < 0$, which is not feasible). Hence, $a > 1$ and the left-hand side of (A7) is an increasing function. Hence, the equation has the only solution $t = 1$, which means $\alpha' = \alpha$ and implies a unique solution of (6). □

Proof of Proposition 3. Let us rewrite (6) as

$$\begin{cases} -4\alpha \ln(s_0) = \ln(y_1), \\ \alpha = \frac{y_2}{y_1} s_0^2, \end{cases}$$

and therefore,

$$\begin{cases} \alpha\left(\ln(\alpha) + \ln\left(\frac{y_1}{y_2}\right)\right) = -\frac{\ln(y_1)}{2}, \\ s_0 = \sqrt{\frac{y_1}{y_2}\alpha}. \end{cases}$$

Denoting $t = \ln(\alpha)$, the first equation can be rewritten as

$$e^t\left(t + \ln\left(\frac{y_1}{y_2}\right)\right) = -\frac{\ln(y_1)}{2},$$

$$e^{t+\ln\left(\frac{y_1}{y_2}\right)}\left(t + \ln\left(\frac{y_1}{y_2}\right)\right) = e^{\ln\left(\frac{y_1}{y_2}\right)} \ln\left(y_1^{-\frac{1}{2}}\right) = \frac{y_1}{y_2} \ln\left(y_1^{-\frac{1}{2}}\right).$$

Hence, by the definition of the Lambert W function, we obtain

$$t = \text{Lambert}W\left(\frac{y_1}{y_2} \ln\left(y_1^{-\frac{1}{2}}\right)\right) - \ln\left(\frac{y_1}{y_2}\right).$$

Finally, (7) follows from $\alpha = e^t$ and $s_0 = (\frac{y_1}{y_2}\alpha)^{\frac{1}{2}}$. □

Proof of Theorem 1. The first statement of the theorem immediately follows from Proposition 3. To investigate properties of the solutions $(\hat{s}_{0j}, \hat{\alpha}_j)$, one has to study properties of $\text{Lambert}W(\frac{x_1}{x_2})$.

Notice that, by the two-dimensional mean value theorem, for any $f : \mathbb{R}^2 \to \mathbb{R}$ that is from $C^1(\mathbb{R}^2)$, it holds

$$f(\mathbf{x}) - f(\mathbf{y}) = \left(\nabla f\left((1-c)\mathbf{x} + c\mathbf{y}\right), \mathbf{x} - \mathbf{y}\right), \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^2,$$

where $c \in [0,1]$, ∇ denotes the gradient and $(\cdot,\cdot)$ is the scalar product in $\mathbb{R}^2$. Therefore,

$$|f(\mathbf{x}) - f(\mathbf{y})| \le \sup_{c\in[0,1]} ||\nabla f\left((1-c)\mathbf{x} + c\mathbf{y}\right)|| \cdot ||\mathbf{x} - \mathbf{y}||.$$

Now, applying this result to the function $f(\mathbf{x}) = \exp(\frac{1}{2}\text{Lambert}W(\frac{x_1}{x_2}))$, $\mathbf{x} = (\ln(\frac{c_2}{\bar{\delta}_{j\cdot}^{(2,a)}}), 2q_j)$, $\mathbf{y} = (4\alpha\ln(s_0), \frac{2\alpha}{s_0^2})$, and noting that the solution $\hat{s}_{0j}$ is given by (8), we obtain

$$|\hat{s}_{0j} - s_0| = \left|e^{\frac{1}{2}\text{Lambert}W\left(\frac{x_1}{x_2}\right)} - e^{\frac{1}{2}\text{Lambert}W\left(\frac{y_1}{y_2}\right)}\right| \le \sup_{c\in[0,1]} \left|\left|\nabla e^{\frac{1}{2}\text{Lambert}W\left(\frac{(1-c)x_1+cy_1}{(1-c)x_2+cy_2}\right)}\right|\right|$$
$$\times \left|\left|\left(\ln\left(\frac{c_2}{\bar{\delta}_{j\cdot}^{(2,a)}}\right) - 4\alpha\ln(s_0), \frac{2c_2}{c_3}\frac{\Delta\bar{\delta}_{j\cdot}^{(2,a)}}{\bar{\delta}_{j\cdot}^{(2,a)}} - \frac{2\alpha}{s_0^2}\right)\right|\right|. \tag{A8}$$

Noting that $(\text{Lambert}W(x))' = \frac{\text{Lambert}W(x)}{x(1+\text{Lambert}W(x))}$, we obtain

$$\nabla e^{\frac{1}{2}\text{Lambert}W\left(\frac{x_1}{x_2}\right)} = e^{\frac{1}{2}\text{Lambert}W\left(\frac{x_1}{x_2}\right)} \frac{\text{Lambert}W\left(\frac{x_1}{x_2}\right)}{2\frac{x_1}{x_2}(1 + \text{Lambert}W(\frac{x_1}{x_2}))}\left(\frac{1}{x_2}, -\frac{x_1}{x_2^2}\right)$$
$$= e^{\frac{1}{2}\text{Lambert}W\left(\frac{x_1}{x_2}\right)} \frac{\text{Lambert}W\left(\frac{x_1}{x_2}\right)}{1 + \text{Lambert}W\left(\frac{x_1}{x_2}\right)}\left(\frac{1}{2x_1}, -\frac{1}{2x_2}\right).$$

By the properties of the adjusted estimates $\mathbf{x} = (\ln(\frac{c_2}{\bar{\delta}_{j\cdot}^{(2,a)}}), 2q_j) \to (4\alpha\ln(s_0), \frac{2\alpha}{s_0^2})$, when $j \to \infty$. As both $4\alpha\ln(s_0)$ and $\frac{2\alpha}{s_0^2}$ are strictly positive real numbers, then x_1 and x_2 are strictly positive values separated from zero for sufficiently large j. Hence, $\frac{1}{x_1}$ and $\frac{1}{x_2}$ are bounded and the above gradient is uniformly bounded for all sufficiently large values of j.

Now, we study the second multiplier in (A8),

$$\left|\left|\left(\ln\left(\frac{c_2}{\bar{\delta}_{j\cdot}^{(2,a)}}\right) - 4\alpha\ln(s_0), \frac{2c_2}{c_3}\frac{\Delta\bar{\delta}_{j\cdot}^{(2,a)}}{\bar{\delta}_{j\cdot}^{(2,a)}} - \frac{2\alpha}{s_0^2}\right)\right|\right| \le \left|\ln\left(\frac{c_2}{\bar{\delta}_{j\cdot}^{(2,a)}s_0^{4\alpha}}\right)\right| + \left|\frac{2c_2s_0^2\Delta\bar{\delta}_{j\cdot}^{(2,a)} - 2\alpha c_3\bar{\delta}_{j\cdot}^{(2,a)}}{c_3s_0^2\bar{\delta}_{j\cdot}^{(2,a)}}\right|. \tag{A9}$$

As $|\ln(x_1) - \ln(y_1)| \le \frac{|x_1-y_1|}{\min(x_1,y_1)}$ for $x_1, y_1 \in \mathbb{R}_+$, then by Proposition 1 and Remark 11, we can estimate the first summand in (A9) as

$$\left|\ln\left(\frac{c_2s_0^{-4\alpha}}{\bar{\delta}_{j\cdot}^{(2,a)}}\right)\right| \le \frac{\left|\bar{\delta}_{j\cdot}^{(2,a)} - c_2s_0^{-4\alpha}\right|}{\min\left(\bar{\delta}_{j\cdot}^{(2,a)}, c_2s_0^{-4\alpha}\right)} \le c_{12}\max\left(r_j, a_j^{-2}\right), \tag{A10}$$

where c_{12} is an almost surely finite random variable.

The second summand in (A9) can be estimated using Propositions 1 and 2 and Remark 11 as

$$\left|\frac{2c_2s_0^2\Delta\bar{\delta}_{j\cdot}^{(2,a)}-2\alpha c_3\bar{\delta}_{j\cdot}^{(2,a)}}{c_3s_0^2\bar{\delta}_{j\cdot}^{(2,a)}}\right| \leq \left|\frac{2c_2s_0^2\left(\Delta\bar{\delta}_{j\cdot}^{(2,a)}-\alpha s_0^{-4\alpha-2}c_3\right)}{c_3s_0^2\bar{\delta}_{j\cdot}^{(2,a)}}\right| + \left|\frac{2\alpha c_3\left(c_2s_0^{-4\alpha}-\bar{\delta}_{j\cdot}^{(2,a)}\right)}{c_3s_0^2\bar{\delta}_{j\cdot}^{(2,a)}}\right|$$
$$\leq c_{13}\left(\max\left(a_j^2r_j, a_j^{-2}\right)+\max\left(r_j, a_j^{-2}\right)\right) \leq 2c_{13}\max\left(a_j^2r_j, a_j^{-2}\right), \quad \text{(A11)}$$

where c_{13} is an almost surely finite random variable.

Putting together the results (A8)-(A11), we obtain $|\hat{s}_{0j}-s_0| \leq c_{10}\max(a_j^2r_j, a_j^{-2})$.

Finally, noting that

$$|\hat{\alpha}_j-\alpha| = \left|\frac{x_2}{2}\mathrm{e}^{\mathrm{LambertW}\left(\frac{x_1}{x_2}\right)}-\frac{y_2}{2}\mathrm{e}^{\mathrm{LambertW}\left(\frac{y_1}{y_2}\right)}\right|$$
$$\leq \frac{1}{2}\left|\mathrm{e}^{\mathrm{LambertW}\left(\frac{x_1}{x_2}\right)}(x_2-y_2)\right|+\frac{1}{2}\left|y_2\left(\mathrm{e}^{\mathrm{LambertW}\left(\frac{x_1}{x_2}\right)}-\mathrm{e}^{\mathrm{LambertW}\left(\frac{y_1}{y_2}\right)}\right)\right|$$

and using the upper bounds in (A10) and (A11), we obtain $|\hat{\alpha}_j-\alpha| \leq c_{11}\max(a_j^2r_j, a_j^{-2})$, which completes the proof. □