THE UNIVERSITY OF
SYDNEY

Economics Working Paper Series

2020 – 09

Nonparametric Inference of Jump Autocorrelation

Simon Kwok

January 2021

Nonparametric Inference of Jump Autocorrelations

Simon Kwok

First Version: August 2020

This Version: January 2021

Abstract

Understanding the jump dynamics of market prices is important for asset pricing and risk management. Despite their analytical tractability, parametric models may impose unrealistic restrictions on the temporal dependence structure of jumps. Using nonparametric technique, we develop an omnibus test for the temporal dependence of jump occurrences in the DGP. It delivers consistent test result under a unified framework that combines and long-span and in-fill asymptotics, and is robust to the jump activity level and the choice of sampling scheme. We also study the use of sample autocorrelogram for visualizing and conducting pointwise inference of the autocorrelation of jump occurrences. We establish asymptotic normality and local power of the inference procedures for a rich set of local alternatives (e.g., self-exciting and/or self-inhibitory jumps). Simulation study confirms the robustness property of the omnibus test and reveals its competitive size and power performance relative to existing tests. In an empirical study on high-frequency stock returns, our procedure uncovers a wide array of autocorrelation profiles of jump occurrences for different stocks in different time periods.

Keywords: jump autocorrelation, self-excited jumps, nonparametric inference, financial contagion, high-frequency returns

JEL classification: C12, C14, C58

1 Introduction

Jumps and diffusion are widely regarded as the key sources of variation in asset returns. Similar to stochastic volatility in diffusion, jumps are found to exhibit a rich variety of dynamic properties that are associated with different market environments. Negative jumps tend to occur more often and in clusters during turbulent periods in financial markets. At the intraday level, the jump dynamics may display seasonality over the course of a trading day (Kashanan et al. (2018)). A deep understanding of the dynamic nature of jumps is key to a comprehensive characterization of asset price formation over time. In particular, learning how jumps are dependent over time is crucial for understanding the mechanism behind some market phenomena (e.g., propagation of financial contagion, limit order arrivals).

The parametric modelling of the temporal dependence of jumps has been around for a few decades. Hawkes (1971) proposes one of the first parametric models for self-exciting jumps. The Hawkes model defines a parametric specification on the conditional jump intensity which is assumed to depend on the timing of past occurrences of jumps. Due to its intuitive and flexible parametric form, variants of Hawkes models have been applied to empirical finance; see Bacry et al. (2015) and Hawkes (2018) for a survey. The popularity of the model lies in its ability to describe the jump-clustering phenomenon found in many financial time series, e.g., market returns during financial contagion and market downturn (Aït-Sahalia, Cacho-Dias and Laeven (2015)), jumps in prices and volatilities (Maneesoonthorn et al. (2017)), extremal returns (Embrechts, Liniger and Lin (2011), Chavez-Demoulin and McGill (2012)).¹

While a parametric model offers analytical tractability and efficiency, it imposes constraints on the jump dynamics; in particular, only a certain kind of temporal dependence structure is allowed for a given kernel function. For example, a stationary Hawkes model with an exponential kernel implies that jumps must be self-exciting and that the associated autocorrelogram must decay exponentially (Hawkes (1971)). The model-implied restrictions may rule out some empirically relevant jump dependence structure.² Without much clue on

¹In some applications, the focus is on modelling the arrival of a series of events (which may be viewed as jumps in a counting process). In market microstructure analysis, the autoregressive duration model (Engle and Russell (1998)) is useful for explaining the correlated timing of transactions, which usually arrive in clusters. To study the interaction between trade and quote arrivals, Bowsher (2007) proposes a sophisticated intensity model that features both intraday and across-day seasonality.

²Various kinds of jump dependence structure are documented in empirical quantitative finance. For example, limit order arrivals may display long-range persistence while obeying stationarity (Bacry et al. (2016)). At the intraday level, self-depressing jumps (as opposed to self-exciting jumps) tend to occur in the opening session of a trading day (Kashanan et al. (2018)).

the precise form of the jump dynamics, the modeller may simply pick a parametric specification for the kernel function out of convenience. This may induce model misspecification risk, which could potentially lead to bias in inference. One answer to the problem is to carry out diagnostic checks on an estimated parametric model in the post-modelling phase, e.g., by checking the model residuals for traces of serial dependence or left-over jumps (Kwok (2013)). A somewhat less arbitrary solution is to employ non-parametric techniques which are informative about the nature of jump dynamics in the underlying data generating process (DGP). The latter solution not only provides guidance for the search of parametric model specification, but also serves as a standalone procedure that directly yields inference results.

Nonparametric inference of jump dynamics has become a vibrant area of research due to increased availability of high-frequency data and recent advance in methodology. While the inference on the first-order properties of jumps is well studied³, there are recent papers dedicated to the study of nonparametric inference of the second-order properties of jumps, in particular their serial dependence structure. To test for self-excitation in jumps, Boswijk, Laeven and Yang (2018; henceforth BLY) proposes a test that captures the co-movements of realized log-price process and spot jump intensity process estimated from realized jumps. Adopting in-fill asymptotics and a pathwise approach, consistency of the test requires an infinite number of jumps over a finite time window (i.e., infinite-activity jumps). By contrast, Corradi, Silvapulle and Swanson (2018; henceforth CSS) tests the hypothesis of having self-excited finite-activity jumps in the population, and delivers consistent inference under long-span asymptotics.

Some practical issues arise from the use of existing tests. For example, pre-testing and pre-estimation are necessary – CSS test requires a pre-testing stage for the presence of jumps, while BLY test requires a precise estimate of the degree of jump activities (BLY). Another concern is on the robustness of test performance when the data are not fully compatible with the asymptotic theory and the underlying assumptions. For a given finite dataset collected over a certain sample horizon and at a certain sampling frequency, it can be difficult to determine whether it fits better to a particular asymptotic scheme. As revealed by our simulation experiments, the performance of existing tests can be unexpectedly sensitive to the choices of sample horizon and sampling frequency, sometimes yielding severely oversized or entirely powerless results.

Motivated by the problems discussed above, we propose a set of procedures that deliver

³The literature on the inference of first-order properties of high-frequency jumps is vast; e.g., see Aït-Sahalia and Jacod (2012) for a survey, and Aït-Sahalia and Jacod (2014) for a textbook treatment.

robust inference on the autocorrelation of jump occurrences for all possible degree of jump activity and for a wide range of empirically relevant sampling schemes. As our main contribution, we develop an omnibus test that jointly detects autocorrelation of jump occurrences over all possible lag orders. Under a unified framework that combines both in-fill and long-span asymptotics, the omnibus test is consistent under the null hypothesis of uncorrelated, stationary jumps with arbitrary form of stochastic volatility and deterministic trends. Furthermore, given the form of the test statistic (a weighted sum of squared autocovariances over all lag orders), the omnibus test exhibits local asymptotic power over a rich set of alternatives, e.g., DGP with self-exciting and/or self-inhibitory jumps, or jumps with cyclical autocorrelation profile. As shown in finite-sample experiments, the test has robust empirical sizes across a range of sampling schemes and for various degree of jump activity, while exhibiting competitive power for different types of autocorrelated jump processes. Furthermore, as a complement of the omnibus test, we examine the use of sample autocorrelogram as a visual tool to uncover the serial correlation structure of stationary jumps (e.g., persistence, seasonality). It also allows for convenient pointwise inference up to some fixed lag order. Our procedures do not require pretesting for the presence of jumps and knowledge of the jump activity level.

We confine ourselves to strictly stationary jump processes in this paper. This temporal stability condition greatly facilitates the development of the asymptotic theory. While this assumption remains restrictive in some applications, the methodology developed under strict stationary will be useful as the basis for extending the robust inference theory to the nonstationary setting.⁴ The stationarity condition can be less restrictive than it appears as it is imposed on the unconditional joint distribution of the jump increment process, while conditional moments (e.g., jump intensity conditional on the past) may be time-varying.

The rest of the paper is organized as follows. Section 2 sets out the DGP and introduces the autocovariance function as a measure of the temporal dependence of jumps. Section 3 introduces the nonparametric inference procedure and discusses the asymptotic theory. Section 4 reports the simulation results and the robustness analysis of the omnibus test. Section 5 presents the empirical application. Section 6 provides the concluding remarks. All proofs are given in the appendices.

⁴It is possible to extend the inference theory without the stationarity assumption. The time-varying mean intensity can be estimated using a nonparametric kernel estimator. The omnibus test can then be applied to the *normalized* residual jump process, which is stationarity under the null of zero autocorrelations.

2 Data Generating Process

Let X_t denote the log-price process defined in a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ which satisfies the usual conditions (Protter (2003)). The following assumption specifies the data generating process of X_t .

Assumption D: The log-price process X_t is a semi-martingale with decomposition

$$\begin{aligned} X_t &= X_0 + B_t + X_t^c + X_t^d \\ &= X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + [\delta 1_{\{\|\delta\| \leq 1\}} * (\mu^X - \nu^X)_t + \delta 1_{\{\|\delta\| > 1\}} * \mu_t^X], \end{aligned}$$

where $\delta = \delta(\omega, t, x)$ is a predictable function, μ^X is the jump measure given by $\mu^X(\omega, t, A) := \int_0^t \int_{x \in A} \varepsilon_{\{s, \Delta X_s(\omega)\}}(ds, dx)$ for $A \subseteq \mathbb{R} \setminus \{0\}$ ($\varepsilon_{\{a\}}$ denotes the Dirac measure at point a), ν^X is the compensator taking the form $\nu^X(\omega, dt, dx) = dt \otimes F_t(\omega, dx)$, and F_t is a locally bounded and predictable process. The pure jump process X_t^d has strictly stationary jump increments $\Delta X_t^d := X_t - X_{t-}$ and adopts one of the following two specifications:

- (1) X_t^d is a finite-activity jump process given by $X_t^d = \int_0^t J_s dN_s$, where N_t is a counting process associated with a stationary simple point process,⁵ and J_t is a strictly stationary process with distribution F_J and is independent of $\mathcal{F}_{t-}$.
- (2) X_t^d is a infinite-activity jump process with compensator $\nu^X(\omega, t, A) = \int_0^t \int_A \lambda_{s-}(\omega) F(dx) ds$, where $F(dx)$ is the Lévy measure given by $F(dx) = \frac{dx}{|x|^{1+\beta}}$ for $0 < \beta < 2$,⁶ and λ_t is a locally bounded process such that $E(\lambda_t) \equiv \bar{\lambda} < \infty$.

Remark 1 *Assumption D places restrictions on the jump process X_t^d . The key restrictions include the strict stationarity condition and the form of the jump measure and compensator. Assumption D encompasses a number of jump processes found in the literature. The DGP in D(1) includes the Hawkes process and the Poisson process as special cases. Examples are*

⁵For small $\Delta > 0$, the conditional law of N_t is characterized by $P(\Delta N_t = 0 | \mathcal{F}_{t-}) = 1 - \lambda_t \Delta + o(\Delta)$, $P(\Delta N_t = 1 | \mathcal{F}_{t-}) = \lambda_t \Delta + o(\Delta)$, and $P(\Delta N_t > 1 | \mathcal{F}_{t-}) = o(\Delta)$, where λ_t is the stochastic intensity defined by $\lambda_t := \frac{P(\Delta N_t \geq 1 | \mathcal{F}_{t-})}{\Delta}$. The unconditional intensity $\bar{\lambda} := E(\lambda_t)$ is a finite constant by stationarity.

⁶The parameter β controls the degree of jump activity and coincides with the Blumenthal-Gettoor index defined by $\beta := \inf\{p : \int_{|x| \leq 1} |x|^p F(dx) < \infty\}$. For the finite-activity jump process in Assumption D(1), β may be set to zero.

found in Aït-Sahalia et al. (2015) and CSS (2018). The DGP in $D(2)$ is based on that in Aït-Sahalia and Jacod (2009a,b), Jing et al. (2012), and BLY (2018).⁷ Autocorrelation of jumps may be accommodated through $\lambda_t(\omega)$, which is viewed as the stochastic intensity in $D(1)$, or as the stochastic scaling factor in $D(2)$.

Remark 2 The strict stationarity condition does not rule out the presence of time-varying conditional intensity, defined as $\lambda_t := \frac{P(|\Delta X_t^d| > 0 | \mathcal{F}_{t-})}{dt}$. For example, the Hawkes process, which assumes that $\lambda_t(\omega) = \lambda_\infty + \int_0^t \varphi(t-s) dN_s(\omega)$ for some kernel function $\varphi(\cdot)$, has stationary jump increments with unconditional intensity $\bar{\lambda} := E(\lambda_t) = \frac{\lambda_\infty}{1 - \int_0^\infty \varphi(u) du}$ if and only if $\int_0^\infty \varphi(u) du < 1$ (Hawkes(1971)).⁸

The strict stationarity condition on the jump increment process (Assumption D) implies that all moments of jump increments are time-invariant whenever they exist. In particular, the process $f(\Delta X_t^d)$ is covariance stationary for some bounded function $f(\cdot)$, i.e., the moments $E(f(\Delta X_t^d))$, $Var(f(\Delta X_t^d))$ and $Cov(f(|\Delta X_t^d|), f(|\Delta X_{t+\ell}^d|))$ exist (by the boundedness of $f(\cdot)$) and do not depend on t (by stationarity). An example of $f(\cdot)$ is the indicator function $f_\epsilon(x) = 1\{|x| > \epsilon\}$ for some $\epsilon > 0$. The quantity $\lambda_\epsilon := \lim_{\Delta \rightarrow 0} \frac{E(f_\epsilon(\Delta X_t^d))}{\Delta t}$ then represents the unconditional intensity of all finite- or infinite-activity jumps with size exceeding ϵ . It is a finite constant over time under Assumption D.

Our goal is to characterize how jump occurrences are temporally dependent. To this end, we introduce the *jump occurrence process*: for $t \geq 0$, $\Delta, \epsilon > 0$ and $\omega \in \Omega_J$,

$$I_t(\Delta, \epsilon; \omega) := 1\{|X_{t+\Delta}^d(\omega) - X_t^d(\omega)| > \epsilon\},$$

where Ω_J is the set of sample paths on which jumps occur ($\Omega_J := \{\omega : |\Delta X_t^d(\omega)| > 0 \text{ for some } t > 0\}$), and $1\{S\}$ is the indicator function which equals one if S occurs, or zero otherwise. Note that $I_t(\Delta, \epsilon; \omega)$ is a binary process which retains only the information about the timing of those jumps with magnitude exceeding ϵ in the time interval $(t, t + \Delta]$; information regarding the jump size is removed (except whether it exceeds ϵ or not).

⁷They consider a slightly more general jump process with a generic truncation function and an additional singular element in the Lévy measure.

⁸If the kernel function of the Hawkes process takes the form $\varphi(x) = \alpha e^{-\zeta x}$, then the Hawkes process is stationary iff $\frac{\alpha}{\zeta} < 1$.

Define the *population autocovariance function* of $I_t(\Delta, \epsilon)$ as follows:

$$\gamma(s; \Delta, \epsilon) := \text{Cov}_{\Omega_J} \{I_t(\Delta, \epsilon), I_{t+s}(\Delta, \epsilon)\}.$$

Conditional on the presence of stationary jumps, the autocovariance function fully captures the pairwise dependence of $I_t(\Delta, \epsilon)$ over time.⁹ In particular, it quantifies the degree of self-excitation or self-inhibition of jumps. E.g., for finite-activity jumps generated by a stationary Hawkes model with exponential kernel $\varphi(x) = \alpha e^{-\zeta x}$ ($0 < \alpha < \zeta$), the autocovariance function decays exponentially: for all small $\Delta, \epsilon > 0$, we have $\gamma(s; \Delta, \epsilon) = A e^{-(\zeta - \alpha)|s|} \Delta^2 + o(\Delta^2)$, where $A > 0$ depends on model parameters (Hawkes (1971)).¹⁰ The same characterization applies to stationary infinite-activity jump processes.

The population autocovariance function $\gamma(s; \Delta, \epsilon)$ is unknown in practice because the pure jump process X_t^d is unobserved. We need to have a way to classify a discretized change as a jump or diffusive increment. Suppose the time series of log-prices $\{X_0, X_1, \dots, X_n\}$ is sampled over the horizon $[0, T]$. Let $\Delta_i^n X := X_i - X_{i-1}$ denote the discretized sequence of log-returns computed over sampling intervals of identical width Δ , so that $\Delta = \frac{T}{n}$. Introduce the bounded functions $g_k(\frac{x}{u}) := \min(1, |\frac{x}{u}|^k)$ for $k > 0$, and $g_\infty(\frac{x}{u}) := 1\{|x| > u\}$. With an appropriately chosen cut-off sequence u_n , we construct the *filtered sequence* $g_k(\frac{\Delta_i^n X}{u_n})$ such that it preserves the increments of X_t^d in the limit (e.g., Aït-Sahalia and Jacod (2009a,b), Jing et al. (2012), and BLY (2018)). If we set $k = \infty$, the filtered sequence records as one all those discretized changes with magnitude exceeding u_n ; and if we set $k > 0$, the sequence is the same except that it additionally records the smaller discretized changes as some value smaller than one. Setting a larger k effectively imposes a heavier discount on smaller increments, thus leaving us with less variation for analyzing smaller jumps, but at the same time lowering the chance of misclassifying a diffusive increment as a jump occurrence. The cut-off sequence u_n needs to be chosen judiciously such that the increments of the diffusive process X_t^c are removed in the limit as $n \rightarrow \infty$. The sequence u_n may or may not vanish to zero depending on the inference method and the sampling scheme. This will be made precise in Section 3 (Assumptions T_L and T_γ).

⁹Indeed, for $s \neq t$, the joint distribution of $I_s(\Delta, \epsilon)$ and $I_t(\Delta, \epsilon)$ is fully characterized by their first two moments (mean, variance, and covariance). This follows from the binary nature of the jump occurrence process.

¹⁰For another example (which violates Assumption D), consider a jump process constructed from a Poisson process with rate $\bar{\lambda}$, except that each Poisson jump is followed by an additional jump one unit of time later. Then, for small $\Delta > 0$, the autocovariance function is given by $\gamma(s) = \bar{\lambda}\Delta + o_p(\Delta)$ for $s = \pm 1$, and zero otherwise.

The following assumption restricts the tuning parameters of the filtered sequence $g_k(\frac{\Delta_i^n X}{u_n})$, including the form of the cutoff sequence u_n and the power k .

Assumption G: The filtered sequence $g_k(\frac{\Delta_i^n X}{u_n})$ is chosen with $k > \frac{1-\beta\varpi}{1-2\varpi}$ and $u_n = c\Delta^\varpi$, where $c > 0$ and $0 < \varpi < \frac{1}{2}$.

Remark 3 Without a priori knowledge of the jump activity level β , the condition on k may be tightened to $k > \frac{1}{1-2\varpi}$, which satisfies Assumption G for all possible values of $\beta \in [0, 2)$.

3 Inference of Autocorrelations of Jump Occurrences

In this section, we develop a joint test for autocorrelations of jump occurrences under weak assumptions on the sampling scheme and jump processes. The test statistic makes use of the filtered sequence as inputs. The test delivers valid size even if no jumps are present in the sample series.¹¹

3.1 Sample Autocorrelogram of Jump Occurrences

We first introduce the sample autocovariance measure for jump occurrences. For notational convenience, we let $g_{n,k,i} := g_k(\frac{\Delta_i^n X}{u_n})$ and $\bar{g}_{n,k} := \frac{1}{n} \sum_{i=1}^n g_{n,k,i}$. Define the *sample autocovariance function* of $g_{n,k,i}$:

$$\hat{\gamma}_{n,k}(\ell) := \frac{1}{n} \sum_{i=1}^{n-\ell} [g_{n,k,i} - \bar{g}_{n,k}][g_{n,k,i+\ell} - \bar{g}_{n,k}]. \quad (1)$$

Our goal is to construct a joint test that is consistent and powerful against autocorrelated jump occurrences over *all* lag orders. The null hypothesis is given by Assumption D₀, stated formally below.

Assumption D₀: Assumption D holds. In addition, the null hypothesis holds:

$$H_0 : \gamma(s; \Delta, \epsilon) = 0 \text{ for all } s \neq 0 \text{ and all } \Delta, \epsilon > 0.$$

¹¹Both methods would conclude that jump occurrences have zero autocorrelations if all time periods contain big jumps (i.e., the filtered jump series $g_{n,k,i}$ is identically one). This problem is likely caused by having an excessively small cutoff u_n and is readily verifiable by the user

The alternative hypothesis is its negation, i.e.,

$$H_a : \gamma(s; \Delta, \epsilon) \neq 0 \text{ for some } s \neq 0 \text{ and some } \Delta, \epsilon > 0.$$

Define the test statistic

$$L(k) := n \sum_{\ell=1}^{n-1} w_h(\ell) \hat{\gamma}_{n,k}^2(\ell),$$

where $w_h(\ell)$ specifies the weights on the squared autocovariances. For notational simplicity, we suppress the dependence of the test statistic on n and h .

Intuitively, the $L(k)$ statistic summarizes all information regarding the autocorrelation of jump occurrences by means of a weighted sum of squared autocovariances of the filtered jump series $g_{k,i}$. The weight decreases with the lag order, reflecting the economic intuition that the association of jumps close to each other is more important than those further apart. The reason for squaring the autocovariance function is to avoid possible cancellation between the positive and negative autocovariances due to self-excited and self-inhibitory jumps. Note that $L(k)$ test reduces to the portmanteau test of Box and Pierce (1970) if the weights are all equal and the sum is truncated up to a finite lag order. As such, $L(k)$ test may be viewed as a generalized Box-Pierce test (Hong (1999)). By allowing for a weighted sum that spans infinite number of terms in the limit, the test is consistent over all possible lag orders.

For valid large-sample inference, the weight function $w(\cdot)$, the tuning parameter h (which controls the decay rate of the weights on squared autocovariances) and the sampling intervals $\Delta = \frac{T}{n}$ need to obey certain restrictions as listed in the following assumption. The assumption accommodates both long-span ($0 < v < 1$, which implies expanding T and shrinking Δ) and fixed-span ($v = 1$, which implies fixed T and shrinking Δ) asymptotics, depending on the jump activity level.

Assumption $\mathbf{T}_L$: (i) $w_h(\ell) = \frac{1}{h} w\left(\frac{\ell}{h}\right)$, where $w(x)$ is defined on $\mathbb{R}_+$ and satisfies $\int_0^\infty w(x)dx = 1$ and $\mathcal{W}_2 = \int_0^\infty w^2(x)dx < \infty$; (ii) for given β and ϖ , there exists $(\nu, \xi) \in \{(\nu, \xi) : 0 < \nu \leq 1, \max[0, 2\nu(1 - \beta\varpi) - 1] < \xi < 1\}$ such that $\Delta = O(n^{-\nu})$ and $h = O(n^\xi)$ as $n \rightarrow \infty$.

Remark 4 Assumption $T_L(i)$ defines the set of valid weight functions that are typical in nonparametric kernel estimation. Assumption $T_L(ii)$ specifies the allowable set of sampling schemes (controlled by ν) and the penalty placed on autocovariances with large lag orders (controlled by ξ). The allowable set of (ν, ξ) depends on the jump activity level β and the

user-specified cut-off parameter ϖ . This is illustrated in Figure 1. This restriction is to ensure the consistency of $L(k)$ test. As an example, suppose that the jumps are from a finite-activity process ($\beta = 0$). Then, Assumption $T_L(ii)$ implies that $\nu < 1$ (i.e., T grows with n); see Figure 1(a). Fixed-span asymptotics (with $\nu = 1$ and T fixed) is prohibited because it would be impossible to capture enough finite-activity jumps over rapidly shrinking intervals and yet achieve consistent estimation of autocorrelations.¹² As another example, suppose that jumps are driven by an infinite-activity process with $\beta = 1$, that ϖ is set to 0.3, and that the sample horizon is fixed ($\nu = 1$). Then the admissible range of ξ is $0.4 < \xi < 1$ (Figure 1(b)). To ensure test consistency, the general rule is that more lags need to be included (i.e., larger ξ , and hence larger h) when the jump activity level is lower (smaller β), the trimming of small jumps is more aggressive (smaller ϖ), or when the sampling intervals shrink more quickly in width (bigger ν).

The following theorem establishes asymptotic normality of the studentized $L(k)$ statistic under the null hypothesis (the subscript 0 indicates that the mean and variance are evaluated under H_0).

Theorem 1 Suppose **Assumptions D_0 , G and T_L** hold. Then for all $r > 0$ and $p > 1$, we have

$$t_{L(k)} := \frac{L(k) - \hat{E}_0[L(k)]}{\sqrt{\widehat{Var}_0[L(k)]}} \xrightarrow{d} N(0, 1),$$

as $n \rightarrow \infty$, where

$$\begin{aligned} \hat{E}_0[L(k)] &= [\widehat{Var}_0(g_{n,k,i})]^2 \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) w_h(\ell), \\ \widehat{Var}_0[L(k)] &= \frac{1}{n} [\widehat{Var}_0(g_{n,2k,i}) + r\Delta^p]^2 \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) w_h^2(\ell) \\ &\quad + 2[\widehat{Var}_0(g_{n,k,i}) + r\Delta^p]^4 \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) \left(1 - \frac{\ell+1}{n}\right) w_h^2(\ell), \\ \widehat{Var}_0(g_{n,k,i}) &= \frac{1}{n-1} \sum_{i=1}^n (g_{n,k,i} - \bar{g}_{n,k})^2, \quad \bar{g}_{n,k} = \frac{1}{n} \sum_{i=1}^n g_{n,k,i}. \end{aligned}$$

¹²Recall that a simple point process has at most a finite number of jumps in a finite time interval.

Remark 5 *The BLY test detects self-excitation of jumps by taking a pathwise approach; hence, the test result is specific to the observed sample paths. By contrast, $L(k)$ test is an inference with respect to the DGP, and strict stationarity condition on the jump increment process (Assumption D) is imposed; in this sense $L(k)$ test is similar to the CSS test. Another key distinction lies in the way temporal dependence of jumps is measured: BLY (2018) studies the co-jumps of the observed log-return series and the estimated spot intensity process, whereas $L(k)$ test directly captures the lead-lag relationships of jump occurrences using the autocovariance measure. The two test statistics summarize the information regarding jump dependence in different ways. Furthermore, compared to BLY test, the cost borne by $L(k)$ test for requiring strict stationarity on jump increments is offset by the advantage of a less restrictive condition on the power parameter k (Assumption G).¹³ We expect that the weaker condition on k induces more variation due to smaller jumps and thus enhances the power of $L(k)$ test. This will be seen in the simulation study in Section 4.*

Remark 6 *The variance estimator of $L(k)$ is composed of two terms: the first term dominates the second one under in-fill asymptotics, whereas the second term dominates the first one under long-span asymptotics. The presence of both terms guarantees robustness of the test to both types of asymptotic schemes. The inclusion of $r\Delta^p$ ensures that the $L(k)$ test is implementable when jumps are absent (as otherwise the denominator of $t_{L(k)}$ would be zero), thereby avoiding the need of a pretest for jumps. It does not affect the asymptotic properties of the test.¹⁴ In our simulation and empirical studies, we will set $r = 1$ and $p = 2$.*

We now study the power of $L(k)$ test by considering local alternatives in which jumps are autocorrelated. The local alternatives and the necessary higher moment restrictions are given in Assumption D_a .

Assumption D_a : Assumption D holds. In addition, there exists a constant $K > 0$ such that the following conditions are satisfied for all n .

¹³For infinite-activity jumps, the BLY test requires that the power be greater than $\frac{1-\varpi\beta}{1/2-\varpi}$ (BLY (2018), Theorem 2), which is twice as high as the minimal k allowed for $L(k)$ test (Assumption G). This is due to the difference in the stochastic orders of $\tilde{g}_{n,k,i}$ (which enters $L(k)$) and $g_{n,k,i}$ (which enters the BLY test statistic). Of course, the demeaned form $\tilde{g}_{n,k,i}$ is meaningful to compute by Assumption D, under which the mean $\bar{g}_{n,k}$ is well defined.

¹⁴For any $r > 0$ and $p > 1$, the extra term $r\Delta^p$ is dominated by $Var_0(g_{n,k,i})$, which is $O(\Delta^{1-\beta\varpi})$ for all k satisfying Assumption G (by lemma 4).

- (i) $\gamma(\ell\Delta) = f(\frac{\ell}{h})\frac{\Delta^{(1-\beta\varpi)/2}}{n^{3/4}h^{\delta/2}}$ for all $0 < \nu \leq 1$, where $0 \leq \delta \leq \frac{1}{2}$ and $f(\cdot)$ satisfies $\int_0^\infty w(x)f^2(x)dx > 0$.
- (ii) $\sum \sum \sum_{j_1, j_2, j_3=1}^n |cum(\tilde{g}_i, \tilde{g}_{j_1}, \tilde{g}_{j_2}, \tilde{g}_{j_3})| \leq K\Delta^{4(1-\beta\varpi)}$ for all $i = 1, \dots, n$.
- (iii) $\sum_{\ell \neq 0} w^2(\ell)Var(\tilde{g}_i^2 \tilde{g}_{i+\ell}^2) \leq K\Delta^{2(1-\beta\varpi)}$ for all $i = 1, \dots, n$, and $\sum_{\ell \neq 0} w^2(\ell)Var(\tilde{g}_i \tilde{g}_{i+\ell} \tilde{g}_j \tilde{g}_{j+\ell}) \leq K\Delta^{4(1-\beta\varpi)}$ for all $i, j = 1, \dots, n$.
- (iv) $\sum \sum_{(i_1, j_1) \neq (i_2, j_2)} |Cov(\tilde{g}_{i_1}^2 \tilde{g}_{j_1}^2, \tilde{g}_{i_2}^2 \tilde{g}_{j_2}^2)| \leq K\Delta^{2(1-\beta\varpi)}$ for all $i_1, j_1 = 1, \dots, n$, and $\sum \sum \sum \sum_{(i_1, j_1, p_1, q_1) \neq (i_2, j_2, p_2, q_2)} |Cov(\tilde{g}_{i_1} \tilde{g}_{j_1} \tilde{g}_{p_1} \tilde{g}_{q_1}, \tilde{g}_{i_2} \tilde{g}_{j_2} \tilde{g}_{p_2} \tilde{g}_{q_2})| \leq K\Delta^{4(1-\beta\varpi)}$ for all $i_1, j_1, p_1, q_1 = 1, \dots, n$.

Remark 7 Condition (i) defines precisely the set of alternatives against which $L(k)$ test exhibits power. We expect that $L(k)$ test is powerful for alternatives in which $\gamma(\ell\Delta)$ alternates in signs as lag order varies; e.g., periodic functions such as $f(x) = \sin(\frac{2\pi x}{\tau})$ clearly satisfy condition (i). For fixed lag order ℓ , as the sample size n increases $\gamma(\ell\Delta)$ diminishes to zero at a certain rate that is determined by how quickly the variance of $L(k)$ vanishes to zero. This diminishing rate of $\gamma(\ell\Delta)$ determines the power of $L(k)$ test. The decay rate of the weighting function (controlled by ξ) plays a key role in affecting the test power.¹⁵ Condition (ii) is similar to Assumption $D_a(\ell)$ (iii) but somewhat stronger as it places restrictions on the dependence of all quadruples of $\tilde{g}_i$. Conditions (iii) and (iv) control the serial dependence of $\tilde{g}_i$ up to the eighth order.

The asymptotic power result of $L(k)$ test is stated in the following theorem.

Theorem 2 Suppose **Assumptions D_a , G and T_L** hold. Then the following statements are true.

- (a) For $0 \leq \delta < \frac{1}{2}$,

$$P[|t_{L(k)}| > c] \longrightarrow 1$$

¹⁵Comparing the local alternatives in Assumption $D_a(i)$ with Assumption $D_a(\ell)(i)$, we see that $\gamma(\ell\Delta)$ diminishes at a faster rate than that of the pointwise $\hat{\rho}_{n,k}(\ell)$ test. Given the nature of $L(k)$ test which aggregates more information of jump autocorrelation over multiple lags, we expect that $L(k)$ test is powerful in more directions than $\hat{\rho}_{n,k}(\ell)$ test.

as $n \rightarrow \infty$, where c is any positive constant.

(b) For $\delta = \frac{1}{2}$,

$$t_{L(k)} = \theta_k + X + o_p(1),$$

as $n \rightarrow \infty$, where X is a zero-mean $O_p(1)$ random variable,

$$\theta_k := \begin{cases} \frac{\tilde{\kappa}}{\sqrt{2W_2}\sigma_k^4} & \text{for } \nu = 0, \\ \frac{\tilde{\kappa}}{\sqrt{W_2}\sigma_{2k}^2} & \text{for } 0 < \nu \leq 1, \end{cases}$$

and $\sigma_k^2 := \lim_{n \rightarrow \infty} \frac{\text{Var}_0(g_{n,k,i})}{\Delta^{1-\beta\varpi}}$.

3.2 Sample Autocorrelogram of Jump Occurrences

As a by-product of our theory, the sample autocovariance function (1) readily lends itself to visual comparison after suitable normalization. We introduce the *lag- ℓ sample autocorrelation* function:

$$\hat{\rho}_{n,k}(\ell) := \frac{\hat{\gamma}_{n,k}(\ell)}{\hat{\gamma}_{n,k}(0) + r\Delta^p},$$

where $r > 0$ and $p > 1$. The term $r\Delta^p$ plays a similar role as that in the variance of $L(k)$ (see Remark 7): it ensures that $\hat{\rho}_{n,k}(\ell)$ remains well-defined when the sample process exhibits no jumps (in which case $\hat{\gamma}_{n,k}(0)$ and $\hat{\gamma}_{n,k}(\ell)$ are both zero but $r\Delta^p$ is strictly positive), while its presence does not asymptotically alter the interpretation of $\hat{\rho}_{n,k}(\ell)$ as the sample autocorrelation of jump occurrences (as $r\Delta^p$ is strictly dominated by $\hat{\gamma}_{n,k}(0)$ asymptotically); consequently, the value of $\hat{\rho}_{n,k}(\ell)$ lies in $(-1, 1)$.

The following assumption imposes mild restrictions on the sampling scheme.

Assumption T_γ : $\Delta = O(n^{-\nu})$ as $n \rightarrow \infty$, where $\nu \in (0, \frac{1}{1-\beta\varpi}) \cap (0, 1]$.

The restriction, which is much milder than Assumption T_L , ensures consistency of the sample autocorrelogram for fixed lag order. In particular, Assumption T_γ prohibits the sampling intervals from shrinking too fast (at a rate of n^{-1}) when jumps are of finite activities ($\beta = 0$); see Remark 4.

We now establish the asymptotic normality of $\hat{\rho}_{n,k}(\ell)$ under the null hypothesis of uncorrelated jump occurrences.

Theorem 3 Suppose *Assumptions D₀, G and T_γ* hold. Then, for any fixed $\ell > 0$, we have

$$\sqrt{n}\hat{\rho}_{n,k}(\ell) \xrightarrow{d} N(0, 1)$$

as $n \rightarrow \infty$.

The above result delivers a convenient large-sample test for the autocorrelations of jump occurrences over a finite range of lags $\ell = 1, \dots, M$, where M is some finite maximum lag order prespecified by the user. The test can be visually implemented by interlacing the sample autocorrelogram plot (i.e., the plot of $\hat{\rho}_{n,k}(\ell)$ against $\ell = 1, \dots, M$) with a pair of parallel lines at levels $\mp z_{\alpha/2}/\sqrt{n}$ representing the critical band ($z_{\alpha/2}$ denotes the standard normal variate associated with a tail probability of $\alpha/2$). Note that the critical values do not depend on the lag order ℓ . We conclude that the lag- ℓ autocorrelation is significant at the level α if $\hat{\rho}_{n,k}(\ell)$ goes beyond the critical band.

Next we examine the power property of the autocorrelation test based on $\hat{\rho}_{n,k}(\ell)$ under a suitable sequence of alternatives. To this end, we impose more explicit structure on the temporal dependence of jump occurrences. The alternatives and restrictions on the temporal dependence structure are given in the following assumption. To simplify notation, we denote $\tilde{g}_i = \tilde{g}_{n,k,i} := g_{n,k,i} - E(g_{n,k,i})$.

Assumption D_a(ℓ): Assumption D holds. In addition, there exists a constant $K > 0$ such that the following conditions are satisfied for all n :

- (i) $\gamma(\ell\Delta) = \frac{K_\ell}{n^\delta}$, where $0 \leq \delta \leq \frac{1}{2}$ and $K_\ell \neq 0$ with $\sum_{\ell>0} |K_\ell| \leq K$.
- (ii) $\sum_{j=1, j \neq i}^n |Cov(\tilde{g}_i^2, \tilde{g}_j^2)| \leq K\Delta^{2(1-\beta\varpi)}$ for all $i = 1, \dots, n$.
- (iii) $\sum_{j=1}^n |Cov(\tilde{g}_i\tilde{g}_{i+\ell}, \tilde{g}_j\tilde{g}_{j+\ell})| \leq K\Delta^{2(1-\beta\varpi)}$ for all $i = 1, \dots, n$.

Remark 8 Condition (i) sets out a sequence of DGP's in which jump occurrences are autocorrelated at lag order ℓ (fixed). The autocovariance function $\gamma(\ell\Delta)$ may remain constant ($\delta = 0$) or diminish to zero ($\delta > 0$) as the sample size n increases, and is assumed to satisfy absolute summability over lags (as implied by $\sum_{\ell>0} |K_\ell| \leq K$). The condition encompasses finite-activity jump processes with finite covariance density (Bartlett (1963)). In particular,

if $K_\ell = Ae^{-(\zeta-\alpha)|\ell|\Delta}\Delta^2$ where $0 < \alpha < \zeta$ and $A > 0$ (depending on α and ζ), the DGP represents a stationary Hawkes process with exponential kernel $\varphi(x) = \alpha e^{-\zeta x}$ (Hawkes (1971)). The condition also accommodates infinite-activity jump processes with autocorrelated jumps. Conditions (ii) and (iii) impose restrictions on the serial dependence of $\tilde{g}_i$ up to the fourth order.

We expect that the autocorrelogram test based on $\hat{\rho}_{n,k}(\ell)$ for fixed lag order $\ell \neq 0$ is powerful against autocorrelated jump occurrences that are separated apart by $\ell\Delta$. This is formally stated in the following theorem; in particular, the result of the local power analysis is reported in part (b).

Theorem 4 Fix $\ell > 0$, and suppose **Assumptions** $D_a(\ell)$, **G** and T_γ hold. Then the following statements are true.

(a) For $0 \leq \delta < \frac{1}{2}$,

$$P[|\sqrt{n}\hat{\rho}_{n,k}(\ell)| > c] \longrightarrow 1$$

as $n \rightarrow \infty$, where c is any positive constant.

(b) For $\delta = \frac{1}{2}$,

$$\sqrt{n}\hat{\rho}_{n,k}(\ell) = \pi_{k,\ell} + U + o_p(1)$$

as $n \rightarrow \infty$, where U is a zero-mean $O_p(1)$ random variable, $\pi_{k,\ell} := \frac{K_\ell}{\sigma_k^2}$, and $\sigma_k^2 := \lim_{n \rightarrow \infty} \frac{\text{Var}_0(g_{n,k,i})}{\Delta^{1-\beta\varpi}}$.

4 Finite-Sample Performance

In this section, we study the finite-sample performance of $L(k)$ test introduced in the previous section.

The sample paths of the log-price process are generated from a stochastic-volatility-with-jumps (SVJ) model given by $dX_t = bdt + \sigma_t dW_t + dX_t^d$. The stochastic volatility process (common to both DGP 1 and 2, given below) is given by $d\sigma_t^2 = \kappa(\theta - \sigma_t^2)dt + \sigma_V \sigma_t dW_t^V + J_t^V dN_t^V$ with long-run variance θ , speed of mean reversion κ , and volatility of volatilities σ_V . The increments of the two standard Brownian motions W_t and W_t^V have a contemporaneous correlation of ρ (i.e., $E(dW_t dW_t^V) = \rho dt$). Jumps in volatilities are modelled by $J_t^V dN_t^V$, where J_t^V is iid exponential with mean μ_{J_v} , and N_t^V is an independent Poisson process with intensity λ_v . The pure jump component X_t^d takes either one of the following baseline specifications:

- DGP 1: X_t^d is an infinite-activity jump process as defined in Assumption D(1), with $0 < \beta < 2$ and the stochastic scaling factor λ_t given by $\lambda_t = \lambda_\infty + \int_0^t \varphi(t-s) dN_s$, where $\varphi(x)$ is some kernel function and $N_t = \sum_{s \leq t} 1_{\{|\Delta X_s| > \epsilon\}}$ with ϵ set to $\epsilon = 5\sqrt{\theta}\Delta^{1/3}$.
- DGP 2: X_t^d is a finite-activity jump process as defined in Assumption D(2), where J_t is *iid* $N(\mu_J, \sigma_J^2)$ and the stochastic intensity λ_t of N_t is modelled by $\lambda_t(\omega) = \lambda_\infty + \int_0^t \varphi(t-s) dN_s(\omega)$ for some kernel function $\varphi(x)$.

The parameter values of the baseline specifications are set as follows.¹⁶ For both DGP 1 and 2: $\kappa = 4.4$, $\theta = 0.014$, $\sigma_v = 0.24$, $\rho = 0.61$, $\mu_{J_v} = 0.018$, $\lambda_v = 1.39$. The kernel function is set to $\varphi(x) = \alpha e^{-\zeta x} 1_{\{x \geq 0\}}$ with $0 \leq \alpha < \zeta = 100$ (except in Figure 4b where $\varphi(x) = 1_{\{0 \leq x \leq \phi\tau\}} \cos(2\pi x/\tau)$ with $\phi \geq 0$ and $\tau = 60$). For DGP 1 only: $\beta = 1.25$, $a > 0$, $\bar{\lambda} = 1$. For DGP 2 only: $\mu_J = -0.03$, $\sigma_J = 0.05$, $\bar{\lambda} = 6.3$.¹⁷ There are a few exceptions: we set $\bar{\lambda} = 100$ in Table 1, Panel B (to ensure enough number of jumps); we set $\mu_J = -0.01$ in Figure 5b (to examine test power under small mean jump sizes); and we modify the SV and jump parameters in Figure 6 (to vary the proportion of variations due to jumps).

The sample horizon $[0, T]$ and the discretization step Δ are set as follows. For DGP 1: $T = 1$ and $\Delta = \frac{1}{20 \times 6.5 \times 60} = \frac{1}{7800}$; for DGP 2: $T = 30 \times 252 = 7560$ and $\Delta = \frac{1}{252}$. An exception is the size experiment in Table 1 and the $L(k)$ distribution plot in Figure 2, in which T and Δ vary. From the log-return series $\Delta_i^n X$, the filtered sequence is obtained as $g_{k,i} := g_k(\frac{\Delta_i^n X}{u_n})$ with $k = 3$ (except in Figure A3 where k may vary). The jump cutoff sequence takes the form $u_n = c\Delta^\varpi$ with $c = 5(\widehat{IV})^{0.5}$ (where $\widehat{IV}$ is the annualized threshold IV estimator of Mancini (2009)) and $\varpi = 0.3$ (except in Figure A4 where c and ϖ may vary). For $L(k)$ test, we adopt the Gaussian kernel $w_h(x) = \frac{1}{\sqrt{2\pi}h} \exp(-\frac{x^2}{2h^2})$, with $h = s.e.(g_{k,i}) \times n^\xi$ with $\xi = 0.7$ (except in Figure A5 where ξ may vary). In all simulation experiments, the number of replications is set to 5,000 and the nominal level is fixed at five percent.

¹⁶The parameters are set in accordance to the empirical estimates obtained in the literature. The parameters of λ_t with exponential kernel are set according to Aït-Sahalia et al. (2015, Table 2, first column). For the parameters of σ_t and J_t (DGP 2 only), which we cannot find in Aït-Sahalia et al. (2015), we refer to the annualized SVSCJ model parameter estimates in Eraker (2004, Table III, last column).

¹⁷Provided that $\alpha < \zeta$, the λ_t process is stationary with unconditional (annualized) jump intensity $\bar{\lambda} = \lambda_\infty / (1 - \frac{\alpha}{\zeta})$.

4.1 Size Performance

Table 1 reports the size performance of $L(k)$ test (with $k = 3$ or 6), which is compared to the competing BLY and CSS tests.¹⁸ Panels A and B present the empirical sizes when jumps are of infinite activity (DGP 1) with Δ fixed and T fixed, respectively. Panels C and D show the corresponding results under the finite-activity jump case (DGP 2). The fixed- Δ scenarios are incompatible with the asymptotic theory of $L(k)$ test – Assumptions T_L and T_γ (which require $v > 0$) are marginally violated. Nonetheless, we include them for practical relevance.

Among the three tests, only $L(k)$ test maintains the most stable sizes across the scenarios. The comparison in Panel B (infinite-activity jumps, fixed T) is not valid for CSS test, and that in Panel C (finite-activity jumps, fixed Δ) is not valid for BLY test (and marginally for $L(k)$ test). This is because the BLY test (resp., CSS test) is based on an in-fill (resp., large T) asymptotics and is applicable to DGP with infinite jumps (resp., finite jumps) only. No inference technique could deliver consistent test results in Panel D (T fixed and Δ shrinking), because consistent estimation of autocorrelations is impossible (see Remark 4). All comparisons that are incompatible with the asymptotic theory of the tests are shaded in gray (and in light gray for borderline incompatibility) in Table 1.

The empirical sizes of $L(k)$ test stay close to the nominal level across all valid settings (Panels A, B and C), although the test becomes undersized in small samples when DGP contains only finite-activity jumps (Panel C). The CSS test has very good sizes (Panel C). The BLY test suffers from a mild oversize problem (Panel B) – the empirical rejection rates remain above 10 percent and show no sign of convergence to the nominal level as the sample grows in size.

Figure 2 displays the empirical distribution of the studentized $L(3)$ statistic when autocorrelation in jumps are absent. The two plots correspond respectively to the settings in the last row of Panels A and B in Table 1 under DGP 1. The standard normal distribution offers a good overall approximation when $T = 50$ and $\Delta = \frac{1}{252}$ (Figure 2(a)). In the case of $T = 1$ and $\Delta = \frac{1}{23400}$, the empirical distribution is right-skewed (Figure 2(b)). The skewness in the latter case stems from the error of approximating the chi-square distribution by $N(0, 1)$ and the rarity of jumps – there are only an average of 68.1 jumps extracted from simulated data despite the long sample length of $n = 23,400$ (compare to an average of 543.1 jumps over

¹⁸The BLY test we implemented is given in (4.2) of Boswijk, Laeven and Yang (2018) under the null hypothesis of “no self-excitation” with function $H = H(6, 1)$, cutoff parameter $\varpi = 0.3$ and bandwidth $k_n = \lfloor \Delta^{-0.7} \rfloor$ (for estimating the spot jump intensity). The CSS test is the “self-excitement test” given in (17) of Corradi, Silvapulle and Swanson (2018).

the sample length of $n = 12,600$ in the first case).

4.2 Power Performance

Next we examine the ability of $L(k)$ test in detecting the presence of autocorrelated jumps in finite samples. Figure 3 shows the power curves of $L(3)$ test for different sample sizes. We examine two hypothetical scenarios which respectively mimic the fixed T and fixed Δ cases. In the first scenario (Figure 3(a)), data are generated from DGP 1 (infinite-activity jumps) and sampled over a fixed horizon (10 days) with an increasing sampling frequency ranging from 4 minutes ($n = 10 \times 6.5 \times 15$) to 30 seconds ($n = 10 \times 6.5 \times 60 \times 2$). In the second scenario (Figure 3(b)), data are generated from DGP 2 (finite-activity jumps) and sampled daily ($\Delta = \frac{1}{252}$) over a growing horizon from five years (with $n = 252 \times 5$) to 40 years (with $n = 252 \times 40$). It is clear from the plots that, in both scenarios, the power increases with the sample size n for any given level of jump excitation.

It is of interest to compare $L(k)$ test with competing ones in the literature. In particular, the BLY test is available to detect self-excitation of jumps when they are of infinite activity. Figure 4 compares the power of $L(3)$ test with that of the BLY test (under the null of no self-excitement). The data are simulated from DGP 1 (infinite-activity jumps) with either an exponentially decaying kernel $\varphi(x) = 100e^{-\zeta x}1_{\{x \geq 0\}}$, or a periodic kernel $\varphi(x) = 1_{\{0 \leq x \leq \phi\tau\}} \cos(\pi x/30)$. The power curves are obtained for different strength of autocorrelation, by increasing either the ratio $\frac{\alpha}{\zeta}$ for the exponential kernel (Figure 4(a)), or the number of cycles ϕ for the periodic kernel (Figure 4(b)). We keep $T = 1$ and $\Delta = \frac{1}{7800}$ throughout the simulations. We notice that the BLY test is oversized and has a lower power than $L(3)$ test as jump occurrences are more autocorrelated. Furthermore, the design of $L(k)$ test (but not the BLY test) prevents the cancellation between positive and negative autocorrelations at different lag orders, which happens when the periodic kernel is in place. This explains the wider gap between the power curves of the two tests in Figure 4(b).

Another competing method is the CSS test designed for detecting self-excitement of finite-activity jumps. For comparison, we implement both the $L(3)$ and CSS tests on data generated from DGP 2. The power curves are plotted in Figure 5. The only difference between the subplots is in the mean jump size: $\mu_J = -0.03$ in Figure 5(a), and $\mu_J = -0.01$ in Figure 5(b). In both cases, the plots reveal that $L(3)$ test outperforms the CSS test by a large extent as the jumps are self-excited (except under weak self-excitement). When the jump size distribution is more symmetrical about zero (Figure 5(b)), the CSS test is

essentially powerless.¹⁹

4.3 Robusness Analysis

In this section we study the effects on test performance of different DGP specifications and different choice of tuning parameters.

The simulation study for the baseline specifications presumes a fixed amount of jumps relative to diffusive variation. We next study the effect on the test power due to a change in the relative contribution of jump and diffusion to log-return variation. We confine to DGP 2 (finite-activity jumps) in which the variance of jump increments is finite. Let p_J denote the proportion of log-return variance due to jumps.²⁰ The model parameter values of four model specifications with different p_J are recorded in Table A1. Figure A1 displays the power curves of $L(3)$ associated with these four model specifications. For a given level of autocorrelation, the test becomes more powerful as jumps contribute more to the log-return variation relative to diffusion.

Somewhat related to the above study, we examine whether the test power reacts to a change in the absolute amount of jumps in the data. Since the validity of $L(k)$ test is conditional on the presence of jumps in the DGP, we expect that the test may break down when jumps are too scarce. To this end, we perform a simulation study on the test under DGP 2 with $T = 30$, $\Delta = \frac{1}{252}$ and varying unconditional jump intensities $\bar{\lambda}$, ranging from one to 20. This means that the data series generated from these model specifications will have an average number of jumps ranging from 30 to 600. The power curves are displayed in Figure A2. We observe that the test power is quite invariant to the jump intensity. As expected, the test becomes less powerful when jumps are rare ($0 < \bar{\lambda} \leq 1$), and does not yield false positive results when no jumps occur ($\bar{\lambda} = 0$).

Recall that $L(k)$ test involves a number of tuning parameters. They include the power k (which determines the discount on small jumps), weight function parameter ξ (which controls the divergence rate of the bandwidth h , and hence the decay rate of the weight function),

¹⁹This limitation is recognized and discussed in CSS (2018), and becomes the motivation for them to propose an alternative test (denoted $\hat{S}_{t^+, \Delta}^\beta$ in their paper) based on the squared return autocorrelation of lag order one. However, when implementing this alternative test in our simulation study we find that it suffers from a serious oversize problem (which CSS (2018) confirms in their Monte Carlo findings); hence we do not report the results of the alternative test.

²⁰The proportion p_J of log-return variance due to jumps, computed under the null hypothesis of no jump autocorrelations, is given by $p_J = \frac{V^d}{V^c + V^d}$, where $V^c = \theta$ and $V^d = \bar{\lambda}(\mu_J^2 + \sigma_J^2)$ represent the annualized long-run variances of the diffusion and jump components, respectively.

and parameters c and ϖ for the cut-off sequence.²¹ We will investigate how a change of these parameters affects the power of the test.

Starting with the robustness on the choice of k , Figure A3 displays the power curves of $L(k)$ test for different values of k under DGP 1 (infinitely-activity jumps with $\beta = 1.25$, Figure A3(a)) and DGP 2 (finite-activity jumps with $\beta = 0$, Figure A3(b)). The empirical rejection rates are plotted against the ratio $\frac{\alpha}{\zeta}$, which measures the extent of autocorrelation of jump occurrences in the DGP. It is clear that the power decreases with k . Recall that having larger k diminishes smaller jumps even more and leads to smaller effective sample size for jumps. On the other hand, we observe that $L(k)$ test is oversized under DGP 2 when k is smaller than three. This is because the choice of k violates Assumption G, which requires that $k > \frac{1-\beta\varpi}{1-2\varpi} = \frac{1}{1-2(0.3)} = 2.5$ under DGP 2. The requirement on k is less stringent when the jump activity level is higher; e.g., for $\beta = 1.25$ we require $k > \frac{1-1.25(0.3)}{1-2(0.3)} = 1.56$. Without precise knowledge of the nature of jump activity, we recommend the user to set $k = 3$ so as to maximize the test power without violating Assumption G.

Suppose the data are commonly generated from the baseline specification of DGP 1 (infinite-activity jumps). We first conduct $L(3)$ test with different cutoff sequence of the form $u_n = c\sqrt{\widehat{IV}}\Delta^\varpi$ by changing c and ϖ . To compute the integrated volatility estimator $\widehat{IV}$ we use the nonparametric threshold estimator of Mancini (2009).²² The power curves (Figure A5) reveal that the test power tends to be higher with a smaller c and/or a larger ϖ . Intuitively, based on the definition of the filtered sequence $g_{k,i}$, there are more jumps to be fed into the test statistic if the sequence u_n vanishes to zero at a faster rate. However, this may come with the risk of mistaking the larger diffusive increments as jumps. The wrong classification may result in an oversized test when the stochastic volatility process is serially correlated. Since the empirical sizes are close to the nominal level, there is no evidence of misclassification for the choice of cutoff sequences considered in Figure A4.

With the same DGP we conduct $L(3)$ test with different values of ξ and obtain the power curves (Figure A5). It is observed that the power tends to increase with ξ for moderate level

²¹Strictly speaking, the tuning parameters also include the form of the filter function $g(\cdot)$, the weight function $w(\cdot)$, the cutoff sequence as a function of Δ , and h as a function of n . Given what we learned from the nonparametric inference, we expect that the test result is less sensitive to the choice of the functional form than to the choice of the bandwidth parameters (c , ϖ and ξ) that control the decay rate of the cutoff sequence and the weight function.

²²We may employ any other consistent integrated volatility estimator that is robust to autocorrelated jumps, e.g., the threshold bipower variation of Corsi et al. (2010). The realized bipower variation of Barndorff-Nielsen and Shephard (2004) tends to overestimate the integrated volatility in the presence of self-excited jumps and hence should not be used.

of autocorrelation, although there is no clear relationship when jump occurrences are weakly autocorrelated. This is as expected because as ξ increases the test statistic includes more lags in autocorrelation; however, a higher ξ reduces the relative importance on autocorrelation with smaller lags. The latter effect may lead to a less sensitive test to autocorrelation over short range.

As a final simulation study, we analyze the effect of jump activity levels on the test performance. We confine to DGP 1 (infinite-activity jumps) with different values of jump activity level ($1 \leq \beta < 2$) that are plausible in high-frequency financial time series. The power curves are shown in Figure A6. We note that the test size is insensitive to β , that the test is powerful against autocorrelated jumps with different activity levels, and that the test power increases with β for a given level of autocorrelation.

5 Empirical Study

In this section, we apply the inference methods to study the dependence structure of jumps (if they exist) for empirical time series of stock returns. Our sample data consist of 10-minute log-returns of a collection of stocks over 2006-2013. The collection (Table A2) contains 38 stocks that become a Dow Jones constituent at some point during the sample period.²³ In addition, we include SPDR S&P 500 ETF Trust (SPY) that approximates the market portfolio. The NBBO filter (see Aït-Sahalia and Jacod (2014)) is employed to remove any price anomalies in the TAQ transaction data. From the NBBO-filtered data we obtain the 10-minute log-returns series.

We first study the jump dynamics for each sample series over the entire sample period. For each sample series, we compute the sample autocorrelation $\hat{\rho}_{n,k}(\ell)$ (for lags $\ell = 1, \dots, 5$), and conduct $L(3)$ test and the BLY test. The results are reported in Table 2.²⁴ All except one stock have significant $\hat{\rho}_{n,k}(\ell)$ for some lags up to the fifth order (shaded in red in Table 2). The temporal dependence is predominantly in the form of jump clustering (as all $\hat{\rho}_{n,k}(\ell)$ are positive), although the magnitude and persistence of autocorrelation vary across differ-

²³We do not include in our sample Visa Inc. (V), which was listed as an IPO on March 19, 2008 and became a constituent stock on September 23, 2013. See Table A2 and the footnotes for further details of our sample.

²⁴The significance level of all tests is set at five percent. To implement our inference procedure, we fix the end time T to one so that the discretization step is $\Delta_n = \frac{1}{n}$, where n is the number of observations ($n \approx 80,000$ on average). The cutoff of the filtered jump series is set to $u_n := 7\sqrt{\hat{V}}\Delta_n^{0.3}$. The Gaussian kernel is used as the weight function with bandwidth $h = n^{0.7} s.e.(g_{n,k,i})$.

ent sample series. Furthermore, 29 out of 39 sampled stocks are detected with significant autocorrelation jointly over all nonzero lags. It turns out that the strongest rejections by $L(k)$ test occur on those stocks that were hit hard during the global financial crisis in 2008 (e.g., AIG, BAC, C, GM, GS). By comparison, the BLY test implemented on our sample yields a rejection on merely four stocks: AIG, AXP, GM and TRV.²⁵

Recall that the pointwise and omnibus tests are different both in the set of targeted alternatives and in their power performance. Having significant autocorrelation for specific lag orders does not guarantee a significant result when tested jointly (e.g., IBM, INTL and MSFT). We also observe that similar estimates of $\hat{\rho}_{n,k}(\ell)$ do not readily translate to similar level of rejection by $L(k)$ test (e.g., compare BAC and UNH).

The temporal dependence of jump occurrences for four selected stocks (AIG, C, HON and MSFT) is represented visually in Figure 6. The selected stocks exhibit a rich variety of jump autocorrelogram profiles that differ in level and shape. The jump dependence in the log-return series of AIG and C tends to be very persistent, while that for the HON series is more short-lived but stronger in magnitude. Although jumps occur in the MSFT series, their occurrences are not temporally correlated over time. Also displayed in Figure 6 are the log-returns series (the cutoff $\pm u_n$ in red dashed line), the filtered jump series $g_{n,3,i}$, and the spot jump intensity (see equation (3.6) of BLY (2018); $k_n = \lfloor \Delta_n^{-0.7} \rfloor$). The plots of $g_{n,3,i}$ show heavy clustering of jumps for AIG and C. The jumps are clustered much more lightly (but still visible, around late 2008) for HON, while those for MSFT appear to scatter randomly without forming clear clusters. These observations conform to the result of our tests which take $g_{n,3,i}$ as input ($L(3)$ test yields rejections for AIG, C and HON). By contrast, the BLY test rejects the null of no self-excitement of jumps only for AIG among the four stocks. The BLY test is designed to measure the co-movements between log-price and spot jump intensity, which appears to be quite strong for AIG as seen from the plots.

The above analysis with $L(3)$ test presumes stationarity in jump dynamics over a long sample horizon spanning eight years. Given the market fluctuations, it is quite plausible that jump stationarity is a strong assumption. Indeed, we observe from Figure 6 that the jump intensity varies substantially over time. In a similar vein, we expect that the jump dependence structure is subject to temporal variation. To investigate empirically how the

²⁵We implement the BLY test with the null hypothesis of no self-exciting jumps (Section 4.1 of BLY (2018)). As suggested in their paper, we adopt the function $H(6, 1)$ and tuning parameter $k_n = \lfloor \Delta_n^{-0.7} \rfloor$ in the test. Other settings are the same as those for our methods: $T = 1$, $\Delta_n = \frac{1}{n}$, same cutoff u_n and same filtered series $g_{n,k,i}$.

degree of jump dependence may vary over time, we split the log-return series by year and then test for autocorrelation of jump occurrences. Table A3 reports the $L(3)$ test and BLY test results split by year. Despite a substantial reduction in sample size, $L(3)$ yields significant results in many instances (especially in 2008), whereas the BLY test reveals only two instances (AIG and GM in 2008) of significant self-excited jumps.²⁶ Interestingly, there are clearly more stocks with significantly autocorrelated jump occurrences in 2008 than in other years. It seems to suggest a positive association between the strength of autocorrelation of jump occurrences and market downturn.

The above empirical result indicates that jumps excite themselves by different degree over cross section and over time. It is of interest to learn about the underlying attributes that drive such heterogeneity. We may get some hints by relating the autocorrelation of jump occurrences to some market variables. To this end, we consider a panel regression using the estimated lag-one $\hat{\rho}_{n,3}(1)$ as the dependent variable. The predictors include sample statistics of the high-frequency data series (Average Price, Annual Log-return, Annual Volatility) and market variables that measure liquidity and size (Bid-offer Spread, Volume, Turnover, Shares Outstanding, Market Capitalization). The regressions (Table A4) reveal that jumps tend to self-excite more with lower stock price denomination and more negative log-returns; in particular, it provides verification for the tendency to have clustered jumps during market downturn. Self-excitement of jumps tend to be more prevalent with wider bid-offer spread and fewer outstanding shares, both of which are associated with illiquid trading. In addition, bigger stocks (with larger market capitalization) appear to have more jumps clustered together.

6 Concluding Remarks

In this paper, we introduce a set of procedures for inferring the autocorrelation of jumps as an important characterization of time series dynamics. We develop an omnibus test of autocorrelation of jump occurrences for all lag orders under a unified analytical framework that combines in-fill and long-span asymptotics. The test delivers consistent inference under the null hypothesis of uncorrelated, stationary jumps of finite- or infinite-activity level, and is robust to the choice of sampling scheme (discretization steps and sample horizon). Being a generalized portmanteau test, it is powerful for a wide range of alternatives, including

²⁶The significance level is set at 10 percent for all tests in Table A3 in light of the reduction of the number of observations ($n \approx 10,000$ on average) after splitting the sample horizon by year.

self-exciting and/or self-inhibitory jumps, jumps with long-range persistence, and jumps with cyclical autocorrelation structure. The finite-sample experiments confirm its robustness property and reveal its competitive size and power performance relative to existing tests. In addition, we advocate the use of autocorrelogram for visualizing the temporal dependence structure and conducting pointwise inference of autocorrelation up to some fixed lag order. The usefulness of our inference procedure is seen in an empirical study of jump dependence dynamics on a sample of high-frequency stock return data. The inference uncovers a rich array of autocorrelation profiles of jump occurrences for different stocks over different time periods.

There are several directions of further research one may explore. As a natural extension, our procedure is adaptable to a multivariate setting for testing cross-autocorrelation of jumps from different processes. This would allow one to test for the presence of mutually exciting and/or inhibitory jumps originated from different sectors or markets. In terms of implementation, the current approach involves pre-testing of jumps. By modifying the null hypothesis and the test, it may be possible to dispense with the pre-testing step. E.g., an alternative test may be based on the cross-autocorrelation of the filtered jumps and log-returns. The test statistic then remains well defined in the absence of jumps, and no pre-testing is necessary. An important extension is to relax the stationarity restriction; in particular, it may allow for time-varying unconditional jump intensity. The latter can be estimated by nonparametric kernel method. One may then test for autocorrelation of “unexpected jumps” by applying the inference procedure developed in this paper to the normalized residual jump occurrence process, which is covariance stationary under the null hypothesis. They are left for future research.

References

- [1] Aït-Sahalia, Y., and J. Jacod (2009a). Testing for jumps in a discretely observed process. *Annal of Statistics*, 37(1), pp.184—222.
- [2] Aït-Sahalia, Y., and J. Jacod (2009b). Estimating the degree of activity of jumps in high frequency data. *Annal of Statistics*, 37(5a), pp.2202—2244.
- [3] Aït-Sahalia, Y., and J. Jacod (2011). Testing whether jumps have finite or infinite activity. *Annal of Statistics*, 39(3), pp.1689—1719.

- [4] Aït-Sahalia, Y., and J. Jacod (2012). Analyzing the spectrum of asset returns: jump and volatility components in high frequency data. *Journal of Economic Literature*, 50(4), pp.1007—1050.
- [5] Aït-Sahalia, Y., and J. Jacod (2014). *High-Frequency Financial Econometrics*. Princeton University Press, New Jersey.
- [6] Aït-Sahalia, Y., J. Cacho-Diaz, and R.J.A. Laeven (2015). Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics*, 117(3), pp.585—606.
- [7] Bacry, E., I. Mastromatteo, and J. Muzy (2015). Hawkes processes in finance. *Market Microstructure and Liquidity*, 1(1), 1550005.
- [8] Bacry, E., T. Jaisson, and J. Muzy (2018). Estimation of slowly decreasing Hawkes kernels: application to high-frequency order book dynamics. *Quantitative Finance*, 16(8), pp.1179—1201.
- [9] Barndorff-Nielsen, O.E., and N. Shephard (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1), pp.1—37.
- [10] Bartlett, M.S. (1963). The spectral analysis of point processes. *Journal of the Royal Statistical Society (Series B)*, 25(2), pp.264—296.
- [11] Boswijk, H.P., R.J.A. Laeven, and X. Yang (2018). Testing for self-excitation in jumps. *Journal of Econometrics*, 203(2), pp.256—266.
- [12] Bowsher, C. (2007). Modelling security market events in continuous time: Intensity based, multivariate point process models. *Journal of Econometrics*, 141(2), pp.876—912.
- [13] Chavez-Demoulin, V., and J.A. McGill (2012). High-frequency financial data modeling using Hawkes processes. *Journal of Banking and Finance*, 36(12), pp.3415—3426.
- [14] Corradi, V., M. Silvapulle, and N.R. Swanson (2018). Testing for jumps and jump intensity path dependence. *Journal of Econometrics*, 204(2), pp.248—267.
- [15] Corsi, F., D. Pirino, and R. Renò (2010). Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics*, 159(2), pp.276—288.

- [16] Embrechts, P., T. Liniger, and L. Lin (2011). Multivariate Hawkes processes: an application to financial data. *Journal of Applied Probability*, 48A, pp.367—378.
- [17] Engle, R., and J.R. Russell (1998). Autoregressive Conditional Duration: a new model for irregularly spaced transaction data. *Econometrica*, 66(5), pp.1127—1162.
- [18] Eraker, B. (2004). Do stock prices and volatility jump? Reconciling evidence from spot and option prices. *Journal of Finance*, 59(3), pp.1367—1403.
- [19] Hawkes, A.G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1), pp.83—90.
- [20] Hawkes, A.G. (2018). Hawkes processes and their applications to finance: a review. *Quantitative Finance*, 18(2), pp.193—198.
- [21] Hong, Y. (1999). Hypothesis testing in time series via the empirical characteristic function: A generalized spectral density approach. *Journal of the American Statistical Association*, 94(448), pp.1201—1220.
- [22] Jing, B., X. Kong, Z. Liu, and P. Mykland (2012). On the jump activity index for semimartingales. *Journal of Econometrics*, 166, pp.213—223.
- [23] Khashanah, K., J. Chen, and A. Hawkes (2018). A slightly depressing jump model: intraday volatility pattern simulation. *Quantitative Finance*, 18(2), pp.213—224.
- [24] Kwok, S. (2013). Diagnostic checks for multivariate parametric intensity models, Working Paper.
- [25] Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics*, 36, pp.270—296.
- [26] Maneesoonthorn, W., C.S. Forbes, and G.M. Martin (2017). Inference on self-exciting jumps in prices and volatilities using high-frequency measures. *Journal of Applied Econometrics*, 32, pp.504—532.
- [27] Protter, P. (2003). *Stochastic Integration and Differential Equations* (Second Edition), Springer.

Figures and Tables

Figure 1: Admissible values of (ν, ζ) for different jump activity levels β : (a) $\beta = 0$; (b) $\beta = 1$; (c) $\beta = 1.5$.

(a)

(b)

Figure 2: Empirical distribution of studentized $L(3)$ statistic in the absence of jump autocorrelation under DGP1 (infinite-activity jumps) with (a) $T=50$, $\Delta=1/252$; and (b) $T=1$, $\Delta=1/23400$.

Figure 3: Power of $L(3)$ test for different sample sizes under (a) DGP 1 (infinite-activity jumps) with fixed $T=1$; and under (b) DGP 2 (finite-activity jumps) with fixed $\Delta=1/252$.

Figure 4: Power of $L(3)$ and BLY tests under DGP 1 (infinite-activity jumps) with (a) exponentially decaying and (b) periodic kernel for λ_t .

Figure 5: Power of $L(3)$ and CSS tests under DGP 2 (finite-activity jumps), with mean jump size equal to (a) $\mu_J = -0.03$ and (b) $\mu_J = -0.01$.

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

Figure 6: autocorrelogram of filtered series, log-returns, filtered series $g_{n,3,i}$, and spot intensity of selected s

Table 1: Empirical sizes of $L(k)$, BLY and CSS tests.

T	Δ	$L(3)$	$L(6)$	BLY	CSS	$L(3)$	$L(6)$	BLY	CSS
Panel A: DGP 1, fixed Δ					Panel C: DGP 2, fixed Δ				
5	1/252	0.0476	0.0466	0.3892	0.0326	0.0192	0.0068	0.1118	0.0390
10	1/252	0.0460	0.0476	0.7010	0.0336	0.0248	0.0118	0.4854	0.0432
20	1/252	0.0546	0.0552	0.9482	0.0362	0.0400	0.0214	0.8090	0.0402
30	1/252	0.0570	0.0536	0.9924	0.0382	0.0426	0.0252	0.9038	0.0420
40	1/252	0.0520	0.0572	1.0000	0.0434	0.0552	0.0354	0.9382	0.0388
50	1/252	0.0490	0.0480	1.0000	0.0384	0.0528	0.0426	0.9630	0.0500
Panel B: DGP 1, fixed T					Panel D: DGP 2, fixed T				
1	1/390	0.0344	0.0322	0.1172	0.0310	0.0108	0.0004	0.0784	0.0284
1	1/780	0.0420	0.0430	0.1318	0.0302	0.0354	0.0260	0.1362	0.0242
1	1/1950	0.0466	0.0478	0.1382	0.0362	0.0476	0.0498	0.1578	0.0132
1	1/3900	0.0476	0.0492	0.1432	0.0328	0.0482	0.0484	0.1702	0.0078
1	1/7800	0.0536	0.0540	0.1376	0.0370	0.0518	0.0530	0.1846	0.0082
1	1/23400	0.0580	0.0564	0.1370	0.0404	0.0586	0.0580	0.2092	0.0108

Note: The table reports the empirical rejection rate (number of rejections out of 5,000 replications) under the null hypothesis of zero jump autocorrelations. The nominal level is set to 5%. The parameter values are detailed in Section 3. DGP 1 is the SVJ model with infinite-activity jumps. DGP 2 is the SVJ model with finite-activity jumps. The sample horizon is $[0, T]$, and the discretization step is Δ . The sets of (T, Δ) combinations that are incompatible (marginally incompatible) with the asymptotic theory of the current test are shaded in gray (light gray). BLY stands for the self-excitation test in (4.2) of Boswijk, Laeven, and Yang (2018). CSS stands for the self-excitement test in Theorem 4 of Corradi, Silvapulle, and Swanson (2018).

Table 2: Inference on autocorrelation of jump occurrences on sampled stocks.

stock	$\rho_{3,\ell}$					$L(3)$	BLY
	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$	$\ell=5$		
AA	0.04	0.10	0.02	0.03	0.02	29.4	0.8
AIG	0.33	0.19	0.20	0.17	0.17	6472.8	5.7
AXP	0.11	0.14	0.03	0.02	0.02	40.8	2.1
BA	0.03	0.07	0.01	0.01	0.00	2.0	1.8
BAC	0.12	0.04	0.03	0.02	0.02	466.1	1.1
C	0.19	0.11	0.09	0.08	0.04	849.0	1.2
CAT	0.04	0.03	0.01	0.02	0.01	1.9	1.8
CSCO	0.01	0.00	0.00	0.00	0.00	-0.2	0.9
CVX	0.09	0.06	0.02	0.04	0.04	3.9	0.8
DD	0.05	0.03	0.01	0.02	0.01	1.2	0.9
DIS	0.13	0.06	0.01	0.02	0.02	10.0	1.2
GE	0.11	0.03	0.02	0.02	0.02	23.0	0.9
GM	0.28	0.14	0.11	0.11	0.10	991.8	-6.1
GS	0.15	0.08	0.07	0.05	0.06	166.1	1.0
HD	0.05	0.10	0.03	0.06	0.02	8.5	0.7
HON	0.38	0.32	0.06	0.05	0.02	94.3	1.1
HPQ	0.02	0.01	0.01	0.00	0.00	2.9	0.6
IBM	0.04	0.02	0.00	0.00	0.01	1.1	0.8
INTC	0.02	0.01	0.00	0.01	0.00	0.1	1.1
JNJ	0.12	0.06	0.04	0.02	0.01	8.5	0.5
JPM	0.06	0.04	0.03	0.02	0.02	23.9	1.7
KFT	0.03	0.01	0.00	0.00	0.00	0.1	1.1
KO	0.01	0.01	0.00	0.01	0.00	0.0	1.0
MCD	0.07	0.07	0.02	0.03	0.01	7.1	0.5
MMM	0.06	0.03	0.01	0.01	0.07	6.2	0.7
MO	0.04	0.02	0.00	0.01	0.00	0.6	1.0
MRK	0.18	0.02	0.01	0.00	0.01	15.3	0.8
MSFT	0.01	0.00	0.00	0.00	0.00	-0.1	1.9
NKE	0.04	0.01	0.00	0.00	0.00	0.8	1.7
PFE	0.07	0.03	0.01	0.01	0.01	2.4	1.1
PG	0.08	0.08	0.02	0.03	0.01	8.9	0.4
SPY	0.04	0.08	0.01	0.02	0.02	1.7	0.9
T	0.08	0.09	0.03	0.04	0.03	3.9	0.9
TRV	0.30	0.11	0.07	0.05	0.02	88.6	2.2
UNH	0.12	0.09	0.01	0.01	0.01	20.0	1.6
UTX	0.05	0.09	0.02	0.02	0.02	5.4	0.8
VZ	0.07	0.06	0.05	0.23	0.04	13.7	0.9
WMT	0.04	0.02	0.01	0.02	0.01	0.5	0.8
XOM	0.07	0.07	0.02	0.03	0.07	3.6	0.9

Note: All sample series consist of 10-minute log-returns over the sample period 2006-2013.

Columns 2-6 contain estimates of autocorrelations of jump occurrences $\rho_{3,\ell}$ up to the fifth lag.

Significant pointwise tests of $\rho_{3,\ell}$ are shaded in red. Column 8 reports the result of $L(3)$ test.

Significant $L(3)$ test are shaded in orange. The last column reports the self-excitation test in (4.2) of Boswijk, Laeven and Yang (2018). Significant BLY tests are shaded in blue. The significance level of all tests is set at five percent.

Appendix Figures and Tables

Figure A1: Power of $L(3)$ test for different proportions of variance due to jumps under DGP 2 (finite-activity jumps).

Figure A2: Power of $L(3)$ test for different jump intensities under DGP 2 (finite-activity jumps).

Figure A3: Power of $L(k)$ test for different values of k under (a) DGP 1 (infinite-activity jumps); and (b) DGP 2 (finite-activity jumps).

Figure A4: Power of $L(3)$ test for different cut-off sequences $u_n = c\Delta^\varpi$: (a) over different ϖ ; (b) over different c .

Figure A5: Power of $L(3)$ test for different tuning parameter $h = O(n^\xi)$.

Figure A6: Power of $L(3)$ test for different jump activity levels β under DGP 1 (infinite-activity jumps).

Table A1: Different model specifications for the simulation study in Figure A3.

Model Spec	p_J	κ	θ	σ_v	ρ	μ_{Jv}	λ_v	μ_J	σ_J	$\bar{\lambda}$
1	0.35	4.4	0.020	0.24	-0.61	0	0	-0.010	0.040	6.3
2	0.54	4.4	0.014	0.24	-0.61	0	0	-0.010	0.050	6.3
3	0.67	4.4	0.010	0.24	-0.61	0	0	-0.014	0.055	6.3
4	0.83	4.4	0.005	0.24	-0.61	0	0	-0.020	0.060	6.3

Note: All model specifications are with reference to DGP 2 with exponential kernel. p_J denotes the proportion of log-return variance due to jumps, computed under the null hypothesis of zero jump autocorrelation. Those parameters not included in the table take the values in the baseline specification of DGP 2 given in Section 4.

Table A2: List of sampled stocks for empirical analysis.

Symbol	Listed Company	Market Capitalization (in trillion \$)	Relative Spread (in percent)	Annual Return (in percent)	Annual Volatility (in percent)
AA	Alcoa Inc.	17.3	1.07	4.5	47.9
AIG	American International Group Inc.	74.5	1.05	0.2	85.6
AXP	American Express Company	57.9	0.76	24.9	43.0
BA	The Boeing Company	56.6	0.92	22.0	30.6
BAC	Bank of America Corporation	144.9	0.81	14.6	63.9
C	Citigroup Inc.	135.3	0.46	-2.1	67.7
CAT	Caterpillar Inc.	47.7	0.65	24.2	35.4
CSCO	Cisco Systems, Inc.	125.1	0.41	15.1	32.9
CVX	Chevron Corporation	179.8	0.46	17.0	28.7
DD	E.I. du Pont de Nemours & Company	40.9	0.80	13.1	30.1
DIS	The Walt Disney Company	70.5	0.72	26.2	29.6
GE	General Electric Company	249.4	0.50	1.9	33.7
GM ¹	General Motors Corporation	27.0	0.97	-28.0	70.2
GS	The Goldman Sachs Group, Inc.	70.0	0.64	20.6	43.9
HD	The Home Depot, Inc.	67.0	0.67	15.9	29.4
HON	Honeywell International Inc.	40.2	0.89	20.6	29.3
HPQ	Hewlett-Packard Company	85.9	0.60	18.3	34.5
IBM	International Business Machines Corp.	173.8	0.56	8.1	22.7
INTC	Intel Corporation	116.5	0.38	11.3	31.3
JNJ	Johnson & Johnson	185.6	0.41	9.8	16.5
JPM	JPMorgan Chase & Co.	156.6	0.63	21.3	48.0
KFT ²	Kraft Foods Inc.	42.6	0.86	10.4	21.1
KO	The Coca-Cola Company	136.6	0.50	10.3	19.5
MCD	McDonald's Corporation	73.8	0.51	20.0	20.0
MMM	3M Company	59.3	0.87	16.9	23.9
MO	Altria Group Inc.	82.4	0.61	23.4	20.4
MRK	Merck & Co., Inc.	102.5	0.60	6.7	27.5
MSFT	Microsoft Corporation	247.4	0.38	10.7	29.1
NKE	Nike Inc.	28.6	1.06	30.2	29.3
PFE	Pfizer Inc.	159.1	0.78	1.9	24.2
PG	The Procter & Gamble Company	189.3	0.48	13.1	18.6
SPY	SPDR S&P 500 ETF Trust	85.2	0.08	10.6	22.1
T	AT&T Inc.	178.2	0.65	2.8	24.1
TRV ³	The Travelers Companies, Inc.	28.1	1.16	28.7	32.5
UNH	UnitedHealth Group Inc.	52.7	0.98	30.6	37.6
UTX	United Technologies Corporation	68.7	0.87	21.0	26.0
VZ	Verizon Communications Inc.	106.6	0.69	6.1	23.5
WMT	Wal-Mart Stores, Inc.	208.1	0.47	8.4	20.0
XOM	Exxon Mobil Corp	395.3	0.33	15.7	26.6

Notes: The sample consists of the S&P 500 ETF (SPY) and 38 stocks that became a Dow Jones constituent stock at some point during 2006-2013. The sample does not include Visa Inc. (V), which was listed as an IPO on March 19, 2008 and became a constituent stock on September 23, 2013. Market Capitalization (in trillion \$) is the average product of shares outstanding and daily closing price. Relative Spread is the ratio of average daily bid-offer spread and average daily closing price. Annual Return is the annualized log-return computed over the sample period. Annual Volatility is the annualized realized volatility computed from daily log-returns over the sample period.

¹GM stopped trading after June 1, 2009 as its share price fell below \$1. It was reissued as an IPO on November 18, 2010.

²KFT spun off to KRFT and MDLZ on October 2, 2012. The prices of KRFT are used in our analysis after the spin-off.

³TRV was renamed from STA on February 27, 2007. STA was first listed on NYSE on April 2, 2004.

Table A3: $L(3)$ test and BLY test on sampled stocks over separate years in 2006-2013.

	2006		2007		2008		2009		2010		2011		2012		2013	
	L(3)	BLY	L(3)	BLY	L(3)	BLY	L(3)	BLY	L(3)	BLY	L(3)	BLY	L(3)	BLY	L(3)	BLY
AA	-0.1	0.3	0.2	0.2	0.3	1.0	0.0	0.4	0.0	0.0	0.0	0.1	0.0	0.2	0.0	0.2
AIG	4.8	0.0	2.2	0.2	492.0	2.5	12.6	0.2	0.0	0.6	0.0	0.2	0.0	0.3	0.0	0.4
AXP	30.9	0.1	0.1	0.0	3.4	0.4	12.9	0.1	0.0	0.1	0.2	0.1	0.2	0.0	0.0	0.2
BA	0.2	0.1	0.3	0.1	0.4	0.6	0.0	0.2	0.2	0.1	0.0	0.2	0.4	0.1	1.4	0.3
BAC	0.0	0.0	0.2	0.1	0.0	1.0	0.3	1.4	0.0	0.2	-0.1	0.9	0.0	-0.1	0.0	0.1
C	0.0	0.0	0.0	0.5	18.3	-0.5	5.2	1.0	0.0	0.3	-0.1	1.0	0.2	0.1	0.0	0.1
CAT	0.0	1.1	0.0	0.7	2.0	0.3	0.0	0.4	0.0	0.2	0.0	0.2	0.1	0.1	0.0	0.5
CSCO	-0.1	1.2	0.0	0.4	1.2	0.0	0.0	0.0	-0.1	1.1	-0.1	1.6	-0.1	1.2	-0.1	1.3
CVX	0.0	0.0	0.0	0.0	1.8	0.2	0.0	0.0	0.0	0.1	0.1	0.1	0.0	0.0	0.0	0.0
DD	2.9	0.1	0.2	0.0	0.4	0.3	1.0	0.0	9.1	0.1	0.1	0.2	0.0	1.0	0.0	0.3
DIS	2.9	0.1	1.3	0.1	7.1	0.1	0.0	0.2	0.1	0.1	4.1	0.8	0.0	0.8	0.0	0.1
GE	9.3	0.1	0.0	0.0	0.8	1.0	2.6	0.2	0.0	0.3	1.5	0.3	0.0	0.0	0.2	0.1
GM	0.0	0.7	0.3	0.2	39.0	2.4	14.5	1.1	0.0	0.0	0.0	0.3	0.0	0.2	0.0	0.1
GS	0.0	0.1	0.2	0.1	2.7	1.0	0.6	0.1	4.8	0.3	0.5	0.1	0.0	0.0	0.1	0.1
HD	0.1	0.1	0.0	0.5	2.5	0.1	0.1	0.1	0.0	0.0	0.1	0.2	0.0	0.3	8.6	0.2
HON	13.4	0.1	0.2	0.2	31.9	0.5	0.0	0.0	0.1	0.2	0.4	0.1	0.5	0.1	0.1	0.1
HPQ	0.0	0.3	0.2	0.1	0.0	1.0	0.6	0.0	0.0	0.4	0.1	0.7	0.0	1.1	-0.1	1.1
IBM	0.0	0.2	0.1	0.1	0.9	0.1	0.1	0.0	0.0	0.5	0.0	0.2	0.0	0.4	-0.1	1.3
INTC	0.0	-0.3	0.0	0.2	0.0	0.2	0.0	0.1	0.6	0.2	0.0	0.3	0.0	0.2	0.0	0.1
JNJ	3.8	0.2	0.0	0.2	4.2	0.2	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
JPM	0.3	0.0	0.1	0.1	0.0	1.0	0.2	0.2	0.1	0.0	0.0	0.3	0.0	0.9	0.0	0.0
KFT	0.3	0.1	1.2	0.1	4.4	0.2	0.0	0.9	30.4	0.2	0.1	0.4	0.0	0.4	0.0	0.2
KO	22.1	0.1	2.4	0.0	0.1	0.2	0.0	0.0	0.0	0.2	0.1	0.1	0.0	1.0	0.0	0.2
MCD	0.0	0.0	1.1	0.1	2.3	0.1	0.0	0.0	0.0	0.1	0.2	0.1	0.0	0.2	0.0	0.1
MMM	0.0	0.9	0.0	0.3	2.2	0.2	0.0	0.0	14.0	0.3	0.0	0.4	0.0	0.1	0.1	0.2
MO	0.0	0.4	0.0	1.0	0.0	1.0	0.0	0.0	9.1	0.9	0.2	0.1	0.0	0.0	0.0	0.0
MRK	0.0	0.1	0.0	0.4	2.1	1.3	0.0	1.1	0.0	0.0	0.1	-0.1	0.0	0.3	0.0	0.5
MSFT	0.0	1.1	0.0	1.1	1.9	0.1	0.0	0.6	0.1	0.1	0.7	0.0	0.0	0.2	-0.1	1.5
NKE	0.0	0.3	0.0	1.0	6.7	0.2	0.0	0.1	0.0	0.2	0.0	0.7	-0.1	1.0	-0.1	1.2
PFE	0.0	1.0	0.1	0.1	3.6	0.3	0.0	0.0	0.1	0.1	0.3	0.1	0.0	0.0	0.9	0.1
PG	1.3	0.2	0.2	0.1	3.3	0.2	0.0	0.0	1.0	0.3	0.9	0.0	0.2	0.2	0.0	0.4
SPY	0.0	0.0	0.0	0.1	0.6	0.3	0.0	0.0	0.0	0.2	0.0	0.1	0.0	0.0	0.0	0.0
T	0.0	0.0	1.0	0.3	2.0	0.2	0.1	0.0	2.2	0.1	0.2	0.1	0.0	0.0	0.0	0.5
TRV	0.9	0.1	9.9	0.1	1.1	1.1	3.1	0.0	0.1	0.0	0.4	0.1	0.1	0.1	0.0	0.2
UNH	0.3	0.1	0.0	0.2	0.9	1.0	0.4	0.2	2.1	0.0	0.0	0.5	0.0	0.5	0.6	0.6
UTX	0.0	0.2	2.8	0.1	3.5	0.2	0.1	0.0	0.5	0.2	0.0	0.5	0.0	0.0	0.0	0.0
VZ	0.2	0.1	0.3	0.1	2.5	0.2	0.2	0.0	0.5	0.2	2.1	0.1	0.0	0.0	0.0	0.2
WMT	0.0	0.1	0.0	0.2	7.0	0.1	0.0	0.2	0.2	0.1	0.4	0.1	0.0	0.8	0.1	0.1
XOM	0.0	0.0	0.0	0.0	1.4	0.3	0.0	0.0	35.3	0.3	0.1	0.0	0.0	0.0	0.0	0.0

Note: All sample series consist of 10-minute log-returns over separate years in 2006-2013. Significant $L(3)$ tests are shaded in orange. Significant BLY tests are shaded in blue. The significance level of all tests is set at 10 percent.

Table A4: Regression of lag-one jump autocorrelation on market variables.

	(1)	(2)	(3)
Price	-0.0004 *** (0.0001)	-0.0004 *** (0.0001)	-0.0007 *** (0.0002)
Bid-offer Spread	0.0439 ** (0.0189)	0.0386 * (0.0207)	0.0483 ** (0.0225)
Yearly Log-return	-0.0301 ** (0.0136)	-0.0339 ** (0.0144)	-0.0325 ** (0.0145)
Yearly Volatility		-0.1240 (0.1630)	-0.1320 (0.1740)
Volume		-9.06×10 ⁻⁵ (5.51×10 ⁻⁵)	1.28×10 ⁻⁴ (1.34×10 ⁻⁴)
Turnover			-9.76×10 ⁻⁷ (1.27×10 ⁻⁶)
Shares Outstanding			-6.74×10 ⁻⁶ ** (3.15×10 ⁻⁶)
Market Capitalization			1.92×10 ⁻⁷ ** (8.52×10 ⁻⁸)
No. of Observations	312	312	312
Adjusted R-squares	0.027	0.028	0.0317
F Statistic	3.9 ***	2.8 **	2.27 **

Note: Pooled regression is run on panel data consisting of 39 stocks and eight years (2006-2013). Dependent variable is the estimated lag-one autocorrelation of the filtered series computed over 10-minute log-returns over separate years. Price, Bid-offer Spread, Volume (in trillion shares) and Shares Outstanding (in trillion shares) are daily quantities averaged over separate years. Yearly Volatility is obtained by the threshold estimator of Mancini (2009). Turnover (in trillion \$) is the average product of Volume and Price. Market Capitalization (in trillion \$) is the average product of Shares Outstanding and Price. Heteroscedasticity-robust standard errors are in parenthesis. *Significant at 10 percent. **Significant at 5 percent. ***Significant at 1 percent.

Nonparametric Inference of Jump Autocorrelation

Simon Kwok

January 2021

1 Technical Appendix

First, some conventions on the notation. We write $A \gg B$ if A dominates B in the sense that $\frac{B}{A} = o_p(1)$ in the limit. We write $A \sim B$ if $\frac{A}{B} = O_p(1)$ in the limit, i.e., A and B have the same stochastic order. We write $A \simeq B$ if $\frac{A}{B} = 1 + o_p(1)$ in the limit. For notational convenience, we suppress the dependence on n and k (in $g_{n,k,i}$, Δ_n , $L(k)$, etc.) unless omitting them might cause confusion.

All the moments in Theorems 1-4 are evaluated conditional on Ω_J , the set in which jumps occur.

We first present some useful intermediate results, followed by their technical proofs.

Lemma 1 *Suppose that the diffusive process X_t^c is as defined in Assumption D. Then, for $c > 0$, $0 < \varpi < \frac{1}{2}$ and small $\Delta > 0$, we have*

$$P\{|\sigma_t \Delta W_t| \geq c\Delta^\varpi\} \leq \sqrt{\frac{2}{\pi}} c^{-1} \Delta^{\frac{1}{2}-\varpi} E\left(\sigma_t \exp\left\{-\frac{c^2 \Delta^{2\varpi-1}}{2\sigma_t^2}\right\}\right).$$

The set $\{|\sigma_t \Delta W_t| \geq c\Delta^\varpi\}$ converges to a null set as $\Delta \rightarrow 0$.

Lemma 2 *Suppose that the pure jump process X_t^d is as defined in Assumption D(2). Then, for $c > 0$, $0 < \varpi < \frac{1}{2}$ and small $\Delta > 0$, we have*

$$P\{|\Delta X_t^d| \geq c\Delta^\varpi\} = \frac{\lambda}{\beta c^\beta} \Delta^{1-\beta\varpi}.$$

In addition, the symmetric difference of the sets $\{|\Delta X_t| > c\Delta^\varpi\}$ and $\{|\Delta X_t^d| > c\Delta^\varpi\}$ converges to a null set as $\Delta \rightarrow 0$.

Lemma 3 Suppose that the pure jump process X_t^d is as defined in Assumption D(1). Then, for all $c > 0$ and $0 < \varpi < \frac{1}{2}$, we have, for small $\Delta > 0$,

$$P\{|\Delta X_t^d| > c\Delta^\varpi\} = P\{|J_t| > c\Delta^\varpi\}P\{\Delta N_t = 1\} + o(\Delta) = O(\Delta).$$

In addition, the symmetric difference of the sets $\{|\Delta X_t| > c\Delta^\varpi\}$ and $\{|\Delta X_t^d| > c\Delta^\varpi\}$ converges to a null set as $\Delta \rightarrow 0$.

Lemma 4 Suppose Assumption D holds. Define $\tilde{g}_k(\frac{\Delta X_t}{c\Delta^\varpi}) := g(\frac{\Delta X_t}{c\Delta^\varpi}) - E[g(\frac{\Delta X_t}{c\Delta^\varpi})]$, where $c > 0$ and $0 < \varpi < \frac{1}{2}$. Then the following statements are valid as $\Delta \rightarrow 0$.

(1) For $0 < k < \frac{1-\beta\varpi}{1-2\varpi}$,

$$\tilde{g}_k\left(\frac{\Delta X_t}{c\Delta^\varpi}\right) \simeq \left|\frac{\sigma_t \Delta W_t}{c\Delta^\varpi}\right|^k 1_{\{|\Delta X_t| \leq c\Delta^\varpi\}} = O_p(\Delta^{k(1/2-\varpi)}),$$

(2) For $k = \frac{1-\beta\varpi}{1-2\varpi}$,

$$\tilde{g}_k\left(\frac{\Delta X_t}{c\Delta^\varpi}\right) \simeq \left|\frac{\sigma_t \Delta W_t}{c\Delta^\varpi}\right|^k 1_{\{|\Delta X_t| \leq c\Delta^\varpi\}} + 1_{\{|\Delta X_t^d| > c\Delta^\varpi\}} - P(|\Delta X_t^d| > c\Delta^\varpi) = O_p(\Delta^{(1-\beta\varpi)/2}).$$

(3) For $k > \frac{1-\beta\varpi}{1-2\varpi}$,

$$\tilde{g}_k\left(\frac{\Delta X_t}{c\Delta^\varpi}\right) \simeq 1_{\{|\Delta X_t^d| > c\Delta^\varpi\}} - P(|\Delta X_t^d| > c\Delta^\varpi) = O_p(\Delta^{(1-\beta\varpi)/2}),$$

and

$$\left[\tilde{g}_k\left(\frac{\Delta X_t}{c\Delta^\varpi}\right)\right]^r \simeq \tilde{g}_k\left(\frac{\Delta X_t}{c\Delta^\varpi}\right), \quad r = 2, 3, \dots$$

Lemma 5 Under Assumption T_L , the following statements are valid for $r \geq 1$ as $n \rightarrow \infty$:

- (i) $\sum_{\ell=1}^{n-1} w_h^r(\ell) = \frac{1}{h^{r-1}} \int_0^\infty w^r(x) dx + o\left(\frac{1}{h^{r-1}}\right),$
- (ii) $\sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) w_h^r(\ell) = \frac{1}{h^{r-1}} \int_0^\infty w^r(x) dx + o\left(\frac{1}{h^{r-1}}\right),$
- (iii) $\sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) \left(1 - \frac{\ell+1}{n}\right) w_h^r(\ell) = \frac{1}{h^{r-1}} \int_0^\infty w^r(x) dx + o\left(\frac{1}{h^{r-1}}\right).$

1.1 Proof of Lemma 1

If $Z \sim N(0, 1)$, then, for any $u > 0$, we have the following upper bound for the tail probability of Z .

$$P\{|Z| > u\} = 2 \int_u^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \leq \frac{1}{u} \sqrt{\frac{2}{\pi}} \int_u^\infty z e^{-\frac{z^2}{2}} dz = \frac{1}{u} \sqrt{\frac{2}{\pi}} e^{-\frac{u^2}{2}}.$$

Since $\frac{\Delta W_t}{c\Delta^\varpi} \stackrel{d}{=} c^{-1}\Delta^{\frac{1}{2}-\varpi}W_1$, we have, by setting $u = c\sigma_t^{-1}\Delta^{\varpi-\frac{1}{2}}$,

$$\begin{aligned} P\{|\sigma_t\Delta W_t| > c\Delta^\varpi | \sigma_t\} &= P\{|W_1| > c\sigma_t^{-1}\Delta^{\varpi-\frac{1}{2}} | \sigma_t\} \\ &\leq \sqrt{\frac{2}{\pi}}c^{-1}\Delta^{\frac{1}{2}-\varpi}\sigma_t \exp\left\{-\frac{c^2\Delta^{2\varpi-1}}{2\sigma_t^2}\right\}. \end{aligned}$$

Taking expectation on both sides yields the first result. The expectation exists because the probability is bounded by one. The second result immediately follows by dominated convergence theorem.

1.2 Proof of Lemma 2

For the assumed pure jump process, we can compute the conditional probability of jump increments. For $0 < u < 1$ and $\Delta > 0$, we have:

$$\begin{aligned} P\left\{|\Delta X_t^d| > u | \mathcal{F}_{t-}\right\} &= \int_t^{t+\Delta} \int_{|x|>u} \nu^X(ds, dx) = \int_t^{t+\Delta} \lambda_{s-} ds \int_u^\infty \frac{dx}{|x|^{1+\beta}} \\ &= \frac{1}{\beta u^\beta} \int_t^{t+\Delta} \lambda_{s-} ds, \end{aligned}$$

Noting that $\lambda := E(\lambda_t)$ exists and $\lambda < \infty$, we obtain, after taking expectations on both sides,

$$P\left\{|\Delta X_t^d| > u\right\} = \frac{\lambda\Delta}{\beta u^\beta}.$$

Substituting $u = c\Delta^\varpi$ yields the first result. To show the second result, we start with the inequality:

$$\left|b_t\Delta + \Delta X_t^d\right| - |\sigma_t\Delta W_t| \leq |\Delta X_t| \leq |\sigma_t\Delta W_t| + \left|b_t\Delta + \Delta X_t^d\right|.$$

It follows that, for any fixed $\epsilon > 0$ and for small enough Δ , we have the following set relations:

$$\begin{aligned} \{|\Delta X_t| > c\Delta^\varpi\} &\subseteq \{|\sigma_t\Delta W_t| > \epsilon c\Delta^\varpi\} \cup \left\{\left|b_t\Delta + \Delta X_t^d\right| > (1-\epsilon)c\Delta^\varpi\right\}, \\ \{|\Delta X_t| > c\Delta^\varpi\} &\supseteq \{|\sigma_t\Delta W_t| < \epsilon c\Delta^\varpi\} \cap \left\{\left|b_t\Delta + \Delta X_t^d\right| > (1+\epsilon)c\Delta^\varpi\right\}. \end{aligned}$$

Note that the drift term $b_t\Delta$ does not appear in the above set relations because $|b_t\Delta| < c\Delta^\varpi$ for all small enough Δ . By lemmas 1 and 2, the tail event of $|\sigma_t\Delta W_t|$ converges faster to the null set than that of $|\Delta X_t^d|$ as $\Delta \rightarrow 0$. Retaining only the dominant sets, we obtain

$$\left\{\left|b_t\Delta + \Delta X_t^d\right| > (1+\epsilon)c\Delta^\varpi\right\} \subseteq \{|\Delta X_t| > c\Delta^\varpi\} \subseteq \left\{\left|b_t\Delta + \Delta X_t^d\right| > (1-\epsilon)c\Delta^\varpi\right\}.$$

Since the above relation holds for all $\epsilon > 0$, the second result follows by letting $\epsilon \downarrow 0$.

1.3 Proof of Lemma 3

We see that

$$\begin{aligned} P\left\{\left|\Delta X_t^d\right| > c\Delta^\varpi \middle| \mathcal{F}_{t-}\right\} &= E[1_{\{|\Delta X_t^d| > c\Delta^\varpi\}} \middle| \mathcal{F}_{t-}] \\ &= 1_{\{J_t > c\Delta^\varpi\}} P\{\Delta N_t = 1 \middle| \mathcal{F}_{t-}\}. \end{aligned}$$

The first result follows by taking expectations on both sides, and noting that J_t is independent of $\mathcal{F}_{t-}$. The second result follows by a similar argument as in lemma 2, meanwhile observing that $P\{|\Delta X_t^d| > c\Delta^\varpi\} = O(\Delta)$ converges at a slower rate than $P\{|\sigma_t \Delta W_t| > c\Delta^\varpi\}$, which decays with Δ exponentially by lemma 1.

1.4 Proof of Lemma 4

Let $\Omega_s := \{|\Delta X_t| \leq c\Delta^\varpi\}$ (the set of small increments) and Ω_s^c be its complement (the set of large increments). We start with the decomposition

$$\tilde{g}_k\left(\frac{\Delta X_t}{c\Delta^\varpi}\right) = \left[\left|\frac{\Delta X_t}{c\Delta^\varpi}\right|^k 1_{\Omega_s} - E\left(\left|\frac{\Delta X_t}{c\Delta^\varpi}\right|^k 1_{\Omega_s}\right)\right] + [1_{\Omega_s^c} - P(\Omega_s^c)] =: G^c + G^d. \quad (1)$$

We want to study the two terms and compare their stochastic orders. To this end, we evaluate their variances. Let us start with $G^c := |\frac{\Delta X_t}{c\Delta^\varpi}|^k 1_{\Omega_s} - E(|\frac{\Delta X_t}{c\Delta^\varpi}|^k 1_{\Omega_s})$. Recall that $\Delta X_t = b_t \Delta + \sigma_t \Delta W_t + \Delta X_t^d$. Over the set Ω_s , the stochastic orders of the drift and diffusive components are $O_p(\Delta)$ and $O_p(\Delta^{1/2})$, respectively.

We now split the proof into two halves. The result is proved first under Assumption D(2), and then under Assumption D(1).

(i) Let us suppose that Assumption D(2) holds. The second moment of the pure jump component over the set Ω_s is evaluated as follows:

$$\begin{aligned} E(|\Delta X_t^d|^2 1_{\Omega_s}) &= \int_t^{t+\Delta} \int_{|x| \leq c\Delta^\varpi} x^2 F(dx) dt = 2\lambda_{t-} \Delta \int_0^{c\Delta^\varpi} \frac{x^2}{x^{1+\beta}} dx \\ &= \frac{2\lambda_{t-} c^{2-\beta}}{2-\beta} \Delta^{1+(2-\beta)\varpi}, \end{aligned}$$

where the first equality makes use of the Lévy measure given in Assumption D(2). We deduce that $|\Delta X_t^d| 1_{\Omega_s}$ is of order $O_p(\Delta^{\frac{1}{2} + \frac{2-\beta}{2}\varpi})$, which for all small Δ is dominated by the diffusive component (which is $O_p(\Delta^{1/2})$) because $\beta < 2$. The increments due to the diffusive component belong to Ω_s for all small Δ because $|\sigma_t \Delta W_t| = O_p(\Delta^{1/2}) < c\Delta^\varpi$ for $\varpi < \frac{1}{2}$. It follows that

$$E\left(\left|\frac{\Delta X_t}{c\Delta^\varpi}\right|^{2k} 1_{\Omega_s}\right) = E\left(\left|\frac{\sigma_t \Delta W_t}{c\Delta^\varpi}\right|^{2k}\right) = O(\Delta^{k(1-2\varpi)}),$$

and hence

$$\text{Var}(G^c) = E \left(\left| \frac{\Delta X_t}{c\Delta^\varpi} \right|^{2k} 1_{\Omega_s} \right) - \left[E \left(\left| \frac{\Delta X_t}{c\Delta^\varpi} \right|^k 1_{\Omega_s} \right) \right]^2 = O(\Delta^{k(1-2\varpi)})$$

for small Δ . As a result, we have

$$G^c := \left| \frac{\Delta X_t}{c\Delta^\varpi} \right|^k 1_{\Omega_s} - E \left(\left| \frac{\Delta X_t}{c\Delta^\varpi} \right|^k 1_{\Omega_s} \right) = O_p \left(\Delta^{k(1/2-\varpi)} \right) \quad (2)$$

for small Δ . In other words, the small increments are predominantly driven by diffusion. Note that the validity of this result does not require $E(\sigma_t) < \infty$ because $|\sigma_t \Delta W_t|$ is bounded from above by $c\Delta^\varpi$ over Ω_s .

Let us turn to $G^d := 1_{\Omega_s^c} - P(\Omega_s^c)$. For small enough Δ , only the pure jump increments belong to the set $\Omega_s^c = \{|\Delta X_t| > c\Delta^\varpi\}$ because the increments due to the drift and diffusive components are smaller than $c\Delta^\varpi$ in magnitude. We deduce from lemma 2 that $P\{\Omega_s^c\} = O(\Delta^{1-\beta\varpi})$. It follows that

$$\text{Var}(G^d) = P\{\Omega_s^c\} - [P\{\Omega_s^c\}]^2 = O(\Delta^{1-\beta\varpi}),$$

and hence

$$G^d := 1_{\Omega_s^c} - P(\Omega_s^c) = O_p(\Delta^{(1-\beta\varpi)/2}), \quad (3)$$

for small Δ . In other words, the big increments are predominantly driven by pure jumps.

Now we compare the stochastic orders of G^c and G^d . The condition $0 < \varpi < \frac{1}{2}$ implies that

$$k \left(\frac{1}{2} - \varpi \right) > 1 - \beta\varpi \iff k > \frac{1 - \beta\varpi}{1 - 2\varpi},$$

with analogous relations for $<$ and $=$. This means that, for small Δ , G^c converges faster (more slowly) to zero than G^d when $k > \frac{1-\beta\varpi}{1-2\varpi}$ ($k < \frac{1-\beta\varpi}{1-2\varpi}$). The two terms converge at the same rate if $k = \frac{1-\beta\varpi}{1-2\varpi}$.

Now we show the last approximation in the lemma. Given that $k > \frac{1-\beta\varpi}{1-2\varpi}$, it follows from the above analysis and the nature of the indicator function ($1_{\Omega_s^c}^r = 1_{\Omega_s^c}$) that

$$\left[\tilde{g}_k \left(\frac{\Delta X_t}{c\Delta^\varpi} \right) \right]^r \simeq [G^d]^r = [1_{\Omega_s^c} - P(\Omega_s^c)]^r \simeq 1_{\Omega_s^c} - P(\Omega_s^c) = \tilde{g}_k \left(\frac{\Delta X_t}{c\Delta^\varpi} \right).$$

(ii) It remains to prove the result under Assumption D(1). Again we start from the decomposition (1). Following the above analysis, the stochastic orders of the drift and diffusive components are $O_p(\Delta)$ and $O_p(\Delta^{1/2})$, respectively. The pure jump component is filtered away because the jumps $\Delta X_t^d = J_t \Delta N_t$, once occurred, are $O_p(1)$ (as J_t is $O_p(1)$) and hence do not belong to the set Ω_s for small enough Δ . It follows that the diffusive component is dominant among the three terms. In other words, the small increments is driven by diffusion.

Next, we turn to large increments. By lemma 3 and the properties of the simple point process, we have, for small Δ ,

$$\begin{aligned} P\{|\Delta X_t^d| \geq c\Delta^\varpi\} &= P\{|\Delta X_t^d| \geq c\Delta^\varpi\} + o(\Delta) \\ &= P(\{\Delta N_t = 1\} \cup \{|J_t| \geq c\Delta^\varpi\}) + o(\Delta) \\ &= O(\Delta). \end{aligned}$$

It follows that $Var(G^d) = P\{|\Delta X_t^d| \geq c\Delta^\varpi\} - [P\{|\Delta X_t^d| \geq c\Delta^\varpi\}]^2 = O(\Delta)$, and hence $G^d = O_p(\Delta^{1/2})$, for small Δ .

Using the above results, and noting that $k(\frac{1}{2} - \varpi) > \frac{1}{2} \iff k > \frac{1}{1-2\varpi}$ (analogous relations for $<$ and $=$), we see that the conclusion remains valid for finite-activity jumps as defined under Assumption D(1). The proof is now completed.

1.5 Proof of Lemma 5

Let $x = \frac{\ell}{h}$ and $\Delta x = \frac{1}{h}$. Then (i) becomes:

$$\sum_{\ell=1}^{n-1} w_h^r(\ell) = \sum_{\ell=1}^{n-1} \frac{1}{h^r} w^r\left(\frac{\ell}{h}\right) = \frac{1}{h^{r-1}} \sum_{\ell=1}^{n-1} w^r(x) \Delta x.$$

Under Assumption T_L , we observe that $h \rightarrow \infty$ and $\frac{h}{n} \rightarrow 0$ as $n \rightarrow \infty$. This gives

$$\sum_{\ell=1}^{n-1} w_h^r(\ell) - \frac{1}{h^{r-1}} \int_0^\infty w^r(x) dx = o\left(\frac{1}{h^{r-1}}\right).$$

Applying the same change of variable to (ii) yields:

$$\sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) w_h^r(\ell) = \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell/h}{n/h}\right) \frac{1}{h^r} w^r\left(\frac{\ell}{h}\right) = \frac{1}{h^{r-1}} \sum_{\ell=1}^{n-1} \left(1 - \frac{x}{n/h}\right) w^r(x) \Delta x.$$

Passing to the limit, we obtain

$$\sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) w_h^r(\ell) - \frac{1}{h^{r-1}} \int_0^\infty w^r(x) dx = o\left(\frac{1}{h^{r-1}}\right).$$

An analogous argument leads to (iii).

1.6 Proof of Theorem 1

To begin, let us approximate $\hat{\gamma}(\ell)$. Let $\tilde{\gamma}(\ell)$ be the infeasible analog of $\hat{\gamma}_\ell$, given by $\tilde{\gamma}(\ell) := \frac{1}{n} \sum_{i=1}^{n-\ell} \tilde{g}_i \tilde{g}_{i+\ell}$, where $\tilde{g}_i := g_i - \mu_g$. The population mean $\mu_g := E(g_i)$ is constant over i by

stationarity. Under Assumption D₀, g_i are uncorrelated for all sufficiently large n , and hence, by lemma 4, we obtain

$$Var(\bar{g} - \mu_g) = Var\left(\frac{1}{n} \sum_{i=1}^n \tilde{g}_i\right) = \frac{1}{n^2} \sum_{i=1}^n Var(\tilde{g}_i) = O\left(\frac{1}{n} \Delta^{1-\beta\varpi}\right).$$

This implies that $\bar{g} - \mu_g = O(\frac{1}{\sqrt{n}} \Delta^{(1-\beta\varpi)/2})$. Noting that $g_i - \bar{g} = (g_i - \mu_g) - (\bar{g} - \mu_g)$, we have

$$\begin{aligned} \hat{\gamma}(\ell) &= \frac{1}{n} \sum_{i=1}^{n-\ell} (g_i - \bar{g})(g_{i+\ell} - \bar{g}) \\ &= \frac{1}{n} \sum_{i=1}^{n-\ell} (g_i - \mu_g)(g_{i+\ell} - \mu_g) - (\bar{g} - \mu_g) \frac{1}{n} \sum_{i=1}^{n-\ell} (g_i - \mu_g) \\ &\quad - (\bar{g} - \mu_g) \frac{1}{n} \sum_{i=1}^{n-\ell} (g_{i+\ell} - \mu_g) + \frac{1}{n} \sum_{i=1}^{n-\ell} (\bar{g} - \mu_g)^2 \\ &= \tilde{\gamma}(\ell) - \left(1 - \frac{\ell}{n}\right) (\bar{g} - \mu_g)^2 - \frac{1}{n} \left(\sum_{i=1}^{\ell} \tilde{g}_i + \sum_{i=n-\ell+1}^n \tilde{g}_i \right) (\bar{g} - \mu_g) \\ &= : (I) + (II) + (III). \end{aligned} \tag{4}$$

The second term is given by

$$(II) = - \left(1 - \frac{\ell}{n}\right) (\bar{g} - \mu_g)^2 \leq (\bar{g} - \mu_g)^2 = O_p\left(\frac{1}{n} \Delta^{1-\beta\varpi}\right)$$

uniformly over ℓ . The third term is

$$(III) = -\frac{1}{n} \left(\sum_{i=1}^{\ell} \tilde{g}_i + \sum_{i=n-\ell+1}^n \tilde{g}_i \right) (\bar{g} - \mu_g) \leq O_p\left(\frac{1}{n^2} \Delta^{1-\beta\varpi}\right),$$

again uniformly over ℓ . Under the null hypothesis, the first term $(I) = \tilde{\gamma}(\ell)$ has mean zero.

Its variance is

$$\begin{aligned} Var[\tilde{\gamma}(\ell)] &= \frac{1}{n^2} \sum_{i=1}^{n-\ell} E\tilde{g}_i^2 E\tilde{g}_{i+\ell}^2 + \frac{1}{n^2} \sum_{i=1}^{n-\ell} \sum_{j=1, j \neq i}^{n-\ell} E\tilde{g}_i E\tilde{g}_j E\tilde{g}_{i+\ell} E\tilde{g}_{j+\ell} \\ &= \frac{1}{n^2} \sum_{i=1}^{n-\ell} E\tilde{g}_i^2 E\tilde{g}_{i+\ell}^2 \\ &= \frac{1}{n} \left(1 - \frac{\ell}{n}\right) [Var(\tilde{g}_i)]^2 \\ &= \frac{1}{n} \left(1 - \frac{\ell}{n}\right) C \Delta^{2(1-\beta\varpi)} \leq \frac{1}{n} C \Delta^{2(1-\beta\varpi)} \end{aligned} \tag{5}$$

for some constant C independent of i (by stationarity). The second-to-last step is by lemma 4. Since the upper bound is independent of ℓ , we see that $(I) = \tilde{\gamma}(\ell)$ is $O_p(\frac{1}{\sqrt{n}}\Delta^{1-\beta\varpi})$, uniformly over ℓ . We thus obtain the decomposition

$$\hat{\gamma}(\ell) = \tilde{\gamma}(\ell) + o_p\left(\frac{\Delta^{1-\beta\varpi}}{n^{1/2}}\right), \quad (6)$$

i.e., $\hat{\gamma}(\ell) \simeq \tilde{\gamma}(\ell)$, uniformly over ℓ .

Consequently, the L statistic can be approximated by $\tilde{L}$, where

$$\tilde{L} := n \sum_{\ell=1}^{n-1} w_h(\ell) [\tilde{\gamma}(\ell)]^2.$$

Formally, $L \simeq \tilde{L}$. From now on, it suffices to derive the asymptotic distribution of $\tilde{L}$ under the null hypothesis. Let us rewrite $\tilde{L}$ as follows:

$$\begin{aligned} \tilde{L} &= n \sum_{\ell=1}^{n-1} w_h(\ell) \left[\frac{1}{n} \sum_{i=1}^{n-\ell} \tilde{g}_i \tilde{g}_{i+\ell} \right]^2 \\ &= \frac{1}{n} \sum_{\ell=1}^{n-1} w_h(\ell) \sum_{i=1}^{n-\ell} \tilde{g}_i^2 \tilde{g}_{i+\ell}^2 + \frac{1}{n} \sum_{\ell=1}^{n-1} w_h(\ell) \sum_{i=1}^{n-\ell} \sum_{j \neq i} \tilde{g}_i \tilde{g}_j \tilde{g}_{i+\ell} \tilde{g}_{j+\ell}. \end{aligned}$$

First, we compute the mean of $\tilde{L}$ under Assumption D_0 . By the stationarity of the jump increment process, we have $E(\tilde{g}_i \tilde{g}_j) = E(\tilde{g}_i)E(\tilde{g}_j) = 0$ and $E(\tilde{g}_i^2 \tilde{g}_j^2) \simeq E(\tilde{g}_i^2)E(\tilde{g}_j^2) = [Var(\tilde{g}_i)]^2$ for $i \neq j$ (by Assumption D_0 and lemma 1.8). This implies that

$$\begin{aligned} E(\tilde{L}) &= \frac{1}{n} \sum_{\ell=1}^{n-1} w_h(\ell) \sum_{i=1}^{n-\ell} E(\tilde{g}_i^2) E(\tilde{g}_{i+\ell}^2) \\ &= [Var(\tilde{g}_i)]^2 \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) w_h(\ell). \end{aligned}$$

To prepare for the variance computation, we obtain the demeaned form of $\tilde{L}$:

$$\begin{aligned} &\tilde{L} - E(\tilde{L}) \\ &= \frac{1}{n} \sum_{\ell=1}^{n-1} w_h(\ell) \sum_{i=1}^{n-\ell} [\tilde{g}_i^2 - E(\tilde{g}_i^2)][\tilde{g}_{i+\ell}^2 - E(\tilde{g}_{i+\ell}^2)] + \frac{1}{n} \sum_{\ell=1}^{n-1} w_h(\ell) \sum_{i=1}^{n-\ell} \sum_{j \neq i} \tilde{g}_i \tilde{g}_j \tilde{g}_{i+\ell} \tilde{g}_{j+\ell} \\ &\quad + \frac{1}{n} \sum_{\ell=1}^{n-1} w_h(\ell) \sum_{i=1}^{n-\ell} [\tilde{g}_i^2 - E(\tilde{g}_i^2)] E(\tilde{g}_{i+\ell}^2) + \frac{1}{n} \sum_{\ell=1}^{n-1} w_h(\ell) \sum_{i=1}^{n-\ell} [\tilde{g}_{i+\ell}^2 - E(\tilde{g}_{i+\ell}^2)] E(\tilde{g}_i^2) \\ &=: A + B + C + D. \end{aligned} \quad (7)$$

By the null hypothesis, which implies zero autocorrelation of $\tilde{g}_i^2 - E(\tilde{g}_i^2)$, we have $E\{\tilde{g}_i^2 -$

$E(\tilde{g}_i^2)[\tilde{g}_j^2 - E(\tilde{g}_j^2)] = 0$ for $i \neq j$. Furthermore, by stationarity, $Var(\tilde{g}_i^2) = E\{[\tilde{g}_i^2 - E(\tilde{g}_i^2)]^2\}$ for all j .

We first compute $Var(A) = E(A^2)$:

$$\begin{aligned} E(A^2) &\simeq \frac{1}{n^2} \sum_{\ell=1}^{n-1} w_h^2(\ell) \sum_{i=1}^{n-\ell} E\{[\tilde{g}_i^2 - E(\tilde{g}_i^2)]^2\} E\{[\tilde{g}_{i+\ell}^2 - E(\tilde{g}_i^2)]^2\} \\ &= [Var(\tilde{g}_i^2)]^2 \frac{1}{n} \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) w_h^2(\ell), \end{aligned} \quad (8)$$

where the first line follows from Assumption D₀ and lemma 1.8. To compute $Var(B) = E(B^2)$, we expand B^2 :

$$B^2 = \frac{1}{n^2} \sum_{\ell_1=1}^{n-1} \sum_{\ell_2=1}^{n-1} w_h(\ell_1) w_h(\ell_2) \sum_{i_1=1}^{n-\ell_1} \sum_{j_1 \neq i_1}^{n-\ell_1} \sum_{i_2=1}^{n-\ell_2} \sum_{j_2 \neq i_2}^{n-\ell_2} \tilde{g}_{i_1} \tilde{g}_{j_1} \tilde{g}_{i_1+\ell_1} \tilde{g}_{j_1+\ell_1} \tilde{g}_{i_2} \tilde{g}_{j_2} \tilde{g}_{i_2+\ell_2} \tilde{g}_{j_2+\ell_2}.$$

We observe that the expectation of the summand is non-zero when: (1) $i_1 = i_2 \neq j_1 = j_2$, $\ell_1 = \ell_2$, and (2) $i_1 = j_2 \neq i_2 = j_1$, $\ell_1 = \ell_2$. Since the two cases are symmetrical, we obtain

$$\begin{aligned} E(B^2) &\simeq \frac{2}{n^2} \sum_{\ell=1}^{n-1} w_h^2(\ell) \sum_{i=1}^{n-\ell} E(\tilde{g}_i^2) E(\tilde{g}_j^2) E(\tilde{g}_{i+\ell}^2) E(\tilde{g}_{j+\ell}^2) \\ &= 2[Var(\tilde{g}_i)]^4 \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) \left(1 - \frac{\ell+1}{n}\right) w_h^2(\ell). \end{aligned} \quad (9)$$

The first line follows from Assumption D₀ and lemma 1.8. Next we compute $Var(C) = E(C^2)$:

$$\begin{aligned} E(C^2) &= \frac{1}{n^2} \sum_{\ell=1}^{n-1} w_h^2(\ell) \sum_{i=1}^{n-\ell} E\{[\tilde{g}_i^2 - E(\tilde{g}_i^2)]^2\} E(\tilde{g}_i^2) \\ &= \frac{1}{n} Var(\tilde{g}_i^2) [Var(\tilde{g}_i)]^2 \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) w_h^2(\ell). \end{aligned}$$

By symmetry, $Var(D)$ is equal to $Var(C)$. By similar arguments, we find that $E(AB) = E(AC) = E(AD) = E(BC) = E(BD) = 0$.

We want to compare the stochastic orders of the terms in the decomposition (7). We deduce from Assumption G that $Var(\tilde{g}_i^2) = O(\Delta^{1-\beta\varpi})$, and we have $\Delta = O(n^{-\nu})$ by As-

sumption T_L . It follows from lemma 5 that

$$A = O_p \left(\frac{\Delta^{1-\beta\varpi}}{\sqrt{nh}} \right) = O_p \left(\frac{1}{n^{1/2+\nu(1-\beta\varpi)} h^{1/2}} \right), \quad (10)$$

$$B = O_p \left(\frac{\Delta^{2(1-\beta\varpi)}}{h^{1/2}} \right) = O_p \left(\frac{1}{n^{2\nu(1-\beta\varpi)} h^{1/2}} \right), \quad (11)$$

$$C = O_p \left(\frac{\Delta^{1.5(1-\beta\varpi)}}{\sqrt{nh}} \right) = O_p \left(\frac{1}{n^{1/2+1.5\nu(1-\beta\varpi)} h^{1/2}} \right),$$

$$D = O_p \left(\frac{1}{n^{1/2+1.5\nu(1-\beta\varpi)} h^{1/2}} \right).$$

Which term is dominant depends on ν . If $\nu = 0$ (fixed intervals), then the dominant term is B , which is $O_p(h^{-1/2})$. If $\nu \in (0, 1]$ (shrinking intervals), then the dominant term is A . The terms C and D are strictly dominated in all cases. It therefore suffices to just consider the terms A and B when approximating the variance of $\tilde{L}$, given by

$$\begin{aligned} & \text{Var}(\tilde{L}) \\ &= E[(\tilde{L} - E(\tilde{L}))^2] \\ &= E(A^2) + E(B^2) + o \left(\min \left(\frac{1}{n^{1+2\nu(1-\beta\varpi)} h}, \frac{1}{n^{4\nu(1-\beta\varpi)} h} \right) \right) \\ &\simeq [\text{Var}(\tilde{g}_i^2)]^2 \frac{1}{n} \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n} \right) w_h^2(\ell) + 2[\text{Var}(\tilde{g}_i)]^4 \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n} \right) \left(1 - \frac{\ell+1}{n} \right) w_h^2(\ell) \\ &\quad + o \left(\min \left(\frac{1}{n^{1+2\nu(1-\beta\varpi)} h}, \frac{1}{n^{4\nu(1-\beta\varpi)} h} \right) \right). \end{aligned} \quad (12)$$

To establish asymptotic normality of $\tilde{L}$, we proceed as follows:

(i) Express $\tilde{L} - E(\tilde{L})$ as $\sum_{i=1}^n Y_i$, where $\{Y_i\}_{i=1}^n$ forms a martingale difference sequence with respect to the natural filtration¹ $\{\mathcal{F}_{ni}\}_{i=1}^n$ generated by the process $\{\tilde{g}_i\}_{i=1}^n$.

(ii) Show asymptotic negligibility²: for any given $\eta > 0$, $s^{-2} \sum_{i=1}^n E(Y_i^2 1_{\{Y_i \geq s\eta\}}) \rightarrow 0$ as $n \rightarrow \infty$, where $s^2 := \text{Var}(\tilde{L}) = E[(\sum_{i=1}^n Y_i^2)]$.

(iii) Show that $s^{-2} \sum_{i=1}^n Y_i^2 \xrightarrow{p} 1$ as $n \rightarrow \infty$.³

(iv) Apply the martingale central limit theorem.

To show (i), we first rewrite $\tilde{L} - E(\tilde{L})$ by switching the order of summations (the sum over ℓ and those over i and j) in (7) and then relabeling the subscripts (replace i by $i - \ell$,

¹The filtration $\{\mathcal{F}_{ni}\}$ is not nested as n increases, unless Δ is fixed (or $\nu = 0$). But since the probability limit in (iii) is a constant, the martingale central limit theorem is applicable. See the remark after Corollary 3.1 on p.59 of Hall and Heyde (1980).

²Here we show the unconditional Lindeberg condition; see the discussion on asymptotic negligibility on p.53-54 of Hall and Heyde (1980).

³See (3.10) and the preceding discussion on p.55 of Hall and Heyde (1980).

and j by $j - \ell$), yielding

$$\begin{aligned}
& \tilde{L} - E(\tilde{L}) \\
&= \sum_{i=2}^n \frac{1}{n} \sum_{\ell=1}^{i-1} w_h(\ell) [\tilde{g}_{i-\ell}^2 - E(\tilde{g}_i^2)] [\tilde{g}_i^2 - E(\tilde{g}_i^2)] + \sum_{i=3}^n \frac{2}{n} \sum_{j=2}^{i-1} \sum_{\ell=1}^{j-1} w_h(\ell) \tilde{g}_{j-\ell} \tilde{g}_{i-\ell} \tilde{g}_j \tilde{g}_i \\
&+ \sum_{i=2}^n \frac{1}{n} \sum_{\ell=1}^{i-1} w_h(\ell) [\tilde{g}_{i-\ell}^2 - E(\tilde{g}_i^2)] E(\tilde{g}_i^2) + \sum_{i=2}^n \frac{1}{n} \sum_{\ell=1}^{i-1} w_h(\ell) [\tilde{g}_i^2 - E(\tilde{g}_i^2)] E(\tilde{g}_i^2).
\end{aligned}$$

By the tower property of conditional expectations, it is straightforward to see that Y_i , defined by $Y_1 = 0$ and, for $i = 2, 3, \dots, n$,

$$\begin{aligned}
Y_i &= \frac{1}{n} \sum_{\ell=1}^{i-1} w_h(\ell) [\tilde{g}_{i-\ell}^2 - E(\tilde{g}_i^2)] [\tilde{g}_i^2 - E(\tilde{g}_i^2)] + 1_{\{i>2\}} \frac{2}{n} \sum_{j=2}^{i-1} \sum_{\ell=1}^{j-1} w_h(\ell) \tilde{g}_{j-\ell} \tilde{g}_{i-\ell} \tilde{g}_j \tilde{g}_i \\
&+ \frac{1}{n} E(\tilde{g}_i^2) \sum_{\ell=1}^{i-1} w_h(\ell) [\tilde{g}_{i-\ell}^2 - E(\tilde{g}_i^2)] + \frac{1}{n} E(\tilde{g}_i^2) \sum_{\ell=1}^{i-1} w_h(\ell) [\tilde{g}_i^2 - E(\tilde{g}_i^2)] \\
&=: Y_{1i} + Y_{2i} + Y_{3i} + Y_{4i},
\end{aligned} \tag{13}$$

is a martingale difference sequence so that Y_i is measurable with respect to $\mathcal{F}_{ni}$ for all $i = 1, 2, \dots, n$ and $E(Y_i | \mathcal{F}_{nj}) = 0$ for all $i = 2, 3, \dots, n$ and $j < i$.

To show (ii), we first note that for any given $\eta > 0$ and $\epsilon > 0$,

$$s^{-2} \sum_{i=1}^n E(Y_i^2 1_{\{Y_i \geq s\eta\}}) = s^{-2} \sum_{i=1}^n E(Y_i^2 1_{\{\eta^{-\epsilon} (\frac{Y_i}{s})^\epsilon \geq 1\}}) \leq \eta^{-\epsilon} s^{-(2+\epsilon)} \sum_{i=1}^n E(Y_i^{2+\epsilon}).$$

It therefore suffices to show (e.g., by setting $\epsilon = 1$) that $s^{-3} \sum_{i=1}^n E(Y_i^3) \rightarrow 0$. Next, by noting that the cross moments are all zero (as they must involve zero-mean factors such as $\tilde{g}_i^2 - E(\tilde{g}_i^2)$ or $\tilde{g}_i$), we obtain

$$\sum_{i=1}^n E(Y_i^3) = \sum_{i=1}^n E(Y_{1i}^3) + \sum_{i=1}^n E(Y_{2i}^3) + \sum_{i=1}^n E(Y_{3i}^3) + \sum_{i=1}^n E(Y_{4i}^3).$$

Its non-zero dominating terms are given as follows:

$$\begin{aligned}
\sum_{i=1}^n E(Y_{1i}^3) &\simeq \sum_{i=2}^n \frac{1}{n^3} \sum_{\ell=1}^{i-1} w_h^3(\ell) E[(\tilde{g}_{i-\ell}^2 - E(\tilde{g}_i^2))^3] E[(\tilde{g}_i^2 - E(\tilde{g}_i^2))^3] \\
&= \frac{1}{n^3} \{E[(\tilde{g}_i^2 - E(\tilde{g}_i^2))^3]\}^2 \sum_{\ell=1}^{n-1} \sum_{i=\ell+1}^n w_h^3(\ell) \\
&= \frac{1}{n^2} \{E[(\tilde{g}_i^2 - E(\tilde{g}_i^2))^3]\}^2 \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) w_h^3(\ell) \\
&= O\left(\frac{\Delta^{2(1-\beta\varpi)}}{n^2 h^2}\right),
\end{aligned}$$

$$\begin{aligned}
\sum_{i=1}^n E(Y_{2i}^3) &\simeq \sum_{i=3}^n \frac{8}{n^3} \sum_{j=2}^{i-1} \sum_{\ell=1}^{j-1} w_h^3(\ell) E(\tilde{g}_{j-\ell}^3) E(\tilde{g}_{i-\ell}^3) E(\tilde{g}_j^3) E(\tilde{g}_i^3) \\
&= \frac{8}{n^3} [E(\tilde{g}_i^3)]^4 \sum_{\ell=1}^{n-1} \sum_{i=\ell+2}^n \sum_{j=\ell+1}^{i-1} w_h^3(\ell) \\
&= \frac{4}{n} [E(\tilde{g}_i^3)]^4 \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) \left(1 - \frac{\ell+1}{n}\right) w_h^3(\ell) \\
&= O\left(\frac{\Delta^{4(1-\beta\varpi)}}{n h^2}\right),
\end{aligned}$$

$$\begin{aligned}
\sum_{i=1}^n E(Y_{3i}^3) &= [E(\tilde{g}_i^2)]^3 \sum_{i=2}^n \frac{1}{n^3} \sum_{\ell=1}^{i-1} w_h^3(\ell) E[(\tilde{g}_{i-\ell}^2 - E(\tilde{g}_i^2))^3] \\
&= \frac{1}{n^3} [E(\tilde{g}_i^2)]^3 E[(\tilde{g}_i^2 - E(\tilde{g}_i^2))^3] \sum_{\ell=1}^{n-1} \sum_{i=\ell+1}^n w_h^3(\ell) \\
&= \frac{1}{n^2} [E(\tilde{g}_i^2)]^3 E[(\tilde{g}_i^2 - E(\tilde{g}_i^2))^3] \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) w_h^3(\ell) \\
&= O\left(\frac{\Delta^{4(1-\beta\varpi)}}{n^2 h^2}\right),
\end{aligned}$$

and $\sum_{i=1}^n E(Y_{4i}^3)$ is the same as $\sum_{i=1}^n E(Y_{3i}^3)$ by symmetry. We thus see that both $\sum_{i=1}^n E(Y_{3i}^3)$ and $\sum_{i=1}^n E(Y_{4i}^3)$ are strictly dominated by $\sum_{i=1}^n E(Y_{2i}^3)$. Using the fact that $\Delta = O(n^{-\nu})$, we observe that the dominant terms of $\sum_{i=1}^n E(Y_i^3)$ are

$$\sum_{i=1}^n E(Y_{1i}^3) + \sum_{i=1}^n E(Y_{2i}^3) = O\left(\frac{1}{n^{2+2\nu(1-\beta\varpi)} h^2}\right) + O\left(\frac{1}{n^{1+4\nu(1-\beta\varpi)} h^2}\right). \quad (14)$$

We consider two separate cases:

(1) Under the condition that $0 \leq \nu < \frac{1}{2(1-\beta\varpi)}$, we deduce from (10)-(11) and (14) that $\sum_{i=1}^n E(Y_{2i}^3) \gg \sum_{i=1}^n E(Y_{1i}^3)$ and $E(B^2) \gg E(A^2)$, and hence

$$\begin{aligned} s^{-3} \sum_{i=1}^n E(Y_i^3) &= \frac{\sum_{i=1}^n E(Y_{2i}^3)}{[E(B^2)]^{3/2}} + o(1) = \frac{O\left(\frac{1}{n^{1+4\nu(1-\alpha\varpi)}h^2}\right)}{O\left(\frac{1}{n^{6\nu(1-\alpha\varpi)}h^{3/2}}\right)} \\ &= O\left(\frac{1}{n^{1-2\nu(1-\alpha\varpi)}h^{1/2}}\right) = O\left(\frac{1}{n^{1-2\nu(1-\alpha\varpi)+\xi/2}}\right), \end{aligned}$$

which tends to zero as $n \rightarrow \infty$ for all ξ .

(2) Under the condition that $\frac{1}{2(1-\beta\varpi)} \leq \nu \leq 1$, we have $\sum_{i=1}^n E(Y_{1i}^3) \gg \sum_{i=1}^n E(Y_{2i}^3)$ and $E(A^2) \gg E(B^2)$, and hence

$$\begin{aligned} s^{-3} \sum_{i=1}^n E(Y_i^3) &= \frac{\sum_{i=1}^n E(Y_{1i}^3)}{[E(A^2)]^{3/2}} + o(1) = \frac{O\left(\frac{1}{n^{2+2\nu(1-\alpha\varpi)}h^2}\right)}{O\left(\frac{1}{n^{3/2+3\nu(1-\alpha\varpi)}h^{3/2}}\right)} \\ &= O\left(\frac{1}{n^{1/2-\nu(1-\alpha\varpi)}h^{1/2}}\right) = O\left(\frac{1}{n^{1/2-\nu(1-\alpha\varpi)+\xi/2}}\right), \end{aligned}$$

which tends to zero as $n \rightarrow \infty$ when $\frac{1}{2} - \nu(1 - \alpha\varpi) + \frac{\xi}{2} > 0$, or when $\xi > 2\nu(1 - \alpha\varpi) - 1$. The latter condition on ξ is guaranteed by Assumption T_L.

Let us proceed to (iii). To verify $s^{-2} \sum_{i=1}^n Y_i^2 \xrightarrow{p} 1$ as $n \rightarrow \infty$, it suffices to show that $E[(s^{-2} \sum_{i=1}^n Y_i^2 - 1)^2] \rightarrow 0$ as $n \rightarrow \infty$ (because L^2 -convergence implies convergence in probability).

To start, we note that

$$E[(s^{-2} \sum_{i=1}^n Y_i^2 - 1)^2] = s^{-4} E[(\sum_{i=1}^n Y_i^2)^2] - 2s^{-2} \sum_{i=1}^n E(Y_i^2) + 1 = s^{-4} E[(\sum_{i=1}^n Y_i^2)^2] - 1,$$

so it suffices to show that $s^{-4} E[(\sum_{i=1}^n Y_i^2)^2] \rightarrow 1$. The expectation is decomposed as:

$$E[(\sum_{i=1}^n Y_i^2)^2] = \sum_{i=1}^n E(Y_i^4) + \sum_{i=1}^n \sum_{i=1, j \neq i}^n E(Y_i^2 Y_j^2) =: M_1 + M_2.$$

We claim that the dominant term is M_2 .

Let us now determine the order of M_1 . Recalling the decomposition (13) of Y_i , and noting that Y_{1i} and Y_{2i} are the dominant terms (i.e., $Y_i \sim Y_{1i} + Y_{2i}$) and have mean zero, the

dominant terms of $M_1 := \sum_{i=1}^n E(Y_i^4)$ are given by

$$\begin{aligned} M_1 &\sim \sum_{i=1}^n E(Y_{1i}^4) + \sum_{i=1}^n E(Y_{2i}^4) + 6 \sum_{i=1}^n E(Y_{1i}^2 Y_{2i}^2). \\ &=: M_{11} + M_{12} + M_{13}. \end{aligned}$$

Now, we note that $M_{11} := M_{111} + M_{112}$, where

$$\begin{aligned} M_{111} &\simeq \sum_{i=1}^n \frac{1}{n^4} \sum_{\ell=1}^{i-1} w_h^4(\ell) E[(\tilde{g}_{i-\ell}^2 - E(\tilde{g}_i^2))^4] E[(\tilde{g}_i^2 - E(\tilde{g}_i^2))^4] \\ &= \frac{1}{n^3} \{E[(\tilde{g}_i^2 - E(\tilde{g}_i^2))^4]\}^2 \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) w_h^4(\ell) \\ &= O\left(\frac{\Delta^{2(1-\beta\varpi)}}{n^3 h^3}\right), \end{aligned}$$

and

$$\begin{aligned} M_{112} &\simeq \sum_{i=1}^n \frac{1}{n^4} \sum_{\ell_1=1}^{i-1} \sum_{\ell_2=1, \ell_2 \neq \ell_1}^{i-1} w_h^2(\ell_1) w_h^2(\ell_2) E[(\tilde{g}_{i-\ell_1}^2 - E(\tilde{g}_i^2))^2] \times \\ &\quad E[(\tilde{g}_{i-\ell_2}^2 - E(\tilde{g}_i^2))^2] E[(\tilde{g}_i^2 - E(\tilde{g}_i^2))^4] \\ &= \frac{1}{n^3} \{E[(\tilde{g}_i^2 - E(\tilde{g}_i^2))^2]\}^2 E[(\tilde{g}_i^2 - E(\tilde{g}_i^2))^4] \times \\ &\quad \sum_{\ell_1=1}^{n-1} \sum_{\ell_2=1, \ell_2 \neq \ell_1}^{n-1} \left(1 - \frac{\min(\ell_1, \ell_2)}{n}\right) w_h^2(\ell_1) w_h^2(\ell_2) \\ &= O\left(\frac{\Delta^{3(1-\beta\varpi)}}{n^3 h^2}\right). \end{aligned}$$

Next, we note that $M_{12} := M_{121} + M_{122} + M_{123} + M_{124}$, where

$$\begin{aligned} M_{121} &\simeq \sum_{i=1}^n \frac{1}{n^4} \sum_{j=1, j \neq i}^n \sum_{\ell=1}^{\max(i,j)-1} w_h^4(\ell) E(\tilde{g}_{i-\ell}^4) E(\tilde{g}_{j-\ell}^4) E(\tilde{g}_i^4) E(\tilde{g}_j^4) \\ &= \frac{1}{n^2} [E(\tilde{g}_i^4)]^4 \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) \left(1 - \frac{\ell+1}{n}\right) w_h^4(\ell) \\ &= O\left(\frac{\Delta^{4(1-\beta\varpi)}}{n^2 h^3}\right), \end{aligned}$$

$$\begin{aligned}
M_{122} &\simeq \sum_{i=1}^n \frac{1}{n^4} \sum_{j_1=1, j_1 \neq i}^n \sum_{j_2=1, j_2 \neq i, j_2 \neq j_1}^n \sum_{\ell_1=1}^{\max(i, j_1, j_2)-1} \sum_{\ell_2=1, \ell_2 \neq \ell_1}^{\max(i, j_1, j_2)-1} w_h^2(\ell_1) w_h^2(\ell_2) \times \\
&\quad E(\tilde{g}_{i-\ell_1}^2) E(\tilde{g}_{i-\ell_2}^2) E(\tilde{g}_{j_1-\ell_1}^2) E(\tilde{g}_{j_2-\ell_2}^2) E(\tilde{g}_{j_1}^2) E(\tilde{g}_{j_2}^2) E(\tilde{g}_i^4) \\
&= \frac{1}{n} [E(\tilde{g}_i^2)]^6 E(\tilde{g}_i^4) \sum_{\ell_1=1}^{n-1} \sum_{\ell_2=1}^{n-1} \left(1 - \frac{\ell_1}{n}\right) \left(1 - \frac{\ell_2}{n}\right) \left(1 - \frac{\ell_1 + \ell_2}{n}\right) w_h^2(\ell_1) w_h^2(\ell_2) \\
&= O\left(\frac{\Delta^{7(1-\beta\varpi)}}{nh^2}\right),
\end{aligned}$$

$$\begin{aligned}
M_{123} &\simeq \sum_{i=1}^n \frac{1}{n^4} \sum_{j_1=1, j_1 \neq i}^n \sum_{j_2=1, j_2 \neq i, j_2 \neq j_1}^n \sum_{\ell=1}^{\max(i, j_1, j_2)-1} w_h^4(\ell) \times \\
&\quad E(\tilde{g}_{i-\ell}^4) E(\tilde{g}_{j_1-\ell}^2) E(\tilde{g}_{j_2-\ell}^2) E(\tilde{g}_{j_1}^2) E(\tilde{g}_{j_2}^2) E(\tilde{g}_i^4) \\
&= \frac{1}{n} [E(\tilde{g}_i^2)]^4 [E(\tilde{g}_i^4)]^2 \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) \left(1 - \frac{\ell+1}{n}\right) \left(1 - \frac{\ell+2}{n}\right) w_h^4(\ell) \\
&= O\left(\frac{\Delta^{6(1-\beta\varpi)}}{nh^3}\right),
\end{aligned}$$

and

$$\begin{aligned}
M_{124} &\simeq \sum_{i=1}^n \frac{1}{n^4} \sum_{j=1, j \neq i}^n \sum_{\ell_1=1}^{\max(i, j)-1} \sum_{\ell_2=1, \ell_2 \neq \ell_1}^{\max(i, j)-1} w_h^2(\ell_1) w_h^2(\ell_2) \times \\
&\quad E(\tilde{g}_{i-\ell_1}^2) E(\tilde{g}_{i-\ell_2}^2) E(\tilde{g}_{j-\ell_1}^2) E(\tilde{g}_{j-\ell_2}^2) E(\tilde{g}_j^4) E(\tilde{g}_i^4) \\
&= \frac{1}{n^2} [E(\tilde{g}_i^2)]^4 [E(\tilde{g}_i^4)]^2 \sum_{\ell_1=1}^{n-1} \sum_{\ell_2=1, \ell_2 \neq \ell_1}^{n-1} \left(1 - \frac{\ell_1}{n}\right) \left(1 - \frac{\ell_2}{n}\right) w_h^2(\ell_1) w_h^2(\ell_2) \\
&= O\left(\frac{\Delta^{6(1-\beta\varpi)}}{n^2 h^2}\right).
\end{aligned}$$

We then turn to $M_{13} := M_{131} + M_{132} + M_{133}$, where

$$\begin{aligned}
M_{131} &= 6 \sum_{i=1}^n E \left\{ \frac{1}{n^2} \sum_{\ell=1}^{\max(i,j)-1} w_h^2(\ell) (\tilde{g}_{i-\ell}^2 - E(\tilde{g}_i^2))^2 (\tilde{g}_i^2 - E(\tilde{g}_i^2))^2 \times \right. \\
&\quad \left. \frac{1}{n^2} w_h^2(\ell) \left(\sum_{j=1, j \neq i, j \neq i-\ell}^n \tilde{g}_{i-\ell}^2 \tilde{g}_{j-\ell}^2 \tilde{g}_i^2 \tilde{g}_j^2 + \tilde{g}_{i-2\ell}^2 \tilde{g}_i^2 \tilde{g}_{i-\ell}^4 \right) \right\} \\
&\simeq \frac{6}{n^2} \{E[\tilde{g}_i^2 (\tilde{g}_i^2 - E(\tilde{g}_i^2))^2]\}^2 [E(\tilde{g}_i^2)]^2 \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) \left(1 - \frac{\ell+1}{n}\right) w_h^4(\ell) \\
&\quad + \frac{6}{n^3} E[\tilde{g}_i^2 (\tilde{g}_i^2 - E(\tilde{g}_i^2))^2] E[\tilde{g}_i^4 (\tilde{g}_i^2 - E(\tilde{g}_i^2))^2] E(\tilde{g}_i^2) \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) w_h^4(\ell) \\
&= O\left(\frac{\Delta^{4(1-\beta\varpi)}}{n^2 h^3}\right) + O\left(\frac{\Delta^{3(1-\beta\varpi)}}{n^3 h^3}\right),
\end{aligned}$$

$$\begin{aligned}
M_{132} &= 6 \sum_{i=1}^n E \left\{ \frac{1}{n^2} \sum_{\ell_1=1}^{\max(i,j)-1} w_h^2(\ell_1) (\tilde{g}_{i-\ell_1}^2 - E(\tilde{g}_i^2))^2 (\tilde{g}_i^2 - E(\tilde{g}_i^2))^2 \times \right. \\
&\quad \left. \frac{1}{n^2} \sum_{j=1, j \neq i, j \neq i-\ell_1}^n \sum_{\ell_2=1, \ell_2 \neq \ell_1}^{\max(i,j)-1} w_h^2(\ell_2) \tilde{g}_{i-\ell_2}^2 \tilde{g}_{j-\ell_2}^2 \tilde{g}_i^2 \tilde{g}_j^2 \right\} \\
&= \frac{6}{n^2} E[\tilde{g}_i^2 (\tilde{g}_i^2 - E(\tilde{g}_i^2))^2] E[(\tilde{g}_i^2 - E(\tilde{g}_i^2))^2] [E(\tilde{g}_i^2)]^3 \times \\
&\quad \sum_{\ell_1=1}^{n-1} \sum_{\ell_2=1, \ell_2 \neq \ell_1}^{n-1} \left(1 - \frac{\ell_1}{n}\right) \left(1 - \frac{\ell_2}{n}\right) w_h^2(\ell_1) w_h^2(\ell_2) \\
&= O\left(\frac{\Delta^{5(1-\beta\varpi)}}{n^2 h^2}\right),
\end{aligned}$$

and

$$\begin{aligned}
M_{133} &= 6 \sum_{i=1}^n E \left\{ \frac{1}{n^2} \sum_{\ell_1=1}^{i-1} w_h^2(\ell_1) (\tilde{g}_{i-\ell_1}^2 - E(\tilde{g}_i^2))^2 (\tilde{g}_i^2 - E(\tilde{g}_i^2))^2 \times \right. \\
&\quad \left. \frac{1}{n^2} \sum_{\ell_2=1, \ell_2 \neq \ell_1}^{i-1} w_h^2(\ell_2) \tilde{g}_{i-\ell_2}^2 \tilde{g}_{i-\ell_1-\ell_2}^2 \tilde{g}_i^2 \tilde{g}_{i-\ell_1}^2 \right\} \\
&\simeq \frac{6}{n^3} \{E[\tilde{g}_i^2(\tilde{g}_i^2 - E(\tilde{g}_i^2))^2]\}^2 [E(\tilde{g}_i^2)]^2 \times \\
&\quad \sum_{\ell_1=1}^{n-1} \sum_{\ell_2=1, \ell_2 \neq \ell_1}^{n-1} \left(1 - \frac{\max(\ell_1 + \ell_2)}{n}\right) w_h^2(\ell_1) w_h^2(\ell_2) \\
&= O\left(\frac{\Delta^{4(1-\beta\varpi)}}{n^3 h^2}\right).
\end{aligned}$$

Now, we evaluate M_2 . Recalling again that $Y_i^2 \sim Y_{1i}^2 + Y_{2i}^2 + 2Y_{1i}Y_{2i}$ (ignoring the dominated terms Y_{3i} and Y_{4i}), and noting that Y_{1i} and Y_{2i} have zero mean, we have

$$\begin{aligned}
M_2 &= \sum_{i=1}^n \sum_{j=1, j \neq i}^n E(Y_i^2 Y_j^2) \simeq \sum_{i=1}^n \sum_{j=1, j \neq i}^n [E(Y_{1i}^2 Y_{1j}^2) + E(Y_{2i}^2 Y_{2j}^2) + E(Y_{1i}^2 Y_{2j}^2)] \\
&= : M_{21} + M_{22} + M_{23}.
\end{aligned}$$

We evaluate each term as follows:

$$\begin{aligned}
M_{21} &= \sum_{i=1}^n \sum_{j=1, j \neq i}^n E(Y_{1i}^2 Y_{1j}^2) \\
&\simeq \frac{1}{n^4} \sum_{i=1}^n \sum_{j=1, j \neq i}^n \sum_{\ell_1=1}^{\max(i,j)-1} \sum_{\ell_2=1, \ell_2 \neq \ell_1}^{\max(i,j)-1} w_h^2(\ell_1) w_h^2(\ell_2) E[(\tilde{g}_{i-\ell_1}^2 - E(\tilde{g}_i^2))^2] \times \\
&\quad E[(\tilde{g}_i^2 - E(\tilde{g}_i^2))^2] E[(\tilde{g}_{j-\ell_2}^2 - E(\tilde{g}_j^2))^2] E[(\tilde{g}_j^2 - E(\tilde{g}_j^2))^2] \\
&= \frac{1}{n^2} \{E[(\tilde{g}_i^2 - E(\tilde{g}_i^2))^2]\}^4 \left[\sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) w_h^2(\ell) \right]^2 \\
&= O\left(\frac{\Delta^{4(1-\beta\varpi)}}{n^2 h^2}\right),
\end{aligned}$$

$$\begin{aligned}
M_{22} &= \sum_{i=1}^n \sum_{j=1, j \neq i}^n E(Y_{2i}^2 Y_{2j}^2) \\
&\simeq \frac{1}{n^4} \sum_{i=1}^n \sum_{j=1, j \neq i}^n \sum_{i'=1, i' \neq i}^n \sum_{j'=1, j' \neq j, j' \neq i'}^n \sum_{\ell=1}^{\max(i,j)-1} \sum_{\ell'=1, \ell' \neq \ell}^{\max(i',j')-1} w_h^2(\ell) w_h^2(\ell') \times \\
&\quad E(\tilde{g}_{i-\ell}^2) E(\tilde{g}_{j-\ell}^2) E(\tilde{g}_i^2) E(\tilde{g}_j^2) E(\tilde{g}_{i'-\ell'}^2) E(\tilde{g}_{j'-\ell'}^2) E(\tilde{g}_{i'}^2) E(\tilde{g}_{j'}^2) \\
&= 4[E(\tilde{g}_i^2)]^8 \left[\sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) \left(1 - \frac{\ell+1}{n}\right) w_h^2(\ell) \right]^2 \\
&= O\left(\frac{\Delta^{8(1-\beta\varpi)}}{h^2}\right),
\end{aligned}$$

and

$$\begin{aligned}
M_{23} &= \sum_{i=1}^n \sum_{j=1, j \neq i}^n E(Y_{1i}^2 Y_{2j}^2) \\
&\simeq \frac{1}{n^4} \sum_{i=1}^n \sum_{j=1, j \neq i}^n \sum_{j'=1, j' \neq j, j' \neq i}^n \sum_{\ell=1}^{\max(i,j)-1} \sum_{\ell'=1, \ell' \neq \ell}^{\max(i,j')-1} w_h^2(\ell_1) w_h^2(\ell_2) \times \\
&\quad E[(\tilde{g}_{i-\ell}^2 - E(\tilde{g}_i^2))^2] E[(\tilde{g}_i^2 - E(\tilde{g}_i^2))^2] E(\tilde{g}_{j-\ell'}^2) E(\tilde{g}_{j'-\ell'}^2) E(\tilde{g}_j^2) E(\tilde{g}_{j'}^2) \\
&= \frac{4}{n} \{E[(\tilde{g}_i^2 - E(\tilde{g}_i^2))^2]\}^2 [E(\tilde{g}_i^2)]^4 \times \\
&\quad \sum_{\ell=1}^{n-1} \sum_{\ell'=1, \ell' \neq \ell}^{n-1} \left(1 - \frac{\ell}{n}\right) \left(1 - \frac{\ell'}{n}\right) \left(1 - \frac{\ell+\ell'}{n}\right) w_h^2(\ell) w_h^2(\ell') \\
&= O\left(\frac{\Delta^{6(1-\beta\varpi)}}{nh^2}\right).
\end{aligned}$$

Comparing the terms M_{1i} ($i = 1, 2, 3$) with M_{2j} ($j = 1, 2, 3$), we observe the following:

(1) When $\nu = 0$ (fixed Δ), we have $M_1 \sim M_{122} = O\left(\frac{\Delta^{7(1-\beta\varpi)}}{nh^2}\right)$, which is dominated by $M_2 \sim M_{22} = O\left(\frac{\Delta^{8(1-\beta\varpi)}}{h^2}\right)$.

(2) When $\nu \in (0, 1]$ (shrinking Δ), we have $M_1 \sim M_{111} + M_{112} = O\left(\frac{\Delta^{2(1-\beta\varpi)}}{n^3 h^3}\right) + O\left(\frac{\Delta^{3(1-\beta\varpi)}}{n^3 h^2}\right)$, while $M_2 \sim M_{21} = O\left(\frac{\Delta^{4(1-\beta\varpi)}}{n^2 h^2}\right)$. Note that

$$\frac{M_{111}}{M_{21}} = O\left(\frac{\Delta^{2(1-\beta\varpi)}/n^3 h^3}{\Delta^{4(1-\beta\varpi)}/n^2 h^2}\right) = O\left(\frac{1}{\Delta^{2(1-\beta\varpi)} n h}\right) = O\left(\frac{1}{n^{-2\nu(1-\beta\varpi)+1+\xi}}\right),$$

which is $o(1)$ because $\xi > \max(2\nu(1-\beta\varpi) - 1, 0)$ by Assumption T_L . Also note that

$$\frac{M_{112}}{M_{21}} = O\left(\frac{\Delta^{3(1-\beta\varpi)}/n^3 h^2}{\Delta^{4(1-\beta\varpi)}/n^2 h^2}\right) = O\left(\frac{1}{\Delta^{1-\beta\varpi} n}\right) = O\left(\frac{1}{n^{-\nu(1-\beta\varpi)+1}}\right),$$

which is $o(1)$ because $\beta < 2$ and $\varpi < \frac{1}{2}$ imply that $1 - \nu(1 - \beta\varpi) > 0$.

It follows that $\frac{M_1}{M_2} = o(1)$, or M_2 dominates M_1 in the limit.

On the other hand, we note from (12) that $s^4 = [Var(\tilde{L})]^2 \sim M_{21} + M_{22} + M_{23}$. We therefore conclude that $s^{-4}E[(\sum_{i=1}^n Y_i^2)^2] \rightarrow 1$, and complete the verification of (iii).

(iv) Finally, we apply the martingale central limit theorem to conclude that $\tilde{L}$ converges (stably) in law to a normal distribution. Since the probability limit of $s^{-2} \sum_{i=1}^n E(Y_i^2 | \mathcal{F}_{n,i-1})$ is a constant, which is trivially measurable with respect to all $\mathcal{F}_{ni}$ for all n , we can dispense with the requirement that the filtration $\mathcal{F}_{ni}$ be nested as n increases. See Corollary 3.1 of Hall and Heyde (1980) and the entailing remarks.

To complete the proof, we note that $\tilde{g}_{k,i}$ is a bounded and uncorrelated sequence under the null hypothesis, and so we have $\widehat{Var}_0(g_{k,i}) \xrightarrow{p} Var(g_{k,i})$. In addition, by the definition of $g_{k,i}$, we observe that $E[(\tilde{g}_{k,i}^2 - E(\tilde{g}_{k,i}^2))^2] = E(\tilde{g}_{k,i}^4) = Var(g_{2k,i})$. By the same token, we have $\widehat{Var}_0(g_{2k,i}) \xrightarrow{p} Var(g_{2k,i}) = E[(\tilde{g}_{k,i}^2 - E(\tilde{g}_{k,i}^2))^2]$. The extra term $r\Delta^p$ in the variance estimator is strictly dominated in the limit and does not affect the asymptotic result. The consistency of $\hat{E}_0(L)$ for $E(L)$ and $\widehat{Var}_0(L)$ for $Var(L)$ immediately follows. The stated result is obtained by rewriting the studentized L statistic as

$$\frac{L - \hat{E}_0(L)}{\sqrt{\widehat{Var}_0(L)}} = \sqrt{\frac{Var(L)}{\widehat{Var}_0(L)}} \left(\frac{L - E(L)}{\sqrt{Var(L)}} + \frac{E(L) - \hat{E}_0(L)}{\sqrt{Var(L)}} \right),$$

and then passing to the limit.

1.7 Proof of Theorem 2

We decompose the normalized L statistic as follows:

$$\begin{aligned} t_L &= \frac{L - \hat{E}_0(L)}{\sqrt{\widehat{Var}_0(L)}} = \sqrt{\frac{Var_0(\tilde{L})}{\widehat{Var}_0(L)}} \left[\frac{\tilde{L} - E(\tilde{L})}{\sqrt{Var_0(\tilde{L})}} + \frac{E(\tilde{L}) - E_0(\tilde{L})}{\sqrt{Var_0(\tilde{L})}} + \frac{E_0(\tilde{L}) - \hat{E}_0(L)}{\sqrt{Var_0(\tilde{L})}} \right] \\ &= (1 + o_p(1)) \left[\frac{\tilde{L} - E(\tilde{L})}{\sqrt{Var_0(\tilde{L})}} + \frac{E(\tilde{L}) - E_0(\tilde{L})}{\sqrt{Var_0(\tilde{L})}} + o_p(1) \right], \\ &=: (1 + o_p(1)) [F_1 + F_2 + o_p(1)], \end{aligned} \tag{15}$$

where the last step is obtained by the consistency of $\hat{E}_0(L)$ and $\widehat{Var}_0(L)$ as proved in Theorem 3. The subscript zero indicates that the population moments are evaluated with respect to the DGP under the null hypothesis. We will complete the proof if we show that:

- (i) F_1 is $O_p(1)$.
- (ii) $F_2 = \theta_k + o(1)$ for $\delta = \frac{1}{2}$; and $F_2 = O(h^{0.5-\delta})$ for $0 \leq \delta < \frac{1}{2}$.

We start by proving (ii). Using the decomposition in (7), we have

$$F_2 := \frac{E(\tilde{L}) - E_0(\tilde{L})}{\sqrt{Var_0(\tilde{L})}} = \frac{E(A + B + C + D)}{\sqrt{Var_0(\tilde{L})}}.$$

We recall from Theorem 3 that

$$Var_0(\tilde{L}) = \frac{2}{h} \mathcal{W}_2[Var_0(g_{n,k,i})]^4 = O\left(\frac{\Delta^{4(1-\beta\varpi)}}{h}\right) \text{ for } \nu = 0; \quad (16)$$

$$Var_0(\tilde{L}) = \frac{1}{nh} \mathcal{W}_2[Var_0(g_{n,2k,i})]^2 = O\left(\frac{\Delta^{2(1-\beta\varpi)}}{nh}\right) \text{ for } \nu > 0, \quad (17)$$

where $\mathcal{W}_2 := \int_0^\infty w^2(x)dx$. Now we take expectations with respect to the DGP under the alternatives in Assumption $D_a(i)$. First, it is readily seen that $E(C) = E(D) = 0$. The expected value of A is given by

$$\begin{aligned} E(A) &= \frac{1}{n} \sum_{\ell=1}^{n-1} w_h(\ell) \sum_{i=1}^{n-\ell} E[(\tilde{g}_i^2 - E(\tilde{g}_i^2))(\tilde{g}_{i+\ell}^2 - E(\tilde{g}_{i+\ell}^2))] \\ &= \frac{1}{n} \sum_{i=1}^{n-1} \sum_{\ell=1}^{n-i} w_h(\ell) Cov(\tilde{g}_i^2, \tilde{g}_{i+\ell}^2) \\ &= \frac{1}{n} \sum_{i=1}^{n-1} \sum_{\ell=1}^{n-i} w_h(\ell) cum(\tilde{g}_i, \tilde{g}_i, \tilde{g}_{i+\ell}, \tilde{g}_{i+\ell}) + \frac{2}{n} \sum_{i=1}^{n-1} \sum_{\ell=1}^{n-i} w_h(\ell) [\gamma(\ell)]^2. \\ &=: A_1 + A_2. \end{aligned}$$

where the third equality makes use of strict stationarity and the property of cumulants applied to zero-mean random variables:

$$cum(\tilde{g}_i, \tilde{g}_i, \tilde{g}_{i+\ell}, \tilde{g}_{i+\ell}) = Cov(\tilde{g}_i^2, \tilde{g}_{i+\ell}^2) - 2[Cov(\tilde{g}_i, \tilde{g}_{i+\ell})]^2.$$

The first term A_1 is bounded from above due to Assumption $D_a(ii)$:

$$A_1 \leq \frac{1}{nh} \sum_{i=1}^{n-1} \sum_{\ell=1}^{n-i} |cum(\tilde{g}_i, \tilde{g}_i, \tilde{g}_{i+\ell}, \tilde{g}_{i+\ell})| \leq \frac{K\Delta^{4(1-\beta\varpi)}}{nh}$$

The second term A_2 is evaluated using Assumption $D_a(i)$. Noting that $\tilde{\mathcal{K}} := \int_0^\infty w(x)f^2(x)dx$ is a finite positive constant under Assumption $D_a(i)$, we have, for $\nu = 0$,

$$A_2 = \frac{1}{h} \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) w\left(\frac{\ell}{h}\right) f^2\left(\frac{\ell}{h}\right) \frac{\Delta^{2(1-\beta\varpi)}}{nh^\delta} \simeq \frac{\tilde{\mathcal{K}}}{nh^\delta};$$

and for $\nu > 0$,

$$A_2 = \frac{1}{h} \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) w\left(\frac{\ell}{h}\right) f^2\left(\frac{\ell}{h}\right) \frac{\Delta^{1-\beta\varpi}}{n^{3/2}h^\delta} \simeq \frac{\tilde{\mathcal{K}}\Delta^{1-\beta\varpi}}{n^{3/2}h^\delta}.$$

We thus see that $A_2 = O(\max\{\frac{1}{nh^\delta}, \frac{\Delta^{1-\beta\varpi}}{n^{3/2}h^\delta}\})$. Using (16)-(17), we deduce that

$$\begin{aligned} \frac{A_1}{\sqrt{Var_0(\tilde{L})}} &= O\left(\max\left\{\frac{\Delta^{2(1-\beta\varpi)}}{n\sqrt{h}}, \frac{\Delta^{3(1-\beta\varpi)}}{\sqrt{nh}}\right\}\right), \\ \frac{A_2}{\sqrt{Var_0(\tilde{L})}} &= O\left(\sqrt{\frac{h}{n}} \frac{1}{nh^\delta}\right), \end{aligned}$$

and hence

$$\frac{E(A)}{\sqrt{Var_0(\tilde{L})}} = o_p(1).$$

To compute $E(B)$, we note that, for all $i \neq j$,

$$\begin{aligned} E(\tilde{g}_i \tilde{g}_j \tilde{g}_{i+\ell} \tilde{g}_{j+\ell}) &= cum(\tilde{g}_i, \tilde{g}_j, \tilde{g}_{i+\ell}, \tilde{g}_{j+\ell}) + Cov(\tilde{g}_i, \tilde{g}_{i+\ell})Cov(\tilde{g}_j, \tilde{g}_{j+\ell}) \\ &\quad + Cov(\tilde{g}_i, \tilde{g}_j)Cov(\tilde{g}_{i+\ell}, \tilde{g}_{j+\ell}) + Cov(\tilde{g}_i, \tilde{g}_{j+\ell})Cov(\tilde{g}_j, \tilde{g}_{i+\ell}) \\ &= cum(\tilde{g}_i, \tilde{g}_j, \tilde{g}_{i+\ell}, \tilde{g}_{j+\ell}) + [\gamma(\ell\Delta)]^2 + [\gamma((j-i)\Delta)]^2 \\ &\quad + \gamma((j+\ell-i)\Delta)\gamma((i+\ell-j)\Delta). \end{aligned}$$

Substituting into the summand of $E(B)$ yields

$$\begin{aligned} E(B) &= \frac{1}{n} \sum_{\ell=1}^{n-1} w_h(\ell) \sum_{i=1}^{n-\ell} \sum_{j \neq i} cum(\tilde{g}_i, \tilde{g}_j, \tilde{g}_{i+\ell}, \tilde{g}_{j+\ell}) \\ &\quad + \frac{1}{n} \sum_{\ell=1}^{n-1} w_h(\ell) \sum_{i=1}^{n-\ell} \sum_{j \neq i} [\gamma(\ell\Delta)]^2 \\ &\quad + \frac{1}{n} \sum_{\ell=1}^{n-1} w_h(\ell) \sum_{i=1}^{n-\ell} \sum_{j \neq i} [\gamma((j-i)\Delta)]^2 \\ &\quad + \frac{1}{n} \sum_{\ell=1}^{n-1} w_h(\ell) \sum_{i=1}^{n-\ell} \sum_{j \neq i} \gamma((j+\ell-i)\Delta)\gamma((i+\ell-j)\Delta) \\ &= : B_1 + B_2 + B_3 + B_4. \end{aligned}$$

The first term of $E(B)$ is

$$\begin{aligned} B_1 &\leq \frac{1}{nh} \sum \sum \sum_{i,j,\ell} |\text{cum}(\tilde{g}_i, \tilde{g}_j, \tilde{g}_{i+\ell}, \tilde{g}_{j+\ell})| \\ &\leq \frac{K\Delta^{4(1-\beta\varpi)}}{nh} \end{aligned}$$

by Assumption $D_a(\text{ii})$. It follows from (16)-(17) that $\frac{B_1}{\sqrt{\text{Var}_0(\tilde{L})}} = O(\max\{\frac{\Delta^{2(1-\beta\varpi)}}{n\sqrt{h}}, \frac{\Delta^{3(1-\beta\varpi)}}{\sqrt{nh}}\}) = o(1)$. The other terms are obtained as follows:

$$\begin{aligned} B_2 &= n \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) w_h(\ell) [\gamma(\ell\Delta)]^2, \\ B_3 &= 2 \sum_{i=1}^{n-1} \sum_{\ell=1}^{n-i} \left(1 - \frac{i}{n}\right) w_h(\ell) [\gamma(i\Delta)]^2, \\ B_4 &= 2 \sum_{i=1}^{n-1} \sum_{\ell=1}^{n-i} \left(1 - \frac{i}{n}\right) w_h(\ell) [\gamma((i+\ell)\Delta)]^2. \end{aligned}$$

Using the assumed form of $\gamma(\ell\Delta)$ in Assumption $D_a(\text{i})$, we evaluate B_2 under two cases. Since $\tilde{\mathcal{K}} = \int_0^\infty w(x)f^2(x)dx > 0$ by Assumption $D_a(\text{i})$, we have, for $\nu = 0$,

$$B_2 = \frac{n}{h} \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) w\left(\frac{\ell}{h}\right) \left[f\left(\frac{\ell}{h}\right) \frac{\Delta^{1-\beta\varpi}}{\sqrt{nh^\delta}} \right]^2 \simeq \frac{\tilde{\mathcal{K}}}{h^\delta}; \quad (18)$$

and for $\nu > 0$,

$$B_2 = \frac{n}{h} \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) w\left(\frac{\ell}{h}\right) \left[f\left(\frac{\ell}{h}\right) \frac{\Delta^{(1-\beta\varpi)/2}}{n^{3/4}h^{\delta/2}} \right]^2 \simeq \frac{\tilde{\mathcal{K}}\Delta^{1-\beta\varpi}}{\sqrt{nh^\delta}}. \quad (19)$$

In both cases, we note that $B_2 > 0$ because $\tilde{\mathcal{K}} > 0$. Using (16)-(17), we deduce that

$$\frac{B_2}{\sqrt{\text{Var}_0(\tilde{L})}} = O(h^{0.5-\delta}).$$

Next we evaluate B_3 . For $\nu = 0$,

$$\begin{aligned} B_3 &= \frac{2}{h} \sum_{i=1}^{n-1} \sum_{\ell=1}^{n-i} \left(1 - \frac{i}{n}\right) w\left(\frac{\ell}{h}\right) \left[f\left(\frac{i}{h}\right) \frac{\Delta^{1-\beta\varpi}}{\sqrt{nh^\delta}} \right]^2 \\ &= \frac{2h}{nh^\delta} \sum_{i=1}^{n-1} \left(1 - \frac{i}{n}\right) f^2\left(\frac{i}{h}\right) \frac{1}{h} \cdot \sum_{\ell=1}^{n-i} w\left(\frac{\ell}{h}\right) \frac{1}{h} \\ &\simeq 2 \frac{h^{1-\delta}}{n} \int_0^\infty f^2(x)dx \int_0^\infty w(x)dx; \end{aligned}$$

and for $\nu > 0$,

$$\begin{aligned}
B_3 &= \frac{2}{h} \sum_{i=1}^{n-1} \sum_{\ell=1}^{n-i} \left(1 - \frac{i}{n}\right) w\left(\frac{\ell}{h}\right) \left[f\left(\frac{i}{h}\right) \frac{\Delta^{(1-\beta\varpi)/2}}{n^{3/4}h^{\delta/2}} \right]^2 \\
&= \frac{2h\Delta^{1-\beta\varpi}}{n^{3/2}h^\delta} \sum_{i=1}^{n-1} \left(1 - \frac{i}{n}\right) f^2\left(\frac{i}{h}\right) \frac{1}{h} \cdot \sum_{\ell=1}^{n-i} w\left(\frac{\ell}{h}\right) \frac{1}{h} \\
&\simeq 2 \frac{h^{1-\delta}}{n} \frac{\Delta^{1-\beta\varpi}}{\sqrt{n}} \int_0^\infty f^2(x) dx \int_0^\infty w(x) dx.
\end{aligned}$$

In both cases, we note that $B_3 > 0$ because $\int_0^\infty f^2(x) dx > 0$.⁴ It follows from (16)-(17) that

$$\frac{B_3}{\sqrt{\text{Var}_0(\tilde{L})}} = O\left(\frac{h}{n} h^{0.5-\delta}\right),$$

which is dominated by $\frac{B_2}{\sqrt{\text{Var}_0(\tilde{L})}}$. The same is true for $\frac{B_4}{\sqrt{\text{Var}_0(\tilde{L})}}$ by a similar argument. Based on the above analysis, we conclude that

$$F_2 = O(h^{0.5-\delta}).$$

In other words, F_2 explodes to $+\infty$ when $0 \leq \delta < \frac{1}{2}$, and $F_2 = O(1)$ when $\delta = \frac{1}{2}$. In the latter case we may evaluate F_2 using (16)-(17) and (18)-(19). For $\nu = 0$,

$$F_2 = \frac{\tilde{\mathcal{K}} \Delta^{2(1-\beta\varpi)}}{\sqrt{2\mathcal{W}_2} [\text{Var}_0(g_{n,k,i})]^2} + o(1);$$

and for $\nu > 0$,

$$F_2 = \frac{\tilde{\mathcal{K}} \Delta^{1-\beta\varpi}}{\sqrt{\mathcal{W}_2} \text{Var}_0(g_{n,2k,i})} + o(1).$$

Define

$$\theta_k = \begin{cases} \frac{\tilde{\mathcal{K}}}{\sqrt{2\mathcal{W}_2} \sigma_k^4} & \text{for } \nu = 0, \\ \frac{\tilde{\mathcal{K}}}{\sqrt{\mathcal{W}_2} \sigma_{2k}^2} & \text{for } 0 < \nu \leq 1. \end{cases}$$

where $\sigma_k^2 := \lim_{n \rightarrow \infty} \frac{\text{Var}_0(g_{n,k,i})}{\Delta^{1-\beta\varpi}}$ is strictly positive and $O(1)$ by lemma 4. We thus obtain $F_2 = \theta_k + o(1)$ when $\delta = \frac{1}{2}$, so (ii) is true.

Next we want to prove (i). It is readily seen that $E(F_1) = 0$. It remains to prove that it

⁴Suppose the contrary that $\int_0^\infty f^2(x) dx = 0$. Then we must have $f(x) \equiv 0$, which contradicts with $\int_0^\infty w(x) f^2(x) dx > 0$ as stated in Assumption D_a(i).

is $O_p(1)$. To this end we decompose $\tilde{L} - E(\tilde{L})$ into mean-zero terms as follows:

$$\begin{aligned}\tilde{L} - E(\tilde{L}) &= \frac{1}{n} \sum_{\ell=1}^{n-1} w_h(\ell) \sum_{i=1}^{n-\ell} [\tilde{g}_i^2 \tilde{g}_{i+\ell}^2 - E(\tilde{g}_i^2 \tilde{g}_{i+\ell}^2)] \\ &\quad + \frac{1}{n} \sum_{\ell=1}^{n-1} w_h(\ell) \sum_{i=1}^{n-\ell} \sum_{j \neq i} [\tilde{g}_i \tilde{g}_j \tilde{g}_{i+\ell} \tilde{g}_{j+\ell} - E(\tilde{g}_i \tilde{g}_j \tilde{g}_{i+\ell} \tilde{g}_{j+\ell})] \\ &=: R + S.\end{aligned}$$

We determine the stochastic orders by computing their variances under the alternative DGP in Assumption D_a(i). The variance of R is

$$\begin{aligned}E(R^2) &= \frac{1}{n^2} \sum_{\ell_1} \sum_{\ell_2} w_h(\ell_1) w_h(\ell_2) \sum_{i_1} \sum_{i_2} Cov(\tilde{g}_{i_1}^2 \tilde{g}_{i_1+\ell_1}^2, \tilde{g}_{i_2}^2 \tilde{g}_{i_2+\ell_2}^2) \\ &=: R_1 + R_2,\end{aligned}$$

where R_1 corresponds to the part with $i_1 = i_2 = i$ and $\ell_1 = \ell_2 = \ell$, and R_2 is the remaining part. By strict stationarity and Assumption D_a(iii)-(iv), they are evaluated as follows:

$$R_1 = \frac{1}{n^2} \sum_{\ell} w_h^2(\ell) \sum_i Var(\tilde{g}_i^2 \tilde{g}_{i+\ell}^2) = \frac{1}{nh} \sum_{\ell} w^2\left(\frac{\ell}{h}\right) Var(\tilde{g}_i^2 \tilde{g}_{i+\ell}^2) \frac{1}{h} \leq \frac{K\Delta^{2(1-\beta\varpi)}}{nh},$$

and

$$R_2 \leq \frac{1}{nh} \sum_{\ell_1} w_h(\ell_1) \sum_{i_2} \sum_{\ell_2} |Cov(\tilde{g}_{i_1}^2 \tilde{g}_{i_1+\ell_1}^2, \tilde{g}_{i_2}^2 \tilde{g}_{i_2+\ell_2}^2)| \leq \frac{K\Delta^{4(1-\beta\varpi)}}{nh}.$$

Since R_1 dominates R_2 , we have $E(R^2) = O(\frac{\Delta^{2(1-\beta\varpi)}}{nh})$. Similarly, the variance of S is

$$\begin{aligned}E(S^2) &= \frac{1}{n^2} \sum_{\ell_1} \sum_{\ell_2} w_h(\ell_1) w_h(\ell_2) \sum_{i_1} \sum_{j_1 \neq i_1} \sum_{i_2} \sum_{j_2 \neq i_2} Cov(\tilde{g}_{i_1} \tilde{g}_{j_1} \tilde{g}_{i_1+\ell_1} \tilde{g}_{j_1+\ell_1}, \tilde{g}_{i_2} \tilde{g}_{j_2} \tilde{g}_{i_2+\ell_2} \tilde{g}_{j_2+\ell_2}) \\ &=: S_1 + S_2,\end{aligned}$$

where S_1 corresponds to the part with $i_1 = i_2 = i$, $j_1 = j_2 = j$ and $\ell_1 = \ell_2 = \ell$, and S_2 is the remaining part. By strict stationarity and Assumption D_a(iii)-(iv), they are evaluated as follows:

$$S_1 = \frac{1}{n^2} \sum_{\ell} w_h^2(\ell) \sum_i \sum_{j \neq i} Var(\tilde{g}_i \tilde{g}_j \tilde{g}_{i+\ell} \tilde{g}_{j+\ell}) = \frac{1}{h} \sum_{\ell} w^2\left(\frac{\ell}{h}\right) Var(\tilde{g}_i \tilde{g}_j \tilde{g}_{i+\ell} \tilde{g}_{j+\ell}) \frac{1}{h} \leq \frac{K\Delta^{4(1-\beta\varpi)}}{h},$$

and

$$S_2 \leq \frac{1}{h} \sum_{\ell_1} w_h(\ell_1) \sum_{i_2} \sum_{j_2} \sum_{\ell_2} |Cov(\tilde{g}_{i_1} \tilde{g}_{j_1} \tilde{g}_{i_1+\ell_1} \tilde{g}_{j_1+\ell_1}, \tilde{g}_{i_2} \tilde{g}_{j_2} \tilde{g}_{i_2+\ell_2} \tilde{g}_{j_2+\ell_2})| \leq \frac{K\Delta^{4(1-\beta\varpi)}}{h}.$$

Since S_1 and S_2 are both $O(\frac{\Delta^{4(1-\beta\varpi)}}{h})$, we have $E(S^2) = O(\frac{\Delta^{4(1-\beta\varpi)}}{h})$. From the above analysis, we observe that $\tilde{L} - E(\tilde{L}) = O_p(\max\{\frac{\Delta^{1-\beta\varpi}}{\sqrt{nh}}, \frac{\Delta^{2(1-\beta\varpi)}}{\sqrt{h}}\})$, which is of the same stochastic order as $\sqrt{\text{Var}_0(\tilde{L})}$. We thus deduce that $F_1 = \frac{\tilde{L} - E(\tilde{L})}{\sqrt{\text{Var}_0(\tilde{L})}} = O_p(1)$ and so (i) is true. The stated result follows immediately.

1.8 Proof of Theorem 3

By (6) in the proof of Theorem 1, it suffices to derive the asymptotic distribution of $\tilde{\gamma}(\ell)$.

Let us compute the first two moments of $\tilde{\gamma}(\ell)$ under the null hypothesis. Since $E(\tilde{g}_i \tilde{g}_{i+\ell}) = E(\tilde{g}_i)E(\tilde{g}_{i+\ell}) = 0$ for all $\ell \neq 0$ under the null hypothesis, the mean is given by

$$E[\tilde{\gamma}(\ell)] = \frac{1}{n} \sum_{i=1}^{n-\ell} E(\tilde{g}_i \tilde{g}_{i+\ell}) = \frac{1}{n} \left(1 - \frac{\ell}{n}\right) E(\tilde{g}_i)E(\tilde{g}_{i+\ell}) = 0.$$

The variance is equal to the second moment (since the mean is zero) and is given by

$$\begin{aligned} \text{Var}[\tilde{\gamma}(\ell)] &= E[\tilde{\gamma}^2(\ell)] \\ &= \frac{1}{n^2} \sum_{i=1}^{n-\ell} E(\tilde{g}_i^2 \tilde{g}_{i+\ell}^2) + \frac{1}{n^2} \sum_{i=1}^{n-\ell} \sum_{j=1, j \neq i}^{n-\ell} E(\tilde{g}_i \tilde{g}_j \tilde{g}_{i+\ell} \tilde{g}_{j+\ell}) \\ &\simeq \frac{1}{n^2} \sum_{i=1}^{n-\ell} E(\tilde{g}_i^2) E(\tilde{g}_{i+\ell}^2) \\ &= \frac{1}{n} \left(1 - \frac{\ell}{n}\right) [E(\tilde{g}_i^2)]^2 \\ &= \frac{1}{n} \left(1 - \frac{\ell}{n}\right) [\text{Var}(g_i)]^2, \end{aligned} \tag{20}$$

which is $O\left(\frac{\Delta^{2(1-\beta\varpi)}}{n}\right)$. The third line follows from part (3) of lemma and Assumption D₀.

This implies that $\tilde{\gamma}(\ell)$ is $O_p\left(\frac{\Delta^{1-\beta\varpi}}{n^{1/2}}\right)$.

Next, we establish the asymptotic normality of $\tilde{\gamma}(\ell)$. To this end, we define $s^2 = \text{Var}[\tilde{\gamma}(\ell)]$ and check the following conditions:

- (i) $\tilde{\gamma}(\ell)$ can be written as $\sum_{i=1}^n Y_i$, where Y_i form a martingale difference sequence: that Y_i is measurable to $\mathcal{F}_i$, and $E(Y_i | \mathcal{F}_j) = 0$ for all $j < i$.
- (ii) for any given $\eta > 0$, $s^{-2} \sum_{i=1}^n E(Y_i^2 1_{\{|Y_i| \geq s\eta\}}) \rightarrow 0$ as $n \rightarrow \infty$.
- (iii) $s^{-2} \sum_{i=1}^n Y_i^2 \xrightarrow{p} 1$.

To show (i), we rewrite $\tilde{\gamma}(\ell)$ as $\tilde{\gamma}(\ell) = \frac{1}{n} \sum_{i=\ell+1}^n \tilde{g}_{i-\ell} \tilde{g}_i$, and construct the process $\{Y_i\}$ by letting $Y_i = 0$ if $i \leq \ell$ and $Y_i = \frac{1}{n} \tilde{g}_{i-\ell} \tilde{g}_i$ if $i > \ell$. It is clear that $\{Y_i; \mathcal{F}_i\}$ is a martingale difference sequence.

To show (ii), it suffices to verify that $s^{-3} \sum_{i=1}^n E(Y_i^3) \rightarrow 0$ as $n \rightarrow \infty$. By the null hypothesis and stationarity, we have

$$\begin{aligned} \sum_{i=1}^n E(Y_i^3) &= \frac{1}{n^3} \sum_{i=1}^n E(\tilde{g}_{i-\ell}^3) E(\tilde{g}_i^3) \\ &= \frac{1}{n^2} \left(1 - \frac{\ell}{n}\right) [E(\tilde{g}_i^3)]^2. \end{aligned}$$

This implies that

$$s^{-3} \sum_{i=1}^n E(Y_i^3) = \frac{1}{n^{0.5}} \left(1 - \frac{\ell}{n}\right)^{-0.5} [E(\tilde{g}_i^3)]^2 [Var(g_i)]^{-3},$$

which tends to zero because $g(\cdot)$ and hence $E(\tilde{g}_i^3)$ are bounded from above, and $Var(g_i) > 0$ on the set Ω_{J_ϵ} .

To show (iii), it suffices to verify that $E[(s^{-2} \sum_{i=1}^n Y_i^2 - 1)^2] \rightarrow 0$ as $n \rightarrow \infty$. We evaluate the expectation

$$\begin{aligned} E[(s^{-2} \sum_{i=1}^n Y_i^2 - 1)^2] &= s^{-4} E\left(\sum_{i=1}^n Y_i^2\right)^2 - 2s^{-2} E\left(\sum_{i=1}^n Y_i^2\right) + 1 \\ &= s^{-4} E\left(\sum_{i=1}^n Y_i^2\right)^2 - 1, \end{aligned}$$

so it suffices to show that $s^{-4} E(\sum_{i=\ell+1}^n Y_i^2)^2 \rightarrow 1$ as $n \rightarrow \infty$. The expectation on the last line can be decomposed as follows

$$E\left(\sum_{i=1}^n Y_i^2\right)^2 = \sum_{i=1}^n E(Y_i^4) + \sum_{i=1}^n \sum_{j=1, j \neq i}^n E(Y_i^2 Y_j^2) =: R_1 + R_2.$$

We evaluate each of the terms. For R_1 , we have

$$R_1 = \sum_{i=1}^n E(Y_i^4) \simeq \frac{1}{n^4} \sum_{i=\ell+1}^n E(\tilde{g}_{i-\ell}^4) E(\tilde{g}_i^4) = \frac{1}{n^3} \left(1 - \frac{\ell}{n}\right) [E(\tilde{g}_i^4)]^2,$$

so that R_1 is $O\left(\frac{\Delta^{2(1-\beta\varpi)}}{n^3}\right)$. For R_2 , we have

$$\begin{aligned}
R_2 &= \sum_{i=1}^n \sum_{j=1, j \neq i}^n E(Y_i^2 Y_j^2) \\
&\simeq \frac{1}{n^4} \sum_{i=\ell+1}^n \sum_{j=\ell+1, j \neq i}^n E(\tilde{g}_{i-\ell}^2) E(\tilde{g}_{j-\ell}^2) E(\tilde{g}_i^2) E(\tilde{g}_j^2) \\
&= \frac{1}{n^4} \sum_{i=\ell+1}^n \sum_{j=\ell+1, j \neq i, j \neq i-\ell}^n E(\tilde{g}_{i-\ell}^2) E(\tilde{g}_{j-\ell}^2) E(\tilde{g}_i^2) E(\tilde{g}_j^2) \\
&\quad + \frac{1}{n^4} \sum_{i=\ell+1}^n E(\tilde{g}_{i-\ell}^4) E(\tilde{g}_{i-2\ell}^2) E(\tilde{g}_i^2) \\
&= \frac{1}{n^2} \left(1 - \frac{\ell}{n}\right) \left(1 - \frac{\ell+1}{n}\right) [E(\tilde{g}_j^2)]^4 + \frac{1}{n^3} \left(1 - \frac{\ell}{n}\right) E(\tilde{g}_j^4) [E(\tilde{g}_j^2)]^2 \\
&=: R_{21} + R_{22}
\end{aligned}$$

where R_{21} is $O\left(\frac{\Delta^{4(1-\beta\varpi)}}{n^2}\right)$, and R_{22} is $O\left(\frac{\Delta^{3(1-\beta\varpi)}}{n^3}\right)$.

Comparing the three terms R_1 , R_{21} and R_{22} , we observe the following: (1) when $\nu = 0$, the dominant term is R_{21} which is $O\left(\frac{\Delta^{4(1-\beta\varpi)}}{n^2}\right)$; (2) when $\nu > 0$, we note that

$$\frac{R_1}{R_{21}} = O\left(\frac{\Delta^{2(1-\beta\varpi)}/n^3}{\Delta^{4(1-\beta\varpi)}/n^2}\right) = O\left(\frac{1}{\Delta^{2(1-\beta\varpi)}n}\right) = O\left(\frac{1}{n^{1-2\nu(1-\beta\varpi)}}\right) = o(1)$$

where the last step follows from $\nu < \frac{1}{2(1-\beta\varpi)}$ by Assumption T_γ , and hence the dominant term is still R_{21} . Noting that $s^{-4}R_{21} \rightarrow 1$ as $n \rightarrow \infty$, we see that (iii) is verified.

By (i)-(iii) and the martingale central limit theorem (Hall and Heyde (1980)), $\tilde{\gamma}(\ell)$ has a limiting normal distribution; i.e.,

$$\frac{\tilde{\gamma}(\ell)}{\sqrt{Var(\tilde{\gamma}(\ell))}} \xrightarrow{d} N(0, 1). \tag{21}$$

Next, we want to show that $\hat{\gamma}(0)$ converges in probability to $E(\tilde{\gamma}(0))$. By (6), it suffices to show that $\frac{\tilde{\gamma}(0)}{E(\tilde{\gamma}(0))} = 1 + o_p(1)$. We compute the first two moments: the mean is

$$E[\tilde{\gamma}(0)] = \frac{1}{n} \sum_{i=1}^n E(\tilde{g}_i^2) = Var(g_i) = O\left(\Delta^{1-\beta\varpi}\right),$$

and the second moment is

$$\begin{aligned} E[\tilde{\gamma}^2(0)] &\simeq \frac{1}{n^2} \sum_{i=1}^n E(\tilde{g}_i^4) + \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1, j \neq i}^n E(\tilde{g}_i^2) E(\tilde{g}_j^2) \\ &= \frac{1}{n} E(\tilde{g}_i^4) + \left(1 - \frac{1}{n}\right) [Var(g_i)]^2, \end{aligned}$$

so that the variance is

$$\begin{aligned} Var[\tilde{\gamma}(0)] &= E[\tilde{\gamma}^2(0)] - [E(\tilde{\gamma}(0))]^2 = \frac{1}{n} E(\tilde{g}_i^4) - \frac{1}{n} [Var(g_i)]^2 \\ &= O\left(\frac{\Delta^{1-\beta\varpi}}{n}\right) + O\left(\frac{\Delta^{2(1-\beta\varpi)}}{n}\right). \end{aligned}$$

This implies that $R_0 := \tilde{\gamma}(0) - E(\tilde{\gamma}(0))$ is $O_p\left(\frac{\Delta^{(1-\beta\varpi)/2}}{n^{1/2}}\right)$. Now we observe that

$$\frac{R_0}{E(\tilde{\gamma}(0))} = O_p\left(\frac{\Delta^{(1-\beta\varpi)/2}/n^{1/2}}{\Delta^{1-\beta\varpi}}\right) = O_p\left(\frac{1}{n^{1/2}\Delta^{(1-\beta\varpi)/2}}\right) = O_p\left(\frac{1}{n^{[1-\nu(1-\beta\varpi)]/2}}\right),$$

which is $o_p(1)$ because $1 - \nu(1 - \beta\varpi) > 0$ by Assumptions G and T_γ . It follows that

$$\frac{\tilde{\gamma}(0)}{E(\tilde{\gamma}(0))} = 1 + \frac{R_0}{E(\tilde{\gamma}(0))} = 1 + o_p(1). \quad (22)$$

Furthermore, recalling from (20), we obtain

$$\frac{\sqrt{nVar(\tilde{\gamma}(\ell))}}{E(\tilde{\gamma}(0))} = \left(1 - \frac{\ell}{n}\right)^{1/2} = 1 + o(1) \quad (23)$$

for fixed ℓ . Combining (21)-(23), we conclude that

$$\sqrt{n} \frac{\tilde{\gamma}(\ell)}{\tilde{\gamma}(0)} = \frac{\tilde{\gamma}(\ell)}{\sqrt{Var(\tilde{\gamma}(\ell))}} \frac{\sqrt{nVar(\tilde{\gamma}(\ell))}}{E(\tilde{\gamma}(0))} \frac{E(\tilde{\gamma}(0))}{\tilde{\gamma}(0)} \xrightarrow{d} N(0, 1).$$

The extra term $r\Delta^p$ in the denominator of $\hat{\rho}_{n,k}(\ell)$ is strictly dominated in the limit and does not affect the asymptotic result. The proof is now completed.

1.9 Proof of Theorem 4

We decompose $\sqrt{n}\hat{\rho}(\ell)$ as follows:

$$\begin{aligned} \sqrt{n}\hat{\rho}(\ell) &= \frac{\hat{\gamma}(\ell)}{\hat{\gamma}(0)/\sqrt{n}} \\ &= \frac{\tilde{\gamma}(0)}{\tilde{\gamma}(0)} \frac{\hat{\gamma}(\ell)}{\tilde{\gamma}(\ell)} \left[\frac{\tilde{\gamma}(\ell) - \gamma(\ell\Delta)}{\tilde{\gamma}(0)/\sqrt{n}} + \frac{\gamma(\ell\Delta)}{\tilde{\gamma}(0)/\sqrt{n}} \right]. \end{aligned}$$

We break down the proof into several parts:

- (i) $\frac{\tilde{\gamma}(0)}{\tilde{\gamma}(0)} = 1 + o_p(1)$.
- (ii) $\frac{\hat{\gamma}(\ell)}{\tilde{\gamma}(\ell)} = 1 + o_p(1)$.
- (iii) $\frac{\tilde{\gamma}(\ell) - \gamma(\ell\Delta)}{\tilde{\gamma}(0)/\sqrt{n}}$ is $O_p(1)$ and has mean zero.
- (iv) for $\delta = \frac{1}{2}$, $\frac{\gamma(\ell\Delta)}{\tilde{\gamma}(0)/\sqrt{n}} = \pi_k + o_p(1)$, where $\pi_k := \frac{K_\ell}{\sigma_k^2} \neq 0$; and for $0 \leq \delta < \frac{1}{2}$, the term diverges to $+\infty$ or $-\infty$.

To prove (i) and (ii), we follow the arguments at the start of the proof of Theorem 1 by analyzing the three terms in (4), now under the alternative hypothesis. For fixed ℓ , the last term (*III*) is strictly dominated by the second term (*II*) and is negligible. To determine the stochastic order of the second term (*II*), we compute

$$\begin{aligned} Var(\mu_g - \bar{g}) &= Var\left(\frac{1}{n} \sum_{i=1}^n \tilde{g}_i\right) = \frac{1}{n^2} \sum_{i=1}^n Var(\tilde{g}_i) + \frac{1}{n^2} \sum_{i=1}^n \sum_{j \neq i} Cov(\tilde{g}_i, \tilde{g}_j) \\ &= \frac{1}{n} Var(\tilde{g}_i) + \frac{1}{n} \sum_{j \neq i} Cov(\tilde{g}_i, \tilde{g}_j). \end{aligned}$$

The first term in the last line is $O(\frac{\Delta^{1-\beta\varpi}}{n})$ by lemma 4. By Assumption $D_a(\ell)$ (i), the second term can be bounded from above as follows: for some $K > 0$,

$$\frac{1}{n} \sum_{j \neq i} Cov(\tilde{g}_i, \tilde{g}_j) \leq \frac{1}{n} \sum_{\ell > 0} |K_\ell| \frac{\Delta^{1-\beta\varpi}}{n^\delta} \leq \frac{K}{n^{1+\delta}} \Delta^{1-\beta\varpi}. \quad (24)$$

We thus see that (*II*) is $O(\frac{\Delta^{1-\beta\varpi}}{n}) = O(\frac{1}{n})$. Now we turn to the first term (*I*) in (4). By Assumption $D_a(\ell)$ (iii) and strict stationarity, the variance of $\tilde{\gamma}(\ell)$ is given by

$$Var[\tilde{\gamma}(\ell)] = \frac{1}{n^2} \sum_{i=1}^{n-\ell} \sum_{j=1}^{n-\ell} Cov(\tilde{g}_i \tilde{g}_{i+\ell}, \tilde{g}_j \tilde{g}_{j+\ell}) \leq \frac{1}{n} \sum_{j=1}^{n-\ell} |Cov(\tilde{g}_i \tilde{g}_{i+\ell}, \tilde{g}_j \tilde{g}_{j+\ell})| \leq \frac{K \Delta^{2(1-\beta\varpi)}}{n}.$$

The bound on $Var[\tilde{\gamma}(\ell)]$ implies

$$\tilde{\gamma}(\ell) - \gamma(\ell\Delta) = O_p\left(\frac{\Delta^{1-\beta\varpi}}{\sqrt{n}}\right) = O_p\left(\frac{1}{\sqrt{n}}\right), \quad (25)$$

which gives $\tilde{\gamma}(\ell) = \gamma(\ell\Delta) + o_p(1)$. By the decomposition (4) and the above analysis, we deduce that $\hat{\gamma}(\ell) = \tilde{\gamma}(\ell) + o_p(1)$, and hence

$$\frac{\hat{\gamma}(\ell)}{\tilde{\gamma}(\ell)} = 1 + o_p(1). \quad (26)$$

Note that the above argument leading to (26) works for all $\ell \geq 0$. We have thus proved (i) and (ii).

Next we prove (iii). The mean of $\tilde{\gamma}(0)$ is

$$\gamma(0) = E[\tilde{\gamma}(0)] = \frac{1}{n} \sum_{i=1}^n E(\tilde{g}_i^2) = \text{Var}(\tilde{g}_i) = O(\Delta^{1-\beta\varpi}).$$

The variance of $\tilde{\gamma}(0)$ is

$$\begin{aligned} \text{Var}[\tilde{\gamma}(0)] &= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \text{Cov}(\tilde{g}_i^2, \tilde{g}_j^2) \\ &= \frac{1}{n^2} \sum_{i=1}^n \text{Var}(\tilde{g}_i^2) + \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1, j \neq i}^n \text{Cov}(\tilde{g}_i^2, \tilde{g}_j^2) \\ &= \frac{1}{n} \text{Var}(\tilde{g}_i^2) + \frac{1}{n} \sum_{j=1, j \neq i}^n \text{Cov}(\tilde{g}_i^2, \tilde{g}_j^2). \end{aligned}$$

The first term in the last line is $O(\frac{\Delta^{1-\beta\varpi}}{n})$. The second term is at most $O(\frac{\Delta^{2(1-\beta\varpi)}}{n})$ by Assumption $D_a(\ell)(ii)$. It follows that $\tilde{\gamma}(0) - \gamma(0) = O_p(\frac{1}{\sqrt{n}})$, which implies $\frac{\tilde{\gamma}(0)}{\gamma(0)} = 1 + o_p(1)$. Note that $\gamma(0) = \text{Var}(1_{\{|\Delta X_i^d| > u_n\}} | \Omega_J) > 0$ for all sufficiently large n . Using (25), we have

$$\frac{\tilde{\gamma}(\ell) - \gamma(\ell\Delta)}{\tilde{\gamma}(0)/\sqrt{n}} = \frac{\tilde{\gamma}(\ell) - \gamma(\ell\Delta)}{\gamma(0)/\sqrt{n}} \frac{\gamma(0)}{\tilde{\gamma}(0)} = O_p(1),$$

which by construction has mean zero; hence (iii) is true.

Finally, we prove (iv). The mean of $\tilde{\gamma}(\ell)$ for fixed $\ell \neq 0$ is given by

$$\gamma(\ell\Delta) = E[\tilde{\gamma}(\ell)] = \frac{1}{n} \sum_{i=1}^{n-\ell} E(\tilde{g}_i \tilde{g}_{i+\ell}) = \left(1 - \frac{\ell}{n}\right) \gamma(\ell\Delta) = \left(1 - \frac{\ell}{n}\right) \frac{K_\ell \Delta^{1-\beta\varpi}}{n^\delta}, \quad (27)$$

which is obtained by Assumption $D_a(\ell)(i)$ and strict stationarity. We therefore deduce that

$$\frac{\gamma(\ell\Delta)}{\tilde{\gamma}(0)/\sqrt{n}} = \frac{\gamma(\ell\Delta)}{\gamma(0)/\sqrt{n}} \frac{\gamma(0)}{\tilde{\gamma}(0)} = \frac{(1 - \frac{\ell}{n}) K_\ell \Delta^{1-\beta\varpi} / n^\delta}{\text{Var}(\tilde{g}_i) / \sqrt{n}} (1 + o_p(1)). \quad (28)$$

For $\delta = \frac{1}{2}$, the fraction is $O(1)$ by lemma 4 and has a non-zero limit $\pi_k := \frac{K_\ell}{\sigma_k^2}$; and for $\delta < \frac{1}{2}$, the fraction diverges to $+\infty$ or $-\infty$ because $K_\ell \neq 0$. We therefore see that (iv) is true. The stated result follows immediately.