Annals of the Institute of Statistical Mathematics (2021) 73:337–367
https://doi.org/10.1007/s10463-020-00751-6

Nonparametric estimation of the kernel function of symmetric stable moving average random functions

Jürgen Kampf[1] · **Georgiy Shevchenko**[2] · **Evgeny Spodarev**[3]

Received: 13 June 2019 / Revised: 31 December 2019 / Published online: 7 April 2020

Abstract
We estimate the kernel function of a symmetric alpha stable ($S\alpha S$) moving average random function which is observed on a regular grid of points. The proposed estimator relies on the empirical normalized (smoothed) periodogram. It is shown to be weakly consistent for positive definite kernel functions, when the grid mesh size tends to zero and at the same time the observation horizon tends to infinity (high-frequency observations). A simulation study shows that the estimator performs well at finite sample sizes, when the integrator measure of the moving average random function is $S\alpha S$ and for some other infinitely divisible integrators.

Keywords High-frequency observations · Moving average random function · Self-normalized periodogram · Stable random function

1 Inverse problem

We consider the problem of estimation of a kernel $f : \mathbb{R} \to \mathbb{R}$, $f \in L^\alpha(\mathbb{R})$ from observations of the symmetric α-stable (for short, $S\alpha S$) stationary (*moving average*) random function

✉ Jürgen Kampf
kampf-juergen@web.de

Georgiy Shevchenko
zhora@univ.kiev.ua

Evgeny Spodarev
evgeny.spodarev@uni-ulm.de

1 Institute of Mathematics, University of Rostock, Ulmenstraße 69/3, 18057 Rostock, Germany

2 Department of Probability Theory, Statistics and Actuarial Mathematics, Taras Shevchenko National University of Kiev, Volodymyrska 64, Kiev 01601, Ukraine

3 Institute of Stochastics, Ulm University, Helmholtzstr. 18, 89081 Ulm, Germany

$$X(t) = \int_{\mathbb{R}} f(t-s)\Lambda(ds), \quad t \in \mathbb{R}, \tag{1}$$

where Λ is a $S\alpha S$ random measure with independent increments and Lebesgue control measure, see, e.g., Samorodnitsky and Taqqu (1994) for more details on $S\alpha S$ moving averages. While the integrator Λ determines the marginal properties of X, the kernel function f forms its dependence structure. The stability index $\alpha \in (0, 2)$ which controls the heaviness of the tails of X is assumed to be known. If α is unknown, then it has to be additionally estimated and used as a plug-in in what follows. There exist several approaches to estimate α from an iid sample of stable random variables, cf., e.g., Koutrouvelis (1980, 1981), McCulloch (1986), Zolotarev (1986, Section 4.3), Zolotarev and Uchaikin (1999, Chapter 9), Gu and Mao (2002), Fan (2006), Chèn (2011) and references in Koblents et al. (2016). Most of them still work in the setting of weakly dependent sample variables, but the estimation quality diminishes. Using these estimators as a plug-in in our estimation of f would surely increase the variance of this estimate, but a careful investigation of that is out of the scope of the present paper.

The class of stochastic processes (1) includes stable CARMA processes, cf. Brockwell and Lindner (2009), and in particular the stable Ornstein–Uhlenbeck process. These processes are popular in econometry and finance, e.g., they have served as (an essential part of) a model for electricity spot and future prices (García et al. 2011; Müller and Seibert 2019) or for the rates of interbank loans (Janczura et al. 2011); see Brockwell (2014) for an overview. An application of stable random fields of the form (1) with index variable $t \in \mathbb{R}^2$ to spatial modeling of the claims of storm insurance in Austria is given in Karcher et al. (2009) and Karcher (2012).

Our aim is to provide a nonparametric estimator for the function f. We assume that the observations are taken at the points $\{t_{k,n}, k = 1, \dots, n\}$, where $t_{k,n} = k\Delta_n$, $\Delta_n \to 0$, $n \to \infty$, and $n\Delta_n \to \infty, n \to \infty$. So we have high-frequency observations, and the observation horizon expands to the whole $\mathbb{R}_+$. In other words, we try to solve the inverse problem

$$\{X(t_{k,n}),\ k = 1, \dots, n,\ n \in \mathbb{N}\} \mapsto f \in L^\alpha(\mathbb{R}). \tag{2}$$

In Mikosch et al. (1995), this problem was solved for a moving average time series X with innovations belonging to the domain of attraction of the stable law. For X being α-stable, $1 < \alpha < 2$, the parametric estimation of f via a minimum contrast method for the first-order madogram of X is performed in Karcher et al. (2009). A nonparametric estimator of a piecewise constant symmetric f based on the covariation of X was proposed in Karcher and Spodarev (2012). However, this procedure is defined recursively and thus errors made at one step influence all following steps.

Problem (2) for random process (1) with square-integrable random measure Λ and causal f, i.e., $\mathrm{supp} f \subseteq \mathbb{R}_+$, was treated in Brockwell et al. (2013). There, a nonparametric estimator for the kernel function f was proposed and its consistency was shown under CARMA assumptions. The estimator made use of the Wold expansion of the sampled process X.

Here, we extend the ideas of the paper Mikosch et al. (1995) and use the empirical (properly normalized) periodogram of the random function X to estimate the symmetric uniformly continuous kernel function f of positive type satisfying some additional assumptions if the stability index $\alpha \in (0, 2)$ is known. The merit of our approach [as compared to the discrete time series method (Mikosch et al. 1995)] is that it is easily extendable to infinitely divisible moving average random fields (cf. the details in Kampf et al. 2019, Remark 6; Section 5.1). This enables applications to heavy-tailed spatial data, cf., e.g., Karcher et al. (2009) and Karcher (2012).

The paper is organized as follows. In the next section, we discuss conditions on f which would guarantee the existence and uniqueness of solution of the problem (2). After introducing the normed smoothed periodogram and the estimator for f in Sect. 3, the weak consistency of the kernel estimation is stated in Sect. 4. There, Theorems 1 and 2 treat the cases of compact and unbounded support of f, respectively. The consistency of the estimation of the L^2-norm of f is treated in Corollary 1. For the ease of reading, proofs are moved to Appendix 1 (Theorems 1 and 3) and Appendix 2 (auxiliary lemmata). The proof of Theorem 2 is omitted and can be found, together with a detailed proof of Theorem 1, in the arXiv version of the present paper Kampf et al. (2019). A simulation study shows the good performance of estimation in Sect. 5. There, the scope of applicability of this estimation method is studied empirically. The estimator performs well also for skewed stable, symmetric infinitely divisible and for Gaussian Λ, whereas it fails to work with some skewed non-stable Λ. We conclude with a summary and conjectures (Sect. 6).

2 Existence and uniqueness of the solution

As most of the inverse problems, the problem (2) of restoring f from observations of X is in general ill posed. Here, we discuss the additional conditions to be imposed onto f to make (2) have a unique solution. These conditions cannot be inferred from real data and have to be postulated as some prior knowledge on the nature of the process under consideration. For instance, if $X(t)$ models temperature, precipitation, air pressure or some other weather feature at a spot t, then the condition (F1) below seems plausible. At the same time, this condition is inacceptable for causal processes, e.g., in finance since it implies the symmetry of the dependence structure of $X(t)$ on observations $X(s)$ in a vicinity $\{s \in \mathbb{R} : |s-t| < \delta\}$ of t making $X(t)$ dependent on the future $X(s)$, $s > t$.

Notice that the spectral representation (1) of X for $0 < \alpha \le 2$ is not unique. However, it is shown in Rosiński (1994, Example 3.2) for $0 < \alpha < 2$ that two functions $f_1, f_2 \in L^\alpha(\mathbb{R})$ fulfilling (1) are connected by $f_2(t) = \pm f_1(t + t_0)$ for almost all $t \in \mathbb{R}$ and for some fixed $t_0 \in \mathbb{R}$. Let $\hat{f}$ be the Fourier transform of f, and let $\hat{f}^{-1}$ be its inverse, whenever these exist. We additionally assume that

(F1) f is *positive semidefinite*.

It follows from Trigub and Bellinsky (2004, 6.2.1) that f is *even* (or *symmetric*), i.e., $f(t) = f(-t)$ for all $t \in \mathbb{R}$. Under the condition (F1), it can be easily shown that $f_1 = f_2$ almost everywhere (for short, a.e.) on $\mathbb{R}$, i.e., f is determined uniquely a.e. on $\mathbb{R}$. In the Gaussian case $\alpha = 2$, the existence of the so-called *canonical kernel* can be shown for a centered purely nondeterministic mean square continuous X, see Hida and Hitsuda (1993, Theorem 3.4). The uniqueness of f cannot be guaranteed. However, under the additional assumption $f \in C(\mathbb{R}) \cap L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ it is unique which can be shown directly by the following covariance-based approach.

Let X in (1) be an infinitely divisible moving average random function with finite second moments, i.e., Λ be an infinitely divisible independently scattered random measure with Lebesgue control measure, $\mathsf{E}\left[\Lambda^2(B)\right] < \infty$ for any bounded Borel set $B \subset \mathbb{R}$, and $f \in C(\mathbb{R}) \cap L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Then, the covariance function of X is given by

$$C(t) = \operatorname{Cov}(X(0), X(t)) = \int_{\mathbb{R}} f(t-s) f(-s)\, \mathrm{d}s, \quad t \in \mathbb{R}.$$

Applying the Fourier transform, we get $\hat{C} = \hat{f}^2$, and hence, the relation

$$f = \widehat{\sqrt{\hat{C}}}^{-1} \quad \text{a.e. on } \mathbb{R} \tag{3}$$

proves the uniqueness of f in the Gaussian case. Assumption (F1) is needed in order to reconstruct f uniquely from the absolute value of its Fourier transform. Indeed, under the condition $f \in C(\mathbb{R}) \cap L^1(\mathbb{R})$ it can easily be shown by the Bochner–Khintchine theorem, see, e.g., Trigub and Bellinsky (2004, 6.2.3) or Akhiezer (1988, p. 54), that (F1) is equivalent to $\hat{f}(\lambda) \geq 0$ for all $\lambda \in \mathbb{R}$, i.e., f being of *positive type*. In turn, to show that f being of positive type implies (F1) one also has to use the inversion formula for Fourier transforms which holds a.e. on $\mathbb{R}$ by Akhiezer (1988, p. 17–18, Corollary 2 and Theorem 2) or by Trigub and Bellinsky (2004, 3.1.10 and 3.1.15). Relation (3) can be used to build a strongly consistent estimator of a symmetric piecewise constant compact supported f if smoothed spectral density estimates are used (cf., e.g., Karcher 2012, § 3.3).

It is worth mentioning that under low-frequency observations, it is in general not possible to identify f in a unique way. Indeed, let $\Delta_n = \Delta$ be constant. Define for any $h \in L^\alpha[-\Delta/2, \Delta/2]$ with $\|h\|_\alpha = 1$ the process

$$X_h(t) = \int_{-\Delta/2}^{\Delta/2} h(t-s) \Lambda(\mathrm{d}s).$$

Then, the observations $\left\{X_h(t_{k,n}), k = 1, \dots, n\right\}$ are iid SαS with scale parameter 1, so their distribution does not depend on h.

The reason why the observation interval should expand infinitely is less obvious. In the Gaussian case, on any finite interval $[0, t]$ it is possible to construct stationary processes such that the corresponding probability measures on $C[0, t]$ are different but the processes have the same distribution. Therefore, one is not able to identify the kernel function (not even the distribution) from observations of the process on a finite interval. However, to the best of our knowledge, there are no such results in the stable case.

3 Estimators

We use the following notation: $a_n = o_P(b_n)$, $n \to \infty$, means $a_n/b_n \xrightarrow{P} 0$, $n \to \infty$; $a_n \overset{P}{\sim} b_n$, $n \to \infty$, means $a_n/b_n \xrightarrow{P} 1$, $n \to \infty$; we write $a_n = O_P(b_n)$, $n \to \infty$, if the sequence $\{a_n/b_n, n \geq 1\}$ is bounded in probability. The symbol C will denote a generic constant, the value of which is not important.

To estimate the function f in (1), we use the *self-normalized (empirical) periodogram* of X, defined as

$$I_{n,X}(\lambda) = \frac{\left|\sum_{j=1}^{n} X(t_{j,n})e^{it_{j,n}\lambda}\right|^2}{\sum_{j=1}^{n} X(t_{j,n})^2}. \tag{4}$$

It is known (Fasen and Fuchs 2013a, Theorem 2.11) that $\Delta_n \cdot I_{n,X}(\lambda)$ converges to a random limit as $n \to \infty$, and so it cannot be a consistent estimator of any deterministic quantity of interest. Thus, following Fasen and Fuchs (2013b) we define its smoothed version. Let $\{m_n, n \geq 1\}$ be a sequence of positive integers such that $m_n \to \infty$ and $m_n = o(n)$, $n \to \infty$. Consider a sequence of filters $\{W_n(m), |m| \leq m_n, n \geq 1\}$ satisfying

(W1) $W_n(m) \geq 0$;
(W2) $\sum_{|m|\leq m_n} W_n(m) = 1$;
(W3) $\max_{|m|\leq m_n} W_n(m) \to 0, n \to \infty$;
(W4) $\sum_{|m|\leq m_n} m^2 W_n(m) = o\left((n\Delta_n)^2\right), n \to \infty$.

In the following, we will denote $W_n^* = \max_{|m|\leq m_n} W_n(m)$, $W_n^{(2)} = \sum_{|m|\leq m_n} m^2 W_n(m)$.

Denote $\nu_n(m,\lambda) = \lambda + m/(n\Delta_n)$, $m = -m_n, \dots, m_n$. Then, a smoothed periodogram is defined as

$$I_{n,X}^s(\lambda) = \sum_{|m|\leq m_n} W_n(m) I_{n,X}(\nu_n(m,\lambda)). \tag{5}$$

Remark 1 The periodogram defined in Fasen and Fuchs (2013b) has the argument $\omega = \lambda/\Delta_n$. This explains why the quantity $\omega + m/n$ rather than $\lambda + m/(n\Delta_n)$ is used in their definition of smoothed periodogram. And as long as the main result of Fasen and Fuchs (2013b) concerns the limit behavior of smoothed periodogram evaluated at $\omega\Delta_n$, it is comparable with our findings.

For the sake of brevity, define the normalized function $g(t) = f(t)/\|f\|_2$, where $\|f\|_2 = \sqrt{\int_{\mathbb{R}} f(x)^2\, dx}$ is the L^2-norm of f whenever it is finite; the Fourier transform of g is

$$\hat{g}(\lambda) = \int_{\mathbb{R}} g(t)\mathrm{e}^{-i\lambda t}\mathrm{d}t, \quad \lambda \in \mathbb{R},$$

whenever it exists. First, we estimate g and $\|f\|_2$ separately. If $\tilde{g}$ and $\widetilde{\|f\|}_2$ are their weakly consistent estimators, then $\tilde{f} = \widetilde{\|f\|}_2 \cdot \tilde{g}$ is a weakly consistent estimator of f.

Considering the fact that $\sqrt{\Delta_n I^s_{n,X}(\lambda)}$ is an estimator for $\hat{g}(\lambda)$ (see, e.g., Theorem 1), it is natural to estimate $g(t)$ by $\frac{1}{2\pi}\int_{\mathbb{R}} \sqrt{\Delta_n I^s_{n,X}(\lambda)}\mathrm{e}^{it\lambda}\,\mathrm{d}\lambda$. However, $\sqrt{\Delta_n I^s_{n,X}(\lambda)} \notin L^1(\mathbb{R})$ a.s. since $I^s_{n,X}$ is $2\pi\Delta_n^{-1}$-periodic. Thus, we put

$$\tilde{g}(t) = \frac{1}{2\pi}\int_{[-a_n,a_n]} \sqrt{\Delta_n I^s_{n,X}(\lambda)}\mathrm{e}^{it\lambda}\,\mathrm{d}\lambda, \quad t \in \mathbb{R}, \tag{6}$$

where $\{a_n, n \geq 1\}$ is a deterministic sequence with the following properties:

(A1) $a_n \to \infty, n \to \infty$;
(A2) $a_n^2 W_n^* \to 0, n \to \infty$;
(A3) $a_n^{3/4} = o((n\Delta_n)^{1/\alpha}), n \to \infty$;
(A4) $a_n^2 \Delta_n \to 0, n \to \infty$;
(A5) $a_n^2 W_n^{(2)} = o\big((n\Delta_n)^2\big), n \to \infty$.

Remark 2 From (W1), (W2) and (W3), it is clear that $\limsup_{n\to\infty} W_n^{(2)} > 0$. Therefore, (A5) implies that $a_n = o(n\Delta_n)$, $n \to \infty$ (this will be used in the future). In particular, (A3) follows from (A5) for $\alpha \leq \frac{4}{3}$. Besides this, the assumptions are rather independent.

Let f satisfy (F1). Further assumptions depend on whether f is compactly supported or not. In the case of compact support, we assume

(F2) $a_n \omega_f(\Delta_n) \to 0, n \to \infty$,

where $\omega_f(\Delta_n) = \sup_{|t-s|<\Delta_n} |f(t) - f(s)|$ is the modulus of continuity of f. Clearly, assumption (F2) implies the uniform continuity of f. Hence, f is bounded, and then $f \in L^p(\mathbb{R})$ for all $p \in (0, \infty]$. In the case of non-compact support, we assume (additionally to (F1)) that for some $a > \max\{2, 1/\alpha\}$

(F2′) $a_n \omega_f(\Delta_n)^{1-1/a} \to 0, n \to \infty$;
(F3′) $f(t) = O(|t|^{-a}), \ |t| \to \infty$;
(F4′) $a_n^{3/4} = o\big(\omega_f(\Delta_n)^{1/(a\alpha)}(n\Delta_n)^{1/\alpha}\big), n \to \infty$.

It follows from (F2′) and (F3′) that f is uniformly continuous and bounded, $f \in L^p(\mathbb{R})$ for $p \in (\frac{1}{a}, \infty]$, e.g, $p = \alpha, 1, 2$. Hence $\hat{f}$ is bounded, too, and moreover, it is square integrable.

Remark 3 The assumptions (F2′)–(F4′) relate the size a_n of "integration window" of the smoothed periodogram used in the estimator $\widetilde{g}$ with the regularity and the rate of decay of *f*. But this does not mean that the latter characteristics should be available a priori: usually the kernel *f* can be assumed to be at least Hölder continuous, so we can choose $a_n = \log n$ or any other slowly varying function at infinity with $a_n \to +\infty$ as $n \to +\infty$. Section 5 further clarifies this by giving explicit examples of kernels and corresponding sequences $\{a_n\}$, $\{W_n\}$, $\{\Delta_n\}$, $\{m_n\}$ satisfying the above assumptions. A problem of the optimal choice of these sequences can be solved numerically by minimizing the upper bounds (8)–(9) of the mean-squared error of $\tilde{g}$. However, these bounds require the a priori knowledge of the modulus of continuity of *f* and the decay rate of the Fourier transform $\hat{g}(\lambda)$ as $|\lambda| \to \infty$ which makes this optimization not realistic in practical data inference.

4 Main results

Here, we state our main results about the weak consistency of the estimates of *g* and *f*.

Theorem 1 *Let f be compactly supported and* (F2), (A1)–(A5), (W1)–(W4) *be satisfied.*

(i) *The following convergence in probability holds:*

$$a_n \cdot \int_{-a_n}^{a_n} (\Delta_n I^s_{n,X}(\lambda) - |\hat{g}(\lambda)|^2)^2 \mathrm{d}\lambda \xrightarrow{P} 0, \quad n \to \infty. \tag{7}$$

(ii) *If additionally* (F1) *is true, then* $\|\tilde{g} - g\|_2 \xrightarrow{P} 0, \quad n \to \infty.$

Remark 4 Using Brockwell and Lindner (2009, Lemma 2.3), it can be shown that in CARMA(p, q) models $|\hat{g}(\lambda)|^2$ coincides with the power transfer function if $p > q + 1$, where p is the order of the autoregressive polynomial and q is the order of the moving average polynomial of the CARMA(p, q) model. Thus, Theorem 1(i) shows that $\Delta_n I^s_{n,X}(\lambda)$ is a weakly consistent estimator for the power transfer function. However, this is already known (Fasen and Fuchs 2013b, Theorem 1) under the weaker assumption $p > q$.

Theorem 2 *The assertion of Theorem 1 holds true also under the assumptions* (F1), (F2′)–(F4′), (A1)–(A5), (W1)–(W4).

Remark 5 Carefully examining the proofs of Theorems 1 and 2, we can bound the rate of convergence in (7) by

$$O_P\big(a_n(n\Delta_n)^{-1} + a_n^2\omega_f(\Delta_n)^2 + a_n^2 W_n^* + a_n^2 W_n^{(2)}(n\Delta_n)^{-2} + a_n^{3/2}(n\Delta_n)^{-2/\alpha} + a_n^4\Delta_n^2\big), \quad n \to \infty$$

in the case of bounded support and by

$$O_P\big(a_n(n\Delta_n)^{-1} + a_n^2\omega_f(\Delta_n)^{2-2/a} + a_n^2 W_n^* + a_n^2 W_n^{(2)}(n\Delta_n)^{-2} + a_n^{3/2}(n\Delta_n)^{-2/\alpha}\omega_f(\Delta_n)^{-2/(a\alpha)} + a_n^4\Delta_n^2\big),$$

in the case of unbounded support.

Consequently (from the first part of the proof),

$$\|\tilde{g}-g\|_2^2 = O_P\Big(a_n^{1/2}(n\Delta_n)^{-1/2} + a_n\omega_f(\Delta_n) + a_n(W_n^*)^{1/2} + a_n(W_n^{(2)})^{1/2}(n\Delta_n)^{-1} + a_n^{3/4}(n\Delta_n)^{-1/\alpha} + a_n^2\Delta_n\Big) + O\left(\int_{\{\lambda:|\lambda|>a_n\}}\hat{g}(\lambda)^2\mathrm{d}\lambda\right),\quad n\to\infty, \tag{8}$$

in the case of bounded support and

$$\|\tilde{g}-g\|_2^2 = O_P\big(a_n^{1/2}(n\Delta_n)^{-1/2} + a_n\omega_f(\Delta_n)^{1-1/a} + a_n(W_n^*)^{1/2} + a_n(W_n^{(2)})^{1/2}(n\Delta_n)^{-1} + a_n^{3/4}(n\Delta_n)^{-1/\alpha}\omega_f(\Delta_n)^{-1/(a\alpha)} + a_n^2\Delta_n\big) + O\left(\int_{\{\lambda:|\lambda|>a_n\}}\hat{g}(\lambda)^2\mathrm{d}\lambda\right),\quad n\to\infty, \tag{9}$$

in the case of unbounded support.

Taking into account the evident relation $\|f\|_2 = \|f\|_\alpha/\|g\|_\alpha$, the estimation of the norm $\|f\|_2$ is reduced to the estimation of $\|f\|_\alpha = \sigma_{X(0)}$, the scale parameter of $X(0)$ (see Samorodnitsky and Taqqu 1994, Property 3.2.2), and $\|g\|_\alpha$. In the literature, there are a number of estimators of scale available, see, e.g., Zolotarev (1986, Chapter 4), (Zolotarev and Uchaikin 1999, Chapter 9), Fan (2006) and Koblents et al. (2016). Among those, we choose the quantile estimator for the sake of its robustness. It is based on the fact that the quantiles of $X(0)$ are equal to those of $S_\alpha(1,0,0)$, multiplied by $\sigma_{X(0)}$. Taking different quantile levels, this can be used to construct a variety of estimators. The most popular choice is quartiles, so that the correspondent estimator is

$$\widetilde{\sigma}_q = \frac{\widetilde{x}_{3/4;n} - \widetilde{x}_{1/4;n}}{x_{3/4} - x_{1/4}}, \tag{10}$$

where $x_{1/4}$ and $x_{3/4}$ are, respectively, the lower and upper quartiles of $S_\alpha(1,0,0)$ and $\widetilde{x}_{1/4;n}$ and $\widetilde{x}_{3/4;n}$ are, respectively, the lower and upper empirical quartiles of the sample $\{X(t_{k,n}), k=1,\dots,n\}$.

It is well known that estimator (10) is a.s. consistent for iid observations, mixing sequences and some linear ergodic processes with or without heavy tails. The proof involves the Bahadur–Kiefer-type representation for the empirical quantiles of $X(0)$, cf. Hesse (1990), Wu (2005), Kulik (2007), Wang et al. (2016) and references therein. For instance, if X is ergodic (cf. Cambanis et al. 1995 for sufficient conditions), its kernel function f is simple (i.e., piecewise constant) and either compactly supported or satisfying condition (F3′) then the a.s. consistency of (10) follows from Hesse (1990, Theorem 1). We believe that it does so also for ergodic X with general kernels f satisfying (F3′) and some additional assumptions, but checking this carefully would blow up the size of this paper. Anyway, the results of our paper are applicable to *any* weakly consistent estimator of scale $\widetilde{\sigma}_{X(0)}$, whatever it is.

Now let us turn to the estimation of $\|g\|_\alpha$. In the case where f is supported by $[-T, T]$ (and T is known a priori), one can use the estimator

$$\widetilde{\|g\|}_{\alpha,T} = \left(\int_{-T}^{T} |\tilde{g}(t)|^\alpha \mathrm{d}t \right)^{1/\alpha}.$$

If the size of support of f is unknown or it has unbounded support, we need a deterministic sequence $\{b_n, n \geq 1\}$ such that

(B1) $b_n \to \infty, n \to \infty$;
(B2) $b_n^{2/\alpha-1} a_n^{1/2} = o\big((n\Delta_n)^{1/2}\big), n \to \infty$;
(B3) $b_n^{2/\alpha-1} a_n W_n^* \to 0, n \to \infty$;
(B4) $b_n^{2/\alpha-1} a_n (W_n^{(2)})^{1/2} = o\big(n\Delta_n\big), n \to \infty$;
(B5) $b_n^{2/\alpha-1} a_n^2 \Delta_n \to 0, n \to \infty$;
(B6) $b_n^{2/\alpha-1} a_n \omega_f(\Delta_n)^q \to 0, n \to \infty$, where $q = 1$ in the compact support case and $q = 1 - 1/a$ in the case of unbounded support;
(B7) $b_n^{2/\alpha-1} a_n^{3/4} = o((n\Delta_n)^{1/\alpha} \omega_f(\Delta_n)^r), n \to \infty$, where $r = 0$ in the compact support case and $r = 1/(a\alpha)$ in the case of unbounded support;
(B8) $b_n^{2/\alpha-1} \int_{\{\lambda : |\lambda| > a_n\}} \hat{g}(\lambda)^2 \mathrm{d}\lambda \to 0, n \to \infty$.

An example of such sequence $\{b_n\}$ is given in Sect. 5. With this in hand, an estimator for $\|g\|_\alpha$ is constructed as

$$\widetilde{\|g\|}_{\alpha,b_n} = \left(\int_{-b_n}^{b_n} |\tilde{g}(t)|^\alpha \mathrm{d}t \right)^{1/\alpha}. \tag{11}$$

Theorem 3

(i) *Let f be supported by* $[-T, T]$ *and the assumptions of Theorem* 1 (ii) *hold. Then*

$$\widetilde{\|g\|}_{\alpha,T} \xrightarrow{P} \|g\|_\alpha, \quad n \to \infty.$$

(ii) *Under the assumptions of Theorem* 1 (ii) *and* (B1)–(B8),

$$\widetilde{\|g\|}_{\alpha,b_n} \xrightarrow{P} \|g\|_\alpha, \quad n \to \infty.$$

(iii) *Under the assumptions of Theorem* 2 *and* (B1)–(B8),

$$\widetilde{\|g\|}_{\alpha,b_n} \xrightarrow{P} \|g\|_\alpha, \quad n \to \infty.$$

Introduce a plug-in estimator $\widetilde{\|f\|}_2 = \widetilde{\sigma}_{X(0)}/\widetilde{\|g\|}_\alpha$ of $\|f\|_2$ where $\widetilde{\sigma}_{X(0)}$ is a scale estimator of $X(0)$ (e.g., $\widetilde{\sigma}_q$) and $\widetilde{\|g\|}_\alpha$ is any of the estimators $\widetilde{\|g\|}_{\alpha,T}$ and $\widetilde{\|g\|}_{\alpha,b_n}$ corresponding to the case of compact or non-compact support of f. Moreover, estimate f by $\tilde{f} := \tilde{g}/\widetilde{\|f\|}_2$.

Corollary 1 *Let $\widetilde{\sigma}_{X(0)}$ be any weakly consistent estimator of scale of $X(0)$. Under the assumptions of Theorems 1 and 3 for compact-supported f (or Theorems 2, 3, otherwise), it holds*

$$\widetilde{\|f\|}_2 \xrightarrow{P} \|f\|_2, \quad n \to \infty,$$

and

$$\|\tilde{f} - f\|_2 \xrightarrow{P} 0, \quad n \to \infty.$$

Remark 6 The above results stay true also for the case of estimation of the kernel function $f : \mathbb{R}^d \to \mathbb{R}$ of a stationary random field $X(t) = \int_{\mathbb{R}^d} f(t-s)\Lambda(ds)$, $t \in \mathbb{R}^d$, where Λ is a homogeneous $S\alpha S$ independently scattered random measure on $\mathbb{R}^d$. See details and numerical experiments in Kampf et al. (2019).

5 Simulation study

In this section, we study the performance and the applicability range of the above estimation method empirically, i.e., by estimating f from Monte Carlo simulations of the trajectories of X. Before that, dwell on the particular choice of the weights W_n and sequences $\{\Delta_n\}$, $\{m_n\}$, $\{a_n\}$ and $\{b_n\}$.

Assumptions (W1)–(W4), (A1)–(A5) and (B1)–(B5) are evidently satisfied, e.g., for

- uniform weights $W_n(m) = \frac{1}{2m_n+1}$,
- $\Delta_n = n^{-\delta}$, $\delta \in (0,1)$,
- $m_n = n^\gamma$, $\gamma \in ((1-\delta)/2, 1-\delta)$,
- $a_n = \log n$,
- $b_n = \left(n^{(1-\delta)/2-\beta} \log^{-1} n\right)^{\alpha/(2-\alpha)}$, $\beta \in \left[\gamma - (1-\delta)/2, (1-\delta)/2\right)$.

Assumptions (F1)–(F2) hold for all positive semidefinite compact supported Lipschitz continuous kernels f. For all Lipschitz continuous functions (F2′) holds. Assumption (F3′) is valid whenever f decays at infinity rapidly enough, e.g., for $f(t) = \mathrm{e}^{-|t|}$, while (F4′) holds for all non-constant functions f provided $\delta < \frac{a}{a+1}$, since then $\omega_f(\Delta) \geq c \cdot \Delta$ for an appropriate constant $c > 0$ and sufficiently small $\Delta > 0$. Assumptions (B6)–(B7) are satisfied when f is Lipschitz continuous and

$$\beta > \min\left\{\frac{1-\delta}{2} - \frac{1-\delta}{\alpha} + \frac{\delta}{\alpha a}, \frac{1-\delta}{2} - \delta q\right\}$$

where $q = 1$ for compactly supported f and $q = 1 - 1/a$, otherwise. This condition can be fulfilled iff $\delta < a/(a+1)$. Assumption (B8) holds true if $\hat{g}(\lambda) = O(\mathrm{e}^{-c\lambda})$, $|\lambda| \to +\infty$, with $c > ((1-\delta)/2 - \beta)/2$.

Now let us study the behavior of our estimator at finite sample size. To simulate the realizations of X, we used the algorithms given in Karcher et al. (2013). For sample size n, we considered a lattice spacing of $\Delta_n = 1/\sqrt{n}$, so that the observation points range from $-\sqrt{n}/2$ to $\sqrt{n}/2 - \Delta_n$. We simulated fields for $\alpha = 0.3, 0.7$ and 1.7 and as kernels we used the triangular, the spherical and the exponential kernels

$$f(t) = \sqrt{3/2}(1 - |t|)\mathbf{1}_{[-1,1]}(t), \tag{12}$$

$$f(t) = \sqrt{70/33}(1 - 1.5|t| + 0.5|t|^3)\mathbf{1}_{[-1,1]}(t), \tag{13}$$

$$f(t) = \exp(-|t|) \tag{14}$$

chosen such that $\|f\|_2 = 1$. These kernels f satisfy conditions (F1)–(F2) and (F1), (F2′)–(F4′), respectively. Indeed, assumption (F1) holds since all these functions are valid covariance functions which are positive semidefinite. One can check that their Fourier transforms are nonnegative also directly, compare (Lantuéjoul 2002, Table 4, p. 245). (F2) and (F2′) follow from Lipschitz continuity of the functions (12)–(14).

As parameters for the estimator we chose $m_n = \lfloor n^{1/4} \rfloor$, uniform weights $W_n(m) = 1/(2m_n + 1)$ and $a_n = \log n$.

5.1 SαS case, $0 < \alpha < 2$

In what follows, we apply our estimation method to SαS moving averages. The results are shown in Fig. 1. Each plot contains the graph of the real kernel function f used to simulate X on $n = 1000$ data points, the mean of 10,000 estimates of f and their (0.025, 0.975)-quantile envelope, i.e., the region containing 95% of all estimated curves of f. We see that the results are quite good. Since an estimator converging to a positive value can, of course, be negative at finite sample size, there is no problem in the fact that the lower envelope attains negative values.

In order to evaluate the tail behavior of our estimators, we report different quantiles of $\tilde{f}(0)$, $\tilde{f}(0.4)$, $\tilde{f}(1)$, $\|\tilde{f}\|_2$ and $\|\tilde{f} - f\|_2$ applied to a $S\alpha S$ moving average random field with $\alpha = 1.7$ and spherical kernel in Table 1. The L_2-norms were approximated by $\|\tilde{f}\|_2 \approx \sqrt{\frac{1}{25}\sum_{k=0}^{1000} \tilde{f}^2(x_k)}$ resp. $\|\tilde{f} - f\|_2 \approx \sqrt{\frac{1}{25}\sum_{k=0}^{1000} (\tilde{f}(x_k) - f(x_k))^2}$, where $x_k := -20 + \frac{k}{25}$, $k = 0, \ldots, 1000$, and the quantiles were estimated by the empirical quantiles of the 10000 simulation runs mentioned above. We see only small differences between the quantiles of different orders, indicating strongly that the distributions are light-tailed if not

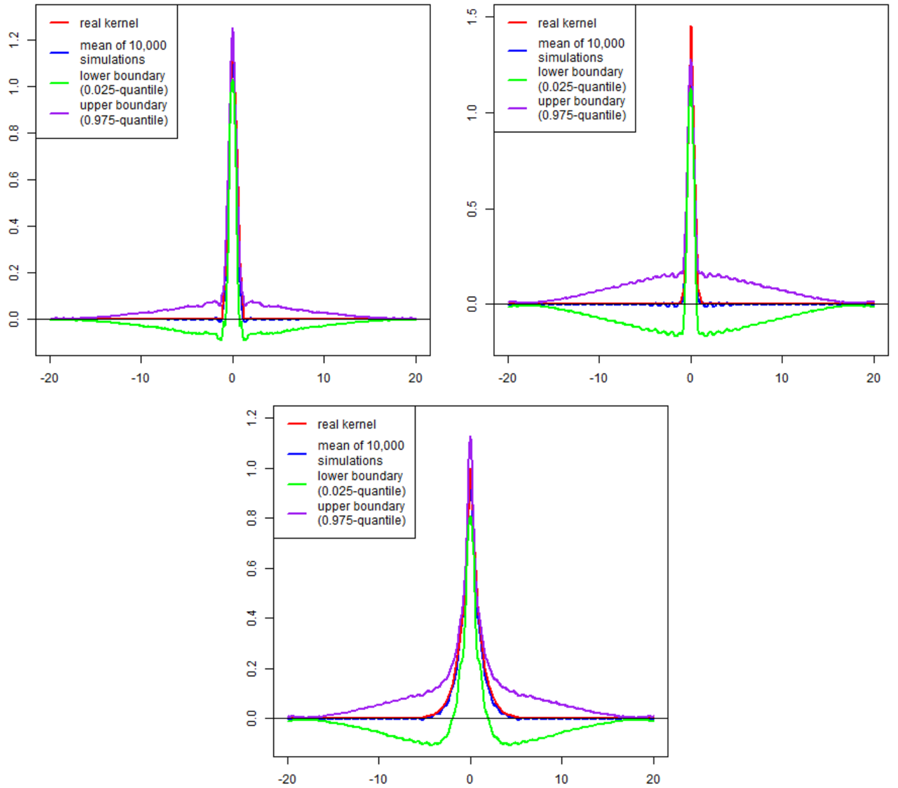

Fig. 1 Estimation results for $S\alpha S$ X with triangular kernel (12), $\alpha = 0.3$ (top left), with spherical kernel (13), $\alpha = 1.7$ (top right) and with exponential kernel (14), $\alpha = 0.7$ (bottom row)

Table 1 Quantiles of the estimator $\tilde{f}$ applied to a $S\alpha S$ field X with triangular kernel and $\alpha = 1.7$

	50%	70%	90%	95%	97%	99%	99.5%
$\tilde{f}(0)$	1.2139	1.2338	1.2592	1.2708	1.2771	1.2952	1.3032
$\tilde{f}(0.4)$	0.6374	0.6532	0.6720	0.6802	0.6855	0.6954	0.7003
$\tilde{f}(1)$	0.0512	0.0802	0.1314	0.1590	0.1745	0.2098	0.2287
$\|\tilde{f}\|_2$	0.9963	0.9969	0.9974	0.9977	0.9978	0.998	0.9982
$\|\tilde{f} - f\|^2$	0.2908	0.3235	0.3745	0.4017	0.4184	0.4505	0.4646

even deterministically bounded. A survey over the quantiles of different orders of $\|\tilde{f} - f\|_2$ for the different settings considered in this section is given in Table 2.

Moreover, we investigated how the estimation improves as more observation points are used. We tried $n = 2000, 4000, 10{,}000, 20{,}000, 40{,}000$, but did only 100 simulations for these larger data sets due to the longer computation time. The results for the $S\alpha S$ moving average field with $\alpha = 1.7$ and spherical kernel are plotted in Fig. 2. We present the mean-squared error of $\tilde{f}(x)$ for $x \in \{0, 0.4, 1\}$ and the L_2-distance $\|\tilde{f} - f\|_2$

Table 2 Survey over the quantiles of $\|\tilde{f}-f\|_2$ in the twelve different settings considered in this section

Setting	50%	70%	90%	95%	97%	99%	99.5%
Figure 1, left	0.0669	0.1347	0.2821	0.3459	0.3874	0.5005	0.5717
Figure 1, right	0.2908	0.3235	0.3745	0.4017	0.4184	0.4505	0.4646
Figure 1, bottom	0.1548	0.2228	0.3248	0.3644	0.3977	0.4719	0.5116
Figure 3, left	0.3353	0.3871	0.5092	0.5797	0.6289	0.7409	0.8149
Figure 3, right	0.6372	0.8169	0.9769	1.0570	1.3848	2.1029	2.6195
Figure 4	0.2728	0.3272	0.4232	0.4820	0.5234	0.5893	0.6309
Figure 5, left	0.2647	0.3127	0.3877	0.4308	0.4597	0.5215	0.5518
Figure 5, right	0.2982	0.3331	0.3861	0.4127	0.4312	0.4684	0.4916
Figure 6, left	0.2681	0.3149	0.3972	0.4422	0.4702	0.5351	0.5651
Figure 6, right	0.2689	0.3202	0.4024	0.4437	0.4722	0.5342	0.5701
Figure 7, left	0.9115	0.9189	0.9294	0.9342	0.9376	0.9440	0.9467
Figure 7, right	0.9220	0.9279	0.9358	0.9397	0.9421	0.9468	0.9493

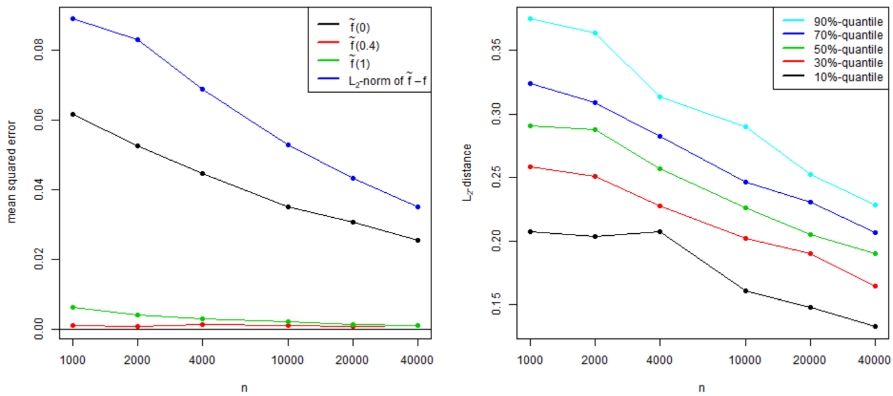

Fig. 2 Convergence of $\tilde{f}$ toward f as the sample size increases. On the left, we plotted the mean-squared error in three different points as well as the expected value of the squared L_2-distance. On the right, we plotted various quantiles of the L_2-distance between $\tilde{f}$ and f

as well as the quantiles of $\|\tilde{f}-f\|_2$. We see that the errors are decreasing with few exceptions which are probably due to simulation errors. For example, we repeated the simulation of the mean-squared errors of $\tilde{f}(0.4)$ for $n = 2000$ and for $n = 4000$ ten times, in order the judge whether this increase is a simulation error. We got five times a higher mean-squared error for $n = 2000$ and five times a higher mean-squared error for $n = 4000$—so there is no incidence for a real increase. In Table 3 we give a survey over the expected values of $\|\tilde{f}-f\|_2^2$ for different sample sizes n. We see that for those ten settings for which the results are already good at sample size $n = 1000$ the mean-squared error decreases with increasing n, while for the last two settings with poor results at sample size $n = 1000$ things get even worse when n is increasing.

In Fig. 1 we concentrated on the estimation of function $g = f$ (which is equivalent to setting $\|f\|_2 = 1$). If the norm of f is unknown, then it has to be estimated

Table 3 Survey over the expected values of $\|\tilde{f} - f\|_2^2$ for different sample sizes n in the twelve different settings considered in this section

Setting	1000	2000	4000	10000	20000	40000
Figure 1, left	0.1247	0.1000	0.1126	0.0821	0.0643	0.0474
Figure 1, right	0.2900	0.2803	0.2561	0.2237	0.2034	0.1831
Figure 1, bottom	0.1707	0.1624	0.1431	0.1123	0.1106	0.0897
Figure 3, left	0.3566	0.3272	0.2842	0.2271	0.2333	0.1931
Figure 3, right	0.6599	0.6456	0.5367	0.4863	0.4614	0.4730
Figure 4	0.2802	0.2553	0.2408	0.1898	0.1952	0.1643
Figure 5, left	0.2648	0.2545	0.2219	0.1893	0.1733	0.1441
Figure 5, right	0.3024	0.2894	0.2703	0.2310	0.2204	0.1903
Figure 6, left	0.2798	0.2664	0.2461	0.2252	0.2039	0.1840
Figure 6, right	0.2816	0.2771	0.2590	0.2265	0.2161	0.1758
Figure 7, left	0.9109	0.9647	1.0144	1.0597	1.0862	1.0915
Figure 7, right	0.9217	0.9731	1.0234	1.0667	1.0925	1.0976

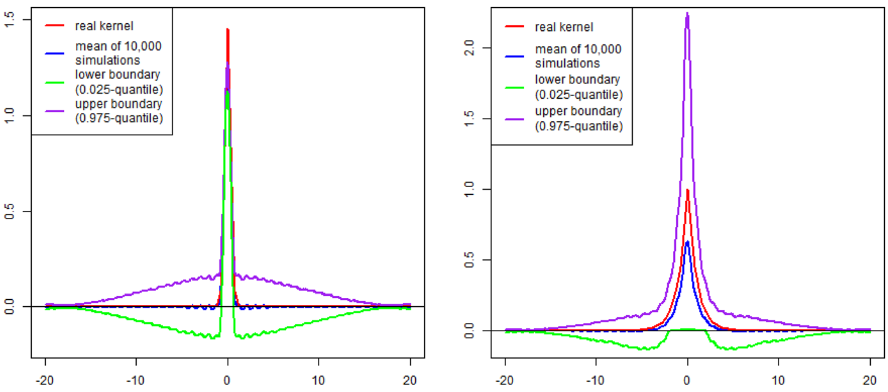

Fig. 3 Estimation results for $S\alpha S$ X with unknown norm of f. Here $\alpha = 1.7$, f is a spherical kernel (13) (left) and $\alpha = 0.7$, f is an exponential kernel (14) (right)

separately, e.g., via relation (10). The same curves as in Fig. 1 are shown for the estimates of f in Fig. 3 for $\alpha = 0.7$ and $\alpha = 1.7$. Not surprisingly, the empirical standard deviation is much higher than for known norm and the performance of the estimators of the norm $\widetilde{\|f\|_2}$ gets better with increasing α.

Numerical experiments with different sampling mesh values Δ_n show that the estimation of f performs well for $\Delta_n \in (0, 0.1]$ (high-frequency framework).

5.2 Beyond the SαS case: Gaussianity, skewness and general infinite divisibility

As shown above, our estimator works well in all cases in which its consistency was proven in Sect. 4. An interesting question is whether it also performs well beyond these cases. Indeed, it does work well for Gaussian ($\alpha = 2$, cf. Fig. 4) and skewed

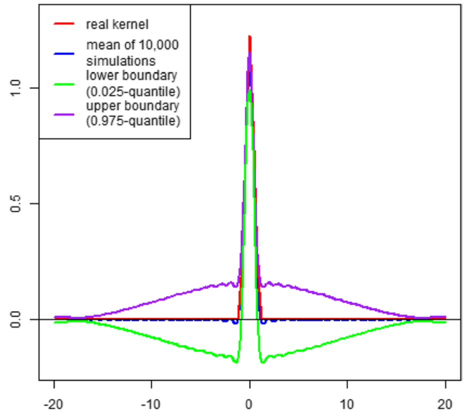

Fig. 4 Estimation results for $S\alpha S$ X and with triangular kernel (12), $\alpha = 2$ (Gaussian case)

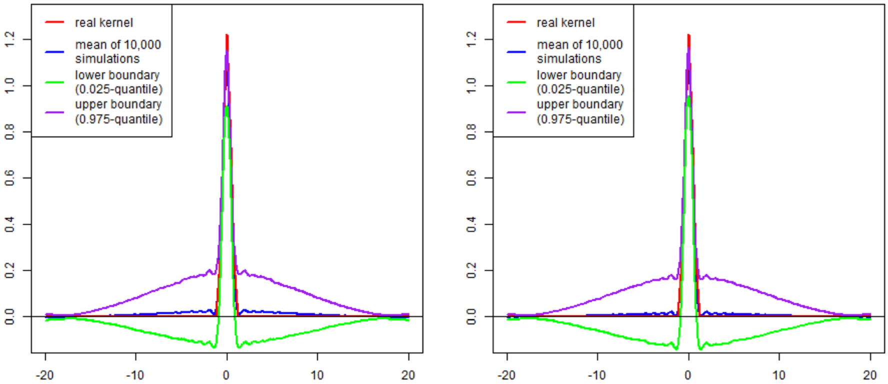

Fig. 5 Estimation results for skewed X with triangular kernel (12), $\alpha = 1.3$ and $\beta = 0.7$ (left), $\beta = -0.5$ (right)

random measures Λ with stability index $\alpha = 1.3$ and skewness intensity $\beta = 0.7, -0.5$, cf. Fig. 5. The parameters of the Gaussian measure Λ were chosen such that $\Lambda(B) \sim N(0, |B|)$ for a bounded Borel subset B.

Finally, we would like to evaluate the performance of our estimator when the integrator Λ is not stable. Since it has to be infinitely divisible, one canonical choice is here of course the Gamma distribution, but we would also like to have a distribution without finite second moment. For this we choose Λ with Lévy density

$$h(x) = \begin{cases} c_1 \frac{|\log x|}{|x|^{p_1}}, & x > \varepsilon, \\ c_2 \frac{|\log(-x)|}{|x|^{p_2}}, & x < -\varepsilon, \\ 0, & |x| \leq \varepsilon \end{cases} \tag{15}$$

for some $\varepsilon \geq 0$, $c_1, c_2 > 0$, $p_1, p_2 > 0$. In more detail, we choose Λ such that for any bounded Borel set $B \subset \mathbb{R}$ we have $\Lambda(B) = \xi(|B|)$ in distribution where $|B|$ is the Lebesgue measure of B and $\xi = \{\xi(t),\ t \geq 0\}$ is the Lévy process given by

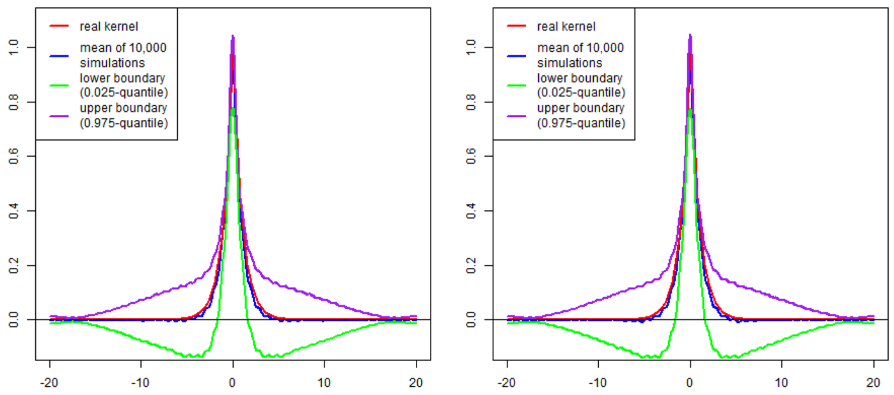

Fig. 6 Estimation results for X with infinitely divisible Λ and exponential kernel (14). Parameters of Lévy density (15) are $c_1 = c_2 = 1$, $p_1 = p_2 = 2.5$ (left) and $p_1 = p_2 = 4$ (right)

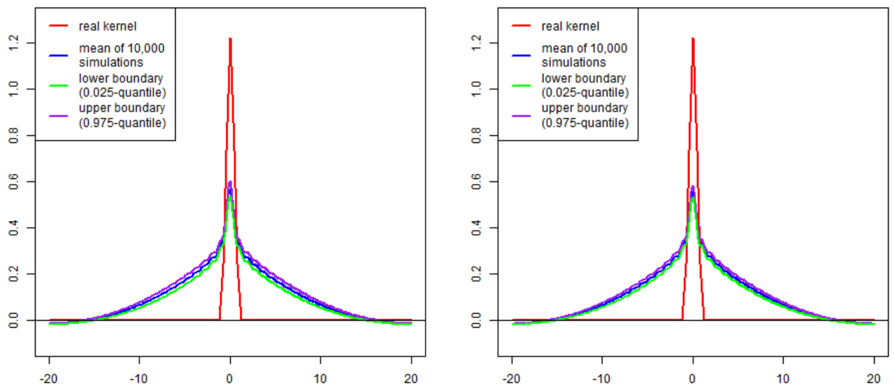

Fig. 7 Estimation results for X with triangular kernel (12), Gamma-distributed Λ (left) and skewed infinitely divisible Λ (right). Parameters of Lévy density (15) are $p_1 = 2.1$, $p_2 = 2.7$ and $c_1 = c_2 = 1$

$$\xi(t) = \int_0^t \int_{\mathbb{R}} xQ(dx, \mathrm{d}s) - t \int_{|x|<1} xh(x)\,\mathrm{d}x, \quad t \geq 0, \tag{16}$$

cf. Sato (2013, Theorem 19.2). Here, Q is a random Poisson measure on $\mathbb{R}_+ \times \mathbb{R}$ with intensity measure $\nu(A, B) = |A| \int_B h(x)\,dx$ for any bounded Borel subset $A \times B \subset \mathbb{R}_+ \times \mathbb{R}$. If $p_1, p_2 \in (0, 3)$ then Λ is not square integrable, cf. Sato (2013, Corollary 25.8). Λ is symmetric iff h is symmetric, i.e., $c_1 = c_2$ and $p_1 = p_2$, cf. Sato (2013, Exercise 18.1). It is known that the distribution of Λ is completely determined by the law of $\xi(1)$. Since the Lévy–Ito representation (16) can be used to generate $\xi(1)$, Karcher et al. (2013) can be used to simulate Λ. We chose $\varepsilon = 0.1$ in order to avoid extremely high jumps.

In the case of Γ-distributed Λ, we set $\Lambda(B) \sim \Gamma(1, |B|)$ for any bounded Borel subset B where a random variable $Y \sim \Gamma(\lambda, p)$ has the density

$$p(x) = \frac{\lambda^p x^{p-1}}{\Gamma(p)} \mathrm{e}^{-\lambda x} \mathbf{1}_{\{x \geq 0\}}.$$

Numerical experiments with non-stable infinitely divisible integrators Λ show that symmetry is an important assumption that cannot be omitted there. Indeed, we saw that the estimation method for f does not work for Gamma-distributed or other unsymmetric non-stable integrators (cf. Fig. 7), but it works well for symmetric infinitely divisible measures Λ with or without finite second moment, compare Fig. 6.

6 Summary and open problems

The preceding section showed the good performance of the high-frequency estimates of a smooth symmetric bounded rapidly decreasing kernel f of positive type for α-stable moving averages X (both skewed and symmetric) in the case $\alpha \in (0, 2]$. Additionally, we verified empirically the applicability of the method to certain non-stable symmetric infinitely divisible integrators Λ. An open problem is to provide rigorous mathematical proofs for this experimental evidence. Recall that we were able to show the consistency of our estimation methods only in the $S\alpha S$ case. Our working hypothesis is that the results of Theorems 1 and 2 stay true for all stable integrators Λ as well as for symmetric infinitely divisible Λ without a finite second moment (at least lying in the domain of attraction of a stable law).

Another open problem is to prove limit theorems for the estimates of g and f in case of $S\alpha S$ Λ. If f is not symmetric (e.g., it is causal) our estimation ansatz fails to work completely, so new ideas are needed here. This is the subject of future research.

Acknowledgements We thank I. Liflyand and V.P. Zastavnyi for the discussion on positive definite functions and their Fourier transforms. We are also grateful to our students L. Palianytsia, O. Stelmakh and B. Ströh for doing numerical experiments in Sect. 5. Finally, we express our gratitude to M. Wendler for drawing our attention to paper Hesse (1990).

Appendix 1: Proofs of Theorems 1, 3

Proof of Theorem 1 We first show how (ii) follows from (i). Notice that $|\hat{g}(\lambda)| = \hat{g}(\lambda)$ for all $\lambda \in \mathbb{R}$, since by (F1) f is of positive type. Since $|\sqrt{a} - \sqrt{b}| \leq \sqrt{|a-b|}$ for $a, b \geq 0$, we get, using the Cauchy–Schwarz inequality,

$$\int_{-a_n}^{a_n} \left(\sqrt{\Delta_n I^s_{n,X}(\lambda)} - \hat{g}(\lambda) \right)^2 \mathrm{d}\lambda \leq \sqrt{2a_n} \cdot$$

$$\sqrt{\int_{-a_n}^{a_n} \left| \Delta_n I^s_{n,X}(\lambda) - \hat{g}(\lambda)^2 \right|^2 \mathrm{d}\lambda} \xrightarrow{P} 0, \quad n \to \infty.$$

Since $\hat{g} \in L^2(\mathbb{R})$, we then have

$$\int_{\mathbb{R}} \left(\mathbf{1}_{[-a_n,a_n]}(\lambda) \cdot \sqrt{\Delta_n I^s_{n,X}(\lambda)} - \hat{g}(\lambda) \right)^2 \mathrm{d}\lambda \xrightarrow{P} 0, \quad n \to \infty.$$

The desired statement now follows from the Plancherel's equality.

Now let us prove (i). Write

$$I^s_{n,X}(\lambda) = \frac{J^s_{n,X}(\lambda)}{S_{n,X}},$$

where $J^s_{n,X}(\lambda) = \sum_{|m|\le m_n} W_n(m) \left| \sum_{j=1}^n X(t_{j,n}) \mathrm{e}^{it_{j,n}\nu_n(m,\lambda)} \right|^2, \quad S_{n,X} = \sum_{j=1}^n X(t_{j,n})^2.$

Let f be supported by $[-T, T]$. We will assume that $N = T/\Delta_n$ is integer: this will simplify the exposition while not harming the rigor. The proof is rather long, so we split it into several steps for better readability. Choose $n \ge 2N + 1$.

Step 1. Denominator

We start with investigating the denominator $S_{n,X}$. First, we study the behavior of a similar expression with f replaced by its discretized version. Specifically, define

$$\begin{aligned} X_n(t_{j,n}) &= \sum_{k=-N}^{N-1} f(t_{k,n}) \Lambda\big(((j-k-1)\Delta_n, (j-k)\Delta_n] \big) \\ &= \int_{\mathbb{R}} f_n(t_{j,n} - s) \Lambda(\mathrm{d}s),\ j = 1, \dots, n, \end{aligned}$$

where $f_n(x) = \sum_{k=-N}^{N-1} f(t_{k,n}) \mathbb{1}_{[t_{k,n}, t_{k+1,n})}(x)$. Denote $\varepsilon_{l,n} = \Lambda\big(((l-1)\Delta_n, l\Delta_n] \big)$, $l \in \mathbb{Z}$. For fixed n, these variables are independent $S\alpha S$ with scale parameter $\Delta_n^{1/\alpha}$.

Decompose

$$\begin{aligned} \sum_{j=1}^n X_n(t_{j,n})^2 &= \sum_{j=1}^n \left(\sum_{l=j-N+1}^{j+N} f(t_{j-l,n}) \varepsilon_{l,n} \right)^2 \\ &= \sum_{j=1}^n \sum_{l=j-N+1}^{j+N} f(t_{j-l,n})^2 \varepsilon_{l,n}^2 \\ &\quad + \sum_{j=1}^n \sum_{\substack{l_1, l_2 = j-N+1 \\ l_1 \ne l_2}}^{j+N} f(t_{j-l_1,n}) f(t_{j-l_2,n}) \varepsilon_{l_1,n} \varepsilon_{l_2,n} \\ &= \left(\sum_{l=N+1}^{n-N} \sum_{j=l-N}^{l+N-1} + \sum_{l=2-N}^{N} \sum_{j=1}^{l+N-1} + \sum_{l=n-N+1}^{n+N} \sum_{j=l-N}^{n} \right) f(t_{j-l,n})^2 \varepsilon_{l,n}^2 \\ &\quad + \sum_{j=1}^n \sum_{\substack{l_1, l_2 = j-N+1 \\ l_1 \ne l_2}}^{j+N} f(t_{j-l_1,n}) f(t_{j-l_2,n}) \varepsilon_{l_1,n} \varepsilon_{l_2,n} \\ &=: S_{1,n} + S_{2,n} + S_{3,n} + S_{4,n}. \end{aligned}$$

We are going to show that the last three terms are negligible. We use the shorthand $E_n = \sum_{l=N+1}^{n-N} \varepsilon_{l,n}^2$, as this will be our benchmark term. Observe that $S_{1,n} = \sum_{k=-N}^{N-1} f(t_{k,n})^2 E_n$. Thanks to the boundedness and uniform continuity of f,

$$\left| S_{1,n} - \frac{1}{\Delta_n} \int_{-T}^{T} f(x)^2 \mathrm{d}x \cdot E_n \right| = O\big(\Delta_n^{-1} \omega_f(\Delta_n) E_n\big), \quad n \to \infty. \tag{17}$$

On the other hand, by Feller (1966, XVII.5, Theorem 3 (i)), we have

$$\frac{E_n}{n^{2/\alpha} \Delta_n^{2/\alpha}} \Rightarrow Z_\alpha, \quad n \to \infty, \tag{18}$$

where Z_α is some positive $\alpha/2$-stable random variable. Therefore, by Slutsky's theorem,

$$\frac{S_{1,n}}{n^{2/\alpha} \Delta_n^{2/\alpha-1}} \Rightarrow Z_\alpha \int_{-T}^{T} f(x)^2 \mathrm{d}x, \quad n \to \infty. \tag{19}$$

Estimating

$$S_{2,n} + S_{3,n} \le \left(\sum_{l=2-N}^{N} + \sum_{l=n-N+1}^{n+N} \right) \varepsilon_{l,n}^2 \sum_{k=-N}^{N-1} f(t_{k,n})^2 =: S_{5,n},$$

similarly to (19), we get

$$\frac{S_{5,n}}{(2N)^{2/\alpha} \Delta_n^{2/\alpha-1}} \Rightarrow Z'_\alpha \int_{-T}^{T} f(x)^2 \mathrm{d}x, \quad n \to \infty. \tag{20}$$

Since $N\Delta_n = T$, we have

$$\begin{aligned} S_{2,n} + S_{3,n} &= O_P(\Delta_n^{-1}) = O_P\big((n\Delta_n)^{-2/\alpha} n^{2/\alpha} \Delta_n^{2/\alpha-1}\big) \\ &= O_P\big(S_{1,n} (n\Delta_n)^{-2/\alpha}\big), n \to \infty. \end{aligned}$$

The term $S_{4,n}$ is estimated using Lemma 3: $S_{4,n} = O_P(N^{3/2} n^{2/\alpha-1/2} \Delta_n^{2/\alpha}) = O_P((n\Delta_n)^{-1/2} S_{1,n})$.

Summing up, we have $\sum_{j=1}^{n} X_n(t_{j,n})^2 = S_{1,n}\big(1 + O_P((n\Delta_n)^{-1/2})\big)$, $n \to \infty$, and $S_{1,n}$ is of order $n^{2/\alpha} \Delta_n^{2/\alpha-1}$, in the sense of (19).

Now we get back to the denominator of $I_{n,X}(\lambda)$. For any positive vanishing sequence $\{\delta_n, n \ge 1\}$, write the following simple estimate:

$$\left|a^2 - b^2\right| \le 2|a(a-b)| + |a-b|^2 \le \delta_n a^2 + (1 + \delta_n^{-1})|a-b|^2. \tag{21}$$

Then, we obtain

$$\left|\sum_{j=1}^{n} X_n(t_{j,n})^2 - S_{n,X}\right| \le \delta_n \sum_{j=1}^{n} X_n(t_{j,n})^2 + (1+\delta_n^{-1}) \sum_{j=1}^{n} \left(X_n(t_{j,n}) - X(t_{j,n})\right)^2.$$

From Lemma 4 it follows that

$$\begin{aligned}\sum_{j=1}^{n} \left(X_n(t_{j,n}) - X(t_{j,n})\right)^2 &= O_P(\|f_n - f\|_\infty^2 n^{2/\alpha} \Delta_n^{2/\alpha-1}) \\ &= O_P(\omega_f(\Delta_n)^2 S_{1,n}), \quad n \to \infty.\end{aligned}$$

Putting $\delta_n = \omega_f(\Delta_n)$ and using (17), we conclude that

$$\Delta_n S_{n,X} = \Delta_n S_{1,n}\left(1 + O_P((n\Delta_n)^{-1/2} + \omega_f(\Delta_n))\right) \overset{P}{\sim} \|f\|_2^2 E_n,\ n \to \infty. \tag{22}$$

Step 2. Whole expression

Thanks to (22),

$$\begin{aligned}&a_n \int_{-a_n}^{a_n} \left|\Delta_n I_{n,X}^s(\lambda) - |\hat{g}(\lambda)|^2\right|^2 \mathrm{d}\lambda \\ &= a_n \int_{-a_n}^{a_n} \left|\frac{\Delta_n^2 J_{n,X}^s(\lambda)}{\Delta_n S_{n,X}} - \frac{\left|\hat{f}(\lambda)\right|^2}{\|f\|_2^2}\right|^2 \mathrm{d}\lambda \\ &\le 2a_n \int_{-a_n}^{a_n} \left|\frac{\Delta_n^2 J_{n,X}^s(\lambda)}{\Delta_n S_{n,X}} - \frac{\left|\hat{f}(\lambda)\right|^2 E_n}{\Delta_n S_{n,X}}\right|^2 \mathrm{d}\lambda + 2a_n \int_{-a_n}^{a_n} \left|\frac{\left|\hat{f}(\lambda)\right|^2 E_n}{\Delta_n S_{n,X}} - \frac{\left|\hat{f}(\lambda)\right|^2}{\|f\|_2^2}\right|^2 \mathrm{d}\lambda \\ &= 2a_n\left[\int_{-a_n}^{a_n} \left|\frac{\Delta_n^2 J_{n,X}^s(\lambda) - \left|\hat{f}(\lambda)\right|^2 E_n}{\Delta_n S_{n,X}}\right|^2 \mathrm{d}\lambda + \left|\frac{\|f\|_2^2 E_n - \Delta_n S_{n,X}}{\Delta_n \|f\|_2^2 S_{n,X}}\right|^2 \int_{-a_n}^{a_n} \left|\hat{f}(\lambda)\right|^4 \mathrm{d}\lambda\right] \\ &= \frac{a_n}{\|f\|_2^4 E_n^2} \int_{-a_n}^{a_n} \left|\Delta_n^2 J_{n,X}^s(\lambda) - \left|\hat{f}(\lambda)\right|^2 E_n\right|^2 \mathrm{d}\lambda + o_P(1), \quad n \to \infty.\end{aligned}$$

Thus, it remains to prove that

$$a_n \int_{-a_n}^{a_n} \left|\Delta_n^2 J_{n,X}^s(\lambda) - \left|\hat{f}(\lambda)\right|^2 E_n\right|^2 \mathrm{d}\lambda = o_P(E_n^2), \quad n \to \infty. \tag{23}$$

Step 3. Numerator

As with the denominator, we start with examining the discretized version of $J_{n,X}^s(\lambda)$:

$$R_n(\lambda) = \sum_{|m|\le m_n} W_n(m) \left| \sum_{j=1}^{n} X_n(t_{j,n}) \mathrm{e}^{it_{j,n}\nu_n(m,\lambda)} \right|^2$$
$$= \sum_{|m|\le m_n} W_n(m) \left| \sum_{j=1}^{n} \sum_{l=j-N+1}^{j+N} f(t_{j-l,n}) \varepsilon_{l,n} \mathrm{e}^{it_{j,n}\nu_n(m,\lambda)} \right|^2.$$

We proceed in three substeps, first considering the following expression

$$R_{1,n}(\lambda) = \sum_{|m|\le m_n} W_n(m) \left| \sum_{l=N+1}^{n-N} \sum_{j=l-N}^{l+N-1} f(t_{j-l,n}) \varepsilon_{l,n} \mathrm{e}^{it_{j,n}\nu_n(m,\lambda)} \right|^2.$$

Step 3a). We shall show

$$a_n \int_{-a_n}^{a_n} \left| \left|\hat{f}(\lambda)\right|^2 E_n - \Delta_n^2 R_{1,n}(\lambda) \right|^2 \mathrm{d}\lambda = o_P(E_n^2), \quad n \to \infty. \tag{24}$$

We have for $\lambda \in [-a_n, a_n]$ that

$$R_{1,n}(\lambda) = \sum_{|m|\le m_n} W_n(m) \left| \sum_{l=N+1}^{n-N} \varepsilon_{l,n} \mathrm{e}^{it_{l,n}\nu_n(m,\lambda)} \sum_{j=l-N}^{l+N-1} f(t_{j-l,n}) \mathrm{e}^{it_{j-l,n}\nu_n(m,\lambda)} \right|^2$$
$$= \sum_{|m|\le m_n} W_n(m) \left| \sum_{k=-N}^{N-1} f(t_{k,n}) \mathrm{e}^{it_{k,n}\nu_n(m,\lambda)} \right|^2 \left| \sum_{l=N+1}^{n-N} \varepsilon_{l,n} \mathrm{e}^{it_{l,n}\nu_n(m,\lambda)} \right|^2$$
$$= F_n(\lambda) \sum_{l=N+1}^{n-N} \varepsilon_{l,n}^2 + \sum_{N+1\le l_1 \ne l_2 \le n-N} a_{l_1,l_2,n}(\lambda) \varepsilon_{l_1,n} \varepsilon_{l_2,n},$$

where

$$F_n(\lambda) = \sum_{|m|\le m_n} W_n(m) \left| \sum_{k=-N}^{N-1} f(t_{k,n}) \mathrm{e}^{it_{k,n}\nu_n(m,\lambda)} \right|^2,$$
$$a_{l_1,l_2,n}(\lambda) = \sum_{|m|\le m_n} W_n(m) \left| \sum_{k=-N}^{N-1} f(t_{k,n}) \mathrm{e}^{it_{k,n}\nu_n(m,\lambda)} \right|^2 \mathrm{e}^{i(l_1-l_2)\Delta_n \nu_n(m,\lambda)} \mathbb{1}_{[-a_n,a_n]}(\lambda).$$

With the help of Lemma 6, we obtain

$$a_n \int_{-a_n}^{a_n} \left| \left|\hat{f}(\lambda)\right|^2 E_n - \Delta_n^2 R_{1,n}(\lambda) \right|^2 \mathrm{d}\lambda$$
$$= O\big(a_n^2 W_n^{(2)} (n\Delta_n)^{-2} + a_n^2 \omega_f(\Delta_n)^2 + a_n^4 \Delta_n^2\big) E_n^2$$
$$+ \ 2a_n \Delta_n^4 \int_{\mathbb{R}} \left| \sum_{N+1 \le l_1 \ne l_2 \le n-N} a_{l_1,l_2,n}(\lambda) \varepsilon_{l_1,n} \varepsilon_{l_2,n} \right|^2 \mathrm{d}\lambda, \quad n \to \infty.$$

By Lemma 2,

$$\int_{\mathbb{R}} \left| \sum_{N+1 \le l_1 \ne l_2 \le n-N} a_{l_1,l_2,n}(\lambda) \varepsilon_{l_1,n} \varepsilon_{l_2,n} \right|^2 \mathrm{d}\lambda = O_P\big(A_n n^{4/\alpha - 2} \Delta_n^{4/\alpha}\big), \quad n \to \infty,$$

where, by Lemma 7,

$$A_n = \int_{-a_n}^{a_n} \sum_{N+1 \le l_1 \ne l_2 \le n-N} \left| a_{l_1,l_2,n}(\lambda) \right|^2 \mathrm{d}\lambda = O\big(a_n W_n^* (K_n^* n)^2\big), \quad n \to \infty,$$

with

$$K_n^* = \sup_{|m| \le m_n} \left| \sum_{k=-N}^{N-1} f(t_{k,n}) \mathrm{e}^{i t_{k,n} v_n(m,\lambda)} \right|^2 \le \left(\sum_{k=-N}^{N-1} \left|f(t_{k,n})\right| \right)^2 \sim \Delta_n^{-2} \|f\|_1^2, \ n \to \infty.$$

Hence,

$$a_n \Delta_n^4 \int_{\mathbb{R}} \left| \sum_{N+1 \le l_1 \ne l_2 \le n-N} a_{l_1,l_2,n}(\lambda) \varepsilon_{l_1,n} \varepsilon_{l_2,n} \right|^2 \mathrm{d}\lambda$$
$$= O_P(a_n^2 W_n^* (n\Delta_n)^{4/\alpha}) = O_P\big(a_n^2 W_n^* E_n^2\big), \ n \to \infty.$$

Combining the estimates, we get (24).

Step 3b). We get

$$a_n \int_{-a_n}^{a_n} \left| \Delta_n^2 R_n(\lambda) - \left|\hat{f}(\lambda)\right|^2 E_n \right|^2 \mathrm{d}\lambda = o_P(E_n^2), \quad n \to \infty. \tag{25}$$

Indeed, write

$$a_n \int_{-a_n}^{a_n} \left| \Delta_n^2 R_n(\lambda) - \left|\hat{f}(\lambda)\right|^2 E_n \right|^2 \mathrm{d}\lambda$$
$$\le 2a_n \int_{-a_n}^{a_n} \Delta_n^4 \left| R_n(\lambda) - R_{1,n}(\lambda) \right|^2 \mathrm{d}\lambda + 2a_n \int_{-a_n}^{a_n} \left| \Delta_n^2 R_{1,n}(\lambda) - \left|\hat{f}(\lambda)\right|^2 E_n \right|^2 \mathrm{d}\lambda$$
$$= 2a_n \Delta_n^4 \int_{-a_n}^{a_n} \left| R_n(\lambda) - R_{1,n}(\lambda) \right|^2 \mathrm{d}\lambda + o_P(E_n^2), \quad n \to \infty.$$

Let us estimate the first expression. Take some positive vanishing sequence $\{\theta_n, n \geq 1\}$, which will be specified later. Using (21), we have

$$\begin{aligned}
\left|R_{1,n}(\lambda) - R_n(\lambda)\right| &\leq \theta_n R_{1,n}(\lambda) + (1+\theta_n^{-1}) \\
&\quad \times \sum_{|m|\leq m_n} W_n(m) \Bigg| \left(\sum_{l=2-N}^{N} \sum_{j=1}^{l+N-1} + \sum_{l=n-N+1}^{n+N} \sum_{j=l-N}^{n} \right) \\
&\quad \varepsilon_{l,n} \mathrm{e}^{it_{l,n}\nu_n(m,\lambda)} f(t_{j-l,n}) \mathrm{e}^{it_{j-l,n}\nu_n(m,\lambda)} \Bigg|^2 \\
&\leq \theta_n R_{1,n}(\lambda) + 2(1+\theta_n^{-1})\big(R_{2,n}(\lambda) + R_{3,n}(\lambda)\big),
\end{aligned}$$

where

$$R_{2,n}(\lambda) = \sum_{|m|\leq m_n} W_n(m) \left| \sum_{l=2-N}^{N} \varepsilon_{l,n} \mathrm{e}^{it_{l,n}\nu_n(m,\lambda)} \sum_{k=1-l}^{N-1} f(t_{k,n}) \mathrm{e}^{it_{k,n}\nu_n(m,\lambda)} \right|^2,$$

$$R_{3,n}(\lambda) = \sum_{|m|\leq m_n} W_n(m) \left| \sum_{l=n-N+1}^{n+N} \varepsilon_{l,n} \mathrm{e}^{it_{l,n}\nu_n(m,\lambda)} \sum_{k=-N}^{n-l} f(t_{k,n}) \mathrm{e}^{it_{k,n}\nu_n(m,\lambda)} \right|^2.$$

Hence,

$$\begin{aligned}
&a_n \int_{-a_n}^{a_n} \left|R_n(\lambda) - R_{1,n}(\lambda)\right|^2 \mathrm{d}\lambda \\
&\quad \leq 2a_n\theta_n^2 \int_{-a_n}^{a_n} R_{1,n}(\lambda)^2 \mathrm{d}\lambda + 16a_n(1+\theta_n^{-1})^2 \int_{-a_n}^{a_n} \big(R_{2,n}(\lambda)^2 + R_{3,n}(\lambda)^2\big)\mathrm{d}\lambda.
\end{aligned}$$

Now

$$\begin{aligned}
R_{2,n}(\lambda) &= \sum_{l=2-N}^{N} \varepsilon_{l,n}^2 \sum_{|m|\leq m_n} W_n(m) \left| \sum_{k=1-l}^{N-1} f(t_{k,n}) \mathrm{e}^{it_{k,n}\nu_n(m,\lambda)} \right|^2 \\
&\quad + \sum_{\substack{l_1,l_2=2-N \\ l_1\neq l_2}}^{N-1} b_{l_1,l_2,n}(\lambda)\varepsilon_{l_1,n}\varepsilon_{l_2,n} \\
&\leq \sum_{l=2-N}^{N} \varepsilon_{l,n}^2 \left(\sum_{k=-N}^{N-1} \left|f(t_{k,n})\right| \right)^2 + \sum_{\substack{l_1,l_2=2-N \\ l_1\neq l_2}}^{N} b_{l_1,l_2,n}(\lambda)\varepsilon_{l_1,n}\varepsilon_{l_2,n} \\
&=: R_{4,n} + R_{5,n}(\lambda),
\end{aligned}$$

where

$$b_{l_1,l_2,n}(\lambda) = \sum_{|m|\le m_n} W_n(m)\mathrm{e}^{i(l_1-l_2)\Delta_n\nu_n(m,\lambda)} \sum_{k_1=1-l_1}^{N-1} \sum_{k_2=1-l_2}^{N-1} f(t_{k_1,n})f(t_{k_2,n})\mathrm{e}^{i(k_1-k_2)\Delta_n\nu_n(m,\lambda)}.$$

Using Lemma 2, we obtain

$$\int_{-a_n}^{a_n} R_{5,n}(\lambda)^2 \mathrm{d}\lambda = O_P(a_n\Delta_n^{-4}), \quad n\to\infty.$$

Further, using (20), $R_{4,n} \sim \Delta_n^{-2}\|f\|_1^2 \sum_{l=2-N}^{N} \varepsilon_{l,n}^2 = O_P(\Delta_n^{-2}),\ n\to\infty$, hence

$$\int_{-a_n}^{a_n} R_{2,n}(\lambda)^2 \,\mathrm{d}\lambda = O_P(a_n\Delta_n^{-4}), n\to\infty.$$

Similarly, $\int_{-a_n}^{a_n} R_{3,n}(\lambda)^2 \,\mathrm{d}\lambda = O_P(a_n\Delta_n^{-4}), n\to\infty$.

Setting $\theta_n = a_n^{1/4}(n\Delta_n)^{-1/\alpha}$, we get by (A3) that

$$\begin{aligned} a_n(1+\theta_n^{-1})^2 \int_{-a_n}^{a_n} \left(R_{2,n}(\lambda)^2 + R_{3,n}(\lambda)^2\right)\mathrm{d}\lambda =& O_P\left(a_n^{3/2}\Delta_n^{-4}(n\Delta_n)^{2/\alpha}\right) \\ =& o_P\left(n^{4/\alpha}\Delta_n^{4/\alpha-4}\right) \end{aligned}$$

as $n\to\infty$. Therefore, we arrive at

$$\begin{aligned} &a_n \int_{-a_n}^{a_n} \left|\Delta_n^2 R_n(\lambda) - \Delta_n^2 R_{1,n}(\lambda)\right|^2 \mathrm{d}\lambda \\ &\le 2a_n^{3/2}(n\Delta_n)^{-2/\alpha}\Delta_n^4 \int_{-a_n}^{a_n} R_{1,n}(\lambda)^2\mathrm{d}\lambda + o_P(n^{4/\alpha}\Delta_n^{4/\alpha}), \quad n\to\infty. \end{aligned}$$

Noting that $o_P(n^{4/\alpha}\Delta_n^{4/\alpha}) = o_P(E_n^2), n\to\infty$, and by Step 3a)

$$\begin{aligned} \Delta_n^4 \int_{-a_n}^{a_n} R_{1,n}(\lambda)^2\mathrm{d}\lambda \le& 2\left(\int_{-a_n}^{a_n} \left|\hat{f}(\lambda)\right|^4 \mathrm{d}\lambda \cdot E_n^2 + \int_{-a_n}^{a_n} \left|\left|\hat{f}(\lambda)\right|^2 E_n - \Delta_n^2 R_{1,n}(\lambda)\right|^2 \mathrm{d}\lambda\right) \\ =& \left(2\int_{-a_n}^{a_n} \left|\hat{f}(\lambda)\right|^4 \mathrm{d}\lambda + o_P(1)\right)E_n^2, \quad n\to\infty, \end{aligned}$$

we get by (A3) $a_n \int_{-a_n}^{a_n} \left|\Delta_n^2 R_n(\lambda) - \Delta_n^2 R_{1,n}(\lambda)\right|^2 \mathrm{d}\lambda = o_P(E_n^2), \quad n\to\infty$, whence (25) follows from (24).

Step 3c). Finally, we have

$$a_n\Delta_n^4 \int_{-a_n}^{a_n} \left|J_{n,X}^s(\lambda) - R_n(\lambda)\right|^2 \mathrm{d}\lambda = o_P(E_n^2), \quad n\to\infty. \tag{26}$$

Using (21) again, write

$$\begin{aligned}\left|J^s_{n,X}(\lambda)-R_n(\lambda)\right| &= \left|R_n(\lambda)-\sum_{|m|\le m_n} W_n(m)\left|\sum_{j=1}^n X(t_{j,n})\mathrm{e}^{it_{j,n}\nu_n(m,\lambda)}\right|^2\right| \\ &\le \delta_n R_n(\lambda)+(1+\delta_n^{-1})\sum_{|m|\le m_n} W_n(m) \\ &\quad \left|\sum_{j=1}^n \big(X_n(t_{j,n})-X(t_{j,n})\big)\mathrm{e}^{it_{j,n}\nu_n(m,\lambda)}\right|^2 \\ &= \delta_n R_n(\lambda)+(1+\delta_n^{-1})\sum_{|m|\le m_n} W_n(m) \\ &\quad \left|\int_{\mathbb{R}}\sum_{j=1}^n \big(f_n(t_{j,n}-s)-f(t_{j,n}-s)\big)\mathrm{e}^{it_{j,n}\nu_n(m,\lambda)}\Lambda(ds)\right|^2 \\ &=: \delta_n R_n(\lambda)+(1+\delta_n^{-1})R_{6,n}(\lambda)\end{aligned} \tag{27}$$

for $\delta_n = a_n^{1/4}\omega_f(\Delta_n)$. Hence,

$$a_n\int_{-a_n}^{a_n}\left|J^s_{n,X}(\lambda)-R_n(\lambda)\right|^2\mathrm{d}\lambda \le 2a_n\delta_n^2\int_{-a_n}^{a_n}R_n(\lambda)^2\mathrm{d}\lambda+2a_n(1+\delta_n^{-1})^2\int_{-a_n}^{a_n}R_{6,n}(\lambda)^2\mathrm{d}\lambda.$$

Define $h_{n,m}(s,\lambda)=\sum_{j=1}^n\big(f_n(t_{j,n}-s)-f(t_{j,n}-s)\big)\mathrm{e}^{it_{j,n}\nu_n(m,\lambda)}\mathbb{1}_{[-a_n,a_n]}(\lambda)$. Note that the summands do not exceed $\omega_f(\Delta_n)$, and at most $2N$ of them are not zero. Hence, $\|h_{n,m}(\cdot,\lambda)\|_\infty \le 2N\omega_f(\Delta_n)$. Applying Lemma 5, we get

$$\begin{aligned}a_n\int_{-a_n}^{a_n}R_{6,n}(\lambda)^2\mathrm{d}\lambda &= O_P(a_n^2N^4\omega_f(\Delta_n)^4(n\Delta_n)^{4/\alpha}) \\ &= O_P(a_n^2\omega_f(\Delta_n)^4n^{4/\alpha}\Delta_n^{4/\alpha-4}),\ n\to\infty.\end{aligned}$$

Recalling that $\delta_n\to 0$ and $a_n^{3/2}\omega_f(\Delta_n)^2\to 0$ as $n\to\infty$ and using (25), we ultimately obtain

$$\begin{aligned}&a_n\Delta_n^4\int_{-a_n}^{a_n}\left|J^s_{n,X}(\lambda)-R_n(\lambda)\right|^2\mathrm{d}\lambda \\ &= O_P\Bigg(a_n^{3/2}(\omega_f(\Delta_n))^2\left(\int_{\mathbb{R}}\left|\hat f(\lambda)\right|^4\mathrm{d}\lambda+o_P(1)\right)E_n^2 \\ &\quad +\Delta_n^4a_n^{-1/2}(\omega_f(\Delta_n))^{-2}a_n^2\omega_f(\Delta_n)^4n^{4/\alpha}\Delta_n^{4/\alpha-4}\Bigg) \\ &= O_p\big(a_n^{3/2}(\omega_f(\Delta_n))^2(E_n^2+n^{4/\alpha}\Delta_n^{4/\alpha})\big)=o_p(E_n^2),\quad n\to\infty.\end{aligned}$$

Combining (25) and (26), we come to (23). □

Proof of Theorem 3

Consider first the case $\alpha\in[1,2)$. (i) Using the triangle inequality and the Hölder inequality, we get

$$\left|\widetilde{\|g\|}_{\alpha,T} - \|g\|_\alpha\right| \le \left(\int_{-T}^{T} |\tilde{g}(t) - g(t)|^\alpha \mathrm{d}t\right)^{1/\alpha}$$
$$\le \left((2T)^{2/\alpha-1} \int_{-T}^{T} |\tilde{g}(t) - g(t)|^2 \mathrm{d}t\right)^{1/2} \xrightarrow{P} 0, \quad n \to \infty.$$

(ii) Similarly to (i),

$$\left|\widetilde{\|g\|}_{\alpha,b_n} - \|g\|_\alpha\right| \le \left(\int_{-b_n}^{b_n} |\tilde{g}(t) - g(t)|^\alpha \mathrm{d}t\right)^{1/\alpha} + \left(\int_{\{t:|t|>b_n\}} |g(t)|^\alpha \mathrm{d}t\right)^{1/\alpha}$$
$$\le \left((2b_n)^{2/\alpha-1} \int_{-b_n}^{b_n} |\tilde{g}(t) - g(t)|^2 \mathrm{d}t\right)^{1/2}$$
$$+ \left(\int_{\{t:|t|>b_n\}} |g(t)|^\alpha \mathrm{d}t\right)^{1/\alpha} \xrightarrow{P} 0, \quad n \to \infty,$$

in view of (8).

The proof of (iii) uses the same ideas and is based on (9).

For $\alpha \in (0,1)$, the proof goes in a similar manner through the triangle inequality for $\|\cdot\|_\alpha^\alpha$. □

Appendix 2: Auxiliary statements

Lemma 1 *Let $(E,\mathcal{E},\nu)$ be a σ-finite measure space, Λ be an independently scattered SαS random measure on E with the control measure ν, and $\{f_t, t \in \mathbf{T}\} \subset L^\alpha(E,\mathcal{E},\nu)$ be a family of functions indexed by some parameter set $\mathbf{T}$, φ be a positive probability density on E. Then*

$$X_t = \int_E f_t(x)\Lambda(\mathrm{d}x), \quad t \in \mathbf{T},$$

has the same finite-dimensional distributions as the almost surely convergent series

$$X_t' = C_\alpha^{1/\alpha} \sum_{k=1}^{\infty} \Gamma_k^{-1/\alpha} \varphi(\xi_k)^{-1/\alpha} f_t(\xi_k)\zeta_k, \quad t \in \mathbf{T},$$

where $\left\{\zeta_k, k \ge 1\right\}$ are iid standard Gaussian random variables, $\left\{\xi_k, k \ge 1\right\}$ are iid random elements of E with density φ, $\Gamma_k = \eta_1 + \cdots + \eta_k$, $\left\{\eta_k, k \ge 1\right\}$ are iid Exp(1) *-distributed random variables, and these three sequences are independent;*

$$C_\alpha = \left(\mathsf{E}\left[\,|g_1|^\alpha\,\right] \int_0^\infty x^{-\alpha} \sin x \,\mathrm{d}x\right)^{-1} = \begin{cases} \frac{(1-\alpha)\sqrt{\pi}}{2^{\alpha/2}\Gamma((\alpha+1)/2)\Gamma(2-\alpha)\cos(\pi\alpha/2)}, & \alpha \ne 1, \\ \sqrt{2/\pi}, & \alpha = 1. \end{cases}$$

Proof The statement follows from Samorodnitsky and Taqqu (1994, Section 3.11) by noting that

$$X_t = \int_E f_t(x)\varphi(x)^{-1/\alpha}M(\mathrm{d}x),$$

where M is an independently scattered SαS random measure on E defined by

$$M(A) = \int_A \varphi^{1/\alpha}(x)\Lambda(x), \quad A \in \mathcal{E},$$

so that the control measure of M has ν-density φ. □

Lemma 2 *Let, for each $n \geq 1$, $\left\{\varepsilon_{m,n}, m = 1, \dots, n\right\}$ be iid SαS random variables with scale parameter σ_n. Let also $\left\{a_{j,l,n},\ 1 \leq j < l \leq n\right\}$ be a collection of measurable functions $a_{j,l,n} : \mathbb{R} \to \mathbb{C}$ such that*

$$A_n = \int_{\mathbb{R}} \sum_{1\leq j<l\leq n} \left|a_{j,l,n}(\lambda)\right|^2 \mathrm{d}\lambda < \infty.$$

Then

$$\int_{\mathbb{R}} \left| \sum_{1\leq j<l\leq n} a_{j,l,n}(\lambda)\varepsilon_{j,n}\varepsilon_{l,n} \right|^2 \mathrm{d}\lambda = O_P(A_n\sigma_n^4 n^{4/\alpha-2}), \quad n \to \infty.$$

Proof W.l.o.g. we can assume that $\sigma_n = 1$. We shall use the LePage series representation. For each $n \geq 1$, the variables $\left\{\varepsilon_{m,n}, m = 1, \dots, n\right\}$ have the same joint distribution as $\{\Lambda([m-1, m]), m = 1, \dots, n\}$, where Λ is an independently scattered SαS random measure on $[0, n]$ with the Lebesgue control measure. By Lemma 1, this distribution coincides with that of

$$\tilde{\varepsilon}_{m,n} = n^{1/\alpha} C_\alpha^{1/\alpha} \sum_{k=1}^{\infty} \Gamma_k^{-1/\alpha} \mathbb{1}_{[m-1,m]}(\xi_k)\zeta_k, \quad m = 1, \dots, n,$$

where $\left\{\Gamma_k, k \geq 1\right\}$ and $\left\{\zeta_k, k \geq 1\right\}$ are as in Lemma 1, $\left\{\xi_k, k \geq 1\right\}$ are iid $U([0, n])$. Since the boundedness in probability relies only on marginal distributions (for fixed n), we can assume that $\varepsilon_{m,n} = \tilde{\varepsilon}_{m,n}$. Let $\Xi_n(\lambda) = \sum_{1\leq j<l\leq n} a_{j,l,n}(\lambda)\varepsilon_{j,n}\varepsilon_{l,n}$. A generic term in the expansion of $|\Xi_n(\lambda)|^2$ has, up to a non-random constant, the form

$$\Gamma_{k_1}^{-1/\alpha}\Gamma_{k_1'}^{-1/\alpha}\Gamma_{k_2}^{-1/\alpha}\Gamma_{k_2'}^{-1/\alpha} \mathbb{1}_{[j_1-1,j_1]}(\xi_{k_1})\mathbb{1}_{[l_1-1,l_1]}(\xi_{k_1'})\mathbb{1}_{[j_2-1,j_2]}(\xi_{k_2})\mathbb{1}_{[l_2-1,l_2]}(\xi_{k_2'})\zeta_{k_1}\zeta_{k_1'}\zeta_{k_2}\zeta_{k_2'}.$$

Recall that $\left\{\zeta_k, k \geq 1\right\}$ are independent and centered. Then, given Γ's and ξ's, such term has a non-zero expectation only if $k_1 = k_2$, $k_1' = k_2'$ or $k_1 = k_2'$, $k_2 = k_1'$ (for $k_1 = k_1'$ it is zero since $j_1 \neq l_1$), so we must also have $j_1 = j_2$, $l_1 = l_2$ or $j_1 = l_2$, $j_2 = l_1$ respectively so that the product of indicators is not zero. The latter, however, is impossible, since $j_1 < l_1$ and $j_2 < l_2$. Consequently, the lemma of Fatou implies

$$\begin{aligned}
&\mathsf{E}\left[\,|\Xi_n(\lambda)|^2 \mid \Gamma\,\right] \\
&\le C_\alpha^{4/\alpha} n^{4/\alpha} \sum_{k\neq k'}^{\infty} \Gamma_k^{-2/\alpha}\Gamma_{k'}^{-2/\alpha} \sum_{1\le j<l\le n}^{n} \left|a_{j,l,n}(\lambda)\right|^2 \mathsf{E}\left[\,\mathbb{1}_{[j-1,j]}(\xi_k)\mathbb{1}_{[l-1,l]}(\xi_{k'})\,\right] \\
&= C_\alpha^{4/\alpha} n^{4/\alpha} \sum_{k\neq k'}^{\infty} \Gamma_k^{-2/\alpha}\Gamma_{k'}^{-2/\alpha} \sum_{1\le j<l\le n}^{n} \left|a_{j,l,n}(\lambda)\right|^2 P\big(\xi_k \in [j-1,j]\big)P\big(\xi_{k'} \in [l-1,l]\big) \\
&\le C_\alpha^{4/\alpha} n^{4/\alpha-2} \sum_{1\le j<l\le n} \left|a_{j,l,n}(\lambda)\right|^2 \left(\sum_{k=1}^{\infty} \Gamma_k^{-2/\alpha}\right)^2.
\end{aligned}$$

Integrating over λ, we get

$$\mathsf{E}\left[\int_{\mathbb{R}} |\Xi_n(\lambda)|^2 \mathrm{d}\lambda \mid \Gamma\right] \le C_\alpha^{4/\alpha} n^{4/\alpha-2} A_n \left(\sum_{k=1}^{\infty} \Gamma_k^{-2/\alpha}\right)^2.$$

By the strong law of large numbers, $\Gamma_k \sim k$, $k \to \infty$, a.s. Therefore, given Γ's, $\int_{\mathbb{R}} |\Xi_n(\lambda)|^2 \mathrm{d}\lambda = O_P(A_n n^{4/\alpha-2})$, $n \to \infty$, whence the required statement follows. □

The following lemma is an immediate corollary of the proof of Lemma 2.

Lemma 3 *Let, for each $n \ge 1$, $\{\varepsilon_{m,n}, m = 1,\dots,n\}$ be iid SαS random variables with scale parameter σ_n. Let also $\{b_{j,l,n},\ 1 \le j < l \le n\}$ be a set of complex numbers with*

$$B_n = \sum_{1\le j<l\le n} \left|b_{j,l,n}\right|^2.$$

Then

$$\sum_{1\le j<l\le n} b_{j,l,n}\varepsilon_{j,n}\varepsilon_{l,n} = O_P(B_n^{1/2}\sigma_n^2 n^{2/\alpha-1}), \quad n \to \infty.$$

In the next two lemmas, $\{\Delta_n, n \ge 1\}$ is some vanishing sequence, $\{N_n, n \ge 1\}$ is a sequence of positive integers such that $N_n \to \infty$, $n \to \infty$, and $N_n = o(n)$, $n \to \infty$. We denote $t_{k,n} = k\Delta_n$, $k \in \mathbb{Z}$, $T_n = N_n\Delta_n$, $n \ge 1$. The proofs of these lemmas are similar to the proof of Lemma 2 and thus omitted. They can be found in the arXiv version of the present paper Kampf et al. (2019).

Lemma 4 *Let $\{h_n, n \ge 1\}$ be a sequence of compactly supported bounded functions such that the bounds of both the function values and the support are uniform in n. Then*

$$\sum_{j=1}^{n} Y_{t_{j,n},n}^2 = O_P\big(\|h_n\|_\infty^2 n^{2/\alpha}\Delta_n^{2/\alpha-1}\big), \quad n \to \infty.$$

Lemma 5 *Let* $\{m_n, n \geq 1\}$ *be a sequence of positive integers such that* $m_n \to \infty$ *as* $n \to \infty$. *For a deterministic sequence* $\{W_n(m), n \geq 1, m = -m_n, \ldots, m_n\}$ *satisfying (W1)–(W2) and continuous functions* $h_{n,m} : [-T_n, n\Delta_n + T_n] \times \mathbb{R} \to \mathbb{C}$, $n \geq 1, m = -m_n, \ldots, m_n$, *define*

$$R_n(\lambda) = \sum_{|m|\leq m_n} W_n(m) \left| \int_{-T_n}^{n\Delta_n+T_n} h_{n,m}(t,\lambda)\Lambda(\mathrm{d}t) \right|^2.$$

Then

$$\int_{\mathbb{R}} R_n(\lambda)^2 \mathrm{d}\lambda = O_P\left(H_n^*(n\Delta_n)^{4/\alpha}\right), \quad n \to \infty,$$

where $H_n^* = \int_{\mathbb{R}} H(\lambda)\,\mathrm{d}\lambda$ *for* $H(\lambda) = \sum_{|m|\leq m_n} W_n(m) \|h_{n,m}(\cdot,\lambda)\|_\infty^4$.

Lemma 6 *Let a bounded uniformly continuous function* $f : \mathbb{R} \to \mathbb{R}$ *with compact support* $[-T, T]$ *and let* $\Delta_n, m_n, W_n(m)$ *and* $\nu_n(m, \lambda)$ *be as defined in Sect.* 1 *or* 3 *fulfilling (W1), (W2) and (W4). Choose a sequence of integers* $(N_n)_{n\in\mathbb{N}}$ *with* $N_n \cdot \Delta_n \sim T$. *Put*

$$F_n(\lambda) = \sum_{|m|\leq m_n} W_n(m) \left| \sum_{k=-N_n}^{N_n-1} f(t_{k,n}) \mathrm{e}^{it_{k,n}\nu_n(m,\lambda)} \right|^2.$$

Then

$$\left| \left|\hat{f}(\lambda)\right|^2 - \Delta_n^2 F_n(\lambda) \right| = O\Big(\left(W_n^{(2)}\right)^{1/2} (n\Delta_n)^{-1} + \omega_f(\Delta_n) + |\lambda|\Delta_n \Big), \quad n \to \infty.$$

We also omit the proof of this and the following lemma and refer the interested reader to Kampf et al. (2019).

Lemma 7 *Let* $\{m_n, n \geq 1\}$ *be a sequence of positive integers such that* $m_n \to \infty$, $m_n = o(n)$, $n \to \infty$, *and let* $\{K_n(m), n \geq 1, m = -m_n, \ldots, m_n\}$ *be a sequence in* $\mathbb{R}$, *and let* $\{W_n(m), n \geq 1, m = -m_n, \ldots, m_n\}$ *be a sequence of filters satisfying* (W1)–(W2). *Then*

$$S_n = \sum_{j_1,j_2=1}^{n} \left| \sum_{|m|\leq m_n} W_n(m) K_n(m) \mathrm{e}^{i(j_1-j_2)m/n} \right|^2 = O(W_n^*(K_n^* n)^2), \ n \to \infty$$

with $W_n^* = \max_{|m|\leq m_n} W_n(m)$, $K_n^* = \max_{|m|\leq m_n} |K_n(m)|$.

References

Akhiezer, N. I. (1988). Lectures on integral transforms, volume 70 of translations of mathematical monographs. Providence, RI: American Mathematical Society. Translated from the Russian by H. H. McFaden.

Brockwell, P. J. (2014). Recent results in the theory and applications of CARMA processes. *Annals of the Institute of Statistical Mathematics*, *66*(4), 647–685.

Brockwell, P. J., Ferrazzano, V., Klüppelberg, C. (2013). High-frequency sampling and kernel estimation for continuous-time moving average processes. *Journal of Time Series Analysis*, *34*(3), 385–404.

Brockwell, P. J., Lindner, A. (2009). Existence and uniqueness of stationary Lévy-driven CARMA processes. *Stochastic Processes and their Applications*, *119*(8), 2660–2681.

Cambanis, S., Podgórski, K., Weron, A. (1995). Chaotic behavior of infinitely divisible processes. *Studia Mathematica*, *115*(2), 109–127.

Chèn, K. (2011). Estimation of the parameters of α-stable distributions. *Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics Series*, (1), 75–81, 127.

Fan, Z. (2006). Parameter estimation of stable distributions. *Communications in Statistics. Theory and Methods*, *35*(1–3), 245–255.

Fasen, V., Fuchs, F. (2013a). On the limit behavior of the periodogram of high-frequency sampled stable CARMA processes. *Stochastic Processes and their Applications*, *123*(1), 229–273.

Fasen, V., Fuchs, F. (2013b). Spectral estimates for high-frequency sampled continuous-time autoregressive moving average processes. *Journal of Time Series Analysis*, *34*(5), 532–551.

Feller, W. (1966). *An introduction to probability theory and its applications*, Vol. II. New York, London, Sydney: Wiley.

García, I., Klüppelberg, C., Müller, G. (2011). Estimation of stable CARMA models with an application to electricity spot prices. *Statistical Modelling. An International Journal*, *11*(5), 447–470.

Gu, J., Mao, S. S. (2002). Estimation of the parameter of stable distributions. *Chinese Journal of Applied Probability and Statistics*, *18*(4), 342–346.

Hesse, C. H. (1990). A Bahadur-type representation for empirical quantiles of a large class of stationary, possibly infinite-variance, linear processes. *The Annals of Statistics*, *18*(3), 1188–1202.

Hida, T., Hitsuda, M. (1993). Gaussian processes, volume 120 of translations of mathematical monographs. Providence, RI: American Mathematical Society. Translated from the 1976 Japanese original by the authors.

Janczura, J., Orzeł, S., Wyłomańska, A. (2011). Subordinated α-stable Ornstein–Uhlenbeck process as a tool for financial data description. *Physica A*, *390*, 4379–4387.

Kampf, J., Shevchenko, G., Spodarev, E. (2019). Nonparametric estimation of the kernel function of symmetric stable moving average random functions (preprint). arXiv:1706.06289.

Karcher, W. (2012). *On infinitely divisible random fields with an application in insurance*. Ph.D. thesis, Ulm University, Ulm, Germany.

Karcher, W., Scheffler, H.-P., Spodarev, E. (2009). Efficient simulation of stable random fields and its applications. In V. Capasso et al. (Eds.), *Stereology and image analysis. Ecs10: Proceedings of the 10th European congress of ISS, the MIRIAM project series* (pp. 63–72). ESCULAPIO Pub. Co., Bologna, Italy.

Karcher, W., Scheffler, H.-P., Spodarev, E. (2013). Simulation of infinitely divisible random fields. *Communications in Statistics. Simulation and Computation*, *42*(1), 215–246.

Karcher, W., Spodarev, E. (2012). *Kernel function estimation of stable moving average random fields*. Ulm University, Ulm (preprint). http://www.uni-ulm.de/fileadmin/website_uni_ulm/mawi.inst.110/forschung/preprints/kernel_estimation_stable.pdf.

Koblents, E., Míguez, J., Rodríguez, M. A., Schmidt, A. M. (2016). A nonlinear population Monte Carlo scheme for the Bayesian estimation of parameters of α-stable distributions. *Computational Statistics & Data Analysis*, *95*, 57–74.

Koutrouvelis, I. A. (1980). Regression-type estimation of the parameters of stable laws. *Journal of the American Statistical Association*, *75*(372), 918–928.

Koutrouvelis, I. A. (1981). An iterative procedure for the estimation of the parameters of stable laws. *Communications in Statistics. B. Simulation and Computation*, *10*(1), 17–28.

Kulik, R. (2007). Bahadur–Kiefer theory for sample quantiles of weakly dependent linear processes. *Bernoulli. Official Journal of the Bernoulli Society for Mathematical Statistics and Probability*, *13*(4), 1071–1090.

Lantuéjoul, C. (2002). *Geostatistical simulation*: *Models and algorithms*, 1st ed. Berlin, Heidelberg: Springer.
McCulloch, J. H. (1986). Simple consistent estimators of stable distribution parameters. *Communications in Statistics. B. Simulation and Computation*, *15*(4), 1109–1136.
Mikosch, T., Gadrich, T., Klüppelberg, C., Adler, R. J. (1995). Parameter estimation for ARMA models with infinite variance innovations. *The Annals of Statistics*, *23*(1), 305–326.
Müller, G., Seibert, A. (2019). Bayesian estimation of stable carma spot models for electricity prices. *Energy Economics*, *78*, 267–277.
Rosiński, J. (1994). On uniqueness of the spectral representation of stable processes. *Journal of Theoretical Probability*, *7*(3), 615–634.
Samorodnitsky, G., Taqqu, M. S. (1994). *Stable non-Gaussian random processes*: *Stochastic models with infinite variance*. Stochastic Modeling. New York: Chapman & Hall.
Sato, K. -I. (2013). Lévy processes and infinitely divisible distributions, volume 68 of Cambridge studies in advanced mathematics. Cambridge: Cambridge University Press. Translated from the 1990 Japanese original, Revised edition of the 1999 English translation.
Trigub, R. M., Bellinsky, E. S. (2004). *Fourier analysis and approximation of functions*. Dordrecht: Kluwer Academic Publishers.
Wang, Y., Yang, W., Hu, S. (2016). The Bahadur representation of sample quantiles for weakly dependent sequences. *Stochastics*, *88*(3), 428–436.
Wu, W. B. (2005). On the Bahadur representation of sample quantiles for dependent sequences. *The Annals of Statistics*, *33*(4), 1934–1963.
Zolotarev, V. M. (1986). One-dimensional stable distributions. Translations of Mathematical Monographs—Vol 65. AMS.
Zolotarev, V. M., Uchaikin, V. V. (1999). *Chance and stability. Stable distributions and their applications. Modern probability and statistics*. Berlin: Walter de Gruyter.