Econometric Theory, **42**, 2026, 1272–1302.
doi:10.1017/S0266466625100170

ESTIMATION OF INTEGRATED VOLATILITY FUNCTIONALS WITH KERNEL SPOT VOLATILITY ESTIMATORS

JOSÉ E. FIGUEROA-LÓPEZ
Washington University in St. Louis

JINCHENG PANG
Washington University in St. Louis

BEI WU
Upstart

For a multidimensional Itô semimartingale, we consider the problem of estimating integrated volatility functionals. Jacod and Rosenbaum (2013, *The Annals of Statistics* 41(3), 1462–1484) studied a plug-in type of estimator based on a Riemann sum approximation of the integrated functional and a spot volatility estimator with a forward uniform kernel. Motivated by recent results that show that spot volatility estimators with general two-sided kernels of unbounded support are more accurate, in this article, an estimator using a general kernel spot volatility estimator as the plug-in is considered. A biased central limit theorem for estimating the integrated functional is established with an optimal convergence rate. Central limit theorems for properly de-biased estimators are also obtained both at the optimal convergence regime for the bandwidth and when applying undersmoothing. Our results show that one can significantly reduce the estimator's bias by adopting a general kernel instead of the standard uniform kernel. Our proposed bias-corrected estimators are found to maintain remarkable robustness against bandwidth selection in a variety of sampling frequencies and functions.

1. INTRODUCTION

In this work, we are concerned with the estimation of integrated volatility (IV) functionals of the form $V_T(g) := \int_0^T g(c_s)ds$, where g is a smooth function and $c_t := \sigma_t\sigma_t^*$ is the spot variance–covariance process of a d-dimensional Itô semimartingale given by

$$dX_t = \mu_t dt + \sigma_t dW_t + dJ_t. \tag{1.1}$$

The authors are truly grateful to both Co-Editors (Dr. Dennis Kristensen and Dr. Peter C. B. Phillips), and two anonymous referees for the numerous suggestions that helped to significantly improve the original manuscript. Research supported in part by the NSF grant: DMS-2413557. Address correspondence to José E. Figueroa-López, Department of Statistics and Data Science, Washington University in St. Louis, St. Louis, MO, USA, e-mail: figueroa-lopez@wustl.edu.

© The Author(s), 2025. Published by Cambridge University Press.

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

Here, $\{W_t\}_{t\geq 0}$ is d-dimensional standard Brownian motion (BM) and $\{J_t\}_{t\geq 0}$ is a pure-jump process of bounded variation. Such functionals have many applications in financial econometrics. In the simplest case, when $d=1$ and $g(x)=x$, this is the well-studied IV or variance $IV_T=\int_0^T \sigma_s^2 ds$, which measures the overall variability or uncertainty latent in the process X during $[0,T]$. Again, when $d=1$ and $g(x)=x^2$, $V_T(g)$ corresponds to the integrated quarticity, which appears in the limiting distribution of the realized variance estimator $\widehat{IV}_T=\sum_{i=1}^n(\Delta_i^n X)^2$ and, hence, it is crucial to construct feasible confidence intervals for IV. Another function appearing in the literature is $g(x)=\log(x)$, when researching the statistical features of a factor Y driving the volatility in a model where σ is assumed to be the exponential of Y. Volatility functionals have also been used in principal component analysis by Aït-Sahalia and Xiu (2019) and in estimating integrated moments for option pricing purposes by Li and Xiu (2016).

One of the fundamental methods to estimate $V_T(g)=\int_0^T g(c_s)ds$ consists of approximating the integral by a Riemman sum and plugging in a suitable spot volatility estimator $\hat{c}$ in place of c:

$$\widehat{V}_T(g)=\Delta_n\sum_{i=1}^{n-k_n+1} g\left(\hat{c}_i^n\right). \tag{1.2}$$

This method was pioneered by Jacod and Rosenbaum (2013), where, as the spot volatility estimator, a local rolling window estimator with uniform weights was used:

$$\hat{c}_i^{n,unif}=\frac{1}{k_n\Delta_n}\sum_{j=0}^{k_n-1}\left(\Delta_{i+j}^n X\right)^2. \tag{1.3}$$

Two interesting facts emerged from their theory. For ease of exposition, we focus on the $d=1$-case in the introduction, though our main results are stated for general dimension $d\geq 1$. First of all, the rate of convergence of the estimation error is unexpectedly of order $\Delta_n^{\frac{1}{2}}$ even though $\hat{c}_i^n$ can only attain the rate $\Delta_n^{\frac{1}{4}}$.[1] Secondly, when $k_n\sim\theta/\sqrt{\Delta_n}$ for some $\theta\in(0,\infty)$, which yields the best convergence rate for the spot volatility estimator $\hat{c}_i^n$, $\frac{1}{\sqrt{\Delta_n}}\left(\widehat{V}_T(g)-V_T(g)\right)$ converges to a mixed Gaussian distribution with four bias terms:

1. a term of the form $-\frac{\theta}{2}\left(g(c_0)+g(c_T)\right)$ due to border effects;
2. a second term of the form $\frac{1}{\theta}\int_0^T g''(c_s)c_s^2 ds$ due to the nonlinearity of g;
3. a third term of the form $-\frac{\theta}{12}\int_0^T g''(\tilde{c}_s)ds$, where $\tilde{c}$ is the volatility of volatility, due to the target error $c_{t_i}-\bar{c}_i^n$, where $\bar{c}_i^n=\frac{1}{k_n\Delta_n}\sum_{j=0}^{k_n-1}c_{t_{i+j}}$;

[1] Intuitively, by Taylor expansion of the estimator $g(\hat{c})$, we can write $\Delta_n\sum_{i=1}^n g(\hat{c}_i^n)-\Delta_n\sum_{i=1}^n g(c_{t_{i-1}})=\Delta_n\sum_{i=1}^n g'(c_{t_{i-1}})\left(\hat{c}_i^n-c_{t_{i-1}}\right)+\frac{1}{2}\Delta_n\sum_{i=1}^n g''(c_{t_{i-1}})\left(\hat{c}_i^n-c_{t_{i-1}}\right)^2+$ higher order terms. The estimation errors in the first term of the right-hand side are "averaged out" so that the convergence rate changes from $n^{-1/4}$ to $n^{-1/2}$.

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

4. a fourth term involving the jump component of the volatility process $\theta \sum_{s \le t} G(c_{s-}, c_s)$ for a suitable function of G depending on g. This error disappears when c is continuous.

In principle, the bias terms can be estimated and eliminated from $\widehat{V}_T(g)$ to obtain a feasible CLT with a centered Gaussian limiting distribution. This strategy was explored by Jacod and Rosenbaum (2015). An alternative solution, however, was adopted in Jacod and Rosenbaum (2013). Specifically, the window length k_n was made to converge to ∞ at a slower rate (undersmoothing) so that $k_n\sqrt{\Delta_n} \to 0$. This corresponds to the case $\theta = 0$ and, hence, the first, third, and fourth bias terms would vanish, while the second term becomes dominant. To keep the convergence rate $\sqrt{\Delta_n}$ of the estimator $\widehat{V}_T(g)$, this second term needs to be estimated and the estimator needs to be de-biased. The resulting estimator takes the form

$$\Delta_n \sum_{i=1}^{n-k_n+1} \left(g\left(\hat{c}_i^n\right) - \frac{1}{k_n} g''(\hat{c}_i)\hat{c}_i^2 \right), \tag{1.4}$$

which indeed achieves the rate $\Delta_n^{\frac{1}{2}}$ with semi-parametrically optimal asymptotic variance. A similar approach to deal with edge effects and biases was also considered in Kristensen (2010), who also proposed kernel estimators and obtained a central limit theorem (though the model considered therein lacks leverage effects). Alternatively, as mentioned before, in the optimal case $\theta \in (0, \infty)$ and employing forward uniform kernels, estimators for all the bias terms were proposed in Jacod and Rosenbaum (2015), where an unbiased estimator with centered limiting distribution was developed.

Li, Liu, and Xiu (2019) took a different approach toward bias correction. They proposed a type of "jackknife" estimator, which is formed as a linear combination of a few uncorrected estimators of the form (1.3) associated with different windows k_n. Since these estimators have similar bias terms but different coefficients depending on the window sizes, the biases can be eliminated by a proper linear combination. Both the "two-scale" jackknife (i.e., when combining only two uncorrected estimators) and Jacod and Rosenbaum's (2013) estimator use undersmoothing, so that the biases due to boundary effects, volatility of volatility, and volatility jumps are rendered asymptotically negligible. One drawback of this approach is that one needs to determine the degree of undersmoothing. In other words, if, for instance, we take $k_n \sim \beta\Delta_n^{\alpha}$ with $\alpha > -1/2$, then one needs to tune both parameters β and α, which is nontrivial. In contrast, a multiscale jackknife estimator, formed as a linear combination of three (or more) estimators, can, in principle, cancel all the biases regardless of the speed of convergence of the window size. However, again, this involves more tuning parameters that need to be carefully calibrated.

Other related work on this topic includes Mykland and Zhang (2009), who proposed using the same method as in Jacod and Rosenbaum (2013), but keeping $k_n = k$ constant. For the quarticity or more generally for estimating $\int_0^T g(c_s)ds$

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

with a power functional $g(x) = x^r$, they obtained a CLT with rate $1/\sqrt{\Delta_n}$ and an asymptotic variance that is not efficient but can arbitrarily approach the optimal variance when k is large. Clément, Delattre, and Gloter (2013) established a locally asymptotically mixed normality (LAMN) condition for the estimation of IV functionals in a continuous diffusion process where the volatility is driven by an independent Itô process. The result therein shows that the optimal rate of convergence is $n^{-1/2}$. Another related result in the same direction is Renault, Sarisoy, and Werker (2017). More recently, Li and Liu (2021) generalized Jacod and Rosenbaum's (2013) result to allow a long-memory component, such as an fBM with Hurst parameter $H > 1/2$. They also established a semiparametric efficiency lower bound, generalizing Renault et al.'s (2017) results. Chen (2019) developed an estimator for $\widehat{V}_T(g)$ in the presence of microstructure noise, based on a forward finite difference approximation of the standard pre-averaging estimator of the IV.

Our work aims to apply a general kernel spot volatility estimator to the plug-in estimator of IV functionals. The general kernel estimator (Fan and Wang, 2008; Kristensen, 2010) is defined as

$$\tilde{c}_i^n := \sum_{j=1}^{n} K_{k_n\Delta_n}\left(t_{j-1} - t_i\right)\left(\Delta_j^n X\right)^2, \tag{1.5}$$

where $K_b(x) := K(x/b)/b$ for some bandwidth $b > 0$ and k_n is an appropriate sequence that needs to be calibrated. Of course, one can see (1.3) as a particular case of (1.5) with $K(x) = \mathbf{1}_{[0,1]}(x)$. Kristensen (2010) developed an asymptotic theory for (1.5) when t_i is a fixed time $\tau \in (0, T)$ and showed asymptotic normality to the spot variance c_τ under certain path-wise Hölder continuity and non-leverage conditions on c. See also Fan and Wang (2008) and Mancini, Mattiussi, and Renò (2015) for other related results. All these results applied undersmoothing, which allowed them to neglect the "target error" coming from approximating the spot volatility by a kernel weighted IV.

In this article, we consider the following version of (1.5):

$$\hat{c}_i^n := \frac{\sum_{j=1}^{n} K_{k_n\Delta_n}\left(t_{j-1} - t_i\right)\left(\Delta_j^n X\right)^2}{\bar{K}_n\left(t_i\right)}, \tag{1.6}$$

with $\bar{K}_n\left(t_i\right) := \Delta_n \sum_{j=1}^{n} K_{k_n\Delta_n}\left(t_{j-1} - t_i\right)$, which is known to be more robust than (1.5) against edge effects (see also Kristensen (2010) for the proposal of other edge-robust alternatives, such as a type of local linear estimator). As a result, we can, in principle, consider the whole range $i = 1, \ldots, n$ in (1.2). Furthermore, we find it more convenient for our proofs to consider the estimator

$$\widehat{V}_T(g) = \Delta_n \sum_{i=1}^{n} \bar{K}_n(t_i) g\left(\hat{c}_i^n\right). \tag{1.7}$$

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

The coefficient $\bar{K}_n(t_i)$ attenuates the contribution of $g\left(\hat{c}_i^n\right)$ for t_i near 0 and T and, thus, serves as an additional edge effect correction. Kristensen (2010) also considered a kernel-based estimator similar to (1.7), but applied an edge effect correction, similar to that in Jacod and Rosenbaum (2015), by eliminating the terms corresponding to i's close to 0 and n, and did not consider the adjustment $\bar{K}_n(t_i)$. Their asymptotic theory also only considered the undersmoothing asymptotic regime mentioned above.

The motivation for considering general kernels comes from Figueroa-López and Li (2020) and Figueroa-López and Wu (2024), where it has been shown, both theoretically and by Monte Carlo experiments, that a general two-sided kernel with unbounded support could significantly reduce the variance of the spot volatility estimator $\hat{c}$. More specifically, Figueroa-López and Li (2020) studied the infill asymptotic behavior of the mean-square error (MSE) of (1.5) for continuous Itô semimartingales and showed that, when setting k_n "optimally" (i.e., in order to minimize the leading order terms of the MSE), the double exponential kernel $K_{exp}(x) = .5e^{-|x|}$ minimizes the resulting MSE over all kernels. This result was established under the absence of leverage effects. Figueroa-López and Wu (2024) extended this result, by first proving a CLT for a general two-sided kernel, with optimal convergence rate and in the presence of jumps and leverage effects. It was then shown that exponential kernels minimize the asymptotic variance of the CLT. It is worth noting that, even when applying undersmoothing, the asymptotic variance could be reduced when choosing a kernel of unbounded support. Specifically, when applying undersmoothing, the asymptotic variance of the spot volatility estimator is proportional to $\|K\|^2 = \int_{-\infty}^{\infty} K^2(x)dx$. When K is constrained to have support on $[0,1]$, the optimal kernel is indeed the uniform kernel $K_{unif}(x) = \mathbf{1}_{[0,1]}(x)$ with $\|K_{unif}\|^2 = 1$.[2] However, picking the kernels $.5e^{-|x|}$ and $e^{-x}\mathbf{1}_{x>0}$ will result in estimators whose asymptotic variances are, respectively, a quarter and a half of that obtained by a uniform kernel $K_{unif}(x) = \mathbf{1}_{[0,1]}(x)$. Similar comments apply when considering only backward looking kernels, such as $\mathbf{1}_{[-1,0)}(x)$ and $e^{-|x|}\mathbf{1}_{x<0}$.

A natural question is whether the estimation accuracy gained when estimating the spot volatility by a general kernel transfers into gains in estimating integrated functionals. Intuitively, recalling again the Taylor expansion $g(\hat{c}) - g(c) = g'(c)\left(\hat{c}-c\right) + \frac{1}{2}g''(c)\left(\hat{c}-c\right)^2 + O_p\left(|\hat{c}-c|^3\right)$, and after taking expectations,[3] we can see that the variance of the spot volatility estimator, which appears in the second term, becomes a bias term of the IV estimator. Therefore, the bias of the estimator $\widehat{V}_T(g)$ could, in principle, be reduced by adopting a general kernel.

As it turns out, the CLT of (1.5) again exhibits several biases, which depend on a rather intricate and nontrivial manner on the kernel. The second bias term,

[2]Indeed, by Jensen's inequality, $K_{unif}(x) = \mathbf{1}_{[0,1]}(x)$ satisfies $\int K_{unif}^2(x)dx = 1 = \left(\int K(x)dx\right)^2 \leq \int K^2(x)dx$, which shows that a uniform kernel $K_{unif}(x)$ minimizes the asymptotic variance when applying undersmoothing.

[3]The argument here is just heuristic. In the proofs of our main results, we employ standard localization arguments (see (A.1) in the Appendix) to justify the boundedness of the expectations of all related sample moments or estimators.

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

involving $\frac{1}{\theta}\int_0^T g''(c_s)c_s^2 ds$, contains the L_2 norm of the kernel and, thus, can be reduced because certain kernels of unbounded support possess smaller L_2 norm. We then propose bias-corrected estimators similar to those in Jacod and Rosenbaum (2015), and derive the corresponding CLT. We show by Monte Carlo experiments that our general kernel estimator typically has superior performance compared with the jackknife estimators of Li et al. (2019). Of course, our estimator is simpler as it does not involve as many tuning parameters, which are hard to calibrate. In fact, our Monte Carlo experiments suggest our estimators are significantly more stable in the sense that they exhibit better performance for most values of the bandwidth, while the jackknife estimator requires accurate tuning of the "bandwidth" to match our estimator's performance. We also observe that even without bias corrections, under certain conditions, the plug-in estimator with general kernel spot volatility estimator has good performance, which could be a result of the increased accuracy of the spot volatility estimator and/or the reduction in bias attained by using a general kernel.

While working on the final stages of writing the present manuscript, we became aware of a recent work by Benvenuti (2021), who also studied a type of plug-in estimator with a general two-sided kernel spot volatility estimator. However, unlike our work, only a suboptimal bandwidth asymptotic regime (i.e., assuming undersmoothing) was considered and only consistency was established.

The rest of this article is organized as follows. Section 2 introduces the framework, assumptions, and main results. Section 3 illustrates the performance of our method via Monte Carlo simulations and compares it to the jackknife estimator of Li et al. (2019). Section 4 provides some concluding remarks. The proofs of the main results of the article are provided in Appendix A. All other proofs of the article are delegated in the Supplemental Material available online.

Notation: We shall use the following notation:

- the set of (nonnegative definite) $d \times d$ real matrices $x = [x^{\ell m}]_{\ell,m=1}^d$ is denoted $(\mathcal{M}_+^d)$ $\mathbb{R}^{d\times d}$;
- for a C^1 function $g : \mathbb{R}^{d\times d} \to \mathbb{R}$, its gradient $\nabla g(x)$ is the $d \times d$ matrix with (ℓ, m) entry $\partial_{\ell m} g(x) := \frac{\partial g(x)}{\partial x^{\ell m}}$;
- for $x, y \in \mathbb{R}^{d\times d}$, its inner product is defined as $x \cdot y = \sum_{\ell,m} x^{\ell m} y^{\ell m}$. Also, x^* is the transpose of x and $\|x\| = \sqrt{\sum_{\ell,m=1}^n (x^{\ell m})^2}$ is the Euclidian norm.

2. THE SETTING, ESTIMATOR, AND MAIN RESULTS

Throughout, we consider a d-dimensional Itô semimartingale $X = [X^1, \ldots, X^d]^* \in \mathbb{R}^{d\times 1}$ of the form:

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s + \int_0^t \int_E \delta(s,z)\mathbb{1}_{\{|\delta(s,z)|\le 1\}}(\mathfrak{p}-\mathfrak{q})(ds,dz) + \int_0^t \int_E \delta(s,z)\mathbb{1}_{\{|\delta(s,z)|>1\}}\mathfrak{p}(ds,dz), \tag{2.1}$$

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

where all stochastic processes ($\mu := \{\mu_t\}_{t\geq 0}$, $\sigma := \{\sigma_t\}_{t\geq 0}$, $W := \{W_t\}_{t\geq 0}$, $\delta := \{\delta(t,z) : t \geq 0, z \in E\}$, $\mathfrak{p} := \{\mathfrak{p}(B) : B \in \mathcal{B}(\mathbb{R}_+ \times E)\}$, etc.) are defined on a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$. Here, $W = [W^1, \ldots, W^d]^*$ is a d-dimensional standard BM adapted to the filtration $\mathbb{F}$, δ is a predictable $\mathbb{R}^d$-valued function on $\Omega \times \mathbb{R}_+ \times E$, and $\mathfrak{p}$ is a Poisson random measure on $\mathbb{R}_+$. We also denote the spot variance–covariance process as $c_t := \sigma_t \sigma_t^*$, which takes values in the set $\mathcal{M}_d^+$ of nonnegative definite $d \times d$ symmetric matrices. As is common in the literature, c is called the spot volatility of the process. We assume c is also an Itô semimartingale following the dynamics:

$$c_t^{\ell m} = c_0^{\ell m} + \int_0^t \tilde{\mu}_s^{\ell m} \mathrm{d}s + \int_0^t \tilde{\sigma}_s^{\ell m} \mathrm{d}W_s, \quad 1 \leq \ell, m \leq d, \tag{2.2}$$

where $W := \{W_t\}_{t\geq 0}$ is the same d-dimensional BM driving the dynamics of X. For each $1 \leq \ell, m \leq d$, $\{\tilde{\mu}_t^{\ell m}\}_{t\geq 0}$ is adapted locally bounded and $\{\tilde{\sigma}_t^{\ell m}\}_{t\geq 0}$ is adapted càdlàg. The assumption (2.2) is common in the literature (see, e.g., the monographs Jacod and Protter (2011) and Aït-Sahalia and Jacod (2014)), and was also assumed in the work of Jacod and Rosenbaum (2013).

We now state the main assumption on the process X.

Assumption 1. The process X follows the dynamics (2.1) with $c_t = \sigma_t \sigma_t^*$ satisfying (2.2) and, for some $r \in [0,1)$, a measurable function $\Gamma_m : E \to \mathbb{R}_+$, a constant $C_m < \infty$, and a localizing sequence of stopping times $(\tau_m)_{m\geq 1}$ such that $\tau_m \to \infty$, we have

$$t \in [0, \tau_m] \Longrightarrow \begin{cases} |\mu_t| + |\sigma_t| + |\tilde{\mu}_t| + |\tilde{\sigma}_t| \leq C_m, \\ |\delta(t,z)| \wedge 1 \leq \Gamma_m(z), \quad \text{where } \int \Gamma_m(z)^r \lambda(dz) < \infty. \end{cases}$$

The parameter r determines the jump activity of the process: the larger r is, the more active or frequent are the small jumps of the process. When $r = 0$, the process exhibits finitely many jumps in any bounded time interval (in that case, we say that the jumps are of finite activity). When $r < 1$, the jump component of the process is of bounded variation, which is a standard constraint for the truncated realized quadratic variation to achieve efficiency (Jacod and Protter, 2011).

Next, we give the conditions on the kernel function K for which we need to define the function

$$L(t) := \int_t^\infty K(u)du \mathbf{1}_{\{t>0\}} - \int_{-\infty}^t K(u)du \mathbf{1}_{\{t\leq 0\}}.$$

As explained in the introduction, we aim to consider two-sided kernels of unbounded support. It is important to remark that this type of kernel presents some challenges to our analysis (see comments after (A.7)).

Assumption 2. Beside the usual condition (a) $\int K(x)dx = 1$, the kernel function $K : \mathbb{R} \to \mathbb{R}$ is assumed to be bounded such that:

b. K is Lipschitz and piecewise C^2 on its support (A, B), where $A < B$, $A \in [-\infty, 0]$, and $B \in [0, \infty]$; K' and K'' are absolutely continuous on their corresponding

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

domains $\mathrm{dom}(K')$, $\mathrm{dom}(K'') \subset (A,B)$, respectively; K''' exists piecewise on $(A,B) \subset \mathbb{R}$;

c. (i) $\int |K(x)x|dx < \infty$; (ii) $K(x)|x|^{2+\epsilon} \to 0$, $K'(x)|x|^{2+\epsilon} \to 0$, with $\epsilon > 1/2$ as $|x| \to \infty$; (iii) $\int |K^{(j)}(x)|dx < \infty$, for $j = 0,1,2,3$; (iv) $\int_0^\infty \int_w^\infty |K'(x)|dxdw < \infty$; (v) $\int K^2(x)dx < \infty$; (vi) $\int_{-\infty}^{\infty} |L(u)u|du < \infty$;

d. (i) $V_{-\infty}^{\infty}(K^{(j)}) := \lim_{m\to\infty} V_{-m}^{m}(K^{(j)}) < \infty$, where $V_{-m}^{m}(K^{(j)})$ is the total variation of $K^{(j)}$ on the interval $[-m,m]$, $j = 0,1,2,3$; (ii) $V_{-\infty}^{\infty}(K^2) < \infty$.

Remark 2.1. Even though the list of assumptions above seems long, it is important to remark that all kernels commonly used in practice satisfy these conditions. The assumptions on the regularity of the high-order derivatives $K^{(j)}$ for $j = 2,3$ are only used to show the following limit:

$$\sqrt{\Delta_n}\sum_{i=1}^{n}\left[\bar{K}_n(t_i) - 1\right] g\left(c_i^n\right) \longrightarrow \theta\int_{-\infty}^{0} L(u)du\, g(c_0) - \theta\int_{0}^{\infty} L(u)du\, g(c_T), \tag{2.3}$$

where recall that $\bar{K}_n(t_i) := \Delta_n \sum_{j=1}^{n} K_{k_n\Delta_n}\left(t_{j-1} - t_i\right)$. This limit is easy to deal with in the case of a uniform kernel because, in that case, $\bar{K}_n(t_i) \equiv 1$ for almost all i except if i is close to 0 or n.

For an arbitrary process $\{U_t\}_{t\geq 0}$, a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$, and a given time span $\Delta_n > 0$, we shall use the notation

$$U_i^n := U_{i\Delta_n}, \qquad \Delta_i^n U := U_i^n - U_{i-1}^n, \qquad \mathcal{F}_i^n := \mathcal{F}_{i\Delta_n}.$$

Stable convergence in law is denoted by $\xrightarrow{st}$ (see Jacod and Protter, 2011, Eq. (2.2.4)). As usual, $a_n \sim c_n$ means that $a_n/c_n \to 1$ as $n \to \infty$.

Throughout, we assume that we sample the process X in (2.1) at discrete times $t_i := t_{i,n} := i\Delta_n$, where $\Delta_n := T/n$ and $T \in (0,\infty)$ is a given fixed time horizon. To estimate the spot volatility c_{t_i}, we adopt a Nadaraya–Watson type of kernel estimator of the form:

$$\hat{c}_i^{n,X} := \frac{\sum_{j=1}^{n} K_{k_n\Delta_n}\left(t_{j-1} - t_i\right)\left(\Delta_j^n X\right)\left(\Delta_j^n X\right)^*}{\bar{K}_n(t_i)}, \tag{2.4}$$

where, again, $K_b(x) := K(x/b)/b$, $k_n \in \mathbb{N}$, $b_n := k_n\Delta_n$ is the bandwidth of the kernel function, and $\bar{K}_n(t_i) := \Delta_n \sum_{j=1}^{n} K_{k_n\Delta_n}\left(t_{j-1} - t_i\right)$. This is a variation of the estimator

$$\tilde{c}_\tau^{n,X} := \sum_{j=1}^{n} K_{k_n\Delta_n}\left(t_{j-1} - \tau\right)\left(\Delta_j^n X\right)\left(\Delta_j^n X\right)^*, \tag{2.5}$$

first proposed in Fan and Wang (2008) and Kristensen (2010), and studied in recent papers, such as Figueroa-López and Li (2020) and Figueroa-López and Wu (2024). The denominator in (2.4) is known to improve the performance (2.5) for values of $\tau = t_i$ near the edge of the estimation interval $[0,T]$. The

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

variation (2.4) was also considered in Kristensen (2010) and Figueroa-López and Li (2020) when t_i is replaced with a fix time $\tau \in (0,T)$. In that case, $\bar{K}_n(\tau) = \Delta_n \sum_{j=1}^n K_{k_n\Delta_n}\left(t_{j-1}-\tau\right) \to 1$ and the asymptotic behavior of (2.4) and (2.5) is equivalent. However, for our estimator (2.9) below, we need to consider $\tau = t_i$ for all $i = 1,\dots,n$ and, thus, it is not direct that we can simply omit $\bar{K}_n(t_i)$. In fact, its presence helps with the analysis (see comments after (A.2)).

To handle jumps in the process X, we also consider a truncated version of (2.4):

$$\hat{c}_i^{n,X,v_n} := \frac{\sum_{j=1}^n K_{k_n\Delta_n}\left(t_{j-1}-t_i\right)\left(\Delta_j X\right)\left(\Delta_j X\right)^* \mathbb{1}_{\{\|\Delta_j X\| \le v_n\}}}{\Delta_n \sum_{j=1}^n K_{k_n\Delta_n}\left(t_{j-1}-t_i\right)}, \tag{2.6}$$

with a suitable sequence of truncation levels $v_n \in (0,\infty]$. More specifically, we consider two types of conditions on v_n. In the case that X is continuous, we take

$$v_n = \alpha\Delta_n^{\varpi}, \quad \text{for some arbitrary fixed } \alpha \in (0,\infty] \text{ and } \varpi < \frac{1}{2}, \tag{2.7}$$

while if X is discontinuous, we take

$$v_n = \alpha\Delta_n^{\varpi}, \quad \text{for some fixed } \alpha \in (0,\infty) \text{ and } \varpi \in \left[\frac{2\ell-1}{2(2\ell-r)}, \frac{1}{2}\right), \tag{2.8}$$

where $\ell \ge 4$ is a positive integer that will be specified below. Note that when $\alpha = \infty$ in (2.7), we recover the non-truncated version (2.4) and, in that case, the value of ϖ is irrelevant. For simplicity, we often omit the dependence on X and v_n on (2.4) and (2.6) and simply write $\hat{c}_i^n$.

Our estimation target is the IV functional, defined as

$$V(g)_t := \int_0^t g(c_s)\,ds, \tag{2.9}$$

for a continuous function $g : \mathcal{M}_d^+ \to \mathbb{R}$. As proposed in Jacod and Rosenbaum (2013), a natural estimator for this integral is given by $\Delta_n \sum_{i=1}^{[t/\Delta_n]} g\left(\hat{c}_i^n\right)$, which can be interpreted as a Riemann sum approximation of (2.9). However, as we shall see, it will be easier to derive the asymptotic properties of the estimator

$$V(g)_t^n := \Delta_n \sum_{i=1}^{[t/\Delta_n]} g\left(\hat{c}_i^n\right)\bar{K}_n(t_i). \tag{2.10}$$

We can also see (2.10) as a weighted average version of $\Delta_n \sum_{i=1}^{[t/\Delta_n]} g\left(\hat{c}_i^n\right)$, which gives less weight to $g\left(\hat{c}_i^n\right)$ for values of t_i near 0 and T, where the estimator $\hat{c}_i^n$ is less accurate due to edge effects.

Our first result characterizes the asymptotic behavior of the estimation error of $V(g)_t^n$ when the bandwidth converges to 0 at an "optimal" rate (cf. Theorem (2.3) below and the remarks before). Its proof can be found in Appendix A.1.

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

THEOREM 2.1. *Let g be a C^3 function such that, for all $x \in \mathbb{R}^{d\times d}$,*

$$|g(x)| \le C\left(1+\|x\|^{\ell}\right), \quad \left|\partial_{p_1q_1,\dots,p_jq_j} g(x)\right| \le C\left(1+\|x\|^{\ell-j}\right), \quad j=1,2,3, \tag{2.11}$$

for $p_1, q_1, \dots, p_j, q_j \in \{1,\dots,d\}$ and for some constants $C>0$ and $\ell \ge 4$. Assume k_n satisfies

$$k_n \sim \frac{\theta}{\sqrt{\Delta_n}} \qquad (i.e.,\ \text{the bandwidth } b_n := k_n\Delta_n \sim \theta\sqrt{\Delta_n}), \tag{2.12}$$

for some $\theta \in (0,\infty)$. Then, under (2.1) and Assumptions 1 and 2, the following assertions hold:

a) *Suppose that X is continuous (i.e., $\delta \equiv 0$ in (2.1)). Then, the estimator (2.10) of the IV functional $V(g)_T$ with $\hat{c}_i^n$ given by (2.6), under the condition (2.7),[4] satisfies the following stable convergence in law, as $n \to \infty$:*

$$\frac{1}{\sqrt{\Delta_n}}\left(V(g)_T^n - V(g)_T\right) \xrightarrow{st} A^1 + A^2 + A^3 + Z, \tag{2.13}$$

where Z is an r.v. defined on an extension $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, (\widetilde{\mathcal{F}}_t)_{t\ge 0}, \widetilde{\mathbb{P}}\right)$ of $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P}\right)$, which, conditionally on $\mathcal{F}$, is a centered Gaussian variable with variance

$$\widetilde{\mathbb{E}}\left(Z^2 \mid \mathcal{F}\right) = \sum_{j,k,l,m} \int_0^T \partial_{jk} g(c_s)\partial_{lm} g(c_s)(c_s^{jl}c_s^{km} + c_s^{jm}c_s^{kl})ds, \tag{2.14}$$

and

$$\begin{aligned}
A^1 &:= -\theta \int_0^\infty L(u)du\, g(c_0) + \theta \int_{-\infty}^0 L(u)du\, g(c_T),\\
A^2 &:= \frac{1}{2\theta}\int_0^T \sum_{p,q,u,v=1}^d \partial^2_{pq,uv} g(c_s)\, \check{c}_s^{pq,uv} ds \int_{-\infty}^\infty K^2(u)du,\\
A^3 &:= \frac{\theta}{2}\int_0^T \sum_{p,q,u,v=1}^d \partial^2_{pq,uv} g(c_s)\tilde{c}_s^{pq,uv} ds\left[\int_{-\infty}^\infty L(u)^2 du + \left(\int_{-\infty}^0 - \int_0^\infty\right) L(u)du\right],
\end{aligned} \tag{2.15}$$

where $\check{c}_s^{pq,uv} := c_s^{pu}c_s^{qv} + c_s^{pv}c_s^{qu}$ and $\tilde{c}_s^{pq,uv} := \sum_{r=1}^d \tilde{\sigma}_s^{pq,r}\tilde{\sigma}_s^{uv,r}$.

b) *When X is discontinuous, we still have (2.13) for the estimator $V(g)_T^n$ with the truncated version (2.6) provided that the stronger condition (2.8) is satisfied.*

Remark 2.2. Note that the bias terms depend on the kernel in a rather intricate manner (especially, the term A^3). The signs of the coefficients are always the same regardless of K. For instance, $\int_{-\infty}^\infty L(u)^2 du + \left(\int_{-\infty}^0 - \int_0^\infty\right) L(u)du$ is negative

[4] In particular, by taking $\alpha = \infty$, we recover the untruncated estimator (2.4).

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

because $|L(u)| \leq 1$ for all u, and $L(u) \leq 0$ ($L(u) \geq 0$) for any $u < 0$ ($u > 0$). Thus, A^3 goes in the opposite direction to A^2.

Remark 2.3. If we use the uniform right-sided kernel $K(x) = \mathbf{1}_{[0,1)}(x)$ in our estimators, as in Jacod and Rosenbaum (2013, 2015), then $\bar{K}_n(t_i) = \Delta_n \sum_{j=1}^{n} K_{b_n}(t_{j-1} - t_i) = 1$, for $0 \leq i \leq n - k_n$. In that case, the estimator in Jacod and Rosenbaum (2015), defined as $V(g)_T^{n,JR} := \Delta_n \sum_{i=1}^{n-k_n-1} g\left(\hat{c}_i^{n,JR}\right)$, where

$$\hat{c}_i^{n,JR} := \frac{1}{k_n \Delta_n} \sum_{j=0}^{k_n-1} \left(\Delta_{i+j}^n X\right)\left(\Delta_{i+j}^n X\right)^* \mathbb{1}_{\left\{\left\|\Delta_{i+j}^n X\right\| \leq v_n\right\}}, \tag{2.16}$$

can be written in terms of our estimator (2.10) as follows:

$$V(g)_T^{n,JR} = V(g)_T^n - \Delta_n \sum_{i=n-k_n-1}^{n} g\left(\hat{c}_i^n\right) \bar{K}(t_i) + \Delta_n g\left(\hat{c}_0^n\right). \tag{2.17}$$

It is then easy to see that Theorem 2.1 implies that

$$\frac{1}{\sqrt{\Delta_n}}\left(V(g)_T^{n,JR} - V(g)_T\right) \xrightarrow{st} \tilde{A}^1 + A^2 + A^3 + Z, \tag{2.18}$$

where $\tilde{A}^1 = -\frac{\theta}{2}[g(c_0) + g(c_T)]$ and, thus, we recover the same result as Jacod and Rosenbaum (2015).

To make this CLT "feasible" in practical work, the bias terms need to be estimated. For A^1, we use that

$$k_n\sqrt{\Delta_n} g\left(\hat{c}_1^n\right) \xrightarrow{\mathbb{P}} \theta g(c_0), \quad k_n\sqrt{\Delta_n} g\left(\hat{c}_n^n\right) \xrightarrow{\mathbb{P}} \theta g(c_T). \tag{2.19}$$

For A^2, we can apply Theorem 2.1 to the functional

$$h(x) := \sum_{p,q,u,v=1}^{d} \partial^2_{pq,uv} g(x)\left(x^{pu}x^{qv} + x^{pv}x^{qu}\right), \tag{2.20}$$

for any $x \in \mathbb{R}^{d\times d}$, to get

$$V(h)_T^n = \Delta_n \sum_{i=1}^{n} h\left(\hat{c}_i^n\right) \bar{K}(t_i) \xrightarrow{\mathbb{P}} \int_0^T \sum_{p,q,u,v=1}^{d} \partial^2_{pq,uv} g(c_s)\, \check{c}_s^{pq,uv} ds. \tag{2.21}$$

The term A^3 is the hardest to estimate since it involves the "volvol" $\tilde{c}$. In the case of a uniform right-sided kernel, $K(x) = \mathbf{1}_{[0,1)}(x)$, Jacod and Rosenbaum (2015) proposed the estimator

$$\widehat{A}^{n,3} := \frac{\sqrt{\Delta_n}}{8} \sum_{i=1}^{n-2k_n+1} \sum_{j,k,l,m} \partial^2_{jk,lm} g(\hat{c}_i^n)\left(\hat{c}_{i+k_n}^{n,jk} - \hat{c}_i^{n,jk}\right)\left(\hat{c}_{i+k_n}^{n,lm} - \hat{c}_i^{n,lm}\right), \tag{2.22}$$

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

with $\hat{c}$ given by (2.16) and showed that $\widehat{A}^{n,3}$ converges to a linear combination of the bias terms A^2 and A^3. The following theorem shows the corresponding result for a general kernel function. The proof of Theorem 2.2 is given in the Supplementary Material.

THEOREM 2.2. *Under the notations and conditions of Theorem 2.1, we have, as* $n \to \infty$,

$$\begin{aligned}\widehat{A}^{n,3} \xrightarrow{\mathbb{P}} & \frac{1}{4\theta}\int_0^T \sum_{j,k,l,m} \partial^2_{jk,lm} g(c_s)\, \check{c}_s^{jk,lm} ds \int_{-\infty}^{\infty} K(z)(K(z)-K(z-1))dz \\ & + \frac{\theta}{4}\int_0^T \sum_{j,k,l,m} \partial^2_{jk,lm} g(c_s)\, \tilde{c}_s^{jk,lm} ds \\ & \times \left(\int_{-\infty}^{\infty} L(z)(L(z)-L(z-1))dz + \frac{1}{2} - \int_{-\infty}^{\infty} K(z)(|z|\wedge 1)dz\right),\end{aligned} \tag{2.23}$$

when $\hat{c}$ *is given by (2.6) with the condition (2.8). Furthermore, when* X *is continuous (i.e.,* $\delta \equiv 0$ *in (2.1)), (2.23) still holds under the weaker condition (2.7) (in particular, we can use the untruncated estimator (2.5) by taking* $\alpha = \infty$*).*

Under the assumption (2.12), based on (2.19), (2.21), and (2.23), we then propose a bias-corrected estimator of the form:

$$\begin{aligned}\widetilde{V}(g)_T^n := & V(g)_T^n + k_n \Delta_n g\left(\hat{c}_1^n\right)\int_0^{\infty} L(u)du - k_n \Delta_n g\left(\hat{c}_n^n\right)\int_{-\infty}^{0} L(u)du \\ & - \frac{\Delta_n}{4}\sum_{i=1}^{n-2k_n+1}\sum_{p,q,u,v=1}^{d} \partial^2_{pq,uv} g\left(\hat{c}_i^n\right)\left(\hat{c}_{i+k_n}^{n,pq} - \hat{c}_i^{n,pq}\right)\left(\hat{c}_{i+k_n}^{n,uv} - \hat{c}_i^{n,uv}\right) C_{K,1} \\ & + \frac{1}{2k_n} V(h)_T^n \left[C_{K,2} - \int K^2(u)du\right],\end{aligned} \tag{2.24}$$

where

$$C_{K,1} := \frac{\int_{-\infty}^{\infty} L(z)^2 dz + \left(\int_{-\infty}^{0} - \int_0^{\infty}\right) L(z)dz}{\int_{-\infty}^{\infty} L(z)(L(z)-L(z-1))dz + \frac{1}{2} - \int_{-\infty}^{\infty} K(z)(|z|\wedge 1)dz},$$

and

$$C_{K,2} := \int_{-\infty}^{\infty} K(z)(K(z)-K(z-1))dz C_{K,1}.$$

Next, we show that indeed $\widetilde{V}(g)_T^n$ enjoys a centered CLT. Its simple proof is given in Appendix A.2.

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

COROLLARY 2.1. *Under the assumptions and notation of Theorem 2.1, as* $n \to \infty$*, we have*

$$\frac{1}{\sqrt{\Delta_n}}\left(\widetilde{V}(g)_T^n - V(g)\right) \xrightarrow{st} Z, \tag{2.25}$$

where $\widetilde{V}(g)_T^n$ *is defined as in (2.24) with* $\hat{c}$ *given by (2.6) with the condition (2.8). Furthermore, when X is continuous, (2.25) holds under the weaker condition (2.7).*

As mentioned above, the bias term A^3 contains the "volvol" $\tilde{c}$ and the estimator of A^3 in (2.23) could introduce extra variance to the asymptotically unbiased estimator (2.24). An alternative way to eliminate the bias is undersmoothing, i.e., through the selection of a bandwidth sequence b_n converging to 0 at a faster rate than $\sqrt{\Delta_n}$ ($b_n \ll \sqrt{\Delta_n}$) or, equivalently, picking k_n such that $\sqrt{\Delta_n} k_n \xrightarrow{n\to\infty} 0$ (meaning that $\theta = 0$ in (2.12)). This is the approach put forward in Jacod and Rosenbaum (2013). In that case, it is expected that the bias terms A^1 and A^3 will vanish and the bias term A^2 will dominate. Following the same idea, we can devise an asymptotically unbiased estimator as follows:

$$\bar{V}(g)_T^n := \Delta_n \sum_{i=1}^{n} \left(g(\hat{c}_i^n) - \frac{1}{2k_n} h(\hat{c}_i^n) \int K^2(u)du \right) \bar{K}(t_i). \tag{2.26}$$

The following result establishes the asymptotic behavior of $\bar{V}(g)_T^n$ (see the proof in the Supplementary Material).

THEOREM 2.3. *Assume* k_n *satisfies*

$$k_n^2 \Delta_n \to 0, \quad \text{and} \quad k_n^3 \Delta_n \to \infty. \tag{2.27}$$

Then, under (2.1) and Assumptions 1 and 2, we have the following stable convergence in law:

$$\frac{1}{\sqrt{\Delta_n}}\left(\bar{V}(g)_T^n - V(g)_T\right) \xrightarrow{st} Z, \tag{2.28}$$

where Z is as in Theorem 2.1 and $\bar{V}(g)_T^n$ *is defined as in (2.26) with* $\hat{c}$ *given by (2.6) under condition (2.8). Furthermore, when X is continuous, (2.28) holds under the weaker condition (2.7).*

3. SIMULATION STUDY

In this section, we analyze the performance of our kernel-based estimators. To isolate the effect of the bandwidth on the estimators' performance, we only consider continuous processes (i.e., $\delta \equiv 0$ in (2.1) and, thus, $X \equiv X'$) and focus on the untruncated estimator (2.4) and the corresponding estimators $V(g)_t^n$ of (2.10)

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

and $\widetilde{V}(g)_T^n$ of (2.24). We consider the same data generating model as in Li et al. (2019):

$$\begin{cases} dX_t = \sigma_t dW_t \\ \sigma_t = exp(-1.6 + F_t), \quad dF_t = -5F_t dt + 2dB_t, \end{cases} \tag{3.1}$$

with $\mathbb{E}[dW_t dB_t] = -0.75dt$. Throughout, we use the function $g(c) = c^2$, which corresponds to the integrated quarticity (for more numerical results, including the analysis of the function $g(c) = \log(c)$, see the Supplementary Material). We simulate data for $T = 5$ and $T = 21$ days, and assume sampling frequencies of 1 minute or 5 minutes, with 6.5 trading hours per day.

We compare our bias-corrected estimator $\widetilde{V}(g)_T^n$ to the Jackknife estimator of Li et al. (2019). For the Jackknife method, we consider the two-scale estimator TS_n, its boundary-adjusted version TS'_n, and the multiscale estimator MS_n as defined in Li et al. (2019, Sect. 3.1). The parameters for the Jackknife estimators are chosen according to Li et al. (2019). By running numerous simulations, we determine that the values chosen by Li et al. (2019) broadly optimize their estimator's performance. For the multiscale estimator MS_n, $(\psi_1, \psi_2, \psi_3) = (-2.5, 8, -4.5)$, and $\left(k_{1,n}, k_{2,n}, k_{3,n}\right) = (15, 30, 45)$ for data sampled every 5 minutes, or $(40, 80, 120)$ for data sampled every minute. For the two-scale estimator, $(\psi_1, \psi_2) = (-1, 2)$ along with the same $\left(k_{1,n}, k_{2,n}\right)$ as above. For our kernel estimator, we use three kernels: a two-sided exponential kernel $K(x) = \frac{1}{2}e^{-|x|}$, a two-sided uniform kernel $K(x) = \frac{1}{2}\mathbf{1}_{[-1,1]}(x)$, and a right-sided uniform kernel $K(x) = \mathbf{1}_{[0,1]}(x)$.[5] We analyze the performance of the simple biased estimator $V(g)_T^n$ in (2.10) and the bias-corrected estimator $\widetilde{V}(g)_T^n$ in (2.24). We also consider the simple biased estimator $V(g)_T^{n,JR}$, defined in Remark 2.3, and the corresponding bias-corrected estimator $\tilde{V}(g)_T^{n,JR}$, proposed in Jacod and Rosenbaum (2015).

The bandwidths of all the kernel estimators are chosen using the iterative method in Figueroa-López and Li (2020, Sect. 5) We iterate the method only once. For a two-sided kernel, the IV of volatility used when updating the bandwidth is estimated by the two-time scale realized volatility (TSRV) estimator proposed in Figueroa-López and Li (2020). For the right-sided uniform kernel, the TSRV estimator no longer applies. We then use the realized variance estimator on the estimated volatility path as a replacement. As finite-sample corrections, in (2.24), at each t_i, we replace $\int_{-\infty}^{\infty} K^2(u)du$ with $\sum_{j=1}^n K_h^2(t_i - t_j)\Delta_n h$, and replace $\int_{-\infty}^{\infty} L(s)ds$ and $\int_{-\infty}^{\infty} L^2(s)ds$ with $\frac{\Delta_n}{b}\sum_{j=1}^n L\left(\frac{t_{j-1}-t_i}{b}\right)$ and $\frac{\Delta_n}{b}\sum_{j=1}^n L^2\left(\frac{t_j-t_i}{b}\right)$, respectively, where $L\left(\frac{t_j-t_i}{b}\right) = -\sum_{l=1}^j K_b(t_{l-1} - t_i)$ if $j \leq i$ and $\sum_{l=j+1}^n K_b(t_{l-1} - t_i)$ if $j > i$.

In Table 1, we report the relative bias and relative root MSEs (RMSEs) in percentage unit for the considered estimators. Our results of the Jackknife estimators are consistent with the ones in Li et al. (2019, Table 1) As shown

[5] As pointed out by a referee, in econometrics, many datasets are cross-sectional and, hence, it may be preferable to use a two-sided kernel. However, in the context of time series, it may be desirable to work with left-sided kernels to estimate the current spot online.

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

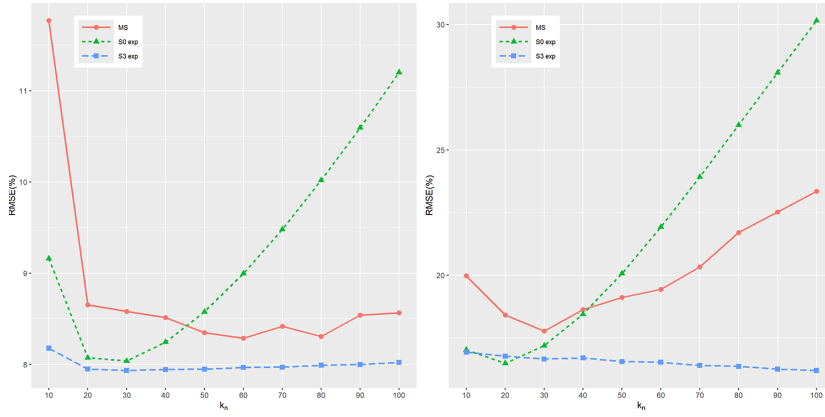
TABLE 1. Bias and relative MSE when $g(c) = c^2$

	$T = 21$ days				$T = 5$ days			
	$\Delta_n = 5$ minutes		$\Delta_n = 1$ minute		$\Delta_n = 5$ minutes		$\Delta_n = 1$ minute	
	BIAS(%)	RMSE(%)	BIAS(%)	RMSE(%)	BIAS(%)	RMSE(%)	BIAS(%)	RMSE(%)
				Exponential kernel				
V_T^n	5.884	9.890	2.345	3.941	13.291	17.627	6.718	8.537
$\widetilde{V}_T^n$	1.557	8.214	0.042	3.336	1.531	16.096	0.949	6.861
				Two-sided uniform kernel				
V_T^n	4.824	9.497	1.821	3.729	11.332	16.557	5.783	7.903
$\widetilde{V}_T^n$	1.804	8.357	0.084	3.352	2.288	15.843	1.183	6.874
				Right-sided uniform kernel				
V_T^n	1.727	8.237	0.208	3.384	8.380	16.339	4.638	7.086
$\widetilde{V}_T^n$	1.083	8.334	−0.267	3.361	1.393	16.509	0.665	6.965
				Jacod & Rosenbaum's estimators				
$V_T^{n,JR}$	5.096	11.127	1.656	4.115	19.714	23.313	8.574	10.459
$\widetilde{V}_T^{n,JR}$	1.795	8.469	0.061	3.370	2.002	16.562	1.139	7.067
				Jackknife estimators				
TS_n	4.365	8.223	2.195	3.435	3.802	14.834	2.991	6.779
TS'_n	1.323	8.162	0.213	3.308	−8.019	20.895	−3.855	8.714
MS_n	1.989	8.220	0.098	3.347	1.618	16.641	1.041	6.930

in Table 1, the kernel estimators exhibit a similar performance to the Jackknife estimators. However, the Jackknife estimators have more tuning parameters, while ours requires only one tuning parameter (namely, θ). As we will show below, the most important advantage of our estimators lies in their stability relative to the choice of the bandwidth. Returning to the results of Table 1, we also observe that in some cases the kernel estimators have smaller relative RMSE compared with the Jackknife estimators, especially when the time span T is short ($T = 5$ days). Intuitively, the superior performance of the spot volatility estimators with exponential kernel (shown in Figueroa-López and Li (2020, Sect. 7)) is expected to be more evident when the time span is relatively short, e.g., five days.

We now investigate how sensitive the estimators are relative to the bandwidth parameter. For the Jackknife estimators, we focus on the boundary adjusted multi-scale estimator MS (which, as shown in Table 1, typically has the best performance among the considered Jackknife estimators), and for the kernel estimators, we first consider an exponential kernel. We try a sequence of bandwidths k_n for the kernel estimator, and a sequence of triplets $(k_{1,n}, k_{2,n}, k_{3,n}) = ([\frac{k_n}{2}], k_n, [\frac{3k_n}{2}])$ and $(\psi_1, \psi_2, \psi_3) = (-2.5, 8, -4.5)$ for MS, which include the parameters chosen in Li et al. (2019) when $k_n = 30$ (see Table 1 therein). For the sake of space, we only consider the frequency of 5 minutes. The results are shown in Figure 1,

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press



(a) $T = 21$ days time horizon. (b) $T = 5$ days time horizon.

FIGURE 1. RMSE(%) versus the bandwidth k_n, for $T = 21$ and 5 days data with 5-minute frequency, $g(c) = c^2$, and 1,000 paths.

where S0 exp denotes the simple biased exponential kernel estimator in (2.10) and S3 exp denotes the bias-corrected estimators with exponential kernel defined in (2.24). We can conclude that the estimator S3 exp exhibits the most stable or consistent performance for either 5 or 21 days data. In other words, the estimator's values are in general relatively invariant to changes in the bandwidth, while the other estimators require more careful tuning of the bandwidth in order to show good performance. We can also observe that the bias-corrected exponential kernel estimator almost always outperforms the boundary-adjusted multi-scale Jackknife estimator MS and the simple biased exponential kernel estimator S0 exp.

4. CONCLUDING REMARKS

In this article, we provide a more flexible class of estimators for functionals of the IV. The new class enjoys enhanced stability with respect to the bandwidth, which allows to obtain more accurate estimates without worrying about optimal bandwidth selections. In other words, our estimators exhibit significantly more stable performance (i.e., less sensitive against the bandwidth) than those using uniform kernels or multiscale debiasing as those proposed in Li et al. (2019). The obatined CLT of (1.5) again exhibits several biases, which depend on a rather intricate and nontrivial manner on the kernel. For future research, we plan to investigate other boundary bias-corrected estimators as proposed in, e.g., Kristensen (2010), estimators that are robust to microstructure noise as in Chen (2019) and Figueroa-López and Wu (2024), and methods for optimally tuning the estimators' parameters as in Figueroa-López and Li (2020).

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

A. PROOFS OF MAIN RESULTS

By virtue of localization, without loss of generality, we assume throughout the proofs that, for some constant A,

$$\Gamma(z) + \int \Gamma(z)^r \lambda(dz) + |\mu_t| + |\sigma_t| + |X_t| + \left|\mu'_t\right| + |\tilde{\sigma}_t| + |\tilde{\mu}_t| \le A, \text{ for any } t \le T, \tag{A.1}$$

(see Jacod and Protter (2011, Sect. 4.4.1 and Eq. (6.2.1)) and Aït-Sahalia and Jacod (2014, Appendix A.5) for details).

As in Jacod and Rosenbaum (2013), we consider the following quantities:

$$\begin{aligned}
\alpha_i^n &:= \left(\Delta_i^n X\right)\left(\Delta_i^n X\right)^* - c_{i-1}^n \Delta_n \in \mathbb{R}^{d\times d}, \\
\beta_i^n &:= \hat{c}_i^{n,X} - c_i^n = \frac{\sum_{j=1}^n K_{k_n\Delta_n}\left(t_{j-1} - t_i\right)\left(\Delta_j^n X\right)\left(\Delta_j^n X\right)^*}{\Delta_n \sum_{j=1}^n K_{k_n\Delta_n}\left(t_{j-1} - t_i\right)} - c_i^n \\
&= \frac{1}{\bar{K}(t_i)} \sum_{j=1}^n K_b\left(t_{j-1} - t_i\right)\left[\alpha_j^n + \left(c_{j-1}^n - c_i^n\right)\Delta_n\right].
\end{aligned} \tag{A.2}$$

The first term is an approximation of the martingale part of $\left(\Delta_i^n X\right)\left(\Delta_i^n X\right)^*$, for which we can derived tight estimates using martingale inequalities (see Jacod and Rosenbaum (2015) and Eqs. (1.17) and (1.18) in the Supplementary Material). The second term is the estimation error of our spot volatility estimator $\hat{c}_i^{n,X}$. The nice decomposition shown therein in terms of α_j^n and $c_{j-1}^n - c_i$ was possible because of the denominator $\bar{K}(t_i)$ in the estimator $\hat{c}_i^{n,X}$ in (2.4). This is one of the main reasons why $\bar{K}(t_i)$ was added. For a uniform kernel, $\bar{K}(t_i)$ equals to 1 for almost all i and is not needed if one considers the truncated version (1.2) as in Jacod and Rosenbaum (2013). However, it seems more natural to include the "standardized" term $\bar{K}(t_i)$ for a general kernel function.

A.1. Proof of Theorem 2.1

As is standard in the literature, we can show that the jumps can be neglected (see Lemma 2.1 in the Supplementary Material). We can then assume without loss of generality the following.

Assumption 3. We have (2.1) with X continuous and $c_t := \sigma_t \sigma_t^*$ satisfies (2.2).

We shall exploit the following decomposition of the estimation error of (2.10):

$$\frac{1}{\sqrt{\Delta_n}}\left(V(g)_T^n - V(g)_T\right) = \sum_{j=1}^5 V_j^n, \tag{A.3}$$

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

where

$$V_1^n = \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \left(g\left(c_i^n\right) - g\left(c_s\right) \right) ds,$$

$$V_2^n = \sqrt{\Delta_n} \sum_{i=1}^{n} \sum_{j=1}^{n} K_b\left(t_{j-1} - t_i\right) \nabla g\left(c_i^n\right) \cdot \alpha_j^n,$$

$$V_3^n = \sqrt{\Delta_n} \sum_{i=1}^{n} \sum_{j=1}^{n} K_b\left(t_{j-1} - t_i\right) \nabla g\left(c_i^n\right) \cdot \left(c_{j-1}^n - c_i^n\right) \Delta_n,$$

$$V_4^n = \sqrt{\Delta_n} \sum_{i=1}^{n} \left[\left(g\left(c_i^n + \beta_i^n\right) - g\left(c_i^n\right) - \nabla g(c_i^n) \cdot \beta_i^n \right] \bar{K}_n\left(t_i\right),$$

$$V_5^n = \sqrt{\Delta_n} \sum_{i=1}^{n} \left[\bar{K}_n\left(t_i\right) - 1 \right] g\left(c_i^n\right). \tag{A.4}$$

Above, recall that $x \cdot y = \sum_{\ell,m} x^{\ell m} y^{\ell m}$ for $x, y \in \mathbb{R}^{d \times d}$ (thus, e.g., $\nabla g(c_i^n) \cdot \beta_i^n := \sum_{\ell,m} \partial_{\ell m} g(c_i^n) \beta_i^{n,\ell m}$).

The first term V_1^n converges to 0 since g is C^2 and c_t is an Itô semimartingale with bounded characteristics (see, e.g., Jacod and Protter, 2011, Eq. (5.3.24)). The second term converges stably in law to Z, while the other terms contribute the bias terms. In the subsequence sections, we show the convergence for each of the terms in the decomposition.

A.1.1. *The Leading Term V_2^n.* In this section, we show

$$V_2^n \overset{st}{\Longrightarrow} Z, \tag{A.5}$$

where Z is defined as in the statement of Theorem 2.1. One of the main technical issues of working with two-sided kernels is the lack of adaptability of the terms in V_2^n. In other words, note that

$$V_2^n = \frac{1}{\sqrt{\Delta_n}} \sum_{j=1}^{n} w_j \cdot \alpha_j^n,$$

where $w_j = \Delta_n \sum_{i=1}^{n} K_b(t_{j-1} - t_i) \nabla g(c_i^n) \in \mathbb{R}^{d \times d}$. In contrast to the w_j in Jacod and Rosenbaum (2015, Sect. 3.6), w_j here is not measurable with respect to $\mathcal{F}_j^n := \mathcal{F}_{t_j}$ since the kernel is in general *two-sided*. Therefore, we cannot directly apply the triangular array CLT of Jacod and Protter (2011, Thm. 2.2.14) To solve this problem, we first separate the left- and right-hand side contribution:

$$V_2^n = \frac{1}{\sqrt{\Delta_n}} \sum_{j=1}^{n} w_j \cdot \alpha_j^n = \frac{1}{\sqrt{\Delta_n}} \sum_{j=1}^{n} w_j^R \cdot \alpha_j^n + \frac{1}{\sqrt{\Delta_n}} \sum_{j=1}^{n} w_j^L \cdot \alpha_j^n,$$

where $w_j^R = \Delta_n \sum_{i=1}^{j-1} K_b\left(t_{j-1} - t_i\right) \nabla g\left(c_i^n\right)$ and $w_j^L = \Delta_n \sum_{i=j}^{n} K_b\left(t_{j-1} - t_i\right) \nabla g\left(c_i^n\right)$ and, then, show that

$$\text{(i)} \ \frac{1}{\sqrt{\Delta_n}} \sum_{j=1}^{n} \bar{w}_j^R \cdot \alpha_j^n \overset{st}{\Longrightarrow} Z, \qquad \text{(ii)} \ \frac{1}{\sqrt{\Delta_n}} \sum_{j=1}^{n} \bar{w}_j^L \cdot \alpha_j^n \longrightarrow 0, \tag{A.6}$$

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

where, with the notation $\bar{K}_j^n(t) := \sum_{i=j}^n \Delta_n K_b(t - t_i)$,

$$\bar{w}_j^R := \Delta_n \sum_{i=1}^{j-1} K_b\left(t_{j-1} - t_i\right) \nabla g\left(c_i^n\right) + \sum_{i=j}^{n} \Delta_n K_b\left(t_{j-1} - t_i\right) \nabla g\left(c_{j-1}^n\right),$$
$$\bar{w}_j^L := \Delta_n \sum_{i=j}^{n} K_b\left(t_{j-1} - t_i\right) \nabla g\left(c_i^n\right) - \sum_{i=j}^{n} \Delta_n K_b\left(t_{j-1} - t_i\right) \nabla g\left(c_{j-1}^n\right). \tag{A.7}$$

For (A.6-i), since $w_j^R, \nabla g\left(c_{j-1}^n\right) \in \mathcal{F}_{j-1}^n$ and $\alpha_j^n \in \mathcal{F}_j^n$, we can follow arguments similar to those in Jacod and Rosenbaum (2015, Sect. 3.6) to apply the triangular array CLT of Jacod and Protter (2011). In particular, we need to show that, for any $\ell, m, p, q \in \{1, \ldots, m\}$,

$$\frac{1}{\Delta_n} \sum_{j=1}^{n} \bar{w}_j^{R,\ell m} \bar{w}_j^{R,pq} \mathbb{E}\left(\alpha_j^{\ell m} \alpha_j^{pq} | \mathcal{F}_{j-1}^n\right) \xrightarrow{\mathbb{P}} \int_0^T \partial_{\ell m} g(c_s)\, \partial_{pq} g(c_s) \left(c_s^{\ell p} c_s^{mq} + c_s^{\ell q} c_s^{mp}\right) ds. \tag{A.8}$$

In light of the third estimate of (1.18) in the Supplementary Material (see also Jacod and Rosenbaum 2015, Eq. (3.22)), it suffices to show that

$$\Delta_n \sum_{j=1}^{n} \bar{w}_j^{R,\ell m} \bar{w}_j^{R,pq} \left(c_{j-1}^{n,\ell p} c_{j-1}^{n,mq} + c_{j-1}^{n,\ell q} c_{j-1}^{n,mp}\right)$$
$$\xrightarrow{\mathbb{P}} \int_0^T \partial_{\ell m} g(c_s)\, \partial_{pq} g(c_s) \left(c_s^{\ell p} c_s^{mq} + c_s^{\ell q} c_s^{mp}\right) ds. \tag{A.9}$$

To this end, first note that, for $|s - t_{j-1}| \le \Delta_n$, $c_{j-1}^n \to c_s$ a.s. If we also have $\bar{w}_j^{R,\ell m} \to \partial_{\ell m} g(c_s)$ a.s., whenever $|s - t_{j-1}| \le \Delta_n$, then (A.9) would follow from dominated convergence theorem. For $\left|s - t_{j-1}\right| \le \Delta_n$, note that

$$\begin{aligned}
\bar{w}_j^{R,\ell m} - \partial_{\ell m} g(c_s) &= \Delta_n \sum_{i=1}^{j-1} \left(\partial_{\ell m} g\left(c_i^n\right) - \partial_{\ell m} g(c_s)\right) K_b\left(t_{j-1} - t_i\right) \\
&\quad + \left(\partial_{\ell m} g\left(c_{j-1}^n\right) - \partial_{\ell m} g(c_s)\right) \Delta_n \sum_{i=j}^{n} K_b\left(t_{j-1} - t_i\right) \\
&\quad + \Delta_n \sum_{i=1}^{n} K_b\left(t_{j-1} - t_i\right) \partial_{\ell m} g(c_s) - \partial_{\ell m} g(c_s) \\
&=: P_1 + P_2 + O(\Delta_n / b),
\end{aligned} \tag{A.10}$$

where the order $O(\Delta_n/b)$ follows from (1.21). Since, for $\left|s - t_{j-1}\right| \le \Delta_n$, $\partial_{\ell m} g\left(c_{j-1}^n\right) \to \partial_{\ell m} g(c_s)$ a.s. and $\Delta_n \sum_{i=j}^n K_b\left(t_{j-1} - t_i\right)$ is bounded, $P_2 \xrightarrow{\mathbb{P}} 0$, as $n \to \infty$. We now deal with P_1. Since c is continuous and $\partial_{\ell m} g$ is locally continuous, for an arbitrary, but fixed, $\varepsilon > 0$, there exists a $\delta = \delta(\varepsilon) \in (0, \min(T - s, s))$ such that $|\partial_{\ell m} g(c_u) - \partial_{\ell m} g(c_s)| < \varepsilon$, whenever

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

$|u-s|<\delta$. Next, consider the decomposition

$$P_1=\Delta_n\left(\sum_{0<s-t_i<\delta}+\sum_{T>s-t_i>\delta}\right)\left(\partial_{\ell m}g\left(c_i^n\right)-\partial_{\ell m}g\left(c_s\right)\right)K_b\left(t_{j-1}-t_i\right)=:P_{11}+P_{12}. \tag{A.11}$$

Since $\partial_{\ell m}g$ is locally bounded and c is bounded, for some $s_{i-1}\in\left(t_{i-1},t_i\right)$, we have

$$\begin{aligned}|P_{12}|&\le C\Delta_n\sum_{i:T>s-t_i\ge\delta}\left|K_b\left(t_{j-1}-t_i\right)\right|\\&\le C\sum_{i:T>s-t_i\ge\delta}\int_{\left(t_{j-1}-t_i\right)/b}^{\left(t_{j-1}-t_{i-1}\right)/b}|K(t)|\,dt\\&\quad+\frac{\Delta_n}{b}\sum_{i:T>s-t_i\ge\delta}\left|K\left(\frac{t_{j-1}-t_i}{b}\right)-K\left(\frac{t_{j-1}-s_{i-1}}{b}\right)\right|\\&\le C\int_{\frac{\delta-\Delta_n}{b}}^{+\infty}|K(t)|\,dt+\frac{\Delta_n}{b}V_{-\infty}^{\infty}(K)\overset{n\to\infty}{\longrightarrow}0,\end{aligned} \tag{A.12}$$

where we used $\int_{\left(t_{j-1}-t_i\right)/h}^{\left(t_{j-1}-t_{i-1}\right)/h}|K(t)|\,dt=\frac{\Delta_n}{b}|K\left(\frac{t_{j-1}-s_{i-1}}{b}\right)|$ for some $s_{i-1}\in\left(t_{i-1},t_i\right)$, and $V_{-\infty}^{\infty}(K)$ is the total variation of K. For P_{11},

$$\begin{aligned}|P_{11}|&\le\Delta_n\sum_{i:0<s-t_i<\delta}\left|\partial_{\ell m}g\left(c_i^n\right)-\partial_{\ell m}g\left(c_s\right)\right|\left|K_b\left(t_{j-1}-t_i\right)\right|\\&\le\varepsilon\sum_{r=0}^{\infty}\frac{\Delta_n}{b}\left|K\left(\frac{\Delta_n r}{b}\right)\right|\overset{n\to\infty}{\longrightarrow}\varepsilon.\end{aligned}$$

Therefore, $\limsup_{n\to\infty}|P_1|\le\varepsilon$ and, since ε is arbitrary, we conclude that $\bar{w}_j^{R,\ell m}\to\partial_{\ell m}g(c_s)$ and, thus, (A.9) and (A.6-i) hold true.

Next, we prove (A.6-ii). Note the expression therein can be written as

$$\begin{aligned}\frac{1}{\sqrt{\Delta_n}}\sum_{j=1}^{n}\bar{w}_j^L\alpha_j^n&=\frac{1}{\sqrt{\Delta_n}}\sum_{j=1}^{n}\sum_{i=j}^{n}\Delta_nK_b\left(t_{j-1}-t_i\right)\sum_{u=j}^{i}\left(\nabla g(c_u)-\nabla g(c_{u-1})\right)\cdot\alpha_j^n\\&=\sum_{\ell,m}\sum_{u=1}^{n}\left(\partial_{\ell m}g(c_u)-\partial_{\ell m}g(c_{u-1})\right)\left[\sum_{j=1}^{u}\sum_{i=u}^{n}\sqrt{\Delta_n}K_b(t_{j-1}-t_i)\alpha_j^{n,\ell m}\right]\\&=\sum_{\ell,m}\int_0^T A_s^{n,\ell m}d\upsilon_s,\end{aligned} \tag{A.13}$$

where $s\to\upsilon_s:=\partial_{\ell m}g(c_s)$ is a real-valued function and $A_s^{n,\ell m}:=\sum_{j=1}^{u}\sum_{i=u}^{n}\sqrt{\Delta_n}K_b(t_{j-1}-t_i)\alpha_j^{n,\ell m}$ for $s\in(t_{u-1},t_u)$. We now proceed to show that $A_s^{n,\ell m}\to0$, for all $s\in(0,T)$. Note that $A_s^{n,\ell m}$ is a triangular array. By the estimates (1.17), (1.18), (1.13), and (1.21) in

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

the Supplementary Material, we have

$$\begin{aligned}
&\mathbb{E}\left|\sum_{j=1}^{u}\sum_{i=u}^{n}\sqrt{\Delta_n}K_b(t_{j-1}-t_i)\mathbb{E}\left[\alpha_j^{n,\ell m}\middle|\mathcal{F}_{j-1}\right]\right| \\
&\leq C\sum_{j=1}^{u}\sum_{i=u}^{n}\sqrt{\Delta_n}|K_b(t_{j-1}-t_i)|\Delta_n^{3/2}\left(\sqrt{\Delta_n}+\mathbb{E}\eta_{j-1}^n\right) \\
&\leq C\sum_{j=1}^{u}\sum_{i=u}^{n}\Delta_n|K_b(t_{j-1}-t_i)|\Delta_n^{3/2}+C\sum_{j=1}^{u}\mathbb{E}\eta_{j-1}^n\Delta_n\to 0. \qquad \textbf{(A.14)}
\end{aligned}$$

Furthermore, recalling that $s\in(t_{u-1},t_u)$ and using (1.17) in the Supplementary Material,

$$\begin{aligned}
&\sum_{j=1}^{u}\left(\sum_{i=u}^{n}\sqrt{\Delta_n}K_b(t_{j-1}-t_i)\right)^2\mathbb{E}\left[\left(\alpha_j^{n,\ell m}\right)^2\middle|\mathcal{F}_{j-1}\right] \\
&\leq C\sum_{j=1}^{u}\left(\sum_{i=u}^{n}\sqrt{\Delta_n}K_b(t_{j-1}-t_i)\right)^2\Delta_n^2 \\
&\leq C\int_0^s\left(\int_s^T K_b(t-v)dv\right)^2dt \\
&\leq Cb\int_{-\infty}^0\left(\int_{-\infty}^x|K(w)|dw\right)^2dx\to 0, \qquad \textbf{(A.15)}
\end{aligned}$$

where we used that $\int_{-\infty}^0\left(\int_{-\infty}^x|K(w)|dw\right)^2dx\leq\int_{-\infty}^0\int_{-\infty}^x|K(w)|dwdx=\int_{-\infty}^0|K(w)|(-w)dw<\infty$ by Assumption 2-c. Therefore, $A_s^n\xrightarrow{\mathbb{P}}0$ and $\int_0^T A_s^n dg'(c_s)\to\int_0^T A_s dg'(c_s)=0$. This finishes the proof of (A.6-(ii).

A.1.2. *The Bias Term V_3^n.* We proceed to prove that

$$\begin{aligned}
V_3^{n,\ell m}&:=\Delta_n^{3/2}\sum_{i=1}^{n}\partial_{\ell m}g\left(c_i^n\right)\sum_{j=1}^{n}K_b\left(t_{j-1}-t_i\right)\left(c_{j-1}^{n,\ell m}-c_i^{n,\ell m}\right) \\
&\xrightarrow{\mathbb{P}}\theta\int_0^T\partial_{\ell m}g\left(c_s\right)dc_s^{lm}\int_{-\infty}^{\infty}L(u)du+\theta\sum_{p,q}\int_0^T\partial_{pq,\ell m}g(c_s)\tilde{c}_s^{\ell m,pq}ds\int_{-\infty}^0L(u)du. \qquad \textbf{(A.16)}
\end{aligned}$$

Define

$$\widetilde{V}_3^{n,\ell m}=\Delta_n^{3/2}\sum_{j=1}^{n}\sum_{i=j}^{n}K_b\left(t_{j-1}-t_i\right)\partial_{\ell m}g\left(c_{j-1}^n\right)\left[c_{j-1}^{n,\ell m}-c_i^{n,\ell m}\right]. \qquad \textbf{(A.17)}$$

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

Similarly to the proof of the leading term V_2^n, we write $V_3^{n,\ell m}$ as

$$V_3^{n,\ell m} = \Delta_n^{3/2} \sum_{j=1}^{n} \left(\sum_{i=j}^{n} + \sum_{i=1}^{j-1} \right) K_b\left(t_{j-1} - t_i\right) \partial_{\ell m} g\left(c_i^n\right) \left[c_{j-1}^{n,\ell m} - c_i^{n,\ell m} \right] \tag{A.18}$$
$$=: L + R,$$

and aim to prove that

$$R + \widetilde{V}_3^{n,\ell m} \to \theta \int_0^T \partial_{\ell m} g\left(c_t\right) dc_t^{lm} \int_{-\infty}^{\infty} L(u) du, \tag{A.19}$$

$$L - \widetilde{V}_3^{n,\ell m} \to \theta \sum_{p,q} \int_0^T \partial_{pq,\,\ell m} g(c_s) \tilde{c}_s^{\ell m, pq} ds \int_{-\infty}^{0} L(u) du. \tag{A.20}$$

We first show (A.19). Writing $c_{j-1}^{n,\ell m} - c_i^{n,\ell m} = \sum_{u=i+1}^{j-1} \Delta_u^n c^{\ell m}$, when $j-1 \geq i$, and $c_{j-1}^{n,\ell m} - c_i^{n,\ell m} = -\sum_{u=j}^{i} \Delta_u^n c^{\ell m}$, when $j-1 < i$, and using the usual convention that $\sum_{j=n+1}^{n} = \sum_{i=1}^{0} = 0$, we have

$$\begin{aligned}
R + \widetilde{V}_3^{n,\ell m} &= \Delta_n^{3/2} \sum_{j=1}^{n} \sum_{i=1}^{j-1} K_b\left(t_{j-1} - t_i\right) \partial_{\ell m} g\left(c_i^n\right) \left[c_{j-1}^{n,\ell m} - c_i^{n,\ell m} \right] \\
&\quad + \Delta_n^{3/2} \sum_{j=1}^{n} \sum_{i=j}^{n} K_b\left(t_{j-1} - t_i\right) \partial_{\ell m} g\left(c_{j-1}^n\right) \left[c_{j-1}^{n,\ell m} - c_i^{n,\ell m} \right] \\
&= \Delta_n^{3/2} \sum_{j=2}^{n} \sum_{i=1}^{j-2} K_b\left(t_{j-1} - t_i\right) \partial_{\ell m} g\left(c_i^n\right) \sum_{u=i+1}^{j-1} \Delta_u^n c^{\ell m} \\
&\quad - \Delta_n^{3/2} \sum_{j=1}^{n} \sum_{i=j}^{n} K_b\left(t_{j-1} - t_i\right) \partial_{\ell m} g\left(c_{j-1}^n\right) \sum_{u=j}^{i} \Delta_u^n c^{\ell m} \\
&= \Delta_n^{3/2} \sum_{u=1}^{n} \Delta_u^n c^{\ell m} \left[\sum_{j=u+1}^{n} \sum_{i=1}^{u-1} K_b\left(t_{j-1} - t_i\right) \partial_{\ell m} g\left(c_i^n\right) \right. \\
&\quad \left. - \sum_{j=1}^{u} \sum_{i=u}^{n} K_b\left(t_{j-1} - t_i\right) \partial_{\ell m} g\left(c_{j-1}^n\right) \right] \\
&=: \int_0^T H(n)_t dc_t^{\ell m},
\end{aligned} \tag{A.21}$$

where, for $t \in ((u-1)\Delta_n, u\Delta_n)$,

$$H(n)_t := \Delta_n^{3/2} \left[\sum_{j=u+1}^{n} \sum_{i=1}^{u-1} K_b\left(t_{j-1} - t_i\right) \partial_{\ell m} g\left(c_i^n\right) - \sum_{j=1}^{u} \sum_{i=u}^{n} K_b\left(t_{j-1} - t_i\right) \partial_{\ell m} g\left(c_{j-1}^n\right) \right]. \tag{A.22}$$

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

Note $H(n)_t$ is adapted to $\mathcal{F}_t$. We first consider the first term in (A.22). It can be shown (see Section 2.1.2 of the Supplementary Material) that

$$\Delta_n^{\frac{3}{2}} \sum_{j=u+1}^{n} \sum_{i=1}^{u-1} K_b\left(t_{j-1}-t_i\right) \partial_{\ell m} g\left(c_i^n\right) - \Delta_n^{-\frac{1}{2}} \int_{t_{u+1}}^{T} \int_0^{t_u} K_b(s-\varpi)\partial_{\ell m} g(c_\varpi) d\varpi\, ds = o_p(1). \tag{A.23}$$

Next, we show that $\Delta_n^{-1/2} \int_{t_{u+1}}^{T} \int_0^{t_u} K_b(s-\varpi)\partial_{\ell m} g(c_\varpi) d\varpi\, ds$ converges to $\theta \partial_{\ell m} g(c_t) \int_0^\infty L(v) dv$. Indeed, since

$$\Delta_n^{-\frac{1}{2}} \int_{t_u}^{t_{u+1}} \int_0^{t_u} |K_b(s-\varpi)\partial_{\ell m} g(c_\varpi)| d\varpi\, ds \le C \Delta_n^{-\frac{1}{2}} \int_{t_u}^{t_{u+1}} \int_{\frac{s-t_u}{b}}^{\frac{s}{b}} |K(v)| dv ds = O(\Delta_n^{\frac{1}{2}}),$$

we only need to consider $\Delta_n^{-1/2} \int_{t_u}^{T} \int_0^{t_u} K_b(s-\varpi)\partial_{\ell m} g(c_\varpi) d\varpi\, ds$. Let us write $\tau = t_u$ for simplicity. We exploit the decomposition

$$\Delta_n^{-\frac{1}{2}} \int_{\tau}^{T} \int_0^{\tau} K_b(s-\varpi)\partial_{\ell m} g(c_\varpi) d\varpi\, ds = \Delta_n^{-\frac{1}{2}} \left(\iint_{S_\delta} + \iint_{S_\delta^c} \right) K_b(s-\varpi)\partial_{\ell m} g(c_\varpi) d\varpi\, ds,$$

where $S_\delta = \{(s,\varpi) : s-\varpi < \delta, s \in (\tau, T), \varpi \in (0,\tau)\}$ for $\delta > 0$ and $S_\delta^c = [\tau, T] \times [0, \tau] - S_\delta$. Letting $w = \frac{s-\varpi}{b}$ first and then $v = \frac{s}{b}$, we can write

$$\begin{aligned}
&\Delta_n^{-1/2} \left| \iint_{S_\delta^c} K_b(s-\varpi)\partial_{\ell m} g(c_\varpi) d\varpi\, ds \right| \\
&= \Delta_n^{-1/2} \left| \int_{\tau}^{\tau+\delta} \int_0^{s-\delta} K_b(s-\varpi)\partial_{\ell m} g(c_\varpi) d\varpi\, ds + \int_{\tau+\delta}^{T} \int_0^{\tau} K_b(s-\varpi)\partial_{\ell m} g(c_\varpi) d\varpi\, ds \right| \\
&\le C \Delta_n^{-1/2} \left[\int_{\tau}^{\tau+\delta} \left(\int_{\delta/b}^{s/b} |K(w)| dw \right) ds + \int_{\tau+\delta}^{T} \left(\int_{(s-\tau)/b}^{s/b} |K(w)| dw \right) ds \right] \\
&\le C \Delta_n^{-1/2} b \left[\int_{\tau/b}^{(\tau+\delta)/b} \left(\int_{\delta/b}^{v} |K(w)| dw \right) dv + \int_{(\tau+\delta)/b}^{T/b} \left(\int_{v-\tau/b}^{v} |K(w)| dw \right) dv \right] \\
&\le C \left[\bar{L}(\delta/b) \frac{\delta}{b} + \bar{L}(\delta/b) \frac{T}{b} \right] \to 0,
\end{aligned} \tag{A.24}$$

for any $\delta > 0$, where $\bar{L}(x) = \int_x^\infty |K(u)| du$ and the last equation uses $\bar{L}(x) x \overset{x \to \infty}{\longrightarrow} 0$ derived from $K(x) x^{2+\epsilon} \to 0$, $\epsilon > 1/2$ as $|x| \to \infty$.

Next, we consider the integral over S_δ. Clearly, for any $\epsilon > 0$, there exists $\delta > 0$ such that $|\partial_{\ell m} g(c_\varpi) - \partial_{\ell m} g(c_\tau)| < \epsilon$, for $|\varpi - \tau| < \delta$. Letting $w = \frac{s-\varpi}{b}$ first and then $v = \frac{s-\tau}{b}$, we can write

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

$$
\begin{aligned}
&\Delta_n^{-1/2}\int_{S_\delta} K_b(s-\varpi)\partial_{\ell m}g(c_\varpi)d\varpi\,ds\\
&=\Delta_n^{-1/2}\int_\tau^{\tau+\delta}\int_{s-\delta}^{\tau} K_b(s-\varpi)\partial_{\ell m}g(c_\varpi)d\varpi\,ds\\
&\le \Delta_n^{-1/2}\int_\tau^{\tau+\delta}\left(\int_{(s-\tau)/b}^{\delta/b} K(w)\left(\partial_{\ell m}g(c_\tau)+\epsilon\right)1_{K(w)>0}dw\right)ds\\
&\quad+\Delta_n^{-1/2}\int_\tau^{\tau+\delta}\left(\int_{(s-\tau)/b}^{\delta/b} K(w)\left(\partial_{\ell m}g(c_\tau)-\epsilon\right)1_{K(w)<0}dw\right)ds\\
&\le b\Delta_n^{-1/2}\left[\int_0^{\delta/b}\int_v^{\delta/b} K(w)\partial_{\ell m}g(c_\tau)dwdv+\epsilon\int_0^\infty\int_v^\infty |K(w)|dwdv\right]. \qquad \textbf{(A.25)}
\end{aligned}
$$

Similarly, we can find a lower bound of the form:

$$
b\Delta_n^{-1/2}\left[\int_0^{\delta/b}\int_v^{\delta/b} K(w)\partial_{\ell m}g(c_\tau)dwdv-\epsilon\int_0^\infty\int_v^\infty |K(w)|dwdv\right]. \qquad \textbf{(A.26)}
$$

Recall $b\Delta_n^{-1/2}\sim\theta$ and $c_\tau=c_{t_u}\overset{n\to\infty}{\longrightarrow}c_t$. Taking $n\to\infty$ and then $\epsilon\to 0$, we have

$$
\Delta_n^{-1/2}\int_\tau^T\int_0^\tau K_b(s-\varpi)\partial_{\ell m}g(c_\varpi)d\varpi\,ds\overset{n\to\infty}{\longrightarrow}\theta\partial_{\ell m}g(c_t)\int_0^\infty L(v)dv. \qquad \textbf{(A.27)}
$$

For the second term in (A.22), similar to (A.23), we can show that

$$
\Delta_n^{\frac{3}{2}}\sum_{j=1}^u\sum_{i=u}^n K_b\left(t_{j-1}-t_i\right)\partial_{\ell m}g\left(c_{j-1}^n\right)-\Delta_n^{-\frac{1}{2}}\int_0^{t_u}\int_{t_u}^T K_b(\varpi-s)\partial_{\ell m}g(c_\varpi)dsd\varpi=o_P(1). \qquad \textbf{(A.28)}
$$

Note that $\Delta_n^{-\frac{1}{2}}\int_0^{t_u}\int_{t_u}^T K_b(\varpi-s)\partial_{\ell m}g(c_\varpi)dsd\varpi$ is the same as the left-hand side in (A.27) but with $K(-w)$ instead of $K(w)$. Therefore,

$$
\Delta_n^{-\frac{1}{2}}\int_0^{t_u}\int_{t_u}^T K_b(\varpi-s)\partial_{\ell m}g(c_\varpi)dsd\varpi\overset{n\to\infty}{\longrightarrow}\theta\partial_{\ell m}g(c_t)\int_0^\infty\int_v^\infty K(-w)dwdv.
$$

We conclude that

$$
H(n)_t\to\theta\partial_{\ell m}g(c_t)\int_{-\infty}^\infty L(v)dv, \qquad \textbf{(A.29)}
$$

where we recall that $L(t):=\int_t^\infty K(u)du\mathbf{1}_{\{t>0\}}-\int_{-\infty}^t K(u)du\mathbf{1}_{\{t\le 0\}}$. Since c is a.s. continuous, (A.19) follows by dominated convergence theorem for stochastic integrals.

Next, we proceed to show (A.20). Note that

$$
L-\widetilde{V}_3^{n,\ell m}=\Delta_n^{3/2}\sum_{j=1}^n\sum_{i=j}^n K_b\left(t_{j-1}-t_i\right)\left(\partial_{\ell m}g\left(c_i^n\right)-\partial_{\ell m}g\left(c_{j-1}^n\right)\right)\left[c_{j-1}^{n,\ell m}-c_i^{n,\ell m}\right]. \qquad \textbf{(A.30)}
$$

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

Let us start by rewriting the inner summation in (A.30). Applying the conclusion in Remark 1.1 in the Supplementary Material with $\tau = t_{j-1}, f(x) = (\partial_{\ell m} g(c^n_{j-1}) - \partial_{\ell m} g(x))(c^{n,\ell m}_{j-1} - x^{\ell m})$, and $K(-x)1_{x>0}$ instead of $K(x)$, we have $f(c^n_{j-1}) = 0$ and

$$\begin{aligned}
&\sum_{i=j}^{n} K_b\left(t_{j-1} - t_i\right)\left(\partial_{\ell m} g\left(c^n_{j-1}\right) - \partial_{\ell m} g\left(c^n_i\right)\right)\left[\left(c^{n,\ell m}_{j-1} - c^{n,\ell m}_i\right)\Delta_n\right] \\
&\quad = \sum_{i=j+1}^{n+1} K_b\left(t_{j-1} - t_{i-1}\right) f(c^n_{i-1})\Delta_n \\
&\quad = \int_{t_{j-1}}^{T} K_b\left(t_{j-1} - s\right) f(c_s) ds + D_1(f) + \sum_{i=j+1}^{n+1} K_b\left(t_{j-1} - t_{i-1}\right) \int_{t_{i-1}}^{t_i} \left(f(c^n_{i-1}) - f(c_s)\right) ds \\
&\quad = \int_{t_{j-1}}^{T} K_b\left(t_{j-1} - s\right) f(c_s) ds + o_p\left(\frac{\Delta_n}{b}\right) + O_p\left(\sqrt{\Delta_n b}\right), \qquad \textbf{(A.31)}
\end{aligned}$$

where the order $O_p(\sqrt{\Delta_n b})$ can be justified since, for $s \in ((i-1)\Delta_n, i\Delta_n]$ and $i > j$,

$$\begin{aligned}
\mathbb{E}\left|f(c^n_{i-1}) - f(c_s)\right| &\le \sqrt{\mathbb{E}(c^{n,\ell m}_{j-1} - c^{\ell m}_s)^2 \mathbb{E}(\partial_{\ell m} g(c^n_{i-1}) - \partial_{\ell m} g(c_s))^2} \\
&\quad + \sqrt{\mathbb{E}(\partial_{\ell m} g(c^n_{j-1}) - \partial_{\ell m} g(c^n_{i-1}))^2 \mathbb{E}(c^{n,\ell m}_{i-1} - c^{\ell m}_s)^2} \\
&\le C\sqrt{\Delta_n (t_i - t_{j-1})},
\end{aligned}$$

for some constant C, and, therefore, for some possibly different constant C,

$$\begin{aligned}
&\sum_{i=j+1}^{n+1} |K_b\left(t_{j-1} - t_{i-1}\right)| \sqrt{\Delta_n |t_{j-1} - t_i|}\Delta_n \le C \int_{t_{j+1}}^{T} |K_b\left(t_{j-1} - s\right)| \sqrt{\Delta_n |t_{j-1} - s|} ds \\
&\quad \le C \int_{-\infty}^{0} |K(u)| \sqrt{\Delta_n |u| b} ds = O_p\left(\sqrt{\Delta_n b}\right).
\end{aligned}$$

Plugging (A.31) back into (A.30), recalling that $f(x) = (\partial_{\ell m} g(c^n_{j-1}) - \partial_{\ell m} g(x))(c^{n,\ell m}_{j-1} - x^{\ell m})$, and applying integration by parts, we get

$$\begin{aligned}
&L - \widetilde{V}^{n,\ell m}_3 \\
&= -\sqrt{\Delta_n} \sum_{j=1}^{n} \left[\int_{t_{j-1}}^{T} K_b\left(t_{j-1} - t\right) \left(\partial_{\ell m} g\left(c^n_{j-1}\right) - \partial_{\ell m} g\left(c_s\right)\right) \left(c^{n,\ell m}_{j-1} - c_s{}^{\ell m}\right) ds \right] + o_p(1)
\end{aligned}$$

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

$$= -\sqrt{\Delta_n}\sum_{j=1}^{n} L\left(\frac{t_{j-1}-T}{b}\right)\left(\partial_{\ell m}g\left(c_{j-1}^n\right)-\partial_{\ell m}g\left(c_T\right)\right)\left(c_{j-1}^{n,\ell m}-c_T^{\ell m}\right)$$

$$+\sqrt{\Delta_n}\sum_{j=1}^{n}\int_{t_{j-1}}^{T} L\left(\frac{t_{j-1}-t}{b}\right)b(j)_s dW_s+\sqrt{\Delta_n}\sum_{j=1}^{n}\int_{t_{j-1}}^{T} L\left(\frac{t_{j-1}-t}{b}\right)a(j)_s ds+o_p(1),$$

$$=: P_1+P_2+P_3+o_p(1), \tag{A.32}$$

where, recalling the notation $\tilde{c}_s^{uv,pq}=\sum_{j=1}^{d}\tilde{\sigma}^{uv,j}\tilde{\sigma}^{pq,j}$,

$$b(j)_s := \sum_{pq}(c_s^{\ell m}-c_{j-1}^{n,pq})\partial^2_{pq,\ell m}g(c_s)\tilde{\sigma}_s^{pq}+(\partial_{\ell m}g(c_s)-\partial_{\ell m}g(c_{j-1}^n))\tilde{\sigma}_s^{\ell m}$$

$$a(j)_s := \sum_{pq}(c_s^{\ell m}-c_{j-1}^{n,pq})\partial^2_{pq,\ell m}g(c_s)\tilde{\mu}_s^{pq}+(\partial_{\ell m}g(c_s)-\partial_{\ell m}g(c_{j-1}^n))\tilde{\mu}_s^{\ell m}$$

$$+\frac{1}{2}\sum_{uv,pq}(c_s^{\ell m}-c_{j-1}^{n,\ell m})\partial^3_{uv,pq,\ell m}g(c_s)\tilde{c}_s^{uv,pq}+\sum_{pq}\partial_{pq,\ell m}g(c_s)\tilde{c}_s^{\ell m,pq}.$$

For the first term, for some ξ in between c_T and c_{j-1}^n, we have

$$\mathbb{E}|P_1| \le \sqrt{\Delta_n}\sum_{j=1}^{n}\left|L\left(\frac{t_{j-1}-T}{b}\right)\right|\mathbb{E}\left[\|\nabla\partial_{\ell m}g(\xi)\|\left\|c_{j-1}^n-c_T^n\right\|^2\right]$$

$$\le C\sqrt{\Delta_n}\sum_{j=1}^{n}\left|L\left(\frac{t_{j-1}-T}{b}\right)\right|\left(T-t_{j-1}\right)$$

$$\le C\frac{b^2}{\sqrt{\Delta_n}}\int_0^T |L(u)|u du\to 0,$$

where we used (1.16-iii) in the Supplementary Material in the second inequality above.

For P_2, let $\zeta_j^n=\sqrt{\Delta_n}\int_{t_{j-1}}^{T}L\left(\frac{t_{j-1}-t}{b}\right)b(j)_s dW_s$ so that $P_2=\sum_{j=1}^{n}\zeta_j^n$. Observe that for $j'<j$, $\mathbb{E}\left[\zeta_{j'}^n\zeta_j^n\right]=\mathbb{E}\left[\zeta_{j'}^n\mathbb{E}\left[\zeta_j^n\mid\mathcal{F}_{j-1}^n\right]\right]=0$ because $\mathbb{E}\left[\zeta_j^n\mid\mathcal{F}_{j-1}^n\right]=0$ and, thus,

$$\mathbb{E}\left[P_2^2\right]=\sum_{j=1}^{n}\mathbb{E}\left[\left(\zeta_j^n\right)^2\right]$$

$$=\Delta_n\sum_{j=1}^{n}\int_{t_{j-1}}^{T}L^2\left(\frac{t_{j-1}-s}{b}\right)\mathbb{E}\left(\|b(j)_s\|^2\right)ds$$

$$\le C\Delta_n\sum_{j=1}^{n}\int_{t_{j-1}}^{T}L^2\left(\frac{t_{j-1}-s}{b}\right)\left(s-t_{j-1}\right)ds$$

$$\le Cb^2\Delta_n\sum_{j=1}^{n}\int_{-\infty}^{0}L^2(u)(-u)du\to 0, \tag{A.33}$$

where in the first inequality, we used that $\mathbb{E}\|b(j)_s\|^2\le C\left|s-t_{j-1}\right|$, which follows from the definition of $b(j)_s$ and (1.16) in the Supplementary Material.

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

For P_3, recall $b \sim \theta\sqrt{\Delta_n}$ and note that $a(j)_s = O_p(\sqrt{|s-t_{j-1}|}) + \sum_{pq} \partial_{pq,\ell m} g(c_s) \tilde{c}_s^{\ell m,pq}$ again due to (1.16) in the Supplementary Material. We then have

$$\begin{aligned}
P_3 &= \sqrt{\Delta_n} \sum_{j=1}^{n} \int_{t_{j-1}}^{T} L\left(\frac{t_{j-1}-s}{b}\right) a(j)_s ds \\
&= o_p(1) + \sqrt{\Delta_n} \sum_{pq} \sum_{j=1}^{n} \int_{t_{j-1}}^{T} L\left(\frac{t_{j-1}-s}{b}\right) \partial_{pq,\ell m} g(c_s) \tilde{c}_s^{\ell m,pq} ds \\
&= o_p(1) + \sqrt{\Delta_n} b \sum_{pq} \sum_{j=1}^{n} \partial_{pq,\ell m} g(c_{t_{j-1}}) \tilde{c}_{t_{j-1}}^{\ell m,pq} \int_{\frac{t_{j-1}-T}{b}}^{0} L(u)\, du \\
&\overset{n\to\infty}{\longrightarrow} \theta \sum_{pq} \int_0^T \partial_{pq,\ell m} g(c_s) \tilde{c}_s^{\ell m,pq} ds \int_{-\infty}^{0} L(u) du, \qquad \textbf{(A.34)}
\end{aligned}$$

since

$$\sqrt{\Delta} \sum_{j=1}^{n} \int_{t_{j-1}}^{T} |L\left(\frac{t_{j-1}-s}{b}\right)| \sqrt{s-t_{j-1}} ds \le C b^{\frac{3}{2}} \Delta^{-\frac{1}{2}} \int_{-\infty}^{0} |L(u)| \sqrt{-u} du = O(\Delta^{\frac{1}{4}}),$$

and, applying again (1.16) in the Supplementary Material,

$$\begin{aligned}
&\mathbb{E}\left| \int_{t_{j-1}}^{T} L\left(\frac{t_{j-1}-s}{b}\right) \partial_{pq,\ell m} g(c_s) \tilde{c}_s^{\ell m,pq} ds - \partial_{pq,\ell m} g(c_{t_{j-1}}) \tilde{c}_{t_{j-1}}^{\ell m,pq} h \int_{\frac{t_{j-1}-T}{b}}^{0} L(u)\, du \right| \\
&\le \int_{t_{j-1}}^{T} \left| L\left(\frac{t_{j-1}-s}{b}\right) \right| \mathbb{E}\left| \partial_{pq,\ell m} g(c_s) \tilde{c}_s^{\ell m,pq} - \partial_{pq,\ell m} g(c_{t_{j-1}}) \tilde{c}_{t_{j-1}}^{\ell m,pq} \right| ds \\
&\le C \int_{t_{j-1}}^{T} \left| L\left(\frac{t_{j-1}-s}{b}\right) \right| \sqrt{s-t_{j-1}} ds \\
&\le Cb \int_{-\infty}^{0} |L(u)| \sqrt{-ub} du = O(b^{3/2}), \qquad \textbf{(A.35)}
\end{aligned}$$

where $\int_{-\infty}^{0} |L(u)| \sqrt{-u} du < \infty$ can be shown using the condition $K(x) = O(|x|^{-2-\epsilon})$, as $|x| \to \infty$, for some $\epsilon > 1/2$. Indeed, for some $M < 0$ and $B < \infty$,

$$\begin{aligned}
\int_{-\infty}^{0} |L(u)| \sqrt{-u} du &\le \int_{M}^{0} |L(u)| \sqrt{-u} du + \int_{-\infty}^{M} \sqrt{-u} \int_{-\infty}^{u} |K(x)| dx du \\
&\le \int_{M}^{0} |L(u)| \sqrt{-u} du + B \int_{-\infty}^{M} \sqrt{-u} \int_{-\infty}^{u} |x|^{-2-\epsilon} dx du < \infty.
\end{aligned}$$

This concludes $L - \widetilde{V}_3^{n,\ell m} \to \theta \sum_{pq} \int_0^T \partial_{pq,\ell m} g(c_s) \tilde{c}_s^{\ell m,pq} ds \int_{-\infty}^{0} L(u) du$ and, hence, (A.16).

A.1.3. *The Bias Term V_4^n.* Recall that

$$V_4^n = \sum_{i=1}^{n} \sqrt{\Delta_n} \left(g\left(c_i^n + \beta_i^n\right) - g\left(c_i^n\right) - \nabla g(c_i^n) \cdot \beta_i^n \right) \bar{K}_n(t_i) =: \sum_{i=1}^{n} \nu_i^n \bar{K}_n(t_i).$$

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

We can show that

$$V_4^n \xrightarrow{\mathbb{P}} \frac{1}{2\theta}\sum_{p,q,u,v}\int_0^T \partial_{pq,uv}g(c_s)\check{c}_s^{pq,uv}ds\int_{-\infty}^{\infty}K^2(u)\,du + \frac{\theta}{2}\sum_{p,q,u,v}\sum_{r=1}^{d}\int_0^T \partial_{pq,uv}g(c_s)\,\tilde{\sigma}_s^{pq,r}\tilde{\sigma}_s^{uv,r}ds\left(\int_{-\infty}^{\infty}L(u)^2\,du\right). \tag{A.36}$$

By Taylor's expansion and (1.9) in the Supplementary Material, for some constants $\ell \geq 3$, $v_i^n = v(1)_i^n + v(2)_i^n$, where

$$v(1)_i^n = \frac{\sqrt{\Delta_n}}{2}\sum_{p,q,u,v}\partial_{pq,uv}g(c_i^n)\beta_i^{n,pq}\beta_i^{n,uv},$$

$$|v(2)_i^n| \leq C\sqrt{\Delta_n}(1+\|\beta_i^n\|^{\ell-3})\|\beta_i^n\|^3.$$

For $v(2)_i^n$, applying (1.20) in the Supplementary Material and recalling that $k_n \sim \frac{\theta}{\sqrt{\Delta_n}}$,

$$\mathbb{E}\left[\sum_{i=1}^{n}|v(2)_i^n|\right] \leq \sum_{i=1}^{n}C\sqrt{\Delta_n}\left(k_n^{-3/2}+(k_n\Delta_n)^{3/2}\right) \leq \sum_{i=1}^{n}C\sqrt{\Delta_n}\Delta_n^{3/4} \to 0.$$

Thus, we only need to consider $v(1)_i^n$. Recalling the decomposition of β_i^n in (A.2), we can write

$$\beta_i^{n,pq}\beta_i^{n,uv} = \bar{K}(t_i)^{-2}\sum_{r=1}^{6}\xi_{r,i}^n, \tag{A.37}$$

where

$$\xi_{1,i}^n := \sum_{j=1}^{n}K_b^2(t_{j-1}-t_i)\,\alpha_j^{n,pq}\alpha_j^{n,uv},$$

$$\xi_{2,i}^n = \Delta_n^2\left(\sum_{j=1}^{n}K_b(t_{j-1}-t_i)\left(c_{j-1}^{n,pq}-c_i^{n,pq}\right)\right)\left(\sum_{j=1}^{n}K_b(t_{j-1}-t_i)\left(c_{j-1}^{n,uv}-c_i^{n,uv}\right)\right)$$

$$\xi_{3,i}^n = \sum_{j=1}^{n}\sum_{l=j+1}^{n}K_b(t_{j-1}-t_i)K_b(t_{l-1}-t_i)\,\alpha_j^{n,uv}\alpha_l^{n,pq},$$

$$\xi_{4,i}^n = \sum_{l=1}^{n-1}\sum_{j=l+1}^{n}K_b(t_{j-1}-t_i)K_b(t_{l-1}-t_i)\,\alpha_j^{n,uv}\alpha_l^{n,pq},$$

$$\xi_{5,i}^n = \sum_{j=1}^{n}\sum_{l=1}^{n}K_b(t_{j-1}-t_i)K_b(t_{l-1}-t_i)\,\alpha_j^{n,uv}\left(c_{l-1}^{n,pq}-c_i^{n,pq}\right)\Delta_n,$$

$$\xi_{6,i}^n = \sum_{j=1}^{n}\sum_{l=1}^{n}K_b(t_{j-1}-t_i)K_b(t_{l-1}-t_i)\,\alpha_j^{n,pq}\left(c_{l-1}^{n,uv}-c_i^{n,uv}\right)\Delta_n.$$

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

Therefore, we only need to consider the convergence of

$$S_{n,r} := S_{n,r}^{(p,q,u,v)} := \sum_{i=1}^{n} \frac{\sqrt{\Delta_n}}{2} \partial_{pq,uv} g(c_i^n) \bar{K}(t_i)^{-1} \xi_{r,i}^n, \quad r = 1, \dots, 6, \tag{A.38}$$

for each p, q, u, v, since

$$V_4^n = \sum_{r=1}^{6} \sum_{p,q,u,v} S_{n,r}^{(p,q,u,v)} + o_p(1).$$

In Section 2.1.3 of the Supplementary Material, we show that

$$S_{n,1} \xrightarrow{n\to\infty} \frac{1}{2\theta} \int_0^T \partial_{pq,uv} g(c_s) \check{c}_s^{pq,uv} ds \int_{-\infty}^{\infty} K^2(u)\, du, \tag{A.39}$$

$$S_{n,2} \xrightarrow{\mathbb{P}} \frac{\theta}{2} \sum_{r=1}^{d} \int_0^T \partial_{pq,uv} g(c_s)\, \tilde{\sigma}_s^{pq,r} \tilde{\sigma}_s^{uv,r} ds \left(\int_{-\infty}^{\infty} L(u)^2\, du \right), \tag{A.40}$$

$$S_{n,r} \to_p 0, \quad \text{for } r = 3, \dots, 6. \tag{A.41}$$

Therefore, we obtain (A.36).

A.1.4. *The Bias Term V_5^n.* In Section 2.14 of the Supplementary Material, we show that

$$V_5^n \to \theta \int_{-\infty}^{0} L(u) du g(c_0) - \theta \int_0^{\infty} L(u) du g(c_T). \tag{A.42}$$

A.1.5. *Eliminating $\theta \int_0^T \sum_{\ell,m=1}^{d} \partial_{\ell m} g(c_s)\, dc_s^{lm}$.* From (A.5), (A.16), (A.36), and Eq. (2.3) in the Supplementary Material, we deduce that

$$\frac{1}{\sqrt{\Delta_n}} \left(V(g)_T^n - V(g)_T \right) \xrightarrow{st} \widetilde{A}^1 + \widetilde{A}^2 + \widetilde{A}^3 + Z, \tag{A.43}$$

with

$$\widetilde{A}^1 := \theta \int_{-\infty}^{0} L(u) du g(c_0) - \theta \int_0^{\infty} L(u) du g(c_T),$$

$$\widetilde{A}^2 := \frac{1}{2\theta} \int_0^T \sum_{p,q,u,v=1}^{d} \partial_{pq,uv}^2 g(c_s)\, \check{c}_s^{pq,uv} ds \int_{-\infty}^{\infty} K^2(u) du,$$

$$\widetilde{A}^3 := \theta \int_0^T \sum_{\ell,m=1}^{d} \partial_{\ell m} g(c_s)\, dc_s^{lm} \int_{-\infty}^{\infty} L(u) du + \theta \int_0^T \sum_{p,q,u,v=1}^{d} \partial_{pq,uv}^2 g(c_s)\, \check{c}_s^{pq,uv} ds \left[\int_{-\infty}^{0} L(u) du + \frac{1}{2} \int_{-\infty}^{\infty} L(u)^2 du \right].$$

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

On the other hand, by Itô's formula, we have

$$g(c_T) = g(c_0) + \int_0^T \sum_{l,m} \partial_{lm} g(c_s)\, dc_s^{lm} + \frac{1}{2}\int_0^T \sum_{p,q,u,v} \partial^2_{pq,uv} g(c_s)\, \tilde{c}_s^{pq,uv} ds.$$

Hence, solving for the second term and plugging in the expression of $\widetilde{A}^3$, we can rewrite the bias $\widetilde{A}^1 + \widetilde{A}^2 + \widetilde{A}^3$ as $A^1 + A^2 + A^3$ with the terms A^j's given as in (2.15).

A.2. Proof of Corollary 2.1

To show the result, we only need to eliminate the bias terms A^1, A^2, A^3 of the CLT (2.13). Note that

$$\begin{aligned} & k_n\sqrt{\Delta_n} g\left(\hat{c}_1^n\right) \int_0^\infty L(u)du - k_n\sqrt{\Delta_n} g\left(\hat{c}_n^n\right) \int_{-\infty}^0 L(u)du \\ & \quad \to \theta g(c_0) \int_0^\infty L(u)du - \theta g(c_T) \int_{-\infty}^0 L(u)du, \end{aligned}$$

which cancels the term A^1. Then, by (2.21) and Theorem 2.2, we have

$$\begin{aligned} & -\frac{\sqrt{\Delta_n}}{4} \sum_{i=1}^{n-2k_n+1} \sum_{p,q,u,v=1}^{d} \partial^2_{pq,uv} g\left(\hat{c}_i^n\right) \left(\hat{c}_{i+k_n}^{n,pq} - \hat{c}_i^{n,pq}\right)\left(\hat{c}_{i+k_n}^{n,uv} - \hat{c}_i^{n,uv}\right) C_{K,1} \\ & = -2C_{K,1}\widehat{A}^{n,3} \\ & \to -\frac{1}{2\theta}\int_0^T \sum_{j,k,l,m} \partial^2_{jk,lm} g(c_s)\, \check{c}_s^{jk,lm} ds C_{K,2} \\ & \quad -\frac{\theta}{2}\int_0^T \sum_{j,k,l,m} \partial^2_{jk,lm} g(c_s)\, \tilde{c}_s^{jk,lm} ds \left[\int_{-\infty}^\infty L(z)^2 dz + \left(\int_{-\infty}^0 - \int_0^\infty\right) L(z)dz\right], \end{aligned}$$

and

$$\begin{aligned} & \frac{1}{2k_n\sqrt{\Delta_n}} V(h)_T^n \left[C_{K,2} - \int K^2(u)du\right] \\ & \quad \to \frac{1}{2\theta}\int_0^T \sum_{p,q,u,v=1}^{d} \partial^2_{pq,uv} g(c_s)\, \check{c}_s^{pq,uv} ds \left[C_{K,2} - \int K^2(u)du\right]. \end{aligned}$$

Supplementary Material

Figueroa-López, J. E., Pang, J. & Wu, B. (2025): Supplement to "Estimation of Integrated Volatility Functionals with Kernel Spot Volatility Estimators," Econometric Theory Supplementary Material. To view, please visit: https://doi.org/10.1017/S0266466625100170.

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press

REFERENCES

Aït-Sahalia, Y., & Jacod, J. (2014). *High-frequency financial econometrics*. Princeton University Press.

Aït-Sahalia, Y., & Xiu, D. (2019). Principal component analysis of high-frequency data. *Journal of the American Statistical Association*, 114(525), 287–303.

Benvenuti, F. (2021). *Essays on high-frequency and financial data analysis* [Ph.D. thesis]. Aarhus University.

Chen, R. Y. (2019). Inference for volatility functionals of multivariate itô semimartingales observed with jump and noise. Tech. rep., Working Paper. arXiv:1810.04725v2.

Clément, E., Delattre, S., & Gloter, A. (2013). An infinite dimensional convolution theorem with applications to the efficient estimation of the integrated volatility. *Stochastic Processes and their Applications*, 123(7), 2500–2521.

Fan, J., & Wang, Y. (2008). Spot volatility estimation for high-frequency data. *Statistics and its Interface*, 1(2), 279–288.

Figueroa-López, J. E., & Li, C. (2020). Optimal kernel estimation of spot volatility of stochastic differential equations. *Stochastic Processes and their Applications*, 130(8), 4693–4720.

Figueroa-López, J. E., & Wu, B. (2024). Kernel estimation of spot volatility with microstructure noise using pre-averaging. *Econometric Theory*, 40(3), 558–607.

Jacod, J., & Protter, P. (2011). *Discretization of processes*. Springer Science & Business Media.

Jacod, J., & Rosenbaum, M. (2013). Quarticity and other functionals of volatility: Efficient estimation. *The Annals of Statistics*, 41(3), 1462–1484.

Jacod, J., & Rosenbaum, M. (2015). Estimation of volatility functionals: The case of a $\sqrt{n}$ window. In P. K. Friz, J. Gatheral, A. Gulisashvili, A. Jacquier, & J. Teichmann (Eds.), *Large deviations and asymptotic methods in finance*. Springer International Publishing.

Kristensen, D. (2010). Nonparametric filtering of the realized spot volatility: A kernel-based approach. *Econometric Theory*, 26(1), 60–93.

Li, J., & Liu, Y. (2021). Efficient estimation of integrated volatility functionals under general volatility dynamics. *Econometric Theory*, 37(4), 664–707.

Li, J., Liu, Y., & Xiu, D. (2019). Efficient estimation of integrated volatility functionals via multiscale jackknife. *The Annals of Statistics*, 47(1), 156–176.

Li, J., & Xiu, D. (2016). Generalized method of integrated moments for high-frequency data. *Econometrica*, 84(4), 1613–1633.

Mancini, C., Mattiussi, V., & Renò, R. (2015). Spot volatility estimation using delta sequences. *Finance & Stochastics*, 19(2), 261–293.

Mykland, P. A., & Zhang, L. (2009). Inference for continuous semimartingales observed at high frequency. *Econometrica*, 77(5), 1403–1445.

Renault, E., Sarisoy, C., & Werker, B. (2017). Efficient estimation of integrated volatility and related processes. *Econometric Theory*, 33, 439–478.

https://doi.org/10.1017/S0266466625100170 Published online by Cambridge University Press