Spot Volatility Estimation Using Delta Sequences[*]

Cecilia Mancini,

Università di Firenze, via delle Pandette 9, 50127, Firenze, Italy

`cecilia.mancini@dmd.unifi.it`

V. Mattiussi

City University, Northampton Square, London EC1V 0HB, UK

`vanessa.mattiussi.1@city.ac.uk`

R. Renò

Università di Siena, Piazza S.Francesco 7, 53100 Siena, Italy

`reno@unisi.it`

July 11, 2012

Abstract. We introduce a class of nonparametric spot volatility estimators based on delta sequences and conceived to include many of the existing estimators in the field as special cases. The full limit theory is first derived when unevenly sampled observations under infill asymptotics and fixed time-horizon are considered, and the state variable is assumed to follow a Brownian semimartingale. We then extend our class of estimators to include Poisson jumps or financial microstructure noise in the observed price process. As an application of our results, we relate the Fourier estimator to a specific delta sequence obtained with the Fejér function. The proposed estimators are applied to data from the S&P500 stock index futures market.

1 Introduction

In the last decade, the larger availability of high-frequency financial data sets has spawned considerable econometric research on integrated volatility, and in particular on realized volatility (reviews on the topic can be found in Barndorff-Nielsen and Shephard [11], and in Bandi and Russell [9]). More recently, the interest has moved to study the variability of the price dynamics at a particular point in time, the so-called instantaneous or spot volatility. The usage of spot volatility estimates is also increasing in financial applications. For example, spot volatility estimates have been shown to be

1

beneficial, with respect to integrated estimates, in estimating infinitesimal cross-moments [8] and in testing for the presence of jumps [12]. Spot volatility is also the crucial ingredient in option pricing with stochastic volatility, where the initial volatility value, in addition to the initial value of the underlying, is needed to price the option.

With this paper, we try to widen and sustain the class of the existing estimators in the field by proposing a method to estimate the spot volatility of a univariate semimartingale which can be adapted in the presence of Poisson jumps or microstructure noise. More importantly, we allow for jumps in the volatility process; in this case, our estimator will converge to the average of spot volatilities observed immediately before and after the eventual jump.

One way of estimating instantaneous volatility consists in assuming that the volatility process is a deterministic function of the observable state variable, and nonparametric techniques can be applied both in the absence (see Florens-Zmirou [19], Bandi and Phillips [6], Renò [47] and Hoffman [23]) and in the presence of jumps in X (see Johannes [30], Bandi and Nguyen [5], and Mancini and Renò [36]). Fully nonparametric methods when volatility is instead a càdlàg process have been studied by Malliavin and Mancino [33, 34] and Kristensen [32] in the absence of jumps, and by Zu and Boswijk [53], Hoffmann, Munk and Schmidt-Hieber [22] and Ogawa and Sanfelici [42] in the absence of jumps but with noisy observations. Related studies include the idea of rolling sample volatility estimators in Foster and Nelson [20], see also Andreou and Ghysels [4], the theory of spot volatility estimation developed in Bandi and Renò [7], and the kernel based methods of Fan and Wang [17], and Mykland and Zhang [40]. In the presence of jumps (but absence of noise), spot volatility have been studied by Jacod and Protter [28], Ngo and Ogawa [41], Aït-Sahalia and Jacod [1] and Dobrev, Andersen and Schaumburg [15]. Alternatives are studied in Alvarez et al. [2], Genon-Catalot et al [21] and Hoffmann [24].

The purpose of our study is to define a large class of non parametric estimators of instantaneous volatility, which includes many of the aforementioned methods. Our intuition suggests that a spot volatility estimator can be written as the convolution of squared price returns with a sequence of functions, known as *delta sequence*, which converges to a Dirac delta function concentrating all the mass around one point (for applications of delta sequences in statistics see, for instance, Watson and Leadbetter [50], and Walter and Blum [49]). In particular, we extend the kernel estimator of Kristensen [32] by proving that a traditional kernel function can be seen as a delta sequence. Our class is shown to be reasonably wide and it includes the Fejér sequence used in the work of Malliavin and Mancino [34] and the indicator function used in the work of Jacod and Protter [28].

The study of the asymptotic theory (see Section 2) reveals that the estimators within the class are, under suitable conditions, normally distributed, when the number of observations diverges to infinity in a fixed interval $[0, T]$ and the maximum interval between the observations (not necessarily equally spaced) shrinks to zero. Our findings are derived under mild assumptions on the driving coefficients of the stochastic differential equation. In Section 3, we allow for microstructure noise in the data

2

and use a two-scale volatility technique, similar to the one in Zhang et al [52], to make our estimator robust against the noise. In addition, we tackle the problem of discontinuities in the return dynamics using a threshold estimator as in Mancini [35] to filter out jumps from the observed price process. In Section 4 we consider the work by Malliavin and Mancino [34] and further investigate the asymptotic behavior of their proposed Fourier estimator of spot σ. Section 5 presents an empirical analysis using high-frequency stock index futures, where the above estimators are applied to detect intraday volatility dynamics. Section 6 concludes.

2   Spot Volatility Estimation in the Basic Setting

In what follows, we will consider a univariate logarithmic price process X_t defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathcal{P})$ satisfying the usual conditions, see e.g. Protter [46]. Our results are based on the set of assumptions outlined below.

Assumption 1. *i) The logarithmic price X_t is the solution of the following stochastic differential equation*

$$dX_t = \mu_t dt + \sigma_t dW_t, \tag{2.1}$$

where the initial condition X_0 is measurable with respect to $\mathcal{F}_0$, W_t is a standard Brownian motion defined on the filtered probability space and μ_t, σ_t are adapted processes with càdlàg paths.
ii) Given a fixed point $\bar{t} \in [0, T]$, let $B_\varepsilon(\bar{t}) = [\bar{t} - \varepsilon, \bar{t} + \varepsilon]$, with fixed $\varepsilon > 0$, and assume that there exist $\Gamma > 0$, a sequence of stopping times $\tau_m \uparrow \infty$ and constants $C_{\bar{t}}^{(m)}$ such that for all m, for $(\omega, s) \in \Omega \times B_\varepsilon(\bar{t}) \cap [0, \tau_m(\omega)]$, a.s.

$$E_{u \wedge s}[|\sigma_u - \sigma_s|^2] \leq C_{\bar{t}}^{(m)} |u - s|^\Gamma, \tag{2.2}$$

where $\mathrm{E}_t[\cdot]$ denotes $\mathrm{E}[\cdot | \mathcal{F}_t]$.

The class of processes for σ_t we wish to estimate point-wise is larger than the class of the processes with differentiable paths, and it includes the important case where σ_t is generated itself by a Brownian motion as in a stochastic volatility model. Indeed, every càdlàg process σ_t is also locally bounded, and for every (possibly jumping) Itô semimartingale σ_t, the requirement (2.2) is satisfied (with $\Gamma = 1$) on every $[0, \tau_m]$ where σ_t is bounded, thanks to the Burkholder-Davis-Gundy (BDG hereafter) inequality (see [26], pag.25). In particular the estimator we are going to propose below is robust to jumps in volatility.

In order to work with irregular sampling, we adapt to our settings the concept of quadratic variation of time defined in Mykland and Zhang [39].

Assumption 2. *The process X_t is observed $n + 1$ times at deterministic instants $0 = t_0 < t_1 < \ldots < t_n = T$, not necessarily equally spaced and with T fixed. We set $\Delta_i = t_i - t_{i-1}$ and*

$\overline{\Delta}_n = \frac{T}{n}$ and assume $\max_{i=1,\dots,n}\{\Delta_i\} = O(\overline{\Delta}_n)$. *The quadratic variation of time up to a given $t \le T$ is defined as $H(t) = \lim_{n\to\infty} H_n(t)$, where*

$$H_n(t) = \frac{1}{\overline{\Delta}_n} \sum_{t_i \le t} (\Delta_i)^2. \tag{2.3}$$

Assuming that the above limit exists, we require that H is Lebesgue-almost surely differentiable in $[0,T]$, with H' such that for some $K \ge 0$ (not depending on i)

$$\left| H'(t_i) - \frac{\Delta_i}{\overline{\Delta}_n} \right| \le K\Delta_i. \tag{2.4}$$

In the special case of equally spaced observations, $\Delta_i = \overline{\Delta}_n$, $H'(t) = 1$ and (2.4) is satisfied with $K = 0$. When the observations are more (less) concentrated around t, then we have $H'(t) < 1$ ($H'(t) > 1$). The assumption $\max_{i=1,\dots,n}\{\Delta_i\} = O(\overline{\Delta}_n)$ is technical, and means that the partition should not vary asymptotically too wildly with respect to the equally spaced partition.

Condition in (2.4) for the partition $\{t_i\}_i$ is different from condition $v)$ in assumption A of [39]. For instance, consider the sequence of partitions where the amplitude of the first $[n/2]$ intervals $]t_{i-1}, t_i]$ is 2Δ and of the remaining $n-[n/2]$ is Δ. Then $\Delta = 2\overline{\Delta}_n/3$ and $H(t) = 4t/3\, I_{\{t \le T_1\}} + (4T/9 + 2t/3) I_{\{t > T_1\}}$ with $T_1 = 2T/3$. This function H is not differentiable in T_1, however, we have that $H'(t_i) - \frac{\Delta_i}{\overline{\Delta}_n} = 0$ for all other points. It follows that our Assumption is satisfied. On the contrary, if we only consider the time points t where H is differentiable, Mykland and Zhang assumption is not fulfilled, since

$$\sup_{t > T_1} \left| \frac{H_n(t) - H_n(t - \sqrt{\overline{\Delta}_n})}{\sqrt{\overline{\Delta}_n}} - H'(t) \right| \to +\infty.$$

Denote the (log-price) return by $\Delta X_i = X_{t_i} - X_{t_{i-1}}$. Our proposed estimator takes the form of a discrete convolution

$$\widehat{\sigma}_{n,f}^2(\overline{t}) = \sum_{i=1}^n f_n(t_{i-1} - \overline{t}) (\Delta X_i)^2, \tag{2.5}$$

where $f_n(\cdot)$ is a given sequence of real functions belonging to the class specified below.

Definition 1. *A sequence $F \doteq \{f_n, n \in \mathbb{N}\}$ of functions $f_n : D \to \mathbb{R}$, with $D \subseteq \mathbb{R}$ being a given set and $0 \in D$ is said to be a delta sequence if, for all processes σ_t satisfying Assumption 1, as $n \to \infty$ (here and throughout all the paper it is intended that the integrals are defined over the intersection with $s \in D$),*

$$\int_0^T f_n(s - \overline{t})\sigma_s^2 ds = (\sigma^2)_{\overline{t}}^{\star} + R_n^{(\sigma^2)}(\overline{t}), \tag{2.6}$$

$$\frac{1}{f_n(0)} \int_0^T f_n^2(s - \overline{t})\sigma_s^2 ds = c_f(\sigma^2)_{\overline{t}}^{\star} + o_p(1), \tag{2.7}$$

$$\frac{1}{f_n^2(0)} \int_0^T f_n^4(s - \overline{t})\sigma_s^2 ds = O_p(f_n(0)), \tag{2.8}$$

where $R_n^{(\sigma^2)}(\overline{t}) = o_p(1)$ and

$$(\sigma^2)_{\overline{t}}^{\star} = \left(\psi_f^+ \sigma_t^2 + \psi_f^- \sigma_{t-}^2 \right) I_{\{t \in]0,T[\}} + \psi_f^- \sigma_{T-}^2 I_{\{t=T\}} + \psi_f^+ \sigma_0^2 I_{\{t=0\}}$$

where $\int_{x<0} f_n(x)dx \to \psi_f^-$ and $\psi_f^+ = 1 - \psi_f^-$ ($\psi_f^- = \psi_f^+ = \frac{1}{2}$ for symmetric delta sequences).

Note that if σ^2 is continuous in $\bar{t} \in]0, T[$, then $(\sigma^2)_{\bar{t}}^{\star} = \sigma_{\bar{t}}^2$. If we instead estimate at boundaries ($\bar{t} = 0$ or $\bar{t} = T$) we have to weight for the exact mass of the delta sequence respectively at the right and at the left of $\bar{t}$.

Condition (2.6) resembles the typical definition of delta sequence in analysis. The technical conditions (2.7) and (2.8) are required to guarantee the existence of a central limit theorem. Delta sequences have been introduced in statistics to estimate the density of a random variable, see e.g. Watson and Leadbetter [50]. The main result of this section is stated under a set of additional conditions that we collect in the following Assumption.

Assumption 3. *We assume that* $F = \{f_n, n \in \mathbb{N}\}$ *is a delta sequence with* $f_n(0) \to +\infty$ *and* $\int_D f_n(x)dx \to 1$, *and that further the functions* f_n *satisfy:*

    i) $\sup_{x \in D} |f_n(x)| \leq C f_n(0)$ *for a suitable constant* C

    ii) f_n *is Lipschitz in a neighborhood of* 0 *with Lipschitz constant* L_n *such that* $L_n \sqrt{\overline{\Delta}_n / f_n(0)} \to 0$; *further, either* $f_n \geq 0$ *or* $\overline{\Delta}_n^{\Gamma/2} \sum_i |f_n(t_{i-1} - \bar{t}) \Delta_i| \to 0$.

    iii) there exists a constant $M_\varepsilon > 0$ *not depending on* n *for which*

$$\sup_{x \in B_\varepsilon^c(0)} |f_n(x)| \leq M_\varepsilon. \tag{2.9}$$

Theorem 2.2 below can be shown to hold also when the condition $L_n \sqrt{\overline{\Delta}_n / f_n(0)} \to 0$ in Assumption 3 *ii)* is replaced with the condition, less stringent but less direct to be verified, $\sum_{i=1}^n \int_{t_{i-1}}^{t_i} |f_n(s - \bar{t}) - f_n(t_{i-1} - \bar{t})| ds \to 0$ for the consistency part, and $\sum_{i=1}^n \int_{t_{i-1}}^{t_i} |f_n(s - \bar{t}) - f_n(t_{i-1} - \bar{t})| ds / \sqrt{\overline{\Delta}_n f_n(0)} \to 0$ for the CLT part. Assumption 3 is not straightforward to verify for a given sequence f_n. For this reason, we specify the following proposition, which involves a set of sufficient conditions using only the features of f_n instead of the features of the process σ_t.

Proposition 2.1. *Consider a sequence of nonnegative functions* $f_n : D \to \mathbb{R}$, *with* $D \subset \mathbb{R}$ *and* $0 \in D$, *such that, as* $n \to \infty$, *conditions* $i) - iii)$ *in Assumption 3 are fulfilled, and further:*

    iv)

$$\int_D f_n(x)dx \to 1 \tag{2.10}$$

    v) there exists a sequence $\varepsilon_n \to 0$ *such that*

$$\int_{-\varepsilon_n}^{\varepsilon_n} f_n(x)dx \to 1 \tag{2.11}$$

    vi)

$$\int_D \frac{f_n^2(x)}{f_n(0)} dx \to c_f \tag{2.12}$$

    where c_f *is a real constant*

Then f_n is a delta sequence.

In condition *iii*) of Assumption 3 we chose to normalize $f_n^2(x)$ by $f_n(0)$, but alternatively $f_n(0)$ can be replaced by any sequence a_n able to deliver similar results, such as $a_n = \int f_n^2(x)dx$.

Some relevant examples of sequences $f_n(x)$ satisfying Assumption 3 are listed below. Other examples can be derived from Walter and Blum [49].

EXAMPLE 1: Kernels

Kernel estimators, used by [32] to estimate spot volatility, can indeed be used to generate a class of delta sequences. Consider a function $K : \mathbb{R} \to \mathbb{R}$ and a positive sequence $h_n \to 0$, and define:

$$f_n(x) = \frac{1}{h_n} K \left(\frac{x}{h_n} \right). \tag{2.13}$$

The sequence h_n is typically called *bandwidth*, and since $f_n(0) = \frac{1}{h_n} K(0)$, we can interpret $f_n(0)$ as the inverse of the bandwidth. In the case in which we write the delta sequence as (2.13), Assumption 3 can be reformulated as follows:

Assumption 3′ (for kernels):

1. $\int_{-\infty}^{+\infty} K(x)dx = 1$ and $\int_{-\infty}^{\infty} K^2(x)dx = c_2$ ($c_f = \frac{c_2}{K(0)}$)

2. $\sup_{x \in \mathbb{R}} |K(x)| \leq CK(0)$

3. $K(x)$ is almost everywhere differentiable and K' is bounded.

4. h_n is such that $\sup_{x \in \mathbb{R}} |K'\left(xh_n^{-1}\right)| \sqrt{\frac{\overline{\Delta}_n}{h_n^3}} \to 0$.

5. $\sup_{x \in B_\varepsilon^c(\bar{t})} \left| \frac{1}{h_n} K\left(\frac{x}{h_n} \right) \right| \leq M_\varepsilon$.

For example, the *Gaussian kernel*:

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

has $c_2 = \frac{1}{2\sqrt{\pi}}$ and $c_f = \frac{1}{\sqrt{2}}$, and Assumption 3′ is readily verified, while the *Epanechnikov kernel*

$$K(x) = \frac{3}{4}(1 - x^2)I_{\{|x| \leq 1\}}$$

has $c_2 = \frac{3}{5}$ and $c_f = \frac{4}{5}$ and also verifies 3′. The *indicator kernel*:

$$K(x) = \frac{1}{2}I_{\{|x| \leq 1\}}$$

also verifies Assumption 3′ and has $c_2 = \frac{1}{2}$ and $c_f = 1$.

EXAMPLE 2: Trigonometric functions

Trigonometric functions used in Fourier analysis are traditional approximants of the Dirac delta, and naturally appear in the construction of the Fourier estimator of Malliavin and Mancino [33]. The first example is the Dirichlet sequence given by $g_n(x) = \frac{1}{2\pi} D_{N_n}(x)$, with domain $[-\pi, \pi]$, where

$$D_N(x) := \sum_{|h| \leq N} e^{ihx} = \frac{\sin\left[\left(N + \frac{1}{2}\right)x\right]}{\sin \frac{x}{2}}, \tag{2.14}$$

and N_n is a diverging sequence. The Dirichlet sequence can be negative at some points. A positive trigonometric example, which will become crucial in Section 4, is given by $f_n(x) = \frac{1}{2\pi} F_{N_n}(x)$ with domain $[-\pi, \pi]$, where $F_N(x)$ is the Fejér sequence

$$F_N(x) := \sum_{|s| \leq N} \left(1 - \frac{|s|}{N+1}\right) e^{isx} = \frac{1}{N+1} \left(\frac{\sin \frac{N+1}{2} x}{\sin \frac{x}{2}}\right)^2, \tag{2.15}$$

and N_n is another diverging sequence. The following properties hold $\forall N$:

$$(i) \quad \frac{1}{2\pi} \int_{-\pi}^{\pi} F_N(x) dx = 1 \qquad\qquad (ii) \quad \frac{1}{2\pi} \int_{-\pi}^{\pi} D_N(x) dx = 1$$

$$(iii) \quad D_N^2(x) = (2N+1) F_{2N}(x) \qquad (iv) \quad \frac{1}{F_N(0)} \int_{-\pi}^{\pi} F_N^2(x) dx = \frac{4\pi}{3},$$

proving that f_n and g_n integrate to 1 and that $c_f = \frac{2}{3}$ and $c_g = 1$. Now, notice that $1/|\sin(x/2)| \leq 1/\sin(\varepsilon/2)$ if $\varepsilon \leq |x| \leq \pi$ with $0 < \varepsilon < \pi$. This easily proves conditions $iv)$ and $vi)$ in Proposition 2.1 for f_n. Moreover, with $\varepsilon = \varepsilon_n \to 0$ we have

$$\int_{\varepsilon_n \leq |x| \leq \pi} F_N(x) dx \leq \frac{1}{N+1} \frac{1}{\sin^2(\varepsilon_n/2)} 2(\pi - \varepsilon_n)$$

which converges to zero if $\epsilon_n^2 N \to \infty$. This proves, together with the remaining trivial conditions in Proposition 2.1, that F_N is a delta sequence.

The following theorem derives the asymptotic distribution of the proposed volatility estimator (2.5). We will use $\mathbf{MN}(0, V)$ to denote a mixed normal distribution with stochastic variance V.

Theorem 2.2. *Let Assumptions 1, 2, 3 hold. If $n, f_n(0) \to \infty$ in such a way that $f_n(0)\overline{\Delta}_n \to 0$, then for any $\bar{t} \in [0, T]$ we have $\widehat{\sigma}_{n,f}^2(\bar{t}) \xrightarrow{P} (\sigma^2)_{\bar{t}}^\star$. If furthermore, $R_n^{(\sigma^2)}(\bar{t}) = o_p\left(\sqrt{f_n(0)\overline{\Delta}_n}\right)$, then*

$$\frac{1}{\sqrt{f_n(0)\overline{\Delta}_n}} \left[\widehat{\sigma}_{n,f}^2(\bar{t}) - (\sigma^2)_{\bar{t}}^\star\right] \xrightarrow{\mathcal{L}-(s)} \mathbf{MN}\left(0, 2c_f H'(\bar{t})(\sigma^4)_{\bar{t}}^\star\right),$$

where the above convergence is stable in law.

A similar result is obtained in Kristensen [32] when $f_n(x)$ is of the form (2.13). On the notion of stable convergence, see Jacod [25].

Remark 1. *(On the validity of a CLT)*
The crucial condition for the validity of a CLT is $R_n^{(\sigma^2)}(\bar{t}) = o_p\left(\sqrt{f_n(0)\overline{\Delta}_n}\right)$. This condition is, however, typically satisfied with suitable choices of the sequence $f_n(0)$ (or, in the kernel case, of the

bandwidth). In the Appendix, we explicitely prove that, for the Gaussian, Epanechnikov and indicator kernel, this condition is fulfilled when $nh_n^{\Gamma+1} \to 0$, and for other kernels the condition can be verified in a similar way. The Fejér sequence is explicitly treated in Proposition 4.1.

Remark 2. *(Small sample correction)*

In small samples, it is advisable to use the estimator

$$\widehat{\widehat{\sigma}}_{n,f}^2(\overline{t}) = \frac{\sum_{i=1}^n f_n(t_{i-1} - \overline{t})(\Delta X_i)^2}{\sum_{i=1}^n f_n(t_{i-1} - \overline{t})\Delta_i}, \tag{2.16}$$

from which it is immediate to derive the same asymptotic results as in Theorem 2.2 given that

$$\sum_{i=1}^n f_n(t_{i-1} - \overline{t})\Delta_i \longrightarrow 1, \quad as \ n \to \infty.$$

Remark 3. *(Choice of the optimal f_n)*

The choice of the optimal sequence f_n relies on usual bias-variance trade-off considerations (see, for example, Fan and Yao [18], or the discussion in Kristensen [32]). From the proof of theorem 2.2, we can see that the bias depends both on the choice of the kernel and of the regularity of σ. For example, in the case in which f_n is the indicator function, we get (see the proof of Remark 1) that the bias is $O\left(f_n(0)^{-\Gamma/2}\right)$ and, given that the variance is $O(f_n(0)\overline{\Delta}_n)$, we get that the optimal choice of $f_n(0)$ is proportional to $\overline{\Delta}_n^{-\frac{1}{1+\Gamma}}$, and the speed of convergence of the spot volatility estimator is $n^{1/4}$. About the choice of the optimal delta sequence, there are almost no results for $\Gamma < 2$, with the exception of [16], who suggests the usage of a double exponential kernel, that is $f_n(x) = f_n(0)e^{-\frac{2|x|}{f_n(0)}}$. To get some insight on this problem, we simulate the model:

$$dX_t = \mu + \sigma_t dW_t^{(1)}$$
$$d\log\sigma_t^2 = \eta dW_t^{(2)} + dJ_t \tag{2.17}$$

where $corr(dW^{(1)}, dW^{(2)}) = \rho$, the jump occurs, exactly, in $\overline{t} = 0.5$, with size normally distributed with mean 1.44 and standard deviation 0.11. We set $\eta = 0.6$, $\rho = -0.25$ and $\mu = 0.06/252$ (these parameters are based on estimates in [8]). We also set $n = 1,000$ and $f_n(0) = 10$ and consider six delta sequences: indicator, triangular, Gaussian, Epanechnikov, Fejér and double exponential. We adopt the correction (2.16). The (relative) root mean square error in estimating $(\sigma^2)^\star_{\overline{t}=0.5}$ is shown in Figure 2.1, showing that there is no substantial difference between the five delta sequences, even if the double exponential kernel seems to present some advantages.

3 Estimation in the presence of microstructure noise/jumps

This section shows that, with proper adjustments, the estimator $\widehat{\sigma}_{n,f}^2(\overline{t})$ can be employed to the analysis of a more general data generating process where prices are affected by microstructure noise or can display a finite number of jumps, two important aspects that play a relevant role in the study of financial time series.
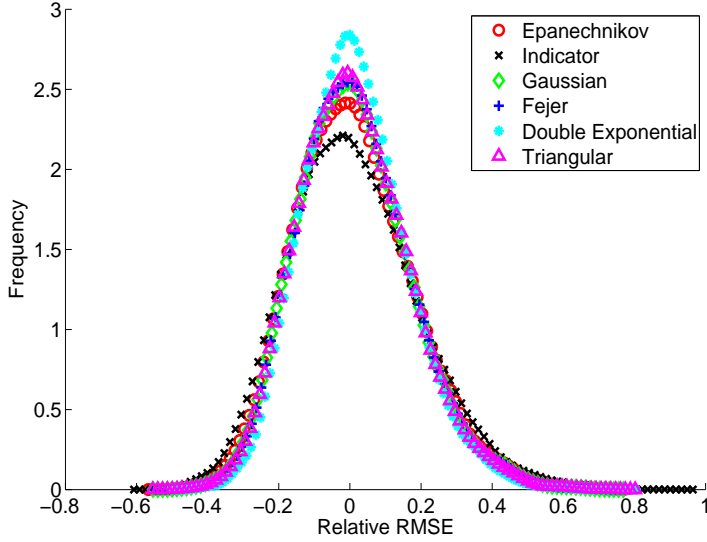
8

Figure 2.1: Relative RMSE distribution for the estimation of $(\sigma^2)^{\star}_{\bar{t}}$ on $1,000$ replications of model (2.17), for the five different delta sequences listed in the legend.

3.1 Robustness to microstructure effects

The following results emphasize the suitability of our theoretical framework to deal with microstructure noise effects in the observed data. For semplicity, consider logarithmic asset prices X_{t_i} which are observed at equispaced discrete times $t_0, \dots, t_n$ and are subject to an observation error due to microstructure noise.

Assumption 4. *Assume that observations are equally spaced ($\Delta_i = \overline{\Delta}_n$). Let*

$$X_{t_i} = Y_{t_i} + \varepsilon_i, \tag{3.1}$$

where $Y(t_i)$ is the unobservable efficient price satisfying Assumption 1, and ε_i denotes the noise component. The noise process $\{\varepsilon_i\}_{0 \le i \le n}$ is i.i.d. and independent of Y with $\mathrm{E}[\epsilon_i] = 0$ and $\mathrm{E}[\epsilon_i^8] < +\infty$.

In what follows, we denote by $V_\varepsilon = \mathrm{E}[\epsilon_i^2]$ and $\kappa_\varepsilon = \mathrm{E}[\epsilon_i^4]$.

Lemma 3.1. *Let Assumptions 3 and 4 hold. If $R_n^{(\sigma^2)}(\bar{t}) = o_p\left(\sqrt{\overline{\Delta}_n f_n(0)}\right)$ and $f_n(0)\overline{\Delta}_n \to 0$ as $n \to \infty$, then*

$$\frac{1}{\sqrt{f_n(0)\overline{\Delta}_n}}\left(\frac{1}{2}\overline{\Delta}_n \widehat{\sigma}^2_{n,f}(\bar{t}) - V_\varepsilon\right) \longrightarrow \mathbf{N}\left(0, \frac{1}{2}c_f\left(\kappa_\varepsilon + V_\varepsilon^2\right)\right), \tag{3.2}$$

where the above convergence is in distribution.

It is immediate to see that the market microstructure-induced bias is given by

$$\mathrm{E}[\widehat{\sigma}^2_{n,f}(\bar{t}) - \sigma^2(\bar{t})] = \frac{2V_\varepsilon}{\overline{\Delta}_n} + o\left(\frac{1}{\overline{\Delta}_n}\right), \tag{3.3}$$

9

which diverges at rate n. However, when appropriately corrected by a factor $\frac{1}{2}\overline{\Delta}_n$, a consistent estimate of the noise variance can be obtained and this is of the form

$$\widehat{V}_\varepsilon = \frac{1}{2}\overline{\Delta}_n\widehat{\sigma}^2_{n,f}(\overline{t}).$$

To obtain a consistent and asymptotically normally distributed estimator of the spot variance, we follow the two-scale approach in Zhang et al [52] and propose an estimator with overlapping prices at the lower frequencies. The idea is to remove the market microstructure noise by subtracting volatility estimated at two different frequencies, leaving the latent volatility unaffected. The approach we are proposing here is not efficient. Efficient estimation could be achieved, for example, using multiscales [51], by smoothing the observed time series via pre-averaging as in Jacod et al [27], or by using autocovariances and a flat-top kernel as in Barndorff-Nielsen et al [10]. Define an integer $\overline{n} < n$ and set

$$\widehat{\sigma}^{2,TS}_{n,\overline{n}}(\overline{t}) = \frac{1}{n}\sum_{i=1}^{n-\overline{n}+1} f_n(t_{i-1} - \overline{t})\left[(X_{t_{i+\overline{n}-1}} - X_{t_{i-1}})^2 - (X_{t_i} - X_{t_{i-1}})^2\right]. \tag{3.4}$$

The following Theorem shows that $\widehat{\sigma}^{2,TS}_{n,\overline{n}}(\overline{t})$ is a consistent and normally distributed estimator in the presence of microstructure noise.

Theorem 3.2. *Let Assumptions 3 and 4 hold. If $n, f_n(0), \overline{n} \to \infty$ in such a way that $\overline{n}f_n(0)\overline{\Delta}_n \to 0$, $R_n^{(\sigma^2)}(\overline{t}) = o_p(1)$ and $L_n\sqrt{\overline{\Delta}_n\overline{n}/f_n(0)} \to 0$, we have $\widehat{\sigma}^{2,TS}_{n,\overline{n}}(\overline{t}) \xrightarrow{p} (\sigma^2)^\star_{\overline{t}}$. Furthermore, if $R_n^{(\sigma^2)}(\overline{t}) = o_p\left(\sqrt{f_n(0)\overline{\Delta}_n\overline{n}}\right)$ and $\overline{n} = c(\overline{\Delta}_n)^{-\frac{2}{3}}$ with $c \in \mathbb{R}$, then*

$$\frac{1}{\sqrt{f_n(0)(\overline{\Delta}_n)^{\frac{1}{3}}}}\left[\widehat{\sigma}^{2,TS}_{n,\overline{n}}(\overline{t}) - (\sigma^2)^\star_{\overline{t}}\right] \xrightarrow{\mathcal{L}-(s)} \mathbf{MN}\left(0, 2c_f\left[V_\varepsilon^2 + c(\sigma^4)^\star_{\overline{t}}\right]\right), \tag{3.5}$$

where the above convergence is stable in law.

Notice that the speed of convergence of the estimator in (3.5) is in line with that obtained in [52] in the case of integrated volatility estimation. Indeed, when estimating integrated volatility in the presence of noise, you get a speed of convergence of $n^{1/6}$ instead of the $n^{1/2}$ that you would get in the absence of noise. With spot volatility, we get $n^{1/6}f_n(0)^{1/2}$ instead of $n^{1/2}f_n(0)^{1/2}$.

3.2 Robustness to jumps

We now consider the case where a finite number of Poisson jumps is added to the stochastic integral driving the state variable dynamics.

Assumption 5. *The adapted process X_t defined on $[0,T]$ satisfies*

$$X_t = Y_t + J_t \tag{3.6}$$

with $dJ_t = c_J(t)dN_t$, where Y_t fulfills Assumption 1, J_t is a doubly stochastic Poisson process, and N_t is a non-explosive Poisson counting process whose intensity is an adapted stochastic process λ_t. The

size of the jumps occurring at times $\tau_1, \ldots, \tau_{N(t)}$ *is given by i.i.d. random variables* $c_J(\tau_j)$ *such that* $\mathcal{P}(\{c_J(\tau_j) = 0\}) = 0 \; \forall t \in [0, T]$.

Following the approach in Mancini [35], we define our estimator to be

$$\widehat{\phi}^2_{n,f}(\bar{t}) = \sum_{i=1}^{n} f_n(t_{i-1} - \bar{t}) \left(\Delta X_i \right)^2 I_{\{(\Delta X_i)^2 \leq \vartheta_n\}}, \tag{3.7}$$

where $I_{\{\cdot\}}$ denotes the indicator function and ϑ_n is a suitable sequence. The aim of the threshold ϑ_n is to disentangle the discontinuous variation induced by the Poisson jumps from the continuous variation induced by the Brownian motion. Asymptotically, this happens when ϑ_n converges to zero slower than the modulus of continuity of the Brownian motion, as specified in the next Theorem. Note that ϑ_n can also be either a function of time or a stochastic process (see Mancini and Renò [36]). Alternative options to (3.7) are the flat kernel estimator in Aït-Sahalia and Jacod [1], or the locally averaged bipower variation proposed by Veraart [48]. Both approaches admit infinite jump activity in the data.

Theorem 3.3. *Let Assumptions 2, 3 and 5 hold. If* $n, f_n(0) \to \infty$ *and* $\vartheta_n \to 0$ *in such a way that* $f_n(0)\overline{\Delta}_n \to 0$, $\vartheta_n / \left(\overline{\Delta}_n \log \frac{1}{\overline{\Delta}_n} \right) \longrightarrow \infty$ *and* $R_n^{(\sigma^2)}(\bar{t}) = o_p(1)$ *we have* $\widehat{\phi}^2_{n,f}(\bar{t}) \xrightarrow{p} \sigma^2(\bar{t})$. *Furthermore, if* $R_n^{(\sigma^2)}(\bar{t}) = o_p\left(\sqrt{f_n(0)\overline{\Delta}_n} \right)$, *then*

$$\frac{1}{\sqrt{f_n(0)\overline{\Delta}_n}} \left[\widehat{\phi}^2_{n,f}(\bar{t}) - (\sigma^2)^{\star}_{\bar{t}} \right] \xrightarrow{\mathcal{L}-(s)} \mathbf{MN}\left(0, 2c_f H'(\bar{t})(\sigma^4)^{\star}_{\bar{t}} \right),$$

where the above convergence is stable in law.

4 Relation to the Fourier estimator

In this section, we analyze the Fourier estimator first introduced in Malliavin and Mancino [33]. In particular, we show that the Fourier estimator can be written as the sum of a delta sequence estimator, when the function $f_n(\cdot)$ is set to be equal to the Fejér sequence (see Example 2), and a zero-mean noisy term.

In the Fourier method, the classical harmonic analysis is combined with stochastic calculus to connect the Fourier transform of the log-price process X_t to the Fourier transform of the volatility function σ^2_t. Specifically, the spot volatility estimator is defined to be

$$\widehat{\sigma}^{2,F}_{n,n',N}(\bar{t}) = \sum_{|k| \leq N} \left(1 - \frac{|k|}{N} \right) \mathcal{H}_{n,n'}(k) e^{ik\tau}, \tag{4.1}$$

where

$$\mathcal{H}_{n,n'}(k) := \frac{T}{2n'+1} \sum_{|s| \leq n'} \mathcal{F}_n(dX)(s) \mathcal{F}_n(dX)(k-s), \tag{4.2}$$

and

$$\mathcal{F}_n(dX)(s) := \frac{1}{T} \sum_{j=1}^{n} e^{-is\tau_{j-1}} \Delta X_j \tag{4.3}$$

is the discrete Fourier transform of dX_t. Here $\tau = 2\pi\bar{t}/T$ and $\tau_i = 2\pi t_i/T$ are rescaled times. Malliavin and Mancino [34] prove that, when $N = n'$,

$$\widehat{\sigma}^{2,F}_{n,n'=N,N}(\bar{t}) \xrightarrow{p} \sigma^2_{\bar{t}},$$

when $n, N \to \infty$. They also provide a weak convergence result for a Lebesgue average of $\widehat{\sigma}^{2,F}_{n,n'=N,N}(\bar{t}) - \sigma^2_t$ on $[0, T]$, but do not provide a central limit theorem for the estimation error of the spot variance (see also Clement and Gloter, [13] for a discussion and a generalization of their results).

In order to apply the Fourier estimator, it is necessary to set the number of coefficients of the price process n' used in the computation of the volatility coefficients, and the number of volatility coefficients N used in the reconstruction of the volatility trajectory. Both n' and N are sequences depending on n. Importantly, here we do not restrict to the choice $n' = N$, suggesting that an higher n' is beneficial. A reference value for equally spaced data is $n' = n/2$, also known as Nyquist frequency (see Priestley [45]).

In what follows, we show that the Fourier estimator does not belong directly to our class but it can be rearranged into the sum of two terms: the volatility estimator $\widehat{\sigma}^2_{n,f}(\bar{t})$, where $f_n(\cdot)$ is a rescaled Fejér sequence, and a cross-product term with zero mean.

Proposition 4.1. *Define $\widehat{\sigma}^2_{n,f}(\bar{t})$ as in (2.5) with $f_n(x) = \frac{1}{2\pi}F_N(x)$, where F_N is defined in Eq. (2.15), with $x \in]-\pi, \pi[$. Assume that Assumptions 1,2 and 3 hold, with the exception of property ii) in Assumption 3. Assume now that, as $n, N \to \infty$, $N/n \to 0$ in such a way that there exists a sequence $\varepsilon_n \to 0$ such that $N^3\varepsilon_n^4/n \to \infty$ and $N/(n\varepsilon_n^\Gamma) \to \infty$. Then, if $\bar{t} \in]0, T[$,*

$$\frac{1}{\sqrt{f_n(0)\overline{\Delta}_n}} \left[\widehat{\sigma}^2_{n,f}(\bar{t}) - (\sigma^2)^\star_{\bar{t}}\right] \xrightarrow{\mathcal{L}-(s)} \mathbf{MN}\left(0, \frac{4}{3}(H'\sigma^4)^\star_{\bar{t}}\right).$$

Proposition 4.2. *The Fourier estimator given in (4.1) is such that*

$$\widehat{\sigma}^{2,F}_{n,n',N}(\bar{t}) = \widehat{\sigma}^2_{n,f}(\bar{t}) + \phi_{n,f,g}(\bar{t})$$

where

$$\widehat{\sigma}^2_{n,f}(\bar{t}) = \sum_{i=1}^{n} f_n(t_{i-1} - \bar{t})(\Delta X_i)^2$$

and

$$\phi_{n,f,g}(\bar{t}) = \frac{1}{g_n(0)}\sum_{i=1}^{n}\sum_{i\neq j=1}^{n} f_n(t_{j-1} - \bar{t})g_n(t_{j-1} - t_{i-1})\Delta X_i \Delta X_j.$$

with $f_n(x) = \frac{1}{T}F_{N-1}(2\pi x/T)$ and $g_n(x) = \frac{1}{T}D_{n'}(2\pi x/T)$. It holds that $E[\phi_{n,f,g}(\bar{t})] = 0$ and, if σ is independent from W_t, the covariance between $\widehat{\sigma}^2_{n,f}(\bar{t})$ and $\phi_{n,f,g}(\bar{t})$ is zero.

Proposition 4.2 shows that, under our set of Assumptions, the Fourier estimator can be redefined using delta sequences with an improvement in terms of the variance. The cross-term typically adds noise and computational burden. Kanatani [31] made a similar remark in the case of the integrated volatility estimation. Note that, in general, the cross-terms might be beneficial to the reduction of the
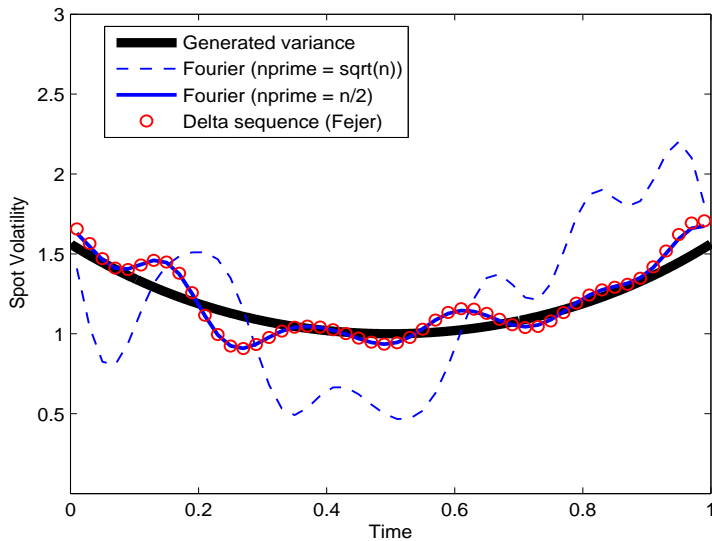
12

Figure 4.1: Fourier estimates of the variance of a single simulated path (the generated variance is the thick solid line) with $n = 2500$ and $N = 8$, in the case $n' = n/2$ (thin solid line) and $n' = \sqrt{n}$ (dashed line). It is clear that a lower n' leads to an higher variance. We also report the estimator (2.5) with the Fejér delta sequence: it is almost identical to the Fourier estimator with $n' = n/2$ but computationally faster.

mean square error in the presence of market microstructure noise, see Mancino and Sanfelici [37, 38], and Barndorff-Nielsen et al [10].

The above findings are clearly illustrated in Figure 4.1 where it is apparent as the trajectory estimated with the Fourier method without cross products can either have a larger noise (dashed line), or perfectly overlap (circles) to the delta sequence estimator in the case in which we choose $n' = \frac{n}{2}$. Further simulation evidence suggests that, in the unequally spaced case, the optimal choice of n' is $\frac{n}{2H'(\bar{t})}$.

5 Empirical application

In this final Section, we apply the proposed estimators (3.4) and (3.7) to a set of market data consisting of high-frequency transactions of the S&P 500 stock index futures. We restrict our attention to year 1999 and to contracts closer to maturity. Transactions are recorded over 251 trading days between 8.30 a.m. to 3.15 p.m. and interpolated to a 5−seconds grid. Every day, we then have a total of 4,860 price returns. For both estimators we use the Epanechnikov kernel with $h = 15$ minutes.

To calculate the low frequency estimator $\overline{\sigma}^2_{n,\bar{n}}(\bar{t})$ on the right-hand side of (3.4), we apply a sub-sampling technique similar to that described in Zhang et al [52] with $\overline{n} = 12$, which corresponds to one-minute returns. In order to avoid the effect of jump dynamics in the observed data, we first remove from the sample all the days characterized by significant price changes using the procedure described
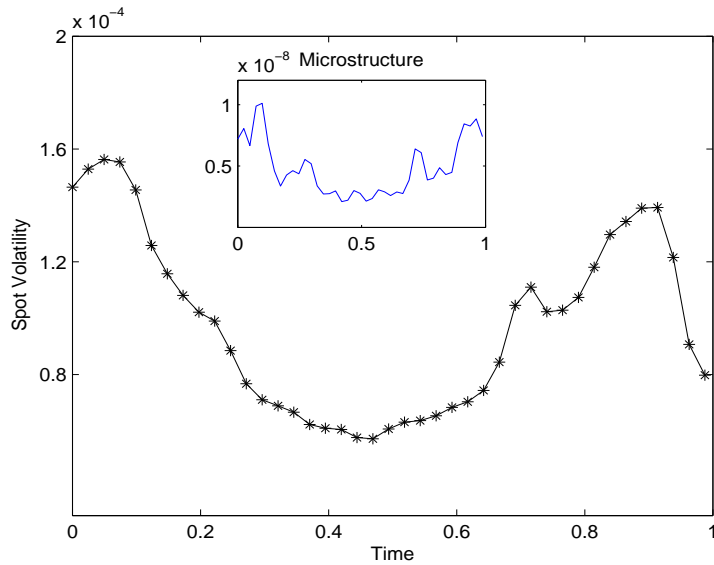
13

Figure 5.1: Intraday spot volatility for the S&P500 stock index futures over one year of data calculated using the two scale estimator (3.4). Days with relevant jump activity are previously removed from the sample. The inset shows the average estimate of the microstructure noise variance V_ϵ.

below. Figure 5.1 plots the estimated intraday spot volatility averaged across days and calculated in daily time units. The well known $U-$shape is clearly detected, as it was already observed in previous studies, see, for instance, Andersen and Bollerslev [3]. The estimate of the microstructure variance V_ε is also provided.

We now turn to the jump-robust estimator, and we use 5-minute returns for computation of spot volatility estimators, to soften the impact of microstructure noise. To show that our threshold estimator $\widehat{\phi}_{n,f}^2(\bar{t})$ is robust to price jumps, we compare it with the original spot volatility estimator (2.5) using a data-set created by removing all days with relevant jump activity. The resulting intraday volatility curves then should be almost identical. To identify the jumps, we employ the $C-Tz$ statistics in Corsi et al [14], After setting the daily significance level of a jump to 99%, a total of 28 days are detected and then excluded from the sample. The top panel in Figure 5.2 shows that the volatility curves obtained with the two aforementioned estimators match almost everywhere, meaning that that $\widehat{\phi}_{n,f}^2(\bar{t})$ is not affected by large price movements and is able to provide robust estimates of the intraday volatility dynamics. We then apply the same estimation procedure to a sample made of the 28 days initially removed; the result in plotted in Figure 5.2, bottom panel. As expected, now the two curves behave quite differently, especially around the market opening time.
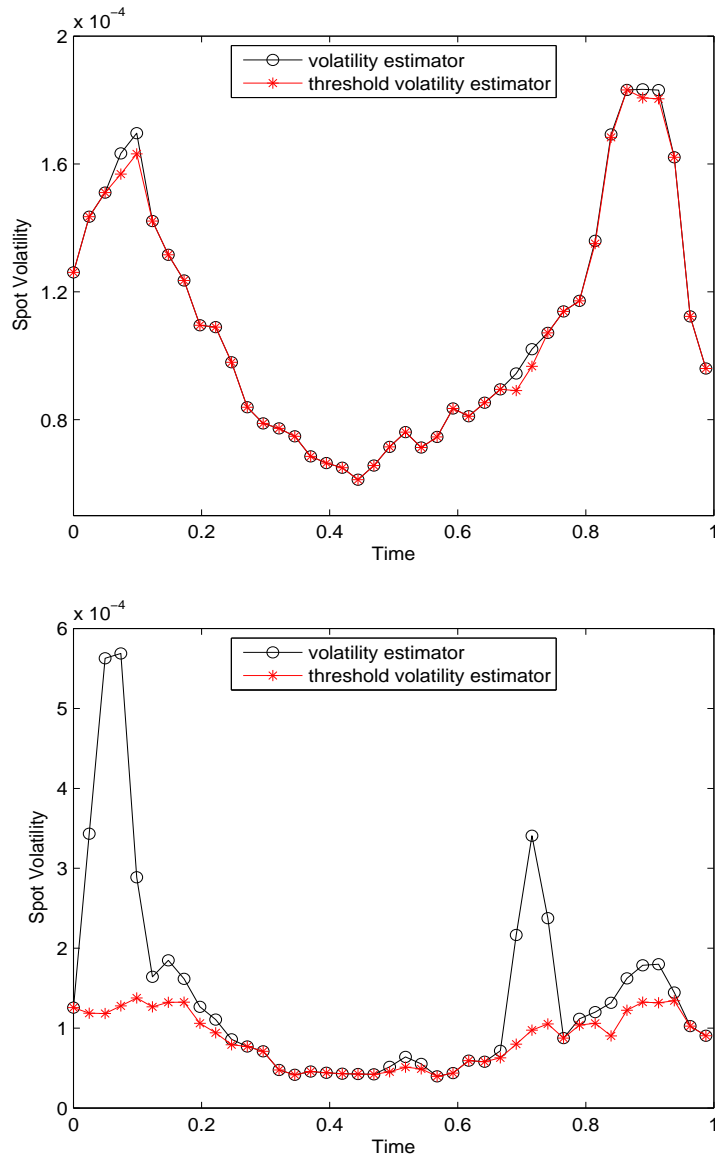
14

Figure 5.2: Intraday spot volatility for the S&P500 stock index futures averaged over one year of data calculated using the original volatility estimator (2.5) and the threshold estimator (3.7) respectively. Top panel: original data-set without relevant jump activity. Bottom panel: sample made of 28 days characterized by large price movements. The significance level of jump detection is set to 99%. The volatility is measured in daily units.

6 Conclusions

We enlarged the class of spot volatility estimators using localizing sequences of functions which converge to a Dirac delta. Under mild hypotheses on the data generating process, we provide an asymptotic theory for the estimators within the class and we propose suitable modifications to assess the effect of microstructure noise or price discontinuities. As a special case, we related the Fourier estimator with the delta sequence obtained with the Fejér sequence, showing that the latter is more efficient in

the case in which the price follows a Brownian semimartingale and there is neither leverage effect nor microstructure noise. We finally applied the resulting estimators to a data-set of high-frequency stock index futures and successfully recovered the traditional U-shaped intraday volatility pattern.

The paper leaves open the possibility of further developments. For example, we would like to study the joint contribution of microstructure noise and jumps, possibly using the techniques in Jacod et al [27] and Podolskij and Vetter [43, 44]. Also, the challenging problem of the optimal choice of the delta sequence is to be addressed. Finally, the asymptotic distribution of the spot volatility obtained with the Fourier estimator is unknown. We leave all these interesting issues for future research.

A   Proofs

In what follows, we will use $\int(...)dx$ to denote an integral over $\mathbb{R}$. C or K indicate a constant which does not depend on i, nor on the sequence $F = \{f_n, n \in \mathbb{N}\}$, but can depend on $\bar{t}$ and the localizing sequence τ_m, and which keeps the same name even when changing from line to line or from one side to another of the inequality. Without loss of generality, we assume $\sigma \geq 0$. Recall that for any Lebesgue-integrable function a, for all $\ell \geq 1$ we have, by Jensen inequality,

$$\frac{1}{\Delta_i} \int_{t_{i-1}}^{t_i} |a_s| ds \leq \left(\frac{1}{\Delta_i} \int_{t_{i-1}}^{t_i} |a_s|^\ell ds \right)^{\frac{1}{\ell}} \tag{A.1}$$

We remark that assuming μ, σ, H' càdlàg entails that they are locally bounded. By a localization procedure similar to that in [29] (section 5.4, p.549), we can assume without loss of generality that they are bounded (as (ω, t) vary within $\Omega \times [0, T]$).

In order to prove Theorem 2.2 we use Lemma A.1 below several times with A being equal to σ^k for some powers $k \in \{0, 1, 2, 3, 4\}$. Note that, by property (2.2) and the boundedness of σ, we have, for $k \geq 2$,

$$E_{u \wedge s}[|\sigma_u - \sigma_s|^k] = E_{u \wedge s}[|\sigma_u - \sigma_s|^2 |\sigma_u - \sigma_s|^{k-2}] \leq C|u - s|^\Gamma. \tag{A.2}$$

and, similarly, we also have, for $k \geq 2$,

$$E_{u \wedge s}[|\sigma_u^2 - \sigma_s^2|^k] \leq C|u - s|^\Gamma. \tag{A.3}$$

For $k = 1$ we instead have, by Jensen inequality, $E_{u \wedge s}[|\sigma_u - \sigma_s|] \leq C|u - s|^{\Gamma/2}$ and $E_{u \wedge s}[|\sigma_u^2 - \sigma_s^2|] \leq C|u - s|^{\Gamma/2}$.

Proposition 2.1. Assume first that $0 < \bar{t} < T$. Using the boundedness of σ^2 and the property (2.10), we can write:

$$R_n^{(\sigma^2)}(\bar{t}) = \int_0^T f_n(s - \bar{t})\sigma_s^2 ds - \left(\int_0^T f_n(s - \bar{t})ds + o(1) \right) (\sigma^2)_{\bar{t}}^\star$$

$$= \int_0^T f_n(s - \bar{t}) \left(\sigma_s^2 - (\sigma^2)_{\bar{t}}^\star \right) ds + o_p(1),$$

and, using (2.2) and property (2.11),

$$\mathrm{E}\left[\left| \int_0^T f_n(s - \bar{t}) \left(\sigma_s^2 - (\sigma^2)_{\bar{t}}^\star \right) ds \right| \right] \leq \int_0^T f_n(s - \bar{t}) \mathrm{E}\left[\left| \sigma_s^2 - (\sigma^2)_{\bar{t}}^\star \right| \right] ds$$

$$= C \int_{|s-\bar{t}| < \varepsilon_n} f_n(s - \bar{t}) \mathrm{E}\left[\left| \sigma_s^2 - (\sigma^2)_{\bar{t}}^\star \right| \right] ds$$

$$+ C \int_{|s-\bar{t}| \geq \varepsilon_n} f_n(s - \bar{t}) \mathrm{E}\left[\left| \sigma_s^2 - (\sigma^2)_{\bar{t}}^\star \right| \right] ds$$

$$\leq C\varepsilon_n^{\Gamma/2} + C \int_{|x| \geq \varepsilon_n} f_n(x)dx \to 0.$$

If instead $\bar{t} = T$, we repeat the same reasoning above using: $\psi_n^- = \int_{s-T<0} f_n(s - T)ds + o(1)$, and we proceed in a similar way if $\bar{t} = 0$. This proves (2.6). To prove (2.7) for f_n, it is thus enough to prove

17

that $g_n(x) = \frac{f_n^2(x)}{c_f f_n(0)}$ satisfies Eqs. (2.10), which is straightforward from property (2.12), and (2.11), which is obtained, using $\sup_x f_n(x) \leq C f_n(0)$ as:

$$0 \leq \int_{|x| \geq \varepsilon_n} g_n(x) dx = \frac{1}{c_f} \int_{|x| \geq \varepsilon_n} f_n(x) \frac{f_n(x)}{f_n(0)} dx \leq C \int_{|x| \geq \varepsilon_n} f_n(x) dx \to 0.$$

To prove (2.8), use the boundedness of σ^2 and (2.12) and write:

$$\int_0^T \frac{f_n^4(t - \bar{t})}{f_n^2(0)} \sigma_s^2 dt \leq C \int_0^T \frac{f_n^2(t - \bar{t})}{f_n(0)} f_n(0) dt \leq C f_n(0)$$

$\square$

Lemma A.1. *i) For a sequence of processes $A^{(n)}$ bounded by the same constant K, if f_n are Lipschitz functions and $\max_i \Delta_i = O(\overline{\Delta}_n)$ then*

$$\int_0^T f_n(s - \bar{t}) A_s^{(n)} ds - \sum_{i=1}^n f_n(t_{i-1} - \bar{t}) \int_{t_{i-1}}^{t_i} A_s^{(n)} ds = O_{a.s.}(L_n \overline{\Delta}_n). \tag{A.4}$$

As a corollary, under (2.6) and if $L_n \overline{\Delta}_n \to 0$ we have $\sum_{i=1}^n f_n(t_{i-1} - \bar{t}) \Delta_i \to 1$, as $n \to \infty$.

ii) Consider a bounded càdlàg process A. If either $f_n \geq 0$ or $\overline{\Delta}_n^{\Gamma/2} \int_0^T |f_n(s - \bar{t})| ds \to 0$ and (2.9) holds, if both f_n and $g_n = f_n^2/(c_f f_n(0))$ satisfy (2.6), under $\max_i \Delta_i = O(\overline{\Delta}_n)$, (2.4) and (2.2) we have

$$\frac{1}{\overline{\Delta}_n} \sum_{i=1}^n f_n(t_{i-1} - \bar{t}) A(t_{i-1}) \Delta_i^2 \xrightarrow[n \to \infty]{p} (H'A)_{\bar{t}}^{\star} \tag{A.5}$$

and

$$\frac{1}{\overline{\Delta}_n f_n(0)} \sum_{i=1}^n f_n^2(t_{i-1} - \bar{t}) A(t_{i-1}) \Delta_i^2 \xrightarrow[n \to \infty]{p} c_f (H'A)_{\bar{t}}^{\star}. \tag{A.6}$$

iii) Under (2.2) and the boundedness of σ, for any $p = 0, 1, 2, 3$ there exists $\alpha > 0$ such that for all $i = 1, \ldots, n$ and for a suitable constant C_p,

$$\mathrm{E}_{i-1} \left[|\sigma_{t_{i-1}} \Delta W_i|^p \left| \int_{t_{i-1}}^{t_i} \left(\sigma_s - \sigma_{t_{i-1}} \right) dW_s \right|^{4-p} \right] \leq$$

$$C_p \Delta_i^2 \sigma_{t_{i-1}}^p (\Delta_i^{\alpha} I_{\{t_{i-1} \in B_\varepsilon(\bar{t})\}} + I_{\{t_{i-1} \notin B_\varepsilon(\bar{t})\}}).$$

iv) Under the same assumptions as for ii) above, for any bounded càdlàg process M and $\alpha > 0$ we have

$$\sum_{i=1}^n \frac{f_n^2(t_{i-1} - \bar{t})}{f_n(0) \overline{\Delta}_n} M_{t_{i-1}} \Delta_i^2 (\Delta_i^{\alpha} I_{\{t_{i-1} \in B_\varepsilon(\bar{t})\}} + I_{\{t_{i-1} \notin B_\varepsilon(\bar{t})\}}) \xrightarrow{p} 0$$

v) Under (2.2) and the boundedness of σ, for any $p = 0, 1$ we have for all $i = 1, \ldots, n$ and for some $\alpha > 0$,

$$\mathrm{E}_{i-1} \left[|\sigma_{t_{i-1}} \Delta W_i|^p \left| \int_{t_{i-1}}^{t_i} \left(\sigma_s - \sigma_{t_{i-1}} \right) dW_s \right|^{2-p} \right] \leq C_p \Delta_i \sigma_{t_{i-1}}^p (\Delta_i^{\alpha} I_{\{t_{i-1} \in B_\varepsilon(\bar{t})\}} +$$

$$+ I_{\{t_{i-1} \notin B_\varepsilon(\bar{t})\}}).$$

vi) Under (2.2) and the boundedness of σ, for $p \in [1, 8]$,

$$\mathrm{E}_{i-1} \left[\left(\int_{t_{i-1}}^{t_i} |\sigma_s^2 - \sigma_{t_{i-1}}^2| ds \right)^p \right] \leq K \Delta_i^p (\Delta_i^{\Gamma/2} I_{\{t_{i-1} \in B_\varepsilon(\bar{t})\}} + I_{\{t_{i-1} \notin B_\varepsilon(\bar{t})\}}).$$

Proof. i) Noting that $\max_i \Delta_i = O(\overline{\Delta}_n)$ implies $\sum_i \Delta_i^2 \leq K\overline{\Delta}_n$, we have

$$\left| \int_0^T f_n(s - \bar{t})A_s ds - \sum_{i=1}^n f_n(t_{i-1} - \bar{t}) \int_{t_{i-1}}^{t_i} A_s ds \right|$$

$$\leq \sum_{i=1}^n \int_{t_{i-1}}^{t_i} |f_n(s - \bar{t}) - f_n(t_{i-1} - \bar{t})| A_s ds$$

$$\leq K \sum_i L_n \int_{t_{i-1}}^{t_i} |s - t_{i-1}| ds \leq K L_n \sum_i \Delta_i^2 \leq K L_n \overline{\Delta}_n.$$

As for the corollary, by (2.6) with $\sigma \equiv 1$ we have $\int_0^T f_n(s-\bar{t})ds \to 1$, as $n \to \infty$. Thus it is sufficient to show that $|\sum_{i=1}^n f_n(t_{i-1} - \bar{t})\Delta_i - \int_0^T f_n(s - \bar{t})ds| \to 0$, which is guaranteed by (A.4) with $A \equiv 1$.

ii) It is enough to show Eq. (A.5), since the assumptions on f_n imply that also $g_n = f_n^2/(c_f f_n(0))$ is Lipschitz with Lipschitz constant $G_n \leq K L_n$ and thus g_n is also satisfying $\overline{\Delta}_n^\Gamma G_n \to 0$ and (2.9). By applying (2.6) we obtain $(H'A)_{\bar{t}}^\star = \int_0^T f_n(s - \bar{t})(H'A)_s ds + o_p(1)$. We now show that this last term has the same asymptotic behavior as $\sum_{i=1}^n f_n(t_{i-1} - \bar{t})A_{t_{i-1}}\Delta_i^2/\overline{\Delta}_n$. In fact,

$$\int_0^T f_n(s - \bar{t})(H'A)_s ds - \sum_{i=1}^n f_n(t_{i-1} - \bar{t})A_{t_{i-1}}\Delta_i^2/\overline{\Delta}_n =$$

$$\sum_i \int_{t_{i-1}}^{t_i} [f_n(s - \bar{t}) - f_n(t_{i-1} - \bar{t})](H'A)_s ds+$$

$$\sum_i \int_{t_{i-1}}^{t_i} f_n(t_{i-1} - \bar{t})[(H'A)_s - A_{t_{i-1}}\Delta_i/\overline{\Delta}_n] ds.$$

Given the assumed boundedness of $H'A$, the first term in the rhs of the display above is a.s. bounded by $K L_n \sum_i \Delta_i^2 \leq K L_n \overline{\Delta}_n \to 0$. As for the second sum above, we can write it as

$$\sum_i \int_{t_{i-1}}^{t_i} f_n(t_{i-1} - \bar{t})H'_s(A_s - A_{t_{i-1}})ds+ \tag{A.7}$$

$$\sum_i \int_{t_{i-1}}^{t_i} f_n(t_{i-1} - \bar{t})(H'_s - \Delta_i/\overline{\Delta}_n)A_{t_{i-1}} ds.$$

Using (A.4) and (2.4), the last term has the same limit as

$$\sum_i \int_{t_{i-1}}^{t_i} f_n(s - \bar{t})(H'_s - \Delta_i/\overline{\Delta}_n)A_{t_{i-1}} ds,$$

which is bounded in absolute value by

$$K \sum_i |f_n(t_{i-1} - \bar{t})|\Delta_i^2 \leq K\overline{\Delta}_n \int_0^T |f_n(s - \bar{t})|ds,$$

which in turn converges to zero by condition $\overline{\Delta}_n f_n(0) \to 0$ and the Lipschitz property of f_n. Now we deal with the term in (A.7) by splitting it into the sum over the indexes i s.t. $t_{i-1} \in B_\varepsilon(\bar{t})$ and the sum of the other terms. Since the càdlàg process A has at most countably many jumps within $[0, T]$ and each jump time has Lebesgue measure 0, we have

$$\sum_{\{i:t_{i-1}\notin B_\varepsilon(\bar{t})\}} \int_{t_{i-1}}^{t_i} f_n(t_{i-1} - \bar{t})H'_s(A_s - A_{t_{i-1}})ds = \tag{A.8}$$

$$\sum_{\{i:t_{i-1}\notin B_\varepsilon(\bar t)\}} \int_{t_{i-1}}^{t_i} f_n(t_{i-1}-\bar t)H'_s(A_{s-}-A_{t_{i-1}})ds.$$

Using then (A.4) and the boundedness of H' and A we obtain that the last sum has the same asymptotic behavior as $\int_{B_\varepsilon^c(\bar t)} \psi_s^{(n)}ds$, where

$$\psi_s^{(n)} \doteq f_n(s-\bar t)H'_s\Big(A_{s-}-\sum_j A_{t_{j-1}}I_{]t_{j-1},t_j]}(s)\Big).$$

Note that, for Lebesgue-almost all s, $\psi_s^{(n)} \to 0$, because any fixed $s \in [0,T]$ belongs to only one interval, say $]t_{\bar j-1},t_{\bar j}]$ and $t_{\bar j-1}$ is always on the left hand side of s, so that $A_{t_{\bar j-1}} \to A_{s-}$, as $n \to \infty$, and thus $\sum_j A_{t_{j-1}}I_{]t_{j-1},t_j]}(s) \to A_{s-}$. Moreover (2.9) and the boundedness of A and H' imply that $|\psi^{(n)}(s)| \le KM_\varepsilon$ which belongs to $L^1(B_\varepsilon^c(\bar t))$. By the dominated convergence theorem we conclude that $\int_0^T \psi_s^{(n)}ds \to 0$ and (A.8) is asymptotically negligible. Finally we show that

$$\sum_{\{i:t_{i-1}\in B_\varepsilon(\bar t)\}} \int_{t_{i-1}}^{t_i} f_n(t_{i-1}-\bar t)H'_s(A_s - A_{t_{i-1}})ds \xrightarrow{P} 0.$$

In fact using (2.2) its $L^1(\Omega)$ norm is dominated by

$$E\Bigg[\sum_{\{i:t_{i-1}\in B_\varepsilon(\bar t)\}} E_{i-1}\Big[\int_{t_{i-1}}^{t_i} |f_n(t_{i-1}-\bar t)||H'_s||A_s - A_{t_{i-1}}|ds\Big]\Bigg] \le$$

$$K \sum_{\{i:t_{i-1}\in B_\varepsilon(\bar t)\}} \int_{t_{i-1}}^{t_i} |f_n(t_{i-1}-\bar t)|ds\, \Delta_i^{\Gamma/2} \le K\overline\Delta_n^{\Gamma/2} \int_0^T |f_n(s-\bar t)|ds \to 0.$$

iii) Using Hölder inequality, the considered conditional expectation is dominated by

$$E_{i-1}^{1/2}[|\sigma_{t_{i-1}}\Delta W_i|^{2p}]E_{i-1}^{1/2}\Bigg[\Big|\int_{t_{i-1}}^{t_i}\sigma_s - \sigma_{t_{i-1}}dW_s\Big|^{8-2p}\Bigg]$$

using then the Burkolder-Davis-Gundy inequality this is less than

$$C_p\sqrt{\sigma_{t_{i-1}}^{2p}\Delta_i^p}\sqrt{E_{i-1}\Bigg[\Big(\int_{t_{i-1}}^{t_i}(\sigma_s - \sigma_{t_{i-1}})^2 ds\Big)^{4-p}\Bigg]}. \tag{A.9}$$

Since $p \le 3$ then by (A.1) with $\ell = 4 - p$,

$$\frac{1}{\Delta_i}\int_{t_{i-1}}^{t_i}(\sigma_s - \sigma_{t_{i-1}})^2 ds \le \Big(\frac{1}{\Delta_i}\int_{t_{i-1}}^{t_i}(\sigma_s - \sigma_{t_{i-1}})^{2(4-p)}ds\Big)^{1/(4-p)},$$

we obtain that

$$E_{i-1}\Bigg[\Big(\int_{t_{i-1}}^{t_i}(\sigma_s - \sigma_{t_{i-1}})^2 ds\Big)^{4-p}\Bigg] \le E_{i-1}\Bigg[\Delta_i^{3-p}\int_{t_{i-1}}^{t_i}(\sigma_s - \sigma_{t_{i-1}})^{8-2p}ds\Bigg]$$

$$\le \Delta_i^{3-p}\int_{t_{i-1}}^{t_i} E_{i-1}[|\sigma_s - \sigma_{t_{i-1}}|^{8-2p}]ds.$$

Using now assumption (2.2) when $t_{i-1} \in B_\varepsilon(\bar t)$ and the boundedness of σ otherwise, the last term is dominated by

$$\Delta_i^{3-p}\Big(I_{\{t_{i-1}\in B_\varepsilon(\bar t)\}}K\Delta_i^{\Gamma/2} + I_{\{t_{i-1}\notin B_\varepsilon(\bar t)\}}\Delta_i\Big)$$

(since $\Delta_i^\Gamma \le \Delta_i^{\Gamma/2}$) and combining this with (A.9) the thesis follows.

    iv) The considered sum is dominated by

$$K\overline{\Delta}_n^\alpha \sum_{i:t_{i-1}\in B_\varepsilon(\bar{t})} \frac{f_n^2(t_{i-1}-\bar{t})}{f_n(0)\overline{\Delta}_n}\Delta_i^2 M_{t_{i-1}} + K \sum_{i:t_{i-1}\notin B_\varepsilon(\bar{t})} \frac{f_n^2(t_{i-1}-\bar{t})}{f_n(0)\overline{\Delta}_n}\Delta_i^2 M_{t_{i-1}}.$$

By (A.6) with $A_t = M_t I_{B_\varepsilon(\bar{t})}(t)$ we obtain that the first sum tends to zero in probability. The same result holds for the second sum, with $A_t = M_t I_{B_\varepsilon^c(\bar{t})}$, as $A(\bar{t}) = A(\bar{t}-) = 0$.

    v) This proceeds exactly as in iii) by substituting $4-p$ each time it appears with $2-p$.

    vi) Using (A.1) with $\ell = p$ we obtain

$$\mathrm{E}_{i-1}\left[\left(\int_{t_{i-1}}^{t_i}|\sigma_s^2 - \sigma_{t_{i-1}}^2|ds\right)^p\right] \le \mathrm{E}_{i-1}\left[\Delta_i^{p-1}\int_{t_{i-1}}^{t_i}|\sigma_s^2 - \sigma_{t_{i-1}}^2|^p ds\right].$$

Since $|\sigma_s^2 - \sigma_{t_{i-1}}^2| = |\sigma_s - \sigma_{t_{i-1}}||\sigma_s + \sigma_{t_{i-1}}| \le K|\sigma_s - \sigma_{t_{i-1}}|$, the last term above goes as follows

$$K\Delta_i^{p-1}\mathrm{E}_{i-1}\left[\int_{t_{i-1}}^{t_i}|\sigma_s - \sigma_{t_{i-1}}|^p ds\right]\left(I_{\{t_{i-1}\in B_\varepsilon(\bar{t})\}} + I_{\{t_{i-1}\notin B_\varepsilon(\bar{t})\}}\right) \le$$

$$K\Delta_i^p(\Delta_i^{\Gamma/2}I_{\{t_{i-1}\in B_\varepsilon(\bar{t})\}} + I_{\{t_{i-1}\notin B_\varepsilon(\bar{t})\}}).$$

<div align="right">□</div>

Theorem 2.2. It is not restrictive to set $\mu_t = 0$. We start by proving the stated convergence in law. Using Eq. (2.6) and then (A.4) and that $\mu \equiv 0$, we have

$$\frac{1}{\sqrt{f_n(0)\overline{\Delta}_n}}\left[\hat{\sigma}_{n,F}^2(\bar{t}) - (\sigma^2)_{\bar{t}}^\star\right]$$

$$= \frac{1}{\sqrt{f_n(0)\overline{\Delta}_n}}\left[\sum_{i=1}^n f_n(t_{i-1}-\bar{t})\Delta X_i^2 - (\sigma^2)_{\bar{t}}^\star\right]$$

$$= \frac{1}{\sqrt{f_n(0)\overline{\Delta}_n}}\left[\sum_{i=1}^n f_n(t_{i-1}-\bar{t})\Delta X_i^2 - \int_0^T f_n(s-\bar{t})\sigma^2(s)ds + R_n^{(\sigma^2)}(\bar{t})\right]$$

$$= \frac{1}{\sqrt{f_n(0)\overline{\Delta}_n}}\left[\sum_{i=1}^n f_n(t_{i-1}-\bar{t})\left(\Delta X_i^2 - \int_{t_{i-1}}^{t_i}\sigma^2(s)ds\right) + O_{a.s.}(L_n\overline{\Delta}_n) + R_n^{(\sigma^2)}(\bar{t})\right]$$

$$= \sum_{i=1}^n U_i + O_{a.s.}\left(L_n\sqrt{\frac{\overline{\Delta}_n}{f_n(0)}}\right) + \frac{R_n^{(\sigma^2)}(\bar{t})}{\sqrt{f_n(0)\overline{\Delta}_n}}, \tag{A.10}$$

where for $i = 1..n$

$$U_i := \frac{f_n(t_{i-1}-\bar{t})}{\sqrt{f_n(0)\overline{\Delta}_n}}\left[\left(\int_{t_{i-1}}^{t_i}\sigma_s dW_s\right)^2 - \int_{t_{i-1}}^{t_i}\sigma_s^2 ds\right]. \tag{A.11}$$

Since we assumed $L_n\sqrt{\frac{\overline{\Delta}_n}{f_n(0)}} \to 0$ and $R_n^{(\sigma^2)}(\bar{t}) = o_p\left(\sqrt{f_n(0)\overline{\Delta}_n}\right)$, the last two terms above tend to zero in probability, and thus it is sufficient we derive a central limit theorem stable in law for

<div align="center">21</div>

$\sum_{i=1}^{n} U_i$ For that we refer to Theorem IX.7.28 in Jacod and Shiryaev [29] ensuring that the following are sufficient conditions

$$(i) \quad \sum_{i=1}^{n} \mathrm{E}_{i-1}[U_i] \xrightarrow{p} 0 \qquad\qquad (iii) \quad \sum_{i=1}^{n} \mathrm{E}_{i-1}[U_i^4] \xrightarrow{p} 0$$

$$(ii) \quad \sum_{i=1}^{n} \mathrm{E}_{i-1}[U_i^2] \xrightarrow{p} V_{\bar{t}} \qquad\qquad (iv) \quad \sum_{i=1}^{n} \mathrm{E}_{i-1}[U_i \Delta H_i] \xrightarrow{p} 0,$$

where $\mathrm{E}_{i-1}[\cdot]$ abbreviates $\mathrm{E}[\cdot|\mathcal{F}_{t_{i-1}}]$ and (iv) has to hold in both the cases where $H = W$ or $H = B$, with B any bounded martingale orthogonal (in the martingale sense) to W. Condition i) is immediately proved using the Itô isometry

$$\sum_{i=1}^{n} \mathrm{E}_{i-1}[U_i] = \sum_{i=1}^{n} \frac{f_n(t_{i-1} - \bar{t})}{\sqrt{f_n(0)\overline{\Delta}_n}} \mathrm{E}_{i-1}\left[\left(\int_{t_{i-1}}^{t_i} \sigma_s dW_s\right)^2 - \int_{t_{i-1}}^{t_i} \sigma_s^2 ds\right] = 0.$$

As for condition (ii), consider

$$\sum_{i=1}^{n} \mathrm{E}_{i-1}\left[U_i^2\right] = \sum_{i=1}^{n} \frac{f_n^2(t_{i-1} - \bar{t})}{f_n(0)\overline{\Delta}_n} \mathrm{E}_{i-1}\left[\left\{\left(\int_{t_{i-1}}^{t_i} \sigma_s dW_s\right)^2 - \int_{t_{i-1}}^{t_i} \sigma_s^2 ds\right\}^2\right] =$$

$$\sum_{i=1}^{n} \frac{f_n^2(t_{i-1}-\bar{t})}{f_n(0)\overline{\Delta}_n} \left\{ \mathrm{E}_{i-1}\left[\left(\int_{t_{i-1}}^{t_i} \sigma_s dW_s\right)^4\right] + \mathrm{E}_{i-1}\left[\left(\int_{t_{i-1}}^{t_i} \sigma_s^2 ds\right)^2\right] \right.$$

$$\left. - 2\mathrm{E}_{i-1}\left[\left(\int_{t_{i-1}}^{t_i} \sigma_s dW_s\right)^2 \left(\int_{t_{i-1}}^{t_i} \sigma_s^2 ds\right)\right] \right\}. \tag{A.12}$$

All the three conditional expectations contain some leading terms, which we need to compute exactly. Basically, for $s \in]t_{i-1}, t_i]$ we write $\sigma_s = \sigma_{t_{i-1}} + (\sigma_s - \sigma_{t_{i-1}})$, we find exact equalities for the expressions containing $\sigma_{t_{i-1}}$ and by using assumption (2.2) we show that the other terms are asymptotically negligible. Now write

$$\mathrm{E}_{i-1}\left[\left(\int_{t_{i-1}}^{t_i} \sigma_s dW_s\right)^4\right] = \mathrm{E}_{i-1}\left[\left(\int_{t_{i-1}}^{t_i} \left[\sigma_{t_{i-1}} + \sigma_s - \sigma_{t_{i-1}}\right] dW_s\right)^4\right]$$

$$= 3\sigma_{t_{i-1}}^4 \Delta_i^2 + \sum_{p=0,1,2,3} c_p \mathrm{E}_{i-1}\left[(\sigma_{t_{i-1}} \Delta W_i)^p \left(\int_{t_{i-1}}^{t_i} \left[\sigma_s - \sigma_{t_{i-1}}\right] dW_s\right)^{4-p}\right],$$

with suitable constants c_p. Using Lemma A.1 iii) and iv), the first term within brackets in (A.12) contributes by

$$\sum_{i=1}^{n} \frac{f_n^2(t_{i-1} - \bar{t})}{f_n(0)\overline{\Delta}_n} 3\sigma_{t_{i-1}}^4 \Delta_i^2 + o_p(1),$$

and thanks to (A.6) this in turn has Plim equal to

$$3c_f (H'\sigma^4)_{\bar{t}}^{\star}.$$

As for the second term within brackets in (A.12) we similarly decompose it as

$$\mathrm{E}_{i-1}\left[\left(\sigma_{t_{i-1}}^2 \Delta_i + \int_{t_{i-1}}^{t_i} \sigma_s^2 - \sigma_{t_{i-1}}^2 ds\right)^2\right]$$

$$= \sigma_{t_{i-1}}^4 \Delta_i^2 + \sum_{q=0,1} c_p \mathrm{E}_{i-1} \left[(\sigma_{t_{i-1}}^2 \Delta_i)^q \left(\int_{t_{i-1}}^{t_i} \sigma_s^2 - \sigma_{t_{i-1}}^2 ds \right)^{2-q} \right].$$

Using Lemma A.1 vi) with $p = 2 - q$, both terms with $q = 0, 1$ are bounded by

$$\sigma_{t_{i-1}}^{2q} \Delta_i^q \Delta_i^{2-q} (\Delta_i^\alpha I_{\{t_{i-1} \in B_\varepsilon(\bar{t})\}} + I_{\{t_{i-1} \notin B_\varepsilon(\bar{t})\}}),$$

for some $\alpha > 0$, which by Lemma A.1 iv) give asymptotically negligible contribution to (A.12), so that the second term within brackets in (A.12) contributes by

$$\sum_{i=1}^n \frac{f_n^2(t_{i-1} - \bar{t})}{f_n(0)\overline{\Delta}_n} \sigma_{t_{i-1}}^4 \Delta_i^2 + o_p(1) \to c_f (H'\sigma^4)_{\bar{t}}^\star$$

by (A.6).

As for the third term within brackets in (A.12) we still decompose it as

$$\mathrm{E}_{i-1} \left[\left(\sigma_{t_{i-1}} \Delta W_i + \int_{t_{i-1}}^{t_i} \left(\sigma_s - \sigma_{t_{i-1}} \right) dW_s \right)^2 \left(\sigma_{t_{i-1}}^2 \Delta_i + \int_{t_{i-1}}^{t_i} \left(\sigma_s^2 - \sigma_{t_{i-1}}^2 \right) ds \right) \right]$$

$$= \sigma_{t_{i-1}}^4 \Delta_i^2 + \sum_{p=0,1} c_p \sigma_{t_{i-1}}^{2+p} \Delta_i \mathrm{E}_{i-1} \left[(\Delta W_i)^p \left(\int_{t_{i-1}}^{t_i} \left(\sigma_s - \sigma_{t_{i-1}} \right) dW_s \right)^{2-p} \right]$$

$$+ \sigma_{t_{i-1}}^2 \mathrm{E}_{i-1} \left[(\Delta W_i)^2 \int_{t_{i-1}}^{t_i} \left(\sigma_s^2 - \sigma_{t_{i-1}}^2 \right) ds \right]$$

$$+ \sum_{q=0,1} c_q \sigma_{t_{i-1}}^q \mathrm{E}_{i-1} \left[(\Delta W_i)^q \left(\int_{t_{i-1}}^{t_i} \left(\sigma_s - \sigma_{t_{i-1}} \right) dW_s \right)^{2-q} \int_{t_{i-1}}^{t_i} \left(\sigma_s^2 - \sigma_{t_{i-1}}^2 \right) ds \right].$$

By Lemma A.1 v) and iv) the terms with $p = 0, 1$ give asymptotically negligible contribution to (A.12). Noting that $\int_{t_{i-1}}^{t_i} |\sigma_s^2 - \sigma_{t_{i-1}}^2| ds \leq K\Delta_i$, the terms with $q = 0, 1$ are reduced to terms of exactly the same type as the ones with $p = 0, 1$ above and thus are asymptotically negligible. Now we deal with the term

$$\sigma_{t_{i-1}}^2 \mathrm{E}_{i-1} \left[(\Delta W_i)^2 \int_{t_{i-1}}^{t_i} \left(\sigma_s^2 - \sigma_{t_{i-1}}^2 \right) ds \right]$$

which, by the Hölder inequality is dominated by

$$\sigma_{t_{i-1}}^2 \sqrt{\mathrm{E}_{i-1}[(\Delta W_i)^4]} \sqrt{\mathrm{E}_{i-1} \left[\left| \int_{t_{i-1}}^{t_i} \left(\sigma_s^2 - \sigma_{t_{i-1}}^2 \right) ds \right|^2 \right]}.$$

Using Lemma A.1 vi) with $p = 2$, we obtain that in turn this is less than

$$K\Delta_i^2 (\Delta_i^\alpha I_{\{t_{i-1} \in B_\varepsilon(\bar{t})\}} + I_{\{t_{i-1} \notin B_\varepsilon(\bar{t})\}}),$$

with a suitable $\alpha > 0$ and by A.1 iv) also the contribution of this term is asymptotically negligible. Therefore the third term within brackets in (A.12) has the same limit in probability as

$$-2 \sum_{i=1}^n \frac{f_n^2(t_{i-1} - \bar{t})}{f_n(0)\overline{\Delta}_n} \sigma_{t_{i-1}}^4 \Delta_i^2 \xrightarrow{P} -2c_f (H'\sigma^4)_{\bar{t}}^\star. \tag{A.13}$$

by lemma A.1 ii). Summing up, the probability limit in condition (ii) is $V_{\bar{t}} = 2c_f(H'\sigma^4)_{\bar{t}}^\star$.

We now deal with the above condition (iii), where we only have to check the negligibility of the fourth conditional moments, and even some rough estimates are sufficient.

$$\sum_{i=1}^n \mathrm{E}_{i-1}\left[U_i^4\right] = \sum_{i=1}^n \frac{f_n^4(t_{i-1}-\bar{t})}{f_n^2(0)\overline{\Delta}_n^2}\, \mathrm{E}_{i-1}\left[\left\{\left(\int_{t_{i-1}}^{t_i}\sigma_s dW_s\right)^2 - \int_{t_{i-1}}^{t_i}\sigma_s^2 ds\right\}^4\right] \le$$

$$\sum_{i=1}^n \frac{f_n^4(t_{i-1}-\bar{t})}{f_n^2(0)\overline{\Delta}_n^2}\, \mathrm{E}_{i-1}\left[\left(\int_{t_{i-1}}^{t_i}\sigma_s dW_s\right)^8 + \left(\int_{t_{i-1}}^{t_i}\sigma_s^2 ds\right)^4\right].$$

As σ_s is assumed wlog to be bounded, by the BDG inequality we have that $\mathrm{E}_{i-1}\left[\left(\int_{t_{i-1}}^{t_i}\sigma_s dW_s\right)^8\right] \le K\Delta_i^4$ and the last sum above is dominated by

$$K\sum_{i=1}^n \frac{f_n^4(t_{i-1}-\bar{t})\Delta_i}{f_n^2(0)}\frac{\Delta_i^3}{\overline{\Delta}_n^2} \le K\overline{\Delta}_n\sum_{i=1}^n \frac{f_n^4(t_{i-1}-\bar{t})\Delta_i}{f_n^2(0)}, \tag{A.14}$$

having used that, by Assumption 2, $\max_i \Delta_i \le K\overline{\Delta}_n$. However

$$\overline{\Delta}_n\left|\sum_{i=1}^n \frac{f_n^4(t_{i-1}-\bar{t})\Delta_i}{f_n^2(0)} - \frac{\int_0^T f_n^4(s-\bar{t})ds}{f_n^2(0)}\right| \xrightarrow{P} 0,$$

as $f_n^4 = g(f_n)$, with $g(x) = x^4$, is also Lipschitz on D with Lipschitz constant $KK_n^3 L_n$ and thus the sum within the last display is bounded by

$$\frac{\overline{\Delta}_n}{f_n^2(0)}\sum_{i=1}^n \int_{t_{i-1}}^{t_i} |f_n^4(t_{i-1}-\bar{t}) - f_n^4(s-\bar{t})|ds \le K\frac{\overline{\Delta}_n^2}{f_n^2(0)}K_n^3 L_n$$

$$\le KL_n\sqrt{\frac{\overline{\Delta}_n}{f_n(0)}}\left(L_n\sqrt{\frac{\overline{\Delta}_n}{f_n(0)}}\right)^3 \to 0.$$

Consequently, using Eq. (2.8) with $\sigma \equiv 1$, the probability limit of (A.14) is the same as

$$\mathrm{plim}\, K\overline{\Delta}_n\frac{\int_0^T f_n^4(s-\bar{t})ds}{f_n^2(0)} \le \mathrm{plim}\, K\overline{\Delta}_n f_n(0) = 0,$$

by assumption.

We finally consider condition (iv), starting from the case $H = B$. Denote $R_t = \int_0^t \sigma_s dW_s$ and $M_t = R_t^2 - \int_0^t \sigma_s^2 ds$. Since B is orthogonal to W we have that $d[M, B] \equiv 0$. In fact $d(R^2)_t = 2R_t dR_t + d[R, R]_t = 2R_t dR_t + \sigma_t^2 dt$, thus $d[M, B]_t = d[R^2, B]_t = 2R_t d[R, B]_t = 2R_t\sigma_t d[W, B]_t \equiv 0$. Therefore also M and W are orthogonal, meaning that $E_{t_{i-1}}[\Delta M_i \Delta B_i] = 0$ for all $i = 1..n$, and condition (iv) is verified.

When instead $H = W$,

$$\sum_{i=1}^n \mathrm{E}_{i-1}\left[U_i \Delta W_i\right] \le$$

$$\leq \sum_{i=1}^{n} \frac{f_n(t_{i-1} - \bar{t})}{\sqrt{f_n(0)\overline{\Delta}_n}} \sqrt{\mathrm{E}_{i-1}\left[\left(\left(\int_{t_{i-1}}^{t_i} \sigma_s dW_s\right)^2 - \int_{t_{i-1}}^{t_i} \sigma_s^2 ds\right)^2\right]} \sqrt{\mathrm{E}_{i-1}\left[(\Delta W_i^2)\right]}.$$

Using the computations done in order to evaluate $E_{i-1}[U_i^2]$ above, we have

$$\mathrm{E}_{i-1}\left[\left(\left(\int_{t_{i-1}}^{t_i} \sigma_s dW_s\right)^2 - \int_{t_{i-1}}^{t_i} \sigma_s^2 ds\right)^2\right]$$

$$\leq 2\sigma_{t_{i-1}}^4 \Delta_i^2 + K\Delta_i^2 \left(\Delta_i^\alpha I_{\{t_{i-1} \in B_\varepsilon(\bar{t})\}} + I_{\{t_{i-1} \notin B_\varepsilon(\bar{t})\}}\right),$$

for a suitable $\alpha > 0$, therefore the last term within the above display is dominated by

$$\sum_{i=1}^{n} \frac{f_n(t_{i-1} - \bar{t})}{\sqrt{f_n(0)\overline{\Delta}_n}} \Delta_i^{\frac{3}{2}} \sqrt{2\sigma_{t_{i-1}}^4 + K(\Delta_i^\alpha I_{\{t_{i-1} \in B_\varepsilon(\bar{t})\}} + I_{\{t_{i-1} \notin B_\varepsilon(\bar{t})\}})} \leq K \sum_{i=1}^{n} \frac{f_n(t_{i-1} - \bar{t})}{\sqrt{f_n(0)}} \Delta_i$$

By (A.5) i) and the assumptions $L_n \overline{\Delta}_n \to 0$ and $f_n(0) \to \infty$, this last has the same probability limit as $\frac{\int_0^T f_n(s - \bar{t}) ds}{f_n(0)}$, which is zero, as $\int_0^T f_n(s - \bar{t}) ds \to 1$ by (2.6), and condition (iv) is verified also when $H = W$ and this completes the proof of the stable convergence of $\hat{\sigma}_{n,f}^2(\bar{t})$.

For the convergence in probability the condition $R_n^{(\sigma^2)}(\bar{t}) = o_p\left(\sqrt{f_n(0)\overline{\Delta}_n}\right)$ is not required. Indeed by multiplying both sides of Eq. (A.10) by $\sqrt{f_n(0)\overline{\Delta}_n}$ we find

$$\hat{\sigma}_{n,F}^2(\bar{t}) - (\sigma^2)_{\bar{t}}^\star = \sum_{i=1}^{n} f_n(t_{i-1} - \bar{t}) \left[\left(\int_{t_{i-1}}^{t_i} \sigma_s dW_s\right)^2 - \int_{t_{i-1}}^{t_i} \sigma_s^2 ds\right]$$

$$+ O_{a.s.}(f_n(0)\overline{\Delta}_n) + R_n^{(\sigma^2)}(\bar{t}).$$

Last two terms are $o_p(1)$ by the assumption $f_n(0)\overline{\Delta}_n \to 0$ and (2.6), while we check the negligibility of the first term by using the law of large numbers for the sum of martingale differences (see e.g. Lemma 4.1 in Jacod [26]). It is sufficient to show that

$$\sum_i E_{i-1}\left[\left(\sqrt{f_n(0)\overline{\Delta}_n} U_i\right)^2\right] = f_n(0)\overline{\Delta}_n \sum \sum_i E_{i-1}[U_i^2] \to 0,$$

which is ensured by $\sum_i E_{i-1}[U_i^2] \to 2c_f(H'\sigma^4)_{\bar{t}}^\star$ obtained with the computations for ii) above and by $f_n(0)\overline{\Delta}_n \to 0$.

We now check the negligibility of the drift in order to reach the consistency and the CLT for $\hat{\sigma}_{n,F}^2(\bar{t})$. If $\mu\nu \equiv 0$, by following (A.10) and substituting ΔX_i we have

$$\frac{\hat{\sigma}_{n,F}^2(\bar{t}) - (\sigma^2)_{\bar{t}}^\star}{\sqrt{f_n(0)\overline{\Delta}_n}} = \sum_{i=1}^{n} U_i + 2\sum_i \frac{f_n(t_{i-1} - \bar{t})}{\sqrt{f_n(0)\overline{\Delta}_n}} \int_{t_{i-1}}^{t_i} \sigma_s dW_s \int_{t_{i-1}}^{t_i} \mu_s ds$$

$$+ \sum_i \frac{f_n(t_{i-1} - \bar{t})}{\sqrt{f_n(0)\overline{\Delta}_n}} (\int_{t_{i-1}}^{t_i} \mu_s ds)^2 + O_{a.s.}(L_n \sqrt{\frac{\overline{\Delta}_n}{f_n(0)}}) + \frac{R_n^{(\sigma^2)}(\bar{t})}{\sqrt{f_n(0)\overline{\Delta}_n}},$$

and we see firstly that again the assumption $\frac{R_n^{(\sigma^2)}(\bar{t})}{\sqrt{f_n(0)\overline{\Delta}_n}} \to_p 0$ is only needed for the CLT and secondly that showing the negligibility of the second and third terms above is sufficient also to state the consistency of $\hat{\sigma}_{n,F}^2(\bar{t})$. To deal with such terms, which are both of the type $\sum_i \xi_i$, we apply Lemma 4.1 in [26], and in both cases we check condition (4.4). Since $\sum_i E_{i-1}[|\xi_i|] \leq \sum_i \sqrt{E_{i-1}[|\xi_i|^2]} \leq K\sqrt{E_{i-1}[|\xi_i|^2]}$, it is sufficient to check that $\sum_i E_{i-1}[\xi_i^2] \to 0$. By the boundedness of μ and BDG inequality and then (A.6) we have

$$\sum_i \frac{f_n^2(t_{i-1}-\bar{t})}{f_n(0)\overline{\Delta}_n} E_{i-1}\left[\left(\int_{t_{i-1}}^{t_i} \sigma_s dW_s\right)^2 \left(\int_{t_{i-1}}^{t_i} \mu_s ds\right)^2\right] \leq K \sum_i \frac{f_n^2(t_{i-1}-\bar{t})}{f_n(0)\overline{\Delta}_n} \Delta_i^3$$

$$\leq K\overline{\Delta}_n \sum_i \frac{f_n^2(t_{i-1}-\bar{t})}{f_n(0)\overline{\Delta}_n} \Delta_i^2 \to_p 0$$

and

$$\sum_i \frac{f_n^2(t_{i-1}-\bar{t})}{f_n(0)\overline{\Delta}_n} E_{i-1}\left[\left(\int_{t_{i-1}}^{t_i} \mu_s ds\right)^4\right] \leq K\overline{\Delta}_n^2 \sum_i \frac{f_n^2(t_{i-1}-\bar{t})}{f_n(0)\overline{\Delta}_n} \Delta_i^2 \to_p 0$$

as desired. $\square$

Remark 1. Assume $\bar{t} \in]0,T[$. For the Gaussian kernel, with ε as in (2.2) we have

$$\frac{R_n^{(\sigma^2)}(\bar{t})}{\sqrt{f_n(0)\overline{\Delta}_n}} = A_n + \frac{1}{\sqrt{f_n(0)\overline{\Delta}_n}} \left[\frac{1}{h_n} \int_{s \in [0,T]:|s-\bar{t}| \leq \varepsilon} K\left(\frac{s-\bar{t}}{h_n}\right) \sigma_s^2 ds - (\sigma^2)_{\bar{t}}^{\star}\right], \qquad (\text{A.15})$$

By changing variable via $x = (s-\bar{t})/h_n$, the second term in (A.15) is written as $B_n + C_n + D_n + E_n$ with

$$B_n = \frac{\int_{-\varepsilon/h_n}^0 K(x)(\sigma_{\bar{t}-h_n|x|}^2 - \sigma_{\bar{t}-}^2) dx}{\sqrt{f_n(0)\overline{\Delta}_n}}, \quad C_n = \frac{\int_0^{\varepsilon/h_n} K(x)(\sigma_{\bar{t}+h_n x}^2 - \sigma_{\bar{t}}^2) dx}{\sqrt{f_n(0)\overline{\Delta}_n}},$$

$$D_n = -\frac{\int_{-\infty}^{-\varepsilon/h_n} K(x)\sigma_{\bar{t}-}^2 dx}{\sqrt{f_n(0)\overline{\Delta}_n}}, \quad E_n = -\frac{\int_{\varepsilon/h_n}^{+\infty} K(x)\sigma_t^2 dx}{\sqrt{f_n(0)\overline{\Delta}_n}}.$$

Using Eq. (2.2), we have:

$$E[|B_n|] \leq \frac{1}{\sqrt{f_n(0)\overline{\Delta}_n}} \int_{-\varepsilon/h_n}^0 K(x)(x h_n)^{\Gamma/2} dx \leq C\sqrt{n h_n^{1+\Gamma}} \to 0.$$

A similar result holds for C_n. Moreover,

$$|A_n| = \frac{1}{h_n \sqrt{f_n(0)\overline{\Delta}_n}} \left|\int_{s \in [0,T]:|s-\bar{t}| > \varepsilon} K\left(\frac{s-\bar{t}}{h}\right) \sigma_s^2 ds\right| =$$

$$\sqrt{\frac{h_n}{K(0)\overline{\Delta}_n}} \left|\int_{-\bar{t}/h_n}^{-\varepsilon/h_n} K(x)\sigma_{h_n x+\bar{t}}^2 dx + \int_{\varepsilon/h_n}^{(T-\bar{t})/h_n} K(x)\sigma_{h_n x+\bar{t}}^2 dx\right| \leq$$

$$\sqrt{n h_n} \int_{-\bar{t}/h_n}^{-\varepsilon/h_n} K(x) dx + \sqrt{n h_n} \int_{\varepsilon/h_n}^{(T-\bar{t})/h_n} K(x) dx.$$

For the first term, on $[-\bar{t}/h_n, -\varepsilon/h_n]$, for small h_n we have $K(x) \leq K(-\varepsilon/h_n) = e^{-\frac{1}{2}(\frac{\varepsilon}{h_n})^2} = o\left(\left(\frac{\varepsilon}{h_n}\right)^4\right) \sim h_n^4$, so that $\sqrt{nh_n} \int_{-\bar{t}/h_n}^{-\varepsilon/h_n} K(x) dx \leq C\sqrt{nh_n^7} \to 0$. The second term can be dealt with in a similar way.

Finally for D_n, the well known inequality

$$\text{for } y > 0: \int_y^{+\infty} K(x) dx \leq \frac{C}{y} K(y)$$

and the boundedness of σ imply

$$|D_n| = \frac{1}{\sqrt{f_n(0)\overline{\Delta}_n}} \int_{-\infty}^{-\varepsilon/h_n} K(x)\sigma_{\bar{t}-}^2 dx \leq C\sqrt{nh_n} \int_{-\infty}^{-\varepsilon/h_n} K(x) dx \leq \sqrt{nh_n} \frac{h_n}{\varepsilon} e^{\frac{-\varepsilon^2}{2h_n^2}}$$

$$\leq C\sqrt{nh_n^4} \to 0.$$

E_n is dealt with similarly.

For the Epanechnikov kernel, $A_n = 0$ for n large enough, and the rest is similar.

For the indicator kernel, for n large enough,

$$\text{E}\left[\frac{|R_n^{(\sigma^2)}(\bar{t})|}{\sqrt{f_n(0)\overline{\Delta}_n}}\right] = \frac{\int_{|x| \leq 1/f_n(0)} \frac{1}{2} f_n(0) \text{E}\left[\left|\sigma_{\bar{t}+x/f_n(0)}^2 - (\sigma^2)_{\bar{t}}^\star\right|\right] dx}{\sqrt{f_n(0)\overline{\Delta}_n}} \leq C \frac{f_n(0)^{-\Gamma/2}}{\sqrt{f_n(0)\overline{\Delta}_n}},$$

providing the same result.

$\square$

Lemma 3.1. The proof is based on that of Theorem 2.2. In what follows, we compute conditional expectations with respect to a new augmented filtration $\mathcal{F}_t^\varepsilon$ obtained by including the observed noise $(\varepsilon_i)_{t_i \leq t}$ for each $t \in [0,T]$. As shown in the proof of Theorem 2.2, it is harmless to set $\mu_t = 0$ so we do it in what follows. Write,

$$\frac{1}{2}\overline{\Delta}_n \widehat{\sigma}_{n,f}^2(\bar{t}) - V_\varepsilon = \frac{1}{2}\overline{\Delta}_n(A_n + B_n + C_n) - V_\varepsilon,$$

where

$$A_n = \sum_{i=1}^n f_n(t_{i-1} - \bar{t})(\Delta Y_i)^2$$

$$B_n = 2\sum_{i=1}^n f_n(t_{i-1} - \bar{t})\Delta Y_i(\varepsilon_i - \varepsilon_{i-1})$$

$$C_n = \sum_{i=1}^n f_n(t_{i-1} - \bar{t})(\varepsilon_i - \varepsilon_{i-1})^2.$$

For the first term A_n, we can simply apply Theorem 2.2 and obtain that $\overline{\Delta}_n A_n = O_p(\overline{\Delta}_n)$. Next, write $B_n = \sqrt{f_n(0)} \sum_{i=1}^n U_{B,i}$ where

$$U_{B,i} := \frac{2}{\sqrt{f_n(0)}} f_n(t_{i-1} - \bar{t})\Delta Y_i(\varepsilon_i - \varepsilon_{i-1}).$$

We have $E_{i-1}[U_{B,i}] = 0$ and, using the independence of the noise,:

$$
\begin{aligned}
\sum_{i=1}^{n} \mathrm{E}_{i-1}\left[U_{B,i}^2\right] &= 4\sum_{i=1}^{n} \frac{f_n^2(t_{i-1} - \bar{t})}{f_n(0)} E_{i-1}\left[(\Delta Y_i)^2\right] E_{i-1}[(\varepsilon_i - \varepsilon_{i-1})^2] \\
&= 4\sum_{i=1}^{n} \frac{f_n^2(t_{i-1} - \bar{t})}{f_n(0)} E_{i-1}\left[(\Delta Y_i)^2\right] (V_\epsilon + \epsilon_{i-1}^2)
\end{aligned}
$$

Now notice that, using the boundedness on σ_s and Eq. (A.2),

$$
\begin{aligned}
\mathrm{E}_{i-1}[(\Delta Y_i)^2] &= \int_{t_{i-1}}^{t_i} \mathrm{E}_{i-1}[\sigma_s^2]ds = \sigma_{t_{i-1}}^2 \overline{\Delta}_n + \int_{t_{i-1}}^{t_i} \mathrm{E}_{i-1}[\sigma_s^2 - \sigma_{t_{i-1}}^2]ds \\
&= \sigma_{t_{i-1}}^2 \overline{\Delta}_n + O_p\left(\overline{\Delta}_n^{1+\Gamma/2}\right)
\end{aligned}
\tag{A.16}
$$

so that we can write:

$$
\begin{aligned}
\sum_{i=1}^{n} \mathrm{E}_{i-1}\left[U_{B,i}^2\right] &= 4\sum_{i=1}^{n} \frac{f_n^2(t_{i-1} - \bar{t})}{f_n(0)} (\sigma_{t_{i-1}}^2 \overline{\Delta}_n + O_p(\overline{\Delta}_n^{1+\Gamma/2}))(V_\varepsilon + \varepsilon_{i-1}^2) \\
&= 8\sum_{i=1}^{n} \frac{f_n^2(t_{i-1} - \bar{t})}{f_n(0)} \sigma_{t_{i-1}}^2 \overline{\Delta}_n V_\varepsilon + o_p(1) + R_{B,n}
\end{aligned}
$$

where the $o_p(1)$ terms is the term multiplying $O_p\left(\overline{\Delta}_n^{1+\Gamma/2}\right)$, while

$$
R_{B,n} = 4\sum_{i=1}^{n} \frac{f_n^2(t_{i-1} - \bar{t})}{f_n(0)} \sigma_{t_{i-1}}^2 \overline{\Delta}_n (\varepsilon_{i-1}^2 - V_\varepsilon) \xrightarrow{p} 0
$$

by a generalized version of the law of large numbers, since $\mathrm{E}[R_{B,n}] = 0$ and

$$
16\sum_{i=1}^{n} \frac{f_n^4(t_{i-1} - \bar{t})}{f_n^2(0)} \mathrm{E}[\sigma_{t_{i-1}}^4]\overline{\Delta}_n^2 \mathrm{E}[(\varepsilon_{i-1}^2 - V_\varepsilon)] = O_p(f_n(0)\overline{\Delta}_n)
$$

by the boundedness of σ_s, the finiteness of the moments of the ϵ and by Eq. (2.8). This proves that $\sum_{i=1}^{n} \mathrm{E}_{i-1}\left[U_{B,i}^2\right] \xrightarrow{p} 8V_\varepsilon c_f(\sigma^2)^\star(\bar{t})$. Similarly, using also the BDG inequality,

$$
\begin{aligned}
\sum_{i=1}^{n} \mathrm{E}_{i-1}\left[U_{B,i}^4\right] &= 16\sum_{i=1}^{n} \frac{f_n^4(t_{i-1} - \bar{t})}{f_n^2(0)} E_{i-1}\left[(\Delta Y_i)^4\right] E_{i-1}[(\varepsilon_i - \varepsilon_{i-1})^4] \\
&\leq 16K\sum_{i=1}^{n} \frac{f_n^4(t_{i-1} - \bar{t})}{f_n^2(0)} E_{i-1}\left[(\Delta Y_i)^2\right]^2 E_{i-1}[(\varepsilon_i - \varepsilon_{i-1})^4] \\
&= 16K\sum_{i=1}^{n} \frac{f_n^4(t_{i-1} - \bar{t})}{f_n^2(0)} \sigma^4(t_{i-1})\Delta_i^2 E_{i-1}[(\varepsilon_i - \varepsilon_{i-1})^4] + o_p(1) \\
&= O_p(f_n(0)\overline{\Delta}_n) \to 0
\end{aligned}
$$

Then

$$
\frac{B_n}{\sqrt{f_n(0)}} \longrightarrow \mathbf{MN}\left(0, 8V_\varepsilon c_f(\sigma^2)^\star(\bar{t})\right),
$$

that is, $\overline{\Delta}_n B_n = O_p(\overline{\Delta}_n \sqrt{f_n(0)})$. Now consider the term $D_n := \frac{1}{2}\overline{\Delta}_n C_n - V_\varepsilon$. Write:

$$
U_{C,i} = \frac{1}{\sqrt{\overline{\Delta}_n f_n(0)}} \frac{1}{2} f_n(t_{i-1} - \bar{t})\overline{\Delta}_n \left[(\varepsilon_i - \varepsilon_{i-1})^2 - 2V_\varepsilon\right]
$$

and notice that, since $\sum_{i=1}^{n} f_n(\bar{t} - t_{i-1})\overline{\Delta}_n - 1 = O(L_n\overline{\Delta}_n) + o(\sqrt{f_n(0)\overline{\Delta}_n})$, we have $D_n = \sqrt{\overline{\Delta}_n f_n(0)} \sum_{i=1}^{n} U_{C,i} + o_p\left(\sqrt{\overline{\Delta}_n f_n(0)}\right)$. We immediately have $\mathrm{E}[U_{C,i}] = 0$, and:

$$
\sum_{i=1}^{n} \mathrm{E}[U_{C,i}^2] = \frac{1}{4} \sum_{i=1}^{n} \frac{1}{f_n(0)} f_n^2(t_{i-1} - \bar{t})\overline{\Delta}_n \left(2\kappa_\varepsilon + 2V_\varepsilon^2\right)
$$
$$
\xrightarrow{p} \frac{1}{2} c_f \left(\kappa_\varepsilon + V_\varepsilon^2\right),
$$

while:

$$
\sum_{i=1}^{n} \mathrm{E}[U_{C,i}^4] = \frac{1}{16} \sum_{i=1}^{n} \frac{1}{f_n(0)^2} f_n^4(t_{i-1} - \bar{t})\overline{\Delta}_n^2 \mathrm{E}\left[\left((\varepsilon_i - \varepsilon_{i-1})^2 - 2V_\varepsilon\right)^4\right]
$$
$$
= O_p(f_n(0)\overline{\Delta}_n)
$$

which proves that $D_n = O_p\left(\sqrt{f_n(0)\overline{\Delta}_n}\right)$, which is the leading order, and also proves the statement in Eq. (3.2). □

Theorem 3.2. Write $Y_i := Y_{t_i}$ and $X_i := X_{t_i}$. We have:

$$
\hat{\sigma}_{n,\bar{n}}^{2,TS}(\bar{t}) = \frac{1}{\bar{n}} \sum_{i=1}^{n-\bar{n}+1} f_n(t_{i-1} - \bar{t})(Y_{i+\bar{n}-1} - Y_{i-1})^2
$$
$$
+ \frac{1}{\bar{n}} \sum_{i=1}^{n-\bar{n}+1} f_n(t_{i-1} - \bar{t})(\epsilon_{i+\bar{n}-1} - \epsilon_{i-1})^2
$$
$$
+ \frac{2}{\bar{n}} \sum_{i=1}^{n-\bar{n}+1} f_n(t_{i-1} - \bar{t})(\epsilon_{i+\bar{n}-1} - \epsilon_{i-1})(Y_{i+\bar{n}-1} - Y_{i-1})
$$
$$
- \frac{1}{\bar{n}} \sum_{i=1}^{n-\bar{n}+1} f_n(t_{i-1} - \bar{t})(Y_i - Y_{i-1})^2
$$
$$
- \frac{1}{\bar{n}} \sum_{i=1}^{n-\bar{n}+1} f_n(t_{i-1} - \bar{t})(\epsilon_i - \epsilon_{i-1})^2
$$
$$
- \frac{2}{\bar{n}} \sum_{i=1}^{n-\bar{n}+1} f_n(t_{i-1} - \bar{t})(\epsilon_i - \epsilon_{i-1})(Y_i - Y_{i-1})
$$
$$
:= A_1 + A_2 + A_3 + B_1 + B_2 + B_3
$$

and define $C_2 = A_2 + B_2$ and $C_3 = A_3 + B_3$.

Start with A_1 and write:

$$
A_1 = \frac{1}{\bar{n}} \sum_{i=1}^{n-\bar{n}-1} f_n(t_{i-1} - \bar{t}) \left(\sum_{j=1}^{\bar{n}} (Y_{i-1+j} - Y_{i-1+j-1})\right)^2
$$
$$
= \frac{1}{\bar{n}} \sum_{i=1}^{n} (Y_i - Y_{i-1})^2 \sum_{j=1}^{\bar{n} \wedge i} f_n(t_{i-j} - \bar{t})
$$
$$
+ \frac{2}{\bar{n}} \sum_{i=1}^{n-\bar{n}+1} f_n(t_{i-1} - \bar{t}) \sum_{j>k\geq 0}^{\bar{n}} (Y_{i-1+j} - Y_{i-1+j-1})(Y_{i-1+k} - Y_{i-1+k-1})
$$
$$
:= a + b
$$

Using the Lipschitz property of f_n:

$$f_n(t_{i-j} - \bar{t}) = f_n(t_{i-1} - \bar{t}) + O(L_n(j-1)\overline{\Delta}_n) \tag{A.17}$$

so that, using Theorem 2.2,:

$$
\begin{aligned}
a &= \sum_{i=1}^{n} f_n(t_{i-1} - \bar{t})(Y_i - Y_{i-1})^2 \left(\frac{\overline{n} \wedge i}{\overline{n}} + \frac{1}{\overline{n}} \sum_{j=1}^{\overline{n} \wedge i} O(L_n(j-1)\overline{\Delta}_n) \right) \\
&= \sum_{i=1}^{n} f_n(t_{i-1} - \bar{t})(Y_i - Y_{i-1})^2 \left(1 + \frac{\overline{n} \wedge i - \overline{n}}{\overline{n}} + O(L_n \overline{n} \overline{\Delta}_n) \right) \\
&= \left((\sigma^2)^\star(\bar{t}) + O_p\left(\sqrt{f_n(0)\overline{\Delta}_n} \right) \right) + (end\ effect) + O_p(L_n \overline{n} \overline{\Delta}_n)
\end{aligned}
$$

The explicit evaluation of the end effect, using the properties of f_n in Assumptions 3 and Eq. (A.2), gives:

$$\sum_{i=1}^{n} f_n(t_{i-1} - \bar{t})(Y_i - Y_{i-1})^2 \frac{\overline{n} \wedge i - \overline{n}}{\overline{n}} = \sum_{i=1}^{\overline{n}-1} f_n(t_{i-1} - \bar{t})(Y_i - Y_{i-1})^2 \frac{i - \overline{n}}{\overline{n}} = O_p(\overline{n}\overline{\Delta}_n) \tag{A.18}$$

Thus we have proved that $a - (\sigma^2)^\star(\bar{t}) = o_p\left(\sqrt{f_n(0)\overline{n}\overline{\Delta}_n} \right)$. For the term b, write, using again the Lipschitz property (A.17), rearranging the terms in the sum and evaluating end effects as in Eq. (A.18),

$$
\begin{aligned}
b &= \frac{2}{\overline{n}} \sum_{i=\overline{n}}^{n} (Y_i - Y_{i-1}) \sum_{j=1}^{\overline{n}-1} f_n(t_{i-j-1} - \bar{t})(Y_{i-1} - Y_{i-j-1}) + (end\ effects) \\
&= \frac{2}{\overline{n}} \sum_{i=\overline{n}}^{n} (Y_i - Y_{i-1}) \sum_{j=1}^{\overline{n}-1} \left(f_n(t_{i-1} - \bar{t}) + O(L_n \overline{n} \overline{\Delta}_n) \right)(Y_{i-1} - Y_{i-j-1}) + o_p\left(\sqrt{f_n(0)\overline{n}\overline{\Delta}_n} \right) \\
&= \frac{2}{\overline{n}} \sum_{i=\overline{n}}^{n} f_n(t_{i-1} - \bar{t})(Y_i - Y_{i-1}) \sum_{j=1}^{\overline{n}-1} \sum_{k=0}^{j-1} (Y_{i-j+k} - Y_{i-j+k-1}) + o_p\left(\sqrt{f_n(0)\overline{n}\overline{\Delta}_n} \right) \\
&= 2 \sum_{i=\overline{n}}^{n} f_n(t_{i-1} - \bar{t})(Y_i - Y_{i-1}) \sum_{j=1}^{\overline{n}-1} \left(1 - \frac{j}{\overline{n}} \right)(Y_{i-j} - Y_{i-j-1}) + o_p\left(\sqrt{f_n(0)\overline{n}\overline{\Delta}_n} \right) \\
&:= \sum_{i=\overline{n}}^{n} b_{n,i} + o_p\left(\sqrt{f_n(0)\overline{n}\overline{\Delta}_n} \right)
\end{aligned}
$$

Now, $E_{i-1}[b_{n,i}] = 0$ and, using Eq. (A.16):

$$
\begin{aligned}
\sum_{i=\overline{n}}^{n} E_{i-1}[b_{n,i}^2] &= 4 \sum_{i=\overline{n}}^{n} f_n^2(t_{i-1} - \bar{t}) E_{i-1}[(Y_i - Y_{i-1})^2] \left(\sum_{j=1}^{\overline{n}-1} \left(1 - \frac{j}{\overline{n}} \right)(Y_{i-j} - Y_{i-j-1}) \right)^2 \\
&= 4 \sum_{i=\overline{n}}^{n} f_n^2(t_{i-1} - \bar{t}) \left(\sigma_{t_{i-1}}^2 \overline{\Delta}_n + O_p(\overline{\Delta}_n^{1+\Gamma/2}) \right) \left(\sum_{j=1}^{\overline{n}-1} \left(1 - \frac{j}{\overline{n}} \right)(Y_{i-j} - Y_{i-j-1}) \right)^2 \\
&= \left(1 + O_p(\overline{\Delta}_n^{\Gamma/2}) \right) \cdot \left[4 \sum_{i=\overline{n}}^{n} f_n^2(t_{i-1} - \bar{t}) \sigma_{t_{i-1}}^2 \overline{\Delta}_n \sum_{j=1}^{\overline{n}-1} \left(1 - \frac{j}{\overline{n}} \right)^2 (Y_{i-j} - Y_{i-j-1})^2 \right.
\end{aligned}
$$

30

$$+8\sum_{i=\overline{n}}^{n} f_n^2(t_{i-1}-\overline{t})\sigma_{t_{i-1}}^2\overline{\Delta}_n \sum_{j>k\geq 0}^{\overline{n}-1}\left(1-\frac{j}{n}\right)\left(1-\frac{k}{n}\right)(Y_{i-j}-Y_{i-j-1})(Y_{i-k}-Y_{i-k-1})\Bigg]$$

$$:=\left(1+O_p(\overline{\Delta}_n^{\Gamma/2})\right)(c+d)$$

Now write

$$(Y_{i-j}-Y_{i-j-1})^2 = \left(\int_{t_{i-j-1}}^{t_{i-j}}\sigma_s dW_s\right)^2 =$$

$$=\left(\int_{t_{i-j-1}}^{t_{i-j}}\left(\sigma_s-\sigma_{t_{i-j-1}}\right)dW_s + \sigma_{t_{i-j-1}}\Delta_{i-j}W_s\right)^2$$

$$=\left(\int_{t_{i-j-1}}^{t_{i-j}}\left(\sigma_s-\sigma_{t_{i-j-1}}\right)dW_s\right)^2$$

$$+2\sigma_{t_{i-j-1}}\Delta_{i-j}W_s\int_{t_{i-j-1}}^{t_{i-j}}\left(\sigma_s-\sigma_{t_{i-j-1}}\right)dW_s + \sigma_{t_{i-j-1}}^2\left(\Delta_{i-j}W_s\right)^2,$$

and, accordingly, $c=c_1+c_2+c_3$. Using the boundedness of σ_s, BDG inequality and property (2.2) it is straightforward to prove that c_1 is $o_p(\overline{n}\overline{\Delta}_n f_n(0))$, and further using the Holder inequality also c_2 is proved to be $o_p(\overline{n}\overline{\Delta}_n f_n(0))$. We then deal with the term c_3. Start by noticing that we can rearrange the sums and neglect end-effects which are given by $O(\overline{n}^2)$ terms of the same kind of the $O_p(n\overline{n})$ terms that we are going to retain, which are the following:

$$c_3 = 4\sum_{i=\overline{n}}^{n} f_n^2(t_{i-1}-\overline{t})\sigma_{t_{i-1}}^2\overline{\Delta}_n \sum_{j=1}^{\overline{n}-1}\left(1-\frac{j}{n}\right)^2\sigma_{t_{i-j-1}}^2\left(\Delta_{i-j}W_s\right)^2$$

$$= 4\sum_{i=\overline{n}-1}^{n-\overline{n}+1}\sigma_{t_i}^2(\Delta_{i+1}W_s)^2\overline{\Delta}_n \sum_{j=1}^{\overline{n}-1}\left(\frac{j}{n}\right)^2\sigma_{t_{i+\overline{n}-j}}^2 f_n^2(t_{i+\overline{n}-j}-\overline{t})$$

Now we use the properties of f_n, the boundedness of σ and Lemma A.1 to show that:

$$c_3 = 4\sum_{i=\overline{n}-1}^{n-\overline{n}+1}\sigma_{t_i}^2(\Delta_{i+1}W_s)^2\overline{\Delta}_n \sum_{j=1}^{\overline{n}-1}\left(\frac{j}{n}\right)^2\sigma_{t_{i+\overline{n}-j}}^2\left(f_n^2(t_{i-1}-\overline{t})+O_p(L_n^2\overline{\Delta}_n\overline{n})\right)$$

$$= 4\sum_{i=\overline{n}-1}^{n-\overline{n}+1}\sigma_{t_i}^2(\Delta_{i+1}W_s)^2\overline{\Delta}_n \sum_{j=1}^{\overline{n}-1}\left(\frac{j}{n}\right)^2\left[\sigma_{t_i}^2+(\sigma_{t_{i+\overline{n}-j}}^2-\sigma_{t_i}^2)\right]f_n^2(t_{i-1}-\overline{t})$$

$$+o_p(\overline{n}\overline{\Delta}_n f_n(0))$$

$$:= c_4 + c_5 + o_p(\overline{n}\overline{\Delta}_n f_n(0))$$

The term $c_4 = 4\sum_{i=\overline{n}-1}^{n-\overline{n}+1}\sigma_{t_i}^2(\Delta_{i+1}W_s)^2\overline{\Delta}_n \sum_{j=1}^{\overline{n}-1}\left(\frac{j}{n}\right)^2\sigma_{t_i}^2 f_n^2(t_{i-1}-\overline{t}) = \sum_{i=0}^{n-1}c_{4,n,i}+end\ effects$ is the leading term since:

$$\frac{1}{\overline{n}\overline{\Delta}_n f_n(0)}\sum E_i[c_{4,n,i}] = 4\sum_{i=\overline{n}-1}^{n-\overline{n}+1}\frac{f_n^2(t_{i-1}-\overline{t})}{f_n(0)}\sigma_{t_i}^4\overline{\Delta}_n\frac{1}{n}\sum_{j=1}^{\overline{n}-1}\left(\frac{j}{n}\right)^2 \xrightarrow{p} \frac{4}{3}c_f(\sigma_{\overline{t}}^4)^\star$$

and, using property 2.9 and the Lipschitz property of f_n^4:

$$\frac{1}{(\overline{n}\overline{\Delta}_n f_n(0))^2} \sum \mathrm{E}_i[c_{4,n,i}^2] \xrightarrow{P} 0$$

which shows that $c_4/(\overline{n}\overline{\Delta}_n f_n(0)) \xrightarrow{P} \frac{4}{3} c_f(\sigma_{\overline{t}}^4)^\star$. The term c_5 is instead negligible, since:

$$\frac{1}{\overline{n}\overline{\Delta}_n f_n(0)} \mathrm{E}[|c_5|] \leq K \sum_{i=\overline{n}-1}^{n-\overline{n}+1} \overline{\Delta}_n \frac{1}{\overline{n}} \sum_{j=1}^{\overline{n}-1} \left(\frac{j}{\overline{n}}\right)^2 \mathrm{E}\left[\left|\sigma_{t_{i+\overline{n}-j}}^2 - \sigma_{t_i}^2\right|\right] \frac{f_n^2(t_{i-1}-\overline{t})}{f_n(0)}$$

$$\leq K(\overline{\Delta}_n \overline{n})^{2\Gamma} \to 0$$

Now consider the d term. Write $d = \sum_{i=\overline{n}}^{n} \delta_{n,i}$ and, using Cauchy-Schwartz, BDG and the boundedness of σ_t,

$$\sum_{i=\overline{n}}^{n} \mathrm{E}_{i-1}[\delta_{n,i}^2] = 64 \sum_{i=\overline{n}}^{n} f_n^4(t_{i-1}-\overline{t})\sigma_{t_{i-1}}^4 \overline{\Delta}_n^2 \cdot$$

$$\cdot \mathrm{E}_{i-1}\left[\left(\sum_{j>k\geq 0}^{\overline{n}-1} \left(1-\frac{j}{\overline{n}}\right)\left(1-\frac{k}{\overline{n}}\right)(Y_{i-j}-Y_{i-j-1})(Y_{i-k}-Y_{i-k-1})\right)^2\right]$$

$$\leq K \sum_{i=\overline{n}}^{n} f_n^4(t_{i-1}-\overline{t})\overline{\Delta}_n^2 \mathrm{E}_{i-1}\left[\left(\sum_{j\geq 0}^{\overline{n}-1}\left(1-\frac{j}{\overline{n}}\right)(Y_{i-j}-Y_{i-j-1})\right)^2\right]\cdot$$

$$\cdot \mathrm{E}_{i-1}\left[\left(\sum_{k\geq 0}^{\overline{n}-1}\left(1-\frac{k}{\overline{n}}\right)(Y_{i-k}-Y_{i-k-1})\right)^2\right]$$

$$\leq 64K \sum_{i=\overline{n}}^{n} f_n^4(t_{i-1}-\overline{t})\overline{\Delta}_n^2 \left(\sum_{j\geq 0}^{\overline{n}-1}\left(1-\frac{j}{\overline{n}}\right)^2 \overline{\Delta}_n\right)\cdot$$

$$\cdot \left(\sum_{k\geq 0}^{\overline{n}-1}\left(1-\frac{k}{\overline{n}}\right)^2 \overline{\Delta}_n\right)$$

$$\leq 64K \sum_{i=\overline{n}}^{n} f_n^4(t_{i-1}-\overline{t})\overline{\Delta}_n^4 \overline{n}^2 = o(f_n^2(0)\overline{\Delta}_n^3 \overline{n}^2) \to 0$$

since $(1-j/\overline{n})^2 \leq 1$. This proves that d is negligible with respect to c. Finally, using the same technique:

$$\sum_{i=\overline{n}}^{n} \mathrm{E}_{i-1}[c_{n,i}^4] = 16 \sum_{i=\overline{n}}^{n} f_n^4(t_{i-1}-\overline{t})\mathrm{E}_{i-1}[(Y_i-Y_{i-1})^4]\left(\sum_{j=1}^{\overline{n}-1}\left(1-\frac{j}{\overline{n}}\right)(Y_{i-j}-Y_{i-j-1})\right)^4$$

$$\leq 16K \sum_{i=\overline{n}}^{n} f_n^4(t_{i-1}-\overline{t})\overline{\Delta}_n^2 \left(\sum_{j=1}^{\overline{n}-1}\left(1-\frac{j}{\overline{n}}\right)(Y_{i-j}-Y_{i-j-1})\right)^4$$

and the expected value of the last term is limited, using again BDG, by

$$16K \sum_{i=\overline{n}}^{n} f_n^4(t_{i-1}-\overline{t})\overline{\Delta}_n^4 \overline{n}^2 = o(f_n(0)^2 \overline{\Delta}_n^3 \overline{n}^2) \to 0,$$

implying that also the limit in probability is 0.

Finally consider $\frac{1}{\sqrt{\overline{n}f_n(0)\overline{\Delta}_n}}\sum_{i=\overline{n}}^{n-1}\mathrm{E}_{i-1}[b_{n,i}\Delta H_i]$ where $H = W$ or $H = B$ as in the proof of Theorem 2.2. When $H = B$, :

$$\sum_{i=\overline{n}}^{n-1}\mathrm{E}_{i-1}[b_{n,i}\Delta B_i] = \sum_{i=\overline{n}}^{n}f_n(t_{i-1}-\overline{t})\mathrm{E}_{i-1}\left[(Y_i-Y_{i-1})\Delta B_i\right]\sum_{j=1}^{\overline{n}-1}\left(1-\frac{j}{n}\right)(Y_{i-j}-Y_{i-j-1})$$

and this converges to zero since, using the same notations as in the proof of Theorem 2.2 $(Y_i - Y_{i-1}) = \Delta M_i$ and $\mathrm{E}_{i-1}[\Delta M_i \Delta B_i] = 0$. When $H = W$, using the same technique as above,:

$$\frac{1}{\sqrt{\overline{n}f_n(0)\overline{\Delta}_n}}\sum_{i=\overline{n}}^{n-1}\mathrm{E}_{i-1}[b_{n,i}\Delta W_i] \leq \sqrt{\frac{1}{\overline{n}f_n(0)\overline{\Delta}_n}\sum_{i=\overline{n}}^{n-1}\mathrm{E}_{i-1}[b_{n,i}^2]\mathrm{E}_{i-1}[(\Delta W_i)^2]}$$

$$= \sqrt{\frac{1}{\overline{n}f_n(0)\overline{\Delta}_n}\sum_{i=\overline{n}}^{n-1}f_n^2(t_{i-1}-\overline{t})\mathrm{E}_{i-1}\left[(Y_i-Y_{i-1})^2\right]\left(\sum_{j=1}^{\overline{n}-1}\left(1-\frac{j}{n}\right)(Y_{i-j}-Y_{i-j-1})\right)^2\overline{\Delta}_n}$$

$$= \sqrt{\frac{1}{f_n(0)}\sum_{i=\overline{n}}^{n-1}f_n^2(t_{i-1}-\overline{t})O_p(\overline{\Delta}_n^4)} \to 0$$

Altogether, this proves that:

$$\frac{\left(A_1-(\sigma^2)_{\overline{t}}^{\star}\right)}{\sqrt{\overline{n}f_n(0)\overline{\Delta}_n}} \longrightarrow \mathcal{MN}\left(0,\frac{4}{3}c_f(\sigma^4)_{\overline{t}}^{\star}\right)$$

stably in law.

Next, consider B_1. By Theorem 2.2, $B_1 = O_p(f_n(0)\overline{\Delta}_n/\overline{n})$ and can be neglected with respect to A_1. Consider now C_2 and C_3. Start with $C_2 := \sum_{i=1}^{n-\overline{n}+1}\alpha_{n,i}$ where

$$\alpha_{n,i} = \frac{1}{n}f_n(\overline{t}-t_{i-1})\left((\epsilon_{i+\overline{n}-1}-\epsilon_{i-1})^2-(\epsilon_i-\epsilon_{i-1})^2\right)$$

Using the fact the the ϵs are iid, we have $\mathrm{E}[\alpha_{n,i}] = 0$ and:

$$\sum_{i=1}^{n-\overline{n}+1}\mathrm{E}[\alpha_{n,i}^2] = \frac{f_n(0)}{\overline{n}^2\overline{\Delta}_n}\sum_{i=1}^{n-\overline{n}+1}\frac{f_n(\overline{t}-t_{i-1})^2}{f_n(0)}\overline{\Delta}_n\left(6V_\epsilon^2+2\kappa_\epsilon\right)$$

which implies, provided $\sum_{i=1}^{n-\overline{n}+1}\mathrm{E}[\alpha_{n,i}^4]\xrightarrow{p}0$ which is readily proved, that

$$\overline{n}\sqrt{\frac{\overline{\Delta}_n}{f_n(0)}}\cdot C_2 \longrightarrow \mathcal{N}\left(0,c_f\left(6V_\epsilon^2+2\kappa_\epsilon\right)\right).$$

We next have $C_3 := \sum_{i=1}^{n-\overline{n}+1}\beta_{n,i}$ where

$$\beta_{n,i} = \frac{2}{n}f_n(\overline{t}-t_{i-1})\left((\epsilon_{i+\overline{n}-1}-\epsilon_{i-1})(Y_{i+\overline{n}-1}-Y_{i-1})-(\epsilon_i-\epsilon_{i-1})(Y_i-Y_{i-1})\right)$$

Again, $\mathrm{E}_{i-1}[\beta_{n,i}] = 0$ and, using the same techniques above and the law of large numbers in the last step:

$$\sum_{i=1}^{n-\overline{n}+1}\mathrm{E}_{i-1}[\beta_{n,i}^2] = \frac{4}{\overline{n}^2}\sum_{i=1}^{n-\overline{n}+1}f_n^2(\overline{t}-t_{i-1})\left[(V_\epsilon+\epsilon_{i-1}^2)\mathrm{E}_{i-1}[(Y_{i+\overline{n}-1}-Y_{i-1})^2]\right.$$

33

$$+ (V_\epsilon + \epsilon_{i-1}^2)\mathrm{E}_{i-1}[(Y_i - Y_{i-1})^2] + 2\epsilon_{i-1}^2\mathrm{E}_{i-1}[(Y_{i+\overline{n}-1} - Y_{i-1})(Y_i - Y_{i-1})]]$$

$$= \frac{4f_n(0)}{\overline{n}^2}\sum_{i=1}^{n-\overline{n}+1}\frac{f_n^2(\overline{t} - t_{i-1})}{f_n(0)}\left[(V_\epsilon + \epsilon_{i-1}^2)\sigma_{t_{i-1}}^2\overline{n}\overline{\Delta}_n\right.$$

$$\left. + (V_\epsilon + \epsilon_{i-1}^2)\sigma_{t_{i-1}}^2\overline{\Delta}_n + 2\epsilon_{i-1}^2\sigma_{t_{i-1}}^2\overline{\Delta}_n\right] + o_p(1)$$

$$= O_p(f_n(0)/\overline{n})$$

which implies that C_2 is the leading term.

Thus, the dominating terms are $(A_1 - (\sigma^2)_{\overline{t}}^*) = O_p((\overline{n}f_n(0)\overline{\Delta}_n)^{1/2})$ and $C_2 = O_p(f_n(0)^{1/2}(\overline{\Delta}_n)^{-1/2}(\overline{n})^{-1})$. The two asymptotic rates coincide when $\overline{n} \sim \overline{\Delta}_n^{-2/3}$ as stated, leading to the desired result. $\qquad\square$

Theorem 3.3. Denote by $X = Y + J$ where Y is a continuous semimartingale. By virtue of Theorem 1 in Mancini [35], for n large enough, we can write, almost surely,

$$\widehat{\phi}_{n,f}^2(\overline{t}) = \sum_{i=1}^n f_n(t_{i-1} - \overline{t})(\Delta Y_i)^2 - \sum_{i=1}^n f_n(t_{i-1} - \overline{t})(\Delta Y_i)^2 I_{\{\Delta N_i \neq 0\}}.$$

Theorem 2.2 can be applied to the first term, while the second term is $O_p(N_T\overline{\Delta}_n f_n(0))$, or equivalently, $o_p\left(\sqrt{\overline{\Delta}_n f_n(0)}\right)$, where N_T is the Poisson counting process and is vanishing in the limit. $\qquad\square$

Lemma A.2. *Result (A.6) continues to hold also when $f_n(x)$ is given by the Fejér sequence.*

Proof. Firstly, remark that for any bounded process A we have

$$\sum_{j=1}^n \int_{t_{j-1}}^{t_j} \left|f_n(s - \overline{t}) - f_n(t_{j-1} - \overline{t})\right| |A_s| ds \leq \tag{A.19}$$

$$C\sum_{j=1}^n \int_{t_{j-1}}^{t_j} \left|\sum_{|k|\leq N}\left(1 - \frac{|k|}{N}\right)\left(e^{ik(s-\overline{t})} - e^{ik(t_{j-1}-\overline{t})}\right)\right| ds$$

$$\sim C\sum_{j=1}^n \int_{t_{j-1}}^{t_j} \left|\sum_{|k|\leq N}\left(1 - \frac{|k|}{N}\right)e^{ik(s-\overline{t})}ik\Delta_j\right| ds = C\overline{\Delta}_n\sum_{j=1}^n \int_{t_{j-1}}^{t_j} f_n(s-\overline{t})Nds \leq CN\overline{\Delta}_n,$$

where $\sim$ denotes asymptotic equivalence. It follows from $\sup_x |f_n(x)| \leq f_n(0)$ that

$$\frac{1}{f_n(0)}\left|\sum_{j=1}^n f_n^2(t_{j-1} - \overline{t})\int_{t_{j-1}}^{t_j} A_s ds - \int_0^T f_n^2(s - \overline{t})A_s ds\right| = \tag{A.20}$$

$$\frac{1}{f_n(0)}\left|\sum_{j=1}^n \int_{t_{j-1}}^{t_j}(f_n^2(t_{j-1} - \overline{t}) - f_n^2(s - \overline{t}))A_s ds\right| =$$

$$\frac{1}{f_n(0)}\left|\sum_{j=1}^n \int_{t_{j-1}}^{t_j}(f_n(t_{j-1} - \overline{t}) - f_n(s - \overline{t}))(f_n(t_{j-1} - \overline{t}) + f_n(s - \overline{t}))A_s ds\right| \leq$$

$$C\sum_{j=1}^n \int_{t_{j-1}}^{t_j}|f_n(t_{j-1} - \overline{t}) - f_n(s - \overline{t})||A_s| ds \leq CN\overline{\Delta}_n.$$

Therefore, for the Fejér sequence in place of (A.4) we will make use of

$$\frac{1}{f_n(0)}\left|\sum_{j=1}^n f_n^2(t_{j-1}-\bar t)\int_{t_{j-1}}^{t_j} A_s ds - \int_0^T f_n^2(s-\bar t)A_s ds\right| \le CN\overline{\Delta}_n \qquad \text{(A.21)}$$

Secondly, for the Fejér sequence there exists a sequence ε_n of positive numbers s.t. $\varepsilon_n^2 N \to \infty$ and $\int_{|x|>\varepsilon_n} f_n(x)dx \to 0$. As $f_n \ge 0$, $\int_{|x|\le\varepsilon_n} f_n(x)dx \to 1, \sup_x |f_n(x)| \le Cf_n(0)$ and $\int_{|x|\le\varepsilon_n} f_n^2(x)/f_n(0)dx \to c_f$, by Proposition 2.1, we obtain that (2.7) holds true. In order to show (A.6) we check that the following tends to 0 in probability:

$$\frac{1}{f_n(0)}\sum_{i=1}^n f_n^2(t_{i-1}-\bar t)A(t_{i-1})\frac{\Delta_i^2}{\overline{\Delta}_n} - c_f(H'A)_{\bar t}^\star =$$

$$\frac{1}{f_n(0)}\sum_{i=1}^n f_n^2(t_{i-1}-\bar t)A(t_{i-1})\frac{\Delta_i^2}{\overline{\Delta}_n} - \frac{1}{f_n(0)}\int_0^T f_n^2(s-\bar t)(H'A)_s ds + o_P(1) =$$

$$A_n + B_n + o_P(1),$$

where

$$A_n = \frac{1}{f_n(0)}\int_0^T \sum_i [f_n^2(s-\bar t)-f_n^2(t_{i-1}-\bar t)]I_{]t_{i-1},t_i]}(s)(H'A)_s ds,$$

$$B_n = \frac{1}{f_n(0)}\int_0^T \sum_i f_n^2(t_{i-1}-\bar t)[(H'A)_s - A_{t_{i-1}}\Delta_i/\overline{\Delta}_n]I_{]t_{i-1},t_i]}(s)ds.$$

A_n coincides with A.20 and tends a.s. to zero. As for B_n, split it as $B_n = B_{1,n}+B_{2,n}$ where

$$B_{1,n} = \frac{1}{f_n(0)}\sum_i \int_{t_{i-1}}^{t_i} f_n^2(t_{i-1}-\bar t)H'_s[A_s - A_{t_{i-1}}]ds,$$

$$B_{2,n} = \frac{1}{f_n(0)}\sum_i \int_{t_{i-1}}^{t_i} f_n^2(t_{i-1}-\bar t)[(H'_s - \Delta_i/\overline{\Delta}_n]A_{t_{i-1}}ds.$$

For $B_{2,n}$ use that: $\frac{f_n^2(t_{i-1}-\bar t)}{f_n(0)} \le |f_n(t_{i-1}-\bar t)| = f_n(t_{i-1}-\bar t)$; assumption (2.4) which implies that, for all $i = 1..n$, $\sup_{s\in]t_{i-1},t_i]}|H'(s) - \Delta_i/\overline{\Delta}_n| \le K\overline{\Delta}_n$; the boundedness of A and note that from A.19 we have $\sum_i f_n(t_{i-1}-\bar t)\Delta_i = \int_0^T f_n(s-\bar t)ds + O_P(N\overline{\Delta}_n) \to 1$, so

$$|B_{2,n}| \le \sum_i \int_{t_{i-1}}^{t_i} f_n(t_{i-1}-\bar t)C\Delta_i ds \le \overline{\Delta}_n \sum_i f_n(t_{i-1}-\bar t)\Delta_i \to 0.$$

For $B_{1,n}$ we proceed analogously as from (A.8) to the end of the proof of ii) by replacing (A.4) with (A.21), and using for the last inequality that $f_n \ge 0$ we end up with $\overline{\Delta}_n^\Gamma \int_0^T f_n(s-\bar t)ds \to 0$.

$\square$

Proposition 4.1. As above, it is harmless to assume that $\mu = 0$. Write,

$$\sigma_{n,F}^2(\bar t) - (\sigma^2)_{\bar t}^\star = TE + DE + ME$$

where

$$TE = \text{truncation error} = \int_0^T f_n(s-\bar t)\sigma_s^2 ds - (\sigma^2)_{\bar t}^\star,$$

$$DE = \text{discretization error} = \sum_{i=1}^{n} f_n(t_{i-1} - \bar{t}) \int_{t_{i-1}}^{t_i} \sigma_s^2 ds - \int_0^T f_n(s - \bar{t})\sigma_s^2 ds,$$

$$ME = \text{martingale error} = \sum_{i=1}^{n} f_n(t_{i-1} - \bar{t})(\Delta X_i)^2 - \sum_{i=1}^{n} f_n(t_{i-1} - \bar{t}) \int_{t_{i-1}}^{t_i} \sigma_s^2 ds.$$

For the martingale error, we do the same as in the proof of Theorem 2.2 to show that

$$\frac{1}{\overline{\Delta}_n f_n(0)} \left(ME - (\sigma^2)_{\bar{t}}^\star \right) \to \mathbf{MN}\left(0, \frac{4}{3}(\sigma^4 H')_{\bar{t}}^\star \right),$$

having also used Lemma A.2.

For the term DE, by (A.21) we have, for a suitable constant C,

$$|DE| \leq CN\overline{\Delta}_n,$$

thus $\sqrt{\frac{n}{N}}DE \to 0$ and DE is negligible.

For the term TE, with no loss of generality we can take $T = \pi$, so that we have $\bar{t} \in (0, \pi)$. Note that, in the case $\Gamma = 1$, for the Fejér sequence choosing e.g. $\varepsilon_n = N^{-\alpha}$ and $N = n^\delta$ with $\alpha > 0$ and $\delta \in (1/2, 1)$ sufficiently close to 1, fulfills the requested requirements for (n, N, ε_n), which also imply the requirements in Proposition (2.1). For large n we then have

$$TE = R_n^{(\sigma^2)} = \int_0^T f_n(s - \bar{t})\sigma_s^2 ds - (\sigma^2)_{\bar{t}}^\star$$

$$= \int_{-\bar{t}}^{\pi-\bar{t}} f_n(x)\sigma_{x+\bar{t}}^2 dx - \int_{-\pi}^{\pi} f_n(x)(\sigma^2)_{\bar{t}}^\star dx$$

$$= TE_1 + TE_2 - TE_3,$$

where

$$TE_1 = \int_{|x|\leq\varepsilon_n} f_n(x)(\sigma_{x+\bar{t}}^2 - (\sigma^2)_{\bar{t}}^\star)dx,$$

$$TE_2 = \int_{\{|x|>\varepsilon_n\}\cap(-\bar{t},\pi-\bar{t})} f_n(x)(\sigma_{x+\bar{t}}^2 - (\sigma^2)_{\bar{t}}^\star)dx,$$

$$TE_3 = \int_{(-\pi,-\bar{t})\cup(\pi-\bar{t},\pi)} f_n(x)(\sigma^2)_{\bar{t}}^\star dx.$$

For large n, $TE_2, TE_3 \leq \int_{|x|>\varepsilon_n} f_n(x)dx$ by the boundedness of σ and it turns out that both TE_2 and TE_3 are $O(1/(N\varepsilon_n^2))$, so that $\sqrt{\frac{n}{N}}(TE_2 - TE_3) = O(\sqrt{n/(N^3\varepsilon_n^4)}) \to 0$. As for TE_1, we take n such that $\varepsilon_n < \varepsilon$, with ε as in (2.2), so that for $|x| \leq \varepsilon_n$ we have $E[|\sigma_{x+\bar{t}}^2 - (\sigma^2)_{\bar{t}}^\star|] \leq C\varepsilon_n^{\Gamma/2}$ and thus

$$\sqrt{\frac{n}{N}}E[|TE_1|] \leq \sqrt{\frac{n}{N}} \int_{|x|\leq\varepsilon_n} f_n(x)E[|\sigma_{x+\bar{t}}^2 - (\sigma^2)_{\bar{t}}^\star|]dx \leq \sqrt{\frac{n\varepsilon_n^\Gamma}{N}} \to 0,$$

which shows that the whole TE is negligible.

$\square$

Proposition 4.2. From Eq. (4.1)-(4.3) we have

$$\widehat{\sigma}_{n,n',N}^{2,F}(\bar{t}) = \sum_{|k|\leq N} \left(1 - \frac{|k|}{N} \right) \left(\frac{T}{2n'+1} \sum_{|s|\leq n'} \mathcal{F}_n(dX)(s)\mathcal{F}_n(dX)(k-s) \right) e^{ik\tau}$$

$$= \frac{1}{T} \sum_{|k| \leq N} \left(1 - \frac{|k|}{N}\right) \left(\frac{1}{2n'+1} \sum_{|s| \leq n'} \sum_{j'=1}^{n} \sum_{j=1}^{n} e^{-is\tau_{j'-1}} e^{-i(k-s)\tau_{j-1}} \Delta X_j \Delta X_{j'}\right) e^{ik\tau}$$

$$= \frac{1}{T} \sum_{|k| \leq N} \left(1 - \frac{|k|}{N}\right) \left[\frac{1}{2n'+1} \sum_{|s| \leq n'} \left(\sum_{j=1}^{n} e^{-ik\tau_{j-1}}(\Delta X_j)^2 + \right.\right.$$

$$\left.\left. + \sum_{j' \neq j=1}^{n} e^{-is\tau_{j'-1}} e^{-i(k-s)\tau_{j-1}} \Delta X_{j'} \Delta X_j\right)\right] e^{ik\tau}$$

$$= \frac{1}{T} \sum_{j=1}^{n} \sum_{|k| \leq N} \left(1 - \frac{|k|}{N}\right) e^{ik(\tau - \tau_{j-1})} (\Delta X_j)^2$$

$$+ \frac{1}{T(2n'+1)} \sum_{j' \neq j=1}^{n} \sum_{|k| \leq N} \left(1 - \frac{|k|}{N}\right) e^{ik(\tau - \tau_{j-1})} \sum_{|s| \leq n'} e^{is(\tau_{j-1} - \tau_{j'-1})} \Delta X_{j'} \Delta X_j.$$

By the definitions of Dirichlet and Fejér sequences as given in equations (2.14) and (2.15), we then get the main statement of the Proposition. □

References

[1] Aït-Sahalia, Y., Jacod, J.: Testing for jumps in a discretely observed process. Annals of Statistics **37**, 184–222 (2009)

[2] Alvarez, A., Panloup, F., Pontier, M., Savy, N.: Estimation of the instantaneous volatility (2009). Working paper

[3] Andersen, T., Bollerslev, T.: Intraday periodicity and volatility persistence in financial markets. Journal of Empirical Finance **4**, 115–158 (1997)

[4] Andreou, E., Ghysels, E.: The impact of sampling frequency and volatility estimators on change-point tests. Journal of Financial Econometrics **2**, 290–318 (2004)

[5] Bandi, F., Nguyen, T.: On the functional estimation of jump-diffusion models. Journal of Econometrics **116**, 293–328 (2003)

[6] Bandi, F., Phillips, P.: Fully nonparametric estimation of scalar diffusion models. Econometrica **71**(1), 241–283 (2003)

[7] Bandi, F., Renò, R.: Nonparametric stochastic volatility (2009). Working Paper

[8] Bandi, F., Renò, R.: Price and volatility co-jumps (2011). Working paper

[9] Bandi, F., Russell, J.: Volatility. In: Handbook of Financial Engineering, pp. 183–222. Elsevier Science (2007)

[10] Barndorff-Nielsen, O., Hansen, P., Lunde, A., Shephard, N.: Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. Econometrica **76**(6), 1481–1536 (2008)

[11] Barndorff-Nielsen, O.E., Shephard, N.: Variation, jumps, market frictions and high frequency data in financial econometrics. In: Advances in Economics and Econometrics. Theory and Applications, Ninth World Congress, pp. 328–372. Cambridge University Press (2007)

[12] Boudt, K., Croux, C., Laurent, S.: Outlyingness weighted quadratic covariation. Journal of Financial Econometrics **9**, 657–684 (2011)

[13] Clément, E., Gloter, A.: Limit theorems in the Fourier transform method for the estimation of multivariate volatility. Stochastic Processes and their Applications **121**, 1097–1124 (2011)

[14] Corsi, F., Pirino, D., Renò, R.: Threshold bipower variation and the impact of jumps on volatility forecasting. Journal of Econometrics **159**, 276–288 (2010)

[15] Dobrev, D., Andersen, T., Schaumburg, E.: Duration-based volatility estimation (2008). Working Paper.

[16] van Eeden, C.: Mean integrated squared error of kernel estimators when the density and its derivative are not necessarily continuous. Annals of the Institute of Statistical Mathematics **37**(1), 461–472 (1985)

[17] Fan, J., Wang, Y.: Spot volatility estimation for high-frequency data. Statistics and Its Interface **1**, 279–288 (2008)

[18] Fan, J., Yao, Q.: Nonlinear time series. Springer (2003)

[19] Florens-Zmirou, D.: On estimating the diffusion coefficient from discrete observations. Journal of Applied Probability **30**, 790–804 (1993)

[20] Foster, D., Nelson, D.: Continuous record asymptotics for rolling sample variance estimators. Econometrica **64**(1), 139–174 (1996)

[21] Genon-Catalot, V., Laredo, C., Picard, D.: Non-paramteric estimation of the diffusion coefficient by wavelets methods. Scandinavian Journal of Statistics **19**, 317–335 (1992)

[22] Hoffman, M., Munk, A., Schmidt-Hieber, J.: Adaptive wavelet estimation of the diffusion coefficient under additive error measurements. Annals of the Institute H. Poincaré (2012). Forthcoming

[23] Hoffmann, M.: L_p estimation of the diffusion coefficient. Bernoulli **5**(3), 447–481 (1999)

[24] Hoffmann, M.: Rate of convergence for parametric estimation in a stochastic volatility model. Stochastic Processes and Applications **97**, 147–170 (2002)

[25] Jacod, J.: On continuous conditional Gaussian martingales and stable convergence in law, Séminaire Proba. XXXI. Springer-Verlag, Berlin, Lecture Notes in Math **1655**, 232–246 (1997)

[26] Jacod, J.: Statistics and high-frequency data (2007). Lecture notes of SEMSTAT course in La Manga. To appear.

[27] Jacod, J., Li, Y., Mykland, P., Podolskij, M., Vetter, M.: Microstructure noise in the continuous case: the pre-averaging approach. Stochastic Processes and their Applications **119**(7), 2249–2276 (2009)

[28] Jacod, J., Protter, P.: Discretization of Processes. Springer-Verlag (2012)

[29] Jacod, J., Shiryaev, A.N.: Limit Theorems for Stochastic Processes. Springer (2003)

[30] Johannes, M.: The statistical and economic role of jumps in continuous-time interest rate models. Journal of Finance **59**, 227–260 (2004)

[31] Kanatani, T.: Integrated volatility measuring from unevenly sampled observations. Economics Bulletin **3**(36), 1–8 (2004)

[32] Kristensen, D.: Nonparametric filtering of the realised spot volatility: a kernel-based approach. Econometric Theory **26**, 60–93 (2010)

[33] Malliavin, P., Mancino, M.: Fourier series method for measurement of multivariate volatilities. Finance & Stochastics **6**(1), 49–61 (2002)

[34] Malliavin, P., Mancino, M.: A Fourier transform method for nonparametric estimation of volatility. Annals of Statistics **37**(4), 1983–2010 (2009)

[35] Mancini, C.: Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. Scandinavian Journal of Statistics **36**(2), 270–296 (2009)

[36] Mancini, C., Renò, R.: Threshold estimation of Markov models with jumps and interest rate modeling. Journal of Econometrics **160**(1), 77–92 (2011)

[37] Mancino, M., Sanfelici, S.: Robustness of Fourier estimator of integrated volatility in the presence of microstructure noise. Computational Statistics and Data Analysis **52**(6), 2966–2989 (2008)

[38] Mancino, M., Sanfelici, S.: Estimating covariance via Fourier method in the presence of asynchronous trading and microstructure noise. Journal of Financial Econometrics **9**(2), 367–408 (2011)

[39] Mykland, P., Zhang, L.: ANOVA for diffusions and Ito processes. Annals of Statistics **34**(4), 1931 (2006)

[40] Mykland, P., Zhang, L.: Inference for volatility-type objects and implications for hedging. Statis-

tics and Its Interface **1**, 255–278 (2008)

[41] Ngo, H.L., Ogawa, S.: A central limit theorem for the functional estimation of the spot volatility. Monte Carlo method and applications **14**(4), 353–380 (2009)

[42] Ogawa, S., Sanfelici, S.: An improved two-step regularization scheme for spot volatility estimation. Economic Notes **40**(3), 107–134 (2011)

[43] Podolskij, M., Vetter, M.: Bipower-type estimation in a noisy diffusion setting. Stochastic Processes and their Applications **119**, 2803–2831 (2009)

[44] Podolskij, M., Vetter, M.: Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. Bernoulli **15**(3), 634–668 (2009)

[45] Priestley, M.: Spectral Time Series Analysis. Wiley (1979)

[46] Protter, P.: Stochastic Integration and Differential Equations. Springer-Verlag (1990)

[47] Renò, R.: Nonparametric estimation of the diffusion coefficient of stochastic volatility models. Econometric Theory **24**(5), 1174–1206 (2008)

[48] Veraart, A.: Inference for the jump part of quadratic variation of Ito semimartingales. Econometric Theory **26**(2), 331–368 (2010)

[49] Walter, G., Blum, J.: Probability density estimation using delta sequences. The Annals of Statistics **6**(2), 328–340 (1979)

[50] Watson, G., Leadbetter, M.: Hazard analysis II. Sankhya, Ser. A **26**, 101–116 (1964)

[51] Zhang, L.: Efficient estimation of stochastic volatility using noisy observations : A multi-scale approach. Bernoulli **12**(6), 1019–1043 (2006)

[52] Zhang, L., Mykland, P.A., Aït-Sahalia, Y.: A tale of two time scales: Determining integrated volatility with noisy high-frequency data. Journal of the American Statistical Association **100**, 1394–1411 (2005)

[53] Zu, Y., Boswijk, P.: Estimating spot volatility with high frequency financial data (2010). Working Paper