INFERENCE ON THE INTRADAY SPOT VOLATILITY FROM HIGH-FREQUENCY ORDER PRICES WITH IRREGULAR MICROSTRUCTURE NOISE

MARKUS BIBINGER,* *Julius-Maximilians-Universität Würzburg*

Abstract

We consider estimation of the spot volatility in a stochastic boundary model with one-sided microstructure noise for high-frequency limit order prices. Based on discrete, noisy observations of an Itô semimartingale with jumps and general stochastic volatility, we present a simple and explicit estimator using local order statistics. We establish consistency and stable central limit theorems as asymptotic properties. The asymptotic analysis builds upon an expansion of tail probabilities for the order statistics based on a generalized arcsine law. In order to use the involved distribution of local order statistics for a bias correction, an efficient numerical algorithm is developed. We demonstrate the finite-sample performance of the estimation in a Monte Carlo simulation.

Keywords: Arcsine law; limit order book; market microstructure; nonparametric boundary model; volatility estimation

2020 Mathematics Subject Classification: Primary 62M09
Secondary 60J65; 60F05

1. Introduction

Time series of intraday prices are typically described as a discretized path of a continuous-time stochastic process. To have arbitrage-free markets the log-price process should be a semimartingale. Risk estimation based on high-frequency data at the highest available observation frequencies has to take microstructure frictions into account. Disentangling these market microstructure effects from the dynamics of the long-run price evolution has led to observation models with additive noise; see, for instance, [2, 13, 19]. The market microstructure noise, modelling among other effects the oscillation of traded prices between bid and ask order levels in an electronic market, is classically a centred (white) noise process with expectation equal to zero. These models can explain many stylized facts of high-frequency data. Having available full limit order books including data of submissions, cancellations, and executions of bid and ask limit orders, however, it is not clear which time series to consider at all. While challenging the concept of one price process it raises the question of whether the information can be exploited more efficiently, in particular to improve risk quantification. The stochastic boundary model considered for limit order prices of an order book has been discussed by [5], [20], and [8, Chapter 1.8]. It preserves the concept of an underlying efficient, semimartingale log-price

which determines the long-run price dynamics and an additive, exogenous noise which models market-specific microstructure frictions. Its key idea is that ask order prices should (in most cases) lie above the unobservable efficient price and bid prices below the efficient price. This leads to observation errors which are irregular in the sense of having non-zero expectation and a distribution with a lower- or upper-bounded support. Considering without loss of generality a model for (best) ask order prices, we obtain lower-bounded observation errors and use local minima for the estimation. Modelling (best) bid prices instead would yield a model with upper-bounded observation errors and local maxima could be used for an analogous estimation. Both can be combined in practice.

It is known that the statistical and probabilistic properties of models with irregular noise are very different than for regular noise and require other methods; see, for instance, [17, 23, 24]. Therefore, our estimation methods and asymptotic theory are quite different compared to the market microstructure literature, while we can still profit from some of the techniques used there. In [5] an estimator for the quadratic variation of a continuous semimartingale, that is, the integrated volatility, was proposed with convergence rate $n^{-1/3}$, based on n discrete observations with one-sided noise. Optimality of the rate was proved in the standard asymptotic minimax sense. The main insight was that this convergence rate is better than the optimal rate, $n^{-1/4}$, under regular market microstructure noise.

A recent strand of literature proposes structural, parametric market microstructure noise models incorporating information based on observed order book quantities as volume or trade types; see [9–11, 18]. Splitting the noise into a parametric function of such quantities and residual noise, a plug-in estimation of integrated volatility can also yield faster convergence rates than in the classical model with uninformative noise. While this effect of improved volatility estimation appears to be a similarity to our work, our viewpoint on market microstructure is quite distinct. We focus on a model with one-sided instead of centred noise, but we neither impose a parametric assumption on the noise, nor do we include additional trading information. Such refinements of a one-sided noise model, as discussed in the mentioned works for the centred noise model, might be of interest for future research when microstructure effects of bid and ask quotes are better understood. This could potentially further improve volatility estimation.

Inference on the spot volatility is one of the most important topics in the financial literature; see, for instance, [4, 22] and the references therein. In this work, we address spot volatility estimation for the model from [5]. Using local minima over blocks of shrinking lengths $h_n \propto n^{-2/3} \propto (nh_n)^{-2}$, the resulting distribution of local minima in [5] became involved and infeasible, such that a central limit theorem for the integrated volatility estimator could not be obtained. Our spot volatility estimator is related to a localized version of the estimator from [5], combined with truncation methods to eliminate jumps of the semimartingale. For the asymptotic theory, however, we follow a different approach choosing blocks of lengths h_n, where $h_n n^{2/3} \to \infty$ slowly. This allows us to establish stable central limit theorems with the best achievable rate, arbitrarily close to $n^{-1/6}$, in the important special case of a semimartingale volatility. We exploit this to construct pointwise asymptotic confidence intervals.

Although the asymptotic theory relies on block lengths that are slightly unbalanced by smoothing out the impact of the noise distribution on the distribution of local minima asymptotically, our numerical study demonstrates that the confidence intervals work well in realistic scenarios with block lengths which optimize the estimation performance. Robustness to different noise specifications is an advantage that is naturally implied by our approach. Our estimator is surprisingly simple: it is a local average of squared differences of block-wise minima times

a constant factor which comes from moments of the half-normal distribution of the minimum of a Brownian motion over the unit time interval. This estimator is consistent. However, the stable central limit theorem at a fast convergence rate requires a subtle bias correction which incorporates a more precise approximation of the asymptotic distribution of local minima. For that purpose, our analysis is based on a generalization of the arcsine law which gives the distribution of the proportion of time over some interval that a Brownian motion is positive. In order to compute the bias-correction function numerically, we introduce an efficient algorithm. Reducing local minima over many random variables to iterated minima of two random variables in each step combined with a convolution step can be interpreted as a kind of dynamic programming approach. It turns out to be much more efficient compared to the natural approximation by a Monte Carlo simulation and is a crucial ingredient of our numerical application. Our convergence rate is much faster than the optimal rate, $n^{1/8}$, for spot volatility estimation under regular noise [4, 14]. The main contribution of this work is to develop the probabilistic foundation for the asymptotic analysis of the estimator and to establish the stable central limit theorems, asymptotic confidence, and a numerically practicable method.

The methods and proof techniques to deal with jumps are inspired by the truncation methods pioneered in [21] and summarized in [15, Chapter 13]. Overall, the strategy and restrictions on jump processes are to some extent similar, while several details under irregular noise using order statistics are rather different compared to settings without noise or with regular centred noise as in [7].

We introduce and further discuss our model in Section 2. Section 3 presents estimation methods and Section 4 asymptotic results. The numerical application is considered in Section 5 and a Monte Carlo simulation study illustrates the appealing finite-sample performance of the method. All proofs are given in Section 6.

2. Model with lower-bounded, one-sided noise and assumptions

Consider an Itô semimartingale

$$X_t = X_0 + \int_0^t a_s \, \mathrm{d}s + \int_0^t \sigma_s \, \mathrm{d}W_s + \int_0^t \int_{\mathbb{R}} \delta(s, z) \mathbf{1}_{\{|\delta(s,z)| \leq 1\}} (\mu - \nu)(\mathrm{d}s, \mathrm{d}z)$$

$$+ \int_0^t \int_{\mathbb{R}} \delta(s, z) \mathbf{1}_{\{|\delta(s,z)| > 1\}} \mu(\mathrm{d}s, \mathrm{d}z), \qquad t \geq 0, \tag{1}$$

with a one-dimensional standard Brownian motion (W_t), defined on some filtered probability space $(\Omega^X, \mathcal{F}^X, (\mathcal{F}_t^X), \mathbb{P}^X)$. For the drift process (a_t) and the volatility process (σ_t) we impose the following quite general assumptions.

Assumption 1. *The processes $(a_t)_{t \geq 0}$ and $(\sigma_t)_{t \geq 0}$ are locally bounded. The volatility process is strictly positive, $\inf_{t \in [0,1]} \sigma_t > 0$, $\mathbb{P}^X$-almost surely. For all $0 \leq t + s \leq 1$, $t \geq 0$, $s \geq 0$, with some constants $C_\sigma > 0$, and $\alpha > 0$,*

$$\mathbb{E}[(\sigma_{t+s} - \sigma_t)^2] \leq C_\sigma s^{2\alpha}. \tag{2}$$

Condition (2) introduces a regularity parameter α, governing the smoothness of the volatility process. The parameter α is crucial, since it will naturally influence the convergence rates of spot volatility estimation. Inequality (2) is less restrictive than α-Hölder continuity, since it does not rule out volatility jumps. For instance, any compound Poisson jump process with a jump size distribution having finite second moments satisfies (2) with $\alpha = \frac{1}{2}$. Since second

moments in (2) of such a process are bounded by a constant times $(s^2 + s)$, i.e. the second moment of a Poisson distribution with parameter s, this readily follows. Similar bounds for more general jump processes are given, for instance, in [15, Section 13]. This is important as empirical evidence for volatility jumps, in particular simultaneous price and volatility jumps, has been reported for intraday high-frequency financial data [6, 28]. The presented theory is, moreover, for general stochastic volatilities, also allowing for rough volatility. Rough fractional stochastic volatility models recently became popular and are used, for instance, in the macroscopic model of [12, 25].

The jump component of (1) is illustrated as in [15] and related literature, where the predictable function δ is defined on $\Omega \times \mathbb{R}_+ \times \mathbb{R}$, and the Poisson random measure μ is compensated by $\nu(\mathrm{d}s, \mathrm{d}z) = \lambda(\mathrm{d}z) \otimes \mathrm{d}s$, with a σ-finite measure λ. We impose the following standard condition with a generalized Blumenthal–Getoor or jump activity index r, $0 \le r \le 2$.

Assumption 2. *Assume that* $\sup_{\omega, x} |\delta(t, x)| / \gamma(x)$ *is locally bounded with a non-negative, deterministic function* γ *which satisfies* $\int_{\mathbb{R}} (\gamma^r(x) \wedge 1) \lambda(\mathrm{d}x) < \infty$.

We use the notation $a \wedge b = \min(a, b)$, and $a \vee b = \max(a, b)$, throughout this paper. Assumption 2 is most restrictive in the case $r = 0$, when jumps are of finite activity. The larger r is, the more general jump components are allowed. We will develop results under mild restrictions on r.

The process (X_t), which can be decomposed into

$$X_t = C_t + J_t, \tag{3}$$

with a continuous component (C_t) and a càdlàg jump component (J_t), provides a model for the latent efficient log-price process in continuous time.

High-frequency (best) ask order prices from a limit order book at times t_i^n, $0 \le i \le n$, on the fix time interval $[0, 1]$ cannot be adequately modelled by discrete recordings of (X_t). Instead, we propose the additive model with lower-bounded, one-sided microstructure noise:

$$Y_i = X_{t_i^n} + \varepsilon_i, \qquad i = 0, \ldots, n, \ \varepsilon_i \overset{\mathrm{iid}}{\sim} F_\eta, \ \varepsilon_i \ge 0. \tag{4}$$

The crucial property of the model is that the support of the noise is lower bounded. It is not that important that this boundary is zero—it could be a different constant, or even a regularly varying function over time. The methods and results presented are robust with respect to such model generalizations. We set the bound equal to zero, which appears to be the most natural choice for limit orders.

Assumption 3. *The independent and identically distributed (i.i.d.) noise* $(\varepsilon_i)_{0 \le i \le n}$ *has a cumulative distribution function (CDF)* F_η *satisfying*

$$F_\eta(x) = \eta x(1 + \mathcal{O}(1)) \quad as \ x \downarrow 0. \tag{5}$$

This is a nonparametric model in that the extreme value index is -1 for the *minimum* domain of attraction close to the boundary. This standard assumption on one-sided noise has already been used in [17, 24] within different frameworks. We do not require assumptions about the maximum domain of attraction, moments, and the tails of the noise distribution. Parametric examples which satisfy (5) are, for instance, the uniform distribution on some interval, the exponential distribution, and the standard Pareto distribution with heavy tails.

The i.i.d. assumption on the noise is crucial, and generalizations to weakly dependent noise will require considerable work and new proof concepts. Heterogeneity, that is, a

time-dependent noise level $\eta(t)$, could be included in our asymptotic analysis under mild assumptions.

3. Construction of spot volatility estimators

We partition the observation interval $[0, 1]$ into h_n^{-1} equispaced blocks, $h_n^{-1} \in \mathbb{N}$, and take local minima on each block. We hence obtain, for $k = 0, \ldots, h_n^{-1} - 1$, the local, block-wise minima

$$m_{k,n} = \min_{i \in \mathcal{I}_k^n} Y_i, \qquad \mathcal{I}_k^n = \{i \in \{0, \ldots, n\} \colon t_i^n \in [kh_n, (k+1)h_n)\}.$$

While h_n^{-1} is an integer, nh_n is in general not integer valued. For a simple interpretation, however, we can think of nh_n as an integer-valued sequence which gives the number of noisy observations per block in the case of equidistant observations. A spot volatility estimator could be obtained as a localized version of the estimator from [5, (2.9)] for the integrated volatility in the analogous model. The idea is that differences $m_{k,n} - m_{k-1,n}$ of local minima estimate differences of efficient prices, and a sum of squared differences can be used to estimate the volatility. However, things are not that simple. To determine the expectation of squared differences of local minima we introduce the function

$$\Psi_n(\sigma^2) = \frac{\pi}{2(\pi - 2)} h_n^{-1} \mathbb{E}\left[\left(\min_{i \in \{0, \ldots, nh_n - 1\}} (\sigma B_{i/n} + \varepsilon_i) - \min_{i \in \{1, \ldots, nh_n\}} (\sigma \tilde{B}_{i/n} + \varepsilon_i)\right)^2\right], \qquad (6)$$

where (B_t) and $(\tilde{B}_t)$ denote two independent standard Brownian motions. In [5], the block length balanced the order of block-wise minimal errors $(nh_n)^{-1}$ under (5) and the order $h_n^{1/2}$ of the movement of the stochastic semimartingale boundary over a block. For $h_n n^{2/3} \to \infty$, Ψ_n tends to the identity function, so we have that

$$\Psi_n(\sigma^2) = \sigma^2 + \mathcal{O}(1) \quad \text{as } n \to \infty. \qquad (7)$$

In this asymptotic regime local minima are mainly determined by local minima of the boundary process, such that the first-order approximation equals (6) when neglecting the noise (ε_i) on the right-hand side. The half-normal distribution of the minimum of a Brownian motion over an interval and its moments then readily yield (7). A formal proof of (7) is contained in Step 3 of the proof of Theorem 1 in Section 6.2. Note that we defined Ψ_n differently than in [5], e.g. in their (A.35), with the additional factor $\pi/(\pi - 2)$. By the simple asymptotic approximation in (7), we do not require Ψ_n^{-1} for a consistent estimator.

When there are no price jumps, a simple consistent estimator for the spot squared volatility σ_τ^2 is given by

$$\hat{\sigma}_{\tau-}^2 = \frac{\pi}{2(\pi - 2)K_n} \sum_{k=(\lfloor h_n^{-1}\tau \rfloor - K_n) \vee 1}^{\lfloor h_n^{-1}\tau \rfloor - 1} h_n^{-1}(m_{k,n} - m_{k-1,n})^2 \qquad (8)$$

for suitable sequences $h_n \to 0$ and $K_n \to \infty$. Using only observations before time τ, the estimator is available online at time $\tau \in (0, 1]$ during a trading day. For τ close to 0, when $\lfloor h_n^{-1}\tau \rfloor \leq K_n$, the factor K_n^{-1} can be adjusted to get an average. Since this is unimportant for asymptotic theory, we keep K_n^{-1} for simple notation. Working with ex post data over the whole interval, instead of using only observations before time τ, we may also use

$$\hat{\sigma}_{\tau+}^2 = \frac{\pi}{2(\pi - 2)K_n} \sum_{k=\lfloor h_n^{-1}\tau \rfloor + 1}^{(\lfloor h_n^{-1}\tau \rfloor + K_n) \wedge (h_n^{-1} - 1)} h_n^{-1}(m_{k,n} - m_{k-1,n})^2, \qquad (9)$$

or an estimator with an average centred around time $\tau \in (0, 1)$. The difference between the two estimators (9) and (8) can be used to infer a possible jump in the volatility process at time $\tau \in (0, 1)$, as in [6].

To construct confidence intervals for the spot volatility, it is useful to also establish a spot quarticity estimator:

$$\widehat{\sigma_{\tau-}^4} = \frac{\pi}{4(3\pi - 8)K_n} \sum_{k=(\lfloor h_n^{-1}\tau \rfloor - K_n) \vee 1}^{\lfloor h_n^{-1}\tau \rfloor - 1} h_n^{-2} (m_{k,n} - m_{k-1,n})^4. \tag{10}$$

A spot volatility estimator that is robust with respect to jumps in (X_t) is obtained with threshold versions of these estimators. We truncate differences of local minima whose absolute values exceed a threshold $u_n = \beta \cdot h_n^\kappa$, $\kappa \in (0, \frac{1}{2})$, with some positive constant β, which leads to

$$\hat{\sigma}_{\tau-}^{2,(\mathrm{tr})} = \frac{\pi}{2(\pi - 2)K_n} \sum_{k=(\lfloor h_n^{-1}\tau \rfloor - K_n) \vee 1}^{\lfloor h_n^{-1}\tau \rfloor - 1} h_n^{-1} (m_{k,n} - m_{k-1,n})^2 \mathbf{1}_{\{|m_{k,n} - m_{k-1,n}| \le u_n\}}, \tag{11}$$

and analogous versions of the estimators (9) and (10).

4. Asymptotic results

We establish asymptotic results for equidistant observations, $t_i^n = i/n$. We begin with the asymptotic theory in a setup without jumps in (X_t).

Theorem 1. (Stable central limit theorem for continuous (X_t).) *Set h_n such that $h_n n^{2/3} \to \infty$ and $K_n = C_K h_n^{\delta - 2\alpha/(1+2\alpha)}$ for arbitrary δ, $0 < \delta < 2\alpha/(1 + 2\alpha)$, and some constant $C_K > 0$. If (X_t) is continuous, i.e. $J_t = 0$ in (3), under Assumptions 1 and 3, the spot volatility estimator (8) is consistent, $\hat{\sigma}_{\tau-}^2 \xrightarrow{\mathbb{P}} \sigma_{\tau-}^2$, and satisfies the stable central limit theorem*

$$K_n^{1/2} (\hat{\sigma}_{\tau-}^2 - \Psi_n(\sigma_{\tau-}^2)) \xrightarrow{\mathrm{st}} \mathcal{N}\left(0, \frac{7\pi^2/4 - 2\pi/3 - 12}{(\pi - 2)^2} \sigma_{\tau-}^4\right). \tag{12}$$

There is only a difference between σ_τ^2 and its left limit $\sigma_{\tau-}^2$ in the case of a volatility jump at time τ. In particular, the estimator is also consistent for σ_τ^2 for any fix $\tau \in (0, 1)$. The convergence rate $K_n^{-1/2}$ gets arbitrarily close to $n^{-2\alpha/(3+6\alpha)}$, which is optimal in our model. The optimal rate is attained, according to [5], for $h_n \propto n^{-2/3}$ and $K_n \propto h_n^{-2\alpha/(1+2\alpha)}$, i.e. $\delta \downarrow 0$. In the important special case when $\alpha = \frac{1}{2}$, for a semimartingale volatility, the rate is arbitrarily close to $n^{-1/6}$. This is much faster than the optimal rate of convergence in the model with additive centred microstructure noise, which is known to be $n^{-1/8}$ [4, 14]. The constant in the asymptotic variance is obtained from several variance and covariance terms including (squared) local minima, and is approximately 2.44. The function Ψ_n was shown to be monotone and invertible in [5], and Ψ_n and its inverse Ψ_n^{-1} can be approximated using Monte Carlo simulations, see Section 5.1. The asymptotic distribution of the estimator does not hinge on the noise level η, which is different to methods for centred noise. Hence, we do not require any pre-estimation of noise parameters and the theory directly extends to a time-varying noise level $\eta(t)$ in (5) under the mild assumption that $0 < \eta(t) < \infty$ for all t. The stable convergence in (12)

is stronger than weak convergence and is important, since the limit distribution is *mixed normal* depending on the stochastic volatility. We refer to [15, Section 2.2.1] for an introduction to stable convergence. For a normalized central limit theorem, we can use the spot quarticity estimator (10).

Proposition 1. (Feasible central limit theorem.) *Under the conditions of Theorem 1, the spot quarticity estimator (10) is consistent, such that we get for the spot volatility estimation the normalized central limit theorem*

$$K_n^{1/2} \frac{\pi - 2}{\sqrt{\widehat{\sigma_\tau^4}_-(7\pi^2/4 - 2\pi/3 - 12)}} \left(\hat\sigma_{\tau-}^2 - \Psi_n(\sigma_{\tau-}^2)\right) \xrightarrow{\mathrm{d}} \mathcal{N}(0, 1).$$

Proposition 1 yields asymptotic confidence intervals for spot volatility estimation. For $q \in (0, 1)$, we have

$$\mathbb{P}\left(\sigma_{\tau-}^2 \in \left[\Psi_n^{-1}\left(\hat\sigma_{\tau-}^2 - \frac{\sqrt{\widehat{\sigma_\tau^4}_-(7\pi^2/4 - 2\pi/3 - 12)}}{\pi - 2} K_n^{-1/2} \Phi^{-1}(1 - q/2)\right),\right.\right.$$

$$\left.\left. \Psi_n^{-1}\left(\hat\sigma_{\tau-}^2 + \frac{\sqrt{\widehat{\sigma_\tau^4}_-(7\pi^2/4 - 2\pi/3 - 12)}}{\pi - 2} K_n^{-1/2} \Phi^{-1}(1 - q/2)\right)\right]\right) \to 1 - q$$

by the monotonicity of Ψ_n^{-1}, with Φ the CDF of the standard normal distribution. Since Ψ_n^{-1} is differentiable by [5, (A.35)] and the derivative is $(\Psi_n^{-1})' = 1 + \mathcal{O}(1)$ by (7), the delta method (for stable convergence) also yields asymptotic confidence intervals and the central limit theorem

$$K_n^{1/2}\left(\Psi_n^{-1}\left(\hat\sigma_{\tau-}^2\right) - \sigma_{\tau-}^2\right) \xrightarrow{\mathrm{st}} \mathcal{N}\left(0, \frac{7\pi^2/4 - 2\pi/3 - 12}{(\pi - 2)^2} \sigma_{\tau-}^4\right). \tag{13}$$

We cannot simply replace $\Psi_n(\sigma_{\tau-}^2)$ in (12) by its first-order approximation $\sigma_{\tau-}^2$, or $\Psi_n^{-1}(\hat\sigma_{\tau-}^2)$ in (13) by $\hat\sigma_{\tau-}^2$, since the biases do not converge to zero sufficiently fast. That is, $(\hat\sigma_{\tau-}^2 - \sigma_{\tau-}^2) = \mathcal{O}_\mathbb{P}(K_n^{-1/2})$ does not hold true in general. Furthermore, if the condition $h_n n^{2/3} \to \infty$ is violated, the central limit theorems do not apply.

Theorem 2. (Stable central limit theorem with jumps in (X_t).) *Set h_n such that $h_n n^{2/3} \to \infty$ and $K_n = C_K h_n^{\delta - 2\alpha/(1+2\alpha)}$ for arbitrary δ, $0 < \delta < 2\alpha/(1 + 2\alpha)$, and some constant $C_K > 0$. Under Assumptions 1, 2, and 3, with*

$$r < \frac{2 + 2\alpha}{1 + 2\alpha},$$

the truncated spot volatility estimator (11) with

$$\kappa \in \left(\frac{1}{2 - r} \frac{\alpha}{2\alpha + 1}, \frac{1}{2}\right)$$

is consistent, $\hat\sigma_{\tau-}^{2,(\mathrm{tr})} \xrightarrow{\mathbb{P}} \sigma_{\tau-}^2$, and satisfies the stable central limit theorem

$$K_n^{1/2}\left(\hat\sigma_{\tau-}^{2,(\mathrm{tr})} - \Psi_n(\sigma_{\tau-}^2)\right) \xrightarrow{\mathrm{st}} \mathcal{N}\left(0, \frac{7\pi^2/4 - 2\pi/3 - 12}{(\pi - 2)^2} \sigma_{\tau-}^4\right).$$

In order to obtain a central limit theorem at (almost) optimal rate, we thus have to impose mild restrictions on the jump activity. For the standard model with a semimartingale volatility, i.e. $\alpha = \frac{1}{2}$, we need $r < \frac{3}{2}$, and for $\alpha = 1$ we have the stronger condition that $r < \frac{4}{3}$. These conditions are equivalent to those of [7, Theorem 1], which gives a central limit theorem for spot volatility estimation under similar assumptions on (X_t), but with a slower rate of convergence for centred microstructure noise. Using a truncated quarticity estimator with the same thresholding again yields a feasible central limit theorem and asymptotic confidence intervals.

Remark 1. From a theoretical point of view we might ponder why we do not work out an asymptotic theory for $h_n \propto n^{-2/3}$ when noise and efficient price both influence the asymptotic distribution of the local minima. However, in this balanced case, the asymptotic distribution is infeasible. For this reason, [5] could not establish a central limit theorem for their integrated volatility estimator. Moreover, their estimator was only implicitly defined depending on the unknown function Ψ_n^{-1}. Even imposing a parametric assumption on the noise as an exponential distribution would not render a feasible limit theory for $h_n \propto n^{-2/3}$—see the discussion in [5]. Choosing h_n such that $h_n n^{2/3} \to \infty$ slowly instead yields a simple, explicit, and consistent estimator and a feasible central limit theorem for spot volatility estimation. In particular, we use Ψ_n only for the bias correction of the simple estimator, while the estimator itself and the (estimated) asymptotic variance do not hinge on Ψ_n. Central limit theorems for spot volatility estimators are in general only available at almost optimal rates, when the variance dominates the squared bias in the mean squared error; see, for instance, Theorem 13.3.3 and the remarks below it in [15]. Therefore, (12) is the best achievable central limit theorem. Moreover, our choice of h_n avoids strong assumptions on the noise that would be inevitable for smaller blocks. Our numerical work will demonstrate that the asymptotic results presented are useful in practice and facilitate efficient inference on the spot volatility. In particular, Section 5.2 revolves around the question of how to choose block lengths in practice.

5. Implementation and simulations

5.1. Monte Carlo approximation of Ψ_n

Although the function Ψ_n from (6) tends to the identity asymptotically, it has a crucial role as a bias correction of our estimator in (12). We can compute the function numerically based on a Monte Carlo simulation. Hence, we have to compute $\Psi_n(\sigma^2)$ as a Monte Carlo mean over many iterations and over a fine grid of values for the squared volatility. Then, we can also numerically invert the function and use $\Psi_n^{-1}(\cdot)$. To make this procedure feasible without too high a computational expense we require an algorithm to efficiently sample from the law of the local minima for some given n and block length h_n.

Consider, for $nh_n \in \mathbb{N}$ with $Z_i \overset{\text{iid}}{\sim} \mathcal{N}(0, 1)$ and the observation errors $(\varepsilon_k)_{k \geq 0}$, the minimum

$$M_1^{nh_n} := \min_{k=1,\ldots,nh_n} \left(\frac{\sigma}{\sqrt{n}} \sum_{i=1}^{k} Z_i + \varepsilon_k \right)$$

for some fixed $\sigma > 0$, and, for $l \in \{0, \ldots, nh_n\}$,

$$M_l^{nh_n} := \min_{k=l,\ldots,nh_n} \left(\frac{\sigma}{\sqrt{n}} \sum_{i=0}^{k} Z_i + \varepsilon_k \right),$$

where we set $Z_0 := 0$. Since

$$\Psi_n(\sigma^2) = \frac{1}{2} \frac{\pi}{\pi - 2} h_n^{-1} \mathbb{E}\big[\big(M_0^{nh_n-1} - M_1^{nh_n}\big)^2\big],$$

with $M_0^{nh_n-1}$ generated independently from $M_1^{nh_n}$, we want to simulate samples distributed as $M_0^{nh_n-1}$ and $M_1^{nh_n}$, respectively. Note that for finite nh_n there is no exact equality between the moments of $M_0^{nh_n-1}$ and $M_1^{nh_n}$, which can be relevant in particular for moderate values of nh_n. As in the simulation of Section 5.2, we implement exponentially distributed observation errors (ε_k), with some given noise level η. In data applications, we can do the same with an estimated noise level

$$\hat{\eta} = \left(\frac{1}{2n} \sum_{i=1}^{n} (Y_i - Y_{i-1})^2\right)^{-1/2} = \eta + \mathcal{O}_{\mathbb{P}}(n^{-1/2}).$$

This estimator works for all noise distributions with finite fourth moments. In view of the discussion of the model in [5], exponentially distributed noise is the most natural example satisfying (5). Simulations with other noise distributions lead to similar results. This is expected, since the estimator only hinges on local minima and their distribution is asymptotically more determined by the Brownian motion than by the noise distribution. To simulate the local minima for given n, h_n, η, and squared volatility σ^2 in an efficient way we use a specific dynamic programming principle. Observe that

$$
\begin{aligned}
M_1^{nh_n} &= \frac{\sigma}{\sqrt{n}} Z_1 + \min\big(\varepsilon_1, M_2^{nh_n}\big) \\
&= \frac{\sigma}{\sqrt{n}} Z_1 + \min\left(\varepsilon_1, \frac{\sigma}{\sqrt{n}} Z_2 + \min\big(\varepsilon_2, M_3^{nh_n}\big)\right) \\
&= \frac{\sigma}{\sqrt{n}} Z_1 + \min\left(\cdots \min\left(\varepsilon_{nh_n-2}, \frac{\sigma}{\sqrt{n}} Z_{nh_n-1} + \min\left(\varepsilon_{nh_n-1}, \frac{\sigma}{\sqrt{n}} Z_{nh_n} + \varepsilon_{nh_n}\right)\right)\cdots\right).
\end{aligned}
$$

In the baseline noise model $\varepsilon_k \overset{iid}{\sim} \text{Exp}(\eta)$, the random variable $(\sigma/\sqrt{n})Z_{nh_n} + \varepsilon_{nh_n}$ has an exponentially modified Gaussian (EMG) distribution. With any fixed noise distribution, we can easily generate realizations from this convolution. A pseudorandom variable distributed as $M_1^{nh_n}$ is now generated following the last transformation in the reverse direction. Algorithmically, this reads

1. Generate $U_{nh_n} \sim \text{EMG}(\sigma^2/n, \eta) \sim \text{Exp}(\eta) + (\sigma/\sqrt{n})\text{Norm}(1)$

2. $U_{nh_n-1} = \min(U_{nh_n}, \text{Exp}(\eta)) + (\sigma/\sqrt{n})\text{Norm}(1)$

3. Iterate until U_1

where the end point U_1 has the target distribution of $M_1^{nh_n}$. In each iteration step, we thus take the minimum of the current state of the process with one independent exponentially distributed random variable and the convolution with one independent normally distributed random variable. To sample from the distribution of $M_0^{nh_n-1}$ instead, we use the same algorithm and just drop the convolution with the normal distribution in the last step.

This algorithm facilitates a many times faster sampling from the distribution of local minima and numerical approximation of Ψ_n compared to running for each value a standard Monte Carlo simulation in that local minima are computed over blocks of length h_n.
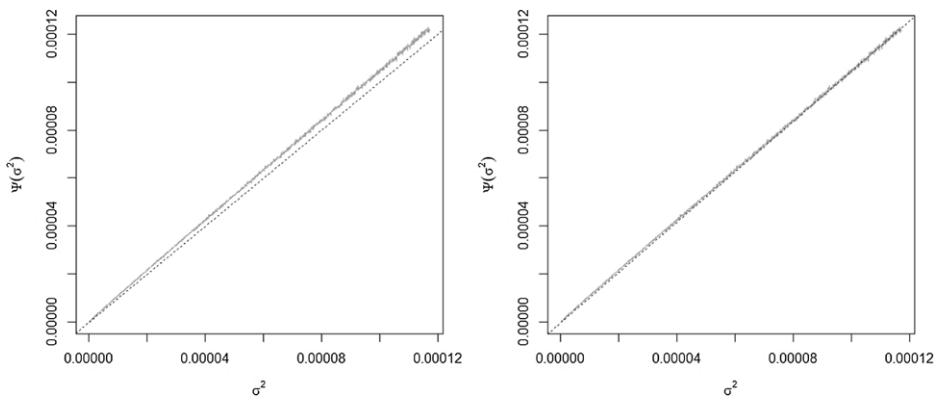
FIGURE 1. Monte Carlo means to estimate $\Psi_n(\sigma^2)$ over a fine grid (interpolated line) for $n = 23\,400$ and $n \cdot h_n = 15$. Left: the dotted line shows the identity function. Right: the dotted line is a linear function with slope 1.046.

Figure 1 plots the result of the Monte Carlo approximation of $\Psi_n(\sigma^2)$ for $n = 23\,400$ and $n \cdot h_n = 15$ on a grid of 1500 values of σ^2. In this case, h_n is quite small, but this configuration turns out to be useful in Section 5.2. We know that $\Psi_n(\sigma^2)$ is monotone, such that the oscillation of the function in Figure 1 is due to the inaccuracy of the Monte Carlo means, although we use $N = 100\,000$ iterations for each grid point. Nevertheless, we can see that the function is rather close to a linear function with slope 1.046 based on a least squares estimate. The left panel of Figure 1 draws a comparison to the identity function which is illustrated by the dotted line, while the right panel draws a comparison to the linear function with slope 1.046. We see that it is crucial to correct for the bias in (12) when using such small values of h_n. Although the function $\Psi_n(\sigma^2)$ is not exactly linear, a simple bias correction dividing estimates by 1.046 is almost as good as using the more precise numerical inversion based on the Monte Carlo approximation. Since the Monte Carlo approximations of $\Psi_n(\sigma^2)$ look close to linear functions in all the cases considered, we report the estimated slopes based on least squares and $N = 100\,000$ Monte Carlo iterations for different values of h_n in Table 1 to summarize concisely the distance between the function $\Psi_n(\sigma^2)$ and the identity. Simulating all iterations for all grid points with our algorithm takes only a few hours with a standard computer.

5.2. Simulation study of estimators

We simulate $n = 23\,400$ observations corresponding to one observation per second over a (NASDAQ) trading day of 6.5 hours. The efficient price process is simulated from the model

$$\mathrm{d}X_t = v_t \sigma_t \, \mathrm{d}W_t,$$

$$\mathrm{d}\sigma_t^2 = 0.0162 \cdot (0.8465 - \sigma_t^2) \, \mathrm{d}t + 0.117 \cdot \sigma_t \, \mathrm{d}B_t,$$

$$v_t = (6 - \sin(3\pi t/4)) \cdot 0.002, \qquad t \in [0, 1].$$

The factor (v_t) generates a typical U-shaped intraday volatility pattern. (W_t, B_t) is a two-dimensional Brownian motion with leverage $\mathrm{d}[W, B]_t = -0.2 \, \mathrm{d}t$. The stochastic volatility component has several realistic features and the simulated model is in line with recent

TABLE 1. Regression slopes to measure the bias of estimator (8) and deviation $\Psi_n(\sigma^2) - \sigma^2$.

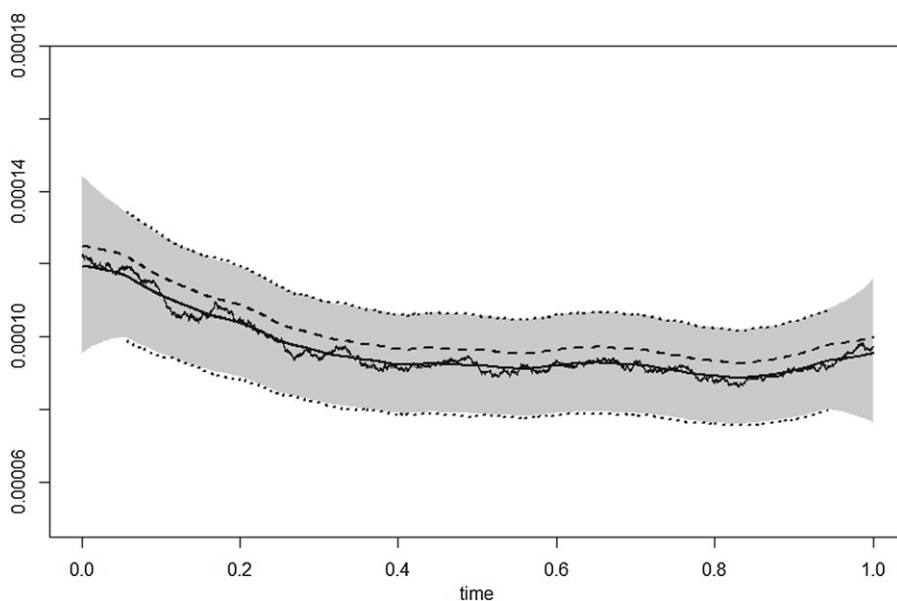| $n \cdot h_n$	10	15	25	39	78	234
h_n^{-1}	2340	1560	936	600	300	100
$h_n \cdot n^{2/3}$	0.350	0.524	0.874	1.36	2.73	8.18
Slope	1.077	1.046	1.025	1.016	1.008	1.003
Approx. bias	7.7%	4.6%	2.5%	1.6%	0.8%	0.3%

FIGURE 2. True and estimated spot volatility with pointwise confidence sets.

literature; see [6] and references therein. We do not include a drift in X_t to avoid introducing another process or more parameters. Any drift evolving within a reasonable range of values will not affect the numerical results presented. Observations with lower-bounded, one-sided microstructure noise are generated by $Y_i = X_{i/n} + \varepsilon_i$, $0 \le i \le n$, with exponentially distributed noise $\varepsilon_i \overset{\text{iid}}{\sim} \text{Exp}(\eta)$, with $\eta = 10\,000$. The noise variance is then rather small, but this is in line with stylized facts of real NASDAQ data such as, for instance, those analyzed in [4, 6]. Note that the noise level estimate is analogous to the one used for regular market microstructure noise. Typical noise levels obtained e.g. for Apple are approximately 15 000, and approximately 4000 for 3M; see the supplement of [4].

Figure 2 shows a fixed path of the squared volatility. We fix this path for the following Monte Carlo simulation and generate new observations of (X_t) and (Y_i) in each iteration according to our model. The dashed line in Figure 2 gives the estimated volatility by the Monte Carlo means over $N = 50\,000$ iterations based on $n \cdot h_n = 15$ observations per block using the non-adjusted estimator (8), but with windows which are centred around the block on which we estimate the spot volatility, i.e. windows centred around the time τ, and with $K_n = 180$. We plot estimates

TABLE 2. Summary statistics of estimation for different values of h_n and K_n. MSD = mean standard deviation, MAB = mean absolute bias, MABC = MAB of bias-corrected estimator.

nh_n	$K_n = 120$			$K_n = 180$			$K_n = 240$		
	MSD	MAB	MABC	MSD	MAB	MABC	MSD	MAB	MABC
10	14.6	7.59	0.73	12.0	7.51	0.90	10.5	7.60	1.13
15	14.4	4.59	0.88	11.8	4.57	1.17	10.3	4.46	1.43
25	14.3	2.56	1.24	11.8	2.63	1.66	10.3	2.86	1.91
78	14.7	2.44	2.52	12.3	3.53	3.42	11.0	4.33	4.16

All values multiplied by a factor of 10^6.

on each block, where the estimates close to the boundaries rely on fewer observations. The solid line gives the bias-corrected volatility estimates using the numerically evaluated function Ψ_n, based on the algorithm from Section 5.1 with $n \cdot h_n = 15$ and $n = 23\,400$. We determined the values $n \cdot h_n = 15$ and $K_n = 180$ as suitable values to obtain a small mean squared error. In fact, the choice of $K_n = 180$ is rather large in favour of a smaller variance that yields a rather smooth estimated spot volatility in Figure 2. The estimated volatility hence appears smoother compared to the true semimartingale volatility, but the intraday pattern is captured well by our estimation. We expect that this is typically an appealing implementation in practice as smaller K_n results in a larger variance. Choosing $K_n = 180$ rather large, we have to use quite small block sizes h_n to control the overall bias of the estimation. Since $h_n \cdot n^{2/3} \approx 0.52$ is small, the bias correction becomes crucial here. Still, our asymptotic results work well for this implementation. This can be seen by the comparison of pointwise empirical 10% and 90% quantiles from the Monte Carlo iterations illustrated by the grey area and the 10% and 90% quantiles of the limit normal distribution with the asymptotic variance from (12). The latter are drawn as dotted lines for the blocks with distance larger than $K_n/2$ from the boundaries, where the variances are of order K_n^{-1}. Close to the boundaries the empirical variances increase due to the smaller number of blocks used for the estimates. Moreover, the bias correction, which is almost identical to dividing each estimate by 1.046, correctly scales the simple estimates which have a significant positive bias for the chosen tuning parameters. Overall, our asymptotic results provide a good finite-sample fit even though we have $h_n \cdot n^{2/3} < 1$ here. Note, however, that $\sigma_t \cdot \eta \approx 100$, and our asymptotic expansion in fact requires that $h_n^{3/2} n \sigma_t \eta$ is large when taking constants into account. Since the simulated scenario uses realistic values, we recommend similar block lengths for applications to real high-frequency financial data. According to the summary statistics in the supplement of [4], some assets exhibit higher noise-to-signal ratios, and for those larger blocks are preferable.

Table 2 summarizes the performance of the estimation along different choices of nh_n and K_n using the following quantities:

MSD: the mean standard deviation of N iterations averaged over all grid points;

MAB: the mean absolute bias of N iterations averaged over all grid points and for the estimator (8) without any bias correction;

MABC: the mean absolute bias of N iterations averaged over all grid points and for the estimator (8) with a simple bias correction dividing estimates by the factors given in Table 1.

All the results are based on $N = 50\,000$ Monte Carlo iterations. First of all, the values used for Figure 2 are not unique minimizers of the mean squared error. Several other combinations given in Table 2 render equally good results. Overall, the performance is comparable within a broad range of block lengths and window sizes. The variances decrease for larger K_n, while the bias increases with larger K_n for fixed h_n. Important for the bias is the total window size, $K_n \cdot h_n$, over which the volatility is approximated by a constant for the estimation. The variance only depends on K_n: changing the block length for fixed K_n does not significantly affect the variance. While the MSD is hence almost constant within the columns of Table 2, the bias after correction, MABC, increases from the top down due to the increasing window size. Without the bias correction two effects interfere for MAB. Larger blocks reduce the systematic bias due to $\Psi_n(\sigma_t^2) - \sigma_t^2$, but the increasing bias due to the increasing window size prevails for $n \cdot h_n = 78$, and the two larger values of K_n.

6. Proofs

6.1. Law of the integrated negative part of a Brownian motion

A crucial lemma for our theory is on an upper bound for the CDF of the integrated negative part of a Brownian motion. We prove a lemma based on a generalization of Lévy's arcsine law by [27]. The result is in line with the conjecture in [16, (261)], where one finds an expansion of the density with a precise constant for the leading term. Denote by f_+ the positive part and by f_- the negative part of some real-valued function f.

Lemma 1. *For a standard Brownian motion* $(W_t)_{t \geq 0}$,

$$\mathbb{P}\left(\int_0^1 (W_t)_- \, \mathrm{d}t \leq x \right) = \mathcal{O}(x^{1/3}), \quad x \to 0.$$

Proof. Observe the equality in distribution $\int_0^1 (W_t)_- \, \mathrm{d}t \stackrel{\mathrm{d}}{=} \int_0^1 (W_t)_+ \, \mathrm{d}t$, such that

$$\mathbb{P}\left(\int_0^1 (W_t)_- \, \mathrm{d}t \leq x \right) = \mathbb{P}\left(\int_0^1 (W_t)_+ \, \mathrm{d}t \leq x \right), \quad x > 0.$$

For any $\varepsilon > 0$, the inequality

$$\int_0^1 (W_t)_+ \, \mathrm{d}t \geq \int_0^1 W_t \cdot \mathbf{1}(W_t > \varepsilon) \, \mathrm{d}t \geq \varepsilon \int_0^1 \mathbf{1}(W_t > \varepsilon) \, \mathrm{d}t$$

leads us to

$$\mathbb{P}\left(\int_0^1 (W_t)_+ \, \mathrm{d}t \leq x \right) \leq \mathbb{P}\left(\varepsilon \int_0^1 \mathbf{1}(W_t > \varepsilon) \, \mathrm{d}t \leq x \right) = \mathbb{P}\left(1 - \int_0^1 \mathbf{1}(W_t \leq \varepsilon) \, \mathrm{d}t \leq x/\varepsilon \right)$$

$$= \mathbb{P}\left(\int_0^1 \mathbf{1}(W_t \leq \varepsilon) \, \mathrm{d}t \geq 1 - x/\varepsilon \right).$$

Using [27, (15) and (16)], we obtain

$$\mathbb{P}\left(\int_0^1 \mathbf{1}(W_t \leq \varepsilon) \, \mathrm{d}t \geq 1 - x/\varepsilon \right) = \frac{1}{\pi} \int_{1-x/\varepsilon}^1 \frac{\exp(-\varepsilon^2/(2u))}{\sqrt{u(1-u)}} \, \mathrm{d}u + 2\Phi(\varepsilon) - 1,$$

with Φ the CDF of the standard normal distribution. Thereby, we obtain

$$\mathbb{P}\left(\int_0^1 (W_t)_+ \, dt \le x\right) \le \frac{1}{\pi} \int_{1-x/\varepsilon}^1 \frac{\exp(-\varepsilon^2/(2u))}{\sqrt{u(1-u)}} \, du + 2 \int_0^\varepsilon \frac{\exp(-u^2/2)}{\sqrt{2\pi}} \, du,$$

and elementary bounds give the upper bound

$$\mathbb{P}\left(\int_0^1 (W_t)_+ \, dt \le x\right) \le \frac{2}{\pi} \sqrt{\frac{x}{\varepsilon}} \frac{1}{\sqrt{1-x/\varepsilon}} + \frac{2\varepsilon}{\sqrt{2\pi}}.$$

Choosing $\varepsilon = x^{1/3}$, we obtain the upper bound

$$\mathbb{P}\left(\int_0^1 (W_t)_+ \, dt \le x\right) \le \frac{2}{\pi} x^{1/3} \frac{1}{\sqrt{1-x^{2/3}}} + \frac{2x^{1/3}}{\sqrt{2\pi}}. \qquad \square$$

6.2. Asymptotics of the spot volatility estimation in the continuous case

Proof of Theorem 1. In the following, we write $A_n \lesssim B_n$ for two real sequences if there exists some $n_0 \in \mathbb{N}$ and a constant K such that $A_n \le K B_n$ for all $n \ge n_0$.

Step 1 In the first step, we prove the approximation

$$\hat{\sigma}_{\tau-}^2 = \frac{\pi}{2(\pi-2)K_n} \sum_{k=(\lfloor h_n^{-1}\tau \rfloor - K_n)\vee 1}^{\lfloor h_n^{-1}\tau \rfloor - 1} h_n^{-1}(m_{k,n} - m_{k-1,n})^2$$

$$= \frac{\pi}{2(\pi-2)K_n} \sum_{k=(\lfloor h_n^{-1}\tau \rfloor - K_n)\vee 1}^{\lfloor h_n^{-1}\tau \rfloor - 1} h_n^{-1}(\tilde{m}_{k,n} - \tilde{m}_{k-1,n}^*)^2 + \mathcal{O}_{\mathbb{P}}(h_n^{\alpha \wedge 1/2})$$

with

$$\tilde{m}_{k,n} = \min_{i \in \mathcal{I}_k^n} (\varepsilon_i + \sigma_{(k-1)h_n}(W_{t_i^n} - W_{kh_n})), \qquad \tilde{m}_{k-1,n}^* = \min_{i \in \mathcal{I}_{k-1}^n} (\varepsilon_i - \sigma_{(k-1)h_n}(W_{kh_n} - W_{t_i^n})).$$

We show that, for $k \in \{1, \ldots, h_n^{-1} - 1\}$,

$$m_{k,n} - m_{k-1,n} = \tilde{m}_{k,n} - \tilde{m}_{k-1,n}^* + \mathcal{O}_{\mathbb{P}}(h_n^{1/2}). \qquad (14)$$

We subtract X_{kh_n} from $m_{k,n}$ and $m_{k-1,n}$, and use that, for all i,

$$(Y_i - X_{kh_n}) - (X_{t_i^n} - (X_{kh_n} + \sigma_{(k-1)h_n}(W_{t_i^n} - W_{kh_n}))) = (\sigma_{(k-1)h_n}(W_{t_i^n} - W_{kh_n}) + \varepsilon_i).$$

This implies that

$$\min_{i \in \mathcal{I}_k^n}(Y_i - X_{kh_n}) - \max_{i \in \mathcal{I}_k^n}(X_{t_i^n} - (X_{kh_n} + \sigma_{(k-1)h_n}(W_{t_i^n} - W_{kh_n}))) \le \min_{i \in \mathcal{I}_k^n}(\sigma_{(k-1)h_n}(W_{t_i^n} - W_{kh_n}) + \varepsilon_i).$$

Changing the roles of $(Y_i - X_{kh_n})$ and $(\sigma_{(k-1)h_n}(W_{t_i^n} - W_{kh_n}) + \varepsilon_i)$, we obtain by the analogous inequalities and the triangle inequality, with $M_t := X_{kh_n} + \int_{kh_n}^t \sigma_{(k-1)h_n} \, \mathrm{d}W_s$, that

$$|m_{k,n} - X_{kh_n} - \tilde{m}_{k,n}| \leq \max_{i \in \mathcal{I}_k^n} |X_{t_i^n} - M_{t_i^n}|$$

$$\leq \sup_{t \in [kh_n, (k+1)h_n]} |X_t - M_t|$$

$$\leq \sup_{t \in [kh_n, (k+1)h_n]} \left| C_t - C_{kh_n} - \int_{kh_n}^t \sigma_{(k-1)h_n} \, \mathrm{d}W_s \right|.$$

We write (C_t) for (X_t) to emphasize continuity, see (3). Then (14) follows from

$$\sup_{t \in [kh_n, (k+1)h_n]} \left| C_t - C_{kh_n} - \int_{kh_n}^t \sigma_{(k-1)h_n} \, \mathrm{d}W_s \right| = \mathcal{O}_{\mathbb{P}}(h_n^{1/2}), \tag{15}$$

and the analogous estimate for $m_{k-1,n}$ and $\tilde{m}_{k-1,n}^*$. We decompose

$$\sup_{t \in [kh_n, (k+1)h_n]} \left| C_t - C_{kh_n} - \int_{kh_n}^t \sigma_{(k-1)h_n} \, \mathrm{d}W_s \right| \leq \sup_{t \in [kh_n, (k+1)h_n]} \left| \int_{kh_n}^t (\sigma_s - \sigma_{(k-1)h_n}) \, \mathrm{d}W_s \right|$$

$$+ \sup_{t \in [kh_n, (k+1)h_n]} \int_{kh_n}^t |a_s| \, \mathrm{d}s.$$

Under Assumption 1, we can assume that (σ_t) and (a_t) are bounded on $[0, 1]$ by the localization from [15, Section 4.4.1]. Using Itô's isometry and Fubini's theorem, we obtain that

$$\mathbb{E}\left[\left(\int_{kh_n}^t (\sigma_s - \sigma_{(k-1)h_n}) \, \mathrm{d}W_s \right)^2 \right] = \mathbb{E}\left[\int_{kh_n}^t (\sigma_s - \sigma_{(k-1)h_n})^2 \, \mathrm{d}s \right] = \int_{kh_n}^t \mathbb{E}[(\sigma_s - \sigma_{(k-1)h_n})^2] \, \mathrm{d}s,$$

such that Assumption 1 yields, for any $t \in [kh_n, (k+1)h_n]$,

$$\mathbb{E}\left[\left(\int_{kh_n}^t (\sigma_s - \sigma_{(k-1)h_n}) \, \mathrm{d}W_s \right)^2 \right] \leq C_\sigma \int_{kh_n}^t (s - (k-1)h_n)^{2\alpha} \, \mathrm{d}s$$

$$\leq C_\sigma (2\alpha + 1)^{-1} (t - (k-1)h_n)^{2\alpha+1} = \mathcal{O}(h_n^{2\alpha+1}).$$

By Doob's martingale maximal inequality and since $\sup_{t \in [kh_n, (k+1)h_n]} \int_{kh_n}^t |a_s| \, \mathrm{d}s = \mathcal{O}_{\mathbb{P}}(h_n)$,

$$\sup_{t \in [kh_n, (k+1)h_n]} \left| C_t - C_{kh_n} - \int_{kh_n}^t \sigma_{(k-1)h_n} \, \mathrm{d}W_s \right| = \mathcal{O}_{\mathbb{P}}(h_n^{(1/2+\alpha)\wedge 1}).$$

We conclude that (15) holds, since $\alpha > 0$. Since

$$h_n^{-1}(m_{k,n} - m_{k-1,n})(m_{k,n} - \tilde{m}_{k,n}) = \mathcal{O}_{\mathbb{P}}(h_n^{\alpha \wedge 1/2}),$$

and analogously for $(m_{k-1,n} - \tilde{m}_{k-1,n}^*)$, we conclude Step 1 by writing

$$(m_{k,n} - m_{k-1,n})^2 - (\tilde{m}_{k,n} - \tilde{m}_{k-1,n}^*)^2 = (m_{k,n} - m_{k-1,n} + \tilde{m}_{k,n} - \tilde{m}_{k-1,n}^*)$$

$$\times (m_{k,n} - \tilde{m}_{k,n} + \tilde{m}_{k-1,n}^* - m_{k-1,n}).$$

Step 2 We bound the bias of the spot volatility estimation using Step 1. For $\lfloor h_n^{-1}\tau \rfloor > K_n$, we obtain from the definition of the function Ψ_n in (6) that

$$\mathbb{E}[\hat{\sigma}_{\tau-}^2 - \Psi_n(\sigma_{\tau-}^2)] = \frac{\pi}{2(\pi-2)K_n} \sum_{k=(\lfloor h_n^{-1}\tau \rfloor - K_n)\vee 1}^{\lfloor h_n^{-1}\tau \rfloor - 1} h_n^{-1}\mathbb{E}[(m_{k,n} - m_{k-1,n})^2] - \mathbb{E}[\Psi_n(\sigma_{\tau-}^2)]$$

$$= \frac{1}{K_n}\frac{\pi}{2(\pi-2)} \sum_{k=(\lfloor h_n^{-1}\tau \rfloor - K_n)\vee 1}^{\lfloor h_n^{-1}\tau \rfloor - 1} h_n^{-1}\mathbb{E}\big[\big(\tilde{m}_{k,n} - \tilde{m}_{k-1,n}^*\big)^2\big] - \mathbb{E}\big[\Psi_n(\sigma_{\tau-}^2)\big] + \mathcal{O}\big(h_n^{\alpha \wedge 1/2}\big)$$

$$= \frac{1}{K_n}\frac{\pi}{2(\pi-2)} \sum_{k=(\lfloor h_n^{-1}\tau \rfloor - K_n)\vee 1}^{\lfloor h_n^{-1}\tau \rfloor - 1} \frac{2(\pi-2)}{\pi}\mathbb{E}\big[\Psi_n\big(\sigma_{(k-1)h_n}^2\big)\big] - \mathbb{E}\big[\Psi_n(\sigma_{\tau-}^2)\big] + \mathcal{O}\big(h_n^{\alpha \wedge 1/2}\big)$$

$$\lesssim \frac{1}{K_n} \sum_{k=(\lfloor h_n^{-1}\tau \rfloor - K_n)\vee 1}^{\lfloor h_n^{-1}\tau \rfloor - 1} \mathbb{E}\big[\sigma_{(k-1)h_n}^2 - \sigma_{\tau-}^2\big] + \mathcal{O}\big(h_n^{\alpha \wedge 1/2}\big)$$

$$\lesssim \frac{1}{K_n} \sum_{k=(\lfloor h_n^{-1}\tau \rfloor - K_n)\vee 1}^{\lfloor h_n^{-1}\tau \rfloor - 1} \mathbb{E}\big[\sigma_{(k-1)h_n} - \sigma_{\tau-}\big] + \mathcal{O}\big(h_n^{\alpha \wedge 1/2}\big)$$

$$\lesssim \frac{1}{K_n} \sum_{k=(\lfloor h_n^{-1}\tau \rfloor - K_n)\vee 1}^{\lfloor h_n^{-1}\tau \rfloor - 1} \big(\mathbb{E}\big[\big(\sigma_{(k-1)h_n} - \sigma_{\tau-}\big)^2\big]\big)^{1/2} + \mathcal{O}\big(h_n^{\alpha \wedge 1/2}\big)$$

$$= \mathcal{O}((K_n h_n)^\alpha) = \mathcal{O}\big(h_n^{\alpha/(1+2\alpha)}\big) = \mathcal{O}\big(K_n^{-1/2}\big).$$

The first $\lesssim$ estimate is in fact an equality up to an additional factor $(1 + \mathcal{O}(1))$, since $\Psi_n'(x) = 1 + \mathcal{O}(1)$ for all $x \geq 0$, exploiting the abovementioned differentiability based on [5, (A.35)]. For the asymptotic upper bounds we used the binomial formula

$$\sigma_{(k-1)h_n}^2 - \sigma_{\tau-}^2 = (\sigma_{(k-1)h_n} - \sigma_{\tau-})(\sigma_{(k-1)h_n} + \sigma_{\tau-}) \leq 2C(\sigma_{(k-1)h_n} - \sigma_{\tau-}),$$

exploiting as in Step 1 that (σ_t) is bounded with some upper bound C, and Hölder's inequality to conclude with (2) from Assumption 1. Finally, we used that $(\alpha \wedge \frac{1}{2}) > \alpha/(2\alpha + 1)$ for all α.

Step 3 For the consistency of $\hat{\sigma}_{\tau-}^2$, we prove that

$$\mathbb{E}\big[\hat{\sigma}_{\tau-}^2 - \sigma_{\tau-}^2\big] = \mathcal{O}(1). \tag{16}$$

This includes a proof of (7). Denote by $\mathbb{P}_{\sigma_{(k-1)h_n}}$ the regular conditional probabilities conditioned on $\sigma_{(k-1)h_n}$, and by $\mathbb{E}_{\sigma_{(k-1)h_n}}$ the expectations with respect to the conditional measures. We obtain by the tower rule that

$$\mathbb{E}\big[h_n^{-1}\big(\tilde{m}_{k,n} - \tilde{m}_{k-1,n}^*\big)^2\big] = \mathbb{E}\big[h_n^{-1}\mathbb{E}_{\sigma_{(k-1)h_n}}\big[\big(\tilde{m}_{k,n} - \tilde{m}_{k-1,n}^*\big)^2\big]\big]$$

$$= \mathbb{E}\big[\mathbb{E}_{\sigma_{(k-1)h_n}}\big[\big(h_n^{-1/2}\tilde{m}_{k,n}\big)^2\big] + \mathbb{E}_{\sigma_{(k-1)h_n}}\big[\big(h_n^{-1/2}\tilde{m}_{k-1,n}^*\big)^2\big]$$

$$- 2\mathbb{E}_{\sigma_{(k-1)h_n}}\big[h_n^{-1/2}\tilde{m}_{k,n}\big]\mathbb{E}_{\sigma_{(k-1)h_n}}\big[h_n^{-1/2}\tilde{m}_{k-1,n}^*\big]\big] \tag{17}$$

by the conditional independence of $\tilde{m}_{k,n}$ and $\tilde{m}_{k-1,n}^*$.

We establish and use an approximation of the tail probabilities of $(\tilde{m}_{k,n})$ and $(\tilde{m}_{k-1,n}^*)$, respectively. For $x \in \mathbb{R}$, we have

$$\mathbb{P}_{\sigma_{(k-1)h_n}}\left(h_n^{-1/2}\min_{i\in\mathcal{I}_k^n}\left(\varepsilon_i + \sigma_{(k-1)h_n}(W_{t_i^n} - W_{kh_n})\right) > x\sigma_{(k-1)h_n}\right)$$

$$= \mathbb{P}_{\sigma_{(k-1)h_n}}\left(\min_{i\in\mathcal{I}_k^n}\left(h_n^{-1/2}(W_{t_i^n} - W_{kh_n}) + h_n^{-1/2}\sigma_{(k-1)h_n}^{-1}\varepsilon_i\right) > x\right)$$

$$= \mathbb{E}_{\sigma_{(k-1)h_n}}\left[\prod_{i=\lfloor knh_n\rfloor+1}^{\lfloor(k+1)nh_n\rfloor}\mathbb{P}\left(\varepsilon_i > h_n^{1/2}\sigma_{(k-1)h_n}\left(x - h_n^{-1/2}(W_{t_i^n} - W_{kh_n})\right) \mid \mathcal{F}^X\right)\right]$$

$$= \mathbb{E}_{\sigma_{(k-1)h_n}}\left[\exp\left(\sum_{i=\lfloor knh_n\rfloor+1}^{\lfloor(k+1)nh_n\rfloor}\log\left(1 - F_\eta\left(h_n^{1/2}\sigma_{(k-1)h_n}\left(x - h_n^{-1/2}(W_{t_i^n} - W_{kh_n})\right)\right)\right)\right)\right]$$

by the tower rule for conditional expectations, and since $\varepsilon_i \overset{\text{iid}}{\sim} F_\eta$. We have

$$W_{t_i^n} - W_{kh_n} = \sum_{j=1}^{i-\lfloor knh_n\rfloor}\tilde{U}_j, \quad \tilde{U}_j \overset{\text{iid}}{\sim} \mathcal{N}(0, n^{-1}),\ j\geq 2,\ \tilde{U}_1 \sim \mathcal{N}(0, t_{\lfloor knh_n\rfloor+1}^n - kh_n),$$

$$U_j = h_n^{-1/2}\tilde{U}_j, \qquad U_j \overset{\text{iid}}{\sim} \mathcal{N}(0, (nh_n)^{-1}),\ j\geq 2,\ U_1 \sim \mathcal{N}(0, h_n^{-1}(t_{\lfloor knh_n\rfloor+1}^n - kh_n)).$$

From (5), and with a first-order Taylor expansion of $z \mapsto \log(1 - z)$, we have

$$\log(1 - F_\eta(y)) \overset{(5)}{=} \log(1 - \eta y(1 + \mathcal{O}(1))) = -\eta y + \mathcal{O}(\eta y) = -\eta y_+ + \mathcal{O}(\eta y)$$

as $y \to 0$, where we add the positive part in the last equality since $F_\eta(y) = 0$ for any $y \leq 0$. We obtain

$$\mathbb{P}_{\sigma_{(k-1)h_n}}\left(h_n^{-1/2}\min_{i\in\mathcal{I}_k^n}\left(\varepsilon_i + \sigma_{(k-1)h_n}(W_{t_i^n} - W_{kh_n})\right) > x\sigma_{(k-1)h_n}\right)$$

$$= \mathbb{E}_{\sigma_{(k-1)h_n}}\left[\exp\left(-h_n^{1/2}\sigma_{(k-1)h_n}\eta\sum_{i=\lfloor knh_n\rfloor+1}^{\lfloor(k+1)nh_n\rfloor}\left(x - \sum_{j=1}^{i-\lfloor knh_n\rfloor}U_j\right)_+(1 + \mathcal{O}(1))\right)\right]$$

$$= \mathbb{E}_{\sigma_{(k-1)h_n}}\left[\exp\left(-h_n^{1/2}nh_n\sigma_{(k-1)h_n}\eta\int_0^1(B_t - x)_-\,\mathrm{d}t(1 + \mathcal{O}(1))\right)\right].$$

In the last equality we used that the Riemann sums tend almost surely to the integral with a standard Brownian motion $(B_t)_{t\geq 0}$ in the integrand. Since the expression in the expectation is bounded, as a product of conditional probabilities, by 1, we conclude with dominated convergence. If $nh_n^{3/2} \to \infty$, we deduce that

$$\mathbb{P}_{\sigma_{(k-1)h_n}}\left(h_n^{-1/2}\min_{i\in\mathcal{I}_k^n}\left(\varepsilon_i+\sigma_{(k-1)h_n}\left(W_{t_i^n}-W_{kh_n}\right)\right)>x\sigma_{(k-1)h_n}\right)$$

$$=\mathbb{P}\left(\inf_{0\le t\le 1}B_t\ge x\right)$$

$$+\mathbb{E}_{\sigma_{(k-1)h_n}}\left[\mathbf{1}\left(\inf_{0\le t\le 1}B_t<x\right)\exp\left(-h_n^{3/2}n\sigma_{(k-1)h_n}\eta\int_0^1(B_t-x)_-\,\mathrm{d}t(1+\mathcal{O}(1))\right)\right]$$

$$=\mathbb{P}\left(\inf_{0\le t\le 1}B_t\ge x\right)+\mathbb{P}\left(\inf_{0\le t\le 1}B_t<x\right)\cdot\mathcal{O}(1). \tag{18}$$

We do not have a lower bound for $\int_0^1(B_t-x)_-\,\mathrm{d}t$. However, using that the first entry time T_x of (B_t) in x, conditional on $\{\inf_{0\le t\le 1}B_t<x\}$, has a continuous conditional density $f(t\mid T_x<1)$, by Lemma 1 and properties of the Brownian motion we obtain, for any $\delta>0$,

$$\mathbb{E}_{\sigma_{(k-1)h_n}}\left[\mathbf{1}\left(\inf_{0\le t\le 1}B_t<x\right)\exp\left(-h_n^{3/2}n\sigma_{(k-1)h_n}\eta\int_0^1(B_t-x)_-\,\mathrm{d}t\right)\right]$$

$$\le\exp\left(-\left(h_n^{3/2}n\right)^\delta\sigma_{(k-1)h_n}\eta\right)\mathbb{P}\left(\inf_{0\le t\le 1}B_t<x\right)$$

$$+\mathbb{P}\left(\inf_{0\le t\le 1}B_t<x,\int_0^1(B_t-x)_-\,\mathrm{d}t\le\left(h_n^{3/2}n\right)^{-1+\delta}\right)$$

$$\le\left(\exp\left(-\left(h_n^{3/2}n\right)^\delta\sigma_{(k-1)h_n}\eta\right)+\int_0^1\mathbb{P}\left(\int_s^1(B_t)_-\,\mathrm{d}t\le\left(h_n^{3/2}n\right)^{-1+\delta}\right)f(s\mid T_x<1)\,\mathrm{d}s\right)$$

$$\times\mathbb{P}\left(\inf_{0\le t\le 1}B_t<x\right)$$

$$\le\left(\exp\left(-\left(h_n^{3/2}n\right)^\delta\sigma_{(k-1)h_n}\eta\right)\right.$$

$$\left.+\int_0^1\mathbb{P}\left((1-s)\int_0^1(B_t)_-\,\mathrm{d}t\le\left(h_n^{3/2}n\right)^{-1+\delta}\right)f(s\mid T_x<1)\,\mathrm{d}s\right)\mathbb{P}\left(\inf_{0\le t\le 1}B_t<x\right).$$

We focus on the second addend of the first factor, since the exponential term decays faster. It is bounded by a constant times

$$\int_0^1\mathbb{P}\left((1-s)\int_0^1(B_t)_-\,\mathrm{d}t\le\left(h_n^{3/2}n\right)^{-1+\delta}\right)\mathrm{d}s$$

$$\le\int_0^{1-b_n}\mathbb{P}\left((1-s)\int_0^1(B_t)_-\,\mathrm{d}t\le\left(h_n^{3/2}n\right)^{-1+\delta}\right)\mathrm{d}s+\int_{1-b_n}^1\mathrm{d}s$$

$$\le\mathbb{P}\left(b_n\int_0^1(B_t)_-\,\mathrm{d}t\le\left(h_n^{3/2}n\right)^{-1+\delta}\right)+b_n=\mathcal{O}\left(\left(h_n^{3/2}nb_n^{-1}\right)^{-(1+\delta)/3}+b_n\right)$$

for any sequence (b_n), $b_n\in(0,1)$, where we used Lemma 1. Choosing a b_n which minimizes the order yields

$$\mathbb{E}_{\sigma_{(k-1)h_n}}\left[\mathbf{1}\left(\inf_{0\le t\le 1}B_t<x\right)\exp\left(-h_n^{3/2}n\sigma_{(k-1)h_n}\eta\int_0^1(B_t-x)_-\,\mathrm{d}t\right)\right]=\mathbb{P}\left(\inf_{0\le t\le 1}B_t<x\right)\cdot R_n$$

almost surely, with a remainder that satisfies $R_n = \mathcal{O}\big((h_n^{3/2}n)^{-(1+\delta)/4}\big)$. From the unconditional Lévy distribution of T_x, $f(s \mid T_x < 1)$ is explicit, but we omit its precise form which does not influence the asymptotic order. Under the condition $nh_n^{3/2} \to \infty$, the minimum of the Brownian motion over the interval hence dominates the noise in the distribution of local minima, different than for the choice $h_n \propto n^{-2/3}$. By the reflection principle,

$$\mathbb{P}\Big(-\inf_{0 \le t \le 1} B_t \ge x\Big) = \mathbb{P}\Big(\sup_{0 \le t \le 1} B_t \ge x\Big) = 2\mathbb{P}(B_1 \ge x) = \mathbb{P}(|B_1| \ge x) \tag{19}$$

for $x \ge 0$.

Using the illustration of moments by integrals over tail probabilities we exploit this, and a completely analogous estimate for $\tilde{m}_{k-1,n}^*$, to approximate conditional expectations. This yields, for all $k \in \{1, \ldots, h_n^{-1} - 1\}$,

$$\mathbb{E}_{\sigma_{(k-1)h_n}}\big[h_n^{-1/2}\tilde{m}_{k,n}\big] = \int_0^\infty \mathbb{P}_{\sigma_{(k-1)h_n}}\big(h_n^{-1/2}\tilde{m}_{k,n} > x\big)\,\mathrm{d}x$$

$$- \int_0^\infty \mathbb{P}_{\sigma_{(k-1)h_n}}\big(-h_n^{-1/2}\tilde{m}_{k,n} > x\big)\,\mathrm{d}x$$

$$= -\int_0^\infty \mathbb{P}_{\sigma_{(k-1)h_n}}\Big(\sigma_{(k-1)h_n}\sup_{0 \le t \le 1} B_t > x\Big)\,\mathrm{d}x + \mathcal{O}_{\mathbb{P}}(1)$$

$$= -\int_0^\infty \mathbb{P}_{\sigma_{(k-1)h_n}}\big(\sigma_{(k-1)h_n}|B_1| > x\big)\,\mathrm{d}x + \mathcal{O}_{\mathbb{P}}(1)$$

$$= -\mathbb{E}_{\sigma_{(k-1)h_n}}\big[\sigma_{(k-1)h_n}|B_1|\big] + \mathcal{O}_{\mathbb{P}}(1)$$

$$= -\sqrt{\frac{2}{\pi}}\sigma_{(k-1)h_n} + \mathcal{O}_{\mathbb{P}}(1).$$

We used (19). An analogous computation yields the same result for $\tilde{m}_{k-1,n}^*$:

$$\mathbb{E}_{\sigma_{(k-1)h_n}}\big[h_n^{-1/2}\tilde{m}_{k-1,n}^*\big] = -\sqrt{\frac{2}{\pi}}\sigma_{(k-1)h_n} + \mathcal{O}_{\mathbb{P}}(1).$$

For the second conditional moments, we obtain, for all $k \in \{1, \ldots, h_n^{-1} - 1\}$,

$$\mathbb{E}_{\sigma_{(k-1)h_n}}\big[h_n^{-1}\big(\tilde{m}_{k,n}\big)^2\big] = 2\int_0^\infty x\mathbb{P}_{\sigma_{(k-1)h_n}}\big(\big|h_n^{-1/2}\tilde{m}_{k,n}\big| > x\big)\,\mathrm{d}x$$

$$= 2\int_0^\infty x\mathbb{P}_{\sigma_{(k-1)h_n}}\Big(\sigma_{(k-1)h_n}\sup_{0 \le t \le 1} B_t > x\Big)\,\mathrm{d}x + \mathcal{O}_{\mathbb{P}}(1)$$

$$= 2\int_0^\infty x\mathbb{P}_{\sigma_{(k-1)h_n}}\big(\sigma_{(k-1)h_n}|B_1| > x\big)\,\mathrm{d}x + \mathcal{O}_{\mathbb{P}}(1)$$

$$= \sigma_{(k-1)h_n}^2 + \mathcal{O}_{\mathbb{P}}(1).$$

The last identity uses the illustration of the second moment of the normal distribution as an integral over tail probabilities. An analogous computation yields

$$\mathbb{E}_{\sigma_{(k-1)h_n}}\big[h_n^{-1}\big(\tilde{m}_{k-1,n}^*\big)^2\big] = \sigma_{(k-1)h_n}^2 + \mathcal{O}_{\mathbb{P}}(1).$$

Inserting the identities for the conditional moments in (17) yields

$$\mathbb{E}\big[h_n^{-1}\big(\tilde{m}_{k,n} - \tilde{m}_{k-1,n}^*\big)^2\big] = 2\left(1 - \frac{2}{\pi}\right)\mathbb{E}\big[\sigma_{(k-1)h_n}^2\big] + \mathcal{O}(1)$$

such that

$$\mathbb{E}\big[\hat{\sigma}_{\tau-}^2 - \sigma_{\tau-}^2\big] = \frac{\pi}{2(\pi - 2)K_n} \sum_{k=(\lfloor h_n^{-1}\tau \rfloor - K_n)\vee 1}^{\lfloor h_n^{-1}\tau \rfloor - 1} h_n^{-1}\mathbb{E}\big[\big(\tilde{m}_{k,n} - \tilde{m}_{k-1,n}^*\big)^2\big] - \mathbb{E}[\sigma_{\tau-}^2] + \mathcal{O}(1)$$

$$= \frac{1}{K_n} \sum_{k=(\lfloor h_n^{-1}\tau \rfloor - K_n)\vee 1}^{\lfloor h_n^{-1}\tau \rfloor - 1} \mathbb{E}\big[\sigma_{(k-1)h_n}^2 - \sigma_{\tau-}^2\big] + \mathcal{O}(1) = \mathcal{O}(1).$$

This proves (16). Since the next step shows that the variance of the estimator tends to zero, consistency holds true.

Step 4 We determine the asymptotic variance of the estimator. Illustrating moments as integrals over tail probabilities, with the analogous approximation as above, we obtain, for all $k \in \{1, \ldots, h_n^{-1} - 1\}$,

$$\mathrm{var}_{\sigma_{(k-1)h_n}}\big(\tilde{m}_{k,n}^2\big) = \mathbb{E}_{\sigma_{(k-1)h_n}}\big[\tilde{m}_{k,n}^4\big] - \big(\mathbb{E}_{\sigma_{(k-1)h_n}}\big[\tilde{m}_{k,n}^2\big]\big)^2$$

$$= 2\sigma_{(k-1)h_n}^4 h_n^2 + \mathcal{O}_{\mathbb{P}}(h_n^2),$$

$$\mathrm{cov}_{\sigma_{(k-1)h_n}}\big(\tilde{m}_{k,n}^2, \tilde{m}_{k,n}\tilde{m}_{k-1,n}^*\big) = \mathbb{E}_{\sigma_{(k-1)h_n}}\big[\tilde{m}_{k,n}^3\big]\mathbb{E}_{\sigma_{(k-1)h_n}}\big[\tilde{m}_{k-1,n}^*\big]$$

$$- \mathbb{E}_{\sigma_{(k-1)h_n}}\big[\tilde{m}_{k,n}^2\big]\mathbb{E}_{\sigma_{(k-1)h_n}}\big[\tilde{m}_{k,n}\big]\mathbb{E}_{\sigma_{(k-1)h_n}}\big[\tilde{m}_{k-1,n}^*\big]$$

$$= \frac{2}{\pi}\sigma_{(k-1)h_n}^4 h_n^2 + \mathcal{O}_{\mathbb{P}}(h_n^2),$$

$$\mathrm{var}_{\sigma_{(k-1)h_n}}\big(\tilde{m}_{k,n}\tilde{m}_{k-1,n}^*\big) = \mathbb{E}_{\sigma_{(k-1)h_n}}\big[\tilde{m}_{k,n}^2\big]\mathbb{E}_{\sigma_{(k-1)h_n}}\big[\big(\tilde{m}_{k-1,n}^*\big)^2\big]$$

$$- \big(\mathbb{E}_{\sigma_{(k-1)h_n}}\big[\tilde{m}_{k,n}\big]\mathbb{E}_{\sigma_{(k-1)h_n}}\big[\tilde{m}_{k-1,n}^*\big]\big)^2$$

$$= \sigma_{(k-1)h_n}^4\left(1 - \frac{4}{\pi^2}\right)h_n^2 + \mathcal{O}_{\mathbb{P}}(h_n^2).$$

We have used the first four moments of the half-normal distribution and their illustration via integrals over tail probabilities. The dependence structure between $\tilde{m}_{k,n}$ and $\tilde{m}_{k,n}^*$ also affects the variance of $\hat{\sigma}_{\tau-}^2$. We perform approximation steps for covariances similar to those for the moments of local minima above, using

$$h_n^{-1} \mathrm{cov}_{\sigma_{(k-1)h_n}}\big(\tilde{m}_{k,n}, \tilde{m}_{k,n}^*\big)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Big(\mathbb{P}_{\sigma_{(k-1)h_n}}\big(h_n^{-1/2}\tilde{m}_{k,n} > x, h_n^{-1/2}\tilde{m}_{k,n}^* > y\big)$$

$$- \mathbb{P}_{\sigma_{(k-1)h_n}}\big(h_n^{-1/2}\tilde{m}_{k,n} > x\big)\mathbb{P}_{\sigma_{(k-1)h_n}}\big(h_n^{-1/2}\tilde{m}_{k,n}^* > y\big) \Big) \, \mathrm{d}x \, \mathrm{d}y$$

$$= \int_0^{\infty} \int_0^{\infty} \Big(\mathbb{P}_{\sigma_{(k-1)h_n}}\Big(\sigma_{(k-1)h_n} \sup_{0 \le t \le 1} B_t > x, \sigma_{(k-1)h_n}\big(\sup_{0 \le t \le 1} B_t - B_1\big) > y\Big)$$

$$- \mathbb{P}_{\sigma_{(k-1)h_n}}\Big(\sigma_{(k-1)h_n} \sup_{0 \le t \le 1} B_t > x\Big)\mathbb{P}_{\sigma_{(k-1)h_n}}\Big(\sigma_{(k-1)h_n}\big(\sup_{0 \le t \le 1} B_t - B_1\big) > y\Big)\Big) \mathrm{d}x \, \mathrm{d}y$$

$$+ \mathcal{O}_{\mathbb{P}}(1).$$

This shows that the joint distribution of $(\tilde{m}_{k,n}, \tilde{m}_{k,n}^*)$ relates to the distribution of the minimum and the difference between the minimum and the endpoint of Brownian motion over an interval, or equivalently the distribution of the maximum and the difference between the maximum and the endpoint. The latter is readily obtained from the joint density of the maximum and the endpoint, which is a well-known result on stochastic processes; see, e.g., [26]. Utilizing this, we obtain, for all $k \in \{1, \ldots, h_n^{-1} - 1\}$,

$$\mathrm{cov}_{\sigma_{(k-1)h_n}}\big(\tilde{m}_{k,n}, \tilde{m}_{k,n}^*\big) = \left(\frac{1}{2} - \frac{2}{\pi}\right) h_n \sigma_{(k-1)h_n}^2\big(1 + \mathcal{O}_{\mathbb{P}}\big(h_n^\alpha\big)\big) + \mathcal{O}_{\mathbb{P}}(h_n).$$

The additional remainder of order h_n^α in probability is due to the different approximations of (σ_t) in $\tilde{m}_{k,n}$ and $\tilde{m}_{k,n}^*$. This implies that, for all $k \in \{1, \ldots, h_n^{-1} - 1\}$,

$$\mathrm{cov}_{\sigma_{(k-1)h_n}}\big(\tilde{m}_{k,n}\tilde{m}_{k-1,n}^*, \tilde{m}_{k+1,n}\tilde{m}_{k,n}^*\big)$$

$$= \big(\mathbb{E}_{\sigma_{(k-1)h_n}}[\tilde{m}_{k,n}\tilde{m}_{k,n}^*] - \mathbb{E}_{\sigma_{(k-1)h_n}}[\tilde{m}_{k,n}]\mathbb{E}_{\sigma_{(k-1)h_n}}[\tilde{m}_{k,n}^*]\big)\mathbb{E}_{\sigma_{(k-1)h_n}}[\tilde{m}_{k-1,n}^*]\mathbb{E}[\tilde{m}_{k+1,n}]$$

$$= \sigma_{(k-1)h_n}^4 \left(\frac{1}{\pi} - \frac{4}{\pi^2}\right) h_n^2 + \mathcal{O}_{\mathbb{P}}\big(h_n^2\big).$$

With analogous steps, we deduce two more covariances which contribute to the asymptotic variance:

$$\mathrm{cov}_{\sigma_{(k-1)h_n}}\big(\tilde{m}_{k,n}^2, \big(\tilde{m}_{k,n}^*\big)^2\big) = -h_n^2 \frac{\sigma_{(k-1)h_n}^4}{2} + \mathcal{O}_{\mathbb{P}}\big(h_n^2\big),$$

$$\mathrm{cov}_{\sigma_{(k-1)h_n}}\big(\big(\tilde{m}_{k,n}^*\big)^2, m_k \tilde{m}_{k-1,n}^*\big) = -h_n^2 \frac{2}{3\pi} \sigma_{(k-1)h_n}^4 + \mathcal{O}_{\mathbb{P}}\big(h_n^2\big).$$

All covariance terms which enter the asymptotic variance are of one of these forms. For the conditional variance given $\sigma_{\tau-}^2$, we obtain

$$\text{var}_{\sigma_{\tau-}^2}\left(\hat{\sigma}_{\tau-}^2\right)$$

$$= \frac{1}{K_n^2}\frac{\pi^2}{4(\pi-2)^2}\left(\sum_{k=(\lfloor h_n^{-1}\tau\rfloor-K_n)\vee 1}^{\lfloor h_n^{-1}\tau\rfloor-1} h_n^{-2}\text{var}_{\sigma_{\tau-}^2}\left(\tilde{m}_{k,n}^2 + (\tilde{m}_{k,n}^*)^2 - 2\tilde{m}_{k,n}\tilde{m}_{k-1,n}^*\right)\right.$$

$$\left. - \sum_{k=(\lfloor h_n^{-1}\tau\rfloor-K_n)\vee 2}^{\lfloor h_n^{-1}\tau\rfloor-1} 4h_n^{-2}\text{cov}_{\sigma_{\tau-}^2}\left(\tilde{m}_{k,n}\tilde{m}_{k-1,n}^*, \tilde{m}_{k-1,n}^2 + (\tilde{m}_{k-1,n}^*)^2 - 2\tilde{m}_{k-1,n}\tilde{m}_{k-2,n}^*\right)\right)$$

$$+ \mathcal{O}_{\mathbb{P}}\left(K_n^{-1}\right)$$

$$= \frac{1}{K_n^2}\frac{\pi^2}{4(\pi-2)^2}\left(\sum_{k=(\lfloor h_n^{-1}\tau\rfloor-K_n)\vee 1}^{\lfloor h_n^{-1}\tau\rfloor-1} h_n^{-2}\left(2\text{var}_{\sigma_{\tau-}^2}\left(\tilde{m}_{k,n}^2\right) + 4\text{var}_{\sigma_{\tau-}^2}\left(\tilde{m}_{k,n}\tilde{m}_{k-1,n}^*\right)\right.\right.$$

$$\left.\left. + 2\text{cov}_{\sigma_{\tau-}^2}\left(\tilde{m}_{k,n}^2, (\tilde{m}_{k,n}^*)^2\right) - 4\text{cov}_{\sigma_{\tau-}^2}\left(\tilde{m}_{k,n}^2, \tilde{m}_{k,n}\tilde{m}_{k-1,n}^*\right) - 4\text{cov}_{\sigma_{\tau-}^2}\left((\tilde{m}_{k,n}^*)^2, \tilde{m}_{k,n}\tilde{m}_{k-1,n}^*\right)\right)\right.$$

$$\left. + \sum_{k=(\lfloor h_n^{-1}\tau\rfloor-K_n)\vee 2}^{\lfloor h_n^{-1}\tau\rfloor-1} 4h_n^{-2}\left(2\text{cov}_{\sigma_{\tau-}^2}\left(\tilde{m}_{k,n}\tilde{m}_{k-1,n}^*, \tilde{m}_{k-1,n}\tilde{m}_{k-2,n}^*\right) - \text{cov}_{\sigma_{\tau-}^2}\left(\tilde{m}_{k,n}\tilde{m}_{k-1,n}^*, \tilde{m}_{k-1,n}^2\right)\right.\right.$$

$$\left.\left. - \text{cov}_{\sigma_{\tau-}^2}\left(\tilde{m}_{k,n}\tilde{m}_{k-1,n}^*, (\tilde{m}_{k-1,n}^*)^2\right)\right)\right) + \mathcal{O}_{\mathbb{P}}\left(K_n^{-1}\right)$$

$$= \frac{1}{K_n}\frac{\pi^2}{4(\pi-2)^2}\sigma_{\tau-}^4\left(8 - \frac{16}{\pi^2} - 1 - \frac{8}{\pi} + \frac{8}{3\pi} + 2\left(\frac{4}{3\pi} - \frac{16}{\pi^2}\right)\right) + \mathcal{O}_{\mathbb{P}}\left(K_n^{-1}\right)$$

$$= \frac{1}{K_n}\frac{1}{(\pi-2)^2}\left(\frac{7\pi^2}{4} - \frac{2\pi}{3} - 12\right)\sigma_{\tau-}^4 + \mathcal{O}_{\mathbb{P}}\left(K_n^{-1}\right).$$

Step 5 For a central limit theorem, the squared bias needs to be asymptotically negligible compared to the variance, which is satisfied for $K_n = \mathcal{O}(h_n^{-2\alpha/(1+2\alpha)})$. By the existence of higher moments of $\tilde{m}_{k,n}$ and $\tilde{m}_{k-1,n}^*$, a Lyapunov-type condition is straightforward, such that asymptotic normality conditional on $\sigma_{\tau-}^2$ is implied by a classical central limit theorem for m-dependent triangular arrays such as the one in [3]. A feasible central limit theorem is implied by this conditional asymptotic normality in combination with $\mathcal{F}^X$-stable convergence. For the stability, we show that the $\alpha_n = K_n^{1/2}\left(\hat{\sigma}_{\tau-}^2 - \sigma_{\tau-}^2\right)$ satisfy

$$\mathbb{E}[Zg(\alpha_n)] \to \mathbb{E}[Zg(\alpha)] = \mathbb{E}[Z]\mathbb{E}[g(\alpha)] \tag{20}$$

for any $\mathcal{F}^X$-measurable bounded random variable Z and continuous bounded function g, where

$$\alpha = \sigma_{\tau-}^2\frac{1}{(\pi-2)}\sqrt{\frac{7\pi^2}{4} - \frac{2\pi}{3} - 12}\,U,$$

with U a standard normally distributed random variable which is independent of $\mathcal{F}^X$. By the above approximations it suffices to prove this for the statistics based on $\tilde{m}_{k,n}$ and $\tilde{m}_{k-1,n}^*$ from

(14), and Z measurable with respect to $\sigma\left(\int_0^t \sigma_s \, dW_s, 0 \le t \le 1\right)$. Set

$$A_n = [\tau - (K_n + 1)h_n, \tau], \quad \tilde{X}(n)_t = \int_0^t \mathbf{1}_{A_n}(s)\sigma_{\lfloor sh_n^{-1}\rfloor h_n} \, dW_s, \quad \bar{X}(n)_t = X_t - \tilde{X}(n)_t.$$

Denote with $\mathcal{H}_n$ the σ-field generated by $\bar{X}(n)_t$ and $\mathcal{F}_0^X$. The sequence $(\mathcal{H}_n)_{n \in \mathbb{N}}$ is isotonic with limit $\bigvee_n \mathcal{H}_n = \sigma\left(\int_0^t \sigma_s \, dW_s, 0 \le t \le 1\right)$. Since $\mathbb{E}[Z \mid \mathcal{H}_n] \to Z$ in $L^1(\mathbb{P})$ as $n \to \infty$, it is enough to show that $\mathbb{E}[Zg(\alpha_n)] \to \mathbb{E}[Z]\mathbb{E}[g(\alpha)]$ for Z being $\mathcal{H}_{n_0}$-measurable for some $n_0 \in \mathbb{N}$. Observe that α_n includes only increments of local minima based on $\tilde{X}(n)_t$, which are uncorrelated from those of $\bar{X}(n)_t$. For all $n \ge n_0$, we hence obtain $\mathbb{E}[Zg(\alpha_n)] = \mathbb{E}[Z]\mathbb{E}[g(\alpha_n)] \to \mathbb{E}[Z]\mathbb{E}[g(\alpha)]$ by a standard central limit theorem. This shows (20), and completes the proof of (12). $\qquad \square$

Proof of Proposition 1. For the quarticity estimator (10), when $\lfloor h_n^{-1}\tau \rfloor > K_n$ we have

$$\mathbb{E}\left[\widehat{\sigma_{\tau-}^4} - \sigma_{\tau-}^4\right] = \frac{\pi}{4(3\pi - 8)K_n} \sum_{k=(\lfloor h_n^{-1}\tau \rfloor - K_n)\vee 1}^{\lfloor h_n^{-1}\tau \rfloor - 1} h_n^{-2}\mathbb{E}\left[\tilde{m}_{k,n}^4 + (\tilde{m}_{k-1,n}^*)^4 - 4\tilde{m}_{k,n}^3\tilde{m}_{k-1,n}^* \right.$$

$$\left. - 4\tilde{m}_{k,n}(\tilde{m}_{k-1,n}^*)^3 + 6\tilde{m}_{k,n}^2(\tilde{m}_{k-1,n}^*)^2\right] - \mathbb{E}[\sigma_{\tau-}^4] + \mathcal{O}\left(h_n^{\alpha \wedge 1/2}\right)$$

$$= \left(\frac{\pi}{4(3\pi - 8)}\left(6 - \frac{16}{\pi} - \frac{16}{\pi} + 6\right) - 1\right)\mathbb{E}[\sigma_{\tau-}^4] + \mathcal{O}(1)$$

$$= \mathcal{O}(1)$$

by using the same moments as in the computation of the asymptotic variance. We can bound its variance by

$$\mathrm{var}\left(\widehat{\sigma_{\tau-}^4}\right) \le \frac{\pi^2}{16(3\pi - 8)^2 K_n^2} 2K_n h_n^{-4} \mathrm{var}\left((\tilde{m}_{k,n} - \tilde{m}_{k-1,n}^*)^4\right) + \mathcal{O}\left(K_n^{-1}\right)$$

$$\le \frac{1}{K_n}\frac{\pi^2}{8(3\pi - 8)^2} h_n^{-4}\mathbb{E}\left[(\tilde{m}_{k,n} - \tilde{m}_{k-1,n}^*)^8\right] + \mathcal{O}\left(K_n^{-1}\right)$$

$$\le \frac{1}{K_n}\frac{\pi^2}{8(3\pi - 8)^2} h_n^{-4} 256\mathbb{E}[\tilde{m}_{k,n}^8] + \mathcal{O}\left(K_n^{-1}\right) = \mathcal{O}(K_n^{-1}),$$

which readily implies Proposition 1. $\qquad \square$

6.3. Asymptotics of the truncated spot volatility estimation with jumps

Proof of Theorem 2. Denote by $D_k^X := m_{k,n} - m_{k-1,n}$, $k = 1, \ldots, h_n^{-1} - 1$, the differences of local minima based on the observations (4), with the general semimartingale (3) with jumps. Denote by $D_k^C := \tilde{m}_{k,n} - \tilde{m}_{k-1,n}^*$, $k = 1, \ldots, h_n^{-1} - 1$, the differences of the unobservable local minima considered in Section 6.2. In particular, the statistics D_k^C are based only on the continuous part (C_t) in (3) such that the jumps are eliminated. Theorem 2 is implied by Proposition 1 if we can show that

$$\frac{\pi}{2(\pi - 2)K_n} \sum_{k=(\lfloor h_n^{-1}\tau \rfloor - K_n)\vee 1}^{\lfloor h_n^{-1}\tau \rfloor - 1} h_n^{-1}\left((D_k^X)^2\mathbf{1}_{\{|D_k^X| \le u_n\}} - (D_k^C)^2\right) = \mathcal{O}_\mathbb{P}\left(h_n^{\alpha/(2\alpha+1)}\right) = \mathcal{O}_\mathbb{P}\left(K_n^{-1/2}\right).$$

We decompose this difference of the truncated estimator, which is based on the available observations with jumps, and the non-truncated estimator, which uses non-available observations without jumps, in the following way:

$$\frac{\pi}{2(\pi-2)K_n} \sum_{k=(\lfloor h_n^{-1}\tau\rfloor-K_n)\vee 1}^{\lfloor h_n^{-1}\tau\rfloor-1} h_n^{-1}\big((D_k^X)^2 \mathbf{1}_{\{|D_k^X|\leq u_n\}} - (D_k^C)^2\big)$$

$$= \frac{\pi}{2(\pi-2)K_n} \sum_{k=(\lfloor h_n^{-1}\tau\rfloor-K_n)\vee 1}^{\lfloor h_n^{-1}\tau\rfloor-1} h_n^{-1}\Big(\mathbf{1}_{\{|D_k^C|>cu_n\}}\big((D_k^X)^2 \mathbf{1}_{\{|D_k^X|\leq u_n\}} - (D_k^C)^2\big)$$

$$+ \mathbf{1}_{\{|D_k^C|\leq cu_n\}}\mathbf{1}_{\{|D_k^X|\leq u_n\}}\big((D_k^X)^2 - (D_k^C)^2\big) - \mathbf{1}_{\{|D_k^C|\leq cu_n\}}\mathbf{1}_{\{|D_k^X|>u_n\}}(D_k^C)^2\Big)$$

with some arbitrary constant $c \in (0, 1)$. Without loss of generality we can set $\beta = 1$ in this proof, i.e. $u_n = h_n^\kappa$. We consider the three addends, which are different error terms, separately by

- large absolute statistics based on the continuous part (C_t);
- non-truncated statistics which contain (small) jumps;
- the truncation also of the continuous parts in the statistics (D_k^X) which exceed the threshold.

The probability $\mathbb{P}(|D_k^C| > cu_n)$ can be bounded using the estimate from (18) and Gaussian tail bounds. Observe that the remainder in (18) is non-negative. This yields that, for some $y > 0$, we have

$$\mathbb{P}\big(h_n^{-1/2}|\tilde{m}_{k,n}| > y\big) \leq \mathbb{P}\Big(\sup_{0\leq t\leq 1} B_t > y\Big),$$

which is intuitive, since the errors (ε_i) are non-negative. We apply the triangular inequality and then Hölder's inequality to the expectation of the absolute first error term and obtain, for any $p \in \mathbb{N}$,

$$\frac{\pi}{2(\pi-2)K_n} \mathbb{E}\Bigg[\bigg|\sum_{k=(\lfloor h_n^{-1}\tau\rfloor-K_n)\vee 1}^{\lfloor h_n^{-1}\tau\rfloor-1} h_n^{-1}\mathbf{1}_{\{|D_k^C|>cu_n\}}\big((D_k^X)^2 \mathbf{1}_{\{|D_k^X|\leq u_n\}} - (D_k^C)^2\big)\bigg|\Bigg]$$

$$\leq \frac{\pi}{2(\pi-2)K_n} \sum_{k=(\lfloor h_n^{-1}\tau\rfloor-K_n)\vee 1}^{\lfloor h_n^{-1}\tau\rfloor-1} h_n^{-1}\mathbb{E}\big[\mathbf{1}_{\{|D_k^C|>cu_n\}}\big|(D_k^X)^2 \mathbf{1}_{\{|D_k^X|\leq u_n\}} - (D_k^C)^2\big|\big]$$

$$\leq \frac{\pi}{2(\pi-2)K_n} \sum_{k=(\lfloor h_n^{-1}\tau\rfloor-K_n)\vee 1}^{\lfloor h_n^{-1}\tau\rfloor-1} h_n^{-1}\big(\mathbb{P}(|D_k^C|>cu_n)2\,(u_n^4 + \mathbb{E}[(D_k^C)^4])\big)^{1/2}$$

$$\leq \frac{\pi}{2(\pi - 2)K_n} \sum_{k=(\lfloor h_n^{-1}\tau \rfloor - K_n)\vee 1}^{\lfloor h_n^{-1}\tau \rfloor - 1} h_n^{-1} \left(\mathbb{P}\left(h_n^{-1/2}|D_k^C| > ch_n^{\kappa-1/2} \right) \right)^{1/2} \sqrt{2}u_n^2$$

$$\leq \frac{\pi}{2(\pi - 2)K_n} \sum_{k=(\lfloor h_n^{-1}\tau \rfloor - K_n)\vee 1}^{\lfloor h_n^{-1}\tau \rfloor - 1} h_n^{-1} \left(2\mathbb{P}\left(|B_1| > \frac{c}{2}h_n^{\kappa-1/2} \right) \right)^{1/2} \sqrt{2}u_n^2$$

$$\leq \frac{\sqrt{2}\pi}{(\pi - 2)K_n} \sum_{k=(\lfloor h_n^{-1}\tau \rfloor - K_n)\vee 1}^{\lfloor h_n^{-1}\tau \rfloor - 1} h_n^{2\kappa-1} \exp\left(-\frac{c^2}{4}h_n^{2\kappa-1} \right)$$

$$= \mathcal{O}\left(h_n^{(-p+1)(2\kappa-1)} \right) = \mathcal{O}\left(h_n^{\alpha/(2\alpha+1)} \right).$$

Since $2\kappa - 1 < 0$ and p is arbitrarily large, we conclude that the first error term is asymptotically negligible. We will use the elementary inequalities

$$D_k^X = \min_{i\in\mathcal{I}_k^n} (C_{i/n} + J_{i/n} + \varepsilon_i) - \min_{i\in\mathcal{I}_{k-1}^n} (C_{i/n} + J_{i/n} + \varepsilon_i)$$

$$\leq \min_{i\in\mathcal{I}_k^n} (C_{i/n} + \varepsilon_i) + \max_{i\in\mathcal{I}_k^n} J_{i/n} - \min_{i\in\mathcal{I}_{k-1}^n} (C_{i/n} + \varepsilon_i) - \min_{i\in\mathcal{I}_{k-1}^n} J_{i/n}$$

$$= D_k^C + \max_{i\in\mathcal{I}_k^n} J_{i/n} - \min_{i\in\mathcal{I}_{k-1}^n} J_{i/n} + \mathcal{O}_{\mathbb{P}}\left(h_n^{\alpha\wedge 1/2} \right),$$

$$D_k^X = \min_{i\in\mathcal{I}_k^n} (C_{i/n} + J_{i/n} + \varepsilon_i) - \min_{i\in\mathcal{I}_{k-1}^n} (C_{i/n} + J_{i/n} + \varepsilon_i)$$

$$\geq \min_{i\in\mathcal{I}_k^n} (C_{i/n} + \varepsilon_i) + \min_{i\in\mathcal{I}_k^n} J_{i/n} - \min_{i\in\mathcal{I}_{k-1}^n} (C_{i/n} + \varepsilon_i) - \max_{i\in\mathcal{I}_{k-1}^n} J_{i/n}$$

$$= D_k^C + \min_{i\in\mathcal{I}_k^n} J_{i/n} - \max_{i\in\mathcal{I}_{k-1}^n} J_{i/n} + \mathcal{O}_{\mathbb{P}}\left(h_n^{\alpha\wedge 1/2} \right).$$

Therefore, we can bound $|D_k^X - D_k^C|$ by

$$\sup_{i\in\mathcal{I}_k^n, j\in\mathcal{I}_{k-1}^n} |J_{\frac{i}{n}} - J_{\frac{j}{n}}| \leq \sup_{s\in[kh_n,(k+1)h_n], t\in[(k-1)h_n,kh_n]} |J_s - J_t|$$

$$\leq \sup_{s\in[kh_n,(k+1)h_n]} |J_s - J_{kh_n}| + \sup_{t\in[(k-1)h_n,kh_n]} |J_{kh_n} - J_t|,$$

and the remainder term of the approximation for the continuous part, which is $\mathcal{O}_{\mathbb{P}}\left(h_n^{\alpha\wedge 1/2} \right)$. Since the compensated small jumps of a semimartingale admit a martingale structure, Doob's inequality for càdlàg L_2-martingales can be used to bound these suprema. Based on these preliminaries, we obtain, for the expected absolute value of the second error term,

$$\frac{\pi}{2(\pi-2)K_n}\mathbb{E}\left[\left|\sum_{k=(\lfloor h_n^{-1}\tau\rfloor-K_n)\vee 1}^{\lfloor h_n^{-1}\tau\rfloor-1} h_n^{-1}\mathbf{1}_{\{|D_k^C|\le cu_n\}}\mathbf{1}_{\{|D_k^X|\le u_n\}}\left((D_k^X)^2-(D_k^C)^2\right)\right|\right]$$

$$\le\frac{\pi}{2(\pi-2)K_n}\sum_{k=(\lfloor h_n^{-1}\tau\rfloor-K_n)\vee 1}^{\lfloor h_n^{-1}\tau\rfloor-1} h_n^{-1}\mathbb{E}\left[\mathbf{1}_{\{|D_k^C|\le cu_n\}}\mathbf{1}_{\{|D_k^X|\le u_n\}}\left|(D_k^X)^2-(D_k^C)^2\right|\right]$$

$$\lesssim\frac{1}{K_n}\sum_{k=(\lfloor h_n^{-1}\tau\rfloor-K_n)\vee 1}^{\lfloor h_n^{-1}\tau\rfloor-1} h_n^{-1}\mathbb{E}\left[\sup_{i\in\mathcal{I}_k^n, j\in\mathcal{I}_{k-1}^n}|J_{i/n}-J_{j/n}|^2\wedge(1+c)^2u_n^2\right]$$

$$\lesssim\frac{1}{K_n}\sum_{k=(\lfloor h_n^{-1}\tau\rfloor-K_n)\vee 1}^{\lfloor h_n^{-1}\tau\rfloor-1} h_n^{-1}\mathbb{E}\left[\sup_{t\in[kh_n,(k+1)h_n]}|J_t-J_{kh_n}|^2\wedge u_n^2\right]$$

$$\lesssim\frac{1}{K_n}\sum_{k=(\lfloor h_n^{-1}\tau\rfloor-K_n)\vee 1}^{\lfloor h_n^{-1}\tau\rfloor-1} h_n^{-1}\mathbb{E}\left[|J_{(k+1)h_n}-J_{kh_n}|^2\wedge u_n^2\right]=\mathcal{O}\left(u_n^{2-r}\right).$$

Applying the elementary inequalities from above, a cross term in the upper bound for $(D_k^X)^2-(D_k^C)^2$ is of smaller order and directly neglected. It can be handled using the Cauchy–Schwarz inequality. In the last step, we adopt a bound on the expected absolute thresholded jump increments from [1, (54)]. For the negligibility of the second error term, we thus get the condition that

$$\kappa(2-r)\ge\frac{\alpha}{1+2\alpha}. \tag{21}$$

Doob's inequality also yields

$$\mathbb{P}\left(\sup_{t\in[kh_n,(k+1)h_n]}|J_t-J_{kh_n}|\ge(1-c)u_n\right)\le\frac{\mathbb{E}[|J_{(k+1)h_n}-J_{kh_n}|^{r\wedge 1}]}{((1-c)u_n)^{r\wedge 1}}+\mathcal{O}(h_n)=\mathcal{O}\left(h_nu_n^{-r}\right).$$

For this upper bound, we decomposed the jumps in the sum of large jumps and the martingale of compensated small jumps, to which we applied Doob's inequality. We derive the following estimate for the expectation of the third (absolute) error term:

$$\frac{\pi}{2(\pi-2)K_n}\sum_{k=(\lfloor h_n^{-1}\tau\rfloor-K_n)\vee 1}^{\lfloor h_n^{-1}\tau\rfloor-1} h_n^{-1}\mathbb{E}\left[\mathbf{1}_{\{|D_k^C|\le cu_n\}}\mathbf{1}_{\{|D_k^X|>u_n\}}(D_k^C)^2\right]$$

$$\le\frac{\pi}{2(\pi-2)K_n}\sum_{k=(\lfloor h_n^{-1}\tau\rfloor-K_n)\vee 1}^{\lfloor h_n^{-1}\tau\rfloor-1} h_n^{-1}\mathbb{E}\left[\mathbf{1}_{\{2\sup_{s\in[(k-1)h_n,(k+1)h_n]}|J_s-J_{kh_n}|\ge(1-c)u_n\}}(D_k^C)^2\right]$$

$$\lesssim\frac{1}{K_n}\sum_{k=(\lfloor h_n^{-1}\tau\rfloor-K_n)\vee 1}^{\lfloor h_n^{-1}\tau\rfloor-1} h_n^{-1}\mathbb{P}\left(\sup_{t\in[kh_n,(k+1)h_n]}|J_t-J_{kh_n}|\ge(1-c)u_n\right)\mathbb{E}\left[(D_k^C)^2\right]$$

$$\lesssim \frac{1}{K_n} \sum_{k=(\lfloor h_n^{-1}\tau \rfloor - K_n)\vee 1}^{\lfloor h_n^{-1}\tau \rfloor -1} \left(\frac{\mathbb{E}[|J_{(k+1)h_n} - J_{kh_n}|^{r\wedge 1}]}{((1-c)u_n)^{r\wedge 1}} + \mathcal{O}(h_n) \right) = \mathcal{O}\big(h_n u_n^{-r}\big).$$

For the negligibility of the third error term, we thus get the condition that

$$1 - \kappa r \geq \frac{\alpha}{1+2\alpha}. \tag{22}$$

Since, under the conditions of Theorem 2, (21) and (22) are satisfied, the proof is finished by the negligibility of all addends in the decomposition above. $\qquad\square$

Acknowledgement

Funding information

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] AÏT-SAHALIA, Y. AND JACOD, J. (2010). Is Brownian motion necessary to model high-frequency data? *Ann. Statist.* **38**, 3093–3128.

[2] AÏT-SAHALIA, Y. AND JACOD, J. (2014). *High-Frequency Financial Econometrics*. Princeton University Press.

[3] BERK, K. N. (1973). A central limit theorem for *m*-dependent random variables with unbounded *m*. *Ann. Prob.* **1**, 352–354.

[4] BIBINGER, M., HAUTSCH, N., MALEC, P. AND REISS, M. (2019). Estimating the spot covariation of asset prices – Statistical theory and empirical evidence. *J. Business Econom. Statist.* **37**, 419–435.

[5] BIBINGER, M., JIRAK, M. AND REISS, M. (2016). Volatility estimation under one-sided errors with applications to limit order books. *Ann. Appl. Prob.* **26**, 2754–2790.

[6] BIBINGER, M., NEELY, C. AND WINKELMANN, L. (2019). Estimation of the discontinuous leverage effect: Evidence from the Nasdaq order book. *J. Econometrics* **209**. 158–184.

[7] BIBINGER, M. AND WINKELMANN, L. (2018). Common price and volatility jumps in noisy high-frequency data. *Electron. J. Statist.* **12**, 2018–2073.

[8] BISHWAL, J. P. N. (2022). *Parameter Estimation in Stochastic Volatility Models*. Springer, Cham.

[9] CHAKER, S. (2017). On high frequency estimation of the frictionless price: The use of observed liquidity variables. *J. Econometrics* **201**. 127–143.

[10] CLINET, S. AND POTIRON, Y. (2019). Testing if the market microstructure noise is fully explained by the informational content of some variables from the limit order book. *J. Econometrics* **209**, 289–337.

[11] CLINET, S. AND POTIRON, Y. (2021). Estimation for high-frequency data under parametric market microstructure noise. *Ann. Inst. Statist. Math.* **73**, 649–669.

[12] EL EUCH, O., FUKASAWA, M. AND ROSENBAUM, M. (2018). The microstructural foundations of leverage effect and rough volatility. *Finance Stoch.* **22**, 241–280.

[13] HANSEN, P. R. AND LUNDE, A. (2006). Realized variance and market microstructure noise. *J. Business Econom. Statist.* **24**, 127–161.

[14] HOFFMANN, M., MUNK A. AND SCHMIDT-HIEBER, J. (2012). Adaptive wavelet estimation of the diffusion coefficient under additive error measurements. *Ann. Inst. H. Poincaré Prob. Statist.* **48**, 1186–1216.

[15] JACOD, J. AND PROTTER, P. (2012). *Discretization of Processes*. Springer, Berlin.

[16] JANSON, S. (2007). Brownian excursion area, Wright's constants in graph enumeration, and other Brownian areas. *Prob. Surv.* **4**, 80–145.

[17] JIRAK M., MEISTER A. AND REISS, M. (2014). Adaptive function estimation in nonparametric regression with one-sided errors. *Ann. Statist.* **42**, 1970–2002.

[18] LI, Y., XIE, S. AND ZHENG, X. (2016). Efficient estimation of integrated volatility incorporating trading information. *J. Econometrics* **195** 33–50.

[19] LI, Z. M. AND LINTON, O. (2022). A ReMeDI for microstructure noise. *Econometrica* **90**, 367–389.

[20] LIU Y., LIU Q., LIU Z. AND DING D. (2017). Determining the integrated volatility via limit order books with multiple records. *Quant. Finance* **17**, 1697–1714.

[21] MANCINI, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scand. J. Statist.* **36**, 270–296.

[22] MANCINI, C., MATTIUSSI, V. AND RENÒ, R. (2015). Spot volatility estimation using delta sequences. *Finance Stoch.* **19**, 261–293.

[23] MEISTER, A. AND REISS, M. (2013). Asymptotic equivalence for nonparametric regression with non-regular errors. *Prob. Theory Relat. Fields* **155**, 201–229.

[24] REISS, M. AND WAHL, M. (2019). Functional estimation and hypothesis testing in nonparametric boundary models. *Bernoulli* **25**, 2597–2619.

[25] ROSENBAUM, M. AND TOMAS, M. (2021). From microscopic price dynamics to multidimensional rough volatility models. *Adv. Appl. Prob.* **53**, 425–462.

[26] SHEPP, L. A. (1979). The joint density of the maximum and its location for a Wiener process with drift. *J. Appl. Prob.* **16**, 423–427.

[27] TAKÁCS, L. (1996). On a generalization of the arc-sine law. *Ann. Appl. Prob.* **6**, 1035–1040.

[28] TAUCHEN, G. AND TODOROV, V. (2011). Volatility jumps. *J. Business Econom. Statist.* **29**, 356–371.