Online Appendix for
"Testing for Jumps in a Discretely Observed Price Process with Endogenous Sampling Times"

Qiyuan Li* Yifan Li† Ingmar Nolte‡ Sandra Nolte§ Shifan Yu¶

This Version: November 1, 2025

This Online Appendix comprises two separate parts. Appendix A collects the proofs for all theoretical results presented in the main text. Appendix B contains additional results for both the Monte Carlo simulations (Section 4) and empirical applications (Section 5).

*Faculty of Business and Economics, University of Hong Kong; Email: qiyuanli@hku.hk.

†Alliance Manchester Business School, University of Manchester; Email: yifan.li@manchester.ac.uk.

‡Lancaster University Management School; Email: i.nolte@lancaster.ac.uk.

§Lancaster University Management School; Email: s.nolte@lancaster.ac.uk.

¶Corresponding author, Oxford-Man Institute of Quantitative Finance, University of Oxford; Email: shifan.yu@eng.ox.ac.uk.

Appendix A Proofs

A.1 Strong Approximation and Some Useful Lemmas

In this section, we establish a strong approximation result in the spirit of Chernozhukov et al. (2013, 2019), which couples the complicated observation scheme under Assumption 2 with the much simpler limiting observation scheme $(\breve{t}_{n,i})$ in Remark 3. Unless specifically stated, we assume $X(\omega)$ to be continuous, i.e., $\omega \in \Omega'$.

In all the sequel, the positive constants K, K', K'' may vary from line to line, but never depend on n, N, and $N^{(c)}$, and the various indices i, j. We use $\langle M, N \rangle$ to denote the quadratic covariation of M and N. When M and N are d- and r-dimensional, respectively, then $\langle M, N \rangle = (\langle M^i, N^j \rangle)_{1 \le i \le d, 1 \le j \le r}$ is a $(d \times r)$-dimensional process, and also $\langle M \rangle \equiv \langle M, M \rangle$.

Similar to the Assumption (S-HON) of Jacod et al. (2019), we impose the following stronger assumption without loss of generality by a standard localization procedure:

Assumption A.1. We have Assumptions 1 and 2 with $\tau_1 = \infty$. Moreover, the function δ and the processes μ, σ, λ, X are bounded, and we have $N \le K\Delta_n^{-1}$ and $\mathbb{E}[\Delta_{n,i}^p] \le K'\Delta_n^p$.

A.1.1 Intrinsic time. With an absolutely continuous time change from the calendar time t to intrinsic time $\tau(t)$:

$$t \to \tau(t) = \int_0^t \sigma_s^2 ds, \tag{A.1}$$

the intrinsic-time counterpart of X adapted to $(\mathcal{F}_t)_{t \ge 0}$ is

$$\widetilde{X}_{\tau(t)} = \widetilde{X}_0 + \int_0^{\tau(t)} \widetilde{\mu}_s ds + \widetilde{W}_{\tau(t)}, \tag{A.2}$$

where $\widetilde{\mu}$ is time-changed processes corresponding to μ in Eq. (1), and $\widetilde{W} = (\widetilde{W}_\tau)_{\tau \ge 0}$ is a Brownian motion evolving in intrinsic time. The relation $\widetilde{X}_{\tau(t)} = X_t$ holds for all t, and the τ-time process $\widetilde{X} = (\widetilde{X}_{\tau(t)})_{t \ge 0}$ is adapted to $(\widetilde{\mathcal{F}}_{\tau(t)})_{t \ge 0}$ with the τ-time σ-algebra satisfying $\widetilde{\mathcal{F}}_{\tau(t)} = \mathcal{F}_t$ (Lemma 1.2, Barndorff-Nielsen and Shiryaev, 2015). Particularly, when X is a calendar-time local martingale, $\widetilde{X}$ is an intrinsic-time Brownian motion (with an initial condition), which is implied by the Dambis-Dubins-Schwarz theorem. Following Mykland and Zhang (2009), the drift can be harmlessly assumed away, as the results on convergence in probability and stable convergence—established in Appendix A.3 and Appendix A.4, respectively—remain valid by a contiguity argument.

A.1.2 Observation schemes. We start with two sequences of observations of $X(\omega)$:

(I) Under Assumption 2: X_{t_i}, for all $i = 0, 1, 2, \ldots, N$,

(II) Equidistant observations in intrinsic time: $\widetilde{X}_{i\Delta_n}$, for all $i = 0, 1, 2, \ldots, N$.

For the ease of notation, we denote $t_i \equiv t_{n,i}$ under Assumption 2, and $\breve{t}_i \equiv \tau^{-1}(i\Delta_n)$. The increments between successive observations are denoted by

$$r_i = X_{t_i} - X_{t_{i-1}} \qquad \text{and} \qquad \breve{r}_i = X_{\breve{t}_i} - X_{\breve{t}_{i-1}}, \tag{A.3}$$

for all $i \in \{1, 2, \ldots, N\}$. Lemma A.1 of Jacod et al. (2017) indicates the sequence (I) is an $(\mathcal{F}_t^n)$-martingale with Gaussian increments. Different from the independent but not identically distributed increments r_i, the increments $\breve{r}_i$ are i.i.d. normal with zero mean and variance Δ_n, which make the sequence (II) a homogenous Gaussian random walk.

Remark A.1. We assume both sequences have the same number $N \equiv N_1^n$ of observations. Assumption A.1 and Eq. (6) indicate that $T_n = \tau^{-1}(N\Delta_n)$ is bounded and $T_n \xrightarrow{\mathbb{P}} 1$ as $n \to \infty$. Moreover, by the triangle inequality and law of iterated expectations, Assumption 2 further implies $\mathbb{E}[|N\Delta_n - \tau(1)|] \le K\Delta_n$, hence $|T - 1| = O_p(\Delta_n)$. That being said, the probability of jump occurrence in the "differenced part" of observation interval is negligible; see more discussions in Section 2.3 of Aït-Sahalia and Jacod (2009). To fix ideas, let $\delta \to 0$, then $\mathbb{P}(X_t'' \neq 0 \text{ for some } t \in [0,T] \triangle [0,1]) \le \mathbb{P}(|T-1| > \delta) + \mathbb{P}(X_t'' \neq 0 \text{ for some } t \in [1-\delta, 1+\delta]) = o(1)$, since we only assume potential jumps in $(0,1)$.

For each sequence of observations, we conduct the PDS with the barrier width $c = m\sqrt{\Delta_n}$. We denote the PDS returns from each sequence of sampled observations by $(r_i^{(c)})_{i\in\{1,2,\ldots,N^{(c)}\}}$ and $(\breve{r}_i^{(c)})_{i\in\{1,2,\ldots,\breve{N}^{(c)}\}}$, respectively. Note that the PDS returns $\breve{r}_i^{(c)}$ are i.i.d., as implied by the strong Markov property of the Gaussian random walk (II) and the symmetric feature of the stopping rule in Eq. (7).

A.1.3 Strong Approximation.

We define two supremum processes $(Y_j)_{1\le j\le N}$ and $(\breve{Y}_j)_{1\le j\le N}$:

$$Y_j = \sup_{1\le i\le j} |X_{t_i}| \qquad \text{and} \qquad \breve{Y}_j = \sup_{1\le i\le j} |X_{\breve{t}_i}|. \tag{A.4}$$

Lemma A.1. For any fixed $1 \le j \le N$, it holds for the supremum processes that

$$|Y_j - \breve{Y}_j| = O_p(j^2 \Delta_n^{1+\kappa/2} \sqrt{L_n}), \tag{A.5}$$

where for the ease of notation, $L_n \equiv \log N \asymp \log(\Delta_n^{-1})$.

Proof. Let $\mathcal{D}_n \equiv \sigma(\Delta_{n,1}, \Delta_{n,2}, \ldots)$ denote the σ-algebra generated by observation times. Note that by the triangle inequality of ℓ_∞-norm

$$\begin{aligned} |Y_j - \breve{Y}_j| &= \left| \max_{1\le i\le j} |X_{t_i}| - \max_{1\le i\le j} |X_{\breve{t}_i}| \right| \\ &\le \max_{1\le i\le j} |X_{t_i} - X_{\breve{t}_i}|. \end{aligned} \tag{A.6}$$

Note that by definition, we have with probability approaching 1,

$$
\begin{aligned}
X_{t_i} - X_{\breve{t}_i} &= \sum_{\ell=1}^{i}\Big(\int_{t_{\ell-1}}^{t_\ell} \sigma_s dW_s - \int_{\breve{t}_{\ell-1}}^{\breve{t}_\ell} \sigma_s dW_s\Big) \\
&= \sum_{\ell=1}^{i}\big(\sigma_{t_{\ell-1}}(W_{t_\ell} - W_{t_{\ell-1}}) - \sigma_{\breve{t}_{\ell-1}}(W_{\breve{t}_\ell} - W_{\breve{t}_{\ell-1}})\big) \\
&+ \sum_{\ell=1}^{i}\Big(\int_{t_{\ell-1}}^{t_\ell} (\sigma_s - \sigma_{t_{\ell-1}})dW_s - \int_{\breve{t}_{\ell-1}}^{\breve{t}_\ell} (\sigma_s - \sigma_{\breve{t}_{\ell-1}})dW_s\Big) \\
&\equiv A_{n,i}^{(1)} + A_{n,i}^{(2)}.
\end{aligned} \tag{A.7}
$$

For the first term, by the maximal inequality of Gaussian variables, we have

$$
\mathbb{E}\Big[\max_{1\le i\le j} |A_{n,i}^{(1)}|\Big|\mathcal{D}_n\Big] \le K\sqrt{L_n \max_{1\le i\le j}\Big|\sum_{\ell=1}^{i}(\Delta_{n,\ell}\lambda_{t_{\ell-1}} - \Delta_n)\Big|}. \tag{A.8}
$$

For the right hand side, note that by the triangle inequality and Assumption 2 (ii),

$$
\max_{1\le i\le j}\Big|\sum_{\ell=1}^{i}\mathbb{E}\big[|\Delta_{n,\ell}\lambda_{t_{\ell-1}} - \Delta_n|\big|\mathcal{F}_{\ell-1}^n\big]\Big| \le Kj\Delta_n^{2+\kappa}. \tag{A.9}
$$

Combining Eq. (A.8) and Eq. (A.9), it follows the law of iterated expectation that

$$
\max_{1\le i\le j} |A_{n,i}^{(1)}| = O_p(j\Delta_n^{1+\kappa/2}\sqrt{L_n}). \tag{A.10}
$$

For the second term in Eq. (A.7), by the maximal inequality, we have

$$
\mathbb{E}\Big[\max_{1\le i\le j} |A_{n,i}^{(2)}|\Big] \le Kj \max_{1\le i\le j} \mathbb{E}[|A_{n,i}^{(2)}|] \le Kj^2\Delta_n^{3/2+\kappa/2}, \tag{A.11}
$$

where the last step is by the Burkholder-Davis-Gundy inequality and smoothness of σ regulated by Assumption 1 (ii). The proof of required statement is completed by the triangle inequality and Eqs. (A.6), (A.7), (A.10) and (A.11).

□

We consider the first sampled observation times for both sequences:

$$
\Pi_1^{(c)} = \inf\{i : |X_{t_i} - X_0| \ge c\} \qquad \text{and} \qquad \breve{\Pi}_1^{(c)} = \inf\{i : |\widetilde{X}_{i\Delta_n} - \widetilde{X}_0| \ge c\}, \tag{A.12}
$$

which means that the $\Pi_1^{(c)}$-th and the $\breve{\Pi}_1^{(c)}$-th observations in (I) and (II), respectively, are the first to breach the symmetric double barrier. Lemma A.2 indicates that the first exit times of both sequences coincide with probability approaching 1 under infill asymptotics.

Lemma A.2. For $c = m\sqrt{\Delta_n}$, let $\overline{N}^{(c)} \equiv N^{(c)} \wedge \breve{N}^{(c)}$.

(i) For all integer $p \geq 1$, $\mathbb{E}[(\breve{\Pi}_1^{(c)})^p] < \infty$.

(ii) The first exit times for both sequences (I) and (II) satisfy

$$\mathbb{P}\Big(\max_{1\leq i\leq \overline{N}^{(c)}} |\Pi_i^{(c)} - \breve{\Pi}_i^{(c)}| \geq 1\Big) \leq K\Delta_n^{\kappa/2}\sqrt{L_n}. \tag{A.13}$$

Proof. (i) Note that $\breve{\Pi}_1^{(c)}$ has the same distribution as the number of steps for a standard Gaussian random walk $(Z_i)_{i=1,2,\ldots}$ to exit the double barrier $(-m, m)$. Let $h = \inf\{\tau : \widetilde{W}_\tau \notin (-m, m)\}$ denote the first exit time of the time-changed Brownian motion $\widetilde{W}$ from $(-m, m)$, then it is clear that $\breve{\Pi}_1^{(c)} - 1 \leq h$ by the continuity of Brownian motion, thus $\mathbb{E}[(\breve{\Pi}_1^{(c)} - 1)^p] \leq \mathbb{E}[h^p]$ for all $p > 0$. The Laplace transform of h is well-known in the literature, see, e.g., Eq. (3.0.1) in Borodin and Salminen (2002): $\mathbb{E}[e^{-\lambda h}] = \cosh^{-1}\sqrt{2\lambda}m$, and its Maclaurin series implies that $\mathbb{E}[h^p] < \infty$ for all integer $p \geq 1$. This completes the proof.

(ii) We start from the first term. By definition, we have

$$\mathbb{P}(\Pi_1^{(c)} \geq k) = \mathbb{P}(Y_k \leq c) \quad \text{and} \quad \mathbb{P}(\breve{\Pi}_1^{(c)} \geq k) = \mathbb{P}(\breve{Y}_k \leq c). \tag{A.14}$$

Let $\epsilon > 0$ be a positive number that can be arbitrarily small but not depend on N, it follows Lemma A.1 and the Markov inequality that

$$\begin{aligned}
\mathbb{P}(\Pi_1^{(c)} - \breve{\Pi}_1^{(c)} \geq 1) &= \sum_{k=1}^N \mathbb{P}(\breve{\Pi}_1^{(c)} = k)\mathbb{P}(\Pi_1^{(c)} > k | \breve{\Pi}_1^{(c)} = k) \\
&\leq \sum_{k=1}^N \mathbb{P}(\breve{\Pi}_1^{(c)} = k)\mathbb{P}(\breve{Y}_k - Y_k > \epsilon) \\
&\leq K\Delta_n^{1+\kappa/2}\sqrt{L_n}\Big[\sum_{k=1}^N k^2\mathbb{P}(\breve{\Pi}_1^{(c)} = k)\Big] \\
&\leq K\Delta_n^{1+\kappa/2}\sqrt{L_n},
\end{aligned} \tag{A.15}$$

where the last line uses $\sum_{k=1}^N k^2\mathbb{P}(\breve{\Pi}_1^{(c)} = k) \leq \mathbb{E}[(\breve{\Pi}_1^{(c)})^2] \leq K$ by Lemma A.2 (i). Similarly, we can also show

$$\mathbb{P}(\breve{\Pi}_1^{(c)} - \Pi_1^{(c)} \geq 1) \leq K\Delta_n^{1+\kappa/2}\sqrt{L_n}. \tag{A.16}$$

Combining above results, we have

$$\mathbb{P}(|\Pi_1^{(c)} - \breve{\Pi}_1^{(c)}| \geq 1) \leq K\Delta_n^{1+\kappa/2}\sqrt{L_n}. \tag{A.17}$$

Let $E_i = \{|\Pi_i^{(c)} - \breve{\Pi}_i^{(c)}| \geq 1 \text{ and } \Pi_j^{(c)} = \breve{\Pi}_j^{(c)} \text{ for all } 1 \leq j \leq i-1\} \in \mathcal{F}_{t_i}$ be the event that the first discrepancy occurs at step i, and $E_1 = \{|\Pi_1^{(c)} - \breve{\Pi}_1^{(c)}| \geq 1\}$. By the renewal property, we have $\Pi_i^{(c)} - \breve{\Pi}_i^{(c)} \stackrel{\mathcal{L}}{=} \Pi_1^{(c)} - \breve{\Pi}_1^{(c)}$ conditional on $\Pi_j^{(c)} = \breve{\Pi}_j^{(c)}$ for all $1 \leq j \leq i-1$. Using the same argument as in driving Eq. (A.17), we can show that $\mathbb{P}(E_i|\mathcal{F}_{t_{i-1}})$ is bounded by $K\Delta_n^{1+\kappa/2}\sqrt{L_n}$. Therefore, it holds that

$$\begin{aligned}\mathbb{P}\Big(\max_{1\le i\le \overline{N}^{(c)}} |\Pi_i^{(c)} - \breve{\Pi}_i^{(c)}| \ge 1\Big) &\le \mathbb{P}\left(\bigcup_{i=1}^{\overline{N}^{(c)}} E_i\right) \le \sum_{i=1}^{\infty} \mathbb{E}\big[\mathbb{1}_{E_i}\mathbb{1}_{\{i\le \overline{N}^{(c)}\}}\big] \\ &= \sum_{i=1}^{\infty} \mathbb{E}\big[\mathbb{E}[\mathbb{1}_{E_i}|\mathcal{F}_{t_{i-1}}]\mathbb{1}_{\{i\le \overline{N}^{(c)}\}}\big] \le K\Delta_n^{1+\kappa/2}\sqrt{L_n}\,\mathbb{E}[\overline{N}^{(c)}] \\ &\le K'\Delta_n^{\kappa/2}\sqrt{L_n},\end{aligned} \tag{A.18}$$

which follows from the fact that $\mathbb{1}_{\{i\le \overline{N}^{(c)}\}} \in \mathcal{F}_{t_{i-1}}$, and $\mathbb{E}[\overline{N}^{(c)}] \le K\Delta_n^{-1}$; see Lemma A.6 (ii). The proof is then completed. □

Lemma A.3. (Strong Approximation for Sampled Returns) It holds that

$$\mathbb{P}\Big(\max_{1\le i\le \overline{N}^{(c)}} |r_i^{(c)} - \breve{r}_i^{(c)}| > K\Delta_n^{1+\kappa/8}\Big) \le K'\Delta_n^{\kappa/8}\sqrt{L_n}. \tag{A.19}$$

Proof. It follows from the maximal inequality of Gaussian variables that

$$\mathbb{E}\Big[\max_{1\le i\le N} |r_i - \breve{r}_i|\Big|\mathcal{D}_n\Big] \le K\sqrt{L_n}\max_{1\le i\le N}\sqrt{\big|\mathbb{E}[\Delta_{n,i}\lambda_{t_{i-1}}|\mathcal{F}_{i-1}^n] - \Delta_n\big|} \le K\Delta_n^{1+\kappa/2}\sqrt{L_n}. \tag{A.20}$$

Let $E_n \equiv \{\Pi_i^{(c)} = \breve{\Pi}_i^{(c)} \text{ for all } 1\le i\le N^{(c)} = \breve{N}^{(c)}\}$, we have $\mathbb{P}(E_n^{\complement}) \le K\Delta_n^{\kappa/2}\sqrt{L_n}$ by Lemma A.2 (ii). Note that by the maximal inequality, we have for any $p>1$,

$$\mathbb{E}\Big[\max_{1\le i\le \breve{N}^{(c)}} |\breve{\Pi}_i^{(c)} - \breve{\Pi}_{i-1}^{(c)}|^p\Big] \le \breve{N}^{(c)}\max_{1\le i\le \breve{N}^{(c)}} \mathbb{E}[|\breve{\Pi}_i^{(c)} - \breve{\Pi}_{i-1}^{(c)}|^p] \le K_p\Delta_n^{-1}, \tag{A.21}$$

where the last step is by Lemma A.2 (i). Taking $p > 4/\kappa$ gives

$$\mathbb{E}\Big[\max_{1\le i\le \breve{N}^{(c)}} |\breve{\Pi}_i^{(c)} - \breve{\Pi}_{i-1}^{(c)}|^2\Big] \le K\Delta_n^{-\kappa/2}. \tag{A.22}$$

Moreover, by Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}\mathbb{E}\Big[\max_{1\le i\le \breve{N}^{(c)}} |r_i^{(c)} - \breve{r}_i^{(c)}|\Big|E_n\Big] &\le \sqrt{\mathbb{E}\Big[\max_{1\le i\le \breve{N}^{(c)}} |\breve{\Pi}_i^{(c)} - \breve{\Pi}_{i-1}^{(c)}|^2\Big]\mathbb{E}\Big[\max_{1\le \ell\le n} |r_\ell - \breve{r}_\ell|^2\Big]} \\ &\le K\Delta_n^{1+\kappa/4}\sqrt{L_n}.\end{aligned} \tag{A.23}$$

Therefore, we have

$$
\begin{aligned}
&\mathbb{P}\Big(\max_{1\le i\le \overline{N}^{(c)}} |r_i^{(c)} - \breve{r}_i^{(c)}| > K\Delta_n^{1+\kappa/8}\Big) \\
\le\ &\mathbb{P}\Big(\max_{1\le i\le \breve{N}^{(c)}} |r_i^{(c)} - \breve{r}_i^{(c)}| > K\Delta_n^{1+\kappa/8}\Big| E_n\Big) + \mathbb{P}(E_n^{\complement}) \\
\le\ &K'(\Delta_n^{\kappa/8}\sqrt{L_n} + \Delta_n^{\kappa/2}\sqrt{L_n}).
\end{aligned}
\tag{A.24}
$$

This completes the proof.

□

Lemma A.3 shows the statistics constructed from sampled returns under observation schemes (I) and (II) are equivalent up to a $\Delta_n^{-1-\kappa/8}$ normalization, which is sufficient for the $c^{-1} \asymp \Delta_n^{-1/2}$ or $\sqrt{\breve{N}^{(c)}} \asymp \Delta_n^{-1/2}$ order in conventional CLT. The requirement is only $\kappa > 0$.

The above type of strong approximation results are similarly used in, e.g., the proof of Theorem 5.1 in Chernozhukov et al. (2013) and the proof of Theorem 4.3 in Chernozhukov et al. (2019). It allows us to focus on the limiting behavior of functionals of $(|\breve{r}_i^{(c)}|/c)^2$, the result can be sufficiently extended to those of $(|r_i^{(c)}|/c)^2$. To fix ideas, consider a possibly multi-dimensional Lipschitz function $f(\cdot)$. Suppose that

$$
\frac{1}{\breve{N}^{(c)}}\sum_{i=1}^{\breve{N}^{(c)}} f\Big(\frac{(\breve{r}_i^{(c)})^2}{c^2}\Big) \xrightarrow{\mathbb{P}} \mu_f, \quad \text{and} \quad \frac{1}{\sqrt{\breve{N}^{(c)}}}\sum_{i=1}^{\breve{N}^{(c)}} \Big(f\Big(\frac{(\breve{r}_i^{(c)})^2}{c^2}\Big) - \mu_f\Big) \xrightarrow{\mathcal{L}} \mathcal{N}(0,\Sigma_f). \tag{A.25}
$$

Let $E_n' \equiv \{\Pi_i^{(c)} = \breve{\Pi}_i^{(c)} \text{ for all } 1 \le i \le \breve{N}^{(c)}\} \cap \{\max_{1\le i\le \breve{N}^{(c)}} |(r_i^{(c)})^2 - (\breve{r}_i^{(c)})^2|/c^2 > K\Delta_n^{1/2+\kappa/16}\}$. Note that $a^2 - b^2 = (a-b)^2 + 2b(a-b)$, it follows from triangle inequality that

$$
\max_{1\le i\le \breve{N}^{(c)}} |(r_i^{(c)})^2 - (\breve{r}_i^{(c)})^2| \le \Big(\max_{1\le i\le \breve{N}^{(c)}} |r_i^{(c)} - \breve{r}_i^{(c)}|\Big)^2 + 2\Big(\max_{1\le i\le \breve{N}^{(c)}} |\breve{r}_i^{(c)}|\Big)\Big(\max_{1\le i\le \breve{N}^{(c)}} |r_i^{(c)} - \breve{r}_i^{(c)}|\Big). \tag{A.26}
$$

Note that $\max_{1\le i\le \breve{N}^{(c)}} |\breve{r}_i^{(c)}| = O_p(\Delta_n^{1/2}\sqrt{L_n}) = o_p(\Delta_n^{1/2-\kappa/16})$ by the maximal inequality of sub-Gaussian variables. Then it follows from Lemma A.2 (ii), A.3, and Eq. (A.26) that $\mathbb{P}(E_n') \ge 1 - K\Delta_n^{\kappa/8}\sqrt{L_n}$. Therefore, for each $\varepsilon > 0$,

$$
\begin{aligned}
&\mathbb{P}\Big(\Big\|\frac{1}{N^{(c)}}\sum_{i=1}^{N^{(c)}} f\Big(\frac{(r_i^{(c)})^2}{c^2}\Big)-\mu_f\Big\|>\varepsilon\Big)\\
&\le \mathbb{P}\Big(\Big\|\frac{1}{\breve{N}^{(c)}}\sum_{i=1}^{\breve{N}^{(c)}} f\Big(\frac{(\breve{r}_i^{(c)})^2}{c^2}\Big)-\mu_f\Big\|>\frac{\varepsilon}{2}\Big)+\mathbb{P}\Big(\Big\|\frac{1}{N^{(c)}}\sum_{i=1}^{N^{(c)}} f\Big(\frac{(r_i^{(c)})^2}{c^2}\Big)-\frac{1}{\breve{N}^{(c)}}\sum_{i=1}^{\breve{N}^{(c)}} f\Big(\frac{(\breve{r}_i^{(c)})^2}{c^2}\Big)\Big\|>\frac{\varepsilon}{2}\Big)\\
&\le \mathbb{P}\Big(\Big\|\frac{1}{\breve{N}^{(c)}}\sum_{i=1}^{\breve{N}^{(c)}} f\Big(\frac{(\breve{r}_i^{(c)})^2}{c^2}\Big)-\mu_f\Big\|>\frac{\varepsilon}{2}\Big)+\mathbb{P}\Big(K\max_{1\le i\le \breve{N}^{(c)}}\frac{|(r_i^{(c)})^2-(\breve{r}_i^{(c)})^2|}{c^2}>\frac{\varepsilon}{2}\Big|E_n'\Big)+\mathbb{P}(E_n'^{\complement})\\
&= \mathbb{P}\Big(\Big\|\frac{1}{\breve{N}^{(c)}}\sum_{i=1}^{\breve{N}^{(c)}} f\Big(\frac{(\breve{r}_i^{(c)})^2}{c^2}\Big)-\mu_f\Big\|>\frac{\varepsilon}{2}\Big)+K\Delta_n^{\kappa/8}\sqrt{L_n}.
\end{aligned}
\tag{A.27}
$$

Let $Z \sim \mathcal{N}(0,\Sigma_f)$. For each $A \subset \mathbb{R}^{\dim(f)}$ and $\varepsilon > 0$, let $A^\varepsilon \equiv \{x \in \mathbb{R}^{\dim(f)} : \inf_{y\in A}\|x-y\| \le \varepsilon\}$ denote the ε-enlargement of A, then we have

$$
\begin{aligned}
&\mathbb{P}\Big(\frac{1}{\sqrt{N^{(c)}}}\sum_{i=1}^{N^{(c)}}\Big(f\Big(\frac{(r_i^{(c)})^2}{c^2}\Big)-\mu_f\Big)\in A\Big)\\
&\le \mathbb{P}\Big(\frac{1}{\sqrt{\breve{N}^{(c)}}}\sum_{i=1}^{\breve{N}^{(c)}}\Big(f\Big(\frac{(\breve{r}_i^{(c)})^2}{c^2}\Big)-\mu_f\Big)\in A^\varepsilon\Big)+\mathbb{P}\Big(\Big\|\frac{1}{\sqrt{N^{(c)}}}\sum_{i=1}^{N^{(c)}} f\Big(\frac{(r_i^{(c)})^2}{c^2}\Big)-\frac{1}{\sqrt{\breve{N}^{(c)}}}\sum_{i=1}^{\breve{N}^{(c)}} f\Big(\frac{(\breve{r}_i^{(c)})^2}{c^2}\Big)\Big\|>\varepsilon\Big)\\
&\le \mathbb{P}\Big(\frac{1}{\sqrt{\breve{N}^{(c)}}}\sum_{i=1}^{\breve{N}^{(c)}}\Big(f\Big(\frac{(\breve{r}_i^{(c)})^2}{c^2}\Big)-\mu_f\Big)\in A^\varepsilon\Big)+\mathbb{P}\Big(K\sqrt{\breve{N}^{(c)}}\max_{1\le i\le \breve{N}^{(c)}}\frac{|(r_i^{(c)})^2-(\breve{r}_i^{(c)})^2|}{c^2}>\varepsilon\Big|E_n'\Big)+\mathbb{P}(E_n'^{\complement})\\
&= \mathbb{P}(Z\in A)+\mathbb{P}(Z\in A^\varepsilon\setminus A)+K\Delta_n^{\kappa/8}\sqrt{L_n}.
\end{aligned}
\tag{A.28}
$$

Taking $\varepsilon \to 0$, the right-hand side becomes $\mathbb{P}(Z \in A) + o(1)$. Similarly, one can show

$$
\mathbb{P}\Big(\frac{1}{\sqrt{N^{(c)}}}\sum_{i=1}^{N^{(c)}}\Big(f\Big(\frac{(r_i^{(c)})^2}{c^2}\Big)-\mu_f\Big)\in A\Big)\ge \mathbb{P}(Z\in A)-o(1),
\tag{A.29}
$$

which is the desired result.

A.1.4 Impact of Small Jumps. Under Assumption 1, we consider the jump component of X in the following form, which is valid as the jumps are of finite variation:

$$
X_t'' = \int_0^t\int_{\mathbb{R}} \delta(s,x)\underline{p}(ds,dx),
\tag{A.30}
$$

where $\delta(\omega, t, x)$ on $\Omega \times \mathbb{R}_+ \times \mathbb{R}$ is predictable, $\underline{p}(dt, dx)$ is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}$ with a compensator $\underline{q}(dt, dx) = dt \otimes \lambda(dx)$, and λ is a σ-finite measure on $\mathbb{R}_+$. Moreover, we have

$$\lim_{u \to 0+} u^r \int_{\{|f_m| \geq u\}} \lambda(dx) \leq \int_{\{|f_m| \geq u\}} |f_m|^r \lambda(dx) < \infty, \tag{A.31}$$

which implies, as $u \to 0$,

$$\lambda(\{x : |f_m(x)| \geq u\}) \equiv \int_{\{|f_m| \geq u\}} \lambda(dx) = O(u^{-r}). \tag{A.32}$$

We split the jumps into "big" and "small" ones by selecting a sequence (u_n) of positive real numbers satisfying:

$$\frac{u_n}{\sqrt{\Delta_n}} \to \infty \qquad \text{and} \qquad u_n \Delta_n^{\beta - 1/2} \to 0, \tag{A.33}$$

for any $0 < \beta \leq 1/2$. Then we rewrite the Itô semimartingale $X = X' + X''$ in Eq. (1) as:

$$X_t = X'_t + \underbrace{\int_0^t \int_{\{|\delta(s,x)| \geq u_n\}} \delta(s,x) \underline{p}(ds, dx)}_{\text{"Big" Jumps: } J^n_{1,t}} + \underbrace{\int_0^t \int_{\{|\delta(s,x)| < u_n\}} \delta(s,x) \underline{p}(ds, dx)}_{\text{"Small" Jumps: } J^n_{2,t}}, \tag{A.34}$$

where the component X'' is partitioned into two n-dependent processes J^n_1 and J^n_2. This "optimal" cutoff level $u_n \asymp \Delta_n^{\varpi}$ with ϖ arbitrarily close to but below $1/2$ separates all jumps that either prevail over or are diluted within Brownian increments.

Next, we show that the existence of small jumps in J^n_2 has no impact on Lemma A.1.

Lemma A.4. For the purely discontinuous process J^n_2 defined in Eq. (A.34), with the sequence (u_n) of thresholds satisfying Eq. (A.33), it holds that for all $p \geq 1$,

$$\sup_{t_{i-1} \leq s \leq t_i} |J^n_{2,s} - J^n_{2,t_{i-1}}|^p = O_p(\Delta_n u_n^{p-r}). \tag{A.35}$$

Proof. Following Assumption A.1, we have Assumption 1 (v) with $\tau_1 = \infty$ without loss of generality by a standard localization procedure, such that $|\delta(\omega, t, x)| \wedge 1 \leq f(x)$ holds uniformly on $\Omega \times \mathbb{R}_+ \times \mathbb{R}$. We start with the notation for a local p-th order variation of small jumps, which resembles the first quantity in Eq. (2.1.35) of Jacod and Protter (2012):

For some $p \geq 1$, we define

$$\widehat{\delta}_{p,i} = \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} ds \int_{\{|\delta(s,x)| < u_n\}} |\delta(s,x)|^p \lambda(dx). \tag{A.36}$$

For all $1 \leq i \leq n$, it holds that

$$\begin{aligned}\mathbb{E}[\widehat{\delta}_{p,i}|\mathcal{F}_{t_{i-1}}] &\leq \frac{1}{t_i - t_{i-1}}\mathbb{E}\left[\int_{t_{i-1}}^{t_i} ds \int_{\{|\delta(s,x)|<u_n\}} |\delta(s,x)|^p \lambda(dx)\middle|\mathcal{F}_{t_{i-1}}\right] \\ &\leq \frac{1}{t_i - t_{i-1}}\mathbb{E}\left[\int_{t_{i-1}}^{t_i} ds \int_{\{|\delta(s,x)|<u_n\}} u_n^{p-r}|\delta(s,x)|^r \lambda(dx)\middle|\mathcal{F}_{t_{i-1}}\right] \\ &\leq u_n^{p-r}\int_{\mathbb{R}} |f(x)|^r \lambda(dx),\end{aligned} \tag{A.37}$$

since $\delta(\omega, t, x)$ is bounded by the deterministic function $f(x)$, and $t_i - t_{i-1}$ is independent of $\mathcal{F}_{t_{i-1}}$. Denote the integral as a constant $C_r = \int_{\mathbb{R}} |f(x)|^r \lambda(dx)$, we have

$$\mathbb{E}[\widehat{\delta}_{p,i}|\mathcal{F}_{t_{i-1}}] \leq C_r u_n^{p-r}. \tag{A.38}$$

Similarly, for another conditional expectation $\mathbb{E}[\widehat{\delta}_{1,i}^p|\mathcal{F}_{t_{i-1}}]$, we have

$$\begin{aligned}\mathbb{E}[\widehat{\delta}_{1,i}^p|\mathcal{F}_{t_{i-1}}] &= \frac{1}{(t_i - t_{i-1})^p}\mathbb{E}\left[\left(\int_{t_{i-1}}^{t_i} ds \int_{\{|\delta(s,x)|<u_n\}} |\delta(s,x)|\lambda(dx)\right)^p \middle|\mathcal{F}_{t_{i-1}}\right] \\ &\leq \frac{1}{(t_i - t_{i-1})^p}\mathbb{E}\left[\left(\int_{t_{i-1}}^{t_i} ds \int_{\{|\delta(s,x)|<u_n\}} u_n^{1-r}|\delta(s,x)|^r \lambda(dx)\right)^p \middle|\mathcal{F}_{t_{i-1}}\right] \\ &\leq u_n^{p(1-r)}\left(\int_{\mathbb{R}} |f(x)|^r \lambda(dx)\right)^p \\ &\leq C_r^p u_n^{p(1-r)}.\end{aligned} \tag{A.39}$$

Then by Lemma 2.1.7 of Jacod and Protter (2012), with the bounds for both $\mathbb{E}[\widehat{\delta}_{p,i}|\mathcal{F}_{t_{i-1}}]$ and $\mathbb{E}[\widehat{\delta}_{1,i}^p|\mathcal{F}_{t_{i-1}}]$ in Eqs. (A.38) and (A.39), respectively, we have for all $p \geq 1$,

$$\begin{aligned}\mathbb{E}\left[\sup_{t_{i-1}\leq s\leq t_i} |J_{2,s}^n - J_{2,t_{i-1}}^n|^p \middle|\mathcal{F}_{t_{i-1}}\right] &\leq K\Big(\Delta_n \mathbb{E}[\widehat{\delta}_{p,i}|\mathcal{F}_{t_{i-1}}] + \Delta_n^p \mathbb{E}[\widehat{\delta}_{1,i}^p|\mathcal{F}_{t_{i-1}}]\Big) \\ &\leq K\Big(C_r \Delta_n u_n^{p-r} + C_r^p \Delta_n^p u_n^{p(1-r)}\Big) \\ &\leq K'\Delta_n u_n^{p-r},\end{aligned} \tag{A.40}$$

where the latter term $KC_r^p\Delta_n^p u_n^{p(1-r)}$ reduces to $K'\Delta_n u_n^{p-r}$ since $1 < 1+\varpi(p-r) < p < p+p\varpi(1-r)$ for some ϖ slightly smaller than 1/2. The desired result in Lemma A.4 follows from the law of iterated expectation and Markov's inequality. □

To examine the impact of small jumps on Lemma A.1, we rewrite the supremum processes into

$$Y_j = \sup_{1\leq i\leq j} |X'_{t_i} + J_{2,t_i}^n| \qquad \text{and} \qquad \breve{Y}_j = \sup_{1\leq i\leq j} |X'_{t_i}|. \tag{A.41}$$

In this case, by the triangle inequality, we have

$$|Y_j - \check{Y}_j| = \Big| \max_{1\le i\le j} |X'_{t_i} + J^n_{2,t_i}| - \max_{1\le i\le j} |X'_{\check{t}_i}| \Big| \le \max_{1\le i\le j} |X'_{t_i} - X'_{\check{t}_i}| + \max_{1\le i\le j} |J^n_{2,t_i}|, \tag{A.42}$$

where, by Lemma A.4,

$$\max_{1\le i\le j} |J^n_{2,t_i}| \le \sum_{i=1}^{j} \sup_{(i-1)\Delta_n \le s\le i\Delta_n} |J^n_{2,s} - J^n_{2,(i-1)\Delta_n}| = O_p(j\Delta_n u_n^{1-r}). \tag{A.43}$$

For some $0<\kappa<1-r$ such that $\kappa/2(1-r)\le \varpi < 1/2$, we have $\max_{1\le i\le j} |J^n_{2,t_i}| = O_p(j\Delta_n^{1+\kappa/2})$. Therefore, for any fixed $1\le j\le N$, the small jumps in J^n_2 do not affect the order of coupling error in Lemma A.1 and any subsequent results in Appendix A.1.3.

A.2 Properties of Functions $h_2(m)$ and $\overline{h}_{2,\epsilon}(m)$

In this section, we prove some properties of the functions $h_2(\cdot)$ and $\overline{h}_{2,\epsilon}(\cdot)$ defined in Eq. (10), which are important for the construction of our test statistic.

Proposition 1. The functions $h_2(\cdot)$ and $\overline{h}_{2,\epsilon}(\cdot)$ are invertible and differentiable with nonvanishing derivatives.

Proof. For the standard Gaussian random walk Z, let $\Pi_1^{(m)} \equiv \min\{n\ge 1 : |Z_n| > m\}$ be the first passage time across $\pm m$, then by definition $Z_1^{(m)} = Z_{\Pi_1^{(m)}}$. We start with $h_{2,\epsilon}(\cdot)$, it follows Fubini's theorem that

$$\begin{aligned} \overline{h}_{2,\epsilon}(m) &= \mathbb{E}\Big[\frac{|Z_1^{(m)}|^2}{m^2} \wedge (1+\epsilon)^2\Big] = \int_0^{(1+\epsilon)^2} \mathbb{P}(|Z_1^{(m)}| > m\sqrt{u})du \\ &= 1 + 2\int_0^{\epsilon} (1+v)\mathbb{P}\big(|Z_1^{(m)}| > m(1+v)\big)dv, \end{aligned} \tag{A.44}$$

where the second line follows from the change of variable $u=(1+v)^2$ and the fact that $\mathbb{P}\big(|Z_1^{(m)}| > m(1+v)\big) = 1$ for $v\in[-1,0)$. As m increases, the annulus $\big(m, m(1+v)\big)$ widens; then it becomes strictly harder for the walk Z to exit $\pm m$ and simultaneously cross $\pm m(1+v)$ at the same first-exit time. Formally, for fixed $v>0$ and $0<\delta<mv$, the absolute continuous increments implies that $|Z_1^{(m)}|$ has a density, and thus

$$\mathbb{P}\big(m(1+v) < |Z_1^{(m)}| < (m+\delta)(1+v)\big) > 0. \tag{A.45}$$

On this event, we have $|Z_1^{(m)}| > m+\delta$ but $|Z_1^{(m+\delta)}| < (m+\delta)(1+v)$, and therefore

$$\mathbb{P}\big(|Z_1^{(m)}| > m(1+v)\big) > \mathbb{P}\big(|Z_1^{(m+\delta)}| > (m+\delta)(1+v)\big), \tag{A.46}$$

i.e., the tail probability $\mathbb{P}\big(|Z_1^{(m)}| > m(1+v)\big)$ is strictly decreasing in m. This further implies $\overline{h}_{2,\epsilon}(\cdot)$ in Eq. (A.44) is also strictly decreasing, the invertibility readily follows.

The differentiability relies on an integral representation of tail probability using renewal identity. Let $(H_n)_{n\geq 1}$ denote the strong ascending ladder heights, with F_H its distribution. By the standard Wiener–Hopf renewal representation for ladder heights (see Chapter XVIII.3 in Feller, 1991), F_H is absolutely continuous with density $f_H(x) = \int_{[0,\infty)} \varphi(x+y)U^-(dy)$ where U^- is the descending ladder height renewal measure. Further denote U_m^+ the renewal measure killed upon hitting $-m$, i.e., $U_m^+(E) \equiv \sum_{n\geq 0} \mathbb{P}\big(\sum_{i=1}^n H_i \in E, \tau_m < \tau_{-m}\big)$ for Borel E. Then by symmetry and the renewal identity (see, e.g., Section 2.6 in Gut, 2009),

$$\mathbb{P}\big(|Z_1^{(m)}| > m(1+v)\big) = 2\int_{[0,m]} \big[1 - F_H\big(m(1+v)-x\big)\big] U_m^+(dx) \equiv \phi_v(m). \tag{A.47}$$

Note that U_m^+ has an atom $U_m^+(\{0\}) = \mathbb{P}(\tau_m < \tau_{-m}) = \frac{1}{2}$, by Theorem VII.1.1 in Asmussen (2003), it admits Stone's decomposition on compact sets, i.e., $U_m^+(dx) = \frac{1}{2}\delta_0(dx) + u_m^+(x)dx$ for a bounded continuous density $u_m^+(\cdot)$. It follows Leibniz rule that

$$\phi_v'(m) = 2\left[u_m^+(m)\big(1 - F_H(mv)\big) - (1+v)\int_{[0,m]} f_H\big(m(1+v)-x\big)U_m^+(dx)\right], \tag{A.48}$$

where we use the fact that $\partial_m U_m^+$ vanishes except for the boundary $x = m$. Combining Eqs. (A.44) and (A.48) gives the differentiability and $\overline{h}'_{2,\epsilon}(m) = 2\int_0^\epsilon (1+v)\phi_v'(m)dv$ which is strictly negative for $m > 0$ and $\epsilon > 0$.

We now turn to $h_2(\cdot)$. Taking $\epsilon \to \infty$ in Eq. (A.44) yields

$$h_2(m) = 1 + 2\int_0^\infty (1+v)\phi_v(m)dv. \tag{A.49}$$

Since for Gaussian increments, the tails of ladder height $1 - F_H(\cdot)$ are exponentially small, by Eq. (A.47) we have $\phi_v(m) \leq 2U_m^+([0,m])\big(1 - F_H(mv)\big) \leq K_m \exp\{-Kv^2m^2\}$. Therefore, $\int_0^\infty 2(1+v)\phi_v(m)dv$ is finite. The rest of the proof follows the similar argument as $\overline{h}_{2,\epsilon}(\cdot)$ and dominated convergence.

□

A.3 Proof of Theorem 1

Following the strong approximation results in Appendix A.1.3, it suffices to consider the limit theorems of the test statistics constructed from $(\breve{r}_i^{(c)})_{1\leq i\leq \breve{N}^{(c)}}$. For ease of notation, we drop the breve mark $(\breve{\ })$ in the subsequent proofs.

Under the null. We shall prove the following three convergence results under $\omega \in \Omega'$:

$$\sum_{i=1}^{N^{(c)}} (r_i^{(c)})^2 \xrightarrow{\mathbb{P}} \tau(1), \qquad \sum_{i=1}^{N^{(c)}} (\overline{r}_i^{(c)})^2 \xrightarrow{\mathbb{P}} \tau(1)\frac{\overline{h}_{2,\epsilon}(m)}{h_2(m)}, \qquad c^2 N^{(c)} \xrightarrow{\mathbb{P}} \frac{\tau(1)}{h_2(m)}. \tag{A.50}$$

To verify these results over the unit interval in Eq. (A.50), we establish a more general result on uniform convergence in probability (u.c.p.). Specifically, for any processes Z^n and Z, where n indexes the stages of statistical experiments, we say that Z^n converges uniformly in probability to Z on compact sets, written as $Z^n \xrightarrow{\text{u.c.p.}} Z$, if and only if $\sup_{0\le t\le T}|Z_t^n - Z_t| \xrightarrow{\mathbb{P}} 0$ for all finite $T>0$.

We start with some notation for clarity: Let $N_{n,t}^{(c)}$ denote the number of sampled observations with the barrier width c_n over $[0,t]$ at stage n. Let $\tau_{n,i}^{(c)} = \tau(t_{n,\Pi_{n,i}^{(c)}})$ denote the intrinsic time (defined in Appendix A.1.1) at the i-th sampled observation, and $\Delta\tau_{n,i}^{(c)} = \tau_{n,i}^{(c)} - \tau_{n,i-1}^{(c)}$ the i-th duration in intrinsic time. We define the discretized filtration $\mathbb{F}^n = (\mathcal{F}^n_{N_{n,t}^{(c)}+1})_{t\ge 0}$, where $\mathcal{F}_i^n = \mathcal{F}_{t_{n,\Pi_{n,i}^{(c)}}} = \widetilde{\mathcal{F}}_{\tau_{n,i}^{(c)}}$.

Lemma A.5 extends Wald's identity to sums of independent, non-negative, and potentially non-identically distributed random variables. Lemma A.6 establishes some properties of the sequence $(\Delta\tau_{n,i}^{(c)})$, the sampled returns $(r_{n,i}^{(c)})$, and the counting process $(N_{n,t}^{(c)})$. Lemma A.7 provides some results for the discretized filtration $\mathbb{F}^n$ that will be used in subsequent proofs.

Lemma A.5. Let $S_n = \sum_{i=1}^n X_i$, where $(X_i)_{i\ge 1}$ is a sequence of independent, non-negative random variables satisfying $\mathbb{E}[X_i] \le C$ for all i. Suppose that N is an integer-valued stopping time with respect to the filtration $\mathcal{G}_i = \sigma(X_1,\dots,X_i)$. Then the expectation of the stopped sum satisfies the bound $\mathbb{E}[S_N] \le C\mathbb{E}[N]$.

Proof. It holds that

$$\mathbb{E}[S_N] = \mathbb{E}\left[\sum_{i=1}^{\infty} X_i \mathbb{1}_{\{i\le N\}}\right] = \sum_{i=1}^{\infty}\mathbb{E}[X_i\mathbb{1}_{\{i\le N\}}] = \sum_{i=1}^{\infty}\mathbb{E}[X_i]\mathbb{E}[\mathbb{1}_{\{i\le N\}}] \le C\sum_{i=1}^{\infty}\mathbb{P}(N\ge i) = C\mathbb{E}[N], \tag{A.51}$$

where the interchange of infinite sum and expectation is valid by Tonelli's theorem since all the summands are non-negative, and X_i is independent from $\mathbb{1}_{\{i\le N\}} \in \mathcal{G}_{i-1}$. This completes the proof of Lemma A.5. □

Lemma A.6. For any finite $t>0$, it holds that

(i) $\mathbb{E}[r_{n,i}^{(c)}] = 0$, and $\mathbb{E}[(r_{n,i}^{(c)})^2] = c_n^2 h_2(m)$;
(ii) $\mathbb{E}[N_{n,t}^{(c)}+1] \le K c_n^{-2}$;
(iii) $\max_{1\le i\le N_{n,t}^{(c)}+1} \Delta\tau_{n,i}^{(c)} = o_p(c_n)$;
(iv) $\mathbb{E}[\Delta\tau_{n,i}^{(c)} r_{n,i}^{(c)} | \mathcal{F}_{i-1}^n] = 0$ for all i and n.

Proof. (i) Since $\mathbb{E}[\tau_{n,1}^{(c)}] < \infty$, the claims follow from Theorem 1 of Shepp (1967).

(ii) We express $\mathbb{E}[N_{n,t}^{(c)}]$ as an infinite sum of probabilities related to $\Delta\tau_{n,i}^{(c)}$:

$$\mathbb{E}[N_{n,t}^{(c)}] = \sum_{k=1}^{\infty}\mathbb{P}(N_{n,t}^{(c)} \ge k) = \sum_{k=1}^{\infty}\mathbb{P}\left(\sum_{i=1}^{k}\Delta\tau_{n,i}^{(c)} \le \tau(t)\right). \tag{A.52}$$

By the Markov property of $\widetilde{X}$, the sequence $(\Delta\tau_{n,i}^{(c)})$ consists of positively-valued, conditionally independent random variables for each n, which satisfy $\Delta\tau_{n,i}^{(c)} \overset{\mathcal{L}}{=} c_n^2\Delta t_i^n$, where (Δt_i^n) is a sequence of i.i.d. random variables, representing the durations for a standard Gaussian random walk to exit the double barrier $(-1,1)$. Specifically, $\Delta t_i^n = \Delta_n(\Pi_{n,i}^{(1)} - \Pi_{n,i-1}^{(1)})$, where $\Pi_{n,i}^{(1)}$ denote the number of steps to breach the barrier for the i-th time. Therefore, for some constant $K>0$, we obtain the bound:

$$\sum_{k=1}^{\infty} \mathbb{P}\left(\sum_{i=1}^{k} \Delta\tau_{n,i}^{(c)} \leq \tau(t)\right) \leq \sum_{k=1}^{\infty} \mathbb{P}\left(\sum_{i=1}^{k} \Delta t_i^n \leq K c_n^{-2}\right) = \mathbb{E}[\mathtt{N}_{n,Kc_n^{-2}}^{(1)}], \tag{A.53}$$

where

$$\mathtt{N}_{n,t}^{(1)} = \sum_{k=1}^{\infty} \mathbb{1}_{\{\sum_{i=1}^{k} \Delta t_i^n \leq t\}} \tag{A.54}$$

is a counting process associated with the standard Gaussian random walk. Next we show that $\mathbb{E}[\mathtt{N}_{n,Kc_n^{-2}}^{(1)}] \leq K c_n^{-2}$. Following an approach similar to the proof of the elementary renewal theorem (see, e.g., Theorem 4.1, Gut, 2009), we censor the durations Δt_i^n for some $a>0$:

$$\Delta \bar{t}_i^n = \begin{cases} \Delta t_i^n, & t_i^n \leq a, \\ a, & t_i^n > a, \end{cases} \tag{A.55}$$

and consider another renewal process with the sequence of durations $(\Delta \bar{t}_i^n)$, and the corresponding counting process

$$\bar{\mathtt{N}}_{n,t}^{(1)} = \sum_{k=1}^{\infty} \mathbb{1}_{\{\sum_{i=1}^{k} \Delta \bar{t}_i^n \leq t\}}, \tag{A.56}$$

which satisfies

$$\mathbb{E}[\mathtt{N}_{n,t}^{(1)}] \leq \mathbb{E}[\bar{\mathtt{N}}_{n,t}^{(1)}], \tag{A.57}$$

for all $t>0$. Note that $\bar{\mathtt{N}}_{n,t}^{(1)}$ is not a stopping time (with respect to the renewal process), while $\bar{\mathtt{N}}_{n,t}^{(1)}+1$ is a stopping time for all $t>0$; see details in Section 2.3 of Gut (2009). Moreover, for any $a>0$ and $t>0$, we have $\sum_{i=1}^{\bar{\mathtt{N}}_{n,t}^{(1)}+1} \Delta \bar{t}_i^n \leq t+a$. For large enough n such that $Kc_n^{-2}>1$, take $t = Kc_n^{-2}$ and by Wald's identity, we have

$$\mathbb{E}\left[\sum_{i=1}^{\bar{\mathtt{N}}_{n,Kc_n^{-2}}^{(1)}+1} \Delta \bar{t}_i^n\right] = \mathbb{E}[\Delta \bar{t}_i^n]\mathbb{E}\big[\bar{\mathtt{N}}_{n,Kc_n^{-2}}^{(1)}+1\big] + \underbrace{a'}_{O(1)} \leq Kc_n^{-2} + a' \quad \Rightarrow \quad \mathbb{E}\big[\bar{\mathtt{N}}_{n,Kc_n^{-2}}^{(1)}+1\big] \leq Kc_n^{-2}. \tag{A.58}$$

Therefore, it follows from Eqs. (A.52), (A.53), (A.57) and (A.58) that $\mathbb{E}[N_{n,t}^{(c)}+1] \leq Kc_n^{-2}$.

(iii) By the maximal inequality, Lemma A.5, and Lemma A.6 (ii), it holds that for some $p>1$,

$$\mathbb{E}\left[\left(\max_{1\leq i\leq N_{n,t}^{(c)}+1} \Delta\tau_{n,i}^{(c)}\right)^p\right] \leq \mathbb{E}\left[\sum_{i=1}^{N_{n,t}^{(c)}+1} (\Delta\tau_{n,i}^{(c)})^p\right] \leq Kc_n^{2p}\mathbb{E}[N_{n,t}^{(c)}+1] \leq Kc_n^{2p-2}. \tag{A.59}$$

Then, by Markov's and Jensen's inequalities, for any $\delta > 0$,

$$\mathbb{P}\left(\max_{1\le i\le N_{n,t}^{(c)}+1} \Delta\tau_{n,i}^{(c)} \ge \delta\right) \le K\mathbb{E}\left[\left(\max_{1\le i\le N_{n,t}^{(c)}+1} \Delta\tau_{n,i}^{(c)}\right)^p\right]^{1/p} \le Kc_n^{2-2/p} = o(c_n). \tag{A.60}$$

(iv) By the Markov property of $\widetilde{X}$, it suffices to show that $\mathbb{E}[\tau_{n,1}^{(c)}\widetilde{X}_{\tau_{n,1}^{(c)}}] = 0$. To justify this result intuitively, we appeal to the reflection principle of Brownian motion. It states that for any sample path of $\widetilde{X}$ stopped at $\widetilde{X}_{\tau_{n,1}^{(c)}}$, there exists a corresponding mirrored path where the process evolves identically up to $\tau_{n,1}^{(c)}$, but with the final position flipped, i.e., $-\widetilde{X}_{\tau_{n,1}^{(c)}}$. As a result, the distribution of $\widetilde{X}_{\tau_{n,1}^{(c)}}$ conditional on $\tau_{n,1}^{(c)}$ is symmetric about zero, which yields $\mathbb{E}[\widetilde{X}_{\tau_{n,1}^{(c)}}|\tau_{n,1}^{(c)}] = 0$, and hence $\mathbb{E}[\tau_{n,1}^{(c)}\widetilde{X}_{\tau_{n,1}^{(c)}}] = 0$ by the law of iterated expectations.

This completes the proof of Lemma A.6.

□

Lemma A.7. The following results hold for each n:

(i) For a $\mathbb{F}$-martingale M, if M is bounded or square-integrable with $\langle M\rangle_t < Ct$ almost surely for some $C < \infty$, then its discretized version $M(n)$ with $M(n)_t = M_{T_n(t)}$ is an $\mathbb{F}^n$-martingale, where $T_n(t) = t_{n,\Pi_{n,i'}^{(c)}}$ with $i' = N_{n,t}^{(c)} + 1$ represents the first sampling time after t.

(ii) The process $\sum_{i=1}^{N_{n,t}^{(c)}+1}(\Delta\tau_{n,i}^{(c)} - h_2(m)c_n^2)$ is an $\mathbb{F}^n$-martingale.

Proof. (i) Note that $N_{n,t}^{(c)} + 1 = \inf\{i \ge 1 : \tau_{n,i}^{(c)} > t\}$ is an $\mathbb{F}^n$-stopping time. For some finite $\overline{T} > 0$, it holds that $M_{T_n(t)\wedge\overline{T}}$ is a discrete-time $\mathbb{F}^n$-martingale by optional stopping. For $0 \le s \le t$, since $N_s^{(c)} + 1 \le N_t^{(c)} + 1$ are both stopping times, the optional sampling theorem implies that

$$\mathbb{E}\left[M_{T_n(t)\wedge\overline{T}}\middle|\mathcal{F}^n_{N_{n,s}^{(c)}+1}\right] = M_{T_n(s)\wedge\overline{T}}. \tag{A.61}$$

To obtain the desired result $\mathbb{E}[M_{T_n(t)}|\mathcal{F}^n_{N_{n,s}^{(c)}+1}] = M_{T_n(s)}$, it remains to prove $\mathbb{E}[|M_{T_n(t)} - M_{T_n(t)\wedge\overline{T}}|] \to 0$ as $\overline{T} \to \infty$.

If M is bounded, then

$$|M_{T_n(t)} - M_{T_n(t)\wedge\overline{T}}| \le 2\left(\sup_{0\le s\le T_n(t)} |M_s|\right)\mathbb{1}_{\{T_n(t)>\overline{T}\}}, \tag{A.62}$$

and the desired result is directly implied by $\mathbb{1}_{\{T_n(t)>\overline{T}\}} \to 0$ as $\overline{T} \to \infty$.

If M is square-integrable with $\langle M\rangle_t < Ct$, it holds that

$$\mathbb{E}[|M_{T_n(t)} - M_{T_n(t)\wedge\overline{T}}|^2] = \mathbb{E}[|\langle M\rangle_{T_n(t)} - \langle M\rangle_{T_n(t)\wedge\overline{T}}|] < C\mathbb{E}[T_n(t)\mathbb{1}_{\{T_n(t)>\overline{T}\}}]. \tag{A.63}$$

Note that $T_n(t) \le t + \max_{1\le i\le N_{n,t}^{(c)}+1} \Delta t_{n,i}^{(c)}$, where $\Delta t_{n,i}^{(c)}$ is the i-th calendar time duration under

PDS. By the smoothness of σ under Assumption 1, Lemma A.5, and Lemma A.6 (iii),

$$\mathbb{E}\left[\max_{1\le i\le N_{n,t}^{(c)}+1}\Delta t_{n,i}^{(c)}\right]\le K\mathbb{E}\left[\max_{1\le i\le N_{n,t}^{(c)}+1}\Delta\tau_{n,i}^{(c)}\right]\le K\mathbb{E}\left[\sum_{i=1}^{N_{n,t}^{(c)}+1}\Delta\tau_{n,i}^{(c)}\right]\le K'c_n\mathbb{E}[N_{n,t}^{(c)}+1]<\infty, \tag{A.64}$$

such that $\mathbb{E}[T_n(t)]<\infty$, and $\lim_{\overline{T}\to\infty}\mathbb{E}[T_n(t)\mathbb{1}_{\{T_n(t)>\overline{T}\}}]=0$ by the dominated convergence theorem. Then the desired result follows from Eq. (A.63).

(ii) By the Markov property in intrinsic time and the first-exit scaling, it holds that $c_n^{-2}\Delta\tau_{n,i}^{(c)}\stackrel{\mathcal{L}}{=}m^{-2}(Z_1^{(m)})^2$ with $\mathbb{E}[m^{-2}(Z_1^{(m)})^2]=h_2(m)$, and thus $\mathbb{E}[\Delta\tau_{n,i}^{(c)}|\mathcal{F}_{i-1}^n]=h_2(m)c_n^2$. Therefore, $M_{n,k}=\sum_{i=1}^k(\Delta\tau_{n,i}^{(c)}-h_2(m)c_n^2)$ is a discrete-time martingale with respect to $\mathcal{F}^n$. For any fixed $n>0$, the stopping time $\mathbb{E}[N_{n,t}^{(c)}+1]<\infty$ by Lemma A.6 (ii), and $\mathbb{E}\big[|M_{n,i}-M_{n,i-1}|\big|\mathcal{F}_{i-1}^n\big]\le\mathbb{E}[\Delta\tau_{n,i}^{(c)}|\mathcal{F}_{i-1}^n]+h_2(m)c_n^2=2h_2(m)c_n^2\le K$, such that the optional sampling theorem implies that

$$\mathbb{E}\left[M_{n,N_{n,t}^{(c)}+1}\middle|\mathcal{F}^n_{N_{n,s}^{(c)}+1}\right]=M_{n,N_{n,s}^{(c)}+1}, \tag{A.65}$$

for $0\le s\le t$. Therefore, $(M_{n,N_{n,t}^{(c)}+1})$ is a discrete-time martingale with respect to $\mathbb{F}^n=(\mathcal{F}^n_{N_{n,t}^{(c)}+1})_{t\ge0}$.

This completes the proof of Lemma A.7. □

Next, we define the following three scaled processes at stage n:

$$V_{1,t}^n=\sum_{i=1}^{N_{n,t}^{(c)}}\zeta_1(r_{n,i}^{(c)})^2,\qquad V_{2,t}^n=\sum_{i=1}^{N_{n,t}^{(c)}}\zeta_2(\overline{r}_{n,i}^{(c)})^2,\qquad V_{3,t}^n=\zeta_3c_n^2N_{n,t}^{(c)}, \tag{A.66}$$

where the scaling factors are given by $(\zeta_1,\zeta_2,\zeta_3)=(1,h_2(m)/\overline{h}_{2,\epsilon}(m),h_2(m))$. Our goal is to show that for each $k=1,2,3$, the process V_k^n satisfies $V_k^n\xrightarrow{\text{u.c.p.}}\tau=(\tau(t))_{t\ge0}$. To establish this, it suffices to show that for any $T>0$,

$$\sup_{t\in[0,T]}|V_{k,t}^n-\tau(t)|=O_p(c_n). \tag{A.67}$$

We begin by proving (A.67) for V_3^n: Define an auxiliary pre-limiting process $U_t^n=\sum_{i=1}^{N_{n,t}^{(c)}}\Delta\tau_{n,i}^{(c)}$. By the triangle inequality,

$$\sup_{t\in[0,T]}|V_{3,t}^n-\tau(t)|\le\sup_{t\in[0,T]}|V_{3,t}^n-U_t^n|+\sup_{t\in[0,T]}|U_t^n-\tau(t)|. \tag{A.68}$$

For the first supremum, we have

$$\sup_{t\in[0,T]}|V_{3,t}^n-U_t^n|=\sup_{t\in[0,T]}\left|\sum_{i=1}^{N_{n,t}^{(c)}}(\zeta_3c_n^2-\Delta\tau_{n,i}^{(c)})\right|\le K\sup_{t\in[0,T]}\left|\sum_{i=1}^{N_{n,t}^{(c)}+1}(\zeta_3c_n^2-\Delta\tau_{n,i}^{(c)})\right|, \tag{A.69}$$

where the inequality holds because adding one more term to the sum cannot decrease the supremum of the process. By Lemma A.6 (iii), it holds that $\mathbb{E}[(\zeta_3c_n^2-\Delta\tau_{n,i}^{(c)})^2]\le Kc_n^4$. Note that $\sum_{i=1}^{N_{n,t}^{(c)}+1}(\zeta_3c_n^2-$

$\Delta\tau_{n,i}^{(c)})$ is an $\mathbb{F}^n$-martingale, i.e., Lemma A.7 (ii). Applying the Burkholder-Davis-Gundy inequality, Lemma A.5, and Lemma A.6 (ii), we have

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|\sum_{i=1}^{N_{n,t}^{(c)}+1}(\zeta_3 c_n^2-\Delta\tau_{n,i}^{(c)})\right|^2\right]\le K\mathbb{E}\left[\sum_{i=1}^{N_{n,t}^{(c)}+1}(\zeta_3 c_n^2-\Delta\tau_{n,i}^{(c)})^2\right]\le K'c_n^4\mathbb{E}[N_{n,t}^{(c)}]\le K''c_n^2. \quad (A.70)$$

Then, by Markov's and Jensen's inequalities, for any $\delta>0$,

$$\mathbb{P}\left(\sup_{t\in[0,T]}\left|\sum_{i=1}^{N_{n,t}^{(c)}+1}(\zeta_3 c_n^2-\Delta\tau_{n,i}^{(c)})\right|>\delta\right)\le\frac{1}{\delta}\mathbb{E}\left[\sup_{t\in[0,T]}\left|\sum_{i=1}^{N_{n,t}^{(c)}+1}(\zeta_3 c_n^2-\Delta\tau_{n,i}^{(c)})\right|^2\right]^{1/2}\le Kc_n, \quad (A.71)$$

which shows that $\sup_{t\in[0,T]}|V_{3,t}^n-U_t^n|=O_p(c_n)$.

For the second term, we have

$$\sup_{t\in[0,T]}|U_t^n-\tau(t)|=\sup_{t\in[0,T]}\left|\tau^n_{N_{n,t}^{(c)}}-\tau(t)\right|\le K\max_{1\le i\le N_{n,t}^{(c)}+1}\Delta\tau_{n,i}^{(c)}=o_p(c_n), \quad (A.72)$$

where the final equation follows from Lemma A.6 (iii). Combining Eqs. (A.69) and (A.72), we conclude that the u.c.p. result in Eq. (A.67) holds for V_3^n.

To prove the u.c.p. result for V_1^n, we write the corresponding supremum process into:

$$\sup_{t\in[0,T]}|V_{1,t}^n-\tau(t)|\le\sup_{t\in[0,T]}|V_{1,t}^n-V_{3,t}^n|+\sup_{t\in[0,T]}|V_{3,t}^n-\tau(t)|, \quad (A.73)$$

thus it suffices to prove that

$$\sup_{t\in[0,T]}|V_{1,t}^n-V_{3,t}^n|=\sup_{t\in[0,T]}\left|\sum_{i=1}^{N_{n,t}^{(c)}}((r_{n,i}^{(c)})^2-\zeta_3 c_n^2)\right|\le K\sup_{t\in[0,T]}\left|\sum_{i=1}^{N_{n,t}^{(c)}+1}((r_{n,i}^{(c)})^2-\zeta_3 c_n^2)\right|=O_p(c_n). \quad (A.74)$$

With a similar martingale argument as for V_3^n (Lemma A.7 (ii)), and $\mathbb{E}\big[\big((r_{n,i}^{(c)})^2-\zeta_3 c_n^2\big)^2\big]\le Kc_n^4$ implied by the Burkholder-Davis-Gundy inequality, we have

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|\sum_{i=1}^{N_{n,t}^{(c)}+1}((r_{n,i}^{(c)})^2-\zeta_3 c_n^2)\right|^2\right]\le K\mathbb{E}\left[\sum_{i=1}^{N_{n,t}^{(c)}+1}((r_{n,i}^{(c)})^2-\zeta_3 c_n^2)^2\right]\le K'c_n^4\mathbb{E}[N_{n,t}^{(c)}]\le K''c_n^2. \quad (A.75)$$

and then, similarly to Eq. (A.71), we conclude that Eq. (A.74), and thus the u.c.p. result in Eq. (A.67), holds for V_1^n. Moreover, the u.c.p. result for V_2^n can be verified with the same steps, and thus omitted here.

Since the u.c.p. results in Eq. (A.67) hold for all three processes V_1^n, V_2^n, and V_3^n, it follows that the three limits in Eq. (A.50) also hold uniformly over the unit interval, such that

$$\frac{V_{1,t}^n}{V_{3,t}^n} = \frac{\sum_{i=1}^{N^{(c)}} (r_i^{(c)})^2}{c^2 N^{(c)}} \xrightarrow{\mathbb{P}} h_2(m) \qquad \text{and} \qquad \frac{V_{2,t}^n}{V_{3,t}^n} = \frac{\sum_{i=1}^{N^{(c)}} (\overline{r}_i^{(c)})^2}{c^2 N^{(c)}} \xrightarrow{\mathbb{P}} \overline{h}_{2,\epsilon}(m). \tag{A.76}$$

The consistency of both $\overline{M}_{c,\epsilon}$ and M_c in Theorem 1 is a direct result from Eq. (A.76) and the continuous mapping theorem.

Under the alternative. We denote by (Λ_t^n) the counting process of all jumps in J_1^n in Eq. (A.34), then Λ_t^n is bounded for each n, and for all n, we have

$$\sum_{0 \le s \le t} |\Delta J_{1,s}^n|^r < \infty, \qquad \text{where } \Delta J_{1,s}^n = J_{1,s}^n - J_{1,s^-}^n, \tag{A.77}$$

which implies for large enough n (such that $u_n \to 0$),

$$u_n^r \Lambda_t^n \le \sum_{0 \le s \le t} |\Delta J_{1,s}^n|^r < \infty, \tag{A.78}$$

with u_n defined in Appendix A.1.4. From Eq. (A.78), we deduce that $\Lambda_t^n = O_p(\Delta_n^{-r\varpi})$ for all fixed t, where ϖ is arbitrarily close to but below 1/2.

When $X(\omega)$ is discontinuous within $(0,1)$, we denote by $\{s_1^n, s_2^n \ldots, s_\Lambda^n\}$ the sequence of all jump times in chronological order, where $\Lambda \equiv \Lambda_1^n(\omega)$ counts the number of all jumps on $(0,1]$. We define

$$k^-(s) = \inf_{0 \le i \le n} \{t_i \ge s : |t_i - s|\} \qquad \text{and} \qquad k^+(s) = \inf_{0 \le i \le n} \{t_i < s : |t_i - s|\} \tag{A.79}$$

as the index of the first observations no earlier than and strictly before s, respectively. We split the sequence of observations $(X_{t_i})_{0 \le i \le N}$ into $\Lambda + 1$ segments with $i = k^+(s_j^n)$ for all $1 \le j \le \Lambda$ as cutoff points. As $N \to \infty$, we have $k^+(s_j^n) - k^+(s_{j-1}^n) \to \infty$ (also, $k^+(s_1^n) \to \infty$), since any intervals of length of order Δ_n mostly contain a single jump of size larger than u_n with probability approaching one, see Section 2.3 of Aït-Sahalia and Jacod (2009).

For each segment $(X_{t_i})_{k^+(s_{j-1}^n) \le i \le k^+(s_j^n)}$, we obtain the PDS returns $(r_i^{(c)})_{N_{j-1}^{(c)}+1 \le i \le N_j^{(c)}}$ with the barrier width $c = m\sqrt{\Delta_n}$. For each $i \in A_n = \{N_1^{(c)}, N_2^{(c)}, \ldots, N_\Lambda^{(c)}\}$, the PDS return $|r_i^{(c)}| \ge u_n \gg c$ contains jumps and will be censored by $\varphi_\epsilon(c)$. For all $i \in A_n^\complement$, the PDS return $r_i^{(c)}$ contains only aggregated Brownian increments. For the censored PDS returns, we have

$$\frac{\sum_{i=1}^{N^{(c)}} (\overline{r}_i^{(c)})^2}{c^2 N^{(c)}} = \frac{\sum_{i \in A_n^\complement} (\overline{r}_i^{(c)})^2}{c^2 N^{(c)}} + \frac{\sum_{i \in A_n} (\overline{r}_i^{(c)})^2}{c^2 N^{(c)}}. \tag{A.80}$$

For the first term above, since the cardinality of A_n is $\Lambda = O_p(\Delta_n^{-r\varpi}) \ll \Delta_n^{-1} \asymp N^{(c)}$, we have

$$\frac{\sum_{i \in A_n^\complement} (\overline{r}_i^{(c)})^2}{c^2 N^{(c)}} = \frac{N^{(c)} - \Lambda}{N^{(c)}} \frac{\sum_{i \in A_n^\complement} (\overline{r}_i^{(c)})^2}{c^2 (N^{(c)} - \Lambda)}, \quad \text{where } \frac{N^{(c)} - \Lambda}{N^{(c)}} \xrightarrow{\mathbb{P}} 1, \tag{A.81}$$

such that it coincides with the limit theorems under the null. For the second term, it holds that

$$\frac{\sum_{i\in A_n}(\overline{r}_i^{(c)})^2}{c^2 N^{(c)}} \leq K\frac{\Lambda}{N^{(c)}} \leq K'\Delta_n^{1-r\varpi}, \tag{A.82}$$

which has no impact on the LLN result. It still vanishes after multiplying by $\sqrt{N^{(c)}} \asymp \Delta_n^{-1/2}$ for any $r \in [0,1)$, and thus does not affect the CLT.

Consider the PDS returns. Under infill asymptotics, we claim that

$$\sum_{i=1}^{N^{(c)}}(r_i^{(c)})^2 = \sum_{i\in A_n^\complement}(r_i^{(c)})^2 + \sum_{i\in A_n}(r_i^{(c)})^2 \xrightarrow{\mathbb{P}} \tau(1) + \sum_{0<s\leq 1}|\Delta J_{1,s}^n|^2, \tag{A.83}$$

where the jump variation is given by

$$\sum_{0<s\leq 1}|\Delta J_{1,s}^n|^2 = \int_0^1 ds\int_{\{|\delta(s,x)|\geq u_n\}}|\delta(s,x)|^2\lambda(dx). \tag{A.84}$$

The convergence $\sum_{i\in A_n^\complement}(r_i^{(c)})^2 \xrightarrow{\mathbb{P}} \tau(1)$ is a direct result from Eq. (A.50), with the cardinality of $A_n^\complement$, $N^{(c)} - \Lambda$, satisfying Eq. (A.81). For the PDS returns with $i \in A_n$, we have

$$\sum_{i\in A_n}|r_i^{(c)}|^2 - \sum_{0<s\leq 1}|\Delta J_{1,s}|^2 \leq \sum_{i\in A_n}||r_i^{(c)}|^2 - |\Delta_i J_1^n|^2|, \tag{A.85}$$

where $\Delta_i J_1^n = J^n_{1,t_{\Pi_i^{(c)}}} - J^n_{1,t_{\Pi_i^{(c)}-1}}$. For all $i \in A_n$, it holds that

$$|r_i^{(c)}| \leq \left|\sum_{i=\Pi_{i-1}^{(c)}}^{\Pi_i^{(c)}}(X'_{t_i} - X'_{t_{i-1}})\right| + |\Delta_i J_1^n| + \left|\sum_{i=\Pi_{i-1}^{(c)}}^{\Pi_i^{(c)}}(J^n_{2,t_i} - J^n_{2,t_i-1})\right|, \tag{A.86}$$

such that

$$||r_i^{(c)}|^2 - |\Delta_i J_1^n|^2| \leq (|r_i^{(c)}| + |\Delta_i J_1^n|)\left(\left|\sum_{i=\Pi_{i-1}^{(c)}}^{\Pi_i^{(c)}}(X'_{t_i} - X'_{t_{i-1}})\right| + \left|\sum_{i=\Pi_{i-1}^{(c)}}^{\Pi_i^{(c)}}(J^n_{2,t_i} - J^n_{2,t_i-1})\right|\right) = O_p(\sqrt{\Delta_n}), \tag{A.87}$$

where, by Lemma A.5 and Lemma A.2 (i),

$$\mathbb{E}\left[\left|\sum_{i=\Pi_{i-1}^{(c)}}^{\Pi_i^{(c)}}(X'_{t_i} - X'_{t_{i-1}})\right|\right] \leq \mathbb{E}\left[\sum_{i=\Pi_{i-1}^{(c)}}^{\Pi_i^{(c)}+1}|X'_{t_i} - X'_{t_{i-1}}|\right] \leq \mathbb{E}[\Pi_1^{(c)} + 1]\max_{1\leq i\leq N}\mathbb{E}[|X'_{t_i} - X'_{t_{i-1}}|] \leq K\sqrt{\Delta_n}, \tag{A.88}$$

and, by Lemma A.4,

$$\mathbb{E}\Bigg[\Bigg|\sum_{i=\Pi_{i-1}^{(c)}}^{\Pi_i^{(c)}}(J^n_{2,t_i}-J^n_{2,t_i-1})\Bigg|\Bigg]\leq\mathbb{E}\Bigg[\sum_{i=\Pi_{i-1}^{(c)}}^{\Pi_i^{(c)}}|J^n_{2,t_i}-J^n_{2,t_i-1}|\Bigg]\leq\mathbb{E}[\Pi_1^{(c)}]\max_{1\leq i\leq N}\mathbb{E}[|J^n_{2,t_i}-J^n_{2,t_i-1}|]$$
$$\leq\mathbb{E}[\Pi_1^{(c)}]\max_{1\leq i\leq N}\mathbb{E}\Big[\sup_{t_{i-1}\leq s\leq t_i}|J^n_{2,s}-J^n_{2,t_{i-1}}|\Big]\leq K\Delta_n^{1+(1-r)\varpi}. \tag{A.89}$$

Therefore, we have

$$\sum_{i\in A_n}|r_i^{(c)}|^2-\sum_{0<s\leq 1}|\Delta J_{1,s}|^2\leq\sum_{i\in A_n}||r_i^{(c)}|^2-|\Delta_i J_1^n|^2|=O_p(\Delta_n^{1/2-r\varpi})=o_p(1), \tag{A.90}$$

which implies the convergence of $\sum_{i\in A_n}(r_i^{(c)})^2$ and thus Eq. (A.83), such that it holds that

$$\frac{\sum_{i=1}^{N^{(c)}}(r_i^{(c)})^2}{c^2N^{(c)}}\xrightarrow{\mathbb{P}}\frac{h_2(m)\langle X,X\rangle_1}{\tau(1)}. \tag{A.91}$$

This completes the proof.

A.4 Proof of Theorem 2

We start with some definitions and notation for clarity in the proof: We follow the design of statistical experiments in the proof of Theorem 1 (Appendix A.3), where $N_{n,t}^{(c)}$ counts the number of sampled observations with the barrier width c_n over $[0,t]$ at stage n, and we define a sequence $(U_{n,i}^{(c)})$ of random variables as scaled intrinsic-time durations, i.e., $U_{n,i}^{(c)}=c_n^{-2}\Delta\tau_{n,i}^{(c)}$.

By Wald's identity, we have $\mathbb{E}[U_{n,i}^{(c)}]=m^{-2}\mathbb{E}[|Z_1^{(m)}|^2]=h_2(m)$ and $\mathbb{E}[(U_{n,i}^{(c)})^2]=h_4(m)$. For simplicity, we denote the scaled cross moments of $U_{n,i}^{(c)}$ with the squared censored and uncensored PDS returns as $\overline{\lambda}(m)=c_n^{-2}\mathbb{E}[U_{n,i}^{(c)}(\overline{r}_{n,i}^{(c)})^2]$ and $\lambda(m)=c_n^{-2}\mathbb{E}[U_{n,i}^{(c)}(r_{n,i}^{(c)})^2]$, respectively. Additionally, we denote the scaled cross moment between the squared censored and uncensored PDS returns as $\overline{\lambda}_r(m)=c_n^{-2}\mathbb{E}[(r_{n,i}^{(c)})^2(\overline{r}_{n,i}^{(c)})^2]=m^{-4}\overline{\rho}_{2,\epsilon}(m)$.

Stable convergence. We state a key theorem from Jacod (1997) to establish $\mathcal{F}$-stable convergence for a sequence of local martingales.[1] We say that Z^n converges $\mathcal{F}$-stably in law to Z, written as $Z^n\xrightarrow{\mathcal{L}-\mathrm{s}}Z$, if for all $\mathcal{F}$-measurable processes Y, we have the joint convergence $(Z^n,Y)\xrightarrow{\mathcal{L}}(Z,Y)$; see more details in Rényi (1963) and Jacod and Protter (2012).

We start with the general setting of Jacod (1997): Let X be a continuous-time local martingale on $(\Omega,\mathcal{F},\mathbb{F}=(\mathcal{F}_t)_{t\geq 0},\mathbb{P})$, and denote by $\mathcal{M}_b$ the set of all bounded martingales on the same basis. A sequence of filtrations $\mathbb{F}^n=(\mathcal{F}_t^n)_{t\geq 0}$ on $(\Omega,\mathcal{F})$ is said to satisfy Property (F) if the following conditions hold for each $n\in\mathbb{N}$:

[1] See also Chapter IX.7 of Jacod and Shiryaev (2003).

Property (F). We have a square-integrable $\mathbb{F}^n$-martingale $X(n)$ and, for each $Z \in \mathcal{M}_b$, a bounded $\mathbb{F}^n$-martingale $Z(n)$, such that for all $t \geq 0$,

(i) $\sup_{n,t,\omega} |Z(n)_t(\omega)| < \infty$;
(ii) $\langle X(n) \rangle_t \xrightarrow{\mathbb{P}} \langle X \rangle_t$;
(iii) For any finite family $(Z^1, \ldots, Z^m) \subset \mathcal{M}_b$, we have the following convergence for the Skorokhod topology on $\mathbb{D}(\mathbb{R}^{d+m})$:

$$(X(n), Z^1(n), \ldots, Z^m(n)) \xrightarrow{\mathbb{P}} (X, Z^1, \ldots, Z^m). \tag{A.92}$$

The following theorem is a simplified version of Theorem IX.7.13 of Jacod and Shiryaev (2003):

Theorem A.1. Assume Property (F). Let H^n denote a sequence of square-integrable $\mathbb{F}^n$-local martingales, and let ΔH^n collects all jumps of H^n. Suppose that there is a C_∞-valued adapted process η starting from zero, such that for all $Z \in \mathcal{M}_b$ orthogonal to X, we have for all $t \geq 0$ and $\varepsilon > 0$:

(i) $\sum_{s \leq t} |\Delta H^n_s|^2 \mathbb{1}_{\{|\Delta H^n_s| > \varepsilon\}} \xrightarrow{\mathbb{P}} 0$;
(ii) $\langle H^n, X(n) \rangle_t \xrightarrow{\mathbb{P}} 0$;
(iii) $\langle H^n, Z(n) \rangle_t \xrightarrow{\mathbb{P}} 0$;
(iv) $\langle H^n \rangle_t \xrightarrow{\mathbb{P}} \eta_t$.

Then it holds that $H^n \xrightarrow{\mathcal{L}-\mathrm{s}} H$, where H is an $\mathcal{F}$-conditional Gaussian martingale on the filtered extension $(\Omega^*, \mathcal{F}^*, (\mathcal{F}^*_t)_{t \geq 0}, \mathbb{P}^*)$ with $\langle H \rangle_t = \eta_t$.

We consider a 3-dimensional $\mathbb{F}^n$-martingale H^n, we aim to show that for $\omega \in \Omega'$:

$$H^n_t = c_n^{-1} \left\{ \sum_{i=1}^{N^{(c)}_{n,t}+1} \begin{pmatrix} (r^{(c)}_{n,i})^2 \\ (\overline{r}^{(c)}_{n,i})^2 \\ c_n^2 \end{pmatrix} - \sum_{i=1}^{N^{(c)}_{n,t}+1} c_n^2 U^{(c)}_{n,i} \begin{pmatrix} \zeta_1^{-1} \\ \zeta_2^{-1} \\ \zeta_3^{-1} \end{pmatrix} \right\} \xrightarrow{\mathcal{L}-\mathrm{s}} H_t, \quad \text{with } \langle H \rangle_t = \frac{\tau(t)}{h_2(m)} \Sigma, \tag{A.93}$$

where $\Sigma = (\sigma_{ij})_{1 \leq i,j \leq 3}$ is symmetric with

$$\sigma_{11} = 2(h_4(m) - \lambda(m)), \tag{A.94}$$

$$\sigma_{22} = \overline{h}_{4,\epsilon}(m) - \frac{2\overline{h}_{2,\epsilon}(m)\overline{\lambda}(m)}{h_2(m)} + \frac{\overline{h}_{4,\epsilon}(m)\overline{h}^2_{2,\epsilon}(m)}{h_2^2(m)}, \tag{A.95}$$

$$\sigma_{33} = \frac{h_4(m) - h_2^2(m)}{h_2^2(m)}, \tag{A.96}$$

$$\sigma_{12} = \overline{\rho}_{2,\epsilon}(m) - \frac{\overline{h}_{2,\epsilon}(m)\lambda(m)}{h_2(m)} - \overline{\lambda}(m) + \frac{h_4(m)\overline{h}_{2,\epsilon}(m)}{h_2(m)}, \tag{A.97}$$

$$\sigma_{13} = \frac{h_4(m) - \lambda(m)}{h_2(m)}, \tag{A.98}$$

$$\sigma_{23} = \frac{h_4(m)\overline{h}_{2,\epsilon}(m)}{h_2^2(m)} - \frac{\overline{\lambda}(m)}{h_2(m)}, \tag{A.99}$$

Note that the martingality of H^n with respect to $\mathbb{F}^n$ can be similarly verified as in Lemma A.7 (ii).

We prove the claimed stable CLT in Eq. (A.93) by verifying the conditions of Theorem A.1. First, we check Property (F) for the discretized filtration $\mathbb{F}^n = (\mathcal{F}^n_{N^{(c)}_{n,t}+1})_{t\geq 0}$. Given the continuous martingale $X \equiv X'$ adapted to $\mathbb{F}$, we define its $\mathbb{F}^n$-discretized version $X(n)$ as in Lemma A.7 (i), which is square-integrable by construction. Pick some $Z \in \mathcal{M}_b$ and consider its discretized version $Z(n)$, then the boundedness of Z ensures that condition (i) of Property (F) holds. Condition (ii) follows from the u.c.p. result of V^n_1 in the proof of Theorem 1 (Appendix A.3). Condition (iii) follows from Proposition VI.6.37 of Jacod and Shiryaev (2003) and Eq. (2.2.13) of Jacod and Protter (2012) by virtue of Lemma A.6 (iii). Therefore, Property (F) is satisfied for the specific PDS-based $\mathbb{F}^n$.

Now we verify the conditions in Theorem A.1 for $H^n = (H^n_1, H^n_2, H^n_3)^\top$. We write

$$H^n_t = \sum_{i=1}^{N^{(c)}_{n,t}+1} \Delta H^n_i, \qquad \text{where } \Delta H^n_i = \begin{pmatrix} \Delta H^n_{1,i} \\ \Delta H^n_{2,i} \\ \Delta H^n_{3,i} \end{pmatrix} = c_n^{-1} \begin{pmatrix} (r^{(c)}_{n,i})^2 - \Delta\tau^{(c)}_{n,i} \\ (\bar{r}^{(c)}_{n,i})^2 - \zeta_2^{-1}\Delta\tau^{(c)}_{n,i} \\ c_n^2 - \zeta_3^{-1}\Delta\tau^{(c)}_{n,i} \end{pmatrix}. \tag{A.100}$$

For Condition (i), we need to show that

$$\sum_{i=1}^{N^{(c)}_{n,t}+1} (\Delta H^n_{k,i})^2 \mathbb{1}_{\{|\Delta H^n_{k,i}|>\varepsilon\}} \xrightarrow{\mathbb{P}} 0, \tag{A.101}$$

for all $t \geq 0$ and $\varepsilon > 0$. This condition corresponds to a classical (conditional) Lindeberg condition, which ensures that the limiting process H has no jumps; see Remark 3 of Podolskij and Vetter (2010). Note that the conditional expectation of the summand can be bounded by:

$$\begin{aligned} \mathbb{E}[(\Delta H^n_{k,i})^2 \mathbb{1}_{\{|\Delta H^n_{k,i}|>\varepsilon\}} | \mathcal{F}^n_{i-1}] &\leq (\mathbb{E}[(\Delta H^n_{k,i})^r | \mathcal{F}^n_{i-1}])^{2/r} (\mathbb{P}(|\Delta H^n_{k,i}| > \varepsilon | \mathcal{F}^n_{i-1}))^{1-2/r} \\ &\leq (\mathbb{E}[(\Delta H^n_{k,i})^r | \mathcal{F}^n_{i-1}])^{2/r} (\varepsilon^{-r}\mathbb{E}[(\Delta H^n_{k,i})^r | \mathcal{F}^n_{i-1}])^{1-2/r} \\ &\leq \varepsilon^{2-r}\mathbb{E}[(\Delta H^n_{k,i})^r | \mathcal{F}^n_{i-1}] \\ &\leq K\mathbb{E}[(\Delta H^n_{k,i})^r | \mathcal{F}^n_{i-1}], \end{aligned} \tag{A.102}$$

for some $r > 2$, by Hölder's and Markov's inequalities. Then we have

$$\mathbb{E}\left[\begin{pmatrix} |\Delta H^n_{1,i}|^r \\ |\Delta H^n_{2,i}|^r \\ |\Delta H^n_{3,i}|^r \end{pmatrix} \middle| \mathcal{F}^n_{i-1} \right] \leq K c_n^r \begin{pmatrix} \mathbb{E}[|c_n^{-2}(r^{(c)}_{n,i})^2 - U^{(c)}_{n,i}|^r | \mathcal{F}^n_{i-1}] \\ \mathbb{E}[|c_n^{-2}(\bar{r}^{(c)}_{n,i})^2 - \zeta_2^{-1}U^{(c)}_{n,i}|^r | \mathcal{F}^n_{i-1}] \\ \mathbb{E}[|1 - \zeta_3^{-1}U^{(c)}_{n,i}|^r | \mathcal{F}^n_{i-1}] \end{pmatrix} \leq K' c_n^r \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \tag{A.103}$$

where all three conditional moments are finite for some $r > 2$, which can be readily verified with the Burkholder-Davis-Gundy inequality, and with finite r-th moment of $U^{(c)}_{n,i}$, i.e., $\mathbb{E}[|U^{(c)}_{n,i}|^r] =$

$c_n^{-2r}\Delta_n^r\mathbb{E}[|\Pi_{n,1}^{(c)}|^r] < \infty$ by Lemma A.2 (ii). By Lemma A.5, we have

$$\mathbb{E}\left[\sum_{i=1}^{N_{n,t}^{(c)}+1} (\Delta H_{k,i}^n)^2 \mathbb{1}_{\{|\Delta H_{k,i}^n|>\varepsilon\}}\right] \le K c_n^{r-2} = o(1), \tag{A.104}$$

which implies Eq. (A.101).

For Condition (ii), we have

$$\begin{aligned}
\mathbb{E}[\Delta H_{1,i}^n r_{n,i}^{(c)}|\mathcal{F}_{i-1}^n] &= c_n^{-1}\mathbb{E}[((r_{n,i}^{(c)})^2 - \Delta\tau_{n,i}^{(c)})r_{n,i}^{(c)}|\mathcal{F}_{i-1}^n] = 0,\\
\mathbb{E}[\Delta H_{2,i}^n r_{n,i}^{(c)}|\mathcal{F}_{i-1}^n] &= c_n^{-1}\mathbb{E}[((\overline{r}_{n,i}^{(c)})^2 - \zeta_2^{-1}\Delta\tau_{n,i}^{(c)})r_{n,i}^{(c)}|\mathcal{F}_{i-1}^n] = 0,\\
\mathbb{E}[\Delta H_{3,i}^n r_{n,i}^{(c)}|\mathcal{F}_{i-1}^n] &= c_n^{-1}\mathbb{E}[(c_n^2 - \zeta_3^{-1}\Delta\tau_{n,i}^{(c)})r_{n,i}^{(c)}|\mathcal{F}_{i-1}^n] = 0,
\end{aligned} \tag{A.105}$$

by Lemma A.6 (iii) and (iv), and also $\mathbb{E}[(\overline{r}_{n,i}^{(c)})^2 r_{n,i}^{(c)}] = 0$.

For a generic martingale $Z \in \mathcal{M}_b$ starting from 0 and orthogonal to X, we define $\widetilde{Z}_{\tau(t)} = Z_t$ as the intrinsic-time counterpart of Z, where $\tau(t) = \langle X\rangle_t$. Then, Condition (iii) can be written into

$$\mathbb{E}\big[\Delta H_{k,i}^n\big(\widetilde{Z}_{\tau_{n,i}^{(c)}} - \widetilde{Z}_{\tau_{n,i-1}^{(c)}}\big)\big|\mathcal{F}_{i-1}^n\big] = 0, \tag{A.106}$$

and it suffices to show that

$$\mathbb{E}\big[\Delta H_{k,1}^n \widetilde{Z}_{\tau_{n,1}^{(c)}}\big] = 0. \tag{A.107}$$

As defined in Appendix A.1.1, the filtration $\widetilde{\mathbb{F}} = (\widetilde{\mathcal{F}}_\tau)_{\tau\ge 0}$ is generated by the Brownian motion $\widetilde{X}$. Since $\tau_{n,1}^{(c)}$ is a $\widetilde{\mathbb{F}}$-stopping time, the increments $\Delta H_{k,1}^n$ in Eq. (A.100) are measurable with respect to $\widetilde{\mathbb{F}}$ with zero mean and finite variance. Hence, by the martingale representation theorem, there exists a predictable process h_k which is adapted to $\widetilde{\mathbb{F}}$, such that

$$\Delta H_{k,1}^n = \int_0^\infty h_{k,s} d\widetilde{X}_s. \tag{A.108}$$

We also have the following integral representation for $\widetilde{Z}_{\tau_{n,1}^{(c)}}$:

$$\widetilde{Z}_{\tau_{n,1}^{(c)}} = \int_0^\infty \mathbb{1}_{\{s\le\tau_{n,1}^{(c)}\}} d\widetilde{Z}_s. \tag{A.109}$$

Therefore, by the Kunita-Watanabe identity, we have

$$\mathbb{E}\big[\Delta H_{k,1}^n \widetilde{Z}_{\tau_{n,1}^{(c)}}\big] = \mathbb{E}\left[\int_0^\infty h_{k,s}\mathbb{1}_{\{s\le\tau_{n,1}^{(c)}\}} d\langle\widetilde{X},\widetilde{Z}\rangle_s\right] = 0, \tag{A.110}$$

since $\langle\widetilde{X},\widetilde{Z}\rangle_{\tau(t)} = \langle X, Z\rangle_t \equiv 0$ as assumed. This result and an iterative conditioning argument lead to Eq. (A.106) and further implies Condition (iii) in Theorem A.1.

Finally, Condition (iv) translates to

$$\sum_{i=1}^{N_{n,t}^{(c)}+1} \mathbb{E}[\langle \Delta H_i^n \rangle | \mathcal{F}_{i-1}^n] \xrightarrow{\mathbb{P}} \frac{\tau(t)}{h_2(m)}\Sigma. \tag{A.111}$$

Note that $c_n^2 N_{n,t}^{(c)} \xrightarrow{\mathbb{P}} \tau(t)/h_2(m)$. We further calculate all cross moments $\mathbb{E}[\Delta H_{k,i}^n \Delta H_{k',i}^n | \mathcal{F}_{i-1}^n]$:

$$\mathbb{E}[(\Delta H_{1,i}^n)^2 | \mathcal{F}_{i-1}^n] = c_n^{-2}\mathbb{E}[((r_{n,i}^{(c)})^2 - c_n^2 U_{n,i}^{(c)})^2 | \mathcal{F}_{i-1}^n] = 2c_n^2(h_4(m) - \lambda(m)), \tag{A.112}$$

$$\begin{aligned} \mathbb{E}[(\Delta H_{2,i}^n)^2 | \mathcal{F}_{i-1}^n] &= c_n^{-2}\mathbb{E}[((\overline{r}_{n,i}^{(c)})^2 - \zeta_2^{-1} c_n^2 U_{n,i}^{(c)})^2 | \mathcal{F}_{i-1}^n] \\ &= c_n^2\left(\overline{h}_{4,\epsilon}(m) - \frac{2\overline{h}_{2,\epsilon}(m)\overline{\lambda}(m)}{h_2(m)} + \frac{\overline{h}_{4,\epsilon}(m)\overline{h}_{2,\epsilon}^2(m)}{h_2^2(m)}\right), \end{aligned} \tag{A.113}$$

$$\mathbb{E}[(\Delta H_{3,i}^n)^2 | \mathcal{F}_{i-1}^n] = c_n^{-2}\mathbb{E}[(c_n^2 - \zeta_3^{-1} c_n^2 U_{n,i}^{(c)})^2 | \mathcal{F}_{i-1}^n] = \frac{c_n^2(h_4(m) - h_2^2(m))}{h_2^2(m)}, \tag{A.114}$$

$$\begin{aligned} \mathbb{E}[\Delta H_{1,i}^n \Delta H_{2,i}^n | \mathcal{F}_{i-1}^n] &= c_n^{-2}\mathbb{E}[((r_{n,i}^{(c)})^2 - c_n^2 U_{n,i}^{(c)})((\overline{r}_{n,i}^{(c)})^2 - \zeta_2^{-1} c_n^2 U_{n,i}^{(c)}) | \mathcal{F}_{i-1}^n] \\ &= c_n^2\left(\overline{\lambda}_r(m) - \frac{\overline{h}_{2,\epsilon}(m)\lambda(m)}{h_2(m)} - \overline{\lambda}(m) + \frac{h_4(m)\overline{h}_{2,\epsilon}(m)}{h_2(m)}\right), \end{aligned} \tag{A.115}$$

$$\mathbb{E}[\Delta H_{1,i}^n \Delta H_{3,i}^n | \mathcal{F}_{i-1}^n] = c_n^{-2}\mathbb{E}[((r_{n,i}^{(c)})^2 - c_n^2 U_{n,i}^{(c)})(c_n^2 - \zeta_3^{-1} c_n^2 U_{n,i}^{(c)}) | \mathcal{F}_{i-1}^n] = \frac{c_n^2(h_4(m) - \lambda(m))}{h_2(m)}, \tag{A.116}$$

$$\begin{aligned} \mathbb{E}[\Delta H_{2,i}^n \Delta H_{3,i}^n | \mathcal{F}_{i-1}^n] &= c_n^{-2}\mathbb{E}[((\overline{r}_{n,i}^{(c)})^2 - \zeta_2^{-1} c_n^2 U_{n,i}^{(c)})(c_n^2 - \zeta_3^{-1} c_n^2 U_{n,i}^{(c)}) | \mathcal{F}_{i-1}^n] \\ &= c_n^2\left(\frac{h_4(m)\overline{h}_{2,\epsilon}(m)}{h_2^2(m)} - \frac{\overline{\lambda}(m)}{h_2(m)}\right). \end{aligned} \tag{A.117}$$

The above calculations verify the result in Eq. (A.111). Therefore, the stable convergence in Eq. (A.93) follows from Theorem A.1 with all conditions satisfied, and it is safe to replace $N_{n,t}^{(c)} + 1$ with $N_{n,t}^{(c)}$ in Eq. (A.93) as the additional term of order $o_p(c_n)$ is asymptotic negligible.

Suppose that $(x_n, y_n, z_n)^\top - (x_0, y_0, z_0)^\top \xrightarrow{\mathcal{L}} \mathcal{MN}(0, \Sigma)$, where $\Sigma = (\sigma_{ij})_{1\le i,j\le 3}$. Consider the function $g(x, y, z) = (x/z, y/z)^\top$. The Jacobian matrix is given by

$$J_g(x, y, z) = \begin{pmatrix} 1/z & 0 & -x/z^2 \\ 0 & 1/z & -x/z^2 \end{pmatrix}. \tag{A.118}$$

By the multivariate delta method, we obtain

$$g(x_n, y_n, z_n)^\top - g(x_0, y_0, z_0)^\top \xrightarrow{\mathcal{L}} \mathcal{MN}(0, \widetilde{\Sigma}), \quad \text{where } \widetilde{\Sigma} = J_g(x_0, y_0, z_0)\Sigma J_g(x_0, y_0, z_0)^\top. \tag{A.119}$$

Explicitly, $\widetilde{\Sigma} = (\widetilde{\sigma}_{ij})_{1\le i,j\le 2}$ is given by

$$\begin{aligned}\widetilde{\sigma}_{11} &= \frac{\sigma_{11}}{z_0^2} - 2\frac{\sigma_{13}x_0}{z_0^3} + \frac{\sigma_{33}x_0^2}{z_0^4},\\ \widetilde{\sigma}_{12} &= \frac{\sigma_{12}}{z_0^2} - \frac{\sigma_{23}x_0}{z_0^3} - \frac{\sigma_{13}y_0}{z_0^3} + \frac{\sigma_{33}x_0y_0}{z_0^4},\\ \widetilde{\sigma}_{22} &= \frac{\sigma_{22}}{z_0^2} - 2\frac{\sigma_{23}y_0}{z_0^3} + \frac{\sigma_{33}y_0^2}{z_0^4}.\end{aligned} \tag{A.120}$$

Using the above result from the multivariate delta method and the joint stable CLT in Eq. (A.101), we can derive the asymptotic distribution of the vector $\left(\sum_{i=1}^{N_{n,t}^{(c)}}(r_{n,i}^{(c)})^2/(c_n^2N_{n,t}^{(c)}), \sum_{i=1}^{N_{n,t}^{(c)}}(\overline{r}_{n,i}^{(c)})^2/(c_n^2N_{n,t}^{(c)})\right)^\top$. When $t=1$, it is the vector $(S_2, \overline{S}_{2,\epsilon})^\top$ defined in Eq. (12). We evaluate each term in Eq. (A.120) with $(x_0, y_0, z_0)^\top = \tau(t)(1, \overline{h}_{2,\epsilon}(m)/h_2(m), 1/h_2(m))^\top$:

$$\begin{aligned}\widetilde{\sigma}_{11} &= \frac{h_2(m)}{\tau(t)}(2(h_4(m)-\lambda(m)) - 2(h_4(m)-\lambda(m)) + h_4(m) - h_2^2(m))\\ &= \frac{h_2(m)}{\tau(t)}(h_4(m) - h_2^2(m)),\end{aligned} \tag{A.121}$$

$$\begin{aligned}\widetilde{\sigma}_{21} &= \frac{h_2(m)}{\tau(t)}\Bigg(\overline{\lambda}_r(m) - \frac{\overline{h}_{2,\epsilon}(m)\lambda(m)}{h_2(m)} - \overline{\lambda}(m) + \frac{h_4(m)\overline{h}_{2,\epsilon}(m)}{h_2(m)} - \left(\frac{h_4(m)\overline{h}_{2,\epsilon}(m)}{h_2(m)} - \overline{\lambda}(m)\right)\\ &\quad - \frac{\overline{h}_{2,\epsilon}(m)}{h_2(m)}(h_4(m)-\lambda(m)) + \frac{\overline{h}_{2,\epsilon}(m)}{h_2(m)}(h_4(m) - h_2^2(m))\Bigg)\\ &= \frac{h_2(m)}{\tau(t)}(\overline{\lambda}_r(m) - h_2(m)\overline{h}_{2,\epsilon}(m)),\end{aligned} \tag{A.122}$$

$$\begin{aligned}\widetilde{\sigma}_{22} &= \frac{h_2(m)}{\tau(t)}\Bigg(\overline{h}_{4,\epsilon}(m) - \frac{2\overline{h}_{2,\epsilon}(m)\overline{\lambda}(m)}{h_2(m)} + \frac{\overline{h}_{4,\epsilon}(m)\overline{h}_{2,\epsilon}^2(m)}{h_2^2(m)} - 2\left(\frac{h_4(m)\overline{h}_{2,\epsilon}^2(m)}{h_2^2(m)} - \frac{\overline{h}_{2,\epsilon}(m)\overline{\lambda}(m)}{h_2(m)}\right)\\ &\quad + \frac{\overline{h}_{2,\epsilon}^2(m)}{h_2^2(m)}(h_4(m) - h_2^2(m))\Bigg)\\ &= \frac{h_2(m)}{\tau(t)}(\overline{h}_{4,\epsilon}(m) - \overline{h}_{2,\epsilon}^2(m)).\end{aligned} \tag{A.123}$$

By the u.c.p. result in Eq. (A.67), we have

$$\sqrt{N_{n,t}^{(c)}}\left(\begin{pmatrix}\sum_{i=1}^{N_{n,t}^{(c)}}(r_{n,i}^{(c)})^2/(c_n^2N_{n,t}^{(c)})\\ \sum_{i=1}^{N_{n,t}^{(c)}}(\overline{r}_{n,i}^{(c)})^2/(c_n^2N_{n,t}^{(c)})\end{pmatrix} - \begin{pmatrix}h_2(m)\\ \overline{h}_{2,\epsilon}(m)\end{pmatrix}\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \widetilde{\Sigma}), \tag{A.124}$$

where

$$\widetilde{\Sigma} = m^{-4}\begin{pmatrix} \mu_4(m) - \mu_2^2(m) & \overline{\rho}_{2,\epsilon}(m) - \mu_2(m)\overline{\mu}_{2,\epsilon}(m) \\ \overline{\rho}_{2,\epsilon}(m) - \mu_2(m)\overline{\mu}_{2,\epsilon}(m) & \overline{\mu}_{4,\epsilon}(m) - \overline{\mu}_{2,\epsilon}^2(m) \end{pmatrix}. \tag{A.125}$$

By the same u.c.p. result and Eq. (6), we have $N_t^n / N_{n,t}^{(c)} \xrightarrow{\mathbb{P}} \mu_2(m)$, and thus

$$\sqrt{N_t^n}\left(\begin{pmatrix} \sum_{i=1}^{N_{n,t}^{(c)}}(r_{n,i}^{(c)})^2/(c_n^2 N_{n,t}^{(c)}) \\ \sum_{i=1}^{N_{n,t}^{(c)}}(\overline{r}_{n,i}^{(c)})^2/(c_n^2 N_{n,t}^{(c)}) \end{pmatrix} - \begin{pmatrix} h_2(m) \\ \overline{h}_{2,\epsilon}(m) \end{pmatrix}\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \mu_2(m)\widetilde{\Sigma}). \tag{A.126}$$

Therefore, for the random vector

$$\begin{pmatrix} h_2^{-1}(\sum_{i=1}^{N_{n,t}^{(c)}}(r_{n,i}^{(c)})^2/(c_n^2 N_{n,t}^{(c)})) \\ \overline{h}_{2,\epsilon}^{-1}(\sum_{i=1}^{N_{n,t}^{(c)}}(\overline{r}_{n,i}^{(c)})^2/(c_n^2 N_{n,t}^{(c)})) \end{pmatrix}, \tag{A.127}$$

the multivariate delta method implies that

$$\sqrt{N_t^n}\left(\begin{pmatrix} h_2^{-1}(\sum_{i=1}^{N_{n,t}^{(c)}}(r_{n,i}^{(c)})^2/(c_n^2 N_{n,t}^{(c)})) \\ \overline{h}_{2,\epsilon}^{-1}(\sum_{i=1}^{N_{n,t}^{(c)}}(\overline{r}_{n,i}^{(c)})^2/(c_n^2 N_{n,t}^{(c)})) \end{pmatrix} - \begin{pmatrix} m \\ m \end{pmatrix}\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \mu_2(m)\Lambda^\top\widetilde{\Sigma}\Lambda), \tag{A.128}$$

where

$$\Lambda = \left(\frac{1}{h_2'(h_2^{-1}(h_2^{-1}(m)))}, \frac{1}{\overline{h}_{2,\epsilon}'(\overline{h}_{2,\epsilon}^{-1}(\overline{h}_{2,\epsilon}(m)))}\right) = \left(\frac{1}{h_2'(m)}, \frac{1}{\overline{h}_{2,\epsilon}'(m)}\right). \tag{A.129}$$

This completes the proof.

A.5 Proof of Proposition 1

Proof. Firstly, we prove that the sequence of pre-averaged returns $(r_i^*)_{1\le i\le N'}$ converges in law to a centered stationary Gaussian process with desired variance under infill asymptotics for each i. We assume $k_n = 2k$ for simplicity, and expand r_i^* in terms of $\Delta_j^N X = X_j - X_{j-1}$ and ε_j:

$$\begin{aligned} r_i^* &= \frac{1}{k_n}\sum_{j=1}^{k}(X_{i+k+j} - X_{i+j}) + \frac{1}{k_n}\sum_{j=1}^{k}(\varepsilon_{i+k+j} - \varepsilon_{i+j}) \\ &= \underbrace{\sum_{j=1}^{k_n} g\Big(\frac{j}{2k}\Big)\Delta_{i+j}^N X}_{A_i} + \underbrace{\frac{1}{k_n}\sum_{j=1}^{k}(\varepsilon_{i+k+j} - \varepsilon_{i+j})}_{B_i}, \end{aligned} \tag{A.130}$$

where $g(s) = s \wedge (1-s)$ is the triangular kernel weighting function. Under Assumption 2 and by the strong approximation result in Eq. (A.20), we deduce that A_i converges in probability to $\sum_{j=1}^{k_n} g\big(\frac{j}{2k}\big)\breve{r}_{i+j}$, which is a linear combination of i.i.d. centered Gaussian random variables. The α-mixing ε with the conditions in Assumption 3 indidates a CLT under weak dependence (Theorem 1.7, Ibragimov, 1962; Theorem 8.3.7, Durrett, 2019), which implies the asymptotic Gaussianity of

B_i. The independence between X and ε implies that r_i^* converges in distribution to a centered Gaussian random variable for all i.

We now identify the limiting law of (r_i^*) by calculating its variance kernel explicitly, which also establishes the stationarity of the limiting Gaussian process. With $\mathrm{Corr}(X_j, \varepsilon_{j'}) = 0$ for any $0 \leq j, j' \leq N$, we have $\mathrm{Var}(r_i^*) = \mathrm{Var}(A_i) + \mathrm{Var}(B_i)$ with

$$\mathrm{Var}(A_i) = \sum_{j=1}^{k_n} g^2\Big(\frac{j}{k_n}\Big)(\Delta_n + o(\Delta_n)) = \frac{k_n \Delta_n}{12} + o(\sqrt{\Delta_n}). \tag{A.131}$$

For the additive noise term, we define the partial sum of ε as

$$S_{n,h} = \sum_{i=1}^{h} \varepsilon_{n+i}, \tag{A.132}$$

and start with the following results for some $\lambda \geq h$:

$$\mathrm{Var}(S_{n,h}) = \sum_{m=1-h}^{h-1} (h - |m|)\Gamma_m = h \sum_{m=1-h}^{h-1} \Big(1 - \Big|\frac{m}{h}\Big|\Big)\Gamma_m, \tag{A.133}$$

$$\begin{aligned}\mathrm{Cov}(S_{n,h}, S_{n+\lambda,h}) = \mathbb{E}[S_{n,h}S_{n+\lambda,h}] &= \sum_{i=0}^{h-1}\sum_{j=0}^{h-1} \mathrm{Cov}(\varepsilon_{n+i}, \varepsilon_{n+\lambda+i+j}) \\ &= \sum_{m=1-h}^{h-1} (h - |m|)\Gamma_{m+\lambda} = h \sum_{m=1-h}^{h-1} \Big(1 - \Big|\frac{m}{h}\Big|\Big)\Gamma_{m+\lambda},\end{aligned} \tag{A.134}$$

where the weight $1 - |m/h|$ is the Bartlett kernel. Therefore, we have

$$\begin{aligned}\mathrm{Var}(B_i) &= \frac{1}{4k^2}\mathrm{Var}(S_{i+k,k} - S_{i,k}) \\ &= \frac{1}{4k^2}\mathrm{Var}(S_{i+k,k}) + \frac{1}{4k^2}\mathrm{Var}(S_{i,k}) - 2\mathrm{Cov}(S_{i+k,k}, S_{i,k}) \\ &= \frac{1}{2k}\sum_{m=1-k}^{k-1}\Big(1 - \Big|\frac{m}{k}\Big|\Big)\Gamma_m - \frac{1}{2k}\sum_{m=1-k}^{k-1}\Big(1 - \Big|\frac{m}{k}\Big|\Big)\Gamma_{m+k}\end{aligned} \tag{A.135}$$

of the order $\sqrt{\Delta_n}$ by the absolute summability of Γ_m, which is implied by the α-mixing property of ε under Assumption 3 (Ibragimov and Linnik, 1971). Since $k_n \asymp \sqrt{N}$, both $\mathrm{Var}(A_i)$ and $\mathrm{Var}(B_i)$ are of the order $\sqrt{\Delta_n}$, such that we can ignore all terms with order smaller than $\sqrt{\Delta_n}$, which yields $\mathrm{Var}(r_i^*) = \mathrm{Var}(A_i) + \mathrm{Var}(B_i) \asymp \sqrt{\Delta_n}$.

With the time-invariant first moment and finite second moment of r_i^* for all time, in order to prove the weak stationarity of (r_i^*), we need to make sure that the autocovariance $\mathrm{Cov}(r_i^*, r_{i+\lambda}^*)$ does not vary with i. Here we firstly deal with the autocovariance of A_i. It suffices to examine the autocovariance for non-negative integer-valued lags λ, as the autocovariance function is always symmetric.

$$\mathrm{Cov}(A_i, A_{i+\lambda}) = \mathbb{E}[A_i A_{i+\lambda}] = \mathbb{E}\left[\sum_{j=1}^{k_n} g\Big(\frac{j}{k_n}\Big)\Delta_{i+j}^N X \sum_{\eta=1}^{k_n} g\Big(\frac{\eta}{k_n}\Big)\Delta_{i+\lambda+\eta}^N X\right]. \tag{A.136}$$

When $\lambda \geq k_n$, $\mathrm{Cov}(A_i, A_{i+\lambda}) = 0$. When $1 \leq \lambda \leq k_n - 1$, we have

$$\begin{aligned}\mathrm{Cov}(A_i, A_{i+\lambda}) &= \mathbb{E}\left[\sum_{j=1}^{k_n-\lambda} g\Big(\frac{j}{k_n}\Big)g\Big(\frac{j+\lambda}{k_n}\Big)(\Delta_{i+\lambda+j}^N X)^2\right] \\ &= \sum_{j=1}^{k_n-\lambda} g\Big(\frac{j}{k_n}\Big)g\Big(\frac{j+\lambda}{k_n}\Big)\mathbb{E}[(\Delta_{i+\lambda+j}^N X)^2] = O(\sqrt{\Delta_n}).\end{aligned} \tag{A.137}$$

For the noise term, we have the lag-λ autocovariance

$$\begin{aligned}&\mathrm{Cov}(B_i, B_{i+\lambda}) = \frac{1}{4k^2}\mathbb{E}[(S_{i+k,k} - S_{i,k})(S_{i+k+\lambda,k} - S_{i+\lambda,k})] \\ &= \frac{1}{4k^2}\left(\mathbb{E}[S_{i+k,k}S_{i+k+\lambda,k}] + \mathbb{E}[S_{i,k}S_{i+\lambda,k}] - \mathbb{E}[S_{i+k,k}S_{i+\lambda,k}] - \mathbb{E}[S_{i,k}S_{i+k+\lambda,k}]\right) \\ &= \frac{1}{2k}\sum_{m=1-k}^{k-1}\Big(1 - \Big|\frac{m}{k}\Big|\Big)\Gamma_{m+\lambda} - \frac{1}{4k}\sum_{m=1-k}^{k-1}\Big(1 - \Big|\frac{m}{k}\Big|\Big)\Gamma_{m+\lambda-k} - \frac{1}{4k}\sum_{m=1-k}^{k-1}\Big(1 - \Big|\frac{m}{k}\Big|\Big)\Gamma_{m+\lambda+k} \\ &= O(\sqrt{\Delta_n}),\end{aligned} \tag{A.138}$$

by the absolute summability of Γ_m. In the limit, both the covariances are finite and time-invariant (not depend on i) for all possible $\lambda \in \mathbb{N}$, which implies the weak stationarity of (r_i^*) in the limit, as desired.

For Step 2, we first demonstrate how the random sign flip eliminates serial correlations in (r_i^*). Let $F(x) = \mathbb{P}(r_i^* \leq x)$ denote the CDF of r_i^*. It is obvious that the product $\delta_i r_i^*$ is a Gaussian random variable with the same distribution:

$$\begin{aligned}\mathbb{P}(\delta_i r_i^* \leq x) &= \mathbb{P}(\delta_i = 1)\mathbb{P}(\delta_i r_i^* \leq x | \delta_i = 1) + \mathbb{P}(\delta_i = -1)\mathbb{P}(\delta_i r_i^* \leq x | \delta_i = -1) \\ &= \frac{1}{2}\mathbb{P}(r_i^* \leq x) + \frac{1}{2}\mathbb{P}(r_i^* \geq -x) = F(x),\end{aligned} \tag{A.139}$$

and the autocovariance function for any $i \in \{1, \ldots, N' - \lambda\}$ satisfies

$$\mathrm{Cov}(\delta_i r_i^*, \delta_{i+\lambda} r_{i+\lambda}^*) = \mathbb{E}[\delta_i \delta_{i+\lambda} r_i^* r_{i+\lambda}^*] = \mathbb{E}[\delta_i]\mathbb{E}[\delta_{i+\lambda}]\mathrm{Cov}(r_i^*, r_{i+\lambda}^*) = 0. \tag{A.140}$$

Next, we establish that, following the uniform random permutation $\pi : \{1, \ldots, N'\} \mapsto \{1, \ldots, N'\}$, any two variables in $(\tilde{r}_i)_{1 \leq i \leq N'}$ are independent when their indices are not sufficiently distant from each other each other in $\{1, \ldots, N'\}$ under infill asymptotics. We start with a formal definition of the local independence for a discrete-time stochastic process: The process $X = (X_i)_{1 \leq i \leq n}$ is said to be locally independent if

$$\begin{aligned} &\lim_{n\to\infty} \sup_{\substack{1\le i,j\le n\\ 1\le |i-j|\le \mathit{\Lambda}_n}} \mathbb{P}(X_i \text{ and } X_j \text{ are dependent}) = 0, \\ \text{or} \quad &\lim_{n\to\infty} \sup_{\substack{1\le i,j\le n\\ 1\le |i-j|\le \mathit{\Lambda}_n}} \{|\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A\in\sigma(X_i), B\in\sigma(X_j)\} = 0, \end{aligned} \tag{A.141}$$

where $\mathit{\Lambda}_n \asymp n^{\varpi}$ for some $\varpi \in (0,1)$, such that X_i is independent to other variables in X whose indices fall within the interval $[i-\mathit{\Lambda}_n, i+\mathit{\Lambda}_n]$. In our case, we need to verify

$$\lim_{n\to\infty} \sup_{\substack{1\le i,j\le N'\\ 1\le |i-j|\le \mathit{\Lambda}_n}} \mathbb{P}(\tilde{r}_i \text{ and } \tilde{r}_j \text{ are dependent}) = 0. \tag{A.142}$$

The fact that $(\varepsilon_i)_{0\le i\le N}$ is α-mixing implies that

$$\alpha(\mathit{\Lambda}_n) = \sup\{|\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A\in\sigma(\varepsilon_i), B\in\sigma(\varepsilon_{i+\mathit{\Lambda}_n})\} \to 0, \tag{A.143}$$

as $n\to\infty$, thus ε_i and ε_j are asymptotically independent if $|i-j|\ge \mathit{\Lambda}_n$.

With the uniform random permutation, we denote

$$\tilde{r}_i = r^*_{\pi(i)}\delta_{\pi(i)} \quad \text{and} \quad \tilde{r}_j = r^*_{\pi(j)}\delta_{\pi(j)} \tag{A.144}$$

where $\pi(i)$, $\pi(j)$ are the corresponding indices of the products before permutation. Therefore, for all $1\le i,j\le N'$ and $1\le |i-j|\le \mathit{\Lambda}_n$, $\tilde{r}_i$ and $\tilde{r}_j$ are independent if the corresponding indices $\pi(i)$ and $\pi(j)$ are sufficiently far apart from one another:

$$\begin{aligned} \mathbb{P}(\tilde{r}_i \text{ and } \tilde{r}_j \text{ are dependent}) &= \mathbb{P}(r^*_{\pi(i)} \text{ and } r^*_{\pi(j)} \text{ are dependent}) \\ &= \mathbb{P}(\sigma(\{\varepsilon_{\pi(i)+\ell} : 0\le \ell\le k_n\}) \text{ and } \sigma(\{\varepsilon_{\pi(j)+\ell} : 0\le \ell\le k_n\}) \text{ are dependent}) \\ &\le 2\mathbb{P}(\pi(i)+1 \le \pi(j) \le \pi(i)+k_n+\mathit{\Lambda}_n) \\ &= \frac{2(k_n+\mathit{\Lambda}_n)}{N'-1} = O(\Delta_n^{\gamma}), \quad \text{where } \gamma = 1-\max\Big\{\frac{1}{2}, \varpi\Big\}. \end{aligned} \tag{A.145}$$

For a sequence of N' variables, the uniform random permutation ensures that each of the $N'!$ possible permutations are equally likely and that each "ball" $r^*_{\pi(i)}\delta_{\pi(i)}$ has an equal chance of being placed into any "box" i, which has become a question of classical probability. As $n\to\infty$, $\tilde{r}_i$ and $\tilde{r}_j$ with $1\le |i-j|\le \mathit{\Lambda}_n$ are asymptotically independent. This completes the proof. □

Appendix B Supplementary Results

B.1 Parameter Choices for Other Tests

For other tests constructed in Sections 4 and 5, we clarify some specific parameter choices:

- LM: For the local realized bipower variation, we consider the window size $K = \sqrt{252N}$, where N is the number of sampled observations.
- ASJ: For the multipower variations constructed on two different sampling intervals δ and $k\delta$, we select $p = 4$ and $k = 2$, which satisfies the requirement.
- CPR: For the auxiliary local variance estimator, we employ the nonparametric filter of length $2L + 1$ with $L = 25$ and a Gaussian kernel, which follows the recommendation in Appendix B of Corsi et al. (2010).
- PZ: We employ the truncated realised power variation with $p = 4$ and the truncation threshold $cN^{-\varpi}$, where c and ϖ follow the recommendation in Section 5 of Podolskij and Ziggel (2010). For the noise-adjusted version, we select the pre-averaging window $k_n = 0.5\lfloor\sqrt{N}\rfloor$.
- LM12: We select the pre-averaging window $k_n = 0.4\lfloor\sqrt{N}\rfloor$, which minimizes the absolute distance between the nominal size and the empirical size with the simulated tick-level noise-contaminated observations.
- ASJL: We select the pre-averaging window $k_n = 0.9\lfloor\sqrt{N}\rfloor$ based on the simulated noise-contaminated data, and the truncation level $C = 5$.

B.2 Jump Detection

As detailed in Section 4.3, we consider the common jump filtering and detection method as a benchmark, which is based on the sequential detection approach of Andersen et al. (2007) and the thresholding technique of Mancini (2009). Particularly, we adjust the threshold parameter k with two types of FWER corrections. Specifically, given N_{spl} tests of null hypotheses (N_{spl} sampled returns) and a family-wise significance level of α, we select the corresponding k for each individual return at α':

- Šidák correction: $\alpha' = 1 - (1 - \alpha)^{1/N_{\text{spl}}}$,
- Bonferroni correction: $\alpha' = \alpha/N_{\text{spl}}$.

Table B.1 presents the finite-sample size and size-adjusted power of the truncation-based filtering technique in the absence of market microstructure noise. The truncation thresholds are determined with (i) the latent true volatility, (ii) the localized tick-by-tick BV, and (iii) the localized pre-averaged BV of Podolskij and Vetter (2009). The threshold parameter k is adjusted with both FWER corrections, and we set $\varpi = 0.5$. Following the same procedure used for the noise case in Table 5, the spot volatility estimates are recursively obtained within a backward-looking tick-time window of 1,800 ticks. The pre-averaging window is chosen to be $\lceil 0.5\sqrt{1800}\rceil$ ticks.

Table B.1: Finite-sample size and power (%) of truncation-based filtering technique in the absence of market microstructure noise

Nominal size: 5%					
Panel A: Normalization with true spot volatility					
		No Jump (with FWER control)			
Ticks	N_{spl}	Šidák	Bonferroni	Moderate Jumps	Large Jumps
1	23400	4.92	4.83	93.13	96.80
5	4680	5.15	4.98	84.91	91.88
15	1560	5.38	5.25	75.12	87.18
30	780	4.96	4.80	67.55	82.46
60	390	5.13	4.99	57.84	76.16
120	195	5.22	5.08	46.50	69.19
180	130	4.93	4.81	40.92	64.71
300	78	5.03	4.90	31.98	56.20
Panel B: Normalization with localized tick-by-tick BV					
		No Jump (with FWER control)			
Ticks	N_{spl}	Šidák	Bonferroni	Moderate Jumps	Large Jumps
1	23400	5.95	5.74	91.21	95.92
5	4680	5.21	5.07	82.41	90.80
15	1560	5.31	5.23	72.45	85.60
30	780	5.06	4.87	64.37	81.08
60	390	5.13	5.01	53.72	74.08
120	195	5.15	5.02	42.28	66.33
180	130	4.89	4.75	36.51	61.57
300	78	4.98	4.89	27.52	52.92
Panel C: Normalization with localized pre-averaged BV					
		No Jump (with FWER control)			
Ticks	N_{spl}	Šidák	Bonferroni	Moderate Jumps	Large Jumps
1	23400	6.02	5.92	91.23	95.85
5	4680	5.87	5.73	82.44	90.72
15	1560	5.51	5.45	72.26	85.50
30	780	4.78	4.68	64.46	81.00
60	390	4.81	4.69	53.53	73.84
120	195	4.93	4.79	41.99	65.94
180	130	4.57	4.52	35.95	61.17
300	78	4.68	4.58	27.52	52.52

This table reports the finite-sample size and size-adjusted power (%) of 10,000 simulations of the truncation-based jump filtering technique in the absence of market microstructure noise. Observations are sampled at various multiples of ticks, where "N_{spl}" stands for the corresponding sampling frequencies. The truncation thresholds are constructed from (i) the latent true volatility, (ii) the localized tick-by-tick BV, and (iii) the localized pre-averaged BV. The threshold parameter k is adjusted with both the Šidák and Bonferroni corrections, and $\varpi = 0.5$.

In addition to the noise-case results in Table 5, we follow the empirical applications of Aït-Sahalia et al. (2025) to consider a broad range of fixed k from 3.5 to 9. Fig. B.1 illustrates the rejection rates under both the null and alternative hypotheses across various frequencies of tick-time sampling.

Furthermore, we extend the comparisons by examining returns sampled at equidistant calendar-time intervals. To estimate the spot volatility for each calendar-time interval and avoid the impact of tick irregularity, we construct the pre-averaged BV from all tick-level price observations within each day. We then adjust these daily volatility estimates to account for intraday volatility pattern for each calendar-time interval, where we follow Aleti et al. (2025) for the time-of-day adjustment

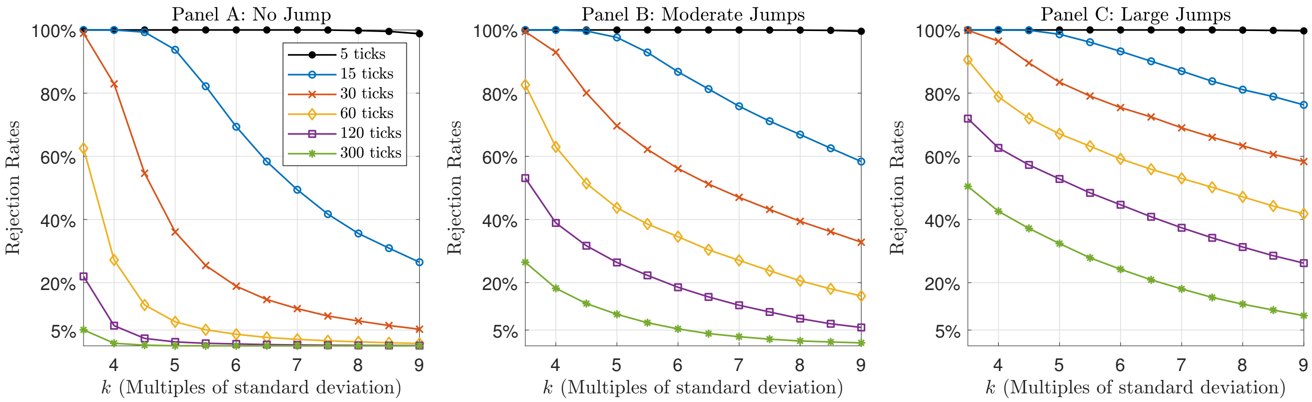

Figure B.1: Rejection rates of 10,000 simulations of the truncation-based jump filtering technique. Observations are sampled at various multiples of ticks. The truncation thresholds are constructed from the localized pre-averaged BV of Podolskij and Vetter (2009) computed within each backward-looking window of 1,800 ticks. The threshold parameter k is varied from 3.5 to 9, corresponding to progressively more stringent cutoffs, and $\varpi = 0.5$.

θ_j for the j-th calendar-time interval:

$$\theta_j = \left(\sum_{i=1}^{M} r_{i,j}^2\right) \Big/ \left(\frac{1}{N_{\text{spl}}}\sum_{i=1}^{M}\sum_{j=1}^{N_{\text{spl}}} r_{i,j}^2\right), \tag{B.1}$$

where $r_{i,j}$ is the return on the asset on day i at interval j, and θ_j is a simple ex ante measure of the fraction of daily RV that arrives at each time-of-day. Given a daily pre-averaged BV estimate BV_i, the spot volatility $\sigma_{i,j}$ is then estimated as $\hat{\sigma}_{i,j} = \sqrt{\text{BV}_i\theta_j}$. Table B.2 presents the finite-sample size and size-adjusted power of the truncation-based filtering technique applied to calendar-time sampled returns, with the threshold parameter k calibrated with both corrections to control a 5% FWER under the null.

Table B.2: Finite-sample size and power (%) of truncation-based filtering technique with calendar-time sampling

Nominal size: 5%					
		Panel A No Jump (with FWER control)		Panel B	Panel C
Int. (sec)	N_{spl}	Šidák	Bonferroni	Moderate Jumps	Large Jumps
5	4680	100.00	100.00	15.58	39.79
15	1560	99.95	99.95	22.49	49.02
30	780	85.80	85.51	30.92	56.68
60	390	40.69	40.27	38.61	62.75
120	195	16.44	16.17	41.29	63.64
180	130	10.41	10.20	39.53	60.79
300	78	7.34	7.14	34.09	55.87

This table reports the finite-sample size and size-adjusted power (%) of 10,000 simulations of the truncation-based jump filtering technique. Observations are sampled at equidistant intervals in calendar time, where "N_{spl}" stands for the corresponding sampling frequencies. The truncation thresholds are constructed from the pre-averaged BV of Podolskij and Vetter (2009), with the intraday volatility seasonality incorporated, and the threshold parameter k is adjusted with both the Šidák and Bonferroni corrections, and $\varpi = 0.5$.

Similarly, we consider some fixed k varying from 3.5 to 9 for the truncation thresholds applied to calendar-time sampled returns. Fig. B.2 illustrates the rejection rates under both the null and alternative hypotheses across various frequencies of calendar-time sampling.

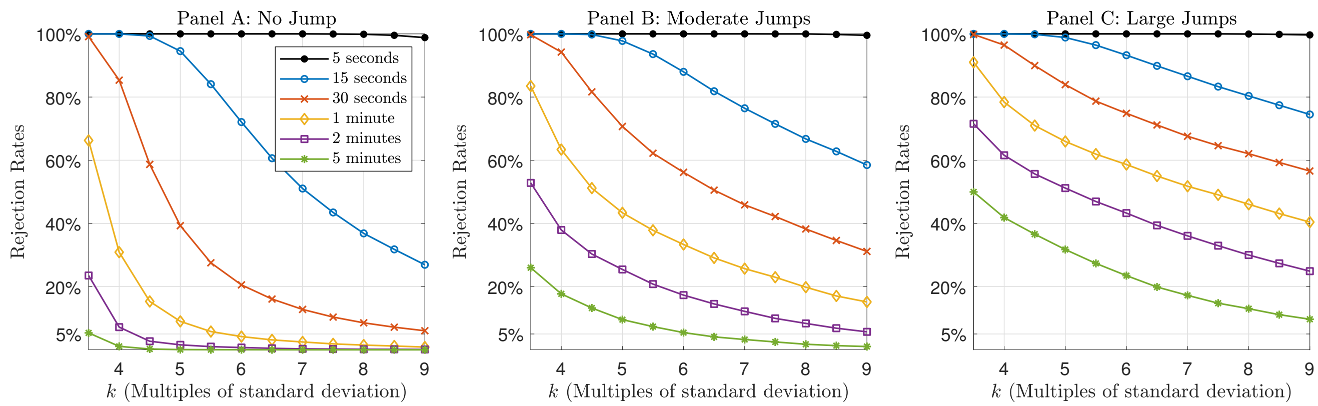

Figure B.2: Rejection rates of 10,000 simulations of the truncation-based jump filtering technique. Observations are sampled at equidistant intervals in calendar time. The truncation thresholds are constructed from the pre-averaged bipower variation of Podolskij and Vetter (2009), with the intraday volatility seasonality incorporated. The threshold parameter k is varied from 3.5 to 9, corresponding to progressively more stringent cutoffs, and $\varpi = 0.5$.

B.3 Simulation Results with Other Noise Specifications

In addition to the simulation results in Section 4, we consider three other specifications for the additive noise that follows Aït-Sahalia et al. (2012) as robustness checks:

(i) Gaussian noise:

$$\varepsilon_i = 2Z_i\sqrt{\frac{\sigma^2_{t_{n,i}}}{n}}, \tag{B.2}$$

where Z_i are i.i.d. draws from a standard normal distribution, see Tables B.3 to B.6.

(ii) Autocorrelated Gaussian noise:

$$\varepsilon_i = 2\omega_i^A\sqrt{\frac{\sigma^2_{t_{n,i}}}{n}}, \tag{B.3}$$

where ω_i^A is an autocorrelated Gaussian defined in Eq. (27), see Tables B.7 to B.10.

(iii) t-distributed noise:

$$\varepsilon_i = 2\omega_i^B\sqrt{\frac{\nu-2}{\nu}}\sqrt{\frac{\sigma^2_{t_{n,i}}}{n}}, \tag{B.4}$$

where ω_i^B are i.i.d. draws from a Student's t distribution with the degree of freedom ν, see Tables B.11 to B.14.

Table B.3: Finite-sample size and power (%) under Gaussian noise

Nominal size: 5%		$\theta = 0.3$				$\theta = 0.4$				$\theta = 0.5$			
			ϵ				ϵ				ϵ		
	$c/\sigma(\tilde{r}_i)$	$N^{(c)}$	0.05	0.07	0.10	$N^{(c)}$	0.05	0.07	0.10	$N^{(c)}$	0.05	0.07	0.10
	3	1785	4.90	5.21	5.35	1783	5.07	5.05	5.76	1783	5.01	5.07	5.70
	4	1099	5.33	5.03	5.49	1098	4.93	5.14	5.81	1098	5.24	5.21	5.62
	5	743	4.71	5.14	5.33	743	5.18	5.28	5.43	742	5.10	5.70	5.27
Panel A	6	535	5.26	5.01	5.49	535	4.63	4.89	5.53	535	4.82	5.47	5.13
No Jump	7	404	4.71	4.99	5.19	404	4.61	5.55	5.62	404	5.17	5.18	5.08
	8	316	4.83	4.79	5.59	315	4.83	5.23	5.73	315	5.08	5.12	5.30
	9	254	5.44	4.80	5.30	253	5.22	5.00	5.38	253	5.20	5.20	5.73
	10	208	4.93	5.31	5.71	208	4.98	5.41	5.68	208	5.18	5.28	5.75
	3	1715	47.74	49.52	50.82	1716	44.73	47.04	49.52	1718	42.85	45.35	47.45
	4	1058	46.68	48.74	50.80	1059	43.38	46.51	48.46	1061	41.71	44.74	46.93
	5	717	45.27	47.51	50.26	718	43.27	45.29	48.09	720	41.23	43.10	45.95
Panel B	6	518	44.45	47.15	49.63	519	43.24	44.57	46.91	520	41.10	42.06	45.13
Moderate Jump	7	392	44.91	46.79	49.45	393	42.51	43.70	46.51	394	40.31	41.96	44.21
	8	307	42.82	46.95	49.09	308	41.76	43.10	45.84	308	39.45	40.95	44.10
	9	247	42.13	45.59	48.45	248	40.81	42.65	46.06	248	38.43	40.97	43.46
	10	203	41.42	44.81	48.32	204	40.32	41.82	46.46	204	38.26	40.44	43.70
	3	1587	69.98	71.41	73.31	1589	68.18	69.74	71.08	1594	67.26	68.39	69.95
	4	979	68.91	71.10	72.74	982	67.81	69.39	70.89	985	65.92	67.49	69.41
	5	665	68.80	70.14	72.20	668	66.75	69.09	70.32	670	65.28	66.87	69.25
Panel C	6	482	67.64	69.69	71.54	485	66.71	68.34	69.77	487	64.99	66.15	68.20
Large Jump	7	365	67.61	69.10	71.27	368	65.47	66.78	69.89	370	63.94	65.80	67.99
	8	287	67.37	68.95	71.00	289	65.11	66.62	68.93	291	64.21	65.45	66.91
	9	232	65.90	69.22	70.84	234	64.78	66.60	68.88	235	63.87	64.83	67.73
	10	191	65.47	68.02	70.71	193	64.23	65.74	68.45	194	63.14	64.76	67.56

This table reports the finite-sample size and size-adjusted power (%) of 10,000 simulations of the test statistic $T_{c,\epsilon}$ at 5% nominal level. All simulated prices are contaminated by the additive Gaussian noise and rounding errors. We utilize the two-step noise reduction method in Section 3.2 to construct the sequence of pseudo-observations with three different pre-averaging windows, i.e., $k_n = \lceil \theta\sqrt{N} \rceil$ with $\theta \in \{0.3, 0.4, 0.5\}$. The observations are sampled with different PDS barrier widths $c = K\sigma(\tilde{r}_i)$, where K ranges from 3 to 10. Different censoring thresholds with $\epsilon \in \{0.05, 0.07, 0.1\}$ are considered. $N^{(c)}$ stands for the average sampling frequencies.

Table B.4: Finite-sample size and power (%) of other tests under Gaussian noise

Nominal size: 5%	Int. (sec)	N_{spl}	BNS	JO	LM	ASJ	CPR	PZ	MinRV	MedRV
Panel A No Jump	5	4680	0.23	1.06	14.02	100.00	0.37	5.59	0.00	0.00
	15	1560	4.93	3.70	22.26	93.73	5.43	9.91	0.91	2.89
	30	780	7.88	5.02	29.32	38.68	8.42	12.55	4.04	6.35
	60	390	7.69	6.23	27.86	13.10	8.26	14.47	5.37	7.14
	120	195	7.49	8.07	17.76	7.10	8.02	16.23	5.71	7.93
	180	130	7.91	9.05	15.11	5.36	8.70	16.12	5.78	8.78
	300	78	7.74	10.98	11.96	4.22	8.70	14.91	5.67	9.12
Panel B Moderate Jump	5	4680	44.28	51.82	69.09	99.76	47.13	66.49	40.11	45.46
	15	1560	40.43	44.90	60.35	92.97	43.88	61.13	37.19	41.85
	30	780	36.17	38.30	51.11	65.25	39.14	52.79	33.48	36.97
	60	390	29.52	30.92	42.23	37.60	32.97	43.63	27.36	31.32
	120	195	21.55	22.20	36.06	22.08	24.92	32.72	21.00	24.32
	180	130	17.48	17.40	30.52	14.55	20.17	26.98	17.02	20.84
	300	78	15.27	11.91	21.62	11.67	17.54	19.96	14.36	16.51
Panel C Large Jump	5	4680	68.50	74.10	84.55	99.83	70.65	82.69	64.96	68.91
	15	1560	65.66	69.37	79.03	95.72	68.29	79.36	62.52	66.47
	30	780	61.28	64.47	73.60	78.16	64.28	74.79	58.50	62.36
	60	390	55.16	57.63	67.45	57.04	58.50	68.70	52.97	57.53
	120	195	46.02	48.22	61.94	36.07	50.16	59.34	44.72	49.98
	180	130	41.61	42.27	56.76	26.63	45.35	53.82	40.37	44.59
	300	78	35.12	33.77	46.86	17.83	39.24	45.27	34.30	38.21

This table reports the finite-sample size and size-adjusted power (%) of 10,000 simulations of 8 classical tests at 5% nominal level: BNS (Barndorff-Nielsen and Shephard, 2006), JO (Jiang and Oomen, 2008), LM (Lee and Mykland, 2008), ASJ (Aït-Sahalia and Jacod, 2009), CPR (Corsi et al., 2010), PZ (Podolskij and Ziggel, 2010), MinRV and MedRV (Andersen et al., 2012). All simulated prices are contaminated by the additive Gaussian noise and rounding errors. All these tests are constructed on observations equidistantly sampled with various intervals in calendar time: 5, 15, 30, 60, 120, 180 and 300 seconds, and "N_{spl}" stands for the sampling frequencies.

Table B.5: Finite-sample size and power (%) of other noise-robust tests under Gaussian noise

Nominal size: 5%	Int. (sec)	N_{spl}	PZ*	LM12	ASJL
Panel A: No Jump	tick	23400	5.10	3.27	5.12
	5	4680	4.93	8.59	8.79
Panel B: Moderate Jump	tick	23400	39.34	24.12	38.06
	5	4680	29.96	18.97	16.88
Panel C: Large Jump	tick	23400	64.18	39.18	62.90
	5	7680	56.03	32.23	41.41

This table reports the finite-sample size and size-adjusted power (%) of 10,000 simulations of 3 noise-robust tests at 5% nominal level: noise-adjusted PZ (Podolskij and Ziggel, 2010), LM12 (Lee and Mykland, 2012), and ASJL (Aït-Sahalia et al., 2012). All simulated prices are contaminated by the additive Gaussian noise and rounding errors. All these tests are constructed on tick-level and 5-second-sampled observations. The tuning parameters for those tests are selected by minimizing the absolute distance between the nominal size and the empirical size with the simulated tick-level noise-contaminated observations.

Table B.6: Finite-sample size and power (%) of truncation-based filtering technique under Gaussian noise

Nominal size: 5%		Panel A No Jump (with FWER control)		Panel B	Panel C
Ticks	N_{spl}	Šidák	Bonferroni	Moderate Jumps	Large Jumps
1	23400	100.00	100.00	74.84	86.80
5	4680	100.00	100.00	72.45	85.36
15	1560	73.56	73.04	66.19	82.02
30	780	28.77	28.23	59.89	78.41
60	390	12.97	12.69	51.74	73.25
120	195	7.89	7.72	40.75	64.72
180	130	6.25	6.12	35.06	59.93
300	78	5.77	5.61	27.01	53.77

This table reports the finite-sample size and size-adjusted power (%) of 10,000 simulations of the truncation-based jump filtering technique. All simulated prices are contaminated by the additive autocorrelated Gaussian noise and rounding errors. Observations are sampled at various multiples of ticks, where "N_{spl}" stands for the corresponding sampling frequencies. The truncation thresholds are constructed from the localized pre-averaged BV of Podolskij and Vetter (2009) computed within each backward-looking window of 1,800 ticks. The threshold parameter k is adjusted with both the Šidák and Bonferroni corrections, and $\varpi = 0.5$.

Table B.7: Finite-sample size and power (%) under autocorrelated Gaussian noise

Nominal size: 5%		$\theta = 0.3$				$\theta = 0.4$				$\theta = 0.5$			
			ϵ				ϵ				ϵ		
	$c/\sigma(\tilde{r}_i)$	$N^{(c)}$	0.05	0.07	0.10	$N^{(c)}$	0.05	0.07	0.10	$N^{(c)}$	0.05	0.07	0.10
Panel A No Jump	3	1785	4.85	5.28	5.39	1784	5.14	5.34	5.69	1783	5.38	5.70	5.84
	4	1099	5.02	5.36	5.34	1099	5.05	4.94	5.32	1097	5.23	5.62	5.69
	5	743	4.84	5.45	4.96	743	4.82	5.58	5.37	743	5.41	5.64	6.07
	6	536	4.71	5.14	5.30	536	4.87	5.32	5.23	535	4.73	5.33	5.63
	7	404	5.21	5.36	5.44	404	5.11	4.91	5.60	404	5.20	5.02	5.45
	8	316	4.74	5.21	5.50	316	4.91	4.79	5.74	316	4.75	5.17	5.41
	9	253	4.56	5.05	5.37	254	4.92	5.16	5.35	253	4.86	5.36	5.32
	10	208	4.87	5.45	5.33	208	5.01	5.48	5.77	208	5.36	5.42	5.75
Panel B Moderate Jump	3	1715	47.02	49.41	52.39	1717	45.14	47.22	49.69	1719	43.27	45.53	48.29
	4	1058	46.22	48.63	51.34	1059	43.97	46.88	48.80	1061	42.70	44.28	47.16
	5	717	45.80	47.83	51.35	719	43.38	45.64	48.35	720	40.70	43.64	45.82
	6	518	45.25	46.85	49.59	519	41.62	45.13	47.71	520	40.83	42.25	45.93
	7	392	43.53	46.57	48.48	393	41.06	44.33	46.80	394	39.47	42.48	45.30
	8	307	43.39	45.85	49.41	308	41.95	43.80	46.44	309	39.25	41.64	44.64
	9	247	43.28	45.83	48.46	248	40.78	43.20	46.57	248	39.04	40.64	45.10
	10	203	42.70	44.97	48.27	204	40.26	41.86	45.96	204	38.51	40.48	43.28
Panel C Large Jump	3	1587	69.39	70.70	72.91	1590	67.87	69.33	71.27	1594	66.82	68.17	70.24
	4	979	68.87	70.46	72.68	983	66.80	69.17	70.63	985	65.84	67.74	69.17
	5	665	67.98	70.11	72.16	668	66.93	68.16	70.27	671	65.21	66.26	68.55
	6	482	68.20	69.38	71.78	485	65.87	67.95	69.81	487	64.66	66.04	67.63
	7	366	66.57	68.81	71.08	368	65.20	67.12	68.95	370	64.53	66.31	67.93
	8	287	67.04	69.14	71.08	289	64.39	66.85	69.28	291	63.97	65.09	67.34
	9	232	66.01	68.47	70.40	234	64.17	66.33	69.04	235	63.15	64.93	67.69
	10	191	66.15	67.95	70.27	193	63.67	65.80	68.23	194	62.39	64.69	67.11

This table reports the finite-sample size and size-adjusted power (%) of 10,000 simulations of the test statistic $T_{c,\epsilon}$ at 5% nominal level. All simulated prices are contaminated by the additive autocorrelated Gaussian noise and rounding errors. We utilize the two-step noise reduction method in Section 3.2 to construct the sequence of pseudo-observations with three different pre-averaging windows, i.e., $k_n = \lceil \theta\sqrt{N} \rceil$ with $\theta \in \{0.3, 0.4, 0.5\}$. The observations are sampled with different PDS barrier widths $c = K\sigma(\tilde{r}_i)$, where K ranges from 3 to 10. Different censoring thresholds with $\epsilon \in \{0.05, 0.07, 0.1\}$ are considered. $N^{(c)}$ stands for the average sampling frequencies.

Table B.8: Finite-sample size and power (%) of other tests under autocorrelated Gaussian noise

Nominal size: 5%	Int. (sec)	N_{spl}	BNS	JO	LM	ASJ	CPR	PZ	MinRV	MedRV
Panel A No Jump	5	4680	0.00	0.72	10.46	100.00	0.00	5.19	0.00	0.00
	15	1560	2.48	2.79	19.26	97.34	2.84	8.11	0.40	1.59
	30	780	5.59	4.27	26.71	47.36	6.32	11.46	2.93	4.99
	60	390	6.84	5.87	26.53	14.89	7.32	13.67	4.93	6.60
	120	195	7.08	7.43	17.00	8.51	7.64	15.53	5.55	7.50
	180	130	7.33	8.53	14.31	5.60	8.09	15.63	5.56	7.98
	300	78	7.92	10.90	12.15	4.35	9.20	15.09	5.73	9.55
Panel B Moderate Jump	5	4680	42.34	49.60	68.05	99.81	45.64	64.41	37.27	42.66
	15	1560	39.11	43.61	59.86	93.84	42.79	60.18	36.30	40.82
	30	780	36.35	37.43	50.10	66.10	40.00	52.66	32.51	37.08
	60	390	28.49	29.46	41.52	39.38	32.06	43.30	26.18	30.62
	120	195	22.16	21.09	34.90	21.09	25.48	32.18	20.48	23.39
	180	130	17.80	16.58	29.83	15.66	20.78	25.61	17.18	19.83
	300	78	13.36	11.19	19.64	10.68	14.91	18.07	12.98	14.99
Panel C Large Jump	5	4680	66.02	71.75	83.05	99.73	68.49	80.74	61.79	66.31
	15	1560	63.75	67.48	78.59	95.40	66.49	78.52	61.00	64.90
	30	780	60.36	62.07	72.40	78.66	63.52	73.58	57.06	61.05
	60	390	53.77	55.08	65.78	56.45	57.05	67.04	51.59	55.26
	120	195	46.55	46.82	60.23	35.58	49.80	57.96	44.44	48.75
	180	130	40.73	41.31	55.31	24.87	44.99	51.75	39.56	44.64
	300	78	33.58	32.74	44.99	16.57	37.28	42.89	32.71	36.46

This table reports the finite-sample size and size-adjusted power (%) of 10,000 simulations of 8 classical tests at 5% nominal level: BNS (Barndorff-Nielsen and Shephard, 2006), JO (Jiang and Oomen, 2008), LM (Lee and Mykland, 2008), ASJ (Aït-Sahalia and Jacod, 2009), CPR (Corsi et al., 2010), PZ (Podolskij and Ziggel, 2010), MinRV and MedRV (Andersen et al., 2012). All simulated prices are contaminated by the additive autocorrelated Gaussian noise and rounding errors. All these tests are constructed on observations equidistantly sampled with various intervals in calendar time: 5, 15, 30, 60, 120, 180 and 300 seconds, and "N_{spl}" stands for the sampling frequencies.

Table B.9: Finite-sample size and power (%) of other noise-robust tests under autocorrelated Gaussian noise

Nominal size: 5%	Int. (sec)	N_{spl}	PZ*	LM12	ASJL
Panel A: No Jump	tick	23400	5.06	2.91	5.19
	5	4680	4.98	8.10	8.92
Panel B: Moderate Jump	tick	23400	38.51	21.87	37.46
	5	4680	29.10	18.91	17.09
Panel C: Large Jump	tick	23400	65.58	39.64	63.69
	5	7680	55.98	32.62	41.88

This table reports the finite-sample size and size-adjusted power (%) of 10,000 simulations of 3 noise-robust tests at 5% nominal level: noise-adjusted PZ (Podolskij and Ziggel, 2010), LM12 (Lee and Mykland, 2012), and ASJL (Aït-Sahalia et al., 2012). All simulated prices are contaminated by the additive autocorrelated Gaussian noise and rounding errors. All these tests are constructed on tick-level and 5-second-sampled observations. The tuning parameters for those tests are selected by minimizing the absolute distance between the nominal size and the empirical size with the simulated tick-level noise-contaminated observations.

Table B.10: Finite-sample size and power (%) of truncation-based detection under autocorrelated Gaussian noise

Nominal size: 5%		Panel A No Jump (with FWER control)		Panel B	Panel C
Ticks	N_{spl}	Šidák	Bonferroni	Moderate Jumps	Large Jumps
1	23400	100.00	100.00	72.29	85.15
5	4680	100.00	100.00	70.59	85.03
15	1560	85.80	85.44	65.00	81.64
30	780	36.99	36.46	58.82	77.33
60	390	15.62	15.28	51.05	72.25
120	195	8.96	8.74	41.05	64.41
180	130	7.06	6.86	34.71	59.46
300	78	5.83	5.73	27.36	52.70

This table reports the finite-sample size and size-adjusted power (%) of 10,000 simulations of the truncation-based jump filtering technique. All simulated prices are contaminated by the additive autocorrelated Gaussian noise and rounding errors. Observations are sampled at various multiples of ticks, where "N_{spl}" stands for the corresponding sampling frequencies. The truncation thresholds are constructed from the localized pre-averaged BV of Podolskij and Vetter (2009) computed within each backward-looking window of 1,800 ticks. The threshold parameter k is adjusted with both the Šidák and Bonferroni corrections, and $\varpi = 0.5$.

Table B.11: Finite-sample size and power (%) under t-distributed noise

Nominal size: 5%		$\theta = 0.3$				$\theta = 0.4$				$\theta = 0.5$			
			ϵ				ϵ				ϵ		
	$c/\sigma(\tilde{r}_i)$	$N^{(c)}$	0.05	0.07	0.10	$N^{(c)}$	0.05	0.07	0.10	$N^{(c)}$	0.05	0.07	0.10
Panel A No Jump	3	1785	4.74	5.28	5.54	1784	5.25	5.01	5.79	1783	5.04	5.38	5.73
	4	1100	5.01	5.04	5.40	1098	5.05	5.00	5.78	1098	5.05	5.10	5.61
	5	743	4.62	4.85	5.29	743	4.77	5.04	5.51	743	4.60	5.31	5.82
	6	536	4.93	5.01	5.35	535	4.67	5.01	5.42	535	4.81	5.44	5.81
	7	404	4.83	5.08	5.22	404	4.81	5.04	5.48	403	5.24	5.58	5.67
	8	316	4.86	5.34	5.27	316	4.91	5.22	5.67	316	4.88	5.15	5.70
	9	254	4.77	5.42	5.24	253	5.13	5.41	5.39	253	4.83	5.01	5.72
	10	208	5.12	5.37	5.64	208	5.27	5.56	5.77	208	4.93	5.62	5.84
Panel B Moderate Jump	3	1716	46.47	48.75	50.70	1718	44.29	46.93	48.39	1718	42.42	45.03	46.94
	4	1058	45.56	48.39	50.84	1060	43.27	45.47	47.73	1061	41.57	43.20	45.28
	5	717	45.09	47.07	50.32	719	42.73	45.48	47.82	720	40.67	42.31	45.21
	6	519	44.78	46.46	48.26	519	41.76	44.18	46.61	521	40.50	42.15	44.40
	7	392	44.00	45.75	48.66	393	40.80	43.60	46.14	394	39.99	41.48	43.88
	8	307	42.56	44.26	48.15	308	40.08	42.59	45.58	308	39.08	41.55	43.25
	9	247	42.79	44.68	48.11	248	39.44	41.88	45.47	248	38.55	41.12	42.51
	10	203	41.21	44.19	46.88	204	39.69	41.62	44.87	204	37.62	39.87	42.92
Panel C Large Jump	3	1585	70.44	71.21	73.23	1589	68.58	70.69	72.01	1592	66.90	68.86	70.80
	4	978	69.91	71.37	73.07	981	68.25	69.64	71.07	985	66.59	68.75	69.66
	5	664	69.36	71.15	72.79	668	67.65	69.30	71.08	670	66.09	67.79	69.56
	6	481	68.75	70.43	72.67	484	66.82	68.64	70.50	487	65.24	66.70	68.72
	7	365	68.08	69.77	71.74	368	65.70	67.87	69.97	370	64.57	66.43	68.83
	8	287	67.67	68.95	70.93	289	65.10	68.12	69.69	291	63.77	66.09	67.83
	9	232	66.89	68.86	71.66	234	64.52	66.96	69.63	235	63.39	66.10	67.81
	10	191	66.21	68.25	70.90	192	64.24	66.28	69.49	194	62.52	64.96	67.43

This table reports the finite-sample size and size-adjusted power (%) of 10,000 simulations of the test statistic $T_{c,\epsilon}$ at 5% nominal level. All simulated prices are contaminated by the additive t-distributed noise and rounding errors. We utilize the two-step noise reduction method in Section 3.2 to construct the sequence of pseudo-observations with three different pre-averaging windows, i.e., $k_n = \lceil \theta\sqrt{N} \rceil$ with $\theta \in \{0.3, 0.4, 0.5\}$. The observations are sampled with different PDS barrier widths $c = K\sigma(\tilde{r}_i)$, where K ranges from 3 to 10. Different censoring thresholds with $\epsilon \in \{0.05, 0.07, 0.1\}$ are considered. $N^{(c)}$ stands for the average sampling frequencies.

Table B.12: Finite-sample size and power (%) of other tests under t-distributed noise

	Int. (sec)	N_{spl}	BNS	JO	LM	ASJ	CPR	PZ	MinRV	MedRV
Panel A No Jump	5	4680	58.90	15.85	100.00	99.92	99.87	99.70	0.12	0.02
	15	1560	12.76	10.10	89.87	92.05	56.23	72.27	0.71	2.19
	30	780	8.89	7.81	62.42	44.91	24.70	40.96	3.06	5.18
	60	390	7.25	7.38	38.66	16.35	11.91	24.07	4.49	7.09
	120	195	7.23	8.22	20.16	7.68	8.87	18.31	5.11	7.38
	180	130	7.37	8.99	16.35	5.38	8.42	17.48	5.30	8.46
	300	78	7.34	10.83	11.47	4.33	8.64	14.86	5.22	8.55
Panel B Moderate Jump	5	4680	37.49	37.05	17.64	99.97	11.25	13.71	39.54	41.12
	15	1560	40.61	38.20	25.74	97.85	21.65	20.95	34.45	37.18
	30	780	36.25	35.77	31.96	70.97	30.82	27.51	31.47	35.67
	60	390	30.11	29.45	36.41	40.48	30.72	35.36	27.12	31.05
	120	195	22.54	21.73	34.11	21.92	24.58	30.89	20.59	23.94
	180	130	17.55	17.67	29.39	14.43	20.16	25.77	16.98	20.66
	300	78	14.49	12.22	22.25	10.45	16.11	20.07	13.74	17.01
Panel C Large Jump	5	4680	62.75	63.82	45.20	99.98	31.59	38.22	63.94	65.60
	15	1560	65.71	64.24	53.45	98.79	47.84	47.73	59.14	61.94
	30	780	61.92	62.02	59.38	81.94	56.98	54.91	56.36	60.86
	60	390	55.55	55.88	63.00	57.74	56.36	62.26	52.38	56.60
	120	195	47.62	47.85	60.48	37.89	51.29	57.77	45.42	50.37
	180	130	41.65	43.02	55.53	25.61	45.02	52.02	40.67	45.06
	300	78	35.10	34.13	47.57	16.34	39.19	45.37	34.32	39.24

This table reports the finite-sample size and size-adjusted power (%) of 10,000 simulations of 8 classical tests at 5% nominal level: BNS (Barndorff-Nielsen and Shephard, 2006), JO (Jiang and Oomen, 2008), LM (Lee and Mykland, 2008), ASJ (Aït-Sahalia and Jacod, 2009), CPR (Corsi et al., 2010), PZ (Podolskij and Ziggel, 2010), MinRV and MedRV (Andersen et al., 2012). All simulated prices are contaminated by the additive t-distributed noise and rounding errors. All these tests are constructed on observations equidistantly sampled with various intervals in calendar time: 5, 15, 30, 60, 120, 180 and 300 seconds, and "N_{spl}" stands for the sampling frequencies.

Table B.13: Finite-sample size and power (%) of other noise-robust tests under t-distributed noise

Nominal size: 5%	Int. (sec)	N_{spl}	PZ*	LM12	ASJL
Panel A: No Jump	tick	23400	5.07	5.46	6.18
	5	4680	4.64	9.24	8.74
Panel B: Moderate Jump	tick	23400	39.40	25.31	37.68
	5	4680	29.26	18.76	17.19
Panel C: Large Jump	tick	23400	65.11	41.81	62.48
	5	7680	55.60	31.85	41.42

This table reports the finite-sample size and size-adjusted power (%) of 10,000 simulations of 3 noise-robust tests at 5% nominal level: noise-adjusted PZ (Podolskij and Ziggel, 2010), LM12 (Lee and Mykland, 2012), and ASJL (Aït-Sahalia et al., 2012). All simulated prices are contaminated by the additive t-distributed noise and rounding errors. All these tests are constructed on tick-level and 5-second-sampled observations. The tuning parameters for those tests are selected by minimizing the absolute distance between the nominal size and the empirical size with the simulated tick-level noise-contaminated observations.

Table B.14: Finite-sample size and power (%) of truncation-based detection under t-distributed noise

Nominal size: 5%					
		Panel A No Jump (with FWER control)		Panel B	Panel C
Ticks	N_{spl}	Šidák	Bonferroni	Moderate Jumps	Large Jumps
1	23400	100.00	100.00	9.95	29.93
5	4680	100.00	100.00	15.54	39.25
15	1560	97.71	97.63	23.69	50.02
30	780	63.20	62.93	29.69	56.28
60	390	26.47	26.23	36.40	61.59
120	195	11.41	11.26	38.29	62.58
180	130	8.19	7.99	32.68	58.53
300	78	5.73	5.48	26.31	52.13

This table reports the finite-sample size and size-adjusted power (%) of 10,000 simulations of the truncation-based jump filtering technique. All simulated prices are contaminated by the additive t-distributed noise and rounding errors. Observations are sampled at various multiples of ticks, where "N_{spl}" stands for the corresponding sampling frequencies. The truncation thresholds are constructed from the localized pre-averaged BV of Podolskij and Vetter (2009) computed within each backward-looking window of 1,800 ticks. The threshold parameter k is adjusted with both the Šidák and Bonferroni corrections, and $\varpi = 0.5$.

B.4 Supplementary Empirical Results

Table B.15 reports the empirical results for 8 other tests. Based on the simulation results in Tables 3 and 4, we select four calendar-time-sampling-based tests: BNS, CPR, MinRV and MedRV, with different sampling intervals: 30, 60, 120, and 300 seconds, and we also construct the noise-robust tests PZ*, LM12 and ASJL from tick-by-tick and 5-second data. Moreover, we consider the truncation-based filtering technique on calendar-time-sampled returns, with the truncation parameter k calibrated with the Šidák correction. For most of the selected stocks, the noise-robust ASJL constructed from tick-level observations obtains comparable results to our PDS-based test.

Table B.15: Empirical rejection rates (%) of other tests for selected NYSE stocks

Test	Int. (sec)	AXP	BA	DIS	IBM	JNJ	JPM	MRK	MCD	PG	WMT
BNS	30	32.02	20.55	20.95	28.46	36.36	32.81	49.80	25.69	43.08	32.41
	60	20.16	11.07	19.37	25.69	28.06	24.51	37.55	26.09	31.23	26.88
	120	17.00	16.21	16.21	22.53	27.67	25.30	25.69	23.32	27.27	27.67
	300	18.58	18.58	15.42	19.76	20.16	17.00	23.32	22.13	22.13	22.13
CPR	30	38.34	32.02	35.57	39.13	47.83	38.34	59.29	33.20	52.57	41.11
	60	28.46	16.60	26.88	33.99	40.32	32.41	45.85	29.25	40.32	35.57
	120	25.30	20.16	21.74	30.83	34.78	32.02	33.99	28.85	32.81	33.60
	300	23.32	23.72	21.34	27.67	29.64	22.13	32.81	28.46	30.43	30.83
MinRV	30	22.53	17.39	15.42	18.18	22.92	21.74	27.67	19.76	26.48	21.34
	60	14.23	9.88	15.42	19.76	21.74	16.21	22.92	20.95	22.13	22.13
	120	14.23	12.65	12.65	18.58	18.18	17.79	18.58	19.76	19.76	21.74
	300	13.04	14.62	13.83	17.39	13.04	11.07	15.81	15.42	16.21	14.23
MedRV	30	30.83	23.72	28.46	31.62	37.15	29.64	40.71	29.64	37.94	32.02
	60	20.55	15.81	22.92	28.46	37.15	27.67	33.60	28.85	32.81	30.04
	120	20.55	18.58	18.58	26.48	29.64	26.48	27.27	26.88	28.85	34.78
	300	18.97	17.00	16.60	24.11	21.74	20.55	23.72	24.90	28.06	25.30
PZ*	tick	7.51	6.32	6.32	5.93	6.32	6.72	7.51	5.14	7.51	4.74
	5	31.23	22.92	19.76	26.09	22.13	23.32	23.72	24.90	30.04	31.62
LM12	tick	12.65	4.35	7.51	9.49	12.65	11.46	12.65	9.49	18.58	11.86
	5	32.02	21.34	30.04	37.55	29.25	27.27	38.74	27.67	40.32	35.18
ASJL	tick	15.81	20.16	13.04	13.83	13.83	13.44	15.02	20.55	13.83	15.02
	5	32.02	20.95	29.25	21.34	26.48	21.74	22.92	26.09	30.04	32.41
Truncation-based Detection	30	69.96	67.98	70.75	63.64	77.87	64.43	80.63	74.31	80.63	75.10
	60	52.96	44.27	49.80	45.45	59.29	45.45	64.43	58.50	65.22	56.13
	120	32.41	36.36	33.99	32.81	40.32	34.78	43.87	36.36	42.69	36.36
	300	21.34	22.92	19.37	20.55	22.53	16.60	21.34	24.11	24.90	24.51

This table reports the proportions of days with jumps for 10 NYSE stocks in 2020, as identified by the following procedures: BNS (Barndorff-Nielsen and Shephard, 2006), CPR (Corsi et al., 2010), PZ (Podolskij and Ziggel, 2010), MinRV and MedRV (Andersen et al., 2012), PZ* (Podolskij and Ziggel, 2010), LM12 (Lee and Mykland, 2012), ASJL (Aït-Sahalia et al., 2012), and the truncation-based filtering technique in the spirit of Andersen et al. (2007) and Mancini (2009). The first four tests, together with the truncation-based filtering, employ observations sampled equidistantly in calendar time (with the last tick interpolation): 30, 60, 120 and 300 seconds. The noise-adjusted PZ*, LM12, and ASJL are constructed from tick-by-tick and 5-second-sampled data. The total number of trading days is 253.

Table B.16 reports the empirical results for other tests constructed from calendar-time-sampled data, with the control of spurious detections using the thresholding methods in Bajgrowicz et al. (2016): (i) the universal threshold $\sqrt{2 \ln 253}$, and (ii) the FDR threshold. We only consider one-sided tests whose limiting distribution is $\mathcal{N}(0, 1)$ under the null, which includes the upper-tailed BNS, CPR, MinRV, MedRV, PZ*, and the lower-tailed ASJL, but excludes the Gumbel-distributed LM12.

Table B.16: Adjusted empirical rejection rate (%) of other tests for selected NYSE stocks

	Test	Int. (sec)	AXP	BA	DIS	IBM	JNJ	JPM	MRK	MCD	PG	WMT
Panel A Universal threshold	BNS	30	24.11	18.18	17.79	23.32	27.67	25.69	33.99	20.95	30.83	22.92
		60	15.81	9.09	15.42	19.37	18.58	18.97	29.25	20.55	24.11	20.16
		120	13.44	13.83	15.02	20.55	22.92	22.13	21.74	17.79	19.37	20.55
		300	15.02	16.21	13.04	15.42	16.21	14.62	19.76	17.79	17.00	20.55
	CPR	30	27.67	27.27	29.64	27.67	31.23	26.88	36.76	23.32	33.99	26.48
		60	22.92	13.83	20.55	25.69	27.67	22.92	31.23	20.95	27.67	24.11
		120	21.34	17.00	19.76	24.90	24.11	27.27	25.69	23.32	23.72	24.51
		300	18.18	20.55	16.60	22.53	23.72	18.18	25.69	22.13	21.74	26.09
	MinRV	30	17.79	16.21	13.44	16.21	19.37	18.18	23.32	16.60	20.55	16.60
		60	11.86	8.30	14.23	15.81	17.79	14.62	20.55	18.97	20.16	18.97
		120	12.25	12.25	12.25	18.18	17.39	17.39	15.02	17.79	16.21	18.97
		300	12.25	13.44	13.44	16.21	12.25	11.07	15.42	13.44	15.02	13.83
	MedRV	30	24.51	21.74	26.09	25.30	26.09	21.74	30.04	22.92	28.46	26.09
		60	17.00	14.23	19.76	23.32	29.64	21.34	24.90	25.69	24.11	22.53
		120	17.39	16.21	17.39	21.34	24.11	24.11	22.53	22.13	22.13	29.64
		300	16.21	15.42	15.02	20.55	18.97	18.97	18.18	20.95	24.51	22.13
	PZ*	tick	5.53	5.53	6.32	3.56	3.56	5.93	5.93	3.95	5.14	3.16
		5	2.77	4.35	1.58	3.95	3.16	2.77	2.77	3.16	3.56	4.74
	ASJL	tick	14.23	18.58	13.04	13.04	13.44	12.25	12.25	19.37	13.44	14.62
		5	26.09	17.39	26.09	18.18	21.34	17.00	18.58	22.53	25.30	24.11
Panel B FDR threshold	BNS	30	14.23	15.02	16.21	14.23	12.25	15.02	9.09	14.62	11.46	11.07
		60	13.04	7.11	12.25	12.65	10.28	14.23	15.02	13.04	14.23	9.09
		120	9.88	13.44	14.62	14.62	13.83	17.39	12.25	15.02	13.83	10.28
		300	13.04	13.04	9.88	12.25	13.83	11.46	12.65	13.44	9.09	16.21
	CPR	30	9.09	15.42	19.37	13.04	11.86	9.09	7.51	12.25	8.30	10.28
		60	14.62	10.67	15.02	15.02	13.04	15.02	12.25	11.46	11.07	9.88
		120	13.83	16.21	16.21	13.04	14.62	16.21	13.04	13.44	13.04	10.67
		300	13.04	16.21	11.46	17.00	13.04	10.28	17.79	15.02	13.44	15.42
	MinRV	30	15.81	14.23	12.65	16.21	12.65	17.39	11.07	14.23	15.02	13.83
		60	10.67	7.51	11.07	12.65	12.65	11.86	13.44	18.58	13.44	14.62
		120	10.67	12.25	12.65	18.18	13.83	16.21	14.23	16.21	14.62	16.21
		300	12.65	11.86	13.44	15.81	12.25	11.07	13.04	13.44	15.02	13.83
	MedRV	30	15.02	19.37	19.76	10.28	16.60	16.60	10.67	16.21	17.39	13.44
		60	13.04	13.04	13.44	11.86	19.76	13.04	13.83	18.18	16.60	11.07
		120	15.42	15.02	13.83	15.42	18.58	21.74	16.21	18.97	12.25	18.58
		300	11.07	13.83	9.49	16.60	13.83	16.60	13.04	13.04	16.60	18.97
	PZ*	tick	5.53	5.53	6.32	3.56	3.56	5.93	5.93	3.95	5.14	3.16
		5	2.77	4.35	1.58	3.95	3.16	2.77	2.77	3.16	3.56	4.74
	ASJL	tick	13.04	16.21	13.04	10.28	13.44	12.25	11.07	17.39	13.44	14.62
		5	15.02	13.44	18.58	15.42	15.42	14.23	12.65	18.58	17.39	14.62

This table reports the proportions of days with jumps for 10 NYSE stocks in 2020, as identified by the following procedures: BNS (Barndorff-Nielsen and Shephard, 2006), CPR (Corsi et al., 2010), PZ (Podolskij and Ziggel, 2010), MinRV and MedRV (Andersen et al., 2012), PZ* (Podolskij and Ziggel, 2010), and ASJL (Aït-Sahalia et al., 2012), with the control of spurious detections using the thresholding methods in Bajgrowicz et al. (2016). The first 4 tests are constructed from observations equidistantly sampled in calendar time (with the last tick interpolation): 30, 60, 120 and 300 seconds. The noise-adjusted PZ* and ASJL are constructed from tick-by-tick and 5-second-sampled data. The total number of trading days is 253.

References

Aït-Sahalia, Y. and Jacod, J. (2009). Testing for jumps in a discretely observed process. *Annals of Statistics*, 37(1):184–222.

Aït-Sahalia, Y., Jacod, J., and Li, J. (2012). Testing for jumps in noisy high frequency data. *Journal of Econometrics*, 168(2):207–222.

Aït-Sahalia, Y., Li, C. X., and Li, C. (2025). So many jumps, so little news. Working Paper.

Aleti, S., Bollerslev, T., and Siggaard, M. (2025). Intraday market return predictability culled from the factor zoo. *Management Science*, forthcoming.

Andersen, T. G., Bollerslev, T., and Dobrev, D. (2007). No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and i.i.d. noise: Theory and testable distributional implications. *Journal of Econometrics*, 138(1):125–180.

Andersen, T. G., Dobrev, D., and Schaumburg, E. (2012). Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics*, 169(1):75–93.

Asmussen, S. (2003). *Applied Probability and Queues*. Springer.

Bajgrowicz, P., Scaillet, O., and Treccani, A. (2016). Jumps in high-frequency data: Spurious detections, dynamics, and news. *Management Science*, 62(8):2198–2217.

Barndorff-Nielsen, O. E. and Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, 4(1):1–30.

Barndorff-Nielsen, O. E. and Shiryaev, A. N. (2015). *Change of Time and Change of Measure*. World Scientific Publishing.

Borodin, A. N. and Salminen, P. (2002). *Handbook of Brownian Motion: Facts and Formulae*. Springer, Second edition.

Chernozhukov, V., Chetverikov, D., and Kato, K. (2013). Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *Annals of Statistics*, 41(6):2786–2819.

Chernozhukov, V., Chetverikov, D., and Kato, K. (2019). Inference on causal and structural parameters using many moment inequalities. *Review of Economic Studies*, 86(5):1867–1900.

Corsi, F., Pirino, D., and Renò, R. (2010). Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics*, 159(2):276–288.

Durrett, R. (2019). *Probability: Theory and Examples*. Cambridge University Press, Fifth edition.

Feller, W. (1991). *An Introduction to Probability Theory and Its Applications, Volume 2*. John Wiley & Sons; Second edition.

Gut, A. (2009). *Stopped Random Walks: Limit Theorems and Applications*. Springer, Second edition.

Ibragimov, I. A. (1962). Some limit theorems for stationary processes. *Theory of Probability & Its Applications*, 7(4):349–382.

Ibragimov, I. A. and Linnik, Y. V. (1971). *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff.

Jacod, J. (1997). On continuous conditional Gaussian martingales and stable convergence in law. *Séminaire de probabilités de Strasbourg*, 31:232–246.

Jacod, J., Li, Y., and Zheng, X. (2017). Statistical properties of microstructure noise. *Econometrica*, 85(4):1133–1174.

Jacod, J., Li, Y., and Zheng, X. (2019). Estimating the integrated volatility with tick observations. *Journal of Econometrics*, 208(1):80–100.

Jacod, J. and Protter, P. (2012). *Discretization of Processes*. Springer.

Jacod, J. and Shiryaev, A. (2003). *Limit Theorems for Stochastic Processes*. Springer, Second edition.

Jiang, G. J. and Oomen, R. C. (2008). Testing for jumps when asset prices are observed with noise – a "swap variance" approach. *Journal of Econometrics*, 144(2):352–370.

Lee, S. S. and Mykland, P. A. (2008). Jumps in financial markets: a new nonparametric test and jump dynamics. *Review of Financial Studies*, 21(6):2535–2563.

Lee, S. S. and Mykland, P. A. (2012). Jumps in equilibrium prices and market microstructure noise. *Journal of Econometrics*, 168(2):396–406.

Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics*, 36(2):270–296.

Mykland, P. A. and Zhang, L. (2009). Inference for continuous semimartingales observed at high frequency. *Econometrica*, 77(5):1403–1445.

Podolskij, M. and Vetter, M. (2009). Bipower-type estimation in a noisy diffusion setting. *Stochastic Processes and their Applications*, 119(9):2803–2831.

Podolskij, M. and Vetter, M. (2010). Understanding limit theorems for semimartingales: A short survey. *Statistica Neerlandica*, 64(3):329–351.

Podolskij, M. and Ziggel, D. (2010). New tests for jumps in semimartingale models. *Statistical Inference for Stochastic Processes*, 13(1):15–41.

Rényi, A. (1963). On stable sequences of events. *Sankhyā, Series A*, 25(3):293–302.

Shepp, L. A. (1967). A first passage problem for the Wiener process. *Annals of Mathematical Statistics*, 38(6):1912–1914.