Extremes (2016) 19:371–403
DOI 10.1007/s10687-016-0246-4

Sup-norm convergence rates for Lévy density estimation

Valentin Konakov[1] · **Vladimir Panov**[1]

Received: 14 May 2015 / Revised: 3 March 2016 / Accepted: 4 March 2016 /
Published online: 11 March 2016

Abstract In this paper, we consider projection estimates for Lévy densities in high-frequency setup. We give a unified treatment for different sets of basis functions and focus on the asymptotic properties of the maximal deviation distribution for these estimates. Our results are based on the idea to reformulate the problems in terms of Gaussian processes of some special type and to further analyze these Gaussian processes. In particular, we construct a sequence of excursion sets, which guarantees the convergence of the deviation distribution to the Gumbel distribution. We show that the exact rates of convergence presented in previous articles on this topic are logarithmic and construct the sequence of accompanying laws, which approximate the deviation distribution with polynomial rate.

Keywords Lévy density · Maximal deviation · Nonparametric inference · Projection estimates

AMS 2000 Subject Classifications 60G51 · 62M99 · 62G05

The article was prepared within the framework of a subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program.

✉ Vladimir Panov
vpanov@hse.ru

Valentin Konakov
vkonakov@hse.ru

[1] National Research University Higher School of Economics, Shabolovka 31, building G, 115162 Moscow, Russia

1 Introduction

Consider a one-dimensional Lévy process X_t with Lévy triplet (μ, σ, ν). Assume that the measure ν has a density $s(\cdot)$, known as Lévy density, that is,

$$\nu(B) = \int_B s(u)du, \quad \forall B \in \mathcal{B}(\mathbb{R}\backslash\{0\}),$$

Assuming that some discrete equidistant observations $X_0, X_\Delta, ..., X_{n\Delta}$ of the process X_t are available, it is natural to ask how one can statistically infer on the Lévy density $s(\cdot)$, or more generally speaking, on the Lévy measure ν. The answer to this question highly depends on the type of the available data. The first situation, known as *high-frequency setup*, is based on the assumption that the time distance between the observations $\Delta = \Delta_n$ depends on n and tends to 0 as $n \to \infty$. Moreover, very often (and in this paper) it is also assumed that the time horizon $T = n\Delta_n \to \infty$ as $n \to \infty$. Non-parametric inference for this case has been considered by Comte and Genon-Catalot (2009, 2010a, 2011), Figueroa-López (2011). The second situation, the so-called *low-frequency setup*, in which Δ is fixed, has been extensively studied by Nickl and Reiss (2012), Gugushvili (2012), Belomestny (2011), Comte and Genon-Catalot (2010a, 2010b), Chen, Delaigle and Hall (2010), Neumann and Reiss (2009), van Es, Gugushvili and Spreij (2007). The essential idea in almost all papers mentioned above is to express the Lévy measure in terms of the characteristic function of X_t and then replace this characteristic function by its natural nonparametric estimator.

It is a worth mentioning that in most papers on this topic, the quality of proposed estimator for $s(\cdot)$ is measured in terms of quadratic risk. More precisely, for a fixed estimate $\hat{s}_n^\circ(x)$, a collection of Lévy processes $\mathcal{T}$ and a window $D = [a, b] \subset \mathbb{R}\backslash\{0\}$, it is common to prove two statements, which present upper and lower bounds for the difference between $\hat{s}_n^\circ(x)$ and the true density function $s(x)$. These two statements are usually formulated as follows: for n large enough,

$$\sup_{\mathcal{T}} \mathbb{E}\left(\hat{s}_n^\circ(x) - s(x)\right)^2 \leq f(n), \qquad \forall x \in D,$$

$$\inf_{\{\hat{s}_n(x)\}} \sup_{\mathcal{T}} \mathbb{E}\left(\hat{s}_n(x) - s(x)\right)^2 \geq g(n), \qquad \forall x \in D,$$

where $\left\{\hat{s}_n(x)\right\}$ is the set of all estimates of the Lévy density $s(x)$, and $f(n)$, $g(n)$ are two functions tending to 0 as $n \to \infty$. If $f(n) \asymp g(n)$, it is usually claimed that the estimate $\hat{s}_n^\circ(x)$ is optimal.

Our research has a slightly different focus. We consider the projection estimator $\hat{s}_n(x)$ defined below by Eq. 9, and analyze asymptotic properties of the distribution of $\sup_{x\in D}\left(\left|\hat{s}_n(x) - s(x)\right| / \sqrt{s(x)}\right)$, under the assumption $\inf_{x\in D} s(x) > 0$. To the best of our knowledge, the unique research in this direction is provided by Figueroa-López (2011), who considered the maximal deviation distribution for projection estimates to the space spanned by Legendre polynomials of orders 0 and 1. We emphasize the main differences between our paper and the paper by Figueroa-López (2011) later in Section 6. For the moment, let us only mention that our setup covers more general classes of estimates - in particular, we provide the proof for Legendre polynomials of any order, as well as for trigonometric basis and wavelets.

The motivation of this research comes from the paper by Konakov and Piterbarg (1984), where the asymptotics of the maximal deviation distribution is proven for the kernel estimates of a density and regression functions. Konakov and Piterbarg (1984) showed that the convergence to asymptotic distribution given in Bickel and Rosenblatt (1973) is very slow (of logarithmic order) and this rate cannot be improved. Moreover, in that paper, it is obtained a sequence of distribution laws, which approximate the maximal deviation distribution with power rate of convergence. Nevertheless, the density and regression problems significantly differ from the estimation of Lévy density, and therefore the techniques of the research by Konakov and Piterbarg (1984) are not applicable to our setup.

The contribution of this paper is twofold. First, we derive the asymptotic behaviour of the maximal deviation distribution for a broad class of projection estimates of the Lévy density. This result can be further applied for constructing asymptotic confidence bands and statistical tests. Second, we show that the rate of convergence to the double exponent distribution is of logarithmic order, and this rate cannot be improved. Finally, we provide the sequence of accompanying laws with power rate of convergence.

The paper is organized as follows. In the next section, we explain our setup and assumptions on the set of basis functions. Section 3 contains a collection of our results. Later on, in Sections 4.1–4.3 we prove and discuss these results separately for different choices of basis functions - trigonometric functions, Legendre polynomials and wavelets. Next, we give a general scheme of construction the asymptotic confidence bands for $s(x)$ in Section 5. Some further discussions can be found in Section 6. Additional proofs are collected in the Appendix.

2 Set-up

2.1 Collections of basis functions

In this paper, we follow the set-up from Figueroa-López (2011), and study the estimation of the Lévy density $s(x)$ over a window D, based on discrete observations of the process on an interval $[0, T]$. We consider a family of finite linear combinations of functions from orthonormal collection $\{\varphi_r(x) : D \to \mathbb{R},\ r = 1..d\}$ with respect to the inner product $< f, g >= \int_D f(x)g(x)dx$:

$$\mathscr{L} = \left\{ \sum_{r=1}^{d} \beta_r \varphi_r(x), \qquad \boldsymbol{\beta} = (\beta_1, ..., \beta_d) \in \mathbb{R}^d \right\}, \tag{1}$$

and later project the Lévy density to the space $\mathscr{L}$ in $L^2(D)$ - metric. In this article, we assume that for any $m \in \mathbb{N}$ there exists a set of normalized bounded functions $\left\{\psi_j^m : D \to \mathbb{R}\right\}_{j=0}^{J}$ supported on $[a, a+\delta)$, where $\delta = (b-a)/m$, such that

$$\left\{\varphi_r(x),\ r = 1..d\right\} = \left\{\psi_j^m\left(x - \delta(p-1)\right) \cdot \mathbb{I}\left\{x \in I_p\right\}, j = 0..J,\ \ p = 1..m\right\},$$
$$\text{where} \qquad I_p := [a + \delta(p-1), a + \delta p).$$

In what follows, it is important how the basis functions $\psi_j^m(x)$ depend on m, or in other words, how these functions depend on $\delta = (b-a)/m$. Below we give an intuition about the dependence. Note that in most examples, basis on $[a, a + \delta)$ is constructed from a basis $\{\widetilde{\psi}_j(x)\}_{j=0..J}$ on some "standard" interval $[\tilde{a}, \tilde{b}]$ by changing the variables:

$$\psi_j^m(x) = \sqrt{\frac{\tilde{b}-\tilde{a}}{\delta}} \cdot \widetilde{\psi}_j\left(\frac{(\tilde{b}-\tilde{a})(x-a)}{\delta} + \tilde{a}\right), \tag{2}$$

and therefore $\psi_j^m(x) = O(\sqrt{m})$ as $m \to \infty$. Some typical examples are listed below.

(i) Trigonometric basis on $[a, a+\delta)$

$$\left\{\psi_j^m(x),\ j = 0..J\right\} = \left\{\chi_0(x) = \frac{1}{\sqrt{\delta}},\ \chi_j(x) = \sqrt{\frac{2}{\delta}} \cos\left(2j\pi(x-a)/\delta\right), \right.$$
$$\tilde{\chi}_j(x) = \sqrt{\frac{2}{\delta}} \sin\left(2j\pi(x-a)/\delta\right),$$
$$\left. j = 1..(J/2)\right\} \tag{3}$$

with even J. In this case, it is natural to define the "standard" interval as $[\tilde{a}, \tilde{b}] = [0, 2\pi]$, and basis on this interval as

$$\left\{\widetilde{\psi}_j(x)\right\} = \left\{\frac{1}{\sqrt{2\pi}}, \quad \frac{1}{\sqrt{\pi}}\cos(jx), \quad \frac{1}{\sqrt{\pi}}\sin(jx)\right\}.$$

(ii) Legendre polynomials, that is

$$\psi_j^m(x) = \sqrt{\frac{2j+1}{\delta}} P_j\Big(\left((x-a-\delta) + (x-a)\right)/\delta\Big), \tag{4}$$

where

$$P_j(x) = \frac{1}{j!2^j}\left[\left(x^2-1\right)^j\right]^{(j)}, \quad |x| \le 1, \qquad j = 0..J.$$

The set of orthonormal polynomials on $[\tilde{a}, \tilde{b}] = [-1, 1]$

$$\left\{\widetilde{\psi}_j(x)\right\} = \left\{\sqrt{(2j+1)/2} \cdot P_j(x), \quad j = 0..J\right\},$$

plays the role of standard basis.

(iii) Wavelets, for instance Haar wavelets

$$\psi_j^m(x) = \left\{ \frac{1}{\sqrt{\delta}}, \frac{1}{\sqrt{\delta}} \Big(\mathbb{I}\{x \in [a, a+\delta/2) \right. \\ \left. - \mathbb{I}\{x \in [a+\delta/2, a+\delta]\}\} \Big) \right\}, \quad (5)$$

where δ is usually taken as 2^{-l} for some $l \in \mathbb{N}$. The role of standard interval is usually given to the interval $[\tilde{a}, \tilde{b}] = [0, 1]$ supplied with two functions

$$\left\{ \widetilde{\psi}_j(x) \right\} = \left\{ 1, \quad \mathbb{I}\{x \in [0, 1/2)\} - \mathbb{I}\{x \in [1/2, 1]\} \right\}.$$

To sum up, basis functions typically depend on δ as it is given by Eq. 2, where the functions $\tilde{\psi}_j$ are bounded and supported on some compact $[\tilde{a}, \tilde{b}]$. For theoretical studies, we assume that the function $\sqrt{\delta}\psi_j^m(x)$ and its total variation are bounded by some absolute constants $\mathcal{C}_1$ and $\mathcal{C}_2$, i.e., for all $j = 0..J,\ m \in \mathbb{N}$,

$$\sqrt{\delta} \cdot \sup_{x \in I_1} |\psi_j^m(x)| \le \mathcal{C}_1, \qquad \sqrt{\delta} \cdot V_{\{I_1\}} \left(\psi_j^m \right) \le \mathcal{C}_2, \quad (6)$$

where by $V_{\{I_1\}}(\psi_j^m)$ we denote the total variation of the function ψ_j^m,

$$V_{\{I_1\}}(\psi_j^m) := \sup_{\|P\| \to 0} \sum_{i=1}^{n} \left| \psi_j^m(x_i) - \psi_j^m(x_{i-1}) \right|,$$

P ranges over the partitions $a = x_0 < x_1 < \dots < x_n = a + \delta$, and $\|P\| = \max_i |x_i - x_{i-1}|$.

2.2 Projection estimates

Consider the L^2-scalar product and L^2-norm in the space of functions $\{g : D \to \mathbb{R}\}$, and introduce the function $\tilde{s}(x)$ as the orthogonal projection of $s(x)$ on $\mathscr{L}$ with respect to this norm:

$$\tilde{s}(x) := \sum_{r=1}^{d} \beta_r \varphi_r(x), \quad (7)$$

where

$$\beta_r = \beta(\varphi_r) = \int_D \varphi_r(x) s(x) dx = \int_D \varphi_r(x) \tilde{s}(x) dx.$$

Returning to the statistical problem, that is, to the problem of statistical estimation of $s(x)$ by the equidistant observations $X_0, X_\Delta, \dots, X_{n\Delta}$, we realize that the main difficulty consists in the estimation of $\beta(\varphi_r)$ for different basis functions φ_r. As it was explained earlier, there exists a crucial difference in the assumptions on the design. It turns out, that in case of the low-frequency setup, this question is not well-understood in the literature. As for the high-frequency setup, estimation of $\beta(\varphi_r)$ has been extensively studied in Figueroa-López (2004) and Wörner (2003), where it is shown that the coefficients β_r can be estimated by

$$\hat{\beta}(\varphi_r) = \frac{1}{n\Delta}\sum_{k=1}^{n}\varphi_r\left(X_\Delta^{(k)}\right), \qquad \text{where} \quad X_\Delta^{(k)} = X_{k\Delta} - X_{(k-1)\Delta}. \tag{8}$$

Next, we can substitute the estimator $\hat{\beta}(\varphi_r)$ in Eq. 7, and get that

$$\hat{s}_n(x) := \sum_{r=1}^{d}\hat{\beta}(\varphi_r)\varphi_r(x) = \frac{1}{n\Delta}\sum_{r=1}^{d}\left[\sum_{k=1}^{n}\varphi_r\left(X_\Delta^{(k)}\right)\right]\varphi_r(x) \tag{9}$$

is a reasonable estimator for the Lévy density $s(x)$.

3 Main results

In this section, we present our results related to the projection estimator $\hat{s}_n(x)$ of the Lévy density $s(x)$. First note that from Corollary 8.9 of Sato (1999), it follows that $\hat{\beta}(\varphi_r)$ defined by Eq. 8 is a consistent estimator of $\beta(\varphi_r)$. For theoretical study, we introduce additional assumptions on the rate of this convergence. As in (Figueroa-López 2011), we assume that the following small-time asymptotic property holds: there exist positive constants q and Δ^0 such that

$$\sup_{x\in D}\left|\frac{1}{\Delta}\mathbb{P}\{X_\Delta \geq x\} - \nu([x, +\infty))\right| < q\Delta, \qquad \forall\, 0 < \Delta < \Delta^0. \tag{10}$$

For instance, this property is fulfilled when $s(\cdot)$ is Lipschitz in an open set containing D and uniformly bounded on $|x| > \varepsilon$ for any positive ε (see Proposition 2.1 from Figueroa-López (2011)).

In this paper, we consider the case of high-frequency data with $T \to \infty$ as $n \to \infty$. Moreover, the parameter m, which indicates the number of intervals in our construction of the set of basis functions, also tends to infinity with n. With no doubt, the rates of growth m, n, T should be somehow coordinated. This can be done in different ways. Mainly for technical reasons, we assume that $T = n^{\varkappa}$ with some $\varkappa > 0$. Since in the high-frequency setup $\Delta = T/n \to 0$, we get that $\varkappa < 1$.

In this research, we focus on the analysis of maximal deviation in terms of

$$\mathscr{D}_n := \sup_{x\in D}\left(\frac{|\hat{s}_n(x) - s(x)|}{\sqrt{s(x)}}\right). \tag{11}$$

Note that throughout the paper we assume that $\inf_{x\in D} s(x) > 0$.

We start the analysis of the distribution of $\mathscr{D}_n$ with a technical result related to the random variable

$$\mathscr{Z}_n := \sup_{x\in D}\left(\frac{|\hat{s}_n(x) - \mathbb{E}\hat{s}_n(x)|}{\sqrt{s(x)}}\right), \tag{12}$$

which is “a random part” of $\mathscr{D}_n$. This result allows to reformulate the problem of finding the asymptotic behaviour of the distribution function of $\mathscr{Z}_n$ in terms of Gaussian process $\Upsilon(x)$ defined by

$$\Upsilon(x) = \Upsilon^{J,m}(x) := \sum_{j=0}^{J} Z_j \psi_j^m(x), \qquad x \in I_1, \tag{13}$$

with i.i.d. standard normal r.v.'s $Z_j,\ j = 0, ..., J$.

Proposition 1 *Let Eqs. 6 and 10 hold, and assume that $s(\cdot)$ is a Lipschitz function in an open set containing D. Denote the r.v.*

$$\zeta = \zeta^{J,m} := \sup_{x\in[a,a+\delta)} \left|\Upsilon^{J,m}(x)\right|.$$

Then there exist positive constants $c_1,\ c_2,\ \lambda_1, \lambda_2$, such that for any $u \in \mathbb{R}$ it holds

$$\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \le u\right\} \le \left[\mathbb{P}\left\{\zeta \le \sqrt{m}\left(u + c_1 n^{-\lambda_1}\right)\right\}\right]^m + c_2 n^{-\lambda_2}, \tag{14}$$

$$\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \le u\right\} \ge \left[\mathbb{P}\left\{\zeta \le \sqrt{m}\left(u - c_1 n^{-\lambda_1}\right)\right\}\right]^m - c_2 n^{-\lambda_2}, \tag{15}$$

provided that $T = n^{\varkappa}$ with $\varkappa \in (0, 1)$ and $m = n^{\gamma}$ with $\gamma \in (\theta/2, \theta)$, where

$$\theta := \min\{\varkappa, \quad (1-\varkappa) - 2\lambda_2\} > 0. \tag{16}$$

Remark 1 Clearly the constant λ_2 may be chosen less than $(1-\varkappa)/2$. Moreover, it follows from the proof that Eqs. 14–15 hold with any $\lambda_1 < (\theta - \gamma)/2$.

Proof The proof is given in Appendix A. We divide the proof into 11 steps to simplify the reading. □

In the next sections, we closely consider the Gaussian processes $\Upsilon^{J,m}(x)$, defined by Eq. 13, where $\{\psi_j^m(x)\}$ are the sets of basis functions listed in Section 2. The next theorem demonstrates the asymptotic behavior of the distribution function of the r.v.

$$\tilde{\zeta} = \tilde{\zeta}^{J,m} := \sup_{x\in[a,a+\delta)} \Upsilon^{J,m}(x).$$

Later on, we will derive from this theorem the asymptotic behavior of $\zeta^{J,m}$.

Theorem 1 *Let u grow as $\delta \to 0$ so that $\sqrt{\delta}u \to \infty$. Then in all cases (i) - (iii) described in Section 2.1, it holds*

$$\mathbb{P}\left\{\tilde{\zeta}^{J,m} \ge u\right\} = \frac{g_1(J)}{\left(\sqrt{\delta}u\right)^k} e^{-g_2(J)\cdot\delta u^2}\left(1 + \tau(\sqrt{\delta}u)\right), \tag{17}$$

where $\tau(x) \to 0$ as $x \to \infty$, and

(i) *in case of trigonometric basis* (3),

$$k=0, \qquad g_1(J)=\left(\frac{2\sum_{j=1}^{J/2} j^2}{J+1}\right)^{1/2}, \qquad g_2(J)=\frac{1}{2(J+1)};$$

(ii) *in case of Legendre polynomials* (4),

$$k=1, \qquad g_1(J)=\sqrt{2}(J+1)/\sqrt{\pi}, \qquad g_2(J)=2^{-1}\,(J+1)^{-2};$$

(iii) *in case of wavelets* (5),

$$k=1, \qquad g_1(J)=2/\sqrt{\pi}, \qquad g_2(J)=1/4.$$

Furthermore, in case (i),

$$\tau(x)=\sqrt{J+1}\left(\sqrt{2\pi}x\cdot g_1(J)\right)^{-1}\left(1+O(x^{-2})\right), \qquad x\to\infty,$$

and in case (iii), $\tau(x)=-2x^{-2}\,(1+o(1))\,,\ x\to\infty.$

Layout of the proof As it was mentioned before, there is no unified approach to find the asymptotics of the distribution of Gaussian process. Since the methodology crucially depends on the properties of covariance function, we separately prove this result for different basis functions, see Sections 4.1–4.3. In the case of trigonometric basis (Section 4.1), we efficiently use the stationarity of the considered Gaussian process, and apply some techniques from Piterbarg (1996). In the case of Legendre polynomials (Section 4.2), we take into account that the variance of the process attains its maximum only in finite number of points, and apply the results from Piterbarg (1996) based on the double sum method. Finally, in case of wavelets (Section 4.3), we directly calculate the asymptotic behaviour of $\mathbb{P}\{\zeta\geq u\}$.

In the sequel, we will use (17) in the following form:

$$\mathbb{P}\left\{\tilde{\zeta}^{J,m}\geq u\right\}=\frac{h_1\,m^{k/2}}{u^k}\exp\left\{-h_2\,u^2/m\right\}\left(1+\breve{\tau}\left(u/\sqrt{m}\right)\right), \tag{18}$$

where $h_1=h_1(J):=g_1(J)\cdot(b-a)^{-k/2}$, $h_2=h_2(J):=g_2(J)\cdot(b-a)$, and $\breve{\tau}(x)=\tau\left(\sqrt{b-a}\cdot x\right)$, and u grows with m so that $u/\sqrt{m}\to\infty$.

Using (18), we derive similar result for the supremum of the absolute value of the Gaussian process, which is formulated below.

Corollary 1 *Let u grow with m so that $u/\sqrt{m}\to\infty$. Then it holds*

$$\mathbb{P}\left\{\zeta^{J,m}\geq u\right\}=2\frac{h_1\,m^{k/2}}{u^k}\exp\left\{-h_2\,u^2/m\right\}\left(1+\breve{\tau}\left(u/\sqrt{m}\right)\right). \tag{19}$$

Proof The proof is given in Konakov and Panov (2016), Appendix A.2. □

Proposition 1 and Theorem 1 yield the following theorem, which shows the asymptotic distribution of the maximal deviation $\mathscr{L}_n$.

Theorem 2 *Let the assumptions of Proposition 1 be fulfilled. Denote for any* $y \in \mathbb{R}$*,*

$$u_m = u_m(y) := \frac{y}{a_m} + \left(b_m - \frac{c_m}{b_m}\right), \tag{20}$$

where

$$a_m := 2h_2 b_m, \quad b_m := \sqrt{\frac{1}{h_2}\log(h_1 m)}, \quad c_m := \frac{k}{2h_2}\log b_m. \tag{21}$$

Then in all cases (i) - (iii) described in Section 2.1, uniformly over compact sets in $y \in \mathbb{R}$*,*

$$\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n(y) \le u_m(y)\right\} = e^{-2e^{-y}}\left(1 - 2e^{-y}R(m)\right),$$

where

$$R(m) := \breve{\tau}(u_m) - \frac{k^2}{16}\frac{(\log\log m)^2}{\log m}(1 + o(1)), \quad as m \to \infty. \tag{22}$$

Proof The proof is given in Appendix B. □

In the next theorem, we get the asymptotic distribution of $\mathscr{D}_n$ from the asymptotic distribution of its "random part" $\mathscr{Z}_n$, see Eqs. 11–12.

Theorem 3 *Let the assumptions of Proposition 1 be fulfilled, in particular,* $T = n^{\varkappa}$ *with* $\varkappa \in (0, 1)$ *and* $m = n^{\gamma}$ *with* $\gamma \in (\theta/2, \theta)$*. Then in all cases (i) - (iii) described in Section 2.1, it holds*

$$\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{D}_n \le u_m(y)\right\} = e^{-2e^{-y}}\left(1 - 2e^{-y}R(m)\right), \tag{23}$$

provided that

$$\frac{\varkappa + \gamma}{2} < \min\{1 - \varkappa, 2\gamma\}. \tag{24}$$

Proof The proof is given in Appendix C. Note that the condition (24) doesn't contradict our previous assumptions on $\varkappa$ and γ. For instance, it is fulfilled with any $\gamma \in (\theta/2, \theta)$ if $\varkappa < 1/3$, because in this case $1 - \varkappa > 2\varkappa > 2\theta > 2\gamma$ due to Eq. 16 and λ_2 may be chosen less than $(1/2 - \varkappa)$. □

According to Theorem 1, the function $\breve{\tau}(x) = \tau\left(\sqrt{b-a} \cdot x\right)$ is known for some sets of basis functions. For instance, in case of trigonometric polynomials, $k = 0$, and

$$R(m) = \left(\frac{(J+1)h_2}{2\pi(b-a)h_1^2}\right)^{1/2} \cdot \frac{1}{\sqrt{\log m}}(1 + o(1)), \qquad m \to \infty.$$

In case of wavelets, $\tau(u_m) = O\left(u_m^{-2}\right) = O(1/\log m)$, $k = 1$, and

$$R(m) = -\frac{1}{16}\frac{(\log\log m)^2}{\log m}(1 + o(1)), \qquad m \to \infty.$$

Therefore, the rates of convergence are typically of logarithmic order. Nevertheless, at least in the case of trigonometric basis, we can also find a sequence of accompanying laws, which approximate the distribution of $\mathscr{D}_n$ with polynomial rate. The next theorem clarifies this point.

Theorem 4 *Consider the case of trigonometric basis. Let the assumptions of Proposition 1 and (24) be fulfilled. Define the sequence of distribution functions*

$$A_m(y) := \begin{cases} \exp\left\{-2\exp\left\{-y - \frac{y^2}{4\log(h_1 m)}\right\} - 2m\left(1 - \Phi\left(u_m(y)\sqrt{2h_2}\right)\right)\right\}, & \text{if } y \geq -b_m^{3/2}, \\ 0, & \text{if } y < -b_m^{3/2}, \end{cases} \tag{25}$$

where $\Phi(\cdot)$ is the cdf of the standard normal distribution. Then there exist some positive constants $\bar{c}$, β, such that for sufficiently large n and for any $y \in \mathbb{R}$,

$$\sup_{y\in\mathbb{R}}\left|\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{D}_n \leq u_m(y)\right\} - A_m(y)\right| \leq \bar{c}\, n^{-\beta}. \tag{26}$$

Proof Proof is given in Appendix D. □

So, Theorem 3 yields that convergence to the Gumbel distribution is quite slow, and therefore we cannot state that for some realistic m the maximal deviation distribution is close to its asymptotic distribution. Such situations are typical for similar types of problems, see, e.g., Hall (1991) and Konakov and Piterbarg (1984). Nevertheless, from Theorem 4, we get that the distance between maximal deviation distribution and the distribution function $A_m(y)$ converges to zero at polynomial rate.

The results of this section can be used for constructing asymptotic confidence bands. For instance, from Theorem 4 it follows that

$$I_{\alpha,m} := \left(-\frac{k_{\alpha,m}^{(+)}}{2} + \sqrt{\frac{\left(k_{\alpha,m}^{(-)}\right)^2}{4} + \hat{s}_n(x)},\quad \frac{k_{\alpha,m}^{(+)}}{2} + \sqrt{\frac{\left(k_{\alpha,m}^{(+)}\right)^2}{4} + \hat{s}_n(x)}\right),$$

is a $(1-\alpha)-$confidence band, that is, for m large enough,

$$\mathbb{P}\{s(x) \in I_{\alpha,m},\ \forall x \in D\} = 1 - \alpha,$$

where $k_{\alpha,m}^{(\pm)} := \sqrt{m/T}\left(q_{\alpha,m}a_m^{-1} + b_m \pm C_\alpha n^{-\beta}a_m^{-1}\right)$, $C_\alpha > 0$, and $q_{\alpha,m}$ is the $(1-\alpha)$- quantile of the distribution function $A_m(\cdot)$, see Section 5 for details.

In the next three sections, we separately consider different choices of basis functions: trigonometric basis, Legendre polynomials and wavelets.

4 Proof of Theorem 1

4.1 Stationary case

Let us consider the case of trigonometric basis,

$$\left\{\psi_j^m(x),\ j=0..J\right\} = \left\{\chi_0(x)=\frac{1}{\sqrt{\delta}},\qquad \chi_j(x)=\sqrt{\frac{2}{\delta}}\cos\left(2j\pi(x-a)/\delta\right),\right.$$
$$\left.\tilde{\chi}_j(x)=\sqrt{\frac{2}{\delta}}\sin\left(2j\pi(x-a)/\delta\right),\qquad j=1..(J/2)\right\}$$

on the interval $[a, a+\delta)$. Changing the variables in (13): $x\to t=(x-a)/\delta$, we get

$$\tilde{\zeta} = \max_{t\in[0,1]}\left\{\frac{Z_0}{\sqrt{\delta}}+\sqrt{\frac{2}{\delta}}\sum_{j=1}^{J/2}\left[Z_j\cos(2\pi jt)+\tilde{Z}_j\sin(2\pi jt)\right]\right\},\qquad t\in[0,1],$$

where all Z_j and $\tilde{Z}_j$ are i.i.d. standard normal r.v.'s.

In the rest of this section, we provide the proof of the Theorem 1 (i). This proof is based on Theorem 15.2 from Piterbarg (2015), which we formulate below.

Proposition 2 *Let $X(t)$, $t\in[0,T]$, be a Gaussian stationary differentiable process with zero mean. Let the following conditions be fulfilled:*

(i) *the covariance function $r(t)$ of $X(t)$ has the following asymptotics*

$$r(t)=1-\frac{1}{2}t^2+\alpha t^4+o(t^4),\qquad t\to 0 \tag{27}$$

with some $\alpha>0$;

(ii) *for any $t_1, t_2\in[0,T]$, $t_1\neq t_2$, the covariance matrix of the vector*

$$\left(X(t_1), X(t_2), X'(t_1), X'(t_2)\right)$$

is non-degenerate.

Then

$$\mathbb{P}\left\{\sup_{t\in[0,T]}X(t)\geq u\right\}=\frac{T}{2\pi}e^{-u^2/2}+(1-\Phi(u))+\rho(u), \tag{28}$$

where $0\leq\rho(u)\leq e^{-u^2(1+\chi)/2}$ for u large enough and some $\chi>0$.

To find the asymptotics of $\mathbb{P}\left\{\tilde{\zeta}\geq u\right\}$, we apply Proposition 2 to the process $\widetilde{\Upsilon}(t)$ defined by

$$\widetilde{\Upsilon}(t) := \frac{1}{\sqrt{J+1}}\left\{Z_0+\sqrt{2}\sum_{j=1}^{J/2}\left[Z_j\cos\left(jt/\sqrt{c}\right)+\tilde{Z}_j\sin\left(jt/\sqrt{c}\right)\right]\right\},$$

where $c = \left(2\sum_{j=1}^{J/2} j^2\right) / (J+1)$. Note that

$$\tilde{\xi} \overset{Law}{=} \sqrt{\frac{J+1}{\delta}} \cdot \sup_{t\in[0,2\pi\sqrt{c}]} \widetilde{\Upsilon}(t),$$

and the process $\widetilde{\Upsilon}(t)$ is a centered Gaussian stationary differentiable process with covariance function $r(t)$ satisfying (27).

Next, we verify the condition (ii) from Proposition 2 for any interval $[0, h\sqrt{c}]$, $\forall h \in (0, 2\pi)$. It is equivalent to the assumption that the following determinant is not equal to zero for any $t \in (0, h\sqrt{c}]$:

$$\mathcal{D} := \begin{vmatrix} 1 & 0 & \langle S(0), S(t)\rangle & \langle S(0), S'(t)\rangle \\ 0 & 1 & \langle S'(0), S(t)\rangle & \langle S'(0), S'(t)\rangle \\ \langle S(0), S(t)\rangle & \langle S(0), S'(t)\rangle & 1 & 0 \\ \langle S'(0), S(t)\rangle & \langle S'(0), S'(t)\rangle & 0 & 1 \end{vmatrix},$$

where $\langle .,. \rangle$ is a usual scalar product in $\mathbb{R}^{2N+1}$, and

$$S(t) = \sqrt{\frac{2}{J+1}}\left(\frac{1}{\sqrt{2}}, \cos\left(\frac{t}{\sqrt{c}}\right), \sin\left(\frac{t}{\sqrt{c}}\right), ..., \cos\left(\frac{Nt}{\sqrt{c}}\right), \sin\left(\frac{Nt}{\sqrt{c}}\right)\right),$$

$$S(0) = \frac{1}{\sqrt{J+1}}\left(1, \sqrt{2}, 0, ..., \sqrt{2}, 0\right),$$

$$S'(t) = \sqrt{\frac{2}{(J+1)c}}\left(0, -\sin\left(\frac{t}{\sqrt{c}}\right), \cos\left(\frac{t}{\sqrt{c}}\right), ..., -N\sin\left(\frac{Nt}{\sqrt{c}}\right), N\cos\left(\frac{Nt}{\sqrt{c}}\right)\right),$$

$$S'(0) = \sqrt{\frac{2}{(J+1)c}}\,(0, 0, 1, 0, 2, ..., 0, N),$$

and $N = J/2$. This verification can be found in Section 4.3 from Konakov and Panov (2016). Moreover, one can show (see pp. 14-15 from Konakov and Panov (2016)) that

$$\lim_{h\to 2\pi} \mathbb{P}\left\{\sup_{t\in[0,h\sqrt{c}]} \widetilde{\Upsilon}(t) \geq u\right\} = \mathbb{P}\left\{\sup_{t\in[0,2\pi\sqrt{c}]} \widetilde{\Upsilon}(t) \geq u\right\}.$$

Since all conditions of Proposition 2 are fulfilled, we get

$$\begin{aligned}\mathbb{P}\left\{\tilde{\xi} \geq u\right\} &= \mathbb{P}\left\{\sqrt{\frac{J+1}{\delta}} \sup_{t\in[0,2\pi\sqrt{c}]} \widetilde{\Upsilon}(t) \geq u\right\} \\ &= \sqrt{c}e^{-\delta u^2/(2(J+1))} + \left(1 - \Phi(u\sqrt{\delta/(J+1)})\right) + \rho(u\sqrt{\delta/(J+1)}),\end{aligned} \tag{29}$$

and the statement of the Theorem 1 (i) follows. Moreover, since

$$1 - \Phi(u) = \frac{1}{\sqrt{2\pi}u}e^{-u^2/2}\left(1 - \frac{1}{u^2} + o\left(\frac{1}{u^2}\right)\right), \qquad u \to \infty, \tag{30}$$

(see, e.g., Michna (2009)), we conclude that

$$\tau(x) := \frac{1-\Phi(x/\sqrt{J+1})}{\sqrt{c}e^{-x^2/(2(J+1))}} = \frac{\sqrt{J+1}}{\sqrt{2\pi}\,g_1(J)x}\left(1+O\left(1/x^2\right)\right), \qquad x \to \infty.$$

This observation completes the proof.

4.2 Legendre polynomials

In this section, we consider the orthogonal Legendre polynomials $P_n(x)$, $-1 \le x \le 1$,defined by the formula

$$P_n(x) = \frac{1}{n!2^n}\left[\left(x^2-1\right)^n\right]^{(n)}, \qquad n = 0, 1, 2, ... \tag{31}$$

The orthonormal Legendre polynomials $\widehat{P}_n(x)$ are defined by

$$\widehat{P}_n(x) = \sqrt{\frac{2n+1}{2}}P_n(x), \qquad -1 \le x \le 1. \tag{32}$$

Recall that we are interested in the asymptotic behaviour of the probability of the event $\left\{\breve{\zeta} \ge u\ c\ \sqrt{\delta/2}\right\}$ as $u \to +\infty$, where

$$\breve{\zeta} = \breve{\zeta}^{J,m} := \sup_{x\in[-1,1]} \breve{\Upsilon}^{J,m}(x), \tag{33}$$

and the Gaussian process $\breve{\Upsilon}^{J,m}(x)$ is defined by

$$\breve{\Upsilon}^{J,m}(x) = \breve{\Upsilon}(x) := c\sum_{j=0}^{J}\widehat{P}_j(x)Z_j, \quad x \in [-1, 1], \tag{34}$$

Z_j are i.i.d. standard normal random variables, and c is a constant that will be defined later. Note that here we slightly change the notation introduced in Section 3.

This section is devoted to the proof of Theorem 1 (ii). The main tool of the proof is Corollary 8.3 from (Piterbarg 1996), which we formulate below as Proposition 3.

Proposition 3 *Let $\breve{\Upsilon}(x), x \in [-1, 1]$, be a centered Gaussian, a.s. continuous process with variance $\sigma^2(x)$ and correlation function $\rho(x, y)$. Let $\sigma^2(x)$ attain its global maximum, which is equal to 1, at distinct points $x_1, ..., x_q$. Let the following conditions be fulfilled:*

1. *for all $j = 1..q$,*

$$\sigma(x) = 1 - A_j\left|x-x_j\right|^{\beta_j}(1+o(1)), \qquad as \quad x \to x_j, \tag{35}$$

with some constants $A_j > 0$, $\beta_j > 0$;

2. *(local homogeneity) for all $j = 1..q$,*

$$\rho(x, y) = 1 - C_j|x-y|^{\alpha_j}(1+o(1)), \qquad as \quad x \to x_j,\ y \to x_j \tag{36}$$

with some constants $C_j > 0$, $\alpha_j \in (0, 2]$;

3. *(Global Hölder condition) there exist some $g > 0$, $G > 0$, such that for all $x, y \in [-1, 1]$*

$$\mathbb{E}\left[\left(\breve{\Upsilon}(x) - \breve{\Upsilon}(y)\right)^2\right] \leq G\,|x-y|^g\,; \tag{37}$$

4. *for any $j_1, j_2 = 1..q,\ \ j_1 \neq j_2$,*

$$\rho(x_{j_1}, x_{j_2}) < 1. \tag{38}$$

Then

$$\mathbb{P}\left\{\sup_{x\in[-1,1]} \breve{\Upsilon}(x) > u\right\} = \left(\sum_{j\in Q} D_j\right) u^b\,(1-\Phi\,(u))\,(1+o(1))\,, \qquad u \to \infty, \tag{39}$$

where $b = \min_{j=1..q}\left(2/\beta_j - 2/\alpha_j\right)_-$, $Q := \left\{j = 1..q :\ \left(2/\beta_j - 2/\alpha_j\right)_- = b\right\}$, and D_j are the asymptotic coefficients calculated by Theorem 8.2 from Piterbarg (1996).

Note that Eq. 39 trivially holds with $b = 0$, $\sum_{j\in Q} D_j = 1$ for the process defined by Eq. 34 if $J = 0$. Below we check the conditions of Proposition 3 for $J \geq 1$.

1. The covariance function $r(x, y)$ of the process $\breve{\Upsilon}(x)$ is equal to

$$r(x, y) = c^2 \sum_{j=0}^{J} \frac{2j+1}{2} P_j(x) P_j(y).$$

Denote by $\sigma^2(x)$ the variance of the process $\Upsilon(x)$, that is,

$$\sigma^2(x) = r(x, x) = c^2 \sum_{j=0}^{J} \frac{2j+1}{2} P_j^2(x). \tag{40}$$

From the Laplace formula (Suetin 2005, p. 128, (1)), one can show that $|P_n(x)| < 1$ for $|x| < 1$, see Section 5.1 from Konakov and Panov (2016) for details.

2. Let us check that Eq. 35 holds with $q = 2$, $x_1 = 1$ and $x_2 = -1$. First note that

$$\sigma^2(1) = \sigma^2(-1) =: \sigma^2_{\max} = c^2 \sum_{j=0}^{J} \frac{2j+1}{2} = \frac{c^2}{2}(J+1)^2. \tag{41}$$

Moreover, since $\sigma_{\max} = 1$, we choose $c = \sqrt{2}/(J+1)$. Next, we calculate the derivative of the standard deviation,

$$\sigma'(x) = \frac{c}{\sqrt{2\sum_{j=0}^{J}(2j+1)\,P_j^2(x)}} \sum_{j=0}^{J} (2j+1)\,P_j(x) P_j'(x).$$

In the neighbourhood of $x = 1$ we have an expansion

$$\sigma(x) = \sigma(1) - \sigma'(1)(1-x) + o(1-x), \tag{42}$$

where $\sigma'(1) = cJ(J+1)(J+2)/(4\sqrt{2}) = J(J+2)/4$. Analogously, in the neighborhood of $x=-1$ we have an expansion

$$\sigma(x) = \sigma(-1) + \sigma'(-1)(x+1) + o(x+1), \tag{43}$$

where

$$\sigma'(-1) = -J(J+2)/4.$$

From Eqs. 42 and 43, we finally conclude that Eq. 35 holds with $\beta_1 = \beta_2 = 1$ and $A_1 = A_2 = J(J+2)/4$.

3. The verification of the condition (36) with $\alpha_1 = \alpha_2 = 1$ is made in Section 5.3 from Konakov and Panov (2016).
4. The condition (37) is fulfilled, because

$$\begin{aligned} d^2(x,y) &= c^2 \cdot \mathbb{E}\left[\left(\sum_{j=0}^{J}\sqrt{\frac{2j+1}{2}}\left(P_j(x)-P_j(y)\right)Z_j\right)^2\right] \\ &= c^2\sum_{j=0}^{J}\frac{2j+1}{2}\left(P_j(x)-P_j(y)\right)^2 \\ &\le \left[c^2\sum_{j=0}^{J}\frac{(2j+1)j^2(j+1)^2}{8}\right](x-y)^2, \end{aligned}$$

where the last inequality follows from $\max_{[-1,1]}\left|P_j'(x)\right| = P_j'(1) = j(j+1)/2$. The proof of this fact follows from Gradshtein and Ryzhik I. (1996), 8.915 (2), and can be found in Section 5.1 from Konakov and Panov (2016).
5. Note also the condition (38) is fulfilled because for $J \ge 1$

$$r(-1,1) = c^2\sum_{j=0}^{J}\frac{2j+1}{2}(-1)^j < c^2\sum_{j=0}^{J}\frac{2j+1}{2} = r(1,1) = 1.$$

6. Applying Proposition 3, we arrive at Eq. 39 with $b=0$, $Q=\{1,2\}$, $D_1 = D_2 = 1$. Taking into account (30), we get

$$\mathbb{P}\left\{\sup_{x\in[-1,1]}\breve{\Upsilon}(x) > u\right\} = \frac{\sqrt{2}}{\sqrt{\pi}u}e^{-u^2/2}\left(1+o(1)\right), \qquad u\to\infty.$$

Finally, we conclude that

$$\begin{aligned} \mathbb{P}\left\{\tilde{\zeta} > u\right\} &= \mathbb{P}\left\{\sup_{x\in[-1,1]}\breve{\Upsilon}(x) > uc\sqrt{\delta/2}\right\} \\ &= \frac{2}{(\sqrt{\pi}c)\cdot(\sqrt{\delta}u)}e^{-c^2\delta u^2/4}\left(1+o(1)\right), \qquad u\to\infty. \end{aligned}$$

This observation completes the proof of Theorem 1 (ii).

4.3 Wavelets

In this section, we consider the case of Haar wavelets, that is,

$$\psi_j(x) = \left\{ \frac{1}{\sqrt{\delta}}, \qquad \frac{1}{\sqrt{\delta}}\Big(\mathbb{I}\{x \in [a, a+\delta/2)\} - \mathbb{I}\{x \in [a+\delta/2, a+\delta]\}\Big)\right\},$$

where $\delta = 2^{-l}$ for some $l \in \mathbb{N}$. For this set of functions,

$$\tilde{\zeta} = 2^{l/2}\,(Z_0 + |Z_1|.)$$

Therefore the distribution of $2^{-l/2}\tilde{\zeta}$ has density

$$p_{Z_0+|Z_1|}(x) = \int_{-\infty}^{x} 2\,\phi(x-y)\,\phi(y)dy = \frac{1}{\sqrt{\pi}}e^{-x^2/4}\Phi\left(x/\sqrt{2}\right),$$

where by $\phi(\cdot)$ we denote the density of the standard normal distribution. This density corresponds to the distribution function

$$F_{Z_0+|Z_1|}(x) = \Phi^2\left(x/\sqrt{2}\right).$$

Next, we apply a Taylor expansion of the function $1 - \Phi(x)$ for large x up to the second order (30), and get that

$$\begin{aligned}\mathbb{P}\left\{\tilde{\zeta} \geq u\right\} &= 1 - \Phi^2\left(\sqrt{\delta}u/\sqrt{2}\right) \\ &= \frac{2}{\sqrt{\pi}\sqrt{\delta}u}e^{-\delta u^2/4}\left(1 - \frac{2}{\delta u^2} + o\left(\frac{1}{\delta u^2}\right)\right).\end{aligned} \tag{44}$$

This completes the proof of Theorem 1 (iii).

5 Asymptotic confidence bands

In this section, we show how Theorem 4 can be used for the construction of the asymptotic confidence bands. In fact, (26) can be rewritten as

$$\mathbb{P}\left\{\mathscr{D}_n \leq \sqrt{\frac{m}{T}}\left(\frac{y}{a_m} + b_m\right)\right\} = A_m(y) + e_n(y). \tag{45}$$

where $|e_n(y)| \leq \bar{c}\; n^{-\beta}$ for any $y \in \mathbb{R}$, $\bar{c} > 0$, $\beta > 0$, a_m and b_m are defined in Eq. 21, and n is sufficiently large. Let us fix some confidence level $\alpha \in (0, 1)$ and denote the $(1-\alpha)$- quantiles of the distribution functions $A_m(\cdot)$ and $A_m(\cdot) + e_n(\cdot)$ by $q_{\alpha,m}$ and $\tilde{q}_{\alpha,m}$ resp., that is,

$$A_m(q_{\alpha,\,m}) = A_m(\tilde{q}_{\alpha,\,m}) + e_n(\tilde{q}_{\alpha,\,m}) = 1 - \alpha.$$

Denote also by q_α the $(1-\alpha)$ - quantile of the distribution function $A(y) := e^{-2e^{-y}}$. It is easy to see that $A_m(y) \to A(y)$, $q_{\alpha,m} \to q_\alpha$, $\tilde{q}_{\alpha,m} \to q_\alpha$ as $m \to \infty$.

Returning now to Eq. 45 and recalling the definition of $\mathscr{D}_n$, see Eq. 11, we conclude that

$$\frac{\left|\hat{s}_n(x) - s(x)\right|}{\sqrt{s(x)}} \le \tilde{k}_{\alpha,m} := \sqrt{\frac{m}{T}}\left(\frac{\tilde{q}_{\alpha,m}}{a_m} + b_m\right), \qquad \forall x \in D \tag{46}$$

with probability $1-\alpha$. Solving (46) as the quadratic inequality with respect to $\sqrt{s(x)}$, we obtain that

$$\tilde{I}_{\alpha,m} := \left(-\frac{\tilde{k}_{\alpha,m}}{2} + \sqrt{\frac{\tilde{k}^2_{\alpha,m}}{4} + \hat{s}_n(x)}, \quad \frac{\tilde{k}_{\alpha,m}}{2} + \sqrt{\frac{\tilde{k}^2_{\alpha,m}}{4} + \hat{s}_n(x)}\right)$$

is a $(1-\alpha)-$confidence band for $s(x)$, that is, $\mathbb{P}\{s(x) \in \tilde{I}_{\alpha,m},\ \forall x \in D\} = 1-\alpha$, provided that m is large enough.

Note that $\tilde{q}_{\alpha,m}$ (which is involved in the last formula through $\tilde{k}_{\alpha,m}$) is a theoretical value, which is not known. To avoid this drawback, we will show that $\tilde{q}_{\alpha,m}$ can be replaced by the quantile $q_{\alpha,m}$, which can be computed because $A_m(y)$ is known explicitly.

Since the function $A_m(y)$ is differentiable with respect to y in all points except $y = -b_m^{3/2}$ (see Eq. 25), we get that

$$e_n(\tilde{q}_{\alpha,m}) = A_m(q_{\alpha,m}) - A_m(\tilde{q}_{\alpha,m}) = A'_m(\theta)\cdot\left(q_{\alpha,m} - \tilde{q}_{\alpha,m}\right),$$

where θ lies between $q_{\alpha,m}$ and $\tilde{q}_{\alpha,m}$ and it is assumed that α is chosen such that $q_\alpha > 0$ and therefore $q_{\alpha,m}, \tilde{q}_{\alpha,m} > -b_m^{3/2}$ for m large enough.

The derivative of $A_m(y)$ is not equal to zero for any $y > -b_m^{3/2}$, and therefore

$$\left|\tilde{q}_{\alpha,m} - q_{\alpha,m}\right| \le \frac{\left|e_n(\tilde{q}_{\alpha,m})\right|}{\inf_{y\in Q}\left|A'_m(y)\right|} \le C_\alpha n^{-\beta},$$

where Q is a small vicinity of q_α, and C_α is a positive constant do not depending on m, because $A'_m(y) \to A'(y)$ as $m \to \infty$ uniformly in Q. Finally, we conclude that for m large enough, the band

$$I_{\alpha,m} := \left(-\frac{k^{(+)}_{\alpha,m}}{2} + \sqrt{\frac{\left(k^{(-)}_{\alpha,m}\right)^2}{4} + \hat{s}_n(x)}, \quad \frac{k^{(+)}_{\alpha,m}}{2} + \sqrt{\frac{\left(k^{(+)}_{\alpha,m}\right)^2}{4} + \hat{s}_n(x)}\right),$$

where

$$k^{(\pm)}_{\alpha,m} := \sqrt{\frac{m}{T}}\left(\frac{q_{\alpha,m}}{a_m} + b_m \pm C_\alpha n^{-\beta} a_m^{-1}\right),$$

covers $s(x)$ with probability larger or equal to $1-\alpha$. Note also that Theorem 3 can be also used for the construction of confidence band for $s(x)$ in the similar way, but the width of this band will be asymptotically larger.

6 Discussion

The main contributions of this paper are Theorems 3 and 4, which give the asymptotic behavior of the maximal deviation distribution for projection estimates. Our research is motivated by the paper (Figueroa-López 2011), which is to the best of our knowledge the unique publication on this topic. Note that in our research, we provide a unified treatment for different sets of basis functions (Legendre polynomials of any order, trigonometric basis, wavelets), whereas the paper (Figueroa-López 2011) is concentrated on the Legendre polynomials of degree 0 and 1 (piecewise constant and piecewise linear functions). Moreover, in comparison to Figueroa-López (2011), our research has a slightly different focus - we show that the rates of convergence of the maximal deviation distribution to the Gumbel distribution are of logarithmic order, and therefore the Gumbel approximation is not appropriate in realistic situations (see Theorems 2 and 3). Finally, we find a sequence of accompanying laws, which approximates the maximal deviation distribution with polynomial rate, see Theorem 4.

The main ingredient of the proof of Theorems 2-4 is formulated as Theorem 1, which reveals the essential difference in asymptotic behaviour of Gaussian processes for different sets of basis functions. An open problem is to prove similar facts for any basis under some mild conditions. The existing theory for Gaussian processes doesn't have a unified remedy for solving such problems, and therefore this issue can be a topic for further research.

Appendix

A Proof of Proposition 1

1. *Preliminary remarks.* For a function $G(\cdot)$ and positive constant h, introduce the notation

$$\mathscr{L}(h, G, d; x) := h \sum_{r=1}^{d} \left[\int_D \varphi_r(u)\, dG(u) \right] \varphi_r(x).$$

Note that

$$\hat{s}_n(x) - \mathbb{E}\hat{s}_n(x) = \mathscr{L}\Big(\sqrt{n}/T, Z_n(F_\Delta(\cdot)), d; x\Big) =: \mathscr{L}_1(x),$$

where $Z_n(\cdot)$ is the empirical process of a uniform on $[0, 1]$ random sample $F_\Delta(X_\Delta^{(k)}),\ k = 1..n$, i.e.,

$$Z_n(x) := \sqrt{n} \left(\frac{1}{n} \sum_{k=1}^{n} \mathbb{I}\left\{ F_\Delta\left(X_\Delta^{(k)}\right) \le x \right\} - \mathbb{P}\left\{ F_\Delta\left(X_\Delta^{(k)}\right) \le x \right\} \right),$$

and $F_t(\cdot)$ is the distribution function of X_t. In our notations, $Z_n(F_\Delta(\cdot))$ is the empirical process for $X_\Delta^{(k)}$, $k = 1..n$,

$$Z_n(F_\Delta(x)) = \sqrt{n}\left(\frac{1}{n}\sum_{k=1}^{n}\mathbb{I}\left\{X_\Delta^{(k)} \le x\right\} - F_\Delta(x)\right).$$

2. *Komlós-Major-Tusnady construction.* Applying Theorem 3 in Komlós et al. (1975), we get that there exists a version of $Z_n(x)$ (denoted below also as $Z_n(x)$ for sake of simplicity) and a sequence of Brownian bridges $B_n(x)$ such that for any $y > 0$ the probability of the event

$$\mathcal{W}_n(y) := \left\{\sup_{x\in[0,1]} |Z_n(x) - B_n(x)| \le C_1\frac{\log(n)}{\sqrt{n}} + y\right\}$$

is larger than $1 - Ke^{-\lambda y\sqrt{n}}$, where $B_n(x) = W_n(x) - xW_n(1)$ with Brownian motions $W_n(x)$, and C_1, K, λ are some positive absolute constants. As usual, we take $y = y^* = \log(n)/\sqrt{n}$, and get that the event $\mathcal{W}_n^* = \mathcal{W}_n(y^*)$ defined by

$$\mathcal{W}_n^* := \left\{\sup_{x\in[0,1]} |Z_n(x) - B_n(x)| \le (C_1 + 1)\,\frac{\log(n)}{\sqrt{n}}\right\} \tag{47}$$

is of probability larger than $1 - K/n^\lambda$. Note that if this statement is fulfilled for $\lambda = \lambda_1$, then it also holds with any $0 < \lambda < \lambda_1$.

3. $\mathscr{L}_1(x) \rightsquigarrow \mathscr{L}_2(x) := \mathscr{L}\left(\sqrt{n}/T, B_n(F_\Delta(\cdot)), d; x\right)$. By the definition of the functional $\mathscr{L}$,

$$\mathscr{L}_1(x) - \mathscr{L}_2(x) = \mathscr{L}\left(\sqrt{n}/T, Z_n(F_\Delta(\cdot)) - B_n(F_\Delta(\cdot)), d; x\right) \tag{48}$$

It is worth mentioning that for a function G,

$$\sup_{x\in D}\left|\mathscr{L}\left(h, G, d; x\right)\right| \le C_2\ h\ m\ w(G, D, \delta), \tag{49}$$

where C_2 is a positive constant depending on (φ_r), and w is the modulus of continuity, i.e.,

$$w(G, D, \delta) := \sup\left\{|G(u) - G(v)| :\ u, v \in D,\ |u - v| < \delta\right\}.$$

In fact, we get for $x \in I_p = [a + \delta(p-1), a + \delta p)$,

$$\begin{aligned}
\left|\mathscr{L}\Big(h, G, d; x\Big)\right| &= \left| h \sum_{j=0}^{J} \left[\int_a^{a+\delta} \psi_j^m(u)\, dG\left(u + \bar{\delta}\right) \right] \psi_j^m(x - \bar{\delta}) \right| \\
&= \left| h \sum_{j=0}^{J} \Big[\psi_j^m(a+\delta) \Big(G(a + \bar{\delta} + \delta) - G(a + \bar{\delta}) \Big) \right. \\
&\qquad \left. - \int_a^{a+\delta} \Big(G(u + \bar{\delta}) - G(a + \bar{\delta}) \Big) d\psi_j^m(u) \Big] \psi_j^m(x - \bar{\delta}) \right| \\
&\leq h \sum_{j=0}^{J} \left(\sup_{x \in I_1} |\psi_j^m(x)| + V_a^{a+\delta}\left(\psi_j^m\right) \right) \sup_{x \in I_1} \left|\psi_j^m(x)\right| \cdot w(G, D, \delta) \\
&\leq C_2\, h\, m\, w(G, D, \delta),
\end{aligned}$$

where $\bar{\delta} := \delta(p-1)$, $C_2 > 0$, and the conditions (6) are used. Combining (47), (48) and (49) we get that on the event $\mathcal{W}_n^*$,

$$\sup_{x \in D} |\mathscr{L}_1(x) - \mathscr{L}_2(x)| \leq C_3 \frac{m \log(n)}{T}, \qquad \text{where} \quad C_3 > 0.$$

4. $\mathscr{L}_2(x) \rightsquigarrow \mathscr{L}_3(x) := \mathscr{L}\Big(\sqrt{n}/T, B_n(1 - F_\Delta(\cdot)), d; x\Big)$. Taking into account that $B_n(x) \stackrel{Law}{=} B_n(1-x)$ in $C([0,1])$, we get

$$\mathscr{L}_2(x) = \mathscr{L}\Big(\sqrt{n}/T, B_n(F_\Delta(\cdot)), d; x\Big) \stackrel{Law}{=} \mathscr{L}\Big(\sqrt{n}/T, B_n(1 - F_\Delta(\cdot)), d; x\Big) = \mathscr{L}_3(x).$$

5. $\mathscr{L}_3(x) \rightsquigarrow \mathscr{L}_4(x) := \mathscr{L}\Big(\sqrt{n}/T, W_n(1 - F_\Delta(\cdot)), d; x\Big)$. Obviously,

$$\mathscr{L}_4(x) - \mathscr{L}_3(x) = \mathscr{L}\Big(\sqrt{n}/T, (1 - F_\Delta(\cdot))\, W_n(1), d; x\Big).$$

Similarly to Step 3, we get with some $C_4, C_5 > 0$

$$\begin{aligned}
\sup_{x \in D} |\mathscr{L}_4(x) - \mathscr{L}_3(x)| &\leq C_2 \frac{\sqrt{n} m}{T} w\Big(1 - F_\Delta(\cdot), D, \delta\Big)\ |W_n(1)| \\
&\leq \frac{\sqrt{n} m}{T} \Big(C_4 \Delta^2 + C_5 \delta \Delta\Big)\ |W_n(1)| \\
&\leq \frac{m}{\sqrt{n}} \Big(C_4 \Delta + C_5 \delta\Big)\ |W_n(1)|,
\end{aligned}$$

where the second inequality holds due to Eq. 10: for any $u < v,\ u, v \in D$,

$$\begin{aligned}
|\mathbb{P}\{X_\Delta \geq u\} - \mathbb{P}\{X_\Delta \geq v\}| &< q\Delta^2 + \nu([u, v])\, \Delta \\
&\leq q\Delta^2 + \max_{x \in [u,v]} |s(x)| \cdot |v - u| \Delta. \qquad (50)
\end{aligned}$$

6. $\mathscr{L}_4(x) \rightsquigarrow \mathscr{L}_5(x) := \mathscr{L}\Big(1/\sqrt{T}, W_n\left(\left(1 - F_\Delta(\cdot)\right)/\Delta\right), d; x\Big)$.
   Applying the self-similarity property of the Brownian motion, we get

$$\mathscr{L}\Big(\sqrt{n}/T, W_n(1 - F_\Delta(\cdot)), d; x\Big) \overset{Law}{=} \mathscr{L}\Big(1/\sqrt{T}, W_n\left(\left(1 - F_\Delta(\cdot)\right)/\Delta\right), d; x\Big).$$

7. $\mathscr{L}_5(x) \rightsquigarrow \mathscr{L}_6(x) := \mathscr{L}\Big(1/\sqrt{T}, W_n\left(\int_\cdot^{+\infty} s(u)du\right), d; x\Big)$.
   Using (49) and the assumption (10), we get

$$\begin{aligned}\sup_{x\in D} |\mathscr{L}_5(x) - \mathscr{L}_6(x)| &\le C_2 \frac{m}{\sqrt{T}} w\left(W_n\Big(\left(1 - F_\Delta(\cdot)\right)/\Delta\Big) - W_n\Big(\int_\cdot^{+\infty} s(u)du\Big), D, \delta\right)\\ &\le C_6 \frac{m}{\sqrt{T}} w(W_n, D, q\Delta), \qquad \text{with } C_6 > 0.\end{aligned}$$

In the paper (Fisher and Nappo 2010), it is proven that

$$\mathbb{E}\left[w(W_n, D, q\Delta)\right] \le C_7 \sqrt{\Delta \log\left(\frac{1}{\Delta}\right)}, \qquad \text{with } C_7 > 0,$$

and therefore due to Chebyshev inequality, it holds with any $\varepsilon > 0$,

$$\begin{aligned}\mathbb{P}\left\{\sup_{x\in D} |\mathscr{L}_5(x) - \mathscr{L}_6(x)| > \varepsilon\right\} &\le \mathbb{P}\left\{w(W_n, D, q\Delta) > \varepsilon \frac{\sqrt{T}}{C_6 m}\right\}\\ &\le C_6 C_7 \frac{m}{\varepsilon\sqrt{T}} \sqrt{\Delta \log\left(\frac{1}{\Delta}\right)}.\end{aligned}$$

8. $\mathscr{L}_6(x) \rightsquigarrow \mathscr{L}_7(x) := \mathscr{L}\left(1/\sqrt{T}, \int_\cdot^{+\infty} \sqrt{s(u)} dW_n(u), d; x\right)$.

The functionals $\mathscr{L}_6(x)$ and $\mathscr{L}_7(x)$ have the same distributions, because

$$W_n\left(\int_\cdot^{+\infty} s(u)du\right) \overset{Law}{=} \int_\cdot^{+\infty} \sqrt{s(u)} dW_n(u)$$

in $C([a, b])$. Recall that

$$\begin{aligned}\mathscr{L}\Big(1/\sqrt{T}, \int_\cdot^{+\infty} \sqrt{s(u)} dW_n(u), d; x\Big) &= -\frac{1}{\sqrt{T}} \sum_{r=1}^{d} \left[\int_D \varphi_r(u)\sqrt{s(u)}\, dW_n(u)\right] \varphi_r(x)\\ &= -\sqrt{\frac{s(x)}{T}} \sum_{r=1}^{d} \left[\int_D \varphi_r(u) \sqrt{\frac{s(u)}{s(x)}}\, dW_n(u)\right] \varphi_r(x).\end{aligned}$$

9. $\mathscr{L}_7(x) \rightsquigarrow \mathscr{L}_8(x) := \mathscr{L}\left(-\sqrt{s(x)/T}, W_n(\cdot), d; x\right)$. Let us show that

$$\sup_{x\in D} |\mathscr{L}_8(x) - \mathscr{L}_7(x)| \le C_8 T^{-1/2} \sup_{x\in D} |W_n(x)|, \qquad \text{with } \ C_8 > 0. \tag{51}$$

In fact, we can represent the difference in Eq. 51 as

$$\mathscr{L}_7(x) - \mathscr{L}_8(x) = \sqrt{\frac{s(x)}{T}} \sum_{p=1}^{m} \sum_{j=0}^{J} G_{j,p}(x) \varphi_{j,p}(x),$$

where $\varphi_{j,p}(x) = \psi_j^m (x - \delta(p-1)) \mathbb{I}\left\{x \in I_p\right\}, j = 0..J, \ p = 1..m$, and

$$G_{j,p}(x) := \int_{I_p} \varphi_{j,p}(u) dW_n(u) - \int_{I_p} \varphi_{j,p}(u) \sqrt{\frac{s(u)}{s(x)}} dW_n(u).$$

Using integration by parts for continuous functions of bounded variation (see Kuo 2006, formula (2.3.1)), we get

$$\begin{aligned} G_{j,p}(x) &= \varphi_{j,p}(a+\delta p) \left(1 - \sqrt{\frac{s(a+\delta p)}{s(x)}}\right) W_n(a+\delta p) \\ &\quad - \varphi_{j,p}(a+\delta(p-1)) \left(1 - \sqrt{\frac{s(a+\delta(p-1))}{s(x)}}\right) W_n(a+\delta(p-1)) \\ &\quad - \int_{I_p} W_n(u) \, d\left(\varphi_{j,p}(u)\left(1 - \sqrt{\frac{s(u)}{s(x)}}\right)\right) \\ &\leq C_9 V_{\{I_p\}} \left(\varphi_{j,p}(\cdot)\left(1 - \sqrt{s(\cdot)/s(x)}\right)\right) \cdot \sup_{I_p} |W_n(u)|, \end{aligned}$$

where $C_9 > 0$ and $V_{\{I_p\}}(g)$ is the total variation of a function $g(\cdot)$. Note that for any $u_1, u_2, x \in I_p$

$$\begin{aligned} &\left| \varphi_{j,p}(u_2) \left(1 - \sqrt{\frac{s(u_2)}{s(x)}}\right) - \varphi_{j,p}(u_1) \left(1 - \sqrt{\frac{s(u_1)}{s(x)}}\right) \right| \\ &\leq \left|\varphi_{j,p}(u_2)\right| \frac{\left|\sqrt{s(u_2)} - \sqrt{s(u_1)}\right|}{\sqrt{s(x)}} + \left|\varphi_{j,p}(u_2) - \varphi_{j,p}(u_1)\right| \frac{\left|\sqrt{s(x)} - \sqrt{s(u_1)}\right|}{\sqrt{s(x)}}, \end{aligned}$$

where

$$\left|\sqrt{s(u_2)} - \sqrt{s(u_1)}\right| = \frac{|s(u_2) - s(u_1)|}{\sqrt{s(u_2)} + \sqrt{s(u_1)}} \leq C_{10} |u_2 - u_1|, \quad \text{with } C_{10} > 0,$$

because the function $s(\cdot)$ is assumed to be Lipschitz. Taking into account (6), we conclude that

$$V_{\{I_p\}} \left(\varphi_{j,p}(\cdot)\left(1 - \sqrt{s(\cdot)/s(x)}\right)\right) \leq C_{11} \left(\sup_{x \in I_p} \left|\varphi_{j,p}(x)\right| + V_{\{I_p\}}\left(\varphi_{j,p}\right)\right) |I_p| \leq \frac{C_{12}}{\sqrt{m}}$$

with some $C_{11}, C_{12} > 0$. Therefore

$$\sup_{I_p} |G_{j,p}(x)| \leq \frac{C_9 C_{12}}{\sqrt{m}} \sup_{D} |W_n(u)|, \qquad \text{with } C_9 > 0,$$

and Eq. 51 follows. Note that

$$\mathbb{P}\left\{\sup_{x\in D}|W_n(x)|>u\right\}\le\mathbb{P}\left\{\sup_{x\in D}W_n(x)>u\right\}+\mathbb{P}\left\{\inf_{x\in D}W_n(x)<-u\right\}$$
$$\le 2\mathbb{P}\left\{\sup_{x\in[0,b]}W_n(x)>u\right\}=4\mathbb{P}\left\{W_n(b)>u\right\},$$

see p. 105 from Revuz and Yor (1999).

10. $\sup_{x\in D}\mathscr{L}_8(x)\rightsquigarrow$ maximum of random variables. The function $\mathscr{L}_8(x)$ is equal to

$$\mathscr{L}_8(x)=-\sqrt{\frac{s(x)}{T}}\sum_{p=1}^{m}\sum_{j=0}^{J}Z_{j,p}\psi_j^m(x-\delta(p-1))\cdot\mathbb{I}\left\{x\in I_p\right\},$$

where the r.v.'s $Z_{j,p}:=\int_{I_p}\psi_j^m(u-\delta(p-1))dW_n(u)$ have standard normal distribution. Therefore,

$$\sqrt{T}\cdot\sup_{x\in D}\left\{\frac{|\mathscr{L}_8(x)|}{\sqrt{s(x)}}\right\}=\max_{p=1..m}\left[\sup_{x\in I_1}\left|\sum_{j=0}^{J}Z_{j,p}\psi_j^m(x)\right|\right]=\max\left\{\zeta_1,\ldots,\zeta_m\right\},\tag{52}$$

where $\zeta_1,..\zeta_m$ are independent copies of the random variable

$$\zeta=\zeta^{J,m}=\sup_{x\in I_1}\left|\sum_{j=0}^{J}Z_j\psi_j^m(x)\right|$$

with i.i.d. standard normal r.v.'s $Z_j,\ j=0..J$.

11. Last step. To complete the proof, we need the following technical lemma.

Lemma A1 *Let $\eta_1,\ldots,\eta_k$ be random variables such that*

$$\mathbb{P}\Big\{|\eta_{i+1}-\eta_i|\le\delta_i\Big\}\ge 1-\gamma_i,\quad i=1..(k-1),$$

for some non-negative $\delta_i,\gamma_i,\ i=1..k$. Denote by F_{η_k} the distribution function of η_k. Then it holds

$$F_{\eta_1}\left(x-\sum_{j=1}^{k-1}\delta_j\right)-\sum_{j=1}^{k-1}\gamma_j\le F_{\eta_k}(x)\le F_{\eta_1}\left(x+\sum_{j=1}^{k-1}\delta_j\right)+\sum_{j=1}^{k-1}\gamma_j.\tag{53}$$

Proof The proof is given in the preprint (Konakov and Panov 2016), where this statement is formulated as Lemma A.1. □

Returning to the proof of Proposition 1, we apply Lemma A1 with

$$\eta_k := \sqrt{\frac{T}{m}} \sup_{x\in D} \left\{ \frac{|\mathscr{L}_k(x)|}{\sqrt{s(x)}} \right\}, \qquad k = 1..7,$$
$$\eta_8 := \frac{1}{\sqrt{m}} \max_{p=1..m} \zeta_p = \sqrt{\frac{T}{m}} \sup_{x\in D} \left\{ \frac{|\mathscr{L}_8(x)|}{\sqrt{s(x)}} \right\},$$

where the last equality follows from Eq. 52. Note that for all $k = 2, .., 8$,

$$|\eta_k - \eta_{k-1}| \le \sqrt{\frac{T}{m}} \sup_{x\in D} \left| \frac{|\mathscr{L}_k(x)| - |\mathscr{L}_{k-1}(x)|}{\sqrt{s(x)}} \right|$$
$$\le \sqrt{\frac{T}{m}} \sup_{x\in D} \left\{ \frac{1}{\sqrt{s(x)}} \right\} \cdot \sup_{x\in D} \left| \mathscr{L}_k(x) - \mathscr{L}_{k-1}(x) \right|.$$

Using the results obtaining on the previous steps of the proof (and changing for simplicity the indexes for constants), we get

$$\delta_1 = C_3 \frac{\sqrt{m}\log n}{\sqrt{T}}, \qquad \gamma_1 = K/n^{\lambda},$$
$$\delta_3 = \sqrt{\frac{T}{m}} \left(C_4 \frac{Tm}{n^{3/2}} + C_5(b-a)\frac{1}{\sqrt{n}} \right) q_n^{(1)}, \qquad \gamma_3 = 2(1 - \Phi(q_n^{(1)})),$$
$$\delta_5 = \sqrt{\frac{T}{m}} \varepsilon, \qquad \gamma_5 = C_6 C_7 \frac{m}{\varepsilon} \sqrt{\frac{1}{n} \log\left(\frac{n}{T}\right)},$$
$$\delta_7 = C_8 \frac{1}{\sqrt{m}} q_n^{(2)}, \qquad \gamma_7 = 4(1 - \Phi(q_n^{(2)}/\sqrt{b})),$$

where the sequences $q_n^{(1)}, q_n^{(2)}$ are tending to ∞ as $n \to \infty$, and all other δ's and γ's are equal to 0.

Let us fix $\varepsilon, q_n^{(1)}, q_n^{(2)}$ such that $\sum_{i=1}^{7} \gamma_i \lesssim n^{-\lambda}$ as $n \to \infty$. More precisely, we take $\varepsilon = C_6 C_7 K^{-1} m n^{\lambda - 1/2} \sqrt{\log(n/T)}$, and motivated by the inequality

$$1 - \Phi(x) \le \frac{1}{x\sqrt{2\pi}} e^{-x^2/2}, \qquad \forall\, x > 0, \tag{54}$$

see, e.g., p.2 in Michna (2009), we choose $q_n^{(1)} = \sqrt{2\lambda \log n}$, and $q_n^{(2)} = \sqrt{2\lambda b \log n}$. Next, we set $T = n^{\varkappa}$. Note that the condition $\varkappa < 1$ guarantees that $\Delta = T/n \to 0$ as $n \to \infty$.

We take now $\lambda < (1-\varkappa)/2$ (see step 2), and fix

$$\theta = \min\{\varkappa, (1-\varkappa) - 2\lambda\} > 0.$$

Assuming that $m \gtrsim n^{\theta/2}$, we conclude that

$$\begin{aligned}
\delta_7 &= \bar{c}_1 \frac{\sqrt{\log n}}{\sqrt{m}} \lesssim \frac{\sqrt{m}\sqrt{\log n}}{n^{\theta/2}} =: \Lambda_{n,m}, \\
\delta_1 &= \bar{c}_2 \frac{\sqrt{m}\log n}{n^{\varkappa/2}} \lesssim \Lambda_{n,m}, \qquad \text{since } \varkappa/2 > \theta/2, \\
\delta_5 &= \bar{c}_3 \frac{\sqrt{m}\cdot\sqrt{\log n}}{n^{1/2-\lambda-\varkappa/2}} \lesssim \Lambda_{n,m}, \qquad \text{since } 1/2-\lambda-\varkappa/2 > \theta/2, \\
\delta_3 &= \bar{c}_4 \frac{\sqrt{m}\cdot\sqrt{\log n}}{n^{3/2(1-\varkappa)}} + \bar{c}_5 \frac{\sqrt{\log n}}{\sqrt{m}\cdot n^{(1-\varkappa)/2}} \lesssim \Lambda_{n,m}, \\
&\qquad \text{since} \quad 3/2(1-\varkappa) > (1-\varkappa)/2 - \lambda > \theta/2 \quad \text{and} \quad \sqrt{\log n/(mn^{1-\varkappa})} \lesssim \delta_7 \lesssim \Lambda_{n,m}.
\end{aligned}$$

with $\bar{c}_k > 0, k = 1..5$. Therefore, $\sum_{i=1}^{7} \delta_i \lesssim \delta_7$ as $n \to \infty$. Applying Lemma A1, we arrive at

$$\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \le x - \tilde{c}_1\Lambda_{n,m}\right\} - \tilde{c}_2 n^{-\lambda} \le \mathbb{P}\left\{\frac{1}{\sqrt{m}} \max_{p=1..m} \zeta_p \le x\right\}$$

$$\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \le x + \tilde{c}_1\Lambda_{n,m}\right\} + \tilde{c}_2 n^{-\lambda} \ge \mathbb{P}\left\{\frac{1}{\sqrt{m}} \max_{p=1..m} \zeta_p \le x\right\}$$

with $\tilde{c}_1, \tilde{c}_2 > 0$. Note that if $m = n^{\gamma}$ with $\gamma \in (\theta/2, \theta)$ then $\Lambda_{n,m} \lesssim n^{-\lambda_1}$ with some $\lambda_1 > 0$. We arrive at the required result (14)–(15) with $\lambda_2 := \lambda$.

B Proof of Theorem 2

From Eq. 14 we get that for any u,

$$\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \le u\right\} \le \left[\breve{F}\left(\sqrt{m}\left(u + c_1 n^{-\lambda_1}\right)\right)\right]^m + c_2 n^{-\lambda_2}, \tag{55}$$

where $\breve{F}(\cdot)$ is the distribution function of the random variable $\zeta^{J,m}$. Substituting $u = u_m - c_1 n^{-\lambda_1}$, we get in the right-hand side of Eq. 55

$$\left[\breve{F}\left(\sqrt{m}u_m\right)\right]^m = e^{m\cdot\log(1-(1-\breve{F}(\sqrt{m}u_m))} = e^{W_m} H_m, \quad m \to \infty,$$

where $W_m := -m \cdot \mathbb{P}\left\{\zeta^{J,m} \ge \sqrt{m}u_m\right\}$, and $H_m := \exp\{-(W_m^2/(2m))(1 + o(1))\}$, $m \to \infty$. Formula (19) yields

$$W_m = -2\frac{h_1\, m}{u_m^k} \exp\left\{-h_2 u_m^2\right\}(1 + \breve{\tau}(u_m)).$$

Since $a_m \to \infty$, $b_m/c_m \to \infty$ as $m \to \infty$, we get

$$W_m = -2\frac{h_1\, m}{b_m^k} \exp\left\{-h_2\cdot\left(2y\frac{b_m}{a_m} + b_m^2 - 2c_m\right)\right\}(1 + R(m)),$$

because

$$\left(1+\frac{y}{a_m b_m}-\frac{c_m}{b_m^2}\right)^{-k}\exp\left\{-h_2\left(\frac{y^2}{a_m^2}-2y\frac{c_m}{a_m b_m}+\frac{c_m^2}{b_m^2}\right)\right\}$$
$$=\left[1+\frac{k^2}{2h_2}\frac{\log b_m}{b_m^2}(1+o(1))\right]\cdot\left[1-\frac{k^2}{4h_2}\left(\frac{\log b_m}{b_m}\right)^2(1+o(1))\right]$$
$$=1-\frac{k^2}{16}\frac{(\log\log m)^2}{\log m}(1+o(1)).$$

Therefore, we get the asympotics for W_m:

$$\begin{aligned} W_m &= -2\exp\left\{-2h_2 y\frac{b_m}{a_m}\right\}\cdot\frac{h_1\, m}{\exp\left\{h_2 b_m^2\right\}}\cdot\frac{\exp\{2h_2 c_m\}}{b_m^k}(1+R(m)) \\ &= -2e^{-y}(1+R(m)). \end{aligned} \tag{56}$$

Note also that Eq. 56 yields $H_m = 1 + l(y)m^{-1}(1+o(1)),\ m\to\infty$ with $l(y) = -2e^{-2y}$. Finally, we get that the following inequality is fulfilled uniformly over a compact subset in $y\in\mathbb{R}$,

$$\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \le u_m(y) - c_1 n^{-\lambda_1}\right\} \le e^{-2e^{-y}}\left(1-2e^{-y}R(m)\right). \tag{57}$$

Note that we do not distinguish $R(m)$ and $R(m)(1+o(1)),\ m\to\infty$, and therefore the equalities like $R(m) = R(m) + cn^{-\lambda}$ are possible. Due to our choice of the sequence $u_m(y)$, we get

$$u_m\big(y\big) - c_1 n^{-\lambda_1} = u_m\big(y - 2c_1h_2\cdot n^{-\lambda_1}b_m\big), \tag{58}$$

where $n^{-\lambda_1}b_m \lesssim R(m)$ as $n, m\to\infty$. Since (57) is fulfilled uniformly over the compact sets, we can apply this inequality with $y - 2c_1h_2\cdot n^{-\lambda_1}b_m$ instead of y. Finally, we arrive at

$$\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \le u_m(y)\right\} \le e^{-2e^{-y}}\left(1-2e^{-y}R(m)\right). \tag{59}$$

Analogously to Eqs. 57 and 59, we derive from Eq. 15 that

$$\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \le u_m(y) + c_1 n^{-\lambda_1}\right\} \ge e^{-2e^{-y}}\left(1-2e^{-y}R(m)\right), \tag{60}$$

and

$$\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \le u_m(y)\right\} \ge e^{-2e^{-y}}\left(1-2e^{-y}R(m)\right). \tag{61}$$

Joint consideration of Eqs. 59 and 61 completes the proof.

C Proof of Theorem 3

First note that for any $x \in \mathbb{R}$

$$\left|\hat{s}_n(x) - s(x)\right| \le \left|\hat{s}_n(x) - \mathbb{E}\hat{s}_n(x)\right| + \left|\mathbb{E}\hat{s}_n(x) - \tilde{s}(x)\right| + |\tilde{s}(x) - s(x)|. \qquad (62)$$

Below we separately consider the second and the third terms in the right-hand side of Eq. 62, see steps 1 and 2 below. Afterwards, we combine our results and arrive at Eq. 23, see step 3.

1. Consider the difference

$$\begin{aligned}\mathbb{E}\hat{s}_n(x) - \tilde{s}(x) &= \sum_{r=1}^{d}\left[\frac{1}{\Delta}\mathbb{E}\left(\varphi_r(X_\Delta)\right) - \int_a^b \varphi_r(u)s(u)du\right]\varphi_r(x)\\ &= \sum_{p=1}^{m}\sum_{j=0}^{J}\left[\frac{1}{\Delta}\mathbb{E}\left(\varphi_{j,p}(X_\Delta)\right) - \int_{I_p}\varphi_{j,p}(u)s(u)du\right]\varphi_{j,p}(x),\end{aligned}$$

where $\varphi_{j,p}(x) = \psi_j^m\left(x - \delta(p-1)\right)\mathbb{I}\left\{x \in I_p\right\},\ j = 0..J,\ \ p = 1..m.$

Note that since $\varphi_{j,p}(\cdot)$ is a function of bounded variation on the compact interval I_p,

$$\mathbb{E}\left[\varphi_{j,p}(X_\Delta)\right] = \varphi_{j,p}(c)\ \mathbb{P}\left\{X_\Delta \ge c\right\} + \int_{I_p}\mathbb{P}\left\{X_\Delta \ge u\right\} d\left(\varphi_{j,p}(u)\right),$$

$$\int_{I_p}\varphi_{j,p}(u)s(u)du = \varphi_{j,p}(c)\ \nu\left([c,\infty)\right) + \int_{I_p}\nu\left([u,\infty)\right) d\left(\varphi_{j,p}(u)\right),$$

where $c = a + \delta(p-1)$. Therefore

$$\left|\frac{1}{\Delta}\mathbb{E}\left[\varphi_{j,p}(X_\Delta)\right] - \int_{I_p}\varphi_{j,p}(u)s(u)du\right| \le \left(\left|\psi_j^m(a)\right| + V_{\{I_1\}}\left(\psi_j^m\right)\right) M_\Delta(I_p),$$

where

$$M_\Delta(I_p) := \sup_{y\in I_p}\left|\frac{1}{\Delta}\mathbb{P}\left\{X_\Delta > y\right\} - \nu\left([y,+\infty)\right)\right|.$$

Applying the small-time asymptotic property (10), we conclude that $M_\Delta(I_p) \le M_\Delta([a,b]) \le q\Delta$, and therefore

$$\sup_{x\in D}\left|\mathbb{E}\hat{s}_n(x) - \tilde{s}(x)\right| \le q n^{\varkappa-1}\cdot\sum_{j=0}^{J}\left(|\psi_j^m(a)| + V_{\{I_1\}}\left(\psi_j^m\right)\right)\sup_{x\in I_1}\left|\psi_j^m(x)\right|.$$

Next, using the conditions (6), we arrive at

$$\sup_{x\in D}\left|\mathbb{E}\hat{s}_n(x) - \tilde{s}(x)\right| \le c_1 n^{\varkappa-1} m, \qquad \text{with } c_1 > 0. \qquad (63)$$

2. Next, we consider the difference

$$\tilde{s}(x) - s(x) = \sum_{p=1}^{m}\sum_{j=0}^{J}\left[\int_{I_p}\varphi_{j,p}(y)s(y)dy\cdot\varphi_{j,p}(x)\right] - s(x).$$

Since in all considered cases $\varphi_{0,p}(x)dx = 1/\sqrt{\delta} \cdot \mathbb{I}\left\{x \in I_p\right\}, \;\; p = 1..m$, we get

$$1 = \sum_{p=1}^{m} \int_{I_p} \varphi_{0,p}(y)dy \cdot \varphi_{0,p}(x), \qquad \forall\, x \in [a,b],$$

and moreover $\int_{I_p} \varphi_{j,p}(y)dy = 0$ for all $j = 1..J, \;\; p = 1..m$. Therefore,

$$\tilde{s}(x) - s(x) = \sum_{p=1}^{m} \sum_{j=0}^{J} \left[\int_{I_p} \varphi_{j,p}(y)\,(s(y) - s(x))\,dy \cdot \varphi_{j,p}(x) \right].$$

Applying the Cauchy-Schwarz inequality for the second summation, we get

$$|\tilde{s}(x) - s(x)| \le \sum_{p=1}^{m} \left(\sum_{j=0}^{J} \left(\int_{I_p} \varphi_{j,p}(y)\,(s(y) - s(x))\,dy \right)^2 \right)^{1/2} \cdot \left(\sum_{j=0}^{J} \varphi_{j,p}^2(x) \right)^{1/2}. \quad (64)$$

Next, we apply the Cauchy-Schwarz inequality for the integral in Eq. 64:

$$|\tilde{s}(x) - s(x)| \le \sum_{p=1}^{m} \left(\int_{I_p} (s(y) - s(x))^2\,dy \cdot \sum_{j=0}^{J} \int_{I_p} \varphi_{j,p}^2(y)dy \right)^{1/2} \cdot \left(\sum_{j=0}^{J} \varphi_{j,p}^2(x) \right)^{1/2}. \quad (65)$$

Note that $\int_{I_p} \varphi_{j,p}^2(y)dy = 1, \;\forall\, j,\, p$, and $\sum_{j=0}^{J} \varphi_{j,p}^2(x) \le \mathcal{C}_1^2 (J+1)\delta^{-1} \cdot \mathbb{I}\{x \in I_p\}$ due to the assumption (6). Furthermore, since the function $s(\cdot)$ is Lipschitz,

$$\int_{I_p} (s(y) - s(x))^2\,dy \le H^2 \delta^3, \qquad \forall\, x \in I_p,$$

where H is a Lipschitz constant. Finally we conclude that

$$|\tilde{s}(x) - s(x)| \le c_2 \delta = c_2 (b-a) m^{-1}, \qquad \text{with } c_2 > 0.$$

3. Using the upper bounds for the second and the third term in Eq. 62, we get

$$\sqrt{\frac{T}{m}} \mathscr{D}_n \le \sqrt{\frac{T}{m}} \mathscr{Z}_n + \check{c} \sqrt{\frac{T}{m}} \max\left\{ n^{\varkappa - 1} m, m^{-1} \right\} = \sqrt{\frac{T}{m}} \mathscr{Z}_n + \check{c}\, \mathcal{G}_{n,m},$$

where $\check{c} = \max\{c_1, c_2(b-a)\}$ and $\mathcal{G}_{n,m} := n^{\varkappa/2} m^{1/2} \max\left\{ n^{\varkappa-1}, m^{-2} \right\}$. On the other side, similar to Eq. 62, we obtain

$$\left|\hat{s}_n(x) - s(x)\right| \ge \left|\hat{s}_n(x) - \mathbb{E}\hat{s}_n(x)\right| - \left|\mathbb{E}\hat{s}_n(x) - \tilde{s}(x)\right| - \left|\tilde{s}(x) - s(x)\right|,$$

and therefore

$$\sqrt{\frac{T}{m}} \mathscr{D}_n \ge \sqrt{\frac{T}{m}} \mathscr{Z}_n - \check{c} \sqrt{\frac{T}{m}} \max\left\{ n^{\varkappa - 1} m, m^{-1} \right\} = \sqrt{\frac{T}{m}} \mathscr{Z}_n - \check{c}\, \mathcal{G}_{n,m}.$$

So, we have proved that

$$\mathbb{P}\left\{ \left| \sqrt{\frac{T}{m}} \mathscr{D}_n - \sqrt{\frac{T}{m}} \mathscr{Z}_n \right| \le \check{c}\, \mathcal{G}_{n,m} \right\} = 1.$$

By Lemma A1, for any $x \in \mathbb{R}$,

$$\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \le x - \check{c}\,\mathcal{G}_{n,m}\right\} \le \mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{D}_n \le x\right\} \le \mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \le x + \check{c}\,\mathcal{G}_{n,m}\right\}. \quad (66)$$

Substituting $x = u_m$ with $u_m = u_m(y) = y/a_m + (b_m - c_m/b_m)$ (see Eq. 20), we get that the left-hand side in Eq. 66 is equal to

$$\begin{aligned}\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \le u_m(y) - \check{c}\,\mathcal{G}_{n,m}\right\} &= \mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \le u_m\big(y - \check{c}\,\mathcal{G}_{n,m}a_m\big)\right\}\\ &= e^{-2e^{-y}}\left(1 - 2e^{-y}R(m)\right),\end{aligned}$$

where we use the fact that according to Eq. 24

$$\mathcal{G}_{n,m}a_m = 2\sqrt{h_2}n^{(\varkappa+\gamma)/2-\min\{1-\varkappa,2\gamma\}}\sqrt{\log(h_1 m)} \lesssim R(m), \qquad n, m \to \infty \quad (67)$$

The same argument holds for the right-hand side of Eq. 66, and the desired result follows.

D Proof of Theorem 4

The main idea of the proof is to show that the "random part" of $\mathscr{D}_n$ satisfies

$$\sup_{y\in\mathbb{R}}\left|\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \le u_m(y)\right\} - A_m(y)\right| \le C_0 n^{-\beta_0}, \quad (68)$$

with $C_0 > 0$, $\beta_0 > 0$ (see steps 1 and 2) and afterwards to study the entire deviation $\mathscr{D}_n$ (step 3).

1. From the proof of Theorem 2 (see Appendix B), we know that

$$\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \le u_m(y) - c_1 n^{-\lambda_1}\right\} \le e^{W_m}\left(1 + \frac{l(y)}{m}(1 + o(1))\right) + c_2 n^{-\lambda_2}, \quad (69)$$

where $l(y) = -2e^{-2y}$ and

$$W_m = -m\mathbb{P}\left\{\zeta^{J,m} \ge \sqrt{m}u_m\right\}, \qquad u_m(y) = b_m + \frac{y}{2h_2 b_m}.$$

Using (29) and notations of Corollary 1, we get

$$W_m = -2h_1 m \exp\left\{-h_2 u_m^2\right\} - 2m\left(1 - \Phi\left(u_m\sqrt{2h_2}\right)\right) + R,$$

where

$$|R| \le m \cdot \exp\left\{-(1+\chi)h_2 u_m^2\right\}. \quad (70)$$

By the definition of u_m,

$$h_1 m \exp\left\{-h_2 u_m^2\right\} = \exp\left\{-y - \frac{y^2}{4\log(h_1 m)}\right\}. \quad (71)$$

Note that for $y \geq -b_m^{3/2}$, for any $\varepsilon > 0$ and sufficiently large m

$$u_m = b_m + \frac{y}{2h_2 b_m} \geq b_m - \frac{1}{2h_2}\sqrt{b_m} \geq (1-\varepsilon) b_m. \tag{72}$$

It follows from Eq. 70 that for sufficiently small $\varepsilon > 0$ and sufficiently large m,

$$\begin{aligned} |R| &\leq m \exp\left\{-(1+\chi) h_2 (1-\varepsilon)^2 b_m^2\right\} \\ &= m \exp\left\{-(1+\chi)(1-\varepsilon)^2 \log(h_1 m)\right\} \leq C_1 m^{-\beta_1}, \qquad C_1 > 0,\ \beta_1 > 0. \end{aligned}$$

Therefore, $e^R = 1 + \Theta_m$, where $\Theta_m \lesssim m^{-\beta_1}$. Finally, we get from Eq. 69 for any $y \geq -b_m^{3/2}$

$$\begin{aligned} &\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \leq u_m - c_1 n^{-\lambda_1}\right\} - A_m(y) \\ &\quad \leq A_m(y)\Theta_m + A_m(y)(1+\Theta_m)\frac{l(y)}{m}(1+o(1)) \leq C_2 m^{-\beta_2}, \qquad m \to \infty \end{aligned}$$

with $C_2 > 0$, $\beta_2 > 0$, $l(y) \asymp e^{-2y}$. Substituting $y + c_1 n^{-\lambda_1}$ instead of y, and taking into account that $m = n^{\gamma}$, we get for n large enough

$$\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \leq u_m\right\} - A_m(y) \leq C_3 n^{-\beta_3} \qquad \text{with } C_3 > 0,\ \beta_3 > 0,$$

because $A_m\left(y + c_1 n^{-\lambda_1}\right) = A_m(y) + C_4 n^{-\beta_4}(1 + o(1))$ with some $C_4, \beta_4 > 0$. Analogously, from the inequality

$$\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \leq u_m + c_1 n^{-\lambda_1}\right\} \geq e^{W_m}\left(1 + \frac{l(y)}{m}(1+o(1))\right) - c_2 n^{-\lambda_2},$$

one can derive that

$$\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \leq u_m\right\} - A_m(y) \geq -C_3 n^{-\beta_3}$$

holds for any n large enough.

2. For $y < -b_m^{3/2}$, we obtain

$$\begin{aligned} W_m &= -m\mathbb{P}\left\{\zeta^{J,m} \geq \sqrt{m}\left(b_m + \frac{y}{2h_2 b_m}\right)\right\} \\ &\leq -m\mathbb{P}\left\{\zeta^{J,m} > \sqrt{m}\left(b_m - \frac{1}{2h_2}\sqrt{b_m}\right)\right\} \\ &= -2h_1 m \exp\left\{-h_2 b_m^2 + b_m^{3/2} - \frac{b_m}{4h_2}\right\}(1+o(1)) \\ &= -2\exp\left\{b_m^{3/2}\right\}(1+o(1)) \\ &= -2\exp\left\{\frac{(\log(h_1 m))^{3/4}}{h_2^{3/4}}\right\}(1+o(1)). \end{aligned}$$

Therefore, similarly to the first step of the proof, we get with for all $y < -b_m^{3/2}$

$$\begin{aligned} \mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \leq u_m - c_1 n^{-\lambda_1}\right\} &\leq e^{W_m}\left(1 + \frac{l(y)}{m}(1+o(1))\right) + c_2 n^{-\lambda_2} \leq C_5 n^{-\beta_5}, \\ \mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \leq u_m + c_1 n^{-\lambda_1}\right\} &\geq e^{W_m}\left(1 + \frac{l(y)}{m}(1+o(1))\right) - c_2 n^{-\lambda_2} \geq -C_5 n^{-\beta_5}, \end{aligned}$$

with some $C_5,\ \beta_5 > 0$, because

$$e^{W_m} \leq e^{-e^{(\log m)^{3/4} h_2^{-3/4}}} \leq e^{-K\log m} = m^{-K}$$

for any $K > 0$ and m large enough. Finally, we conclude that

$$\sup_{y < -b_m^{3/2}} \left| \mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \leq u_m(y)\right\} - A_m(y)\right| \leq C_6 n^{-\beta_6},$$

with some $C_6,\ \beta_6 > 0$.

3. It follows from (66) that

$$\begin{aligned} \mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{D}_n \leq u_m(y)\right\} &\leq \mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \leq u_m(y + \check{c}\,\mathcal{G}_{n,m} a_m)\right\}, \\ \mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{D}_n \leq u_m(y)\right\} &\geq \mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{Z}_n \leq u_m(y - \check{c}\,\mathcal{G}_{n,m} a_m)\right\}, \end{aligned}$$

where $\mathcal{G}_{n,m}a_m \lesssim n^{-\beta_7}$ (see (67)) with $\beta_7 > 0$. Applying (73) to the right-hand sides of these inequalities, we get

$$\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{D}_n \leq u_m(y)\right\} \leq A_m(y + \check{c}\,\mathcal{G}_{n,m}a_m) + C_0 n^{-\beta_0} \leq A_m(y) + \bar{c}n^{-\beta},$$

$$\mathbb{P}\left\{\sqrt{\frac{T}{m}}\mathscr{D}_n \leq u_m(y)\right\} \geq A_m(y - \check{c}\,\mathcal{G}_{n,m}a_m) - C_0 n^{-\beta_0} \geq A_m(y) - \bar{c}n^{-\beta},$$

where $\bar{c} > 0$, $\beta > 0$. This observation completes the proof of Theorem 4.

References

Belomestny, D.: Statistical inference for time-changed Lévy processes via composite characteristic function estimation. Ann. Stat. **39**(4), 2205–2242 (2011)

Bickel, P., Rosenblatt, M.: On some global measures of the deviations of density function estimates. Ann. Stat. **1**(6), 1071–1095 (1973)

Chen, S.X., Delaigle, A., Hall, P.: Nonparametric estimation for a class of Lévy processes. J. Econ. **157**, 257–271 (2010)

Comte, F., Genon-Catalot, V.: Nonparametric estimation for pure jump Lévy processes based on high frequency data. Stoch. Process. Appl. **119**, 4088–4123 (2009)

Comte, F., Genon-Catalot, V.: Nonparametric estimation for pure jump irregularly sampled or noisy Lévy processes. Statistica Neerlandica **64**, 290–313 (2010a)

Comte, F., Genon-Catalot, V.: Nonparametric adaptive estimation for pure jump Lévy processes. Ann. l'Inst. Henri Poincaré - Prob. Stat. **46**(3), 595–617 (2010b)

Comte, F., Genon-Catalot, V.: Estimation for Lévy processes from high frequency data within a long time interval. Ann. Stat. **39**, 803–837 (2011)

Figueroa-López, J.E.: Nonparametric Estimation of Lévy Processes with a View Towards Mathematical Finance. PhD thesis, Georgia Institute of Technology (2004)

Figueroa-López, J.E.: Sieve-based confidence intervals and bands for Lévy densities. Bernoulli **17**(2), 643–670 (2011)

Fisher, M., Nappo, G.: On the moments of the modulus of continuity of Ito Processes. Stochast. Process. Appl. **28**(1), 103–122 (2010)

Gradshtein, I., Ryzhik I.: Table of integrals, Series and Products. Academic Press (1996)

Gugushvili, S.: Nonparametric inference for discretely sampled Lévy processes. Ann. l'Inst. Henri Poincaré - Probab. Stat. **48**(1), 282–307 (2012)

Hall, P.: On convergences rates of suprema. Probab. Theory Relat. Fields **89**(447-455) (1991)

Komlós, J., Major, P., Tusnády, G.: An approximation of partial sums of independent rv's and the sample DF. Z. Wahrscheinlichkeitstheorie Verw Geb. **32**, 111–131 (1975)

Konakov, V., Piterbarg, V.: On the convergence rates of maximal deviation distribution for kernel regression estimates. J. Multivar. Anal. **15**, 279–294 (1984)

Konakov, V., Panov, V.: Convergence rates of maximal deviation distribution for projection estimates of Lévy densities (2016). arXiv:1411.4750v3

Kuo, H.-H.: Introduction to stochastic integration. Springer (2006)

Michna, Z.: Remarks on Pickands theorem (2009). arXiv:0904.3832v1

Neumann, M., Reiss, M.: Nonparametric estimation for Lévy processes from low-frequency observations. Bernoulli **15**(1), 223–248 (2009)

Nickl, R., Reiss, M.: A Donsker theorem for Lévy measures. J. Funct. Anal. **253**, 3306–3332 (2012)

Piterbarg, V.I.: Asymptotic Methods in the Theory of Gaussian Processes and Fields. AMS, Providence (1996)

Piterbarg, V.I.: Twenty Lectures about Gaussian Processes. Atlantic Financial Press, London (2015)

Revuz, D., Yor, M.: Continuous martingales and Brownian motion. Springer-Verlag (1999)
Sato, K.: Lévy processes and infinitely divisible distributions. Cambridge University Press (1999)
Suetin, P.: Classical orthogonal polynomials (in Russian). Fizmatlit (2005)
Van Es, B., Gugushvili, S., Spreij, P.: A kernel type nonparametric density estimator for decompounding. Bernoulli **13**, 672–694 (2007)
Wörner, J.: Variational sums and power variations: a unifying approach to model selection and estimation in semimartingale models. Stat. Decis. **21**, 47–68 (2003)