Supplementary Material to “Dependent Microstructure Noise and Integrated Volatility Estimation from High-Frequency Data”

Z. Merrick Li*

Erasmus University Rotterdam

University of Amsterdam

and Tinbergen Institute

Roger J. A. Laeven†

Amsterdam School of Economics

University of Amsterdam, EURANDOM

and CentER

Michel H. Vellekoop‡

Amsterdam School of Economics

University of Amsterdam

June 11, 2018

Abstract

Sections A–G in this appendix contain detailed technical proofs of our results. In Sections H and I, we provide additional Monte Carlo simulation and empirical results.

*Corresponding author. University of Amsterdam, Amsterdam School of Economics, PO Box 15867, 1001 NJ Amsterdam, The Netherlands. Email: Z.Merrick.Li@gmail.com. Phone: +31 (0)20 5254252.

†University of Amsterdam, Amsterdam School of Economics, PO Box 15867, 1001 NJ Amsterdam, The Netherlands. Email: R.J.A.Laeven@uva.nl. Phone: +31 (0)20 5254219.

‡University of Amsterdam, Amsterdam School of Economics, PO Box 15867, 1001 NJ Amsterdam, The Netherlands. Email: M.H.Vellekoop@uva.nl. Phone: +31 (0)20 5254210.

Appendix

In the proofs that follow the constants C and $\delta \in (0,1)$ may vary from line to line. We add a subscript q if they depend on some parameter q.

A Proof of Proposition 3.1

Proof. Adopting the standard localization procedure (see e.g., Jacod and Protter (2011) for further details), we may assume that the processes a and σ are bounded by constants $C_a, C_\sigma > 0$. This yields for any such continuous Itô semimartingale X and stopping times $S \le T$ that

$$\mathbb{E}\left(|X_T - X_S|^p \,|\mathcal{F}_S\right) \le C_p \mathbb{E}\left(T - S \,|\mathcal{F}_S\right), \quad \forall p \ge 2. \tag{A.1}$$

Let $\Delta_n = 1/n$. For any process V, we write $\Delta^n_{i,j} V := V^n_{i+j} - V^n_i$, $j = 1, 2, \ldots, n-i$. Then, for the log-price process Y,

$$[Y,Y]^j_n := \sum_{i=0}^{n-j} (\Delta^n_{i,j} Y)^2 = \sum_{i=0}^{n-j} (\Delta^n_{i,j} X)^2 + 2 \sum_{i=0}^{n-j} \Delta^n_{i,j} X \; \Delta^n_{i,j} U + \sum_{i=0}^{n-j} (\Delta^n_{i,j} U)^2. \tag{A.2}$$

We now analyze the asymptotic properties of the three components on the right-hand side of (A.2):

(i) First note that $\sum_{i=0}^{n-j} (\Delta^n_{i,j} X)^2 / j \xrightarrow{\mathbb{P}} [X,X]$, where $[X,X]$ is the quadratic variation of X.

(ii) By the independence of X and U, we have

$$\sum_{i=0}^{n-j} \mathbb{E}\Big(\big(\Delta^n_{i,j} X \; \Delta^n_{i,j} U\big)^2\Big) = \sum_{i=0}^{n-j} \mathbb{E}\Big(\big(\Delta^n_{i,j} X\big)^2\Big) \mathbb{E}\Big(\big(\Delta^n_{i,j} U\big)^2\Big) \le Cj. \tag{A.3}$$

The last inequality follows from the fact that U has bounded moments and from an application of (A.1). Next,

$$\begin{aligned}
&\sum_{i,i':i<i'} \mathbb{E}\big(\Delta^n_{i,j} X \; \Delta^n_{i,j} U \; \Delta^n_{i',j} X \; \Delta^n_{i',j} U\big) \\
=& \sum_{i,i':i<i'} \mathbb{E}\big(\Delta^n_{i,j} X \; \Delta^n_{i',j} X\big) \mathbb{E}\big(\Delta^n_{i,j} U \; \Delta^n_{i',j} U\big) \\
\le& Cj\Delta_n \left(\sum_{i,i':i+j<i'} \mathbb{E}\big(\Delta^n_{i,j} U \; \Delta^n_{i',j} U\big) + \sum_{i,i':i+j \ge i' > i} \mathbb{E}\big(\Delta^n_{i,j} U \; \Delta^n_{i',j} U\big) \right) \\
\le& Cj^2.
\end{aligned} \tag{A.4}$$

The first inequality follows from the Cauchy-Schwarz inequality and (A.1). To see the second inequality, we apply the Cauchy-Schwarz inequality, Lemma VIII 3.102 of Jacod and Shiryaev

(2003) (hereafter abbreviated as JS-Lemma), and the fact that $v > 2$ to obtain

$$\begin{aligned}\sum_{i,i':i+j<i'} \mathbb{E}(\Delta^n_{i,j}U\ \Delta^n_{i',j}U) &= \sum_{i,i':i+j<i'} \mathbb{E}(\Delta^n_{i,j}U\ \mathbb{E}\left(\Delta^n_{i',j}U\,|\mathcal{F}_{(i+j)\Delta_n}\right)) \\ &\le C\sum_i \sum_{i':i+j<i'} \sqrt{\mathbb{E}\left(\left(\mathbb{E}\left(\Delta^n_{i',j}U\,|\mathcal{F}_{(i+j)\Delta_n}\right)\right)^2\right)} \qquad \text{(A.5)} \\ &\le C\sum_i \sum_{i':i+j<i'} (i'-(i+j))^{-v/2} \le C\Delta_n^{-1}.\end{aligned}$$

Eqns. (A.3) and (A.4) imply that $\mathbb{E}\left(\left(\sum_{i=0}^{n-j} \Delta^n_{i,j}X\ \Delta^n_{i,j}U\right)^2\right) \le Cj^2$, thus

$$\sum_{i=0}^{n-j} \Delta^n_{i,j}X\ \Delta^n_{i,j}U = O_p(j). \qquad \text{(A.6)}$$

(iii) Turning to the last sum of (A.2), let $\nu_j := \mathbb{E}\big((U^n_{i+j} - U^n_i)^2\big) = 2(\mathbf{Var}(U) - \gamma(j))$. For $i > j$, we obtain the following in a similar way in which we derived (A.5):

$$\left|\mathbf{Cov}\big((U^n_j - U^n_0)^2, (U^n_{i+j} - U^n_i)^2\big)\right| \le C(i-j)^{-v/2},$$

which implies

$$\mathbb{E}\left(\left(\sum_{i=0}^{n-j}\left((\Delta^n_{i,j}U)^2 - \nu_j\right)\right)^2\right) \le C\Delta_n^{-1}j. \qquad \text{(A.7)}$$

For any fixed j, any j_n satisfying $\Delta_n j_n \to 0, j_n \to \infty$, we have by (A.6), (A.7) and (4) that

$$\begin{aligned}&\widehat{\langle Y,Y\rangle}_n(j) - (\mathbf{Var}(U) - \gamma(j)) = O_p\left(\sqrt{\Delta_n j}\right); \\ &\widehat{\langle Y,Y\rangle}_n(j_n) - \mathbf{Var}(U) = O_p\left(\max\left\{\sqrt{\Delta_n j_n}, j_n^{-v/2}\right\}\right).\end{aligned} \qquad \text{(A.8)}$$

Now the stated results follow from (6).

□

B Proof of Proposition 3.2

Proof. Let $k = \lfloor \frac{n}{j} \rfloor$. We will adopt the square bracket notation in (A.2) for X and U as well. By Itô's isometry, we have

$$\begin{aligned}\mathbb{E}_\sigma\left([X,X]^j_{kj-1}\right) &= \sum_{i=0}^{j-1}\int_{i\Delta_n}^{((k-1)j+i)\Delta_n} \sigma_s^2 \mathrm{d}s = \sum_{i=0}^{j-1}\left(\int_0^{kj\Delta_n}\sigma_s^2\mathrm{d}s - \int_0^{i\Delta_n}\sigma_s^2\mathrm{d}s - \int_{((k-1)j+i)\Delta_n}^{kj\Delta_n}\sigma_s^2\mathrm{d}s\right) \\ &= j\int_0^{kj\Delta_n}\sigma_s^2\mathrm{d}s + O_p(j^2\Delta_n).\end{aligned}$$

Hence, we have

$$\mathbb{E}_\sigma\left([X,X]_n^j\right) = j\int_0^1 \sigma_s^2 \mathrm{d}s + O_p(j^2\Delta_n),$$

where the stochastic orders follow from the regularity conditions of the volatility path at 0 and 1. Furthermore, it is immediate that $\mathbb{E}_\sigma\left([U,U]_n^j\right) = 2(n-j+1)(\mathbf{Var}(U)-\gamma(j))$. Thus, we have, by the independence of X and U,

$$\mathbb{E}_\sigma\left(\widehat{\langle Y,Y\rangle}_n(j)\right) = \frac{j\int_0^1 \sigma_s^2 \mathrm{d}s}{2(n-j+1)} + \mathbf{Var}(U) - \gamma(j) + O_p(j^2\Delta_n^2).$$

□

C Proof of Proposition 4.1

Proof of Proposition 4.1. Recall that

$$\begin{aligned}\bar{U}_m^n &= \frac{1}{k_n+1}\sum_{i=(2m-2)k_n}^{(2m-1)k_n}\left(U_{i+k_n}^n - U_i^n\right)\\ &= \frac{1}{k_n+1}\left(\sum_{i=(2m-1)k_n}^{2mk_n} U_i^n - \sum_{i=(2m-2)k_n}^{(2m-1)k_n} U_i^n\right).\end{aligned}$$

Also recall that U is symmetrically distributed around 0, whence $\bar{U}_m^n$ is equal to the following in distribution:

$$\bar{U}_m^n \overset{d}{=} \frac{1}{k_n+1}\left(\sum_{i=(2m-2)k_n}^{2mk_n} U_i^n\right) + O_p(\sqrt{\Delta_n}). \tag{C.1}$$

Since $v>2$, we have $\sigma_U^2<\infty$, and an application of Corollary VIII 3.106 of Jacod and Shiryaev (2003) yields

$$\frac{1}{\sqrt{2k_n+1}}\sum_{i=(2m-2)k_n}^{2mk_n} U_i^n \xrightarrow{\mathcal{L}} \mathcal{N}\left(0,\sigma_U^2\right),$$

whence

$$n^{1/4}\bar{U}_m^n \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, 2\sigma_U^2/c\right).$$

□

D Proof of Proposition 4.2

Proof. For any fixed j, (A.8) implies $\widehat{\gamma(j)}_n - \gamma(j) = O_p\left(\max\left\{\sqrt{\Delta_n j_n}, j_n^{-v/2}\right\}\right)$. Therefore,

$$\widehat{\sigma_U^2} - \sum_{j=-i_n}^{i_n} \gamma(j) = O_p\left(\max\left\{\sqrt{\Delta_n j_n i_n^2}, j_n^{-v/2} i_n\right\}\right).$$

Now the result follows given that $\Delta_n j_n^3 \to 0, i_n \le j_n, i_n \to \infty, v > 2$. □

E Proof of Theorem 4.1

The proof of this theorem basically follows Podolskij and Vetter (2009), but we need to deal with generally dependent noise.

First, we introduce some notation:

$$\beta_m^n := n^{1/4}\left(\sigma_{\frac{m-1}{M_n}} \overline{W}_m^n + \bar{U}_m^n\right); \tag{E.1}$$

$$\xi_m^n := n^{1/4}\bar{Y}_m^n - \beta_m^n; \tag{E.2}$$

$$\eta_m^n := \frac{n^{r/4}}{2c}\mathbb{E}\left(\left|\bar{Y}_m^n\right|^r \,\middle|\mathcal{F}_{\frac{m-1}{M_n}}\right); \tag{E.3}$$

$$\widetilde{\eta_m^n} := \frac{\mu_r}{2c}\left(\frac{2c}{3}\sigma^2_{\frac{m-1}{M_n}} + \frac{2}{c}\sigma_U^2\right)^{\frac{r}{2}}; \tag{E.4}$$

$$\mathrm{PAV}^n := \sum_{m=1}^{M_n} \eta_m^n; \tag{E.5}$$

$$\widetilde{\mathrm{PAV}}^n := \sum_{m=1}^{M_n} \widetilde{\eta_m^n}. \tag{E.6}$$

Then, we state the following lemma:

Lemma E.1. *For any $q > 0$, there is some constant $C_q > 0$ (depending on q), such that $\forall m$:*

$$\mathbb{E}(|\xi_m^n|^q) + \mathbb{E}\left(\left|n^{1/4}\bar{X}_m^n\right|^q\right) < C_q; \tag{E.7}$$

and the following holds for $q \in (0, 2r+\varepsilon)$ with ε as defined in Theorem 4.1:

$$\mathbb{E}(|\beta_m^n|^q) + \mathbb{E}\left(\left|n^{1/4}\bar{Y}_m^n\right|^q\right) < C_q. \tag{E.8}$$

Proof of Lemma E.1. The boundedness of moments of ξ_m^n and $n^{1/4}\bar{X}_m^n$ (which don't depend on the noise) follows from Lemma 1 in Podolskij and Vetter (2009).

Now we show the boundedness of $\mathbb{E}\left(|n^{1/4}\bar{Y}_m^n|^q\right)$ for $0 < q < 2r + \varepsilon$. We note (see Proposition 3.8 in White (2000)) that there is some C_q so that the following is true:

$$\mathbb{E}\left(\left|n^{1/4}\bar{Y}_m^n\right|^q\right) \le C_q\left(\mathbb{E}\left(\left|n^{1/4}\bar{X}_m^n\right|^q\right) + \mathbb{E}\left(\left|n^{1/4}\bar{U}_m^n\right|^q\right)\right).$$

Boundedness of $\mathbb{E}\Big(\big|n^{1/4}\bar{X}_m^n\big|^q\Big)$ has already been established, while $\mathbb{E}\Big(\big|n^{1/4}\bar{U}_m^n\big|^q\Big)$ is bounded by Proposition 4.1 and a well known fact that convergence in distribution implies convergence in moments under uniformly bounded moments condition, see, e.g., Theorem 4.5.2 of Chung (2001). A similar proof holds for $\mathbb{E}(|\beta_m^n|^q)$. □

Proof of Theorem 4.1. We present the proof in several steps.

(i) We first prove that

$$\mathrm{PAV}(Y,r)_n \; - \frac{1}{M_n}\mathrm{PAV}^n \xrightarrow{\mathbb{P}} 0. \tag{E.9}$$

First, recall our choice of $M_n = \left\lfloor \frac{\sqrt{n}}{2c} \right\rfloor$. Next, observe that the difference on the left-hand side of (E.9) is in fact a sum of *martingale differences*:

$$\begin{aligned}&\mathrm{PAV}(Y,r)_n \; - \frac{1}{M_n}\mathrm{PAV}^n\\ =&\sum_{m=1}^{M_n}\frac{1}{\sqrt{n}}\left(\left|n^{\frac14}\bar{Y}_m^n\right|^r - \mathbb{E}\left(\left|n^{\frac14}\bar{Y}_m^n\right|^r \Big|\mathcal{F}_{\frac{m-1}{M_n}}\right)\right).\end{aligned}$$

In light of Lemma 2.2.11 in Jacod and Protter (2011), it suffices to show that

$$\frac{1}{n}\sum_{m=1}^{M_n}\mathbb{E}\left(\left|n^{\frac14}\bar{Y}_m^n\right|^{2r}\Big|\mathcal{F}_{\frac{m-1}{M_n}}\right) \xrightarrow{\mathbb{P}} 0. \tag{E.10}$$

But this follows from the boundedness established in Lemma E.1 and the choice of M_n.

(ii) Next, we prove that

$$\frac{1}{M_n}\mathrm{PAV}^n - \frac{1}{M_n}\widetilde{\mathrm{PAV}}^n \xrightarrow{\mathbb{P}} 0. \tag{E.11}$$

To prove this, we proceed in several steps:

(a) We first note that the error of approximating $n^{1/4}\bar{Y}_m^n$ by β_m^n, denoted by ξ_m^n in (E.2), is small in the sense that

$$\frac{1}{M_n}\sum_{m=1}^{M_n}\mathbb{E}\Big(|\xi_m^n|^2\Big) \to 0. \tag{E.12}$$

For a detailed proof, see Podolskij and Vetter (2009). (Note that our assumptions on the noise process are different from Podolskij and Vetter (2009), but the noise terms don't appear in ξ_m^n.)

(b) Next, define the approximation error

$$\zeta_m^n := \frac{\left|n^{1/4}\bar{Y}_m^n\right|^r - |\beta_m^n|^r}{2c}.$$

We note that this error is also small:

$$\frac{1}{M_n}\sum_{m=1}^{M_n}\mathbb{E}(|\zeta_m^n|) \to 0, \tag{E.13}$$

which follows from

$$\frac{1}{M_n}\sum_{m=1}^{M_n}\mathbb{E}\Big(|\zeta_m^n|^2\Big)\to 0. \tag{E.14}$$

This, in turn, can be proved following Podolskij and Vetter (2009). (E.13) then follows, and it implies

$$\frac{1}{M_n}\sum_{m=1}^{M_n}\mathbb{E}\left(\zeta_m^n\,\middle|\mathcal{F}_{\frac{m-1}{M_n}}\right)\xrightarrow{\mathbb{P}}0, \tag{E.15}$$

by the Markov inequality.

(c) Now we show the following:

$$\mathbb{E}\left(|\beta_m^n|^r\,\middle|\mathcal{F}_{\frac{m-1}{M_n}}\right)=\mu_r\left(\frac{2c}{3}\sigma^2_{\frac{m-1}{M_n}}+\frac{2\sigma_U^2}{c}\right)^{\frac{r}{2}}+o_p(1), \tag{E.16}$$

which holds uniformly in m. Recall that $r\geq 2$ is an even integer. Let $r_n\to\infty$ but $r_n=o(n^{1/2})$. Denote

$$\begin{aligned}\overline{\beta}^n_{m-1,r_n}&=\frac{n^{1/4}}{k_n+1}\left(\sum_{i=(2m-2)k_n}^{(2m-2)k_n+r_n}\sigma_{\frac{m-1}{M_n}}\left(W^n_{i+k_n}-W^n_i\right)+\left(U^n_{i+k_n}-U^n_i\right)\right)\\&=:n^{1/4}\left(\sigma_{\frac{m-1}{M_n}}\overline{W}^n_{m-1,r_n}+\overline{U}^n_{m-1,r_n}\right);\\\overline{\beta}^n_{r_n,m}&=\frac{n^{1/4}}{k_n+1}\left(\sum_{i=(2m-2)k_n+r_n+1}^{(2m-1)k_n}\sigma_{\frac{m-1}{M_n}}\left(W^n_{i+k_n}-W^n_i\right)+\left(U^n_{i+k_n}-U^n_i\right)\right)\\&=:n^{1/4}\left(\sigma_{\frac{m-1}{M_n}}\overline{W}^n_{r_n,m}+\overline{U}^n_{r_n,m}\right).\end{aligned}$$

Then, we have $\beta^n_m=\overline{\beta}^n_{m-1,r_n}+\overline{\beta}^n_{r_n,m}$. Furthermore, by our construction, $\overline{\beta}^n_{m-1,r_n}=o_p(1)$ and $\overline{\beta}^n_{r_n,m}$ has the same asymptotic distribution as β^n_m, which can be derived from the asymptotic distributions of $n^{1/4}\overline{U}^n_m$ and $n^{1/4}\overline{W}^n_m$, and the independence assumption between X and U. By the Mean Value Theorem, we have

$$\mathbb{E}\left((\beta^n_m)^r-\left(\overline{\beta}^n_{r_n,m}\right)^r\,\middle|\mathcal{F}_{\frac{m-1}{M_n}}\right)=\mathbb{E}\left(r\left(\overline{\beta}^n_{r_n,m}\right)^{r-1}\left(\overline{\beta}^n_{m-1,r_n}\right)\,\middle|\mathcal{F}_{\frac{m-1}{M_n}}\right)+o_p(1).$$

The moment conditions and an application of Cauchy-Schwarz inequality yields

$$\mathbb{E}\left(\left(\overline{\beta}^n_{r_n,m}\right)^{r-1}\left(\overline{\beta}^n_{m-1,r_n}\right)\,\middle|\mathcal{F}_{\frac{m-1}{M_n}}\right)=o_p(1).$$

Thus,

$$\mathbb{E}\left((\beta^n_m)^r\,\middle|\mathcal{F}_{\frac{m-1}{M_n}}\right)=\mathbb{E}\left(\left(\overline{\beta}^n_{r_n,m}\right)^r\,\middle|\mathcal{F}_{\frac{m-1}{M_n}}\right)+o_p(1). \tag{E.17}$$

For any $l\leq r$, define $\overline{U}^{n,l}_{r_n,m}:=\left(n^{1/4}\overline{U}^n_{r_n,m}\right)^l$, and let

$$C_l:=\mathbb{E}\bigg(\Big(\mathbb{E}\left(\overline{U}^{n,l}_{r_n,m}\,\middle|\mathcal{F}_{\frac{m-1}{M_n}}\right)-\mathbb{E}\Big(\overline{U}^{n,l}_{r_n,m}\Big)\Big)^2\bigg).$$

By the JS-Lemma, we have $C_l \leq C r_n^{-v}$. Let

$$\Lambda_l := \frac{\mathbb{E}\left(\overline{U}^{n,l}_{r_n,m} \middle| \mathcal{F}_{\frac{m-1}{M_n}}\right) - \mathbb{E}\left(\overline{U}^{n,l}_{r_n,m}\right)}{\sqrt{C_l}};$$

note that $\mathbb{E}(\Lambda_l^2) = 1$. Thus,

$$\mathbb{E}\left(\overline{U}^{n,l}_{r_n,m} \middle| \mathcal{F}_{\frac{m-1}{M_n}}\right) = \mathbb{E}\left(\overline{U}^{n,l}_{r_n,m}\right) + \sqrt{C_l}\Lambda_l. \tag{E.18}$$

Therefore, we can substitute the conditional moments by the unconditional moments and we obtain the following ($C_r^k = \frac{r!}{k!(r-k)!}$ denotes the binomial coefficient):

$$\begin{aligned}
&\mathbb{E}\left(\left(\overline{\beta}^n_{r_n,m}\right)^r \middle| \mathcal{F}_{\frac{m-1}{M_n}}\right) \\
=&\mathbb{E}\left(\sum_{k=0}^r C_r^k \sigma^k_{\frac{m-1}{M_n}} \left(n^{1/4}\overline{W}^n_{r_n,m}\right)^k \left(n^{1/4}\overline{U}^n_{r_n,m}\right)^{r-k} \middle| \mathcal{F}_{\frac{m-1}{M_n}}\right) \\
=&\sum_{k=0}^r C_r^k \sigma^k_{\frac{m-1}{M_n}} \mathbb{E}\left(\left(n^{1/4}\overline{W}^n_{r_n,m}\right)^k \middle| \sigma_{\frac{m-1}{M_n}}\right) \mathbb{E}\left(\left(n^{1/4}\overline{U}^n_{r_n,m}\right)^{r-k} \middle| \mathcal{F}_{\frac{m-1}{M_n}}\right) \\
=&\mathbb{E}\left(\left(\overline{\beta}^n_{r_n,m}\right)^r \middle| \sigma_{\frac{m-1}{M_n}}\right) + \sum_{k=0}^r C_k^r \sigma^k_{\frac{m-1}{M_n}} \mathbb{E}\left(\left(n^{1/4}\overline{W}^n_{r_n,m}\right)^k \middle| \sigma_{\frac{m-1}{M_n}}\right) \sqrt{C_{r-k}}\Lambda_{r-k}.
\end{aligned}$$

Clearly, the last term is $o_p(1)$, and together with (E.17), we have

$$\begin{aligned}
\mathbb{E}\left((\beta^n_m)^r \middle| \mathcal{F}_{\frac{m-1}{M_n}}\right) &= \mathbb{E}\left(\left(\overline{\beta}^n_{r_n,m}\right)^r \middle| \sigma_{\frac{m-1}{M_n}}\right) + o_p(1) \\
&= \mu_r \left(\frac{2c}{3}\sigma^2_{\frac{m-1}{M_n}} + \frac{2\sigma_U^2}{c}\right)^{\frac{r}{2}} + o_p(1).
\end{aligned} \tag{E.19}$$

The last equality is a consequence of the asymptotic distribution of β^n_m.

(d) Now (E.11) follows from (E.15) and (E.19).

(iii) Following Proposition 2.2.8 in Jacod and Protter (2011), we see that the Riemann approximation converges:

$$\frac{1}{M_n}\sum_{m=1}^{M_n} \widetilde{\mathrm{PAV}}^n \xrightarrow{\mathbb{P}} \mathrm{PAV}(Y,r). \tag{E.20}$$

Recall that we already proved that

$$\mathrm{PAV}(Y,r)_n - \frac{1}{M_n}\mathrm{PAV}^n \xrightarrow{\mathbb{P}} 0; \quad \text{and} \quad \frac{1}{M_n}\mathrm{PAV}^n - \frac{1}{M_n}\widetilde{\mathrm{PAV}}^n \xrightarrow{\mathbb{P}} 0;$$

in previous steps. Now it is immediate to conclude that

$$\mathrm{PAV}(Y,r)_n \xrightarrow{\mathbb{P}} \mathrm{PAV}(Y,r).$$

This finalizes the proof of Theorem 4.1.

□

F Robustness to Irregular Sampling

In this section, we show that the consistency results for integrated volatility in Theorem 4.1 and Corollary 4.1 can be extended to irregular sampling times for the case $r = 2$, by adapting the approach in Appendix C of Christensen et al. (2014) to allow for serially dependent noise in our general setting (recall $Y_i^n = X_{t_i^n} + U_i^n$). Let $f : [0,1] \mapsto [0,1]$ be a strictly increasing map with Lipschitz continuous first order derivatives. Let $f(0) = 0$ and $f(1) = 1$. Suppose that the observation times are $\{t_i^n = f(i/n) : 0 \leq i \leq n\}$. Let $C'_f = \max_{x\in[0,1]} |f'(x)|$. Note that $C'_f < \infty$ by the continuity of f'.

First, we note that the asymptotic results related to the noise process we derived so far still hold under irregular sampling, because the noise is indexed by i rather than by t_i in our setting. The proof then proceeds in several steps:

1. We first provide the analogs of Lemma E.1 and step (i) in the proof of Theorem 4.1. Assume $q \geq 1$. Then,

$$
\begin{aligned}
\mathbb{E}(|\xi_m^n|^q) &= \mathbb{E}\left(\left|\frac{n^{1/4}}{k_n+1}\sum_{i=(2m-2)k_n}^{(2m-1)k_n} X_{i+k_n}^n - X_i^n - \sigma_{t^n_{(2m-2)k_n}}\left(W_{i+k_n}^n - W_i^n\right)\right|^q\right)\\
&\leq \frac{n^{\frac{q}{4}}}{k_n+1}\sum_{i=(2m-2)k_n}^{(2m-1)k_n}\mathbb{E}\left(\left|X_{i+k_n}^n - X_i^n - \sigma_{t^n_{(2m-2)k_n}}\left(W_{i+k_n}^n - W_i^n\right)\right|^q\right)\\
&= \frac{n^{\frac{q}{4}}}{k_n+1}\sum_{i=(2m-2)k_n}^{(2m-1)k_n}\mathbb{E}\left(\left|\int_{t_i^n}^{t_{i+k_n}^n}\left(\alpha_s \mathrm{d}s + \left(\sigma_s - \sigma_{t^n_{(2m-2)k_n}}\right)\mathrm{d}W_s\right)\right|^q\right)\\
&\leq C_\alpha (C'_f)^q n^{-\frac{q}{4}} + \frac{C_q n^{\frac{q}{4}}}{k_n+1}\sum_{i=(2m-2)k_n}^{(2m-1)k_n}\mathbb{E}\left(\left|\int_{t_i^n}^{t_{i+k_n}^n}\left(\sigma_s - \sigma_{t^n_{(2m-2)k_n}}\right)\mathrm{d}W_s\right|^q\right)\\
&\leq C + \frac{C_q n^{\frac{q}{4}}}{k_n+1}\sum_{i=(2m-2)k_n}^{(2m-1)k_n}\mathbb{E}\left(\left(\int_{t_i^n}^{t_{i+k_n}^n}\left|\sigma_s - \sigma_{t^n_{(2m-2)k_n}}\right|^2 \mathrm{d}s\right)^{q/2}\right)\\
&\leq C.
\end{aligned}
$$

The second inequality follows from the boundedness of α and C'_f. The third inequality is an application of the Burkholder-Davis-Gundy inequality. The last inequality follows from the fact that σ is bounded. Similarly, we can prove that $\mathbb{E}\left(\left|n^{1/4}\bar{X}_m^n\right|^q\right)$ is bounded. For $q \in (0,1)$, the result is immediate using Jensen's inequality. Now the boundedness of $\mathbb{E}\left(\left|n^{1/4}\bar{Y}_m^n\right|^q\right)$, $q \in (0, 2r+\varepsilon)$, is obvious as the asymptotic distribution of the pre-averaged noise (which is indexed by i) does not change under irregular sampling.

2. Next, we prove the analog of step (ii) item (a) in the proof of Theorem 4.1. We have that

$$\begin{aligned}
&\mathbb{E}\Big(|\xi_m^n|^2\Big) \\
\le &\sum_{i=(2m-2)k_n}^{(2m-1)k_n} \frac{\mathbb{E}\Big(\Big|n^{\frac{1}{4}}\left(\left(X_{i+k_n}^n - X_i^n\right) - \sigma_{t^n_{(2m-2)k_n}}\left(W_{i+k_n}^n - W_i^n\right)\right)\Big|^2\Big)}{k_n+1} \\
= &\sum_{i=(2m-2)k_n}^{(2m-1)k_n} \frac{\mathbb{E}\Big(\Big|n^{\frac{1}{4}}\left(\int_{t_i^n}^{t_{i+k_n}^n}\alpha_s \mathrm{d}s + \int_{t_i^n}^{t_{i+k_n}^n}\left(\sigma_s - \sigma_{t^n_{(2m-2)k_n}}\right)\mathrm{d}W_s\right)\Big|\Big)^2}{k_n+1} \\
\le &\sum_{i=(2m-2)k_n}^{(2m-1)k_n} \frac{2\mathbb{E}\Big(n^{\frac{1}{2}}\left(\int_{t_i^n}^{t_{i+k_n}^n}\alpha_s \mathrm{d}s\right)^2 + n^{1/2}\int_{t_i^n}^{t_{i+k_n}^n}\left(\sigma_s - \sigma_{t^n_{(2m-2)k_n}}\right)^2\mathrm{d}s\Big)}{k_n+1} \\
\le &\frac{{C_f'}^2 C_\alpha}{\sqrt{n}} + 2n^{1/2}\mathbb{E}\left(\int_{t^n_{(2m-2)k_n}}^{t^n_{2mk_n}}\left(\sigma_s - \sigma_{t^n_{(2m-2)k_n}}\right)^2\mathrm{d}s\right).
\end{aligned}$$

The second inequality is due to the Cauchy's inequality and Itô's isometry. The third inequality is a consequence of the boundedness of $\alpha, |f'|$ and our choice of k_n; it is obtained by taking i to be the lower and upper bound. Now we have

$$\begin{aligned}
\frac{1}{M_n}\sum_{m=1}^{M_n}\mathbb{E}\Big(|\xi_m^n|^2\Big) &\le O(1/\sqrt{n}) + \frac{2n^{1/2}}{M_n}\sum_{m=1}^{M_n}\mathbb{E}\left(\int_{t^n_{(2m-2)k_n}}^{t^n_{2mk_n}}\left(\sigma_s - \sigma_{t^n_{(2m-2)k_n}}\right)^2\mathrm{d}s\right) \\
&= O(1/\sqrt{n}) + 4c\int_0^1 \mathbb{E}\left(\left(\sigma_s - \sigma_{\frac{\lfloor M_n s\rfloor}{M_n}}\right)^2\right)\mathrm{d}s.
\end{aligned}$$

Since $\sigma_{\frac{\lfloor M_n s\rfloor}{M_n}} \to \sigma_s$-a.s., and σ is bounded, upon applying Lebesgue's Dominated Convergence Theorem, we obtain the analog of (E.12). We note that the analog of item (b) of step (ii) in the proof of Theorem 4.1 is directly obtained because (6.10) in Podolskij and Vetter (2009) holds.

3. We now provide the analog of (E.19). First, we note that all the steps in proving (E.19) hold except those pertaining to the conditional variance of the pre-averaging Brownian motion. Next, we show that

$$\mathbf{Var}\Big(n^{1/4}\overline{W}_m^n\Big) = f'((2m-2)k_n/n)\frac{2c}{3} + o(1).$$

By the Lipschitz continuity of f' we obtain:

$$\begin{aligned}
&\mathbf{Var}\left(\sum_{i=2(m-1)k_n}^{(2m-1)k_n}(W^n_{i+k_n}-W^n_i)\right)\\
=&\sum_{i=2(m-1)k_n}^{(2m-1)k_n}\mathbf{Var}(W^n_{i+k_n}-W^n_i)+\sum_{i\neq j}\mathbf{Cov}(W^n_{i+k_n}-W^n_i,W^n_{j+k_n}-W^n_j)\\
=&\sum_{i=(2m-2)k_n}^{(2m-1)k_n}(t^n_{i+k_n}-t^n_i)+2\sum_{i=(2m-2)k_n}^{(2m-1)k_n-1}\sum_{j>i}(t^n_{i+k_n}-t^n_j)\\
=&\sum_{i=(2m-2)k_n}^{(2m-1)k_n}\left(f'(i/n)\frac{k_n}{n}+o(k_n/n)\right)+2\sum_{i=(2m-2)k_n}^{(2m-1)k_n-1}\sum_{j>i}\left(f'(j/n)\frac{i+k_n-j}{n}+o(k_n/n)\right)\\
=&f'\left(\frac{(2m-2)k_n}{n}\right)\sum_{i=(2m-2)k_n}^{(2m-1)k_n}\left(\frac{k_n}{n}+o(k_n/n)\right)\\
&+2\sum_{i=(2m-2)k_n}^{(2m-1)k_n-1}\sum_{j>i}\left(\frac{i+k_n-j}{n}+o(k_n/n)\right)\\
=&f'\left(\frac{(2m-2)k_n}{n}\right)\frac{2c^3\sqrt{n}}{3}+o(\sqrt{n}).
\end{aligned}$$

Now the analog of (E.19) (with $r=2$) is

$$\mathbb{E}\left((\beta^n_m)^2\,\Big|\,\mathcal{F}_{t^n_{(2m-2)k_n}}\right)=\left(f'\left(\frac{(2m-2)k_n}{n}\right)\sigma^2_{f\left(\frac{(2m-2)k_n}{n}\right)}\frac{2c}{3}+\frac{2\sigma_U^2}{c}\right)+o_p(1). \tag{F.1}$$

4. Finally, Riemann integrability yields the analog of (E.20):

$$\mathrm{PAV}(Y,2)_n\xrightarrow{\mathbb{P}}\int_0^1\left(f'(s)\sigma^2_{f(s)}\frac{2c}{3}+\frac{2\sigma_U^2}{c}\right)\mathrm{d}s=\int_0^1\left(\frac{2c}{3}\sigma_t^2+\frac{2\sigma_U^2}{c}\right)\mathrm{d}t.$$

The last equality is due to the change of variable $f(s)=t$.

G Proof of Theorem 4.2

We will first prove three lemmas. Then Theorem 4.2 follows as a consequence.

Lemma G.1. *We have that*

$$\mathbb{E}\left((\beta^n_m)^2\,\Big|\,\mathcal{F}_{\frac{m-1}{M_n}}\right)=\left(\frac{2c}{3}\sigma^2_{\frac{m-1}{M_n}}+\frac{2}{c}\sigma_U^2\right)+o_p(n^{-1/4}). \tag{G.1}$$

Proof. Let r_n satisfy

$$r_n\asymp n^{\vartheta},\quad \frac{1}{4v}<\vartheta<\frac{1}{4}. \tag{G.2}$$

To simplify notation, we let $s^n_m:=(2m-2)k_n+r_n$, and we recall our earlier notation used in the proof of Theorem 4.1:

$$\begin{aligned}
\overline{\beta}^n_{m-1,r_n} &= \frac{n^{1/4}}{k_n+1}\left(\sum_{i=(2m-2)k_n}^{(2m-2)k_n+r_n} \sigma_{\frac{m-1}{M_n}}\left(W^n_{i+k_n}-W^n_i\right)+\left(U^n_{i+k_n}-U^n_i\right)\right) \\
&=: n^{1/4}\left(\sigma_{\frac{m-1}{M_n}}\overline{W}^n_{m-1,r_n}+\overline{U}^n_{m-1,r_n}\right); \\
\overline{\beta}^n_{r_n,m} &= \frac{n^{1/4}}{k_n+1}\left(\sum_{i=(2m-2)k_n+r_n+1}^{(2m-1)k_n} \sigma_{\frac{m-1}{M_n}}\left(W^n_{i+k_n}-W^n_i\right)+\left(U^n_{i+k_n}-U^n_i\right)\right) \\
&=: n^{1/4}\left(\sigma_{\frac{m-1}{M_n}}\overline{W}^n_{r_n,m}+\overline{U}^n_{r_n,m}\right),
\end{aligned}$$

where $\overline{\beta}^n_{m-1,r_n}+\overline{\beta}^n_{r_n,m}=\beta^n_m$. The proof consists of three steps:

1. We start by showing that

$$\mathbb{E}\left((\beta^n_m)^2\,\Big|\mathcal{F}_{\frac{m-1}{M_n}}\right)-\mathbb{E}\left(\left(\overline{\beta}^n_{r_n,m}\right)^2\Big|\mathcal{F}_{\frac{m-1}{M_n}}\right)=o_p(n^{-1/4}). \tag{G.3}$$

To prove (G.3), we first prove that

$$\mathbb{E}\left(\left(\overline{\beta}^n_{m-1,r_n}\right)^2\Big|\mathcal{F}_{\frac{m-1}{M_n}}\right)=o_p(n^{-1/4}). \tag{G.4}$$

For this purpose, we show the following for any $k\leq i<j$:

$$\mathbb{E}\Big(\Big|\mathbb{E}\left(U^n_iU^n_j\,\Big|\mathcal{F}_{\frac{k}{n}}\right)\Big|\Big)\leq C\,(j-i)^{-v/2}. \tag{G.5}$$

To see this, we apply JS-Lemma to obtain that

$$c_{ij}:=\mathbb{E}\bigg(\left(\mathbb{E}\left(U^n_j\,\Big|\mathcal{F}_{\frac{i}{n}}\right)\right)^2\bigg)\leq C\,(j-i)^{-v}.$$

Then,

$$\mathbb{E}\Big(\Big|\mathbb{E}\left(U^n_iU^n_j\,\Big|\mathcal{F}_{\frac{k}{n}}\right)\Big|\Big)\leq\sqrt{C\,(j-i)^{-v}}\,\mathbb{E}\left(\left|\mathbb{E}\left(U^n_i\frac{\mathbb{E}\left(U^n_j\,\Big|\mathcal{F}_{\frac{i}{n}}\right)}{\sqrt{c_{ij}}}\,\Big|\mathcal{F}_{\frac{k}{n}}\right)\right|\right).$$

Now applying the Cauchy-Schwarz inequality and using the fact that the variance of noise is bounded, we obtain (G.5). From (G.5) and some simple algebra we find that

$$\mathbb{E}\left(\left(\sum_{i=(2m-2)k_n}^{s^n_m}\sigma_{\frac{m-1}{M_n}}\left(W^n_{i+k_n}-W^n_i\right)\right)^2\Big|\mathcal{F}_{\frac{m-1}{M_n}}\right)$$

is asymptotically much smaller than

$$\mathbb{E}\left(\left(\sum_{i=(2m-2)k_n}^{s_m^n}(U_{i+k_n}^n-U_i^n)\right)^2\Big|\mathcal{F}_{\frac{m-1}{M_n}}\right)=O_p(r_n)=o_p(n^{1/4}),\tag{G.6}$$

whence (G.4) holds.

Next, we prove that

$$\mathbb{E}\left(\left(\overline{\beta}_{r_n,m}^n\right)\left(\overline{\beta}_{m-1,r_n}^n\right)\Big|\mathcal{F}_{\frac{m-1}{M_n}}\right)=o_p(n^{-1/4}).\tag{G.7}$$

(Note that the left-hand side of (G.3) is equal to the left-hand side of (G.4) plus twice the left-hand side of (G.7)). To show that

$$\frac{n^{1/2}}{(k_n+1)^2}\mathbb{E}\left(\left(\sum_{i=(2m-2)k_n}^{s_m^n}U_{i+k_n}^n-U_i^n\right)\left(\sum_{i=s_m^n+1}^{(2m-1)k_n}U_{i+k_n}^n-U_i^n\right)\Big|\mathcal{F}_{\frac{m-1}{M_n}}\right)=o_p(n^{-1/4}),$$

we first evaluate

$$\begin{aligned}&\frac{n^{1/2}}{(k_n+1)^2}\left|\mathbb{E}\left(\left(\sum_{i=(2m-2)k_n}^{s_m^n}U_{i+k_n}^n\right)\left(\sum_{j=s_m^n+1}^{(2m-1)k_n}U_{j+k_n}^n\right)\Big|\mathcal{F}_{\frac{m-1}{M_n}}\right)\right|\\ \le&\frac{n^{1/2}}{(k_n+1)^2}\sum_{i=(2m-2)k_n}^{s_m^n}\sum_{j=s_m^n+1}^{(2m-1)k_n}\left|\mathbb{E}\left(U_{i+k_n}^nU_{j+k_n}^n\Big|\mathcal{F}_{\frac{m-1}{M_n}}\right)\right|.\end{aligned}$$

Now apply (G.5) and by the fact that $v>4$, we have

$$\begin{aligned}\sum_{i=(2m-2)k_n}^{s_m^n}\sum_{j=s_m^n+1}^{(2m-1)k_n}\mathbb{E}\left(\left|\mathbb{E}\left(U_{i+k_n}^nU_{j+k_n}^n\Big|\mathcal{F}_{\frac{m-1}{M_n}}\right)\right|\right)&\overset{(G.5)}{\le}\sum_{i=(2m-2)k_n}^{s_m^n}\sum_{j=s_m^n+1}^{(2m-1)k_n}C(j-i)^{-v/2}\\&\le C\sum_{\ell=1}^{r_n}\ell^{1-\frac{v}{2}}\le C.\end{aligned}$$

Similarly, we can prove that the other three cross products have the same order. It is also easy to verify that

$$\frac{\sqrt{n}}{(k_n+1)^2}\mathbb{E}\left(\sum_{i=(2m-2)k_n}^{s_m^n}(W_{i+k_n}^n-W_i^n)\sum_{j=s_m^n+1}^{(2m-1)k_n}(W_{j+k_n}^n-W_j^n)\right)=O(r_n/\sqrt{n}).$$

Now (G.7) is proved and consequently (G.3) follows from (G.4) and (G.7).

2. Next, we prove that

$$\mathbb{E}\left(\left(\overline{\beta}_{r_n,m}^n\right)^2\Big|\mathcal{F}_{\frac{m-1}{M_n}}\right)-\mathbb{E}\left(\left(\overline{\beta}_{r_n,m}^n\right)^2\Big|\sigma_{\frac{m-1}{M_n}}\right)=o_p(n^{-1/4}).\tag{G.8}$$

For this purpose, we note that

$$\frac{(k_n+1)^2}{\sqrt{n}}\left|\mathbb{E}\left(\left(\overline{\beta}^n_{r_n,m}\right)^2\Big|\mathcal{F}_{\frac{m-1}{M_n}}\right)-\mathbb{E}\left(\left(\overline{\beta}^n_{r_n,m}\right)^2\Big|\sigma_{\frac{m-1}{M_n}}\right)\right|$$
$$=\left|\left(\mathbb{E}\left(\left(\sum_{i=s^n_m+1}^{(2m-1)k_n}\left(U^n_{i+k_n}-U^n_i\right)\right)^2\Big|\mathcal{F}_{\frac{m-1}{M_n}}\right)-\mathbb{E}\left(\left(\sum_{i=s^n_m+1}^{(2m-1)k_n}\left(U^n_{i+k_n}-U^n_i\right)\right)^2\right)\right)\right|.$$

Applying again the JS-Lemma, we find that

$$\mathbb{E}\left(\left(\overline{\beta}^n_{r_n,m}\right)^2\Big|\mathcal{F}_{\frac{m-1}{M_n}}\right)-\mathbb{E}\left(\left(\overline{\beta}^n_{r_n,m}\right)^2\Big|\sigma_{\frac{m-1}{M_n}}\right)=O_p(r_n^{-v}),$$

whence (G.8) follows from (G.2).

3. Finally, we show that

$$\mathbb{E}\left(\left(\overline{\beta}_{r_n,m}\right)^2\Big|\sigma_{\frac{m-1}{M_n}}\right)=\left(\frac{2c}{3}\sigma^2_{\frac{m-1}{M_n}}+\frac{2}{c}\sigma^2_U\right)+o_p(n^{-1/4}). \tag{G.9}$$

This follows from the following equalities, which are straightforward:

$$\mathbb{E}\left(\left(\frac{n^{1/4}}{k_n+1}\left(\sum_{i=s^n_m+1}^{(2m-1)k_n}\sigma_{\frac{m-1}{M_n}}\left(W^n_{i+k_n}-W^n_i\right)\right)\right)^2\Big|\sigma_{\frac{m-1}{M_n}}\right)=\frac{2c}{3}\sigma^2_{\frac{m-1}{M_n}}+o_p(n^{-1/4}),$$

$$\mathbb{E}\left(\left(\frac{n^{1/4}}{k_n+1}\left(\sum_{i=s^n_m+1}^{(2m-1)k_n}\left(U^n_{i+k_n}-U^n_i\right)\right)\right)^2\Big|\sigma_{\frac{m-1}{M_n}}\right)=\frac{2\sigma^2_U}{c}+o_p(n^{-1/4}).$$

Now (G.1) follows from (G.3), (G.8) and (G.9), and the proof is complete. □

Lemma G.2. *Let*

$$L_n:=n^{-1/4}\sum_{m=1}^{M_n}\left((\beta^n_m)^2-\mathbb{E}\left((\beta^n_m)^2\Big|\mathcal{F}_{\frac{m-1}{M_n}}\right)\right).$$

Then, we have the following stable convergence in law:

$$L_n\xrightarrow{\mathcal{L}-s}\sqrt{\frac{1}{c}}\int_0^1\left(\frac{2c}{3}\sigma^2_s+\frac{2\sigma^2_U}{c}\right)\mathrm{d}W'_s, \tag{G.10}$$

where W' is a standard Wiener process independent of $\mathcal{F}$.

Proof. Let $\theta^n_m:=n^{-1/4}\left((\beta^n_m)^2-\left(\frac{2c}{3}\sigma^2_{\frac{m-1}{M_n}}+\frac{2}{c}\sigma^2_U\right)\right)$. Then,

$$L_n=\sum_{m=1}^{M_n}\theta_m+o_p(1),$$

by Lemma G.1. We also have

$$\sum_{m=1}^{M_n} \mathbb{E}\left(\theta_m^n \,\middle|\, \mathcal{F}_{\frac{m-1}{M_n}}\right) \xrightarrow{\mathbb{P}} 0, \tag{G.11}$$

again by Lemma G.1 and

$$\begin{aligned}\sum_{m=1}^{M_n} \mathbb{E}\left((\theta_m^n)^2 \,\middle|\, \mathcal{F}_{\frac{m-1}{M_n}}\right) =& \frac{1}{2cM_n} \sum_{m=1}^{M_n} \mathbb{E}\left((\beta_m^n)^4 \,\middle|\, \mathcal{F}_{\frac{m-1}{M_n}}\right) + \frac{1}{2cM_n} \sum_{m=1}^{M_n} \left(\frac{2c}{3}\sigma^2_{\frac{m-1}{M_n}} + \frac{2\sigma_U^2}{c}\right)^2 \\ & - \frac{1}{cM_n} \sum_{m=1}^{M_n} \mathbb{E}\left((\beta_m^n)^2 \,\middle|\, \mathcal{F}_{\frac{m-1}{M_n}}\right) \left(\frac{2c}{3}\sigma^2_{\frac{m-1}{M_n}} + \frac{2\sigma_U^2}{c}\right).\end{aligned}$$

Now it follows from (E.16) and a Riemann approximation that

$$\sum_{m=1}^{M_n} \mathbb{E}\left((\theta_m^n)^2 \,\middle|\, \mathcal{F}_{\frac{m-1}{M_n}}\right) \xrightarrow{\mathbb{P}} \frac{1}{c} \int_0^1 \left(\frac{2c}{3}\sigma_u^2 + \frac{2\sigma_U^2}{c}\right)^2 \mathrm{d}u. \tag{G.12}$$

Next, denote $\overline{\Delta_m^n V} = V^n_{(2m-1)k_n} - V^n_{2(m-1)k_n}$, for any process V. We will show that

$$\sum_{m=1}^{M_n} \mathbb{E}\left(\theta_m^n \overline{\Delta_m^n N} \,\middle|\, \mathcal{F}^n_{2(m-1)k_n}\right) \xrightarrow{\mathbb{P}} 0, \tag{G.13}$$

for any bounded martingale N defined on the same probability space, where $\mathcal{F}_i^n = \mathcal{F}_{i/n}$ whence $\mathcal{F}^n_{2(m-1)k_n} = \mathcal{F}_{\frac{m-1}{M_n}}$. To complete the proof, it is convenient to specify the respective probability spaces as follows. (We can always extend the probability space — whether the noise process and the efficient price process are defined on the same probability space or not — see e.g., the detailed arguments in Jacod et al. (2017).) The efficient price process lives on $(\Omega', \mathcal{F}', (\mathcal{F}'_t)_{t\in\mathbb{R}}, \mathbb{P}')$. The noise process $(U_i)_{i\in\mathbb{N}}$ is defined on $(\Omega'', \mathcal{F}'', (\mathcal{F}''_i)_{i\in\mathbb{N}}, \mathbb{P}'')$, where the filtration is defined by $\mathcal{F}''_i = \sigma(U_j, j \le i, j \in \mathbb{N})$ and $\mathcal{F}'' = \bigvee_{i\in\mathbb{N}} \mathcal{F}''_i$. Let

$$\Omega = \Omega' \times \Omega'', \quad \mathcal{F} = \mathcal{F}' \otimes \mathcal{F}'', \quad \mathbb{P}\left(\mathrm{d}\omega', \mathrm{d}\omega''\right) = \mathbb{P}'(\mathrm{d}\omega')\, \mathbb{P}''\left(\mathrm{d}\omega''\right). \tag{G.14}$$

For a realization of observation times $(t_i^n)_{0\le i\le n}$, we introduce $\mathcal{F}_i^n = \mathcal{F}'_{t_i^n} \otimes \mathcal{F}''_i$.

According to Jacod et al. (2009) and the proof of Theorem IX 7.28 of Jacod and Shiryaev (2003) it suffices to consider martingales in $\mathcal{N}^0$ or $\mathcal{N}^1$, where $\mathcal{N}^0$ is the set of all bounded martingales on $(\Omega', \mathcal{F}', \mathbb{P}')$, orthogonal to W, and $\mathcal{N}^1$ is the set of all martingales having a limit $N_\infty = f(Y_{t_1}, \ldots, Y_{t_q})$, where f is any bounded Borel function on $\mathbb{R}^q$, $t_1 < \ldots < t_q$ and $q \ge 1$.

First, let $N \in \mathcal{N}^0$ and let $\widetilde{\mathcal{F}}'_t = \bigcap_{s>t} \mathcal{F}'_s \otimes \mathcal{F}''$. Then, for any $t > \frac{m-1}{M_n}$, $\overline{\theta}_m^n(t) := \mathbb{E}\left(\theta_m^n \,\middle|\, \widetilde{\mathcal{F}}'_t\right)$, conditional on $\sigma_{\frac{m-1}{M_n}}$, is a martingale with respect to the filtration generated by $\{W_t - W_{\frac{m-1}{M_n}} | t > \frac{m-1}{M_n}\}$. By the martingale representation theorem, we have $\overline{\theta}_m^n(t) = \overline{\theta}_m^n(\frac{m-1}{M_n}) + \int_{\frac{m-1}{M_n}}^t \gamma_u \mathrm{d}W_u$ for some predictable process γ. Now it follows from the orthogonality of W, N and the martingale property of N that

$$\mathbb{E}\left(\theta_m^n \overline{\Delta_m^n N} \,\middle|\, \widetilde{\mathcal{F}}'_{\frac{m-1}{M_n}}\right) = \mathbb{E}\left(\left(\theta_m^n - \overline{\theta}_m^n\left(\frac{m-1}{M_n}\right)\right) \overline{\Delta_m^n N} + \overline{\theta}_m^n\left(\frac{m-1}{M_n}\right) \overline{\Delta_m^n N} \,\middle|\, \widetilde{\mathcal{F}}'_{\frac{m-1}{M_n}}\right) = 0,$$

which leads to

$$\mathbb{E}\left(\theta_m^n \overline{\Delta_m^n N} \,\middle|\, \mathcal{F}^n_{2(m-1)k_n}\right) = 0, \tag{G.15}$$

since $\mathcal{F}_t \subset \widetilde{\mathcal{F}}'_t$.

Next, assume that $N \in \mathcal{N}^1$. It can be shown (see Jacod et al. (2009)) that there exists some $\hat{f}_t$ such that $t \in [t_l, t_{l+1})$, $N_t = \hat{f}_t(Y_{t_0}, Y_{t_1}, \dots, Y_{t_l})$ with $t_0 = 0, t_{q+1} = \infty$, and such that it is measurable in $(Y_{t_1}, \dots, Y_{t_l})$. Hence, $\overline{\Delta_m^n N} = 0$ if it does not cover any of the points $t_1, \dots, t_{q+1}$. But such intervals (to compute $\overline{\Delta_m^n N}$) that contain any of $t_1, \dots, t_{q+1}$ are at most finite in number. Furthermore, by the boundedness of N and the conditional Cauchy-Schwarz inequality, we have the following:

$$\mathbb{E}\left(\left|\theta_m^n \overline{\Delta_m^n N}\right| \,\middle|\, \mathcal{F}^n_{2(m-1)k_n}\right) \le \sqrt{\mathbb{E}\left((\theta_m^n)^2 \,\middle|\, \mathcal{F}^n_{2(m-1)k_n}\right)}\sqrt{\mathbb{E}\left((\overline{\Delta_m^n N})^2 \,\middle|\, \mathcal{F}^n_{2(m-1)k_n}\right)} = O_p(n^{-1/4}).$$

Now (G.13) follows since there are at most finitely many such intervals.

The following is also trivial:

$$\mathbb{E}\left(\theta_m^n \overline{\Delta_m^n W} \,\middle|\, \mathcal{F}^n_{2(m-1)k_n}\right) = 0, \tag{G.16}$$

since θ_m^n is an even functional of U and W and (U, W) are distributed symmetrically.

From (E.19), we know that $(\theta_m^n)^2 \mathbf{1}_{\{|\theta_m^n| > \varepsilon\}} = o_p(n^{-1/2})$ for any $\varepsilon > 0$. We then have

$$\sum_{m=1}^{M_n} \mathbb{E}\left((\theta_m^n)^2 \mathbf{1}_{\{|\theta_m^n| > \varepsilon\}} \,\middle|\, \mathcal{F}^n_{2(m-1)k_n}\right) \xrightarrow{\mathbb{P}} 0. \tag{G.17}$$

Now the proof is complete in view of (G.11), (G.12), (G.13), (G.16) and (G.17), and Theorem IX 7.28 of Jacod and Shiryaev (2003). □

Lemma G.3. *We have that*

$$\sum_{m=1}^{M_n} \left(\bar{Y}_m^n\right)^2 - \frac{1}{\sqrt{n}} \sum_{m=1}^{M_n} \left(\beta_m^n\right)^2 = o_p(n^{-1/4}). \tag{G.18}$$

Proof. Denote

$$\widetilde{Y}_m^n = \sigma_{\frac{m-1}{M_n}} \overline{W}_m^n + \bar{U}_m^n. \tag{G.19}$$

Then,

$$\mathbb{E}\left(\left|\sum_{m=1}^{M_n} \left(\bar{Y}_m^n\right)^2 - \frac{1}{\sqrt{n}} \sum_{m=1}^{M_n} \left(\beta_m^n\right)^2\right|\right) \le \sum_{m=1}^{M_n} \sqrt{\mathbb{E}\left(\left(\bar{Y}_m^n - \widetilde{Y}_m^n\right)^2\right)}\sqrt{\mathbb{E}\left(\left(\bar{Y}_m^n + \widetilde{Y}_m^n\right)^2\right)}.$$

Since $\sqrt{\mathbb{E}\left(\left(\bar{Y}_m^n + \widetilde{Y}_m^n\right)^2\right)} = O(n^{-1/4})$, the result follows if

$$\sum_{m=1}^{M_n} \sqrt{\mathbb{E}\left(\left(\bar{Y}_m^n - \widetilde{Y}_m^n\right)^2\right)} \to 0. \tag{G.20}$$

But this follows directly from Lemma 7.8 in Barndorff-Nielsen et al. (2006). □

Proof of Theorem 4.2. Now the proof of Theorem 4.2 is complete in view of (G.10) and (G.18), and our consistency result in (22). □

H Simulation Study under Stochastic Volatility

In this section, we provide additional simulation results in the presence of stochastic volatility. We simulate the microstructure noise process employing various combinations of dependence structure and sampling frequency.

We assume that the efficient log-price is generated by the following dynamics:

$$\mathrm{d}X_t = -\delta(X_t - \mu_1)\mathrm{d}t + \sigma_t \mathrm{d}W_t, \qquad \mathrm{d}\sigma_t^2 = \kappa\left(\mu_2 - \sigma_t^2\right)\mathrm{d}t + \gamma\sigma_t \mathrm{d}B_t,$$

where B is a standard Brownian motion and its quadratic covariation with the standard Brownian motion W is ϱt. We set the parameters as follows: $\delta = 0.5$, $\mu_1 = 1.6, \kappa = 5/252, \mu_2 = 0.04/252, \gamma = 0.05/252$, and $\varrho = -0.5$. We employ the same noise process as in (34). We set $\mathbb{E}\big(V^2\big) = 1.9 \times 10^{-7}$, and $\mathbb{E}\big(\epsilon^2\big) = 1.3 \times 10^{-7}$. Note that these parameters are slightly different from those in Section 6, which were based on Aït-Sahalia et al. (2011). They are chosen to mimic the results of our empirical studies.

Figure H.1 presents the estimates of the second moments of noise. Clearly, the bias correction can be important, potentially yielding significantly improved results. Turning to the estimation of the integrated volatility using $\widehat{\mathrm{IV}}_{\mathrm{step1}}$, $\widehat{\mathrm{IV}}_n$, $\widehat{\mathrm{IV}}_{\mathrm{step2}}$ and $\widehat{\mathrm{IV}}_{\mathrm{step3}}$, we observe from Table H.1 similar results under stochastic volatility as in our previous simulation studies that assumed deterministic volatility: the two-step estimators of the integrated volatility have much smaller bias and only slightly larger standard deviations when noise is dependent. One more iteration of bias corrections further improves the performance when noise is serially correlated. They also deliver reliable estimates when noise turns out to be independent.

I Empirical Study of Transaction Data for General Electric

We collect 2,721,475 transaction prices of General Electric (GE) over the month January 2011. On average there are 5.8 observations per second. In contrast to the analysis of Citigroup transaction prices in Sections 7.2 and 7.3, bias correction plays a very pronounced role here. Despite the high data frequency, the finite sample bias can be very significant if the underlying noise-to-signal ratio is small

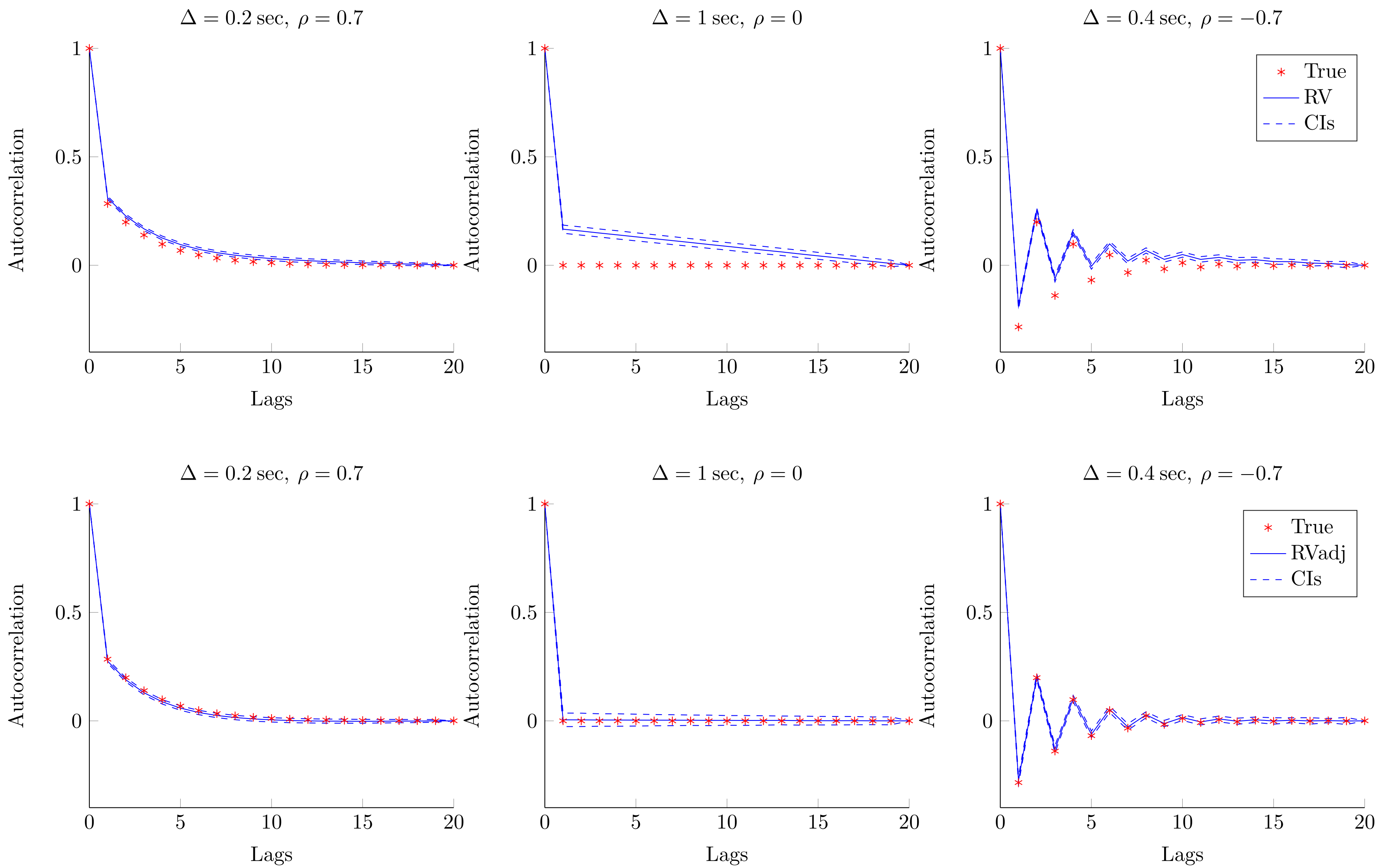

Figure H.1: Realized volatility (RV) and bias corrected realized volatility (RVadj) estimators of the autocorrelations of noise in the presence of stochastic volatility against the number of lags j, averaged over 1,000 simulated samples. Top panel: RV estimators of the autocorrelations of noise (solid). Bottom panel: RV estimators of the autocorrelations of noise with finite sample bias correction (solid). The true autocorrelations are displayed in stars and the 95% simulated confidence intervals are dashed. From the left to the right, the three combinations of ρ, Δ_n mimic transaction time sampling, regular time sampling (at 1 sec scale), and tick time sampling. The tuning parameters are set as follows: $j_n = 20$, $i_n = 10$ and $c = 0.2$.

ρ, Δ_n	$\rho = 0.7,\ \Delta_n = 0.2$ sec	$\rho = 0,\ \Delta_n = 1$ sec	$\rho = -0.7,\ \Delta_n = 0.4$ sec
$\widehat{\mathrm{IV}}_{\mathrm{step1}} - \int_0^1 \sigma_t^2 \mathrm{d}t$	5.02e-5 (1.10e-5)	4.33e-7 (1.32e-5)	-1.50e-5 (9.97e-6)
$\widehat{\mathrm{IV}}_n - \int_0^1 \sigma_t^2 \mathrm{d}t$	-1.64e-5 (1.09e-5)	-7.82e-5 (1.18e-5)	-3.17e-5 (9.77e-6)
$\widehat{\mathrm{IV}}_{\mathrm{step2}} - \int_0^1 \sigma_t^2 \mathrm{d}t$	4.32e-6 (1.20e-5)	9.94e-7 (1.79e-5)	-3.15e-6 (1.17e-5)
$\widehat{\mathrm{IV}}_{\mathrm{step3}} - \int_0^1 \sigma_t^2 \mathrm{d}t$	-2.32e-7 (1.21e-5)	1.27e-6 (2.06e-5)	-8.05e-7 (1.21e-5)

Table H.1: Estimation of the integrated volatility in the presence of stochastic volatility and under various combinations of noise dependence structure and sampling frequency. We report the means of the bias of the four integrated volatility estimators: $\widehat{\mathrm{IV}}_{\mathrm{step1}} - \int_0^1 \sigma_t^2 \mathrm{d}t, \widehat{\mathrm{IV}}_n - \int_0^1 \sigma_t^2 \mathrm{d}t$, $\widehat{\mathrm{IV}}_{\mathrm{step2}} - \int_0^1 \sigma_t^2 \mathrm{d}t$ and $\widehat{\mathrm{IV}}_{\mathrm{step3}} - \int_0^1 \sigma_t^2 \mathrm{d}t$, based on 1,000 simulations with standard deviations between parentheses. From the left to the right, the three combinations of ρ, Δ_n mimic transaction time sampling, regular time sampling (at 1 sec scale), and tick time sampling. The tuning parameters are set as follows: $j_n = 20$, $i_n = 10$ and $c = 0.2$.

(recall Remark 3.3). This is indeed the case as Figure I.1 reveals: compared with Citigroup, the data frequency of the General Electric sample is typically lower but the noise-to-signal ratio is also (much) smaller. While the data frequency is immediately available, the noise-to-signal ratio is latent. Therefore, one should always be wary to rely solely on asymptotic theory in practice.

The top panel of Figure I.2 shows that both the realized volatility (RV) and local averaging (LA) estimators indicate that the noise is strongly autocorrelated, while the bias corrected realized volatility (BCRV) estimator reveals that the noise is only weakly dependent. Such a pattern also appears in our simulation study, where we have seen that it is the finite sample bias that induces this discrepancy. The bottom panel of Figure I.2 plots two estimators of the integrated volatility, $\widehat{\mathrm{IV}}_n$ and $\widehat{\mathrm{IV}}_{\mathrm{step2}}$, to illustrate that the finite sample bias correction is particularly essential. If one would solely rely on asymptotic theory, then one would end up with much lower estimates and narrow confidence intervals that may well exclude the true values!

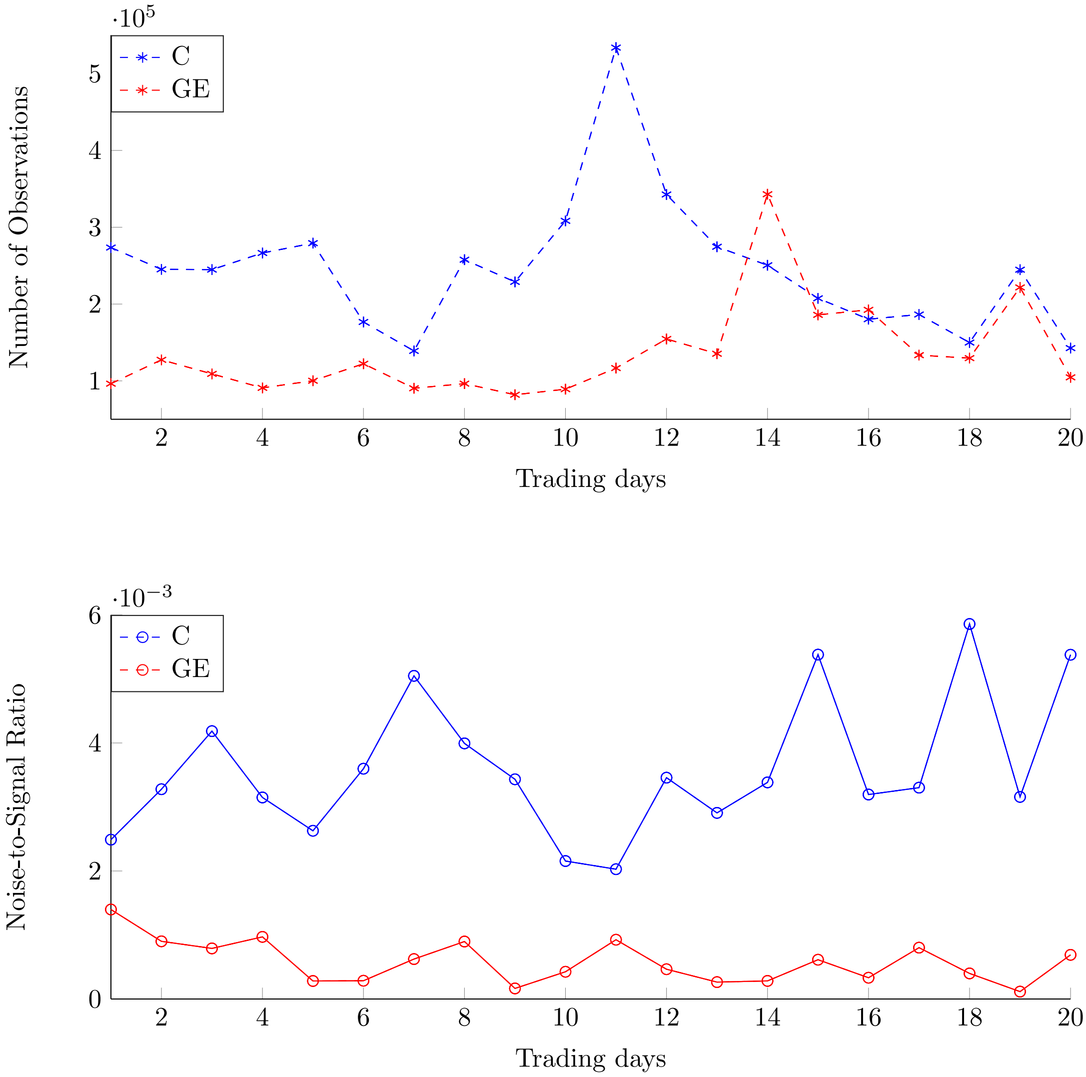

Figure I.1: Number of daily observations of transaction prices (top panel) and noise-to-signal ratio (bottom panel) for Citigroup (C) and General Electric (GE). Sample period: January, 2011, consisting of 20 trading days. In the bottom panel, the noise-to-signal ratio, $\frac{\sigma_U^2}{\int_0^1 \sigma_s^2 \mathrm{d}s}$, is estimated by $\frac{\hat{\sigma}^2_{U,\text{step2}}}{\widehat{\text{IV}}_{\text{step2}}}$, where $\hat{\sigma}^2_{U,\text{step2}}$ and $\widehat{\text{IV}}_{\text{step2}}$ are defined in (32) and (33), respectively. We set $j_n = 30$, $i_n = 15$ and $c = 0.2$.

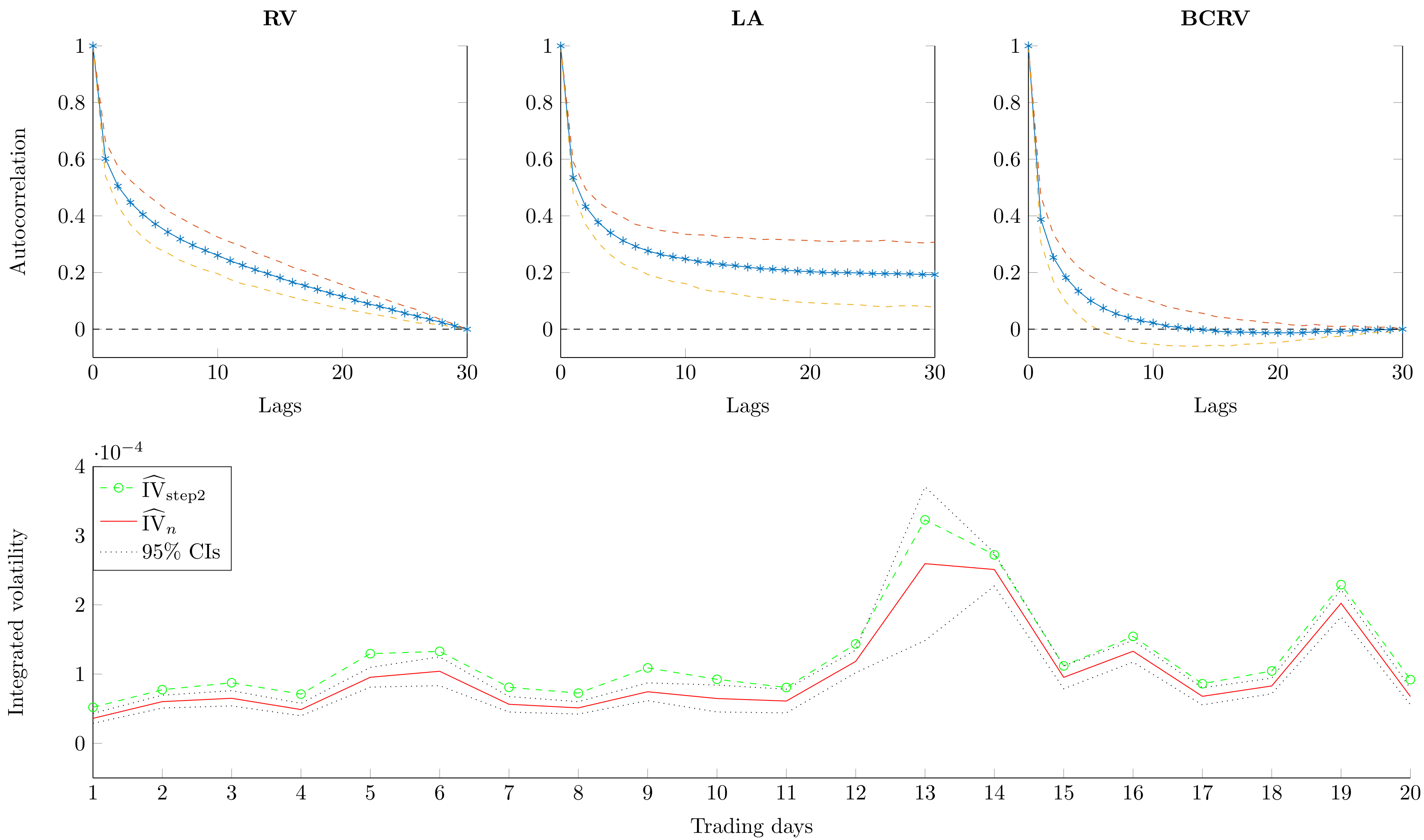

Figure I.2: Autocorrelations of noise and integrated volatility based on transaction data for General Electric (GE). Sample period: January, 2011, consisting of 20 trading days. On average there are 5.8 observations per second in the sample. Top panel: From the left to the right, we display the realized volatility (RV), local averaging (LA), and the bias corrected realized volatility (BCRV) estimators of the autocorrelations of noise against the number of lags j. The three estimators are applied to and next averaged over each of the 20 trading days. The stars indicate the means of the 20 estimates. The dashed lines are 2 standard deviations away from the mean. Bottom panel: Estimation of the integrated volatility. The estimators $\widehat{\mathrm{IV}}_{\mathrm{step2}}$ and $\widehat{\mathrm{IV}}_n$ are given by (33) and (23). The asymptotic confidence intervals (CIs) are based on the limit distribution in Theorem 4.2. We set $j_n = 30$, $i_n = 15$ and $c = 0.2$.

References

Aït-Sahalia, Y., P. A. Mykland, and L. Zhang (2011): "Ultra high frequency volatility estimation with dependent microstructure noise," *Journal of Econometrics*, 160, 160–175.

Barndorff-Nielsen, O. E., S. E. Graversen, J. Jacod, M. Podolskij, and N. Shephard (2006): "A central limit theorem for realised power and bipower variations of continuous semimartingales," in *From Stochastic Calculus to Mathematical Finance*, ed. by Y. Kabanov, R. Lipster, and J. Stoyanov, Springer.

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.

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.

Jacod, J. and P. E. Protter (2011): *Discretization of Processes*, vol. 67, Springer Science & Business Media.

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

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

White, H. (2000): *Asymptotic Theory for Econometricians*, Academic Press.