*Econometrica Supplementary Material*

# SUPPLEMENT TO "STATISTICAL PROPERTIES OF MICROSTRUCTURE NOISE"

(*Econometrica*, Vol. 85, No. 4, July 2017, 1133–1174)

JEAN JACOD
Institut de Mathématiques de Jussieu, CNRS–UMR7586 and Université Pierre et Marie Curie–P6

YINGYING LI
Department of Information Systems, Business Statistics and Operations Management and Department of Finance, Hong Kong University of Science and Technology

XINGHUA ZHENG
Department of Information Systems, Business Statistics and Operations Management, Hong Kong University of Science and Technology

## APPENDIX B: MORE RESULTS ON NUMERICAL STUDIES

### B.1. *Supplementary to Section 4.1*

IN THIS SUBSECTION, WE REPORT RESULTS ABOUT AUTOCOVARIANCES; those about autocorrelations are given in the main article. Figure 14 compares the estimates of scaled autocovariances based on Theorem 3.1 with the infeasible estimates based on the noise process and the theoretical values. The estimates are based on one simulated path. More specifically, we use red dashed curve to report the estimates of the scaled autocovariances based on Theorem 3.1, namely,

$$\widehat{R}(j)^n_t := U(\mathbf{j})^n_t / N^n_t \quad \text{for } \mathbf{j} = (0, j), j = 0, 1, \dots, \tag{B.1}$$

as in (3.9) with $t = 1$; blue dotted curve to report the *infeasible* estimates based on the noise process $\varepsilon^n_i = \gamma_{T(n,i)}\chi_i$, namely, the autocovariances estimated from $(\varepsilon^n_i)$ (which are not observed in practice); black solid curve to report the theoretical values, that is, $R(j)_1 := \mathbf{R}(\mathbf{j})_1$ as in (3.4). The theoretical values $R(j)_1$ are, with $\psi_0 = 1$ and $\psi_j$, $j = 1, 2, \dots$ as in (4.39),

$$R(j)_1 = \sigma^2 \sum_{i=0}^{M-j} \psi_i \psi_{i+j} \int_0^1 \gamma^2_s \alpha_s \, ds, \quad \text{for } j = 0, 1 \dots, M. \tag{B.2}$$

Figure 14 shows that for autocovariances, our estimates are also comparable to the infeasible estimates based on the noise process, both of which are close to the theoretical values.

To examine the CLT in Theorem 3.8, we plot the normal quantile-quantile plots of

$$\frac{\sqrt{N^n_1}}{\sqrt{\widehat{\mathcal{Z}}'(\mathbf{j}, \mathbf{j})^n_1}} \big(\widehat{R}(j)^n_1 - R(j)_1\big) \tag{B.3}$$

Jean Jacod: jean.jacod@gmail.com
Yingying Li: yyli@ust.hk
Xinghua Zheng: xhzheng@ust.hk

 DOI: 10.3982/ECTA13085

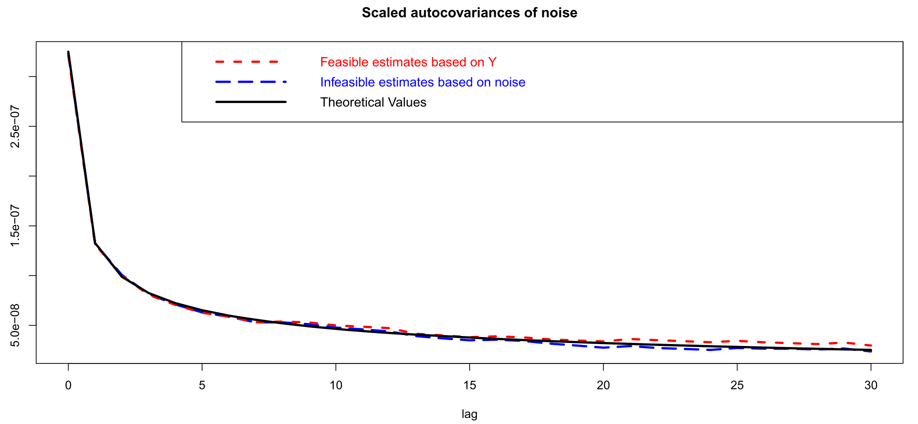


FIGURE 14.—Estimates of scaled autocovariances. The feasible estimates based on the observed prices are compared with the infeasible estimates based on unobservable noise process, and with the theoretical values.

with $\widehat{\mathcal{Z}}'(\mathbf{j},\mathbf{j})^n_1$ given by the second formula (3.32), since here $\overline{\alpha}\equiv 1$ (with $k'_n$ chosen to be 3). Analogously to Figure 2, we show the plots for lags 1, 4, and 7 in Figure 15. The normality is again supported.

## B.2. *Supplementary to Section 4.2*

### B.2.1. *Colored Noise With Rounding*

The setting considered is the same as in Section 4.1, namely, when the processes $X$, $\gamma$, and $\chi$ are specified by equations (4.38)–(4.41), *except* that the observed prices are given by

$$\widetilde{S}^n_i=\big[\exp\big(X_{T(n,i)}+\varepsilon^n_i\big)/0.01\big]\times 0.01, \tag{B.4}$$

that is, the (contaminated) observed prices in Section 4.1 further rounded to cents. A sample path of $(\widetilde{S}^n_i)$ is given in Figure 16. One can clearly see the rounding effect from Figure 16. The proportion of flat trading, namely, zero intraday (observed) returns for this

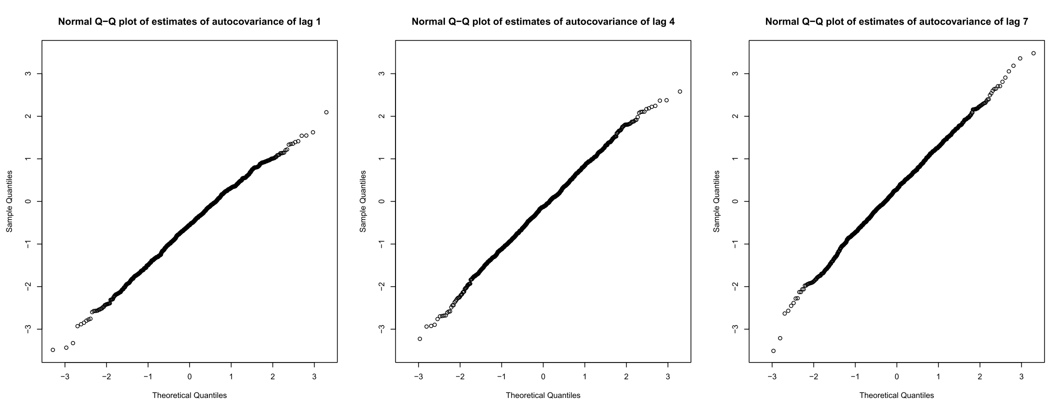


FIGURE 15.—Normal QQ-plots of (B.3) for lags 1, 4, and 7, based on 1000 replications.

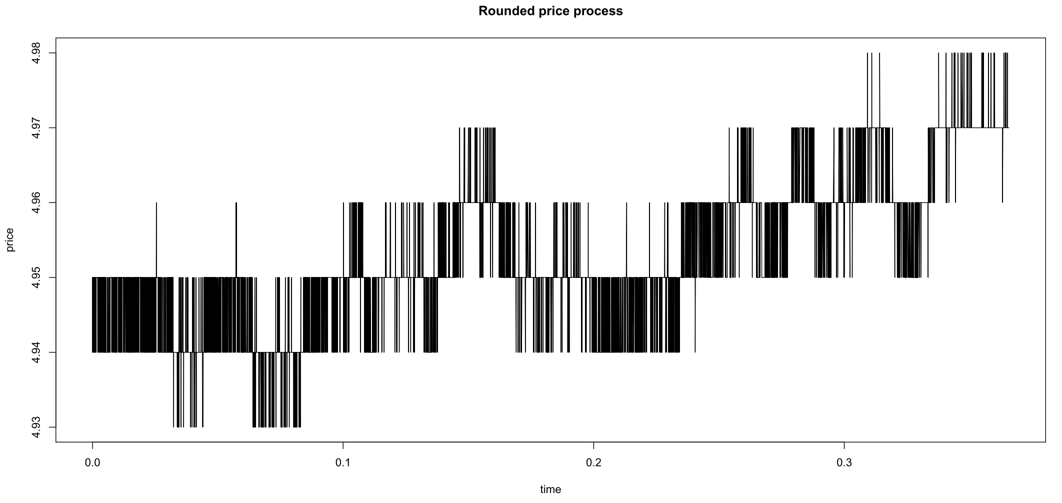


FIGURE 16.—A sample path of the rounded prices $(\widetilde{S}_i^n)$ as in (B.4).

particular sample path is 75%. Different sample paths have similar proportions of flat trading.

Under such a setting, the "noise" is

$$\widetilde{\varepsilon}_i^n = \log\bigl(\widetilde{S}_i^n\bigr) - X_{T(n,i)}.$$

Our goal is to estimate the autocovariance and autocorrelation of the noise $(\widetilde{\varepsilon}_i^n)$ based on the observations $\widetilde{S}_i^n$, or equivalently, based on

$$\widetilde{Y}_i^n := \log\bigl(\widetilde{S}_i^n\bigr).$$

To do so, we apply the same estimators as in Section 4.1, namely, $\widehat{R}(j)_1^n := U(\mathbf{j})_1^n / N_1^n$ for $\mathbf{j} = (0, j)$, $j = 0, 1, \dots$ as in (3.9) for estimating the autocovariances, and $\widehat{r}(j)_1^n$ as in (3.12) for estimating the autocorrelations. The tuning parameter $k_n$ is chosen to be 6 based on the heuristic criterion in Section 5.1.2. We compare these estimates with those estimated from the noise $(\widetilde{\varepsilon}_i^n)$. Figure 17 shows the comparison results for autocovariances based on one random sample path; those for autocorrelations are given in the main article. We see from Figure 17 that our estimates are close to the infeasible estimates based on the noise, indicating that our method works well even in the presence of rounding.

B.2.2. *White Noise With Rounding*

For the white noise with rounding case, we let the process $\chi$ be a sequence consisting of i.i.d. $N(0, \sigma_0^2)$ random variables. The proportion of flat trading is about 72%, slightly lower than the proportion in Section B.2.1 where $\chi$ is colored. The estimators are also the same as before, except that the tuning parameter $k_n$ is reduced to 3, again based on the heuristic criterion in Section 5.1.2. This smaller choice of $k_n$ is due to the weaker dependence in the noise; see also the discussions in Remark 3.6. Figure 18 reports the estimation results for autocovariances. Figure 18 shows that our estimators also work well in this case.

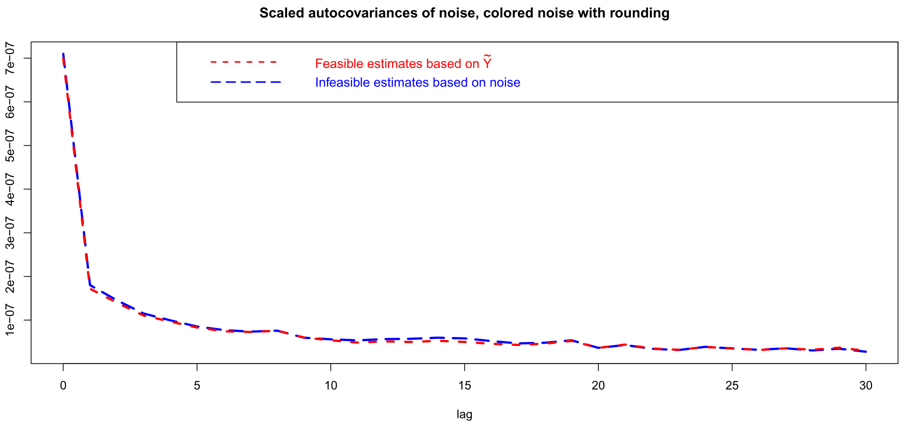


FIGURE 17.—Estimates of autocovariances for the case of colored noise with rounding. The feasible estimates based on the observed prices are compared with the infeasible estimates based on the noise.

## B.3. *Supplementary to Section 5*

### B.3.1. *Results for Citigroup (C) on May 06, 2010*

In this section, we analyze the same stock, Citigroup (C), but focus on a special day, May 06, 2010, the day on which a flash crash occurred. The number of transactions on that day is about 524,000.

*Choosing $k_n$.* Following the heuristic criterion in Section 5.1.2, $k_n$ is chosen to be 60, which yields the plot in Figure 19, analogous to Figure 6.

*Diurnal Features in the Noise.* We start with the size of noise. Figure 20 shows the estimated sizes of noise and estimated volatility during different half-hour intervals.

We see that when the volatility is unusually high, so is the size of noise.

Next, we examine the dependence during different half-hour intervals. Figure 21 gives the estimated autocorrelations during different half-hour intervals.

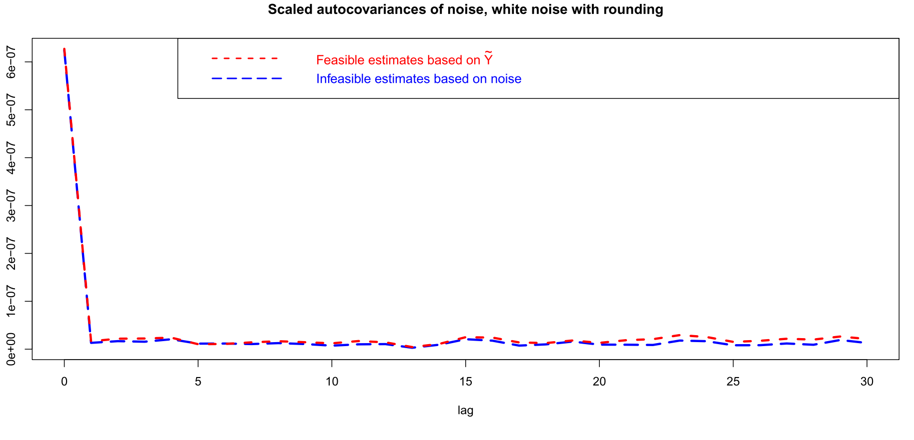


FIGURE 18.—Estimates of autocovariances for the case of white noise with rounding. The feasible estimates based on the observed prices are compared with the infeasible estimates based on the (unobservable) noise.

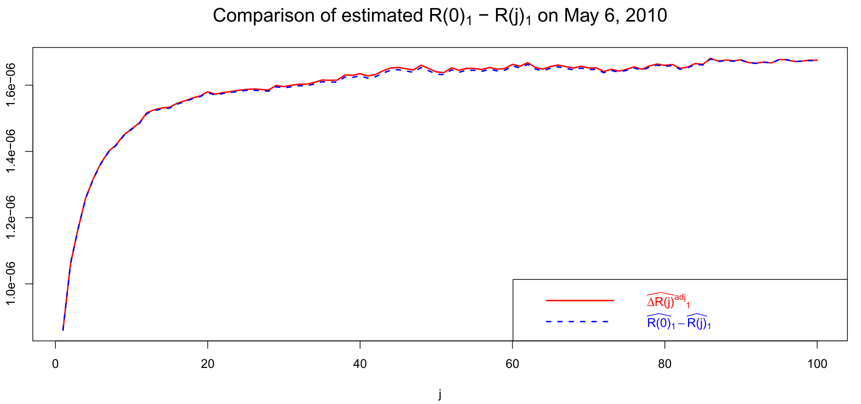


FIGURE 19.—Comparison of the two estimates of $R(0)_1 - R(j)_1$, one based on $\widehat{\Delta R(j)}^{n,\text{adj}}$, the other based on $(\widehat{R}(0)_1^n - \widehat{R}(j)_1^n)$, for Citigroup (C) on May 06, 2010.

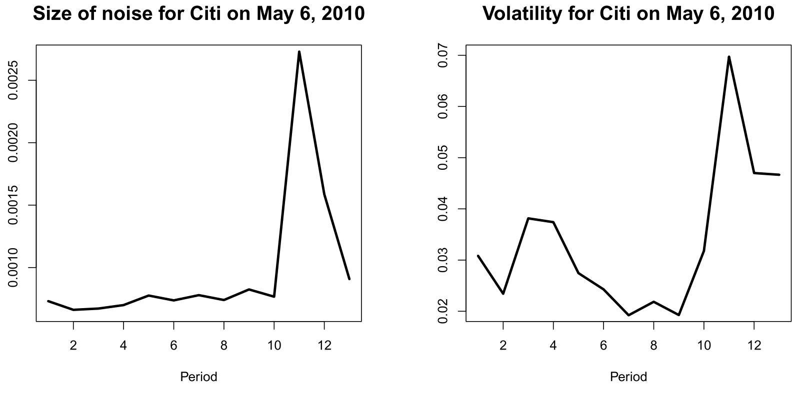


FIGURE 20.—Estimated sizes of noise and volatilities for Citigroup (C) during 13 half-hour intervals on May 06, 2010.

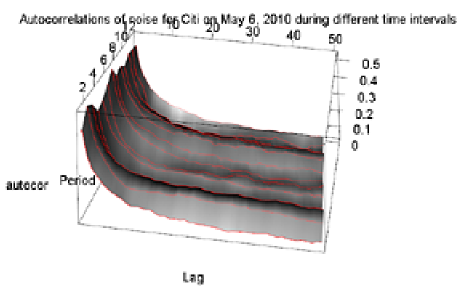


FIGURE 21.—Estimated autocorrelations of noise up to lag 50 for Citigroup (C) during different half-hour intervals on May 06, 2010.

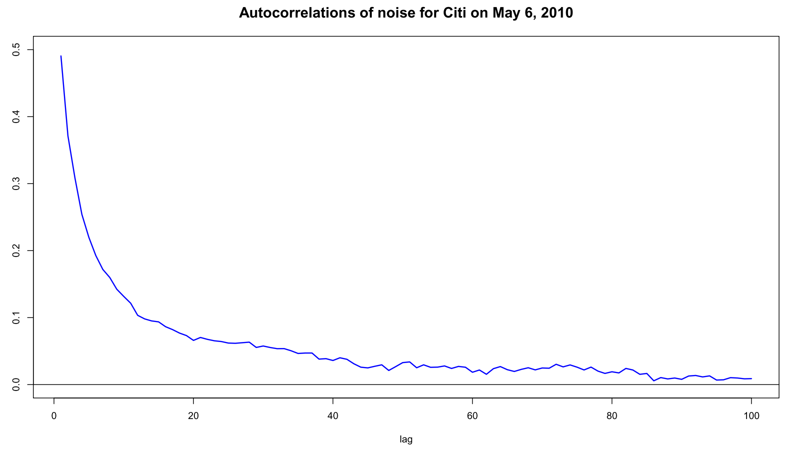


FIGURE 22.—Estimated autocorrelations of noise up to lag 100 for Citigroup (C) stock on May 06, 2010.

*Estimating the Autocorrelations of Noise.* Finally, we estimate the autocorrelations using the whole-day data. Figure 22 shows similar features to Figure 11, namely, the autocorrelations are positive and slowly decay to 0.

B.3.2. *Results for Intel (INTC)*

In this section, we briefly report the results for Intel (INTC) during the same period of January 2011. The average number of daily transactions is about 143,500.

*Choosing $k_n$.* Analogously to Figure 6, we have the plot in Figure 23 by using $k_n = 20$.

*Diurnal Features in the Noise.* We first examine the size of noise. Figure 24 shows the estimated sizes of noise during different half-hour intervals.

The average curve suggests that the noise tends to be larger towards the end of trading hours. As to the volatility, we have the plots in Figure 25.

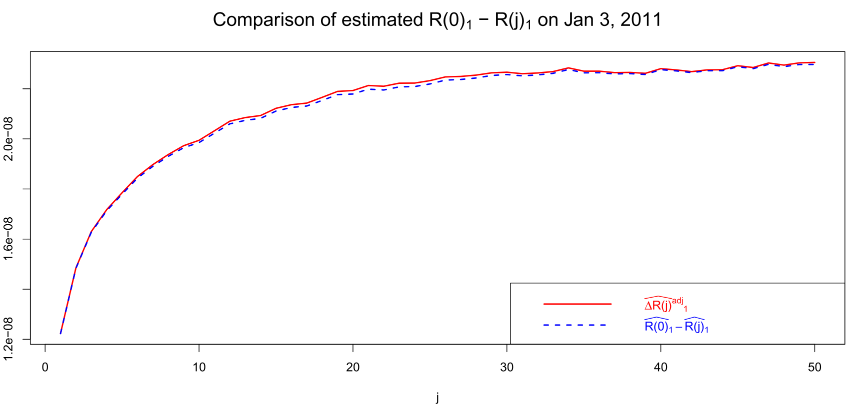


FIGURE 23.—Comparison of the two estimates of $R(0)_1 - R(j)_1$, one based on $\widehat{\Delta R(j)}^{n,\mathrm{adj}}$, the other based on $(\widehat{R}(0)_1^n - \widehat{R}(j)_1^n)$, for Intel (INTC) stock, on January 3, 2011.

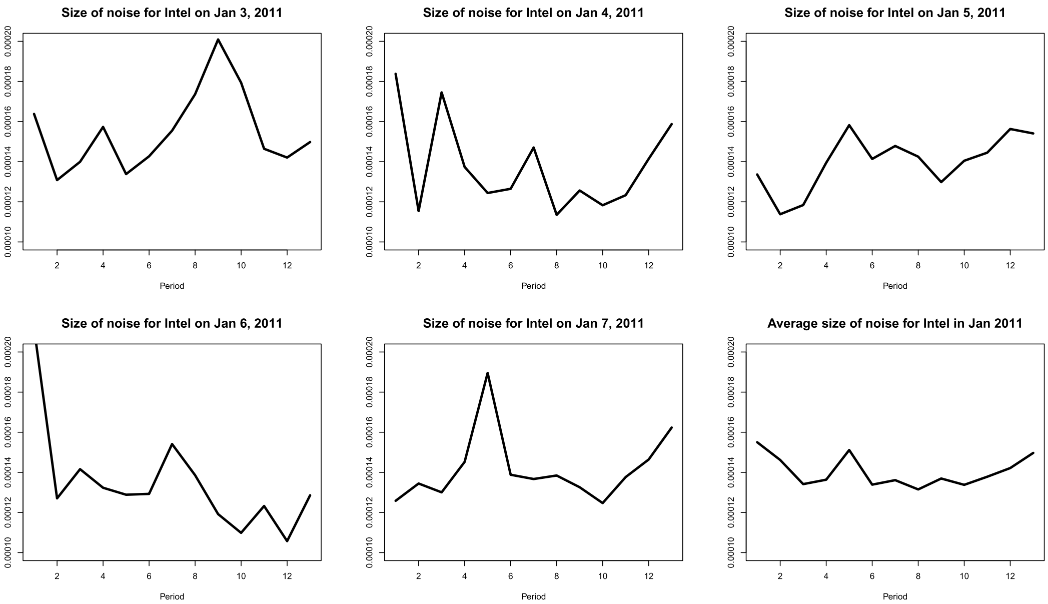


FIGURE 24.—Estimated sizes of noise for Intel (INTC) stock, during 13 half-hour intervals and for different trading days. The first five curves plot the estimates for the first five trading days in January 2011. The lower-right curve plots the sizes of noise during different half-hour intervals averaged over the 20 trading days in January 2011.

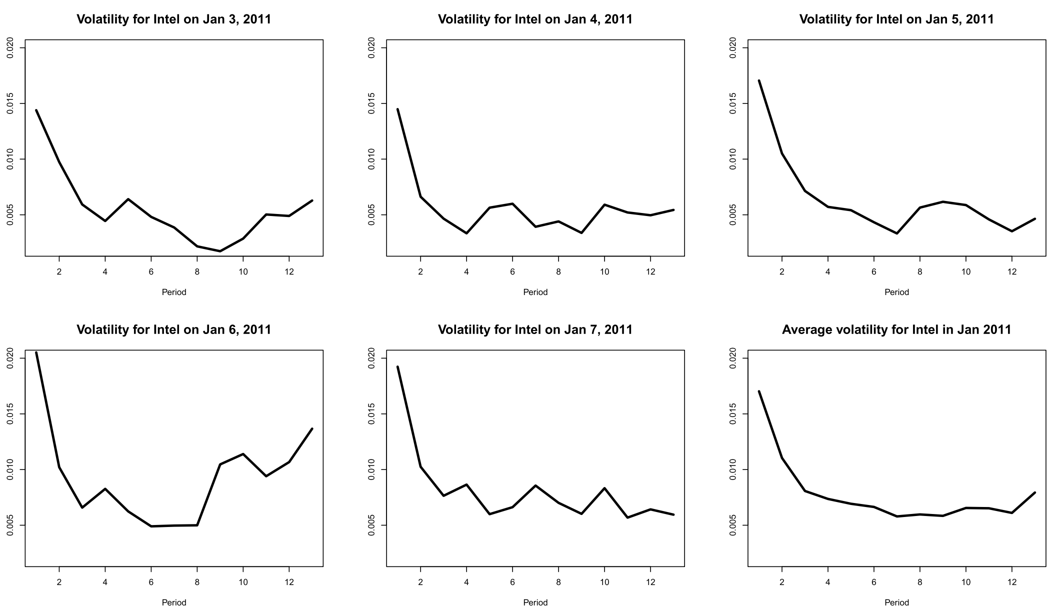


FIGURE 25.—Estimated volatilities for Intel (INTC) stock, during 13 half-hour intervals and for different trading days. The first five curves plot the estimates for the first five trading days in January 2011. The lower-right curve plots the estimated half-hour volatilities during different half-hour intervals averaged over the 20 trading days in January 2011.

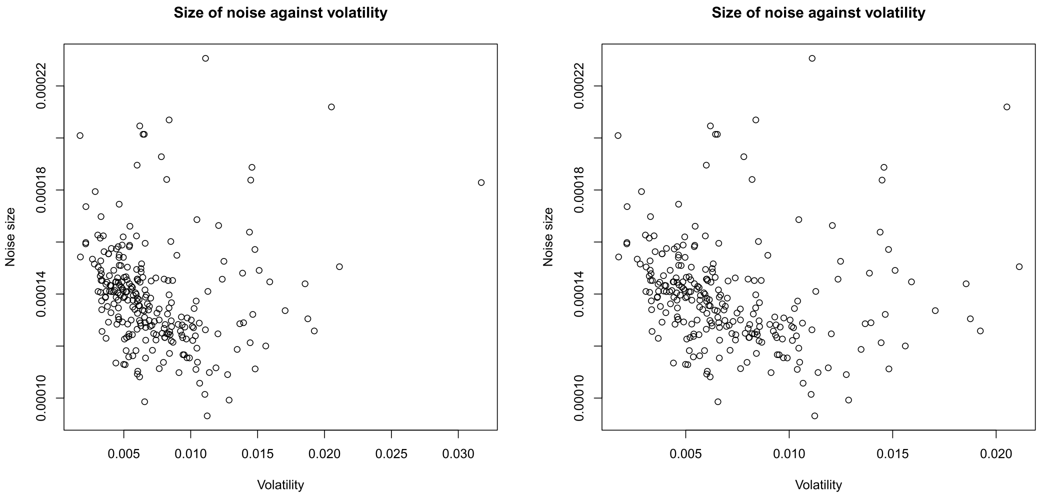


FIGURE 26.—Scatterplots of the estimated sizes of noise against volatilities during different half-hour intervals for Intel (INTC) stock, in January 2011. The right panel is a zoomed-in version of the left panel with the two outliers on its far right removed.

Again, we see a U-shaped pattern.

Finally, we plot the estimated size of noise against volatility during different half-hour intervals in Figure 26.

We observe a similar pattern to Figure 9. The correlations in two plots are $-0.07$ and $-0.13$, respectively, the former being insignificant and the latter being significant at the 5% level.

Next, we examine the dependence during different half-hour intervals. Figure 27 shows the estimated autocorrelations during different half-hour intervals.

*Estimating the Autocorrelations of Noise.* Finally, the estimated autocorrelations using the whole-day data are given in Figure 28. Again, we observe similar features to Figure 11, namely, the autocorrelations are positive and slowly decay to 0.

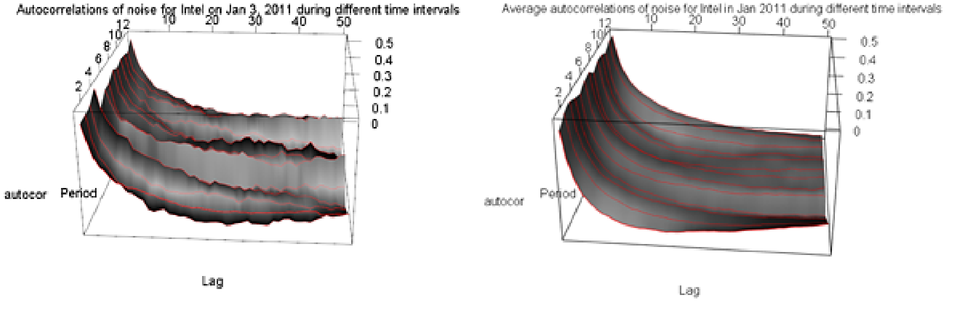


FIGURE 27.—Estimated autocorrelations of noise up to lag 50 for Intel (INTC) stock, during different half-hour intervals. Each red curve represents the estimates during one half-hour interval. Left: autocorrelations on January 3, 2011; right: autocorrelations averaged over the 20 trading days in January 2011.

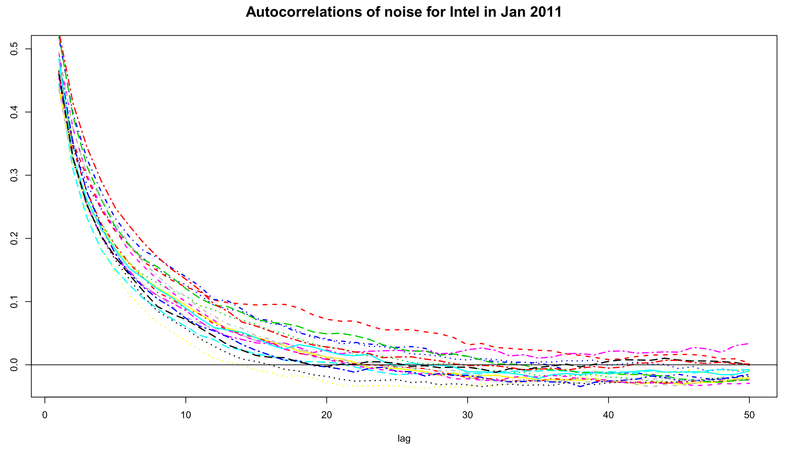


FIGURE 28.—Estimated autocorrelations of noise for Intel (INTC) stock in January 2011. Each curve is for one trading day, and plots the estimated autocorrelations of lags up to 50.

# APPENDIX C: TECHNICAL DETAILS OF THE PROOFS

## C.1. *The Sampling Scheme and the Localization Procedure*

PROOF OF LEMMA A.1: For any $i \geq 0$, we let $(\mathcal{F}_t^{n,i})$ be the smallest filtration containing $(\mathcal{F}_t)$ and for which $T(n,0), \dots, T(n,i)$ are stopping times, so $\mathcal{F}_t^{n,0} = \mathcal{F}_t$ and $\mathcal{F}_t^{n,\infty} = \mathcal{F}_t^n$. We will prove by induction on $i$ the property $(\mathrm{P}_i)$: any $(\mathcal{F}_t)$-martingale is an $(\mathcal{F}_t^{n,i})$-martingale.

The proof is based on a "concrete" description of $\mathcal{F}_t^{n,i+1}$ in terms of $\mathcal{F}_t^{n,i}$ and $\Delta(n,i+1)$. Obviously, $\mathcal{F}_t^{n,i+1}$ contains the class $\mathcal{H}_t^n$ of all sets $B \in \mathcal{F}$ such that $B \cap \{T(n,i+1) > t\} = B' \cap \{T(n,i+1) > t\}$ and $B \cap \{T(n,i+1) \leq t\} = B'' \cap \{T(n,i+1) \leq t\}$ for some $B' \in \mathcal{F}_t^{n,i}$ and $B'' \in \mathcal{F}_t^{n,i} \vee \sigma(\Delta(n,i+1))$. It is easy checking that $\mathcal{H}_t^n$ is a $\sigma$-field containing $\mathcal{F}_t^{n,i}$ and increasing with $t$ and that $T(n,i+1)$ is an $(\mathcal{H}_t^n)$-stopping time. Thus, we indeed have $\mathcal{F}_t^{n,i+1} = \mathcal{H}_t^n$.

Suppose $(\mathrm{P}_i)$ for some $i \geq 0$, and let $M$ be an $(\mathcal{F}_t)$-martingale, hence an $(\mathcal{F}_t^{n,i})$-martingale. Let $s > t \geq 0$ and $B \in \mathcal{F}_t^{n,i+1}$, with which we associate $B', B''$ as above, and observe that $B'' = \{\omega : (\omega, \Delta(n,i+1)(\omega)) \in \overline{B}\}$ for some $\in \mathcal{F}_t^{n,i} \otimes \mathcal{R}$-measurable subset $\overline{B}$ of $\Omega \times \mathbb{R}$ ($\mathcal{R}$ is the Borel $\sigma$-field of $\mathbb{R}$). We have $\mathbb{E}((M_s - M_t)1_B) = a + a'$, where

$$a = \mathbb{E}\big((M_s - M_t)1_{\{\Delta(n,i+1) \leq t - T(n,i)\}} 1_{\overline{B}}\big(\cdot, \Delta(n,i+1)\big)\big),$$
$$a' = \mathbb{E}\big((M_s - M_t)1_{B' \cap \{\Delta(n,i+1) > t - T(n,i)\}}\big).$$

Now, we apply (ii) of Assumption (O), plus the obvious fact that $\mathcal{F}_{T(n,i)}^{n,i} = \mathcal{F}_{T(n,i)}^{n,i}$. By conditioning on this $\sigma$-field, since $M_s - M_t$ is $\mathcal{F}_\infty$-measurable, and with $F$ denoting the $\mathcal{F}_{T(n,i)}^n$-conditional law of $\Delta(n,i)$, we obtain

$$a = \mathbb{E}\bigg(\int_0^{t - T(n,i)} \mathbb{E}\big((M_s - M_t)1_{\overline{B}}(\cdot, x) \mid \mathcal{F}_{T(n,i)}^{n,i}\big) F(dx)\bigg).$$

Since $1_{\overline{B}}(\cdot, x)$ is $\mathcal{F}_t^{n,i}$-measurable and $M$ is an $(\mathcal{F}_t^{n,i})$-martingale, the inner conditional expectation above vanishes, and $a = 0$. An analogous argument yields $a' = 0$, and thus $M$ is an $(\mathcal{F}_t^{n,i+1})$-martingale.

Therefore, $(\mathrm{P}_i)$ implies $(\mathrm{P}_{i+1})$ and, since $(\mathrm{P}_0)$ obviously holds true, we see that in fact $(\mathrm{P}_i)$ holds for all $i$.

Now, proving the claim is easy. First, it is enough to prove it when $M$ is a bounded $(\mathcal{F}_t)$-martingale. From what precedes it is an $(\mathcal{F}_t^{n,i})$-martingale for all $i$, hence the stopped process $M_t^i = M_{t\wedge T(n,i)}$ as well. Since $\mathcal{F}_t^n = \mathcal{F}_t^{n,i}$ in restriction to the set $\{t \le T(n,i)\}$ and $M^i$ is constant in time after $T(n,i)$, we deduce that $M^i$ is an $(\mathcal{F}_t^n)$-martingale. Since $T(n,i) \to \infty$ as $i \to \infty$, it follows that $M$ is an $(\mathcal{F}_t^n)$-local martingale, hence an $(\mathcal{F}_t^n)$-martingale because it is bounded. *Q.E.D.*

PROOF OF LEMMA A.2: By a classical localization procedure, it is no restriction to assume that $\tau_1 = \infty$ in Assumption (O), and also that $V$ is bounded.

(a) We first prove (2.4); we set

$$A_t^n = \Delta_n N_t^n, \qquad L_t^n = T\big(n, [t/\Delta_n]\big) = \sum_{i=1}^{[t/\Delta_n]} \Delta(n,i), \qquad S_t^n = \sum_{i=1}^{[t/\Delta_n]} \frac{\Delta(n,i)}{\alpha_{T(n,i-1)}}.$$

We have $S_t^n - t = \sum_{i=1}^{[t/\Delta_n]} \zeta_i^n + (\Delta_n[t/\Delta_n] - t)$, where $\zeta_i^n = \Delta(n,i)/\alpha_{T(n,i-1)} - \Delta_n$, and (2.3) implies $|\mathbb{E}(\zeta_i^n \mid \mathcal{F}_{T(n,i-1)})| \le K\Delta_n^{3/2+\kappa}$ and $\mathbb{E}((\zeta_i^n)^2 \mid \mathcal{F}_{T(n,i-1)}) \le K\Delta_n^{2+\kappa}$, whereas $\zeta_i^n$ is $\mathcal{F}_{T(n,i)}$-measurable. Therefore, we deduce $S_t^n \xrightarrow{\mathbb{P}} t$, hence $S_t^n \overset{\text{u.c.p.}}{\Longrightarrow} t$ as well, from

$$\mathbb{E}\big((S_t^n - t)^2\big) \le 2\Delta_n^2 + Kt\Delta_n^{1+\kappa} + Kt^2\Delta_n^{1/2+3\kappa/2}.$$

By the subsequence principle, for (2.4) it is enough to show that any infinite sequence $n_k$ contains a subsequence $n_k'$ such that $A_t^{n_k'} \to A_t$ for all $t$, for all $\omega$ outside a null set. Since from any subsequence one can extract a further subsequence such that $S_t^n \to t$ holds, outside a null set again, locally uniformly in time, it is enough to show that if $S_t^n(\omega) \to t$ locally uniformly in $t$ for some given $\omega$, then $A_t^n(\omega) \to A_t(\omega)$.

So below, we assume $S_t^n(\omega) \to t$ locally uniformly, and omit to mention $\omega$. The definitions of $L^n$ and $S^n$ imply $L_t^n = \int_0^t \alpha_{H_{s-}^n}\, dS_s^n$. Equation (2.4) yields $L_{t+s}^n - L_t^n \le \kappa_1(S_{t+s}^n - S_t^n)$; hence, by Ascoli's theorem, from any subsequence we can extract a further subsequence $n'$ such that $L^{n'}$ converges locally uniformly to a continuous nondecreasing limit $L$. Picking any $\varepsilon > 0$, we let $t_1 < t_2 < \cdots$ be the times at which $t \mapsto \alpha_t$ has a jump of size bigger than $\varepsilon$, and set $B_t = \bigcup_{i\ge1}((t_i - \varepsilon, t_i + \varepsilon] \cap [0,t])$ and $B_t' = [0,t] \setminus B_t$. The modulus of continuity $w_t(\rho)$ of $\alpha_s^\varepsilon = \alpha_s - \sum_{i\ge1} \Delta\alpha_{t_i} 1_{\{t_i \le s\}}$ on $[0,t]$ satisfies $\limsup_{\rho\to0} w_t(\rho) \le \varepsilon$, whereas $L_s^{n'} \to L_s$ locally uniformly, so $\limsup_{n'} \sup_{s\in B_t'} |\alpha_{L_{s-}^{n'}} - \alpha_{L_s}| \le \varepsilon$. Thus, for $n'$ large enough, $|L_t^{n'} - \int_0^t \alpha_{L_s}\, dS_s^{n'}| \le 2\varepsilon S_t^{n'} + K\int_{B_t} dS_s^{n'}$, which in turn goes to $2\varepsilon t + K\int_{B_t} ds \le K\varepsilon$. Since $\varepsilon$ is arbitrarily small, we get $L_t^{n'} - \int_0^t \alpha_{L_s}\, dS_s^{n'} \to 0$. Another application of $S_s^n \to s$ for all $s$ yields $\int_0^t \alpha_{L_s}\, dS_s^{n'} \to \int_0^t \alpha_{L_s}\, ds$. Thus $L_t = \int_0^t \alpha_{L_s}\, ds$, so $L$ is strictly increasing and its inverse $L^{-1}$ is $A$, as defined by (2.4). Therefore, $L$ is uniquely determined and the original sequence $L^n$ converges to $L = A^{-1}$.

Now, the definitions of $A_t^n$ and $L_t^n$ imply that they are right-continuous inverses one from the other, hence $A_t^n \to L_t^{-1} = A_t$, and the proof of (2.4) is complete.

For (a), it remains to deduce (A.2) from (2.4), and without loss of generality we assume $d = 1$. If $V_t^* = \sup_{s\le t} |V_s|$, we have $|H_t'^n - \int_0^t V_s\, dA_s^n| \le u_n\Delta_n V_t^*$, which goes to 0 by our assumptions on $u_n$. The property $A^n_t \xrightarrow{\mathbb{P}} A_t$ implies $\int_0^t V_s\,dA^n_s \xrightarrow{\mathbb{P}} \int_9^t V_s\,dA_s$ because $V$ is càdlàg and $A$ is continuous. We deduce the convergence (A.2) for each $t$, and the local uniform convergence easily follows, again because $A$ is continuous.

(b) Set

$$H'^n_t = H(V)^n_t = \frac{1}{\sqrt{\Delta_n}} \sum_{i=0}^{N^n_t} V_{T(n,i)}\big(\alpha_{T(n,i)}\Delta(n,i+1) - \Delta_n\big), \qquad H''^n_t = H^n_t - H'^n_t.$$

If $u_n > 0$, we have $\Delta(n,i) \le K\Delta_n^\rho$, hence $|H''^n_t| \le K u_n \Delta_n^{\rho-1/2}$ because $V$ and $\alpha$ are bounded. Then $u_n\Delta_n^{\rho-1/2} \to 0$ yields $H''^n \overset{\text{u.c.p.}}{\Longrightarrow} 0$. Henceforth, it is enough to show the convergence of $H'^n$, or equivalently suppose that $u_n \equiv 0$.

With $\zeta^{n,j}_i = \frac{1}{\sqrt{\Delta_n}} V^j_{T(n,i)}(\alpha_{T(n,i)}\Delta(n,i+1) - \Delta_n)$, we have $H^{n,j}_t = \sum_{i=0}^{N^n_t} \zeta^{n,j}_i$, and $\zeta^n_i$ is $\mathcal{F}^n_{T(n,i+1)}$-measurable. Hence by Theorem IX-7-28 of Jacod and Shiryaev (2003), it suffices to prove the following convergences in probability, for all $t > 0$ and all bounded $(\mathcal{F}_t)$-martingales $M$ and with $\Delta^n_i M = M_{T(n,i+1)} - M_{T(n,i)}$:

$$\begin{aligned}
&\sum_{i=1}^{N^n_t} \big|\mathbb{E}\big(\zeta^{n,j}_i \mid \mathcal{F}^n_{T(n,i)}\big)\big| \xrightarrow{\mathbb{P}} 0, \qquad \sum_{i=1}^{N^n_t} \mathbb{E}\big(\zeta^{n,j}_i \zeta^{n,m}_i \mid \mathcal{F}^n_{T(n,i)}\big) \xrightarrow{\mathbb{P}} \int_0^t V^j_s V^m_s \alpha_s \overline{\alpha}_s\,ds, \\
&\sum_{i=1}^{N^n_t} \mathbb{E}\big(\big|\zeta^{n,j}_i\big|^4 \mid \mathcal{F}^n_{T(n,i)}\big) \xrightarrow{\mathbb{P}} 0, \qquad \sum_{i=1}^{N^n_t} \mathbb{E}\big(\zeta^{n,j}_i \Delta^n_i M \mid \mathcal{F}^n_{T(n,i)}\big) \xrightarrow{\mathbb{P}} 0.
\end{aligned} \tag{C.1}$$

The first and third parts of (C.1) readily follow from (2.4) and $A^n_t \xrightarrow{\mathbb{P}} A_t$. Next,

$$\big|\mathbb{E}\big(\zeta^{n,j}_i \zeta^{n,m}_i \mid \mathcal{F}^n_{T(n,i)}\big) - \Delta_n \overline{\alpha}_{T(n,i)} V^j_{T(n,i)} V^m_{T(n,i)}\big| \le K\Delta_n^{1+\kappa},$$

so the second part of (C.1) follows from (A.2) applied to $\overline{\alpha}V^jV^m$. For the last part, since $\Delta(n,i+1)$ is $\mathcal{F}^n_{T(n,i)}$-conditionally independent of $\mathcal{F}_\infty$, whereas $M$ is $\mathcal{F}_\infty$-measurable and an $(\mathcal{F}^n_t)$-martingale by the previous lemma, its left side is in fact identically vanishing. *Q.E.D.*

PROOF OF LEMMA A.3: (1) Assume for a moment the existence of a localizing sequence $\theta_m$ of stopping times and, for each $m$, of a semimartingale $X(m)$, a sampling scheme $(\Delta^m(n,i) : n, i \ge 1)$, and a noise process $(\varepsilon^{m,n}_i : n \ge 1, i \ge 0)$, with which we associate $N^{m,n}_t$ and $T^m(n,i)$ as in (2.2), $\alpha^m$ and $\overline{\alpha}^m$ as in (O), and $\gamma^m$ as in (2.7), such that:

$$\begin{aligned}
&\text{(a) For each } m\text{, the family } \big[X^m, \big(\Delta^m(n,i)\big), \big(\varepsilon^{m,n}_i\big)\big] \text{ satisfies (SHON)}, \\
&\text{(b) For each } m\text{, we have } \lim_n \mathbb{P}\big(\Omega^{m,n}\big) = 1\text{, where} \\
&\qquad \Omega^{m,n} = \big\{T^m(n,i) = T(n,i) \text{ for all } i \text{ with } T(n,i) < \theta_m\big\}, \\
&\text{(c) We have } X^m_t = X_t \text{ if } t < \theta_m \text{ and } \varepsilon^{m,n}_i = \varepsilon^n_i \text{ if } T^m(n,i) = T(n,i) < \theta_m.
\end{aligned} \tag{C.2}$$

Each result of Section 3 is the convergence (in probability or stably in law) of a sequence of statistics $Z^n_t$ toward some limit $Z_t$, with $Z^n_t$ based on the restriction of $[X, (\Delta(n,i)), (\varepsilon^n_i)]$ to some time interval $[0,t]$, and the limit $Z_t$ or its $\mathcal{F}$-conditional law is always some function $g_t(\alpha, \overline{\alpha}, \gamma)$ only depending on the restriction of $\alpha, \overline{\alpha}, \gamma$ to $[0,t]$.

Now, compute the same statistic, say $Z(m)^n_t$, with $[X^m, (\Delta^m(n,i)), (\varepsilon^{m,n}_i)]$. Equation (C.2)-(a) and the assumptions of the lemma imply that $Z(m)^n_t$ converges to a limit $Z(m)_t$ which is characterized by $g_t(\alpha^m, \overline{\alpha}^m, \gamma^m)$ with the same $g_t$ as above. Equation (C.2)-(b,c) implies that $Z(m)^n_t = Z^n_t$ and also $g_t(\alpha^m, \overline{\alpha}^m, \gamma^m) = g_t(\alpha, \overline{\alpha}, \gamma)$ on the set $\Omega'^n_{m,t} = \Omega^n_m \cap \{t < \theta_m\}$, and

$$\lim_m \liminf_n \mathbb{P}\big((\Omega'^n_{m,t})^c\big) \le \lim_m \mathbb{P}(t \ge \theta_m) + \lim_m \limsup_n \mathbb{P}\big((\Omega^n_m)^c\big) = 0.$$

We deduce first that, in restriction to $\{t < \theta_m\}$, we have $Z(m)_t = Z$ in case of convergence in probability, or $Z(m)_t$ and $Z_t$ have the same $\mathcal{F}$-conditional laws. Second, it follows that indeed $Z^n_t$ converges to $Z_t$ in the appropriate sense, and the claim is proved.

Therefore, it remains to show that we can find $\theta_m$ and $[X^m, (\Delta^m(n,i)), (\varepsilon^{m,n}_i)]$ satisfying (C.2). This is achieved through several steps.

(2) By the classical localization procedure (see, e.g., Section 4.4.4 of Jacod and Protter (2012)), there are a localizing sequence $(\theta^1_m)$ and processes $X'^m, \alpha^m, \overline{\alpha}^m, \gamma^m$ satisfying (K) with $1/\alpha^m$ bounded, such that $X'^m, \alpha^m, \overline{\alpha}^m, \gamma^m$ coincide with $X', \alpha, \overline{\alpha}, \gamma$ on $[0, \theta^1_m)$.

Second, the process $X''^m_t = \sum_{i=1}^m \Gamma_i 1_{\{|\Gamma_i| \le m, 0 < S_i \le t\}}$ coincides with $X''$ on $[0, \theta^2_m)$, with the localizing sequence $\theta^2_m = S_m \wedge \inf(S_i : i \ge 1, |\Gamma_i| > m)$, and satisfies (ii) of (SHON).

The localizing sequence ($\theta_m$ in (C.2)) will be $\theta_m = m \wedge \tau_m \wedge \theta^1_m \wedge \theta^2_m$, and $X^m = X'^m + X''^m$ and $\varepsilon^{m,n}_i = \gamma^m_{T(n,i)} \chi_i$ with the same sequence $\chi_i$ as in (N): we have $X^m_t = X_t$ and $\lambda^m_t = \alpha_t$ and $\overline{\alpha}^m_t = \overline{\alpha}_t$ and $\gamma^m_t = \gamma_t$ for $t < \theta_m$, and also $\varepsilon^{m,n}_i = \varepsilon^n_i$ if $T(n,i) < \theta_m$, whereas $X^m, \alpha^m, \overline{\alpha}^m, \gamma^m, \varepsilon^{m,n}_i$ satisfy (i) and (ii) of (SHON).

(3) The construction of a sampling scheme $\Delta^m(n,i)$ satisfying (O) with the processes $\alpha^m, \overline{\alpha}^m$ and $\Delta^m(n,i) \le K\Delta^q_n$ identically and satisfying (C.2)-(ii) is more delicate.

Note that $1/\alpha^m \sqrt{\overline{\alpha}_m} \ge C$ for a constant $C > 0$. Upon enlarging the space if necessary, we have a sequence $(\Phi_i)$ of i.i.d. variables, independent of $\mathcal{F}_\infty$ and of all $\Delta(n,i)$'s and $\chi_i$'s, and which are centered with variance 1 and with support in $(-C/2, C']$ for another constant $C' > 0$. Set also $I_n = \inf(i : \Delta(n,i) > \Delta^\rho_n \text{ or } T(n,i) \ge \tau_m)$. The two sequences $\Delta^m(n,i), T^m(n,i)$ are defined by induction on $i$, as follows: we start with $T^m(n,0) = 0$ and set

$$\Delta^m(n,i) = \begin{cases} \Delta(n,i) & \text{if } i < I_n, \\ \dfrac{\Delta_n}{\alpha^m_{T^m(n,i-1)}} + \Delta_n \Phi_i \sqrt{\overline{\alpha}^m_{T^m(n,i-1)}} & \text{if } i \ge I_n, \end{cases}$$
$$T^m(n,i) = T^m(n,i-1) + \Delta^m(n,i).$$

Since $-C/2 \le \Phi_i \le C'$ and $1/\alpha^m$ and $\overline{\alpha}^m$ are bounded, we have $\frac{\Delta_n}{C''} \le \Delta^m(n,i) \le C''\Delta_n$ for $i \ge I_n$, so the previous induction defines a new sampling scheme with which we associate the filtration $(\mathcal{F}^{m,n}_t)$ as in (O). In this step, we show that this new sampling scheme satisfies (SHON) with the associated processes $\alpha^m, \overline{\alpha}^m$ in (2.3).

Since $\Delta(n,i) < \Delta^\rho_n$ if $i < I_n$ and $\Delta^m(n,i) \le C''\Delta_n$ otherwise, we obviously have $\Delta^m(n,i) \le K\Delta^\rho_n$ for some constant $K$. By construction, $T^m(n,i) = T(n,i)$ when $i < I_n$, so the restrictions of the $\sigma$-fields $\mathcal{F}^{m,n}_{T^m(n,i)}$ and $\mathcal{F}^n_{T(n,i)}$ to the set $\{i < I_n\}$ coincide, and this set is $\mathcal{F}^n_{T(n,i)}$-measurable. Hence, for any $\mathcal{R} \otimes \mathcal{F}^{m,n}_{T^m(n,i)}$-measurable function $f \ge 0$ on $\mathbb{R} \times \Omega$ and any $\mathcal{F}_\infty$-measurable variable $Z \ge 0$, the variable $B = \mathbb{E}(Zf(\Delta^m(n,i)) \mid \mathcal{F}^{m,n}_{T^m(n,i-1)})$ takes the form

$$B=\begin{cases}\mathbb{E}\big(Zf\big(\Delta(n,i),.\big)1_{\{\Delta(n,i)<\Delta_n^\rho\}}\mid\mathcal{F}^n_{T(n,i-1)}\big)\\ \quad+\mathbb{E}\left(Z1_{\{\Delta(n,i)\geq\Delta_n^\rho\}}f\left(\frac{\Delta_n}{\alpha^m_{T(n,i-1)}}+\Delta_n\Phi_i\sqrt{\overline{\alpha}^m_{T(n,i-1)}},\cdot\right)\,\Big|\,\mathcal{F}^n_{T(n,i-1)}\right) & \text{on } \{i-1<I_n\},\\ \mathbb{E}\left(Zf\left(\frac{\Delta_n}{\alpha^m_{T^m(n,i-1)}}+\Delta_n\Phi_i\sqrt{\overline{\alpha}^m_{T^m(n,i-1)}},\cdot\right)\,\Big|\,\mathcal{F}^{m,n}_{T^m(n,i-1)}\right) & \text{on } \{i-1\geq I_n\}.\end{cases}$$

Since the original scheme satisfies (ii) of (O) and $\Phi$ is independent of $\mathcal{F}_\infty$ and of $\mathcal{F}^{m,n}_{T^m(n,i-1)}$, we deduce that $B$ is the product of $\mathbb{E}(Z\mid\mathcal{F}^{m,n}_{T^m(n,i-1)})$ and $\mathbb{E}(f(\Delta^m(n,i),\cdot)\mid\mathcal{F}^{m,n}_{T^m(n,i-1)})$, so $\Delta^m(m,i)$ and $\mathcal{F}_\infty$ are $f^{m,n}_{T^m(n,i-1)}$-conditionally independent.

Since $\Phi_i$ is bounded, centered with variance 1, the above formula with $Z=1$ and $f(x,\omega)=x$ or $f(x,\omega)=(x\alpha^m_{T(n,i-1)}(\omega)-\Delta_n)^2$ or $f(x,\omega)=|x|^p$ shows us that (iii) of (O) holds with $\alpha^m,\overline{\alpha}^m$ on the set $\{i-1\geq I_n\}$ (with $\tau_1=\infty$). To see that it holds also on the complement $\{i-1<I_n\}$, and because $T(n,i-1)<\tau_m$ on this set and this property holds by hypothesis for the original scheme, it is clearly enough to prove that, for any $r,q\geq 0$,

$$\mathbb{E}\big(\Delta(n,i)^q1_{\{\Delta(n,i)\geq\Delta_n^\rho\}}\mid\mathcal{F}^n_{T(n,i)}\big)\leq K_{r,q}\Delta_n^{q+r}.\tag{C.3}$$

Using Markov's inequality and the third part of (O)-(iii), we see that the left side above is smaller than $\kappa_{m,q+p(1-\rho)}\Delta_n^{q+p(1-\rho)}$, and upon taking $p\geq r/(1-\rho)$. Therefore, the scheme $(\Delta^m(n,i))$ satisfies (SHON).

(4) It remains to show (b) of (C.2). By our definition of our new sampling scheme and of $\theta_m$, this will be implied by the property $\mathbb{P}(B_n)\to 1$, where $B_n=\{\Delta(n,i)<\Delta_n^\rho$ for $i=1,\dots,N^n_m+1\}$. Applying (C.3) with $q=0$ and $r=3$, we get

$$\mathbb{P}\big((B_n)^c\big)\leq\mathbb{P}\big(N^n_m>1/\Delta_n^2\big)+\sum_{i=1}^{[1/\Delta_n^2]}\mathbb{P}\big(\Delta(n,i)\geq\Delta_n^\rho\big)\leq\mathbb{P}\big(N^n_m>1/\Delta_n^2\big)+K\Delta_n$$

and $\mathbb{P}(N^n_m>1/\Delta_n^2)\to 0$ as $n\to\infty$ by (2.4). This completes the proof. *Q.E.D.*

### C.2. *Some Facts About Stationary Sequences*

The proof of Theorem A.4 is rather involved. We begin with some notation. By (A.11) and the Cauchy–Schwarz inequality, $\sum_{i\geq 1}\mathbb{E}(\|\mathbb{E}(\xi^n_i\mid\mathcal{H}_m)\|)<\infty$ for any $m$, whereas $V$ is bounded, so the following $d$-dimensional variables $U^n_m$ and $M^n_m$ are well defined, componentwise, as

$$U^{n,j}_m=\sqrt{\Delta_n}\sum_{i=(m-w_n)^+}^{\infty}V^j_{T(n,i)}\mathbb{E}\big(\xi^{n,j}_i\mid\mathcal{H}_m\big),$$

$$M^{n,j}_m=\sqrt{\Delta_n}\sum_{i=0}^{\infty}V^j_{T(n,i)}\big(\mathbb{E}\big(\xi^{n,j}_i\mid\mathcal{H}_m\big)-\mathbb{E}\big(\xi^{n,j}_i\mid\mathcal{H}_0\big)\big)$$

(recall $\mathcal{H}_m=\mathcal{F}^0\otimes\mathcal{G}_m$), and we write $\overline{M}^n_m$ for the same variable as $M^n_m$, with $\xi^n$ substituted with $\xi$. We also consider the $\mathcal{F}^0$-measurable variables $\nu_n(t)=(N^n_t+w_n-u_n+1)^+$. Since $\xi^n_i$ is $\mathcal{G}_{i+w_n}$-measurable, we have $\mathbb{E}(\xi^n_i\mid\mathcal{H}_{\nu_n(t)})=\xi^n_i$ when $0\leq i\leq N^n_t-u_n$, hence

$$G^n_t=M^n_{\nu_n(t)}+U^n_0-U^n_{\nu_n(t)}.\tag{C.4}$$

LEMMA C.1: *Under the assumptions of Theorem* A.4, *for any* $t > 0$ *we have*

$$U^n_{\nu_n(t)} \xrightarrow{\mathbb{P}} 0, \qquad U^n_0 \xrightarrow{\mathbb{P}} 0, \tag{C.5}$$

*and*

$$M^n_{\nu_n(t)} - \overline{M}^n_{\nu_n(t)} \xrightarrow{\mathbb{P}} 0. \tag{C.6}$$

PROOF: (1) We only prove the first convergence in (C.5), the second one being similar. We have $U^n_{\nu_n(t)} = B_n + C_n$, where

$$B_n = \sqrt{\Delta_n} \sum_{i=1+\nu_n(t)}^{\infty} V^j_{T(n,i)} \mathbb{E}\big(\xi^{n,j}_i \mid \mathcal{H}_{\nu_n(t)}\big), \qquad C_n = \sqrt{\Delta_n} \sum_{i=(\nu_n(t)-w_n)^+}^{\nu_n(t)} V^j_{T(n,i)} \mathbb{E}\big(\xi^{n,j}_i \mid \mathcal{H}_{\nu_n(t)}\big).$$

Since $V$ is bounded, we deduce from (A.11), $v > 1$, and the Cauchy–Schwarz inequality that

$$\mathbb{E}\big(\|B_n\|^2 \mid \mathcal{F}^0\big) \leq K\Delta_n \sum_{i,j=1+\nu_n(t)}^{\infty} \mathbb{E}\big(\big\|\mathbb{E}\big(\xi^n_i \mid \mathcal{H}_{\nu_n(t)}\big)\big\| \big\|\mathbb{E}\big(\xi^n_j \mid \mathcal{H}_{\nu_n(t)}\big)\big\| \mid \mathcal{F}^0\big) \leq K\Delta_n,$$

hence $B_n \xrightarrow{\mathbb{P}} 0$. Next, we have $\mathbb{E}(\|\xi^n_i\|^2) \leq K$ by (A.10), hence $\mathbb{E}(\|C_n\|^2)$ is obviously smaller than $K\Delta_n w_n^2$ because the sum defining $C_n$ contains at most $w_n$ terms. Since $\Delta_n w_n^2 \to 0$, we deduce $C_n \xrightarrow{\mathbb{P}} 0$, hence the first convergence in (C.5).

(2) Now we turn to (C.6). Setting $\xi'^n_i = \xi^n_i - \xi_i$, we have $M^n_{\nu_n(t)} - \overline{M}^n_{\nu_n(t)} = \sum_{k=1}^{\nu_n(t)} \eta^n_k$, where

$$\eta^{n,j}_k = \sqrt{\Delta_n} \sum_{i\geq 0} V^j_{T(n,i)} \big(\mathbb{E}\big(\xi'^{n,j}_i \mid \mathcal{H}_k\big) - \mathbb{E}\big(\xi'^{n,j}_i \mid \mathcal{H}_{k-1}\big)\big)$$

is a martingale increment, relative to the discrete time filtration $(\mathcal{H}_k)_{k\geq 0}$. Therefore,

$$\begin{aligned}
\mathbb{E}\big(\big(M^{n,j}_{\nu_n(t)} - \overline{M}^{n,j}_{\nu_n(t)}\big)^2 \mid \mathcal{F}^0\big) &= \sum_{k=1}^{\nu_n(t)} \mathbb{E}\big(\big(\eta^{n,j}_k\big)^2 \mid \mathcal{F}^0\big) \\
&= \Delta_n \sum_{k=1}^{\nu_n(t)} \mathbb{E}\bigg(\sum_{i,l\geq 0} V^j_{T(n,i)} V^j_{T(n,l)} \big(\mathbb{E}\big(\xi'^{n,j}_i \mid \mathcal{H}_k\big)\mathbb{E}\big(\xi'^{n,j}_l \mid \mathcal{H}_k\big) \\
&\quad - \mathbb{E}\big(\xi'^{n,j}_i \mid \mathcal{H}_{k-1}\big)\mathbb{E}\big(\xi'^{n,j}_l \mid \mathcal{H}_{k-1}\big)\big) \,\Big|\, \mathcal{F}^0\bigg).
\end{aligned}$$

The double series $\sum_{i,l}$ above is absolutely convergent (almost surely), so we may permute the summation over $(i, l)$, the one over $k$, and the conditional expectation $\mathbb{E}(\cdot \mid \mathcal{F}^0)$, to get

$$\begin{aligned}
\mathbb{E}\big(\big(M^{n,j}_{\nu_n(t)} - \overline{M}^{n,j}_{\nu_n(t)}\big)^2 \mid \mathcal{F}^0\big) &= D^n_{\nu_n(t)} - D^n_0, \quad \text{where} \\
D^n_k &= \Delta_n \sum_{i,l\geq 0} V^j_{T(n,i)} V^j_{T(n,l)} \alpha^n_{i,l,k}, \\
\alpha^n_{i,l,k} &= \mathbb{E}\big(\mathbb{E}\big(\xi'^{n,j}_i \mid \mathcal{H}_k\big)\mathbb{E}\big(\xi'^{n,j}_l \mid \mathcal{H}_k\big) \mid \mathcal{F}^0\big).
\end{aligned}$$

If $\rho_n = \mathbb{E}(\|\xi^n - \xi\|^2)$, an application of (A.8), (A.9) and the Cauchy–Schwarz inequality gives us for $i \le l$:

$$|\alpha^n_{i,l,k}| \le \begin{cases} K\rho_n/(i-k)^v(l-k)^v & \text{if } k < i, \\ K\rho_n/(l-k)^v & \text{if } 0 \le i \le k < l, \\ K\rho_n/(l-i-w_n)^v & \text{if } i + w_n < l \le k, \\ K\rho_n & \text{otherwise} \end{cases}$$

(use the property $\mathbb{E}(\xi'^n_i \mid \mathcal{H}_k) = \xi'^n_i$ when $i \le k - w_n$). Thus, since $\alpha^n_{i,l,k} = \alpha^n_{l,i,k}$, we get $\sup_{k\le m} |D^n_k| \le K\Delta_n\rho_n(1 + m + w_n m)$, hence, recalling (A.5),

$$\mathbb{E}\big(\big(M^{n,j}_{\nu_n(t)} - \overline{M}^{n,j}_{\nu_n(t)}\big)^2 \mid \mathcal{F}^0\big) \le K\rho_n(Ct+1)(1+w_n) \quad \text{on the set } \Omega^n_t.$$

Since $w_n\rho_n \to 0$ and $\mathbb{P}(\Omega^n_t) \to 1$, we readily deduce (C.6). *Q.E.D.*

Next, we observe that

$$\overline{M}^n_m = \sum_{k=1}^m \zeta^n_k, \quad \text{where}$$
$$\zeta^{n,j}_k = \sqrt{\Delta_n}\sum_{i=0}^\infty V^j_{T(n,i)}\beta^j_{i,k}, \qquad \beta^j_{i,k} = \mathbb{E}\big(\xi^j_i \mid \mathcal{H}_k\big) - \mathbb{E}\big(\xi^j_i \mid \mathcal{H}_{k-1}\big). \tag{C.7}$$

LEMMA C.2: *Under the assumptions of Theorem* A.4, *for all* $t > 0$ *and* $\varepsilon > 0$ *we have*

$$\sum_{k=1}^{\nu_n(t)} \mathbb{E}\big(\zeta^{n,j}_k \zeta^{n,m}_k \mid \mathcal{H}_{k-1}\big) \xrightarrow{\mathbb{P}} a^{jm}\int_0^t V^j_s V^m_s\, dA_s,$$
$$\sum_{k=1}^{\nu_n(t)} \mathbb{E}\big(\|\zeta^n_k\|^2 1_{\{\|\zeta^n_k\|>\varepsilon\}} \mid \mathcal{H}_{k-1}\big) \xrightarrow{\mathbb{P}} 0. \tag{C.8}$$

PROOF: (1) The second convergence is easy to prove. Since $\beta_{i,k} = \beta_{i-k,0}\circ\theta^k$ (for all $i, k \in \mathbb{Z}$), the variables $\widetilde{\beta}_k = \sum_{i\in\mathbb{Z}} \|\beta_{i,k}\|$ satisfy $\widetilde{\beta}_k = \widetilde{\beta}_0 \circ \theta^k$. A priori, $\widetilde{\beta}_k$ could be infinite; however, $\beta_{i,k} = 0$ when $i < k - w$ by (A.10), so (A.11) for $\xi$ implies $\mathbb{E}((\widetilde{\beta}_k)^2) < \infty$. The obvious estimate $\|\zeta^n_k\| \le D\sqrt{\Delta_n}\widetilde{\beta}_k$ for some $D > 0$ and stationarity yield

$$\mathbb{E}\big(\mathbb{E}\big(\|\zeta^n_k\|^2 1_{\{\|\zeta^n_k\|>\varepsilon\}} \mid \mathcal{H}_{k-1}\big) \mid \mathcal{F}^0\big) \le D^2\Delta_n\mathbb{E}\big(\widetilde{\beta}^2_k 1_{\{\widetilde{\beta}_k > \varepsilon/(D\sqrt{\Delta_n})\}}\big) = D^2\Delta_n\overline{\gamma}(\varepsilon)_n \quad \text{where}$$
$$\overline{\gamma}(\varepsilon)_n = \mathbb{E}\big((\widetilde{\beta}_0)^2 1_{\{\widetilde{\beta}_0 > \varepsilon/(D\sqrt{\Delta_n})\}}\big).$$

Now, $\overline{\gamma}(\varepsilon)_n \to 0$ as $n \to \infty$, because $\mathbb{E}(\widetilde{\beta}^2_0) < \infty$, and the second part of (C.8) follows since $\Delta_n\nu_n(t) \xrightarrow{\mathbb{P}} A_t$ by (2.4) and $\Delta_n(u_n + w^2_n) \to 0$.

(2) By virtue of the square-integrability of $\widetilde{\beta}_k$, the $d$-dimensional variables $\overline{\beta}_k = \sum_{i\ge 0}\beta_{i,k}$ are well-defined, square-integrable, and also $\overline{\beta}_k = \overline{\beta}_{w+1}\circ\theta^{k-w-1}$ for all $k \ge w+1$ (this fails when $1 \le k \le w$). In this step, we show that

$$\big|\mathbb{E}\big(\overline{\beta}^j_k\overline{\beta}^m_k\big)\big| \le K, \quad \text{and if } k \ge w+1, \text{ then } \mathbb{E}\big(\overline{\beta}^j_k\overline{\beta}^m_k\big) = a^{jm}, \tag{C.9}$$

with $a^{jm}$ given by (A.12). The first estimate follows from $\|\overline{\beta}_k\| \le \widetilde{\beta}_k = \widetilde{\beta}_0 \circ \theta^k$ and $\widetilde{\beta}_0 \in \mathbb{L}^2$. For the second property, by polarization it is enough to show it in the one-dimensional case $d = 1$, so below we omit $j, m$. Then $\beta_{l,k} = \mathbb{E}(\xi_l \mid \mathcal{G}_k) - \mathbb{E}(\xi_l \mid \mathcal{G}_{k-1})$ is $\mathcal{G}_k$-measurable with vanishing $\mathcal{G}_{k-1}$-conditional mean, and $\xi_{i+1}\mathbb{E}(\xi_{l+1} \mid \mathcal{G}_k) = (\xi_i \mathbb{E}(\xi_l \mid \mathcal{G}_{k-1})) \circ \theta$, hence

$$\begin{aligned}\mathbb{E}(\beta_{i,k}\beta_{l,k}) &= \mathbb{E}\big(\mathbb{E}(\xi_i \mid \mathcal{G}_k)\beta_{l,k} - \mathbb{E}(\xi_i \mid \mathcal{G}_{k-1})\beta_{l,k}\big) = \mathbb{E}(\xi_i \beta_{l,k}) \\ &= \mathbb{E}\big(\mathbb{E}(\xi_i - \xi_{i+1} \mid \mathcal{G}_k)\mathbb{E}(\xi_l \mid \mathcal{G}_k)\big) + \mathbb{E}\big(\mathbb{E}(\xi_{i+1} \mid \mathcal{G}_k)\mathbb{E}(\xi_l - \xi_{l+1} \mid \mathcal{G}_k)\big).\end{aligned}$$

Therefore, for any $L > 2k$ we have

$$\sum_{i,l=0}^{L} \mathbb{E}(\beta_{i,k}\beta_{l,k}) = \sum_{l=0}^{L} \mathbb{E}\big(\mathbb{E}(\xi_0 \mid \mathcal{G}_k)\mathbb{E}(\xi_l + \xi_{l+1} \mid \mathcal{G}_k)\big) - \sum_{l=0}^{L} \mathbb{E}\big(\mathbb{E}(\xi_{L+1} \mid \mathcal{G}_k)\mathbb{E}(\xi_l + \xi_{l+1} \mid \mathcal{G}_k)\big).$$

By (A.8), the $l$th summand in the last sum above is smaller in absolute value than $K/L^v$ always, and than $K/L^v l^v$ when $l > 2k$. Since $v > 1$, by letting $L \to \infty$ we obtain that

$$\mathbb{E}\big((\overline{\beta}_k)^2\big) = \sum_{l=0}^{\infty} \mathbb{E}\big(\mathbb{E}(\xi_0 \mid \mathcal{G}_k)\mathbb{E}(\xi_l + \xi_{l+1} \mid \mathcal{G}_k)\big) = \mathbb{E}\big(\xi_0^2\big) + 2\sum_{l=1}^{\infty} \mathbb{E}(\xi_0 \xi_l),$$

the last equality following from the fact that $k \ge w + 1$, hence $\xi_0$ is $\mathcal{H}_k$-measurable. The right side above is (A.12) in the one-dimensional case, and thus the last part of (C.9) holds.

(3) In this step, we set $a_k'^{jm} = \mathbb{E}(\overline{\beta}_k^j \overline{\beta}_k^m \mid \mathcal{G}_{k-1})$ and prove that

$$B_n := \sum_{k=1}^{\nu_n(t)} \big(\mathbb{E}\big(\zeta_k^{n,j}\zeta_k^{n,m} \mid \mathcal{H}_{k-1}\big) - \Delta_n V^j_{T(n,k-1)} V^m_{T(n,k-1)} a_k'^{jm}\big) \stackrel{\text{u.c.p.}}{\Longrightarrow} 0. \qquad \text{(C.10)}$$

Letting $\eta_k^n$ be the $k$th summand above, we see that $\eta_k^n = \Delta_n \sum_{i,l \ge 0} \eta(i,l)_k^n$, where

$$\eta(i,l)_k^n = \big(V^j_{T(n,i)} V^m_{T(n,l)} - V^j_{T(n,k-1)} V^m_{T(n,k-1)}\big)\mathbb{E}\big(\beta^j_{i,k}\beta^m_{l,k} \mid \mathcal{G}_{k-1}\big).$$

As seen before, $\beta_{i,k} = 0$ if $i < k - w$ and $\mathbb{E}(\|\beta_{i,k}\|^2)$ is smaller than $K/(i-k)^{2v}$ if $i > k$ and than $K$ always. Moreover, (A.5) and (A.6) imply $\mathbb{E}(\|V_{T(n,u)} - V_{T(n,v)}\|^2) \le K|v-u|\Delta_n^\rho$, whereas $V$ is bounded; hence by (A.11) and the Cauchy–Schwarz inequality, we obtain for $i \le l$:

$$\mathbb{E}\big(\big|\eta(i,l)_k^n\big|\big) \le \begin{cases} 0 & \text{if } i < k - w, \\ K\dfrac{1 \wedge \sqrt{\Delta_n^\rho\big((l+1-k) \vee (k-1-i)\big)}}{(l-k)^v} & \text{if } k - w \le i \le k < l, \\ K\dfrac{1 \wedge \sqrt{\Delta_n^\rho(l+1-k)}}{(i-k)^v(l-k)^v} & \text{if } i > k, \\ K\Delta_n^{\rho/2} & \text{if } k - w \le i \le l \le k, \end{cases}$$

and similar estimates hold for $l \le i$. Since one can always assume $v \in (1, 3/2)$, in which case $\sum_{i \ge 1}(1 \wedge \sqrt{i\Delta_n^\rho})/i^v \le K\Delta_n^{\rho(v-1)}$, we get $\mathbb{E}(|\eta_k^n|) \le K\Delta_n^{1+\rho(v-1)}$. Therefore, for any

$\eta, C > 0$,

$$\mathbb{P}(B_n > \eta) \le \mathbb{P}\big((\Omega_t^n)^c\big) + \frac{K}{\eta} \sum_{k=1}^{[(Ct+1)/\Delta_n]+w_n+1} \mathbb{E}\big(\big|\eta_k^n\big|\big) \le \mathbb{P}\big((\Omega_t^n)^c\big) + \frac{K}{\eta}\Delta_n^{\rho(v-1)}.$$

Since $\mathbb{P}((\Omega_t^n)^c) \to 0$, (C.10) follows.

(4) By the previous step, in order to get the first part of (C.8), we are left to show

$$\Delta_n \sum_{k=1}^{\nu_n(t)} V^j_{T(n,k-1)} V^m_{T(n,k-1)} a_k^{\prime jm} \xrightarrow{\mathbb{P}} a^{jl} \int_0^t V_s^j V_s^m \, dA_s. \tag{C.11}$$

The left side above is the integral of the càglàd function $s \mapsto V_s^j V_s^m$ with respect to the (random) measure $F_t^{n,jm}(ds) = \Delta_n \sum_{k=1}^{\nu_n(t)} a_k^{\prime jm} \delta_{T(n,k-1)}(ds)$, where $\delta_x$ stands for the delta measure at $x$, so it is enough to show that $F_t^{n,jm}$ converges in probability to the measure $a^{jm} 1_{[0,t]}(s)\, dA_s$, for the weak topology on the set of (signed) finite measures on $\mathbb{R}_+$. To this aim, it is enough to prove the following convergence of the cumulative distribution functions:

$$s \le t \quad \Rightarrow \quad \Delta_n \sum_{k=1}^{\nu_n(s)} a_k^{\prime jm} \xrightarrow{\mathbb{P}} a^{jm} A_s \tag{C.12}$$

(this is obvious when $m = j$, in which case $F_t^{n,jm}$ is a positive measure; when $m \neq j$, the absolute value of $F_t^{n,jm}$ is dominated by $\frac{1}{2}(F_t^{n,jj} + F_t^{n,mm})$, so again (C.12) is enough).

We recall that $\overline{\beta}_k = \overline{\beta}_{w+1} \circ \theta^{k-w-1}$ when $k > w$, implying $a'_k = a'_{w+1} \circ \theta^{k-w+1}$, so the ergodic theorem and (C.9) and $|a_k^{\prime jm}| \le K$ if $k \le w$ imply that

$$\frac{1}{L} \sum_{k=1}^{L} a_k^{\prime jm} \to a^{jm} \quad \mathbb{P}^{(1)}\text{-a.s., as } L \to \infty.$$

Since $\Delta_n \nu_n(s) \xrightarrow{\mathbb{P}} A_s$, we readily deduce (C.12). *Q.E.D.*

PROOF OF THEOREM A.4: The proof heavily relies on Jacod and Shiryaev (2003), abbreviated as [JS] below.

(1) Let the pair $(G_t, H_t)$ be as in the statement of the theorem. It can be realized as the value at time $t$ of a process $(G, H)$ defined on an extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ of $(\Omega, \mathcal{F}_\infty, \mathbb{P})$, and which, conditionally on $\mathcal{F}_\infty$, is a continuous centered Gaussian martingale, with the covariance structure given by (A.15) (for all $t$).

The $\mathcal{F}_\infty$-stable convergence $(G_t^n, H_t^n) \to (G_t, H_t)$ amounts to having, for any bounded $\mathcal{F}_\infty$-measurable variable $Y$, any continuous bounded function $f$ on $\mathbb{R}^d$, and any $u \in \mathbb{R}^d$:

$$\mathbb{E}\big(Yf\big(H_t^n\big)e^{iu\cdot G_t^n}\big) \to \widetilde{\mathbb{E}}\big(Yf(H_t)e^{iu\cdot G_t}\big)$$

($u \cdot v$ is the scalar product on $\mathbb{R}^d$). In view of (C.4) and Lemma C.1, this is implied by

$$\mathbb{E}\big(Yf\big(H_t^n\big)e^{iu\cdot \overline{M}^n_{\nu_n(t)}}\big) \to \widetilde{\mathbb{E}}\big(Yf(H_t)e^{iu\cdot G_t}\big). \tag{C.13}$$

(2) Equation (A.15) and Lemma C.2 and the fact that the $\zeta_k^n$'s are martingale increments relative to the filtration $(\mathcal{H}_k)$ imply, with the help of Theorems VIII.3.22 and

VIII.5.14 of [JS], that

$$\overline{M}^n_{\nu_n(t)} \text{ converge } \mathcal{F}^0\text{-stably in law to } G_t. \tag{C.14}$$

On the other hand, Lemma A.2 and our assumption on $u_n$ imply

$$H^n_t \text{ converge } \mathcal{F}_\infty\text{-stably in law to } H_t. \tag{C.15}$$

Equations (C.14) and (C.15) are not enough for us; we need a joint convergence. To this end, we need to revisit the convergence (C.14). As for all special semimartingales, we can write $e^{iu\cdot\overline{M}^n_{\nu_n(t)}}$ uniquely as a product $g(u)^n_t Z(u)^n_t$ of a predictable càdlàg process with finite variation $g(u)^n$ (called $G^n(u)$ in [JS]) and a martingale $Z(u)^n$, both relative to the filtration $(\mathcal{H}_{\nu_n(t)})_{t\geq 0}$, both starting at 1 at time 0, and both complex-valued. Analogously, $e^{iu\cdot G_t} = g(u)_t Z(u)_t$, with $g(u)$ predictable with finite variation and $Z(u)$ a martingale, relative to the smallest filtration $(\widetilde{\mathcal{F}}^{(1)}_t)$ to which $G$ is adapted and such that $\mathcal{F}^0 \subset \widetilde{\mathcal{F}}^{(1)}_0$.

We do not need to recall the explicit form of $g(u)^n$ and $g(u)$ (although $g(u)$ takes the simple form $g(u)_t = \exp(-\frac{1}{2}\sum_{j,m=1}^d u^j u^m a^{jm} \int_0^t V^j_s V^m_s \alpha_s\, ds))$, but only that, according to the proof of Theorem VIII.2.4 of [JS], we do have, for all $t$ and by Lemma C.2:

$$g(u)^n_t \xrightarrow{\mathbb{P}} g(u)_t. \tag{C.16}$$

Observe that $2\varepsilon \leq |g(u)_t| \leq \frac{1}{2\varepsilon}$ for some constant $\varepsilon \in (0, \frac{1}{2})$ (depending on $t$ and $u$). The $(\mathcal{H}_{\nu_n(t)})$-stopping times $R_n = \inf\{s : |g(u)^n_s| \leq \varepsilon \text{ or } |g(u)^n_s| \geq \frac{1}{\varepsilon}\}$ satisfy $\mathbb{P}(R_n \leq t) \to 0$ by (C.16) and are predictable, so there are $(\mathcal{H}_{\nu_n(t)})$-stopping times $S_n < R_n$ such that $\mathbb{P}(S_n \leq t) \to 0$ as well, whereas $\varepsilon \leq |g(u)^n_s| \leq \frac{1}{\varepsilon}$ and thus $\varepsilon \leq |Z(u)^n_s| \leq \frac{1}{\varepsilon}$ for all $s \leq S_n$. Then, partly reproducing the proof of Theorem VIII.5.16 of [JS], for any uniformly bounded sequence $Y_n$ of $\mathcal{F}^0$-measurable variables, one can write

$$\begin{aligned}
\big|\widetilde{\mathbb{E}}\big(Y_n\big(e^{iu\cdot\overline{M}^n_{\nu_n(t)}} - e^{iu\cdot G_t}\big)\big)\big| &\leq K\mathbb{P}(S_n \leq t) + \big|\widetilde{\mathbb{E}}(Y_n\big(g(u)^n_{t\wedge S_n} Z(u)^n_{t\wedge S_n} - g(u)_t Z(u)_t\big)\big| \\
&\leq K\mathbb{P}(S_n \leq t) + \big|\widetilde{\mathbb{E}}\big(Y_n\big(g(u)^n_{t\wedge S_n} - g(u)_t\big) Z(u)^n_{t\wedge S_n}\big)\big| \\
&\quad + \big|\widetilde{\mathbb{E}}\big(Y_n g(u)_t\big(Z(u)^n_{t\wedge S_n} - Z(u)_t\big)\big)\big|.
\end{aligned}$$

Since $\mathbb{E}(Z(u)^n_{t\wedge S_n} \mid \mathcal{F}^0) = \widetilde{\mathbb{E}}(Z(u)_t \mid \mathcal{F}^0) = 1$ and $Y_n g(u)_t$ is $\mathcal{F}^0$-measurable, the last term above vanishes. We have $|Y_n Z(u)^n_{t\wedge S_n}| \leq K/\varepsilon$, hence

$$\big|\widetilde{\mathbb{E}}\big(Y_n\big(e^{iu\cdot\overline{M}^n_{\nu_n(t)}} - e^{iu\cdot G_t}\big)\big)\big| \leq K\mathbb{P}(S_n \leq t) + \frac{K}{\varepsilon}\mathbb{E}\big(\big|g(u)^n_{t\wedge S_n} - g(u)_t\big|\big).$$

Using again (C.16), plus $\mathbb{P}(S_n \leq t) \to 0$, we deduce

$$\widetilde{\mathbb{E}}(Y_n\big(e^{iu\cdot\overline{M}^n_{\nu_n(t)}} - e^{iu\cdot G_t}\big) \to 0. \tag{C.17}$$

(3) Now, we are in a position to prove (C.13). First, the characterization of $G_t$ gives us $\widetilde{\mathbb{E}}(e^{iu\cdot G_t} \mid \mathcal{F}) = g(u)_t$. Next, apply (C.17) with $Y_n = Yf(H^n_t)$ to get

$$\mathbb{E}\big(Yf\big(H^n_t\big)e^{iu\cdot\overline{M}^n_{\nu_n(t)}}\big) - \mathbb{E}\big(Yf\big(H^n_t\big)e^{iu\cdot G_t}\big) \to 0.$$

Finally, $Y' = Yg(u)_t$ is bounded $\mathcal{F}_\infty$-measurable, hence (C.15) yields

$$\widetilde{\mathbb{E}}\big(Yf\big(H^n_t\big)e^{iu\cdot G_t}\big) = \mathbb{E}\big(Yf\big(H^n_t\big)g(u)_t\big) \to \widetilde{\mathbb{E}}\big(Yf(H_t)g(u)_t\big) = \widetilde{\mathbb{E}}\big(Yf(H_t)e^{iu\cdot G_t}\big),$$

where the last equality comes from the independence of $H_t$ and $G_t$, conditionally on $\mathcal{F}$. We then deduce (C.13), and the theorem is proved. *Q.E.D.*

PROOF OF LEMMA A.5: We use a unified approach, by setting

$$\begin{aligned}
&\text{case 1:}\quad F^n = G'^{n,1}, \qquad \xi_i'^n = \xi_i^{n,1}, \qquad h_n = w_n^1,\\
&\text{case 2:}\quad F^n = G'^{n,2}, \qquad \xi_i'^n = \xi_{i+w_n^1+k_n}^{n,2}, \qquad h_n = w_n^2,\\
&\text{case 3:}\quad F^n = G''^n, \qquad \xi_i'^n = \xi_i^{n,1}\xi_{i+w_n^1+k_n}^{n,2}, \qquad h_n = k_n + w_n^1 + w_n^2.
\end{aligned}$$

In all cases, $\mathbb{E}(|\xi_i'^n|^2) \leq K$, and $\xi_i'^n$ is centered in cases 1 and 2, whereas in case 3, (A.9) yields $|\mathbb{E}(\xi_i'^n)| \leq K/k_n^v$. Then another application of (A.9) also yields

$$l > h_n \quad \Rightarrow \quad \left|\mathbb{E}\left(\xi_i'^n \xi_{i+l}'^n \mid \mathcal{F}^0\right)\right| \leq \begin{cases} K(l-h_n)^{-v} & \text{in cases 1 and 2,}\\ Kk_n^{-2v} + K(l-h_n)^{-v} & \text{in case 3.}\end{cases}$$

Since $V$ is bounded and $N_t^n$ is $\mathcal{F}^0$-measurable and independent of all $\xi_i'^n$'s, we deduce

$$\mathbb{E}\left(\left|F_t^n\right|^2 \mid \mathcal{F}^0\right) \leq \sum_{i,j=0}^{N_t^n} \left|\mathbb{E}\left(\xi_i'^n \xi_j'^n\right)\right| \leq \begin{cases} Kh_nN_t^n + N_t^n \displaystyle\sum_{m=1}^{N_t^n} m^{-v} & \text{in cases 1,2,}\\ Kh_nN_t^n + Kk_n^{-2v}\left(N_t^n\right)^2 + N_t^n \displaystyle\sum_{m=1}^{N_t^n} m^{-v} & \text{in case 3.}\end{cases}$$

The result is now obvious. *Q.E.D.*

### C.3. *Further Auxiliary Results*

PROOF OF LEMMA A.6: We set $\delta_i'^n = \mathbb{E}(\delta_i^n \mid \widetilde{\mathcal{H}}_i^n)$ and $\delta_i''^n = \delta_i^n - \delta_i'^n$ and, for $j = 0, \dots, w_n - 1$,

$$B_l'^n = \sum_{i=0}^{l} \delta_i'^n, \qquad B''(j)_l^n = \sum_{i=1}^{[(l-j)/w_n]} \delta_{j+(i-1)w_n}''^n,$$

so $B_l^n = B_l'^n + \sum_{j=0}^{w_n-1} B''(j)_l^n$. We clearly have $\mathbb{E}(\sup_{l\leq k}|B_l'^n|) \leq (k+1)a_n$. The summands $\delta_{j+(i-1)w_n}''^n$ are martingale increments, relative to the filtration $(\widetilde{\mathcal{H}}_{j+iw_n}^n)_{i\geq 0}$, hence

$$\mathbb{E}\left(\sup_{l\leq k}\left|B''(j)_l^n\right|^2\right) \leq 4\sum_{i=1}^{[(k-j)/w_n]} \mathbb{E}\left(\left(\delta_i''^n\right)^2\right) \leq 4\sum_{i=1}^{[(k-j)/w_n]} \mathbb{E}\left(\left(\delta_i^n\right)^2\right) \leq \frac{Kka_n'}{w_n}$$

by Doob's inequality, so $\mathbb{E}(\sup_{l\leq k}|B''(j)_l^n|) \leq K\sqrt{ka_n'/w_n}$. The left side of (A.17) being smaller than $\mathbb{E}(\sup_{l\leq k}|B_l'^n|) + \sum_{j=0}^{w_n-1}\mathbb{E}(\sup_{l\leq k}|B''(j)_l^n|)$, we deduce the result. *Q.E.D.*

PROOF OF LEMMA A.7: With $\overline{V}=V_t\alpha_t$, we have

$$J(V)^n_t=D^n_t+B^n_{N^n_t-u_n},\qquad D^n_t=\frac{1}{\sqrt{\Delta_n}}\int_{T(n,N^n_t-u_n)}^t (V_s-V_{T(n,N^n_s)})\alpha_s\,ds,$$

$$B^n_k=\sum_{i=0}^k\delta^n_i,\qquad \delta^n_i=\frac{1}{\sqrt{\Delta_n}}\int_{T(n,i)}^{T(n,i+1)}(\overline{V}_s-\overline{V}_{T(n,i)})\,ds.$$

Equation (A.5) and the boundedness of $V$ and $\alpha$ imply $|D^n_t|\le Ku_n\Delta_n^{\rho-1/2}$, which goes to 0. In view of (A.5), $B^n_{N^n_t-u_n}\overset{\text{u.c.p.}}{\Longrightarrow}0$ follows from the property $\mathbb{E}(\sup_{j\le[D/\Delta_n]}|B^n_j|)\to 0$ for any constant $D$. This in turn follows from Lemma A.6 applied to $\delta^n_i$, the assumptions of this lemma being fulfilled with $\widetilde{\mathcal{H}}^n_i=\mathcal{F}^n_{T(n,i)}$ and $w_n=1$ and, by virtue of (A.7), $a_n=K\Delta_n^{3/2}$ and $a'_n=K\Delta_n^2$. This completes the proof. *Q.E.D.*

PROOF OF LEMMA A.8: Since $|r(m)|\le K(1\wedge m^{-v})$, we have for all $j$:

$$\mathbb{E}\big((\overline{\chi}^n_j)^2\big)=\frac{1}{k_n^2}\sum_{0\le i,l<k_n-1}r(i-l)\le\frac{2}{k_n}\sum_{m=0}^{k_n-1}\big|r(m)\big|\le\frac{K}{k_n}\left(1+\sum_{m=1}^{k_n-1}\frac{1}{m^v}\right)\le Kf_v(k_n),\tag{C.18}$$

with $f_v(k_n)$ as in the statement of the lemma. This yields the first claim. Since all moments of $\overline{\chi}^n_j$ are bounded in $j$ and $n$, by Hölder's inequality we also get, for $p>2$ and $\varepsilon>0$:

$$\mathbb{E}\big(\big|\overline{\chi}^n_j\big|^p\big)\le K_{p,\varepsilon}f_v(k_n)^{1-\varepsilon}.\tag{C.19}$$

Next, we denote by $\mathcal{Q}$ the set of all non-empty subsets $Q$ of $\{1,\dots,q\}$, the complement of $Q$ being denoted as $Q^c$, and its cardinal is $|Q|$. We have

$$U_n:=\prod_{m=1}^q\big(\chi_{j_m}-\overline{\chi}^n_{\mu+(2m-1)k_n}\big)-\prod_{m=1}^q\chi_{j_m}=\sum_{Q\in\mathcal{Q}}(-1)^{|Q|}V_n(Q),$$

where

$$V_n(Q)=\prod_{\ell\in Q^c}\chi_{j_\ell}\prod_{\ell\in Q}\overline{\chi}^n_{\mu+(2\ell-1)k_n}.$$

We fix $Q\in\mathcal{Q}$ and let $r_0=\max Q$ and $Q'=Q\setminus\{r_0\}$. We have

$$V_n(Q)=V'_n(Q)\overline{\chi}^n_{\mu+(2r_0-1)k_n},\quad\text{where } V'_n(Q)=\prod_{\ell\in Q^c}\chi_{j_\ell}\prod_{\ell\in Q'}\overline{\chi}^n_{\mu+(2\ell-1)k_n}.$$

The variable $V'_n(Q)$ is $\mathcal{G}_{\mu+(2r_0-2)k_n}$-measurable, with all moments bounded in $n$, whereas $\overline{\chi}^n_{\mu+2(r_0-1)k_n}$ is $\mathcal{G}^{\mu+(2r_0-1)k_n}$-measurable, centered, and satisfies (C.19). Then (C.18), (C.19), (A.9), and the property that $k_n^{-v}f_v(k_n)^{1/2-\varepsilon}\le Kf_v(k_n)$ for $\varepsilon\in(0,1/2]$ yield

$$\big|\mathbb{E}\big(V_n(Q)\big)\big|\le Kk_n^{-v}\sqrt{f_v(k_n)},\qquad \mathbb{E}\big(\big|V_n(Q)\big|^2\big)\le K_{p,\varepsilon}f_v(k_n).$$

Since (3.1) and (3.26) yield $\mathbf{r}(k_n;\mathbf{j})-\mathbf{r}(\mathbf{j})=\mathbb{E}(U_n)$, by summing up the estimates above over all $Q\in\mathcal{Q}$, we readily get (A.19). *Q.E.D.*

PROOF OF LEMMA A.9: (1) We will focus on the second part of (A.21), the first part being analogous, and in fact slightly simpler. We take $\mathbf{j}, \mathbf{j}'$ as above, and set for any integer $k$:

$$\overline{U}_k^n = \sum_{i=1}^{k} \big(\widehat{Y}(\mathbf{j})_i^n \widehat{Y}(\mathbf{j}')_{i+\mu+(2q+1)k_n}^n - \gamma_{T(n,i)}^{q''} \widehat{\chi}(\mathbf{j})_i^n \widehat{\chi}(\mathbf{j}')_{i+\mu+(2q+1)k_n}^n\big). \tag{C.20}$$

Suppose for a moment that we have

$$\mathbb{E}\Big(\sup_{k\le j} \big|\overline{U}_k^n\big|\Big) \le K_{r,q,q'} \big(j\big(k_n \Delta_n^\rho\big)^{1/r} + j^{1/2} k_n^{(1+r)/2r} \Delta_n^{\rho/2r}\big). \tag{C.21}$$

Observe that, with $\zeta_i^n = |\widehat{\chi}(\mathbf{j})_i^n \widehat{\chi}(\mathbf{j}')_{i+\mu+(2q+1)k_n}^n|$ and $u_n' = u_n \vee (\mu'' + (2q''+1)k_n))$, we have

$$\big|U^4\big(\mathbf{j},\mathbf{j}'\big)_t^n - U'^4\big(\mathbf{j},\mathbf{j}'\big)_t^n\big| \le K R_t^n + \sup_{k\le N_t^n} \big|\overline{U}_k^n\big|, \qquad R_t^n = \sum_{i=N_t^n-u_n'}^{N_t^n} \zeta_i^n.$$

Since the variables $\zeta_i^n$ have a finite moment (not depending on $i$) and are independent of $\mathcal{F}^0$, by conditioning first on this $\sigma$-field we see that $\mathbb{E}(R_t^n) \le K_{q''} k_n$. On the other hand, on the set $\Omega_t^n$ we have $N_t^n \le (1+Ct)/\Delta_n$, so the left side of (A.21) is smaller than $Kk_n$ plus the bound in (C.21) evaluated at $j = [(1+Ct)/\Delta_n]$. Since both $k_n$ and $k_n^{(1+r)/2r} \Delta_n^{(\rho-r)/2r}$ are smaller than $Kk_n^{1/r} \Delta_n^{\rho/r-1}$ by (A.4), we deduce the second part of (A.21).

(2) The proof of (C.21) goes through several steps, the first one being devoted to some estimates. Set for $u, l, w \in \mathbb{N}$:

$$\zeta(m)_{i,u,l}^n = \begin{cases} X_{T(n,i+u)} - \overline{X}_{i+l}^n + (\gamma_{T(n,i+u)} - \gamma_{T(n,i)})\chi_{i+u} \\ \quad - \dfrac{1}{k_n} \displaystyle\sum_{j=0}^{k_n-1} (\gamma_{T(n,i+l+j)} - \gamma_{T(n,i)})\chi_{i+l+j} & \text{if } m=1, \\ \gamma_{T(n,i)}\big(\chi_{i+u} - \overline{\chi}_{i+l}^n\big) & \text{if } m=2. \end{cases}$$

Upon using (A.5) and (A.6) with $V = X'$ (since here $X = X_0 + X'$) or with $V = \gamma$, the independence of $\mathcal{F}^0$ and $\mathcal{G}$, and the fact that $\chi_i$ has moments of all orders, we get for $p \ge 2$:

$$\mathbb{E}\big(\big|\zeta(1)_{i,u,l}^n\big|^p\big) \le K_p \Delta_n^\rho (u+l+k_n), \qquad \mathbb{E}\big(\big|\zeta(2)_{i,u,l}^n\big|^p\big) \le K_p. \tag{C.22}$$

Moreover, (A.6) and the independence of $\mathcal{F}^0$ and $\mathcal{G}$ yield

$$\big|\mathbb{E}\big(\zeta(1)_{i,u,l}^n \mid \mathcal{F}_{T(n,i)}^n \otimes \mathcal{G}\big)\big| \le K\Delta_n^\rho (u+l+k_n)\left(1 + |\chi_{i+u}| + \frac{1}{k_n}\sum_{m=0}^{k_n-1} |\chi_{i+l+m}|\right). \tag{C.23}$$

(3) Since $Y_{i+u}^n - \overline{Y}_{i+l}^n = \sum_{m=1}^4 \zeta(m)_{i,u,l}^n$, and with the notation

$$1 \le \ell \le q \quad \Rightarrow \quad u_\ell^n = j_\ell, \qquad l_\ell^n = \mu + (2\ell-1)k_n,$$
$$q < \ell \le q'' \quad \Rightarrow \quad u_\ell^n = \mu + (2q+1)k_n + j_{\ell-q}', \qquad l_\ell^n = \mu'' + 2(\ell+q)k_n,$$

a simple calculation shows us that $\overline{U}^n_k = \sum_{i=0}^k \xi^n_i$, where

$$\xi^n_i = \prod_{\ell=1}^{q''}\Bigg(\sum_{m=1}^{2} \zeta(m)^n_{i+2+2k_n,u^n_\ell,l^n_\ell}\Bigg) - \prod_{\ell=1}^{q''} \zeta(2)^n_{i+2+2k_n,u^n_\ell,l^n_\ell}.$$

So, if $\mathcal{Q}$ is the set of all partitions $Q = (Q_1, Q_2)$ of $\{1, \dots, q''\}$ such that $Q_1 \neq \emptyset$, we have

$$\xi^n_i = \sum_{Q\in\mathcal{Q}} \eta(Q)^n_i, \quad \text{where}$$

$$\eta(Q)^n_i = \eta(Q_1, 1)^n_i \eta(Q_2, 2)^n_i \quad \text{and} \quad \eta(Q_m, m)^n_i = \prod_{\ell\in Q_j} \zeta(m)^n_{i+2+2k_n,u^n_\ell,l^n_\ell}.$$

It remains to prove that, for each $Q \in \mathcal{Q}$, the sequence $M^n_k = \sum_{i=0}^k \eta(Q)^n_i$ satisfies (C.21).

(4) We start with the case where $Q_1$ is a singleton, say $Q_1 = \{\ell\}$ for some $\ell \in \{1, \dots, q''\}$, so $\eta(Q)^n_i = \zeta(1)^n_{i,u^n_\ell,l^n_\ell}\eta(Q_2, 2)^n_i$. By (C.22), (C.23), successive conditioning, Hölder's inequality, and the facts that $\eta(Q_2, 2)^n_i$ is $\mathcal{G}$-measurable with bounded moments of all orders and $u^n_\ell + l^n_\ell \le K_{q''} k_n$, the assumptions of Lemma A.6 are satisfied by the variables $\delta^n_i = \eta(Q)^n_i$, with $w_n = \mu'' + (2q'' + 1)k_n$ and $\widetilde{\mathcal{H}}^n_i = \mathcal{H}^n_i$ and, for any $r > 1$,

$$a_n \le K_{q,q'} k_n\Delta^\rho_n, \qquad a'_n \le K_{r,q,q'}\big(k_n\Delta^\rho_n\big)^{1/r}.$$

Then, since $k_n\Delta^\rho_n \to 0$, (A.17) implies that $M(Q)^n$ satisfies (C.21).

In all other cases of $Q \in \mathcal{Q}$, there are at least two distinct integers $\ell$ and $\ell'$ in $Q_1$. Then $\eta(Q)^n_i = \zeta(1)^n_{i+2+2k_n,u^n_\ell,l^n_\ell}\zeta(1)^n_{i+2+2k_n,u^n_{\ell'},l^n_{\ell'}}\zeta'^n_i$, where $\zeta'^n_i$ has finite moments of all order by (C.22). Then (C.24), (C.22), and Hölder's inequality yield $\mathbb{E}(|\eta(Q)^n_i|) \le K_r(k_n\Delta^\rho_n)^{1/r}$. Thus (C.21) holds again. This completes the proof. *Q.E.D.*

PROOF OF LEMMA A.10: (1) We begin by proving that

$$\big|\mathbb{E}\big(\widehat{\Delta}^n_i \mid \mathcal{F}^n_{T(n,i)}\big) - \overline{\alpha}_{T(n,i)}\big| \le K\phi'_n, \qquad \mathbb{E}\big(\big(\widehat{\Delta}^n_i\big)^2 \mid \mathcal{F}^n_{T(n,i)}\big) \le K. \tag{C.24}$$

Set

$$\lambda^{1,n}_i = \frac{\Delta(n, i+1)}{\Delta_n} - \frac{1}{\alpha_{T(n,i)}}, \qquad \lambda^{2,n}_i = \frac{T(n, i+k_n) - T(n, i)}{k_n\Delta_n} - \frac{1}{\alpha_{T(n,i)}},$$

$$\lambda^{3,n}_i = \frac{T(n, i+1+k_n) - T(n, i+1)}{k_n\Delta_n} - \frac{1}{\alpha_{T(n,i)}}, \qquad \lambda^{4,n}_i = \big(\lambda^{1,n}_i - \lambda^{3,n}_i\big)^2,$$

$$\lambda^{5,n}_i = \frac{k_n\Delta_n}{\big(T(n, i+k_n) - T(n, i)\big) \vee \phi_n}.$$

First, (2.3) with $\tau_1 = \infty$ yields

$$\big|\mathbb{E}\big(\lambda^{1,n}_i \mid \mathcal{F}^n_{T(n,i)}\big)\big| \le K\Delta_n^{1/2+\kappa}, \qquad \mathbb{E}\big(\big(\lambda^{1,n}_i\big)^4 \mid \mathcal{F}^n_{T(n,i)}\big) \le K,$$
$$\left|\mathbb{E}\big(\big(\lambda^{1,n}_i\big)^2 \mid \mathcal{F}^n_{T(n,i)}\big) - \frac{\overline{\alpha}_{T(n,i)}}{\alpha^2_{T(n,i)}}\right| \le K\Delta^\kappa_n. \tag{C.25}$$

Next, we have

$$\lambda_i^{1,n} = \sum_{m=1}^{3} \frac{1}{k_n} \sum_{j=0}^{k_n-1} \zeta_i^{m,j,n}, \qquad \zeta_i^{m,j,n} = \begin{cases} \mathbb{E}\big(\lambda_{i+j}^{1,n} \mid \mathcal{F}_{T(n,i+j)}^n\big) & \text{if } m=1, \\ \lambda_{i+j}^{1,n} - \mathbb{E}\big(\lambda_{i+j}^{1,n} \mid \mathcal{F}_{T(n,i+j)}^n\big) & \text{if } m=2, \\ 1/\alpha_{T(n,i+j)} - 1/\alpha_{T(n,i)} & \text{if } m=1. \end{cases}$$

The process $1/\alpha$ satisfies (K), hence (A.6), (C.25), and (3.5) yield for $r = 1, 2$:

$$\begin{aligned} &\big|\mathbb{E}\big(\lambda_i^{r,n} \mid \mathcal{F}_{T(n,i)}^n\big)\big| \le K\big(\Delta_n^{1/2+\kappa} + k_n\Delta_n\big), \qquad \mathbb{E}\big(\big(\lambda_i^{r,n}\big)^4 \mid \mathcal{F}_{T(n,i)}^n\big) \le K, \\ &\mathbb{E}\big(\big(\lambda_i^{r,n}\big)^2 \mid \mathcal{F}_{T(n,i)}^n\big) \le K\left(\Delta_n^{1+2\kappa} + \frac{1}{k_n} + k_n\Delta_n\right) \le \frac{K}{k_n}. \end{aligned} \tag{C.26}$$

Upon expanding the square in the definition of $\lambda_i^{4,n}$, this for $r = 2$ and (C.25) and also (A.6) for the process $\overline{\alpha}/\alpha^2$ yield

$$\left|\mathbb{E}\big(\lambda_{i+k_n}^{4,n} \mid \mathcal{F}_{T(n,i)}^n\big) - \frac{\overline{\alpha}_{T(n,i)}}{\alpha_{T(n,i)}^2}\right| \le K\left(\Delta_n^\kappa + \frac{1}{\sqrt{k_n}}\right), \qquad \mathbb{E}\big(\big(\lambda_{i+k_n}^{4,n}\big)^2 \mid \mathcal{F}_{T(n,i)}^n\big) \le K. \tag{C.27}$$

Observe that $\lambda_i^{5,n} = \alpha_{T(n,i)}/(1 + \alpha_{T(n,i)}\lambda_i^{2,n}) \vee b_n$, where $b_n = \phi_n\alpha_{T(n,i)}/k_n\Delta_n$. Let $D$ be a constant such that $\alpha_t \le D$, and $p = 2, 4$. Expanding $x \mapsto \frac{1}{(1+x)^p}$ around 0 for $|x| \le \frac{1}{2}$, and using (C.26) with $r = 2$, we get (since $\phi_n/k_n\Delta_n \to 0$ by (3.14), we can assume $b_n < \frac{1}{2}$):

$$\begin{aligned} &\big|\mathbb{E}\big(\big(\lambda_i^{5,n}\big)^p \mid \mathcal{F}_{T(n,i)}^n\big) - \alpha_{T(n,i)}^p\big| \\ &\quad\le \frac{\alpha_{T(n,i)}^p}{b_n^p}\mathbb{P}\big(\big|\lambda_i^{2,n}\big| > 1/2D \mid \mathcal{F}_{T(n,i)}^n\big) + K\alpha_{T(n,i)}^{p+1}\mathbb{E}\big(\big(\lambda_i^{4,n}\big)^p \mid \mathcal{F}_{T(n,i)}^n\big) \\ &\quad\le \frac{(2D)^2(k_n\Delta_n)^p}{\phi_n^p}\mathbb{E}\big(\big(\lambda_i^{4,n}\big)^2 \mid \mathcal{F}_{T(n,i)}^n\big) + K\mathbb{E}\big(\big(\lambda_i^{4,n}\big)^p \mid \mathcal{F}_{T(n,i)}^n\big) \le \begin{cases} K\dfrac{k_n\Delta_n^2}{\phi_n^2} & \text{if } p=2, \\ K\dfrac{k_n^3\Delta_n^4}{\phi_n^4} + K & \text{if } p=4. \end{cases} \end{aligned}$$

Since $\widehat{\Delta}_i^n = (\lambda_{i+k_n}^{4,n})^2(\lambda_i^{5,n})^2$ and $k_n^3\Delta_n^4/\phi_n^4 \to 0$, the above estimate and (C.27) yield (C.24).

(2) Now we turn to our claim. The assumptions of Lemma A.6, for the sequence $\delta_i^n = \Delta_n V_{T(n,i)}(\widehat{\Delta}_i^n - \overline{\alpha}_{T(n,i)}^n)$, are satisfied with $\widetilde{\mathcal{H}}_i^n = \mathcal{F}_{T(n,i)}^n$ and $w_n = 2 + 2k_n$ and $a_n = K\phi_n'$ and $a_n' = K$, by (C.24). Hence (A.22) readily follows from (A.17). *Q.E.D.*

PROOF OF LEMMA A.11: (1) We prove (3.19) only, the other two claims being proved analogously (in a slightly simpler way). The second part of (3.19) is an obvious consequence of the first part and of (2.4), so we focus on the first part, and we let $\mathbf{j}, \mathbf{j}' \in \mathcal{J}^+$ with $\mu = \mu(\mathbf{j})$ and $\mu' = \mu(\mathbf{j}')$ and $q'' = q + q'$ and $\mu'' = \mu + \mu'$. By virtue of (A.2), (A.4), and Lemma A.8, it is enough to show that

$$B_t^n := \Delta_n U^3\big(\mathbf{j}, \mathbf{j}'\big)_t^n - \Delta_n \mathbf{r}(k_n; \mathbf{j}) r\big(k_n; \mathbf{j}'\big) \sum_{i=0}^{N_t^n - 1 - \mu'' - (2q''+3)k_n} \overline{\alpha}_{T(n,i)} \gamma_{T(n,i)}^{q''} \xrightarrow{\mathbb{P}} 0.$$

With the same notation $\widehat{\chi}(\mathbf{j})_i^n$ as in (A.20), we set

$$\overline{U}_k^n = \sum_{i=0}^k \overline{\delta}_i^n, \qquad U_k^n = \sum_{i=0}^k \delta_i^n, \qquad U_k'^n = \sum_{i=0}^k \delta_i'^n, \quad \text{where}$$
$$\overline{\delta}_i^n = \widehat{\Delta}_i^n \big(\widehat{Y}(\mathbf{j})_{i+2+2k_n}^n \widehat{Y}(\mathbf{j}')_{i+2+\mu+(2q+3)k_n}^n - \gamma_{T(n,i)}^{q''} \widehat{\chi}(\mathbf{j})_{i+2+2k_n}^n \widehat{\chi}(\mathbf{j}')_{i+2+\mu+(2q+3)k_n}^n\big),$$
$$\delta_i^n = \widehat{\Delta}_i^n \gamma_{T(n,i)}^{q''} \big(\widehat{\chi}(\mathbf{j})_{i+2+2k_n}^n \widehat{\chi}(\mathbf{j}')_{i+2+\mu+(2q+3)k_n}^n - \mathbf{r}(k_n;\mathbf{j}) r(k_n;\mathbf{j}')\big),$$
$$\delta_i'^n = \mathbf{r}(k_n;\mathbf{j}) r(k_n;\mathbf{j}') \gamma_{T(n,i)}^{q''} \big(\widehat{\Delta}_i^n - \overline{\alpha}_{T(n,i)}\big).$$

We will prove that, for any $r > 1$,

$$\begin{aligned} &\mathbb{E}\Big(\sup_{k\le j} \big|\overline{U}_k^n\big|\Big) \le K_r\big(j\big(k_n\Delta_n^\rho\big)^{1/r} + j^{1/2} k_n^{(1+r)/2r} \Delta_n^{\rho/2r}\big), \\ &\mathbb{E}\Big(\sup_{k\le j} \big|U_k^n\big|\Big) \le K\big(jk_n^{-v} + \sqrt{jk_n}\big), \\ &\mathbb{E}\Big(\sup_{k\le j} \big|U_k'^n\big|\Big) \le K\big(j\phi_n' + \sqrt{jk_n}\big). \end{aligned} \tag{C.28}$$

Suppose indeed that (C.28) holds. We have $|B_t^n| \le \Delta_n \sup_{k\le[(1+Ct)/\Delta_n]}(|U_k^n| + |U_k'^n|)$ in restriction to the set $\Omega_t^n$, whose probability goes to 1. Substituting $j$ with $[(1+Ct)/\Delta_n]$ in (C.28), and since $\phi_n' \to 0$ and under our assumptions, we readily deduce $B_t^n \xrightarrow{\mathbb{P}} 0$.

(2) We are thus left to proving (C.28). The third estimate is (A.22) with $V = \gamma^{q''}$. The first one is proved exactly as (C.21), once noticed that $\overline{U}_k^n$ here has the same structure as in (C.20), with, in each summand, a shift by $2+2k_n$ of the indices and the additional multiplicative term $\widehat{\Delta}_i^n$ which satisfies (C.24).

Finally, with the notation $\widehat{\chi}'(\mathbf{j})_i^n = \widehat{\chi}((\mathbf{j})_i^n - \mathbf{r}(k_n;\mathbf{j})$, we have

$$\delta_i^n = \widehat{\Delta}_i^n \gamma_{T(n,i)}^{q''} \big(\widehat{\chi}(\mathbf{j})_{i+2+2k_n}^n \widehat{\chi}'(\mathbf{j}')_{i+2+\mu+(2q+2)k_n}^n + \mathbf{r}(k_n;\mathbf{j}')\widehat{\chi}'(\mathbf{j})_{i+2+2k_n}^n\big).$$

Since the variables $\widehat{\chi}'(\mathbf{j})_i^n$ are centered by definition of $\mathbf{r}(k_n;\mathbf{j}')$, we deduce from (A.8) and the facts that $\widehat{\chi}(\mathbf{j})_{i+2+2k_n}^n$ is $\mathcal{G}_{i+2+\mu+(2q+1)k_n}$-measurable and $\widehat{\chi}'(\mathbf{j}')_{i+2+\mu+(2q+2)k_n}^n$ is $\mathcal{G}^{i+2+\mu+(2q+2)k_n}$-measurable that $\mathbb{E}(|\mathbb{E}(\delta_i^n \mid \mathcal{H}_i^n)|) \le K/k_n^v$. Then $\delta_i^n$ satisfies the assumptions of Lemma A.6, with $\widetilde{\mathcal{H}}_i^n = \mathcal{F}_{T(n,i)}^n$ and $w_n = 1 + \mu'' + (2q''+3)k_n$ and $a_n' = K$ and $a_n = K/k_n^v$, and the second part of (C.28) follows from (A.17). *Q.E.D.*

## REFERENCES


JACOD, J., AND P. PROTTER (2012): *Discretization of Processes*. Berlin: Springer-Verlag. [12]

JACOD, J., AND A. SHIRYAEV (2003): *Limit Theorems for Stochastic Processes* (Second Ed.). Berlin: Springer-Verlag. [11,17]


---


*Co-editor Lars Peter Hansen handled this manuscript.*

*Manuscript received 29 December, 2014; final version accepted 13 February, 2017; available online 2 March, 2017.*