Optimal estimation of the supremum and occupation times of a self-similar Lévy process[*]

Jevgenijs Ivanovs and Mark Podolskij

Department of Mathematics, Aarhus University, e-mail: jevgenijs.ivanovs@math.au.dk

Department of Mathematics, The University of Luxembourg, e-mail:
mark.podolskij@uni.lu

Abstract: In this paper we present new theoretical results on optimal estimation of certain random quantities based on high frequency observations of a Lévy process. More specifically, we investigate the asymptotic theory for the conditional mean and conditional median estimators of the supremum/infimum of a linear Brownian motion and a strictly stable Lévy process. Another contribution of our article is the conditional mean estimation of the local time and the occupation time of a linear Brownian motion. We demonstrate that the new estimators are considerably more efficient compared to the classical estimators studied in e.g. [6, 14, 29, 30, 38]. Furthermore, we discuss pre-estimation of the parameters of the underlying models, which is required for practical implementation of the proposed statistics.

MSC2020 subject classifications: Primary 62M05, 62G20, 60F05; secondary 62G15, 60G18, 60G51.
Keywords and phrases: Conditioning to stay positive, local time, Lévy processes, occupation time, optimal estimation, self-similarity, supremum, weak limit theorems.

1. Introduction

During the past decades the increasing availability of high frequency data in economics and finance has led to an immense progress in high frequency statistics. In particular, high frequency functionals of Itô semimartingales have received a great deal of attention in the statistical and probabilistic literature, where the focus has been on estimation of quadratic variation, realised jumps and related (random) quantities. A detailed discussion of numerous high frequency methods and their applications to finance can be found in the monographs [1, 31].

Despite large amount of literature on high frequency statistics, the question of *optimality* has rarely been addressed. To fix ideas we consider a stochastic process $(X_t)_{t\in[0,1]}$ with a known law and an associated random quantity

$Q = F((X_t)_{t \in [0,1]})$, where F is a measurable functional. The major problem of interest is outlined by the following question:

Given observations $(X_{i/n})_{i \in [0:n]}$, what is the optimal estimator of the random variable Q and its asymptotic properties as $n \to \infty$?

Let us stress that we are interested in estimating Q for the underlying $(X_t)_{t \in [0,1]}$ and not just its law.

Of course, the formulated problem is hard to address in full generality. But even for particular model classes the assessment of optimality is far from trivial, which is mainly due to the randomness of Q. Indeed, the classical methods such as minimax theory, Le Cam theory or Cramér-Rao bounds, do not apply in this setting. There are only a few results in the literature that discuss optimality in high frequency statistics. In [21] the authors apply the infinite dimensional version of local asymptotic mixed normality to obtain lower efficiency bounds for estimation of integrated functionals of volatility in the setting of diffusion models with a particular structure. In particular, their result shows that the standard estimator of the quadratic variation, the realised volatility, is indeed asymptotically efficient for the considered class of models. In a later paper [22] similar lower bounds have been obtained in the framework of certain jump diffusions. The paper [38] discusses estimation of occupation times for continuous diffusion models and the authors prove that $n^{3/4}$ is the optimal rate of convergence in the case of Brownian motion (however, they do not discuss efficiency bounds). The articles [3, 4, 5] investigate estimation of integral functionals $Q = \int_0^1 f(X_s) ds$ for various Markovian and non-Markovian models. The main focus here is on deriving error bounds and weak limit theorems for Riemann sum type estimators, which heavily depend on the smoothness of f. In several settings they also prove rate optimality in the case of Brownian motion.

The aim of our paper is to study optimal estimation of extrema, local time and occupation time of certain Lévy processes. Accurate estimation of these random functionals is important for numerous applications. For instance, supremum is a key quantity in insurance, queueing, financial mathematics, optimal stopping and various applied domains such as environmental science where maximal level of pollution is often of interest. It is noted that our theory can also be used in Monte Carlo simulation of extrema via discretization, but this is not our main focus since much better algorithms exist [17]; see also [27] for exact simulation of the supremum of a stable process. These algorithms, however, can not handle, e.g., the diameter of the range of X, whereas our estimators still apply. Accurate estimation of local times is required in a number of statistical methods including estimation of the volatility coefficient in a diffusion model [24], estimation of the skewed Brownian motion [34] and estimation of the reflected fractional Brownian motion [28], just to name a few.

The estimation of the aforementioned random quantities has been studied in several papers. The standard estimator of the supremum of a stochastic process is given by the maximum of its high frequency observations. In the setting of a linear Brownian motion the corresponding non-central limit theorem has been proven in [6]; their result has been later extended in [29] to the class of Lévy

processes satisfying certain regularity assumption. Statistical inference for local times has been investigated in [14, 30], who showed asymptotic mixed normality for kernel type estimators in the framework of continuous SDEs. Finally, [5, 38] discussed the estimation of the occupation time measure via Riemann sums.

In this paper we show that the standard estimators proposed in the literature are indeed rate optimal, but they are not asymptotically efficient. Instead, we consider the conditional mean and conditional median estimators, which turn out to be manageable in some important cases. It is well known that the conditional mean $\mathbb{E}[Q|(X_{i/n})_{i\in[0:n]}]$ is the optimal L^2-predictor when $\mathbb{E}[Q^2] < \infty$. In many cases considered below, however, the random variable Q will not have a finite second moment. Then we use the conditional median estimator $\mathrm{med}[Q|(X_{i/n})_{i\in[0:n]}]$, which is optimal in L^1 sense given that $\mathbb{E}[|Q|] < \infty$. Additionally, we still do consider the conditional mean which is a very natural estimator even when the second moment is infinite. Importantly, it is optimal with respect to the Bregman distance: $D(x,y) = \phi(x) - \phi(y) - \phi'(y)(x-y)$ with ϕ being a strictly convex differentiable function [8]. It is only required here that $\mathbb{E}[|Q|]$ and $\mathbb{E}[|\phi(Q)|]$ are finite. We often have $Q \geq 0$ and $\mathbb{E}[Q^p] < \infty$ for some $p > 1$, and hence we may take $\phi(x) = x^p$ to produce an optimality statement for the conditional mean estimator. Finally, the conditional median is optimal with respect to $D(x,y) = (1_{\{x \geq y\}} - 1/2)(g(x) - g(y))$ for an increasing function g which in our case can be taken as $g(x) = x^p$ for $p > 0$, see [26] and references therein.

In the case of supremum, the conditional mean and median estimators have a rather explicit and simple form, but their performance assessment is not a trivial task. Importantly, self-similarity of X (up to measure change) is the key property when evaluating such estimators and establishing the corresponding weak limit theory. Thus we consider the following two classes of processes: (i) linear Brownian motions and (ii) non-monotone self-similar Lévy processes. In the case of local/occupation time we only work with the class (i) of linear Brownian motions and focus on the conditional mean estimators exclusively, which is dictated by the structure of the problem and the tools currently available. Importantly, our conditional mean estimator of the local time fits the framework of [30] and yields an asymptotically optimal statistic in some large class in the case of continuous SDEs, see Remark 2. We find that our new optimal estimators are considerably more efficient than the standard ones and that they do have narrower confidence intervals. In the case of supremum, this is illustrated by a numerical study. Furthermore, we discuss several modifications of our statistics including pre-estimation of unknown parameters of the underlying model.

This paper is structured as follows. §2 is devoted to the supremum and its conditional mean and median estimators with the corresponding weak limit theory in the case of a self-similar Lévy process with a known law. Here we also treat the case of a linear Brownian motion, and comment on the conditional mean estimator of the range diameter. In §3 we present the conditional mean estimators of the local time and occupation time together with the asymptotic

theory in the case of a linear Brownian motion. Then in §4 we study modified statistics based on pre-estimation of the unknown parameters of the model. In particular, we show that reasonable pre-estimation of the model parameters does not affect the asymptotic theory. Furthermore, the effect of truncation of the potentially infinite product involved in the construction of the supremum estimators is discussed, and some comments concerning a general Lévy process are given. Numerical illustrations for the case of supremum are presented in §5, where both a linear Brownian motion and a one-sided stable processes are considered. The proofs are collected in Appendix A and Appendix B for the supremum and local/occupation time, respectively. The former also requires some additional theory for Lévy processes conditioned to stay positive which is given in Appendix C.

2. Optimal estimation of supremum for a self-similar Lévy process

In this section we assume that $(X_t)_{t\geq 0}$ is a non-monotone $1/\alpha$-self-similar Lévy process, i.e.

$$(X_{ut})_{t\geq 0} \stackrel{d}{=} u^{1/\alpha}(X_t)_{t\geq 0} \qquad \text{for all } u > 0,$$

where necessarily $\alpha \in (0,2]$. Assuming that the law of X (or its parameters) is known, we focus on optimal estimation of the supremum and infimum of X on the interval $[0,1]$ from high-frequency observations. The case $\alpha \in (0,2)$ corresponds to a strictly α-stable process, whereas for $\alpha = 2$ we have a scaled Brownian motion, and the respective simplified expressions for the statistics and their limits can be found in §2.4. In fact, §2.4 considers a more general setting of a linear Brownian motion, which is not self-similar but becomes such under Girsanov change of measure. Some further results concerning estimation of infimum and the range diameter are given in §2.5.

We introduce the notation

$$\overline{X}_t := \sup_{s\leq t} X_s \qquad \text{and} \qquad \underline{X}_t := \inf_{s\leq t} X_s$$

to denote the running supremum and infimum process, respectively. Furthermore, the time of supremum will often be needed, and thus we define

$$\tau_t := \inf\{s \in (0,t] : X_{s-} \vee X_s = \overline{X}_t\}.$$

The standard distribution free estimator of $\overline{X}_1$ is given by the empirical maximum of the observed data:

$$M_n := \max_{i\in[0:n]} X_{i/n}.$$

We remark, however, that M_n is always downward biased. Finally, estimation of the infimum amounts to estimation of the supremum of $-X$, and thus no additional theory is needed. The joint estimation of supremum and infimum is discussed in §2.5.

In the following we will often use the notion of *stable convergence*. We recall that a sequence of random variables $(Y_n)_{n\in\mathbb{N}}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ is said to converge stably with limit Y $(Y_n \xrightarrow{d_{st}} Y)$ defined on an extension $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$ of the original probability space $(\Omega, \mathcal{F}, \mathbb{P})$, iff for any bounded, continuous function g and any bounded $\mathcal{F}$-measurable random variable Z it holds that

$$\mathbb{E}[g(Y_n)Z] \to \overline{\mathbb{E}}[g(Y)Z], \quad \text{as } n \to \infty.$$

The notion of stable convergence is due to Rényi [39]. We also refer to [2] for properties of this mode of convergence.

2.1. Preliminaries

We will now review the asymptotic theory for the estimator M_n, which will be useful for studying conditional mean and median estimators. In order to state the limit theorem for M_n, we need to introduce an auxiliary process $(\xi_t)_{t\in\mathbb{R}}$. It is defined as the following functional weak limit:

$$(\overline{X}_T - X_{\tau_T+t})_{t\in\mathbb{R}} \xrightarrow{d} (\xi_t)_{t\in\mathbb{R}} \qquad \text{as } T \to \infty, \tag{1}$$

where $T > 0$ is a deterministic time horizon, see [9]. Here and in the following it is tacitly assumed that the left hand side is ∞ when $\tau_T + t \notin [0, T]$. The functional convergence is always with respect to the Skorokhod J_1 topology, unless specified otherwise. It may be useful to think of ξ as the process X seen from its supremum as the time horizon tends to infinity.

It is well known that $(\xi_t)_{t\geq 0}$ and $(\xi_{(-t)-})_{t\geq 0}$ are independent finite Feller processes starting at 0. Various representations of these processes exist and a number of important properties have been established, see e.g. [19] and references therein. The latter process when started at a positive level is often referred to as X *conditioned to stay positive* (the negative of the former is X *conditioned to stay negative*); here conditioning is understood in a certain limiting sense. The law of the limiting process ξ is not explicit except when X is a Brownian motion and then both parts of ξ are Bessel processes of order 3 scaled by σ, the standard deviation of X_1. In all cases ξ inherits self-similarity from X, and hence both parts (when started from positive values) are positive self-similar Markov processes admitting Lamperti representation studied in detail in [16].

Due to self-similarity of the process X it holds that

$$\xi_t^{(n)} := n^{1/\alpha} \left(\overline{X}_1 - X_{\tau_1+\frac{t}{n}}\right)_{t\in\mathbb{R}} \xrightarrow{d} (\xi_t)_{t\in\mathbb{R}} \qquad \text{as } n \to \infty, \tag{2}$$

where again $\xi_t^{(n)} = \infty$ when $\tau_1 + \frac{t}{n} \notin [0, 1]$. In other words, the process ξ arises from zooming-in on X at its supremum point. We refer the reader to [6, 29] for the case of a linear Brownian motion and a general Lévy process, respectively.

The following result is an instructive application of the convergence in (2). It is a particular case of [29, Thm. 5] extending the result of [6] for Brownian motion.

Theorem 1. *For a non-monotone $1/\alpha$-self-similar Lévy process X we obtain the stable convergence as $n \to \infty$:*

$$V^{(n)} := n^{1/\alpha}(\overline{X}_1 - M_n) \xrightarrow{d_{st}} V := \min_{j \in \mathbb{Z}} \xi_{j+U} \tag{3}$$

where ξ and the standard uniform random variable U are mutually independent, and independent of $\mathcal{F}$.

Let us mention the underlying intuition, which will be important to understand our main result in Theorem 2 given below. Note the identity

$$n^{1/\alpha}(\overline{X}_1 - M_n) = \min_{j \in \mathbb{Z}} \xi^{(n)}_{j+\{n\tau_1\}}$$

where $\{x\}$ stands for the fractional part of x. The random time τ_1 has a density [18] and thus according to [31, 33]

$$\{n\tau_1\} \xrightarrow{d_{st}} U,$$

which together with (2) hint at (3). It is noted that the convergence in (2) is, in fact, stable with ξ being independent of $\mathcal{F}$. Intuitively, zooming-in at the supremum makes the values of X at some fixed times irrelevant. We stress that this only provides intuition and the proof is far from being complete, see [29] and also [13] providing the necessary corrections.

2.2. Optimal estimators

Let us proceed to construct our optimal estimators given by the conditional mean and median. For this purpose we introduce the conditional distribution of $\overline{X}_1$ given the terminal value X_1 via

$$F(x, y) := \mathbb{P}(\overline{X}_1 \leq x | X_1 = y).$$

We choose a version continuous in y which is, in fact, jointly continuous in (x, y) as will be shown in Lemma 3 below. By self-similarity we also have

$$F_{1/n}(x, y) := \mathbb{P}(\overline{X}_{1/n} \leq x | X_{1/n} = y) = F(n^{1/\alpha}x, n^{1/\alpha}y).$$

Next, consider the conditional distribution of $\overline{X}_1 - M_n$ given the observations:

$$
\begin{aligned}
H_n(x) &:= \mathbb{P}\left(\overline{X}_1 - M_n \leq x | X_{j/n}, j \in [1:n]\right) \\
&= \prod_{j=0}^{n-1} F_{1/n}\left(x + M_n - X_{j/n}, X_{\frac{j+1}{n}} - X_{\frac{j}{n}}\right) \\
&= \prod_{j=0}^{n-1} F\left(n^{1/\alpha}(x + \Delta_j^n), n^{1/\alpha}(\Delta_j^n - \Delta_{j+1}^n)\right) \qquad \text{for all } x \geq 0,
\end{aligned}
$$

where $\Delta_j^n := M_n - X_{j/n}$ and the second line follows from the stationarity and independence of increments. We note that $H_n(x)$ is continuous and strictly increasing in $x \geq 0$. Finally, we introduce the conditional mean and conditional median estimators of $\overline{X}_1$:

$$\overline{T}_n^{\mathrm{mean}} := \mathbb{E}[\overline{X}_1 | X_{j/n}, j \in [1:n]] = M_n + \int_0^\infty (1 - H_n(x))\mathrm{d}x, \qquad (4)$$

$$\overline{T}_n^{\mathrm{med}} := \mathrm{med}[\overline{X}_1 | X_{j/n}, j \in [1:n]] = M_n + H_n^{-1}(1/2), \qquad (5)$$

where in the first line we use the integrated tail formula. Interestingly, $\overline{T}_n^{\mathrm{mean}} < \infty$ even when $\mathbb{E}\overline{X}_1 = \infty$, see Remark 4. When evaluating our statistics defined in (4) and (5) we need access to the function $F(x, y)$. This function, however, is explicit only in the Brownian case analyzed in §2.4 and is semi-explicit in the case of one-sided jumps, see Proposition 4. Thus, in the case of general strictly stable process one needs to assess F numerically, which may necessitate truncation of the product in the definition of H_n. Such modifications are discussed in §4.2.

2.3. Limit theory

We start by noting that H_n is a random probability measure and $H_n \xrightarrow{\mathrm{d}} \delta_{\overline{X}_1}$ $\mathbb{P}$-almost surely, whereas $H_n(xn^{-1/\alpha})$ has a non-trivial limit. Observe that $\xi_{j+\{n\tau_1\}}^{(n)}$ is the rescaled distance of the jth observation following τ_1 from the supremum. Thus

$$H_n(xn^{-1/\alpha}) = \prod_{j \in \mathbb{Z}} F\left(x + \xi_{j+\{n\tau_1\}}^{(n)} - V^{(n)}, \xi_{j+\{n\tau_1\}}^{(n)} - \xi_{j+1+\{n\tau_1\}}^{(n)}\right),$$

where we tacitly assume that the factors with $\xi_\cdot^{(n)} = \infty$ evaluate to 1. In view of Theorem 1 it is intuitive that the limit is

$$H(x) := \prod_{j \in \mathbb{Z}} F\left(x + \xi_{j+U} - V, \xi_{j+U} - \xi_{j+1+U}\right), \qquad (6)$$

where the random quantities U, ξ and V are defined in Theorem 1. By substitution we obtain the identities

$$\overline{T}_n^{\mathrm{mean}} = M_n + n^{-1/\alpha} \int_0^\infty (1 - H_n(n^{-1/\alpha}x))\mathrm{d}x, \qquad (7)$$

$$\overline{T}_n^{\mathrm{med}} = M_n + n^{-1/\alpha} H_n(n^{-1/\alpha}\cdot)^{-1}(1/2), \qquad (8)$$

which suggest the asymptotic behaviour of our estimators defined in (4) and (5). We formalise this in one of our main results:

Theorem 2. *Assume that X is a non-monotone $1/\alpha$-self-similar Lévy process. Then the random function H is continuous and strictly increasing with $H(0) = 0$ and $H(\infty) = 1$ $\mathbb{P}$-a.s. and*

$$\left(n^{1/\alpha}(\overline{X}_1 - M_n), (H_n(xn^{-1/\alpha}))_{x \geq 0}\right) \xrightarrow{d_{st}} (V, (H(x))_{x \geq 0}) \qquad (9)$$

with respect to the uniform topology, where V and $H(x)$ are defined in (3) and (6), respectively. Furthermore, our estimators satisfy

$$n^{1/\alpha}(\overline{X}_1 - \overline{T}_n^{\mathrm{mean}}) \xrightarrow{d_{st}} V - \int_0^\infty (1 - H(x))\mathrm{d}x, \quad \text{when } \alpha \in (1, 2], \quad (10)$$

$$n^{1/\alpha}(\overline{X}_1 - \overline{T}_n^{\mathrm{med}}) \xrightarrow{d_{st}} V - H^{-1}(1/2), \quad (11)$$

where the limit random variables are finite.

It is noted that the proof of this result is far from trivial, since it requires precise understanding of the tail function $1 - F(x, y)$ for large x and the rate of growth of $\xi_t^{(n)}$ as $t \to \infty$ (uniformly in n) among other things. The identities (7) and (8) show that the statistics $\overline{T}_n^{\mathrm{mean}}$ and $\overline{T}_n^{\mathrm{med}}$ are first order equivalent to the standard estimator M_n, and the knowledge of the distribution of X only enters through the $n^{-1/\alpha}$-order term. This fact will prove to be important in Section 4, where the parameters of the law of X will need to be estimated.

Recall that $\mathbb{E}\overline{X}_1^p < \infty$ for $p \in (0, \alpha)$. Moreover, all moments of $\overline{X}_1$ are finite when X is a Brownian motion or a strictly α-stable process with no positive jumps. In the latter cases the conditional mean estimator is optimal in L^2 sense. In the case $\alpha \in (1, 2]$ the conditional median is optimal in L^1 sense and the conditional mean is optimal with respect to the above mentioned Bregman distance $D(x, y) = x^p - y^p - py^{p-1}(x - y)$, where $p \in (1, \alpha)$. Finally, the conditional median is optimal with respect to the loss function $D(x, y) = (1_{\{x \geq y\}} - 1/2)(x^p - y^p)$ for $p \in (0, \alpha)$ and any α.

Interestingly, all the expressions in Theorem 2 stay the same if the process X is replaced by its negative $-X$, see Proposition 3. In particular, in the spectrally-positive case the difference $\overline{X}_1 - \overline{T}_n^{\mathrm{mean}}$ has moments of all orders even though each term has infinite second moment, see also Remark 4 below.

2.4. Linear Brownian motion

Consider a linear Brownian motion X with drift parameter $\mu \in \mathbb{R}$ and scale parameter $\sigma > 0$, which is self-similar (and hence Theorem 2 applies) only when $\mu = 0$. Nevertheless, X can be obtained from a scaled Brownian motion by Girsanov change of measure and, in particular, the conditional distribution $\mathbb{P}(\overline{X}_{1/n} \leq x | X_{1/n} = y)$ does not depend on μ, see §A.4.1. Hence our estimators have exactly the same form as in the case of $\mu = 0$, see §2.2. Furthermore, the conditional distribution function F is explicit in this case and is given by

$$F(x, y) = 1 - \exp\left(-2x(x - y)/\sigma^2\right) \quad \text{for } x > y_+ = y \vee 0,$$

which follows from [42] or earlier sources, see also [15, 1.1.8]. Thus

$$H_n(x) = \prod_{i=0}^{n-1} \left(1 - \exp(-2(x + \Delta_i^n)(x + \Delta_{i+1}^n)n/\sigma^2)\right)$$

and the estimators are then defined by (4) and (5).

Interestingly, also the limit theorem has exactly the same form. The main reason for this is that the limit in (2) does not depend on μ either, see [6]. In the following result we prefer to choose the scaling $\sqrt{n}/\sigma$ rather than $\sqrt{n}$ so that the respective quantities correspond to the standard Brownian motion.

Corollary 1. *For a linear Brownian motion X with drift parameter μ and scale $\sigma > 0$ we have*

$$\frac{\sqrt{n}}{\sigma}(\overline{X}_1 - \overline{T}_n^{\mathrm{mean}}) \xrightarrow{d_{st}} V - \int_0^\infty (1 - H(x))\mathrm{d}x, \tag{12}$$

$$\frac{\sqrt{n}}{\sigma}(\overline{X}_1 - \overline{T}_n^{\mathrm{med}}) \xrightarrow{d_{st}} V - H^{-1}(1/2), \tag{13}$$

where $V = \min_{j \in \mathbb{Z}} \xi_{j+U}$ and

$$H(x) = \prod_{j \in \mathbb{Z}} \left(1 - \exp\left(-2(x + \xi_{j+U} - V)(x + \xi_{j+1+U} - V)\right)\right)$$

with ξ being the two-sided Bessel process of order 3 and U a standard uniform, which are mutually independent and independent of $\mathcal{F}$.

Additionally, we now show that (12) extends to convergence of moments.

Lemma 1. *For a linear Brownian motion X and any $p > 0$ we have*

$$\mathbb{E}\left[\left(\frac{\sqrt{n}}{\sigma}\left|\overline{X}_1 - \overline{T}_n^{\mathrm{mean}}\right|\right)^p\right] \to \mathbb{E}\left[\left|V - \int_0^\infty (1 - H(x))\mathrm{d}x\right|^p\right] < \infty.$$

As a consequence of Lemma 1 we obtain an asymptotic expansion of the mean squared error:

$$\mathbb{E}[(\overline{X}_1 - \overline{T}_n^{\mathrm{mean}})^2] \sim \frac{\sigma^2}{n}\mathbb{E}\left[\left(V - \int_0^\infty (1 - H(x))\mathrm{d}x\right)^2\right].$$

2.5. Joint estimation of supremum and infimum

Consider the process $-X_t$ and the associated conditional mean estimator $\underline{T}_n^{\mathrm{mean}}$ of its supremum $\sup_{t \in [0,1]}(-X_t) = -\underline{X}_1$, which is the negative of the infimum of X. According to Proposition 3 there is the symmetry:

$$(-\underline{X}_1) - \underline{T}_n^{\mathrm{mean}} \overset{d}{=} \overline{X}_1 - \overline{T}_n^{\mathrm{mean}}$$

for all n, and so also the asymptotic theory is the same. Furthermore, we have the following joint convergence (linear Brownian motion included with $\alpha = 2$ and then the limit corresponds to the case $\mu = 0$):

Corollary 2. *For $\alpha \in (1, 2]$ it holds that*

$$n^{1/\alpha}(\overline{X}_1 - \overline{T}_n^{\mathrm{mean}}, -\underline{X}_1 - \underline{T}_n^{\mathrm{mean}}) \xrightarrow{d_{st}} (L, L'),$$

where L' and L are identically distributed, mutually independent, and independent of $\mathcal{F}$. Their common distribution is the limiting law in (10).

This, for example, readily yields the limit result for the conditional mean estimator of the range diameter $\overline{X}_1 - \underline{X}_1$.

3. Optimal estimation of local time and occupation time for a linear Brownian motion

In this section X denotes a linear Brownian motion with drift parameter $\mu \in \mathbb{R}$ and scale $\sigma > 0$, and $(L_t(x))_{t \geq 0}$ denotes the corresponding local time process at the level $x \in \mathbb{R}$, which is a continuous increasing process given as the almost sure limit:

$$L_t(x) := \lim_{\epsilon \downarrow 0} \frac{1}{2\epsilon} \int_0^t 1_{(x-\epsilon, x+\epsilon)}(X_s) \mathrm{d}s.$$

Furthermore, $(O_t(x))_{t \geq 0}$ stands for the occupation time in the interval (x, ∞):

$$O_t(x) := \int_0^t 1_{(x,\infty)}(X_s) ds = \int_x^\infty L_t(y) \mathrm{d}y \quad \text{a.s.} \tag{14}$$

Our aim here is to establish limit theorems for the conditional mean estimators of $L_t(x)$ and $O_t(x)$.

3.1. Basic formulae

An important role will be played by the functions

$$g(x, z) := \mathbb{E}^0[L_1(x)|X_1 = z],$$

$$G(x, z) := \mathbb{E}^0[O_1(x)|X_1 = z] = \int_x^\infty g(y, z) \mathrm{d}y,$$

where $\mathbb{E}^0$ corresponds to the law of the standard Brownian motion. Both functions g and G have explicit formulae in terms of the density φ and survival function $\overline{\Phi}$ of the standard normal distribution. Some basic observations and these formulae are collected in the following result.

Lemma 2. *There are the identities*

$$\mathbb{E}[L_t(x)|X_t = z] = \frac{\sqrt{t}}{\sigma} g\left(\frac{x}{\sigma\sqrt{t}}, \frac{z}{\sigma\sqrt{t}}\right),$$

$$\mathbb{E}[O_t(x)|X_t = z] = t G\left(\frac{x}{\sigma\sqrt{t}}, \frac{z}{\sigma\sqrt{t}}\right).$$

Moreover, the functions g and G are bounded on $\mathbb{R}^2$ and satisfy $g(x,z) = g(-x,-z), G(x,z) = 1 - G(-x,-z)$. For $x \geq 0$ we have the formulae

$$z < x: \quad g(x,z) = \overline{\Phi}(2x - z)/\varphi(z),$$

$$G(x,z) = \frac{1}{2}\exp(-2x(x-z)) - (2x - z)\frac{\overline{\Phi}(2x-z)}{2\varphi(z)},$$

$$z \geq x: \quad g(x,z) = \overline{\Phi}(z)/\varphi(z),$$

$$G(x,z) = \frac{1}{2} + (z - 2x)\frac{\overline{\Phi}(z)}{2\varphi(z)}.$$

3.2. Estimators and the limit theory

The conditional mean estimators of L_t and O_t are easily derived using stationarity and independence of increments of X together with Lemma 2:

$$\widehat{L}_t(x) = \mathbb{E}[L_t(x)|(X_{i/n})_{i\geq 1}] \tag{15}$$

$$= \frac{1}{\sigma\sqrt{n}}\sum_{i=1}^{\lfloor nt \rfloor} g\left(\frac{\sqrt{n}}{\sigma}(x - X_{\frac{i-1}{n}}), \frac{\sqrt{n}}{\sigma}\Delta_i^n X\right) + O_{\mathbb{P}}(n^{-1/2}),$$

$$\widehat{O}_t(x) = \mathbb{E}[O_t(x)|(X_{i/n})_{i\geq 1}]$$

$$= \frac{1}{n}\sum_{i=1}^{\lfloor nt \rfloor} G\left(\frac{\sqrt{n}}{\sigma}(x - X_{\frac{i-1}{n}}), \frac{\sqrt{n}}{\sigma}\Delta_i^n X\right) + O_{\mathbb{P}}(n^{-1}),$$

where $\Delta_i^n X = X_{\frac{i}{n}} - X_{\frac{i-1}{n}}$. It is noted that the lower order terms can be written down explicitly (they are 0 when tn is an integer), but we keep them implicit, because they do not have an influence on the limit theorem presented below.

Theorem 3. *Assume that X is a linear Brownian motion with drift parameter $\mu \in \mathbb{R}$ and scale $\sigma > 0$. Then for any $x \in \mathbb{R}$ we have the functional stable convergence:*

$$n^{\frac{1}{4}}\left(\widehat{L}_t(x) - L_t(x)\right) \xrightarrow{d_{st}} \frac{v_l}{\sqrt{\sigma}}W_{L_t(x)}, \tag{16}$$

$$n^{\frac{3}{4}}\left(\widehat{O}_t(x) - O_t(x)\right) \xrightarrow{d_{st}} v_o\sqrt{\sigma}W_{L_t(x)}, \tag{17}$$

where W is a Brownian motion independent of $\mathcal{F}$ and

$$v_l^2 = \int_{\mathbb{R}} \mathbb{E}^0\left[g(y, X_1) - L_1(y)\right]^2 \mathrm{d}y = 2\frac{3\log(1 + \sqrt{2}) - \sqrt{2}}{3\sqrt{\pi}} \approx 0.4626,$$

$$v_o^2 = \int_{\mathbb{R}} \mathbb{E}^0[G(y, X_1) - O_1(y)]^2 \mathrm{d}y = \frac{13\sqrt{2} - 15\log(1 + \sqrt{2})}{45\sqrt{\pi}} \approx 0.065.$$

Importantly, our conditional mean estimator (15) is a particular example of a more general class of statistics investigated in [30] in the context of continuous

diffusion processes. The expression for v_l in [30] is rather lengthy and hard to evaluate, because of the generality assumed therein. In our case, $g(x, X_1) = \mathbb{E}[L_1(x)|X_1]$ is the conditional expectation and, in fact, a rather short direct proof can be given yielding the constant v_l^2 at the same time, see Appendix B.

Remark 1. The above v_l^2 can be compared to $\frac{8}{3\sqrt{\pi}}(\sqrt{2}-1) \approx 0.6232$ obtained when instead of the optimal $g(x,z)$ one uses the kernel $\hat{g}(x) = \int_{\mathbb{R}}(|x+u| - |x|)\varphi(u)\mathrm{d}u$ depending on x only, see [30, (1.27)]. The corresponding estimator (for $\sigma = 1$) is $\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor nt \rfloor} \hat{g}(\sqrt{n}(x - X_{\frac{i-1}{n}}))$, which does not take the increment following $X_{\frac{i-1}{n}}$ into account.

Remark 2. Consider the class of continuous SDEs defined via the equation

$$dX_t = \mu(X_t)dt + \sigma(X_t)dB_t,$$

where B is a standard Brownian motion and $\sigma \in C^1(\mathbb{R}), \mu \in C(\mathbb{R})$ are such that the above SDE has a unique strong solution. In [30] the author considers statistics of the form

$$L(h;x)_t^n = \frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor nt \rfloor} h\left(\sqrt{n}(x - X_{\frac{i-1}{n}}), \sqrt{n}\Delta_i^n X\right).$$

When $\sigma > 0$ and $|h(y,z)| \le \tilde{h}(y)\exp(a|z|)$ with $\tilde{h}$ bounded and satisfying $\int_{\mathbb{R}} |y|^r \tilde{h}(y)\mathrm{d}y < \infty$ for some $r > 3$, the stable convergence

$$n^{1/4}\left(L(h;x)_t^n - c_h(x)L_t(x)\right) \xrightarrow{d_{st}} v_h(x)W_{L_t(x)}$$

holds, see [30, Theorem 1.2]. Furthermore, the positive constant $v_h(x)$ (and the proof of stable convergence) stems from the simpler model $X_t = \sigma(x)B_t$. Hence, we can conclude that our estimator $\widehat{L}_t(x)$ is asymptotically optimal within the class of statistics $L(h;x)_t^n$ in the general setting of continuous SDEs. We believe that the restriction to the class $L(h;x)_t^n$ is not required and $\widehat{L}_t(x)$ is asymptotically efficient for continuous SDEs. Furthermore, when the function σ is unknown the coefficient $\sigma(x)$ can be estimated with a $n^{1/3}$-accuracy [24] and we can build a feasible statistic without affecting the asymptotic theory (cf. Proposition 2 below).

4. Some modifications of the proposed statistics

The main goal of this section is to show that the above developed theory also applies in the setting when the law of X is not known, but a consistent estimator of the parameters is available. Furthermore, we construct certain simplified estimators of the supremum in order to cope with potential numerical issues.

4.1. Unknown parameters

The main results of Theorem 2 and Theorem 3 above assume that the law of the process X is known, which is hard to accept in practice. At most, we are willing to assume that the process X belongs to some parametric class, and we distinguish between the following two:

(i) Linear Brownian motion with drift parameter $\mu \in \mathbb{R}$ and scale $\sigma > 0$, where for convenience we set $\alpha = 2$. As we remarked earlier neither the statistics nor the limits in Corollary 1 and Theorem 3 depend on μ, which, in fact, can not be estimated consistently on finite interval of time. Hence, the only parameter of interest is $\theta = \sigma$.

(ii) Non-monotone self-similar Lévy process which is naturally parameterized [43, §I.5] by a triplet $\theta = (\alpha, \rho, \lambda)$, where $\rho = \mathbb{P}(X_1 > 0)$ is the positivity parameter and $\lambda = \mathbb{E}[\log(|X_1|)]$ is related to the scale. It is noted that $\rho \in [1 - 1/\alpha, 1/\alpha]$ for $\alpha \in (1, 2]$, and $\rho \in (0, 1)$ for $\alpha \in (0, 1]$ which excludes monotone processes. This parametrization, unlike the one with skewness parameter, is continuous in the sense that convergence of parameters holds iff the processes converge.

Suppose now that we have a consistent estimator θ_n of the true parameter θ. Feasible estimators for supremum, local time and occupation time are now obtained via the plug-in approach. In particular, we have

$$\widetilde{T}_n^{\mathrm{mean}} = M_n + \int_0^\infty \left(1 - H_n^{\theta_n}(x)\right) \mathrm{d}x, \quad \widetilde{T}_n^{\mathrm{med}} = M_n + (H_n^{\theta_n})^{-1}(1/2),$$

where $H_n^{\theta_n}(x) = \prod_{j=0}^{n-1} F_{\theta_n}(n^{1/\alpha_n}(x + \Delta_j^n), n^{1/\alpha_n}(\Delta_j^n - \Delta_{j+1}^n))$, and

$$\widetilde{L}_t(x) = \frac{1}{\sigma_n \sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} g\left(\frac{\sqrt{n}}{\sigma_n}(x - X_{\frac{i-1}{n}}), \frac{\sqrt{n}}{\sigma_n} \Delta_i^n X\right) + O_{\mathbb{P}}(n^{-1/2}),$$

$$\widetilde{O}_t(x) = \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} G\left(\frac{\sqrt{n}}{\sigma_n}(x - X_{\frac{i-1}{n}}), \frac{\sqrt{n}}{\sigma_n} \Delta_i^n X\right) + O_{\mathbb{P}}(n^{-1}).$$

The construction of estimators θ_n of the unknown parameter θ for models (i) and (ii) is a well understood problem in the statistical literature. In particular, in class (i) the maximum likelihood estimator of σ is given by

$$\sigma_n^2 = \sum_{i=1}^{n} (\Delta_i^n X)^2$$

and it holds that $\sqrt{n}(\sigma_n^2 - \sigma^2) \xrightarrow{\mathrm{d}} \mathcal{N}(0, 2\sigma^4)$. Numerous theoretical results on parametric estimation of model (ii) can be found in e.g. [35]. Since the maximum likelihood estimator of θ is not explicit, we rather propose to use the following statistics:

$$\alpha_n = \frac{q \log(2)}{\log\left(\sum_{i=2}^{n} |X_{i/n} - X_{(i-2)/n}|^q\right) - \log\left(\sum_{i=1}^{n} |X_{i/n} - X_{(i-1)/n}|^q\right)},$$

$$\rho_n = \frac{1}{n} \sum_{i=1}^{n} 1_{\{\Delta_i^n X > 0\}}, \qquad \lambda_n = \frac{1}{n} \sum_{i=1}^{n} \log(n^{1/\alpha_n} |\Delta_i^n X|),$$

where $q \in (-1/2, 0)$. Additionally, we need to ensure that our parameters are legal, and in particular α_n, when larger than 1, is truncated at $(\rho_n \vee (1 - \rho_n))^{-1}$. Due to self-similarity of X and the law of large numbers we have that

$$\frac{\sum_{i=2}^{n} |X_{i/n} - X_{(i-2)/n}|^q}{\sum_{i=1}^{n} |X_{i/n} - X_{(i-1)/n}|^q} \xrightarrow{\mathbb{P}} 2^{q/\alpha},$$

which gives the idea behind the construction of α_n. Indeed, all estimators are weakly consistent and since $\mathbb{E}[|X_1|^{2q}] < \infty$ for $q \in (-1/2, 0)$ we easily conclude that

$$\alpha_n - \alpha = O_{\mathbb{P}}(n^{-1/2}), \qquad \rho_n - \rho = O_{\mathbb{P}}(n^{-1/2}), \qquad \lambda_n - \lambda = O_{\mathbb{P}}(n^{-1/2} \log(n)).$$

The proposed estimators are not efficient, but they suffice for our purposes, see Proposition 1 below.

It turns out that the limit theory presented in Theorem 2 and Theorem 3 continues to hold under a rather weak assumption on a consistent estimator θ_n of θ; in particular, this assumption is satisfied by estimators we proposed above. In other words, the difference between the modified and original estimators is negligible in the right sense.

Proposition 1. *Consider parametric class (i) with $\theta = \sigma$ and $\sigma_n \xrightarrow{\mathbb{P}} \sigma$ or (ii) with $\theta = (\alpha, \rho, \lambda)$ and $\theta_n \xrightarrow{\mathbb{P}} \theta$, $(\alpha_n - \alpha) \log n \xrightarrow{\mathbb{P}} 0$. Then*

$$n^{1/\alpha}(\overline{T}_n^{\mathrm{mean}} - \widetilde{T}_n^{\mathrm{mean}}) \xrightarrow{\mathbb{P}} 0, \qquad \text{for } \alpha \in (1, 2],$$
$$n^{1/\alpha}(\overline{T}_n^{\mathrm{med}} - \widetilde{T}_n^{\mathrm{med}}) \xrightarrow{\mathbb{P}} 0.$$

Moreover, the limit distributions in (10) *and* (11) *are continuous in θ.*

This shows that the estimators $\widetilde{T}_n^{\mathrm{mean}}$ and $\widetilde{T}_n^{\mathrm{med}}$ are asymptotically efficient in the sense that they are asymptotically equivalent to the respective optimal estimators relying on the knowledge of true parameters. In class (ii) the true α is not known, but in view of Proposition 1 the assumption $(\alpha_n - \alpha) \log n \xrightarrow{\mathbb{P}} 0$ guarantees that

$$n^{1/\alpha_n}(\overline{X}_1 - \widetilde{T}_n^{\mathrm{mean}}) \xrightarrow{d_{st}} V - \int_0^{\infty} (1 - H(x)) \mathrm{d}x, \qquad \text{when } \alpha \in (1, 2],$$
$$n^{1/\alpha_n}(\overline{X}_1 - \widetilde{T}_n^{\mathrm{med}}) \xrightarrow{d_{st}} V - H^{-1}(1/2).$$

Furthermore, the limit distributions are well approximated by their analogues corresponding to parameter θ_n, and so we may construct asymptotic confidence intervals for the estimators $\widetilde{T}_n^{\mathrm{mean}}$ and $\widetilde{T}_n^{\mathrm{med}}$.

With respect to local/occupation time we have the following result.

Proposition 2. *Consider class (i) and assume that*

$$n^{1/4}(\sigma - \sigma_n) \xrightarrow{\mathbb{P}} 0. \tag{18}$$

Then for any $x \in \mathbb{R}, T > 0$ it holds that

$$n^{1/4} \sup_{t \le T} \left| \widehat{L}_t(x) - \widetilde{L}_t(x) \right| \xrightarrow{\mathbb{P}} 0, \qquad n^{3/4} \sup_{t \le T} \left| \widehat{O}_t(x) - \widetilde{O}_t(x) \right| \xrightarrow{\mathbb{P}} 0.$$

This again shows that the estimators $\widetilde{L}_t(x)$ and $\widetilde{O}_t(x)$ are asymptotically efficient, and provides the respective asymptotic confidence bounds. Condition (18) is quite expected in the case of local times since $n^{1/4}$ is the corresponding rate of convergence in (16), but it is surprising that this condition is also sufficient to conclude the asymptotic efficiency of $\widetilde{O}_t(x)$. Roughly speaking, the reason for condition (18) to be sufficient in the latter case is that partial derivatives of G correspond to the local time asymptotics thus changing the convergence rate from $n^{3/4}$ to $n^{1/4}$. We refer to §B.3 for more details.

4.2. Truncation of products in supremum estimators

Here we return to the assumption that the law of X is known. Consider supremum estimators defined in §2.2 in terms of the conditional distribution function $H_n(x)$. When the number n of observations is large, it may be desirable to reduce the number of terms in the product defining $H_n(x)$, in order to avoid numerical issues and to speed-up the calculations. This is especially true when X is not a linear Brownian motion and so the function F is not explicit.

Intuitively, we may want to keep the terms which are formed from the observations closest to the maximum. Thus, we let $H_n(x; k)$ for $k \in \mathbb{N}_+$ be the analogue of $H_n(x)$, but such that the product has at most $2k$ terms and, in particular, the indices j are chosen such that $0 \vee (I_n - k) \le j \le (I_n + k - 1) \wedge (n - 1)$ with I_n being the index of the maximal observation. Define $\overline{T}_{n,k}^{\mathrm{mean}}$ and $\overline{T}_{n,k}^{\mathrm{med}}$ as before but using $H_n(x; k)$ instead of H_n.

Letting $I \in \mathbb{Z}$ be the unique number satisfying $\xi_{I+U} = V$ (it achieves the minimum V in (3)), we define

$$H(x; k) := \prod_{I-k \le j \le I+k-1} F\left(x + \xi_{j+U} - V, \xi_{j+U} - \xi_{j+1+U}\right), \quad x \ge 0.$$

We now have the limit result analogous to Theorem 2:

Corollary 3. *For any $\alpha \in (0, 2]$ it holds that*

$$n^{1/\alpha}(\overline{X}_1 - \overline{T}_{n,k}^{\mathrm{mean}}) \xrightarrow{d_{st}} V - \int_0^\infty (1 - H(x; k)) \mathrm{d}x,$$

$$n^{1/\alpha}(\overline{X}_1 - \overline{T}_{n,k}^{\mathrm{med}}) \xrightarrow{d_{st}} V - H^{-1}(1/2; k).$$

It turns out that we do not need to exclude $\alpha \in (0,1]$ in the case of modified conditional mean estimator, because the number of terms is kept finite. In the Brownian case and $k = 1$ we arrive at the following simple formulas

$$
\begin{aligned}
H_n(x;1) &= \left(1 - 1_{\{I_n \geq 1\}} \exp\left[-2x(x + \Delta_{I_n-1}^n)n/\sigma^2\right]\right) \\
&\quad \times \left(1 - 1_{\{I_n < n\}} \exp\left[-2x(x + \Delta_{I_n+1}^n)n/\sigma^2\right]\right), \\
H(x;1) &= \left(1 - \exp\left[-2x(x + \xi_{I-1+U} - V)/\sigma^2\right]\right) \\
&\quad \times \left(1 - \exp\left[-2x(x + \xi_{I+1+U} - V)/\sigma^2\right]\right).
\end{aligned}
$$

Remark 3. It is likely that Theorem 2 can be generalized to an arbitrary Lévy process satisfying the following weak regularity condition:

$$
(a_u X_{t/u})_{t \geq 0} \xrightarrow{\mathrm{d}} (\widehat{X}_t)_{t \geq 0} \qquad \text{as } u \to \infty,
$$

for some positive function a_u and necessarily self-similar Lévy process $\widehat{X}$. Importantly, the general versions of (2) and (3) are proven in [29]; here the limiting objects correspond to $\widehat{X}$.

There are, however, two very serious difficulties. Firstly, joint convergence does not necessarily imply convergence of the conditional distributions. Thus, one needs to use the underlying structure to show that

$$
F_{1/n}(x/a_n, y/a_n) = \mathbb{P}(a_n \overline{X}_{1/n} \leq x | a_n X_{1/n} = y) \to \widehat{F}(x,y).
$$

Secondly, the proof of uniform negligibility of truncation in §A.3.2 crucially depends on X being self-similar. This part may be notoriously hard for a general Lévy process X.

5. Numerical illustration of the limit laws

In this section we perform some numerical experiments in order to illustrate the limit laws in Theorem 2 and Theorem 3. For simplicity we take X to be a standard Brownian motion and, additionally, a one-sided stable process in supremum estimation which is motivated by the semi-explicit formula for the function F in Proposition 4. All the densities are obtained from $10,000$ independent samples using standard kernel estimates. The number of samples is reduced to $1,000$ in the case of the one-sided stable process.

5.1. Supremum estimation for Brownian motion

Consider a standard Brownian motion X and the limiting random variable V in (3), as well as $V_{\mathrm{mean}} := V - \int_0^\infty (1 - H(x))\mathrm{d}x$ and $V_{\mathrm{med}} := V - H^{-1}(1/2)$ in (10) and (11), respectively. Recall that all of these quantities are explicit, see also Corollary 1, but they all depend on infinitely many observations $\xi_{j+U}, j \in \mathbb{Z}$ of the two-sided Bessel process ξ of order 3. We approximate these quantities by setting $\xi_{j+U} = \infty$ for $j < -50$ or $j \geq 50$, which effectively amounts to
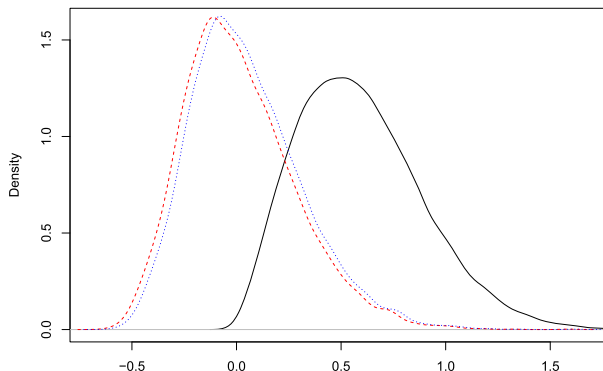
F<small>IG</small> 1. *Simulated densities of V (solid black), V_{mean} (dashed red) and V_{med} (dotted blue) in the Brownian case*

T<small>ABLE</small> 1
Some statistics in the Brownian case

	V	V_{mean}	V_{med}	V_{shift}	V^1_{mean}
L^2-norm	0.66	0.26	0.27	0.30	0.29
L^1-norm	0.59	0.21	0.21	0.24	0.22
95% conf. int. length	1.14	1.03	1.03	1.14	1.06

considering 100 epochs centered around 0; choosing twice as many epochs had negligible effect on the results below. The resulting densities are depicted in Figure 1. In Table 1 we report the L^2-norm, the L^1-norm, and the narrowest 95%-confidence interval length for each of the limiting distributions. It is noted that, indeed, V_{mean} has the smallest L^2-norm and V_{med} has the smallest L^1-norm, and the respective distributions are very similar.

Observe that the main problem of the standard estimator M_n is that it is downward biased and so V is not centered. This, however, can be easily remedied since according to [6]

$$\mathbb{E}V = -\zeta\left(\frac{1}{2}\right)\frac{1}{\sqrt{2\pi}} \approx 0.5826,$$

where ζ is the Riemann zeta function. In other words, we may consider an asymptotically centered estimator $M_n + \frac{1}{\sqrt{n}}\mathbb{E}V$, which leads to $V_{\mathrm{shift}} := V - \mathbb{E}V$. Finally, we also consider the truncated conditional mean estimator $\overline{T}_{n,1}^{\mathrm{mean}}$ based on $H(x;1)$, which is a product of two terms and thus only moderately more complicated to evaluate as compared to M_n, see §4.2. The respective limit is denoted by V^1_{mean}. Relative comparison of the latter two together with V_{mean} is provided in Figure 2, see also Table 1.

In conclusion, the conditional mean and conditional median estimators are very similar to each other and considerably better than the standard estimator M_n in terms of L^2-norm and L^1-norm. Nevertheless, the other simple estimators discussed above are only slightly worse than the optimal ones.
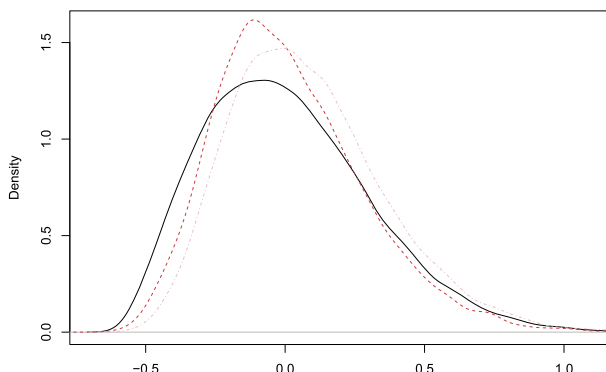
FIG 2. *Simulated densities of* V_{shift} *(solid black),* V_{mean} *(dashed red) and* V_{mean}^{1} *(dot-dashed pink) in the Brownian case*

TABLE 2
Some statistics in the stable case

	V	V_{mean}	V_{med}	V_{shift}
L^2-norm	0.87	0.41	0.42	0.43
L^1-norm	0.75	0.32	0.32	0.34
95% conf. int. length	1.45	1.42	1.42	1.45

5.2. Supremum for one-sided stable process

Here we consider a strictly stable Lévy process with $\alpha = 1.8$, standard scale and only negative jumps present, i.e., the skewness parameter is $\beta = -1$. Note that the results in the opposite case $\beta = 1$ must be similar according to Proposition 3. The conditional distribution function F is numerically evaluated using the expressions in Proposition 4, see Figure 3a.

In this case we perform a number of approximations. Firstly, simulation of ξ is not obvious (unlike the Brownian case) and so we approximate the limiting object $(V, H(x))$ by $(n^{1/\alpha}(\overline{X}_1 - M_n), H_n(xn^{-1/\alpha}))$ with $n = 300$, see (9). Instead of scaling $\overline{X}_1, M_n, \Delta_i$ with $n^{1/\alpha}$ we perform the simulation of the process X on the interval $[0, n]$, which is allowed by self-similarity of X. Furthermore, X is simulated on the grid with step-size $1/m$ for $m = 300$, which yields an approximation of $\overline{X}_n$ further corrected by the easily computable asymptotic mean error $m^{-1/\alpha}\mathbb{E}V$, see [7]. Next, we take (at most) 30 terms in the product defining H_n based on the observations closest to the maximum, that is we replace it by $H_n(\cdot; 15)$ defined in §4.2. Finally, $\int_0^\infty (1 - H(x))\mathrm{d}x$ is approximated using the trapezoidal rule with step size 0.1 and truncation at $x = 3$, see Figure 3b; the same approximation is used in calculation of the inverse.

The results are presented in Figure 4 and Table 2. They are quite similar to the results in the Brownian case.
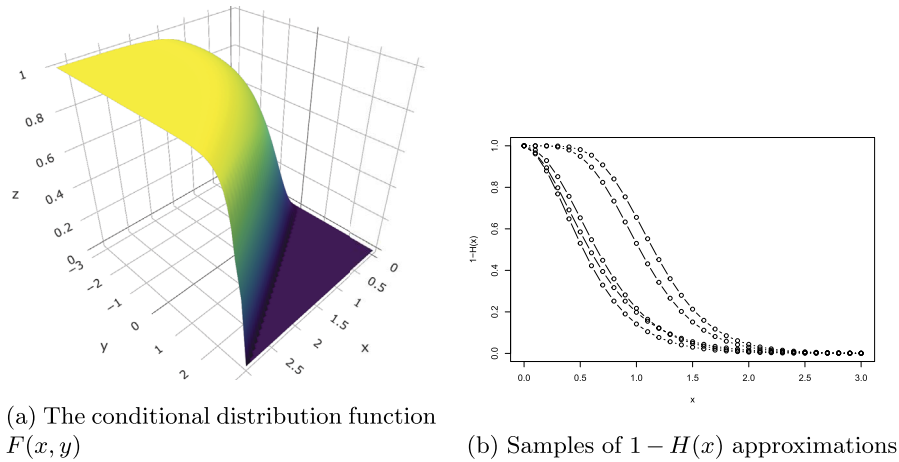
(a) The conditional distribution function $F(x, y)$

(b) Samples of $1 - H(x)$ approximations

Fig 3. *The stable case*

5.3. Local time and occupation time for Brownian motion

Let again X be a standard Brownian motion and choose $x = 0, t = 1$. We use $\widehat{L}_1(0)$ with $n = 10,000$ as a substitute for the true $L_1(0)$, which then allows to sample (approximately) from the limit distribution in (16). Next, we use the same sample path to construct $\widehat{L}_1(0)$ with $n = 100$, which allows to sample from the pre-limit expression in (16). Finally, we also take a standard estimator $\frac{1}{2\sqrt{n}} \# \{i \in [0 : n-1] : |X_{i/n}| < \frac{1}{\sqrt{n}}\}$ for $n = 100$. The respective densities are depicted in Figure 5. The ratio of variances for $n = 100$ is $1 : 1.64$, which can be compared to $1 : 1.35$ for the more advanced estimator mentioned in Remark 1 (here we use the exact expressions of the limits).

We perform a similar procedure for the occupation time in $(0, \infty)$. Here the standard estimator is $\frac{1}{n} \# \{i \in [0 : n-1] : X_{i/n} \geq 0\}$. The respective densities are given in Figure 6, and we see a very substantial improvement. The ratio of variances for $n = 100$ is $1 : 2.64$.

Appendix A: Proofs for supremum estimation

In the following all positive constants will be denoted by c although the may change from line to line.

A.1. Duality

In this section we establish a duality result for a general Lévy process X. Even though it is not needed for the proofs, we present this duality, because it explains certain structure in the main results. To this end, consider the process $X'_t = -X_t$ and the associated quantities $\overline{X'}_1, M'_n, H'_n(x), F'_t(x, y)$, see §2.2.
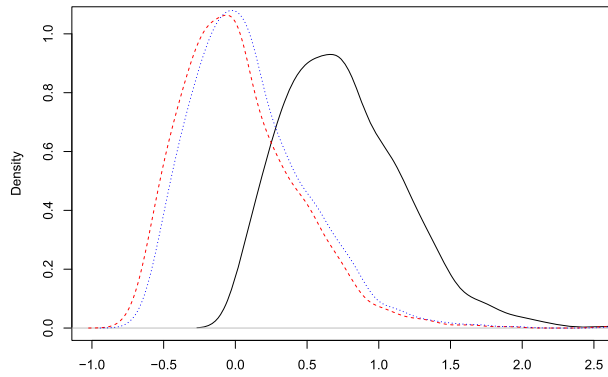
FIG 4. *Simulated densities of V (solid black), V*$_{\mathrm{mean}}$ *(dashed red) and V*$_{\mathrm{med}}$ *(dotted blue) in the stable case*
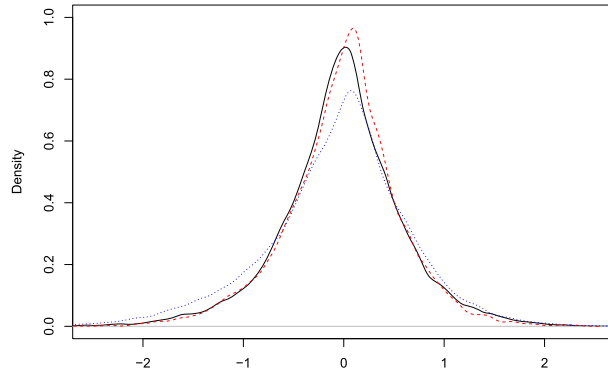

FIG 5. *Local time: the densities of the limit (solid black) and pre-limit (dashed red) in* (16) *for n = 100, as well as pre-limit for the standard estimator (dotted blue)*
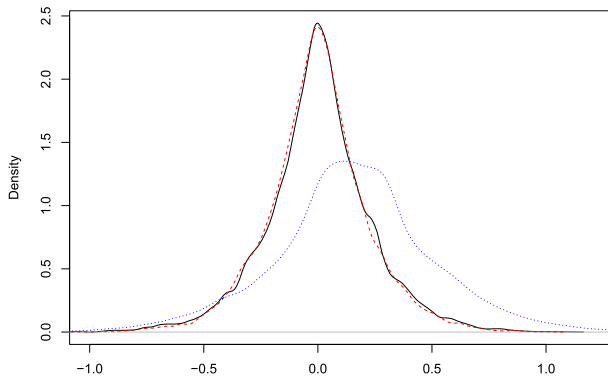

FIG 6. *Occupation time: the densities of the limit (solid black) and pre-limit (dashed red) in* (17) *for n = 100, as well as pre-limit for the standard estimator (dotted blue)*

Proposition 3. *Let X be an arbitrary Lévy process. Then*

$$\left(\overline{X'}_1 - M_n', (H_n'(x))_{x \geq 0}\right) \overset{d}{=} \left(\overline{X}_1 - M_n, (H_n(x))_{x \geq 0}\right).$$

Furthermore, $F_t'(x, y) = F_t(x - y, -y)$.

Proof. Take $(X_t'')_{t \in [0,1]} := (X_{(1-t)-} - X_1)_{t \in [0,1]}$ having the law of $(-X_t)_{t \in [0,1]}$ (standard time-reversal). Then $\overline{X''}_1 - M_n'' = \overline{X}_1 - X_1 - (M_n - X_1) = \overline{X}_1 - M_n$, because X does not jump at j/n almost surely. Letting x_j be the observation of $X_{j/n}$ we find that

$$H_n(x) = \mathbb{P}(\overline{X''}_1 - M_n'' \leq x | X_{k/n} - X_1 = x_k - x_n \; \forall k \in [1 : n - 1], X_1 = x_n)$$
$$= \mathbb{P}(\overline{X''}_1 - M_n'' \leq x | X_{(n-k)/n}'' = x_{n-k}'' \; \forall k \in [1 : n - 1], X_1'' = x_n'') = H_n'(x).$$

Finally,

$$F'(x, y) = \mathbb{P}(\overline{X''}_1 \leq x | X_1'' = y) = \mathbb{P}(\overline{X}_1 - X_1 \leq x | X_1 = -y)$$
$$= \mathbb{P}(\overline{X}_1 \leq x - y | X_1 = -y) = F(x - y, -y)$$

and the same reasoning works for $F_t'(x, y)$ when time-reverting at t. $\square$

In view of Proposition 3, the errors $\overline{X}_1 - \overline{T}_n^{\text{mean}}$ and $\overline{X}_1 - \overline{T}_n^{\text{med}}$ have the same distribution as the respective errors for the process $-X$. Thus, the corresponding limit results must stay the same when the skewness parameter β is flipped to the opposite. In the proofs we may safely assume that $\beta \geq 0$, say.

A.2. On the function F in the stable case

Before starting the proof of the main result we establish some basic properties of the conditional probability $F(x, y)$ in the case of a strictly α-stable process when it is not explicit. Throughout this subsection we assume that X is a strictly α-stable process with skewness parameter $\beta \in [-1, 1]$. Note that the boundary values $\beta = -1$ and $\beta = 1$ correspond to spectrally negative and spectrally positive processes, respectively; in both cases we must have $\alpha \in (1, 2)$, because we have excluded monotone processes.

It is well known [41, p. 88] that X_t has a continuous strictly positive bounded density, call it $f_t(x)$. Moreover, by self-similarity

$$f_t(x) = t^{-1/\alpha} f(t^{-1/\alpha} x) \qquad \text{with } f = f_1. \tag{19}$$

Furthermore, $f(x) \sim c x^{-\alpha-1}$ as $x \to \infty$ when $\beta \neq -1$, and otherwise it decays faster than an exponential function [41, Eq. (14.37)].

Let us define the first passage times $\tau_x^\pm = \inf\{t \geq 0 : \pm X_t > \pm x\}$ above and below a given level x.

Lemma 3. *The function $F(x,y)$ is jointly continuous. Moreover, $F(x,y) = 0$ for $x \leq y_+$, and otherwise*

$$\overline{F}(x,y)f(y) := (1 - F(x,y))f(y)$$

$$= \mathbb{E}\left[(1 - \tau_x^+)^{-1/\alpha} f((1 - \tau_x^+)^{-1/\alpha}(y - X_{\tau_x^+})); \tau_x^+ < 1\right] \qquad (20)$$

$$= \mathbb{E}\left[(1 - \tau_{y-x}^-)^{-1/\alpha} f((1 - \tau_{y-x}^-)^{-1/\alpha}(y - X_{\tau_{y-x}^-})); \tau_{y-x}^- < 1\right], \tag{21}$$

where $\mathbb{E}[Y; A] = \mathbb{E}[Y 1_A]$.

Proof. Assume for the moment that $x > y_+$. By time reversal (or from Proposition 3) we get

$$\mathbb{P}(\overline{X}_1 > x, X_1 \in \mathrm{d}y) = \mathbb{P}(\underline{X}_1 < y - x, X_1 \in \mathrm{d}y). \tag{22}$$

Using the strong Markov property we find that

$$\iint_{t \in (0,1), z \geq x} \mathbb{P}(\tau_x^+ \in \mathrm{d}t, X_{\tau_x^+} \in \mathrm{d}z) f_{1-t}(y - z)$$

is a version of the density of the measure on the left of (22). This expression coincides with (20) according to (19). Similarly, (21) is a version of the density of the measure on the right of (22), and hence both expressions coincide for almost all y.

Next, we show that the expressions in (20) and (21) are jointly continuous on $x > y_+$, and thus must coincide on this domain. We do this for the first expression only, since the other can be treated in the same way. By the basic properties of Lévy processes [10] we see that $\tau_x^+ \neq 1$ and $(\tau_x^+, X_{\tau_x^+})$ is continuous on an event of probability 1. Hence we only need to show that the dominated convergence theorem applies. Choose an arbitrary sequence (x', y') converging to (x, y) with $x > y_+$. Now $X_{\tau_{x'}^+} - y' > x' - y' > \epsilon$ for some $\epsilon > 0$ (further down in the sequence). Note that $f(-x) \leq cx^{-\alpha-1}$ for some $c > 0$ and all $x > 0$; in the spectrally positive case the decay is even faster. Hence the term under the expectation is bounded by $c(1 - \tau_{x'}^+)\epsilon^{-\alpha-1}$ and we are done.

It is left to show that either one of (20) and (21) converges to $f(y')$ as $x \to x', y \to y'$ with $x < y_+$ and $x' = y'_+$ (the boundary of the domain); this would imply $F(x,y) \to 0$. In the case $y' < 0$ use (20) and the above reasoning, while for $x' > 0$ use (21). It is left to analyze the case of $x' = y' = 0$. Note that (20) is lower bounded by the same expression with the indicator replaced by the indicator of $\tau_x^+ < 1/2$. But now the dominated convergence theorem applies and yields the limit $f(0)$. The upper bound is $f(y)$ by construction, and the limit is again $f(0)$. The proof is thus complete. □

We are now ready to provide some bounds on $\overline{F}(x,y)$. In the one-sided cases the bounds can be considerably improved, but this is not needed in this work and so we prefer a simpler statement.

Lemma 4. *There exists a constant $c > 0$ such that for all $x \geq y_+$:*

$$\overline{F}(x,y) \leq cx^{-\alpha}(x-y)^{-\alpha-1}(|y| \vee 1)^{\alpha+1}, \qquad \beta \in (-1,1),$$

$$\overline{F}(x,y) \leq c\exp(-(x-y_+)), \qquad\qquad \beta = \pm 1.$$

Proof. Suppose that $\beta \in (-1,1)$. We know that $f(x) < c|x|^{-\alpha-1}$. According to (20) we then have

$$\overline{F}(x,y)f(y) \leq c\mathbb{E}[(X_{\tau_x^+} - y)^{-\alpha-1}(1-\tau_x^+); \tau_x^+ < 1] \leq c(x-y)^{-\alpha-1}\mathbb{P}(\overline{X}_1 > x).$$

It is left to recall that $\mathbb{P}(\overline{X}_1 > x) \sim cx^{-\alpha}$ as $x \to \infty$ when $\beta \neq -1$ [12, 23].

Assume that $\beta = -1$. In this case $f(x) \sim ax^b\exp(-ux^v)$ with $a, u > 0$ and $v > 1$ as $x \to \infty$, see [41, Eq. (14.37)]. Furthermore, the asymptotics of $\mathbb{P}(\overline{X}_1 > x)$ has a similar form [12, Prop. 3b]. Observe that $f(Ax + B) < cA^{-1}f(x)$ for all $A \geq 1, B \geq 0, x \geq 1$. Hence, from (21) we get the bound

$$\overline{F}(x,y) \leq cf(x)/f(y),$$

which for $y > 0$ leads to the claimed bound $c\exp(-(x-y))$. For $y \leq 0$ we find from (20) that

$$\overline{F}(x,y) \leq c(x-y)^{-\alpha-1}\exp(-x)/f(y),$$

which readily implies the bound $c\exp(-x)$. Similar analysis yields the bound in the case $\beta = 1$. □

It is noted that we may also derive a bound

$$\overline{F}(x,y) \leq cx^{-\alpha-1}(x-y)^{-\alpha}(|y| \vee 1)^{\alpha+1}$$

for $\beta \in (-1,1)$ by using (21) instead of (20). This bound is better when $y > 0$ and worse when $y < 0$. For our purpose any of these bounds is sufficient.

Finally, we derive a semi-explicit expression of $F(x,y)$ in the one-sided case. This expression is in terms of the density f.

Proposition 4. *In the spectrally one-sided cases we have for all $x > y_+$:*

$\beta = -1$:

$$\overline{F}(x,y) = \frac{x}{f(y)}\int_0^1(1-t)^{-1/\alpha}t^{-1/\alpha-1}f(xt^{-1/\alpha})f((y-x)(1-t)^{-1/\alpha})\mathrm{d}t,$$

$\beta = 1$:

$$\overline{F}(x,y) = \frac{x-y}{f(y)}\int_0^1(1-t)^{-1/\alpha}t^{-1/\alpha-1}f(x(1-t)^{-1/\alpha})f((y-x)t^{-1/\alpha})\mathrm{d}t.$$

Proof. It is known that $\mathbb{P}(\overline{X}_1 \in \mathrm{d}x) = \alpha f(x)$ for $x > 0$, when X is spectrally negative, see [36]. Moreover,

$$\mathbb{P}(\tau_x^+ < t) = \mathbb{P}(\overline{X}_t > x) = \mathbb{P}(\overline{X}_1 > xt^{-1/\alpha})$$

yielding that $\mathbb{P}(\tau_x^+ \in \mathrm{d}t) = xt^{-1/\alpha-1}f(xt^{-1/\alpha})\mathrm{d}t$. Plugging this into Lemma 3 yields the result. Finally, (21) follows from $\overline{F}'(x,y) = \overline{F}(x-y,-y)$, see Proposition 3. □

Remark 4. Note that $\mathbb{E}(\overline{X}_1|X_1 = y) < \infty$ for all $y \in \mathbb{R}$, even in the cases $\alpha \in (0, 1]$ where $\mathbb{E}\overline{X}_1 = \infty$. This follows from Lemma 4 showing that for fixed y we have a bound $\overline{F}(x, y) \leq cx^{-2\alpha - 1}$. Thus conditional moments of order up to $1 + 2\alpha$ exist. In the spectrally-positive case we even have $\mathbb{E}(\exp(\lambda \overline{X}_1)|X_1 < b) < \infty$ for any $b < \infty, \lambda > 0$ (Lemma 4 gives only $\lambda < 1$ though).

A.3. Proof of Theorem 2

In the following we frequently use the inequality

$$|\prod_{j \in \mathbb{Z}} a_j - \prod_{j \in \mathbb{Z}} b_j| \leq \sum_{j \in \mathbb{Z}} |a_j - b_j| \text{ when } a_j, b_j \in (0, 1). \tag{23}$$

Let $I^{(n)} = \lceil \tau n \rceil$ be the index of the first observation to the right of the supremum time, and put

$$u_i^{(n)} = \begin{cases} n^{1/\alpha}(\overline{X}_1 - X_{(i+I^{(n)})/n}), & i + I^{(n)} \in [0, n] \\ \infty, & \text{otherwise.} \end{cases}$$

In other words, $u_i^{(n)}$ are the rescaled distances from the supremum to the observations indexed with respect to the time of supremum. Now we can represent the quantities appearing in (9) as follows:

$$V^{(n)} := n^{1/\alpha}(\overline{X}_1 - M_n) = \min_{i \in \mathbb{Z}} u_i^{(n)},$$

$$H^{(n)}(x) := H_n(xn^{-1/\alpha}) = \prod_{i \in \mathbb{Z}} F(x + u_i^{(n)} - V^{(n)}, u_i^{(n)} - u_{i+1}^{(n)}),$$

where by convention $F = 1$ if either of $u_i^{(n)}, u_{i+1}^{(n)}$ is infinite. According to [29] (or [6] in the case of Brownian motion) we have the following weak convergence for every $k > 0$:

$$\left((u_i^{(n)})_{|i| \leq k}, V^{(n)}\right) \xrightarrow{d_{st}} \left((\xi_{i+U})_{|i| \leq k}, \min_{i \in \mathbb{Z}} \xi_{i+U}\right). \tag{24}$$

Intuitively, this limit can be understood as arising from (2) together with the fact that $\{n\tau\}$ converges to an independent uniform on $(0, 1)$. This explains (of course, only intuitively) the form of the result in Theorem 2.

A.3.1. Convergence of the truncated versions

Let $H_k^{(n)}$ be the same as $H^{(n)}$, but with the product running over $|i| \leq k$:

$$H_k^{(n)}(x) = \prod_{|i| \leq k} F(x + u_i^{(n)} - V^{(n)}, u_i^{(n)} - u_{i+1}^{(n)}),$$

where again $F = 1$ when the index is out of range. We also define the analogous object formed from the limiting quantities:

$$H_k^{(\infty)}(x) = \prod_{|i| \leq k} F\left(x + \xi_{j+U} - V, \xi_{j+U} - \xi_{j+1+U}\right).$$

Note that $H_k^{(n)}(x), H_k^{(\infty)}(x)$ are continuous and strictly increasing in $x \geq 0$ which is inherited from $F(x, y)$. Furthermore, $H_k^{(n)}(\infty) = H_k^{(\infty)}(\infty) = 1$ and their value at 0 is not necessarily 0. In the following the inverse of an increasing function f is defined as usual: $f^{-1}(q) = \inf\{s : f(s) \geq q\}$.

Lemma 5. *For any $k \in \mathbb{N}$ as $n \to \infty$ we have the functional stable convergence*

$$(V^{(n)}, H_k^{(n)}(x)_{x \geq 0}) \xrightarrow{d_{st}} (V, H_k^{(\infty)}(x)_{x \geq 0})$$

with respect to the uniform topology. Moreover,

$$\left(V^{(n)}, \int_0^\infty (1 - H_k^{(n)}(x))\mathrm{d}x\right) \xrightarrow{d_{st}} \left(V, \int_0^\infty (1 - H_k^{(\infty)}(x))\mathrm{d}x\right),$$

where the limit variables are finite almost surely.

Proof. In view of (24) we only need to establish the continuity of the respective maps. Consider $(2k+1)$-dimensional vectors $a^{(n)}$ and $b^{(n)}$ converging to some vectors a and b, respectively, where the entries of $a^{(n)}$ and a are non-negative and the entries of a, b are finite. Observe using (23) that

$$\sup_{x \geq 0} \left| \prod_{|i| \leq k} F\left(x + a_i^{(n)}, b_i^{(n)}\right) - \prod_{|i| \leq k} F\left(x + a_i, b_i\right) \right|$$
$$\leq \sum_{|i| \leq k} \sup_{x \geq 0} \left| F\left(x + a_i^{(n)}, b_i^{(n)}\right) - F\left(x + a_i, b_i\right) \right| \to 0,$$

where convergence of F is uniform in $x \geq 0$ since the limit function is continuous and non-decreasing in $x \geq 0$, and is upper bounded (Polya's theorem). Thus, the first statement is now proven.

Concerning the second statement, we find that

$$\left| \int_0^\infty (1 - \prod_{|i| \leq k} F(x + a_i^{(n)}, b_i^{(n)}))\mathrm{d}x - \int_0^\infty (1 - \prod_{|i| \leq k} F(x + a_i, b_i))\mathrm{d}x \right|$$
$$\leq \sum_{|i| \leq k} \int_0^\infty \left| \overline{F}(x + a_i^{(n)}, b_i^{(n)}) - \overline{F}(x + a_i, b_i) \right| \mathrm{d}x$$

and it is left to show that each summand converges to 0, i.e. that the dominated convergence theorem applies. According to Lemma 4 both $\overline{F}(x + a_i^{(n)}, b_i^{(n)})$ and $\overline{F}(x + a_i, b_i)$ are bounded by $c(1 \wedge x^{-2\alpha-1})$, because of monotonicity of $\overline{F}$ in

the first argument and the fact that $b_i^{(n)} \to b_i < \infty$; the decay is even faster in the case of $\beta = \pm 1$ or when X is a Brownian motion. The proof of the second statement is now complete, since finiteness of the limit is shown in the same way. $\qquad \square$

A.3.2. Uniform negligibility of truncation

Showing that truncation at a finite k is uniformly negligible (in the sense of [11, Thm. 3.2]) is the crux of the proof. Firstly, we will need the following representation-in-law of the sequences $u_i^{(n)}$, which builds on [9] and self-similarity of X.

Lemma 6. *There exists a process $\tilde\xi$ having the law of ξ and a sequence of random variables τ_n such that $(\tau_n)_{n>0}$ and $(n - \tau_n)_{n>0}$ are non-negative non-decreasing sequences and the following is true: Let*

$$\tilde{u}_i^{(n)} := \tilde\xi_{i+1-\{\tau_n\}}$$

for all $i \in [-\lceil\tau_n\rceil, n - \lceil\tau_n\rceil]$, and otherwise $\tilde u_i^{(n)} := \infty$. Then $(u_i^{(n)})_{i\in\mathbb{Z}} \stackrel{\mathrm{d}}{=} (\tilde u_i^{(n)})_{i\in\mathbb{Z}}$ for all $n \in \mathbb{N}_+$.

Proof. By self-similarity $(n^{1/\alpha} X_{t/n})_{t\in[0,n]}$ has the same law as $(X_t)_{t\in[0,n]}$. According to [9], the law of the latter process when seen from the supremum, see (1), coincides with a certain process $\tilde\xi$ killed outside of the interval $[-\tau_n, n - \tau_n]$, where $\tau_n = \int_0^n \mathbf{1}_{\{X_t>0\}}\mathrm{d}t$. It is noted that $\tilde\xi$ is constructed using juxtaposition of the excursions of X in half-lines according to their signs, and it does not depend on n. Clearly, τ_n and $n - \tau_n = \int_0^n \mathbf{1}_{\{X_t\leq0\}}\mathrm{d}t$ are non-decreasing sequences going to $+\infty$, and the laws of $\tilde\xi$ and ξ defined by (1) coincide. It is now left to recall the definition of $u_i^{(n)}$. $\qquad \square$

We will also need asymptotic bounds on the process ξ, which can be read of [25, Cor. 3.3] or [20], see also [37] for the Brownian case.

Lemma 7. *For any $p_-, p_+ > 0$ such that $p_- < 1/\alpha < p_+$ it holds that*

$$\lim_{t\to\infty} \xi_t/t^{p_-} = \infty, \qquad \lim_{t\to\infty} \xi_t/t^{p_+} = 0 \qquad almost\ surely.$$

In particular, the probability of the event

$$E_{T,p_\pm} := \{\forall t \geq T : \xi_t \in [t^{p_-}, t^{p_+}]\}$$

tends to 1 as $T \to \infty$.

The following result establishes convergence of certain series, which is only needed for the case of a stable process with two-sided jumps.

Lemma 8. *Assume that $\beta \in (-1, 1)$ and consider*

$$D_t = \sup_{h\in[0,1]} |\xi_{t+1+h} - \xi_{t+h}|.$$

Then there exist $p_\pm > 0$ such that $p_- < 1/\alpha < p_+$ and the following series are convergent for any $T > 0$:

$$\alpha \in (1,2): \qquad\qquad \sum_{i \geq 1} i^{-2\alpha p_-} \mathbb{E}[D_i^{\alpha+1}; E_{T,p_\pm}] < \infty, \qquad (25)$$

$$\alpha \in (0,1]: \qquad\qquad \sum_{i \geq 1} i^{-2\alpha p_- - p_-} \mathbb{E}[D_i^{\alpha+1}; E_{T,p_\pm}] < \infty.$$

Proof. Assume that $\alpha \in (1,2)$. Let us show that there exists a natural number k and

$$0 = \delta_0 < \delta_1 < \cdots < \delta_{k-1} < 1 < \delta_k$$

such that $\delta_j(\alpha+1)/\alpha - \delta_{j-1} < 1$ for all $j = 1, \ldots, k$. The jth inequality reads as $\delta_j < \psi(\delta_{j-1})$ with $\psi(u) = (1+u)\alpha/(1+\alpha)$ being a continuous function such that $\psi(u) > u$ iff $u < \alpha$. Note that it is sufficient to pick the smallest k such that $\psi^{(k)}(0) > 1$, where the latter denotes kth iterate. To see that such k exists, simply observe that $\psi^{(k)}(0)$ converges to $\alpha > 1$ as $k \to \infty$.

Choose $p_\pm$ close enough to $1/\alpha$ so that $\delta_{k-1} < \alpha p_- < 1 < \alpha p_+ < \delta_k$ and

$$\delta_j(\alpha+1)/\alpha - \delta_{j-1} < 2\alpha p_- - 1, \qquad \text{for all } j = 1, \ldots, k. \qquad (26)$$

According to Proposition 5 in §C, for any $i > T$ we have

$$\mathbb{P}\left(\{D_i \geq i^{\delta_{j-1}/\alpha}\} \cap E_{T,p_\pm}\right) \leq c i^{-\delta_{j-1}},$$

because $\xi_i > i^{p_-} > i^{\delta_{j-1}/\alpha}$ on the respective event. Now for any $j = 1, \ldots, k$ we have

$$\sum_i i^{-2\alpha p_-} \mathbb{E}\left[D_i^{\alpha+1}; \{D_i \in [i^{\delta_{j-1}/\alpha}, i^{\delta_j/\alpha})\} \cap E_{T,p_\pm}\right]$$

$$\leq \sum_i i^{-2\alpha p_-} i^{\delta_j(\alpha+1)/\alpha} \mathbb{P}\left(\{D_i \geq i^{\delta_{j-1}/\alpha}\} \cap E_{T,p_\pm}\right)$$

$$\leq c \sum_i i^{-2\alpha p_- + \delta_j(\alpha+1)/\alpha - \delta_{j-1}} < \infty$$

according to (26). Summing up over $j = 1, \ldots, k$ completes the proof of (25), because on the event $E_{T,p_\pm}$ we have $D_i < (i+2)^{p_+} < i^{\delta_k/\alpha}$ for $i > T$ large enough. Moreover, the first interval $[1, i^{\delta_1/\alpha})$ can be replaced by $[0, i^{\delta_1/\alpha})$ without any change required.

Next, assume that $\alpha \in (0,1]$ and choose $\delta_1 < \alpha p_- < 1 < \alpha p_+ < \delta_2$. Similarly, to the above calculation we find that it is sufficient to additionally guarantee that

$$-2\alpha p_- - p_- + \delta_j(\alpha+1)/\alpha - \delta_{j-1} < -1, \qquad j = 1, 2.$$

This is always possible when $\delta_2 < 1 + \delta_1\alpha/(\alpha+1)$. $\qquad\qquad \square$

We are now ready to establish that truncation is indeed uniformly negligible:

Lemma 9. *For any $\epsilon > 0$ we have*

$$\lim_{k \to \infty} \sup_n \mathbb{P}(\|H^{(n)} - H_k^{(n)}\|_\infty > \epsilon) = 0, \tag{27}$$

$$\lim_{k \to \infty} \sup_n \mathbb{P}\left(\left|\int_0^\infty (H^{(n)}(x) - H_k^{(n)}(x))\mathrm{d}x\right| > \epsilon\right) = 0, \quad \text{for } \alpha \in (1, 2]. \tag{28}$$

Moreover, almost surely it holds that

$$\sup_{x \geq 0}\left|1 - \prod_{|j|>k} F\left(x + \xi_{j+U} - V, \xi_{j+U} - \xi_{j+1+U}\right)\right| \to 0, \tag{29}$$

$$\int_0^\infty (1 - H_k^{(\infty)}(x))\mathrm{d}x \to \int_0^\infty (1 - H(x))\mathrm{d}x < \infty, \quad \text{for } \alpha \in (1, 2].$$

as $k \to \infty$.

Proof. We start by showing (27). Using (23) we find that

$$\|H^{(n)} - H_k^{(n)}\|_\infty \leq \sup_{x \geq 0} \sum_{|i|>k} \overline{F}(x + u_i^{(n)} - V^{(n)}, u_i^{(n)} - u_{i+1}^{(n)}),$$

where the summand is 0 when either of $u_i^{(n)}, u_{i+1}^{(n)}$ is infinite. By monotonicity of F in the first argument, and the fact that $\mathbb{P}(V^{(n)} > v)$ can be made arbitrarily small by choosing large enough v (recall that $V^{(n)} \xrightarrow{\mathrm{d}} V$), it is sufficient to show that

$$\sup_n \mathbb{P}\left(\sum_{i>k} \overline{F}(\tilde{u}_i^{(n)} - v, \tilde{u}_i^{(n)} - \tilde{u}_{i+1}^{(n)}) > \epsilon\right) \to 0, \tag{30}$$

where we have replaced $u_i^{(n)}$ by $\tilde{u}_i^{(n)}$ having the same law as defined in Lemma 6. Note also that the sum here runs over $i > k$ since the other part ($i < -k$) can be handled in the same way.

Choose $p_\pm$ with $p_- < 1/\alpha < p_+$ such that the conclusion of Lemma 8 is satisfied when $\alpha \in (0, 2), \beta \in (0, 1)$. Note that we may restrict to the event $\tilde{E}_{T,p_\pm}$ for a large enough $T > 0$, see Lemma 7; that is, we have $t^{p_-} \leq \tilde{\xi}_t \leq t^{p_+}$ for all $t > T$.

First, assume that $\alpha \in (0, 1)$ and $\beta \in (-1, 1)$. According to Lemma 4 we have the bound (this bound is 0 when $\tilde{u}_i^{(n)}$ or $\tilde{u}_{i+1}^{(n)}$ is infinite)

$$\begin{aligned}
&\overline{F}(\tilde{u}_i^{(n)} - v, \tilde{u}_i^{(n)} - \tilde{u}_{i+1}^{(n)}) \\
&\leq c(\tilde{u}_i^{(n)} - v)^{-\alpha}(\tilde{u}_{i+1}^{(n)} - v)^{-\alpha-1}(1 \vee \tilde{D}_i)^{\alpha+1} \\
&\leq c i^{-p_-(2\alpha+1)}(1 + \tilde{D}_i^{\alpha+1}),
\end{aligned}$$

where $\tilde{D}_i \leq \sup_n |\tilde{u}_i^{(n)} - \tilde{u}_{i+1}^{(n)}|$, see the definition of $\tilde{u}_i^{(n)}$ in Lemma 6. Now (30) follows by Markov's inequality from

$$\sum_{i>k} i^{-p_-(2\alpha+1)} \mathbb{E}[\tilde{D}_i^{\alpha+1}; \tilde{E}_{T,p_\pm}] < \infty,$$

see Lemma 8. In the case $\beta = \pm 1$ and $\alpha = 2$ the above becomes

$$\sum_{i>k} \exp(-ci^{p_-}) < \infty, \qquad \sum_{i>k} \exp(-ci^{2p_-}) < \infty,$$

respectively, which is obviously true.

Next we show (28). With respect to the second statement we only need to show that

$$\sup_n \mathbb{P} \left(\sum_{i>k} \int_0^\infty \overline{F}(x + \tilde{u}_i^{(n)} - v, \tilde{u}_i^{(n)} - \tilde{u}_{i+1}^{(n)}) dx > \epsilon \right) \to 0,$$

In the case $\alpha \in (1,2), \beta \in (-1,1)$ the upper of Lemma 3 reads

$$\overline{F}(x + \tilde{u}_i^{(n)} - v, \tilde{u}_i^{(n)} - \tilde{u}_{i+1}^{(n)}) < c(x + i^{p_-})^{-2\alpha-1}(1 + D_i^{\alpha+1}),$$

for $i > T$. Integrating over $x \geq 0$ we get the bound $ci^{-2\alpha p_-}(1 + D_i^{\alpha+1})$ and the proof is again completed by the Markov's inequality and Lemma 8. In the case $\beta = \pm 1$ the bound is

$$\sum_{i>k} \int_0^\infty \exp(-(x + i^{p_-} - v)) dx < \infty,$$

and a similar bound holds for $\alpha = 2$.

Finally, similar (but simpler) arguments show that there is convergence in probability in (29). But the product is monotone for each $x \geq 0$. Thus we have uniform convergence almost surely. For $\alpha \in (1,2]$ we find using above arguments that the integral $\int_0^\infty (1 - H(x)) dx$ is finite almost surely. Now the dominated convergence theorem applies. □

Proof of Theorem 2. Let us show the stated properties of H. It is clear that $H(x), x \geq 0$ is non-decreasing and takes values in $[0,1]$. Moreover, $H(0) = 0$ since one of the terms in the product is 0. Observe that (29) implies convergence of $H_k^{(\infty)}$ to H uniformly in $x \geq 0$ on the set of probability 1. Thus H is continuous and $H(\infty) = 1$, because the same is true about $H_k^{(\infty)}$. Finally, H is strictly monotone, since $H(x) > 0$ for every $x > 0$ which follows from positivity of $H_k^{(\infty)}$ and (29).

Stable convergence statements in (9) and (10) follow from Lemma 5 and Lemma 9 by means of [11, Thm. 3.2] extended to the setting of stable convergence. Concerning (11) we apply Skorokhod's representation theorem to the sequence $H^{(n)}$ (the underlying space of continuous functions with a limit at ∞ is indeed separable, as it can be time-changed into the space of continuous functions on $[0,1]$). The inverse H^{-1} is continuous and finite on $(0,1)$ and hence we have convergence of respective inverses [40, Prop. 0.1]. □

A.4. Related results

Here we provide the proofs (or just the main ingredients) of the results related to Theorem 2.

A.4.1. Linear Brownian motion

Proof of Corollary 1. Firstly, note that the scaling σ can be indeed taken out as in (12) and (13). This is true in general, because we may always rescale the process and the corresponding observations before the analysis. Thus we may assume that $\sigma = 1$ in the following.

Now suppose that $\mu \neq 0$ and so X is not self-similar. Recall that the estimators are the same as in the case $\mu = 0$. Furthermore, according to [6] the convergence in (24) is still true, where the limit variables are defined in terms of the Bessel process of order 3. The main difficulty is that Lemma 6 is no longer true and the proof of uniform negligibility of truncation fails.

By Girsanov's theorem, we may introduce arbitrary drift using exponential change of measure $\mathrm{d}\mathbb{P}'/\mathrm{d}\mathbb{P} = \exp(aX_1 + b)$ with appropriately chosen constants $a, b \in \mathbb{R}$. But then

$$\mathbb{P}'(\|H^{(n)} - H_k^{(n)}\|_\infty > \epsilon) = \mathbb{E}[\exp(aX_1 + b); \|H^{(n)} - H_k^{(n)}\|_\infty > \epsilon]$$
$$\leq \exp(|a|c + b)\mathbb{P}(\|H^{(n)} - H_k^{(n)}\|_\infty > \epsilon) + \mathbb{P}'(|X_1| > c),$$

where $c > 0$ is arbitrary. But as $k \to \infty$ the $\limsup_n$ of this expression converges to $\mathbb{P}'(|X_1| > c)$, which can be made arbitrarily small. Thus (27) holds for an arbitrary linear Brownian motion, and the same argument works for (28). □

Proof of Lemma 1. It is only required to show that $\mathbb{E}[\left(\sqrt{n}|\overline{X}_1 - \overline{T}_n^{\mathrm{mean}}|\right)^p]$ is bounded for an arbitrarily large p and all n. Furthermore, we may again restrict our attention to a driftless Brownian motion by change of measure and Cauchy-Schwarz inequality. The fact that $\mathbb{E}[\exp(\theta V^{(n)})]$ for any θ is bounded was established in [6], and so it is sufficient to show that

$$\mathbb{E}\left[\left(\int_0^\infty (1 - H_n(xn^{-1/2}))\mathrm{d}x\right)^p\right]$$
$$\leq \mathbb{E}\left[\left(\sum_i \int_0^\infty \overline{F}(x + u_i^{(n)} - V^{(n)}, u_i^{(n)} - u_{i+1}^{(n)})\mathrm{d}x\right)^p\right]$$

is bounded. The right-hand side is increased by pulling the sum out. Using the explicit expression for $\overline{F}$ we see that it is left to consider

$$\sum_{i \geq 1} \mathbb{E}\left[\left(\int_0^\infty \exp\left(-2(x + u_i^{(n)} \wedge u_{i+1}^{(n)} - V^{(n)})^2\right)\mathrm{d}x\right)^p\right]$$
$$\leq c\sum_{i \geq 1} \mathbb{E}\left[\exp\left(-p(u_i^{(n)} \wedge u_{i+1}^{(n)} - V^{(n)})\right)\right],$$

where we used that $\overline{\Phi}(4x) < c\exp(-x)$. Moreover, $V^{(n)}$ can be dropped out, because of Cauchy-Schwarz inequality and boundedness of $\mathbb{E}[\exp(pV^{(n)})]$. Finally, use Lemma 6 to get the bound:

$$\sum_{i \geq 1} \mathbb{E}[\exp(-p \min_{t \in [i, i+2]} \xi_t)].$$

The above is bounded by

$$\sum_{i \geq 1} \mathbb{E}[\exp(-p\xi_i/2)] + \sum_{i \geq 1} \mathbb{P}(\min_{t \in [i, i+2]} \xi_t < \xi_i/2).$$

The first sum is finite, because the inequality between arithmetic and quadratic means, $\sqrt{a^2 + b^2 + c^2} \geq (|a| + |b| + |c|)/\sqrt{3}$, and the definition of Bessel-3 process imply that the respective terms are bounded by $\mathbb{E}[\exp(-p\sqrt{i}|Z|/(2\sqrt{3}))^3]$ where Z is standard normal. By Tauberian theorem this quantity behaves as $\ell i^{-3/2}$ for large i with ℓ being a positive constant, and the first sum is indeed finite. The second sum can be treated using the arguments from Appendix C. In particular, we can show that $\mathbb{P}_x^{\uparrow}(\underline{X}_2 < x/2) < c\exp(-x)$, and hence we are left to consider $\sum_i \exp(-\xi_i/2)$ again. The proof is now complete. □

A.4.2. *Joint estimation: proof of Corollary 2*

The only new ingredient needed is the joint convergence of sequences in (24) corresponding to the processes X and $-X$ to their respective limits which are independent. Similar result appears in [7, Lem. 1] and only a minor adaptation is needed.

A.4.3. *On simplified estimators: proof of Corollary 3*

We only need to show that the analogue of (24) is true, where we take the respective $2k + 1$ elements in the vectors on the left. One can not apply the continuous mapping theorem for the infinite sequences though. We consider truncated sequences, apply the continuous mapping theorem, and then show uniform negligibility of truncation. The latter follows from the fact that

$$\lim_{T \to \infty} \sup_n \mathbb{P}\left(\sup_{|t| > T} \xi_t^{(n)} < a\right) = 0$$

for any $a > 0$, which readily follows from the representation of $\xi^{(n)}$ as in Lemma 6 in the self-similar case.

A.4.4. *Unknown parameters: proof of Proposition 1*

We will show that $n^{1/\alpha}(\widetilde{T}_n^{\mathrm{mean}} - \overline{T}_n^{\mathrm{mean}}) \xrightarrow{\mathbb{P}} 0$ when $\alpha \in (1, 2]$, and the same is true for the conditional median estimator for all $\alpha \in (0, 2]$. The proof of continuity of the limit disributions follows similar steps, see also [29] for the convergence of the respective processes ξ. The above readily translates into

$$\int_0^\infty (H_n^{\theta_n}(xn^{-1/\alpha}) - H_n(xn^{-1/\alpha}))\mathrm{d}x \xrightarrow{\mathbb{P}} 0, \qquad \sup_{x \geq 0} |H_n^{\theta_n}(xn^{-1/\alpha}) - H(x)| \xrightarrow{\mathbb{P}} 0,$$

respectively. We focus on the class of strictly stable Lévy processes (the proof for the class (i) is similar but easier) and let X^n be the process with parameters θ_n. Furthermore we write F^n and f^n for the analogues of conditional distribution F and density f.

We claim that it is sufficient to establish that F^n converges to F continuously, i.e.

$$F^n(x_n, y_n) \to F(x, y) \qquad \text{for any } (x_n, y_n) \to (x, y), \text{ s.t. } x > y_+. \tag{31}$$

For this note that α_n is arbitrarily close to α with high probability, and thus the arguments from the proof of Theorem 2 apply essentially without a change.

Thus we are left to prove (31) by reexamining the proof of Lemma 3. Firstly, we observe that $(X_{\tau_{x_n}^+}, \tau_{x_n}^+)\mathbf{1}_{\{\tau_{x_n}^+ < \infty\}}$ under $\mathbb{P}_{\theta_n}$ weakly converges to the respective quantity under $\mathbb{P}$, which follows by the (generalized) continuous mapping theorem and weak convergence of the Lévy processes. Secondly, the function

$$g_n(t, x, y) := f_{1-t}^n(y - x) = (1 - t)^{-1/\alpha_n} f^n((1 - t)^{-1/\alpha_n}(y - x))$$

converges to the obviously defined $g(t, x, y)$ continuously on the domain $t \in (0, 1), x \geq 0, y \in \mathbb{R}$, which follows from continuous convergence of the density f^n of X_1^n, see Lemma 10 below. Hence we have weak convergence of the quantity under the expectation in (20), and so it is left to show that the respective quantities are bounded. Lemma 10 completes the proof.

Lemma 10. *There is the uniform convergence:* $\sup_{x \in \mathbb{R}} |f^n(x) - f(x)| \to 0$ *as* $n \to \infty$. *Moreover, for any* $\epsilon > 0$ *it holds that*

$$\sup_n \sup_{t \in (0,1), x \geq \epsilon} f_t^n(x) < \infty.$$

Proof. The characteristic function of X_t^n is given by $\exp(-c_n^{\pm}|z|^{\alpha_n} t)$ according to $\pm z > 0$ with $c_n^{\pm}$ being a complex constant with positive real part (converging to $c^{\pm}$), see [43, Thm. C.4]. Thus by inversion formula we have

$$\sup_{x \in \mathbb{R}} |f^n(x) - f(x)| \leq \frac{1}{2\pi} \int |\exp(-c_n^{\pm}|z|^{\alpha_n} t) - \exp(-c^{\pm}|z|^{\alpha} t)| \mathrm{d}z,$$

but this converges to 0 by the dominated convergence theorem, since the real parts of $c_n^{\pm}$ are positive and bounded away from 0.

With respect to the second statement we need to show that

$$\int_1^\infty \exp(-izx - c_n z^{\alpha_n} t) \mathrm{d}z$$

is bounded for all $t \in (0, 1), x \geq \epsilon$ and all n, where $c_n = c_n^+$; the integral over $(-\infty, -1]$ is hadled in the same way, whereas the rest is clearly bounded by 2. Using integration by parts we find that it is sufficient to show that

$$\int_1^\infty \frac{c_n t \alpha_n z^{\alpha_n - 1}}{ix} \exp(-izx - c_n z^{\alpha_n} t) \mathrm{d}z$$

is bounded, or equivalently the boundedness of

$$\int_1^\infty \alpha_n z^{\alpha_n - 1} t \exp(-r_n z^{\alpha_n} t) \mathrm{d}z = \int_t^\infty \exp(-r_n z) \mathrm{d}z \leq \frac{1}{r_n},$$

where $r_n = \Re(c_n)$. But $r_n \to r > 0$ and we are done.   □

Appendix B: Proofs for local and occupation times

Here X denotes a linear Brownian motion with drift parameter $\mu \in \mathbb{R}$ and scale $\sigma > 0$.

Proof of Lemma 2. The fact that $\mathbb{E}\left[L_t(x)|X_t = z\right]$ does not depend on μ follows readily by applying exponential change of measure, for example. Thus we may assume that $\mu = 0$ and consider the process σX with X being the standard Brownian motion. Using self-similarity of X we find

$$\left(\frac{1}{2\epsilon} \int_0^t 1_{(x-\epsilon, x+\epsilon)}(\sigma X_s)\mathrm{d}s, \sigma X_t\right) = \left(\frac{t}{2\epsilon} \int_0^1 1_{(x-\epsilon, x+\epsilon)}(\sigma X_{ts})\mathrm{d}s, \sigma X_t\right)$$

$$\stackrel{d}{=} \left(\frac{t}{2\epsilon} \int_0^1 1_{(x-\epsilon, x+\epsilon)}(\sigma\sqrt{t}X_s)\mathrm{d}s, \sigma\sqrt{t}X_1\right)$$

and we readily find the stated expression for $\mathbb{E}\left[L_t(x)|X_t = z\right]$ from the definition of L. For further reference let us also note that

$$(L_t(x), X_t) \stackrel{d}{=} (\sqrt{t}L_1(x/\sqrt{t}), \sqrt{t}X_1) \qquad \text{under } \mathbb{P}^0. \tag{32}$$

The formula for $\mathbb{E}\left[O_t(x)|X_t = z\right]$ is obtained similarly, or directly from (14).

Next, we note that $g(x, z) = g(-x, -z), G(x, z) = 1 - G(-x, -z)$ follow easily from symmetry, and so we assume in the following that $x \geq 0$. From [15, 1.3.8] we find

$$g(x, z) = \exp(z^2/2) \int_0^\infty y(|z - x| + |x| + y) \exp(-(|z - x| + |x| + y)^2/2)\mathrm{d}y$$

which indeed evaluates to the given expression. Next, we recall the Mill's ratio: $\overline{\Phi}(z)/\varphi(z) \sim 1/z$ as $z \to \infty$. Hence

$$g(x, z) \sim \frac{1}{|z - x| + |x|} \exp(-(|z - x| + |x|)^2/2 + z^2/2) \qquad \text{as } |x| \vee |z| \to \infty, \tag{33}$$

showing that $g(x, z)$ is bounded since $|z - x| + |x| \geq |z|$.

Finally, $G(x, z)$ is clearly bounded by 1 and the given formulae are found from the occupation density formula $G(x, z) = \int_x^\infty g(y, z)\mathrm{d}y$, see (14).   □

B.1. Local time

Proof of (16). Firstly, we may replace t by $\lfloor tn \rfloor / n$ on the left hand side of (16), see [30, Rem. 2]. The result would follow from [30, Thm. 2.1] if we show that g satisfies condition [30, (B-r)] for some $r > 3$. But this follows from the bound $g(x, z) < c \exp(-2|x| + 2|z|)$ for all $x, z \in \mathbb{R}$, see the proof of Lemma 13.

Now we have the stated convergence, but the constant in front of the limit needs to be identified. The expressions in [30] are lengthy and non-trivial to evaluate, because of the generality assumed therein. In our case, $g(x, X_1) = \mathbb{E}[L_1(x) | X_1]$ is the conditional expectation and, in fact, a rather short direct proof can be given yielding the constant.

Direct Proof: As in [30] we observe that it is sufficient to consider the case $\mu = 0$, which can be extended to an arbitrary μ using change of measure argument. Importantly, $\widehat{L}_t(x)$ is a functional of X and this functional does not depend on μ. Next, consider a standard Brownian motion $X_t^0 = X_t / \sigma$ and assume that our result is proven for X^0. Noting that $L_t(x) = \frac{1}{\sigma} L_t^0(x/\sigma)$ as well as $\widehat{L}_t(x) = \frac{1}{\sigma} \widehat{L}_t^0(x/\sigma)$ we find that

$$
n^{1/4} \left(\widehat{L}_t(x) - L_t(x) \right) = \frac{1}{\sigma} n^{1/4} \left(\widehat{L}_t^0(x/\sigma) - L_t^0(x/\sigma) \right)
$$
$$
\xrightarrow{d_{st}} \frac{v_l}{\sigma} W_{L_t^0(x/\sigma)} = \frac{v_l}{\sigma} W_{\sigma L_t(x)}.
$$

It is left to replace the process $W_{\sigma t}$ by $\sqrt{\sigma} W_t$ having the same law. Thus we may assume in the following that X is a standard Brownian motion.

Let $S_t^n = \sum_{i=1}^{\lfloor tn \rfloor} \xi_{in}$ be the pre-limiting object, where

$$
\xi_{in} = n^{-1/4} \left(g(\sqrt{n}(x - X_{\frac{i-1}{n}}), \sqrt{n} \Delta_i^n X) - \sqrt{n} L_{[\frac{i-1}{n}, \frac{i}{n}]}(x) \right),
$$

$\Delta_i^n X = X_{i/n} - X_{(i-1)/n}$ and $L_{[a,b]}(x)$ denotes the local time at x in the interval $[a, b]$. Firstly, observe using the scaling property (32) that

$$
h_1(x) := \mathbb{E} \left(g(x, \sqrt{n} X_{1/n}) - \sqrt{n} L_{1/n}(x/\sqrt{n}) \right) = \mathbb{E} \left(g(x, X_1) - L_1(x) \right) = 0.
$$

Thus we have

$$
\mathbb{E}[\xi_{in} | \mathcal{F}_{\frac{i-1}{n}}] = n^{-1/4} h_1(\sqrt{n}(x - X_{\frac{i-1}{n}})) = 0,
$$

and similarly we find that

$$
\mathbb{E}[\xi_{in}^2 | \mathcal{F}_{\frac{i-1}{n}}] = n^{-1/2} h_2(\sqrt{n}(x - X_{\frac{i-1}{n}})),
$$
$$
\mathbb{E}[\xi_{in} \Delta_i^n X | \mathcal{F}_{\frac{i-1}{n}}] = n^{-3/4} h_3(\sqrt{n}(x - X_{\frac{i-1}{n}})) = 0,
$$
$$
\mathbb{E}[\xi_{in}^4 | \mathcal{F}_{\frac{i-1}{n}}] = n^{-1} h_4(\sqrt{n}(x - X_{\frac{i-1}{n}})),
$$

where $h_i(y) = \mathbb{E}(g(y, X_1) - L_1(y))^i$ for $i = 2, 4$, and $h_3(y) = \mathbb{E}[(g(y, X_1) - L_1(y)) X_1] = 0$.

Let us show that h_i for $i = 2, 4$ are bounded and in $L^1(\mathbb{R})$. By Minkowski's and Jensen's inequality we have the bound $h_i(y) \leq 2^i \mathbb{E}[L_1(y)^i]$. Using additivity of L we deduce that

$$\mathbb{E}[L_1(y)^i] \leq \mathbb{P}(\tau_y < 1)\mathbb{E}[L_1(0)^i],$$

where the latter moment is finite and τ_y is the first passage time of X into the level y. Finally, note that

$$\int_0^\infty \mathbb{P}(\tau_y < 1)\mathrm{d}y = \int_0^\infty \mathbb{P}(\overline{X}_1 > y)\mathrm{d}y = \mathbb{E}\overline{X}_1 < \infty$$

and hence by symmetry $h_i(y)$ are integrable. Thus according to [30, Thm. 1.1] we have

$$n^{-1/2} \sum_{i=1}^{\lfloor nt \rfloor} h_i(\sqrt{n}(x - X_{\frac{i-1}{n}})) \xrightarrow{\mathbb{P}} L_t(x) \int h_i(x)\mathrm{d}x, \qquad i = 2, 4,$$

where the convergence is uniform on compact intervals of time. This immediately yields that

$$\sum_{i=1}^{\lfloor nt \rfloor} \mathbb{E}[\xi_{in}^2 | \mathcal{F}_{\frac{i-1}{n}}] \xrightarrow{\mathbb{P}} v_l^2 L_t(x), \qquad \sum_{i=1}^{\lfloor nt \rfloor} \mathbb{E}[\xi_{in} \Delta_i^n X | \mathcal{F}_{\frac{i-1}{n}}] = 0,$$

$$\sum_{i=1}^{\lfloor nt \rfloor} \mathbb{E}[\xi_{in}^2 1_{\{|\xi_{in}| > \epsilon\}} | \mathcal{F}_{\frac{i-1}{n}}] \leq \epsilon^{-2} \sum_{i=1}^{\lfloor nt \rfloor} \mathbb{E}[\xi_{in}^4 | \mathcal{F}_{\frac{i-1}{n}}] \xrightarrow{\mathbb{P}} 0 \qquad \text{for any } \epsilon > 0.$$

Finally, let N be a continuous bounded martingale orthogonal to X, i.e. $[X, N] = 0$. For $t \geq (i-1)/n$ define the process $M_t = \mathbb{E}[\xi_{in} | \mathcal{F}_t]$. Then the martingale representation theorem implies the existence of a progressively measurable process η^n such that

$$M_t = \int_{\frac{i-1}{n}}^t \eta_s^n dX_s.$$

Since $[X, N] = 0$ we conclude that

$$\mathbb{E}[\Delta_i^n N \xi_{in} | \mathcal{F}_{\frac{i-1}{n}}] = \mathbb{E}[\Delta_i^n N \Delta_i^n M | \mathcal{F}_{\frac{i-1}{n}}] = 0.$$

The result now follows from [32, Thm. IX.7.28]. Moreover, we have a simple expression for $v_l^2 = \int h_2(y)\mathrm{d}y$ which is evaluated in Lemma 11 below. $\qquad \square$

It is left to calculate v_l^2, which is the integrated reduction in variance when $L_1(y)$ is replaced by its conditional mean $\mathbb{E}[L_1(y)|X_1]$:

Lemma 11. *For a standard Brownian motion we have*

$$\int_{\mathbb{R}} \mathbb{E}[(g(y, X_1) - L_1(y))^2]\mathrm{d}y = 2\frac{3\log(1 + \sqrt{2}) - \sqrt{2}}{3\sqrt{\pi}}.$$

Proof. Recalling that $g(y, X_1) = \mathbb{E}[L_1(y)|X_1]$ we find

$$\int_{\mathbb{R}} \left(\mathbb{E}[L_1^2(y)] - \mathbb{E}[g^2(y, X_1)] \right) \mathrm{d}y = 2 \int_0^\infty \left(\mathbb{E}[L_1^2(y)] - \mathbb{E}[g^2(y, X_1)] \right) \mathrm{d}y.$$

According to [15, 1.3.4] we calculate

$$\int_0^\infty \mathbb{E}[L_1^2(y)]\mathrm{d}y = \int_0^\infty \int_0^\infty x^2 \sqrt{\frac{2}{\pi}} \exp(-(x+y)^2/2)\mathrm{d}x\mathrm{d}y = \frac{2}{3}\sqrt{\frac{2}{\pi}},$$

and

$$\int_0^\infty \mathbb{E}[g^2(y, X_1)]\mathrm{d}y = \int_0^\infty \int_{\mathbb{R}} \overline{\Phi}^2(|z - y| + y)/\varphi(z)\mathrm{d}z\mathrm{d}y = \frac{\sqrt{2} - \log(1 + \sqrt{2})}{\sqrt{\pi}},$$

where in both cases we first integrate in $y > 0$. Combine these formulae to get the result. □

B.2. Occupation time

Proof of (17). We may assume that $\mu = 0$ and let $X_t^0 = X_t/\sigma$. Supposing that the result is true for X^0 we get

$$n^{\frac{3}{4}} \left(\widehat{O}_t(x) - O_t(x) \right) = n^{\frac{3}{4}} \left(\widehat{O}_t^0(x/\sigma) - O_t^0(x/\sigma) \right) \rightarrow v_o W_{L_t^0(x/\sigma)} = v_o W_{\sigma L_t(x)}$$

and so we assume that X is a standard Brownian motion in the following.

Letting

$$\xi_{in} = n^{-\frac{1}{4}} \left(G \left(\sqrt{n}(x - X_{\frac{i-1}{n}}), \sqrt{n}\Delta_i^n X \right) - n \int_{\frac{i-1}{n}}^{\frac{i}{n}} 1_{(x,\infty)}(X_s)ds \right)$$

and using

$$(nO_{1/n}(x/\sqrt{n}), \sqrt{n}X_{1/n}) \overset{d}{=} (O_1(x), X_1)$$

we find that

$$\mathbb{E}[\xi_{in}^2|\mathcal{F}_{\frac{i-1}{n}}] = n^{-1/2}h_2(\sqrt{n}(x - X_{\frac{i-1}{n}})),$$
$$\mathbb{E}[\xi_{in}\Delta_i^n X|\mathcal{F}_{\frac{i-1}{n}}] = 0,$$
$$\mathbb{E}[\xi_{in}^4|\mathcal{F}_{\frac{i-1}{n}}] = n^{-1}h_4(\sqrt{n}(x - X_{\frac{i-1}{n}})),$$

where $h_j(y) = \mathbb{E}[G(y, X_1) - O_1(y)]^j$ for $j = 2, 4$.

It is left to prove that h_j are bounded and in $L^1(\mathbb{R})$ for $j = 2, 4$. The result then follows from [30, Thm. 1.1] and [32, Thm. IX.7.28] as for the local time. It would be sufficient to show the same property for $\mathbb{E}[(O_1(y) - c_y)^j]$ where c_y is arbitrary, because $G(y, X_1) - c_y$ is the conditional expectation of $O_1(y) - c_y$ given X_1. When $y \geq 0$ we take $c_y = 0$ and observe that $\mathbb{E}[O_1(y)^j] \leq \mathbb{P}(\tau_y < 1)$ which is bounded and integrable over $[0, \infty)$, see the local time case. When $y < 0$ we take $c_y = 1$ and observe that $\mathbb{E}[(1 - O_1(y))^j] \leq \mathbb{P}(\tau_y < 1)$ and the same conclusion is true. The proof is complete upon calculation of v_o^2 which is given in Lemma 12 below. □

Lemma 12. *For a standard Brownian motion we have*

$$\int_{\mathbb{R}} \mathbb{E}\left[G(y, X_1) - O_1(y)\right]^2 \mathrm{d}y = \frac{13\sqrt{2} - 15\log(1 + \sqrt{2})}{45\sqrt{\pi}}.$$

Proof. Note that

$$\int_{\mathbb{R}} \mathbb{E}\left[G(y, X_1) - O_1(y)\right]^2 \mathrm{d}y = 2\int_0^\infty (\mathbb{E}O_1(y)^2 - \mathbb{E}G(y, X_1)^2)\mathrm{d}y,$$

because for $y < 0$ the integrand can be rewritten as $\mathbb{E}[(1 - O_1(y))^2] - \mathbb{E}[(1 - G(y, X_1))^2]$ corresponding to the occupation time in $(-\infty, y)$ and its conditional expectation, and it is left to apply symmetry.

The density of the occupation time $O_1(y)$ is given in [15, 1.4.4] and reads as

$$\frac{1}{\pi\sqrt{x(1 - x)}}\exp\left(-\frac{y^2}{2(1 - x)}\right), \qquad x \in (0, 1).$$

Thus we find $\int_0^\infty \mathbb{E}[O_1(y)^2]\mathrm{d}y = \frac{\sqrt{2}}{5\sqrt{\pi}}$ by integrating in y first.

Similar trick works in the calculation of

$$\int_0^\infty \mathbb{E}[G^2(y, X_1)]\mathrm{d}y = \frac{\sqrt{2} + 3\log(1 + \sqrt{2})}{18\sqrt{\pi}}.$$

Combination of these expressions yields the result. $\qquad\qquad\square$

B.3. Unknown parameters

Let us define $g_\sigma(x, z) = \frac{1}{\sigma}g(x/\sigma, z/\sigma)$ together with $G_\sigma(x, z) = G(x/\sigma, z/\sigma)$.

Lemma 13. *For any $\sigma_0 > 0$ there exist constants $\epsilon \in (0, \sigma_0)$ and $c, a > 0$ such that*

$$\sup_{\sigma \in [\sigma_0 - \epsilon, \sigma_0 + \epsilon]}\left|\frac{\partial g_\sigma(x, z)}{\partial \sigma}\right| \vee \left|\frac{\partial G_\sigma(x, z)}{\partial \sigma}\right| \leq c\exp(a(|z| - |x|))$$

for all $x, z \in \mathbb{R}$.

Proof. Recall that $g(x, z) = g(-x, -z), G(x, z) = 1 - G(-x, -z)$ and so we may assume that $x \geq 0$. Furthermore, it is sufficient to establish the stated property for $\partial g_\sigma/\partial\sigma$. This is so, because $G_\sigma(x, z) = \int_x^\infty g_\sigma(y, z)\mathrm{d}y$, the derivative $\partial g_\sigma/\partial\sigma$ is continuous in σ away from 0 and integrable in $y \geq 0$. Hence

$$\frac{\partial G_\sigma(x, z)}{\partial \sigma} = \int_x^\infty \frac{\partial g_\sigma(y, z)}{\partial \sigma}\mathrm{d}y \leq c\int_x^\infty \exp(-ay)\mathrm{d}y\exp(a|z|)$$

and the bound follows.

It is sufficient to establish the bound for $x \geq 0$:

$$\left|\frac{\partial g(x/\sigma, z/\sigma)}{\partial \sigma}\right| \leq c(\exp(a(|z| - x)) \wedge 1)$$

locally uniformly in $\sigma > 0$. This is so, because $g(x/\sigma, z/\sigma)/\sigma^2$ satisfies the analogous bound, see (33).

Writing x', z' for x/σ, z/σ, respectively, we find from Lemma 2 for $z \geq x$ that

$$\partial g(x/\sigma, z/\sigma)/\partial\sigma = z'\frac{\varphi(z') - z'\overline{\Phi}(z')}{2\sigma\varphi(z')} =: h(z').$$

By L'Hôpitale and Mill's ratio this quantity tends to 0 as $z' \to \infty$, and thus this quantity is bounded for all $z \geq x \geq 0$ locally uniformly in $\sigma > 0$.

Next, we consider $z < x$ where

$$\partial g(x/\sigma, z/\sigma)/\partial\sigma = \frac{(2x' - z')\varphi(2x' - z') - z'^2\overline{\Phi}(2x' - z')}{2\sigma\varphi(z')}$$

$$= \frac{2x'(x' - z')}{\sigma(2x' - z')}\exp(-2x'(x' - z')) + \frac{z'^2}{(2x' - z')^2}h(2x' - z')\exp(-2x'(x' - z')).$$

Note that $2x' - z' > (x' - z') \vee |z'|$ and so the above terms stay bounded when $2x' - z' \to 0$ implying that $x', z' \to 0$. Moreover, $z'^2/(2x' - z')^2$ is bounded and so it is left to consider $(1 + 2x'(x' - z'))\exp(-2x'(x' - z'))$ as $x' \to \infty$. For $x' > z' + 1$ this is bounded by $c\exp(-x')$ and otherwise by c, which is sufficient. □

Proof of Proposition 2. Observe that

$$n^{1/4}\sup_{t \leq T}\left|\widehat{L}_t(x) - \widetilde{L}_t(x)\right|$$

$$\leq n^{-1/4}\sum_{i=1}^{\lfloor nT \rfloor}\left|g_\sigma(\sqrt{n}(x - X_{\frac{i-1}{n}}), \sqrt{n}\Delta_i^n X) - g_{\sigma_n}(\sqrt{n}(x - X_{\frac{i-1}{n}}), \sqrt{n}\Delta_i^n X)\right|.$$

According to (18) we may assume that $n^{1/4}|\sigma_n - \sigma| < h$ for an arbitrary $h > 0$ and all large n. By mean value theorem and Lemma 13 we have an upper bound

$$n^{-1/4}\sum_{i=1}^{\lfloor nT \rfloor}|\sigma_n - \sigma|\tilde{g}(\sqrt{n}(x - X_{\frac{i-1}{n}}), \sqrt{n}\Delta_i^n X)$$

$$\leq hn^{-1/2}\sum_{i=1}^{\lfloor nT \rfloor}\tilde{g}(\sqrt{n}(x - X_{\frac{i-1}{n}}), \sqrt{n}\Delta_i^n X),$$

where $\tilde{g}(x, z) = c\exp(-a|x| + a|z|)$. But $\tilde{g}$ verifies condition (B-0) in [30] and thus our upper bound converges to hL in probability, where L is a certain finite random variable, see [30, Thm. 1.1]. The proof is complete since $h > 0$ can be arbitrarily small. The corresponding proof for the occupation time follows exactly the same arguments. □

Appendix C: On X conditioned to stay positive

Throughout this section we assume that $\alpha \in (0,2)$ and $\beta \neq \pm 1$. Let us recall that $(\xi_{(-t)-})_{t \geq 0}$ is a Feller process and, as usual, we denote its law when started from $x > 0$ by $\mathbb{P}_x^{\uparrow}((X_t)_{t \geq 0} \in \cdot)$. Such a process can be seen as X conditioned to stay positive in a certain limiting sense, see [19, 16] for the basic properties of this process. The law of $(\xi_t)_{t \geq 0}$ is then $(-X)$ conditioned to stay positive, and the following bound holds without a change.

Proposition 5. *There exists a constant $c > 0$ such that for all $x, v > 0$ with $x > v$ we have*

$$\mathbb{P}_x^{\uparrow}\left(\sup_{h \in [0,1]} |X_{1+h} - X_h| > v\right) < cv^{-\alpha}.$$

The proof will be at the end of this section. Let us note that the restriction $x > v$ can not be removed in the above bound. We start with a simpler result where $h = 0$:

Lemma 14. *There exists $c > 0$ such that for all $x > v > 0$ we have*

$$\mathbb{P}_x^{\uparrow}(|X_1 - x| > v) < cv^{-\alpha}.$$

Proof. Let $\rho = \mathbb{P}(X_1 < 0)$ be the negativity parameter. Recall the semigroup of the conditioned process [16]:

$$\mathbb{P}_x^{\uparrow}(X_1 \in \mathrm{d}y) = \frac{y^{\alpha\rho}}{x^{\alpha\rho}}\mathbb{P}_x(X_1 \in \mathrm{d}y, \underline{X}_1 > 0).$$

Hence

$$
\begin{aligned}
\mathbb{P}_x^{\uparrow}(|X_1 - x| > v) &= \frac{1}{x^{\alpha\rho}}\mathbb{E}_x[X_1^{\alpha\rho}; |X_1 - x| > v, \underline{X}_1 > 0] \\
&\leq \frac{1}{x^{\alpha\rho}}\mathbb{E}[(X_1 + x)^{\alpha\rho}; |X_1| > v, X_1 > -x] \\
&= \frac{1}{x^{\alpha\rho}}\left(\int_v^{\infty} (x+y)^{\alpha\rho}f(y)\mathrm{d}y + \int_{-x}^{-v} (x+y)^{\alpha\rho}f(y)\mathrm{d}y\right).
\end{aligned}
$$

Recall that $f(y) \leq c|y|^{-\alpha-1}$ as $y \to \pm\infty$, and hence the first integral is upper bounded by

$$2^{\alpha\rho}c\int_x^{\infty} y^{\alpha\rho-\alpha-1}\mathrm{d}y + (2x)^{\alpha\rho}c\int_v^x y^{-\alpha-1}\mathrm{d}y \leq cx^{\alpha\rho}v^{-\alpha}$$

and the second has a similar bound. The result now follows. $\qquad \square$

The following is an immediate consequence of the Doob's h-transform representation of the kernel; here $h(x) = x^{\alpha\rho}$.

Lemma 15. *For any $B \in \mathcal{F}_1$ it holds that*

$$\mathbb{P}_x^{\uparrow}(B, X_1 \in \mathrm{d}y) = \mathbb{P}_x^{\uparrow}(X_1 \in \mathrm{d}y)\mathbb{P}_x(B|\underline{X}_1 > 0, X_1 = y)$$

Proof. For $0 < t_1 < \cdots < t_k < 1$ we have

$$
\begin{aligned}
&\mathbb{P}_x^\uparrow(X_{t_1} \in \mathrm{d}x_1, \ldots, X_{t_k} \in \mathrm{d}x_k, X_1 \in \mathrm{d}y) \\
&= \frac{h(x_1)}{h(x)}\mathbb{P}_x(X_{t_1} \in \mathrm{d}x_1, \underline{X}_{t_1} > 0) \times \cdots \times \frac{h(y)}{h(x_k)}\mathbb{P}_{x_k}(X_{1-t_k} \in \mathrm{d}y, \underline{X}_{1-t_k} > 0) \\
&= \frac{h(y)}{h(x)}\mathbb{P}_x(X_{t_1} \in \mathrm{d}x_1, \ldots, X_{t_k} \in \mathrm{d}x_k, X_1 \in \mathrm{d}y, \underline{X}_1 > 0) \\
&= \mathbb{P}_x^\uparrow(X_1 \in \mathrm{d}y)\mathbb{P}_x(X_{t_1} \in \mathrm{d}x_1, \ldots, X_{t_k} \in \mathrm{d}x_k | X_1 = y, \underline{X}_1 > 0)
\end{aligned}
$$

and the result follows. $\qquad\square$

Lemma 16. *There exists $c > 0$ such that for all $x > v > 0$ we have*

$$
\mathbb{P}_x^\uparrow(\overline{X}_1 - x > v) < cv^{-\alpha}, \qquad \mathbb{P}_x^\uparrow(x - \underline{X}_1 > v) < cv^{-\alpha}.
$$

Proof. We only show the first statement, since the second follows the same arguments. According to Lemma 15 we find that

$$
\mathbb{P}_x^\uparrow(\overline{X}_1 - x > v) = \int \mathbb{P}_x^\uparrow(X_1 \in x + \mathrm{d}y)\mathbb{P}_x(\overline{X}_1 - x > v | \underline{X}_1 > 0, X_1 = x + y).
$$

We may restrict the integration to the interval $[-v/2, v/2]$ in view of Lemma 14. Thus it is sufficient to establish that

$$
\mathbb{P}(\overline{X}_1 > v | \underline{X}_1 > -x, X_1 = y) < cv^{-\alpha}
$$

for all $x > v$ and $y \in [-v/2, v/2]$. But the quantity on the left is upper bounded by

$$
\overline{F}(v, y)/\mathbb{P}(\underline{X}_1 > -x | X_1 = y),
$$

where $\overline{F}(v, y) \leq cv^{-\alpha}$ according to Lemma 4; for bounded v the result is obvious. Finally, observe that $\mathbb{P}(\underline{X}_1 > -x | X_1 = y)$ is bounded away from 0; here we may use Lemma 4 applied to the process $-X$. The proof is complete. $\qquad\square$

Proof of Proposition 5. Observe that the quantity of interest is upper bounded by

$$
\mathbb{P}_x^\uparrow(\overline{X}_2 - x > v/2 \text{ or } x - \underline{X}_2 > v/2).
$$

Hence the bound follows from Lemma 16, which also holds for time 2 instead of 1; use e.g. self-similarity here. $\qquad\square$

Acknowledgments

References

[1] Y. Aït-Sahalia and J. Jacod. *High-frequency financial econometrics.* Princeton University Press, 2014.

[2] D. J. Aldous and G. K. Eagleson. On mixing and stability of limit theorems. *Ann. Probability*, 6(2):325–331, 1978. MR0517416

[3] R. Altmeyer. Central limit theorems for discretized occupation time functionals. *Preprint arXiv:1909.00474*, 2019.

[4] R. Altmeyer. Estimating occupation time functionals. *Preprint arXiv:1706.03418*, 2019.

[5] R. Altmeyer and J. Chorowski. Estimation error for occupation time functionals of stationary Markov processes. *Stochastic Process. Appl.*, 128(6):1830–1848, 2018. MR3797645

[6] S. Asmussen, P. Glynn, and J. Pitman. Discretization error in simulation of one-dimensional reflecting Brownian motion. *Ann. in Appl. Probab.*, pages 875–896, 1995. MR1384357

[7] S. Asmussen and J. Ivanovs. Discretization error for a two-sided reflected Lévy process. *Queueing Syst.*, 89(1-2):199–212, 2018. MR3803961

[8] A. Banerjee, X. Guo, and H. Wang. On the optimality of conditional expectation as a Bregman predictor. *IEEE Trans. Inform. Theory*, 51(7):2664–2669, 2005. MR2246384

[9] J. Bertoin. Splitting at the infimum and excursions in half-lines for random walks and Lévy processes. *Stochastic Process. Appl.*, 47(1):17–35, 1993. MR1232850

[10] J. Bertoin. *Lévy processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996. MR1406564

[11] P. Billingsley. *Convergence of probability measures.* John Wiley & Sons, 1999. 2nd edition. MR1700749

[12] N. H. Bingham. Maxima of sums of random variables and suprema of stable processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 26:273–296, 1973. MR0415780

[13] K. Bisewski and J. Ivanovs. Zooming-in on a Lévy process: Failure to observe threshold exceedance over a dense grid. *Preprint arXiv:1904.06162*, 2019. MR4161123

[14] A. N. Borodin. On the character of convergence to Brownian local time. I, II. *Probab. Theory Relat. Fields*, 72(2):231–250, 251–277, 1986. MR0836277

[15] A. N. Borodin and P. Salminen. *Handbook of Brownian motion—facts and formulae.* Probability and its Applications. Birkhäuser Verlag, Basel, second edition, 2002. MR1912205

[16] M. E. Caballero and L. Chaumont. Conditioned stable Lévy processes and the Lamperti representation. *Journal of Applied Probability*, 43(4):967–983, 2006. MR2274630

[17] J. G. Cázares, A. Mijatović, and G. U. Bravo. Geometrically convergent simulation of the extrema of Lévy processes. *Preprint arXiv:1810.11039*, 2018. MR4032169

[18] L. Chaumont. On the law of the supremum of Lévy processes. *Ann. Probab.*,

41(3A):1191–1217, 2013. MR3098676

[19] L. Chaumont and R. A. Doney. On Lévy processes conditioned to stay positive. *Electron. J. Probab.*, 10(28):948–961, 2005. MR2164035

[20] L. Chaumont and J. C. Pardo. The lower envelope of positive self-similar Markov processes. *Electron. J. Probab.*, 11:no. 49, 1321–1341, 2006. MR2268546

[21] E. Clément, S. Delattre, and A. Gloter. An infinite dimensional convolution theorem with applications to the efficient estimation of the integrated volatility. *Stochastic Process. Appl.*, 123(7):2500–2521, 2013. MR3054534

[22] E. Clément, S. Delattre, and A. Gloter. Asymptotic lower bounds in estimating jumps. *Bernoulli*, 20(3):1059–1096, 2014. MR3217438

[23] R. A. Doney and M. S. Savov. The asymptotic behavior of densities related to the supremum of a stable process. *Ann. Probab.*, 38(1):316–326, 2010. MR2599201

[24] D. Florens-Zmirou. On estimating the diffusion coefficient from discrete observations. *J. Appl. Probab.*, 30(4):790–804, 1993. MR1242012

[25] S. Fourati. Inversion de l'espace et du temps des processus de Lévy stables. *Probab. Theory Related Fields*, 135(2):201–215, 2006. MR2218871

[26] T. Gneiting. Making and evaluating point forecasts. *J. Amer. Statist. Assoc.*, 106(494):746–762, 2011. MR2847988

[27] J. I. González Cázares, A. Mijatović, and G. U. Bravo. Exact simulation of the extrema of stable processes. *Adv. in Appl. Probab.*, 51(4):967–993, 2019. MR4032169

[28] Y. Hu and C. Lee. Drift parameter estimation for a reflected fractional Brownian motion based on its local time. *J. Appl. Probab.*, 50(2):592–597, 2013. MR3102502

[29] J. Ivanovs. Zooming in on a Lévy process at its supremum. *The Annals of Applied Probability*, 28(2):912–940, 2017. MR3784492

[30] J. Jacod. Rates of convergence to the local time of a diffusion. *Annales de l'I.H.P. (B)*, 34(4):505–544, 1998. MR1632849

[31] J. Jacod and P. Protter. *Discretization of processes*, volume 67 of *Stochastic Modelling and Applied Probability*. Springer, Heidelberg, 2012. MR2859096

[32] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1987. MR0959133

[33] P. Kosulajeff. Sur la répartition de la partie fractionnaire d'une variable. *Math. Sbornik*, 2(5):1017–1019, 1937.

[34] A. Lejay. Estimation of the bias parameter of the skew random walk and application to the skew Brownian motion. *Stat. Inference Stoch. Process.*, 21(3):539–551, 2018. MR3846991

[35] H. Masuda. Parametric estimation of Lévy processes. In *Lévy Matters IV*, pages 179–286. Springer, 2015. MR3364258

[36] Z. Michna. Explicit formula for the supremum distribution of a spectrally negative stable process. *Electron. Commun. Probab.*, 18:no. 10, 6, 2013. MR3033593

[37] M. Motoo. Proof of the iterated logarithm through diffusion equation. *Ann. Inst. Statist. Math.*, 10:21–28, 1959. MR0097866

[38] H.-L. Ngo and S. Ogawa. On the discrete approximation of occupation time of diffusion processes. *Electron. J. Stat.*, 5:1374–1393, 2011. MR2842909

[39] A. Rényi. On stable sequences of events. *Sankhyā Ser. A*, 25:293 302, 1963. MR0170385

[40] S. I. Resnick. *Extreme values, regular variation and point processes*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2008. Reprint of the 1987 original. MR2364939

[41] K.-I. Sato. *Lévy processes and infinitely divisible distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2013. MR3185174

[42] L. A. Shepp. The joint density of the maximum and its location for a Wiener process with drift. *J. Appl. Probab.*, 16(2):423–427, 1979. MR0531776

[43] V. M. Zolotarev. *One-dimensional stable distributions*, volume 65 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1986. Translated from the Russian by H. H. McFaden, Translation edited by Ben Silver. MR0854867