UNIVERSITY OF CAMBRIDGE
Faculty of Economics

Institute for New Economic Thinking

CAMBRIDGE WORKING PAPERS IN ECONOMICS
CAMBRIDGE-INET WORKING PAPERS

Robust Estimation of Integrated Volatility

Z. Merrick Li
University of Cambridge

Oliver Linton
University of Cambridge

Abstract

We introduce a new method to estimate the integrated volatility (IV) based on noisy high-frequency data. Our method employs the ReMeDI approach introduced by Li and Linton (2021a) to estimate the moments of the 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 endogenous, and the noise process is nonstationary, autocorrelated and dependent on the efficient price. We derive the limit distribution for the proposed estimators under infill asymptotics in a general setting. Our simulation and empirical studies demonstrate the robustness, accuracy and computational efficiency of our estimators compared to several alternatives recently proposed in the literature.

Reference Details

2115	Cambridge Working Papers in Economics
2021/08	Cambridge-INET Working Paper Series
Published	23 February 2021
Websites	www.econ.cam.ac.uk/cwpe
	www.inet.econ.cam.ac.uk/working-papers

Robust Estimation of Integrated Volatility*

Z. Merrick Li† Oliver Linton‡

February 23, 2021

Abstract

We introduce a new method to estimate the integrated volatility (IV) based on noisy high-frequency data. Our method employs the ReMeDI approach introduced by Li and Linton (2021a) to estimate the moments of the 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 endogenous, and the noise process is nonstationary, autocorrelated and dependent on the efficient price. We derive the limit distribution for the proposed estimators under infill asymptotics in a general setting. Our simulation and empirical studies demonstrate the robustness, accuracy and computational efficiency of our estimators compared to several alternatives recently proposed in the literature.

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 such as the *integrated volatility* (IV) 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), provides a consistent estimator of IV. This result dates back to Jacod (2018)[1] and Jacod and

*The project is partially sponsored by the Keynes Fund (JHUL).

†Corresponding author. Faculty of Economics, University of Cambridge, Austin Robinson Building, Sidgwick Avenue, Cambridge, CB3 9DD, United Kingdom. Email: z.merrick.li@gmail.com.

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

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

Protter (1998), and it is introduced to econometrics by Andersen et al. (2003). 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; Da and Xiu, 2020), the pre-averaging method (Podolskij and Vetter, 2009; Jacod et al., 2009, 2010; Li, 2013; Jacod et al., 2019), and the realized kernel (Hansen and Lunde, 2006; Barndorff-Nielsen et al., 2008; Varneskov, 2017). 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 the 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 in the scale of the noise. 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 (Bogousslavsky, 2016; Hendershott et al., 2018). 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 execution of limit orders (Parlour, 1998). Thus, microstructure noise is highly autocorrelated. Third, the size of high-frequency 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 suggest that a flexible modelling and robust estimation

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

of microstructure noise are of great concern.

This paper introduces a robust estimation of the integrated volatility (IV). 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 microstructure noise is serially dependent, nonstationary and could be dependent with the efficient price. We develop a new IV estimator using the ReMeDI estimators (Li and Linton, 2021a) of moments of noise[3] to correct the bias of a pre-averaging type estimator. The estimator inherits the great finite sample properties of the ReMeDI estimators. Our extensive simulation studies show that it performs very well in samples where the data generating process follows different specifications; in particular, it provides accurate estimates when the noise-to-signal ratio (the ratio of the variance of noise and the integrated volatility) varies in a large range. Moreover, the estimator is computationally very efficient. This is an advantage when dealing with the massive high-frequency datasets currently available.

Many recent papers study the estimation of IV with general modelling of microstructure noise. Kalnina and Linton (2008) consider microstructure noise with a time-varying scale; Aït-Sahalia et al. (2011) show that the TSRV (and MSRV) remain valid when the noise is autocorrelated; Hautsch and Podolskij (2013) study q-dependent noise; Li et al. (2020) use a variant of the realized variance to estimate second moments of serially dependent noise and develop a consistent estimator of the IV. The econometric models in the aforementioned papers are quite restrictive compared to two recent studies by Jacod et al. (2019) and Da and Xiu (2020), where noise is allowed to exhibit general serial dependence, and has random scales and could be dependent on the efficient prices, the observation times are random and could be endogenous. The proposed estimators under the general setup are quite useful—they can be directly applied to tick data while earlier methods usually need subsampling thereby discarding a substantial amount of data to satisfy the restrictive assumptions.

We adopt the same general framework as in Jacod et al. (2019) and Da and Xiu (2020). Our estimator is essentially a pre-averaging type estimator; thus, we share a lot of similarities with Jacod et al. (2019). Nevertheless, there are some key differences. First, we employ a different technology to estimate the moments of noise, which is essential to make bias correction to obtain consistent estimators of IV. Jacod et al. (2019) use the Local Averaging (LA) estimators (Jacod et al., 2017) while we employ the ReMeDI method (Li and Linton, 2021a). Both the LA and ReMeDI methods can estimate arbitrary moments of noise but the ReMeDI approach is more robust to model specifications, sampling frequencies, etc. The accuracy of noise estimation translates into an accurate estimation of IV, as we demonstrated in the extensive simulation studies. Second, we have a different estimator of the asymptotic variance. In addition to the use of different noise moments estimator, we also try to avoid a direct estimation of higher-order moments of noise to reduce estimation errors. Third, we have an explicit rule to select the optimal bandwidth of the pre-averaging estimators. Moreover, our estimator is quite efficient in practice—it typically takes less than 5% of the computational time

[3]The way ReMeDI works is explained with greater specificity below.

used by the Jacod et al. (2019) estimator. In terms of finite sample performance, our estimator is comparable to the Quasi-Maximum-Likelihood-Estimator (QMLE) (Da and Xiu, 2020). Both estimators give very accurate estimates of IVs under each specification, especially when noise is relatively small. The most significant advantage of our estimator is that it is computationally very efficient. It usually takes less than 0.5% computational time used by QMLE. 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.

The rest of the paper proceeds as follows. Section 2 discusses the model settings. Section 3 introduces the new IV estimator. Sections 4 and 5 present the simulation and empirical studies. Section 6 concludes the paper. Some additional simulation and empirical studies, as well as the mathematical proofs are in the Appendix.

2 Model Setting

We follow the general setup in Jacod et al. (2019) that allows for general Itô semimartingale efficient price, nonstationary and serially dependent microstructure noise and random observation scheme.

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:

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

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 $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$ and the predictable compensator of μ is $\nu(\mathrm{d}s, \mathrm{d}z) = \mathrm{d}s \otimes \lambda(\mathrm{d}z)$ 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 dependent 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*

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

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:

$$X_t := X_0 + \int_0^t b_s \mathrm{d}\mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s + (\delta \mathbf{1}_{\{|\delta|\leq 1\}}) \star (\mu - \nu)_t + (\delta \mathbf{1}_{\{|\delta|>1\}}) \star \mu_t. \tag{3}$$

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 $r = 2$.*

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

$$C_t := \int_0^t \sigma_s^2 \mathrm{d}s.$$

2.2 Random observation schemes

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 $T(n,0) = 0, T(n,i) \to \infty$ as $i \to \infty$, 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

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

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 $\Delta_n \to 0$ as $n \to \infty$. 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 $T(n,i-1)$. Specifically, 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) $\Delta_n N_t^n \to A_t := \int_0^t \alpha_s \mathrm{d}s \; \forall t > 0$.

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

(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 $i, n, m, \rho > 1/4$:

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

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,* see the discussion in Jacod et al. (2017).

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.

Definition 2.1. *Let $\{\chi_i\}_{i\in\mathbb{Z}}$ be a sequence of stationary random variables defined on a probability space $(\Omega^{(1)}, \mathcal{G}, \mathbb{P}^{(1)})$. 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 $k \in \mathbb{N}_+$:*

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

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

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*

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

where γ is a nonnegative Itô semimartingale satisfying Assumption (H-r) with $r = 2$. 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 $K > 0, v > 3$.

Remark 2.1. *The noise represented in (7) has a multiplicative form. The process γ captures the stochastic scale of noise. It is a continuous time process and dependent on the* calendar time. *This process also captures some endogeneity since γ can be a functions of the efficient price, e.g., γ could be dependent on 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.

3 When ReMeDI meets Pre-averaging

Now we introduce the new IV estimator. 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 $\int_0^1 g^2(s)\mathrm{d}s > 0$.[4] Let $\{h_n\}_{n\in\mathbb{N}^*}$ be a sequence of integers.

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

We introduce the following notations for the pre-averaging method.

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

For any processes V, let $\overline{V}_i^n := \sum_{j=1}^{h_n-1} g_j^n \Delta_{i+j}^n V = -\sum_{j=0}^{h_n-1} \overline{g}_j^n V_{i+j}^n$.

3.2 The ReMeDI method

We use the ReMeDI approach (Li and Linton, 2021a) to estimate the moments of noise, which will be subsequently subtracted off to get a consistent estimator of the integrated volatility. For any process V and its discretized observations $\{V_i^n\}_i$, integers i, k, ℓ, we introduce a multi-difference operator $\Delta_{i,\ell}^{n,k}$ such that $\Delta_{i,\ell}^{n,k} V := (V_{i+\ell}^n - V_{i+\ell+k}^n)(V_i^n - V_{i-k}^n)$.

Given three sequences of integers $\{d_n\}_n, \{k_n\}_n, \{\ell_n\}_n$ with $d_n \geq \ell_n + k_n$, we introduce

$$
r(V;\ell)_{i,d_n}^n := \frac{1}{d_n}\sum_{d=i+k_n+1}^{i+d_n+k_n} \Delta_{d,\ell}^{n,k_n} V; \quad R(V)_{i,d_n}^n := \textstyle\sum_{|\ell|\leq \ell_n} r(V;|\ell|)_{i,d_n}^n. \tag{8}
$$

In the sequel, we will show that $r(V;\ell)_{i,d_n}^n$ and $R(V)_{i,d_n}^n$ provide local estimators of the autocovariances and long-run variances of microstructure noise when $V = Y$, which will be used to correct the bias in the pre-averaging estimators.

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

We propose a pre-averaging method coupled with the ReMeDI bias correction as follows:

$$
\widehat{C}_t^n := \frac{1}{h_n \phi_0^n}\sum_{i=0}^{N_t^n - h_n} (\overline{Y}_i^n)^2 \mathbf{1}_{\{|\overline{Y}_i^n| \leq u_n\}} - \frac{1}{h_n^2 \phi_0^n}\sum_{i=k_n}^{N_t^n - h_n}\sum_{|\ell|\leq \ell_n} \overline{\phi}_\ell^n \Delta_{i,\ell}^{n,k_n} Y. \tag{9}
$$

Note that the tuning parameter u_n is introduced to truncate the jumps of the efficient price process (Mancini, 2001).

3.4 Large Sample Properties of PaReMeDI

We next present the main result of the paper regarding the large sample properties of our estimator and a feasible CLT, which requires a consistent estimator of the asymptotic variance.

We introduce the following processes and coefficients:

$$\begin{cases} V_t^{n,1} := \sum_{i=0}^{N_t^n-h_n} \left(\overline{Y}_i^n\right)^4 \mathbf{1}_{\{|\overline{Y}_i^n|<u_n\}}; \\ V_t^{n,2} := \frac{1}{h_n}\sum_{i=0}^{N_t^n-h_n} (\overline{Y}_i^n)^2 \mathbf{1}_{\{|\overline{Y}_i^n|<u_n\}} R(Y)_{i,h_n}^n; \\ V_t^{n,3} := \frac{1}{h_n^2}\sum_{i=0}^{N_t^n-k_n-\ell_n} (R(Y)_{i,h_n}^n)^2. \end{cases} \qquad \begin{cases} K_{g,1} := \frac{\Phi_{00}}{3}; \\ K_{g,2} := \left(\Phi_{01}\phi(0) - \Phi_{00}\overline{\phi}(0)\right); \\ K_{g,3} := \Phi_{11}\phi^2(0) - 2\Phi_{01}\overline{\phi}(0)\phi(0) + \Phi_{00}\overline{\phi}^2(0). \end{cases}$$

We use the following estimator for the asymptotic variance:

$$\Sigma_t^n := \frac{4}{h_n\phi^4(0)}\left(K_{g,1}V_t^{n,1} + 2K_{g,2}V_t^{n,2} + K_{g,3}V_t^{n,3}\right).$$

To implement the new estimation method we need several tuning parameters: h_n controls the pre-averaging bandwidth; k_n is the tuning parameter of the ReMeDI approach to estimate the moments of microstructure noise; ℓ_n is the truncation parameter to estimate the long run variance of noise; and u_n truncates jumps of the efficient price. We require the tuning parameters to satisfy the following asymptotic conditions for the theory to hold:[5]

$$\begin{cases} h_n = \lfloor \frac{\theta}{\sqrt{\Delta_n}} \rfloor; \quad k_n \asymp \Delta_n^{-\frac{2}{11}}; \quad \ell_n \asymp \Delta_n^{-\frac{1}{7}}; \\ u_n \asymp \Delta_n^{\varpi/2}, \text{ where } \frac{1}{4(2-r)} < \varpi < \frac{2[v]-3}{8([v]-1)}, \; r < \frac{2[v]-4}{2[v]-3}. \end{cases} \tag{10}$$

Theorem 3.1. *Under the Assumptions (H-X-r), (O-ρ, ρ'), (N-v) and the asymptotic conditions (10), we have for any $t > 0$ that the sequence $(\widehat{C}_t^n - C_t)/\Delta_n^{1/4}$ converges $\mathcal{F}_\infty$-stably in law to a limiting variable*

$$\mathcal{U}_t := \int_0^t \beta_s \mathrm{d}B_s \tag{11}$$

defined on an extension of the original space, where B is a standard Brownian motion that is independent of $\mathcal{F}$ and

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

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

3.5 On the Implementation of PaReMeDI

To implement PaReMeDI in practice, we need to select several tuning parameters, recall (10). The key parameter that affects the performance of PaReMeDI is the pre-averaging bandwidth h_n—it not only effects the finite-sample performance but also the limiting distribution of PaReMeDI via θ. In fact, we can find the *optimal* θ as the value that minimizes the asymptotic

[5]The parameter $r \in [0,2]$ controls the bound on the degree of activity of jumps, see Assumption (H-r) for details.

variance $\Sigma_t := \int_0^t \beta_s^2 ds$ of our estimator. The optimal θ is given explicitly by

$$\theta^* := \left(\frac{\Phi_{01} R \int_0^t \sigma_s^2 \gamma_s^2 ds + \sqrt{\Phi_{01}^2 R^2 (\int_0^t \sigma_s^2 \gamma_s^2 ds)^2 + 3\Phi_{00}\Phi_{11} R^2 \int_0^t \sigma_s^4/\alpha_s ds \int_0^t \gamma_s^4 \alpha_s ds}}{\Phi_{00} \int_0^t \sigma_s^4/\alpha_s ds} \right)^{\frac{1}{2}}. \quad (13)$$

Note that in our general settings of the efficient price, the microstructure noise and the observation scheme, θ^* will be random. Thus, the implementation using θ^* requires an estimate of θ^*. For a given θ_0, we let

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

Then a consistent estimator of the optimal θ^* is given by

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

To implement the optimal choice of h_n (or θ) established in (13), one needs to estimate many parameters. This slows down the estimation procedure and brings in more variability, and may fail to provide a proper guidance to select h_n in a finite sample.

However, some insights can be obtained if we consider a setting where the noise is stationary, volatility is constant, and the observation scheme is regular. That is, $\gamma_s \equiv K_\gamma, \sigma_s \equiv K_\sigma, \alpha_s \equiv 1$, for all s. In this special case, the optimal choice of θ becomes $\theta'^* := K_\Phi K_\gamma \sqrt{R}/K_\sigma$, where $K_\Phi := \sqrt{\left(\Phi_{01} + \sqrt{\Phi_{01}^2 + 3\Phi_{00}\Phi_{11}}\right)/\Phi_{00}}$ 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 larger (or small) θ if noise is relatively large (or small).

Now let $r(V;\ell)_t^n := \frac{1}{N_t^n} \sum_{i=k_n}^{N_t^n - \ell - k_n} \Delta_{i,\ell}^{n,k_n} Y$; $R(Y)_t^n := \sum_{\ell:|\ell| \le \ell_n} r(V;\ell)_t^n$. Intuitively, we have the convergence $R(Y)_t^n \xrightarrow{\mathbb{P}} K_\gamma^2 R$. Hence we have the following proxy of θ'^*:

$$\widehat{\theta'}^* = \frac{K_\Phi \sqrt{R(Y)_t^n}}{\sqrt{\widehat{C}_t^n}}, \quad (14)$$

which provides a guidance to select θ whence h_n in practice.

4 Simulation Studies

In this section, we adopt the simulation approach to explore the finite sample performance of the PaReMeDI estimator of IV. The overall goal is to compare PaReMeDI with the IV estimators proposed by Da and Xiu (2020) and Jacod et al. (2019), where the theoretical settings are similar to ours.

4.1 Model settings

The efficient price is allowed to have stochastic volatility with jumps in both the price level and the volatility process:

$$\begin{aligned} &\mathrm{d}X_t = \kappa_1(\mu_1 - X_t)\mathrm{d}t + \sigma_t \mathrm{d}W_{1,t} + \xi_{1,t}\mathrm{d}N_t, \quad \mathrm{d}\sigma_t^2 = \kappa_2(\mu_2 - \sigma_t^2)\mathrm{d}t + \eta\sigma_t \mathrm{d}W_{2,t} + \xi_{2,t}\mathrm{d}N_t; \\ &\mathbf{Corr}(W_1, W_2) = v, \quad \xi_{1,t} \sim \mathcal{N}(0, \mu_2/10), \quad N_t \sim \mathrm{Poi}(\lambda); \quad \xi_{2,t} \sim \mathrm{Exp}(\delta), \end{aligned} \tag{15}$$

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). In this section, we set $\xi_{1,t} \equiv 0$,[6] the Appendix A contains further simulation studies where the efficient price exhibits jumps.

The stationary component of the microstructure noise follows an AR(1) process with Gaussian innovations

$$\chi_{i+1} = \varrho\chi_i + e_i, \quad e_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}\left(0, 1 - \varrho^2\right), \quad |\varrho| < 1. \tag{16}$$

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

$$\alpha_t = (1 + \cos(2\pi t))/2; \quad \gamma_t = K_\gamma \gamma_t', \quad \mathrm{d}\gamma_t' = -\rho_\gamma(\gamma_t' - \mu_t)\mathrm{d}t + \sigma_\gamma \mathrm{d}W_t. \tag{17}$$

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 $n = 23,400$, $\rho_\gamma = 10$, $\mu_t = 1 + 0.1\cos(2\pi t)$, $\sigma_\gamma = 0.1$. 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.

4.2 PaReMeDI, PaLA and QMLE

In this section, we compare the performances of three estimators developed in the same setting: the pre-averaging estimator with ReMeDI bias correction (PaReMeDI), the pre-averaging estimator with Local Averaging (LA) bias correction (PaLA) (Jacod et al., 2019), and the quasi-maximum-likelihood estimator (QMLE) (Da and Xiu, 2020).

Table 1 reports the performances of the three estimators under various specifications for the noise dynamics and noise-to-signal ratios. Since the QMLE method proposed in Da and Xiu (2020) does not separate the jumps component from the quadratic variation, we eliminate jumps from the efficient price, so the focus is the IV for all three estimators.

[6] We would like to compare our volatility estimator with QMLE (Da and Xiu, 2020), which does not explicitly truncate jumps. Therefore, we neglect jumps from the efficient price to present a fair comparison.

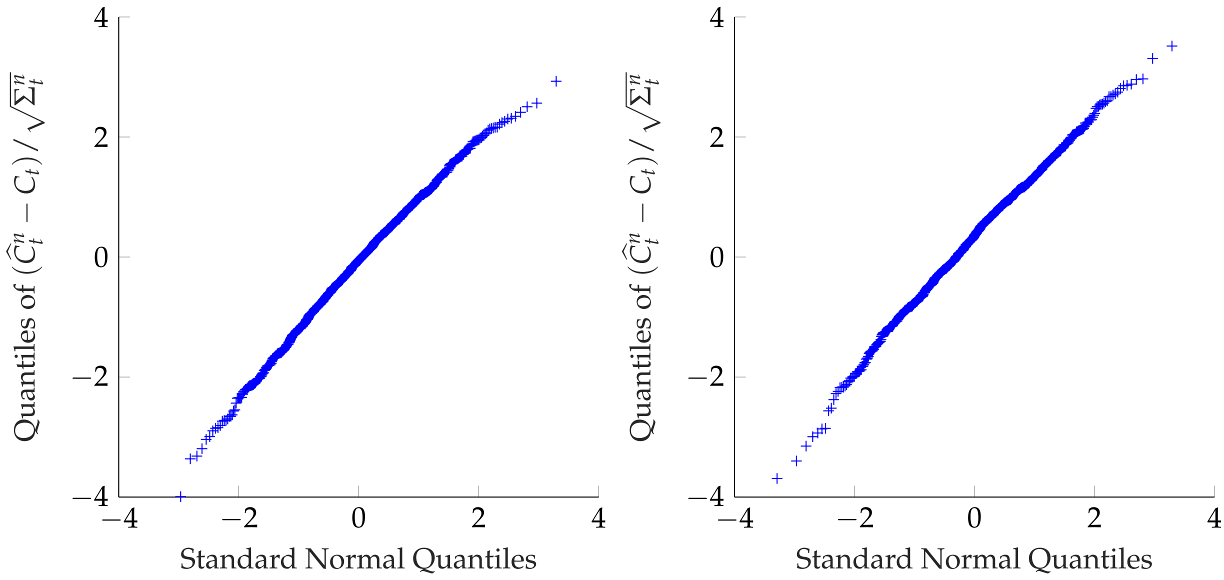

Figure 1: QQ Plot of $(\widehat{C}_t^n - C_t) / \sqrt{\Sigma_t^n}$ versus Standard Normal. The model is specified by (15), (16) and (17). We follow the numerical settings of the parameters as in Section 4.1. The scale of noise is fixed at $K\gamma = 10^{-4}$, the AR(1) coefficient of noise $\varrho = 0.3$ in the left panel and $\varrho = 0.8$ for the right panel. The tuning parameters are set as follows: $k_n = 8, \ell_n = 10, \theta = 0.3$.

Da and Xiu (2020) claim that the QMLE works well for a wide range of noise-to-signal ratios. This is confirmed by our simulation results—the estimator has small bias and standard deviation in all specifications when the variance of noise varies from 10^{-6} to 10^{-10}. Our PaReMeDI estimates are in line with the QMLE. The striking difference is the computational efficiency—the PaReMeDI typically uses less than 0.5% of the computational time used by QMLE!

The PaLA estimator has a smaller standard deviation than both the PaReMeDI and QMLE. The finite sample bias of the LA method when estimating the moments of noise is partially responsible for this reduction in standard deviation. Jacod et al. (2017) show that the LA method elicits a finite sample bias that is a fraction of the IV. Thus, when the estimated second moments of noise is subtracted off from the pre-averaging statistics (recall (9)), the IV estimates become smaller by a fraction that is equal to the finite sample bias of the LA estimates. However, the bias of LA is transmitted to a bias of the PaLA, and the bias is substantial that it dominates the RMSE of PaLA in most scenarios. A large bias and small standard deviation intend to conclude that the PaLA estimates are statistically very different from other estimates or thresholds. It may, therefore lead to unreliable statistical inferences.

4.3 Examine the feasible CLT

Figure 1 presents the QQ-plot for the standardized statistics $(\widehat{C}_t^n - C_t) / \sqrt{\Sigma_t^n}$. It provides the numerical evidence supporting the limit distribution in Theorem 3.1.

	PaReMeDI	PaLA	QMLE	specifications
Bias	0.2460	0.7060	0.2139	
std	2.3116	1.2823	2.2097	$\varrho = 0.3$
RMSE	2.3247	1.4638	2.2200	$K_\gamma = 10.0$
cpu time%	0.2924	7.8355	91.8721	
Bias	0.2549	2.7429	0.1782	
std	2.6632	1.9350	2.4456	$\varrho = 0.3$
RMSE	2.6754	3.3567	2.4521	$K_\gamma = 5.0$
cpu time%	0.3750	10.1660	89.4590	
Bias	0.1227	6.4554	0.1195	
std	2.2076	1.8907	2.2160	$\varrho = 0.3$
RMSE	2.2111	6.7266	2.2192	$K_\gamma = 0.5$
cpu time%	0.3821	10.4779	89.1400	
Bias	0.1019	6.4325	0.1192	
std	2.1920	1.9331	2.1916	$\varrho = 0.3$
RMSE	2.1943	6.7167	2.1948	$K_\gamma = 0.1$
cpu time%	0.3888	10.6673	88.9439	
Bias	2.5590	2.4513	1.3149	
std	2.9920	2.5645	2.7767	$\varrho = 0.8$
RMSE	3.9371	3.5476	3.0722	$K_\gamma = 10.0$
cpu time%	0.3164	8.3389	91.3447	
Bias	1.3020	1.1851	0.8126	
std	2.6618	0.8520	2.5899	$\varrho = 0.8$
RMSE	2.9631	1.4596	2.7143	$K_\gamma = 5.0$
cpu time%	0.3742	9.9675	89.6583	
Bias	0.1348	6.4036	0.1370	
std	2.1615	1.8087	2.1653	$\varrho = 0.8$
RMSE	2.1657	6.6541	2.1696	$K_\gamma = 0.5$
cpu time%	0.3936	10.7188	88.8877	
Bias	0.1102	6.5184	0.1283	
std	2.1003	1.8207	2.1024	$\varrho = 0.8$
RMSE	2.1032	6.7679	2.1063	$K_\gamma = 0.1$
cpu time%	0.3935	10.7285	88.8780	

Table 1: Volatility estimation by the Pre-averaging-ReMeDI (PaReMeDI) method, the Pre-averaging-Local-Averaging (PaLA) method, and the Quasi-Maximum-Likelihood-Estimator (QMLE). All numeric results are multiplied by 10^4. The models of the efficient price, microstructure noise and the observation scheme are specified in Section 4.1. The AR(1) coefficient $\varrho \in \{0.3, 0.8\}$ and the noise scale $K\gamma \in \{10^{-3}, 5\times10^{-4}, 5\times10^{-5}, 10^{-5}\}$. The number of simulation is 1000. The tuning parameters of the PaReMeDI are set as follows: $k_n = 10, \ell_n = \lfloor N_t^{n\,\frac{1}{7}} \rfloor$, and for PaLA $k_n = 6, k'_n = \lfloor N_t^{n\,\frac{1}{8}} \rfloor$. θ are selected according to (14) for both methods. For the QMLE (with AIC selection criterion), the optimal q is selected within $\{5, 6, 7, 8, 9, 10\}$.

5 Empirical Studies

This section presents a simple empirical study. We use intraday transaction prices for the S&P 500 index ETF (ticker: SPY) obtained from the Trade and Quote (TAQ) database for January 2015. There are 20 trading days. The index is quite liquid, with 20.1 transactions per second on average. The average returns is -3.6×10^{-9} with a standard deviation 1.06×10^{-4}.

The top panel of Figure 2 reports the estimates of the three estimators based on transaction prices for each trading day. The PaReMeDI and QMLE generate virtually identical estimates of the IV. All estimates are within the 95% confidence intervals of PaReMeDI, whence they are statistically indistinguishable. The PaLA estimator has similar " volatility trend" but gives smaller estimates. This is in line with our earlier analysis that PaLA only estimates a fraction of the IV due to the presence of a finite sample bias of the LA estimator. One exception occurs on the 5th trading day, where the three estimates give almost the same IV estimate. The bottom panel of Figure 2 provides an explanation, where the estimates[7] of the noise-to-signal ratio $R\int_0^t \gamma_s^2 \alpha_s \mathrm{d}s / C_t$ are plotted. The noise scale is very large compared to the IV on the 5th trading day.[8] Thus the finite sample bias (induced by IV) of LA becomes less significant; consequently, the LA estimates of the second moments of noise become more accurate and the PaLA also benefits from the smaller finite sample bias. In the Appendix B, we repeat the estimation on two individual stocks and we find similar results.

As in the simulation studies, the PaReMeDI approach is much faster than PaLA and QMLE. To see the computational time in "real terms", we report that it takes 9.1, 215.3 and 762.8 seconds for PaReMeDI, PaLA and QMLE to perform the 20 estimations of IV.[9]

Therefore, we conclude that PaReMeDI and QMLE have similar and reliable estimates of integrated volatility. The performance of PaLA, however, depends on the characteristics of the underlying noise and efficient price processes, in particular the noise-to-signal ratio. Moreover, the computational efficiency of PaReMeDI is very significant. This is a great advantage to deal with high-frequency datasets with enormous sample sizes.

6 Conclusion

We introduce a new integrated volatility estimator using noisy high-frequency data. The estimator is applicable in a broad setting of microstructure noise and observation schemes. Compared to alternative estimators recently proposed, our estimator provides accurate and robust estimates, and is computationally efficient. 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

[7] The estimator is given by $\left(\mathrm{ReMeDI}(Y;0)_t^n + 2\sum_{\ell=1}^{\ell_n} \mathrm{ReMeDI}(Y;\ell)_t^n\right) / \widehat{C}_t^n$, where $\mathrm{ReMeDI}(Y;\ell)_t^n := \frac{1}{N_t^n} \sum_{i=k_n}^{N_t^n - k_n - \ell} \Delta_{i,\ell}^{n,k_n} Y$.

[8] This is caused by several extreme outliers occurred on that day.

[9] All our numerical results are obtained by running Matlab (version 9.7) on a MacBook Pro (6-core, Intel Core i7 @ 2.60GHz, 16 Gigabytes of RAM).

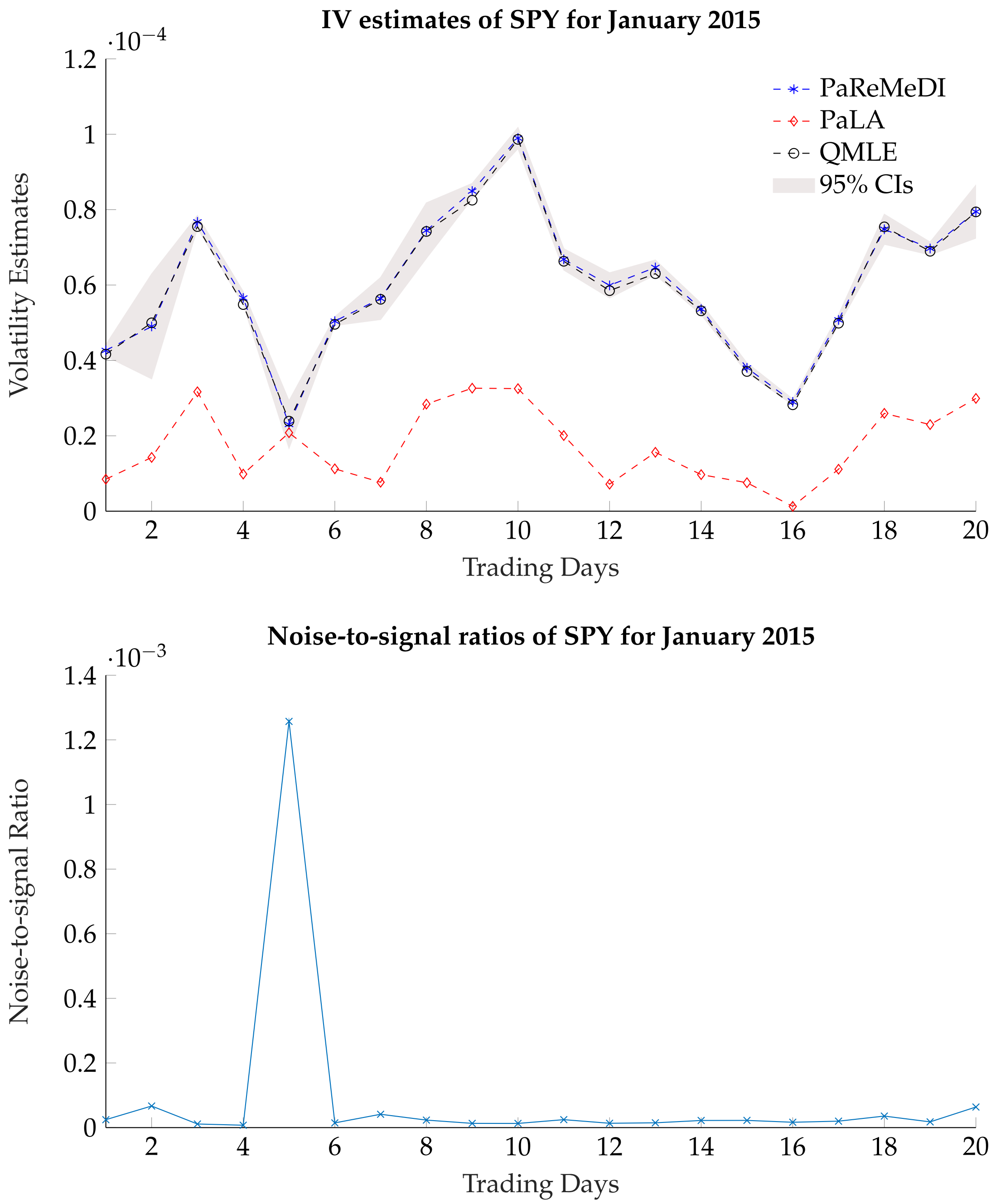

Figure 2: Volatility estimates and Noise-to-signal ratios of S&P 500 index ETF (ticker: SPY) for January, 2015. The tuning parameters are set as follows: For PaReMeDI, $k_n = 10, \ell_n = \lfloor N_t^{n\,\frac{1}{7}} \rfloor$, and for PaLA $k_n = 6, k_n' = \lfloor N_t^{n\,\frac{1}{8}} \rfloor$; θ are selected according to (14) for both methods. For the QMLE (with AIC selection criterion), the optimal q is selected within $\{5, 6, 7, 8, 9, 10\}$.

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.

Appendix A Additional Simulation Studies

We adopt the same simulation settings in Section 4 except that we allow for jumps in the price level. Table 2 reports the estimates by PaLA and PaReMeDI with the same truncation level for jumps. Similar to the simulation results in Section 4, PaLA has quite significant bias hence large RMSE in all scenarios. While PaReMeDI has small RMSE, it is uses much less computing time.

	Bias	std	RMSE	cpu time%	specifications
PaReMeDI	1.0831	0.8607	1.3834	3.2909	$\varrho = 0.3$
PaLA	1.8607	0.3579	1.8948	96.7091	$K_\gamma = 10.0$
PaReMeDI	0.3715	1.7255	1.7651	3.2192	$\varrho = 0.3$
PaLA	4.2463	3.1601	5.2931	96.7808	$K_\gamma = 5.0$
PaReMeDI	0.1483	2.1415	2.1466	2.8180	$\varrho = 0.3$
PaLA	13.9810	6.4837	15.4113	97.1820	$K_\gamma = 0.5$
PaReMeDI	0.1275	2.0717	2.0757	2.7982	$\varrho = 0.3$
PaLA	13.7745	6.3441	15.1652	97.2018	$K_\gamma = 0.1$
PaReMeDI	1.4312	0.3243	1.4675	3.2531	$\varrho = 0.8$
PaLA	1.6613	0.2140	1.6750	96.7469	$K_\gamma = 10.0$
PaReMeDI	0.7892	1.2894	1.5117	3.1587	$\varrho = 0.8$
PaLA	2.0356	1.1158	2.3214	96.8413	$K_\gamma = 5.0$
PaReMeDI	0.1583	1.9737	1.9801	2.9857	$\varrho = 0.8$
PaLA	13.7957	6.0469	15.0628	97.0143	$K_\gamma = 0.5$
PaReMeDI	0.1366	2.2146	2.2188	2.8479	$\varrho = 0.8$
PaLA	14.0045	6.9747	15.6452	97.1521	$K_\gamma = 0.1$

Table 2: Volatility estimation by the Pre-averaging-ReMeDI (PaReMeDI) method and the Pre-averaging-Local-Averaging (PaLA). All numeric results are multiplied by 10^4. The models of the efficient price, microstructure noise and the observation scheme are specified in Section 4.1. The AR(1) coefficient $\varrho \in \{0.3, 0.8\}$ and the noise scale $K\gamma \in \{10^{-3}, 5\times 10^{-4}, 5\times 10^{-5}, 10^{-5}\}$. The number of simulation is 1000. The tuning parameters of the PaReMeDI are set as follows: $k_n = 10, \ell_n = \lfloor N_t^{n\frac{1}{7}} \rfloor$, and for PaLA $k_n = 6, k_n' = \lfloor N_t^{n\frac{1}{8}} \rfloor$. θ are selected according to (14) for both methods and the jump truncation level for both methods are set $u_n = 5\mu_2$.

Appendix B Additional Empirical Studies

We estimate the integrated volatility of two individual stocks General Electric (ticker: GE) and Intel Corporation (ticker: INTC) for January, 2015. The estimates are reported in Figure 3. We observe similar patterns as in Section 5, where we use the transaction prices for the S&P 500 index ETF. PaReMeDI and QMLE yield almost identical estimates while PaLA departs from both estimates by a large margin. However, the differences are smaller when the noise-to-signal ratios are large. This is consistent with our previous analysis in Section 5 that PaLA becomes more accurate when noise scale is large.

Appendix C 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, we will first show in Lemma C.3 that the ReMeDI estimators correct the asymptotic bias that appears in the pre-averaged statistics. Next, aided by several lemmas, we will show that after proper scaling, the proposed asymptotic variance estimator $\widehat{\Sigma}_t^n$ provides a consistent estimator of the asymptotic variance of $\widehat{C}_t^n$. 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.

Assumption (S-HON). *Assume Assumptions (H-X-r), (O-ρ, ρ') (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*

$$N_t^n \leq Kt\Delta_n^{-1}, \ \left|\mathbb{E}\left(\Delta(n,i) - \Delta_n \alpha_{T(n,i-1)}^{-1} \,\middle|\, \mathcal{F}_{T(n,i-1)}\right)\right| \leq K\Delta_n^{1+\rho}, \ \mathbb{E}\left(\Delta(n,i)^\kappa \,\middle|\, \mathcal{F}_{T(n,i-1)}\right) \leq K\Delta_n^\kappa. \quad (18)$$

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

$$X_t^{\text{c}} := X_0 + \int_0^t b_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s, \quad X_c^{\text{d}} := \int_{[0,t]\times E} \delta(s,z)\mu(s,z). \quad (19)$$

In the sequel, $p > 1$ will be an integer. We introduce the following notations

$$\begin{aligned}
&\mathcal{H}_i^n := \mathcal{F}_\infty \otimes \mathcal{G}_{i-h_n}, \quad \mathcal{K}_i^n := \mathcal{F}_i^n \otimes \mathcal{G}_{i-h_n}, \quad \mathcal{H}(p)_j^n := \mathcal{K}_{j(p+2)h_n}^n, \\
&\mathbf{j} := (j_1, \ldots, j_d),\ j_\ell \in \mathbb{N}, 1 \leq \ell \leq d, \quad \overline{\boldsymbol{r}}(\mathbf{j})_n := \mathbb{E}\left(\prod_{\ell=1}^d \overline{\chi}_\ell^n\right), \\
&\mathcal{H}'(p)_j^n := \mathcal{K}_{(j(p+2)+p)h_n}^n, \quad \widehat{c}_i^n := \sum_{j\in\mathbb{Z}} (g_j^n)^2 \Delta_{i+j}^n C, \quad \widehat{X}_i^{\text{c},n} := \left(\overline{X}_i^{\text{c},n}\right)^2 - \widehat{c}_i^n, \\
&\widehat{\varepsilon}_i^n := (\overline{\varepsilon}_i^n)^2 - (\gamma_i^n)^2 \overline{\boldsymbol{r}}(0,0)_n, \quad \widehat{X^{\text{c}}\varepsilon}_i^n := \overline{X}_i^{c,n} \widehat{\varepsilon}_i^n, \\
&Z_i^n := (\overline{Y}_i^{\text{c},n})^2 - \widehat{c}_i^n - (\gamma_i^n)^2 \overline{\boldsymbol{r}}(0,0)_n = \widehat{X}_i^{\text{c},n} + \widehat{\varepsilon}_i^n + 2\widehat{X^{\text{c}}\varepsilon}_i^n, \quad \zeta(p)_i^n := \sum_{j=i}^{i+ph_n-1} Z_j^n, \\
&\eta(p)_j^n := \frac{1}{h_n \phi_0^n} \zeta(p)_{(j-1)(p+2)h_n}^n, \quad \overline{\eta}(p)_j^n := \mathbb{E}\left(\eta(p)_j^n \,\middle|\, \mathcal{H}(p)_{j-1}^n\right), \\
&\eta'(p)_j^n := \frac{1}{h_n \phi_0^n} \zeta(2)_{(j-1)(p+2)h_n + ph_n}^n, \quad \overline{\eta}'(p)_j^n := \mathbb{E}\left(\eta'(p)_j^n \,\middle|\, \mathcal{H}'(p)_{j-1}^n\right).
\end{aligned}$$

Next, we define several processes:

$$F(p)_t^n := \sum_{j=1}^{J_n(p,t)} \overline{\eta}(p)_j^n, \quad M(p)_t^n := \sum_{j=1}^{J_n(p,t)} \left(\eta(p)_j^n - \overline{\eta}(p)_j^n\right),$$
$$F'(p)_t^n := \sum_{j=1}^{J_n(p,t)} \overline{\eta}'(p)_j^n, \quad M'(p)_t^n := \sum_{j=1}^{J_n(p,t)} \left(\eta'(p)_j^n - \overline{\eta}'(p)_j^n\right),$$
$$\widehat{C}(p)_t^n := \frac{1}{h_n\phi_0^n} \sum_{i=N_t^n-h_n+2}^{\ell_n(p,t)} Z_i^n, \quad \widehat{C}_t^{n,1} := \frac{1}{h_n\phi_0^n} \sum_{i=0}^{N_t^n-h_n} \widehat{c}_i^n - C_t,$$
$$\widehat{C}_t^{n,2} := \frac{1}{h_n\phi_0^n} \sum_{i=0}^{N_t^n-h_n} \left((\overline{Y}_i^n)^2 \mathbf{1}_{\{|\overline{Y}_i^n| \le u_n\}} - (\overline{Y}_i^{c,n})^2\right),$$
$$\widehat{C}_t^{n,3} := \frac{\overline{\boldsymbol{r}}(0,0)_n}{h_n\phi_0^n} \sum_{i=0}^{N_t^n-h_n} (\gamma_i^n)^2 - \frac{1}{h_n^2\phi_0^n} \sum_{|\ell|\le\ell_n} \sum_{i=k_n}^{N_t^n-h_n} \overline{\phi}_\ell^n \Delta_{i,\ell}^{n,k_n} Y.$$

For all $p > 1$, we have

$$\widehat{C}_t^n - C_t = M(p)_t^n + M'(p)_t^n + F(p)_t^n + F'(p)_t^n - \widehat{C}(p)_t^n + \sum_{j=1}^{3} \widehat{C}_t^{n,j}. \tag{20}$$

Lemma C.1. *Under Assumption (N-v), we have*

$$\sqrt{h_n}\overline{\chi}_i^n \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \overline{\phi}(0)R\right). \tag{21}$$

Moreover, we have

$$h_n\overline{\boldsymbol{r}}(0,0)_n \to \overline{\phi}(0)R; \quad h_n^{3/2}\overline{\boldsymbol{r}}(0,0,0)_n \to 0; \quad h_n^2\overline{\boldsymbol{r}}(0,0,0,0)_n \to 3\overline{\phi}^2(0)R^2. \tag{22}$$

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

$$\sum_{k\in\mathbb{N}^*} k^{\frac{2}{r-2}}\rho_k < \infty, \tag{23}$$

where $\{\rho_k\}$ are the ρ-mixing coefficients introduced in Definition 2.1, and r is a positive real such that $\mathbb{E}(|\chi_i|^r) < \infty$. Assumption (N-$v$) implies r can be arbitrarily large, thus (23) holds since $v > 3$. The rest of the proof of (21) follows from the proof of Lemma A.1 in Li et al. (2020). (22) follows from the condition $\mathbb{E}(|\chi_i|^r) < \infty, r > 4$ and the limit distribution (21); 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 C.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 $|j-i| \le d_n, |\ell| \le \ell_n$, we have for any $q \ge 2$,*

$$\mathbb{E}\left(\left|\Delta_{i,\ell}^{n,k_n}Y - (\gamma_j^n)^2\Delta_{i,\ell}^{n,k_n}\chi\right|^q\right) \le K\left((d_n+\ell_n+k_n)\Delta_n + ((d_n+\ell_n+k_n)\Delta_n)^{\frac{1}{z}}\right). \tag{24}$$

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

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

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

Lemma C.3. *Assume $v > 3$, under the asymptotic conditions (10), for all t we have $\Delta_n^{-\frac{1}{4}} \widehat{C}_t^{n,3} \xrightarrow{\mathbb{P}} 0$.*

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

$$|\bar{\boldsymbol{r}}(0,0)_n - \bar{\boldsymbol{r}}(0,0)_{\ell_n}| \leq \frac{K}{h_n \ell_n^{v-1}} \Rightarrow h_n\, |\bar{\boldsymbol{r}}(0,0)_n - \bar{\boldsymbol{r}}(0,0)_{\ell_n}| = o(\Delta_n^{1/4}), \tag{25}$$

since $\ell_n \asymp \Delta_n^{-1/7}$ and $v > 3$. Moreover,

$$\left| \frac{\bar{\boldsymbol{r}}(0,0)_n}{h_n \phi_0^n} \sum_{i=0}^{k_n - 1} (\gamma_i^n)^2 \right| \leq K k_n / (h_n)^2 = o(\Delta_n^{1/4}), \tag{26}$$

(25) and (26) imply that it suffices to prove

$$\frac{\bar{\boldsymbol{r}}(0,0)_{\ell_n}}{h_n \phi_0^n} \sum_{i=k_n}^{N_t^n - h_n} (\gamma_i^n)^2 - \frac{1}{h_n^2 \phi_0^n} \sum_{|\ell| \leq \ell_n} \sum_{i=k_n}^{N_t^n - h_n} \overline{\phi}_\ell^n \Delta_{i,\ell}^{n,k_n} Y = o_p(\Delta_n^{1/4}).$$

To get this, we note

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

The first estimate follows from Lemma C.2 and the Cauchy-Schwarz inequality, and the second one follows from the boundedness of γ, the Cauchy-Schwarz inequality and the limit distribution established in Theorem C.1 of Li and Linton (2021b) (hereafter Theorem LL-C.1) such that

$$\mathbb{E}\left(\left(\sum_{i=k_n}^{N_t^n - h_n} \left((r(\ell) - \Delta_{i,\ell}^{n,k_n} \chi) \right) / h_n^2 \right)^2 \right) \leq K \Delta_n.$$

Note that our choice of k_n is to satisfy the asymptotic conditions in that theorem. For z close to 1, we see that both bounds are of order $o(\Delta_n^{1/4})$. This finishes the proof. ∎

Now we prove

$$V_t^{n,1} \xrightarrow{\mathbb{P}} 3\theta^2\phi^2(0)\int_0^t \frac{\sigma_s^4}{\alpha_s}\mathrm{d}s + 6\phi(0)\overline{\phi}(0)R\int_0^t \sigma_s^2\gamma_s^2\mathrm{d}s + \frac{3\overline{\phi}^2(0)R^2}{\theta^2}\int_0^t \gamma_s^4\mathrm{d}A_s; \tag{27}$$

$$V_t^{n,2} \xrightarrow{\mathbb{P}} \phi(0)R\int_0^t \sigma_s^2\gamma_s^2\mathrm{d}s + \frac{\overline{\phi}(0)R^2}{\theta^2}\int_0^t \gamma_s^4\mathrm{d}A_s; \tag{28}$$

$$V_t^{n,3} \xrightarrow{\mathbb{P}} \frac{R^2}{\theta^2}\int_0^t \gamma_s^4\mathrm{d}A_s. \tag{29}$$

Proof of (27). Now let $G_s^{n,i} := \sum_{j=1}^{h_n-1} g_j^n \mathbf{1}_{\{(T(n,i+j-1),T(n,i+j)]\}}(s)$, and denote

$$\widehat{G}_{j,j'}^{n,i} := \int_0^\infty G_u^{n,i+j}G_u^{n,i+j'}\mathrm{d}u, \quad \overline{G}_{j,j'}^{n,i} := \int_0^\infty G_u^{n,i+j}G_u^{n,i+j'}\mathrm{d}u\int_0^u G_s^{n,i+j}G_s^{n,i+j'}\mathrm{d}s.$$

We have $\left(\overline{Y}_i^{c,n}\right)^4 = \sum_{k=1}^{10}\delta_i^{n,k}$, where

$$\begin{aligned}
&\delta_i^{n,1} = \left(\overline{X}_i^{c,n}\right)^4 - 6\left(\sigma_i^n\right)^4\overline{G}_{0,0}^{n,i}; \quad \delta_i^{n,2} = 4\overline{X}_i^{c,n}\left((\overline{\varepsilon}_i^n)^3 - (\gamma_i^n)^3\overline{\boldsymbol{r}}(0,0,0)_n\right);\\
&\delta_i^{n,3} = 4\overline{X}_i^{c,n}\left(\gamma_i^n\right)^3\overline{\boldsymbol{r}}(0,0,0)_n; \quad \delta_i^{n,4} = 6\left(\overline{X}_i^{c,n}\right)^2\left((\overline{\varepsilon}_i^n)^2 - (\gamma_i^n)^2\overline{\boldsymbol{r}}(0,0)_n\right);\\
&\delta_i^{n,5} = 4\left(\overline{X}_i^{c,n}\right)^3\overline{\varepsilon}_i^n; \quad \delta_i^{n,6} = 6\left(\left(\overline{X}_i^{c,n}\right)^2 - \widehat{G}_{0,0}^{n,i}(\sigma_i^n)^2\right)(\gamma_i^n)^2\overline{\boldsymbol{r}}(0,0)_n;\\
&\delta_i^{n,7} = (\overline{\varepsilon}_i^n)^4 - (\gamma_i^n)^4\overline{\boldsymbol{r}}(0,0,0,0)_n; \quad \delta_i^{n,8} = 6\left(\sigma_i^n\right)^4\overline{G}_{0,0}^{n,i};\\
&\delta_i^{n,9} = 6\widehat{G}_{0,0}^{n,i}\left(\sigma_i^n\right)^2\left(\gamma_i^n\right)^2\overline{\boldsymbol{r}}(0,0)_n; \quad \delta_i^{n,10} = (\gamma_i^n)^4\overline{\boldsymbol{r}}(0,0,0,0)_n.
\end{aligned}$$

By (JLZ-A.22), we have $\mathbb{E}\left(\left|\mathbb{E}\left(\delta_i^{n,1}\,|\mathcal{K}_i^n\right)\right|\right) \le K\Delta_n^{\frac{5}{4}}$; apply (JLZ-A.25) and (JLZ-A.17), we have $\mathbb{E}\left(\left|\mathbb{E}\left(\delta_i^{n,2}\,|\mathcal{K}_i^n\right)\right|\right) \le K\Delta_n^{\frac{3}{2}}$, $\mathbb{E}\left(\left|\mathbb{E}\left(\delta_i^{n,4}\,|\mathcal{K}_i^n\right)\right|\right) \le K\Delta_n^{\frac{5}{4}}$; by the boundedness of γ, the facts that $|\overline{\boldsymbol{r}}(0,0,0)_n| \le K\Delta_n^{\frac{3}{4}}$, $|\overline{\boldsymbol{r}}(0,0)_n| \le K\Delta_n^{\frac{1}{2}}$ (apply (22)), the independence of $\mathcal{G},\mathcal{F}_\infty$ and (JLZ-A.6), we have $\mathbb{E}\left(\left|\mathbb{E}\left(\delta_i^{n,3}\,|\mathcal{K}_i^n\right)\right|\right) \le K\Delta_n^{\frac{5}{4}}$; (JLZ-A.26) implies $\mathbb{E}\left(\left|\mathbb{E}\left(\delta_i^{n,5}\,|\mathcal{K}_i^n\right)\right|\right) \le K\Delta_n^{\frac{2v+3}{4}}$; (JLZ-A.19) and (JLZ-A.21), the independence of $\mathcal{G},\mathcal{F}_\infty$ and Lemma C.1 imply $\mathbb{E}\left(\left|\mathbb{E}\left(\delta_i^{n,6}\,|\mathcal{K}_i^n\right)\right|\right) \le K\Delta_n^{\frac{3}{2}}$; (JLZ-A.25) implies $\mathbb{E}\left(\left|\mathbb{E}\left(\delta_i^{n,7}\,|\mathcal{K}_i^n\right)\right|\right) \le K\Delta_n^{\frac{5}{4}}$. Thus, we have

$$\sum_{i=0}^{N_t^n-h_n}\mathbb{E}\left(\delta_i^{n,k}\,|\mathcal{K}_i^n\right) \xrightarrow{\mathbb{P}} 0, \text{ for } k=1,\ldots,7. \tag{30}$$

Next, we have the following:

$$\mathbb{E}\left(\left(\sigma_i^n\right)^4\overline{G}_{0,0}^{n,i}\,|\mathcal{F}_i^n\right) = \frac{\theta^2(\sigma_i^n)^4\Delta_n\phi^2(0)}{2(\alpha_i^n)^2} + O_p(\Delta_n^{1+\rho}), \tag{31}$$

$$\mathbb{E}\left(\widehat{G}_{0,0}^{n,i}\left(\sigma_i^n\right)^2\left(\gamma_i^n\right)^2\,|\mathcal{F}_i^n\right) = \frac{\theta(\sigma_i^n\gamma_i^n)^2\Delta_n^{\frac{1}{2}}\phi(0)}{\alpha_i^n} + O_p(\Delta_n^{\frac{1}{2}+\rho}), \tag{32}$$

$$\boldsymbol{r}(0,0)_n = \frac{\overline{\phi}(0)R\sqrt{\Delta_n}}{\theta} + o(\Delta_n^{1/2}) \quad \overline{\boldsymbol{r}}(0,0,0,0)_n = \frac{3\overline{\phi}^2(0)R^2\Delta_n}{\theta^2} + o(\Delta_n), \tag{33}$$

where we obtain (32) and (31) by (JLZ-A.38) and (JLZ-A.39). (33) follows from Lemma C.1. By Lemma (JLZ-A.11), we have

$$\begin{aligned}
&\sum_{i=0}^{N_t^n-h_n} \mathbb{E}\left(\delta_i^{n,8}\,|\mathcal{K}_i^n\right) \xrightarrow{\mathbb{P}} 3\theta^2\phi^2(0)\int_0^t \frac{\sigma_s^4}{\alpha_s}\mathrm{d}s;\\
&\sum_{i=0}^{N_t^n-h_n} \mathbb{E}\left(\delta_i^{n,9}\,|\mathcal{K}_i^n\right) \xrightarrow{\mathbb{P}} 6\phi(0)\overline{\phi}(0)R\int_0^t \sigma_s^2\gamma_s^2\mathrm{d}s;\\
&\sum_{i=0}^{N_t^n-h_n} \mathbb{E}\left(\delta_i^{n,10}\,|\mathcal{K}_i^n\right) \xrightarrow{\mathbb{P}} \frac{3\overline{\phi}^2(0)R^2}{\theta^2}\int_0^t \gamma_s^4\alpha_s\mathrm{d}s.
\end{aligned} \tag{34}$$

Let $\Theta_i^n := (\overline{Y}_i^{c,n})^4 - \mathbb{E}\left((\overline{Y}_i^{c,n})^4\,|\mathcal{K}_i^n\right)$ 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 $\mathbb{E}\left((\Theta_i^n)^2\right) \leq \mathbb{E}\left((\overline{Y}_i^{c,n})^8\right) \leq K\Delta_n^2$ (by (JLZ-A.28)). Apply Lemma A.6 in Jacod et al. (2017), we have

$$\mathbb{E}\left(\left|\sum_{i=0}^{N_t^n-h_n}\Theta_i^n\right|\right) \leq K\sqrt{\Delta_n}. \tag{35}$$

Next, (JLZ-A.33) implies

$$\mathbb{E}\left(\left|\sum_{i=0}^{N_t^n-h_n}\widetilde{\Theta}_i^n\right|\right) \leq K\Delta_n^{\frac{\eta}{2}} \tag{36}$$

for some $\eta > 0$, 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 (30), (34), (35) and (36), we have (27). ∎

Proof of (28). We denote (recall $R(Y)_{i,h_n}^n$ and $r(\chi;|\ell|)_{i,h_n}^n$ are defined in (8))

$$\begin{aligned}
&\Xi_t^{n,1} := \frac{1}{h_n}\sum_{i=k_n}^{N_t^n-h_n}\left(\overline{Y}_i^{c,n}\right)^2 R(Y)_{i,h_n}^n; \quad \Xi_t^{n,2} := \frac{1}{h_n}\sum_{i=k_n}^{N_t^n-h_n}\left(\overline{Y}_i^{c,n}\right)^2(\gamma_i^n)^2\sum_{|\ell|\leq\ell_n} r(\chi;|\ell|)_{i,h_n}^n;\\
&\Xi_t^{n,3} := \frac{1}{h_n}\sum_{i=k_n}^{N_t^n-h_n}\left(\overline{Y}_i^{c,n}\right)^2(\gamma_i^n)^2\sum_{|\ell|\leq\ell_n} r(\ell); \quad \Xi_t^{n,4} := \frac{1}{h_n}\sum_{i=k_n}^{N_t^n-h_n}\left(\overline{Y}_i^{c,n}\right)^2(\gamma_i^n)^2 R.
\end{aligned}$$

First, we prove

$$\Xi_t^{n,4} \xrightarrow{\mathbb{P}} \phi(0)R\int_0^t \sigma_s^2\gamma_s^2\mathrm{d}s + \frac{\overline{\phi}(0)R^2}{\theta^2}\int_0^t \gamma_s^4\mathrm{d}A_s, \tag{37}$$

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

$$\frac{1}{h_n}\sum_{i=k_n}^{N_t^n-h_n}\left(\overline{X}_i^{c,n}\right)^2(\gamma_i^n)^2 R \xrightarrow{\mathbb{P}} \phi(0)R\int_0^t \sigma_s^2\gamma_s^2\mathrm{d}s; \tag{38}$$

$$\frac{1}{h_n}\sum_{i=k_n}^{N_t^n-h_n}(\overline{\varepsilon}_i^n)^2(\gamma_i^n)^2 R \xrightarrow{\mathbb{P}} \frac{\overline{\phi}(0)R^2}{\theta^2}\int_0^t \gamma_s^4\mathrm{d}A_s; \tag{39}$$

$$\frac{1}{h_n}\sum_{i=k_n}^{N_t^n-h_n}\overline{X}_i^{c,n}\overline{\varepsilon}_i^n\left(\gamma_i^n\right)^2 R \xrightarrow{\mathbb{P}} 0. \tag{40}$$

By (JLZ-A.19), (JZL-A.21), (JLZ-A.25) and (JLZ-A.38), we have

$$\frac{1}{h_n}\sum_{i=k_n}^{N_t^n-h_n}\mathbb{E}\left(\left(\overline{X}_i^{c,n}\right)^2(\gamma_i^n)^2 R\,|\mathcal{K}_i^n\right) = \phi(0)R\sum_{i=k_n}^{N_t^n-h_n}(\gamma_i^n\sigma_i^n)^2/\alpha_i^n + O(\Delta_n^\rho);$$
$$\frac{1}{h_n}\sum_{i=k_n}^{N_t^n-h_n}\mathbb{E}\left((\overline{\varepsilon}_i^n)^2(\gamma_i^n)^2 R\,|\mathcal{K}_i^n\right) = \frac{\Delta_n\overline{\phi}(0)R^2}{\theta^2}\sum_{i=k_n}^{N_t^n-h_n}(\gamma_i^n)^4 + o_p(1).$$

(JLZ-A.43) implies

$$\frac{1}{h_n}\sum_{i=k_n}^{N_t^n-h_n}\mathbb{E}\left(\left(\overline{X}_i^{c,n}\right)^2(\gamma_i^n)^2 R\,|\mathcal{K}_i^n\right) \xrightarrow{\mathbb{P}} \phi(0)R\int_0^t\sigma_s^2\gamma_s^2 \mathrm{d}s;$$
$$\frac{1}{h_n}\sum_{i=k_n}^{N_t^n-h_n}\mathbb{E}\left((\overline{\varepsilon}_i^n)^2(\gamma_i^n)^2 R\,|\mathcal{K}_i^n\right) \xrightarrow{\mathbb{P}} \frac{\overline{\phi}(0)R^2}{\theta^2}\int_0^t\gamma_s^4 \mathrm{d}A_s.$$

Let $\Gamma_i^{n,1} := (\overline{X}_i^{c,n}\gamma_i^n)^2R - \mathbb{E}\left((\overline{X}_i^{c,n}\gamma_i^n)^2R\,|\mathcal{K}_i^n\right)$, $\Gamma_i^{n,2} := (\overline{\varepsilon}_i^{c,n}\gamma_i^n)^2R - \mathbb{E}\left((\overline{\varepsilon}_i^{c,n}\gamma_i^n)^2R\,|\mathcal{K}_i^n\right)$. Then it's immediate that $\mathbb{E}\left(\left|\mathbb{E}\left(\Gamma_i^{n,k}\,|\mathcal{K}_i^n\right)\right|\right) = 0, \mathbb{E}\left((\Gamma_i^{n,k})^2\right) \le K\Delta_n$ 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 $\frac{1}{h_n}\sum_{i=k_n}^{N_t^n-h_n}\Gamma_i^{n,k} \xrightarrow{\mathbb{P}} 0$, for $k=1,2$. This proves (38) and (39).

To see the last convergence in (40), we denote $a_i^n = \Delta_n^{-\frac{1}{2}}\overline{X}_i^{c,n}\overline{\varepsilon}_i^n\left(\gamma_i^n\right)^2$. Since χ has mean zero, by the independence of $\mathcal{G}, \mathcal{F}_\infty$ and successive conditioning, we have $\mathbb{E}(a_i^n) = 0$. Next, we have by (JLZ-A.19) and (JLZ-A.25) the following:

$$\mathbb{E}\left((a_i^n)^2\right) = \Delta_n^{-1}\left(\mathbb{E}\left(\left((\overline{\varepsilon}_i^n)^2 - (\gamma_i^n)^2\boldsymbol{r}(0,0)_n\right)(\overline{X}_i^n)^2(\gamma_i^n)^4\right) + \mathbb{E}\left((\gamma_i^n)^6(\overline{X}_i^n)^2\boldsymbol{r}(0,0)_n\right)\right),$$

which is bounded by a constant. Now the convergence follows from the law of large numbers. This finishes the proof of (37).

Now we denote

$$\Xi_t^{n,1} - \Xi_t^{n,2} = \frac{1}{h_n}\sum_{i=k_n}^{N_t^n-h_n}(\overline{Y}_i^{c,n})^2\sum_{|\ell|\le\ell_n}\sum_{k=i+k_n}^{i+h_n-k_n-\ell}\zeta_{k,\ell}^{n,i} \text{ where } \zeta_{k,\ell}^{n,i} := \frac{1}{h_n}\left(\Delta_{k,\ell}^{n,k_n}Y - (\gamma_i^n)^2\Delta_{k,\ell}^{n,k_n}\chi\right).$$

Let $\varrho_{i,\ell}^n := \sum_{k=i+k_n}^{i+h_n-k_n-\ell}\zeta_{k,\ell}^{n,i}$. By Lemma C.2, we have $\mathbb{E}\left((\varrho_{i,\ell}^n)^2\right) \le K(\Delta_n(k_n+h_n+\ell_n))^{1/z}$. Let $\varrho_i^n := \sum_{|\ell|\le\ell_n}\varrho_{i,|\ell|}^n$, we have $\mathbb{E}\left((\varrho_i^n)^2\right) \le K\ell_n^2(\Delta_n(h_n+\ell_n+k_n))^{\frac{1}{z}}$. By Cauchy-Schwarz inequality and (JLZ-A.28), we have

$$\mathbb{E}\left(\left|\Xi_t^{n,1} - \Xi_t^{n,2}\right|\right) \le K\ell_n(\Delta_n(k_n+h_n+\ell_n))^{\frac{1}{2z}} \to 0, \tag{41}$$

for z close to 1.

Let $R_{i,\ell_n}^{n,\chi} := \sum_{|\ell|\le\ell_n}r(\chi;|\ell|)_{i,h_n}^n$, according to the limit distribution in Theorem LL-C.1, we have $\mathbb{E}\left((r(\chi;|\ell|)^n_{i,h_n}-r(\ell))^2\right) \leq Kh_n^{-1}$. Thus, $\mathbb{E}\left((R^{n,\chi}_{i,\ell_n}-\sum_{|\ell|\leq\ell_n} r(\ell))^2\right) \leq K\ell_n^2/h_n$. Apply Cauchy-Schwarz inequality, we get

$$\mathbb{E}\left(|\Xi_t^{n,2}-\Xi_t^{n,3}|\right) \leq K\ell_n\Delta_n^{3/4}. \tag{42}$$

It's immediate that

$$\mathbb{E}\left(|\Xi_t^{n,3}-\Xi_t^{n,4}|\right) \leq \frac{K}{\ell_n^{v-1}} \tag{43}$$

Let $r(\chi;|\ell|)^n_{i,h_n}=r(|\ell|)+e^n_{|\ell|}$, then Theorem LL-C.1 implies $\mathbb{E}\left((e^n_\ell)^2\right)\leq K/h_n$. Thus

$$\mathbb{E}\left((E^n_{\ell_n})^2\right) \leq K\ell_n^2/h_n, \tag{44}$$

where $E^n_{\ell_n} := \sum_{|\ell|\leq\ell_n} e^n_{|\ell|}$. Therefore,

$$\mathbb{E}\left(\left(\sum_{|\ell|\leq\ell_n} r(\chi;|\ell|)^n_{i,h_n}\right)^2\right) \leq K\left(R^2+\mathbb{E}\left((E^n_{\ell_n})^2\right)\right) \leq K \Rightarrow \mathbb{E}\left((R(Y)^n_{i,h_n})^2\right) \leq K. \tag{45}$$

The last implication follows from Lemma C.2. Then (JLZ-A.33) and Cauchy-Schwarz inequality yield for some $\eta>0$ that $\mathbb{E}\left(\left|\left((\overline{Y}_i^n)^2\mathbf{1}_{\{|\overline{Y}_i^n|\leq u_n\}}-(\overline{Y}_i^{c,n})^2\right)R^n_{i,h_n}\right|\right) \leq K\Delta_n^{\frac{1}{2}+\frac{\eta}{4}}$, which further implies

$$\mathbb{E}\left(\left|\Xi_t^{n,1}-V_t^{n,2}\right|\right) \leq K\Delta_n^{\eta/4}. \tag{46}$$

We proved (28) by (37), (41), (42), (43) and (46). ∎

Proof of (29). We have by (45) and Cauchy-Schwarz inequality that

$$\mathbb{E}\left(\left|(R(Y)^n_{i,h_n})^2-(\gamma_i^n)^4\left(R^{n,\chi}_{i,\ell_n}\right)^2\right|\right) \leq K\sqrt{\mathbb{E}\left(\left(R(Y)^n_{i,h_n}-(\gamma_i^n)^2R^{n,\chi}_{i,\ell_n}\right)^2\right)} \leq K\ell_n(\Delta_n(k_n+h_n+\ell_n))^{\frac{1}{2z}}. \tag{47}$$

Similarly, we have

$$\begin{aligned}
&\mathbb{E}\left(\left|\left(R^{n,\chi}_{i,\ell_n}\right)^2-\left(\sum_{|\ell|\leq\ell_n} r(|\ell|)\right)^2\right|\right) \leq K\sqrt{\mathbb{E}\left((E_i^n)^2\right)} \overset{(44)}{\leq} K\ell_n h_n^{-\frac{1}{2}};\\
&\mathbb{E}\left(\left|R^2-\left(\sum_{|\ell|\leq\ell_n} r(|\ell|)\right)^2\right|\right) \leq K\ell_n^{-(v-1)}.
\end{aligned} \tag{48}$$

By (47) and (48), we have (29). ∎

Proof of Theorem 3.1. The stable convergence to $\mathcal{U}_t$ follows from Theorem 3.1 in Jacod et al. (2019) and our Lemma C.3, which states that the ReMeDI correction of noise is valid.

To see the convergence to a standard normal variable, it suffices to have $K_{g,1}V_t^{n,1}+2K_{g,2}V_t^{n,2}+K_{g,3}V_t^{n,3} \xrightarrow{\mathbb{P}} \frac{\theta\phi^4(0)}{4}\int_0^t \beta_s^2 ds$, which follows from (27), (28) and (29). ∎

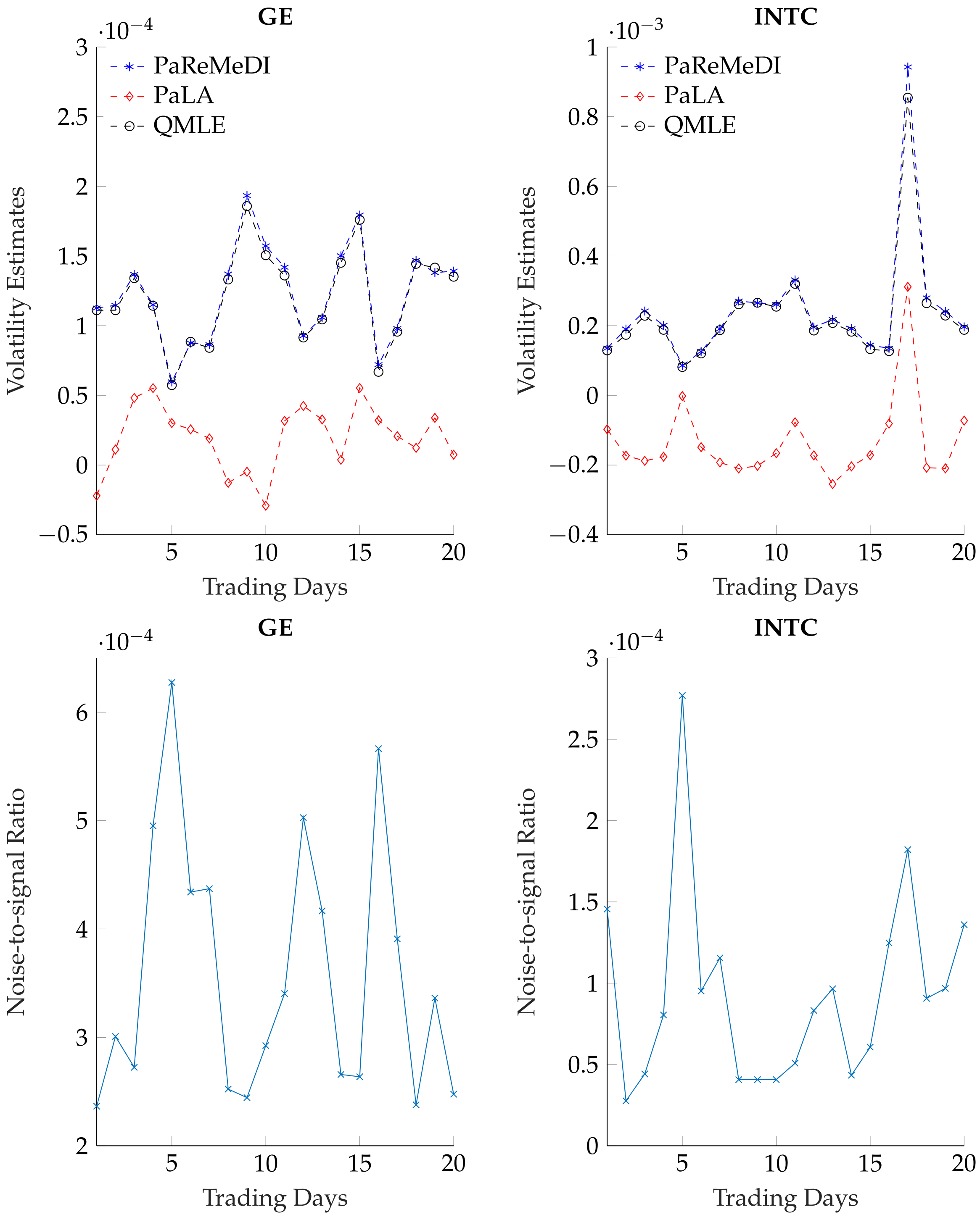

Figure 3: Volatility estimates (top panel) and Noise-to-signal ratios (bottom panel) of General Electric (ticker: GE) and Intel Corporation (ticker: INTC) for January, 2015. The tuning parameters are set as follows: For PaReMeDI, $k_n = 10, \ell_n = \lfloor N_t^{n\,\frac{1}{7}} \rfloor$, and for PaLA $k_n = 6, k'_n = \lfloor N_t^{n\,\frac{1}{8}} \rfloor$; θ are selected according to (14) for both methods. For the QMLE (with AIC selection criterion), the optimal q is selected within $\{5, 6, 7, 8, 9, 10\}$.

References

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., T. BOLLERSLEV, F. X. DIEBOLD, AND P. LABYS (2003): "Modeling and forecasting realized volatility," *Econometrica*, 71, 579–625.

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.

BOGOUSSLAVSKY, V. (2016): "Infrequent rebalancing, return autocorrelation, and seasonality," *Journal of Finance*, 71, 2967–3006.

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.

CHUNG, K. L. (2001): *A Course in Probability Theory*, Academic press.

DA, R. AND D. XIU (2020): "When moving-average models meet high-frequency data: Uniform inference on volatility," Tech. rep.

DELBAEN, F. AND W. SCHACHERMAYER (1994): "A general version of the fundamental theorem of asset pricing," *Mathematische Annalen*, 300, 463–520.

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. X. PRAZ, AND M. S. SEASHOLES (2018): "Asset price dynamics with limited attention," Tech. rep.

HO, T. AND H. R. STOLL (1981): "Optimal dealer pricing under transactions and return uncertainty," *Journal of Financial Economics*, 9, 47–73.

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 A. N. SHIRYAEV (2003): *Limit Theorems for Stochastic Processes*, vol. 288, Springer-Verlag Berlin.

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.

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, 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 (2021a): "A ReMeDI for microstructure noise," https://papers.ssrn.com/abstract=3788658.

——— (2021b): "Supplementary material for 'A ReMeDI for microstructure noise'," https://papers.ssrn.com/abstract=3788734.

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.

MANCINI, C. (2001): "Disentangling the jumps of the diffusion in a geometric jumping Brownian motion," *Giornale dell'Istituto Italiano degli Attuari*, 64, 19–47.

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.

RIO, E. (1997): "About the Lindeberg method for strongly mixing sequences," *ESAIM: Probability and Statistics*, 1, 35–61.

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. (2017): "Estimating the quadratic variation spectrum of noisy asset prices using generalized flat-top realized kernels," *Econometric Theory*, 33, 1457.

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 volatitility with noisy high-frequency data," *Journal of the American Statistical Association*, 100, 1394–1411.