

CAMBRIDGE WORKING PAPERS IN ECONOMICS
CAMBRIDGE-INET WORKING PAPERS

Robust Estimation of Integrated and Spot Volatility

Z. Merrick	Oliver
Li	Linton
Chinese University of Hong Kong	University of Cambridge

Abstract

We introduce a new method to estimate the integrated volatility (IV) and the spot volatility (SV) based on noisy high-frequency data. Our method employs the ReMeDI approach introduced by Li and Linton (2022a) to estimate the moments of microstructure noise and thereby eliminate their influence, and the pre-averaging method to target the volatility parameter. The method is robust: it can be applied when the efficient price exhibits stochastic volatility and jumps, the observation times are random, and the noise process is nonstationary, autocorrelated, asymptotically vanishing and dependent on the efficient price. We derive the limit distributions for the proposed estimators under the infill asymptotics in a general setting. Our extensive simulation studies demonstrate the robustness, accuracy and computational efficiency of our estimators compared to several alternative estimators recently proposed in the literature. Empirically, we show that neglecting the complexities of noise and the random observation times yields substantial biases in volatility estimation and may lead to a different intraday volatility pattern.

Reference Details

2115	Cambridge Working Papers in Economics
2021/08	Cambridge-INET Working Paper Series

Published	23 February 2021
Revised	9 August 2022

Websites	www.econ.cam.ac.uk/cwpe
	www.inet.econ.cam.ac.uk/working-papers

Robust Estimation of Integrated and Spot Volatility

Z. Merrick Li* Oliver Linton†

August 9, 2022

Abstract

We introduce a new method to estimate the integrated volatility (IV) and the spot volatility (SV) based on noisy high-frequency data. Our method employs the ReMeDI approach introduced by [Li and Linton \(2022a\)](#) to estimate the moments of microstructure noise and thereby eliminate their influence, and the pre-averaging method to target the volatility parameter. The method is robust: it can be applied when the efficient price exhibits stochastic volatility and jumps, the observation times are random, and the noise process is nonstationary, autocorrelated, asymptotically vanishing and dependent on the efficient price. We derive the limit distributions for the proposed estimators under the infill asymptotics in a general setting. Our extensive simulation studies demonstrate the robustness, accuracy and computational efficiency of our estimators compared to several alternative estimators recently proposed in the literature. Empirically, we show that neglecting the complexities of noise and the random observation times yields substantial biases in volatility estimation and may lead to a different intraday volatility pattern.

1 Introduction

The past two decades or so have seen the emergence of high-frequency trading (HFT) that operates at astonishing time scales. HFT yields a vast quantity of transaction data, which is in principle good for the accurate measurement of economic parameters or processes, such as the *integrated volatility* (IV), or the entire volatility process, i.e., the *spot volatility* (SV) of financial returns. On the other hand, this data can be quite noisy, and one needs a coherent strategy for dealing with this noise. This paper introduces a robust estimation method that accounts for many salient features of high-frequency data. In particular, it can be directly applied to prices that are available at the *highest* possible frequency and takes account of microstructure noise (measurement error) of a general form.

The estimation of IV becomes straightforward if the stock price is sampled from a semimartingale—the sum of the squared log-returns, usually called the *realized volatility* (RV),

*Corresponding author. Department of Economics, the Chinese University of Hong Kong, Esther Lee Building, Shatin, New Territories, Hong Kong SAR. Email: merrickli@cuhk.edu.hk

†Faculty of Economics, University of Cambridge, Austin Robinson Building, Sidgwick Avenue, Cambridge, CB3 9DD, United Kingdom. Email: obl20@cam.ac.uk.

provides a consistent estimator of IV. This result dates back to [Jacod \(2018\)](#)¹ and [Jacod and Protter \(1998\)](#), and it is introduced to econometrics by [Andersen et al. \(2003\)](#). As a consequence, the estimation of SV becomes simple as well: the difference of the IV estimates scaled by the difference of the observation times provides a simple yet efficient estimator of SV. In reality, however, high-frequency prices are often perceived of as being “noisy” in a sense that the observed price process deviates from the semimartingale “efficient price”². The deviation, or “noise” reflects some market imperfectness, such as transaction costs, the presence of minimal ticks, informational effects, inventory risk, etc. The presence of microstructure noise motivates the following model

$$Y = X + \varepsilon, \tag{1}$$

where X is the semimartingale efficient price and ε represents the aforementioned market microstructure effects. Since both X and ε are latent, the inference and estimation of either component become challenging.

Several “de-noise” methods have been proposed in order to make statistical inference on the parameters of X . Without being exhaustive, here we mention the methods of Two-Scale (TSRV) and Multi-Scale Realized Volatility (MSRV) ([Zhang et al., 2005](#); [Zhang, 2006](#)), the maximum likelihood estimators ([Aït-Sahalia et al., 2005](#); [Xiu, 2010](#); [Shephard and Xiu, 2017](#)), the pre-averaging method ([Podolskij and Vetter, 2009](#); [Jacod et al., 2009, 2010](#); [Li, 2013](#)), and the realized kernel ([Hansen and Lunde, 2006](#); [Barndorff-Nielsen et al., 2008](#)). Intuitively, the statistical assumptions imposed on ε will affect the estimation and inference of the parameters of X , since both are latent and only their sum is observable. Most papers quoted above have very restrictive assumptions on microstructure noise, often assuming it is an i.i.d. process.

However, such simple assumptions are often contradicted by empirical evidence, theoretical motivations and many practical concerns about the characteristics of high-frequency data. Empirical studies ([Chan and Lakonishok, 1995](#); [Hasbrouck, 1993](#); [Madhavan et al., 1997](#); [Wood et al., 1985](#)) reveal that microstructure noise may have prominent intraday patterns, typically a U-shape or reverse J-shape. There is also a large theoretical literature seeking to characterize the economic mechanisms that govern the dynamic properties of microstructure noise, including the modelling of the order flow reversal due to a market maker’s risk aversion ([Grossman and Miller, 1988](#); [Campbell et al., 1993](#)) or inventory controls ([Ho and Stoll, 1981](#); [Hendershott and Menkveld, 2014](#)), and the presence of inattentive (or infrequent) traders ([Bogouslavsky, 2016](#); [Hendershott et al., 2022](#)). However, high-frequency data, in particular tick-by-tick data, has several prominent features that confront researchers. First, the transaction times are random, thus the prices are irregularly spaced. Second, transactions are often clustered on one side of the market as a consequence of order splitting or the execution of limit orders ([Parlour, 1998](#)). Thus, microstructure noise could be highly autocorrelated. Third, the size of high-frequency

¹The classic paper was written in 1994. It gets published in the *Journal of Financial Econometrics* recently.

²It is well known that the stock price follows a semimartingale if no arbitrage is allowed, see [Delbaen and Schachermayer \(1994\)](#).

data is huge. For example, the quote data sizes from the Trade and Quote (TAQ) database are close to levels of several terabytes per month after 2007. These considerations call for a flexible model and a robust estimation strategy to deal with microstructure noise.

This paper introduces a robust estimation strategy of IV and SV. Our theoretical setup is based on two seminal works by [Jacod et al. \(2017, 2019\)](#), where the observation times are random and possibly endogenous, the noise is serially dependent, nonstationary and could be dependent of the efficient price. While [Jacod et al. \(2017, 2019\)](#) only consider noise of order $O_p(1)$, we allow noise to be shrinking as the data frequency increases. The “small noise” asymptotics are considered by [Kalnina and Linton \(2008\)](#), [Da and Xiu \(2021b\)](#), see also Chapter 7 of [Aït-Sahalia and Jacod \(2014\)](#) and Chapter 16 of [Jacod and Protter \(2011\)](#). We develop new volatility estimators using the ReMeDI estimators ([Li and Linton, 2022a](#)) of the moments of noise to correct the bias of a pre-averaging type estimator. The bias term consists of the long-run variance (LRV) of noise. It turns out that the ReMeDI method is flexible in estimating both the spot and the integrated LRV of noise, thus providing a correction method that works in the estimation of both IV and SV. We derive the limit distributions under the infill asymptotics for the proposed estimators, and we demonstrate the inconsistency if one fails to account for the serial dependence in noise and randomness in the observation scheme. What is particularly intriguing is the role played by the random observation scheme in SV estimation: The spot estimator may even be inconsistent if the irregular observation scheme is misspecified as a regular one.³

The estimators inherit the excellent finite sample properties of the ReMeDI estimators. Our extensive simulation studies show that compared to several alternative estimators, our estimators perform very well in finite samples where the data generating process follows different specifications. Moreover, the estimators are also quite robust to the choices of tuning parameters and are computationally very efficient.⁴ We apply our estimators to estimate the IV and SV of two individual stocks, and we find that substantial biases in SV and IV estimation will emerge if one neglects the complexities of noise and randomness of the observation times. In particular, we find a (moderate) U-shaped intraday volatility. But the U-shape disappears if the randomness and irregularities of the observation times are not explicitly treated.

Many recent papers study the estimation of IV with a general framework for microstructure noise. [Kalnina and Linton \(2008\)](#) consider microstructure noise with a time-varying scale; [Aït-Sahalia et al. \(2011\)](#) show that the TSRV and the MSRV remain valid when the noise is autocorrelated; [Hautsch and Podolskij \(2013\)](#); [Christensen et al. \(2013\)](#) study q -dependent noise; [Li et al. \(2020\)](#) use a variant of RV to estimate the second moments of serially dependent noise and develop a consistent estimator of the IV. [Ikeda \(2015, 2016\)](#) propose the two-scale realized kernel (TSRK) to estimate IV in the presence of serially correlated noise and random durations between observations. [Varneskov \(2016, 2017\)](#) employ the flat-top realized kernel

³This is in contrast with the IV estimation: Many IV estimators preserve the consistency in the presence of random and irregular sampling times, though the limiting variance will be different, see Remark 4.2 in [Li et al. \(2020\)](#).

⁴For example, to perform the same estimation, our estimator usually uses less than 5% and 0.5% of the CPU time used by the estimators in [Jacod et al. \(2019\)](#) and [Da and Xiu \(2021b\)](#), respectively.

(FRK) to estimate IV, in a setting that is similar to [Jacod et al. \(2019\)](#). Both the TSRK and FRK allow for correlations between the efficient price and noise.⁵ Moreover, [Varneskov \(2016\)](#) also proposes to use TSRK to correct the bias of a pre-averaging type estimator. Compared to the above estimators of IV, the key advantage of our IV estimator is that it is still valid when the noise effect is asymptotically diminishing. As a consequence, it converges faster than the usual rate $n^{1/4}$ in the presence of shrinking noise. This is an advantage in dealing with recent high-frequency data where it is found that the scale of noise is much smaller, see, e.g., [Da and Xiu \(2021a\)](#).

The most two recent papers closest in spirit are [Jacod et al. \(2019\)](#) and [Da and Xiu \(2021b\)](#). Specifically, we adopt and generalize the framework of [Jacod et al. \(2019\)](#), and both our estimator and the one proposed in [Jacod et al. \(2019\)](#) are pre-averaging type estimators. Nevertheless, there are several key differences. First, we use different technologies to remove the noise effect: while we use the ReMeDI approach ([Li and Linton, 2022a](#)), [Jacod et al. \(2019\)](#) employ the local averaging (LA) method ([Jacod et al., 2017](#)). Second, we develop a different asymptotic variance estimator⁶ that yields accurate approximations in finite samples, see our numerical evidence in Section 6.5. Third, theoretically and empirically, our estimators work well with small noise. Lastly, our estimator is more robust to the choice of tuning parameters and model specifications. In terms of finite sample performances, our estimator is comparable to the Quasi-Maximum-Likelihood-Estimator (QMLE) ([Da and Xiu, 2021b](#)). Both estimators give very accurate estimates of IVs under each specification in the simulation studies, especially when the noise is relatively small. The most significant advantage of our estimator is that it is computationally very efficient. Moreover, our asymptotic variance estimator has a faster convergence rate. We should also mention that the QMLE has other advantages over our approach: it is \sqrt{n} -consistent when noise is absent. Moreover, it always yields a positive estimate of IV.

We also contribute to the literature on SV estimation. A partial list of related papers are [Foster and Nelson \(1996\)](#), [Fan and Wang \(2008\)](#), [Kalnina and Linton \(2008\)](#), [Malliavin and Mancino \(2009\)](#), [Kristensen \(2010\)](#), [Jacod and Rosenbaum \(2013\)](#), [Zu and Boswijk \(2014\)](#), see also some recent papers, e.g., [Bibinger and Winkelmann \(2018\)](#), [Bollerslev et al. \(2021\)](#), [Li et al. \(2022\)](#). The spot estimation strategies are also used in other problems rather than the SV estimation, see, e.g., [Aït-Sahalia and Jacod \(2009\)](#), [Jacod and Todorov \(2009\)](#), [Andersen et al. \(2021\)](#). To the best of our knowledge, our SV estimator is the first one that works with a serially dependent noise process and a random observation scheme, a framework that well suits the massive *tick data* of high-frequency asset prices. Thus, the proposed SV estimator can be directly applied to tick data. While the richness of the datasets certainly improves the accuracy of the spot estimation, it poses further challenges. In particular, we explain and demonstrate that the SV estimator will be inconsistent if one ignores the randomness

⁵The pre-averaging estimators usually assume the noise is uncorrelated with the efficient price, although some degree of higher order dependence is allowed, see, e.g., [Jacod et al. \(2009\)](#), [Jacod et al. \(2019\)](#).

⁶We use the ReMeDI approach to estimate the moments of noise that appear in the asymptotic variance. Moreover, we also avoid a direct estimation of higher-order moments of noise to reduce estimation errors.

of the sampling scheme. Our estimator is carefully designed to deal with such technicalities underpinning these datasets. Therefore, it is particularly useful in the context of SV estimation.

The rest of the paper proceeds as follows. Section 2 discusses the model settings. Section 3 introduces the new IV and SV estimators. Section 4 presents the limiting theorems of the estimators, and Section 5 discusses the tuning parameters selection and some implementation issues. Sections 6 and 7 present the simulation and empirical studies. Section 8 concludes the paper. Some additional simulation studies, as well as mathematical proofs, are in the supplementary materials of this paper (Li and Linton, 2022b).

2 Model Setting

We follow the general setup in Jacod et al. (2019) that allows for a general Itô semimartingale efficient price process, a nonstationary and serially dependent microstructure noise process and a random observation scheme. We also extend their setting by allowing for asymptotically vanishing noise.

Let Z be a generic Itô semimartingale that is defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ with the Grigelionis representation:

$$Z_t := Z_0 + \int_0^t b_s^Z ds + \int_0^t \sigma_s^Z dW_s^Z + \left(\delta^Z \mathbf{1}_{\{|\delta^Z| \leq 1\}} \right) \star (\mu - \nu)_t + \left(\delta^Z \mathbf{1}_{\{|\delta^Z| > 1\}} \right) \star \mu_t, \quad (2)$$

where W^Z, μ are Wiener process and a Poisson random measure on $\mathbb{R}_+ \times E$ with (E, \mathcal{E}) a measurable Polish space on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ and the predictable compensator of μ is $\nu(ds, dz) = ds \otimes \lambda(dz)$ for some given σ -finite measure on (E, \mathcal{E}) , see Jacod and Shiryaev (2003) for detailed introduction of the last two integrals. The processes b^Z, σ^Z are optional. The function δ^Z on $\Omega \times \mathbb{R}_+ \times E$ is predictable. For any Itô semimartingale Z , we could impose an assumption that depends on some $r \in [0, 2]$:

Assumption (H-r). *There is a sequence of stopping times $\{\tau_n\}$, a sequence of reals $\{w_n\}$, and for each n a deterministic nonnegative function Γ_n^Z on E satisfying*

$$|b^Z(\omega)| \leq w_n, \quad |\sigma_t^Z(\omega)| < w_n, \quad |\delta^Z(\omega, t, z)|^r \wedge 1 \leq \Gamma_n^Z(z)$$

for all (ω, t, z) satisfying $t \leq \tau_n(\omega)$.

2.1 The efficient price

The efficient price is an Itô semimartingale and its Grigelionis representation is as follows:

$$X_t := X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + (\delta \mathbf{1}_{\{|\delta| \leq 1\}}) \star (\mu - \nu)_t + (\delta \mathbf{1}_{\{|\delta| > 1\}}) \star \mu_t. \quad (3)$$

We further assume the following regularity conditions.

Assumption (H-X-r). The coefficients of the efficient price X satisfy Assumption (H-r) with $r \in [0, 1)$; the processes b, σ are Itô semimartingale whose coefficients satisfy Assumption (H-r) with $r = 2$.

Remark 2.1. The parameter r in Assumption (H-X-r) restricts the degree of jump activities: finite activity for $r = 0$, summable jumps on each finite interval for $r \leq 1$ and no restrictions for $r = 2$. We will separate jumps from the volatility parameters of the efficient price X using a truncation method, and we need to restrict $r \in [0, 1)$ for X .

This is a general class of processes that allows for stochastic volatility and jumps, and it includes most models used in finance to characterize security prices. The parameters of interest in this paper are the Integrated Volatility (IV) and Spot Volatility (SV) of the efficient price:

$$C_t := \int_0^t \sigma_s^2 ds, \quad c_t := \sigma_t^2.$$

2.2 The random observation scheme

Now we describe the general observation scheme. For each positive integer n , let $\{T(n, i) : i \in \mathbb{N}^*\}$ be the set of observation times, which is a sequence of strictly increasing finite stopping times with $T(n, 0) = 0, T(n, i) \rightarrow \infty$ as $i \rightarrow \infty$, where \mathbb{N}^* is the set of nonnegative integers. We denote the (random) number of observations upon time t and the spacing of successive observations by

$$N_t^n := \sum_{i \geq 1} \mathbf{1}_{\{T(n, i) \leq t\}}, \quad \Delta(n, i) := T(n, i) - T(n, i - 1). \quad (4)$$

In the sequel, for any process V , we denote $V_i^n := V_{T(n, i)}$, $\Delta_i^n V := V_i^n - V_{i-1}^n$, $\mathcal{F}_i^n := \mathcal{F}_{T(n, i)}$.

Let Δ_n be the time lag between observations in a regular sampling scheme that satisfies $\Delta_n \rightarrow 0$ as $n \rightarrow \infty$. An empirical proxy of Δ_n is the average duration of observations. Let α be another nonnegative Itô semimartingale. It serves as the ‘‘observation density’’ process that relates the real observation scheme to the (possibly latent) regular observation scheme. Specifically, $\alpha_{T(n, i-1)} \Delta(n, i) \approx \Delta_n$ conditional on the information set upon time $T(n, i - 1)$. Formally, we assume

Assumption (O- ρ, ρ'). α is an Itô semimartingale satisfying Assumption (H-r) with $r = 2$, and $\alpha_t > 0$, $\alpha_{t-} > 0$ for all $t > 0$. We further assume

(i) $\Delta_n N_t^n \rightarrow A_t := \int_0^t \alpha_s ds \forall t > 0$.

(ii) For all $s, t > 0$, the sequence $\Delta_n^{\frac{1}{2} + \rho'} \left(N_t^n - N_{(t-s\Delta_n^{\rho'})_+}^n \right)$ is bounded in probability for some $1/4 \leq \rho' < 1/2$.

(iii) For any $\kappa \geq 2$, there are a sequence $(\tau(\kappa)_m)_{m \geq 1}$ of stopping times increasing to infinity and real numbers $(w(\kappa)_m)$ such that we have for all $i, n, m, \rho > 1/4$:

$$T(n, i - 1) \leq \tau(\kappa)_m \Rightarrow \begin{cases} |\mathbb{E}(\alpha_{T(n, i-1)} \Delta(n, i) | \mathcal{F}_{T(n, i-1)}) - \Delta_n| \leq w(\kappa)_m \Delta_n^{1+\rho}, \\ \mathbb{E}(|\alpha_{T(n, i-1)} \Delta(n, i)|^\kappa | \mathcal{F}_{T(n, i-1)}) \leq w(\kappa)_m \Delta_n^\kappa. \end{cases} \quad (5)$$

The observation times framework is very general, and includes, e.g., *regular sampling scheme, time-changed regular sampling scheme, modulated Poisson sampling scheme, and predictably-modulated random walk sampling scheme*. However, the observation scheme does requires that the durations are conditionally independent of X , see the discussion on p.1136-1137 in [Jacod et al. \(2017\)](#).⁷

2.3 Microstructure noise

We now introduce the setting of the microstructure noise. The microstructure noise has a random scale and could exhibit some degree of serial dependence of an unknown form. Moreover, the scales of noise could be diminishing as the data frequency increases.

Definition 2.1. Let $\{\chi_i\}_{i \in \mathbb{Z}}$ be a sequence of stationary random variables defined on a probability space $(\Omega^{(1)}, \mathcal{G}, \mathbb{P}^{(1)})$, where \mathbb{Z} is the set of integers. The probability space has discrete filtrations $\mathcal{G}_p := \sigma\{\chi_k : p \geq k\}$, $\mathcal{G}^q := \sigma\{\chi_k : q \leq k\}$ satisfying $\mathcal{G}^{-\infty} = \mathcal{G}_\infty = \mathcal{G}$. For any $k \in \mathbb{N}^*$, we define the following mixing coefficients for $k \in \mathbb{N}^*$:

$$\rho_k := \sup \{ |\mathbb{E}(V_h V_{k+h})| : \mathbb{E}(V_k) = \mathbb{E}(V_{k+h}) = 0, \|V_h\|_2 \leq 1, \|V_{k+h}\|_2 \leq 1, V_h \in \mathcal{G}_h, V_{k+h} \in \mathcal{G}^{k+h} \}. \quad (6)$$

The sequence $\{\chi_i\}_{i \in \mathbb{Z}}$ is ρ mixing if $\rho_k \rightarrow 0$ as $k \rightarrow \infty$.

Assumption (N- v). Let $\{\chi_i\}_{i \in \mathbb{Z}}$ be a stationary and strongly mixing random sequence with mixing coefficients $\{\rho_k\}_{k \in \mathbb{N}^*}$ on some probability space $(\Omega^{(1)}, \mathcal{G}, \mathbb{P}^{(1)})$. At stage n , the noise at time $T(n, i)$ is given by

$$\varepsilon_i^n = \Delta_n^\eta \cdot \gamma_{T(n,i)} \cdot \chi_i, \quad (7)$$

where γ is an Itô semimartingale satisfying Assumption (H- r) with $r = 2$ and $\gamma_t > 0$ for all t . We further assume $\{\chi_i\}_{i \in \mathbb{Z}}$ is centred at 0 with variance 1 and finite moments of all orders, independent of $\mathcal{F}_\infty := \bigvee_{t>0} \mathcal{F}_t$. The mixing coefficients satisfy $\rho_k \leq Kk^{-v}$ for some $K > 0, v > 0$. The shrinking index η satisfies $\eta \in [0, \frac{1}{6})$.

Remark 2.2. The noise represented in (7) has a multiplicative form. Δ_n^η controls the shrinking effect with $\eta = 0$ corresponding to the non-vanishing noise that is well studied in the literature. The process γ captures the stochastic scale of noise. It is a continuous time process and dependent on the calendar time. This process could also be dependent of the efficient price, e.g., γ could be a function of the volatility of the efficient price so that both the noise and the efficient price may exhibit some diurnal features that are well documented in the literature. On the other hand, the χ process is essentially a discrete-time process that characterizes the serial dependence of noise. In our context, the discreteness reflects the ticks of high-frequency prices. Therefore, the serial dependence of noise is in tick times.

⁷Li et al. (2014) discuss a stronger form of endogeneity under which volatility estimators like RV may have a bias.

3 The Estimators

Now we introduce the new IV and SV estimators. This method is based on the pre-averaging method with a ReMeDI correction of the impact of microstructure noise.

3.1 The Pre-averaging method

The kernel g on \mathbb{R} is continuous piecewise C^1 with a piecewise Lipschitz derivative g' , and $s \notin (0, 1) \Rightarrow g(s) = 0$, we also assume $\int_0^1 g^2(s)ds > 0$.⁸ Let $\{h_n\}_{n \in \mathbb{N}^*}$ be a sequence of integers. We introduce the following notations for the pre-averaging method.

$$\begin{aligned} g_i^n &:= g(i/h_n), \quad \tilde{g}_i^n := g_{i+1}^n - g_i^n, \quad \phi_j^n := \frac{1}{h_n} \sum_{i \in \mathbb{Z}} g_i^n g_{i-j}^n, \quad \tilde{\phi}_j^n := h_n \sum_{i \in \mathbb{Z}} \tilde{g}_i^n \tilde{g}_{i-j}^n; \\ \phi(s) &:= \int_{\mathbb{R}} g(u)g(u-s)du, \quad \tilde{\phi}(s) := \int_{\mathbb{R}} g'(u)g'(u-s)du; \\ \Phi_{00} &:= \int_0^1 \phi^2(s)ds, \quad \Phi_{01} := \int_0^1 \phi(s)\tilde{\phi}(s)ds, \quad \Phi_{11} := \int_0^1 \tilde{\phi}^2(s)ds. \end{aligned}$$

For any processes V , let $\bar{V}_i^n := \sum_{j=1}^{h_n-1} g_j^n \Delta_{i+j}^n V$.

3.2 The ReMeDI method

We use the ReMeDI approach (Li and Linton, 2022a) to estimate the moments of noise, which will be subsequently subtracted off to get a consistent estimator of the integrated volatility. The ReMeDI approach employs the noisy returns on non-overlapping intervals to estimate the second moments of noise. The intuition is that the efficient returns are not predictable; thus, the autocovariances of the efficient returns will be negligible. As a consequence, the autocovariances of the noisy returns will be a good proxy of the autocovariances of the noise differences. By tuning the distance and length of the non-overlapping intervals, we can estimate the variance and autocovariances (of all orders) of the noise; see Section 3 of Li and Linton (2022a) for the intuition of the ReMeDI method.

3.3 The Pre-averaging–ReMeDI (PaReMeDI) estimator of IV and SV

We propose a pre-averaging method coupled with the ReMeDI bias correction to estimate the volatility parameters. We call this approach the PaReMeDI method. Given four sequences of integers $\{h_n\}_n, \{k_n\}_n, \{\ell_n\}_n, \{l_n\}_n$ and a sequence of reals $\{u_n\}_n$, the PaReMeDI estimator of IV is given by

$$\hat{C}_t^n := \frac{1}{h_n \phi_0^n} \sum_{i=0}^{N_t^n - h_n} (\bar{Y}_i^n)^2 \mathbf{1}_{\{|\bar{Y}_i^n| \leq u_n\}} - \frac{1}{h_n^2 \phi_0^n} \sum_{i=k_n}^{N_t^n - h_n} \sum_{|\ell| \leq \ell_n} \tilde{\phi}_\ell^n (Y_{i+\ell}^n - Y_{i+\ell+k_n}^n)(Y_i^n - Y_{i-k_n}^n), \quad (8)$$

⁸The simulation and empirical studies, we will use the triangular kernel: $g(x) = x \wedge (1 - x)$.

and the PaReMeDI estimator of SV (at time t) is given by

$$\tilde{c}_t^n := \frac{1}{s_t^n h_n \phi_0^n} \sum_{i=N_t^n}^{N_t^n + l_n - h_n} (\bar{Y}_i^n)^2 \mathbf{1}_{\{|\bar{Y}_i^n| \leq u_n\}} - \frac{1}{s_t^n h_n^2 \phi_0^n} \sum_{i=N_t^n + k_n}^{N_t^n + l_n - h_n} \sum_{|\ell| \leq \ell_n} \tilde{\phi}_\ell^n (Y_{i+\ell}^n - Y_{i+\ell+k_n}^n)(Y_i^n - Y_{i-k_n}^n), \quad (9)$$

where $s_t^n := T(n, N_t^n + l_n) - T(n, N_t^n)$.

The second terms in (8) and (9) are the ReMeDI corrections of the noise effect, which are based on the autocovariances of noisy returns. They are the estimators of the moments of noise that appear in the limit of the pre-averaged statistics \bar{Y}_i^n . The moments, up to scaling by $\tilde{\phi}_i^n$, are the (integrated and spot) long-run variances of the noise process.

Remark 3.1 (Tuning parameters of the PaReMeDI estimators). *The PaReMeDI estimators have several tuning parameters that warrant some explanation. h_n controls the pre-averaging bandwidth; k_n is the tuning parameter of the ReMeDI estimators to estimate the moments of the microstructure noise; l_n controls the size of the local window to perform spot estimation; u_n controls the levels to truncate the jumps of the efficient price process (Mancini, 2001); ℓ_n is the lag truncation parameter to estimate the long-run variance of noise. The regularity conditions imposed on the tuning parameters will be introduced in the next section alongside the limiting theorems.*

3.4 The estimators of the asymptotic variances

The next section presents the large sample properties of our estimators and the associated feasible CLTs, which require consistent estimators of the asymptotic variances. This subsection develops the estimators of the asymptotic variances.

We introduce some notations first. Given three sequences of integers $\{d_n\}_n, \{k_n\}_n, \{\ell_n\}_n$ with $d_n \geq \ell_n + k_n$, two integers $\ell, i \in \mathbb{N}^*$, and a process V , we introduce

$$r(V; \ell)_{i, d_n}^n := \frac{1}{d_n} \sum_{d=i+k_n+1}^{i+d_n+k_n} (V_{d+\ell}^n - V_{d+\ell+k_n}^n)(V_d^n - V_{d-k_n}^n); \quad R(V)_{i, d_n}^n := \sum_{|\ell| \leq \ell_n} r(V; |\ell|)_{i, d_n}^n. \quad (10)$$

Intuitively, $r(V; \ell)_{i, d_n}^n$ and $R(V)_{i, d_n}^n$ are the local estimators of the autocovariances and long-run variances of microstructure noise when $V = Y$.

Let $\{\tilde{h}_n\}_n$ and $\{s_n\}_n$ be two sequences such that $\tilde{h}_n \sim \Delta_n h_n^2, s_n \sim \Delta_n l_n$. Let $\tilde{N}_t^n := N_t^n + l_n, \hat{N}_t^n := \tilde{N}_t^n - \ell_n - 2k_n - h_n$. We introduce the following processes and coefficients:

$$\begin{aligned} V_t^{n,1} &:= \frac{1}{\tilde{h}_n} \sum_{i=0}^{N_t^n - h_n + 1} (\bar{Y}_i^n)^4 \mathbf{1}_{\{|\bar{Y}_i^n| < u_n\}}, & \tilde{V}_t^{n,1} &:= \frac{1}{s_n \tilde{h}_n} \sum_{i=N_t^n}^{\tilde{N}_t^n - h_n} (\bar{Y}_i^n)^4 \mathbf{1}_{\{|\bar{Y}_i^n| < u_n\}}; \\ V_t^{n,2} &:= \frac{1}{\tilde{h}_n h_n} \sum_{i=0}^{N_t^n - h_n - 2k_n - \ell_n} (\bar{Y}_i^n)^2 \mathbf{1}_{\{|\bar{Y}_i^n| < u_n\}} R(Y)_{i, h_n}^n, & \tilde{V}_t^{n,2} &:= \frac{1}{s_n \tilde{h}_n h_n} \sum_{i=N_t^n}^{\hat{N}_t^n} (\bar{Y}_i^n)^2 \mathbf{1}_{\{|\bar{Y}_i^n| < u_n\}} R(Y)_{i, h_n}^n; \\ V_t^{n,3} &:= \frac{1}{\tilde{h}_n h_n^2} \sum_{i=0}^{N_t^n - h_n - 2k_n - \ell_n} (R(Y)_{i, h_n}^n)^2, & \tilde{V}_t^{n,3} &:= \frac{1}{s_n \tilde{h}_n h_n^2} \sum_{i=N_t^n}^{\hat{N}_t^n} (R(Y)_{i, h_n}^n)^2. \end{aligned}$$

We use the following estimators for the asymptotic variances:

$$\begin{aligned}\Sigma_t^n &:= \frac{4\tilde{h}_n}{\phi^4(0)h_n} \left(K_{g,1}V_t^{n,1} + 2K_{g,2}V_t^{n,2} + K_{g,3}V_t^{n,3} \right), \\ \tilde{\Sigma}_t^n &:= \frac{4h_n}{\phi^4(0)l_n} \left(K_{g,1}\tilde{V}_t^{n,1} + 2K_{g,2}\tilde{V}_t^{n,2} + K_{g,3}\tilde{V}_t^{n,3} \right).\end{aligned}$$

where

$$K_{g,1} := \frac{\Phi_{00}}{3}; K_{g,2} := \Phi_{01}\phi(0) - \Phi_{00}\tilde{\phi}(0); K_{g,3} := \Phi_{11}\phi^2(0) - 2\Phi_{01}\tilde{\phi}(0)\phi(0) + \Phi_{00}\tilde{\phi}^2(0).$$

4 Asymptotic Properties of the Estimators

We present the limit theorems of our proposed estimators in this section. Throughout this section, we assume h_n , the bandwidth of the pre-averaging estimator, and u_n , the truncation level of jumps satisfy

$$h_n = \theta\Delta_n^{\eta-\frac{1}{2}} \left(1 + o\left(\Delta_n^{-(2\eta+1)/4}\right) \right), u_n \asymp (\Delta_n h_n)^\varpi, \text{ where } \frac{1}{4(2-r)} < \varpi < \frac{2[v]-3}{8([v]-1)}, r < \frac{2[v]-4}{2[v]-3}.$$

Recall that the parameter η controls the asymptotic order of noise and satisfies $\eta \in [0, 1/6)$;⁹ v restricts the decaying rate of the autocovariances of noise (recall Assumption **(N- v)**);¹⁰ and r is the jump activity index. We will also assume the three sequences of integers $\{l_n\}_n, \{\ell_n\}_n, \{k_n\}_n$ satisfy

$$l_n \asymp \Delta_n^{-l}, \ell_n \asymp \Delta_n^{-\ell}, k_n \asymp \Delta_n^{-k}.$$

The regularity conditions imposed on the exponents l, ℓ and k will be discussed later.

Theorem 4.1. *Let Assumptions **(H-X-r)**, **(O- ρ, ρ')**, **(N- v)** hold. Assume*

$$v > \frac{4}{1-6\eta}, \ell \in \left(\frac{1+2\eta}{4(v-1)}, k \right), k \in \left(\ell \sqrt{\frac{1+2\eta}{2(v-1)}}, \frac{1-6\eta}{6} \right). \quad (11)$$

Let $\eta_n \sim \Delta_n^{-\frac{1}{4}-\frac{\eta}{2}}$. Then, we have the following \mathcal{F}_∞ -stable convergence in law for any $t > 0$

$$\eta_n(\hat{C}_t^n - C_t) \xrightarrow{\mathcal{L}_s-\mathcal{F}_\infty} \mathcal{U}_t,$$

where $\mathcal{U}_t := \int_0^t \beta_s dB_s$ is defined on an extension $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ of the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and B is a standard Brownian motion that is independent of \mathcal{F} , and

$$\beta_s^2 = \frac{4}{\phi^2(0)} \left(\Phi_{00} \frac{\theta\sigma_s^4}{\alpha_s} + 2\Phi_{01} \frac{\sigma_s^2\gamma_s^2}{\theta} R + \Phi_{11} \frac{\gamma_s^4\alpha_s}{\theta^3} R^2 \right), \quad R := \sum_{\ell \in \mathbb{Z}} \mathbb{E}(\chi_i \chi_{i+\ell}). \quad (12)$$

⁹The same condition is imposed on shrinking noise in Chapter 7 of [Ait-Sahalia and Jacod \(2014\)](#) for i.i.d. noise.

¹⁰ $[v]$ is the integer part of v .

Moreover, the sequence $\frac{\tilde{c}_t^n - c_t}{\sqrt{\Sigma_t^n}}$ converges \mathcal{F}_∞ -stably in law to an $\mathcal{N}(0, 1)$ variable that is independent of \mathcal{F} .

Remark 4.1 (Convergence rate with small noise). Note that when the noise is shrinking, i.e., $\eta > 0$, the convergence rate is faster than the optimal rate $\Delta_n^{-1/4}$ when the noise has a constant scale. Since $\eta < 1/6$, the convergence rate can be arbitrarily close to $\Delta_n^{-1/3}$.

Theorem 4.2. Let Assumptions (H-X-r), (O- ρ, ρ'), (N-v) hold. Assume

$$v > 4, l \in \left(\frac{1}{2} - \eta, \frac{3}{4} - \frac{\eta}{2} \right), \ell \in \left(\frac{1 + 2\eta}{8(v-1)}, k \right), k \in \left(\ell \vee \frac{1 + 2\eta}{4(v-1)}, \frac{1 - 2\eta}{6} \right). \quad (13)$$

Let $\tilde{\eta}_n \sim l_n^{\frac{1}{2}} \Delta_n^{\frac{1}{4} - \frac{\eta}{2}}$. We have the following finite-dimensional \mathcal{F}_∞ -stable convergence in law,

$$\tilde{\eta}_n (\tilde{c}_t - c_t) \xrightarrow{\mathcal{L}_{\mathcal{F}}-s} Z_t;$$

where the process Z is defined on an extension $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ of the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and conditionally on \mathcal{F}_∞ , is a Gaussian white noise with conditional variance

$$\mathbb{E}(Z_t^2 | \mathcal{F}) = \tilde{\beta}_t^2 = \frac{4}{\phi^2(0)} \left(\Phi_{00} \frac{\theta \sigma_t^4}{\alpha_t^2} + 2\Phi_{01} \frac{\sigma_t^2 \gamma_t^2}{\theta \alpha_t} R + \Phi_{11} \frac{\gamma_t^4}{\theta^3} R^2 \right).$$

Moreover, we have the finite-dimensional \mathcal{F}_∞ -stable convergence in law for the sequence $\frac{\tilde{c}_t^n - c_t}{\sqrt{\Sigma_t^n}}$ to an $\mathcal{N}(0, 1)$ variable that is independent of \mathcal{F} .

Remark 4.2 (Random observations and SV estimation). We may consider the following estimator of SV at time t (recall $s_n \sim \Delta_n l_n$):

$$\tilde{c}_t^n := \frac{1}{s_n h_n \phi_0^n} \sum_{i=N_t^n}^{\tilde{N}_t^n - h_n} (\bar{Y}_i^n)^2 \mathbf{1}_{\{|\bar{Y}_i^n| \leq u_n\}} - \frac{1}{s_n h_n^2 \phi_0^n} \sum_{i=N_t^n + k_n}^{\tilde{N}_t^n - h_n} \sum_{|\ell| \leq \ell_n} \tilde{\phi}_\ell^n (Y_{i+\ell}^n - Y_{i+\ell+k_n}^n) (Y_i^n - Y_{i-k_n}^n). \quad (14)$$

Note that the estimator \tilde{c}_t^n is scaled by the average time span of the local blocks $l_n \Delta_n$; such estimator often appears in the literature of SV estimation; see, [Ait-Sahalia and Jacod \(2014\)](#) and [Jacod and Protter \(2011\)](#) and references therein, see also the recent work by [Li et al. \(2022\)](#). However, our estimator \tilde{c}_t^n is scaled by the real time duration s_t^n . The two estimators coincide with each other if the observation scheme is regular, i.e., $\alpha_t \equiv 1 \forall t$. In the presence of stochastic and irregular observation times, the two estimators have different stochastic limits. Under the same conditions of [Theorem 4.2](#), we have the following finite dimensional \mathcal{F}_∞ -stable convergence in law:

$$\tilde{\eta}_n \left(\tilde{c}_t^n - \frac{c_t}{\alpha_t} \right) \xrightarrow{\mathcal{L}_{\mathcal{F}}-s} \frac{Z_t}{\alpha_t}. \quad (15)$$

Therefore, the SV estimators should be normalized or scaled by the real-time durations in the presence of random and irregular observation times. Otherwise, the estimators may be inconsistent. Specifically, when transactions occur more (less) often than in a regular scheme, i.e., $\alpha_t > 1$ ($\alpha_t < 1$), \tilde{c}_t^n tends to

underestimate (overestimate) c_t . More numerical and empirical evidence of the caused inconsistency is provided in Section 6 and Section 7.

Remark 4.3 (Optimal convergence rate). *The convergence rate of our estimator is $\Delta_n^{-\frac{1}{8}-\frac{\eta}{4}+\iota}$ for any $\iota > 0$. Thus, $\Delta_n^{-\frac{1}{8}-\frac{\eta}{4}}$ serves as the upper bound of the optimal convergence rate,¹¹ and the upper bound can be obtained if we set $l_n \asymp \Delta_n^{-\frac{3}{4}+\frac{\eta}{2}}$. However, the limiting process will be slightly different—there will be an additional white noise process Z_t^ι independent of Z_t , with its limiting variance depending on the volatility of the volatility (VoV), see the comments after Theorem 13.3.3 in [Jacod and Protter \(2011\)](#). Therefore, to make the limit theorem feasible in practice, we need to construct an estimator of the limiting variance of the SV estimator at the cost of imposing additional assumptions on VoV. This is beyond the scope of this paper, and we leave it for future work.*

5 On the implementation of PaReMeDI

To implement the PaReMeDI estimators, we need to select several tuning parameters, recall (11) and (13). This section provides some guidances on the selection of the tuning parameters in practice.

5.1 The optimal selection of θ

As we will see in our extensive simulation studies that the key parameter that affects the performance of PaReMeDI is the pre-averaging bandwidth h_n —it not only affects the finite-sample performance but also the limiting distribution via θ . Therefore, the selection of θ is of great practical concern.

Let's first consider the estimation of IV. We can find the *optimal* θ as the value that minimizes the asymptotic variance $\Sigma_t := \int_0^t \beta_s^2 ds$ of our estimator. The optimal θ is given explicitly by¹²

$$\theta^* := \left(\frac{\Phi_{01} R \int_0^t \sigma_s^2 \gamma_s^2 ds + \sqrt{\Phi_{01}^2 R^2 (\int_0^t \sigma_s^2 \gamma_s^2 ds)^2 + 3\Phi_{00}\Phi_{11} R^2 \int_0^t \sigma_s^4 / \alpha_s ds \int_0^t \gamma_s^4 \alpha_s ds}}{\Phi_{00} \int_0^t \sigma_s^4 / \alpha_s ds} \right)^{\frac{1}{2}}.$$

The implementation the optimal selection rule requires an estimate of θ^* . For a given θ_0 , we let

$$\Theta_t^{n,1} := \theta_0^2 V_t^{n,3}; \quad \Theta_t^{n,2} := \frac{V_t^{n,2} - \tilde{\phi}(0) V_t^{n,3}}{\phi(0)}; \quad \Theta_t^{n,3} := \frac{V_t^{n,1} - 6\tilde{\phi}(0) V_t^{n,2} + 3\tilde{\phi}^2(0) V_t^{n,3}}{3\theta_0^2 \phi^2(0)}.$$

A consistent estimator of the optimal θ^* is given by

$$\hat{\theta}^* := \left(\frac{\Phi_{01} \Theta_t^{n,2} + \sqrt{\Phi_{01}^2 (\Theta_t^{n,2})^2 + 3\Phi_{00}\Phi_{11} \Theta_t^{n,1} \Theta_t^{n,3}}}{\Phi_{00} \Theta_t^{n,3}} \right)^{1/2}.$$

¹¹For example, if the noise is shrinking with η close to $\frac{1}{6}$ in our setting, the convergence rate can be close to $\Delta_n^{-\frac{1}{6}}$.

¹²See also the discussion in [Jacod and Mykland \(2015\)](#) in a simpler setting.

The implementation of $\hat{\theta}^*$ requires the local estimations of many parameters that may bring in more variability, thus may not provide a proper guidance to select h_n in a finite sample. However, we can make some compromise between estimation complexity and parsimony.

Consider a simpler setting where the noise is stationary, the volatility process is constant, and the observation scheme is regular, i.e., $\gamma_s \equiv K_\gamma, \sigma_s \equiv K_\sigma, \alpha_s \equiv 1$ for all s . In this special case, the optimal choice of θ becomes $\theta'^* := K_\Phi K_\gamma \sqrt{R}/K_\sigma$, where $K_\Phi := \sqrt{(\Phi_{01} + \sqrt{\Phi_{01}^2 + 3\Phi_{00}\Phi_{11}})/\Phi_{00}}$ is a constant determined by the choice of g . Note that $K_\gamma \sqrt{R}/K_\sigma$ is the noise-to-signal ratio. Thus, the selection rule set by θ'^* is very intuitive: one should choose a larger (or small) θ if noise is relatively large (or small).

Let $R(Y)_t^n := R(Y)_{0, N_t^n - \ell_n - k_n}^n$ (see (10) for the notation of $R(Y)_{i, d_n}^n$). Based on the convergence that $R(Y)_t^n \xrightarrow{\mathbb{P}} K_\gamma^2 R$, we propose the following proxy of θ'^* :

$$\hat{\theta}'^* = \frac{K_\Phi \sqrt{R(Y)_t^n}}{\sqrt{\hat{C}_t^n}}, \quad (16)$$

which provides a simple rule to select θ whence h_n in practice.

We can also develop a spot version of the optimal θ when the noise is stationary and the observation scheme is regular. The optimal θ is given by $\theta_t'^* := K_\Phi K_\gamma \sqrt{R}/\sigma_t$. Thus, an estimator of $\theta_t'^*$ is given by

$$\tilde{\theta}_t'^* := \frac{K_\Phi \sqrt{\tilde{R}(Y)_t^n}}{\sqrt{\tilde{c}_t^n}}, \quad (17)$$

where $\tilde{R}(Y)_t^n := R(Y)_{N_t^n, \ell_n - \ell_n - k_n}^n$.

Remark 5.1. *To implement the optimal selection rule in practice, we initially set $\theta = 0.5$ to get some preliminary estimates of \hat{C}_t^n and \tilde{c}_t^n , which will be used to calculate the optimal θ via (16) and (17).¹³ The optimal θ calculated in this way may be extremely small or large, depending on the underlying noise process and the initial estimates. As a consequence, the pre-averaging bandwidth would be extreme as well. To avoid such extreme scenarios, we restrict the range of θ in the implementation of PaReMeDI: we set $\theta \in [0.05, 10]$ for the IV estimation and $\theta \in [0.2, 1]$ in the SV estimation.*

Remark 5.2 (Nonnegative Estimates). *Another practical issue related to the choice of θ is the negative estimate of volatility. The PaReMeDI estimators are not guaranteed to always yield a positive estimate of IV or SV, although we do not find any negative estimates in our simulation and empirical studies. An estimate that is negative would mainly signal that there are some anomalies in the microstructure noise, e.g., extremely large or small scales or strong autocorrelation patterns. We propose two remedies. First, one could proceed with PaReMeDI, but select a large θ instead of the optimal one. A large θ has an over-smoothing effect and the anomalies of noise will be smoothed away, see more numerical experiments*

¹³We keep other tuning parameters, e.g., k_n, ℓ_n fixed while we are calculating the optimal θ based on an initial choice $\theta = 0.5$. Our numerical studies show that the performances of the PaReMeDI estimators are very robust to the choices and combinations of other parameters.

in the next section. Second, one could just ignore the noise and use the realized volatility (RV) instead. RV yields accurate estimates, especially if the scale of noise is small.

5.2 On the choices of other tuning parameters

Our simulation studies (see next section) show the PaReMeDI estimators are less sensitive to the choices of k_n . Thus, we recommend using some mild choices of k_n , e.g., we fix $k_n = 10$ in the empirical studies. Moreover, we could combine the ReMeDI estimates with different k_n s and use the average estimates to correct the biases of the pre-averaging method. Intuitively, using multiple k_n s would improve the efficiency of the ReMeDI estimators and would reduce the errors in bias-correction for the PaReMeDI method. However, we do not further explore the method in this paper.

The selection of the lag truncation parameter ℓ_n in the estimation of the LRV is well studied in the literature (Andrews, 1991). In this paper, however, we will only require $\ell_n \leq k_n$ as required by our asymptotic conditions in (11) and (13). In our simulation studies, we find $\ell_n = \lfloor k_n/2 \rfloor$ will produce accurate estimates.

We also find that our SV estimators are very robust to the choices of the bandwidth l_n . In our empirical studies, we find the differences in the estimates using different l_n s are minor. However, we observe sizeable differences if one treats the random observation times as regular.

6 Simulation Studies

In this section, we adopt the simulation approach to explore the finite sample performance of the PaReMeDI estimators of IV and SV. We will compare the PaReMeDI estimators with several other competing estimators recently studied in the literature. Each estimator employs different combinations of the tuning parameters, and the performances are studied in a range of empirically relevant scenarios where the noise scales and autocorrelation patterns vary. Thus, the numerical study directly speaks to the robustness of the estimators to the choices of the tuning parameters. Moreover, it also provides some guidance for selecting the tuning parameters in empirical research.

6.1 Model settings

We normalize the interval of observations to $[0, 1]$ without loss of generality. The efficient price is allowed to have stochastic volatility with jumps in both the price level and the volatility process:

$$\begin{aligned} dX_t &= \kappa_1(\mu_1 - X_t)dt + \sigma_t dW_{1,t} + \xi_{1,t} dN_t, & d\sigma_t^2 &= \kappa_2(\mu_2 - \sigma_t^2)dt + \eta\sigma_t dW_{2,t} + \xi_{2,t} dN_t; \\ \text{Corr}(W_1, W_2) &= \nu, & \xi_{1,t} &\sim \mathcal{N}(0, \mu_2/10), & N_t &\sim \text{Poi}(\lambda); & \xi_{2,t} &\sim \text{Exp}(\delta), \end{aligned} \quad (18)$$

where we set

$$\kappa_1 = 0.5; \mu_1 = 3.6; \kappa_2 = 5/252; \mu_2 = 0.04/252; \eta = 0.05/252; v = -0.5; \lambda = 1; \delta = \eta.$$

The setting of jumps is motivated by some empirical facts that jumps in price levels and volatility tend to occur together (Todorov and Tauchen, 2011; Bibinger and Winkelmann, 2018). In this section, we set $\xi_{1,t} \equiv 0$,¹⁴ the supplementary materials of this paper (Li and Linton, 2022b) contains additional simulation studies where the efficient price exhibits jumps.

The stationary component of the microstructure noise follows an AR(1) process with innovations following a t -distribution with 5 degrees of freedom:

$$\chi_{i+1} = \varrho\chi_i + e_i, \quad |\varrho| < 1. \quad (19)$$

$\{T(n, i)\}_i$ follow an inhomogeneous Poisson process with rate $n\alpha_t$ where the processes α and γ satisfy

$$\alpha_t = 1 + \frac{\cos(2\pi t)}{2}; \quad \gamma_t = K_\gamma \gamma'_t, \quad d\gamma'_t = -\rho_\gamma(\gamma'_t - \mu_t)dt + \sigma_\gamma dW_{1,t}. \quad (20)$$

The setting is to mimic the empirical facts that trades tend to cluster in the beginning and end of a trading day, and microstructure noise has an approximately U-shape. We set $n = 23400$, $\rho_\gamma = 10$, $\mu_t = 1 + 0.1 \cos(2\pi t)$, $\sigma_\gamma = 0.1$. Note that ϱ, K_γ control the serial dependence of noise and the signal-to-noise ratio, which are essential in the model specifications. We will let the two variables vary to examine the robustness of the estimators.

6.2 FRK, PaReMeDI, QMLE and TSRK

We compare the performances of four estimators developed in similar settings: The Flat-top Realized Kernels (FRK) (Varneskov, 2016, 2017), the Pre-averaging estimator with ReMeDI bias correction (PaReMeDI), the quasi-maximum-likelihood estimator (QMLE) (Da and Xiu, 2021b), and the Two-scales Realized Kernels (TSRK) (Ikeda, 2015, 2016). The next subsection will compare PaReMeDI with a class of pre-averaging estimators, including the pre-averaging estimator with local averaging correction (PaLA).

We consider four scales of noise and three autocorrelation patterns. Thus, we have a wide range of specifications to examine the performances of the estimators. We allow for different combinations of the tuning parameters for the nonparametric estimators (FRK, PaReMeDI and TSRK), and we use both the AIC and BIC rules to select the MA(q) models for QMLE.

Table 1 reports the relative biases of the estimators. All estimators are performing very well in all scenarios except for large ($K_\gamma = 5 \times 10^{-4}$) and strongly persistent ($\varrho = 0.8$) noise. Even with such extreme specifications, the biases could be very small with certain combinations of the tuning parameters for the nonparametric estimators, e.g., FRK and TSRK with large

¹⁴We would like to compare our volatility estimator with QMLE (Da and Xiu, 2021b), which does not explicitly truncate jumps. Therefore, we neglect jumps from the efficient price to present a fair comparison.

bandwidths. One may observe that the PaReMeDI estimator with the optimal θ selected via (16) has slightly larger biases compared to the other estimators. It is not surprising since the optimal θ is not set to minimize the finite sample biases. In fact, the biases can be further reduced when we fix θ at certain values; see the detailed discussion in the next subsection. However, the advantage with the optimal θ is that the PaReMeDI estimator has smaller root mean squared relative errors (RMSRE). Table 2 reports the RMSRE for all the estimators. FRK (with certain choices of the tuning parameters) and QMLE perform well on relatively large noise. Both QMLE and PaReMeDI have very small RMSRE when noise is relatively small.¹⁵ In terms of computational efficiency, FRK and PaReMeDI are faster than QMLE: both estimators typically use less than 0.5% of the computational time used by QMLE. This is an advantage to deal with the massive high-frequency datasets in the empirical research.

6.3 Pre-averaging methods with various bias-correction methods

We now conduct an extensive evaluation of a group of pre-averaging estimators with different methods to remove the effect of microstructure noise. We consider four correction methods: the ReMeDI method, the local averaging (LA) estimator, the TSRK correction proposed in Varneskov (2016), and the realized volatility (RV) estimator of the variances of i.i.d. noise that is widely used in the literature, see, e.g., Jacod et al. (2009).

Table 3 reports the relative biases of each estimator. The pre-averaging method with LA (PaLA) corrections works well when the bandwidth parameter θ is large, and the tuning parameter of LA k_n is small. It is as expected as a larger θ has an over-smoothing effect in practice and the noise is almost smoothed away irrespective of its scale or serial correlation pattern (see also the discussions in Barndorff-Nielsen et al. (2008) and Varneskov (2016)), and a smaller k_n yields a smaller finite sample bias in estimating the moments of noise using the LA method (Jacod et al., 2017). However, PaLA is very sensitive to θ : a smaller θ yields a significant and negative bias. The intuition is that the estimated noise moments, including the bias if there is any, account for a more significant variation in the pre-averaged noisy returns when θ is small. Since LA has a positive bias,¹⁶ it over-corrects the noise effect and induces a negative bias for PaLA. It is interesting to contrast PaLA with the traditional pre-averaging method (Jacod et al., 2009) that employs a realized volatility (RV) correction of the (possibly misspecified) i.i.d. noise. While it is clear in Table 3 that there is a significant bias when noise is misspecified, a direct comparison of the LA and RV corrections reveals that the finite sample bias often overwhelms the biases caused by misspecifications.

Now we compare the relative biases of the ReMeDI correction with other methods. The ReMeDI correction yields a very small bias and is quite robust to the choices of the tuning parameters. The TSRK also works well except for large noise and small θ . The ReMeDI correction with an optimal choice of θ using the selection rule (16) (ReMeDI* in Table 3) has larger biases compared to the ReMeDI corrections with fixed θ s. But the gain of the slightly

¹⁵We find the noise scales in our empirical data are closer to being moderate or small.

¹⁶The bias is a fraction of $k_n IV$, see the discussion in Section 3.4 of Jacod et al. (2017); see also the numerical studies by Da and Xiu (2021a), where the authors show that LA is very sensitive to noise-to-signal ratios.

larger bias is the reduced root-mean-squared-relative-errors (RMSRE) reported in Table 4. Measured by RMSRE, the ReMeDI correction with an optimal selection of θ outperforms other estimators, except for large and strongly correlated noise with very small tuning parameters.

6.4 SV estimation

We consider four estimators of the SV: the PaReMeDI estimator $\tilde{c}_t^n(\theta)$ (9) with a given θ ; the PaReMeDI estimator $\tilde{c}_t^n(\theta^*)$ with an optimally selected θ^* via (17); the PaReMeDI estimator $\tilde{c}_t^n(\theta)$ (14) with a given θ , the pre-averaging estimator of SV with a realized volatility correction of the i.i.d. noise, denoted by $\hat{c}_t^n(\theta)$. The estimator $\tilde{c}_t^n(\theta)$ ignores the effects of random sampling while $\hat{c}_t^n(\theta)$ treats noise as a simple i.i.d. process. We consider different combinations of tuning parameters for each estimator. We estimate SV at $t = 5$ min, $t = 100$ min and $t = 200$ min when $\alpha_t > 1$, $\alpha_t \approx 1$ and $\alpha_t < 1$; respectively.

Table 5 reports the relative biases for all SV estimators. We first discuss the left panel with large noise ($K_\gamma = 5 \times 10^{-4}$). Similar to IV estimation, all estimators have larger biases when the noise scale is large and the serial correlation is strong ($\rho = 0.8$). When the large noise becomes i.i.d., the biases of $\tilde{c}_t^n(\theta)$ caused by misspecification are much reduced. It is also interesting to note that the bias of $\tilde{c}_t^n(\theta)$ has a clear pattern when $\theta = 1$: it is negative at $t = 5$ min and substantially positive at $t = 200$, and remains small at $t = 100$ min, and the pattern is more pronounced for small noise ($K_\gamma = 5 \times 10^{-5}$) as one can observe in the right panel of Table 5. Such systematic and persistent bias is predicted by our asymptotic theory (15): An SV estimator that neglects the irregular and stochastic observation scheme tends to overestimate (underestimate) the SV when transactions are observed more (less) often than in a regular observation scheme.

Now we turn to the RMSRE of the estimators presented in Table 6. Compared to $\tilde{c}_t^n(\theta)$ and $\tilde{c}_t^n(\theta^*)$, $\hat{c}_t^n(\theta)$ has large RMSREs when $t = 5$ min and $t = 200$ min. Thus, the misspecification of the observation schemes directly translates into significant RMSREs. The RMSRE of $\tilde{c}_t^n(\theta)$ is small for small noise, but heightened significantly when the noise scale becomes large and the serial dependence becomes strong. Measured by RMSRE, our PaReMeDI estimator $\tilde{c}_t^n(\theta^*)$ with the optimal choice of θ is quite accurate and robust.

6.5 Examine the feasible CLTs

Figure 1 presents the QQ-plot for the standardized statistics $(\hat{C}_t^n - C_t)/\sqrt{\Sigma_t^n}$. The plots clearly demonstrate that the limit distribution provides an accurate fit in finite samples, and the fit is quite robust to model specifications and combinations of the tuning parameters. Our novel asymptotic variance estimators are at least partially responsible for such robustness. The supplementary material Li and Linton (2022b) contains further simulation studies on the feasible CLTs of the IV and SV estimators under various specifications, see Section A.1 in Li and Linton (2022b).

Specifications		$K_\gamma = 5 \times 10^{-4}$			$K_\gamma = 10^{-4}$			$K_\gamma = 5 \times 10^{-5}$			$K_\gamma = 10^{-5}$		
		$\varrho = 0$	0.3	0.8	$\varrho = 0$	0.3	0.8	$\varrho = 0$	0.3	0.8	$\varrho = 0$	0.3	0.8
FRK	$b_n = 25, f_n = 5$	0.0035	0.0042	0.6159	0.0026	0.0006	0.0232	-0.0017	-0.0019	0.0066	-0.0008	0.0001	-0.0038
	$b_n = 25, f_n = 10$	0.0039	0.0027	0.1997	0.0031	0.0012	0.0058	-0.0016	-0.0027	0.0021	-0.0007	-0.0001	-0.0043
	$b_n = 25, f_n = 15$	0.0032	0.0047	0.0678	0.0029	0.0021	-0.0006	-0.0010	-0.0030	-0.0000	-0.0005	-0.0000	-0.0049
	$b_n = 50, f_n = 5$	0.0033	0.0042	0.2238	0.0026	0.0012	0.0055	-0.0008	-0.0025	0.0017	-0.0003	0.0001	-0.0046
	$b_n = 50, f_n = 10$	0.0033	0.0039	0.0735	0.0024	0.0012	-0.0016	-0.0003	-0.0027	-0.0002	-0.0001	0.0001	-0.0050
	$b_n = 50, f_n = 15$	0.0030	0.0041	0.0248	0.0019	0.0008	-0.0044	0.0003	-0.0027	-0.0011	0.0002	-0.0000	-0.0052
	$b_n = 75, f_n = 5$	0.0031	0.0037	0.1114	0.0015	0.0003	-0.0003	-0.0000	-0.0025	0.0001	0.0001	-0.0003	-0.0052
	$b_n = 75, f_n = 10$	0.0029	0.0034	0.0364	0.0010	-0.0001	-0.0038	0.0003	-0.0026	-0.0008	0.0003	-0.0005	-0.0056
	$b_n = 75, f_n = 15$	0.0027	0.0033	0.0121	0.0003	-0.0007	-0.0052	0.0006	-0.0025	-0.0013	0.0005	-0.0008	-0.0060
PaReMeDI	$k_n = 5, \ell_n = 3$	0.0536	0.0472	1.1219	0.0627	0.1045	0.1704	0.0147	0.0277	0.0442	-0.0003	0.0011	-0.0007
	$k_n = 10, \ell_n = 5$	0.0582	0.0398	0.2731	0.0544	0.0861	0.0807	0.0126	0.0243	0.0344	-0.0003	0.0007	-0.0017
	$k_n = 15, \ell_n = 8$	0.0745	0.0540	0.0867	0.0340	0.0574	0.0513	0.0043	0.0130	0.0200	-0.0023	-0.0013	-0.0036
TSRK	$G_n = 12, H_n = 76$	0.0001	0.0256	0.2534	0.0018	0.0017	0.0061	-0.0000	-0.0020	0.0019	0.0001	0.0000	-0.0047
	$G_n = 12, H_n = 153$	-0.0007	0.0058	0.0676	-0.0012	-0.0028	-0.0019	0.0006	-0.0017	-0.0013	0.0005	-0.0026	-0.0076
	$G_n = 12, H_n = 306$	-0.0045	-0.0010	0.0128	-0.0045	-0.0064	-0.0058	0.0012	-0.0037	-0.0043	0.0025	-0.0068	-0.0089
	$G_n = 29, H_n = 76$	-0.0004	0.0249	0.2636	0.0018	0.0014	0.0065	0.0000	-0.0020	0.0021	0.0002	0.0000	-0.0047
	$G_n = 29, H_n = 153$	-0.0009	0.0054	0.0703	-0.0013	-0.0030	-0.0018	0.0006	-0.0017	-0.0012	0.0005	-0.0027	-0.0076
	$G_n = 29, H_n = 306$	-0.0046	-0.0011	0.0135	-0.0046	-0.0064	-0.0058	0.0012	-0.0037	-0.0043	0.0025	-0.0068	-0.0090
	$G_n = 56, H_n = 76$	-0.0000	0.0260	0.2653	0.0018	0.0016	0.0066	-0.0001	-0.0020	0.0020	0.0001	0.0000	-0.0048
	$G_n = 56, H_n = 153$	-0.0013	0.0060	0.0709	-0.0014	-0.0031	-0.0018	0.0005	-0.0017	-0.0015	0.0005	-0.0028	-0.0078
	$G_n = 56, H_n = 306$	-0.0048	-0.0010	0.0136	-0.0046	-0.0065	-0.0059	0.0012	-0.0038	-0.0045	0.0025	-0.0069	-0.0090
QMLE	aic	-0.0070	-0.0021	0.2589	0.0055	0.0040	0.0508	-0.0001	0.0007	0.0172	-0.0005	0.0013	-0.0028
	bic	-0.0138	-0.0122	0.2815	0.0046	0.0048	0.0980	0.0009	0.0013	0.0285	-0.0011	0.0004	-0.0022

Table 1: The relative biases of FRK, PaReMeDI, TSRK and QMLE. The noise scale parameter K_γ is selected in $\{5 \times 10^{-4}, 10^{-4}, 5 \times 10^{-5}, 10^{-5}\}$. The AR(1) coefficient of the stationary noise is given by $\varrho \in \{0, 0.3, 0.8\}$. The flatness and bandwidth tuning parameters of FRK are selected from $(f_n, b_n) \in \{5, 10, 15\} \times \{25, 50, 75\}$. The jittering bandwidth is fixed at $j_n = 5$. The ReMeDI tuning parameter k_n is selected in $\{5, 10, 15\}$, and the lag truncation parameter ℓ_n is set via $\ell_n = \lceil k_n/2 \rceil$. θ is selected by the optimal rule in (16). The two bandwidths of TSRK are selected by $G_n \in \{[n^{1/4}], [n^{1/3}], [n^{2/5}]\}$, $H_n \in \{[\sqrt{n}/2], [\sqrt{n}], 2[\sqrt{n}]\}$ with $n = 23400$. The optimal moving average order q is selected within $\{5, 6, 7, 8, 9, 10\}$. The biases are obtained by taking the averages of 1000 simulations.

Specifications		$K_\gamma = 5 \times 10^{-4}$			$K_\gamma = 10^{-4}$			$K_\gamma = 5 \times 10^{-5}$			$K_\gamma = 10^{-5}$		
		$\varrho = 0$	0.3	0.8	$\varrho = 0$	0.3	0.8	$\varrho = 0$	0.3	0.8	$\varrho = 0$	0.3	0.8
FRK	$b_n = 25, f_n = 5$	0.1022	0.0989	0.7288	0.0994	0.1012	0.1017	0.1096	0.1026	0.0928	0.0971	0.0956	0.0976
	$b_n = 25, f_n = 10$	0.1156	0.1152	0.2786	0.1152	0.1197	0.1157	0.1303	0.1231	0.1127	0.1146	0.1151	0.1164
	$b_n = 25, f_n = 15$	0.1289	0.1278	0.1798	0.1287	0.1358	0.1306	0.1421	0.1418	0.1299	0.1311	0.1344	0.1321
	$b_n = 50, f_n = 5$	0.1182	0.1152	0.2943	0.1198	0.1251	0.1213	0.1321	0.1302	0.1193	0.1210	0.1233	0.1216
	$b_n = 50, f_n = 10$	0.1320	0.1298	0.1748	0.1337	0.1405	0.1368	0.1462	0.1470	0.1351	0.1370	0.1390	0.1365
	$b_n = 50, f_n = 15$	0.1456	0.1432	0.1692	0.1465	0.1546	0.1510	0.1585	0.1619	0.1487	0.1518	0.1526	0.1500
	$b_n = 75, f_n = 5$	0.1372	0.1336	0.2006	0.1387	0.1456	0.1416	0.1514	0.1515	0.1387	0.1434	0.1438	0.1417
	$b_n = 75, f_n = 10$	0.1506	0.1467	0.1728	0.1518	0.1596	0.1547	0.1642	0.1657	0.1513	0.1579	0.1570	0.1553
	$b_n = 75, f_n = 15$	0.1636	0.1586	0.1813	0.1643	0.1730	0.1664	0.1758	0.1786	0.1628	0.1717	0.1691	0.1681
PaReMeDI	$k_n = 5, \ell_n = 3$	0.1948	0.1640	1.5066	0.0932	0.1427	0.2046	0.0559	0.0674	0.0773	0.0556	0.0533	0.0552*
	$k_n = 10, \ell_n = 5$	0.2372	0.1932	0.4154	0.0957	0.1363	0.1409	0.0832	0.0776	0.0804	0.0674	0.0633	0.0668
	$k_n = 15, \ell_n = 8$	0.3726	0.3871	0.2084	0.1075	0.1340	0.1334	0.1040	0.1055	0.0949	0.0967	0.0855	0.0959
TSRK	$G_n = 12, H_n = 76$	0.1296	0.1278	0.3271	0.1298	0.1362	0.1322	0.1419	0.1418	0.1298	0.1337	0.1346	0.1324
	$G_n = 12, H_n = 153$	0.1822	0.1763	0.2086	0.1832	0.1968	0.1813	0.1914	0.1975	0.1786	0.1957	0.1847	0.1878
	$G_n = 12, H_n = 306$	0.2487	0.2537	0.2690	0.2457	0.2820	0.2540	0.2548	0.2807	0.2588	0.2817	0.2543	0.2750
	$G_n = 29, H_n = 76$	0.1349	0.1296	0.3382	0.1301	0.1363	0.1327	0.1425	0.1416	0.1302	0.1339	0.1342	0.1326
	$G_n = 29, H_n = 153$	0.1849	0.1782	0.2116	0.1851	0.1988	0.1831	0.1932	0.1994	0.1806	0.1978	0.1861	0.1898
	$G_n = 29, H_n = 306$	0.2499	0.2549	0.2703	0.2468	0.2833	0.2551	0.2560	0.2820	0.2602	0.2830	0.2554	0.2764
	$G_n = 56, H_n = 76$	0.1284	0.1273	0.3390	0.1284	0.1349	0.1305	0.1407	0.1400	0.1283	0.1323	0.1332	0.1311
	$G_n = 56, H_n = 153$	0.1876	0.1805	0.2120	0.1849	0.1998	0.1828	0.1931	0.1995	0.1807	0.1984	0.1862	0.1904
	$G_n = 56, H_n = 306$	0.2523	0.2572	0.2724	0.2485	0.2857	0.2570	0.2578	0.2841	0.2625	0.2853	0.2573	0.2789
QMLE	aic	0.1126	0.1068	0.3982	0.0852	0.0847	0.1159	0.0860	0.0876	0.0832	0.0828	0.0812	0.0815
	bic	0.1240	0.1082	0.4963	0.0751	0.0743	0.1370	0.0695	0.0749	0.0772	0.0711	0.0674	0.0706

Table 2: The root-mean-squared-relative-errors (RMSRE) of FRK, PaReMeDI, TSRK and QMLE. The noise scale parameter K_γ is selected in $\{5 \times 10^{-4}, 10^{-4}, 5 \times 10^{-5}, 10^{-5}\}$. The AR(1) coefficient of the stationary noise is given by $\varrho \in \{0, 0.3, 0.8\}$. The flatness and bandwidth tuning parameters of FRK are selected from $(f_n, b_n) \in \{5, 10, 15\} \times \{25, 50, 75\}$. The jittering bandwidth is fixed at $j_n = 5$. The ReMeDI tuning parameter k_n is selected in $\{5, 10, 15\}$, and the lag truncation parameter ℓ_n is set via $\ell_n = \lfloor k_n/2 \rfloor$. θ is selected by the optimal rule in (16). The two bandwidths of TSRK are selected by $G_n \in \{[n^{1/4}], [n^{1/3}], [n^{2/5}]\}$, $H_n \in \{[\sqrt{n}/2], [\sqrt{n}], 2[\sqrt{n}]\}$ with $n = 23400$. The optimal moving average order q is selected within $\{5, 6, 7, 8, 9, 10\}$. The number of simulations is 1000.

Specifications		$K_\gamma = 5 \times 10^{-5}$			$K_\gamma = 5 \times 10^{-4}$		
		$\varrho = 0$	0.3	0.8	$\varrho = 0$	0.3	0.8
LA	$k_n = 5, \ell_n = 5, \theta = 0.3$	-0.3710	-0.3717	-0.3680	-0.3721	-0.3697	-0.1247
	$k_n = 5, \ell_n = 5, \theta = 0.8$	-0.0609	-0.0635	-0.0624	-0.0639	-0.0632	-0.0094
	$k_n = 5, \ell_n = 5, \theta = 1.5$	-0.0231	-0.0258	-0.0269	-0.0276	-0.0271	-0.0120
	$k_n = 5, \ell_n = 10, \theta = 0.3$	-0.5633	-0.5638	-0.5612	-0.5646	-0.5617	-0.3644
	$k_n = 5, \ell_n = 10, \theta = 0.8$	-0.1062	-0.1088	-0.1079	-0.1092	-0.1084	-0.0647
	$k_n = 5, \ell_n = 10, \theta = 1.5$	-0.0374	-0.0401	-0.0413	-0.0419	-0.0414	-0.0294
	$k_n = 10, \ell_n = 5, \theta = 0.3$	-0.7661	-0.7657	-0.7634	-0.7664	-0.7653	-0.6164
	$k_n = 10, \ell_n = 5, \theta = 0.8$	-0.1229	-0.1254	-0.1245	-0.1258	-0.1253	-0.0866
	$k_n = 10, \ell_n = 5, \theta = 1.5$	-0.0413	-0.0440	-0.0451	-0.0457	-0.0453	-0.0346
	$k_n = 10, \ell_n = 10, \theta = 0.3$	-1.1647	-1.1634	-1.1627	-1.1645	-1.1638	-1.1070
	$k_n = 10, \ell_n = 10, \theta = 0.8$	-0.2168	-0.2191	-0.2186	-0.2195	-0.2192	-0.2010
$k_n = 10, \ell_n = 10, \theta = 1.5$	-0.0709	-0.0736	-0.0748	-0.0753	-0.0749	-0.0707	
ReMeDI	$k_n = 5, \ell_n = 5, \theta = 0.3$	-0.0024	-0.0032	0.0040	0.0023	0.0093	0.5042
	$k_n = 5, \ell_n = 5, \theta = 0.8$	-0.0031	-0.0056	-0.0040	-0.0058	-0.0048	0.0861
	$k_n = 5, \ell_n = 5, \theta = 1.5$	-0.0062	-0.0089	-0.0098	-0.0106	-0.0100	0.0158
	$k_n = 5, \ell_n = 10, \theta = 0.3$	-0.0025	-0.0033	0.0036	0.0025	0.0089	0.4577
	$k_n = 5, \ell_n = 10, \theta = 0.8$	-0.0031	-0.0057	-0.0041	-0.0057	-0.0049	0.0749
	$k_n = 5, \ell_n = 10, \theta = 1.5$	-0.0062	-0.0089	-0.0099	-0.0106	-0.0101	0.0122
	$k_n = 10, \ell_n = 5, \theta = 0.3$	-0.0025	-0.0036	0.0017	0.0028	0.0078	0.2769
	$k_n = 10, \ell_n = 5, \theta = 0.8$	-0.0031	-0.0057	-0.0044	-0.0057	-0.0050	0.0502
	$k_n = 10, \ell_n = 5, \theta = 1.5$	-0.0062	-0.0089	-0.0099	-0.0106	-0.0101	0.0052
	$k_n = 10, \ell_n = 10, \theta = 0.3$	-0.0024	-0.0038	0.0008	0.0027	0.0079	0.1965
	$k_n = 10, \ell_n = 10, \theta = 0.8$	-0.0031	-0.0057	-0.0046	-0.0057	-0.0050	0.0308
$k_n = 10, \ell_n = 10, \theta = 1.5$	-0.0062	-0.0089	-0.0100	-0.0106	-0.0101	-0.0009	
ReMeDI*	$k_n = 5, \ell_n = 5$	0.0147	0.0280	0.0471	0.0535	0.0419	0.7974
	$k_n = 5, \ell_n = 10$	0.0146	0.0278	0.0467	0.0520	0.0395	0.5844
	$k_n = 10, \ell_n = 5$	0.0126	0.0243	0.0344	0.0582	0.0398	0.2731
	$k_n = 10, \ell_n = 10$	0.0099	0.0205	0.0314	0.0641	0.0395	0.1507
TSRK	$G_n = 12, H_n = 76, \theta = 0.3$	-0.0011	-0.0016	0.0049	0.1229	0.1701	0.5945
	$G_n = 12, H_n = 76, \theta = 0.8$	-0.0029	-0.0054	-0.0038	0.0116	0.0193	0.1040
	$G_n = 12, H_n = 76, \theta = 1.5$	-0.0061	-0.0088	-0.0098	-0.0056	-0.0031	0.0213
	$G_n = 12, H_n = 153, \theta = 0.3$	-0.0010	-0.0016	0.0047	0.1228	0.1698	0.5850
	$G_n = 12, H_n = 153, \theta = 0.8$	-0.0029	-0.0054	-0.0038	0.0116	0.0192	0.1027
	$G_n = 12, H_n = 153, \theta = 1.5$	-0.0061	-0.0088	-0.0098	-0.0056	-0.0031	0.0209
	$G_n = 29, H_n = 76, \theta = 0.3$	-0.0006	-0.0020	0.0018	0.1228	0.1621	0.3427
	$G_n = 29, H_n = 76, \theta = 0.8$	-0.0028	-0.0054	-0.0042	0.0116	0.0182	0.0686
	$G_n = 29, H_n = 76, \theta = 1.5$	-0.0061	-0.0088	-0.0099	-0.0056	-0.0034	0.0112
	$G_n = 29, H_n = 153, \theta = 0.3$	-0.0004	-0.0018	0.0011	0.1224	0.1610	0.3141
$G_n = 29, H_n = 153, \theta = 0.8$	-0.0028	-0.0054	-0.0044	0.0115	0.0180	0.0646	
$G_n = 29, H_n = 153, \theta = 1.5$	-0.0061	-0.0088	-0.0099	-0.0057	-0.0034	0.0101	
RV	$\theta = 0.3$	-0.0052	-0.0046	0.0039	-0.0045	0.1343	0.7711
	$\theta = 0.8$	-0.0034	-0.0058	-0.0039	-0.0063	0.0143	0.1289
	$\theta = 1.5$	-0.0063	-0.0089	-0.0098	-0.0107	-0.0045	0.0283

Table 3: The relative biases of the pre-averaging estimators of IV with various debiasing methods. The bandwidth parameter of the pre-averaging estimator θ is selected from $\{0.3, 0.8, 1.5\}$. The tuning parameter of the LA and ReMeDI estimators k_n is selected from $\{5, 10\}$. The lag truncation parameter ℓ_n is selected from $\{5, 10\}$. The two bandwidths of TSRK are selected by $G_n \in \{[n^{1/4}], [n^{1/3}]\}$, $H_n \in \{[\sqrt{n}/2], [\sqrt{n}]\}$ with $n = 23400$. The method ReMeDI* follows the optimal rule of selecting θ in (16). The RV method refers to the realized volatility estimator of the variance of iid noise. The biases are obtained by taking the averages of 1000 simulations.

Specifications		$K_\gamma = 5 \times 10^{-5}$			$K_\gamma = 5 \times 10^{-4}$		
		$\varrho = 0$	0.3	0.8	$\varrho = 0$	0.3	0.8
LA	$k_n = 5, \ell_n = 5, \theta = 0.3$	0.6968	0.7366	0.6948	0.6945	0.6520	0.4778
	$k_n = 5, \ell_n = 5, \theta = 0.8$	0.2070	0.2123	0.1893	0.2001	0.1836	0.1812
	$k_n = 5, \ell_n = 5, \theta = 1.5$	0.2286	0.2446	0.2223	0.2221	0.2166	0.2332
	$k_n = 5, \ell_n = 10, \theta = 0.3$	1.0471	1.1112	1.0542	1.0471	0.9902	0.8598
	$k_n = 5, \ell_n = 10, \theta = 0.8$	0.2618	0.2730	0.2473	0.2593	0.2369	0.2274
	$k_n = 5, \ell_n = 10, \theta = 1.5$	0.2346	0.2512	0.2283	0.2299	0.2209	0.2379
	$k_n = 10, \ell_n = 5, \theta = 0.3$	1.4201	1.5057	1.4328	1.4168	1.3483	1.2738
	$k_n = 10, \ell_n = 5, \theta = 0.8$	0.2847	0.2971	0.2708	0.2824	0.2588	0.2501
	$k_n = 10, \ell_n = 5, \theta = 1.5$	0.2362	0.2527	0.2298	0.2318	0.2220	0.2391
	$k_n = 10, \ell_n = 10, \theta = 0.3$	2.1527	2.2860	2.1814	2.1496	2.0538	2.0992
	$k_n = 10, \ell_n = 10, \theta = 0.8$	0.4343	0.4588	0.4268	0.4356	0.4045	0.4134
	$k_n = 10, \ell_n = 10, \theta = 1.5$	0.2578	0.2758	0.2520	0.2566	0.2402	0.2606
ReMeDI	$k_n = 5, \ell_n = 5, \theta = 0.3$	0.1148	0.1100	0.0999	0.1038	0.0996	0.5988
	$k_n = 5, \ell_n = 5, \theta = 0.8$	0.1724	0.1732	0.1560	0.1590	0.1532	0.1977
	$k_n = 5, \ell_n = 5, \theta = 1.5$	0.2250	0.2406	0.2191	0.2164	0.2150	0.2326
	$k_n = 5, \ell_n = 10, \theta = 0.3$	0.1170	0.1117	0.1016	0.1048	0.1014	0.5470
	$k_n = 5, \ell_n = 10, \theta = 0.8$	0.1729	0.1736	0.1563	0.1593	0.1537	0.1919
	$k_n = 5, \ell_n = 10, \theta = 1.5$	0.2252	0.2407	0.2192	0.2165	0.2152	0.2324
	$k_n = 10, \ell_n = 5, \theta = 0.3$	0.1265	0.1200	0.1097	0.1115	0.1098	0.3521
	$k_n = 10, \ell_n = 5, \theta = 0.8$	0.1742	0.1747	0.1573	0.1603	0.1548	0.1821
	$k_n = 10, \ell_n = 5, \theta = 1.5$	0.2254	0.2410	0.2193	0.2167	0.2154	0.2323
	$k_n = 10, \ell_n = 10, \theta = 0.3$	0.1317	0.1258	0.1153	0.1157	0.1145	0.2726
	$k_n = 10, \ell_n = 10, \theta = 0.8$	0.1756	0.1762	0.1586	0.1616	0.1562	0.1780
	$k_n = 10, \ell_n = 10, \theta = 1.5$	0.2258	0.2414	0.2196	0.2171	0.2158	0.2327
ReMeDI*	$k_n = 5, \ell_n = 5$	0.0572	0.0699	0.0819	0.1952	0.1581	1.1294
	$k_n = 5, \ell_n = 10$	0.0572	0.0698	0.0817	0.2075	0.1578	0.8939
	$k_n = 10, \ell_n = 5$	0.0832	0.0776	0.0804	0.2372	0.1932	0.4154
	$k_n = 10, \ell_n = 10$	0.0862	0.0792	0.0815	0.3148	0.2462	0.2913
TSRK	$G_n = 12, H_n = 76, \theta = 0.3$	0.1160	0.1113	0.1016	0.1732	0.2096	0.7010
	$G_n = 12, H_n = 76, \theta = 0.8$	0.1728	0.1737	0.1564	0.1593	0.1553	0.2090
	$G_n = 12, H_n = 76, \theta = 1.5$	0.2251	0.2408	0.2192	0.2162	0.2150	0.2332
	$G_n = 12, H_n = 153, \theta = 0.3$	0.1170	0.1122	0.1019	0.1735	0.2097	0.6901
	$G_n = 12, H_n = 153, \theta = 0.8$	0.1733	0.1743	0.1570	0.1598	0.1558	0.2086
	$G_n = 12, H_n = 153, \theta = 1.5$	0.2253	0.2410	0.2194	0.2164	0.2152	0.2334
	$G_n = 29, H_n = 76, \theta = 0.3$	0.1326	0.1304	0.1194	0.1814	0.2099	0.4197
	$G_n = 29, H_n = 76, \theta = 0.8$	0.1761	0.1775	0.1599	0.1627	0.1585	0.1922
	$G_n = 29, H_n = 76, \theta = 1.5$	0.2260	0.2417	0.2200	0.2171	0.2159	0.2334
	$G_n = 29, H_n = 153, \theta = 0.3$	0.1435	0.1422	0.1278	0.1888	0.2155	0.3909
$G_n = 29, H_n = 153, \theta = 0.8$	0.1787	0.1805	0.1626	0.1657	0.1613	0.1930	
$G_n = 29, H_n = 153, \theta = 1.5$	0.2271	0.2429	0.2212	0.2182	0.2170	0.2345	
RV	$\theta = 0.3$	0.1107	0.1058	0.0958	0.0992	0.1730	0.9017
	$\theta = 0.8$	0.1719	0.1728	0.1555	0.1587	0.1537	0.2257
	$\theta = 1.5$	0.2249	0.2405	0.2190	0.2164	0.2148	0.2339

Table 4: The root-mean-squared-relative-errors (RMSRE) of the pre-averaging estimators of IV with various debiasing methods. The bandwidth parameter of the pre-averaging estimator θ is selected from $\{0.3, 0.8, 1.5\}$. The tuning parameter of the LA and ReMeDI estimators k_n is selected from $\{5, 10\}$. The lag truncation parameter ℓ_n is selected from $\{5, 10\}$. The two bandwidths of TSRK are selected by $G_n \in \{[n^{1/4}], [n^{1/3}]\}$, $H_n \in \{[\sqrt{n}/2], [\sqrt{n}]\}$ with $n = 23400$. The method ReMeDI* follows the optimal rule of selecting θ in (16). The RV method refers to the realized volatility estimator of the variance of i.i.d. noise. The number of simulations is 1000.

Specifications		$K_\gamma = 5 \times 10^{-4}$						$K_\gamma = 5 \times 10^{-5}$					
		$\rho = 0$			$\rho = 0.8$			$\rho = 0$			$\rho = 0.8$		
Estimators	time	t=5 min	100 min	200 min	t=5 min	100 min	200 min	t=5 min	100 min	200 min	t=5 min	100 min	200 min
	$\tilde{c}_t^n(\theta)$	$\theta = 0.5, l_n = 300, k_n = 5$	2.9929	1.2551	0.4644	10.8252	4.4852	1.6201	0.0277	0.0014	-0.0037	0.0999	0.0379
$\theta = 0.5, l_n = 300, k_n = 10$		3.2846	1.4325	0.5042	9.7096	4.1235	1.4464	0.0420	-0.0072	-0.0036	0.0894	0.0360	0.0055
$\theta = 0.5, l_n = 600, k_n = 5$		0.8507	0.3618	0.1570	6.5638	2.6499	1.0369	0.0202	0.0077	0.0122	0.0610	0.0285	0.0139
$\theta = 0.5, l_n = 600, k_n = 10$		0.9729	0.4378	0.1755	4.2781	1.7073	0.6728	0.0307	0.0114	0.0117	0.0308	0.0245	0.0097
$\theta = 0.5, l_n = 900, k_n = 5$		0.4420	0.1975	0.1091	5.6127	2.2338	0.9295	0.0386	0.0251	0.0420	0.0673	0.0347	0.0397
$\theta = 0.5, l_n = 900, k_n = 10$		0.5035	0.2356	0.1178	2.5477	0.9937	0.4379	0.0447	0.0249	0.0430	0.0304	0.0261	0.0339
$\theta = 1.0, l_n = 300, k_n = 5$		0.1976	0.0233	-0.0124	4.8267	1.9964	0.6946	-0.0249	-0.0486	-0.0380	0.0078	-0.0152	-0.0366
$\theta = 1.0, l_n = 300, k_n = 10$		0.2785	0.0833	0.0132	2.0806	0.8354	0.2739	-0.0156	-0.0407	-0.0453	-0.0338	-0.0258	-0.0472
$\theta = 1.0, l_n = 600, k_n = 5$		0.0375	0.0161	0.0102	3.6247	1.4515	0.5655	-0.0060	-0.0162	-0.0100	0.0078	-0.0034	-0.0136
$\theta = 1.0, l_n = 600, k_n = 10$		0.0816	0.0477	0.0207	1.5125	0.5704	0.2309	-0.0018	-0.0109	-0.0101	-0.0222	-0.0073	-0.0164
$\tilde{c}_t^m(\theta)$	$\theta = 0.5, l_n = 300, k_n = 5$	1.6577	1.3229	1.8756	6.8496	4.6818	4.1494	-0.3183	0.0396	0.9551	-0.2682	0.0753	0.9826
	$\theta = 0.5, l_n = 300, k_n = 10$	1.8437	1.5084	1.9567	6.1170	4.3088	3.8080	-0.3089	0.0308	0.9552	-0.2751	0.0739	0.9778
	$\theta = 0.5, l_n = 600, k_n = 5$	0.2340	0.4462	1.2385	4.0405	2.8822	2.9421	-0.3208	0.0746	0.9577	-0.2913	0.0960	0.9623
	$\theta = 0.5, l_n = 600, k_n = 10$	0.3150	0.5274	1.2745	2.5168	1.8804	2.2372	-0.3139	0.0788	0.9565	-0.3115	0.0920	0.9538
	$\theta = 0.5, l_n = 900, k_n = 5$	-0.0370	0.3043	1.0982	3.4190	2.5303	2.6493	-0.3069	0.1196	0.9703	-0.2861	0.1309	0.9678
	$\theta = 0.5, l_n = 900, k_n = 10$	0.0052	0.3462	1.1147	1.3717	1.1777	1.7188	-0.3028	0.1196	0.9719	-0.3108	0.1216	0.9566
	$\theta = 1.0, l_n = 300, k_n = 5$	-0.2025	0.0558	0.9430	2.8675	2.1068	2.3315	-0.3532	-0.0122	0.8879	-0.3292	0.0208	0.8947
	$\theta = 1.0, l_n = 300, k_n = 10$	-0.1473	0.1203	0.9945	1.0462	0.9011	1.5056	-0.3470	-0.0035	0.8738	-0.3567	0.0098	0.8727
	$\theta = 1.0, l_n = 600, k_n = 5$	-0.3081	0.0795	0.9555	2.0814	1.6084	2.0300	-0.3382	0.0493	0.9146	-0.3268	0.0623	0.9089
	$\theta = 1.0, l_n = 600, k_n = 10$	-0.2785	0.1133	0.9763	0.6738	0.6719	1.3821	-0.3354	0.0551	0.9140	-0.3468	0.0583	0.9035
$\tilde{c}_t^n(\theta)$	$\theta = 1.0, l_n = 900, k_n = 5$	-0.3185	0.1143	0.9427	1.5964	1.3440	1.8145	-0.3197	0.0984	0.9347	-0.3173	0.1025	0.9247
	$\theta = 1.0, l_n = 900, k_n = 10$	-0.3069	0.1285	0.9483	0.5385	0.6436	1.3332	-0.3182	0.0988	0.9367	-0.3305	0.0974	0.9197
	$\theta = 0.5, l_n = 300$	-0.3673	-0.1732	-0.1431	13.4635	5.5619	1.9213	-0.0974	-0.1014	-0.1025	0.0303	-0.0446	-0.0800
	$\theta = 0.5, l_n = 600$	-0.1499	-0.0622	-0.0353	10.9911	4.4271	1.6716	-0.0300	-0.0359	-0.0312	0.0663	0.0047	-0.0166
	$\theta = 0.5, l_n = 900$	-0.3075	-0.1154	-0.0340	9.2429	3.6339	1.4606	0.0013	-0.0052	0.0122	0.0761	0.0202	0.0198
	$\theta = 1.0, l_n = 300$	-0.3662	-0.1927	-0.1133	7.9083	3.2430	1.1320	-0.0529	-0.0669	-0.0606	0.0169	-0.0251	-0.0516
$\tilde{c}_t^n(\theta^*)$	$\theta = 1.0, l_n = 600$	-0.0995	-0.0414	-0.0188	5.5581	2.2251	0.8497	-0.0175	-0.0263	-0.0207	0.0179	-0.0063	-0.0192
	$\theta = 1.0, l_n = 900$	-0.0495	-0.0072	0.0105	4.2302	1.6658	0.6881	0.0110	-0.0014	0.0160	0.0273	0.0065	0.0131
	$l_n = 300, k_n = 5$	0.3130	0.2264	0.3038	5.0829	2.2021	0.8854	0.1077	0.0347	0.0011	0.0815	0.0070	-0.0270
	$l_n = 300, k_n = 10$	0.7391	0.3956	0.2191	2.1537	0.9083	0.3683	0.0093	-0.0430	-0.0957	-0.0360	-0.0798	-0.1173
	$l_n = 600, k_n = 5$	0.0712	0.1002	0.1758	4.1150	1.7924	0.8531	0.1161	0.0522	0.0174	0.1103	0.0431	0.0121
	$l_n = 600, k_n = 10$	0.2572	0.3541	0.2243	1.5173	0.5941	0.2868	0.0377	-0.0255	-0.0496	-0.0107	-0.0356	-0.0569
$\tilde{c}_t^n(\theta^*)$	$l_n = 900, k_n = 5$	0.0460	0.0887	0.1395	3.6760	1.6687	0.8681	0.0766	0.0444	0.0476	0.1279	0.0566	0.0531
	$l_n = 900, k_n = 10$	0.0948	0.1964	0.1630	1.3042	0.5220	0.2831	0.0221	-0.0238	-0.0121	0.0181	-0.0132	-0.0204

Table 5: The relative biases of SV estimators. The tuning parameter k_n of ReMeDI is either 5 or 10; the lag truncation parameter $\ell_n = \lfloor k_n/2 \rfloor$; the bandwidth of estimating SV $l_n \in \{300, 600, 900\}$; the bandwidth parameter of the pre-averaging method $\theta \in \{0.5, 1\}$. The biases are obtained by taking the averages of 1000 simulations.

Specifications		$K_\gamma = 5 \times 10^{-4}$						$K_\gamma = 5 \times 10^{-5}$					
		$\rho = 0$			$\rho = 0.8$			$\rho = 0$			$\rho = 0.8$		
Estimators	time	t=5 min	100 min	200 min	t=5 min	100 min	200 min	t=5 min	100 min	200 min	t=5 min	100 min	200 min
$\tilde{c}_t^n(\theta)$	$\theta = 0.5, l_n = 300, k_n = 5$	10.2533	1.5110	0.6538	35.6338	4.9207	1.5314	0.8429	0.4020	0.3958	1.0273	0.4228	0.4342
	$\theta = 0.5, l_n = 300, k_n = 10$	11.7552	1.7442	0.7559	32.6313	4.6333	1.4838	0.9451	0.4929	0.4627	1.2789	0.5095	0.4975
	$\theta = 0.5, l_n = 600, k_n = 5$	3.2248	0.5925	0.4217	21.4422	2.9061	1.0650	0.9232	0.3864	0.3480	0.9319	0.3723	0.3964
	$\theta = 0.5, l_n = 600, k_n = 10$	3.9983	0.7074	0.5278	14.4671	1.9820	0.8586	1.0315	0.4641	0.4159	1.0409	0.4408	0.4550
	$\theta = 0.5, l_n = 900, k_n = 5$	2.1451	0.4865	0.3840	18.2908	2.4454	0.9455	0.9543	0.3702	0.3415	0.9380	0.3419	0.3768
	$\theta = 0.5, l_n = 900, k_n = 10$	2.2003	0.5354	0.4636	8.8474	1.2827	0.6741	1.0811	0.4556	0.4128	1.0376	0.4432	0.4518
	$\theta = 1.0, l_n = 300, k_n = 5$	2.8836	0.7013	0.5328	17.3046	2.5104	0.9632	0.9662	0.5009	0.4832	1.2212	0.5361	0.5595
	$\theta = 1.0, l_n = 300, k_n = 10$	2.4928	0.8144	0.6505	8.8697	1.4571	0.8491	1.2232	0.6251	0.6398	1.5022	0.6716	0.7418
	$\theta = 1.0, l_n = 600, k_n = 5$	1.5653	0.4805	0.4523	12.3772	1.7450	0.7809	0.9800	0.4372	0.3931	0.9903	0.4201	0.4424
	$\theta = 1.0, l_n = 600, k_n = 10$	1.3754	0.5385	0.5494	6.3593	1.0390	0.7204	1.1059	0.5302	0.4820	1.1606	0.5101	0.5249
	$\theta = 1.0, l_n = 900, k_n = 5$	1.1914	0.4394	0.4081	9.6949	1.3854	0.6659	1.0062	0.4091	0.3702	0.9465	0.3955	0.4126
	$\theta = 1.0, l_n = 900, k_n = 10$	1.2310	0.4856	0.4656	5.3041	0.9275	0.6091	1.0905	0.4709	0.4235	1.0455	0.4640	0.4663
$\tilde{c}_t^m(\theta)$	$\theta = 0.5, l_n = 300, k_n = 5$	6.0482	1.5908	2.4133	22.5707	5.1358	4.2378	1.0436	0.4169	1.6931	1.1593	0.4600	1.7817
	$\theta = 0.5, l_n = 300, k_n = 10$	7.0334	1.8248	2.5347	20.6606	4.8320	4.0787	1.1067	0.5060	1.8200	1.2940	0.5542	1.8803
	$\theta = 0.5, l_n = 600, k_n = 5$	1.5936	0.6808	1.7830	13.2549	3.1703	3.2578	1.1150	0.4152	1.6641	1.2115	0.4203	1.7312
	$\theta = 0.5, l_n = 600, k_n = 10$	2.1423	0.8043	1.8190	8.6570	2.1834	2.7963	1.1586	0.5051	1.6863	1.3226	0.4861	1.7729
	$\theta = 0.5, l_n = 900, k_n = 5$	1.1855	0.5806	1.6792	11.2048	2.7858	2.9448	1.0872	0.4232	1.6353	1.1886	0.4092	1.6675
	$\theta = 0.5, l_n = 900, k_n = 10$	1.1588	0.6372	1.6854	4.9886	1.5025	2.3332	1.1336	0.5186	1.6757	1.3087	0.5071	1.7353
	$\theta = 1.0, l_n = 300, k_n = 5$	2.0112	0.7287	1.7341	10.5098	2.6428	2.9504	1.1735	0.5087	1.7069	1.4392	0.5686	1.8204
	$\theta = 1.0, l_n = 300, k_n = 10$	1.7401	0.8497	1.7870	5.1488	1.5406	2.4661	1.2629	0.6412	1.8034	1.6476	0.7019	2.0551
	$\theta = 1.0, l_n = 600, k_n = 5$	1.4226	0.5262	1.5693	7.2729	1.9309	2.6034	1.1837	0.4634	1.6343	1.3466	0.4562	1.7038
	$\theta = 1.0, l_n = 600, k_n = 10$	1.2843	0.5945	1.6251	3.5408	1.1569	2.2836	1.2289	0.5693	1.6880	1.4729	0.5468	1.7873
	$\theta = 1.0, l_n = 900, k_n = 5$	1.2523	0.4954	1.5383	5.5141	1.6202	2.3333	1.1383	0.4566	1.6015	1.2939	0.4477	1.6506
	$\theta = 1.0, l_n = 900, k_n = 10$	1.2498	0.5475	1.5620	2.8411	1.0888	2.0802	1.1675	0.5256	1.6417	1.3699	0.5171	1.7067
$\tilde{c}_t^n(\theta)$	$\theta = 0.5, l_n = 300$	5.9620	0.9513	0.5103	43.4563	5.8910	1.6880	0.7300	0.3587	0.3591	0.7258	0.3471	0.3662
	$\theta = 0.5, l_n = 600$	2.2998	0.4921	0.3731	35.0967	4.6301	1.4869	0.8080	0.3385	0.2860	0.7698	0.2926	0.3205
	$\theta = 0.5, l_n = 900$	1.7085	0.4105	0.3330	29.3890	3.7951	1.3026	0.8649	0.3355	0.2934	0.8622	0.2822	0.3241
	$\theta = 1.0, l_n = 300$	2.4579	0.6153	0.4689	26.1134	3.5931	1.1303	0.8396	0.4219	0.4143	0.9903	0.4502	0.4684
	$\theta = 1.0, l_n = 600$	1.1813	0.4258	0.4193	18.0055	2.4169	0.9001	0.9190	0.4035	0.3540	0.8968	0.3764	0.4027
	$\theta = 1.0, l_n = 900$	0.9961	0.4069	0.3861	13.6466	1.8301	0.7518	0.9704	0.3922	0.3488	0.9060	0.3691	0.3906
$\tilde{c}_t^n(\theta^*)$	$l_n = 300, k_n = 5$	5.0396	2.2299	1.6223	18.8126	2.9399	1.2368	1.1042	0.4344	0.3391	1.0523	0.3723	0.3859
	$l_n = 300, k_n = 10$	13.0478	2.3067	1.2867	9.5457	1.7432	1.0133	1.1026	0.5307	0.5111	1.2701	0.5062	0.5176
	$l_n = 600, k_n = 5$	1.7906	1.0327	0.7855	14.6056	2.3326	1.1077	0.9590	0.3319	0.2657	0.9200	0.2810	0.3062
	$l_n = 600, k_n = 10$	4.9159	1.5242	0.8405	6.3915	1.1738	0.7840	1.0240	0.4222	0.3804	1.0351	0.3889	0.3929
	$l_n = 900, k_n = 5$	1.4584	0.5897	0.5083	12.9737	2.1177	1.0434	0.8944	0.3357	0.2980	0.9471	0.2773	0.3151
	$l_n = 900, k_n = 10$	1.7322	0.7971	0.5636	5.3267	0.9845	0.6751	0.9507	0.3817	0.3294	0.9388	0.3570	0.3714

Table 6: The root-mean-squared-relative-errors (RMSRE) of SV estimators. The tuning parameter k_n of ReMeDI is either 5 or 10; the lag truncation parameter $\ell_n = \lfloor k_n/2 \rfloor$; the bandwidth of estimating SV $l_n \in \{300, 600, 900\}$; the bandwidth parameter of the pre-averaging method $\theta \in \{0.5, 1\}$. The number of simulations is 1000.

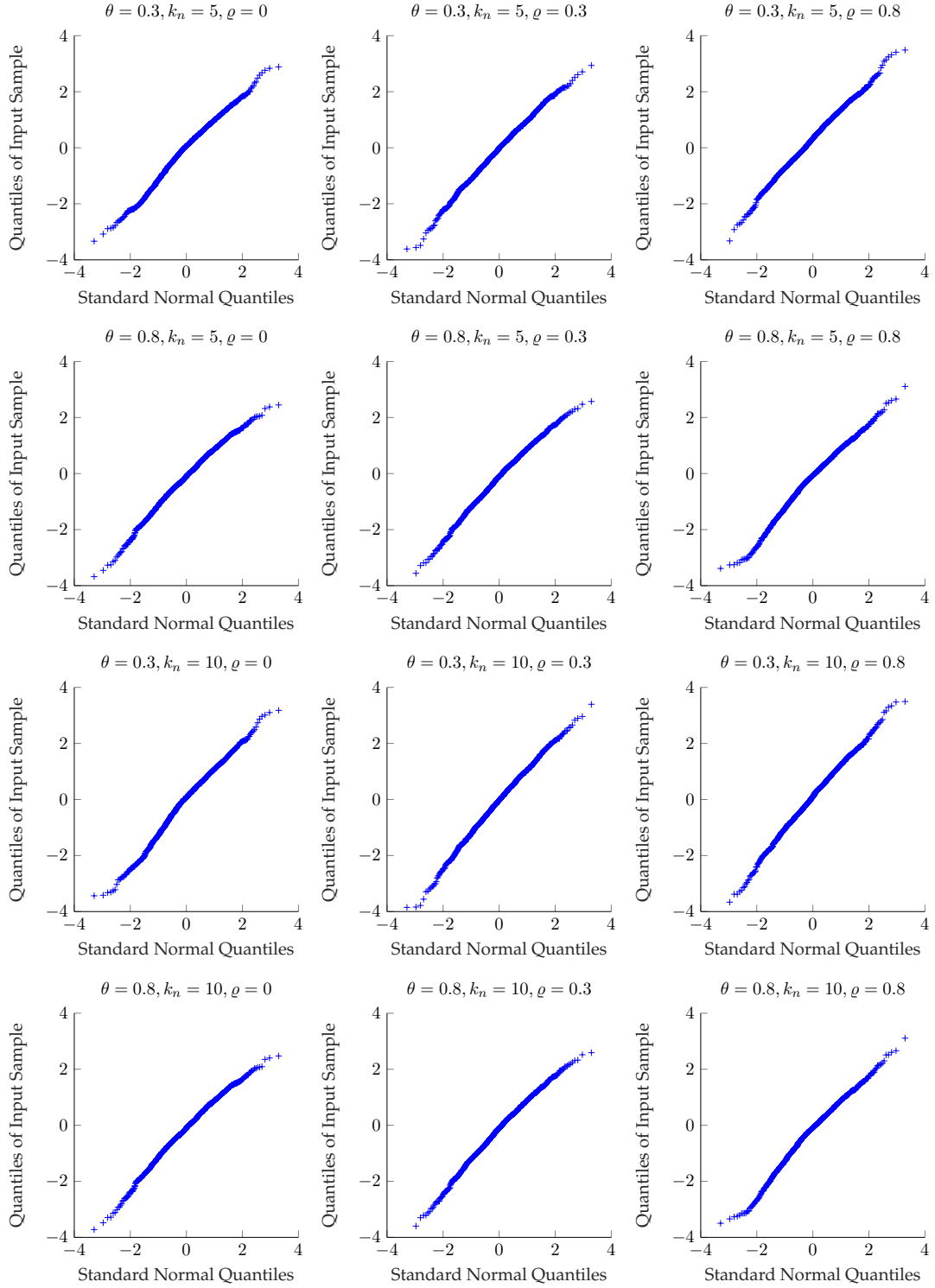


Figure 1: QQ Plot of $(\widehat{C}_t^n - C_t)/\sqrt{\Sigma_t^n}$ versus Standard Normal. The scale of noise is fixed at $K_\gamma = 10^{-4}$, the AR(1) coefficient of noise $\rho \in \{0, 0.3, 0.8\}$. The tuning parameters are selected as follows: $\theta \in \{0.3, 0.8\}, k_n \in \{5, 10\}, \ell_n = \lfloor k_n/2 \rfloor$. The number of simulations is 1000.

7 Empirical Studies

We use the intraday transaction prices of two individual stocks, the Coca-Cola Co (KO) and the General Electric Company (GE). The transaction prices are obtained from the Trade and Quote (TAQ) database for January 2015. There are 20 trading days.

7.1 IV estimation

Figure 2 presents and compares the IV estimates of several estimators. The top panel of Figure 2 plots the estimates of FRK, PaReMeDI, QMLE and TSRK.¹⁷ The estimates are very close to each other on most of the trading days. Specifically, the PaReMeDI and QMLE generate virtually identical estimates of the IV; most estimates are within the 95% confidence intervals of PaReMeDI, whence they are statistically indistinguishable.¹⁸ The bottom panel of Figure 2 illustrates the sensitivity of PaLA to the choice of θ . When θ is small, PaLA tends to underestimate the IVs due to the errors in the LA corrections. This is consistent with our observations in the simulation studies. When θ becomes large, the over-smoothing effect dominates the errors in the LA corrections and the estimates become closer to the PaReMeDI estimates, which we believe to provide accurate proxies of the true IVs.

It is worth mentioning that there are several trading days where the discrepancies among the estimators are large (the 2nd and 19th trading days for KO and the 15th trading day for GE). However, QMLE and PaReMeDI still conform to each other closely. One possible explanation is that both estimators have a quite small root mean squared errors, as they often do in the simulation studies.

7.2 SV estimation

We estimate SV by three estimators studied in the simulation section: \hat{c}_t^n defined in (9), \tilde{c}_t^n defined in (14) and \hat{c}_t^n , where \hat{c}_t^n is the pre-averaging estimator of SV with a realized volatility correction of the i.i.d. noise, see Chapter 8 of Aït-Sahalia and Jacod (2014) and the recent work by Li et al. (2022).¹⁹

The average spot volatility estimates are presented in Figure 3. In the top panel, we observe the two SV estimates \hat{c}_t^n and \tilde{c}_t^n are almost in line with each other, indicating that an i.i.d. noise assumption is acceptable for the transaction prices of KO in the sample. However, \hat{c}_t^n exceeds \tilde{c}_t^n in almost each minute in the bottom panel for GE. The systematic differences suggest the inadequacy of the simple i.i.d. assumption of noise. Both estimators display a weakly U-shaped or reverse J-shaped pattern for the intraday volatility that is well documented in the

¹⁷In this empirical study, we do not consider the robustness of jumps.

¹⁸We notice that Oh and Kim (2021) study the U.S.—China trade war based on a volatility contagion model, and the model is estimated by QMLE. The PaReMeDI estimator is used as an alternative measure of IV, and it is found that the QMLE and PaReMeDI estimates are quite close.

¹⁹Note that Aït-Sahalia and Jacod (2014) and Li et al. (2022) only consider the regular observation scheme. The estimator \hat{c}_t^n used here is normalized by the differences of real observation times, recall the discussion in Remark 4.2. Thus, the persistent differences of \hat{c}_t^n and \tilde{c}_t^n —if there are any—are caused by the misspecification of i.i.d. noise.

literature (Andersen and Bollerslev, 1997; Wood et al., 1985).

Now we compare \tilde{c}_t^n and $\tilde{c}_t^{\prime n}$ and examine the consequences of ignoring the stochastic observation schemes. As is evident in the top panel of Figure 3, the differences of \tilde{c}_t^n and $\tilde{c}_t^{\prime n}$ are quite persistent: the latter is noticeably lower near the opening and near the close compared with the volatility during the middle of the day. The SV estimates for GE exhibit a similar pattern, albeit the differences are less substantial. The differences tie in closely with our theoretical analysis and the numerical studies—the observation density α_t plays a key role to cause such a difference. We plot the average number of transactions in each minute in Figure 4. The plot serves as a proxy of the process α_t of each stock. Indeed, we observe a clear U-pattern for KO, and an inverse J-pattern for GE. Thus, our findings align well with the well-documented U-shape in volatility and trading volume in the literature.²⁰ However, we should be cautious—we will get a different intraday volatility pattern if we do not account for the irregular observation scheme.

Figure 4 represents the average patterns over 20 trading days. It is of interest to directly compare the daily estimates of SV at different trading times by different estimators. Figure 5 displays the SV estimates by \tilde{c}_t^n and $\tilde{c}_t^{\prime n}$ at 9:35, 11:55 and 15:55. Some immediate observations confirm the patterns we find in Figure 4: the estimates by $\tilde{c}_t^{\prime n}$ are almost always smaller than the estimates by \tilde{c}_t^n when the market opens and closes. However, the two estimates around noon are indistinguishable.

8 Conclusion

We introduce a new class of IV and SV estimators using noisy high-frequency data. The estimators are applicable in a broad setting of microstructure noise and observation schemes. Compared to alternative estimators recently proposed, our estimators provides accurate and robust estimates, and are computationally efficient. We also demonstrate the consequences of neglecting the complexities of the noise component and the observation scheme. The paper also initiates several future research projects. Since the pre-averaging method coupled with ReMeDI correction can be directly applied to tick data, it opens the possibility to test for jumps in tick data in the spirit of Aït-Sahalia et al. (2012); it is an interesting approach since jumps should be identified in samples of the highest possible frequencies (Christensen et al., 2014). Asymptotic efficiency is not discussed in details. Adaptive estimator (Jacod and Mykland, 2015) can be developed to improve efficiency. We leave these open questions for future research.

²⁰The results and the patterns are robust to the choices of tuning parameters.

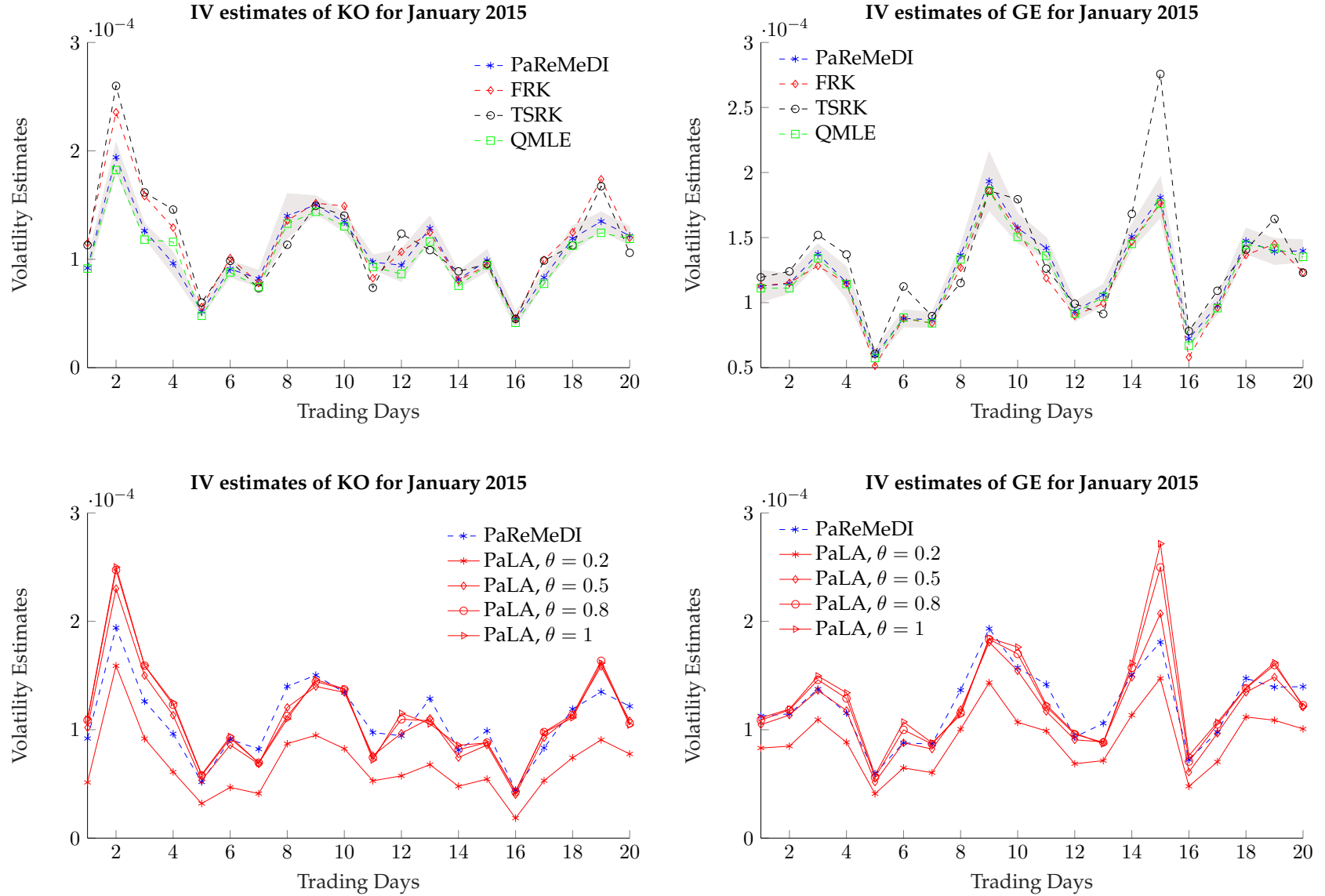


Figure 2: Integrated Volatility (IV) estimates of KO and GE for January, 2015. The tuning parameters are set as follows: For PaReMeDI, the tuning parameter of ReMeDI is $k_n = 10$ and the number of autocovariances included in the estimation of the long-run variance $\ell_n = 5$; θ are selected according to (16). For FRK, we set $f_n = 15, b_n = 75, j_n = 5$. For PaLA, we set $k_n = 6, \ell_n = 5, \theta \in \{0.2, 0.5, 0.8, 1\}$. For TSRK, we set $G_n = [n^{1/3}], H_n = [n^{1/2}]$ where n is the number of observations. For the QMLE (with AIC selection criterion), the optimal q is selected within $\{5, 6, 7, 8, 9, 10\}$. The shaded areas in the top panel represent the 95% confidence intervals of the PaReMeDI estimators.

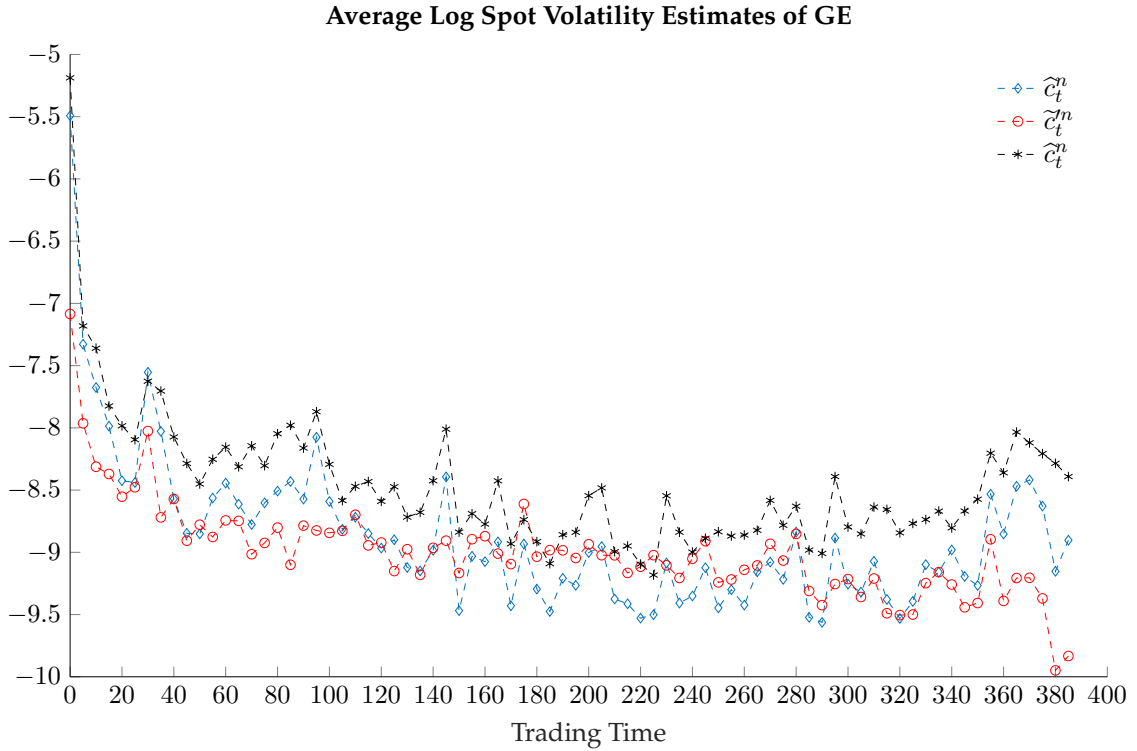
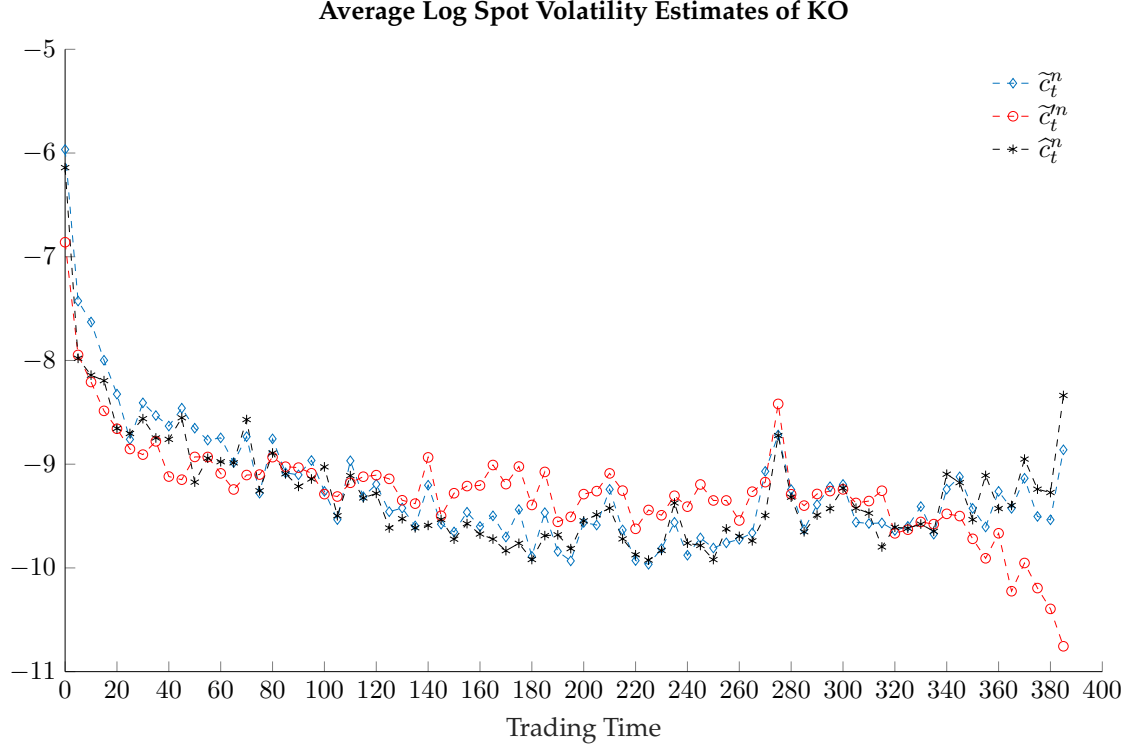


Figure 3: Average logarithmic spot volatility estimates for KO (top panel) and GE (bottom panel). The average is taken over 20 trading days in January 2015. The trading time 0 (and 390) corresponds to 9:30 (and 16:00). SV is estimated at each minute from 9:30 to 16:00. The estimators \tilde{c}_t^n and \tilde{c}_t^n are defined in (9) and (14). We set $k_n = 10, \ell_n = 5, l_n = 300$ (for KO) and $l_n = 600$ (for GE). The choices of l_n is approximately equal to the number of observations in 2 minutes for each stock. θ is selected via (17). \tilde{c}_t^n is the pre-averaging estimator of SV with a realized volatility correction of the i.i.d. noise.

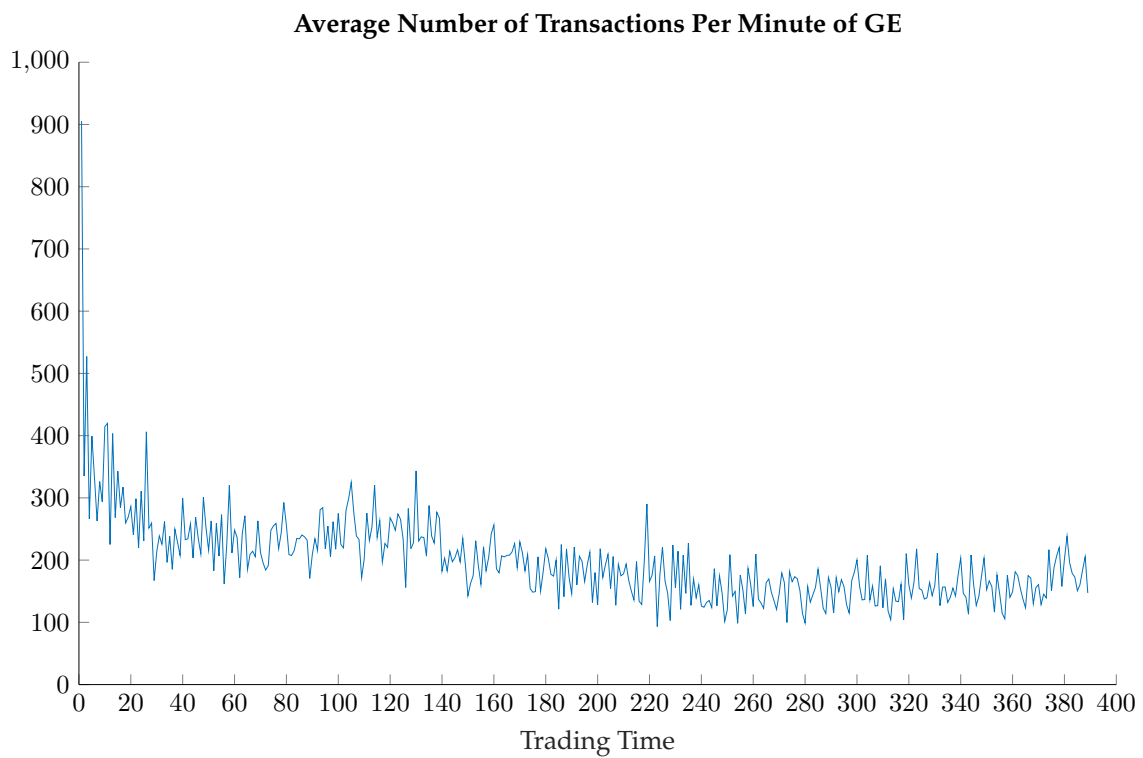
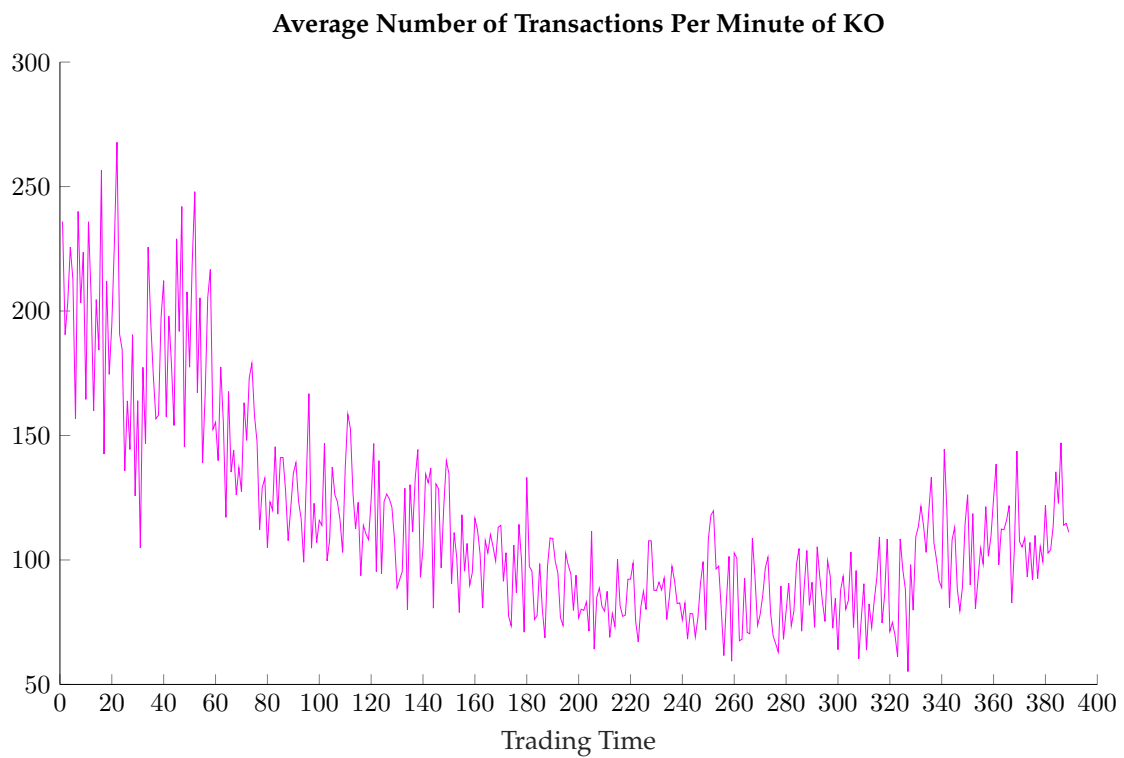


Figure 4: The Average number of transactions per minute for KO (top panel) and GE (bottom panel). The averages are taken over 20 trading days in January 2015.

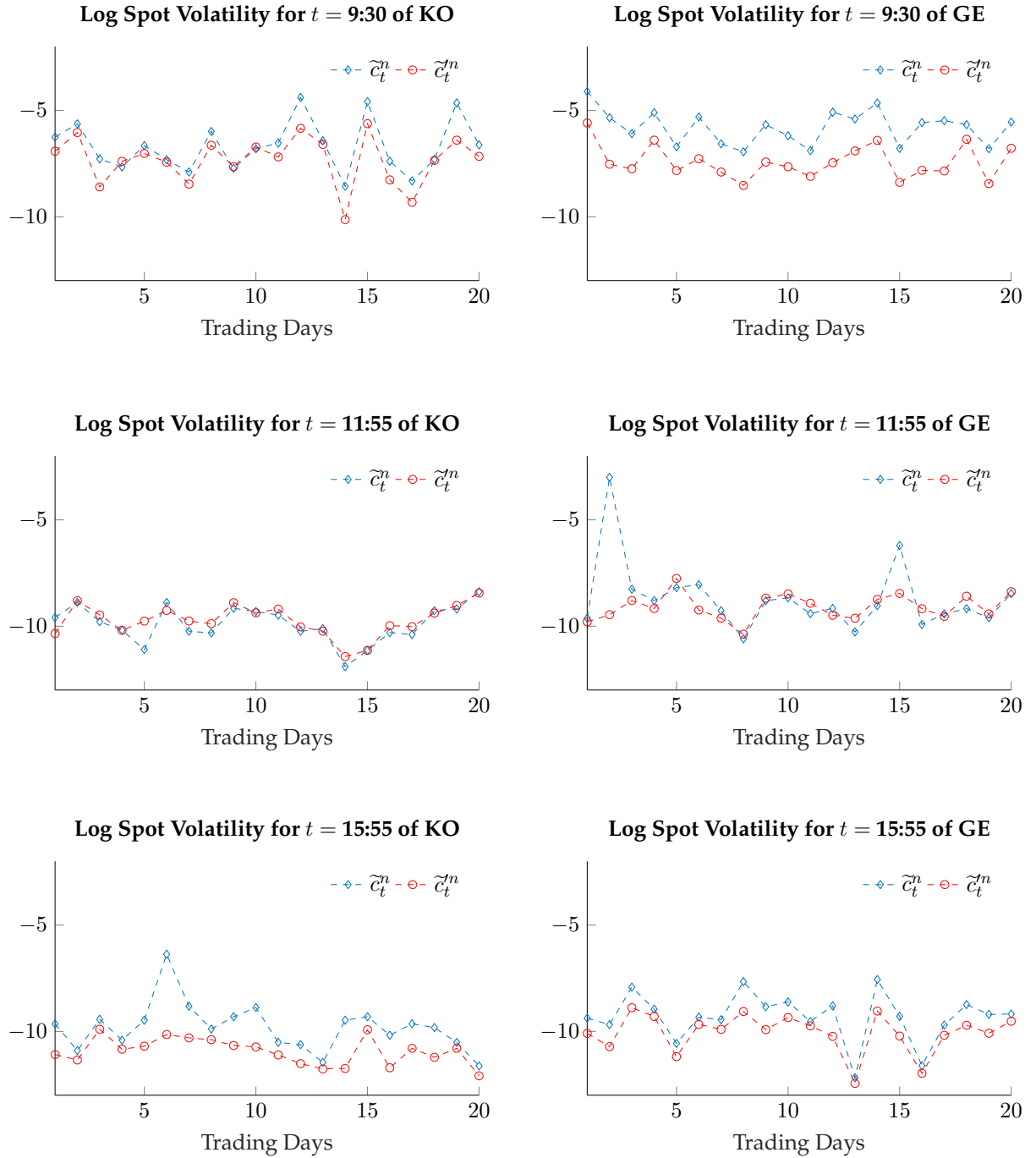


Figure 5: Logarithmic spot volatility estimates at 9:30 (top panel), 11:55 (middle panel) and 15:55 (bottom panel) of KO (left panel) and GE (right panel). The estimators \tilde{c}_t^n and \tilde{c}_t^n are defined in (9) and (14). We set $k_n = 10$, $\ell_n = 5$, $l_n = 300$ (for KO) and $l_n = 600$ (for GE). The choices of l_n are approximately equal to the number of observations in 2 minutes for each stock. θ is selected via (17).

References

- AÏT-SAHALIA, Y. AND J. JACOD (2009): "Testing for jumps in a discretely observed process," *Annals of Statistics*, 184–222.
- (2014): *High-Frequency Financial Econometrics*, Princeton University Press.
- AÏT-SAHALIA, Y., J. JACOD, AND J. LI (2012): "Testing for jumps in noisy high frequency data," *Journal of Econometrics*, 168, 207–222.
- AÏT-SAHALIA, Y., P. A. MYKLAND, AND L. ZHANG (2005): "How often to sample a continuous-time process in the presence of market microstructure noise," *Review of Financial Studies*, 18, 351–416.
- (2011): "Ultra high frequency volatility estimation with dependent microstructure noise," *Journal of Econometrics*, 160, 160–175.
- ANDERSEN, T. G. AND T. BOLLERSLEV (1997): "Intraday periodicity and volatility persistence in financial markets," *Journal of Empirical Finance*, 4, 115–158.
- ANDERSEN, T. G., T. BOLLERSLEV, F. X. DIEBOLD, AND P. LABYS (2003): "Modeling and forecasting realized volatility," *Econometrica*, 71, 579–625.
- ANDERSEN, T. G., M. THYRSGAARD, AND V. TODOROV (2021): "Recalcitrant betas: Intraday variation in the cross-sectional dispersion of systematic risk," *Quantitative Economics*, 12, 647–682.
- ANDREWS, D. W. (1991): "Heteroskedasticity and autocorrelation consistent covariance matrix estimation," *Econometrica*, 817–858.
- BARNDORFF-NIELSEN, O. E., P. R. HANSEN, A. LUNDE, AND N. SHEPHARD (2008): "Designing realized kernels to measure the ex post variation of equity prices in the presence of noise," *Econometrica*, 76, 1481–1536.
- BIBINGER, M. AND L. WINKELMANN (2018): "Common price and volatility jumps in noisy high-frequency data," *Electronic Journal of Statistics*, 12.
- BOGOUSLAVSKY, V. (2016): "Infrequent rebalancing, return autocorrelation, and seasonality," *Journal of Finance*, 71, 2967–3006.
- BOLLERSLEV, T., J. LI, AND Z. LIAO (2021): "Fixed-k inference for volatility," *Quantitative Economics*, 12, 1053–1084.
- CAMPBELL, J. Y., S. J. GROSSMAN, AND J. WANG (1993): "Trading volume and serial correlation in stock returns," *Quarterly Journal of Economics*, 108, 905–939.
- CHAN, L. K. AND J. LAKONISHOK (1995): "The behavior of stock prices around institutional trades," *Journal of Finance*, 50, 1147–1174.

- CHRISTENSEN, K., R. C. OOMEN, AND M. PODOLSKIJ (2014): “Fact or friction: Jumps at ultra high frequency,” *Journal of Financial Economics*, 114, 576–599.
- CHRISTENSEN, K., M. PODOLSKIJ, AND M. VETTER (2013): “On covariation estimation for multivariate continuous Itô semimartingales with noise in non-synchronous observation schemes,” *Journal of Multivariate Analysis*, 120, 59–84.
- DA, R. AND D. XIU (2021a): “Disentangling Autocorrelated Intraday Returns,” *Chicago Booth Research Paper*.
- (2021b): “When Moving-Average Models Meet High-Frequency Data: Uniform Inference on Volatility,” *Econometrica*, 89, 2787–2825.
- DELBAEN, F. AND W. SCHACHERMAYER (1994): “A general version of the fundamental theorem of asset pricing,” *Mathematische Annalen*, 300, 463–520.
- FAN, J. AND Y. WANG (2008): “Spot volatility estimation for high-frequency data,” *Statistics and Its Interface*, 1, 279–288.
- FOSTER, D. P. AND D. B. NELSON (1996): “Continuous Record Asymptotics for Rolling Sample Variance Estimators,” *Econometrica*, 139–174.
- GROSSMAN, S. J. AND M. H. MILLER (1988): “Liquidity and market structure,” *Journal of Finance*, 43, 617–633.
- HANSEN, P. R. AND A. LUNDE (2006): “Realized variance and market microstructure noise,” *Journal of Business & Economic Statistics*, 24, 127–161.
- HASBROUCK, J. (1993): “Assessing the quality of a security market: A new approach to transaction-cost measurement,” *Review of Financial Studies*, 6, 191–212.
- HAUTSCH, N. AND M. PODOLSKIJ (2013): “Preaveraging-Based Estimation of Quadratic Variation in the Presence of Noise and Jumps: Theory, Implementation, and Empirical Evidence,” *Journal of Business & Economic Statistics*, 31, 165–183.
- HENDERSHOTT, T. AND A. J. MENKVELD (2014): “Price pressures,” *Journal of Financial Economics*, 114, 405–423.
- HENDERSHOTT, T., A. J. MENKVELD, R. PRAZ, AND M. SEASHOLES (2022): “Asset price dynamics with limited attention,” *Review of Financial Studies*, 35, 962–1008.
- HO, T. AND H. R. STOLL (1981): “Optimal dealer pricing under transactions and return uncertainty,” *Journal of Financial Economics*, 9, 47–73.
- IKEDA, S. S. (2015): “Two-scale realized kernels: A univariate case,” *Journal of Financial Econometrics*, 13, 126–165.

- (2016): “A bias-corrected estimator of the covariation matrix of multiple security prices when both microstructure effects and sampling durations are persistent and endogenous,” *Journal of Econometrics*, 193, 203–214.
- JACOD, J. (2018): “Limit of random measures associated with the increments of a Brownian semimartingale,” *Journal of Financial Econometrics*, 16, 526–569.
- JACOD, J., Y. LI, P. A. MYKLAND, M. PODOLSKIJ, AND M. VETTER (2009): “Microstructure noise in the continuous case: the pre-averaging approach,” *Stochastic Processes and their Applications*, 119, 2249–2276.
- JACOD, J., Y. LI, AND X. ZHENG (2017): “Statistical properties of microstructure noise,” *Econometrica*, 85, 1133–1174.
- (2019): “Estimating the integrated volatility with tick observations,” *Journal of Econometrics*, 208, 80–100.
- JACOD, J. AND P. A. MYKLAND (2015): “Microstructure noise in the continuous case: Approximate efficiency of the adaptive pre-averaging method,” *Stochastic Processes and their Applications*, 125, 2910–2936.
- JACOD, J., M. PODOLSKIJ, AND M. VETTER (2010): “Limit theorems for moving averages of discretized processes plus noise,” *Annals of Statistics*, 38, 1478–1545.
- JACOD, J. AND P. PROTTER (1998): “Asymptotic error distributions for the Euler method for stochastic differential equations,” *Annals of Probability*, 267–307.
- JACOD, J. AND P. E. PROTTER (2011): *Discretization of Processes*, vol. 67, Springer Science & Business Media.
- JACOD, J. AND M. ROSENBAUM (2013): “Quarticity and other functionals of volatility: efficient estimation,” *Annals of Statistics*, 41, 1462–1484.
- JACOD, J. AND A. N. SHIRYAEV (2003): *Limit Theorems for Stochastic Processes*, vol. 288, Springer-Verlag Berlin.
- JACOD, J. AND V. TODOROV (2009): “Testing for common arrivals of jumps for discretely observed multidimensional processes,” *Annals of Statistics*, 37, 1792–1838.
- KALNINA, I. AND O. LINTON (2008): “Estimating quadratic variation consistently in the presence of endogenous and diurnal measurement error,” *Journal of Econometrics*, 147, 47–59.
- KRISTENSEN, D. (2010): “Nonparametric filtering of the realized spot volatility: A kernel-based approach,” *Econometric Theory*, 26, 60–93.
- LI, J. (2013): “Robust Estimation and Inference for Jumps in Noisy High Frequency Data: A Local-to-Continuity Theory for the Pre-Averaging Method,” *Econometrica*, 81, 1673–1693.

- LI, Y., G. LIU, AND Z. ZHANG (2022): “Volatility of volatility: Estimation and tests based on noisy high frequency data with jumps,” *Journal of Econometrics*, 229, 422–451.
- LI, Y., P. A. MYKLAND, E. RENAULT, L. ZHANG, AND X. ZHENG (2014): “Realized volatility when sampling times are possibly endogenous,” *Econometric theory*, 30, 580–605.
- LI, Z. M., R. J. LAEVEN, AND M. H. VELLEKOOP (2020): “Dependent microstructure noise and integrated volatility estimation from high-frequency data,” *Journal of Econometrics*, 215, 536–558.
- LI, Z. M. AND O. LINTON (2022a): “A ReMeDI for microstructure noise,” *Econometrica*, 90, 367–389.
- (2022b): “Supplementary Materials for “Robust Estimation of Integrated and Spot Volatility”,” .
- MADHAVAN, A., M. RICHARDSON, AND M. ROOMANS (1997): “Why do security prices change? A transaction-level analysis of NYSE stocks,” *Review of Financial Studies*, 10, 1035–1064.
- MALLIAVIN, P. AND M. E. MANCINO (2009): “A Fourier transform method for nonparametric estimation of multivariate volatility,” *Annals of Statistics*, 37, 1983–2010.
- MANCINI, C. (2001): “Disentangling the jumps of the diffusion in a geometric jumping Brownian motion,” *Giornale dell’Istituto Italiano degli Attuari*, 64, 19–47.
- OH, M. AND D. KIM (2021): “Effect of the US–China Trade War on Stock Markets: A Financial Contagion Perspective,” *arXiv preprint arXiv:2111.09655*.
- PARLOUR, C. A. (1998): “Price dynamics in limit order markets,” *Review of Financial Studies*, 11, 789–816.
- PODOLSKIJ, M. AND M. VETTER (2009): “Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps,” *Bernoulli*, 15, 634–658.
- SHEPHARD, N. AND D. XIU (2017): “Econometric analysis of multivariate realised QML: efficient positive semi-definite estimators of the covariation of equity prices.” *Journal of Econometrics*, 201, 19–42.
- TODOROV, V. AND G. TAUCHEN (2011): “Volatility jumps,” *Journal of Business & Economic Statistics*, 29, 356–371.
- VARNESKOV, R. T. (2016): “Flat-top realized kernel estimation of quadratic covariation with nonsynchronous and noisy asset prices,” *Journal of Business & Economic Statistics*, 34, 1–22.
- (2017): “Estimating the quadratic variation spectrum of noisy asset prices using generalized flat-top realized kernels,” *Econometric Theory*, 33, 1457–1501.

- WOOD, R. A., T. H. MCINISH, AND J. K. ORD (1985): "An investigation of transactions data for NYSE stocks," *Journal of Finance*, 40, 723–739.
- XIU, D. (2010): "Quasi-maximum likelihood estimation of volatility with high frequency data," *Journal of Econometrics*, 159, 235–250.
- ZHANG, L. (2006): "Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach," *Bernoulli*, 12, 1019–1043.
- ZHANG, L., P. A. MYKLAND, AND Y. AÏT-SAHALIA (2005): "A tale of two time scales: determining integrated volatility with noisy high-frequency data," *Journal of the American Statistical Association*, 100, 1394–1411.
- ZU, Y. AND H. P. BOSWIJK (2014): "Estimating spot volatility with high-frequency financial data," *Journal of Econometrics*, 181, 117–135.

Supplementary Materials for “Robust Estimation of Integrated and Spot Volatility”

Z. Merrick Li* Oliver Linton†

August 8, 2022

A Additional Simulation Studies

A.1 Examine the feasible CLTs

Figure A.1 replicates Figure 1 in Li and Linton (2022a) with a different scale of noise $K_\gamma = 5 \times 10^{-5}$. The plots strongly support the robustness of the feasible CLT of the IV estimator. Figure A.2 presents the QQ plot of the standardized SV estimators: $(\tilde{c}_t^n - c_t)/\sqrt{\tilde{\Sigma}_t^n}$. We observe some distortions: the standardized statistics have larger dispersions at $t = 5$ and $t = 100$.¹ One possible explanation is that both the volatility and the noise have symmetric U-patterns; thus, the spot volatility and the spot noise scale at $t = 5$ and $t = 100$ are more prominent than $t = 200$ (recall that the maximal $t = 390$).

A.2 Robustness to jumps

Now we include jumps in our simulation design by setting $\xi_{1,t} \sim \mathcal{N}(0, \mu_2/10)$, recall the simulation setting in Section 6.1 of Li and Linton (2022a). We set the truncation levels at $w_n = 10\mu_2, w_n = 20\mu_2$. We compare with the Medium Blocked Realized Kernels (MBRK) (Varneskov, 2016, 2017) that are robust to finite-activity jumps. Table A.1 and Table A.2 report the relative bias and RMSRE. We observe that both estimators work well to eliminate the effects of jumps in IV estimation. MBRK are quite robust to the choices of the tuning parameters. PaReMeDI tends to have smaller biases and RMSREs except for the estimator with $\theta = 1.5$.

*Corresponding author. Department of Economics, the Chinese University of Hong Kong, Esther Lee Building, Shatin, New Territories, Hong Kong SAR. Email: merrickli@cuhk.edu.hk

†Faculty of Economics, University of Cambridge, Austin Robinson Building, Sidgwick Avenue, Cambridge, CB3 9DD, United Kingdom. Email: obl20@cam.ac.uk

¹We obtain similar QQ plots when the parameters ρ, K_γ, θ and l_n vary.

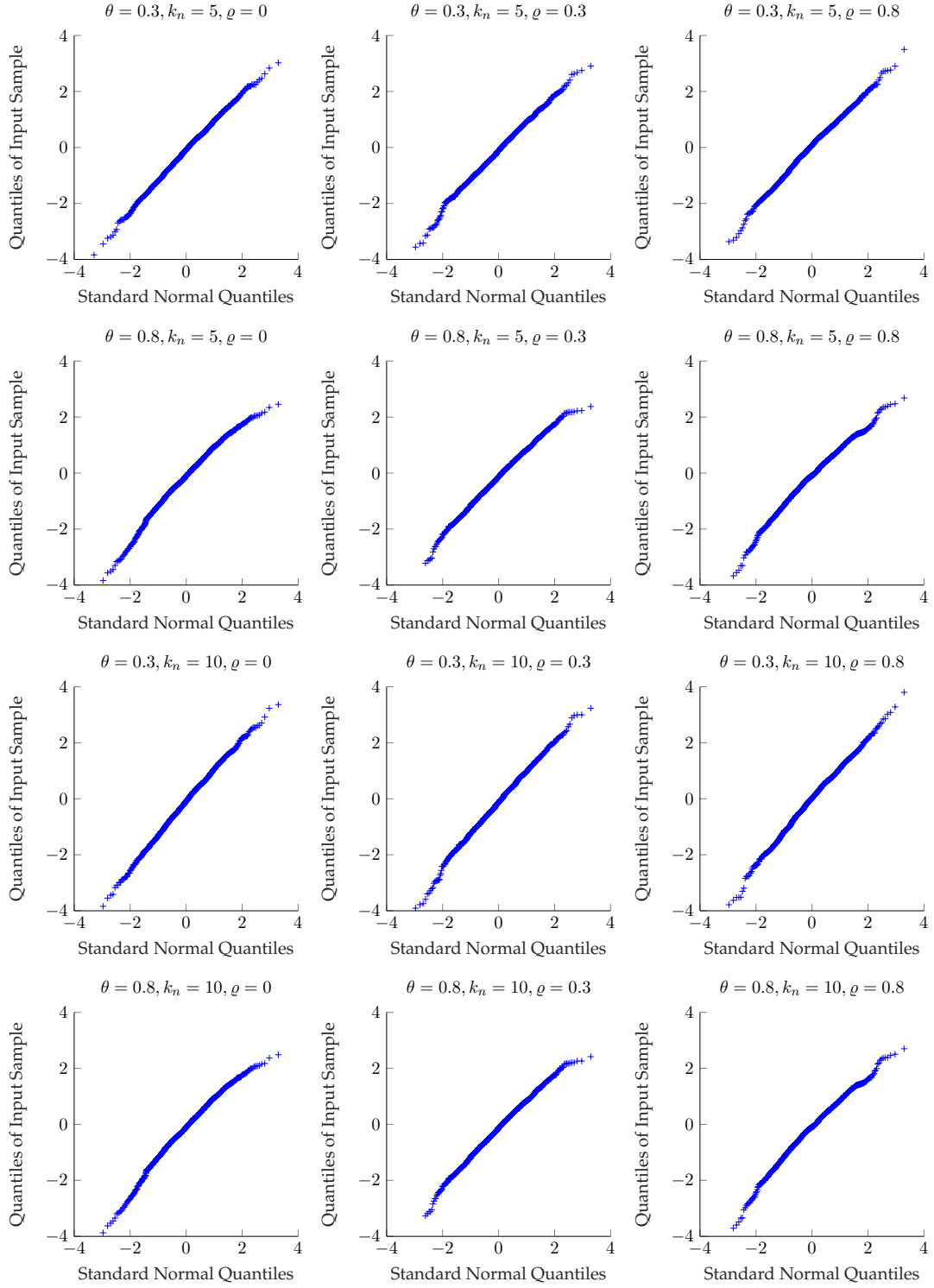


Figure A.1: QQ Plot of $(\widehat{C}_t^n - C_t)/\sqrt{\widehat{\Sigma}_t^n}$ versus Standard Normal. The scale of noise is fixed at $K_\gamma = 5 \times 10^{-5}$, the AR(1) coefficient of noise $\rho \in \{0, 0.3, 0.8\}$. The tuning parameters are selected as follows: $\theta \in \{0.3, 0.8\}, k_n \in \{5, 10\}, \ell_n = \lfloor k_n/2 \rfloor$. The number of simulations is 1000.

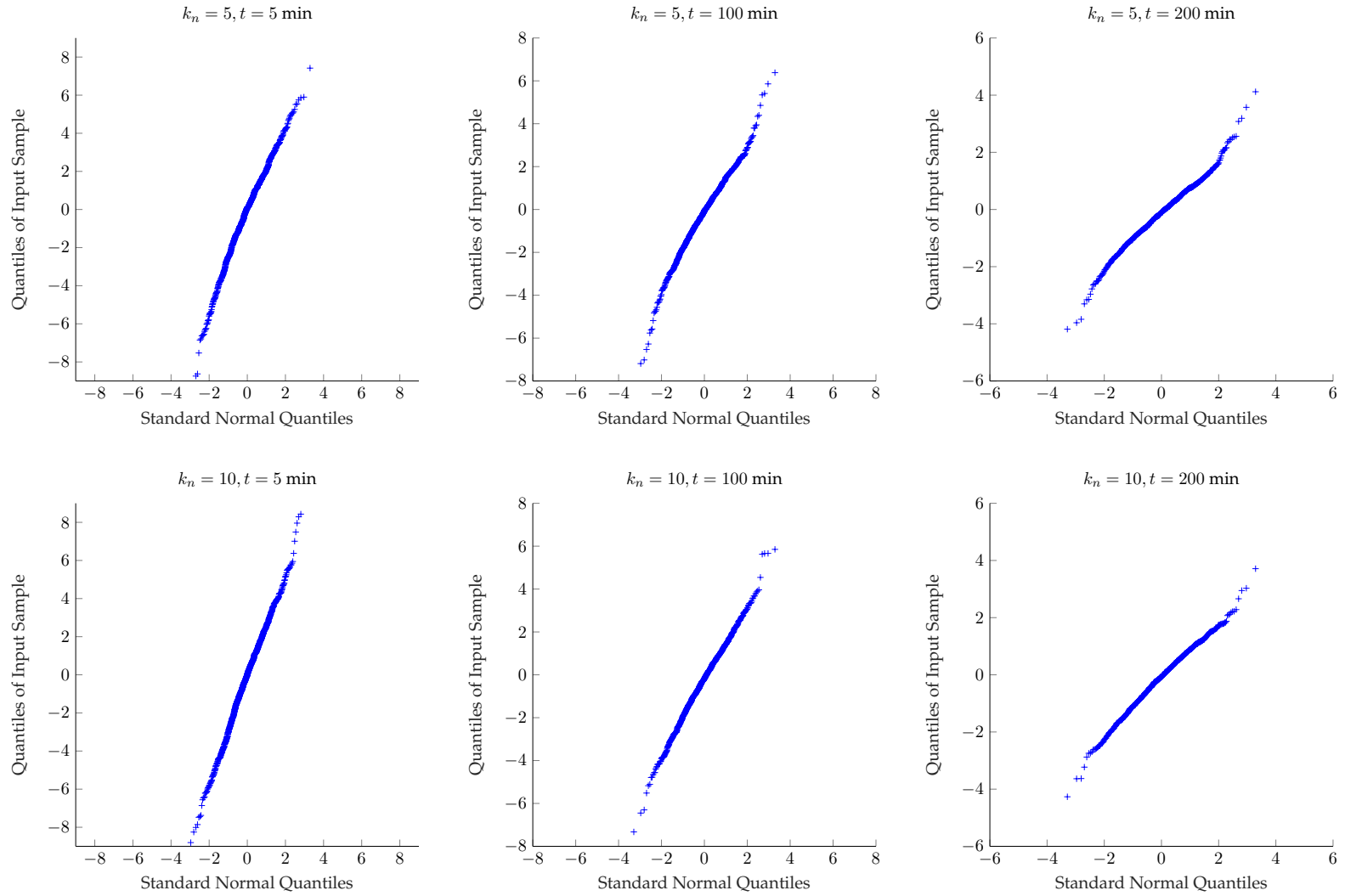


Figure A.2: QQ Plot of $(\tilde{c}_t^n - c_t)/\sqrt{\tilde{\Sigma}_t^n}$ versus Standard Normal for $t \in \{5, 100, 200\}$. The scale of noise is fixed at $K_\gamma = 5 \times 10^{-5}$, the AR(1) coefficient of noise $\rho = 0.5$. The tuning parameters are selected as follows: $\theta = 0.5, k_n \in \{5, 10\}, \ell_n = \lfloor k_n/2 \rfloor, l_n = 600$. The number of simulations is 1000.

Specifications		$\gamma = 5 \times 10^{-4}$			$\gamma = 5 \times 10^{-5}$		
		$\rho = 0$	0.3	0.8	$\rho = 0$	0.3	0.8
PaReMeDI	$k_n = 5, \ell_n = 5, \theta = 0.3, w_n = 10\mu_2$	0.0193	0.0184	0.0226	0.0225	0.0261	0.5468
	$k_n = 5, \ell_n = 5, \theta = 0.3, w_n = 20\mu_2$	0.0642	0.0641	0.0692	0.0711	0.0748	0.5991
	$k_n = 5, \ell_n = 5, \theta = 0.8, w_n = 10\mu_2$	-0.0252	-0.0249	-0.0272	-0.0301	-0.0272	0.0717
	$k_n = 5, \ell_n = 5, \theta = 0.8, w_n = 20\mu_2$	0.0572	0.0585	0.0584	0.0594	0.0614	0.1615
	$k_n = 5, \ell_n = 5, \theta = 1.5, w_n = 10\mu_2$	-0.1122	-0.1119	-0.1174	-0.1243	-0.1198	-0.0801
	$k_n = 5, \ell_n = 5, \theta = 1.5, w_n = 20\mu_2$	0.0429	0.0464	0.0434	0.0440	0.0468	0.0808
	$k_n = 5, \ell_n = 10, \theta = 0.3, w_n = 10\mu_2$	0.0192	0.0183	0.0221	0.0227	0.0262	0.4992
	$k_n = 5, \ell_n = 10, \theta = 0.3, w_n = 20\mu_2$	0.0642	0.0641	0.0687	0.0713	0.0749	0.5515
	$k_n = 5, \ell_n = 10, \theta = 0.8, w_n = 10\mu_2$	-0.0252	-0.0249	-0.0273	-0.0301	-0.0272	0.0602
	$k_n = 5, \ell_n = 10, \theta = 0.8, w_n = 20\mu_2$	0.0572	0.0585	0.0583	0.0594	0.0614	0.1500
	$k_n = 5, \ell_n = 10, \theta = 1.5, w_n = 10\mu_2$	-0.1122	-0.1119	-0.1174	-0.1243	-0.1198	-0.0838
	$k_n = 5, \ell_n = 10, \theta = 1.5, w_n = 20\mu_2$	0.0429	0.0464	0.0434	0.0440	0.0468	0.0772
	$k_n = 10, \ell_n = 5, \theta = 0.3, w_n = 10\mu_2$	0.0191	0.0182	0.0203	0.0233	0.0259	0.3121
	$k_n = 10, \ell_n = 5, \theta = 0.3, w_n = 20\mu_2$	0.0641	0.0640	0.0669	0.0718	0.0745	0.3644
	$k_n = 10, \ell_n = 5, \theta = 0.8, w_n = 10\mu_2$	-0.0252	-0.0249	-0.0276	-0.0300	-0.0273	0.0346
	$k_n = 10, \ell_n = 5, \theta = 0.8, w_n = 20\mu_2$	0.0572	0.0585	0.0581	0.0595	0.0614	0.1243
	$k_n = 10, \ell_n = 5, \theta = 1.5, w_n = 10\mu_2$	-0.1122	-0.1119	-0.1175	-0.1243	-0.1198	-0.0910
	$k_n = 10, \ell_n = 5, \theta = 1.5, w_n = 20\mu_2$	0.0429	0.0464	0.0433	0.0441	0.0468	0.0699
	$k_n = 10, \ell_n = 10, \theta = 0.3, w_n = 10\mu_2$	0.0190	0.0181	0.0193	0.0233	0.0260	0.2283
	$k_n = 10, \ell_n = 10, \theta = 0.3, w_n = 20\mu_2$	0.0640	0.0639	0.0659	0.0718	0.0747	0.2806
$k_n = 10, \ell_n = 10, \theta = 0.8, w_n = 10\mu_2$	-0.0253	-0.0249	-0.0278	-0.0300	-0.0272	0.0143	
$k_n = 10, \ell_n = 10, \theta = 0.8, w_n = 20\mu_2$	0.0572	0.0585	0.0578	0.0595	0.0614	0.1041	
$k_n = 10, \ell_n = 10, \theta = 1.5, w_n = 10\mu_2$	-0.1122	-0.1120	-0.1176	-0.1243	-0.1198	-0.0974	
$k_n = 10, \ell_n = 10, \theta = 1.5, w_n = 20\mu_2$	0.0429	0.0464	0.0433	0.0441	0.0468	0.0635	
MBRK	$b_n = 25, f_n = 5, L_n = 300$	-0.0490	-0.0522	-0.0483	0.2441	0.1851	0.6816
	$b_n = 25, f_n = 5, L_n = 500$	-0.0476	-0.0497	-0.0447	0.1315	0.0948	0.6261
	$b_n = 25, f_n = 5, L_n = 1000$	-0.0667	-0.0667	-0.0621	0.0244	0.0037	0.5403
	$b_n = 25, f_n = 10, L_n = 300$	-0.0568	-0.0594	-0.0594	0.2397	0.1800	0.2973
	$b_n = 25, f_n = 10, L_n = 500$	-0.0520	-0.0542	-0.0530	0.1314	0.0914	0.2406
	$b_n = 25, f_n = 10, L_n = 1000$	-0.0687	-0.0684	-0.0680	0.0239	0.0021	0.1699
	$b_n = 25, f_n = 15, L_n = 300$	-0.0647	-0.0666	-0.0687	0.2337	0.1734	0.1807
	$b_n = 25, f_n = 15, L_n = 500$	-0.0565	-0.0587	-0.0588	0.1292	0.0888	0.1190
	$b_n = 25, f_n = 15, L_n = 1000$	-0.0706	-0.0709	-0.0712	0.0230	0.0017	0.0507
	$b_n = 50, f_n = 5, L_n = 300$	-0.0643	-0.0665	-0.0667	0.2295	0.1670	0.3026
	$b_n = 50, f_n = 5, L_n = 500$	-0.0563	-0.0585	-0.0563	0.1253	0.0847	0.2521
	$b_n = 50, f_n = 5, L_n = 1000$	-0.0707	-0.0711	-0.0694	0.0216	-0.0008	0.1837
	$b_n = 50, f_n = 10, L_n = 300$	-0.0727	-0.0738	-0.0760	0.2230	0.1604	0.1593
	$b_n = 50, f_n = 10, L_n = 500$	-0.0609	-0.0626	-0.0621	0.1227	0.0812	0.1101
	$b_n = 50, f_n = 10, L_n = 1000$	-0.0732	-0.0735	-0.0726	0.0203	-0.0016	0.0476
	$b_n = 50, f_n = 15, L_n = 300$	-0.0809	-0.0810	-0.0840	0.2161	0.1548	0.1108
	$b_n = 50, f_n = 15, L_n = 500$	-0.0655	-0.0663	-0.0668	0.1191	0.0777	0.0627
	$b_n = 50, f_n = 15, L_n = 1000$	-0.0756	-0.0759	-0.0747	0.0184	-0.0022	0.0038
	$b_n = 75, f_n = 5, L_n = 300$	-0.0792	-0.0798	-0.0818	0.2163	0.1545	0.1851
	$b_n = 75, f_n = 5, L_n = 500$	-0.0647	-0.0655	-0.0654	0.1184	0.0771	0.1397
	$b_n = 75, f_n = 5, L_n = 1000$	-0.0751	-0.0753	-0.0737	0.0180	-0.0028	0.0785
	$b_n = 75, f_n = 10, L_n = 300$	-0.0872	-0.0873	-0.0903	0.2096	0.1478	0.1085
	$b_n = 75, f_n = 10, L_n = 500$	-0.0691	-0.0694	-0.0709	0.1148	0.0731	0.0656
	$b_n = 75, f_n = 10, L_n = 1000$	-0.0775	-0.0775	-0.0763	0.0158	-0.0040	0.0097
$b_n = 75, f_n = 15, L_n = 300$	-0.0951	-0.0942	-0.0982	0.2027	0.1416	0.0802	
$b_n = 75, f_n = 15, L_n = 500$	-0.0735	-0.0732	-0.0757	0.1107	0.0693	0.0390	
$b_n = 75, f_n = 15, L_n = 1000$	-0.0799	-0.0796	-0.0784	0.0136	-0.0050	-0.0139	

Table A.1: The relative biases of the jump-robust PaReMeDI and MBRK estimators of IV. The noise scale parameter K_γ is selected in $\{5 \times 10^{-4}, 5 \times 10^{-5}\}$. The AR(1) coefficient of the stationary noise is given by $\rho \in \{0, 0.3, 0.8\}$. The tuning parameters of PaReMeDI are set as follows: $k_n \in \{5, 10\}$, $\ell_n \in \{5, 10\}$, $\theta \in \{0.3, 0.8, 1.5\}$, $w_n \in \{10\mu_2, 20\mu_2\}$. For MBRK, the flatness and bandwidth tuning parameters are selected from $(f_n, b_n) \in \{5, 10, 15\} \times \{25, 50, 75\}$. The jittering bandwidth is fixed at $j_n = 5$. The block sizes $L_n \in \{300, 500, 1000\}$. The number of simulation is 1000.

Specifications		$\gamma = 5 \times 10^{-4}$			$\gamma = 5 \times 10^{-5}$		
		$\rho = 0$	0.3	0.8	$\rho = 0$	0.3	0.8
ReMeDI	$k_n = 5, \ell_n = 5, \theta = 0.3, w_n = 10\mu_2$	0.1199	0.1094	0.1176	0.1186	0.1288	0.6528
	$k_n = 5, \ell_n = 5, \theta = 0.3, w_n = 20\mu_2$	0.1861	0.1764	0.1799	0.1898	0.2025	0.7516
	$k_n = 5, \ell_n = 5, \theta = 0.8, w_n = 10\mu_2$	0.3416	0.3379	0.3493	0.3584	0.3635	0.3463
	$k_n = 5, \ell_n = 5, \theta = 0.8, w_n = 20\mu_2$	0.2245	0.2117	0.2133	0.2077	0.2220	0.2873
	$k_n = 5, \ell_n = 5, \theta = 1.5, w_n = 10\mu_2$	0.6640	0.6698	0.6827	0.6816	0.6930	0.6691
	$k_n = 5, \ell_n = 5, \theta = 1.5, w_n = 20\mu_2$	0.2524	0.2503	0.2504	0.2382	0.2407	0.2614
	$k_n = 5, \ell_n = 10, \theta = 0.3, w_n = 10\mu_2$	0.1212	0.1109	0.1189	0.1197	0.1302	0.5985
	$k_n = 5, \ell_n = 10, \theta = 0.3, w_n = 20\mu_2$	0.1872	0.1774	0.1808	0.1906	0.2034	0.6976
	$k_n = 5, \ell_n = 10, \theta = 0.8, w_n = 10\mu_2$	0.3416	0.3380	0.3494	0.3584	0.3637	0.3466
	$k_n = 5, \ell_n = 10, \theta = 0.8, w_n = 20\mu_2$	0.2248	0.2120	0.2136	0.2079	0.2223	0.2777
	$k_n = 5, \ell_n = 10, \theta = 1.5, w_n = 10\mu_2$	0.6640	0.6698	0.6827	0.6816	0.6931	0.6710
	$k_n = 5, \ell_n = 10, \theta = 1.5, w_n = 20\mu_2$	0.2525	0.2504	0.2505	0.2383	0.2408	0.2598
	$k_n = 10, \ell_n = 5, \theta = 0.3, w_n = 10\mu_2$	0.1275	0.1179	0.1246	0.1265	0.1382	0.3914
	$k_n = 10, \ell_n = 5, \theta = 0.3, w_n = 20\mu_2$	0.1918	0.1826	0.1854	0.1964	0.2078	0.4911
	$k_n = 10, \ell_n = 5, \theta = 0.8, w_n = 10\mu_2$	0.3419	0.3382	0.3496	0.3584	0.3642	0.3491
	$k_n = 10, \ell_n = 5, \theta = 0.8, w_n = 20\mu_2$	0.2256	0.2129	0.2144	0.2088	0.2230	0.2579
	$k_n = 10, \ell_n = 5, \theta = 1.5, w_n = 10\mu_2$	0.6640	0.6698	0.6827	0.6816	0.6932	0.6747
	$k_n = 10, \ell_n = 5, \theta = 1.5, w_n = 20\mu_2$	0.2527	0.2505	0.2506	0.2384	0.2410	0.2570
	$k_n = 10, \ell_n = 10, \theta = 0.3, w_n = 10\mu_2$	0.1315	0.1225	0.1283	0.1306	0.1439	0.3052
	$k_n = 10, \ell_n = 10, \theta = 0.3, w_n = 20\mu_2$	0.1950	0.1864	0.1888	0.2004	0.2113	0.4040
$k_n = 10, \ell_n = 10, \theta = 0.8, w_n = 10\mu_2$	0.3422	0.3384	0.3498	0.3584	0.3648	0.3531	
$k_n = 10, \ell_n = 10, \theta = 0.8, w_n = 20\mu_2$	0.2266	0.2141	0.2154	0.2100	0.2240	0.2442	
$k_n = 10, \ell_n = 10, \theta = 1.5, w_n = 10\mu_2$	0.6640	0.6698	0.6827	0.6815	0.6933	0.6780	
$k_n = 10, \ell_n = 10, \theta = 1.5, w_n = 20\mu_2$	0.2530	0.2508	0.2509	0.2386	0.2412	0.2549	
MBRK	$b_n = 25, f_n = 5, L_n = 300$	0.1412	0.1442	0.1387	0.3124	0.2570	0.8066
	$b_n = 25, f_n = 5, L_n = 500$	0.1439	0.1400	0.1389	0.1984	0.1784	0.7496
	$b_n = 25, f_n = 5, L_n = 1000$	0.1686	0.1654	0.1552	0.1418	0.1410	0.6455
	$b_n = 25, f_n = 10, L_n = 300$	0.1608	0.1673	0.1634	0.3157	0.2594	0.3768
	$b_n = 25, f_n = 10, L_n = 500$	0.1640	0.1582	0.1621	0.2088	0.1915	0.3259
	$b_n = 25, f_n = 10, L_n = 1000$	0.1848	0.1790	0.1743	0.1559	0.1583	0.2581
	$b_n = 25, f_n = 15, L_n = 300$	0.1758	0.1865	0.1858	0.3177	0.2636	0.2693
	$b_n = 25, f_n = 15, L_n = 500$	0.1812	0.1739	0.1827	0.2178	0.2012	0.2235
	$b_n = 25, f_n = 15, L_n = 1000$	0.2001	0.1927	0.1895	0.1675	0.1770	0.1947
	$b_n = 50, f_n = 5, L_n = 300$	0.1692	0.1783	0.1756	0.3070	0.2488	0.3759
	$b_n = 50, f_n = 5, L_n = 500$	0.1729	0.1660	0.1697	0.2058	0.1877	0.3295
	$b_n = 50, f_n = 5, L_n = 1000$	0.1930	0.1864	0.1788	0.1558	0.1619	0.2632
	$b_n = 50, f_n = 10, L_n = 300$	0.1866	0.1967	0.1958	0.3071	0.2536	0.2416
	$b_n = 50, f_n = 10, L_n = 500$	0.1903	0.1807	0.1881	0.2147	0.1984	0.2095
	$b_n = 50, f_n = 10, L_n = 1000$	0.2093	0.1999	0.1935	0.1682	0.1796	0.1842
	$b_n = 50, f_n = 15, L_n = 300$	0.2050	0.2138	0.2125	0.3058	0.2589	0.2138
	$b_n = 50, f_n = 15, L_n = 500$	0.2056	0.1950	0.2020	0.2230	0.2080	0.1958
	$b_n = 50, f_n = 15, L_n = 1000$	0.2235	0.2125	0.2056	0.1813	0.1936	0.1913
	$b_n = 75, f_n = 5, L_n = 300$	0.1982	0.2048	0.2034	0.2997	0.2513	0.2612
	$b_n = 75, f_n = 5, L_n = 500$	0.1973	0.1874	0.1920	0.2159	0.2006	0.2314
	$b_n = 75, f_n = 5, L_n = 1000$	0.2145	0.2054	0.1986	0.1740	0.1838	0.1959
	$b_n = 75, f_n = 10, L_n = 300$	0.2172	0.2221	0.2209	0.2983	0.2556	0.2126
	$b_n = 75, f_n = 10, L_n = 500$	0.2124	0.2018	0.2063	0.2242	0.2105	0.1997
	$b_n = 75, f_n = 10, L_n = 1000$	0.2280	0.2181	0.2116	0.1884	0.1981	0.1922
	$b_n = 75, f_n = 15, L_n = 300$	0.2365	0.2380	0.2363	0.2963	0.2601	0.2099
$b_n = 75, f_n = 15, L_n = 500$	0.2272	0.2159	0.2186	0.2319	0.2197	0.2067	
$b_n = 75, f_n = 15, L_n = 1000$	0.2407	0.2308	0.2241	0.2029	0.2113	0.2080	

Table A.2: The Root-mean-squared-relative-errors (RMSRE) of the jump-robust PaReMeDI and MBRK estimators of IV. The noise scale parameter K_γ is selected in $\{5 \times 10^{-4}, 5 \times 10^{-5}\}$. The AR(1) coefficient of the stationary noise is given by $\rho \in \{0, 0.3, 0.8\}$. The tuning parameters of PaReMeDI are set as follows: $k_n \in \{5, 10\}, \ell_n \in \{5, 10\}, \theta \in \{0.3, 0.8, 1.5\}, w_n \in \{10\mu_2, 20\mu_2\}$. For MBRK, the flatness and bandwidth tuning parameters are selected from $(f_n, b_n) \in \{5, 10, 15\} \times \{25, 50, 75\}$. The jittering bandwidth is fixed at $j_n = 5$. The block sizes $L_n \in \{300, 500, 1000\}$. The number of simulation is 1000.

B Technical Proofs

We will follow the scheme of the proofs in [Jacod et al. \(2019\)](#), and we add a prefix ‘‘JLZ’’ to cite their results. In the sequel, K will be a generic constant independent of n , which may change from line to line or even within one line. We write it K_{par} if it depends on some parameter par . The variable Ψ_{par}^n is a nonnegative \mathcal{G} -measurable variable with $\mathbb{E}((\Psi_{par}^n)^2) \leq 1$.

Using a classical localization procedure, it suffices to have the following stronger version of Assumptions [\(H-X-r\)](#), [\(O- \$\rho, \rho'\$ \)](#) and [\(N-v\)](#) in [Li and Linton \(2022a\)](#):

Assumption (S-HON). Assume Assumptions [\(H-X-r\)](#), [\(O- \$\rho, \rho'\$ \)](#) (with $\tau_1 = \infty$) and [\(N-v\)](#) hold. Assumption [\(H-r\)](#) hold for $b, \sigma, \alpha, \gamma$ with $r = 2$; the function δ and the processes $b, \sigma, \alpha, 1/\alpha, \gamma, X$ are bounded and

$$N_t^n \leq Kt\Delta_n^{-1}, \left| \mathbb{E} \left(\Delta(n, i) - \Delta_n \alpha_{T(n, i-1)}^{-1} \mid \mathcal{F}_{T(n, i-1)} \right) \right| \leq K\Delta_n^{1+\rho}, \mathbb{E} \left(\Delta(n, i)^\kappa \mid \mathcal{F}_{T(n, i-1)} \right) \leq K\Delta_n^\kappa. \quad (\text{B.1})$$

We decompose X into the continuous and discontinuous parts: $X = X^c + X^d$ with

$$X_t^c := X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s, \quad X_t^d := \int_{[0, t] \times E} \delta(s, z) \mu(s, z). \quad (\text{B.2})$$

In the sequel, $i, j, j', p > 1$ are integers. We introduce the following notations

$$\begin{aligned} G_s^{m, i} &:= \sum_{j=1}^{h_n-1} g_j^n \mathbf{1}_{\{(T(n, i+j-1), T(n, i+j))\}}(s), \quad \widehat{G}_{j, j'}^{n, i} := \int_0^\infty G_u^{m, i+j} G_u^{m, i+j'} du, \\ m(j)_p^n &:= (j-1)(p+2)h_n, \quad \eta_n := \Delta_n^{-\frac{1}{4} - \frac{\eta}{2}}, \quad \mathbf{0}_m = \underbrace{(0, \dots, 0)}_m, \\ \overline{G}_{j, j'}^{n, i} &:= \int_0^\infty G_u^{m, i+j} G_u^{m, i+j'} du \int_0^u G_s^{m, i+j} G_s^{m, i+j'} ds, \quad J_n(p, t) := 1 + \left[\frac{N_t^n}{(p+2)h_n} \right], \\ I_n(p, t) &:= J_n(p, t)(p+2)h_n - 1; \quad \mathcal{H}_i^n := \mathcal{F}_\infty \otimes \mathcal{G}_{i-h_n}, \quad \mathcal{K}_i^n := \mathcal{F}_i^n \otimes \mathcal{G}_{i-h_n}, \\ \mathcal{H}(p)_j^n &:= \mathcal{K}_{m(j+1)_p^n}^n, \quad \mathbf{j} := (j_1, \dots, j_d), \quad j_\ell \in \mathbb{N}, 1 \leq \ell \leq d, \quad \overline{\mathbf{r}}(\mathbf{j})_n := \mathbb{E} \left(\prod_{\ell=1}^d \overline{X}_\ell^n \right), \\ \mathcal{H}'(p)_j^n &:= \mathcal{K}_{m(j+1)_p^n + ph_n}^n, \quad \widehat{\mathcal{C}}_i^n := \sum_{j \in \mathbb{Z}} (g_j^n)^2 \Delta_{i+j}^n C, \quad \widehat{X}_i^{c, n} := (\overline{X}_i^{c, n})^2 - \widehat{\mathcal{C}}_i^n, \\ \widehat{\varepsilon}_i^n &:= (\overline{\varepsilon}_i^n)^2 - (\gamma_i^n \Delta_n^\eta)^2 \overline{\mathbf{r}}(\mathbf{0}_2)_n, \quad \widehat{X}_i^{c, n} \varepsilon_i^n := \overline{X}_i^{c, n} \varepsilon_i^n, \\ Z_i^n &:= (\overline{Y}_i^{c, n})^2 - \widehat{\mathcal{C}}_i^n - (\gamma_i^n \Delta_n^\eta)^2 \overline{\mathbf{r}}(\mathbf{0}_2)_n = \widehat{X}_i^{c, n} + \widehat{\varepsilon}_i^n + 2\widehat{X}_i^{c, n} \varepsilon_i^n, \quad \zeta(p)_i^n := \sum_{j=i}^{i+ph_n-1} Z_j^n, \\ \eta(p)_j^n &:= \frac{1}{h_n \phi_0^n} \zeta(p)_{m(j)_p^n}^n, \quad \overline{\eta}(p)_j^n := \mathbb{E} \left(\eta(p)_j^n \mid \mathcal{H}(p)_{j-1}^n \right), \\ \eta'(p)_j^n &:= \frac{1}{h_n \phi_0^n} \zeta(2)_{m(j)_p^n + ph_n}^n, \quad \overline{\eta}'(p)_j^n := \mathbb{E} \left(\eta'(p)_j^n \mid \mathcal{H}'(p)_{j-1}^n \right). \end{aligned}$$

For any process V , we defined $\Delta_{i, \ell}^{n, k} := (V_{i+\ell}^n - V_{i+\ell+k}^n)(V_i^n - V_{i-k}^n)$. Next, we define several

processes:

$$\begin{aligned}
F(p)_t^n &:= \sum_{j=1}^{J_n(p,t)} \bar{\eta}(p)_j^n, & M(p)_t^n &:= \sum_{j=1}^{J_n(p,t)} (\eta(p)_j^n - \bar{\eta}(p)_j^n), \\
F'(p)_t^n &:= \sum_{j=1}^{J_n(p,t)} \bar{\eta}'(p)_j^n, & M'(p)_t^n &:= \sum_{j=1}^{J_n(p,t)} (\eta'(p)_j^n - \bar{\eta}'(p)_j^n), \\
\widehat{C}(p)_t^n &:= \frac{1}{h_n \phi_0^n} \sum_{i=N_t^n - h_n + 2}^{I_n(p,t)} Z_i^n, & \widehat{C}_t^{n,1} &:= \frac{1}{h_n \phi_0^n} \sum_{i=0}^{N_t^n - h_n + 1} \widehat{c}_i^n - C_t, \\
\widehat{C}_t^{n,2} &:= \frac{1}{h_n \phi_0^n} \sum_{i=0}^{N_t^n - h_n + 1} \left((\bar{Y}_i^n)^2 \mathbf{1}_{\{|\bar{Y}_i^n| \leq u_n\}} - (\bar{Y}_i^{c,n})^2 \right), \\
\widehat{C}_t^{n,3} &:= \frac{\bar{\mathbf{r}}(\mathbf{0}_2)_n}{h_n \phi_0^n} \sum_{i=0}^{N_t^n - h_n + 1} (\gamma_i^n \Delta_n^\eta)^2 - \frac{1}{h_n^2 \phi_0^n} \sum_{|\ell| \leq \ell_n} \sum_{i=k_n}^{N_t^n - h_n + 1} \widetilde{\phi}_\ell^n \Delta_{i,\ell}^{n,k_n} Y.
\end{aligned}$$

For all $p > 1$, we have

$$\widehat{C}_t^n - C_t = M(p)_t^n + M'(p)_t^n + F(p)_t^n + F'(p)_t^n - \widehat{C}(p)_t^n + \sum_{j=1}^3 \widehat{C}_t^{n,j}.$$

Lemma B.1. *Under Assumption (N- v) with $v > 3$, we have the following convergence in law*

$$\sqrt{h_n} \bar{\chi}_i^n \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \widetilde{\phi}(0)R\right). \tag{B.3}$$

Moreover, we have

$$h_n \bar{\mathbf{r}}(\mathbf{0}_2)_n \rightarrow \widetilde{\phi}(0)R; \quad h_n^{3/2} \bar{\mathbf{r}}(\mathbf{0}_3)_n \rightarrow 0; \quad h_n^2 \bar{\mathbf{r}}(\mathbf{0}_4)_n \rightarrow 3\widetilde{\phi}^2(0)R^2. \tag{B.4}$$

Proof. First, we need to check the following condition according to [Rio \(1997\)](#)

$$\sum_{k \in \mathbb{N}^*} k^{\frac{2}{r-2}} \rho_k < \infty, \tag{B.5}$$

where $\{\rho_k\}$ are the ρ -mixing coefficients introduced in Definition 2.1 of [Li and Linton \(2022a\)](#), and r is a positive real such that $\mathbb{E}(|\chi_i|^r) < \infty$. Assumption (N- v) implies r can be arbitrarily large, thus (B.5) holds since $v > 3$. The rest of the proof of (B.3) follows from the proof of Lemma A.1 in [Li et al. \(2020\)](#). (B.4) follows from the condition $\mathbb{E}(|\chi_i|^w) < \infty, w > 4$ and the limit distribution (B.3); that is, convergence in distribution implies convergence in moments when some higher order moments of χ are bounded, see, e.g., Theorem 4.5.2 in [Chung \(2001\)](#). ■

Lemma B.2. *Given a fixed integer i , a real $z > 1$ and two sequences of integers $\{d_n\}_n, \{\ell_n\}_n$ and two*

integers j, ℓ satisfying $|j - i| \leq d_n, |\ell| \leq \ell_n$, we have for any $q \geq 2$,

$$\mathbb{E}\left(\left|\Delta_{i,\ell}^{n,k_n} Y - (\Delta_n^\eta \gamma_j^n)^2 \Delta_{i,\ell}^{n,k_n} \chi\right|^q\right) \leq K \Delta_n^{q\eta} \left((d_n + \ell_n + k_n) \Delta_n + ((d_n + \ell_n + k_n) \Delta_n)^{\frac{1}{2}} \right). \quad (\text{B.6})$$

Proof. We rewrite $\Delta_{i,\ell}^{n,k_n} Y = (\eta_i^{n,1} + \tilde{\eta}_{i,j}^{n,1})(\eta_i^{n,2} + \tilde{\eta}_{i,j}^{n,2})$, where

$$\begin{cases} \eta_i^{n,1} := X_{i+\ell}^n - X_{i+k_n+\ell}^n + \Delta_n^\eta (\gamma_{i+\ell}^n - \gamma_j^n) \chi_{i+\ell} - \Delta_n^\eta (\gamma_{i+k_n+\ell}^n - \gamma_j^n) \chi_{i+k_n+\ell}; \\ \eta_i^{n,2} := X_i^n - X_{i-k_n}^n + \Delta_n^\eta (\gamma_i^n - \gamma_j^n) \chi_i - \Delta_n^\eta (\gamma_{i-k_n}^n - \gamma_j^n) \chi_{i-k_n}; \\ \tilde{\eta}_{i,j}^{n,1} := \Delta_n^\eta \gamma_j^n (\chi_{i+\ell} - \chi_{i+k_n+\ell}), \quad \tilde{\eta}_{i,j}^{n,2} := \Delta_n^\eta \gamma_j^n (\chi_i - \chi_{i-k_n}). \end{cases}$$

For $s \in \{1, 2\}, w \geq 2$, we have $\mathbb{E}(|\eta_i^{n,s}|^w) \leq K \Delta_n (d_n + \ell_n + k_n)$ by (JLZ-A.6), the independence of \mathcal{F}_∞ and \mathcal{G} and the fact that χ has bounded moments of all orders. Next, we have the estimate that $\mathbb{E}(|\tilde{\eta}_{i,j}^{n,s}|^w) \leq K \Delta_n^{w\eta}$ by the boundedness of γ and that χ has bounded moments of all orders. Since $\Delta_{i,\ell}^{n,k_n} Y - (\Delta_n^\eta \gamma_j^n)^2 \Delta_{i,\ell}^{n,k_n} \chi = \eta_i^{n,1} \eta_i^{n,2} + \eta_i^{n,1} \tilde{\eta}_{i,j}^{n,2} + \tilde{\eta}_{i,j}^{n,1} \eta_i^{n,2}$, the result follows from Hölder's inequality. ■

Lemma B.3. Let $p \geq 1$ and $q \geq 2$, and $\mathbf{j} = (j_1, \dots, j_q)$ where $j_k \in \mathbb{N}^*$ and $\max_k j_k \leq p h_n$. U is any \mathcal{F}_∞ -measurable variable. We have the following estimates

$$\left| \mathbb{E} \left(\prod_{m=1}^q \bar{\varepsilon}_{i+j+j_m}^n - (\Delta_n^\eta \gamma_i^n)^q \bar{\mathbf{r}}(\mathbf{j})_n \mid \mathcal{K}_i^n \right) \right| \leq K_{p,q} \Delta_n^{\eta q} \left(\frac{\Psi_{i,\mathbf{j}}^n}{h_n^v} + \frac{\Delta_n}{h_n^{\lfloor \frac{q}{2} \rfloor - 1 \wedge (v-1)}} \right); \quad (\text{B.7})$$

$$\left| \mathbb{E} \left(U \left(\prod_{m=1}^q \bar{\varepsilon}_{i+j+j_m}^n - (\Delta_n^\eta \gamma_i^n)^q \bar{\mathbf{r}}(\mathbf{j})_n \right) \mid \mathcal{K}_i^n \right) \right| \leq K_{p,q} \Delta_n^{\eta q} \left(\frac{\Psi_{i,\mathbf{j}}^n}{h_n^v} + \frac{\sqrt{\Delta_n}}{h_n^{\lfloor \frac{q-1}{2} \rfloor + \frac{1}{2} \wedge (v-\frac{1}{2})}} \right) \sqrt{\mathbb{E}(U^2 \mid \mathcal{F}_i^n)}; \quad (\text{B.8})$$

$$\left| \mathbb{E} \left(U \bar{\varepsilon}_{i+j}^n \mid \mathcal{K}_i^n \right) \right| \leq \frac{K_p \Psi_{i,\mathbf{j}}^n \Delta_n^\eta}{h_n^v} \mathbb{E}(|U| \mid \mathcal{F}_i^n); \quad (\text{B.9})$$

$$\mathbb{E}((\bar{\varepsilon}_i^n)^q \mid \mathcal{K}_i^n) \leq K_q \frac{\Psi_i^n \Delta_n^{\eta q}}{h_n^{v \wedge \frac{q}{2}}} \text{ for even integer } q; \quad (\text{B.10})$$

$$\mathbb{E}((\bar{Y}_i^n)^q \mid \mathcal{K}_i^n) \leq K_q \left(\frac{\Psi_i^n \Delta_n^{\eta q}}{h_n^{v \wedge \frac{q}{2}}} + (h_n \Delta_n)^{\frac{q}{2}} \right) \text{ for even integer } q. \quad (\text{B.11})$$

Proof. We decompose $U \left(\prod_{m=1}^q \bar{\varepsilon}_{i+j+j_m}^n - (\Delta_n^\eta \gamma_i^n)^q \bar{\mathbf{r}}(\mathbf{j})_n \right) =: B_n + UC_n$, where

$$\begin{aligned} B_n &= (-1)^q U \sum_{l_1, \dots, l_q=0}^{h_n-1} \prod_{m=1}^q \tilde{g}_{l_m}^n \Delta_n^\eta \gamma_{i+l_m}^n \left(\prod_{m=1}^q \chi_{i+l_m} - \mathbf{r}(\mathbf{l}') \right); \\ C_n &= (-1)^q \sum_{l_1, \dots, l_q=0}^{h_n-1} \left(\prod_{m=1}^q \gamma_{i+l_m}^n - (\gamma_i^n)^q \right) \mathbf{r}(\mathbf{l}') \prod_{m=1}^q \tilde{g}_{l_m}^n \Delta_n^\eta \end{aligned}$$

with $\mathbf{l}' = (l'_1, \dots, l'_q)$, where $l'_m = j + j_m + l_m$. Next, we have by the mixing property that $\mathbb{E}(|\mathbb{E}(\prod_{m=1}^q \chi_{i+l'_m} - \mathbf{r}(\mathbf{l}') \mid \mathcal{G}_{i-h_n})|) \leq K/h_n^v$, thus

$$|\mathbb{E}(B_n \mid \mathcal{K}_i^n)| \leq K_p \Psi_{i,\mathbf{j}}^n \mathbb{E}(|U| \mid \mathcal{F}_i^n) \Delta_n^{\eta q} / h_n^v. \quad (\text{B.12})$$

Note that the estimate in (B.12) is also valid when $q = 1$, and this immediately leads to (B.9).

Let $\zeta(\mathbf{j})_i^n := \prod_{m=1}^q \gamma_{i+j_m}^n - (\gamma_i^n)^q$ for any $\mathbf{j} = (j_1, \dots, j_q)$. Next, we have

$$\begin{aligned} |\mathbb{E}(C_n | \mathcal{F}_i^n)| &\leq \frac{K_q (\Delta_n^\eta)^q}{h_n^q} \sum_{l_1, \dots, l_q=0}^{h_n-1} |\mathbf{r}(\mathbf{l}')| |\mathbb{E}(\zeta(\mathbf{l}')_i^n | \mathcal{F}_i^n)| \\ &\leq \frac{K_{p,q} \Delta_n^{1+\eta q}}{h_n^{q-1}} \mathcal{L}_{q,(2p+1)h_n} \leq \frac{K_{p,q} \Delta_n^{1+\eta q}}{h_n^{\lceil \frac{q-1}{2} \rceil \wedge (v-1)}}, \end{aligned} \quad (\text{B.13})$$

where the notation $\mathcal{L}_{q,k}$ is defined in the proof of Lemma JLZ-A.3. Now (B.7) follows from (B.12) and (B.13). On the other hand, we have

$$\begin{aligned} \mathbb{E} \left(\left(\sum_{l_1, \dots, l_q=0}^{h_n-1} \zeta(\mathbf{l}')_i^n \mathbf{r}(\mathbf{l}') \right)^2 | \mathcal{F}_i^n \right) &= \mathbb{E} \left(\sum_{l_1, \dots, l_q=0}^{h_n-1} \sum_{\ell_1, \dots, \ell_q=0}^{h_n-1} \zeta(\ell') \mathbf{r}(\ell') \zeta(\mathbf{l}')_i^n \mathbf{r}(\mathbf{l}') | \mathcal{F}_i^n \right) \\ &\leq K_q \Delta_n h_n \sum_{l_1, \dots, l_q=0}^{h_n-1} \sum_{\ell_1, \dots, \ell_q=0}^{h_n-1} |\mathbf{r}(\ell')| |\mathbf{r}(\mathbf{l}')| \\ &\leq K_q \Delta_n h_n \mathcal{L}_{q,(2p+1)h_n}^2, \end{aligned}$$

which yields

$$\mathbb{E}(C_n^2 | \mathcal{F}_i^n) \leq \frac{K_q \Delta_n^{1+2\eta q}}{h_n^{2q-1}} \mathcal{L}_{q,(2p+1)h_n}^2 \leq \frac{K_q \Delta_n^{1+2\eta q}}{h_n^{(2\lceil \frac{q-1}{2} \rceil + 1) \wedge (2v-1)}}. \quad (\text{B.14})$$

(B.8) readily follows from (B.14) and an application of Cauchy-Schwarz inequality.

Now suppose q is an even integer. By (JLZ-A.25) with $j = 0, j_m = 0 \forall m = 1, \dots, q$ and the fact that $|\bar{\mathbf{r}}(\mathbf{j})_n| \leq \frac{K_{p,q}}{h_n^{v \wedge \lceil (q+1)/2 \rceil}}$ (JLZ-A.23), we have

$$\mathbb{E}((\bar{\varepsilon}_i^n)^q | \mathcal{K}_i^n) \leq K_{p,q} \Delta_n^{\eta q} \left(\frac{\Psi_i^n}{h_n^v} + \frac{\Delta_n}{h_n^{(\frac{q}{2}-1) \wedge (v-1)}} + \frac{1}{h_n^{v \wedge \lceil (q+1)/2 \rceil}} \right) \leq K_q \frac{\Psi_i^n \Delta_n^{\eta q}}{h_n^{v \wedge \frac{q}{2}}}.$$

This proves (B.10). (B.11) follows immediately from (B.10) and (JLZ-A.17). ■

Lemma B.4. For $j, j' \in \{0, \dots, ph_n\}$ and an even integer $q \geq 2$, and $v > 1 + \frac{2}{1-2\eta}$ where $\eta \in [0, 1/6)$, we have

$$|\mathbb{E}(\widehat{\varepsilon}_i^n | \mathcal{K}_i^n)| \leq K \Psi_i^n \Delta_n^{1+2\eta}; \quad (\text{B.15})$$

$$\left| \mathbb{E}(\widehat{X}^c \varepsilon_i^n | \mathcal{K}_i^n) \right| \leq K \Psi_i^n \Delta_n^{\frac{1}{2} + \eta} h_n^{\frac{1}{2} - v}; \quad (\text{B.16})$$

$$|\mathbb{E}((\widehat{\varepsilon}_i^n)^q | \mathcal{K}_i^n)| \leq K_q \Psi_i^n \Delta_n^{2\eta q} h_n^{-(v \wedge q)}; \quad (\text{B.17})$$

$$\left| \mathbb{E}(\widehat{X}^c \varepsilon_i^n)^4 | \mathcal{K}_i^n \right| \leq K \Delta_n^{4\eta + 2}; \quad (\text{B.18})$$

$$|\mathbb{E}(\widehat{\varepsilon}_{i+j}^n \widehat{\varepsilon}_{i+j'}^n - (\Delta_n^\eta \gamma_i^n)^4 (\bar{\mathbf{r}}(\mathbf{j}')_n - \bar{\mathbf{r}}(\mathbf{0}_2)_n^2) | \mathcal{K}_i^n)| \leq K \Delta_n^{\frac{5}{4} + \frac{5\eta}{2}}, \text{ where } \mathbf{j}' = (j, j, j', j'); \quad (\text{B.19})$$

$$\left| \mathbb{E}(\widehat{X}^c \varepsilon_{i+j}^n \widehat{X}^c \varepsilon_{i+j'}^n | \mathcal{K}_i^n) - (\sigma_i^n \Delta_n^\eta \gamma_i^n)^2 \bar{\mathbf{r}}(j, j')_n \widehat{G}_{j, j'}^{n, i} \right| \leq K \Delta_n^{\frac{5}{4} + \frac{5\eta}{2}}. \quad (\text{B.20})$$

Proof. Recall that we have $\widehat{\varepsilon}_i^n = (\bar{\varepsilon}_i^n)^2 - (\Delta_n^\eta \gamma_i^n)^2 \bar{\mathbf{r}}(\mathbf{0}_2)_n$. (B.15) follows directly from (B.7) and

the fact $v > 2/(1 - 2\eta)$.

By first conditioning on \mathcal{H}_i^n , we have $|\mathbb{E}(\bar{\varepsilon}_i^n | \mathcal{H}_i^n)| \leq K\Psi_i^n \Delta_n^\eta / h_n^v$. On the other hand, we have by (JLZ-A.17) that $\mathbb{E}((\bar{X}_i^{c,n})^2 | \mathcal{K}_i^n) \leq Kh_n \Delta_n$. Now we have by first conditioning on \mathcal{H}_i^n , and then apply Cauchy-Schwarz inequality, we obtain (B.16).

Since $(\hat{\varepsilon}_i^n)^q = \sum_{m=0}^q C_m^q (-1)^m (\gamma_i^n \Delta_n^\eta)^{2m} \bar{\mathbf{r}}(\mathbf{0}_2)_n^m (\bar{\varepsilon}_i^n)^{2(q-m)}$, (B.10) implies

$$\left| \mathbb{E} \left((\bar{\varepsilon}_i^n)^{2(q-m)} | \mathcal{K}_i^n \right) \right| \leq K_q \Psi_i^n \Delta_n^{2\eta(q-m)} h_n^{-(v \wedge (q-m))}.$$

Since $|\bar{\mathbf{r}}(\mathbf{0}_2)_n| \leq Kh_n^{-1}$, we have

$$|\mathbb{E}((\hat{\varepsilon}_i^n)^q | \mathcal{K}_i^n)| \leq K_q \Psi_i^n \Delta_n^{2\eta q} h_n^{-m - (v \wedge (q-m))} \leq K_q \Psi_i^n \Delta_n^{2\eta q} h_n^{-(v \wedge q)}.$$

This proves (B.17).

Let

$$\mathfrak{X}(1)_i^n := (\bar{X}_i^{c,n})^4 ((\bar{\varepsilon}_i^n)^4 - \Delta_n^{4\eta} (\gamma_i^n)^4 \bar{\mathbf{r}}(\mathbf{0}_4)_n), \quad \mathfrak{X}(2)_i^n := (\bar{X}_i^{c,n})^4 \Delta_n^{4\eta} (\gamma_i^n)^4 \bar{\mathbf{r}}(\mathbf{0}_4)_n.$$

Then we have $(\widehat{X}_i^c \varepsilon_i^n)^4 = \mathfrak{X}(1)_i^n + \mathfrak{X}(2)_i^n$. Next, we can use (B.8) and a simple estimate that $|\mathbb{E}((\bar{X}_i^{c,n})^8 | \mathcal{F}_i^n)| \leq K(h_n \Delta_n)^4$ to get $|\mathbb{E}(\mathfrak{X}(1)_i^n | \mathcal{K}_i^n)| \leq K(\Delta_n h_n)^2 \Delta_n^{4\eta} \left(\frac{\Psi_i^n}{h_n^v} + \frac{\sqrt{\Delta_n}}{h_n^{3/2}} \right)$. Since $\bar{\mathbf{r}}(\mathbf{0}_4)_n \leq Kh_n^{-2}$ and $|\mathbb{E}((\bar{X}_i^{c,n})^4 | \mathcal{F}_i^n)| \leq K(h_n \Delta_n)^2$, we have $|\mathbb{E}(\mathfrak{X}(2)_i^n | \mathcal{K}_i^n)| \leq K\Delta_n^{4\eta+2}$. This finishes the proof of (B.18).

Next, we denote $\hat{\varepsilon}_{i+j}^n \hat{\varepsilon}_{i+j'}^n - (\Delta_n^\eta \gamma_i^n)^4 (\bar{\mathbf{r}}(\mathbf{j}')_n - \bar{\mathbf{r}}(\mathbf{0}_2)_n^2) =: \sum_{k=1}^4 \mathfrak{A}(k)_{i,j,j'}^n$, where

$$\begin{aligned} \mathfrak{A}(1)_{i,j,j'}^n &:= (\bar{\varepsilon}_{i+j}^n \bar{\varepsilon}_{i+j'}^n)^2 - (\Delta_n^\eta \gamma_i^n)^4 \bar{\mathbf{r}}(\mathbf{j}')_n; \\ \mathfrak{A}(2)_{i,j,j'}^n &:= -\bar{\mathbf{r}}(\mathbf{0}_2)_n (\Delta_n^\eta \gamma_{i+j'}^n)^2 ((\bar{\varepsilon}_{i+j}^n)^2 - (\Delta_n^\eta \gamma_i^n)^2 \bar{\mathbf{r}}(\mathbf{0}_2)_n); \\ \mathfrak{A}(3)_{i,j,j'}^n &:= -\bar{\mathbf{r}}(\mathbf{0}_2)_n (\Delta_n^\eta \gamma_{i+j}^n)^2 ((\bar{\varepsilon}_{i+j'}^n)^2 - (\Delta_n^\eta \gamma_i^n)^2 \bar{\mathbf{r}}(\mathbf{0}_2)_n); \\ \mathfrak{A}(4)_{i,j,j'}^n &:= \bar{\mathbf{r}}(\mathbf{0}_2)_n^2 (\Delta_n^\eta)^4 ((\gamma_{i+j}^n)^2 - (\gamma_i^n)^2)((\gamma_{i+j'}^n)^2 - (\gamma_i^n)^2). \end{aligned}$$

We have the following estimates

$$|\mathbb{E}(\mathfrak{A}(1)_{i,j,j'}^n | \mathcal{K}_i^n)| \leq K_{p,q} \Delta_n^{4\eta} \left(\frac{\Psi_i^n}{h_n^v} + \frac{\Delta_n}{h_n} \right); \quad (\text{B.21})$$

$$|\mathbb{E}(\mathfrak{A}(2)_{i,j,j'}^n | \mathcal{K}_i^n)| + |\mathbb{E}(\mathfrak{A}(3)_{i,j,j'}^n | \mathcal{K}_i^n)| \leq K_{p,q} \frac{\Delta_n^{4\eta}}{h_n} \left(\frac{\Psi_i^n}{h_n^v} + \sqrt{\frac{\Delta_n}{h_n}} \right); \quad (\text{B.22})$$

$$|\mathbb{E}(\mathfrak{A}(4)_{i,j,j'}^n | \mathcal{K}_i^n)| \leq K \Delta_n^{1+4\eta} / h_n. \quad (\text{B.23})$$

(B.21) and (B.22) follow from (B.7) and (B.8); (B.23) follows from Cauchy-Schwarz inequality.

Now (B.19) follows immediately.

To prove (B.20), we first introduce

$$\mathfrak{B}(1)_{i,j,j'}^n := \mathbb{E} \left(\int_0^\infty \bar{X}_u^{c,n,i+j} dB_u^{n,i+j'} + \int_0^\infty \bar{X}_u^{c,n,i+j'} dB_u^{n,i+j} | \mathcal{F}_i^n \right);$$

$$\begin{aligned}\mathfrak{B}(2)_{i,j,j'}^n &:= \mathbb{E} \left((\sigma_i^n)^2 \int_0^\infty (G_u^n)^2 du - \int_0^\infty (\sigma_u G_u^n)^2 du \mid \mathcal{F}_i^n \right), \text{ where } (G_u^n)^2 := G_u^{n,i+j} G_u^{n,i+j'}; \\ \mathfrak{C}(1)_{i,j,j'}^n &:= \overline{X}_{i+j}^{c,n} \overline{X}_{i+j'}^{c,n} (\overline{\varepsilon}_{i+j}^n \overline{\varepsilon}_{i+j'}^n - (\gamma_i^n \Delta_n^\eta)^2 \overline{\mathbf{r}}(j, j')_n); \\ \mathfrak{C}(2)_{i,j,j'}^n &:= \left(\overline{X}_{i+j}^{c,n} \overline{X}_{i+j'}^{c,n} - (\sigma_i^n)^2 \widehat{G}_{j,j}^{n,i} \right) (\gamma_i^n \Delta_n^\eta)^2 \overline{\mathbf{r}}(j, j')_n.\end{aligned}$$

Then, we have then by Itô's formula that

$$\mathbb{E} \left(\overline{X}_{i+j}^{c,n} \overline{X}_{i+j'}^{c,n} - (\sigma_i^n)^2 \widehat{G}_{j,j}^{n,i} \mid \mathcal{F}_i^n \right) = \mathfrak{B}(1)_{i,j,j'}^n + \mathfrak{B}(2)_{i,j,j'}^n.$$

By the first and last part of (JLZ-A.17), we have $|\mathfrak{B}(1)_{i,j,j'}^n| \leq K(h_n \Delta_n)^{3/2}$. According to the proof of Lemma (JLZ-A.6), we have $|\mathfrak{B}(2)_{i,j,j'}^n| \leq K(h_n \Delta_n)^{3/2}$. It leads to $|\mathbb{E}(\mathfrak{C}(2)_{i,j,j'}^n \mid \mathcal{F}_i^n)| \leq K \Delta_n^{\frac{5\eta}{2} + \frac{5}{4}}$. By (B.8), we have $|\mathbb{E}(\mathfrak{C}(1)_{i,j,j'}^n \mid \mathcal{F}_i^n)| \leq K \Delta_n^{\frac{5\eta}{2} + \frac{5}{4}}$. Now (B.20) is proved. ■

Lemma B.5. Assume $\eta \in [0, \frac{1}{6}]$, $v > 1 + \frac{2}{1-2\eta}$. We have for any $p \geq 2$ that

$$|\mathbb{E}(\zeta(p)_i^n \mid \mathcal{K}_i^n)| \leq K_p \Psi_{i,p}^n \Delta_n^{\frac{1}{2} + 2\eta}; \quad (\text{B.24})$$

$$|\mathbb{E}((\zeta(p)_i^n)^4 \mid \mathcal{K}_i^n)| \leq K_p \Psi_{i,p}^n \Delta_n^{12\eta - 2 - (\eta - \frac{1}{2})(v \wedge 4)}; \quad (\text{B.25})$$

$$|\mathbb{E}((\zeta(p)_i^n)^2 - 4(\sigma_i^n)^4 \rho(p, 1)_i^n - (\gamma_i^n \Delta_n^\eta)^4 \rho(p, 2)_i^n - 4(\sigma_i^n \Delta_n^\eta \gamma_i^n)^2 \rho(p, 3)_i^n \mid \mathcal{K}_i^n)| \leq K_p \Psi_{i,p}^n \Delta_n^{\frac{5\eta}{2} + \frac{1}{4}}, \quad (\text{B.26})$$

$$\text{where } \rho(p, 1) := \sum_{j,j'=0}^{ph_n-1} \overline{G}_{j,j'}^{n,i}, \rho(p, 2) := \sum_{j,j'=0}^{ph_n-1} (\overline{\mathbf{r}}(\mathbf{j}') - \overline{\mathbf{r}}(\mathbf{0}_2)_n)^2, \rho(p, 3)_i^n := \sum_{j,j'=0}^{ph_n-1} \overline{\mathbf{r}}(j, j')_n \widehat{G}_{j,j'}^{n,i}.$$

Proof. The first part of (JLZ-A.19), (B.15) and (B.16) imply $|\mathbb{E}(Z_i \mid \mathcal{K}_i^n)| \leq K \Delta_n^{1+\eta}$. (B.24) now follows immediately. (B.25) can be derived from the second part of (JLZ-A.19), (B.17) and (B.18), and the simple fact that $4\eta + 2 \geq 8\eta + (\frac{1}{2} - \eta)(v \wedge 4)$.

To see (B.26), we need the following estimates

$$\left| \sum_{j,j'=0}^{ph_n-1} \mathbb{E} \left(\widehat{X}_{i+j}^{c,n} \widehat{X}_{i+j'}^{c,n} \mid \mathcal{K}_i^n \right) - 4(\sigma_i^n)^4 \overline{G}_{j,j'}^{n,i} \right| \leq K_p \Delta_n^{\frac{1}{4} + \frac{9\eta}{2}}; \quad (\text{B.27})$$

$$\left| \sum_{j,j'=0}^{ph_n-1} \mathbb{E} \left(\widehat{\varepsilon}_{i+j}^n \widehat{\varepsilon}_{i+j'}^n \mid \mathcal{K}_i^n \right) - (\Delta_n^\eta \gamma_i^n)^4 (\overline{\mathbf{r}}(\mathbf{j}')_n - \overline{\mathbf{r}}(\mathbf{0}_2)_n)^2 \right| \leq K_p \Delta_n^{\frac{1}{4} + \frac{9\eta}{2}}; \quad (\text{B.28})$$

$$\left| \sum_{j,j'=0}^{ph_n-1} \mathbb{E} \left(\widehat{X}_{i+j}^{c,n} \widehat{X}_{i+j'}^{c,n} \mid \mathcal{K}_i^n \right) - (\sigma_i^n \Delta_n^\eta \gamma_i^n)^2 \overline{\mathbf{r}}(j, j')_n \widehat{G}_{j,j'}^{n,i} \right| \leq K_p \Delta_n^{\frac{1}{4} + \frac{9\eta}{2}}; \quad (\text{B.29})$$

$$\left| \sum_{j,j'=0}^{ph_n-1} \mathbb{E} \left(\widehat{X}_{i+j}^{c,n} \widehat{\varepsilon}_{i+j'}^n \mid \mathcal{K}_i^n \right) \right| \leq K_p \Psi_{i,j,j'}^n \Delta_n^{\frac{1}{4} + \frac{9\eta}{2}}; \quad (\text{B.30})$$

$$\left| \sum_{j,j'=0}^{ph_n-1} \mathbb{E} \left(\widehat{X}_{i+j}^{c,n} \widehat{X}_{i+j'}^{c,n} \mid \mathcal{K}_i^n \right) \right| \leq K_p \Psi_{i,j,j'}^n \Delta_n^{\frac{3}{4} + \frac{5\eta}{2}}; \quad (\text{B.31})$$

$$\left| \sum_{j,j'=0}^{ph_n-1} \mathbb{E} \left(\widehat{\varepsilon}_{i+j'}^n \widehat{X}_{i+j}^{c,n} \mid \mathcal{K}_i^n \right) \right| \leq K_p \Psi_{i,j,j'}^n \Delta_n^{\frac{1}{4} + \frac{7\eta}{2}}. \quad (\text{B.32})$$

(B.27) follows directly from Lemma JLZ-A.2 and our choice of h_n ; (B.28) and (B.29) are consequences of (B.19) and (B.20); (B.30) follows directly from (B.8) and (JLZ-A.19) (the second part). Since $\widehat{X}_{i+j}^{c,n} \widehat{X}_{i+j'}^{c,n} = \widehat{X}_{i+j}^{c,n} \overline{X}_{i+j'}^{c,n} \overline{\varepsilon}_{i+j'}^n$, an application of (B.9) plus Cauchy-Schwarz

inequality yield (B.31). To see (B.32), we have $\widehat{\varepsilon}_{i+j}^n \widehat{X}^c \varepsilon_{i+j}^n = \overline{X}_{i+j}^{c,n} \sum_{k=1}^3 \mathfrak{D}(k)_{i,j,j'}^n$, where

$$\begin{aligned} \mathfrak{D}(1)_{i,j,j'}^n &:= (\widehat{\varepsilon}_{i+j}^n)^2 \overline{\varepsilon}_{i+j}^n - (\gamma_i^n \Delta_n^\eta)^3 r_{3,j,j'}; \\ \mathfrak{D}(2)_{i,j,j'}^n &:= ((\gamma_i^n)^3 - (\gamma_{i+j}^n)^3) (\Delta_n^\eta)^3 r_{3,j,j'}; \\ \mathfrak{D}(3)_{i,j,j'}^n &:= \Delta_n^{2\eta} (\gamma_{i+j}^n)^2 (\Delta_n^\eta r_{3,j,j'} \gamma_{i+j}^n - \mathbf{r}(\mathbf{0}_2)_n \overline{\varepsilon}_{i+j}^n), \text{ where } r_{3,j,j'} := \overline{\mathbf{r}}(\mathbf{0}_2, |j-j'|)_n. \end{aligned}$$

Note that $|r_{3,j,j'}| \leq K h_n^{-2}$. Now (B.8) yields $\left| \mathbb{E} \left(\mathfrak{D}(1)_{i,j,j'}^n \overline{X}_{i+j}^{c,n} | \mathcal{K}_i^n \right) \right| \leq K \Psi_{i,j,j'}^n \Delta_n^{2\eta + \frac{3}{2}}$; Cauchy-Schwarz inequality and (JLZ-A.17) give $\left| \mathbb{E} \left(\mathfrak{D}(2)_{i,j,j'}^n \overline{X}_{i+j}^{c,n} | \mathcal{K}_i^n \right) \right| \leq K \Psi_{i,j,j'}^n \Delta_n^{2\eta + \frac{3}{2}}$; (B.9) leads to $\left| \mathbb{E} \left(\mathfrak{D}(3)_{i,j,j'}^n \overline{X}_{i+j}^{c,n} | \mathcal{K}_i^n \right) \right| \leq K \Psi_{i,j,j'}^n \Delta_n^{\frac{3\eta}{2} + \frac{5}{4}}$. ■

Lemma B.6. For $v > 2, \eta \in [0, 1/6)$, we have for any fixed $p \geq 1$ that

$$\eta_m F(p)_t^n \xrightarrow{\mathbb{P}} 0; \quad (\text{B.33})$$

$$\eta_m F'(p)_t^n \xrightarrow{\mathbb{P}} 0; \quad (\text{B.34})$$

$$\eta_m \widehat{C}(p)_t^n \xrightarrow{\mathbb{P}} 0; \quad (\text{B.35})$$

$$\eta_m \widehat{C}_t^{n,1} \xrightarrow{\mathbb{P}} 0; \quad (\text{B.36})$$

$$\eta_m \widehat{C}_t^{n,2} \xrightarrow{\mathbb{P}} 0; \quad (\text{B.37})$$

$$\mathbb{E}((\eta_m M'(p)_t^n)^2) \leq Kt/p. \quad (\text{B.38})$$

Proof. Note that $J_n(p, t) \leq \frac{K_p t}{h_n \Delta_n}$, we have by (B.24) that $\mathbb{E}(|F(p)_t^n|) \leq K \sqrt{\Delta_n}$, which leads to $\mathbb{E}(|\eta_m F(p)_t^n|) \leq K/\sqrt{h_n} \rightarrow 0$. The same applies to $F'(p)_t^n$. The proofs of (B.33) and (B.34) are complete.

By (JLZ-A.17), we have

$$\mathbb{E} \left((\widehat{X}_i^{c,n})^2 | \mathcal{F}_i^n \right) \leq K (h_n \Delta_n)^2 \leq K \Delta_n^{1+2\eta}. \quad (\text{B.39})$$

Next, by (B.17) and (B.18), we have the following estimates

$$\mathbb{E} \left((\widehat{\varepsilon}_i^n)^2 + (\widehat{X}^c \varepsilon_i^n)^2 | \mathcal{F}_i^n \right) \leq K \Psi_i^n \Delta_n^{1+2\eta}. \quad (\text{B.40})$$

Thus, (B.39) and (B.40) lead to the estimate that $|\mathbb{E}(\zeta_i^n | \mathcal{K}_i^n)| \leq K \Psi_i^n \Delta_n^{1+2\eta}$, where $\zeta_i^n := (\widehat{X}_i^{c,n})^2 + (\widehat{\varepsilon}_i^n)^2 + (\widehat{X}^c \varepsilon_i^n)^2$. For any integer w , we have $\mathbb{E}(A(w)_s^n) \leq K w \Delta_n^{1+2\eta}$, where $A(w)_s^n = \sum_{i=N_s^n+1}^{N_s^n+w} \zeta_i^n$. To prove (B.35), it suffices to show that the sequences $\{s_n\}, \{w_n\}$ constructed in the proof of Lemma JLZ-A.13 satisfy

$$w_n \Delta_n^{1+2\eta} \rightarrow 0, \mathbb{P}(B_n) \rightarrow 1, \text{ where } B_n = \{N_{s_n}^n \leq N_t^n - h_n, N_t^n + (p+3)h_n \leq N_{s_n}^n + w_n\}. \quad (\text{B.41})$$

Then, (B.37) follows immediately from the following inequalities

$$\mathbb{P}\left((\eta_n \widehat{C}(p)_i^n)^2 > \varepsilon\right) \leq \mathbb{P}(B_n^c) + \frac{Kp}{\varepsilon} \mathbb{E}(A(w)_s^n) \leq \mathbb{P}(B_n^c) + \frac{Kpw_n \Delta_n^{1+2\eta}}{\varepsilon}.$$

A sufficient condition for (B.41) to hold is $\frac{1}{2} + \rho' < 1 + 2\eta$, and indeed it holds since $\rho' < 1/2$.

(B.36) follows immediately from the proof of Lemma A.14 in Jacod et al. (2019) since $h_n \Delta_n^{a-1/4} \rightarrow 0$ where $a \in [1/2 + \rho', (\kappa - 1)/\kappa)$. To see (B.37), we note that Lemma JLZ-A.6 with $w = z = 1$ yields $\mathbb{E}(|\overline{Y}_i^n|^2 - |\overline{Y}_i^{c,n}|^2|) \leq K(h_n \Delta_n)^{\frac{3}{2} + \delta}$ for some $\delta > 0$. Now it follows that $\mathbb{E}\left(|\eta_n \widehat{C}_t^{n,2}|\right) \leq K(h_n \Delta_n)^\delta \rightarrow 0$. (B.26) implies $\mathbb{E}((\zeta(2)_i^n)^2 | \mathcal{K}_i^n) \leq K \Delta_n^{4\eta}$, and it leads to (B.38). ■

Proposition B.1. Assume $\eta \in [0, \frac{1}{8})$, $v > 1 + \frac{2}{1-2\eta}$. For any fixed $p \geq 2$, the sequence of processes $\eta_n M(p)^n$ converges \mathcal{F}_∞ -stably in law to the process

$$Y(p)_t = \int_0^t \beta(p)_s dB_s,$$

where B is as in Theorem 4.1 in Li and Linton (2022a) and $\beta(p)_t$ is the square root of

$$\beta(p)_t^2 := \frac{4}{\phi^2(0)} \left(\frac{\theta(p\Phi_{00} - \overline{\Phi}_{00})\sigma_t^4}{(p+2)\alpha_t} + \frac{2(p\Phi_{01} - \overline{\Phi}_{01})R\sigma_t^2\gamma_t^2}{\theta(p+2)} + \frac{R^2(p\Phi_{11} - \overline{\Phi}_{11})\alpha_t\gamma_t^4}{\theta^3(p+2)} \right).$$

Proof. Let $\widehat{\eta}(p)_j^n = \eta(p)_j^n - \overline{\eta}(p)_j^n$, and $\Delta(V, p)_j^n := V_{j(p+2)h_n}^n - V_{(j-1)(p+2)h_n}^n$ for any process V . $\mathcal{M} = \mathcal{M}_1 \cup \{W\}$, where \mathcal{M}_1 is the set of all bounded martingales orthogonal to W . By the standard limit theorem, see, e.g., Jacod and Shiryaev (2003) Theorem IX 7.28, we need to prove the following

$$t > 0 \Rightarrow \sum_{j=1}^{J_n(p,t)} \mathbb{E}((\eta_n \widehat{\eta}(p)_j^n)^2 | \mathcal{H}(p)_{j-1}^n) \xrightarrow{\mathbb{P}} \int_0^t \beta(p)_s^2 ds; \quad (\text{B.42})$$

$$t > 0 \Rightarrow \sum_{j=1}^{J_n(p,t)} \mathbb{E}((\eta_n \widehat{\eta}(p)_j^n)^4 | \mathcal{H}(p)_{j-1}^n) \xrightarrow{\mathbb{P}} 0; \quad (\text{B.43})$$

$$t > 0, V \in \mathcal{M} \Rightarrow \sum_{j=1}^{J_n(p,t)} \mathbb{E}(\eta_n \widehat{\eta}(p)_j^n \Delta(V, p)_j^n | \mathcal{H}(p)_{j-1}^n) \xrightarrow{\mathbb{P}} 0. \quad (\text{B.44})$$

Proof of (B.42). Note that (B.24) gives $(\overline{\eta}(p)_j^n)^2 \leq K\Psi_i^n \Delta_n^{1+4\eta}/h_n^2 \leq K\Psi_i^n \Delta_n^{2+2\eta}$. On the other hand, we have $J_n(p, t) \leq \frac{K_p t}{h_n \Delta_n}$, thus $\sum_{j=1}^{J_n(p,t)} \mathbb{E}((\eta_n \overline{\eta}(p)_j^n)^2 | \mathcal{H}(p)_{j-1}^n) \leq K_p t \Delta_n \rightarrow 0$. Therefore, we will replace $\widehat{\eta}(p)_j^n$ by $\eta(p)_j^n$ to prove (B.42).

(B.26) and Lemma JLZ-A.9 imply

$$\left| \mathbb{E} \left((\zeta(p)_{m(j)_p^n}^n)^2 \middle| \mathcal{K}_{m(j)_p^n}^n \right) - 4 \left(\frac{(\sigma_{m(j)_p^n}^n)^2 h_n^2 \Delta_n}{\alpha_{m(j)_p^n}^n} \right)^2 (p\Phi_{00} - \bar{\Phi}_{00}) - 4(\gamma_{m(j)_p^n}^n \Delta_n^\eta)^4 (p\Phi_{11} - \bar{\Phi}_{11}) R^2 \right. \\ \left. - 8 \frac{R(\sigma_{m(j)_p^n}^n \gamma_{m(j)_p^n}^n h_n)^2 \Delta_n^{2\eta+1} (p\Phi_{01} - \bar{\Phi}_{01})}{\alpha_{m(j)_p^n}^n} \right| \leq K_p \Psi_{j,p}^n \Delta_n^{\frac{5\eta}{2} + \frac{1}{4}}. \quad (\text{B.45})$$

Note that the contribution of the error terms is given by $\eta_n^2 \Delta_n^{\frac{5\eta}{2} + \frac{1}{4}} J_n(p, t) / h_n^2 \asymp \Delta_n^{\frac{1}{4} - \frac{3\eta}{2}} \rightarrow 0$ since $\eta < 1/6$. Now by (JLZ-A.43), and $\phi_0^n \rightarrow \phi(0)$, we have (B.42).

Proof of (B.43). For the same reasoning as in the previous proof, it suffices to replace $\widehat{\eta}(p)_j^n$ by $\eta(p)_j^n$. (B.25) yields $\mathbb{E} \left((\eta(p)_j^n)^4 \middle| \mathcal{H}(p)_{j-1}^n \right) \leq K_p \Psi_{j,p}^n \Delta_n^{8\eta + (\frac{1}{2} - \eta)(v \wedge 4)}$. Since $J_n(p, t) \leq \frac{K_p t}{h_n \Delta_n}$, $\mathbb{E} \left(\Psi_{j,p}^n \right) \leq 1$, we have

$$\mathbb{E} \left(\sum_{j=1}^{J_n(p,t)} \mathbb{E} \left((\eta_m \widehat{\eta}(p)_j^n)^4 \middle| \mathcal{H}(p)_{j-1}^n \right) \right) \leq K_p \Delta_n^{5\eta + (v \wedge 4)(\frac{1}{2} - \eta) - \frac{3}{2}} \rightarrow 0.$$

This finishes the proof of (B.43).

Proof of (B.44). Since V is a martingale, we have $\mathbb{E} \left(\Delta(V, p)_j^n \middle| \mathcal{H}(p)_{j-1}^n \right) = 0$. Thus, it suffices to prove

$$\frac{\eta_n}{h_n} \sum_{j=1}^{J_n(p,t)} \mathbb{E} \left(\zeta(p)_{m(j)_p^n}^n \Delta(V, p)_j^n \middle| \mathcal{H}(p)_{j-1}^n \right) \xrightarrow{\mathbb{P}} 0. \quad (\text{B.46})$$

Let

$$Z(k)_j^n = \sum_{i=m(j)_p^n}^{m(j)_p^n + ph_n - 1} z(k)_i^n, \text{ where} \quad (\text{B.47}) \\ z(1)_i^n = 2D_1^{n,i}, \quad z(2)_i^n = 2(D_2^{n,i} + D_3^{n,i} + D_4^{n,i}), \\ z(3)_i^n = \widehat{\varepsilon}_i^n, \quad z(4)_i^n = \widehat{X}^c \varepsilon_i^n.$$

Note that the $D_\ell^{n,i}$, $\ell = 1, 2, 3, 4$ are defined in the proof of Lemma JLZ-A.2. To prove (B.46), it is left to show $k = 1, 2, 3, 4$ that

$$A(k)_n := \frac{\eta_n}{h_n} \sum_{j=1}^{J_n(p,t)} \mathbb{E} \left(Z(k)_j^n \Delta(V, p)_j^n \middle| \mathcal{H}(p)_{j-1}^n \right) \xrightarrow{\mathbb{P}} 0, \quad (\text{B.48})$$

since $\zeta(p)_{m(j)_p}^n = \sum_{k=1}^4 Z(k)_j^n$. First, we note that for i that is in the range of defining the variable $Z(1)_j^n$, we have $\mathbb{E} \left(z(1)_i^n \Delta(V, p)_j^n \middle| \mathcal{H}(p)_{j-1}^n \right) = 2 \sum_{\ell=1}^3 \mathbb{E} (H_\ell^n | \mathcal{F}_i^n)$, where $H_\ell^n, \ell = 1, 2, 3$ are defined on p.98 of [Jacod et al. \(2019\)](#), and their analysis implies

$$\left| \sum_{\ell=1}^3 \mathbb{E} (H_\ell^n | \mathcal{F}_i^n) \right| \leq K_p \left(h_n \Delta_n^{\frac{3}{2}} \vee (h_n \Delta_n)^2 \vee h_n^2 \Delta_n^{\frac{3}{2} + \rho} \vee h_n^3 \Delta_n^{\frac{5}{2}} \right) \leq K_p \Delta_n^{(1+\eta) \wedge (2\eta + \rho + \frac{1}{2})}, \quad (\text{B.49})$$

which further implies $|A(1)_n| \leq K_p \Delta_n^{(1+\eta) \wedge (2\eta + \rho + \frac{1}{2}) - (\frac{3}{2}\eta + \frac{3}{4})} \rightarrow 0$, since $\rho > \frac{1}{4} - \frac{\eta}{2}$ and $\eta < \frac{1}{4}$. [\(B.48\)](#) is proved for $k = 1$.

Next, Cauchy-Schwarz inequality implies

$$|\mathbb{E} (z(k)_j^n \Delta(V, p)_j^n | \mathcal{H}(p)_{j-1}^n)| \leq K_p \Gamma(k)_j^n \sqrt{\mathbb{E} \left((\Delta(V, p)_j^n)^2 | \mathcal{H}(p)_{j-1}^n \right)}. \quad (\text{B.50})$$

Now we figure out $\Gamma(k)_j^n$ for $k = 2, 3, 4$.

(1) $k = 2$. Using the estimates in [\(JLZ-S.1\)](#), we have $\Gamma(2)_j^n = (h_n \Delta_n)^{\frac{3}{2}} \leq K \Delta_n^{\frac{3\eta}{2} + \frac{3}{4}}$.

(2) $k = 3$. [\(B.8\)](#) yields $\Gamma(3)_j^n = \Psi_{j,p}^n \Delta_n^{\frac{3\eta}{2} + \frac{3}{4}}$.

(3) $k = 4$. [\(B.9\)](#) leads to the estimate $\Gamma(3)_j^n = \Psi_{j,p}^n \Delta_n^{\frac{v}{2} + \frac{1}{4} - (v - \frac{3}{2})\eta} \leq \Psi_{j,p}^n \Delta_n^{\frac{3\eta}{2} + \frac{3}{4}}$.

Next, let $\beta_n := \mathbb{E} \left(\sum_{j=1}^{J_n(p,t)} (\Delta(V, p)_j^n)^2 \right)$, $\delta(k)_n := \frac{\eta_n^2}{h_n^2} \mathbb{E} \left(\sum_{j=1}^{J_n(p,t)} \left(\sum_{i=m(j)_p}^{m(j)_p + ph_n - 1} \Gamma(k)_i^n \right)^2 \right)$. We have by [\(B.50\)](#) that $\delta(k)_n \leq K_{p,t} \Delta_n^{\eta + \frac{1}{2}}$, and $\beta_n \leq \mathbb{E}((V_\infty - V_0)^2) < \infty$ since V is a square-integrable martingale. Since $\mathbb{E}(A(k)_n^2) \leq K_p \delta(k)_n \beta_n$ by Cauchy-Schwarz inequality, we see that the proof of [\(B.48\)](#) for $k = 2, 3, 4$ is complete. This finishes the proof of [\(B.46\)](#). ■

Lemma B.7. *Given two sequences of integers d_n, k_n and a fixed integer ℓ , we have*

$$\mathbb{E} \left(\left(r(\chi, \ell)_{i, d_n}^{n, k_n} - r(\ell) \right)^2 \right) \leq K \left(\frac{k_n}{d_n} \vee \frac{1}{(k_n + \ell)^v} \right)$$

where $r(\chi, \ell)_{i, d_n}^{n, k_n} := \frac{1}{d_n} \sum_{d=i+k_n-1}^{i+d_n+k_n} \Delta_{d, \ell}^{n, k_n} \chi$, and $\Delta_{d, \ell}^{n, k_n} \chi = (\chi_{d+\ell} - \chi_{d+\ell+k_n})(\chi_d - \chi_{d-k_n})$.

Proof. Let $r(\ell; k_n) := \mathbb{E} \left(\Delta_{d, \ell}^{n, k_n} \chi \right)$. The mixing property of the χ series implies

$$\mathbb{E} \left(\left(r(\chi, \ell)_{i, d_n}^{n, k_n} - r(\ell; k_n) \right)^2 \right) \leq K \frac{k_n}{d_n}.$$

Moreover, Lemma S2 of [Li and Linton \(2022b\)](#) implies $|r(\ell) - r(\ell; k_n)| \leq K(k_n + \ell)^{-v}$. Now the result follows. ■

Lemma B.8. *Assume $\eta \in [0, 1/6)$, $v > \frac{4}{1-6\eta}$, $\ell_n \asymp \Delta_n^{-\ell}$, $k_n \asymp \Delta_n^{-k}$, with*

$$\ell \in \left(\frac{1+2\eta}{4(v-1)}, k \right), \quad k \in \left(\ell \vee \frac{1+2\eta}{2(v-1)}, \frac{1}{6} - \eta \right). \quad (\text{B.51})$$

For all t , we have $\eta_n \widehat{C}_t^{n,3} \xrightarrow{\mathbb{P}} 0$.

Proof. Let $\bar{\mathbf{r}}(\mathbf{0}_2)_{\ell_n} := \frac{1}{h_n} \sum_{|\ell| \leq \ell_n} \widetilde{\phi}_\ell^n r(\ell)$. Note that $\bar{\mathbf{r}}(\mathbf{0}_2)_n = \frac{1}{h_n} \sum_{\ell \in \mathbb{Z}} \widetilde{\phi}_\ell^n r(\ell)$, thus we have

$$|\bar{\mathbf{r}}(\mathbf{0}_2)_n - \bar{\mathbf{r}}(\mathbf{0}_2)_{\ell_n}| \leq \frac{K}{h_n \ell_n^{v-1}} \Rightarrow \Delta_n^{2\eta-1} |\bar{\mathbf{r}}(\mathbf{0}_2)_n - \bar{\mathbf{r}}(\mathbf{0}_2)_{\ell_n}| / h_n = o(1/\eta_n). \quad (\text{B.52})$$

Moreover,

$$\left| \frac{\Delta_n^{2\eta} \bar{\mathbf{r}}(\mathbf{0}_2)_n}{h_n \phi_0^n} \sum_{i=0}^{k_n-1} (\gamma_i^n)^2 \right| \leq K \Delta_n^{2\eta} k_n / (h_n)^2 = o(1/\eta_n). \quad (\text{B.53})$$

(B.52) and (B.53) imply that it suffices to prove

$$\frac{\Delta_n^{2\eta} \bar{\mathbf{r}}(\mathbf{0}_2)_{\ell_n}}{h_n \phi_0^n} \sum_{i=k_n}^{N_t^n - h_n + 1} (\gamma_i^n)^2 - \frac{1}{h_n^2 \phi_0^n} \sum_{|\ell| \leq \ell_n} \sum_{i=k_n}^{N_t^n - h_n + 1} \widetilde{\phi}_\ell^n \Delta_{i,\ell}^{n,k_n} Y = o_p(1/\eta_n).$$

To get this, we note

$$\begin{cases} \frac{1}{\phi_0^n} \sum_{|\ell| \leq \ell_n} \sum_{i=k_n}^{N_t^n - h_n + 1} \frac{1}{h_n^2} \mathbb{E} \left(\left| (\Delta_n^\eta \gamma_i^n)^2 \Delta_{i,\ell}^{n,k_n} \chi - \Delta_{i,\ell}^{n,k_n} Y \right| \right) \leq K \ell_n \Delta_n^{-\eta} ((k_n + \ell_n) \Delta_n)^{\frac{1}{2z}}, \\ \frac{1}{\phi_0^n} \sum_{|\ell| \leq \ell_n} \mathbb{E} \left(\left| \sum_{i=k_n}^{N_t^n - h_n + 1} \frac{1}{h_n^2} \left((\Delta_n^\eta \gamma_i^n)^2 (r(\ell) - \Delta_{i,\ell}^{n,k_n} \chi) \right) \right| \right) \leq K (\ell_n \sqrt{\Delta_n k_n}) \vee \sqrt{k_n^{-v+1}}. \end{cases} \quad (\text{B.54})$$

The first estimate follows from Lemma B.2 and the Cauchy-Schwarz inequality, and the second one follows from the boundedness of γ , the Cauchy-Schwarz inequality and Lemma B.7 $\mathbb{E} \left(\left(\Delta_n \sum_{i=k_n}^{N_t^n - h_n + 1} \left((r(\ell) - \Delta_{i,\ell}^{n,k_n} \chi) \right) \right)^2 \right) \leq K (\Delta_n k_n \vee (k_n + \ell)^{-v})$. For z close to 1, we see that both bounds are of order $o(1/\eta_n)$. This finishes the proof. ■

Lemma B.9. *Assume*

$$\eta \in [0, 1/6), v > 3, k_n \asymp \Delta_n^{-k}, \ell_n \asymp \Delta_n^{-\ell}, h_n \asymp \theta \Delta_n^{\eta - \frac{1}{2}}, \ell > 0, k \in \left(\ell, \frac{1-2\eta}{6} \right). \quad (\text{B.55})$$

We have

$$V_t^{n,1} \xrightarrow{\mathbb{P}} 3\phi^2(0) \int_0^t \frac{\sigma_s^4}{\alpha_s} ds + \frac{6\phi(0)\widetilde{\phi}(0)R}{\theta^2} \int_0^t \sigma_s^2 \gamma_s^2 ds + \frac{3\widetilde{\phi}^2(0)R^2}{\theta^4} \int_0^t \gamma_s^4 dA_s; \quad (\text{B.56})$$

$$V_t^{n,2} \xrightarrow{\mathbb{P}} \frac{\phi(0)R}{\theta^2} \int_0^t \sigma_s^2 \gamma_s^2 ds + \frac{\widetilde{\phi}(0)R^2}{\theta^4} \int_0^t \gamma_s^4 dA_s; \quad (\text{B.57})$$

$$V_t^{n,3} \xrightarrow{\mathbb{P}} \frac{R^2}{\theta^4} \int_0^t \gamma_s^4 dA_s. \quad (\text{B.58})$$

Proof of (B.56). We first have the following decomposition that $(\bar{Y}_i^{c,n})^4 = \sum_{k=1}^{10} \delta_i^{n,k}$, where

$$\begin{aligned} \delta_i^{n,1} &= (\bar{X}_i^{c,n})^4 - 6(\sigma_i^n)^4 \bar{G}_{0,0}^{n,i}; & \delta_i^{n,2} &= 4\bar{X}_i^{c,n} \left((\bar{\varepsilon}_i^n)^3 - (\Delta_n^\eta \gamma_i^n)^3 \bar{\mathbf{r}}(\mathbf{0}_3)_n \right); \\ \delta_i^{n,3} &= 4\bar{X}_i^{c,n} (\Delta_n^\eta \gamma_i^n)^3 \bar{\mathbf{r}}(\mathbf{0}_3)_n; & \delta_i^{n,4} &= 6(\bar{X}_i^{c,n})^2 \left((\bar{\varepsilon}_i^n)^2 - (\Delta_n^\eta \gamma_i^n)^2 \bar{\mathbf{r}}(\mathbf{0}_2)_n \right); \end{aligned}$$

$$\begin{aligned}
\delta_i^{n,5} &= 4 (\bar{X}_i^{c,n})^3 \bar{\varepsilon}_i^n; & \delta_i^{n,6} &= 6 \left((\bar{X}_i^{c,n})^2 - \widehat{G}_{0,0}^{n,i} (\sigma_i^n)^2 \right) (\Delta_n^\eta \gamma_i^n)^2 \bar{\mathbf{r}}(\mathbf{0}_2)_n; \\
\delta_i^{n,7} &= (\bar{\varepsilon}_i^n)^4 - (\Delta_n^\eta \gamma_i^n)^4 \bar{\mathbf{r}}(\mathbf{0}_4)_n; & \delta_i^{n,8} &= 6 (\sigma_i^n)^4 \bar{G}_{0,0}^{n,i}; \\
\delta_i^{n,9} &= 6 \widehat{G}_{0,0}^{n,i} (\sigma_i^n)^2 (\Delta_n^\eta \gamma_i^n)^2 \bar{\mathbf{r}}(\mathbf{0}_2)_n; & \delta_i^{n,10} &= (\Delta_n^\eta \gamma_i^n)^4 \bar{\mathbf{r}}(\mathbf{0}_4)_n.
\end{aligned}$$

(JLZ-A.22) implies $\mathbb{E} \left(\left| \mathbb{E} \left(\delta_i^{n,1} | \mathcal{K}_i^n \right) \right| \right) \leq K \Delta_n^{\frac{5}{4} + \frac{5\eta}{2}}$. Apply (B.8) and (JLZ-A.17), we have

$$\mathbb{E} \left(\left| \mathbb{E} \left(\delta_i^{n,2} | \mathcal{K}_i^n \right) \right| \right) \leq K \Delta_n^{\frac{3}{2} + 2\eta}, \quad \mathbb{E} \left(\left| \mathbb{E} \left(\delta_i^{n,4} | \mathcal{K}_i^n \right) \right| \right) \leq K \Delta_n^{\frac{5}{4} + \frac{5\eta}{2}}.$$

By the boundedness of γ , the fact that $|\bar{\mathbf{r}}(\mathbf{0}_3)_n| \leq K h_n^{\frac{3}{2}}$, and the estimate that $|\mathbb{E}(\bar{X}_i^n | \mathcal{K}_i^n)| \leq K h_n \Delta_n$, we have $\mathbb{E} \left(\left| \mathbb{E} \left(\delta_i^{n,3} | \mathcal{K}_i^n \right) \right| \right) \leq K \Delta_n^{\frac{5}{4} + \frac{5\eta}{2}}$. Next, (B.9) implies $\mathbb{E} \left(\left| \mathbb{E} \left(\delta_i^{n,5} | \mathcal{K}_i^n \right) \right| \right) \leq K \Delta_n^{\frac{5\eta}{2} + \frac{3}{4} + (\frac{1}{2} - \eta)v}$. By (JLZ-A.19), (JLZ-A.21), the boundedness of γ , and Lemma B.1, we have $\mathbb{E} \left(\left| \mathbb{E} \left(\delta_i^{n,6} | \mathcal{K}_i^n \right) \right| \right) \leq K \Delta_n^{\frac{3}{2} + 2\eta}$; and (B.8) implies $\mathbb{E} \left(\left| \mathbb{E} \left(\delta_i^{n,7} | \mathcal{K}_i^n \right) \right| \right) \leq K \Delta_n^{\frac{3}{2} + 2\eta}$. Thus, we have

$$\frac{1}{\widetilde{h}_n} \sum_{i=0}^{N_t^n - h_n + 1} \mathbb{E} \left(\delta_i^{n,k} | \mathcal{K}_i^n \right) \xrightarrow{\mathbb{P}} 0, \quad \text{for } k = 1, \dots, 7. \quad (\text{B.59})$$

Next, we have

$$\left| \mathbb{E} \left(\widehat{G}_{0,0}^{n,i} (\sigma_i^n)^2 (\Delta_n^\eta \gamma_i^n)^2 | \mathcal{F}_i^n \right) - \frac{\theta (\sigma_i^n \gamma_i^n)^2 \Delta_n^{\frac{1}{2} + 3\eta} \phi(0)}{\alpha_i^n} \right| \leq K \Delta_n^{3\eta + \frac{1}{2} + \rho}, \quad (\text{B.60})$$

$$\left| \mathbb{E} \left((\sigma_i^n)^4 \bar{G}_{0,0}^{n,i} | \mathcal{F}_i^n \right) - \frac{\theta^2 (\sigma_i^n)^4 \Delta_n^{1+2\eta} \phi^2(0)}{2(\alpha_i^n)^2} \right| \leq K \Delta_n^{1+2\eta + \rho}, \quad (\text{B.61})$$

$$\mathbf{r}(\mathbf{0}_2)_n = \frac{\widetilde{\phi}(0) R \Delta_n^{\frac{1}{2} - \eta}}{\theta} + o(\Delta_n^{\frac{1}{2} - \eta}), \quad \bar{\mathbf{r}}(\mathbf{0}_4)_n = \frac{3\widetilde{\phi}^2(0) R^2 \Delta_n^{1-2\eta}}{\theta^2} + o(\Delta_n^{1-2\eta}), \quad (\text{B.62})$$

where we obtain (B.60) and (B.61) by and (JLZ-A.38) and (JLZ-A.39). (B.62) follows from Lemma B.1. By Lemma JLZ-A.11, we have

$$\begin{aligned}
\frac{1}{\widetilde{h}_n} \sum_{i=0}^{N_t^n - h_n + 1} \mathbb{E} \left(\delta_i^{n,8} | \mathcal{K}_i^n \right) &\xrightarrow{\mathbb{P}} 3\phi^2(0) \int_0^t \frac{\sigma_s^4}{\alpha_s} ds; \\
\frac{1}{\widetilde{h}_n} \sum_{i=0}^{N_t^n - h_n + 1} \mathbb{E} \left(\delta_i^{n,9} | \mathcal{K}_i^n \right) &\xrightarrow{\mathbb{P}} \frac{6\phi(0)\widetilde{\phi}(0)R}{\theta^2} \int_0^t \sigma_s^2 \gamma_s^2 ds; \\
\frac{1}{\widetilde{h}_n} \sum_{i=0}^{N_t^n - h_n + 1} \mathbb{E} \left(\delta_i^{n,10} | \mathcal{K}_i^n \right) &\xrightarrow{\mathbb{P}} \frac{3\widetilde{\phi}^2(0)R^2}{\theta^4} \int_0^t \gamma_s^4 \alpha_s ds.
\end{aligned} \quad (\text{B.63})$$

Let $\Theta_i^n := (\bar{Y}_i^{c,n})^4 - \mathbb{E} \left((\bar{Y}_i^{c,n})^4 | \mathcal{K}_i^n \right)$ which is $\mathcal{K}_{i+2h_n}^n$ measurable, with $\mathbb{E} \left(\left| \mathbb{E} \left(\Theta_i^n | \mathcal{K}_i^n \right) \right| \right) = 0$ and $\mathbb{E} \left((\Theta_i^n)^2 \right) \leq \mathbb{E} \left((\bar{Y}_i^{c,n})^8 \right) \leq K \Delta_n^{(4\eta+2) \wedge (8\eta + (\frac{1}{2} - \eta)v)}$ by (B.11). Apply Lemma A.6 in Jacod et al.

(2017), we have

$$\frac{1}{\widetilde{h}_n} \mathbb{E} \left(\left| \sum_{i=0}^{N_t^n - h_n + 1} \Theta_i^n \right| \right) \leq K \sqrt{\Delta_n^{(\eta + \frac{1}{2}) \wedge (5\eta + (\frac{1}{2} - \eta)v - \frac{3}{2})}} \rightarrow 0. \quad (\text{B.64})$$

Next, (JLZ-A.33) implies

$$\frac{1}{\widetilde{h}_n} \mathbb{E} \left(\left| \sum_{i=0}^{N_t^n - h_n + 1} \widetilde{\Theta}_i^n \right| \right) \leq K \Delta_n^{(\eta + \frac{1}{2})(2 + \delta) - 1 - 2\eta} \rightarrow 0, \quad (\text{B.65})$$

for some $\delta > 0$, where $\widetilde{\Theta}_i^n := (\overline{Y}_i^n)^4 \mathbf{1}_{\{|\overline{Y}_i^n| \leq u_n\}} - (\overline{Y}_i^{c,n})^4$. In view of (B.59), (B.63), (B.64) and (B.65), we have (B.56). ■

Proof of (B.57). Let $\overline{N}_t^n := N_t^n - h_n - 2k_n - \ell_n$. We denote²

$$\begin{aligned} \Xi_t^{n,1} &:= \frac{1}{h_n \widetilde{h}_n} \sum_{i=k_n}^{\overline{N}_t^n} (\overline{Y}_i^{c,n})^2 R(Y)_{i,h_n}^n; & \Xi_t^{n,2} &:= \frac{1}{h_n \widetilde{h}_n} \sum_{i=k_n}^{\overline{N}_t^n} (\overline{Y}_i^{c,n})^2 (\Delta_n^\eta \gamma_i^n)^2 R(\chi)_{i,h_n}^n; \\ \Xi_t^{n,3} &:= \frac{1}{h_n \widetilde{h}_n} \sum_{i=k_n}^{\overline{N}_t^n} (\overline{Y}_i^{c,n})^2 (\Delta_n^\eta \gamma_i^n)^2 \sum_{|\ell| \leq \ell_n} r(\ell); & \Xi_t^{n,4} &:= \frac{1}{h_n \widetilde{h}_n} \sum_{i=k_n}^{\overline{N}_t^n} (\overline{Y}_i^{c,n})^2 (\Delta_n^\eta \gamma_i^n)^2 R. \end{aligned}$$

First, we prove

$$\Xi_t^{n,4} \xrightarrow{\mathbb{P}} \frac{\phi(0)R}{\theta^2} \int_0^t \sigma_s^2 \gamma_s^2 ds + \frac{\widetilde{\phi}(0)R^2}{\theta^4} \int_0^t \gamma_s^4 dA_s, \quad (\text{B.66})$$

which can be derived by the following three convergence in probability:

$$\frac{1}{h_n \widetilde{h}_n} \sum_{i=k_n}^{\overline{N}_t^n} (\overline{X}_i^{c,n})^2 (\Delta_n^\eta \gamma_i^n)^2 R \xrightarrow{\mathbb{P}} \frac{\phi(0)R}{\theta^2} \int_0^t \sigma_s^2 \gamma_s^2 ds; \quad (\text{B.67})$$

$$\frac{1}{h_n \widetilde{h}_n} \sum_{i=k_n}^{\overline{N}_t^n} (\overline{\varepsilon}_i^n)^2 (\Delta_n^\eta \gamma_i^n)^2 R \xrightarrow{\mathbb{P}} \frac{\widetilde{\phi}(0)R^2}{\theta^4} \int_0^t \gamma_s^4 dA_s; \quad (\text{B.68})$$

$$\frac{1}{h_n \widetilde{h}_n} \sum_{i=k_n}^{\overline{N}_t^n} \overline{X}_i^{c,n} \overline{\varepsilon}_i^n (\Delta_n^\eta \gamma_i^n)^2 R \xrightarrow{\mathbb{P}} 0. \quad (\text{B.69})$$

By (JLZ-A.19), (JLZ-A.21), (JLZ-A.38) for the first estimate, Lemma B.1 and (B.7) for the second estimate, we have

$$\left| \frac{1}{h_n \widetilde{h}_n} \sum_{i=k_n}^{\overline{N}_t^n} \mathbb{E} \left((\overline{X}_i^{c,n})^2 (\Delta_n^\eta \gamma_i^n)^2 R | \mathcal{K}_i^n \right) - \Delta_n \frac{\phi(0)R}{\theta^2} \sum_{i=k_n}^{\overline{N}_t^n} (\gamma_i^n \sigma_i^n)^2 / \alpha_i^n \right| \leq K \Delta_n^\rho;$$

²Recall $R(Y)_{i,h_n}^n$ and $r(\chi; |\ell|)_{i,h_n}^n$ are defined in (10) of Li and Linton (2022a)

$$\frac{1}{h_n \tilde{h}_n} \sum_{i=k_n}^{\bar{N}_t^n} \mathbb{E} \left((\bar{\varepsilon}_i^n)^2 (\gamma_i^n)^2 R | \mathcal{K}_i^n \right) = \frac{\Delta_n \tilde{\phi}(0) R^2}{\theta^4} \sum_{i=k_n}^{\bar{N}_t^n} (\gamma_i^n)^4 + o_p(1).$$

Now Lemma [JLZ-A.11](#) implies

$$\begin{aligned} \frac{1}{h_n \tilde{h}_n} \sum_{i=k_n}^{\bar{N}_t^n} \mathbb{E} \left((\bar{X}_i^{c,n})^2 (\Delta_n^\eta \gamma_i^n)^2 R | \mathcal{K}_i^n \right) &\xrightarrow{\mathbb{P}} \frac{\phi(0) R}{\theta^2} \int_0^t \sigma_s^2 \gamma_s^2 ds; \\ \frac{1}{h_n \tilde{h}_n} \sum_{i=k_n}^{\bar{N}_t^n} \mathbb{E} \left((\bar{\varepsilon}_i^n)^2 (\gamma_i^n)^2 R | \mathcal{K}_i^n \right) &\xrightarrow{\mathbb{P}} \frac{\tilde{\phi}(0) R^2}{\theta^4} \int_0^t \gamma_s^4 dA_s. \end{aligned}$$

Let

$$\Gamma_i^{n,1} := (\bar{X}_i^{c,n} \gamma_i^n)^2 R - \mathbb{E} \left((\bar{X}_i^{c,n} \gamma_i^n)^2 R | \mathcal{K}_i^n \right), \quad \Gamma_i^{n,2} := (\bar{\varepsilon}_i^n \gamma_i^n)^2 R - \mathbb{E} \left((\bar{\varepsilon}_i^n \gamma_i^n)^2 R | \mathcal{K}_i^n \right).$$

Then it's immediate that $\mathbb{E} \left(\left| \mathbb{E} \left(\Gamma_i^{n,k} | \mathcal{K}_i^n \right) \right| \right) = 0$, $\mathbb{E} \left((\Gamma_i^{n,k})^2 \right) \leq K \Delta_n^{1+2\eta}$ and $\Gamma_i^{n,k}$ are $\mathcal{K}_{i+2k_n}^n$ -measurable for $k = 1, 2$. Then Lemma A.6 in [Jacod et al. \(2017\)](#) implies $\frac{1}{h_n} \sum_{i=k_n}^{\bar{N}_t^n - h_n + 1} \Gamma_i^{n,k} \xrightarrow{\mathbb{P}} 0$, for $k = 1, 2$. This proves [\(B.67\)](#) and [\(B.68\)](#).

Let $a_i^n := (\bar{X}_i^{c,n} \bar{\varepsilon}_i^n) (\gamma_i^n)^2$, $a_{ij}^n := -\bar{X}_i^{c,n} (\gamma_i^n)^2 \tilde{g}_j^n \gamma_{i+j}^n \chi_{i+j}$. Then, we have $\sum_{j=1}^{h_n-1} a_{ij}^n = a_i^n$. Since γ is bounded, \mathcal{G} and \mathcal{F}_∞ are independent, and $|\tilde{g}_j^n| \leq K/h_n$, we have the first inequality below; an application of Cauchy-Schwarz inequality and the mixing property of χ yield the second inequality:

$$|\mathbb{E} (a_{ij}^n | \mathcal{K}_i^n)| \leq \frac{K \Psi_i^n}{h_n} |\mathbb{E} (\bar{X}_i^n \gamma_{i+j}^n | \mathcal{F}_i^n)| |\mathbb{E} (\chi_{i+j} | \mathcal{G}_{i-h_n})| \leq \frac{K \Psi_i^n \Delta_n^{\frac{3}{4} - \frac{\eta}{2}}}{(j+h_n)^v}.$$

And it further yields $\mathbb{E} (|\mathbb{E} (a_i^n | \mathcal{K}_i^n)|) \leq K \Delta_n^{\frac{3}{4} - \frac{\eta}{2}} h_n^{1-v}$. Next, by the independence of \mathcal{G} and \mathcal{F}_∞ , the boundedness of γ , [\(JLZ-A.17\)](#) and Lemma [B.1](#), we have

$$\mathbb{E} ((a_i^n)^2) \leq K \mathbb{E} ((\bar{X}_i^n)^2) \mathbb{E} ((\bar{\chi}_i^n)^2) \leq K \Delta_n.$$

Since a_i^n is $\mathcal{K}_{i+2h_n}^n$ -measurable, we can apply Lemma A.6 of [Jacod et al. \(2017\)](#) to get

$$\frac{1}{h_n} \mathbb{E} \left(\sum_{i=k_n}^{\bar{N}_t^n} |a_i^n| \right) \leq \frac{K}{h_n^v \Delta_n^{(1+2\eta)/4} \wedge \sqrt{h_n}} \rightarrow 0. \quad (\text{B.70})$$

The proof of [\(B.69\)](#) is complete.

Now we denote

$$\Xi_t^{n,1} - \Xi_t^{n,2} = \frac{1}{h_n \tilde{h}_n} \sum_{i=k_n}^{\bar{N}_t^n} (\bar{Y}_i^{c,n})^2 \sum_{|\ell| \leq \ell_n} \sum_{k=i+k_n+1}^{i+h_n+k_n} \zeta_{k,\ell}^{n,i},$$

where $\zeta_{k,\ell}^{n,i} := \left(\Delta_{k,\ell}^{n,k_n} Y - (\gamma_i^n \Delta_n^\eta)^2 \Delta_{k,\ell}^{n,k_n} \chi \right) / h_n$. Let $\varrho_{i,\ell}^n := \sum_{k=i+k_n+1}^{i+h_n+k_n} \zeta_{k,\ell}^{n,i}$. By Lemma [B.2](#),

we have $\mathbb{E}\left((\varrho_{i,\ell}^n)^2\right) \leq K\Delta_n^{2\eta}(\Delta_n(k_n + h_n + \ell_n))^{1/z}$. Let $\varrho_i^n := \sum_{|\ell|\leq\ell_n} \varrho_{i,|\ell|}^n$, we have $\mathbb{E}\left((\varrho_i^n)^2\right) \leq K(\ell_n\Delta_n^\eta)^2(\Delta_n(h_n + \ell_n + k_n))^{\frac{1}{z}}$. By Cauchy-Schwarz inequality and (B.11), we have under (B.55) that

$$\mathbb{E}\left(\left|\Xi_t^{n,1} - \Xi_t^{n,2}\right|\right) \leq K\ell_n\Delta_n^{-\eta}(\Delta_n(k_n + h_n + \ell_n))^{\frac{1}{2z}} \rightarrow 0, \quad (\text{B.71})$$

for z close to 1.

Lemma B.7 leads to $\mathbb{E}\left((r(\chi; |\ell|)_{i,h_n}^n - r(\ell))^2\right) \leq K\left(\frac{k_n}{h_n} \vee \frac{1}{k_n^v}\right)$. Thus,

$$\mathbb{E}\left(\left(R(\chi)_{i,h_n}^n - \sum_{|\ell|\leq\ell_n} r(\ell)\right)^2\right) \leq K\ell_n^2\left(\frac{k_n}{h_n} \vee \frac{1}{k_n^v}\right). \quad (\text{B.72})$$

Apply Cauchy-Schwarz inequality, we get from (B.55)

$$\mathbb{E}\left(\left|\Xi_t^{n,2} - \Xi_t^{n,3}\right|\right) \leq K\ell_n\sqrt{\frac{k_n}{h_n} \vee \frac{1}{k_n^v}} \rightarrow 0. \quad (\text{B.73})$$

Next, it's immediate that

$$\mathbb{E}\left(\left|\Xi_t^{n,3} - \Xi_t^{n,4}\right|\right) \leq \frac{K}{\ell_n^{v-1}}. \quad (\text{B.74})$$

Note that $R(Y)_{i,h_n}^n = \frac{1}{h_n} \sum_{|\ell|\leq\ell_n} \sum_{d=i+k_n+1}^{i+k_n+h_n} \sum_{j=1}^3 \mathfrak{J}(\ell, j)_{i,d}^n$, where

$$\begin{aligned} \mathfrak{J}(\ell, 1)_{i,d}^n &:= \Delta_{d,\ell}^{n,k_n} Y - (\Delta_n^\eta \gamma_i^n)^2 \Delta_{d,\ell}^{n,k_n} \chi; \\ \mathfrak{J}(\ell, 2)_{i,d}^n &:= (\Delta_n^\eta \gamma_i^n)^2 \left(\Delta_{d,\ell}^{n,k_n} \chi - r(\ell)\right); \\ \mathfrak{J}(\ell, 3)_{i,d}^n &:= (\Delta_n^\eta \gamma_i^n)^2 r(\ell). \end{aligned}$$

Lemma B.2 yields $\mathbb{E}\left(\left(\mathfrak{J}(\ell, 1)_{i,d}^n\right)^2\right) \leq K\Delta_n^{2\eta}(\ell_n + h_n + k_n)\Delta_n^{\frac{1}{z}}$; Lemma B.7 and the boundedness of γ imply $\mathbb{E}\left(\left(\sum_{d=i+k_n+1}^{i+k_n+h_n} \mathfrak{J}(\ell, 2)_{i,d}^n\right)^2\right) \leq Kh_n k_n \Delta_n^{4\eta}$. We also have the simple estimate that $\mathbb{E}\left(\left(\sum_{|\ell|\leq\ell_n} \mathfrak{J}(\ell, 3)_{i,d}^n\right)^2\right) \leq K\Delta_n^{4\eta} \ell_n \sum_{|\ell|\leq\ell_n} \ell^{-2v} \leq K\Delta_n^{4\eta}$. The three estimates yield

$$\mathbb{E}\left(\left(R(Y)_{i,h_n}^n\right)^2\right) \leq K\left(\ell_n^2\Delta_n^{2\eta}(\ell_n + h_n + k_n)\Delta_n^{\frac{1}{z}} \vee \Delta_n^{4\eta} \vee k_n\ell_n^2 h_n^{-1} \Delta_n^{4\eta}\right). \quad (\text{B.75})$$

Using Lemma JLZ-A.6, Cauchy-Schwarz inequality, and (B.75) (with z close to 1), we have

$$\mathbb{E}\left(\left|\Xi_t^{n,1} - V_t^{n,2}\right|\right) \leq K\Delta_n^{(\eta+\frac{1}{2})\delta/2} \rightarrow 0, \quad (\text{B.76})$$

under (B.55).

We proved (B.57) by (B.66), (B.71), (B.73), (B.74) and (B.76). ■

Proof of (B.58). Let $R_{\ell_n} := \sum_{|\ell| \leq \ell_n} r(|\ell|)$. Let $V_t^{n,3} =: \sum_{j=1}^4 \mathfrak{E}(j)_t^n$, where

$$\begin{aligned}\mathfrak{E}(1)_t^n &:= \frac{1}{\tilde{h}_n h_n^2} \sum_{i=0}^{\bar{N}_t^n} (R(Y)_{i,h_n}^n)^2 - (\Delta_n^\eta \gamma_i^n)^4 (R(\chi)_{i,h_n}^n)^2; \\ \mathfrak{E}(2)_t^n &:= \frac{1}{\tilde{h}_n h_n^2} \sum_{i=0}^{\bar{N}_t^n} (\Delta_n^\eta \gamma_i^n)^4 \left((R(\chi)_{i,h_n}^n)^2 - R_{\ell_n}^2 \right); \\ \mathfrak{E}(3)_t^n &:= \frac{1}{\tilde{h}_n h_n^2} \sum_{i=0}^{\bar{N}_t^n} (\Delta_n^\eta \gamma_i^n)^4 (R_{\ell_n}^2 - R^2); \\ \mathfrak{E}(4)_t^n &:= \frac{1}{\tilde{h}_n h_n^2} \sum_{i=0}^{\bar{N}_t^n} (\Delta_n^\eta \gamma_i^n)^4 R^2.\end{aligned}$$

First, we note $\mathbb{E} \left(\left| R(Y)_{i,h_n}^n + (\Delta_n^\eta \gamma_i^n)^2 R(\chi)_{i,h_n}^n \right| \right) \leq K \Delta_n^{2\eta}$ by (B.72) and (B.75). Thus,

$$\begin{aligned}\mathbb{E} \left(\left| (R(Y)_{i,h_n}^n)^2 - (\Delta_n^\eta \gamma_i^n)^4 (R(\chi)_{i,h_n}^n)^2 \right| \right) &\leq K \Delta_n^{2\eta} \sqrt{\mathbb{E} \left(\left(R(Y)_{i,h_n}^n - (\Delta_n^\eta \gamma_i^n)^2 R(\chi)_{i,h_n}^n \right)^2 \right)} \\ &\leq K \Delta_n^{4\eta} \ell_n (\Delta_n (k_n + h_n + \ell_n))^{\frac{1}{2z}}.\end{aligned}\tag{B.77}$$

The second inequality follows from Lemma B.2. For z close to 1, we have $\mathbb{E}(|\mathfrak{E}(1)_t^n|) \rightarrow 0$, whence $\mathfrak{E}(1)_t^n \xrightarrow{\mathbb{P}} 0$. Similarly, we have

$$\mathbb{E} \left(\left| (R(\chi)_{i,h_n}^n)^2 - R_{\ell_n}^2 \right| \right) \leq K \ell_n \sqrt{\frac{k_n}{h_n} \vee \frac{1}{k_n^v}}; \quad |R^2 - R_{\ell_n}^2| \leq K \ell_n^{-(v-1)},\tag{B.78}$$

and the estimates yield $\mathfrak{E}(2)_t^n + \mathfrak{E}(3)_t^n \xrightarrow{\mathbb{P}} 0$ under (B.55). By (JLZ-A.43), we have $\mathfrak{E}(4)_t^n \xrightarrow{\mathbb{P}} \frac{R^2}{\theta^4} \int_0^t \gamma_s^4 dA_s$. Now (B.58) is proved. ■

Proof of Theorem 4.1 in Li and Linton (2022a). Now Proposition B.1, Lemma B.6 and Lemma B.8 leads to

$$\lim_{p \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\left| \eta_n(\hat{C}_t^n - C_t - M(p)^n) \right| > \varepsilon \right) = 0$$

for any $\varepsilon > 0$. A standard argument, see, e.g., Theorem 3.1 in Jacod et al. (2019) shows that $\beta(p)_s(\omega)^2 \rightarrow \beta_s(\omega)$ for all s and ω , and $\beta(p)_s^2 \leq K$, which yields $Y(p)_t \xrightarrow{\mathbb{P}} Y_t$.

To see the convergence to a standard normal variable, it suffices to have $K_{g,1} V_t^{n,1} + 2K_{g,2} V_t^{n,2} + K_{g,3} V_t^{n,3} \xrightarrow{\mathbb{P}} \frac{\phi^4(0)}{4\theta} \int_0^t \beta_s^2 ds$, which follows from (B.56), (B.57) and (B.58). ■

We now introduce some additional notations that will be used to prove results about the SV estimators.

$$\tilde{N}_t^n := N_t^n + l_n; \quad m(j)_{p,t}^n := N_t^n + m(j)_p^n; \quad s_n \asymp \Delta_n l_n;$$

$$\begin{aligned}
\tilde{\mathcal{H}}(p)_j^n &:= \mathcal{K}_{m(j+1)_{p,t}^n}^n; & \tilde{J}_{n,p} &:= 1 + \left\lceil \frac{l_n}{(p+2)h_n} \right\rceil; \\
\tilde{\mathcal{H}}'(p)_j^n &:= \mathcal{K}_{m(j+1)_{p,t+ph_n}^n}^n; & \tilde{I}_n(p,t) &:= N_t^n + (p+2)h_n\tilde{J}_{n,p} - 1; \\
\tilde{\eta}(p)_j^n &:= \frac{1}{s_n h_n \phi_0^n} \zeta(p)_{m(j)_{p,t}^n}^n; & \bar{\eta}(p)_j^n &:= \mathbb{E} \left(\tilde{\eta}(p)_j^n \mid \tilde{\mathcal{H}}(p)_{j-1}^n \right); \\
\tilde{\eta}'(p)_j^n &:= \frac{1}{s_n h_n \phi_0^n} \zeta(2)_{m(j)_{p,t+ph_n}^n}^n; & \bar{\eta}'(p)_j^n &:= \mathbb{E} \left(\tilde{\eta}'(p)_j^n \mid \tilde{\mathcal{H}}'(p)_{j-1}^n \right); \\
\tilde{F}(p)_t^n &:= \sum_{j=1}^{\tilde{J}_{n,p}} \bar{\eta}(p)_j^n; & \tilde{M}(p)_t^n &:= \sum_{j=1}^{\tilde{J}_{n,p}} (\tilde{\eta}(p)_j^n - \bar{\eta}(p)_j^n); \\
\tilde{F}'(p)_t^n &:= \sum_{j=1}^{\tilde{J}_{n,p}} \bar{\eta}'(p)_j^n; & \tilde{M}'(p)_t^n &:= \sum_{j=1}^{\tilde{J}_{n,p}} (\tilde{\eta}'(p)_j^n - \bar{\eta}'(p)_j^n); \\
\tilde{c}(p)_t^n &:= \frac{1}{s_n h_n \phi_0^n} \sum_{i=\tilde{N}_t^n - h_n + 2}^{\tilde{I}_n(p,t)} Z_i; & \tilde{c}_t^{n,1} &:= \frac{1}{s_n h_n \phi_0^n} \sum_{i=N_t^n}^{\tilde{N}_t^n - h_n + 1} \tilde{c}_i^n - \frac{c_t}{\alpha_t}; \\
\tilde{c}_t^{n,2} &:= \frac{1}{s_n h_n \phi_0^n} \sum_{i=N_t^n}^{\tilde{N}_t^n - h_n + 1} \left((\bar{Y}_i^n)^2 \mathbf{1}_{\{|\bar{Y}_i^n| < u_n\}} - (\bar{Y}_i^{c,n})^2 \right); \\
\tilde{c}_t^{n,3} &:= \frac{\bar{\mathbf{r}}(\mathbf{0}_2)_n}{s_n h_n \phi_0^n} \sum_{i=N_t^n}^{\tilde{N}_t^n - h_n + 1} (\gamma_i^n \Delta_n^\eta)^2 - \frac{1}{s_n h_n^2 \phi_0^n} \sum_{|\ell| \leq \ell_n} \sum_{i=N_t^n + k_n}^{\tilde{N}_t^n - h_n + 1} \tilde{\phi}_\ell^n \Delta_{i,\ell}^{n,k_n} Y.
\end{aligned}$$

For all $p > 1$, we have

$$\tilde{c}_t^n - \frac{c_t}{\alpha_t} = \tilde{M}(p)_t^n + \tilde{M}'(p)_t^n + \tilde{F}(p)_t^n + \tilde{F}'(p)_t^n - \tilde{c}(p)_t^n + \sum_{j=1}^3 \tilde{c}_t^{n,j},$$

where $\tilde{c}_t^{n,j}$ is defined in (14) of Li and Linton (2022a).

Lemma B.10. *Let $l_n \rightarrow \infty, h_n \rightarrow \infty, p_n \rightarrow \infty$ and $l_n/h_n \rightarrow \infty, (l_n \vee p_n)\Delta_n \rightarrow 0$. Then, we have*

$$\frac{s_t^n}{s_n} \xrightarrow{\mathbb{P}} \frac{1}{\alpha_t}; \tag{B.79}$$

$$\frac{1}{\tilde{J}_{n,p}} \sum_{j=1}^{\tilde{J}_{n,p}} V_{m(j)_{p,t}^n}^n \xrightarrow{\mathbb{P}} V_t; \tag{B.80}$$

$$\frac{1}{p_n} \sum_{j=1}^{p_n} V_{N_t^n + j}^n \xrightarrow{\mathbb{P}} V_t. \tag{B.81}$$

if V is one of the processes $X, b, \sigma, \sigma^2, \gamma, \alpha, 1/\alpha$, or any power of them.

Proof. Denote $x_{j,t}^n := \frac{\alpha_t \Delta(n,j)}{\Delta_n}$ for $j = N_t^n + 1, \dots, \tilde{N}_t^n$. We first note from (5) in Li and Linton (2022a) that

$$\mathbb{E} \left(\Delta(n,j) \alpha_{j-1}^n / \Delta_n - 1 \mid \mathcal{F}_{j-1}^n \right) \leq K \Delta_n^\rho.$$

Then, we have by the Cauchy-Schwarz inequality, (JLZ-A.6), (5) in Li and Linton (2022a), and successive conditioning, we have

$$|\mathbb{E}(\Delta(n, j)(\alpha_t - \alpha_{j-1}^n)/\Delta_n | \mathcal{F}_t)| \leq K\sqrt{l_n \Delta_n}.$$

The two estimates immediately lead to

$$|\mathbb{E}(x_{j,t}^n - 1 | \mathcal{F}_t)| \leq K(\Delta_n^\rho \vee \sqrt{l_n \Delta_n}). \quad (\text{B.82})$$

Next, apply the estimate in (JLZ-S.14), we have

$$|\mathbb{E}(\Delta(n, j)\Delta(n, j')\alpha_t^2/\Delta_n^2 - 1)| \leq K(\Delta_n^\rho \vee \sqrt{l_n \Delta_n}),$$

which, together with (B.82), yields

$$|\mathbb{E}((x_{j,t}^n - 1)(x_{j',t}^n - 1) | \mathcal{F}_t)| \leq K(\Delta_n^\rho \vee \sqrt{l_n \Delta_n}) \rightarrow 0.$$

We thus have

$$\mathbb{E}\left(\left(\sum_{j=N_t^n+1}^{\tilde{N}_t^n} (x_{j,t}^n - 1)\right)^2 | \mathcal{F}_t\right) / l_n^2 \rightarrow 0.$$

The convergence in (B.79) follows from an application of the Markov inequality.

Now we show (B.80) and (B.81). First, we have the following estimates

$$\begin{aligned} |\mathbb{E}\left(\left(V_{m(j)_{p,t}}^n - V_{N_t^n}^n\right)\left(V_{m(j')_{p,t}}^n - V_{N_t^n}^n\right) | \mathcal{F}_{N_t^n}^n\right)| &\leq K_p h_n \Delta_n \sqrt{j j'}, \\ |\mathbb{E}\left(\left(V_{N_t^n+j}^n - V_{N_t^n}^n\right)\left(V_{N_t^n+j'}^n - V_{N_t^n}^n\right) | \mathcal{F}_{N_t^n}^n\right)| &\leq K \Delta_n \sqrt{j j'}, \end{aligned}$$

which can be obtained by an application of the Cauchy-Schwarz inequality and (JLZ-A.6).

They yield

$$\begin{aligned} \mathbb{E}\left(\left(\sum_{j=1}^{\tilde{J}_{n,p}} (V_{m(j)_{p,t}}^n - V_{N_t^n}^n) / \tilde{J}_{n,p}\right)^2\right) &\leq K_p l_n \Delta_n \rightarrow 0; \\ \mathbb{E}\left(\left(\sum_{j=1}^{p_n} (V_{N_t^n+j}^n - V_{N_t^n}^n) / p_n\right)^2\right) &\leq K p_n \Delta_n \rightarrow 0. \end{aligned}$$

On the other hand, we have $V_{N_t^n}^n \xrightarrow{\mathbb{P}} V_t$ since $t \in [T(n, N_t^n), T(n, N_t^n + 1))$. Now the proof of (B.80) is complete. ■

Lemma B.11. Let $\eta \in [0, 1/6)$, $v > 1 + \frac{2}{1-2\eta}$, and $h_n \asymp \Delta_n^{-\frac{1}{2}+\eta}$, $l_n \asymp \Delta_n^{-l}$, $l \in (\frac{1}{2} - \eta, \frac{3}{4} - \frac{\eta}{2})$. We have

$$\tilde{\eta}_n \tilde{F}(p)_t^n \xrightarrow{\mathbb{P}} 0; \quad (\text{B.83})$$

$$\tilde{\eta}_n \tilde{F}'(p)_t^n \xrightarrow{\mathbb{P}} 0; \quad (\text{B.84})$$

$$\tilde{\eta}_n \tilde{c}(p)_t^n \xrightarrow{\mathbb{P}} 0; \quad (\text{B.85})$$

$$\tilde{\eta}_n \tilde{c}_t^{n,1} \xrightarrow{\mathbb{P}} 0; \quad (\text{B.86})$$

$$\tilde{\eta}_n \tilde{c}_t^{n,2} \xrightarrow{\mathbb{P}} 0; \quad (\text{B.87})$$

$$\mathbb{E}\left(\left(\tilde{\eta}_n \tilde{M}'(p)_t^n\right)^2\right) \leq K_t/p. \quad (\text{B.88})$$

Proof. Note that $\tilde{J}_{n,p} \leq K_p l_n / h_n$, (B.24) implies $\tilde{\eta}_n \mathbb{E}\left(\left|\tilde{F}(p)_t^n\right|\right) \leq K \Delta_n^{\frac{3}{4}-\frac{\eta}{2}} \sqrt{l_n} \rightarrow 0$, thus (B.83) follows. (B.84) can be proved similarly. (JLZ-A.19), (B.17) and (B.18) yield the estimate that $\mathbb{E}\left(\left(\tilde{\eta}_n \tilde{c}(p)_t^n\right)^2\right) \leq K h_n / l_n \rightarrow 0$, whence (B.85).

Denote $\tilde{c}_t^{n,1} := \frac{1}{s_t^n h_n \phi_0^n} \sum_{i=N_t^n}^{\tilde{N}_t^n - h_n + 1} \tilde{c}_i^n - c_t$. We have

$$\tilde{c}_t^{n,1} = \frac{1}{s_t^n h_n \phi_0^n} \sum_{j=N_t^n+1}^{N_t^n+l_n-1} \Delta_j^n C \sum_{l=1 \vee (j+h_n-N_t^n-l_n)}^{(j-N_t^n) \wedge (h_n-1)} (g_l^n)^2 - c_t = \frac{1}{s_t^n} \sum_{j=N_t^n+1}^{N_t^n+l_n} \Delta_j^n C - c_t + \Gamma_t^n,$$

where $\Gamma_t^n \leq K h_n \Delta_n^a$ on the set $\{\bar{M}_t^n \leq \Delta_n^a\}$ with $\bar{M}_t^n := \max\{\Delta(n, i) : i = 1, \dots, \tilde{N}_t^n\}$. By first conditioning on $\mathcal{F}_{N_t^n}^n \vee \sigma(\Delta(n, i) : i = N_t^n + 1, \dots, \tilde{N}_t^n)$, and upon using (A.6) of Jacod et al. (2017), we have

$$\mathbb{E}\left(\left(\frac{1}{s_t^n} \int_{T(n, N_t^n)}^{T(n, \tilde{N}_t^n)} (c_s - c_{N_t^n}^n) ds\right)^2 \middle| \mathcal{F}_{N_t^n}^n\right) \leq K \Delta_n l_n.$$

Using the fact that $\mathbb{P}(\{\bar{M}_t^n \leq \Delta_n^a\}) \rightarrow 1$ (see the proof of Lemma JLZ-A.13), and the convergence that $\tilde{\eta}_n (c_{N_t^n}^n - c_t) \xrightarrow{\mathbb{P}} 0$ since $t \in [T(n, N_t^n), T(n, N_t^n + 1))$, we have that $\tilde{\eta}_n \tilde{c}_t^{n,1} \xrightarrow{\mathbb{P}} 0$. (B.86) follows immediately in view of (B.79).

Next, by successive conditioning and an application of Lemma JLZ-A.6, we get

$$\mathbb{E}\left(\left|\tilde{\eta}_n \tilde{c}_t^{n,2}\right|\right) \leq K \tilde{\eta}_n (\Delta_n h_n)^{\frac{1}{2}+\delta} \rightarrow 0,$$

which yields (B.87) (δ is a positive real according to Lemma JLZ-A.6). Finally, (B.88) follows directly from (B.26), the fact that $\tilde{J}_{n,p} \leq \frac{K_l l_n}{(p+2)h_n}$. ■

Lemma B.12. Assume $\eta \in [0, 1/6)$, $v > \frac{6-4\eta}{3-10\eta}$, and $\ell_n \asymp \Delta_n^{-\ell}$, $k_n \asymp \Delta_n^{-k}$, $l_n \asymp \Delta_n^{-l}$ with

$$\ell \in \left(\frac{1+2\eta}{8(v-1)}, k\right), \quad k \in \left(\frac{1+2\eta}{4(v-1)} \vee \ell, \frac{1}{4} - \frac{5\eta}{6}\right), \quad l \in \left(\frac{1}{2} - \eta, \frac{3}{4} - \frac{\eta}{2}\right). \quad (\text{B.89})$$

Then, we have $\tilde{\eta}_n \tilde{c}_t^{n,3} \xrightarrow{\mathbb{P}} 0$.

Proof. In this proof, we will directly use notations and several estimates that appear in the proof of Lemma B.8. We divide the proof into two steps. In the first step, we show it suffices

to prove

$$\tilde{\eta}_n \tilde{c}_t^{n,3} \xrightarrow{\mathbb{P}} 0, \text{ where } \tilde{c}_t^{n,3} := \frac{\Delta_n^{2\eta} \bar{\mathbf{r}}(\mathbf{0}_2)_{\ell_n}}{s_n h_n \phi_0^n} \sum_{i=N_t^n+k_n}^{\tilde{N}_t^n-h_n} (\gamma_i^n)^2 - \frac{1}{s_n h_n^2 \phi_0^n} \sum_{|\ell| \leq \ell_n} \sum_{i=N_t^n+k_n}^{\tilde{N}_t^n-h_n} \tilde{\phi}_\ell^n \Delta_{i,\ell}^{n,k_n} Y.$$

In the second step, we will prove the above convergence in probability.

Let $\tilde{c}_t^{n,3} - \tilde{c}_t^{n,3} = c(1)_t^{n,3} + c(2)_t^{n,3}$, where

$$c(1)_t^{n,3} := \frac{\Delta_n^{2\eta} (\bar{\mathbf{r}}(\mathbf{0}_2)_n - \bar{\mathbf{r}}(\mathbf{0}_2)_{\ell_n})}{s_n h_n \phi_0^n} \sum_{i=N_t^n}^{\tilde{N}_t^n-h_n} (\gamma_i^n)^2; \quad c(2)_t^{n,3} := \frac{\Delta_n^{2\eta} \bar{\mathbf{r}}(\mathbf{0}_2)_{\ell_n}}{s_n h_n \phi_0^n} \sum_{i=N_t^n}^{N_t^n+k_n-1} (\gamma_i^n)^2.$$

Since $|\bar{\mathbf{r}}(\mathbf{0}_2)_n - \bar{\mathbf{r}}(\mathbf{0}_2)_{\ell_n}| \leq \frac{K}{h_n \ell_n^{v-1}}$, we have $\mathbb{E}\left(|c(1)_t^{n,3}|\right) \leq K \ell_n^{1-v}$, and

$$\tilde{\eta}_n \mathbb{E}\left(|c(1)_t^{n,3}|\right) \leq K \tilde{\eta}_n \ell_n^{1-v} \rightarrow 0,$$

since $\eta < 1/6$ and $\ell > \frac{1}{6(v-1)}$. Next, we have $\mathbb{E}\left(|c(2)_t^{n,3}|\right) \leq K k_n / l_n$. Thus,

$$\tilde{\eta}_n \mathbb{E}\left(|c(2)_t^{n,3}|\right) \leq K \tilde{\eta}_n k_n / l_n \rightarrow 0.$$

On the other hand, we can apply the same analysis to obtain (B.54) to get

$$\mathbb{E}\left(\left|\Delta_{i,\ell}^{n,k_n} Y - (\Delta_n^\eta \gamma_i^n)^2 r(\ell)\right|\right) \leq K \left(\Delta_n^\eta ((k_n + \ell_n) \Delta_n)^{\frac{1}{2z}} \vee \Delta_n^{2\eta} \sqrt{\Delta_n k_n} \vee (k_n + \ell)^{-v}\right),$$

for any $z > 1$, and it leads to the following estimate that

$$\mathbb{E}\left(|\tilde{c}_t^{n,3}|\right) \leq K \left(\ell_n \Delta_n^{-\eta} ((k_n + \ell_n) \Delta_n)^{\frac{1}{2z}} \vee \ell_n \sqrt{\Delta_n k_n} \vee k_n^{-(v-1)}\right).$$

Thus, for z close to 1, we have $\tilde{\eta}_n \mathbb{E}\left(|\tilde{c}_t^{n,3}|\right) \rightarrow 0$ under (B.89).

This finishes the proof. ■

Proposition B.2. Assume $v > 4$, for any fixed $p \geq 2$, we have the finite-dimensional convergence in law for the sequence of processes $\tilde{\eta}_n \tilde{M}(p)_t^n$ to the limiting process $Z(p)_t$

$$\tilde{\eta}_n \tilde{M}(p)_t^n \xrightarrow{\mathcal{L}_f-s} Z(p)_t,$$

which conditional on \mathcal{F}_∞ is Gaussian white noise with the following conditional variance

$$\tilde{\beta}(p)_t^2 := \frac{4}{\phi^2(0)} \left(\frac{\theta(p\Phi_{00} - \bar{\Phi}_{00})\sigma_t^4}{(p+2)\alpha_t^2} + \frac{2(p\Phi_{01} - \bar{\Phi}_{01})R\sigma_t^2\gamma_t^2}{\alpha_t\theta(p+2)} + \frac{R^2(p\Phi_{11} - \bar{\Phi}_{11})\gamma_t^4}{\theta^3(p+2)} \right).$$

Proof. Let $\widehat{\eta}(p)_j^n = \widetilde{\eta}(p)_j^n - \widetilde{\eta}(p)_j^n$. By the standard limit theorem, we need to prove the following

$$\sum_{j=1}^{\widetilde{J}_{n,p}} \mathbb{E} \left((\widetilde{\eta}_n \widehat{\eta}(p)_j^n)^2 \mid \widetilde{\mathcal{H}}(p)_{j-1}^n \right) \xrightarrow{\mathbb{P}} \widetilde{\beta}(p)_t^2; \quad (\text{B.90})$$

$$\sum_{j=1}^{\widetilde{J}_{n,p}} \mathbb{E} \left((\widetilde{\eta}_n \widehat{\eta}(p)_j^n)^4 \mid \widetilde{\mathcal{H}}(p)_{j-1}^n \right) \xrightarrow{\mathbb{P}} 0; \quad (\text{B.91})$$

$$V \in \mathcal{M} \Rightarrow \sum_{j=1}^{\widetilde{J}_{n,p}} \mathbb{E} \left(\widetilde{\eta}_n \widehat{\eta}(p)_j^n \Delta(V, p)_j^n \mid \widetilde{\mathcal{H}}(p)_{j-1}^n \right) \xrightarrow{\mathbb{P}} 0. \quad (\text{B.92})$$

Note that $(\widetilde{\eta}(p)_j^n)^2 \leq K \Psi_i^n \Delta_n^{2+2\eta}$ and $\widetilde{J}_{n,p} \leq K_{p,t} l_n / h_n$ give the estimate that

$$\sum_{j=1}^{\widetilde{J}_{n,p}} \mathbb{E} \left((\widetilde{\eta}_n \widetilde{\eta}(p)_j^n)^2 \mid \widetilde{\mathcal{H}}(p)_{j-1}^n \right) \leq K_{p,t} \Delta_n^3 l_n^2 \rightarrow 0.$$

Whence, we can replace $\widehat{\eta}(p)_j^n$ by $\widetilde{\eta}(p)_j^n$ to prove (B.90). Now (B.90) follows immediately from (B.45) and Lemma B.10. By the same reasoning above, we will replace $\widehat{\eta}(p)_j^n$ by $\widetilde{\eta}(p)_j^n$ to prove (B.91). Since $v > 4$, $\widetilde{J}_{n,p} \leq K l_n / h_n$, we have by (B.25) that

$$\sum_{j=1}^{\widetilde{J}_{n,p}} \mathbb{E} \left((\widetilde{\eta}_n \widetilde{\eta}(p)_j^n)^4 \mid \widetilde{\mathcal{H}}(p)_{j-1}^n \right) \leq K h_n / l_n \rightarrow 0.$$

To prove (B.92), it suffices to show that

$$\frac{1}{\Delta_n^{2\eta} \widetilde{\eta}_n} \sum_{j=1}^{\widetilde{J}_{n,p}} \mathbb{E} \left(\zeta(p)_{m(j)_p^n}^n \Delta(V, p)_j^n \mid \widetilde{\mathcal{H}}(p)_{j-1}^n \right) \xrightarrow{\mathbb{P}} 0. \quad (\text{B.93})$$

Adopt a similar notation as in (B.47) by replacing $m(j)_p^n$ by $m(j)_{p,t}^n$ and $Z(k)_j^n$ by $Z(k)_{j,t}^n$, we can prove the following convergence that is equivalent to (B.93):

$$\widetilde{A}(k)_{n,t} := \frac{1}{\Delta_n^{2\eta} \widetilde{\eta}_n} \sum_{j=1}^{\widetilde{J}_{n,p}} \mathbb{E} \left(Z(k)_{j,t}^n \Delta(V, p)_j^n \mid \widetilde{\mathcal{H}}(p)_{j-1}^n \right) \xrightarrow{\mathbb{P}} 0, \quad k = 1, 2, 3, 4. \quad (\text{B.94})$$

Apply the estimate in (B.49), we have $\widetilde{A}(1)_{n,t} \leq K_{p,t} \Delta_n^{(1-\eta) \wedge (\rho + \frac{1}{2})} / \widetilde{\eta}_n \rightarrow 0$. Next, we have $\widetilde{\delta}(k)_{n,t} \leq K_{p,t} / \widetilde{\eta}_n^2$ where $\widetilde{\delta}(k)_{n,t} := \frac{1}{\Delta_n^{4\eta} \widetilde{\eta}_n^2} \mathbb{E} \left(\sum_{j=1}^{\widetilde{J}_{n,p}} \left(\sum_{i=m(j)_{p,t}^n}^{m(j)_{p,t}^n + p h_n - 1} \Gamma(k)_i^n \right)^2 \right)$ with $\Gamma(k)_i^n$ defined in (B.50). Now Cauchy-Schwarz inequality implies

$$\mathbb{E} \left(\widetilde{A}(k)_{n,t}^2 \right) \leq K_p \widetilde{\delta}(k)_{n,t} \mathbb{E} \left((V_\infty - V_0)^2 \right) \rightarrow 0$$

for $k = 2, 3, 4$. This finishes the proof of (B.94), thus (B.92). ■

Lemma B.13. Assume $v > 4$, $\eta \in [0, \frac{1}{6})$, and k_n, ℓ_n, l_n, h_n satisfy the following asymptotic conditions

$$k_n \asymp \Delta_n^{-k}, \ell_n \asymp \Delta_n^{-\ell}, h_n \asymp \theta \Delta_n^{\eta - \frac{1}{2}}, \ell > 0, k \in \left(\ell, \frac{1-2\eta}{6} \right), l \in \left(\frac{1}{2} - \eta, \frac{3}{4} - \frac{\eta}{2} \right).$$

We have

$$\tilde{V}_t^{n,1} \xrightarrow{\mathbb{P}} \frac{3\phi^2(0)\sigma_t^4}{\alpha_t^2} + \frac{6\phi(0)\tilde{\phi}(0)R\sigma_t^2\gamma_t^2}{\theta^2\alpha_t} + \frac{3\tilde{\phi}^2(0)R^2\gamma_t^4}{\theta^4}; \quad (\text{B.95})$$

$$\tilde{V}_t^{n,2} \xrightarrow{\mathbb{P}} \frac{\phi(0)R\sigma_t^2\gamma_t^2}{\theta^2\alpha_t} + \frac{\tilde{\phi}(0)R^2\gamma_t^4}{\theta^4}; \quad (\text{B.96})$$

$$\tilde{V}_t^{n,3} \xrightarrow{\mathbb{P}} \frac{R^2\gamma_t^4}{\theta^4}; \quad (\text{B.97})$$

Proof of (B.95). In the sequel, we will follow the proofs of Lemma B.9. We will directly use many notations therein without further references.

Using the same decomposition as we did in the proof of (B.56) that $(\bar{Y}_i^{c,n})^4 = \sum_{k=1}^{10} \delta_i^{n,k}$, and the various estimates therein, we have

$$\mathbb{E} \left(\left| \mathbb{E} \left(\delta_i^{n,k} \mid \mathcal{K}_i^n \right) \right| \right) \leq K \Delta_n^{\frac{5}{4} + \frac{5\eta}{2}} \text{ for } k = 1, 2, \dots, 7.$$

Thus,

$$\frac{1}{s_n \tilde{h}_n} \sum_{i=N_t^n}^{\tilde{N}_t^n - h_n + 1} \mathbb{E} \left(\delta_i^{n,k} \mid \mathcal{K}_i^n \right) \xrightarrow{\mathbb{P}} 0, \text{ for } k = 1, 2, \dots, 7.$$

Using the estimates in (B.60), (B.61) and (B.62), we have

$$\begin{aligned} \frac{1}{s_n \tilde{h}_n} \sum_{i=N_t^n}^{\tilde{N}_t^n - h_n + 1} \mathbb{E} \left(\delta_i^{n,8} \mid \mathcal{K}_i^n \right) &\xrightarrow{\mathbb{P}} \frac{3\phi^2(0)\sigma_t^4}{\alpha_t^2}, \\ \frac{1}{s_n \tilde{h}_n} \sum_{i=N_t^n}^{\tilde{N}_t^n - h_n + 1} \mathbb{E} \left(\delta_i^{n,9} \mid \mathcal{K}_i^n \right) &\xrightarrow{\mathbb{P}} \frac{6\phi(0)\tilde{\phi}(0)R\sigma_t^2\gamma_t^2}{\theta^2\alpha_t}, \\ \frac{1}{s_n \tilde{h}_n} \sum_{i=N_t^n}^{\tilde{N}_t^n - h_n + 1} \mathbb{E} \left(\delta_i^{n,10} \mid \mathcal{K}_i^n \right) &\xrightarrow{\mathbb{P}} \frac{3\tilde{\phi}^2(0)R^2\gamma_t^4}{\theta^4}. \end{aligned}$$

Next, we have the estimate $\mathbb{E}((\Theta_i^n)^2) \leq K \Delta_n^{(4\eta+2) \wedge (8\eta + (\frac{1}{2}-\eta)v)}$, and Lemma A.6 of Jacod et al. (2017) implies (recall $v > 4$)

$$\frac{1}{s_n \tilde{h}_n} \mathbb{E} \left(\left| \sum_{i=N_t^n}^{\tilde{N}_t^n - h_n + 1} \Theta_i^n \right| \right) \leq K \sqrt{h_n (1 \wedge \Delta_n^{(1/2-\eta)v + 4\eta - 2}) / l_n} \rightarrow 0.$$

(JLZ-A.33) yields $\frac{1}{s_n \tilde{h}_n} \mathbb{E} \left(\left| \sum_{i=N_t^n}^{\tilde{N}_t^n - h_n + 1} \tilde{\Theta}_i^n \right| \right) \leq K \Delta_n^{\delta(\eta+1/2)} \rightarrow 0$, where $\tilde{\Theta}_i^n := (\bar{Y}_i^n)^4 \mathbf{1}_{\{|\bar{Y}_i^n| \leq u_n\}}$

$(\bar{Y}_i^{c,n})^4$. Now the proof of (B.95) is complete. ■

Proof of (B.96). Let

$$\begin{aligned}\tilde{\Xi}_t^{n,1} &:= \frac{1}{s_n h_n \tilde{h}_n} \sum_{i=N_t^n}^{\hat{N}_t^n} (\bar{Y}_i^{c,n})^2 R(Y)_{i,h_n}^n; & \tilde{\Xi}_t^{n,2} &:= \frac{1}{s_n h_n \tilde{h}_n} \sum_{i=N_t^n}^{\hat{N}_t^n} (\bar{Y}_i^{c,n})^2 (\Delta_n^\eta \gamma_i^n)^2 R(\chi)_{i,h_n}^n; \\ \tilde{\Xi}_t^{n,3} &:= \frac{1}{s_n h_n \tilde{h}_n} \sum_{i=N_t^n}^{\hat{N}_t^n} (\bar{Y}_i^{c,n})^2 (\Delta_n^\eta \gamma_i^n)^2 \sum_{|\ell| \leq \ell_n} r(\ell); & \tilde{\Xi}_t^{n,4} &:= \frac{1}{s_n h_n \tilde{h}_n} \sum_{i=N_t^n}^{\hat{N}_t^n} (\bar{Y}_i^{c,n})^2 (\Delta_n^\eta \gamma_i^n)^2 R.\end{aligned}$$

By (JLZ-A.19), (JLZ-A.21), (JLZ-A.38), and (B.7), we have

$$\begin{aligned}\frac{1}{s_n h_n \tilde{h}_n} \sum_{i=N_t^n}^{\hat{N}_t^n} \left| \mathbb{E} \left((\bar{X}_i^{c,n} \Delta_n^\eta \gamma_i^n)^2 R | \mathcal{K}_i^n \right) - \frac{(\sigma_i^n \Delta_n^\eta \gamma_i^n)^2 h_n \Delta_n \phi(0)}{\alpha_i^n} \right| &\leq K \Delta_n^\rho; \\ \frac{1}{s_n h_n \tilde{h}_n} \sum_{i=N_t^n}^{\hat{N}_t^n} \mathbb{E} \left((\bar{\varepsilon}_i^n \Delta_n^\eta \gamma_i^n)^2 R | \mathcal{K}_i^n \right) &= \frac{\tilde{\phi}(0) R^2}{\theta^4 l_n} \sum_{i=N_t^n}^{\hat{N}_t^n} (\gamma_i^n)^4 + o_p(1).\end{aligned}$$

By Lemma B.10, we have

$$\begin{aligned}\frac{1}{s_n h_n \tilde{h}_n} \sum_{i=N_t^n}^{\hat{N}_t^n} \mathbb{E} \left((\bar{X}_i^{c,n} \Delta_n^\eta \gamma_i^n)^2 R | \mathcal{K}_i^n \right) &\xrightarrow{\mathbb{P}} \frac{R \phi(0) \sigma_t^2 \gamma_t^2}{\theta^2 \alpha_t}; \\ \frac{1}{s_n h_n \tilde{h}_n} \sum_{i=N_t^n}^{\hat{N}_t^n} \mathbb{E} \left((\bar{\varepsilon}_i^n \Delta_n^\eta \gamma_i^n)^2 R | \mathcal{K}_i^n \right) &\xrightarrow{\mathbb{P}} \frac{\tilde{\phi}(0) R^2 \gamma_t^4}{\theta^4}.\end{aligned}\tag{B.98}$$

On the other hand, Lemma A.6 of Jacod et al. (2017) gives

$$\frac{1}{s_n h_n \tilde{h}_n} \mathbb{E} \left(\left| \sum_{i=N_t^n}^{\hat{N}_t^n} \delta_i^n \right| \right) \leq K \sqrt{h_n / l_n},$$

where

$$\delta_i^n := (\bar{X}_i^{c,n} \Delta_n^\eta \gamma_i^n)^2 R - \mathbb{E} \left((\bar{X}_i^{c,n} \Delta_n^\eta \gamma_i^n)^2 R | \mathcal{K}_i^n \right) \text{ or } \delta_i^n := (\bar{\varepsilon}_i^n \Delta_n^\eta \gamma_i^n)^2 R - \mathbb{E} \left((\bar{\varepsilon}_i^n \Delta_n^\eta \gamma_i^n)^2 R | \mathcal{K}_i^n \right).$$

Thus, we have

$$\frac{1}{s_n h_n \tilde{h}_n} \sum_{i=N_t^n}^{\hat{N}_t^n} \left((\bar{X}_i^{c,n} \Delta_n^\eta \gamma_i^n)^2 R - \mathbb{E} \left((\bar{X}_i^{c,n} \Delta_n^\eta \gamma_i^n)^2 R | \mathcal{K}_i^n \right) \right) \xrightarrow{\mathbb{P}} 0;\tag{B.99}$$

$$\frac{1}{s_n h_n \tilde{h}_n} \sum_{i=N_t^n}^{\hat{N}_t^n} \left((\bar{\varepsilon}_i^n \Delta_n^\eta \gamma_i^n)^2 R - \mathbb{E} \left((\bar{\varepsilon}_i^n \Delta_n^\eta \gamma_i^n)^2 R | \mathcal{K}_i^n \right) \right) \xrightarrow{\mathbb{P}} 0.\tag{B.100}$$

Use the same approach to get (B.70), we have for any integer k ,

$$\frac{1}{s_n h_n \tilde{h}_n} \sum_{i=k}^{k+l_n-h_n-2k_n-\ell_n} E_{i,k}^n \leq K(\Delta_n^{\frac{\eta}{2}-\frac{1}{4}} h_n^{-\nu} \vee \sqrt{h_n/l_n}) \rightarrow 0,$$

where $E_{i,k}^n := \mathbb{E}(|\bar{X}_i^{c,n} \bar{\varepsilon}_i^n (\Delta_n^\eta \gamma_i^n)^2 R| | N_t^n = k)$. Now we have

$$\frac{1}{s_n h_n \tilde{h}_n} \sum_{i=N_t^n}^{\hat{N}_t^n} \bar{X}_i^{c,n} \bar{\varepsilon}_i^n (\Delta_n^\eta \gamma_i^n)^2 R \xrightarrow{\mathbb{P}} 0, \quad (\text{B.101})$$

since

$$\frac{1}{s_n h_n \tilde{h}_n} \sum_{i=N_t^n}^{\hat{N}_t^n} \mathbb{E}(|\bar{X}_i^{c,n} \bar{\varepsilon}_i^n (\Delta_n^\eta \gamma_i^n)^2 R|) = \sum_{k=0}^{\infty} \frac{1}{s_n h_n \tilde{h}_n} \sum_{i=k}^{k+l_n-h_n-2k_n-\ell_n} \mathbb{P}(N_t^n = k) E_{i,k}^n \rightarrow 0.$$

(B.98), (B.99), (B.100) and (B.101) lead to

$$\tilde{\Xi}_t^{n,4} \xrightarrow{\mathbb{P}} \frac{R\phi(0)\sigma_t^2\gamma_t^2}{\theta^2\alpha_t} + \frac{\tilde{\phi}(0)R^2\gamma_t^4}{\theta^4}. \quad (\text{B.102})$$

Since

$$\tilde{\Xi}_t^{n,1} - \tilde{\Xi}_t^{n,2} = \frac{1}{s_n h_n \tilde{h}_n} \sum_{i=N_t^n}^{\hat{N}_t^n} (Y_i^{c,n})^2 \varrho_i^n,$$

we now use the estimate that $\mathbb{E}((\varrho_i^n)^2) \leq K\ell_n^2 \Delta_n^{2\eta} (\Delta_n(h_n + k_n + \ell_n))^{\frac{1}{z}}$ for any $z > 1$, (B.11) and Cauchy-Schwarz inequality, we have

$$\mathbb{E}\left(|\tilde{\Xi}_t^{n,1} - \tilde{\Xi}_t^{n,2}|\right) \leq K\Delta_n^{-\eta}\ell_n(\Delta_n(h_n + k_n + \ell_n))^{\frac{1}{2z}} \rightarrow 0, \quad (\text{B.103})$$

for z close to 1 since $\ell < 1/6$. Next, (B.72) gives

$$\mathbb{E}\left(|\tilde{\Xi}_t^{n,2} - \tilde{\Xi}_t^{n,3}|\right) \leq K\ell_n \sqrt{\frac{k_n}{h_n} \vee \frac{1}{k_n^\nu}} \rightarrow 0. \quad (\text{B.104})$$

We also have a simple estimate that

$$\mathbb{E}\left(|\tilde{\Xi}_t^{n,3} - \tilde{\Xi}_t^{n,4}|\right) \leq \frac{K}{\ell_n^{\nu-1}} \rightarrow 0. \quad (\text{B.105})$$

By (B.75), Lemma J LZ-A.6 and Cauchy-Schwarz inequality, we have

$$\mathbb{E}\left(|\tilde{\Xi}_t^{n,1} - \tilde{V}_t^{m,2}|\right) \leq K(h_n \Delta_n)^{\delta/2} \rightarrow 0, \quad (\text{B.106})$$

where $\delta > 0$ according to Lemma J LZ-A.6. Now the proof of (B.96) is complete in view of (B.102), (B.103), (B.104), (B.105), and (B.106). ■

Proof of (B.97). Let $\tilde{V}_t^{m,3} =: \sum_{j=1}^4 \tilde{\mathfrak{E}}(j)_t^n$, where

$$\begin{aligned}\tilde{\mathfrak{E}}(1)_t^n &:= \frac{1}{s_n \tilde{h}_n h_n^2} \sum_{i=N_t^n}^{\hat{N}_t^n} (R(Y)_{i,h_n}^n)^2 - (\Delta_n^\eta \gamma_i^n)^4 (R(\chi)_{i,h_n}^n)^2; \\ \tilde{\mathfrak{E}}(2)_t^n &:= \frac{1}{s_n \tilde{h}_n h_n^2} \sum_{i=N_t^n}^{\hat{N}_t^n} (\Delta_n^\eta \gamma_i^n)^4 \left((R(\chi)_{i,h_n}^n)^2 - R_{\ell_n}^2 \right); \\ \tilde{\mathfrak{E}}(3)_t^n &:= \frac{1}{s_n \tilde{h}_n h_n^2} \sum_{i=N_t^n}^{\hat{N}_t^n} (\Delta_n^\eta \gamma_i^n)^4 (R_{\ell_n}^2 - R^2); \\ \tilde{\mathfrak{E}}(4)_t^n &:= \frac{1}{s_n \tilde{h}_n h_n^2} \sum_{i=N_t^n}^{\hat{N}_t^n} (\Delta_n^\eta \gamma_i^n)^4 R^2.\end{aligned}$$

By (B.77), we have

$$\mathbb{E} \left(\left| \tilde{\mathfrak{E}}(1)_t^n \right| \right) \leq K \ell_n (\Delta_n (\ell_n + k_n + h_n))^{\frac{1}{2z}} \rightarrow 0$$

for z close to 1. Next,

$$\tilde{\mathfrak{E}}(2)_t^n + \tilde{\mathfrak{E}}(3)_t^n \xrightarrow{\mathbb{P}} 0$$

follows immediately from (B.78); and $\tilde{\mathfrak{E}}(4)_t^n \xrightarrow{\mathbb{P}} \frac{R^2 \gamma_t^4}{\theta^4}$ is a direct consequence of Lemma B.10. ■

Proof of Theorem 4.2 in Li and Linton (2022a). Apply the standard arguments, see, e.g., p.98 of Jacod et al. (2019), we derive the following convergence from Proposition B.2:

$$\tilde{\eta}_n \left(\tilde{c}_t^n - \frac{c_t}{\alpha_t} \right) \xrightarrow{\mathcal{L}_f^{-s}} \frac{Z_t}{\alpha_t}.$$

Now the result follows directly from (B.79). The convergence to a standard normal distribution follows immediately from Lemma B.13. ■

References

- CHUNG, K. L. (2001): *A Course in Probability Theory*, Academic press.
- JACOD, J., Y. LI, AND X. ZHENG (2017): “Statistical properties of microstructure noise,” *Econometrica*, 85, 1133–1174.
- (2019): “Estimating the integrated volatility with tick observations,” *Journal of Econometrics*, 208, 80–100.
- JACOD, J. AND A. N. SHIRYAEV (2003): *Limit Theorems for Stochastic Processes*, vol. 288, Springer-Verlag Berlin.
- LI, Z. M., R. J. LAEVEN, AND M. H. VELLEKOOP (2020): “Dependent microstructure noise and integrated volatility estimation from high-frequency data,” *Journal of Econometrics*, 215, 536–558.
- LI, Z. M. AND O. LINTON (2022a): “Robust Estimation of Integrated and Spot Volatility,” .
- (2022b): “Supplement to ‘A ReMeDI for microstructure noise’,” *Econometrica Supplemental Material*.
- RIO, E. (1997): “About the Lindeberg method for strongly mixing sequences,” *ESAIM: Probability and Statistics*, 1, 35–61.
- VARNESKOV, R. T. (2016): “Flat-top realized kernel estimation of quadratic covariation with nonsynchronous and noisy asset prices,” *Journal of Business & Economic Statistics*, 34, 1–22.
- (2017): “Estimating the quadratic variation spectrum of noisy asset prices using generalized flat-top realized kernels,” *Econometric Theory*, 33, 1457–1501.