CrossMark

Asymptotic properties of the realized skewness and related statistics

Yuta Koike[1,2,3,4] · **Zhi Liu**[5]

Abstract The recent empirical works have pointed out that the realized skewness, which is the sample skewness of intraday high-frequency returns of a financial asset, serves as forecasting future returns in the cross section. Theoretically, the realized skewness is interpreted as the sample skewness of returns of a discretely observed semimartingale in a fixed interval. The aim of this paper is to investigate the asymptotic property of the realized skewness in such a framework. We also develop an estimation theory for the limiting characteristic of the realized skewness in a situation where measurement errors are present and sampling times are stochastic.

Keywords High-frequency data · Itô semimartingale · Jumps · Microstructure noise · Realized skewness · Stochastic sampling

✉ Zhi Liu
  liuzhi@umac.mo

1   Graduate School of Mathematical Sciences, University of Tokyo, Tokyo 153-8914, Japan

2   Department of Business Administration, Graduate School of Social Sciences, Tokyo Metropolitan University, Tokyo 100-0005, Japan

3   The Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan

4   CREST, Japan Science and Technology Agency, Kawaguchi 333-0012, Japan

5   Department of Mathematics, University of Macau, Avenida da Universidade, Taipa, Macau 999078, China

🙋 Springer

1 Introduction

In the past decades, with widely available high-frequency financial data, statistical inference for stochastic processes has significantly been developed. Among others, inference for the quadratic variation of a semimartingale using high-frequency data is particularly of interest in the literature, due to its important applications in finance, namely measuring the fluctuation of security markets; see Jacod and Protter (1998), Jacod (2008), Andersen et al. (2005), Bandi and Russell (2006) and references therein.

In practice, the quadratic variation of a semimartingale is important in finance because it can be considered as a realized measure of the variance of short-period returns. Besides, higher moments rather than the variance, in particular the third moment and the fourth moment which appear in measuring the skewness and kurtosis of assets, have attracted vast attention in finance, see Bakshi et al. (2003), Friend and Westerfield (1980), Martellini and Ziemann (2009), Harvey and Siddique (1999, 2000), Mitton and Vorkink (2007), Kozhan et al. (2013), among others. By using high-frequency data, the efficiency of estimating the quadratic variation has substantially improved. Thus, a natural question is whether we can achieve some improvements by using high-frequency data in the inferences for higher-order realized moments. In the empirical aspect, recently Amaya et al. (2015) have showed strong evidence that the sample skewness of intraday high-frequency returns, which is called the *realized skewness* in the paper[1], serves as predicting future equity returns in the cross section. More precisely, they have found that if a stock's realized skewness averaged over a week is relatively higher (resp. lower) than other stocks' ones (e.g., more than the 90% quantile (resp. less than the 10% quantile) of all stocks' ones), then the stock's return in the next week tends to be negative (resp. positive). They have also confirmed that this empirical finding is robust across various implementations. The asymptotic property of the realized skewness is briefly discussed in Amaya et al. (2015) as well. The aim of this paper is to investigate this point more deeply. Specifically, suppose that the dynamics of the log price process of an asset is modeled by an Itô semimartingale $X = (X_t)_{t \geq 0}$ and we have discrete observation data $\{X_{i\Delta_n}\}_{i=0}^{\lfloor T/\Delta_n \rfloor}$ on the interval $[0, T]$, where Δ_n is a positive number tending to 0 as $n \to \infty$. Then the realized skewness is given by

$$\text{RDSkew} = \frac{\lfloor T/\Delta_n \rfloor \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} (\Delta_i^n X)^3}{\left\{ \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} (\Delta_i^n X)^2 \right\}^{3/2}}.$$

Amaya et al. (2015) have pointed out that

$$\sum_{i=1}^{\lfloor T/\Delta_n \rfloor} (\Delta_i^n X)^2 \to^P [X, X]_T \quad \text{and} \quad \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} (\Delta_i^n X)^3 \to^P \sum_{0 \leq s \leq T} (\Delta X_s)^3,$$

---

[1] Neuberger (2012) uses the term realized skewness for a different concept.

where $\Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}$, $\to^P$ denotes convergence in probability, $[X, X]$ denotes the quadratic variation process of X and $\Delta X_s = X_s - X_{s-}$. Hence, the appropriately scaled realized skewness, $\mathrm{RDSkew}/\lfloor T/\Delta_n \rfloor$, is a consistent estimator for the following quantity:

$$\frac{\sum_{0 \leq s \leq T}(\Delta X_s)^3}{([X, X]_T)^{3/2}}. \tag{1}$$

In this paper, we aim at deriving the asymptotic distribution of the estimation error

$$\frac{1}{\lfloor T/\Delta_n \rfloor}\mathrm{RDSkew} - \frac{\sum_{0 \leq s \leq T}(\Delta X_s)^3}{([X, X]_T)^{3/2}}. \tag{2}$$

We shall remark that the asymptotic property of the statistic of the form

$$\sum_{i=1}^{\lfloor T/\Delta_n \rfloor} g(\Delta_i^n X), \tag{3}$$

where g is a function on $\mathbb{R}$ satisfying some smoothness condition, is well studied in the literature. To our knowledge, the most general condition to derive the asymptotic distribution of the above statistic is given by Theorem 5.1.2 from Jacod and Protter (2012), which requires that g is of class C^2 and satisfies $g(0) = 0$, $g'(0) = 0$ and $g''(x) = o(|x|)$ as $x \to 0$. Unfortunately, this condition is not satisfied by the cubic function $g(x) = x^3$, so this theorem is not applicable to deriving the asymptotic distribution of Eq. (2). Kinnebrock and Podolskij (2008) proved the result for $g(x) = x^3$ when X is continuous. One aim of this paper is to fill in this gap.

Another important issue in high-frequency financial econometrics is to take account of microstructure noise and randomness of observation times: At ultra high frequencies, asset prices are usually modeled as discrete observations of a semimartingale with observation noise, which is referred to as microstructure noise, because such data typically exhibit several empirical properties which are inconsistent with the semimartingale assumption. In addition, "raw" high-frequency financial data are usually recorded at certain event times such as transaction times or order arrival times, which would be random and depend on observed values. See Chapters 7 and 9 of Aït-Sahalia and Jacod (2014) and references therein for more details on this topic. This paper also deals with this issue. Namely, we construct a consistent estimator for quantity Eq. (1) and develop an associated asymptotic distribution theory when microstructure noise is present and the sampling scheme is stochastic. To accomplish this, we study the asymptotic property of the "pre-averaged" version of the statistic Eq. (3) with the cubic function $g(x) = x^3$. Here, "pre-averaging" is a denoising scheme which enables us to systematically adapt functionals of semimartingale increments [such as Eq. (3)] to the case that microstructure noise is present. The method was originally introduced in Podolskij and Vetter (2009) and subsequently generalized in Jacod et al. (2009), and many theoretical results on it are now available in the literature. In particular, under mild regularity assumptions, Theorem 16.3.1 of Jacod

and Protter (2012) provides the asymptotic distribution of the pre-averaged version of the statistic Eq. (3) (in the equidistant case) when g is a linear combination of positively homogeneous C^2 functions with degree (strictly) bigger than 3. Here, a function $f : \mathbb{R} \to \mathbb{R}$ is said to be positively homogeneous with degree $w \geq 0$ if $f(\alpha x) = \alpha^w f(x)$ for any $\alpha \geq 0$ and any $x \in \mathbb{R}$. Hence, the condition on the function g again rules out the cubic function. We thus need to perform an additional analysis to cover the cubic function. We will also show that the randomness of observation times has no essential impact on the asymptotic distribution of the pre-averaged version of the statistic Eq. (3) with the cubic function $g(x) = x^3$. This kind of phenomenon has already been observed in Koike (2016, 2017) for the pre-averaged version of the realized volatility. It contrasts the non-noisy case because the randomness of observation times can cause non-trivial modification of the asymptotic distribution of the realized volatility as illustrated in Fukasawa (2010), Li et al. (2014), Bibinger and Vetter (2015) and Vetter and Zwingmann (2017) for example.

The remainder of this paper is arranged as follows. Section 2 investigates the asymptotic property of statistic Eq. (3) and derives the asymptotic distribution of the realized skewness. Section 3 develops an estimation theory in a situation with microstructure noise and stochastic sampling times. Section 4 is devoted to the proofs.

2 The asymptotic distribution of the realized skewness

On a filtered probability space $\mathcal{B} = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$, we consider a stochastic process $(X_t)_{t \geq 0}$ of the form

$$
X_t = X_0 + \int_0^t b_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}B_s
$$
$$
+ \int_0^t \int_{|\delta(s,z)| \leq 1} \delta(s,z)(\mu - \nu)(\mathrm{d}s, \mathrm{d}z) + \int_0^t \int_{|\delta(s,z)| > 1} \delta(s,z)\mu(\mathrm{d}s, \mathrm{d}z),
$$

where the drift process b is $(\mathcal{F}_t)$-progressively measurable, the spot volatility process σ is $(\mathcal{F}_t)$-adapted and càdlàg, B is a standard Brownian motion, μ is a Poisson random measure on $\mathbb{R}^+ \times E$ with predictable compensator $\nu(\mathrm{d}t, \mathrm{d}z) = \mathrm{d}t\lambda(\mathrm{d}z)$ and λ being a σ-finite measure on a Polish space $(E, \mathcal{E})$, and δ is a predictable function on $\Omega \times \mathbb{R}^+ \times E$.

We impose the following standard structural assumption:

[H] There are a sequence (τ_k) of stopping times increasing to infinity and a sequence (γ_k) of deterministic nonnegative measurable functions on E such that $\int \gamma_k(z)^2 \lambda(dz) < \infty$ and $|\delta(\omega, t, z)| \wedge 1 \leq \gamma_k(z)$ for all k and all (ω, t, z) with $t \leq \tau_k(\omega)$.

Let us assume that we observe the process X at equidistant discrete points $\{i\Delta_n\}_{i=0}^{\lfloor T/\Delta_n \rfloor}$ for some $T > 0$, where Δ_n is a sequence of positive numbers tending to zero as $n \to \infty$. We develop a central limit theorem for the non-normalized increments of X

$$\mathcal{V}_T^n(X, g) := \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} g(\Delta_i^n X)$$

for a function $g : \mathbb{R} \to \mathbb{R}$ satisfying some smoothness condition. If g is continuous and satisfies $g(x) = o(x^2)$ as $x \to 0$, it is known that

$$\mathcal{V}_T^n(X, g) \to^P \sum_{0 \leq s \leq T} g(\Delta X_s)$$

as $n \to \infty$; see e.g., Theorem 3.3.1 from Jacod and Protter (2012). If further g is of class C^2 and satisfies $g(0) = g'(0) = 0$ and $g''(x) = o(|x|)$ as $x \to 0$, a central limit theorem for $\mathcal{V}_T^n(X, g)$ is also known under Assumption [H] [see Theorem 5.1.2 of Jacod and Protter (2012)]. This condition is, however, not sufficient to allow the cubic function $g(x) = x^3$, which is crucial for deriving the asymptotic distribution of the realized skewness. Motivated by this reason, in the following we relax this condition to incorporate such a function.

We will use the notion of *stable convergence* denoted by $\to^S$. Here we briefly describe it before the main theorems. Let $(\mathcal{X}, \mathcal{A}, \mathbb{P})$ be a probability space and assume that we have a random element Z_n taking values in a Polish space S and defined on an extension $(\mathcal{X}_n, \mathcal{A}_n, \mathbb{P}_n)$ of $(\mathcal{X}, \mathcal{A}, \mathbb{P})$ for each $n \in \mathbb{N} \cup \{\infty\}$. In this setup, the sequence Z_n is said to *converge stably in law* to Z_∞ if $\mathbb{E}_n[Uf(Z_n)] \to \mathbb{E}_\infty[Uf(Z_\infty)]$ for any $\mathcal{A}$-measurable bounded random variable U and any bounded continuous function f on S. The most important property of this notion is the following: For each $n \in \mathbb{N}$, let V_n be a real-valued variable on $(\mathcal{X}_n, \mathcal{A}_n, \mathbb{P}_n)$, and suppose that the sequence V_n converges in probability to a variable V on $(\mathcal{X}, \mathcal{A}, \mathbb{P})$. Then we have $(Z_n, V_n) \to^{d_s} (Z_\infty, V)$ for the product topology on the space $S \times \mathbb{R}$, provided that $Z_n \to^S Z$. We refer to Section 2.2.1 of Jacod and Protter (2012) for more detailed discussions.

We need some ingredients to describe the limiting random variables appearing in the central limit theorems below. Consider an auxiliary space $(\Omega', \mathcal{F}', P')$ supporting a standard normal variable U^0, two sequences $(U_q)_{q \geq 1}$, $(U_q')_{q \geq 1}$ of standard normal variables, and a sequence $(\kappa_q)_{q \geq 1}$ of variables uniformly distributed on $(0, 1)$, all of these being mutually independent. Then we introduce the extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ of $(\Omega, \mathcal{F}, P)$ by putting $\widetilde{\Omega} = \Omega \times \Omega'$, $\widetilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}'$ and $\widetilde{P} = P \times P'$. Now let $(T_q)_{q \geq 1}$ be a sequence of stopping times exhausting the jumps of X. Namely, $\{s \geq 0 : \Delta X_s(\omega) \neq 0\} = \{T_q(\omega) : q \geq 1, T_q(\omega) < \infty\}$ for almost all ω and $T_q \neq T_{q'}$ if $q \neq q'$ and $T_q < \infty$. It is well-known that such a sequence always exists as long as X is càdlàg and adapted; see Proposition I-1.32 of Jacod and Shiryaev (2003). For any C^1 function $g : \mathbb{R} \to \mathbb{R}$ such that $g'(x) = o(|x|)$ as $x \to 0$, we define the random variable $\overline{\mathcal{V}}_T(X, g)$ by

$$\overline{\mathcal{V}}_T(X, g) = \sum_{q:T_q \leq T} g'(\Delta X_{T_q}) R_q,$$

where $R_q = \sqrt{\kappa_q} \sigma_{T_q -} U_q + \sqrt{1 - \kappa_q} \sigma_{T_q} U_q'$. From Proposition 5.1.1 of Jacod and Protter (2012) the variable $\overline{\mathcal{V}}_T(X, g)$ is well-defined and its $\mathcal{F}$-conditional law does

not depend on the choice of the exhausting sequence (T_q). For any integer $r \geq 2$ we also define the random variable $\mathcal{Z}_T(X, r)$ by $\mathcal{Z}_T(X, r) = \overline{\mathcal{V}}_T(X, g_r)$, where the function g_r is defined by $g_r(x) = x^r$.

In order to derive the asymptotic distribution of the realized skewness, we also need to consider the *realized volatility* of X, i.e., $\mathrm{RV}_T^n(X) = \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} (\Delta_i^n X)^2$. Under Assumption [H], the following central limit theorem for $\mathrm{RV}_T^n(X)$ is known (e.g., Theorem 5.4.2 of Jacod and Protter (2012)):

$$\frac{1}{\sqrt{\Delta_n}} (\mathrm{RV}_T^n(X) - [X, X]_T) \to^{\mathcal{S}} \sqrt{2IQ_T} U^0 + \mathcal{Z}_T(X, 2)$$

as $n \to \infty$, where $IQ_T := \int_0^T \sigma_s^4 \mathrm{d}s$ is the so-called *integrated quarticity*. Here our aim is to develop a joint central limit theorem for the bivariate variables $(\mathrm{RV}_T^n(X), \mathcal{V}_T^n(X, g))$. In the following, the variables $\mathcal{Z}_T(X, 2)$ and $\overline{\mathcal{V}}_T(X, g)$ are defined with respect to the same auxiliary sequence R_q.

Theorem 1 *Let g be a real-valued C^2 function on $\mathbb{R}$ satisfying $g(0) = g'(0) = 0$ and $g''(x) = O(|x|)$ as $x \to 0$. Under Assumption [H], the random variables*

$$\frac{1}{\sqrt{\Delta_n}} \left(\mathrm{RV}_T^n(X) - [X, X]_T, \mathcal{V}_T^n(X, g) - \mathcal{V}_T^n(X^c, g) - \sum_{0 \leq s \leq T} g(\Delta X_s) \right)$$

converge stably in law to

$$\left(\sqrt{2IQ_T} U^0 + \mathcal{Z}_T(X, 2), \overline{\mathcal{V}}_T(X, g) \right)$$

as $n \to \infty$, where X^c denotes the continuous martingale part of X, i.e., $X_t^c = \int_0^t \sigma_s \mathrm{d}B_s$.

We prove this result in Sect. 4.1.

Remark 1 (i) If in addition $g''(x) = o(|x|)$ as $x \to 0$, it can easily be seen that $\mathcal{V}_T^n(X^c, g) = o_P(\sqrt{\Delta_n})$ as $n \to \infty$, hence the theorem is a special case of Theorem 5.5.1 from Jacod and Protter (2012) once we note that $[X, X]_{\lfloor T/\Delta_n \rfloor \Delta_n} - [X, X]_T = o_P(\sqrt{\Delta_n})$ under the assumptions of the theorem.

(ii) If the probability limit $\mathcal{V}_T(X^c, g) := P\text{-}\lim \mathcal{V}_T^n(X^c, g)/\sqrt{\Delta_n}$ exists, by using the properties of stable convergence we can deduce a central limit theorem for $\frac{1}{\sqrt{\Delta_n}} (\mathcal{V}_T^n(X, g) - \sum_{0 \leq s \leq T} g(\Delta X_s))$ (in this case $\mathcal{V}_T(X^c, g)$ appears as the $\mathcal{F}$-conditional mean of the limiting variable).

(iii) If g is positively homogeneous, i.e., there exists a constant w such that $g(\alpha x) = \alpha^w g(x)$ for any $\alpha \geq 0$ and any $x \in \mathbb{R}$, the probability limit of $\mathcal{V}_T^n(X^c, g)/\sqrt{\Delta_n}$ can be derived from e.g., Theorem 3.4.1 of Jacod and Protter (2012). In the following, we give two examples of such a case as corollaries.

(iv) In general, the variables $\mathcal{V}_T^n(X^c, g)/\sqrt{\Delta_n}$ may not converge in probability (even in law, indeed); see the next proposition (we prove it in Sect. 4.2).

Proposition 1 *For every* $a \in \mathbb{R}$, *define the function* $g_a : \mathbb{R} \to \mathbb{R}$ *by* $g_a(x) = |x|^3 \sin(2a \log|x|)$.

(a) *For all* $a \in \mathbb{R}$, g_a *is a* C^2 *function and satisfies* $g_a(0) = g'_a(0) = 0$ *and* $g''_a(x) = O(|x|)$ *as* $x \to 0$.

(b) *Suppose that* $\sigma_s = 1$ *for all* $s \in [0, T]$. *There is a real number* $a \neq 0$ *such that the variables* $\mathcal{V}^n_T(X^c, g_a)/\sqrt{\Delta_n}$ *do not converge in law with* $\Delta_n = \exp(-n\pi/a)$.

If we assume $g(x) = |x|^3$, we obtain a generalization of case 2 from Section 1.4.2 of Jacod and Protter (2012) (p. 20 of that book) where X is assumed to be a scaled Brownian motion with a linear drift plus a compound Poisson process to a situation where X is a more general Itô semimartingale:

Corollary 1 *Under Assumption [H]*,

$$\frac{1}{\sqrt{\Delta_n}} \left(\sum_{i=1}^{\lfloor T/\Delta_n \rfloor} |\Delta^n_i X|^3 - \sum_{0 \leq s \leq T} |\Delta X_s|^3 \right) \to^{\mathcal{S}} \frac{2\sqrt{2}}{\sqrt{\pi}} \int_0^T |\sigma_s|^3 \mathrm{d}s$$

$$+ 3 \sum_{q: T_q \leq T} \mathrm{sign}(\Delta X_{T_q})(\Delta X_{T_q})^2 R_q$$

as $n \to \infty$, *where* $\mathrm{sign}(x) = 1$ *if* $x \geq 0$; *otherwise* $\mathrm{sign}(x) = -1$.

If we consider $g(x) = x^3$, the following joint central limit theorem for the realized volatility and the cubic power variation is obtained:

Corollary 2 *Under Assumption [H], the variables*

$$\frac{1}{\sqrt{\Delta_n}} \left(\mathrm{RV}^n_T(X) - [X, X]_T, \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} (\Delta^n_i X)^3 - \sum_{0 \leq s \leq T} (\Delta X_s)^3 \right)$$

converge stably in law to

$$\left(\sqrt{2IQ_T}U^0 + \mathcal{Z}_T(X, 2), \mathcal{Z}_T(X, 3) \right)$$

as $n \to \infty$.

We can use Corollary 2 to derive the asymptotic distribution of the realized skewness: combining Corollary 2 with the delta method for stable convergence [Proposition 2(ii) of Podolskij and Vetter (2010)], we obtain the following result:

Theorem 2 *Under Assumption [H], the variables*

$$\frac{1}{\sqrt{\Delta_n}} \left(\frac{1}{\lfloor T/\Delta_n \rfloor} RDSkew - \frac{\sum_{0 \leq s \leq T}(\Delta X_s)^3}{([X, X]_T)^{3/2}} \right)$$

converge stably in law to

$$G_T := \frac{[X,X]_T^{3/2} \mathcal{Z}_T(X,3) - \frac{3}{2}\sqrt{[X,X]_T} \sum_{0 \leq s \leq T} (\Delta X_s)^3 \left\{ \sqrt{2IQ_T}U^0 + \mathcal{Z}_T(X,2) \right\}}{[X,X]_T^3}$$

as $n \to \infty$.

3 Microstructure noise and stochastic sampling

It is widely recognized that modeling raw high-frequency financial data as *direct* observations of an Itô semimartingale X is unrealistic. One common approach to deal with this issue is to assume that we observe the process X with some measurement errors (referred to as *microstructure noise*) rather than X itself; see Chapter 7 of Aït-Sahalia and Jacod (2014) and references therein. Also, raw high-frequency financial data are typically recorded at stochastic sampling times, so the assumption that we observe data at equidistant sampling times is not applicable. Motivated by these reasons, in this section we consider an observed model which takes account of microstructure noise and stochastic sampling times, and develop an asymptotic theory for estimating Eq. (1) under such a situation.

Let us introduce the precise mathematical description of our model. We denote by $t_0^n, t_1^n, \ldots$ the observation times which are assumed to be $(\mathcal{F}_t)$-stopping times and satisfy $t_i^n \uparrow \infty$ as $i \to \infty$. We also assume that

$$r_n(t) := \sup_{i \geq 0} (t_i^n \wedge t - t_{i-1}^n \wedge t) \to^P 0$$

as $n \to \infty$ for any $t \in \mathbb{R}_+$, with setting $t_{-1}^n = 0$ for notational convenience.

The observed process Y is contaminated by some noise:

$$Y_t = X_t + \epsilon_t.$$

The noise process ϵ implicitly depends on $n \in \mathbb{N}$ and is defined on a very good filtered extension $\mathcal{B}_n = (\Omega_n, \mathcal{F}^n, (\mathcal{F}_t^n)_{t \geq 0}, P_n)$ of $\mathcal{B}$ (see p. 36 of Jacod and Protter (2012) for the definition of very good filtered extensions). ϵ is an $(\mathcal{F}_t^n)$-optional process and, conditionally on $\mathcal{F}$, the sequence $(\epsilon_{t_i^n})_{i=0}^\infty$ is independent and the (conditional) distribution of $\epsilon_{t_i^n}$ is given by $Q_{t_i^n(\omega)}(\omega, \mathrm{d}u)$ for each $i = 0, 1, \ldots$, where for each $t \geq 0$ $Q_t(\omega, \mathrm{d}u)$ denotes a transition probability from $(\Omega, \mathcal{F}_t)$ to $\mathbb{R}$. We assume that the $Q_t(\omega, \mathrm{d}u)$'s satisfy the following condition:

$$\left. \begin{array}{l} \int u \, Q_t(\omega, \mathrm{d}u) = 0 \text{ for every } t \geq 0, \\ \text{the process } (Q_t(\cdot, A))_{t \geq 0} \text{ is } (\mathcal{F}_t)\text{-progressively measurable for any Borel set } A \text{ of } \mathbb{R}. \end{array} \right\} \tag{4}$$

A concrete construction of such a noise process can be found in Section 2 of Koike (2016).

3.1 Construction of estimators

As was pointed out by Liu et al. (2014), the realized skewness RDSkew is no longer a consistent estimator for Eq. (1) in the presence of microstructure noise even after appropriate scaling. Hence we modify the realized skewness by the *pre-averaging procedure*, which is a general scheme to remove the effects of microstructure noise from observation data; see Podolskij and Vetter (2009) and Chapter 16 of Jacod and Protter (2012) for example.

First, we choose a sequence k_n of positive integers and a number $\theta \in (0, \infty)$ such that $k_n = \theta \Delta_n^{-1/2} + o(\Delta_n^{-1/4})$ as $n \to \infty$. Here, unlike the previous section, Δ_n is an auxiliary sequence which is not observable: It typically corresponds to the expected duration between successive observation times. We also choose a continuous function $g : [0, 1] \to \mathbb{R}$ which is piecewise C^1 with a piecewise Lipschitz derivative g' and satisfies $g(0) = g(1) = 0$ and $\int_0^1 g(x)^3 dx \neq 0$. After that, for any process V we define the variables

$$\overline{V}_i = \sum_{p=1}^{k_n-1} g\left(\frac{p}{k_n}\right) \left(V_{t_{i+p}^n} - V_{t_{i+p-1}^n}\right), \quad i = 0, 1, \ldots.$$

In the following, we develop a central limit theorem for the Pre-averaged Realized Volatility

$$\mathrm{PRV}_T^n = \frac{1}{\psi_2 k_n} \sum_{i=0}^{N_T^n - k_n + 1} \left(\overline{Y}_i\right)^2 - \frac{\psi_1}{2\psi_2 k_n} \sum_{i=1}^{N_T^n} (Y_{t_i^n} - Y_{t_{i-1}^n})^2, s$$

and the Pre-averaged realized Cubic power Variation

$$\mathrm{PCV}_T^n = \frac{1}{\psi_3 k_n} \sum_{i=0}^{N_T^n - k_n + 1} \left(\overline{Y}_i\right)^3,$$

where $N_T^n = \max\{i : t_i^n \leq T\}$ and

$$\psi_1 = \int_0^1 g'(x)^2 dx, \qquad \psi_2 = \int_0^1 g(x)^2 dx, \qquad \psi_3 = \int_0^1 g(x)^3 dx.$$

For the equidistant sampling case $t_i^n = i\Delta_n$, it is known that

$$\mathrm{PRV}_T^n \to^P [X, X]_T, \qquad \mathrm{PCV}_T^n \to^P \sum_{0 \leq s \leq T} (\Delta X_s)^3$$

as $n \to \infty$ from Theorems 16.2.1 and 16.6.1 of Jacod and Protter (2012). Therefore, we may expect that

$$\frac{\text{PCV}_T^n}{\left(\text{PRV}_T^n\right)^{3/2}}$$

is a consistent estimator for Eq. (1). Our aim is to derive the asymptotic distribution of the above statistic.

3.2 Asymptotic results

3.2.1 Notation

We write $X^n \xrightarrow{ucp} X$ for processes X^n and X to express shortly that $\sup_{0 \le t \le T} |X_t^n - X_t| \to^P 0$. ϖ denotes some (fixed) positive constant. We denote by $(\mathcal{G}_t)$ the smallest filtration containing $(\mathcal{F}_t)$ such that $\mathcal{G}_0$ contains the σ-field generated by μ, i.e., the σ-field generated by all the variables $\mu(A)$, where A ranges all measurable subsets of $\mathbb{R}_+ \times E$.

For any real-valued bounded measurable functions u, v on $[0, 1]$, we define the function $\phi_{u,v}$ on $[0, 1]$ by

$$\phi_{u,v}(y) = \int_y^1 u(x - y)v(x)\mathrm{d}x.$$

Then we put

$$\Phi_{22} = \int_0^1 \phi_{g,g}(y)^2\mathrm{d}y, \qquad \Phi_{12} = \int_0^1 \phi_{g,g}(y)\phi_{g',g'}(y)\mathrm{d}y, \qquad \Phi_{11} = \int_0^1 \phi_{g',g'}(y)^2\mathrm{d}y$$

and

$$\Phi_{3+} = \int_0^1 \phi_{g,g^2}(y)^2\mathrm{d}y, \qquad \Phi_{3-} = \int_0^1 \phi_{g^2,g}(y)\mathrm{d}y,$$

$$\Phi'_{3+} = \int_0^1 \phi_{g',g^2}(y)^2\mathrm{d}y, \qquad \Phi'_{3-} = \int_0^1 \phi_{g^2,g'}(y)^2\mathrm{d}y,$$

$$\Phi_{23+} = \int_0^1 \phi_{g,g}(y)\phi_{g,g^2}(y)\mathrm{d}y, \qquad \Phi_{23-} = \int_0^1 \phi_{g,g}(y)\phi_{g^2,g}(y)\mathrm{d}y,$$

$$\Phi'_{23+} = \int_0^1 \phi_{g',g}(y)\phi_{g',g^2}(y)\mathrm{d}y, \qquad \Phi'_{23-} = \int_0^1 \phi_{g,g'}(y)\phi_{g^2,g'}(y)\mathrm{d}y.$$

We define the process α by $\alpha(\omega)_t = \int u^2 Q_t(\omega, \mathrm{d}u)$.

3.2.2 Assumptions

We enumerate the assumptions which are imposed to derive our limit theorem.

[A1] It holds that

$$r_n(t) = o_p(\Delta_n^\xi) \tag{5}$$

as $n \to \infty$ (note that $t_{-1}^n = 0$ by convention) for every $t > 0$ and every $\xi \in (0, 1)$. Moreover, for each n we have a $(\mathcal{G}_t)$-progressively measurable positive-valued process G_t^n and a random subset $\mathcal{N}^n$ of $\mathbb{Z}_+$ satisfying the following conditions:

(i) $\{(\omega, p) \in \Omega \times \mathbb{Z}_+ : p \in \mathcal{N}^n(\omega)\}$ is a measurable set of $\Omega \times \mathbb{Z}_+$. Moreover, there is a constant $\kappa \in (0, \frac{1}{2})$ such that $\#(\mathcal{N}^n \cap \{p : t_p^n \le t\}) = O_p(\Delta_n^{-\kappa})$ as $n \to \infty$ for every $t > 0$.

(ii) $E[\Delta_n^{-1}(t_{p+1}^n - t_p^n)|\mathcal{G}_{t_p^n}] = G_{t_p^n}^n$ for every n and every $p \in \mathbb{Z}_+ \setminus \mathcal{N}^n$.

(iii) There is a càdlàg $(\mathcal{F}_t)$-adapted positive-valued process G such that

(iii-a) $\Delta_n^{-\varpi}(G^n - G) \xrightarrow{ucp} 0$,

(iii-b) $G_{t-} > 0$ for every $t > 0$,

(iii-c) G is an Itô semimartingale of the form

$$G_t = G_0 + \int_0^t \widehat{b}_s \mathrm{d}s + \int_0^t \widehat{\sigma}_s \mathrm{d}W_s$$
$$+ \int_0^t \widehat{\sigma}_s' \mathrm{d}\widehat{W}_s$$
$$+ \int_0^t \int_{|\widehat{\delta}(s,z)| \le 1} \widehat{\delta}(s, z)(\mu - \nu)(\mathrm{d}s, \mathrm{d}z)$$
$$+ \int_0^t \int_{|\widehat{\delta}(s,z)| > 1} \widehat{\delta}(s, z)\mu(\mathrm{d}s, \mathrm{d}z),$$

where $\widehat{b}_s$ is a locally bounded and $(\mathcal{F}_t)$-progressively measurable real-valued process, $\widehat{\sigma}_s$ and $\widehat{\sigma}_s'$ are càdlàg $(\mathcal{F}_t)$-adapted processes, $\widehat{W}_s$ is an $(\mathcal{F}_t)$-standard Wiener process independent of W, and $\widehat{\delta}$ is an $(\mathcal{F}_t)$-predictable real-valued function on $\Omega \times \mathbb{R}_+ \times E$ such that there is a sequence $(\widehat{\rho}_j)$ of $(\mathcal{F}_t)$-stopping times increasing to infinity and, for each j, a deterministic nonnegative function $\widehat{\gamma}_j$ on E satisfying $\int \widehat{\gamma}_j(z)^2 \wedge 1\lambda(\mathrm{d}z) < \infty$ and $|\widehat{\delta}(\omega, t, z)| \le \widehat{\gamma}_j(z)$ for all (ω, t, z) with $t \le \widehat{\rho}_j(\omega)$.

Furthermore, $\max_{p=0,1,\ldots,N_T^n} E\left[\Delta_n^{-1}(t_{p+1}^n - t_p^n)|\mathcal{F}_{t_p^n}\right]$ is tight as $n \to \infty$ for every $t > 0$.

[A2] The volatility process σ is an Itô semimartingale of the form

$$\sigma_t = \sigma_0 + \int_0^t \widetilde{b}_s \mathrm{d}s + \int_0^t \widetilde{\sigma}_s \mathrm{d}W_s + \int_0^t \widetilde{\sigma}_s' \mathrm{d}\widetilde{W}_s$$
$$+ \int_0^t \int_{|\widetilde{\delta}(s,z)| \le 1} \widetilde{\delta}(s, z)(\mu - \nu)(\mathrm{d}s, \mathrm{d}z)$$
$$+ \int_0^t \int_{|\widetilde{\delta}(s,z)| > 1} \widetilde{\delta}(s, z)\mu(\mathrm{d}s, \mathrm{d}z)$$

where $\widetilde{b}_s$ is a locally bounded and $(\mathcal{F}_t)$-progressively measurable process, $\widetilde{\sigma}_s$ and $\widetilde{\sigma}'_s$ are càdlàg $(\mathcal{F}_t)$-adapted processes, $\widetilde{W}_s$ is an $(\mathcal{F}_t)$-standard Wiener process independent of W, and $\widetilde{\delta}$ is an $(\mathcal{F}_t)$-predictable function on $\Omega \times \mathbb{R}_+ \times E$. Moreover, for each j there is an $(\mathcal{F}_t)$-stopping time ρ_j, a bounded $(\mathcal{F}_t)$-progressively measurable process $b(j)_s$, a deterministic nonnegative function γ_j on E, and a constant Λ_j such that $\rho_j \uparrow \infty$ as $j \to \infty$ and, for each j,

(i) $b(\omega)_s = b(j)(\omega)_s$ if $s < \rho_j(\omega)$,
(ii) $E\left[|b(j)_{t_1} - b(j)_{t_2}|^2 | \mathcal{F}_{t_1 \wedge t_2}\right] \leq \Lambda_j E\left[|t_1 - t_2|^\varpi | \mathcal{F}_{t_1 \wedge t_2}\right]$ for any $(\mathcal{F}_t)$-stopping times t_1 and t_2 bounded by j,
(iii) $\int \left\{\gamma_j(z)^2 \wedge 1\right\} \lambda(\mathrm{d}z) < \infty$ and $|\delta(\omega, t, z)| \vee |\widetilde{\delta}(\omega, t, z)| \leq \gamma_j(z)$ for all (ω, t, z) with $t \leq \rho_j(\omega)$,
(iv) $E\left[|\delta(t_1 \wedge \rho_j, z) - \delta(t_2 \wedge \rho_j, z)|^2 | \mathcal{F}_{t_1 \wedge t_2}\right] \leq \Lambda_j \gamma_j(z)^2 E\left[|t_1 - t_2|^\varpi | \mathcal{F}_{t_1 \wedge t_2}\right]$ for any $(\mathcal{F}_t)$-stopping times t_1 and t_2 bounded by j.

[A3] Q_t's satisfy Eq. (4) and there is a sequence $(\rho'_j)_{j \geq 1}$ of $(\mathcal{F}_t)$-stopping times increasing to infinity such that

$$\sup_{\omega \in \Omega, t < \rho'_j(\omega)} \int u^6 Q_t(\omega, \mathrm{d}u) < \infty.$$

Moreover, for each j there is a bounded càdlàg $(\mathcal{F}_t)$-adapted process $\alpha(j)_t$ and a constant Λ'_j such that

(i) $\alpha(j)(\omega)_t = \alpha(\omega)_t$ if $t < \rho'_j(\omega)$,
(ii) $E\left[|\alpha(j)_{t_1} - \alpha(j)_{t_2}|^2 | \mathcal{F}_{t_1 \wedge t_2}\right] \leq \Lambda'_j E\left[|t_1 - t_2|^\varpi | \mathcal{F}_{t_1 \wedge t_2}\right]$ for any $(\mathcal{F}_t)$-stopping times t_1 and t_2 bounded by j.

[A4] A regular conditional probability of P given $\mathcal{H}$ exists for any sub-σ-field $\mathcal{H}$ of $\mathcal{F}$.

Remark 2 Assumptions [A1]–[A4] are almost identical to the ones imposed to prove a central limit theorem for PRV_T^n in Koike (2016). A few differences appear in [A1] and [A3]: We add an additional (mild) assumption on the tightness of random variables

$$\max_{p=0,1,\ldots,N_T^n} E\left[\Delta_n^{-1}(t_{p+1}^n - t_p^n) | \mathcal{F}_{t_p^n}\right].$$

to [A1] (note that this assumption automatically follows from [A1](ii)–(iii) if $\mathcal{N}^n = \emptyset$). In the meantime, [A3] requires the finiteness of the sixth moment of the noise process.

Remark 3 Since Assumption [A1] contains some non-standard aspects compared with common ones used in the literature, we briefly make some comments on it (see Remark 3.2 of Koike (2016) for more details).

1. A typical example satisfying this assumption is the *restricted discretization scheme* introduced in Chapter 14 of Jacod and Protter (2012), where the t_p^n's are modeled as

$$t_p^n = t_{p-1}^n + \theta_{t_{p-1}^n}^n \varepsilon_p^n, \qquad p = 1, 2, \ldots,$$

with θ^n being a càdlàg $(\mathcal{F}_t)$-adapted process, $(\varepsilon_p^n)_{p \geq 1}$ being a sequence of i.i.d. positive variables independent of b, σ, δ, W, μ, and such that $E[\varepsilon_p^n] = 1$ and $E[(\varepsilon_p^n)^r] < \infty$ for every $r > 0$, and $t_0^n = 0$. An appropriate construction of the filtration $(\mathcal{F}_t)$ allows us to assume the independence between ϵ_p^n and $\mathcal{F}_{t_{p-1}^n}$ for all n, p. In this case we have [A1](i)–(ii) by setting $\mathcal{N}^n = \emptyset$ and $G^n = \Delta_n^{-1}\theta^n$. Then, [A1](iii) corresponds to (a weaker version of) Assumption (E) of Jacod and Protter (2012), and Eq. (5) follows from Lemma 14.1.5 of Jacod and Protter (2012) (note that the last condition on tightness is automatically satisfied once [A1](iii) holds true since $\mathcal{N}^n = \emptyset$).

2. The main reason why we introduce an involved assumption compared with the standard ones is that our assumption does not rule out the dependence between ε_p^n's and X (see Example 4.1 of Koike (2016) for instance). However, we remark that [A1] rules out some kind of dependence between the observation times and the jumps of the observed process because we take the conditional expectation given $\mathcal{G}_{t_p^n}$'s instead of $\mathcal{F}_{t_p^n}$'s in [A1](ii). In fact, our assumption does not allow the case that t_p^n's are given by hitting times of a pure-jump Lévy process whose jump measure is μ (note that in this case the Lévy process must have infinite activity jumps due to Eq. (5)). We, however, note that our assumption does not exclude such a dependence completely; see Example 4.2 of Koike (2016) for instance.

3. The set $\mathcal{N}^n$ can be interpreted as a set of exceptional indices p for which the equation $E[\Delta_n^{-1}(t_{p+1}^n - t_p^n)|\mathcal{G}_{t_p^n}] = G_{t_p^n}^n$ does not hold true. This additional complexity is useful to ensure the stability of Assumption [A1] under the localization used in the proof (see Lemma 6.1 of Koike (2016). Non-empty $\mathcal{N}^n$ also excludes some trivial exceptions of [A1] with $\mathcal{N}^n = \emptyset$, such as $t_0^n = \log n / n$ and $t_p^n = t_{p-1}^n + \Delta_n$ for $p \geq 1$.

3.2.3 Results

Theorem 3 *Suppose that [A1]–[A4] are satisfied. Then*

$$\left(\Delta_n^{-1/4} \left(\mathrm{PRV}_T^n - [X, X]_T \right), \Delta_n^{-1/4} \left(\mathrm{PCV}_T^n - \sum_{0 \leq s \leq T} (\Delta X_s)^3 \right) \right)^* \to^{\mathcal{S}} \Gamma_T^{1/2} \zeta$$

as $n \to \infty$, where ζ is a bivariate standard normal variable which is defined on an extension of $\mathcal{B}$ and independent of $\mathcal{F}$, and Γ_T is the $\mathbb{R}^2 \otimes \mathbb{R}^2$-valued variable given by

$$\Gamma_T = \begin{bmatrix} \Gamma_T^c + \overline{\Gamma}_T^{11} & \overline{\Gamma}_T^{12} \\ \overline{\Gamma}_T^{12} & \overline{\Gamma}_T^{22} \end{bmatrix}$$

with

$$\Gamma_T^c = \frac{4}{\psi_2^2} \int_0^T \left[\Phi_{22}\theta\sigma_s^4 G_s + 2\frac{\Phi_{12}}{\theta}\sigma_s^2\alpha_s + \frac{\Phi_{11}}{\theta^3}\frac{\alpha_s^2}{G_s} \right] ds,$$

$$\overline{\Gamma}_T^{11} = \frac{4}{\psi_2^2} \sum_{0 \le s \le T} (\Delta X_s)^2 \left\{ \Phi_{22} \theta \left(\sigma_s^2 G_s + \sigma_{s-}^2 G_{s-} \right) + \frac{\Phi_{12}}{\theta} (\alpha_s + \alpha_{s-}) \right\},$$

$$\overline{\Gamma}_T^{12} = \frac{6}{\psi_2 \psi_3} \sum_{0 \le s \le T} (\Delta X_s)^3 \left\{ \theta \left(\Phi_{23+} \sigma_s^2 G_s + \Phi_{23-} \sigma_{s-}^2 G_{s-} \right) \right.$$

$$\left. + \theta^{-1} \left(\Phi'_{23+} \alpha_s + \Phi'_{23-} \alpha_{s-} \right) \right\},$$

$$\overline{\Gamma}_T^{22} = \frac{9}{\psi_3^2} \sum_{0 \le s \le T} (\Delta X_s)^4 \left\{ \theta \left(\Phi_{3+} \sigma_s^2 G_s + \Phi_{3-} \sigma_{s-}^2 G_{s-} \right) + \theta^{-1} \left(\Phi'_{3+} \alpha_s + \Phi'_{3-} \alpha_{s-} \right) \right\}.$$

A proof of the above result is given in Sect. 4.3. Combining the above result with the delta method for stable convergence, we obtain the asymptotic distribution of $\mathrm{PCV}_T^n / \left(\mathrm{PRV}_T^n \right)^{3/2}$ as follows:

Theorem 4 *Under the assumptions of Theorem 3, we have*

$$\Delta_n^{-1/4} \left(\frac{\mathrm{PCV}_T^n}{\left(\mathrm{PRV}_T^n \right)^{3/2}} - \frac{\sum_{0 \le s \le T} (\Delta X_s)^3}{([X, X]_T)^{3/2}} \right) \to^{\mathcal{S}} \sqrt{d_{1,T}^2 (\Gamma_T^c + \overline{\Gamma}_T^{11}) + 2 d_{1,T} d_{2,T} \overline{\Gamma}_T^{12} + d_{2,T}^2 \overline{\Gamma}_T^{22}} \times \zeta$$

as $n \to \infty$, *where* ζ *is a standard normal variable which is defined on an extension of* $\mathcal{B}$ *and independent of* $\mathcal{F}$, *and*

$$d_{1,T} = -\frac{3}{2} \frac{\sum_{0 \le s \le T} (\Delta X_s)^3}{([X, X]_T)^{5/2}}, \qquad d_{2,T} = \frac{1}{([X, X]_T)^{3/2}}.$$

Remark 4 The stable convergence result on the first component $\Delta_n^{-1/4} \left(\mathrm{PRV}_T^n - [X, X]_T \right)$ in Theorem 3 is a special case of Theorem 3.1 from Koike (2016). In contrast, the stable convergence result on the second component

$$\Delta_n^{-1/4} \left(\mathrm{PCV}_T^n - \sum_{0 \le s \le T} (\Delta X_s)^3 \right)$$

in Theorem 3 is new even in the equidistant sampling case $t_i^n = i \Delta_n$. As is remarked in Introduction, Theorem 16.3.1 of Jacod and Protter (2012) deals with the asymptotic distribution of the statistic

$$\frac{1}{k_n} \sum_{i=0}^{N_T^n - k_n + 1} f(\overline{Y}_i)$$

in the equidistant sampling case for a C^2 function $f : \mathbb{R} \to \mathbb{R}$ which is a linear combinations of positively homogeneous functions with degree (strictly) bigger than 3. Since the cubic function $f(x) = x^3$ does not have this property, Theorem 16.3.1 of Jacod and Protter (2012) is not applicable to deriving the asymptotic distribution of PCV_T^n.

Remark 5 The main difference between the asymptotic distributions given in Theorems 1–2 and Theorems 3–4 is that the former ones are in general not $\mathcal{F}$-conditionally Gaussian due to the additional randomness caused by the uniform variables U_q, U_q', while the latter ones are always $\mathcal{F}$-conditionally Gaussian. This is a byproduct of the pre-averaging procedure and commonly observed in the literature of pre-averaging estimators for functionals of jumps.

Remark 6 If X is continuous, Theorem 3 implies $\Delta_n^{-1/4} \mathrm{PCV}_T^n \to^P 0$ as $n \to \infty$, hence we will need a larger scaling factor than $\Delta_n^{-1/4}$ to obtain a non-degenerate asymptotic distribution of PCV_T^n. To the best of our knowledge, nothing is known about the non-trivial asymptotic distribution of PCV_T^n even in the equidistant sampling setting. An analogy to the non-noisy case suggests that the proper scaling factor is $\Delta_n^{-1/2}$ and $\Delta_n^{-1/2} \mathrm{PCV}_T^n$ would converge stably in law to a mixed normal distribution with a nonzero conditional mean (cf. Example 6 of Kinnebrock and Podolskij (2008) for the equidistant sampling setting and Theorem 2 of Li et al. (2014) for an irregular sampling setting).

Remark 7 We note that the main reason why the asymptotic distribution of the realized skewness obtained in the previous section is centered is because we have $\sqrt{\Delta_n} \sum_{i=1}^{N_T^n} (X_{t_i^n}^c - X_{t_{i-1}^n}^c)^3 \to^P 0$ as $n \to \infty$ *as long as* $t_i^n = i\Delta_n$. This is in general not true when we consider more general sampling schemes as (t_i^n) such that they depend on the process X^c. In particular, Assumption [A1] does not rule out situations where the variables $\sqrt{\Delta_n} \sum_{i=1}^{N_T^n} (X_{t_i^n}^c - X_{t_{i-1}^n}^c)^3$ converge in probability to some nonzero random variable as $n \to \infty$; see e.g., Example 3.2 from Koike (2017) and Example 5 from Li et al. (2014). In such a situation, we conjecture that the asymptotic distribution of the realized skewness estimator would be no longer centered. In contrast, Theorem 4 tells us that the estimator $\mathrm{PCV}_T^n / (\mathrm{PRV}_T^n)^{3/2}$ is asymptotically centered even under [A1]. This is another byproduct of the pre-averaging procedure.

Remark 8 In the absence of noise, the simultaneous presence of jumps and the randomness of observation times typically adds more complexity to the asymptotic distribution of statistics of the form Eq. (3) more complex, as seen in Bibinger and Vetter (2015), Vetter and Zwingmann (2017) and Martin and Vetter (2016). In this sense, the result of Theorem 3 again contrasts with non-noisy cases because the estimators PRV_T^n and PCV_T^n are asymptotically mixed normal even with stochastic sampling times. This is also a byproduct of the pre-averaging procedure.

4 Proofs

4.1 Proof of Theorem 1

First of all, a standard localization procedure, described in detail in Lemma 4.49 of Jacod and Protter (2012), for instance, allows us to assume that there are a positive constant A and a nonnegative deterministic measurable function γ on E such that

$$\left.\begin{array}{ll} |b(\omega)_t| \leq A, \quad |\sigma(\omega)_t| \leq A, \quad |X(\omega)_t| \leq A, \quad |X^c(\omega)_t| \leq A, \\ |\delta(\omega, t, z)| \leq \gamma(z) \leq A, \quad \int \gamma(z)^2 \lambda(dz) \leq A. \end{array}\right\} \quad (6)$$

The strategy of the proof is the same as the one used in the proof of Theorem 5.1.2 from Jacod and Protter (2012), and we divide the proof into several steps. For the part corresponding to Steps 1–3 of Jacod and Protter (2012)'s proof, we can adopt almost the same argument as the original one, hence it is just briefly sketched in Step 2. In the remainder steps, we will need an argument which is somewhat different from theirs.

Throughout the discussions, for random variables X and Y which may depend on the parameters n, m, i, $X \lesssim Y$ means that there exists a (non-random) constant $K > 0$ independent of n, m, i such that $X \leq KY$ a.s.

Step 1) We begin with introducing some notations. For each $m \in \mathbb{N}$, set $A_m = \{z : \gamma(z) > 1/m\}$. Noting that $\nu(A_m) < \infty$, we denote by $(S(m, j))_{j \geq 1}$ the successive jump times of the Poisson process $(\mu((0, t] \times (A_m \setminus A_{m-1})))_{t \geq 0}$. Let $(S_p)_{p \geq 1}$ be a reordering of the double sequence $(S(m, j))$, and we denote by $\mathcal{P}_m$ the set of all indices p such that $S_p = S(m', j)$ for some $j \geq 1$ and some $m' \leq m$. In the light of Proposition 5.1.1 from Jacod and Protter (2012), we may assume that $(S_p) = (T_q)$ without loss of generality.

Set

$$b(m)_t = b_t - \int_{A_m \cap \{z : |\delta(t,z)| \leq 1\}} \delta(t, z) \lambda(dz), \quad C(m)_t = X_0 + \int_0^t b(m)_s ds + X_t^c,$$

$$X(m)_t = C(m)_t + \int_0^t \int_{A_m^c} \delta(s, z)(\mu - \nu)(ds, dz),$$

$$X'(m)_t = X_t - X(m)_t = \int_0^t \int_{A_m} \delta(s, z) \mu(ds, dz), \quad (7)$$

$$X''(m)_t = X(m)_t - X_t^c = X_0 + \int_0^t b(m)_s ds + \int_0^t \int_{A_m^c} \delta(s, z)(\mu - \nu)(ds, dz),$$

$$Y(m)_t = X_t - \int_0^t \int_{A_m^c} \delta(s, z)(\mu - \nu)(ds, dz) = C(m)_t + X'(m)_t,$$

and denote by $\Omega_n(T, m)$ the set of all ω such that each interval $[0, T] \cap ((i-1)\Delta_n, i\Delta_n]$ contains at most one jump of $X'(m)(\omega)$. Note that the notations here are consistent with those from Eq.(5.1.10) of Jacod and Protter (2012). Since $X'(m)$ has finitely many jumps on $[0, T]$, it holds that $P(\Omega_n(T, m)) \to 1$ as $n \to \infty$. We also set $i_p^n = \lceil S_p/\Delta_n \rceil$ so that S_p is in $((i_p^n - 1)\Delta_n, i_p^n \Delta_n]$. Here, for a real number x, $\lceil x \rceil$ denotes the minimum integer l satisfying $l \geq x$. Then we define

$$R(n, p) = \frac{1}{\sqrt{\Delta_n}} \left(\Delta_{i_p^n}^n X - \Delta X_{S_p} \right),$$

$$\zeta_p^n = \frac{1}{\sqrt{\Delta_n}} \left(g(\Delta_{i_p^n}^n X) - g(\Delta X_{S_p}) - g(\sqrt{\Delta_n} R(n, p)) \right)$$

and $Y_T^n(m) = \sum_{p \in \mathcal{P}_m : S_p \leq \Delta_n \lfloor T/\Delta_n \rfloor} \zeta_p^n$. Moreover, for any semimartingale S, we set

$$\overline{\mathcal{V}}_T^n(S, g) = \frac{1}{\sqrt{\Delta_n}} \left(\mathcal{V}_T^n(S, g) - \sum_{0 \leq s \leq T} g(\Delta S_s) \right),$$

$$\overline{Z}_T^n(S) = \frac{1}{\sqrt{\Delta_n}} \left(\mathrm{RV}_T^n(S) - [S, S]_{\lfloor T/\Delta_n \rfloor \Delta_n} \right).$$

Step 2) First we fix m. With p fixed, the sequence $R(n, p)$ is tight due to Proposition 4.4.10 of Jacod and Protter (2012). Therefore, we have $g(\sqrt{\Delta_n} R(n, p))/\sqrt{\Delta_n} \to^P 0$ as $n \to \infty$ because $g(x) = O(x^2)$ as $x \to 0$. On the other hand, repeated applications of the fundamental theorem of calculus yield

$$g(\Delta_{i_p}^n X) - g(\Delta X_{S_p}) = g'(\Delta X_{S_p}) \sqrt{\Delta_n} R(n, p)$$
$$+ \int_0^{\sqrt{\Delta_n} R(n,p)} \int_0^u g''(\Delta X_{S_p} + v) \mathrm{d}v \mathrm{d}u.$$

Since X is bounded and g'' is continuous, we have

$$\frac{1}{\sqrt{\Delta_n}} \int_0^{\sqrt{\Delta_n} R(n,p)} \int_0^u g''(\Delta X_{S_p} + v) \mathrm{d}v \mathrm{d}u \to^P 0$$

as $n \to \infty$. Consequently, we conclude that $\zeta_p^n - g'(\Delta X_{S_p}) R(n, p) \to^P 0$ as $n \to \infty$. On the other hand, an argument similar to the proof of Lemma 5.4.10 from Jacod and Protter (2012) implies that $\left(\overline{Z}_T^n(C(m)), (R(n, p))_{p \geq 1} \right) \to^S \left(\sqrt{2IQ_T} U^0, (R_p)_{p \geq 1} \right)$ as $n \to \infty$. Hence it holds that

$$\left(\overline{Z}_T^n(C(m)), \left(\Delta X_{S_p} R(n, p) \right)_{p \geq 1}, \left(\zeta_p^n \right)_{p \geq 1} \right) \to^S \left(\sqrt{2IQ_T} U^0, \left(\Delta X_{S_p} R_p \right)_{p \geq 1}, \left(\zeta_p \right)_{p \geq 1} \right)$$

as $n \to \infty$, where $\zeta_p = g'(\Delta X_{S_p}) R_p$. Since the set $\{S_p : p \in \mathcal{P}_m\} \cap [0, T]$ is finite, it follows that

$$\left(\overline{Z}_T^n(C(m)), \mathcal{Z}_T^n(m), Y^n(m)_T \right) \to^S \left(\sqrt{2IQ_T} U^0, \mathcal{Z}_T(X'(m), 2), \overline{\mathcal{V}}_T(X'(m), g) \right)$$

as $n \to \infty$, where $\mathcal{Z}_T^n(m)_T = 2 \sum_{p \in \mathcal{P}_m : S_p \leq \Delta_n \lfloor T/\Delta_n \rfloor} \Delta X_{S_p} R(n, p)$. Furthermore, the same argument as the last part of the proof of Lemma 5.4.10 from Jacod and Protter (2012) implies that $\overline{Z}_T^n(C(m)) + \mathcal{Z}_T^n(m)_T - \overline{Z}_T^n(Y(m)) \to^P 0$ as $n \to \infty$. Consequently, we conclude that

$$\left(\overline{Z}_T^n(Y(m)), Y_T^n(m) \right) \to^S \left(\sqrt{2IQ_T} U^0 + \mathcal{Z}_T(X'(m), 2), \overline{\mathcal{V}}_T(X'(m), g) \right)$$

as $n \to \infty$.

Next we vary m. We can prove

$$\overline{\mathcal{V}}_T(X'(m), g) \to^P \overline{\mathcal{V}}_T(X, g)$$

as $m \to \infty$ by the same argument as the proof of Eq.(5.1.16) from Jacod and Protter (2012) and

$$\left.\begin{array}{l} \lim_{m\to\infty} \limsup_{n\to\infty} P\left(\left|\overline{Z}_T^n(Y(m)) - \overline{Z}_T^n(X)\right| > \eta\right) \to 0, \\ \mathcal{Z}_T(X'(m), 2) \to^P \mathcal{Z}_T(X, 2) \quad \text{as} \quad m \to \infty \end{array}\right\}$$

for any $\eta > 0$ by the same argument as in the proof of Lemma 5.4.12 from Jacod and Protter (2012) where our $Y(m)$ is denoted by $X(m)$.

Now, noting that $[X, X]_{\lfloor T/\Delta_n \rfloor \Delta_n} - [X, X]_T = o_P(\sqrt{\Delta_n})$ and the decomposition

$$\overline{\mathcal{V}}_T^n(X, g) = \overline{\mathcal{V}}_T^n(X(m), g) + Y_T^n(m)$$

holding on the set $\Omega_n(T, m)$ as well as $\lim_{n\to\infty} P(\Omega_n(T, m)) = 1$ for every $m \in \mathbb{N}$, the proof of the theorem is completed once we show that

$$\lim_{m\to\infty} \limsup_{n\to\infty} P\left(\Omega_n(T, m) \cap \left\{\left|\overline{\mathcal{V}}_T^n(X(m), g) - \frac{1}{\sqrt{\Delta_n}}\mathcal{V}_T^n(X^c, g)\right| > \eta\right\}\right) = 0 \tag{8}$$

for any $\eta > 0$.

Step 3) We begin by showing three inequalities used in the proof. The first and the second ones are elementary: if $\rho \in (0, 2]$, the Lyapunov and Doob inequalities as well as (6) yield

$$E\left[\sup_{(i-1)\Delta_n \leq t \leq i\Delta_n} \left|\int_{(i-1)\Delta_n}^t \int_{A_m^c} \delta(s, z)(\mu - \nu)(ds, dz)\right|^\rho\right] \leq \left(4\Delta_n \overline{\gamma}_m\right)^{\rho/2},$$

where $\overline{\gamma}_m := \int_{A_m^c} \gamma(z)^2 \lambda(dz)$. Therefore, noting that $\left|\int_{A_m \cap \{z \mid |\delta(t,z)| \leq 1\}} \delta(t, z)\nu(dz)\right| \leq Am$, there exists a positive constant K_ρ such that

$$E\left[\sup_{(i-1)\Delta_n \leq t \leq i\Delta_n} |X''(m)_t - X''(m)_{(i-1)\Delta_n}|^\rho\right] \leq K_\rho \left\{(m\Delta_n)^\rho + \left(\Delta_n \overline{\gamma}_m\right)^{\rho/2}\right\} \tag{9}$$

for every i, n, m. On the other hand, for every $\rho \geq 1$ there exists a constant K_ρ' such that

$$E\left[\sup_{(i-1)\Delta_n \leq t \leq i\Delta_n} \left|X_t^c - X_{(i-1)\Delta_n}^c\right|^\rho\right] \leq K_\rho' \Delta_n^{\rho/2} \tag{10}$$

for every i, n, due to the Burkholder–Davis–Gundy inequality and (6).

Now we prove the third one. By using integration by parts repeatedly we obtain

$$
\begin{aligned}
\Delta_i^n X^c \left(\Delta_i^n X''(m)\right)^2 = {} & 2 \int_{(i-1)\Delta_n}^{i\Delta_n} (X_s^c - X_{(i-1)\Delta_n}^c)(X''(m)_{s-} - X''(m)_{(i-1)\Delta_n}) dX''(m)_s \\
& + \int_{(i-1)\Delta_n}^{i\Delta_n} \int_{A_m^c} (X_s^c - X_{(i-1)\Delta_n}^c)\delta(s,z)^2 \mu(ds,dz) \\
& + \int_{(i-1)\Delta_n}^{i\Delta_n} (X''(m)_{s-} - X''(m)_{(i-1)\Delta_n})^2 dX_s^c \\
=: {} & 2\mathbb{I}_i^n + \mathbb{II}_i^n + \mathbb{III}_i^n .
\end{aligned}
$$

First consider $\mathbb{I}_i^n$. We decompose it as

$$
\begin{aligned}
\mathbb{I}_i^n = {} & \int_{(i-1)\Delta_n}^{i\Delta_n} (X_s^c - X_{(i-1)\Delta_n}^c)(X''(m)_{s-} - X''(m)_{(i-1)\Delta_n}) b(m)_s \, ds \\
& + \int_{(i-1)\Delta_n}^{i\Delta_n} \int_{A_m^c} (X_s^c - X_{(i-1)\Delta_n}^c)(X''(m)_{s-} - X''(m)_{(i-1)\Delta_n})\delta(s,z)(\mu - \nu)(ds,dz) \\
=: {} & \mathbb{I}_i^{n,1} + \mathbb{I}_i^{n,2} .
\end{aligned}
$$

The Schwarz inequality and (9)–(10) yield

$$
\begin{aligned}
E\left[\left|\mathbb{I}_i^{n,1}\right|\right] \leq {} & Cm \int_{(i-1)\Delta_n}^{i\Delta_n} E\left[\left|(X_s^c - X_{(i-1)\Delta_n}^c)(X''(m)_{s-} - X''(m)_{(i-1)\Delta_n})\right|\right] ds \\
\lesssim {} & m\Delta_n^{3/2}\left(m\Delta_n + \sqrt{\overline{\gamma}_m \Delta_n}\right).
\end{aligned}
$$

On the other hand, since integration by parts implies that

$$
\begin{aligned}
& (X_s^c - X_{(i-1)\Delta_n}^c)(X''(m)_{s-} - X''(m)_{(i-1)\Delta_n}) \\
& = \int_{(i-1)\Delta_n}^{s-} (X_u^c - X_{(i-1)\Delta_n}^c) dX''(m)_u + \int_{(i-1)\Delta_n}^{s} (X''(m)_{u-} - X''(m)_{(i-1)\Delta_n}) dX_u^c,
\end{aligned}
$$

the Doob inequality, (6) and (9)–(10) imply that

$$
E\left[\sup_{(i-1)\Delta_n \leq s \leq i\Delta_n} \left|(X_s^c - X_{(i-1)\Delta_n}^c)(X''(m)_s - X''(m)_{(i-1)\Delta_n})\right|^2\right] \lesssim \Delta_n^2 \left(m^2 \Delta_n + \overline{\gamma}_m\right),
$$

hence the Lyapunov and Doob inequalities yield

$$
\begin{aligned}
E\left[\left|\mathbb{I}_i^{n,2}\right|\right] \leq {} & \left\{4\overline{\gamma}_m \int_{(i-1)\Delta_n}^{i\Delta_n} E\left[\left|(X_s^c - X_{(i-1)\Delta_n}^c)(X''(m)_{s-} - X''(m)_{(i-1)\Delta_n})\right|^2\right] ds\right\}^{1/2} \\
\lesssim {} & \Delta_n^{3/2}\left(m\sqrt{\overline{\gamma}_m \Delta_n} + \overline{\gamma}_m\right).
\end{aligned}
$$

Consequently, it holds that $E\left[|\mathbb{I}_i^n|\right] \lesssim \Delta_n^{3/2}\left(m^2\Delta_n + m\sqrt{\overline{\gamma}_m\Delta_n} + \overline{\gamma}_m\right)$. On the other hand, since ν is the compensator of μ, we have

$$E\left[|\mathbb{II}_i^n|\right] \leq E\left[\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{A_m^c}\left|X_s^c - X_{(i-1)\Delta_n}^c\right|\delta(s,z)^2\nu(ds,dz)\right] \lesssim \Delta_n^{3/2}\overline{\gamma}_m$$

by (6) and (10), whereas the Davis inequality, (6) and (9) imply that

$$E\left[|\mathbb{III}_i^n|\right] \lesssim E\left[\left\{\int_{(i-1)\Delta_n}^{i\Delta_n}(X''(m)_{s-} - X''(m)_{(i-1)\Delta_n})^4\sigma_s^2ds\right\}^{1/2}\right]$$

$$\lesssim \Delta_n^{1/2}E\left[\sup_{(i-1)\Delta_n\leq s\leq i\Delta_n}(X''(m)_s - X''(m)_{(i-1)\Delta_n})^2\right]$$

$$\lesssim \Delta_n^{3/2}\left(m^2\Delta_n + \overline{\gamma}_m\right).$$

After all, there exists a positive constant K'' such that

$$E\left[\left|\Delta_i^n X^c\left(\Delta_i^n X''(m)\right)^2\right|\right] \leq K''\Delta_n^{3/2}\left(m^2\Delta_n + m\sqrt{\overline{\gamma}_m\Delta_n} + \overline{\gamma}_m\right) \qquad (11)$$

for every i, n, m.

Step 4) Setting $k(x,y) = g(x+y) - g(x) - g(y)$, we have

$$\overline{\mathcal{V}}_T^n(X(m),g) - \frac{1}{\sqrt{\Delta_n}}\mathcal{V}_T^n(X^c,g) = \frac{1}{\sqrt{\Delta_n}}\sum_{i=1}^{\lfloor T/\Delta_n\rfloor}k(\Delta_i^n X^c, \Delta_i^n X''(m)) + \overline{\mathcal{V}}_T^n(X''(m),g)$$

because $X(m) = X^c + X''(m)$. Therefore, the proof is completed once we verify the following equations for any $\eta > 0$:

$$\lim_{m\to\infty}\limsup_{n\to\infty}P\left(\Omega_n(T,m)\cap\left\{\frac{1}{\sqrt{\Delta_n}}\left|\sum_{i=1}^{\lfloor T/\Delta_n\rfloor}k(\Delta_i^n X^c, \Delta_i^n X''(m))\right| > \eta\right\}\right) = 0,$$
$$(12)$$

$$\lim_{m\to\infty}\limsup_{n\to\infty}P\left(\Omega_n(T,m)\cap\left\{\left|\overline{\mathcal{V}}_T^n(X''(m),g)\right| > \eta\right\}\right) = 0. \qquad (13)$$

This step is devoted to the proof of (12). First, since $g(0) = g'(0) = 0$ and $g''(x) = O(|x|)$ as $x \to 0$, there exists a positive constant α such that

$$|x| \leq 6A \Rightarrow |g(x)| \leq \alpha|x|^3, \quad |g'(x)| \leq \alpha|x|^2, \quad |g''(x)| \leq \alpha|x|. \qquad (14)$$

Next, using the fundamental theorem of calculus repeatedly, we have

$$k(x,y) = \int_0^y\int_0^x g''(v+u)dvdu,$$

hence (14) yields

$$|x| + |y| \le 6A \Rightarrow |k(x, y)| \le \alpha \left(|x|^2 |y| + |x||y|^2 \right). \tag{15}$$

Since $|\Delta_i^n X^c| \le 2A$ and $|\Delta_i^n X(m)| \le |\Delta_i^n X| + |\Delta_i^n X'(m)| \le 3A$ on $\Omega_n(T, m)$ due to (6), we obtain

$$E\left[\left| \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} k(\Delta_i^n X^c, \Delta_i^n X''(m)) \right| 1_{\Omega_n(T,m)} \right]$$

$$\lesssim \Delta_n^{-3/2} \left\{ \Delta_n \left(m\Delta_n + \sqrt{\overline{\gamma}_m \Delta_n} \right) + \Delta_n^{3/2} \left(m^2 \Delta_n + m\sqrt{\overline{\gamma}_m \Delta_n} + \overline{\gamma}_m \right) \right\}$$

by (15), the Schwarz inequality and (9)–(11). Therefore, noting that $\overline{\gamma}_m \to 0$ as $m \to \infty$ because of (6) and the dominated convergence theorem and that $\Delta X(m)_s = \Delta X''(m)_s$ as well as $\Delta X_s^c = 0$ for all $s \ge 0$, (12) has been shown.

Step 5) Now we prove (13) and complete the proof of the theorem. First, set $\phi(x, y) = k(x, y) - g'(x)y$. If $|y| > |x|$, (14) and (15) yield

$$|x| \le 5A, |y| \le A \Rightarrow |\phi(x, y)| \le 3\alpha |x||y|^2, \tag{16}$$

whereas repeated applications of the fundamental theorem of calculus imply that $\phi(x, y) = \int_0^y \int_0^u g''(x + v) dv du - g(y)$, hence, if $|y| \le |x|$, by (14) we have

$$|x| \le 5A, |y| \le A \Rightarrow |\phi(x, y)| \le \alpha(|x| + |y|)|y|^2 + \alpha|y|^3 \le 3\alpha|x||y|^2,$$

and thus (16) holds true.

Next, for any i, an application of Itô's formula to the process $\Xi(m, i)_t = \int_0^t 1_{\{s > (i-1)\Delta_n\}} dX''(m)_s$ and the function g yields

$$g(\Xi(m, i)_t) = \int_{(i-1)\Delta_n}^t g'(\Xi(m, i)_{s-}) dX''(m)_s$$

$$+ \sum_{(i-1)\Delta_n < s \le t} \{ \phi(\Xi(m, i)_{s-}, \Delta X(m)_s) + g(\Delta X(m)_s) \}$$

for any $t > (i - 1)\Delta_n$. Therefore, noting that $\sum_{(i-1)\Delta_n < s \le t} g(\Delta X(m)_s)$ is well-defined by assumption, for any $t > (i - 1)\Delta_n$ we have

$$g(\Xi(m, i)_t) - \sum_{(i-1)\Delta_n < s \le t} g(\Delta X(m)_s)$$

$$\int_{(i-1)\Delta_n}^t g'(\Xi(m, i)_{s-}) rmdX''(m)_s + \int_{(i-1)\Delta_n}^t \int_{A_m^c} \phi(\Xi(m, i)_{s-}, \delta(s, z)) \mu(ds, dz)$$

$$= \int_{(i-1)\Delta_n}^t a(n, m, i)_u du + \int_{(i-1)\Delta_n}^t \int_{A_m^c} k(\Xi(m, i)_{s-}, \delta(s, z))(\mu - \nu)(ds, dz)$$

$$=: A(n, m, i)_t + M(n, m, i)_t, \tag{17}$$

where $a(n, m, i)_u = g'(\Xi(m, i)_u)b(m)_u + \int_{A_m^c} \phi(\Xi(m, i)_u, \delta(u, z))\lambda(dz)$. Note that $A(n, m, i)$ and $M(n, m, i)$ are well-defined due to (6) and (15)–(16).

Let us set $T(n, m, i) = \inf\{s > (i-1)\Delta_n : |\Xi(m, i)_s| > 5A\}$. On the set $\Omega_n(T, m)$ we have $|\Xi(m, i)_s| \leq 5A$ for all $s \leq T$ and $i \leq \lfloor T/\Delta_n \rfloor$ due to the decomposition $X''(m) = X - X^c - X'(m)$ and (6), hence $T(n, m, i) > i\Delta_n$. Thus, in view of (17), we have

$$
P\left(\Omega_n(T, m) \cap \left\{\left|\overline{\mathcal{V}}_T^n(X''(m), g)\right| > \eta\right\}\right)
$$

$$
\leq P\left(\frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} |A(n, m, i)_{(i\Delta_n)\wedge T(n,m,i)}| > \frac{\eta}{2}\right)
$$

$$
+ P\left(\frac{1}{\sqrt{\Delta_n}} \left|\sum_{i=1}^{\lfloor T/\Delta_n \rfloor} M(n, m, i)_{(i\Delta_n)\wedge T(n,m,i)}\right| > \frac{\eta}{2}\right). \tag{18}
$$

Therefore, in order to prove (13) it suffices to show that[2]

$$
\left.
\begin{aligned}
&\lim_{m\to\infty} \limsup_{n\to\infty} \frac{1}{\sqrt{\Delta_n}} E\left[\sum_{i=1}^{\lfloor T/\Delta_n \rfloor} |A(n, m, i)_{(i\Delta_n)\wedge T(n,m,i)}|\right] = 0, \\
&\lim_{m\to\infty} \limsup_{n\to\infty} \frac{1}{\Delta_n} E\left[\sum_{i=1}^{\lfloor T/\Delta_n \rfloor} \langle M(n, m, i)\rangle_{(i\Delta_n)\wedge T(n,m,i)}\right] = 0,
\end{aligned}
\right\} \tag{19}
$$

where we apply the Lenglart inequality to derive the convergence of the second term in the right side of Eq. (18) from the second convergence of Eq. (19) (for this application we need to introduce the stopping time $T(n, m, i)$, which enables us to drop the indicator $1_{\Omega_n(T,m)}$). Recall that $|b(m)| \leq (1 + m)A$ and $|\delta(s, z)| \leq \gamma(z) \leq A$ due to (6), so we have for $(i - 1)\Delta_n \leq u < T(n, m, i)$ (then $|\Xi(m, i)_u| \leq 5A$):

$$
|a(n, m, i)_u| \lesssim m|\Xi(m, i)_u|^2 + \overline{\gamma}_m|\Xi(m, i)_u|,
$$

$$
\int_{A_m^c} k(\Xi(m, i)_u, \delta(u, z))^2\lambda(dz) \lesssim \overline{\gamma}_m|\Xi(m, i)_u|^2
$$

by (6), (14)–(15) and (16). Combining these estimates with (9) as well as the fact that $\overline{\gamma}_m \to 0$ as $m \to \infty$, we conclude that (19) holds true. $\qquad\square$

4.2 Proof of Proposition 1

We can easily check that g_a is a C^2 function and we have

$$
g_a'(x) = x|x| \{3\sin(2a\log|x|) + 2a\cos(2a\log|x|)\}
$$

---

[2] Recall that, for a locally square-integrable martingale M such that $M_0 = 0$, $\langle M \rangle$ denotes the predictable quadratic variation of M, i.e., the predictable increasing process such that $M^2 - \langle M \rangle$ is a local martingale (such a process always exists and is unique; see e.g., Theorem 4.2 from Chapter I of Jacod and Shiryaev (2003).

and

$$g_a''(x) = |x| \left\{ 6 \sin(2a \log |x|) + 10a \cos(2a \log |x|) - 4a^2 \sin(2a \log |x|) \right\}$$

for any $x \in \mathbb{R}$. Hence claim (a) holds true.

Next we prove claim (b). In the following we denote by $\mathfrak{N}$ the standard normal density. First we show that there is a real number a such that

$$\int_0^\infty x^3 \sin(2a \log x) \mathfrak{N}(x) dx \neq 0. \tag{20}$$

In fact, substituting $y = \log x$, we have

$$\int_0^\infty x^3 \sin(2a \log x) \mathfrak{N}(x) dx = \int_{-\infty}^\infty e^{4y} \mathfrak{N}(e^y) \sin(2ay) dy.$$

Since the function $\mathbb{R} \ni y \mapsto e^{4y} \mathfrak{N}(e^y) \in \mathbb{R}$ is square integrable and not even, the imaginary part of its Fourier transform is not identical to zero. Hence Eq. (20) holds true for some $a \in \mathbb{R}$.

Now we show that the variables $\mathcal{V}_T^n(X^c, g_a)/\sqrt{\Delta_n}$ do not converge in law with $\Delta_n = \exp(-n\pi/a)$ if a satisfies Eq. (20) (note that such an a must not be zero). To obtain a contradiction, suppose that the variables $\mathcal{V}_T^n(X^c, g_a)/\sqrt{\Delta_n}$ converge in law to some random variable Z as $n \to \infty$. Since we have

$$\mathrm{Var}\left[\frac{1}{\sqrt{\Delta_n}} \mathcal{V}_T^n(X^c, g_a) \right] = \frac{1}{\Delta_n} \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} \mathrm{Var}\left[g_a(\Delta_i^n X^c) \right]$$

$$\leq \frac{1}{\Delta_n} \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} E\left[(\Delta_i^n X^c)^6 \right] = O(\Delta_n)$$

and

$$\frac{1}{\sqrt{\Delta_n}} |E[\mathcal{V}_T^n(X^c, g_a)]| \leq \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} E\left[|\Delta_i^n X^c|^3 \right] = O(1),$$

we obtain

$$\sup_{n \in \mathbb{N}} E\left[\left| \frac{1}{\sqrt{\Delta_n}} \mathcal{V}_T^n(X^c, g_a) \right|^2 \right] < \infty.$$

Therefore, the variables $\mathcal{V}_T^n(X^c, g_a)/\sqrt{\Delta_n}$ are uniformly integrable, and thus Theorem 3.5 of Billingsley (1999) implies that Z is integrable and

$$E\left[\frac{1}{\sqrt{\Delta_n}} \mathcal{V}_T^n(X^c, g_a) \right] \to E[Z] \tag{21}$$

as $n \to \infty$. In the meantime, we have

$$
E\left[\frac{1}{\sqrt{\Delta_n}}V_T^n(X^c, g_a)\right] = \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} E\left[g_a(\Delta_i^n X^c)\right]
$$

$$
= \frac{2T}{\Delta_n^{3/2}} \int_0^\infty g_a(\sqrt{\Delta_n}x)\mathfrak{N}(x)dx + O(\Delta_n)
$$

$$
= 2T \int_0^\infty x^3 \sin\left(2a \log(\sqrt{\Delta_n}x)\right)\mathfrak{N}(x)dx + O(\Delta_n)
$$

as $n \to \infty$. Hence, in view of Eq. (21), the sequence

$$
c_n := \int_0^\infty x^3 \sin\left(2a \log(\sqrt{\Delta_n}x)\right)\mathfrak{N}(x)dx, \qquad n = 1, 2, \ldots
$$

converges as $n \to \infty$. Using the identity

$$
\sin\left(2a \log(\sqrt{\Delta_n}x)\right) = \sin(a \log \Delta_n) \cos(2a \log x) + \cos(a \log \Delta_n) \sin(2a \log x),
$$

we can rewrite c_n as

$$
c_n = \sin(a \log \Delta_n) \int_0^\infty x^3 \cos(2a \log x)\,\mathfrak{N}(x)dx
$$

$$
+ \cos(a \log \Delta_n) \int_0^\infty x^3 \sin(2a \log x)\,\mathfrak{N}(x)dx.
$$

Since $\Delta_n = \exp(-n\pi/a)$, we obtain

$$
c_n = (-1)^n \int_0^\infty x^3 \sin(2a \log x)\,\mathfrak{N}(x)dx.
$$

Therefore, the sequence c_n does not converge due to Eq. (20), a contradiction. □

4.3 Proof of Theorem 3

4.3.1 Localization

As in Sect. 4.1, a standard localization argument allows us to replace the assumptions [A2]–[A3] by the following strengthened versions:

[SA2] We have [A2], and the processes X_t, b_t, σ_t, $\widetilde{b}_t$ and $\widetilde{\sigma}_t$ are bounded. Moreover, there are a constant Λ and a nonnegative bounded function γ on E such that

$\int \gamma(z)^2 \lambda(\mathrm{d}z) < \infty$ and $|\delta(\omega, t, z)| \vee |\tilde{\delta}(\omega, t, z)| \leq \gamma(z)$ and

$$E\left[|b_{t_1} - b_{t_2}|^2|\mathcal{F}_{t_1 \wedge t_2}\right] \leq \Lambda E\left[|t_1 - t_2|^{\varpi}|\mathcal{F}_{t_1 \wedge t_2}\right],$$

$$E\left[|\delta(t_1, z) - \delta(t_2, z)|^2|\mathcal{F}_{t_1 \wedge t_2}\right] \leq \Lambda \gamma(z)^2 E\left[|t_1 - t_2|^{\varpi}|\mathcal{F}_{t_1 \wedge t_2}\right]$$

for any bounded $(\mathcal{F}_t)$-stopping times t_1 and t_2.

[SA3] The process $\int u^6 Q_t(\mathrm{d}z)$ is bounded and there is a constant Λ' such that

$$E\left[|\alpha_{t_1} - \alpha_{t_2}|^2|\mathcal{F}_{t_1 \wedge t_2}\right] \leq \Lambda' E\left[|t_1 - t_2|^{\varpi}|\mathcal{F}_{t_1 \wedge t_2}\right]$$

for any bounded $(\mathcal{F}_t)$-stopping times t_1 and t_2. Moreover, α_t is càdlàg.

In the following we fix a constant $\xi \in (0, 1)$ such that

$$\xi > \frac{11}{12} \vee \frac{2 + \varpi}{2(1 + \varpi)}, \tag{22}$$

and we set $\bar{r}_n = \Delta_n^{\xi}$. By a similar argument to Section 6.1.1 of Koike (2016), we can further replace the assumption [A1] by the following strengthened version:

[SA1] We have [A1], and for every n it holds that

$$\sup_{i \geq 0}(t_i^n - t_{i-1}^n) \leq \bar{r}_n. \tag{23}$$

4.3.2 Notation and estimates

We use the same notation as in Sect. 4.1 with the following change for the definition of the set $\Omega_n(T, m)$: For $m, n \in \mathbb{N}$, we denote by $\Omega_n(T, m)$ the set on which $k_n - 1 \leq N_{S_p-}^n \leq N_T^n - k_n$ for all $p \in \mathcal{P}_m$ such that $S_p \leq T$, and $|S_{p_1} - S_{p_2}| > k_n \bar{r}_n$ for any $p_1, p_2 \in \mathcal{P}_m$ such that $p_1 \neq p_2$ and $S_{p_1}, S_{p_2} < \infty$. We have $\lim_n P(\Omega_n(T, m)) = 1$. We additionally define the processes $B(m)$ and $Z(m)$ by $B(m)_t = \int_0^t b(m)_s \mathrm{d}s$ and $Z(m)_t = \int_0^t \int_{A_m^c} \delta(s, z)(\mu - \nu)(\mathrm{d}s, \mathrm{d}z)$, respectively. We also define the 2×2 symmetric matrix valued variable $\overline{\Gamma}(m)_T$ by

$$\overline{\Gamma}(m)_T^{11} = \frac{4}{\psi_2^2} \sum_{p \in \mathcal{P}_m, S_p \leq T} (\Delta X_{S_p})^2 \left\{ \Phi_{22}\theta \left(\sigma_{S_p}^2 G_{S_p} + \sigma_{S_p-}^2 G_{S_p-} \right) + \frac{\Phi_{12}}{\theta} (\alpha_s + \alpha_{s-}) \right\},$$

$$\overline{\Gamma}(m)_T^{12} = \frac{6}{\psi_2 \psi_3} \sum_{p \in \mathcal{P}_m, S_p \leq T} (\Delta X_{S_p})^3 \left\{ \theta \left(\Phi_{23+}\sigma_{S_p}^2 G_{S_p} + \Phi_{23-}\sigma_{S_p-}^2 G_{S_p-} \right) \right.$$
$$\left. + \theta^{-1} \left(\Phi_{23+}' \alpha_{S_p} + \Phi_{23-}' \alpha_{S_p-} \right) \right\},$$

$$\overline{\Gamma}(m)_T^{22} = \frac{9}{\psi_3^2} \sum_{p \in \mathcal{P}_m, S_p \leq T} (\Delta X_{S_p})^4 \left\{ \theta \left(\Phi_{3+}\sigma_{S_p}^2 G_{S_p} + \Phi_{3-}\sigma_{S_p-}^2 G_{S_p-} \right) \right.$$
$$\left. + \theta^{-1} \left(\Phi_{3+}' \alpha_{S_p} + \Phi_{3-}' \alpha_{S_p-} \right) \right\}.$$

We set $g_p^n = g(p/k_n)$ for $p = 0, 1, \ldots, k_n$ and $\Delta(g)_p^n = g_{p+1}^n - g_p^n$ for $p = 0, 1, \ldots, k_n - 1$. We also set $I_i = [t_{i-1}^n, t_i^n)$ and $\overline{I}_i = [t_{i-1}^n, t_{i+k_n-1}^n)$ for $i = 0, 1, \ldots$.

For every $i \geq 0$ we define the process $\overline{g}_i^n$ by $\overline{g}_i^n(s) = \sum_{p=1}^{k_n-1} g_p^n 1_{I_{i+p}}(s)$. For any semimartingale V, we define the process $\overline{V}_{i,t}$ by $\overline{V}_{i,t} = \int_0^t \overline{g}_i^n(s-)dV_s$. Note that $\overline{V}_i = \overline{V}_{i,t_{i+k_n-1}^n}$.

Recall that, for random variables X and Y which may depend on the parameters n, m, i, $X \lesssim Y$ means that there exists a (non-random) constant $K > 0$ independent of n, m, i such that $X \leq KY$ a.s. In addition, if K possibly depends on m, we write $X \lesssim_m Y$ instead.

[SA2] and Eq. (23) yield

$$E\left[\sup_{s \in \overline{I}_i} \left|\overline{X'(m)}_{i,s}\right| \,\big|\, \mathcal{F}_{t_{i-1}^n}\right] \lesssim m E\left[|\overline{I}_i| \,\big|\, \mathcal{F}_{t_{i-1}^n}\right] \lesssim m k_n \overline{r}_n \tag{24}$$

and

$$\left|\overline{B(m)}_i\right| \lesssim m \left|\overline{I}_i\right| \lesssim m k_n \overline{r}_n. \tag{25}$$

Here, $|\cdot|$ denotes the Lebesgue measure. The BDG inequality, [SA2] and Eq. (23) yield

$$E\left[\sup_{s \in \overline{I}_i} \left|\overline{X}_{i,s}^c\right|^r \,\big|\, \mathcal{F}_{t_{i-1}^n}\right] \lesssim E\left[|\overline{I}_i|^{r/2} \,\big|\, \mathcal{F}_{t_{i-1}^n}\right] \lesssim (k_n \overline{r}_n)^{r/2} \qquad \text{for any } r > 0 \tag{26}$$

and

$$E\left[\sup_{s \in \overline{I}_i} \left|\overline{Z(m)}_{i,s}\right|^2 \,\big|\, \mathcal{F}_{t_{i-1}^n}\right] \lesssim \overline{\gamma}_m E\left[|\overline{I}_i| \,\big|\, \mathcal{F}_{t_{i-1}^n}\right] \lesssim \overline{\gamma}_m k_n \overline{r}_n, \tag{27}$$

where $\overline{\gamma}_m = \int_{A_m^c} \gamma(z)^2 \lambda(dz)$. Note that $\overline{\gamma}_m \to 0$ as $m \to \infty$ by the dominated convergence theorem. The BDG inequality and [SA3] yield

$$E\left[|\overline{\epsilon}_i|^r \,|\, \mathcal{F}\right] + E\left[|\overline{\epsilon}_i|^r \,|\, \mathcal{F}_{t_{i-1}^n}\right] \leq K_r k_n^{-r/2} \qquad \text{for any } r \in [2, 6]. \tag{28}$$

Finally, Lemma 6.1 of Koike (2017) implies that

$$N_T^n = O_p(\Delta_n^{-1}) \tag{29}$$

as $n \to \infty$ for every $t > 0$.

4.3.3 Main body

For each $m \in \mathbb{N}$, we consider the following decomposition of $\Delta_n^{-1/4}$ $\left(\mathrm{PCV}_T^n - \sum_{0 \le s \le T} (\Delta X_s)^3\right)$:

$$
\Delta_n^{-1/4} \left(\mathrm{PCV}_T^n - \sum_{0 \le s \le T} (\Delta X_s)^3 \right)
$$

$$
= \frac{\Delta_n^{-1/4}}{\psi_3 k_n} \sum_{i=0}^{N_T^n - k_n + 1} \left\{ \left(\overline{X(m)}_i + \overline{\epsilon}_i \right)^3 - \left(\overline{X(m)}_i \right)^3 \right\}
$$

$$
+ \Delta_n^{-1/4} \left\{ \frac{1}{\psi_3 k_n} \sum_{i=0}^{N_T^n - k_n + 1} \left(\overline{X(m)}_i \right)^3 - \sum_{0 \le s \le T} (\Delta X(m)_s)^3 \right\}
$$

$$
+ 3 \frac{\Delta_n^{-1/4}}{\psi_3 k_n} \sum_{i=0}^{N_T^n - k_n + 1} \left(\overline{X(m)}_i + \overline{\epsilon}_i \right)^2 \overline{X'(m)}_i
$$

$$
+ 3 \frac{\Delta_n^{-1/4}}{\psi_3 k_n} \sum_{i=0}^{N_T^n - k_n + 1} \left(\overline{C(m)}_i + \overline{\epsilon}_i \right) \left(\overline{X'(m)}_i \right)^2
$$

$$
+ 3 \frac{\Delta_n^{-1/4}}{\psi_3 k_n} \sum_{i=0}^{N_T^n - k_n + 1} \overline{Z(m)}_i \left(\overline{X'(m)}_i \right)^2
$$

$$
+ \Delta_n^{-1/4} \left\{ \frac{1}{\psi_3 k_n} \sum_{i=0}^{N_T^n - k_n + 1} \left(\overline{X'(m)}_i \right)^3 - \sum_{0 \le s \le T} (\Delta X'(m)_s)^3 \right\}
$$

$$
=: \mathbb{I}_n(m) + \mathbb{II}_n(m) + \mathbb{III}_n(m) + \mathbb{IV}_n(m) + \mathbb{V}_n(m) + \mathbb{VI}_n(m).
$$

Recall that the process $Y(m)$ is defined in Eq. (7) and corresponds to the process obtained by removing "small" jumps from the process X. Since we have

$$
\lim_{m \to \infty} \limsup_{n \to \infty} P_n \left(\Delta_n^{-1/4} \left| \left(\mathrm{PRV}_T^n - [X, X]_T \right) - \left(\mathrm{PRV}(m)_T^n - [Y(m), Y(m)]_T \right) \right| > \eta \right) = 0
$$

for any $\eta > 0$ by Proposition 6.3 from Koike (2016), where

$$
\mathrm{PRV}(m)_T^n = \frac{1}{\psi_2 k_n} \sum_{i=0}^{N_T^n - k_n + 1} \left(\overline{Y(m)}_i + \overline{\epsilon}_i \right)^2 - \frac{\psi_1}{2\psi_2 k_n} \sum_{i=1}^{N_T^n} (Y_{t_i^n} - Y_{t_{i-1}^n})^2,
$$

and we obviously have $\overline{\Gamma}(m)_T \to^P \overline{\Gamma}_T$ as $m \to \infty$, the proof is completed once we show the following convergences for any $\eta > 0$ due to Proposition 2.2.4 of Jacod and Protter (2012):

$$\lim_{m\to\infty} \limsup_{n\to\infty} P_n\left(|\mathbb{I}_n(m)| > \eta\right) = 0, \tag{30}$$

$$\lim_{m\to\infty} \limsup_{n\to\infty} P_n\left(|\mathbb{III}_n(m)| > \eta\right) = 0, \tag{31}$$

$$\mathbb{III}_n(m) \to^P 0 \qquad \text{for any } m \in \mathbb{N}, \tag{32}$$

$$\left(\Delta_n^{-1/4}\left(\text{PRV}(m)_T^n - [Y(m), Y(m)]_T\right), \mathbb{IV}_n(m)\right)^* \to^{\mathcal{S}} \Gamma(m)_T^{1/2}\zeta, \tag{33}$$

$$\lim_{m\to\infty} \limsup_{n\to\infty} P_n\left(|\mathbb{V}_n(m)| > \eta\right) = 0, \tag{34}$$

$$\mathbb{VI}_n(m) \to^P 0 \qquad \text{for any } m \in \mathbb{N}, \tag{35}$$

where ζ is a bivariate standard normal variable which is defined on an extension of $\mathcal{B}$ and independent of $\mathcal{F}$, and

$$\Gamma(m)_T = \begin{bmatrix} \Gamma_T^c + \overline{\Gamma}(m)_T^{11} & \overline{\Gamma}(m)_T^{12} \\ \overline{\Gamma}(m)_T^{12} & \overline{\Gamma}(m)_T^{22} \end{bmatrix}.$$

Proof of (35) On the set $\Omega_n(T, m)$ we have

$$\mathbb{VI}_n(m) = \frac{\Delta_n^{-1/4}}{\psi_3} \sum_{p\in\mathcal{P}_m: S_p \leq T} \left\{\left(\frac{1}{k_n}\sum_{i=1}^{k_n-1}(g_i^n)^3\right) - \psi_3\right\}(\Delta X_{S_p})^3.$$

Since $\frac{1}{k_n}\sum_{i=1}^{k_n-1}(g_i^n)^3 = \psi_3 + O(k_n^{-1})$ by the Lipschitz continuity of g and $\lim_n P(\Omega_n(T, m)) = 1$, we obtain the desired result. □

Proof of (30) We decompose the target quantity as

$$\mathbb{I}_n(m) = \frac{\Delta_n^{-1/4}}{\psi_3 k_n} \sum_{i=0}^{N_T^n - k_n + 1} \left[3\left(\overline{Z(m)}_i\right)^2 \overline{\epsilon}_i \right.$$
$$\left. + \left\{\left(\overline{X(m)}_i + \overline{\epsilon}_i\right)^3 - \left(\overline{X(m)}_i\right)^3 - 3\left(\overline{Z(m)}_i\right)^2 \overline{\epsilon}_i\right\}\right]$$
$$=: \mathbb{I}_n^{(1)}(m) + \mathbb{I}_n^{(2)}(m).$$

It suffices to prove

$$\lim_{m\to\infty} \limsup_{n\to\infty} P_n\left(\left|\mathbb{I}_n^{(l)}(m)\right| > \eta\right) = 0 \tag{36}$$

for $l = 1, 2$. First, by Eqs. (28) and (27) we have

$$E\left[\left|\mathbb{I}_n^{(1)}(m)\right|\right] \lesssim \frac{1}{k_n} E\left[\sum_{i=0}^{N_T^n+1}\left(\overline{Z(m)}_i\right)^2\right] \lesssim \frac{\overline{\gamma}_m}{k_n} E\left[\sum_{i=0}^{N_T^n+1}|\overline{I}_i|\right] \lesssim \overline{\gamma}_m,$$

where we use the following inequality to obtain the final upper bound:

$$\sum_{i=0}^{N_T^n+1} |\bar{I}_i| = \sum_{i=0}^{N_T^n+1} \sum_{p=0}^{k_n-1} |I_{i+p}| = \sum_{p=0}^{k_n-1} \sum_{i=0}^{N_T^n-k_n+1} |I_{i+p}| + \sum_{i=N_T^n-k_n+2}^{N_T^n+1} \sum_{p=0}^{k_n-1} |I_{i+p}|$$

$$\leq k_n T + k_n(k_n-1)\bar{r}_n \lesssim k_n.$$

Hence Eq. (36) holds true for $l = 1$.

Next we consider the case $l = 2$, and we start with some preliminary results. First we note that

$$\sup_{N_T^n-k_n+1<i\leq N_T^n+1} |\bar{V}_i| = O_p(\sqrt{k_n \bar{r}_n}) \qquad (37)$$

for $V \in \{X(m), Z(m)\}$. In fact, summation by parts yields $\bar{V}_i = -\sum_{p=0}^{k_n-1} \Delta(g)_p^n (V_{t_{i+p}^n} - V_{t_i^n})$, and $\sup_{|h|\leq h_0} |V_{t+h} - V_t| = O_p(\sqrt{h_0})$ as $h_0 \downarrow 0$ by [SA2] and the Doob inequality, hence Eq. (37) holds true by Eq. (23). Next, for any $K > 0$ we define the $(\mathcal{F}_t)$-stopping time R_K^n by

$$R_K^n = \inf\left\{s : n^{-1}N_s^n > K\right\}. \qquad (38)$$

Since $\Delta N_s^n \leq 1$ for every s, it holds that

$$N_{s\wedge R_K^n}^n \leq Kn + 1 \qquad (39)$$

for all $s \geq 0$. Moreover, by Eq. (29) we also have

$$\limsup_{K\to\infty} \limsup_{n\to\infty} P\left(R_K^n \leq t\right) = 0. \qquad (40)$$

Now we turn to the proof of Eq. (36) for the case $l = 2$. For each m, set

$$\zeta(m)_i^n = \frac{\Delta_n^{-1/4}}{\psi_3 k_n}\left\{\left(\overline{X(m)}_i + \bar{\epsilon}_i\right)^3 - \left(\overline{X(m)}_i\right)^3 - 3\left(\overline{Z(m)}_i\right)^2 \bar{\epsilon}_i\right\}, \qquad i = 0, 1, \ldots.$$

Then, by Eqs. (28) and (37) we have $\mathbb{I}_n^{(2)}(m) = \sum_{i=0}^{N_T^n+1} \zeta(m)_i^n + o_p(1)$ as $n \to \infty$. Therefore, by the Markov inequality and Eq. (40) it is enough to prove $\sum_{i=0}^{N_{T\wedge R_K^n}^n+1} \zeta(m)_i^n = o_p(1)$ as $n \to \infty$ for any fixed $K > 0$ and $m \in \mathbb{N}$.

Since integration by parts yields $\overline{X_i^c Z(m)}_i = \int_{\overline{I}_i} \overline{X}_{i,s-}^c \, d\overline{Z(m)}_{i,s} + \int_{\overline{I}_i} \overline{Z(m)}_{i,s-} \, d\overline{X}_{i,s}^c$, we have

$$E\left[\left|\overline{X_i^c Z(m)}_i\right|^2 \big| \mathcal{F}_{t_{i-1}^n}\right] \lesssim E\left[\left(\sup_{s\in\overline{I}_i}\left|\overline{X}_{i,s-}^c\right|^2 + \sup_{s\in\overline{I}_i}\left|\overline{Z(m)}_{i,s-}\right|^2\right)|\overline{I}_i|\right] \lesssim (k_n\bar{r}_n)^2 \tag{41}$$

by [SA2], Eqs. (23) and (26)–(27). Moreover, we can rewrite $\zeta(m)_i^n$ as

$$\zeta(m)_i^n = \frac{\Delta_n^{-1/4}}{\psi_3 k_n}\left\{3\left(\left(\overline{C(m)}_i\right)^2 + 2\overline{C(m)}_i\overline{Z(m)}_i\right)\bar{\epsilon}_i + 3\overline{X(m)}_i\,(\bar{\epsilon}_i)^2 + (\bar{\epsilon}_i)^3\right\}.$$

Hence, using the relation $C(m)_t = X_0 + B(m)_t + X_t^c$ and estimates Eqs. (25)–(28) and (41), we obtain $E[|\zeta(m)_i^n|^2|\mathcal{F}_{t_{i-1}^n}] \lesssim_m \bar{r}_n^2$. Therefore, noting that $\zeta(m)_i^n$ is $\mathcal{F}_{t_{i+k_n-1}^n}$-measurable, we have

$$E\left[\left|\sum_{i=0}^{N_{T\wedge R_K^n}^n+1}\left(\zeta(m)_i^n - E[\zeta(m)_i^n|\mathcal{F}_{t_{i-1}^n}]\right)\right|^2\right] \lesssim_m \Delta_n^{-1}k_n\bar{r}_n^2 = o(1).$$

Hence it holds that $\sum_{i=0}^{N_{T\wedge R_K^n}^n+1}\zeta(m)_i^n = \sum_{i=0}^{N_{T\wedge R_K^n}^n+1}E[\zeta(m)_i^n|\mathcal{F}_{t_{i-1}^n}] + o_p(1)$. Now, since $E[\bar{\epsilon}_i|\mathcal{F}] = 0$, we can decompose $\sum_{i=0}^{N_{T\wedge R_K^n}^n+1}E[\zeta(m)_i^n|\mathcal{F}_{t_{i-1}^n}]$ as

$$\sum_{i=0}^{N_{T\wedge R_K^n}^n+1}E[\zeta(m)_i^n|\mathcal{F}_{t_{i-1}^n}] = \frac{\Delta_n^{-1/4}}{\psi_3 k_n}\sum_{i=0}^{N_{T\wedge R_K^n}^n+1}\left\{3E[\overline{X(m)}_i\,(\bar{\epsilon}_i)^2\,|\mathcal{F}_{t_{i-1}^n}] + E[(\bar{\epsilon}_i)^3\,|\mathcal{F}_{t_{i-1}^n}]\right\}$$
$$=: \mathbb{A}_{1,n} + \mathbb{A}_{2,n}$$

(we drop the index m because we fix it here). First we consider $\mathbb{A}_{1,n}$. We can rewrite it as

$$\mathbb{A}_{1,n} = \frac{3\Delta_n^{-1/4}}{\psi_3 k_n}\sum_{i=0}^{N_{T\wedge R_K^n}^n+1}\sum_{p=0}^{k_n-1}(\Delta(g)_p^n)^2 E[\overline{X(m)}_i\alpha_{t_{i+p}^n}|\mathcal{F}_{t_{i-1}^n}].$$

Since we have $E[\overline{X(m)}_i\alpha_{t_{i-1}^n}|\mathcal{F}_{t_{i-1}^n}] = E[\overline{B(m)}_i\alpha_{t_{i-1}^n}|\mathcal{F}_{t_{i-1}^n}]$, it holds that

$$\left|E[\overline{X(m)}_i\alpha_{t_{i+p}^n}|\mathcal{F}_{t_{i-1}^n}]\right| \leq E\left[\left|\overline{X(m)}_i(\alpha_{t_{i+p}^n} - \alpha_{t_{i-1}^n})\right||\mathcal{F}_{t_{i-1}^n}\right] + E\left[\left|\overline{B(m)}_i\alpha_{t_{i-1}^n}\right||\mathcal{F}_{t_{i-1}^n}\right]$$
$$\lesssim_m (k_n\bar{r}_n)^{(1+\varpi)/2} + k_n\bar{r}_n$$

by the Schwarz inequality, [SA3] and Eqs. (25)–(27). Therefore, we obtain $\mathbb{A}_{1,n} = O_p\left(\Delta_n^{-\frac{1}{4}+\left(\frac{1+\varpi}{2}\wedge 1\right)\left(\xi-\frac{1}{2}\right)}\right) = o_p(1)$ by Eq. (22) after distinguishing the cases $\varpi \geq 1$ and $\varpi < 1$. Next, let us consider $\mathbb{A}_{2,n}$. For any nonnegative integers p, q, r, $E[\epsilon_{t^n_{i+p}}\epsilon_{t^n_{i+q}}\epsilon_{t^n_{i+r}}|\mathcal{F}^n_{t^n_{i-1}}]$ does not vanish only if $p = q = r$, hence we have

$$
\mathbb{A}_{2,n} = \frac{\Delta_n^{-1/4}}{\psi_3 k_n} \sum_{i=0}^{N^n_{T\wedge R^n_K}+1} \sum_{p=0}^{k_n-1} (\Delta(g)^n_p)^3 E\left[(\epsilon_{t^n_{i+p}})^3|\mathcal{F}^n_{t^n_{i-1}}\right]
$$
$$
= O_p(\Delta_n^{-1/4}k_n^{-1}\cdot\Delta_n^{-1}k_n^{-2}) = o_p(1).
$$

Consequently, we conclude that $\sum_{i=0}^{N^n_{T\wedge R^n_K}+1}\zeta(m)^n_i = o_p(1)$ and the proof is completed. □

Proof of (33) The proof is analogous to that of Proposition 6.2 of Koike (2016), which is based on Propositions 6.4–6.7 of that paper. So we omit it. □

We recall that, for a locally square-integrable martingale M such that $M_0 = 0$, $\langle M\rangle$ denotes the predictable quadratic variation of M (see also footnote 2). The next inequality plays a key role in the remaining proof:

Lemma 1 *We have*

$$
E\left[\sup_{\tau_1\leq t\leq\tau_2}|M_t|\,|\mathcal{F}_{\tau_1}\right] \leq 3E\left[\sqrt{\langle M\rangle_{\tau_2}}|\mathcal{F}_{\tau_1}\right]
$$

for any stopping time τ_1, τ_2 such that $\tau_1 \leq \tau_2$ and for any locally square-integrable martingale M such that $M_0 = 0$.

Proof This result is a direct consequence of Theorem 5 from Chapter 1, Section 9 of Liptser and Shiryayev (1989). □

Proof of (32) Set $L(m) = X^c + Z(m)$. Then we have

$$
|\mathbb{III}_n(m)| \leq 6\frac{\Delta_n^{-1/4}}{\psi_3 k_n} \sum_{i=0}^{N^n_T-k_n+1}\left\{\left(\overline{L(m)}_i\right)^2 + \left(\overline{B(m)}_i + \overline{\epsilon}_i\right)^2\right\}\left|\overline{X'(m)}_i\right|
$$
$$
=: \mathbb{III}_n^{(1)} + \mathbb{III}_n^{(2)},
$$

so it suffices to prove $\mathbb{III}_n^{(l)} \to^P 0$ as $n \to \infty$ for $l = 1, 2$ (note that we drop the index m because it is fixed in this part). First, Eqs. (25), (28) and (24) yield

$$
E\left[\left|\mathbb{III}_n^{(2)}\right|\right] \lesssim_m \frac{\Delta_n^{-1/4}}{k_n^2}E\left[\sum_{i=0}^{N^n_T+1}\left|\overline{X'(m)}_i\right|\right] \lesssim_m \Delta_n^{-1/4}k_n^{-1} = o(1),
$$

hence we have $\mathbb{III}_n^{(2)} \to^P 0$ as $n \to \infty$.

To prove $\mathrm{III}_n^{(1)} \to^P 0$ as $n \to \infty$, it suffices to show that there is a constant K (which may depend on m) such that

$$E\left[\left|\left(\overline{L(m)}_i\right)^2 \overline{X'(m)}_i\right| \,\middle|\, \mathcal{F}_{t_{i-1}^n}\right] \leq K\,(k_n \bar{r}_n)^2 \tag{42}$$

for any i, n because of the Lenglart inequality, Eq. (29) and the fact that $\xi > 7/8$. To prove Eq. (42), we consider the following decomposition of $\left(\overline{L(m)}_i\right)^2 \overline{X'(m)}_i$, which is obtained by applying integration by parts repeatedly (note that $[L(m), X'(m)] \equiv 0$ by construction):

$$\left(\overline{L(m)}_i\right)^2 \overline{X'(m)}_i = \int_{\overline{I}_i} \left(\overline{L(m)}_{i,s-}\right)^2 d\overline{X'(m)}_{i,s}$$

$$+ 2 \int_{\overline{I}_i} \overline{X'(m)}_{i,s-} \overline{L(m)}_{i,s-} d\overline{L(m)}_{i,s}$$

$$+ \int_{\overline{I}_i} \overline{X'(m)}_{i,s-} d[\overline{L(m)}_{i,\cdot}]_s$$

$$=: \mathbf{A}_i^{n,1} + 2\mathbf{A}_i^{n,2} + \mathbf{A}_i^{n,3}.$$

Then, it is enough to show that

$$E\left[\left|\mathbf{A}_i^{n,l}\right| \,\middle|\, \mathcal{F}_{t_{i-1}^n}\right] \lesssim_m (k_n \bar{r}_n)^2 \tag{43}$$

for every $l = 1, 2, 3$. First, we have

$$E\left[\left|\mathbf{A}_i^{n,1}\right| \,\middle|\, \mathcal{F}_{t_{i-1}^n}\right] \lesssim_m E\left[\sup_{s \in \overline{I}_i} \left(\overline{L(m)}_{i,s}\right)^2 |\overline{I}_i| \,\middle|\, \mathcal{F}_{t_{i-1}^n}\right]$$

$$\lesssim_m k_n \bar{r}_n E\left[\sup_{s \in \overline{I}_i} \left(\overline{L(m)}_{i,s}\right)^2 \,\middle|\, \mathcal{F}_{t_{i-1}^n}\right] \lesssim_m (k_n \bar{r}_n)^2$$

by [SA2], Eqs. (23) and (26)–(27), so Eq. (43) holds true for $l = 1$. Next, Lemma 1 and Eq. (23) yield

$$E\left[\left|\mathbf{A}_i^{n,2}\right| \,\middle|\, \mathcal{F}_{t_{i-1}^n}\right] \leq 3E\left[\left\{\int_{\overline{I}_i} \left|\overline{X'(m)}_{i,s-} \overline{L(m)}_{i,s-}\right|^2 d\langle\overline{L(m)}_{i,\cdot}\rangle_s\right\}^{1/2} \,\middle|\, \mathcal{F}_{t_{i-1}^n}\right]$$

$$\lesssim E\left[\sup_{s \in \overline{I}_i} \left|\overline{X'(m)}_{i,s} \overline{L(m)}_{i,s}\right| \sqrt{|\overline{I}_i|} \,\middle|\, \mathcal{F}_{t_{i-1}^n}\right]$$

$$\lesssim \sqrt{k_n \bar{r}_n} E\left[\sup_{s \in \overline{I}_i} \left|\overline{X'(m)}_{i,s} \overline{L(m)}_{i,s}\right| \,\middle|\, \mathcal{F}_{t_{i-1}^n}\right].$$

Now, integration by parts, Lemma 1, [SA2], Eqs. (23), (24) and (26)–(27) imply that

$$
E\left[\sup_{s\in\bar{I}_i}\left|\overline{X'(m)}_{i,s}\,\overline{L(m)}_{i,s}\right|\,\big|\mathcal{F}_{t^n_{i-1}}\right]
$$

$$
\lesssim_m E\left[\sup_{s\in\bar{I}_i}\left|\overline{X'(m)}_{i,s}\right|\sqrt{|\bar{I}_i|}\big|\mathcal{F}_{t^n_{i-1}}\right] + E\left[\sup_{s\in\bar{I}_i}\left|\overline{L(m)}_{i,s}\right|\,|\bar{I}_i|\,\big|\mathcal{F}_{t^n_{i-1}}\right] \lesssim_m (k_n\bar{r}_n)^{3/2}.
$$

Consequently, (43) holds true for $l = 2$. Finally, by [SA2], Eqs. (23) and (24) we have

$$
E\left[\left|\mathbf{A}^{n,3}_i\right|\,\big|\mathcal{F}_{t^n_{i-1}}\right] \lesssim_m E\left[\sup_{s\in\bar{I}_i}\left|\overline{X'(m)}_{i,s}\right|\,|\bar{I}_i|\,\big|\mathcal{F}_{t^n_{i-1}}\right] \lesssim (k_n\bar{r}_n)^2,
$$

hence Eq. (43) holds true for $l = 3$. This completes the proof. □

Proof of (34) Define the processes $B'(m)$ and $Z'(m)$ by $B'(m)_t = \int_0^t \int_{A_m} \delta(s, z)\mathrm{d}s\,\lambda(\mathrm{d}z)$ and $Z'(m)_t = X'(m)_t - B'(m)_t = \int_0^t \int_{A_m} \delta(s, z)(\mu - \nu)(\mathrm{d}s, \mathrm{d}z)$. Since we have

$$
|\mathbb{V}_n(m)| \le 3\frac{\Delta_n^{-1/4}}{\psi_3 k_n} \sum_{i=0}^{N^n_T-k_n+1}\left\{\left|\overline{Z(m)}_i\left(\overline{B'(m)}_i\right)^2\right| + \left|\overline{Z(m)}_i\left(\overline{Z'(m)}_i\right)^2\right|\right\}
$$

$$
=: \mathbb{V}^{(1)}_n(m) + \mathbb{V}^{(2)}_n(m),
$$

it suffices to prove

$$
\lim_{m\to\infty}\limsup_{n\to\infty} P_n\left(\mathbb{V}^{(l)}_n(m) > \eta\right) = 0 \tag{44}
$$

for $l = 1, 2$.

We have

$$
E\left[\left|\overline{Z(m)}_i\left(\overline{B'(m)}_i\right)^2\right|\,\big|\mathcal{F}_{t^n_{i-1}}\right] \lesssim_m (k_n\bar{r}_n)^2 E\left[\left|\overline{Z(m)}_i\right|\,\big|\mathcal{F}_{t^n_{i-1}}\right] \lesssim_m (k_n\bar{r}_n)^{5/2}
$$

by [SA2], Eqs. (23) and (27). Therefore, Eq. (44) holds true for $l = 1$ by the Lenglart inequality, Eq. (29) and the fact that $\xi > 4/5$.

Now we prove Eq. (44) for $l = 2$. We start with introducing a further localization procedure for the observation times. For each $K \in \mathbb{N}$, we define the sequence $(t^n_i(K))^\infty_{i=-1}$ inductively by $t^n_{-1}(K) = 0$ and

$$
t^n_i(K) = \begin{cases} t^n_i, & \text{if } \max_{j=1,\dots,i} E\left[\Delta_n^{-1}|I_j|\big|\mathcal{F}_{t^n_{j-1}}\right] \le K \text{ and } \Delta_n N^n_{t^n_{i-1}} \le K, \\ t^n_{i-1}(K) + \Delta_n, & \text{otherwise.} \end{cases}
$$

By construction $t_i^n(K)$ is an $(\mathcal{F}_t)$-stopping time for every i. Moreover, by [A1] and Eq. (29) we have $\sup_n P(t_i^n \neq t_i^n(K)$ for some $i \in \{0, 1, \ldots, N_T^n\}) \to 0$ as $K \to \infty$. Consequently, it suffices to show that

$$\lim_{m \to \infty} \limsup_{n \to \infty} P_n\left(\mathbb{V}_n^{(2)}(m) > \eta, t_i^n = t_i^n(K) \text{ for any } i \in \{0, 1, \ldots, N_T^n\}\right) = 0 \quad (45)$$

for any fixed $K \in \mathbb{N}$.

Set $\widetilde{I}_i = [t_{i-1}^n(K), t_i^n(K))$ and define the process $\widetilde{g}_i^n$ by $\widetilde{g}_i^n(s) = \sum_{p=1}^{k_n-1} g_p^n 1_{\widetilde{I}_{i+p}}(s)$. For any semimartingale V, we define the process $\widetilde{V}_{i,t}$ by $\widetilde{V}_{i,t} = \int_0^t \widetilde{g}_i^n(s-) dV_s$. Then, to prove Eq. (45) it is enough to show that

$$\lim_m \limsup_n E\left[\frac{\Delta_n^{-1/4}}{k_n} \sum_{i=0}^{\widetilde{N}_T^n+1} \left|\widetilde{Z(m)}_{i,t_{i+k_n-1}^n(K)} \left(\widetilde{Z'(m)}_{i,t_{i+k_n-1}^n(K)}\right)^2\right|\right] = 0, \quad (46)$$

where $\widetilde{N}_T^n = \max\{i : t_i^n(K) \leq T\}$. Note that $\widetilde{N}_T^n \leq (K+T)\Delta_n^{-1}$ by construction.

Set $\widetilde{I}_i^+ = [t_{i-1}^n(K), t_{i+k_n-1}^n(K))$. To prove Eq. (46), we consider the following decomposition, which is obtained by applying integration by parts repeatedly (note that $[Z(m), Z'(m)] \equiv 0$ by construction):

$$\widetilde{Z(m)}_{i,t_{i+k_n-1}^n(K)} \left(\widetilde{Z'(m)}_{i,t_{i+k_n-1}^n(K)}\right)^2 = \int_{\widetilde{I}_i^+} \left(\widetilde{Z'(m)}_{i,s-}\right)^2 d\widetilde{Z(m)}_{i,s}$$

$$+ 2\int_{\widetilde{I}_i^+} \widetilde{Z(m)}_{i,s-} \widetilde{Z'(m)}_{i,s-} d\widetilde{Z'(m)}_{i,s}$$

$$+ \int_{\widetilde{I}_i^+} \widetilde{Z(m)}_{i,s-} d[\widetilde{Z'(m)}_{i,\cdot}]_s$$

$$=: \mathbf{A}(m)_i^{n,1} + 2\mathbf{A}(m)_i^{n,2} + \mathbf{A}(m)_i^{n,3}.$$

Then, it is enough to show that

$$\lim_m \limsup_n E\left[\frac{\Delta_n^{-1/4}}{k_n} \sum_{i=0}^{\widetilde{N}_T^n+1} \left|\mathbf{A}(m)_i^{n,l}\right|\right] = 0 \quad (47)$$

for every $l = 1, 2, 3$. First, we consider the case $l = 1$. Lemma 1 and [SA2] yield

$$E\left[\left|\mathbf{A}(m)_i^{n,1}\right| \bigg| \mathcal{F}_{t_{i-1}^n(K)}\right] \lesssim \sqrt{\gamma_m} E\left[\sup_{s \in \widetilde{I}_i^+} \left(\widetilde{Z'(m)}_{i,s}\right)^2 \sqrt{|\widetilde{I}_i^+|} \bigg| \mathcal{F}_{t_{i-1}^n(K)}\right]$$

$$\lesssim \sqrt{\gamma_m} E\left[\sup_{s \in \widetilde{I}_i^+} \left(\widetilde{Z'(m)}_{i,s}\right)^2 \sqrt{\left|\sum_{p=1}^{k_n-1} \left(|\widetilde{I}_{i+p}| - E\left[|\widetilde{I}_{i+p}| \big| \mathcal{F}_{T_{i+p-1}^K}\right]\right)\right|} \bigg| \mathcal{F}_{t_{i-1}^n(K)}\right]$$

$$+ \sqrt{\gamma_m} E \left[\sup_{s \in \widetilde{I}_i^+} \left(\widetilde{Z'(m)}_{i,s} \right)^2 \sqrt{\sum_{p=1}^{k_n-1} E \left[|\widetilde{I}_{i+p}| \big| \mathcal{F}_{T_{i+p-1}^K} \right] \big| \mathcal{F}_{t_{i-1}^n}(K)} \right]$$

$$=: \sqrt{\gamma_m} \left(\mathbf{B}(m)_i^{n,1} + \mathbf{B}(m)_i^{n,2} \right).$$

It suffices to prove

$$\lim_m \limsup_n E \left[\frac{\Delta_n^{-1/4}}{k_n} \sum_{i=0}^{\widetilde{N}_T^n+1} \sqrt{\gamma_m} \mathbf{B}(m)_i^{n,j} \right] = 0 \tag{48}$$

for $j = 1, 2$. By the Hölder and BDG inequalities and Eq. (23), we have

$$\mathbf{B}(m)_i^{n,1} \lesssim (k_n \bar{r}_n^2)^{1/4} \left\{ E \left[\sup_{s \in \widetilde{I}_i^+} \left(\widetilde{Z'(m)}_{i,s} \right)^{2p} \big| \mathcal{F}_{t_{i-1}^n}(K) \right] \right\}^{1/p}$$

for any $p \in (1, 2]$. Therefore, the Novikov inequality (Theorem 1 of Novikov (1975)) implies that

$$\mathbf{B}(m)_i^{n,1} \lesssim (k_n \bar{r}_n^2)^{1/4} (k_n \bar{r}_n)^{1/p}$$

for any $p \in (1, 2]$. Now, we can take $p \in (1, \frac{\xi - \frac{1}{2}}{\frac{7}{8} - \frac{\xi}{2}})$ because $\xi > \frac{11}{12}$, hence the above inequality yields Eq. (48) for $j = 1$. On the other hand, the construction of $(t_i^n(K))$ and the Doob inequality imply that

$$\mathbf{B}(m)_i^{n,2} \lesssim \sqrt{k_n \Delta_n} E \left[\sup_{s \in \widetilde{I}_i^+} \left(\widetilde{Z'(m)}_{i,s} \right)^2 \big| \mathcal{F}_{t_{i-1}^n}(K) \right]$$

$$\lesssim \sqrt{k_n \Delta_n} E \left[|\widetilde{I}_i^+| \big| \mathcal{F}_{t_{i-1}^n}(K) \right] \lesssim (k_n \Delta_n)^{3/2},$$

hence Eq. (48) also holds true for $j = 2$.

Next consider the case $l = 2$. Lemma 1, [SA2] and the Schwarz inequality yield

$$E \left[\left| \mathbf{A}(m)_i^{n,2} \right| \big| \mathcal{F}_{t_{i-1}^n}(K) \right] \lesssim E \left[\sup_{s \in \widetilde{I}_i^+} \left| \widetilde{Z(m)}_{i,s} \widetilde{Z'(m)}_{i,s} \right| \sqrt{|\widetilde{I}_i^+|} \big| \mathcal{F}_{t_{i-1}^n}(K) \right]$$

$$\lesssim \sqrt{E \left[\sup_{s \in \widetilde{I}_i^+} \left| \widetilde{Z(m)}_{i,s} \widetilde{Z'(m)}_{i,s} \right|^2 \big| \mathcal{F}_{t_{i-1}^n}(K) \right] E \left[|\widetilde{I}_i^+| \big| \mathcal{F}_{t_{i-1}^n}(K) \right]}.$$

Noting $[Z(m), Z'(m)] \equiv 0$ by construction, we obtain the following identity for $s \in \widetilde{I}_i^+$ by applying integration by parts:

$$\widetilde{Z(m)}_{i,s} \widetilde{Z'(m)}_{i,s} = \int_{t_{i-1}^n(K)}^s \widetilde{Z(m)}_{i,u-} d\widetilde{Z'(m)}_{i,u} + \int_{t_{i-1}^n(K)}^s \widetilde{Z'(m)}_{i,u-} d\widetilde{Z(m)}_{i,u}.$$

Therefore, the Doob inequality yields

$$E\left[\sup_{s \in \widetilde{I}_i^+} \left|\widetilde{Z(m)}_{i,s} \widetilde{Z'(m)}_{i,s}\right|^2 |\mathcal{F}_{t_{i-1}^n(K)}\right]$$

$$\lesssim E\left[\left(\sup_{s \in \widetilde{I}_i^+} \widetilde{Z(m)}_{i,s}^2 + \overline{\gamma}_m \sup_{s \in \widetilde{I}_i^+} \widetilde{Z'(m)}_{i,s}^2\right) |\widetilde{I}_i^+| \,|\mathcal{F}_{t_{i-1}^n(K)}\right].$$

Hence, by an analogous argument to the proof of the case $l = 1$ we obtain

$$E\left[\sup_{s \in \widetilde{I}_i^+} \left|\widetilde{Z(m)}_{i,s} \widetilde{Z'(m)}_{i,s}\right|^2 |\mathcal{F}_{t_{i-1}^n(K)}\right] \lesssim \overline{\gamma}_m (k_n \Delta_n)^2.$$

Consequently, we conclude that

$$E\left[\left|\mathbf{A}(m)_i^{n,2}\right| |\mathcal{F}_{t_{i-1}^n(K)}\right] \lesssim \sqrt{\overline{\gamma}_m}(k_n \Delta_n)^{3/2},$$

and thus we obtain Eq. (47) for $l = 2$.

Finally consider the case $l = 3$. Since [SA2] yields

$$E\left[\left|\mathbf{A}(m)_i^{n,3}\right| |\mathcal{F}_{t_{i-1}^n(K)}\right] \lesssim E\left[\sup_{s \in \widetilde{I}_i^+} \left|\widetilde{Z(m)}_{i,s}\right| |\widetilde{I}_i^+| \,|\mathcal{F}_{t_{i-1}^n(K)}\right],$$

we can again apply an analogous argument to the proof of the case $l = 1$, and thus Eq. (47) holds true for $l = 3$. This completes the proof. $\square$

Proof of (31) We decompose the target quantity as

$$\mathbb{III}_n(m) = \frac{\Delta_n^{-1/4}}{\psi_3 k_n} \sum_{i=0}^{N_T^n - k_n + 1} \left(\overline{C(m)}_i\right)^3 + \frac{3\Delta_n^{-1/4}}{\psi_3 k_n} \sum_{i=0}^{N_T^n - k_n + 1} \left(\overline{C(m)}_i\right)^2 \overline{Z(m)}_i$$

$$+ \frac{3\Delta_n^{-1/4}}{\psi_3 k_n} \sum_{i=0}^{N_T^n - k_n + 1} \overline{C(m)}_i \left(\overline{Z(m)}_i\right)^2$$

$$+ \Delta_n^{-1/4} \left\{ \frac{1}{\psi_3 k_n} \sum_{i=0}^{N_T^n - k_n + 1} \left(\overline{Z(m)}_i \right)^3 - \sum_{0 \leq s \leq T} (\Delta X(m)_s)^3 \right\}$$

$$=: \mathbb{II}_n^{(1)}(m) + \mathbb{II}_n^{(2)}(m) + \mathbb{II}_n^{(3)}(m) + \mathbb{II}_n^{(4)}(m).$$

It suffices to prove

$$\lim_{m \to \infty} \limsup_{n \to \infty} P_n \left(\left| \mathbb{II}_n^{(l)}(m) \right| > \eta \right) = 0 \tag{49}$$

for every $l = 1, 2, 3, 4$.

Since Proposition 4.1 of Koike (2017) yields $\mathbb{II}_n^{(1)}(m) \to^P 0$ as $n \to \infty$ for every m, Eq. (49) holds true for $l = 1$. Moreover, we can prove Eq. (49) for $l = 2, 3$ analogously to the proof of Eq. (34). So it remains to prove Eq. (49) for $l = 4$. Applying integration by parts repeatedly, we can decompose the target quantity as

$$\mathbb{II}_n^{(4)}(m) = \frac{3 \Delta_n^{-1/4}}{\psi_3 k_n} \sum_{i=0}^{N_T^n - k_n + 1} \int_{\overline{I}_i} \left(\overline{Z(m)}_{i,s-} \right)^2 d\overline{Z(m)}_{i,s}$$

$$+ \frac{3 \Delta_n^{-1/4}}{\psi_3 k_n} \sum_{i=0}^{N_T^n - k_n + 1} \int_{\overline{I}_i} \overline{Z(m)}_{i,s-} d[\overline{Z(m)}_{i,\cdot}]_s$$

$$+ \Delta_n^{-1/4} \left\{ \frac{1}{\psi_3 k_n} \sum_{i=0}^{N_T^n - k_n + 1} \left[\overline{Z(m)}_{i,\cdot}, [\overline{Z(m)}_{i,\cdot}] \right]_{t_{i+k_n-1}^n} - \sum_{0 \leq s \leq T} (\Delta X(m)_s)^3 \right\}$$

$$=: \mathbb{A}_n^{(1)}(m) + \mathbb{A}_n^{(2)}(m) + \mathbb{A}_n^{(3)}(m),$$

hence it is enough to prove

$$\lim_{m \to \infty} \limsup_{n \to \infty} P_n \left(\left| \mathbb{A}_n^{(l)}(m) \right| > \eta \right) = 0 \tag{50}$$

for every $l = 1, 2, 3$. For $l = 1, 2$, Eq. (50) can be shown analogously to the proof of Eq. (44) for $l = 2$. On the other hand, since we have $[\overline{Z(m)}_{i,\cdot}]_s = \sum_{p=1}^{k_n-1} (g_p^n)^2 \sum_{t_{i+p-1}^n < u \leq s} (\Delta X(m)_u)^2$ for $s \in \overline{I}_i$ and $\Delta Z(m) = \Delta X(m)$, we obtain

$$\mathbb{A}_n^{(3)}(m) = \Delta_n^{-1/4} \left\{ \frac{1}{\psi_3 k_n} \sum_{p=1}^{N_T^n} \left(\sum_{i=(p-k_n+1)_+}^{(p-1) \wedge (N_T^n - k_n + 1)} (g_{p-i}^n)^3 \right) \right.$$

$$\left. \sum_{t_{p-1}^n < s \leq t_p^n} (\Delta X(m)_s)^3 - \sum_{0 \leq s \leq T} (\Delta X(m)_s)^3 \right\}.$$

Now, since we have $\sum_{(t-h)_+ < s \le t} |\Delta X(m)_s|^3 = O_p(h)$ as $h \downarrow 0$ by [SA2], we can deduce that $\mathbb{A}_n^{(3)}(m) \to 0$ as $n \to \infty$ for every m, so Eq. (50) holds true for $l = 3$. $\square$

References

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

Amaya, D., Christoffersen, P., Jacobs, K., Vasquez, A. (2015). Does realized skewness predict the cross-section of equity returns? *Journal of Financial Economics*, *118*, 135–167.

Andersen, T. G., Bollerslev, T., Meddahi, N. (2005). Correcting the errors: Volatility forecast evaluation using high-frequency data and realized volatilities. *Econometrica*, *73*, 279–296.

Bakshi, G., Kapadia, N., Madan, D. (2003). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies*, *16*, 101–143.

Bandi, F. M., Russell, J. R. (2006). Separating microstructure noise from volatility. *Journal of Financial Economics*, *79*, 655–692.

Bibinger, M., Vetter, M. (2015). Estimating the quadratic covariation of an asynchronously observed semi-martingale with jumps. *Annals of the Institute of Statistical Mathematics*, *67*, 707–743.

Billingsley, P. (1999). *Convergence of probability measures* (2nd ed.). New York: Wiley.

Friend, I., Westerfield, R. (1980). Co-skewness and capital asset pricing. *The Journal of Finance*, *35*, 897–913.

Fukasawa, M. (2010). Realized volatility with stochastic sampling. *Stochastic Processes and Their Applications*, *120*, 829–852.

Harvey, C. R., Siddique, A. (1999). Autoregressive conditional skewness. *Journal of Financial and Quantitative Analysis*, *34*, 465–487.

Harvey, C. R., Siddique, A. (2000). Conditional skewness in asset pricing tests. *The Journal of Finance*, *55*, 1263–1295.

Jacod, J. (2008). Asymptotic properties of realized power variations and related functionals of semimartingales. *Stochastic Processes and Their Applications*, *118*, 517–559.

Jacod, J., Protter, P. (1998). Asymptotic error distributions for the Euler method for stochastic differential equations. *Annals of Probability*, *26*, 267–307.

Jacod, J., Protter, P. (2012). *Discretization of processes*. New York: Springer.

Jacod, J., Shiryaev, A. N. (2003). *Limit theorems for stochastic processes*. New York: Springer.

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

Kinnebrock, S., Podolskij, M. (2008). A note on the central limit theorem for bipower variation of general functions. *Stochastic Processes and Their Applications*, *118*, 1056–1070.

Koike, Y. (2016). Quadratic covariation estimation of an irregularly observed semimartingale with jumps and noise. *Bernoulli*, *22*, 1894–1936.

Koike, Y. (2017). Time endogeneity and an optimal weight function in pre-averaging covariance estimation. *Statistical Inference for Stochastic Processes*, *20*, 15–56.

Kozhan, R., Neuberger, A., Schneider, P. (2013). The skew risk premium in the equity index market. *Review of Financial Studies*, *26*, 2174–2203.

Li, Y., Mykland, P. A., Renault, E., Zhang, L., Zheng, X. (2014). Realized volatility when sampling times are possibly endogenous. *Econometric Theory*, *30*, 580–605.

Liptser, R. S., Shiryayev, A. N. (1989). *Theory of martingales*. New York: Springer.

Liu, Z., Wang, K., Liu, J. (2014) . Realized skewness at high frequency and the link to a conditional market premium. SSRN: http://ssrn.com/abstract=2224216.

Martellini, L., Ziemann, V. (2009). Improved estimates of higher-order comoments and implications for portfolio selection. *Review of Financial Studies*, *23*(4), 1465–1502.

Martin, O., Vetter, M. (2016) Testing for simultaneous jumps in case of asynchronous observations. arXiv:1606.07246.

Mitton, T., Vorkink, K. (2007). Equilibrium underdiversification and the preference for skewness. *Review of Financial Studies*, *20*(4), 1255–1288.

Neuberger, A. (2012). Realized skewness. *Review of Financial Studies*, *25*, 3423–3455.

Novikov, A. (1975). On discontinuous martingales. *Theory of Probability and Its Applications*, *20*, 11–26.

Podolskij, M., Vetter, M. (2009). Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Bernoulli*, *15*, 634–658.

Podolskij, M., Vetter, M. (2010). Understanding limit theorems for semimartingales: A short survey. *Statistica Neerlandica*, *64*, 329–351.

Vetter, M., Zwingmann, T. (2017). A note on central limit theorems for quadratic variation in case of endogenous observation times. *Electronic Journal of Statistics*, *11*, 963–980.