The Annals of Statistics
2022, Vol. 50, No. 4, 2072–2095
https://doi.org/10.1214/22-AOS2176

A CLT FOR SECOND DIFFERENCE ESTIMATORS WITH AN APPLICATION TO VOLATILITY AND INTENSITY

BY EMIL A. STOLTENBERG[1,a], PER A. MYKLAND[2,b] AND LAN ZHANG[3,c]

[1]*Department of Data Science and Analytics, BI Norwegian Business School,* [a]*emilstoltenberg@gmail.com*
[2]*Department of Statistics, The University of Chicago,* [b]*mykland@pascal.uchicago.edu*
[3]*Department of Finance, University of Illinois at Chicago,* [c]*lanzhang@uic.edu*

In this paper, we introduce a general method for estimating the quadratic covariation of one or more spot parameter processes associated with continuous time semimartingales, and present a central limit theorem that has this class of estimators as one of its applications. The class of estimators we introduce, that we call Two-Scales Quadratic Covariation (TSQC) estimators, is based on sums of increments of second differences of the observed processes, and the intervals over which the differences are computed are rolling and overlapping. This latter feature lets us take full advantage of the data, and, by sufficiency considerations, ought to outperform estimators that are based on only one partition of the observational window. Moreover, a two-scales approach is employed to deal with asymptotic bias terms in a systematic manner, thus automatically giving consistent estimators without having to work out the form of the bias term on a case-to-case basis. We highlight the versatility of our central limit theorem by applying it to a novel leverage effect estimator that does not belong to the class of TSQC estimators. The principal empirical motivation for the present study is that the discrete times at which a continuous time semimartingale is observed might depend on features of the observable process other than its level, such as its spot-volatility process. As an application of the TSQC estimators, we therefore show how it may be used to estimate the quadratic covariation between the spot-volatility process and the intensity process of the observation times, when both of these are taken to be semimartingales. The finite sample properties of this estimator are studied by way of a simulation experiment, and we also apply this estimator in an empirical analysis of the Apple stock. Our analysis of the Apple stock indicates a rather strong correlation between the spot volatility process of the log-prices process and the times at which this stock is traded and hence observed.

1. Introduction. With an increasing availability of high frequency data, the ambition level as to what can be estimated with reasonable precision has, naturally, also been raised. This paper concerns the estimation of the quadratic covariation of various spot parameter processes associated with continuous time semimartingales, which are observed at discrete times over a finite interval of time. The main result of the paper is a central limit theorem that applies to a class of such estimators. Estimation of the quadratic covariation associated with spot parameter processes is, for example, important for learning about the (hyper-) parameters governing the spot parameter processes, for example, volatility of volatility; or for learning about possible dependencies between concurrently observed semimartingale processes; or for estimating the possible dependency between the observation times and various spot parameter processes associated with the observable process. The motivation for the present paper is

Received December 2020; revised November 2021.
MSC2020 subject classifications. Primary 60F05, 60G44, 62M09, 62M10; secondary 60G42, 62G20, 62P20.
Key words and phrases. Asynchronous times, central limit theorem, consistency, convergence rates, counting processes, endogenous observation times, high frequency, intensity, irregular times, microstructure, observed asymptotic variance, overlapping intervals, rolling intervals, sufficiency, two-scales estimation.

an example of the latter, namely the estimation of the quadratic covariation between the volatility of a continuous semimartingale process, and the intensity processes governing the observation times of this process.

To fix ideas, consider a typical analysis of high frequency data: Based on n discrete time observations $X_{t_1}, \dots, X_{t_n}$ of a continuous semimartingale process X_t one seeks to estimate an integrated parameter Θ,

$$\Theta_T = \int_0^T \theta_s \, \mathrm{d}s,$$

where θ_t is a spot parameter process such as volatility, leverage effect, an instantaneous regression coefficient, or the like. The canonical example is the case where $\theta_t = \sigma_t^2$ is the spot-volatility process associated with an Itô process of the form $\mathrm{d}X_t = \mu_t \, \mathrm{d}t + \sigma_t \, \mathrm{d}W_t$, where W_t is a standard Wiener process, and the problem is to estimate the integrated volatility $\int_0^T \sigma_s^2 \, \mathrm{d}s$ over one or consecutive intervals of time. This example goes back to the research on realised volatility by Andersen et al. (2001), Barndorff-Nielsen and Shephard (2002), Jacod and Protter (1998), Zhang, Mykland and Aït-Sahalia (2005), and others. The econometric interest in investigating nonparametric estimates of this type grew out of the study of volatility clustering by Engle (1982) and Bollerslev (1986). For further references, see Jacod and Protter (2012), Mykland and Zhang (2012) and Aït-Sahalia and Jacod (2014).

The general setup and results of this paper take the following form. Let α_t and β_t be spot parameter processes (potentially the same) associated with one or more semimartingale processes observed at discrete times over a finite interval of time $[0, T]$. In Mykland and Zhang (2017a), an estimator of the quadratic covariation $[\alpha, \beta]_T$ was introduced, and it was shown that this estimator is consistent. In the present paper, we derive the convergence rate for such estimators, and prove a general central limit theorem that, under some regularity conditions, applies to a wide range of estimators based on the second differencing of estimators of integrated spot processes. This central limit theorem is the main theoretical novelty of the paper. As mentioned, our main example of the use of this estimator is the problem of estimating the quadratic covariation between the spot volatility of a semimartingale process, and the intensity process of its observation times. This type of endogenous time problem exists in real applications, but is often overlooked. We also sketch how our estimation methods and the central limit theorem can be applied to a novel estimator of the leverage effect, and how it may be used to gain efficiency in spot volatility estimation.

The paper proceeds as follows. In Section 2, we first describe the model and state our most important assumptions, subsequently we provide a heuristic derivation of the stochastic quantities that are important for the theory that follows. Section 2.2 contains the consistency results and introduces the two-scales estimator of $[\alpha, \beta]_T$. These consistency results generalise the findings in Mykland and Zhang (2017a). In Section 3, we present the main theoretical contribution of the paper, namely a central limit theorem for triangular array rolling quadratic variations based on second differencing of estimators of integrated spot processes. The proof of this theorem is deferred to Appendix F of the Supplementary Material (Stoltenberg, Mykland and Zhang (2022)). Section 3.2 also contains an important corollary to the effect that the observed asymptotic variance (observed AVAR) developed in Mykland and Zhang (2017a) yields consistent estimates of the variance of the limiting distribution that appears in the central limit theorem. In Section 3.3, we provide an example of how the central limit theorem can be applied to a leverage effect estimator. In Section 4, we specialise the theory developed in the preceding sections to the problem of estimating the quadratic covariation between the spot volatility process of a continuous time semimartingale, and the intensity process of its observation times. This is the volatility-intensity problem. In Section 4.2, we investigate the finite sample properties of our estimator by way of a simulation study, while Section 4.3 contains an empirical analysis of the Apple stock observed over 21 trading days in January 2018.

In Section 4.4, we round off by discussing how the volatility-intensity relationship can be used to gain efficiency in the estimation of the spot volatility process. Most technical matters and long proofs can be found in the appendices of the Supplementary Material (Stoltenberg, Mykland and Zhang (2022)). Appendix B of the Supplementary Material also contains a stable central limit theorem for càdlàg martingales, as well as a corollary with some alternative conditions that might be easier to check in applications.

2. The general setup and problem. In this section, we first present the setting for our estimation procedures, define some key quantities, provide a heuristic overview of some important results, and explain what type of estimators our central limit theorem applies to. Subsequently, in Section 2.2, we provide a more formal presentation, and state the main consistency results of the paper.

2.1. *Setup and basic insights.* We suppose that one or more semimartingale processes X_t are observed at high frequency over a finite interval of time $[0, T]$. The semimartingales X_t are typically contaminated by microstructure noise, so what we observe is $Y_{t_i} = X_{t_i} + \varepsilon_{t_i}$, for $t_1, \dots, t_n$ time points, where ε_{t_i} is microstructure noise. Based on these data, we form estimators $\widehat{\Theta}^n$ and $\widehat{\Lambda}^n$, which are consistent for $\Theta_t = \int_0^t \theta_s \,\mathrm{d}s$ and $\Lambda_t = \int_0^t \lambda_s \,\mathrm{d}s$, respectively, where the spot parameter processes θ_t and λ_t are also assumed to be semimartingales. Our results continue to hold when θ_t and λ_t are replaced by sequences $\theta_t^{(n)}$ and $\lambda_t^{(n)}$ of semimartingale processes, but to ease the notation we drop the superscript n for the time being. The spot parameter processes θ_t may be the spot volatility of the continuous part X^c of the process $\mathrm{d}X_t = \sigma_s \,\mathrm{d}W_s + \mathrm{d}t$-terms $+$ jumps, with W_t a standard Wiener process, that is $\theta_t = \sigma_t^2$; it may be the instantaneous leverage effect, $\theta_t = \mathrm{d}[X^c, \sigma^2]_t/\mathrm{d}t$; or the instantaneous volatility of volatility, $\theta_t = \mathrm{d}[\sigma^2, \sigma^2]_t/\mathrm{d}t$; or the stochastic intensity process governing the frequency of the observation times, etc.

As usual, $[X, Y]_t$ denotes the continuous time quadratic covariation of two semimartingales X and Y from time zero to t (Jacod and Shiryaev (2003), page 51). Semimartingales are defined in, for example, Jacod and Shiryaev (2003), Definition I.4.21, page 43. The continuous time predictable quadratic covariation of two locally square integrable martingales M and N from time zero to t is denoted $\langle M, N\rangle_t$ (Jacod and Shiryaev (2003), Theorem I.4.2, page 38).

DEFINITION 1. We assume that all our semimartingales are càdlàg (right continuous with left limits), and that all data generating and latent processes live on the same filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ with $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$, and that this filtered space satisfies the 'usual conditions' (Jacod and Shiryaev (2003), Definitions I.1.2–I.1.3, page 2). When necessary, we will also invoke sequences of filtrations $\mathbb{F}^n = (\mathcal{F}_t^n)_{0\leq t\leq T}$ on $(\Omega, \mathcal{F}, P)$, that is $\mathcal{F}_T^n \subseteq \mathcal{F}$ for all n.

For the proof of the main central limit theorem of the paper, Theorem 3.2, we will need a few additional technical conditions on the structure of the filtered probability space. Also, note that in most cases, parameter processes are indexed by n, and we shall assume that the limiting processes (as $n \to \infty$, when they exist) are continuous, cf. Remark 3 in Section 3.1.

We now turn to the construction of our estimator. Divide the time interval $[0, T]$ into B_n blocks $(t_{i-1}^n, t_i^n]$, of equal length, with $t_0^n = 0$ and $t_{B_n}^n = T$. Set $\Delta_n = T/B_n$, and for convenience, assume that $t_i^n = i\Delta_n$ for $i = 1, \dots, B_n$. Since we shall permit rolling and overlapping intervals, let K_n be an integer no greater than $B_n/2$. From now on, we drop the index n from the t_i^n, B_n, and K_n when it does not cause confusion. For any real functions Θ_t and Λ_t, define

$$\mathrm{QV}_{B,K}(\Theta, \Lambda)_T = \frac{1}{K} \sum_{i=K}^{B-K} (\Theta_{(t_i,t_{i+K}]} - \Theta_{(t_{i-K},t_i]})(\Lambda_{(t_i,t_{i+K}]} - \Lambda_{(t_{i-K},t_i]}), \tag{2.1}$$

where $\Theta_{(s,t]} = \Theta_t - \Theta_s$, and write $\mathrm{QV}_{B,K}(\Theta)_T = \mathrm{QV}_{B,K}(\Theta,\Theta)_T$. For $l = 1, \dots, 2K$, the notation $i \equiv l[2K]$ means that

$$i = 2Kj + l \quad \text{for } K \le i \le B - K,$$

with j an increasing sequence of integers. The basic building block for all the estimators we present is the rolling quadratic covariation

$$\mathrm{QV}_{B,K}(\widehat{\Theta}^n, \widehat{\Lambda}^n)_T = \frac{1}{K} \sum_{i=K}^{B-K} (\widehat{\Theta}^n_{(t_i,t_{i+K}]} - \widehat{\Theta}^n_{(t_{i-K},t_i]})(\widehat{\Lambda}^n_{(t_i,t_{i+K}]} - \widehat{\Lambda}^n_{(t_{i-K},t_i]}),$$

where $\widehat{\Theta}^n$ and $\widehat{\Lambda}^n$ are consistent estimators of the integrated spot processes $\Theta_t = \int_0^t \theta_s \,\mathrm{d}s$ and $\Lambda_t = \int_0^t \lambda_s \,\mathrm{d}s$, respectively. It is important to keep in mind that $\mathrm{QV}_{B,K}(\Theta,\Theta)_T$ and $\mathrm{QV}_{B,K}(\widehat{\Theta}^n, \widehat{\Lambda}^n)_T$ are defined on the discrete grid $\{0, \Delta_n, 2\Delta_n, \dots, T\}$, as opposed to the continuous time quadratic covariation $[X,Y]_t$.

To see how $\mathrm{QV}_{B,K}(\widehat{\Theta}^n, \widehat{\Lambda}^n)_T$ is used to estimate $[\theta,\lambda]_T$, we here present a heuristic analysis, to be made precise in the subsequent section. Under the assumption that $\widehat{\Theta}^n_t$ can be expressed as a sum of $\Theta_t = \int_0^t \theta_s \,\mathrm{d}s$, an error martingale, and terms associated with the edge effects, we can write,

$$\mathrm{QV}_{B,K}(\widehat{\Theta}^n, \widehat{\Lambda}^n)_T = \mathrm{QV}_{B,K}(\Theta,\Lambda)_T + \text{estimation error},$$

where the 'estimation error' might contain terms that are not asymptotically negligible, and that we deal with using so-called two-scale constructions (Mykland, Zhang and Chen (2019), Zhang, Mykland and Aït-Sahalia (2005)). We return to this issue shortly. From Mykland and Zhang ((2017a), Theorem 1, page 203) we have the 'Integral-to-Spot Device', that is,

$$\frac{\mathrm{QV}_{B,K}(\Theta,\Lambda)}{(K\Delta_n)^2} = \frac{2}{3}\left(1 - \frac{1}{K^2}\right)[\theta,\lambda]_T + \frac{1}{K^2}\int_0^T \left[\left(\frac{t^* - t}{\Delta_n}\right)^2 + \left(\frac{t - t_*}{\Delta_n}\right)^2\right] \mathrm{d}[\theta,\lambda]_t + o_p(1),$$

as $K\Delta_n \to 0$, and $t_* = \max\{i\Delta_n : i\Delta_n < t\}$ and $t^* = \min\{i\Delta_n : i\Delta_n \ge t\}$. The key ingredient for proving this theorem is an application of Theorem 2 in Mykland and Zhang ((2017a), page 206), from which we obtain

$$\begin{aligned}
\frac{\mathrm{QV}_{B,K}(\Theta,\Lambda)}{(K\Delta_n)^2} &= \frac{1}{K}\sum_{i=K}^{B-K}\left(\int_{t_i}^{t_{i+K}} \frac{t_{i+K} - s}{K\Delta_n}\,\mathrm{d}\theta_s + \int_{t_{i-K}}^{t_i} \frac{s - t_{i-K}}{K\Delta_n}\,\mathrm{d}\theta_s\right) \\
&\quad \times \left(\int_{t_i}^{t_{i+K}} \frac{t_{i+K} - s}{K\Delta_n}\,\mathrm{d}\lambda_s + \int_{t_{i-K}}^{t_i} \frac{s - t_{i-K}}{K\Delta_n}\,\mathrm{d}\lambda_s\right) \qquad (2.2)\\
&= \frac{1}{K}\sum_{l=1}^{2K}\sum_{i\equiv l[2K]} \int_{t_{i-K}}^{t_{i+K}} f_s^{(l,n)}\,\mathrm{d}\theta_s \int_{t_{i-K}}^{t_{i+K}} f_s^{(l,n)}\,\mathrm{d}\lambda_s,
\end{aligned}$$

where $f_s^{(l,n)}$ for $l = 1, \dots, 2K$ are the functions

$$f_s^{(l,n)} = \sum_{i\equiv l[2K], K\le i\le B-K} \left(\frac{t_{i+K} - s}{K\Delta_n} I\{t_i \le s < t_{i+K}\} + \frac{s - t_{i-K}}{K\Delta_n} I\{t_{i-K} \le s < t_i\}\right). \qquad (2.3)$$

The central limit theorem we present in Section 3 concerns quantities of the type

$$\frac{1}{2K}\sum_{l=1}^{2K}\left\{\sum_{i\equiv l[2K]} \int_{t_{i-K}}^{t_{i+K}} f_{s-}^{(l,n)}\,\mathrm{d}\alpha_s^{(n)} \int_{t_{i-K}}^{t_{i+K}} g_{s-}^{(l,n)}\,\mathrm{d}\beta_s^{(n)} - \int_0^T f_{s-}^{(l,n)} g_{s-}^{(l,n)}\,\mathrm{d}[\alpha^{(n)}, \beta^{(n)}]_s\right\},$$

when $K\Delta_n \to 0$ and $K \to \infty$ as $n \to \infty$. The functions $f_s^{(l,n)}$ and $g_s^{(l,n)}$ are bounded and deterministic, while $\alpha^{(n)}$ and $\beta^{(n)}$ are sequences of semimartingale processes. We see that the right-hand side of (2.2) is a special case of this type of quantity, and so are the nonnegligible terms contained in the 'estimation error' referred to above.

2.2. *Consistency.* Suppose that $\Theta_t^{(n)} = \int_0^t \theta_s^{(n)}\,\mathrm{d}s$ and $\Lambda_t^{(n)} = \int_0^t \lambda_s^{(n)}\,\mathrm{d}s$ are two integrated spot-processes, and that $\theta_t^{(n)}$ and $\lambda_t^{(n)}$ are sequences of semimartingales adapted to $\mathbb{F}^n$ (or $\mathbb{F}$ in the case that $\theta_t^{(n)} = \theta_t$ or $\lambda_t^{(n)} = \lambda$ for all n), and that both sequences are tight with respect to convergence in law relative to the Skorokhod topology on the space $\mathbb{D}[0,T]$ of càdlàg functions on $[0,T]$. If $\theta_t^{(n)}$ and $\lambda_t^{(n)}$ depend on n, we assume that the pair converges in probability to limiting semimartingales θ_t and λ_t, respectively, and that $[\theta^{(n)}, \lambda^{(n)}]$ converges in probability to $[\theta, \lambda]$.

The two spot-processes might be associated with the same underlying semimartingale (in which case we can have $\theta^{(n)} = \lambda^{(n)}$ for all n), or with two different semimartingales concurrently observed. In the latter case, the sampling times can be asynchronous, and the total number of observations may differ. To not overburden the notation, however, we assume that the number of observations are the same for both processes, and equals n. We are given the estimators $\widehat{\Theta}_t^n$ and $\widehat{\Lambda}_t^n$ of $\Theta_t^{(n)}$ and $\Lambda_t^{(n)}$, respectively. Both $\widehat{\Theta}_t^n$ and $\widehat{\Lambda}_t^n$ are consistent and admit representations of the type $\widehat{\Theta}_t^n = \Theta_t + M_{n,t}^\theta + e_{n,t}^\theta - \tilde{e}_{n,0}^\theta$, in terms of a semimartingale $M_{n,t}^\theta$ and edge effects $e_{n,t}^\theta$ and $\tilde{e}_{n,0}^\theta$ associated with phasing in and phasing out the estimator, respectively. For $s < t$, we write $\widehat{\Theta}_{(s,t]}^n = \widehat{\Theta}_t^n - \widehat{\Theta}_s^n$. This means that for $s < t$ the estimators can be represented as

$$
\begin{aligned}
\widehat{\Theta}_{(s,t]}^n - \Theta_{(s,t]} &= M_{n,t}^\theta - M_{n,s}^\theta + e_{n,t}^\theta - e_{n,s}^\theta, \\
\widehat{\Lambda}_{(s,t]}^n - \Lambda_{(s,t]} &= M_{n,t}^\lambda - M_{n,s}^\lambda + e_{n,t}^\lambda - e_{n,s}^\lambda.
\end{aligned}
\tag{2.4}
$$

The assumption, implicit in (2.4), that the edge effect of phasing in an estimator at $s < t$ is the same as the edge effect associated with phasing out an estimator at t. This is exact in the (usual) case of additive estimators (Mykland and Zhang (2017a), Section 5.1, page 215). The results that follow extend with little effort to situations where the edge effects in the two ends of the interval behave differently.

DEFINITION 2 (Stable convergence). We say that a sequence $Z_n = (Z_{n,t})_{0\le t\le T}$ of martingales converges stably in law to $Z = (Z_t)_{0\le t\le T}$ with respect to $\mathcal{G} \subseteq \mathcal{F}$ if (i) Z is measurable with respect to $\widetilde{\mathcal{G}}$ belonging to an extension $(\widetilde{\Omega}, \widetilde{\mathcal{G}}, \widetilde{P})$ of $(\Omega, \mathcal{G}, P)$; and (ii) $\mathrm{E}Yf(Z_n) \to \widetilde{\mathrm{E}}Yf(Z)$ for every bounded and continuous f and all $\mathcal{G}$-measurable bounded random variables Y. We then write $Z_n \Rightarrow Z$ stably.

The notion of an adapted càdlàg sequence being predictably uniformly tight, P-UT in the following, is defined in Jacod and Shiryaev (2003), Definition VI.6.1, page 377.

CONDITION 1. Assume that (2.4) holds, and that there are $\alpha > 0$ and $\beta > 0$ such that, as $n \to \infty$,

$$
n^\alpha M_n^\theta \Rightarrow L^\theta \quad \text{stably} \quad \text{and} \quad n^\beta M_n^\lambda \Rightarrow L^\lambda \quad \text{stably},
$$

with respect to a σ-algebra $\mathcal{G} \subseteq \mathcal{F}$. Both $n^\alpha M_{n,t}^\theta$ and $n^\beta M_{n,t}^\lambda$ are P-UT, and the quadratic variations $[L^\theta, L^\theta]_T$ and $[L^\lambda, L^\lambda]_T$ are measurable with respect to $\mathcal{G}$.

REMARK 1. The requirements of Condition 1 are likely to be satisfied in applications, but they are stronger than what we need for the present purposes. For the consistency results of this section, we only need the weaker Condition 5 of Mykland and Zhang (2017b), page 7. A sequence of semimartingales fulfils this condition if it is tight and P-UT.

THEOREM 2.1 (Consistency of the covariance estimator). *Assume that $\widehat{\Theta}_t^n$ and $\widehat{\Lambda}_t^n$ satisfy (2.4) and Condition 1. Let $K = K_n$ be positive integers, assume that $K\Delta_n \to 0$, and that the edge effects $e_{n,t}^\theta$ and $e_{n,t}^\lambda$ are $o_p((K\Delta_n)^{1/2}n^{-\alpha})$ and $o_p((K\Delta_n)^{1/2}n^{-\beta})$, respectively. Then*

$$\mathrm{QV}_{B,K}(\widehat{\Theta}^n, \widehat{\Lambda}^n)_T = 2[M_n^\theta, M_n^\lambda]_T + \frac{2}{3}(K\Delta_n)^2[\theta^{(n)}, \lambda^{(n)}]_T + o_p\big((K\Delta_n)^2\big) + o_p\big(n^{-(\alpha+\beta)}\big),$$

as $n \to \infty$.

PROOF. The proof follows with trivial adjustments from Mykland and Zhang (2017a), Theorem 3, page 208. A brief sketch of the proof along with some remarks on the edge effects are given in Appendix D of the Supplementary Material. □

In Appendix D of the Supplementary Material, we also provide conclusions to the above theorem with slightly more stringent restrictions on the edge effects. Corresponding results for all combinations of assumptions on the edge effects can be deduced from these results.

We now turn to estimation of the quadratic covariation $[\theta, \lambda]$. As will become clear, how one ought to estimate $[\theta, \lambda]$ depends on the convergence rates of the error martingales M_n^θ and M_n^λ, that is the α and β required for $n^\alpha M_n^\theta$ and $n^\beta M_n^\lambda$ to satisfy Condition 1. From the conclusion of Theorem 2.1 we see that, provided $K\Delta_n$ is of order $n^{-\alpha\wedge\beta}$, then

$$\frac{\mathrm{QV}_{B,K}(\widehat{\Theta}^n, \widehat{\Lambda}^n)_T}{(K\Delta_n)^2} = \frac{2[M_n^\theta, M_n^\lambda]_T}{(K\Delta_n)^2} + \frac{2}{3}[\theta^{(n)}, \lambda^{(n)}]_T + o_p(1).$$

By Condition 1 the quadratic covariation of the error martingales $[M_n^\theta, M_n^\lambda]$ is $O_p(n^{-(\alpha+\beta)})$, consequently,

$$(K\Delta_n)^{-2}[M_n^\theta, M_n^\lambda] = O_p\big((K\Delta_n)^{-1}n^{-\alpha\vee\beta}\big),$$

which tends to zero in probability as $n \to \infty$ provided $\alpha \neq \beta$. We summarise this in a lemma.

LEMMA 2.2. *Assume that Condition 1 holds. Suppose that $\alpha \neq \beta$, that $K \to \infty$ and $\Delta_n \to 0$ such that $K\Delta_n$ is of order $n^{-\alpha\wedge\beta}$ as $n \to \infty$, then*

$$\frac{3}{2}\frac{\mathrm{QV}_{B,K}(\widehat{\Theta}^n, \widehat{\Lambda}^n)_T}{(K\Delta_n)^2} = [\theta, \lambda]_T + o_p(1). \tag{2.5}$$

PROOF. By Condition 1 this is direct from the two displays above. □

Notice that the conclusion of Lemma 2.2 continues to hold when $\alpha = \beta$ provided $(K\Delta_n)^{-1}M_n^\theta$ and $(K\Delta_n)^{-1}M_n^\lambda$ are asymptotically orthogonal (see Jacod and Shiryaev (2003), Proposition I.4.15, page 41, for the notion of local martingales being orthogonal). Also note that when $\alpha = \beta$ one might choose $K\Delta_n$ such that $K\Delta_n n^\alpha \to \infty$, at the cost of a slower rate of convergence.

The estimation problem is harder when $(K\Delta_n)^{-2}[M_n^\theta, M_n^\lambda]_T$ is not asymptotically negligible. This occurs, for example, when one seeks to estimate $[\theta, \theta]_T$, such that the convergence rates α and β of Condition 1 are equal. As an estimator of the quadratic covariation $[\theta, \lambda]_T$ in such situations we propose the *Two-Scales Quadratic Covariation* (TSQC) estimator. It is given by

$$\mathrm{TSQC}_{B,K_1,K_2}(\widehat{\Theta}^n, \widehat{\Lambda}^n)_T = \frac{3}{2}\frac{\mathrm{QV}_{B,K_2}(\widehat{\Theta}^n, \widehat{\Lambda}^n) - \mathrm{QV}_{B,K_1}(\widehat{\Theta}^n, \widehat{\Lambda}^n)}{(K_2^2 - K_1^2)\Delta_n^2}, \tag{2.6}$$

where $K_2 > K_1$ are user specified sequences of integers (tuning parameters) tending to infinity as n tends to infinity. Since K_2 and K_1 must be of the same order (they are both sequences in n), a natural choice is $K_2 = \gamma K_1$ for some integer $\gamma \geq 2$, with γ fixed and independent of n. We first present a consistency result, and then return to the central limit theory for this estimator at the end of Section 3.

COROLLARY 2.3 (Consistency of the TSQC estimator). *Assume that the conditions of Theorem 2.1 are in force, and that $\alpha = \beta$. Let $K_2 = \gamma K_1$, for some fixed integer $\gamma \geq 2$, be positive integers tending to infinity such that $K_1 \Delta_n = O(n^{-\alpha})$. Then,*

$$\mathrm{TSQC}_{B,K_1,K_2}(\widehat{\Theta}^n, \widehat{\Lambda}^n)_T = [\theta, \lambda]_T + o_p(1),$$

as $n \to \infty$.

PROOF. From Theorem 2.1, we have that for $j = 1, 2$

$$\frac{\mathrm{QV}_{B,K_j}(M_n^\theta, M_n^\lambda)_T}{(\gamma^2 - 1)K_1^2\Delta_n^2} = 2\big[n^\alpha M_n^\theta, n^\alpha M_n^\lambda\big]_T + O_p(K_{n,j}\Delta_n),$$

so when $K_2 = \gamma K_1$ for some $\gamma \geq 2$, then

$$\frac{\mathrm{QV}_{B,K_2}(M_n^\theta, M_n^\lambda)_T - \mathrm{QV}_{B,K_1}(M_n^\theta, M_n^\lambda)_T}{(\gamma^2 - 1)K_1^2\Delta_n^2} = o_p(1).$$

On the other hand,

$$\frac{\mathrm{QV}_{B,K_2}(\Theta, \Lambda)_T}{(\gamma^2 - 1)K_1^2\Delta_n^2} = \frac{\gamma^2}{(\gamma^2 - 1)K_2}\sum_{l=1}^{2K_2}\sum_{i \equiv l[2K_2]}\int_{t_{i-K_2}}^{t_{i+K_2}} f_s^{(l,n)}\,\mathrm{d}\theta_s \int_{t_{i-K_2}}^{t_{i+K_2}} f_s^{(l,n)}\,\mathrm{d}\lambda_s,$$

so that by Mykland and Zhang ((2017b), Theorem 7, page 1),

$$\begin{aligned}
&\frac{\mathrm{QV}_{B,K_2}(\Theta, \Lambda)_T - \mathrm{QV}_{B,K_1}(\Theta, \Lambda)_T}{(\gamma^2 - 1)K_1^2\Delta_n^2} \\
&\quad = \frac{\gamma^2}{(\gamma^2 - 1)K_2}\sum_{l=1}^{2K_2}\int_0^T \big(f_s^{(l,n)}\big)^2\,\mathrm{d}[\theta, \lambda]_s \\
&\qquad - \frac{1}{(\gamma^2 - 1)K_1}\sum_{l=1}^{2K_1}\int_0^T \big(f_s^{(l,n)}\big)^2\,\mathrm{d}[\theta, \lambda]_s + O_p\big((K_1\Delta_n)^{1/2}\big),
\end{aligned}$$

where the $f_s^{(l,n)}$ are as in (2.3) with K_2 and K_1 in place of K. Thus,

$$\mathrm{TSQC}_{B,K_1,K_2}(\widehat{\Theta}^n, \widehat{\Lambda}^n)_T = [\theta^{(n)}, \lambda^{(n)}]_T + ((\gamma^2 - 1)K_1^2\Delta_n^2)^{-1} o_p(n^{-2\alpha}) + o_p(1),$$

as $K_1\Delta_n \to 0$ with $K_1 \to \infty$, and the result follows because $K_1\Delta_n$ is of order $n^{-\alpha}$. □

REMARK 2. The conclusion of Corollary 2.3 is still valid when $\alpha \neq \beta$ provided $K_1\Delta_n = O(n^{-\alpha\wedge\beta})$. But if the convergence rates are known and different one would, as already mentioned, rather use the estimator in (2.5). There might be situations, however, where the convergence rates α and β are not known exactly, but known to lie in some interval, say $\alpha, \beta \in [r_1, r_2]$. In that case, one sets $K_1\Delta_n = O(n^{-r_1})$, and the conclusion of Corollary 2.3 holds.

3. Central limit theory. The consistency results of the previous section are extensions of theory developed in Mykland and Zhang (2017a). That paper, however, did not establish limiting normality for the estimators presented, and it is to this topic we now turn.

In a first part, we present a theorem on the convergence rate of triangular array rolling quadratic covariations as approximations to quadratic covariations of spot processes. We then present the central limit theorem for such approximations. Both these results supplement the consistency result of Mykland and Zhang (2017b), Theorem 7, page 1. The proofs of these two theorems are provided in the Supplementary Material. As an example of the use of this theorem, and to show its versatility, we show how it can be applied to a novel estimator of the leverage effect. In Section 3.2, we present theory for the TSQC estimator. In particular, we show that the observed asymptotic variance of Mykland and Zhang (2017a) can be applied to estimate the asymptotic variance of this estimator. This is important because analytical expressions for the limiting variance of the TSQC may be hard to derive. This point is further discussed in Section 3.2 (see also Mykland and Zhang (2017a), pages 198–200).

3.1. *Convergence rate and CLT for rolling quadratic variations.* Introduce the processes

$$\alpha_t^{(l,n)} = \int_0^t f_{s-}^{(l,n)}\, d\alpha_s^{(n)} \quad \text{and} \quad \beta_t^{(l,n)} = \int_0^t g_{s-}^{(l,n)}\, d\beta_s^{(n)} \quad \text{for } l = 1, \dots, 2K, \tag{3.1}$$

where $\alpha_t^{(n)}$ and $\beta_t^{(n)}$ are sequences of semimartingales, and $f_t^{(l,n)}$ and $g_t^{(l,n)}$ are deterministic càdlàg functions bounded by 1 (there is nothing special about 1 here, and it suffices that they are bounded by a constant). We denote $\mathscr{F}$ and $\mathscr{G}$ countable collections $f_\cdot^{(l,n)}, l = 1, \dots, 2K, n = 1, 2, \dots$ and $g_\cdot^{(l,n)}, l = 1, \dots, 2K, n = 1, 2, \dots$, of such functions (see Appendix A in the Supplementary Material). The $f_\cdot^{(l,n)}$ in (2.3) constitute one example of such a class of functions. In Mykland and Zhang ((2017b), Theorem 7, page 1), it was shown that, with $K\Delta_n \to 0$ as $n \to \infty$,

$$\frac{1}{2K}\sum_{l=1}^{2K}\sum_{i\equiv l[2K]}\bigl(\alpha_{t_{i+K}}^{(l,n)} - \alpha_{t_{i-K}}^{(l,n)}\bigr)\bigl(\beta_{t_{i+K}}^{(l,n)} - \beta_{t_{i-K}}^{(l,n)}\bigr) = \frac{1}{2K}\sum_{l=1}^{2K}\bigl[\alpha^{(l,n)}, \beta^{(l,n)}\bigr]_T + o_p(1). \tag{3.2}$$

In this section, we study the rate of convergence and present a central limit theorem for the approximation in (3.2). Such statements will help with the assessment of the accuracy and with optimal calibration of the TSQC estimators, as well as other rolling intervals estimators that depend on approximations such as the one in (3.2).

Conditions A.1 and A.2 of the Supplementary Material, that are invoked in the next theorem, say that $\alpha_t^{(n)}$ and $\beta_t^{(n)}$ are sequences (in n) of Itô semimartingales that are $O_p(1)$ with respect to convergence in law relative to the Skorokhod topology on $\mathbb{D}[0, T]$. Note that, in this section, n is an index that need not be the sample size.

THEOREM 3.1 (Rate of convergence). *Suppose that $\alpha_t^{(n)}$ and $\beta_t^{(n)}$ satisfy Conditions A.1 and A.2 of the Supplementary Material, that $f_\cdot^{(l,n)} \in \mathscr{F}$ and $g_\cdot^{(l,n)} \in \mathscr{G}$, and that $\alpha_t^{(l,n)}$ and $\beta_t^{(l,n)}$ are as defined in (3.1). Then, with $K\Delta_n \to 0$ as $n \to \infty$,*

$$\frac{1}{2K}\sum_{l=1}^{2K}\Biggl\{\sum_{i\equiv l[2K]}\bigl(\alpha_{t_{i+K}}^{(l,n)} - \alpha_{t_{i-K}}^{(l,n)}\bigr)\bigl(\beta_{t_{i+K}}^{(l,n)} - \beta_{t_{i-K}}^{(l,n)}\bigr) - \bigl[\alpha^{(l,n)}, \beta^{(l,n)}\bigr]_T\Biggr\} = O_p\bigl((K\Delta_n)^{1/2}\bigr).$$

PROOF. See Appendix E of the Supplementary Material. □

Let the error term in the approximation in (3.2) be

$$
\begin{aligned}
Z_n(t) = \frac{1}{2K}\sum_{l=1}^{2K}\Big\{ & \sum_{t_{i+K}\le t, i\equiv l[2K]} \big(\alpha_{t_{i+K}}^{(l,n)} - \alpha_{t_{i-K}}^{(l,n)}\big)\big(\beta_{t_{i+K}}^{(l,n)} - \beta_{t_{i-K}}^{(l,n)}\big) \\
& + \big(\alpha_t^{(l,n)} - \alpha_{t_{*,l}}^{(l,n)}\big)\big(\beta_t^{(l,n)} - \beta_{t_{*,l}}^{(l,n)}\big) - \big[\alpha^{(l,n)}, \beta^{(l,n)}\big]_t\Big\},
\end{aligned} \tag{3.3}
$$

where

$$
t_{*,l} = t_{*,l}(t) = \max\{t_{i+K} : t_{i+K} \le t, i \equiv l[2K]\} \quad \text{for } l = 1, \dots, 2K. \tag{3.4}
$$

Notice that $Z_n(t)$ is interpolated into a continuous time semimartingale. The errors are only defined at discrete times, but the interpolation error is asymptotically negligible, and consequently we only need to prove the central limit theorem for the interpolated process, which is done by applying Theorem B.1 of the Supplementary Material.

For the notion of an $\mathcal{F}$-conditional Gaussian martingale, see Jacod and Shiryaev (2003), Definition II.7.4, page 129, or Jacod (1997), page 233. We write ν_α^n for the compensator of the jump process μ_α^n associated with a sequence (in n) of semimartingale process $\alpha^{(n)}$ (see Jacod and Shiryaev (2003), Chapter II.1). We can now state the main result of the paper.

THEOREM 3.2 (CLT for triangular array rolling quadratic variations). *Let $\alpha^{(n)}$ and $\beta^{(n)}$ be sequences of semimartingales, and write $\alpha_t^{(n)} = \alpha_0^{(n)} + \int_0^t \zeta_s^{(n)}\,\mathrm{d}s + \bar{\alpha}_t^{(n)}$ and $\beta_t^{(n)} = \beta_0^{(n)} + \int_0^t \eta_s^{(n)}\,\mathrm{d}s + \bar{\beta}_t^{(n)}$, where $\bar{\alpha}_t^{(n)}$ and $\bar{\beta}_t^{(n)}$ are the martingale parts of the two decompositions. Let Z_n be as defined in (3.3). Suppose that Conditions A.1–A.4 in Appendix A of the Supplementary Material hold; that $\mathrm{d}\langle\bar{\alpha}^{(n)}, \bar{\alpha}^{(n)}\rangle_t/\mathrm{d}t$, $\mathrm{d}\langle\bar{\beta}^{(n)}, \bar{\beta}^{(n)}\rangle_t/\mathrm{d}t$, and $\mathrm{d}\langle\bar{\alpha}^{(n)}, \bar{\beta}^{(n)}\rangle_t/\mathrm{d}t$ are locally continuous in mean square; and that for all $\varepsilon > 0$, the Lindeberg conditions*

$$
\int_{|x|>\varepsilon} x^2 \nu_\alpha^n([0,T]\times \mathrm{d}x) \xrightarrow{p} 0 \quad \text{and} \quad \int_{|x|>\varepsilon} x^2 \nu_\beta^n([0,T]\times \mathrm{d}x) \xrightarrow{p} 0, \tag{3.5}
$$

as $n \to \infty$ are satisfied for both processes. Set

$$
\begin{aligned}
\kappa_s^{(n)} = \frac{1}{4K^3\Delta_n}\sum_{l_1=1}^{2K}\sum_{l_2=1}^{2K}\int_{t_{*,l_1}\vee t_{*,l_2}}^{s} \Big\{ & f_{u-}^{(l_1,n)} f_{u-}^{(l_2,n)} g_{s-}^{(l_1,n)} g_{s-}^{(l_2,n)}\,\mathrm{d}\big[\bar{\alpha}^{(n)}, \bar{\alpha}^{(n)}\big]_u \frac{\mathrm{d}\langle\bar{\beta}^{(n)}, \bar{\beta}^{(n)}\rangle_s}{\mathrm{d}s}[2] \\
& + f_{u-}^{(l_1,n)} g_{u-}^{(l_2,n)} g_{s-}^{(l_1,n)} f_{s-}^{(l_2,n)}\,\mathrm{d}\big[\bar{\alpha}^{(n)}, \bar{\beta}^{(n)}\big]_u \frac{\mathrm{d}\langle\bar{\beta}^{(n)}, \bar{\alpha}^{(n)}\rangle_s}{\mathrm{d}s}[2]\Big\},
\end{aligned}
$$

where the $f^{(l,n)}$ and $g^{(l,n)}$ belong to $\mathscr{F}$ and $\mathscr{G}$, respectively, as described in Definition A.2 in Appendix A of the Supplementary Material. Assume that there is an $\mathcal{F}$-measurable process κ_s such that when $K\Delta_n \to 0$ and $K \to \infty$ as $n \to \infty$,

$$
\int_0^t \kappa_s^{(n)}\,\mathrm{d}s \xrightarrow{p} \int_0^t \kappa_s\,\mathrm{d}s \quad \textit{for each } t \in [0,T].
$$

Then $(K\Delta_n)^{-1/2} Z_n$ converges stably in law to an $\mathcal{F}$-conditional Gaussian martingale $\mathscr{Z}$ with quadratic variation $\langle\mathscr{Z}, \mathscr{Z}\rangle_t = \int_0^t \kappa_s\,\mathrm{d}s$, when $K\Delta_n \to 0$ and $K \to \infty$ as $n \to \infty$.

PROOF. The full proof is given in Appendix F of the Supplementary Material. □

The notation "[2]" means that we sum over two terms, the one given and the corresponding one where, in $\kappa^{(n)}$, f and g and $\alpha^{(n)}$ and $\beta^{(n)}$ have changed place. As an example, $a_1 b_2 + a_2 b_1 = a_1 b_2[2]$. A full expression for $\kappa_s^{(n)}$ is given on page 23 of the Supplementary Material. The notion of a process being locally continuous in mean square is defined in Definition A.3 of the Supplementary Material.

REMARK 3. It is important to note that condition (3.5) ensures that the limiting processes α, β (if they exist) and $\mathscr{Z}$ are continuous. The pre-limiting processes need not be continuous, and may have jumps. Condition (3.5) thus also rules out situations where the limiting processes α and β have co-jumps (in the limit) with their volatility. It is known that the realised volatility estimator does not have a limiting conditionally Gaussian distribution if the price process co-jumps with its volatility (see Jacod and Protter (2012), Theorem 5.4.2, page 162), but rather a doubly mixed normal distribution (for further references see the end of Section 3.3). How Theorem 3.2 generalises if condition (3.5) is relaxed and one or both processes α and β are allowed to co-jump with their volatility, is a topic for further research.

REMARK 4. Theorem 3.2 is formulated in terms of general functions $f^{(l,n)}$ and $g^{(l,n)}$ belonging to classes $\mathscr{F}$ and $\mathscr{G}$ as defined in Definition A.2 of the Supplementary Material. Basically, these functions need to be deterministic and bounded (for convenience we require them to be bounded by 1). In our main application of the theorem, however, the functions $f^{(l,n)}$ and $g^{(l,n)}$ are as defined in (2.3). An example where the $f^{(l,n)}$ functions and the $g^{(l,n)}$ functions differ is given in Section 3.3.

The above theorem is stated in terms of a univariate process Z_n. A multivariate version of Theorem 3.2 only requires notational modifications to the proof of said theorem. Let $\alpha^{(j,n)}$, $\beta^{(j,n)}$ be sequences (in n) of semimartingales for $j = 1, \dots, p$. Analogously to (3.1), set $\alpha_t^{(l,j,n)} = \int_0^t f_{s-}^{(l,j,n)}\,\mathrm{d}\alpha_s^{(j,n)}$ and $\beta_t^{(l,j,n)} = \int_0^t g_{s-}^{(l,j,n)}\,\mathrm{d}\beta_s^{(j,n)}$, where $f_\cdot^{(l,j,n)} \in \mathscr{F}$ and $g_\cdot^{(l,j,n)} \in \mathscr{G}$, with $\mathscr{F}$ and $\mathscr{G}$ as defined in Definition A.2 of the Supplementary Material. Let $t_{*,l}$ be as in (3.4), and form the processes $Z_{n,l}^{(j)}(t) = (2K)^{-1}\sum_{l=1}^{2K}\sum_{t_{i+K}\le t; i\equiv l[2K]}(\alpha_{t_{i+K}}^{(l,j,n)} - \alpha_{t_{i-K}}^{(l,j,n)})(\beta_{t_{i+K}}^{(l,j,n)} - \beta_{t_{i-K}}^{(l,j,n)}) + (\alpha_t^{(l,j,n)} - \alpha_{t_{*,l}}^{(l,j,n)})(\beta_t^{(l,j,n)} - \beta_{t_{*,l}}^{(l,j,n)}) - [\alpha^{(l,j,n)}, \beta^{(l,j,n)}]_t$ for $j = 1, \dots, p$, and set

$$\mathscr{Z}_n(t) = (K\Delta_n)^{-1/2}\sum_{l=1}^{2K}\begin{pmatrix} Z_{n,l}^{(1)}(t) \\ \vdots \\ Z_{n,l}^{(p)}(t)\end{pmatrix}. \tag{3.6}$$

The quadratic variation of $\mathscr{Z}_n(t)$ is $\langle \mathscr{Z}_n, \mathscr{Z}_n\rangle_t = \int_0^t k_s^{(n)}\,\mathrm{d}s$ say, where $k_s^{(n)}$ is a $p \times p$ matrix with elements $(K\Delta_n)^{-1}\sum_{l_1=1}^{2K}\sum_{l_2=1}^{2K}\langle Z_{n,l_1}^{(i)}, Z_{n,l_2}^{(j)}\rangle_s$ for $i, j = 1, \dots, p$. It follows directly from the proof of Theorem 3.2 that each of these elements of the matrix $k_s^{(n)}$ are of the same form as the $\kappa_s^{(n)}$ defined in Theorem 3.2.

COROLLARY 3.3. *Let $\mathscr{Z}_n$ be as defined in (3.6), and let $k_t^{(n)}$ be defined via $\langle \mathscr{Z}_n, \mathscr{Z}_n\rangle_t = \int_0^t k_s^{(n)}\,\mathrm{d}s$. Suppose that $\alpha^{(j,n)}$ and $\beta^{(j,n)}$, as well as $f_\cdot^{(l,j,n)}$, and $g_\cdot^{(l,j,n)}$ for $j = 1, \dots, p$ satisfy the conditions of Theorem 3.2, and suppose that there is a $p \times p$ matrix process k_t such that $\int_0^t k_s^{(n)}\,\mathrm{d}s \to_p \int_0^t k_s\,\mathrm{d}s$ for all $t \in [0, T]$. Then $\mathscr{Z}_n$ converges stably in law to a p-dimensional $\mathcal{F}$-conditional Gaussian martingale with quadratic variation $\int_0^t k_s\,\mathrm{d}s$.*

PROOF. This is direct from the proof of Theorem 3.2. □

A consequence of this corollary is that the sequence $(K\Delta_n)^{-1/2}U_{n,K}(t)$, where

$$U_{n,K}(t) = \frac{\mathrm{QV}_{B,K}(\widehat{\Theta}^n, \widehat{\Lambda}^n)_t}{(K\Delta_n)^2} - \frac{2[M_n^\theta, M_n^\lambda]_t}{(K\Delta_n)^2} - \frac{2}{3}[\theta, \lambda]_t, \tag{3.7}$$

converges stably in law to a mean zero Gaussian martingale, provided $\theta^{(n)}$, $\lambda^{(n)}$, $n^\alpha M_n^\theta$, and $n^\beta M_n^\lambda$ satisfy the conditions imposed in Corollary 3.3. To see this, for $l = 1, \dots, 2K$,

$K \leq B/2$, with $t_{*,l}$ as in (3.4), and general semimartingales θ and λ, define the interpolated processes

$$\mathbb{Z}_K^{(l)}(f,\theta,g,\lambda)_t = \sum_{i\equiv l[2K]} \int_{t_{i-K}}^{t_{i+K}} f_{s-}^{(l,n)}\,\mathrm{d}\theta_s \int_{t_{i-K}}^{t_{i+K}} g_{s-}^{(l,n)}\,\mathrm{d}\lambda_s + \int_{t_{*,l}}^{t} f_{s-}^{(l,n)}\,\mathrm{d}\theta_s \int_{t_{*,l}}^{t} g_{s-}^{(l,n)}\,\mathrm{d}\lambda_s - \int_0^t f_{s-}^{(l,n)} g_{s-}^{(l,n)}\,\mathrm{d}[\theta,\lambda]_s,$$

and families of functions $f = (f_{\cdot}^{(l,n)})_{1\leq l\leq 2K, n\geq 1}$ and $g = (g_{\cdot}^{(l,n)})_{1\leq l\leq 2K, n\geq 1}$ belonging to the classes $\mathscr{F}$ and $\mathscr{G}$, respectively. Set

$$\mathbb{Z}_K(f,\theta,g,\lambda)_t = \frac{1}{2K}\sum_{l=1}^{2K} \mathbb{Z}_K^{(l)}(f,\theta,g,\lambda)_t,$$

and let the functions $f_s^{(l,n)}$ be as defined in (2.3), and introduce

$$g_s^{(l,n)} = \sum_{K\leq i\leq B-K, i\equiv l[2K]} \big(I\{t_i \leq s < t_{i+K}\} - I\{t_{i-K}\leq s < t_i\}\big), \tag{3.8}$$

for $l = 1,\dots,2K$. Then (3.7) can be written $U_{n,K}(t) = \bar{U}_{n,K}(t) + o_p((K\Delta_n)^{1/2})$, where

$$\begin{aligned}\bar{U}_{n,K}(t) = {}& \mathbb{Z}_K(f,\theta,f,\lambda)_t + \frac{n^{-\beta}}{K\Delta_n}\mathbb{Z}_K\big(f,\theta,g,n^{\beta}M_n^{\lambda}\big)_t \\ &+ \frac{n^{-\alpha}}{K\Delta_n}\mathbb{Z}_K\big(g,n^{\alpha}M_n^{\theta},f,\lambda\big)_t + \frac{n^{-(\alpha+\beta)}}{(K\Delta_n)^2}\mathbb{Z}_K\big(g,n^{\alpha}M_n^{\theta},g,n^{\beta}M_n^{\lambda}\big)_t.\end{aligned} \tag{3.9}$$

From this representation, we see that when $K\Delta_n = O(n^{-\alpha\wedge\beta})$, Corollary 3.3 combined with the Cramér–Wold device (Billingsley (1995), page 382), and the fact that sums of C-tight sequences are C-tight (Jacod and Shiryaev (2003), Corollary VI.3.33, page 353) entail that the sequence $(K\Delta_n)^{-1/2}U_{n,K}$ defined via (3.7), converges stably in law. From the expression in (3.9), we also see that when $\alpha = \beta$ and $K\Delta_n$ is of order $n^{-\alpha}$, then all four terms in the sum will contribute to the asymptotic variance, while only two of the terms contribute to the asymptotic variance when the convergence rates differ.

3.2. *Uncertainty of the* TSQC. To compute the uncertainty associated with TSQC estimators, we use the *observed asymptotic variance* of Mykland and Zhang (2017a). It should be noted that using the observed asymptotic variance, instead of some other approach, is just a question of *how* we estimate the asymptotic variance. The observed asymptotic variance is an extremely useful approach to variance estimation as it allows us to circumvent the derivation of an explicit expression for the asymptotic variance of the TSQC estimators. Using the observed asymptotic variance is akin to, in likelihood estimation, using the observed information or, for many types of data, bootstrapping the variance.

For the observed asymptotic variance to be consistent, it is sufficient that the sequence of error martingales associated with the estimator whose uncertainty one wants to compute is tight and P-UT (see Condition 5 in Mykland and Zhang (2017b), page 7). When constructing confidence intervals and conducting tests, however, we cannot do with merely consistency, but need a central limit theorem. The sequence for which we are going to use the observed asymptotic variance to compute its uncertainty, is

$$(K_1\Delta_n)^{-1/2}\big(\mathrm{TSQC}_{B,K_1,K_2}(\widehat{\Theta}^n,\widehat{\Lambda}^n)_t - [\theta,\lambda]_t\big). \tag{3.10}$$

Under an assumption on the convergence in probability of the quadratic covariations of QV_{B,K_1} and QV_{B,K_2} (properly normalised) for $K_2 = \gamma K_1$, stable asymptotic normality ensues from our previous results.

COROLLARY 3.4 (Uncertainty of the TSQC). *Assume that $\theta^{(n)}$ and $\lambda^{(n)}$, as well as the error martingales $n^\alpha M_n^\theta$ and $n^\beta M_n^\lambda$ satisfy the conditions imposed on $\alpha^{(n)}$ and $\beta^{(n)}$ in Theorem 3.2; that the conditions of Lemma 2.3 are in force, and that $\langle U_{n,K_1}, U_{n,K_2}\rangle_t$ converges in probability to a continuous process for each $t \in [0,T]$. Then the sequence in (3.10) is C-tight and P-UT, and converges stably in law to an $\mathcal{F}$-conditional Gaussian martingale.*

PROOF. Since $K_1\Delta_n$ is of the same order as $n^{-\alpha\wedge\beta}$ (so that the factors $n^{-\alpha}/(K_1\Delta_n)$, $n^{-\beta}/(K_1\Delta_n)$, and $n^{-(\alpha\wedge\beta)}/(K_1\Delta_n)^2$ either tend to zero or to one) it follows from Theorem 3.2 that all the error martingales on the right-hand side of (3.9) are C-tight and P-UT. Since sums of C-tight sequences are C-tight (Jacod and Shiryaev (2003), Corollary VI.3.33, page 353), and sums of sequences that are P-UT are P-UT (Jacod and Shiryaev (2003), VI.6.4, page 377), U_{n,K_j}, $j=1,2$ are C-tight and P-UT, and so is the sequence in (3.10). Combining this with the assumption about the quadratic covariation $\langle U_{n,K_1}, U_{n,K_2}\rangle_t$, Theorem B.1 of the Supplementary Material gives the result. □

Before we proceed to Section 4 and the problem that motivated the present study, we showcase the applicability of Theorem 3.2 by considering a novel estimator of the leverage effect, that does not belong to the class of estimators introduced in (2.1).

3.3. *Leverage effect estimation.* In this section, we introduce a rolling intervals estimator of the leverage effect, and sketch how Theorem 3.2 can be used to derive the limiting distribution of this estimator. Since the goal of this section is not leverage effect estimation *per se*, but rather to show an application of Theorem 3.2, we limit ourselves to a continuous process model, without microstructure noise.

Assume that the process $X_t = X_0 + \int_0^t \sigma_s\,\mathrm{d}W_s$ is observed at the discrete and equidistant times $0 = t_{0,n} < t_{1,n} < \cdots < t_{n-1,n} < t_{n,n} = T$, meaning that there is no microstructure noise; W_t is a one dimensional Wiener process, and σ^2 is a locally bounded Itô process which may or may not be correlated with W_t. Write $\theta = \sigma^2$. The leverage effect is the spot process $\mathrm{d}[\theta, X]_t/\mathrm{d}t$. A natural estimator of $[X,X]_T$ is the realised volatility $\widehat{\Theta}_T^n = \sum_{t_{i+1,n}\le T}(X_{t_{i+1,n}} - X_{t_{i,n}})^2$ (see the references in the Introduction). Define $M_{n,t}$ via $\widehat{\Theta} = [X,X]_t + M_{n,t}$. It can then be shown that $n^{1/2}M_n$ converges stably in law to a normal distribution with (random) variance $2T\int_0^T \sigma_s^4\,\mathrm{d}s$ (Mykland and Zhang (2012), Corollary 2.30, page 154), hence $n^{1/2}M_n$ satisfies Condition 1. In analogy with (2.1), consider

$$\widehat{[\theta, X]}_T^{n,K} = \frac{1}{K}\sum_{i=K}^{n-K}\big(\widehat{\Theta}^n_{(t_{i,n},t_{i+K,n}]} - \widehat{\Theta}^n_{(t_{i-K,n},t_{i,n}]}\big)(X_{t_{i+K,n}} - X_{t_{i-K,n}}), \tag{3.11}$$

where, due to the equidistant sampling times, we take $B = n$. Notice that this estimator is different from the QV of (2.1), as it is only the first differences of the X process that enter in (3.11). Similarly to Lemma 2.2, we have that

$$(K\Delta_n)^{-1}\widehat{[\theta, X]}_T^{n,K} = [\theta, X]_T + o_p(1).$$

Let $f_s^{(l,n)}$ be as defined in (2.3), and let $g_s^{(l,n)}$ be as defined in (3.8), for $l = 1,\dots,2K$. Let $M_t^{(l,n)} = \int_0^t g_s^{(l,n)}\,\mathrm{d}M_{n,s}$, and define the two continuous time martingales

$$\begin{aligned} U_{n,l}(t) = \sum_{t_{i+K}\le t:\, i\equiv l[2K]} &\big(\theta_{t_{i+K}}^{(l,n)} - \theta_{t_{i-K}}^{(l,n)}\big)(X_{t_{i+K}} - X_{t_{i-K}}) \\ &+ \big(\theta_t^{(l,n)} - \theta_{t_{*,l}}^{(l,n)}\big)(X_t - X_{t_{*,l}}) - \big[\theta^{(l,n)}, X\big]_t, \end{aligned}$$

and

$$V_{n,l}(t) = \sum_{t_{i+K} \le t : i \equiv l[2K]} \big(M^{(l,n)}_{t_{i+K}} - M^{(l,n)}_{t_{i-K}}\big)\big(X_{t_{i+K}} - X_{t_{i-K}}\big) \\ + \big(M^{(l,n)}_{t} - M^{(l,n)}_{t_{*,l}}\big)\big(X_{t} - X_{t_{*,l}}\big) - \big[M^{(l,n)}, X\big]_t,$$

where $t_{*,l}$ is as in (3.4). Then $(K\Delta_n)^{-1}\widehat{[\theta, X]}^{n,K}_T - K^{-1}\sum_{l=1}^{2K}[\theta^{(l,n)}, X]_T$ is asymptotically equivalent to $Z_n = K^{-1}\sum_{l=1}^{2K}\{U_{n,l}(T) + V_{n,l}(T)\}$. Its predictable quadratic variation is

$$\langle Z_n, Z_n\rangle_T = \frac{1}{K^2}\sum_{l_1=1}^{2K}\sum_{l_2=1}^{2K}\big\{\langle U_{n,l_1}, U_{n,l_2}\rangle_T + \langle V_{n,l_1}, V_{n,l_2}\rangle_T + 2\langle U_{n,l_1}, V_{n,l_2}\rangle_T\big\}.$$

Provided that the processes involved satisfy the assumptions of Theorem 3.2, we see how the development so far leads to a central limit theorem for the leverage effect estimator of (3.11). In particular, $(K\Delta_n)^{-1/2}Z_n$ converges stably in law to a Gaussian martingale with (random) asymptotic variance of the form $\int_0^T (a_s + b_s + 2c_s)\,\mathrm{d}s$.

Estimators of the leverage effect have been studied previously by Wang and Mykland (2014), Kalnina and Xiu (2017), Aït-Sahalia et al. (2017), to mention some. The latter paper considers situations where the price and the volatility may jump together, so-called co-jumps (for more on co-jumps, see Li, Todorov and Tauchen (2017a), Li, Todorov and Tauchen (2017b) and Jacod and Todorov (2010)). In such cases, the leverage effect consists of two parts: (i) the correlation between the continuous parts of the price and the volatility processes, and (ii) the inclination of the two processes of jumping at the same time. How the leverage effect estimator of (3.11) generalises to more complicated data structures, for example involving co-jumps, microstructure noise, and nonequidistant sampling times, are topics we are currently exploring.

4. Volatility and intensity. In this section, we turn to the application that motivated the current paper, namely the estimation of the quadratic covariation between the volatility process of a continuous time semimartingale, and the intensity process of the observation times.

When estimating parameters associated with a continuous time process that is only observed at discrete times, simplifying assumptions are often imposed on the relation between the observation times and the underlying process. The observation times are typically either taken as fixed and equidistant, or they are governed by a stochastic process postulated to be independent of the observable process (see, e.g., Aït-Sahalia and Jacod (2014), Chapter 9, for a discussion). We refer to both cases as ‘exogenous times’. In many settings the assumption of exogenous times is violated, the case of high frequency financial data being, at least in some cases, a pertinent example. Decisions to buy or sell a given security may, in part, be determined by features of that security, and since it is only at the times at which transactions are conducted that we get a glimpse of the continuous processes ticking in the background (modulo microstructure noise), one would expect that the observation times may be correlated with transaction-igniting features of the underlying process.

In recent years, much progress has been made when the assumption of exogenous times is relaxed. In Li, Zhang and Zheng (2013) and Li et al. (2014), the realised volatility estimator is studied in the presence of endogenous observation times, and it is shown that a ‘bias’ term appears in the limiting distribution of this estimator. This ‘bias’ term is of the same order as the process tending (stably) to a normal limit, and is thus not a bias term in the traditional sense. The reasons for caring about it have to do with efficiency considerations, and not with the estimation being off-the-target in an expected value sense. Jacod, Li and Zheng

(2019) construct an estimator of the integrated volatility in the presence of microstructure noise, jumps, and endogenous times. Other papers have dealt with consistency and central limit theorems under irregular and random times (see Renault and Werker (2011), Hayashi, Jacod and Yoshida (2011), Fukasawa and Rosenbaum (2012), Potiron and Mykland (2017)). Common for all the above papers is that the endogeneity of the observation times comes about because the times depend on the efficient price process itself, as opposed to latent spot parameter processes governing the evolution of the efficient price process. The tools developed in Sections 2 and 3 allow us to statistically study situations where the observation times might depend on underlying nonobservable features of the efficient price process, such as its spot-volatility process, the associated volatility of volatility, the leverage effect, and so on. To assess the direction and magnitude of such correlations, we can use the TSQC estimator of (2.6), and also a correlation estimator based on the TSQC. In this section, we first present some theory specific to the volatility-intensity covariance estimation, then, in Section 4.2 we perform a simulation study to assess the finite sample behaviour of our estimators, while Section 4.3 contains an empirical study of the Apple stock over 21 trading days in January 2018.

4.1. *A model for volatility-intensity covariance estimation.* For a given frequency of observations, indexed by $n \geq 1$, the succesive observations occur at times $0 = T_{n,0} < T_{n,1} < \cdots \leq T$, where $(T_{n,i})_{n\geq 1}$ is a sequence of finite stopping times. Define the sequence of counting processes $N_{n,t} = \sum_{i\geq 1} I\{T_{n,i} \leq t\}$. We are going to assume (in Condition 2) that, for observation frequency n, the inter-observational lags $T_{n,i} - T_{n,i-1}$ are of the same order as $1/n$, and moreover, that $n^{-1}N_{n,t}$ has a possibly random probability limit when n goes to infinity (see Li et al. (2014), Jacod, Li and Zheng (2017), and Jacod, Li and Zheng (2019) for similar constructions). Based on the $N_{n,T}$ observations of X_t, we form an estimator $\widehat{\Theta}_t^n$ of $\Theta_t = \int_0^t \theta_s \, \mathrm{d}s$, where the spot parameter process θ_s is itself assumed to be a semimartingale, and assume that $\widehat{\Theta}_t^n$ is consistent for Θ_t. In the following, we think of θ_t as the spot-volatility process σ^2, and Θ_t as the integrated volatility $\int_0^t \sigma_s^2 \, \mathrm{d}s$. The counting process $N_{n,t}$ can be decomposed as $N_{n,t} = M_{n,t} + \Lambda_{n,t}$, in terms of a martingale $M_{n,t}$ and an increasing and predictable process $\Lambda_{n,t}$. We assume that the latter process is absolutely continuous, so that $\Lambda_{n,t} = \int_0^t \lambda_{n,s} \, \mathrm{d}s$, and that $\lambda_{n,t}$, called the intensity process, is itself a semimartingale. The process we seek to estimate is then $[\theta, \lambda]_t$ over one or consecutive observation windows.

Since X is followed over the finite interval $[0, T]$, where T is fixed, our arguments are based on asymptotics as the observation frequency gets higher, that is $\max_{i\geq 1}|T_{n,i} - T_{n,i-1}| \to 0$, so-called infill asymptotics. To let the number of observations $N_{n,T}$ tend to infinity, and at the same time get a finite limit for the intensities of the observation times, we impose the following condition.

CONDITION 2. There is a nonnegative semimartingale λ_t such that $n^{-1}\Lambda_{n,t} \xrightarrow{p} \Lambda_t := \int_0^t \lambda_s \, \mathrm{d}s$, for all $t \in [0, T]$.

One may think of $1/n$ as proportional to the expected distance between two observation times, or n as being proportional to the expected number of observations per period. The point is that Condition 2 allows us to develop asymptotic theory in terms of $N_{n,T}$ for the estimators we construct. This construction is similar to that previously employed by Li, Zhang and Zheng (2013); and by Jacod, Li and Zheng (2019), Assumption (O-ρ, ρ'), page 82.

Suppose that the estimator $\widehat{\Theta}_t^n$ satisfies the decomposition in (2.4), and that its error process martingale $M_{n,t}^{\theta}$ obeys Condition 1. We return to the assumptions on the edge effects in due time. Define $\widetilde{\Lambda}_t^n = n^{-1}N_{n,t}$. The counting process $N_{n,t}$ simply counts the transactions and is hence observable, whereas n is a nonobservable abstraction introduced so that the asymptotic theory developed in the two preceding sections generalises to volatility-intensity estimation. This means that $\widetilde{\Lambda}_t^n$ is a rescaling of an estimator. For the (finite sample) empirical applications of our estimator, the index n will turn out to be immaterial.

REMARK 5. We emphasize that n does not need to be observed for the developments in this section to be valid. We need n to exist in the sense of Condition 2, but otherwise n is a notational convenience that permits us to state results more simply, and n is in this sense always only a scaling. For example, $\widetilde{\Lambda}_t^n \xrightarrow{p} \Lambda_t$ can be restated as $\widehat{\Lambda}_{n,t} = \Lambda_{n,t}(1 + o_p(1))$, where $\widehat{\Lambda}_{n,t} = N_{n,t}$.

Notice that there are no edge effects associated with $\widetilde{\Lambda}_t^n$, so (2.4) becomes $\widetilde{\Lambda}_t^n = n^{-1}\Lambda_{t,n} + M_{n,t}^\lambda$, where $M_{n,t}^\lambda = n^{-1}(N_{n,t} - \Lambda_{t,n})$ is a martingale sequence. Moreover, as $n \to \infty$,

$$n[M_n^\lambda, M_n^\lambda]_t = n^{-1} N_{n,t} \xrightarrow{p} \Lambda_t, \tag{4.1}$$

by Condition 2. The convergence in (4.1) combined with the fact that Λ_t is increasing and continuous, yield

$$n^{1/2} M_{n,t}^\lambda \Rightarrow \int_0^t \lambda_s^{1/2}\, \mathrm{d}W_s' \quad \text{stably,}$$

where W_s' is a Wiener process defined on an extension of the original probability space (see Theorem B.1 in the Supplementary Material). Set $L_t^\lambda = \int_0^t \lambda_s^{1/2}\, \mathrm{d}W_s'$, and we have the first part of Condition 1. For Theorem 2.1 to be applicable, the sequence of martingales $n^{1/2} M_{n,t}^\lambda$ must also be P-UT.

LEMMA 4.1. *Assume Condition 2. Then $n^{1/2} M_{n,t}^\lambda$ is P-UT.*

PROOF. That $n\langle M_n^\lambda, M_n^\lambda\rangle_t = n^{-1}\int_0^t \lambda_{n,s}\, \mathrm{d}s$ ensures that $n\langle M_n^\lambda, M_n^\lambda\rangle_t$ is tight (Jacod and Shiryaev (2003), Proposition VI.3.26, page 351). Being a counting process martingale, the jumps $n^{1/2}|\Delta M_{n,t}^\lambda| \leq 1$, and Jacod and Shiryaev (2003), Proposition VI.6.13, page 379, gives the result. □

In the absence of edge effects on the part of $\widetilde{\Lambda}_t^n$, $\mathrm{QV}(\widehat{\Theta}^n, \widetilde{\Lambda}^n)$ can be decomposed as (cf. the decomposition in equation (D.1) of Appendix D in the Supplementary Material),

$$\mathrm{QV}(\widehat{\Theta}^n, \widetilde{\Lambda}^n) = \overline{\mathrm{QV}}(\widehat{\Theta}^n, \widetilde{\Lambda}^n) + O_p\big(\mathrm{QV}(\widetilde{\Lambda}^n)^{1/2} R_{n,k}(\Theta)^{1/2}\big), \tag{4.2}$$

by the Cauchy–Schwarz inequality, where

$$\overline{\mathrm{QV}}(\widehat{\Theta}^n, \widetilde{\Lambda}^n) = \mathrm{QV}(\Theta, \Lambda_n/n) + \mathrm{QV}\big(M^\theta, \Lambda_n/n\big) + \mathrm{QV}\big(\Theta, M_n^\lambda\big) + \mathrm{QV}\big(M^\theta, M_n^\lambda\big),$$

and $R_{n,K}(\Theta) = K^{-1}\sum_{i=K}^{B-K}(e_{n,t_{i+K}}^\theta - e_{n,t_i}^\theta - (e_{n,t_i}^\theta - e_{n,t_{i-K}}^\theta))^2$.

COROLLARY 4.2. *Suppose that $\widehat{\Theta}_t^n$ satisfies Condition 1 in Section 2.2, that $(\lambda_{n,t}/n)_{n\geq 1}$ is tight and P-UT, and that $e_{n,t}^\theta$ are $o_p((K\Delta_n)^{1/2} n^{-\alpha})$. Then, as $K\Delta_n \to 0$*

$$\overline{\mathrm{QV}}(\widehat{\Theta}^n, \widetilde{\Lambda}^n)_T = 2[M_n^\theta, M_n^\lambda]_T + \frac{2}{3}(K\Delta_n)^2[\theta, \lambda_n/n]_T + o_p\big((K\Delta_n)^2\big) + o_p\big(n^{-\alpha} n^{-1/2}\big).$$

PROOF. By Lemma 4.1, the sequence $\widetilde{\Lambda}_t^n = n^{-1}\Lambda_{n,t} + M_{n,t}^\lambda$ satisfies Condition 1, and the conditions on $(\lambda_{n,t}/n)_{n\geq 1}$ ensure that Theorem 7 in Mykland and Zhang (2017b) is applicable. The second part of Theorem 2.1 then gives the result. □

We have that $\mathrm{QV}(\widetilde{\Lambda}^n) = O_p(K\Delta_n + n^{-1/2})$, which via (4.2) shows how differing restrictions on the edge effects associated with the integrated volatility estimator give differing conclusions about $\mathrm{QV}(\widehat{\Theta}^n, \widetilde{\Lambda}^n)$ (see the discussion in Appendix D of the Supplementary Material). If we assume that the edge effects associated with $\widehat{\Theta}^n_t$ are $o_p((K\Delta_n)^{3/4}n^{-\alpha})$, which is not unrealistic when working with two-scales estimators and pre-averaged observations (see Zhang, Mykland and Aït-Sahalia (2005) and Mykland, Zhang and Chen (2019)), then the conclusion of Corollary 4.2 is

$$\begin{aligned}\mathrm{QV}\big(\widehat{\Theta}^n, \widetilde{\Lambda}^n\big)_T &= 2\big[M_n^\theta, M_n^\lambda\big]_T + \frac{2}{3}(K\Delta_n)^2[\theta, \lambda_n/n]_T \\ &\quad + O_p\big((K\Delta_n)^{5/2}\big) + O_p\big((K\Delta_n)^{1/2}n^{-\alpha}n^{-1/2}\big).\end{aligned}$$

Since $[\theta, \lambda_n/n]_T \to_p [\theta, \lambda]_T$, Corollary 2.3 entails that $\mathrm{TSQC}_{B,K_1,K_2}(\widehat{\Theta}^n, \widetilde{\Lambda}^n)$ is consistent. With the definitions in Remark 5, $\mathrm{TSQC}_{B,K_1,K_2}(\widehat{\Theta}^n, \widehat{\Lambda}^n)$ is also consistent. Also, consider the process $\rho_t(\cdot,\cdot)$, given by

$$\rho(\theta, \lambda)_t = \frac{[\theta, \lambda]_t}{([\theta, \theta]_t[\lambda, \lambda]_t)^{1/2}}.$$

Notice that $0 \le \rho(\theta, \lambda)_t \le 1$ for all t due to the Kunita–Watanabe inequality (Protter (2004), Theorem II.25, page 69). For each t we see that $\rho(\theta, \lambda_n)_t = \rho(\theta, \lambda_n/n)_t \xrightarrow{p} \rho(\theta, \lambda)_t$ by the continuous mapping theorem, which means that the coefficient $\rho(\theta, \lambda)_t$ can be consistently estimated using the estimators $\widehat{\Theta}^n_t$ and $\widehat{\Lambda}^n_t$, the latter simply defined as $\widehat{\Lambda}^n_t = N_{n,t}$. In particular, define

$$\rho_{\mathrm{TSQC}}\big(\widehat{\Theta}^n, \widehat{\Lambda}^n\big)_T = \frac{\mathrm{TSQC}(\widehat{\Theta}^n, \widehat{\Lambda}^n)_T}{(\mathrm{TSQC}(\widehat{\Theta}^n)_T\mathrm{TSQC}(\widehat{\Lambda}^n)_T)^{1/2}},$$

and note that $\rho_{\mathrm{TSQC}}(\widehat{\Theta}^n, \widehat{\Lambda}^n)_T = \rho_{\mathrm{TSQC}}(\widehat{\Theta}^n, \widetilde{\Lambda}^n)_T$, from which consistency of this estimator follows. When $\widehat{\Theta}^n$ and $\widetilde{\Lambda}^n$ have different convergence rates, as in Lemma 2.2, another consistent estimator for $\rho(\theta, \lambda)_t$ is $\mathrm{QV}_{B,K_2}(\widehat{\Theta}^n, \widetilde{\Lambda}^n)/(\mathrm{TSQC}(\widehat{\Theta}^n)_T\mathrm{TSQC}(\widehat{\Lambda}^n)_T)^{1/2}$. Since the speed at which $\mathrm{QV}_{B,K}$ converges is governed by the inferior convergence rate, there is, however, not that much to be gained in using this latter estimator, potentially apart from some less fine tuning of the K_1 and K_2 parameters. These two estimators of $\rho(\theta, \lambda)_t$ have a similar flavour to them, but are different from, the first-order correlation estimator introduced in Barndorff-Nielsen and Shephard (2004), Sections 3.1-3.2, pages 899–903.

In Section 4.2, we study the performance of various TSQC-estimators on simulated data. Before proceeding to the simulations and the empirical application, we provide an example of a simple model satisfying the above assumptions.

EXAMPLE 1 (A volatility-intensity model). Suppose that we observe samples from the process $X_t = X_0 + \int_0^t \sigma_s \, \mathrm{d}W_s$, where the spot volatility and the intensity both follow CIR-processes (Cox, Ingersoll and Ross (1985)) given by,

$$\begin{aligned}&\mathrm{d}\sigma_t^2 = \kappa\big(\alpha - \sigma_t^2\big)\,\mathrm{d}t + \gamma\sigma_t\,\mathrm{d}Z_t, \quad \sigma_0^2 = \alpha, \\ &\mathrm{d}\lambda_{n,t} = \beta_n(\xi_n - \lambda_{n,t})\,\mathrm{d}t + \nu_n\lambda_{n,t}^{1/2}\,\mathrm{d}B_t, \quad \lambda_{n,0} = \xi^n,\end{aligned} \tag{4.3}$$

where Z_t and B_t are Wiener processes such that $\mathrm{corr}(Z_t, B_t) = \rho$, and W_t is a Wiener process that may or may not be correlated with Z_t. The parameters κ, α and γ as well as β_n, ξ_n and ν_n are positive and we assume that the Feller condition (Feller (1951)) holds for both the volatility and the intensity, that is $2\kappa\alpha \ge \gamma^2$, and $2\beta_n\xi_n \ge \nu_n^2$ for all $n \ge 1$. In this model, the dependency between σ_t^2 and $\lambda_{n,t}$ is introduced by the correlation between Z_t and B_t. Suppose that $\xi^n = n\xi$, $\nu_n = \sqrt{n}\nu$ and that $0 < \beta \le \beta_n \to \infty$ as $n \to \infty$. Then, for each $t \in [0, T]$, we have that $n^{-1}\Lambda_{n,t} \xrightarrow{p} \xi t$, and that $n^{-1}[\sigma^2, \lambda_n]_t \to_p [\sigma^2, \lambda]_t = \rho\gamma\nu\xi^{1/2}\int_0^t \sigma_s \,\mathrm{d}s$ as $n \to \infty$. The expression on the right is an estimand for the TSQC estimator. See Appendix C in the Supplementary Material for details related to this example. In the next section the model of Example 1 is used as the basis for a simulation study.

4.2. *Simulations*. To investigate the finite sample properties of the TSQC estimators we simulated 1000 datasets from the model of Example 1. The initial observations (i.e., at time zero) for the volatility and intensity processes were sampled from a Gamma distribution with parameters $(2\kappa\alpha/\gamma^2, 2\kappa/\gamma^2)$ and a Gamma distribution with parameters $(2\beta_n\xi_n/\nu_n^2, 2\beta_n/\nu_n^2)$, respectively. The parameter values were $\alpha = 2.172, \kappa = 2.345, \gamma = 1.000$ (volatility model), $\xi_n = n6.912$, $\beta_n = n^{1/4}0.305$, $\nu = \sqrt{n}1.000$, with $n = 3500$. The microstructure noise was taken as additive on the efficient price and independent of the three underlying Wiener processes, that is, we observe

$$Y_{t_i} = X_{t_i} + \varepsilon_{t_i},$$

where the ε_{t_i} were independent mean zero normals with standard deviation 0.001, independent of W, Z and B. These three process were all Wiener processes, W was independent of Z and B, while Z and B were dependent with correlation $\rho = 0.765$. All the processes were simulated over the unit interval, and the average sample size over the 1000 simulations were 24114. This corresponds roughly to one observation per second over one trading day. In other words, our simulation study is *not* of the ultra high frequency type (for example, in the empirical application of Section 4.3 the daily sample size is about ten times as big as in these simulations).

For each simulated dataset we estimated the quadratic covariation between the volatility and the intensity $[\sigma^2, \lambda]$, the volatility of volatility $[\sigma^2, \sigma^2]$, and the volatility of the intensity $[\lambda, \lambda]$. The first two of these estimands were estimated both using the observations Y_{t_i} (i.e., the efficient price contaminated by microstructure noise), and also assuming that the efficient price X_{t_i} is observed directly (to estimate $[\lambda, \lambda]$ only the observations times are used, and these are not affected by microstucture noise). In the latter case, the integrated volatility was simply estimated using the realised volatility $\widehat{\Theta} = \sum_{t_{i+1}\leq 1}(X_{t_{i+1}} - X_{t_i})^2$, while for the noisy observations Y_{t_i} we estimated the integrated volatility using the Two-Scales Realised Volatility (TSRV) of Zhang, Mykland and Aït-Sahalia (2005). The TSRV we employed is given by

$$\widehat{\Theta} = \left\{\left(1 - \frac{K - J + 1/3}{N}\right)(K - J)\right\}^{-1}\left\{K[\bar{Y}, \bar{Y}]^{(K)} - J[\bar{Y}, \bar{Y}]^{(J)}\right\}, \tag{4.4}$$

where $\bar{Y}$ are pre-averaged observations, and $[\bar{Y}, \bar{Y}]^{(K)} = K^{-1}\sum_{i=1}^{N-K}(\bar{Y}_{i+K} - \bar{Y}_i)^2$, where N are the number of 'observations' of $\bar{Y}$, and K is a tuning parameter chosen by the user (Mykland, Zhang and Chen (2019), equation (17), page 106, for this construction). For the pre-averaging, we used a rolling average of five adjacent observations, and set the tuning parameters in the TSRV to $K = 10$ and $J = 5$. All three estimands were estimated using the $\mathrm{TSQC}_{B,K_1,K_2}$ estimator of (2.6), with $B = 10^5$, and tuning parameters $K_2 = 2K_1$ with $K_1 = 400$. The uncertainty of our TSQC estimates were in all cases computed using the two-scales observed AVAR introduced in Mykland and Zhang ((2017a) Definition 4, equation (26)), page 209, (with $K_1 = 2K_2$ and $K_2 = 500$, using the notation of the cited paper). The applicability of the two-scales observed AVAR to the TSQC estimators is discussed in Section 3.2. Figure 1 contains histograms of the approximate pivots

$$\frac{\mathrm{TSQC}(\cdot,\cdot) - \text{estimand}}{\sqrt{\mathrm{TSAVAR}}},$$

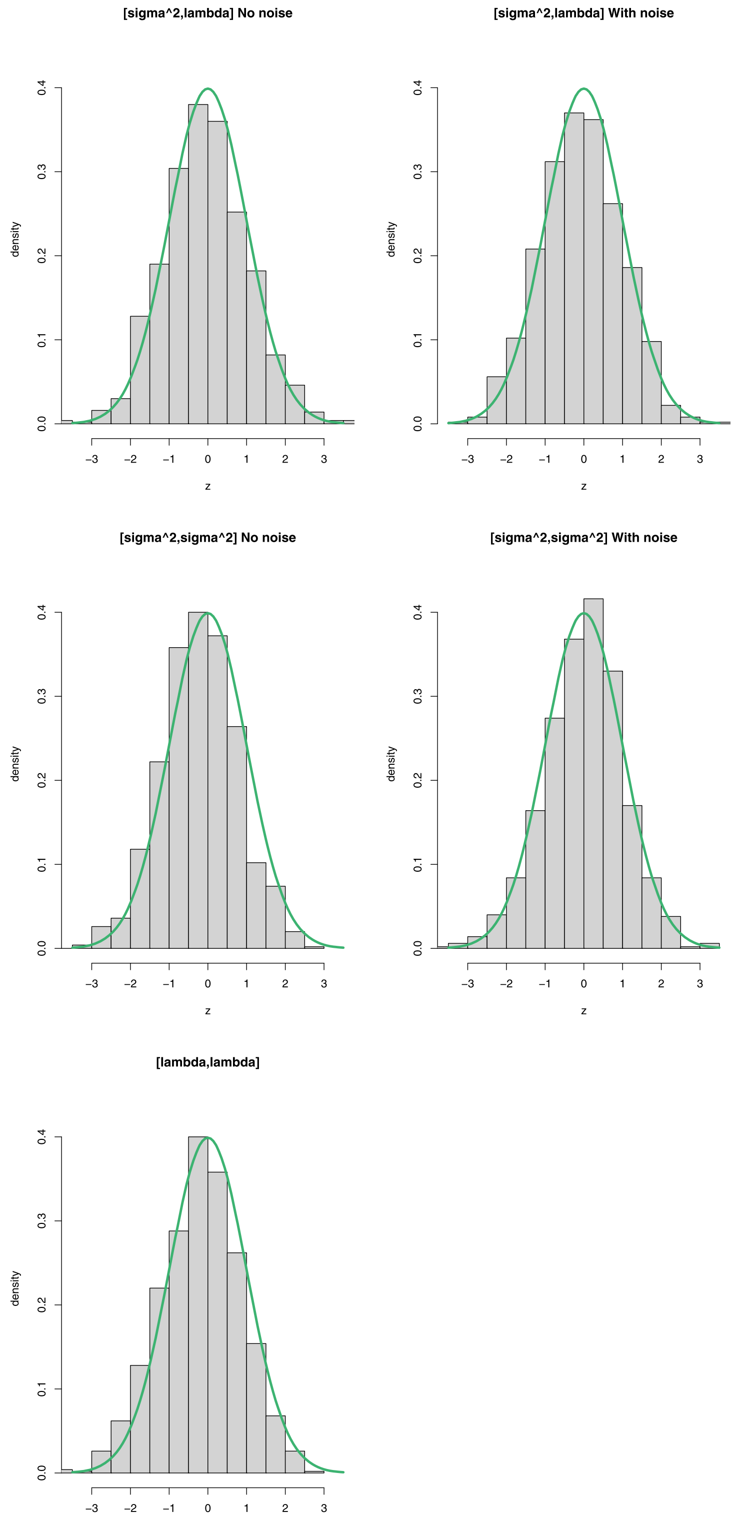

FIG. 1. 1000 *simulated datasets with an average of* 24114 *observations per dataset. The parameter values are as described in Section* 4.2. *The green curves indicate the density of the standard normal distribution.*

TABLE 1

Estimates based on 1000 *Monte Carlo replications. The average sample size over these replications were* 24114. *The bias and the root mean squared error* (*rmse*) *are the average of the* 1000 *replicates of the nonnormalised deviances* TSQC$(\cdot,\cdot)$ − estimand

Estimand	Truth	Bias	Rmse	90% Cov. prob.	95% Cov. prob.
$[\sigma^2, \lambda]$ (no noise)	2.917	−0.153	28.357	0.874	0.933
$[\sigma^2, \lambda]$ (with noise)		−0.961	178.088	0.889	0.943
$[\sigma^2, \sigma^2]$ (no noise)	2.146	−0.731	24.802	0.895	0.951
$[\sigma^2, \sigma^2]$ (with noise)		2.448	1211.322	0.876	0.939
$[\lambda, \lambda]$	6.889	−1.343	129.210	0.877	0.930

with the density of the standard normal distribution indicated by green curves. Summary statistics from the simulations are given in Table 1. The bias and the root mean squared error of our estimators are computed based on the nonnormalised deviances TSQC − estimand. The take-away from the histograms in Figure 1 and the summary statistics reported in Table 1 is that for the quantities here estimated the normal approximation to the normalised TSQC estimator works well even in what are, in a high frequency econometrics context, rather small samples.

4.3. *An empirical application.* In the empirical study, we analyse features of the Apple stock as traded over a period of 21 trading days in January 2018. All transactions registered in the U.S. National Market System conducted between 9:45 am–3:45 pm Eastern Standard Time are included. The reason for choosing this window is to avoid abnormal trading activity during the opening and closing of the New York Stock Exchange, and to avoid those pre- and post-market hours during which the trading frequency is low (Wang and Mykland (2014), page 205). The Apple stock data is recorded down to the nanosecond (10^{-9} seconds), and for the period under study the mean number of transactions over a trading day during the time window we use was 203924, which is about nine transactions per second. After some data cleaning, the data was pre-averaged and the TSRV estimator of Zhang, Mykland and Aït-Sahalia (2005) was used to estimate the integrated volatility. As documented by Jacod, Li and Zheng (2017), a positive serial correlation of the microstructure noise can be an issue when the data are sampled at ultra high frequencies (as here). In this application, the serial correlation should, however, for the most part be taken care of by our choice of tuning parameters in the TSRV estimator combined with the rather wide windows over which the data are pre-averaged. Specifically, we used $J = 20$ and $K = 40$ (see equation (4.4)), choices that filters out most dependence in the noise (see, e.g., Aït-Sahalia, Mykland and Zhang (2011), Section 4, pages 165–167), and a pre-averaging window of, on average, 12 observations (see Mykland, Zhang and Chen (2019), page 106).

The cumulative intensity of the observation times was estimated by $10^{-6}N_t$, where N_t counts the number of transactions conducted from 9:45 am to 9:45 am plus t. Besides making the plots more aesthetically pleasing, the number 10^{-6} plays no role. We used the TSQC estimator for daily estimation of the volatility-intensity covariance matrix and the two transformations thereof, $\rho(\sigma^2, \lambda)_t$ and β_t. The estimates of $\rho(\sigma^2, \lambda)_t$ are time-varying and lie between 0.5 and 0.8 for most of the days under study, indicating that the two processes are indeed correlated. To estimate the (pointwise) confidence bands of our TSQC estimators we employed the observed asymptotic variance. As discussed in Section 3.2, the applicability of the observed asymptotic variance is ensured by Corollary 3.4. Figure 2 summarises the results of this empirical application.

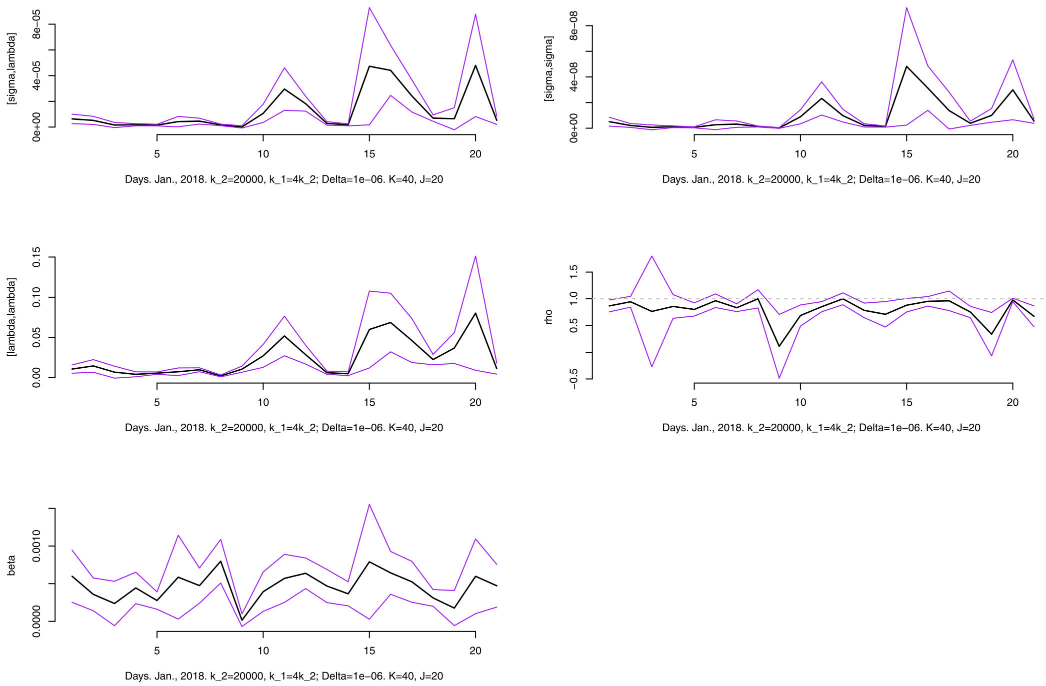

FIG. 2. *The Apple stock January* 2.-31., 2018. *Daily estimates of* $[\sigma^2, \sigma^2]_T$, $[\sigma^2, \lambda]_T$ *and* $[\lambda, \lambda]_T$, *as well as the parameters* β_T *and* ρ_T. *The TSRV was used as the estimator of the integrated volatility. The purple lines are pointwise* 95 *percent confidence bands computed using the observed asymptotic variance of* Mykland and Zhang (2017*a*), *along with the delta method. In the plot with the daily estimates of* ρ_T, *the value* 1 *is indicated by the dashed grey line.*

4.4. *Using the volatility-intensity relationship to gain efficiency.* We have seen above that $d\theta_t = \beta_t \, d\lambda_t + dZ_t$ and $d\theta_{n,t} = \beta_{n,t} \, d\lambda_{n,t} + dZ_{n,t}$, where, in the latter equation, there is no normalisation by n, hence the two equations are equivalent, and, once again, one can calculate as if n were known. This is an ANOVA decomposition along the lines of Mykland and Zhang (2006), but in this case, θ and λ are unobserved. The process β is estimated as above in this paper. The quantities θ and λ can be estimated as spot (instantaneous) quantities, as in Mykland and Zhang (2008).

When microstructure is present in prices, but not in the observation times (as is the usual understanding), then $\widehat{\Lambda}_{n,t}$ has a faster rate of convergence than $\widehat{\Theta}_{n,t}$, and hence this is also true for $\widehat{\lambda}_{n,t}$ and $\widehat{\theta}_{n,t}$. The construction in Mykland and Zhang (2008) uses $\widehat{\theta}_{n,t} = (\widehat{\Theta}_{n,t} - \widehat{\Theta}_{n,t-h_{n,\theta}})/h_{n,\theta}$, and similarly for $\widehat{\lambda}_{n,t}$, where $h_{n,\theta}$ and $h_{n,\lambda}$ are chosen to be (at least rate-) optimal, by the use of a variance-variance tradeoff. This leads to the rates for $\widehat{\theta}_{n,t}$ and $\widehat{\lambda}_{n,t}$ to be $n^{-1/8}$ and $n^{-1/4}$, respectively (when a rate optimal estimator of volatility is used, such as the S-TSRV which is used in this paper, or the multiscale estimator of Zhang (2006); see also Bibinger and Mykland (2016) for the multivariate case and the connection to realised kernels, as well as the references therein). Finally, Lemma 2.2 and Theorem 3.1 provide for $\widehat{\beta}_{n,t}$ to have a rate of convergence of $n^{-1/4}$, thus

$$\int_0^t \widehat{\beta}_{n,s} \, d\widehat{\lambda}_{n,s} - \int_0^t \beta_{n,s} \, d\lambda_{n,s} = O_p(n^{-1/4}).$$

If, as in our data, the residual Z_t (or $\widehat{Z}_{n,t}$) is small, the question naturally occurs whether to prefer $\widehat{\theta}_{n,t}$, with a low rate of convergence, or $\int_0^t \widehat{\beta}_{n,s} \, d\widehat{\lambda}_{n,s}$ with a much better rate of convergence, but with a bias of Z_t (or $Z_{n,t}$). The conventional asymptotics-based answer to this question is that a slow convergence rate of $O_p(n^{-1/8})$ is preferable to a much better convergence rate $O_p(n^{-1/4})$ to a limit with an $O_p(1)$ bias. In other words, pick $\widehat{\theta}_{n,t}$, even if Z_t is small.

This answer is uncomfortable, and has already caused some degree of argument in connection with volatility estimation, where there is an argument over whether intra-day estimators are always preferable, or whether to draw on longer time periods. Assumptions of stationarity will not help, and longer time periods are usually introduced by drawing on more highly specified models, such as ARCH and GARCH type models, going back to the seminal papers of Engle (1982) and Bollerslev (1986). There is a huge literature in this area; see, for example, the survey by Engle (1995).

Another path is to express "$Z_{n,t}$ is small" by a triangular array asymptotic regime whereby $Z_{n,t} = o_p(1)$ as $n \to \infty$. Triangular array asymptotic regimes are often used close to a singularity; see, for example, Chan and Wei (1987) and Phillips (1987) in the context of time series close to the unit root. In this context, it is often referred to as 'local to unity asymptotics'. Under this regime, one can augment the estimate of $d\theta$ by adding an estimate of $\beta_t \,\mathrm{d}\lambda_t$, giving rise to an estimate of the form

$$\breve{\theta}_{n,t} = c_n \widehat{\theta}_{n,t} + (1 - c_n) \int_0^t \widehat{\beta}_{n,t} \,\mathrm{d}\widehat{\lambda}_{n,t}. \tag{4.5}$$

The tuning parameter c_n should then be chosen to minimise the (random) mean squared error in $\breve{\theta}$, and in any case, $c_n \to 0$, thus improving the rate of convergence. A proper analysis of (4.5) would require an assessment of the mean squared error of $\breve{\theta}_{n,t}$, which would presumably involve the estimation of $[Z_n, Z_n]_t$, which brings us back to the ANOVA problem of Mykland and Zhang (2006), but now with latent variables everywhere. This is beyond the scope of the present paper.

If a reasonable solution can be found, similar methods may apply to a number of estimators that involve the estimation of spot volatility, such as leverage effect (see Section 3.3), volatility of volatility (in this paper, and also Vetter (2015) and Mykland and Zhang (2017a)), as well as regression, and ANOVA (Mykland and Zhang (2009), Section 4.2, pages 1424–1426, Zhang (2012), Section 4, pages 268–273, Reiß, Todorov and Tauchen (2015), and the references therein).

5. Conclusion. This paper introduces a consistent estimator of the quadratic covariation between two nonobservable spot-process semimartingales, derives the convergence rates of this estimator, and presents a central limit theorem for such estimators. The main theoretical contribution of the paper is this central limit theorem, a theorem that is applicable to a wide range of estimators based on triangular arrays of rolling quadratic covariations and second differencing of estimators of integrated spot processes.

As recognised in much recent literature on estimation in high frequency data, the assumption of exogenous observation times is often untenable, and one typically allows for dependency between the observation times and the price process. In this paper, we have considered possible dependencies between the observation times and nonobservable spot-processes associated with the price process, of which the spot volatility is a prime example. A simulation study shows that the estimators perform well with decent amounts of data. The empirical study of the Apple stock indicates that the observation times and the volatility process of this stock are positively correlated.

Acknowledgments. Emil A. Stoltenberg would like to thank the Fulbright Foundation for financial support; The University of Chicago for their hospitality; Sylvie Bendier Decety for her kindness; Nils Lid Hjort for his encouragement; the statistics section at the Department of Mathematics, University of Oslo; and the PharmaTox Strategic Research Initiative at the Faculty of Mathematics and Natural Sciences, University of Oslo.

Funding. The authors are grateful for the detailed comments of an anonymous referee and the Editor. The authors would also like to thank the United States National Science Foundation under grants DMS 17-13118 and DMS-2015530 (Zhang), and DMS 17-13129 and DMS-2015544 (Mykland).

SUPPLEMENTARY MATERIAL

A CLT for second difference estimators with an application to volatility and intensity (DOI: 10.1214/22-AOS2176SUPP; .pdf). Appendix A of the supplement contains definitions and conditions that are used throughout the article. Appendix B presents a stable central limit theorem for càdlàg martingales, as well as a corollary giving conditions that may replace the Lindeberg condition of the theorem. Appendix C gives details on Example 1. Appendix D provides some background for Theorem 2.1 of the main text, while Appendix E contains the full proof of Theorem 3.1 of the main text. Appendix F contains the full proof of the central limit theorem, Theorem 3.2, of the main text.

REFERENCES

AÏT-SAHALIA, Y. and JACOD, J. (2014). *High-Frequency Financial Econometrics*. Princeton Univ. Press, Princeton.

AÏT-SAHALIA, Y., MYKLAND, P. A. and ZHANG, L. (2011). Ultra high frequency volatility estimation with dependent microstructure noise. *J. Econometrics* **160** 160–175. MR2745875 https://doi.org/10.1016/j.jeconom.2010.03.028

AÏT-SAHALIA, Y., FAN, J., LAEVEN, R. J. A., WANG, C. D. and YANG, X. (2017). Estimation of the continuous and discontinuous leverage effects. *J. Amer. Statist. Assoc.* **112** 1744–1758. MR3750896 https://doi.org/10.1080/01621459.2016.1240082

ANDERSEN, T. G., BOLLERSLEV, T., DIEBOLD, F. X. and LABYS, P. (2001). The distribution of realized exchange rate volatility. *J. Amer. Statist. Assoc.* **96** 42–55. MR1952727 https://doi.org/10.1198/016214501750332965

BARNDORFF-NIELSEN, O. E. and SHEPHARD, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **64** 253–280. MR1904704 https://doi.org/10.1111/1467-9868.00336

BARNDORFF-NIELSEN, O. E. and SHEPHARD, N. (2004). Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica* **72** 885–925. MR2051439 https://doi.org/10.1111/j.1468-0262.2004.00515.x

BIBINGER, M. and MYKLAND, P. A. (2016). Inference for multi-dimensional high-frequency data with an application to conditional independence testing. *Scand. J. Stat.* **43** 1078–1102. MR3573677 https://doi.org/10.1111/sjos.12230

BILLINGSLEY, P. (1995). *Probability and Measure*, 3rd ed. *Wiley Series in Probability and Mathematical Statistics*. Wiley, New York. MR1324786

BOLLERSLEV, T. (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31** 307–327. MR0853051 https://doi.org/10.1016/0304-4076(86)90063-1

CHAN, N. H. and WEI, C. Z. (1987). Asymptotic inference for nearly nonstationary AR(1) processes. *Ann. Statist.* **15** 1050–1063. MR0902245 https://doi.org/10.1214/aos/1176350492

COX, J. C., INGERSOLL, J. E. JR. and ROSS, S. A. (1985). A theory of the term structure of interest rates. *Econometrica* **53** 385–407. MR0785475 https://doi.org/10.2307/1911242

ENGLE, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50** 987–1007. MR0666121 https://doi.org/10.2307/1912773

ENGLE, R. F. (1995). *ARCH*: *Selected Readings*. Oxford Univ. Press, New York, NY.

FELLER, W. (1951). Two singular diffusion problems. *Ann. of Math.* (2) **54** 173–182. MR0054814 https://doi.org/10.2307/1969318

FUKASAWA, M. and ROSENBAUM, M. (2012). Central limit theorems for realized volatility under hitting times of an irregular grid. *Stochastic Process. Appl.* **122** 3901–3920. MR2971719 https://doi.org/10.1016/j.spa.2012.08.005

HAYASHI, T., JACOD, J. and YOSHIDA, N. (2011). Irregular sampling and central limit theorems for power variations: The continuous case. *Ann. Inst. Henri Poincaré Probab. Stat.* **47** 1197–1218. MR2884231 https://doi.org/10.1214/11-AIHP432

JACOD, J. (1997). On continuous conditional Gaussian martingales and stable convergence in law. In *Séminaire de Probabilités*, *XXXI*. *Lecture Notes in Math*. **1655** 232–246. Springer, Berlin. MR1478732 https://doi.org/10.1007/BFb0119308

JACOD, J., LI, Y. and ZHENG, X. (2017). Statistical properties of microstructure noise. *Econometrica* **85** 1133–1174. MR3681766 https://doi.org/10.3982/ECTA13085

JACOD, J., LI, Y. and ZHENG, X. (2019). Estimating the integrated volatility with tick observations. *J. Econometrics* **208** 80–100. MR3906962 https://doi.org/10.1016/j.jeconom.2018.09.006

JACOD, J. and PROTTER, P. (1998). Asymptotic error distributions for the Euler method for stochastic differential equations. *Ann. Probab.* **26** 267–307. MR1617049 https://doi.org/10.1214/aop/1022855419

JACOD, J. and PROTTER, P. (2012). *Discretization of Processes*. *Stochastic Modelling and Applied Probability* **67**. Springer, Heidelberg. MR2859096 https://doi.org/10.1007/978-3-642-24127-7

JACOD, J. and SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **288**. Springer, Berlin. MR1943877 https://doi.org/10.1007/978-3-662-05265-5

JACOD, J. and TODOROV, V. (2010). Do price and volatility jump together? *Ann. Appl. Probab.* **20** 1425–1469. MR2676944 https://doi.org/10.1214/09-AAP654

KALNINA, I. and XIU, D. (2017). Nonparametric estimation of the leverage effect: A trade-off between robustness and efficiency. *J. Amer. Statist. Assoc.* **112** 384–396. MR3646579 https://doi.org/10.1080/01621459.2016.1141687

LI, J., TODOROV, V. and TAUCHEN, G. (2017a). Jump regressions. *Econometrica* **85** 173–195. MR3611769 https://doi.org/10.3982/ECTA12962

LI, J., TODOROV, V. and TAUCHEN, G. (2017b). Robust jump regressions. *J. Amer. Statist. Assoc.* **112** 332–341. MR3646575 https://doi.org/10.1080/01621459.2016.1138866

LI, Y., ZHANG, Z. and ZHENG, X. (2013). Volatility inference in the presence of both endogenous time and microstructure noise. *Stochastic Process. Appl.* **123** 2696–2727. MR3054542 https://doi.org/10.1016/j.spa.2013.04.002

LI, Y., MYKLAND, P. A., RENAULT, E., ZHANG, L. and ZHENG, X. (2014). Realized volatility when sampling times are possibly endogenous. *Econometric Theory* **30** 580–605. MR3205607 https://doi.org/10.1017/S0266466613000418

MYKLAND, P. A. and ZHANG, L. (2006). ANOVA for diffusions and Itô processes. *Ann. Statist.* **34** 1931–1963. MR2283722 https://doi.org/10.1214/009053606000000452

MYKLAND, P. A. and ZHANG, L. (2008). Inference for volatility-type objects and implications for hedging. *Stat. Interface* **1** 255–278. MR2476743 https://doi.org/10.4310/SII.2008.v1.n2.a4

MYKLAND, P. A. and ZHANG, L. (2009). Inference for continuous semimartingales observed at high frequency. *Econometrica* **77** 1403–1445. MR2561071 https://doi.org/10.3982/ECTA7417

MYKLAND, P. A. and ZHANG, L. (2012). The econometrics of high-frequency data. In *Statistical Methods for Stochastic Differential Equations*. *Monogr. Statist. Appl. Probab.* **124** 109–190. CRC Press, Boca Raton, FL. MR2976983 https://doi.org/10.1201/b12126-3

MYKLAND, P. A. and ZHANG, L. (2017a). Assessment of uncertainty in high frequency data: The observed asymptotic variance. *Econometrica* **85** 197–231. MR3611770 https://doi.org/10.3982/ECTA12501

MYKLAND, P. A. and ZHANG, L. (2017b). Supplement to "Assessment of uncertainty in high frequency data: The observed asymptotic variance". *Econometrica* **85**. Available at https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA12501.

MYKLAND, P. A., ZHANG, L. and CHEN, D. (2019). The algebra of two scales estimation, and the S-TSRV: High frequency estimation that is robust to sampling times. *J. Econometrics* **208** 101–119. MR3906963 https://doi.org/10.1016/j.jeconom.2018.09.007

PHILLIPS, P. C. B. (1987). Towards a unified asymptotic theory for autoregression. *Biometrika* **74** 535–547. MR0909357 https://doi.org/10.1093/biomet/74.3.535

POTIRON, Y. and MYKLAND, P. A. (2017). Estimation of integrated quadratic covariation with endogenous sampling times. *J. Econometrics* **197** 20–41. MR3598643 https://doi.org/10.1016/j.jeconom.2016.10.004

PROTTER, P. E. (2004). *Stochastic Integration and Differential Equations*, 2nd ed. *Applications of Mathematics* (*New York*) **21**. Springer, Berlin. MR2020294

REISS, M., TODOROV, V. and TAUCHEN, G. (2015). Nonparametric test for a constant beta between Itô semi-martingales based on high-frequency data. *Stochastic Process. Appl.* **125** 2955–2988. MR3343284 https://doi.org/10.1016/j.spa.2015.02.008

RENAULT, E. and WERKER, B. J. M. (2011). Causality effects in return volatility measures with random times. *J. Econometrics* **160** 272–279. MR2745883 https://doi.org/10.1016/j.jeconom.2010.03.036

STOLTENBERG, E. A., MYKLAND, P. A. and ZHANG, L. (2022). Supplement to "A CLT for second difference estimators with an application to volatility and intensity." https://doi.org/10.1214/22-AOS2176SUPP

VETTER, M. (2015). Estimation of integrated volatility of volatility with applications to goodness-of-fit testing. *Bernoulli* **21** 2393–2418. MR3378471 https://doi.org/10.3150/14-BEJ648

WANG, C. D. and MYKLAND, P. A. (2014). The estimation of leverage effect with high-frequency data. *J. Amer. Statist. Assoc.* **109** 197–215. MR3180557 https://doi.org/10.1080/01621459.2013.864189

ZHANG, L. (2006). Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Bernoulli* **12** 1019–1043. MR2274854 https://doi.org/10.3150/bj/1165269149

ZHANG, L. (2012). Implied and realized volatility: Empirical model selection. *Ann. Finance* **8** 259–275. MR2922798 https://doi.org/10.1007/s10436-010-0168-0

ZHANG, L., MYKLAND, P. A. and AÏT-SAHALIA, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *J. Amer. Statist. Assoc.* **100** 1394–1411. MR2236450 https://doi.org/10.1198/016214505000000169