LIMIT THEOREMS FOR THE EMPIRICAL DISTRIBUTION FUNCTION OF SCALED INCREMENTS OF ITÔ SEMIMARTINGALES AT HIGH FREQUENCIES

By Viktor Todorov and George Tauchen

Northwestern University and Duke University

We derive limit theorems for the empirical distribution function of "devolatilized" increments of an Itô semimartingale observed at high frequencies. These " devolatilized" increments are formed by suitably rescaling and truncating the raw increments to remove the effects of stochastic volatility and "large" jumps. We derive the limit of the empirical cdf of the adjusted increments for any Itô semimartingale whose dominant component at high frequencies has activity index of $1 < \beta \leq 2$, where $\beta = 2$ corresponds to diffusion. We further derive an associated CLT in the jump-diffusion case. We use the developed limit theory to construct a feasible and pivotal test for the class of Itô semimartingales with non-vanishing diffusion coefficient against Itô semimartingales with no diffusion component.

1. Introduction. The standard jump-diffusion model used for modeling many stochastic processes is an Itô semimartingale given by the following differential equation

$$dX_t = \alpha_t dt + \sigma_t dW_t + dY_t, \tag{1.1}$$

where α_t and σ_t are processes with càdlàg paths, W_t is a Brownian motion and Y_t is an Itô semimartingale process of pure-jump type (i.e., semimartingale with zero second characteristic, Definition II.2.6 in [10]).

At high-frequencies, provided σ_t does not vanish, the dominant component of X_t is its continuous martingale component and at these frequencies the increments of X_t in (1.1) behave like scaled and independent Gaussian random variables. That is, for each fixed t, we have the following convergence

$$\frac{1}{\sqrt{h}}(X_{t+sh} - X_t) \xrightarrow{\mathcal{L}} \sigma_t \times (B_{t+s} - B_s), \quad \text{as } h \to 0 \text{ and } s \in [0,1], \tag{1.2}$$

where B_t is a Brownian motion and the above convergence is for the Skorokhod topology, see e.g., Lemma 1 of [19]. There are two distinctive features

AMS 2000 subject classifications: Primary 62F12, 62M05; secondary 60H10, 60J75.

Keywords and phrases: Itô semimartingale, Kolmogorov-Smirnov test, high-frequency data, stochastic volatility, jumps, stable process.

of the convergence in (1.2). The first is the scaling factor of the increments on the left side of (1.2) is the square-root of the length of the high-frequency interval, a feature that has been used in developing tests for presence of diffusion. The second distinctive feature is that the limiting distribution of the (scaled) increments on the right side of (1.2) is mixed Gaussian (the mixing given by σ_t). Both these features of the local Gaussianity result in (1.2) for models in (1.1) have been key in the construction of essentially all nonparametric estimators of functionals of volatility. Examples include the jump-robust Bipower Variation of [5, 6] and the many other alternative measures of powers of volatility summarized in the recent book of [9]. Another important example is the general approach of [15] (see also [14]) where estimators of functions of volatility are formed by utilizing directly (1.2) and working as if volatility is constant over a block of decreasing length.

Despite the generality of the jump-diffusion model in (1.1), however, there are several examples of stochastic processes considered in various applications that are not nested in the model in (1.1). Examples include pure-jump Itô semimartingales (i.e., the model in (1.1) with $\sigma_t = 0$ and jumps present), semimartingales contaminated with noise or more generally non-semimartingales. In all these cases, both the scaling constant on the left side of (1.2) as well as the limiting process on the right side of (1.2) change. Our goal in this paper, therefore, is to derive a limit theory for a feasible version of the local Gaussianity result in (1.2) based on high-frequency record of X. An application of the developed limit theory is a feasible and pivotal test based on Kolmogorov-Smirnov type distance for the class of Itô semimartingales with non-vanishing diffusion component.

The result in (1.2) implies that the high-frequency increments are approximately Gaussian but the key obstacle of testing directly (1.2) is that the (conditional) variance of the increments, σ_t^2, is unknown and further is approximately constant only over a short interval of time. Therefore, on a first step we split the high-frequency increments into blocks (with length that shrinks asymptotically to zero as we sample more frequently) and form local estimators of volatility over the blocks. We then scale the high-frequency increments within each of the blocks by our local estimates of the volatility. This makes the scaled high-frequency increments approximately i.i.d. centered normal random variables with unit variance. To purge further the effect of "big" jumps, we then discard the increments that exceed a time-varying threshold (that shrinks to zero asymptotically) with time-variation determined by our estimator of the local volatility. We derive a (functional) Central Limit Theorem (CLT) for the convergence of the empirical cdf of the scaled high-frequency increments, not exceeding the threshold, to the cdf of a standard normal random variable. The rate of convergence can be made arbitrary close to $\sqrt{n}$, by appropriately choosing the rate of increase of the block size, where n is the number of high-frequency observations within the time interval. This is achieved despite the use of the block estimators of volatility, each of which can estimate the spot volatility σ_t at a rate no faster than $n^{1/4}$.

We further derive the limit behavior of the empirical cdf described above in two possible alternatives to the model (1.1). The first is the case where X_t does not contain a diffusive component, i.e., the second term in (1.1) is absent. Models of these type have received a lot of attention in various fields, see e.g., [3, 4], [13], [11] and [22]. The second alternative to (1.1) is the case in which the Itô semimartingale is distorted with measurement error. In each of these two cases, the empirical cdf of the scaled high-frequency increments below the threshold converges to a cdf of a distribution different from the standard normal law. This is the stable distribution in the pure-jump case and the distribution of the noise in the case of Itô semimartingale observed with error.

The paper is organized as follows. In Section 2 we introduce the formal setup and state the assumptions needed for our theoretical results. In Section 3 we construct our statistic and in Sections 4 and 5 we derive its limit behavior. In Section 6 we construct the statistic using alternative local estimator of volatility and derive its limit behavior in the jump-diffusion case. Section 7 constructs a feasible test for local Gaussianity using our limit theory and in Sections 8 and 9 we apply the test on simulated and real financial data respectively. The proofs are given in Section 10.

2. Setup.

We start with the formal setup and assumptions. We will generalize the setup in (1.1) to accommodate also the alternative hypothesis in which X can be of pure-jump type. Thus, the generalized setup we consider is the following. The process X is defined on a filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and has the following dynamics

$$dX_t = \alpha_t dt + \sigma_{t-} dS_t + dY_t, \tag{2.1}$$

where α_t, σ_t and Y_t are processes with càdlàg paths adapted to the filtration and Y_t is of pure-jump type. S_t is a stable process with a characteristic function, see e.g., [17], given by

$$\log\left[\mathbb{E}(e^{iuS_t})\right] = -t|cu|^{\beta}\left(1 - i\gamma \mathrm{sign}(u)\Phi\right), \quad \Phi = \begin{cases} \tan(\pi\beta/2) & \text{if } \beta \neq 1, \\ -\frac{2}{\pi}\log|u|, & \text{if } \beta = 1, \end{cases} \tag{2.2}$$

where $\beta \in (0, 2]$ and $\gamma \in [-1, 1]$. When $\beta = 2$ and $c = 1/2$ in (2.2), we recover our original jump-diffusion specification in (1.1) in the introduction. When $\beta < 2$, X is of pure-jump type. Y_t in (2.1) will play the role of a "residual" jump component at high frequencies (see assumption A2 below). We note that Y_t can have dependence with S_t (and α_t and σ_t), and thus X_t does not "inherit" the tail properties of the stable process S_t, e.g., X_t can be driven by a tempered stable process whose tail behavior is very different from that of the stable process.

Throughout the paper we will be interested in the process X over an interval of fixed length and hence without loss of generality we will fix this interval to be $[0, 1]$. We collect our basic assumption on the components in X next.

Assumption A. X_t *satisfies (2.1).*

A1. $|\sigma_t|^{-1}$ *and* $|\sigma_{t-}|^{-1}$ *are strictly positive on* $[0, 1]$. *Further, there is a sequence of stopping times* T_p *increasing to infinity and for each* p *a bounded process* $\sigma_t^{(p)}$ *satisfying* $t < T_p \implies \sigma_t = \sigma_t^{(p)}$ *and a positive constant* K_p *such that*

$$\mathbb{E}\left(|\sigma_t^{(p)} - \sigma_s^{(p)}|^2 |\mathcal{F}_s\right) \leq K_p|t - s|, \quad \textit{for every} \quad 0 \leq s \leq t \leq 1. \tag{2.3}$$

A2. *There is a sequence of stopping times* T_p *increasing to infinity and for each* p *a process* $Y_t^{(p)}$ *satisfying* $t < T_p \implies Y_t = Y_t^{(p)}$ *and a positive constant* K_p *such that*

$$\mathbb{E}\left(|Y_t^{(p)} - Y_s^{(p)}|^q |\mathcal{F}_s\right) \leq K_p|t - s|, \quad \textit{for every} \quad 0 \leq s \leq t \leq 1, \tag{2.4}$$

and for every $q > \beta'$ *where* $\beta' < \beta$.

The assumption in (2.3) can be easily verified for Itô semimartingales which is the typical way of modeling σ_t, but it is also satisfied for models outside of this class. The condition in (2.4) can be easily verified for pure-jump Itô semimartingales, see e.g., Corollary 2.1.9 of [9].

Remark 1. *Our setup in (2.1) (together with assumption A) includes the more parsimonious pure-jump models for* X *of the form* $\int_0^t \sigma_{s-} dL_s$ *and* L_{T_t} *where* T_t *is absolute continuous time-change process and* L_t *is a Lévy process with no diffusion component and Lévy density of the form* $\frac{A_+ 1_{\{x>0\}} + A_- 1_{\{x<0\}}}{|x|^{1+\beta}} + \nu'(x)$ *for* $|\nu'(x)| \leq \frac{K}{|x|^{1+\beta'}}$ *when* $|x| < x_0$ *for some* $x_0 > 0$ *(and assumptions for* σ_t *and the density of the time change as in A1 above). We refer to [19] and its supplementary appendix where this is shown.*

Under assumption A, we can extend the local Gaussianity result in (1.2) to

$$h^{-1/\beta}(X_{t+sh} - X_t) \xrightarrow{\mathcal{L}} \sigma_t \times (S'_{t+s} - S'_t), \quad \text{as } h \to 0 \text{ and } s \in [0,1], \tag{2.5}$$

for every t and where S'_t is a Lévy process identically distributed to S_t and the convergence in (2.5) being for the Skorokhod topology, see e.g., Lemma 1 of [19]. That is, the local behavior of the increments of the process is like that of a stable process in the more general setting of (2.1).

For deriving the CLT for our statistic (in the case of the jump-diffusion model in (1.1)), we need a stronger assumption which we state next.

Assumption B. *X_t satisfies (2.1) with $\beta = 2$, i.e., $S_t = W_t$.*

B1. *The process Y_t is of the form*

$$Y_t = \int_0^t \int_E \delta^Y(s,x)\mu(ds,dx), \tag{2.6}$$

where μ is Poisson measure on $\mathbb{R}_+ \times E$ with Lévy measure $\nu(dx)$ and $\delta^Y(t,x)$ is some predictable function on $\Omega \times \mathbb{R}_+ \times E$.

B2. *$|\sigma_t|^{-1}$ and $|\sigma_{t-}|^{-1}$ are strictly positive on $[0,1]$. Further, σ_t is an Itô semimartingale having the following representation*

$$\sigma_t = \sigma_0 + \int_0^t \widetilde{\alpha}_u du + \int_0^t \widetilde{\sigma}_u dW_u + \int_0^t \widetilde{\sigma}'_u dW'_u + \int_0^t \int_E \delta^\sigma(s,x)\mu(ds,dx), \tag{2.7}$$

where W'_t is a Brownian motion independent from W_t; $\widetilde{\alpha}_t$, $\widetilde{\sigma}_t$ and $\widetilde{\sigma}'_t$ are processes with càdlàg paths and $\delta^\sigma(t,x)$ is a predictable function on $\Omega \times \mathbb{R}_+ \times E$.

B3. *$\widetilde{\sigma}_t$ and $\widetilde{\sigma}'_t$ are Itô semimartingales with coefficients with càdlàg paths and further jumps being integrals of some predictable functions, $\delta^{\widetilde{\sigma}}$ and $\delta^{\widetilde{\sigma}'}$, with respect to the jump measure μ.*

B4. *There is a sequence of stopping times T_p increasing to infinity and for each p a deterministic nonnegative function $\gamma_p(x)$ on $\mathbb{E}$, satisfying $\nu(x : \gamma_p(x) \neq 0) < \infty$ and such that $|\delta^Y(t,x)| \wedge 1 + |\delta^\sigma(t,x)| \wedge 1 + |\delta^{\widetilde{\sigma}}(t,x)| \wedge 1 + |\delta^{\widetilde{\sigma}'}(t,x)| \wedge 1 \leq \gamma_p(x)$ for $t \leq T_p$.*

The Itô semimartingale restriction on σ_t (and its coefficients) is satisfied in most applications. Similarly, we allow for general time-dependence in the jumps in X which encompasses most cases in the literature. B4 is the strongest assumption and it requires the jumps to be of finite activity.

3. Empirical CDF of the "Devolatilized" High Frequency Increments.

Throughout the paper we assume that X is observed on the equidistant grid $0, \frac{1}{n}, ..., 1$ with $n \to \infty$. In the derivation of our statistic we will suppose that S_t is a Brownian motion and then in the next section we will derive its behavior under the more general case when S_t is a stable process. The result in (1.2) suggests that the high-frequency increments $\Delta_i^n X = X_{\frac{i}{n}} - X_{\frac{i-1}{n}}$ are approximately Gaussian with conditional variance given by the value of the process σ_t^2 at the beginning of the increment. Of course, the stochastic volatility σ_t is not known and varies over time. Hence to test for the local Gaussianity of the high-frequency increments we first need to estimate locally σ_t and then divide the high-frequency increments by this estimate. To this end, we divide the interval $[0, 1]$ into blocks each of which contains k_n increments, for some deterministic sequence $k_n \to \infty$ with $k_n/n \to 0$. On each of the blocks our local estimator of σ_t^2 is given by

$$\widehat{V}_j^n = \frac{\pi}{2}\frac{n}{k_n - 1}\sum_{i=(j-1)k_n+2}^{jk_n} |\Delta_{i-1}^n X||\Delta_i^n X|, \quad j = 1, ..., \lfloor n/k_n \rfloor. \tag{3.1}$$

$\widehat{V}_j^n$ is the Bipower Variation proposed by [5, 6] for measuring the quadratic variation of the diffusion component of X. We note that an alternative measure of σ_t can be constructed using the so-called Truncated Variation. It turns out, however, that while the behavior of the two volatility measures in the case of the jump-diffusion model (1.1) is the same, it differs in the case when S_t is stable with $\beta < 2$. Using Truncated Variation will lead to degenerate limit of our statistic, unlike the case of using the Bipower Variation estimator in (3.1). For this reason we prefer the latter in our analysis but later in Section 6 we also derive in the jump-diffusion case the behavior of the statistic when Truncated Variation is used in its construction.

We use the first m_n increments on each block, with $m_n \leq k_n$, to test for local Gaussianity. The case $m_n = k_n$ amounts to using all increments in the block and we will need $m_n < k_n$ for deriving feasible CLT-s later on. Finally, we remove the high-frequency increments that contain "big" jumps. The total number of increments used in our statistic is thus given by

$$N^n(\alpha, \varpi) = \sum_{j=1}^{\lfloor n/k_n \rfloor} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1\left(|\Delta_i^n X| \leq \alpha\sqrt{\widehat{V}_j^n} n^{-\varpi}\right), \tag{3.2}$$

where $\alpha > 0$ an $\varpi \in (0, 1/2)$. We note that here we use a time-varying threshold in our truncation to account for the time-varying σ_t.

The scaling of every high-frequency increment will be done after adjusting

$\widehat{V}_j^n$ to exclude the contribution of that increment in its formation

$$\widehat{V}_j^n(i) = \begin{cases} \frac{k_n-1}{k_n-3}\widehat{V}_j^n - \frac{\pi}{2}\frac{n}{k_n-3}|\Delta_i^n X||\Delta_{i+1}^n X|, \text{ for } i = (j-1)k_n+1, \\ \frac{k_n-1}{k_n-3}\widehat{V}_j^n - \frac{\pi}{2}\frac{n}{k_n-3}\left(|\Delta_{i-1}^n X||\Delta_i^n X| + |\Delta_i^n X||\Delta_{i+1}^n X|\right), \\ \quad \text{for } i = (j-1)k_n+2, ..., jk_n-1, \\ \frac{k_n-1}{k_n-3}\widehat{V}_j^n - \frac{\pi}{2}\frac{n}{k_n-3}|\Delta_{i-1}^n X||\Delta_i^n X|, \text{ for } i = jk_n. \end{cases} \tag{3.3}$$

With this, we define

$$\widehat{F}_n(\tau) = \frac{1}{N^n(\alpha,\varpi)}\sum_{j=1}^{\lfloor n/k_n \rfloor}\sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1\left\{\frac{\sqrt{n}\Delta_i^n X}{\sqrt{\widehat{V}_j^n(i)}} \leq \tau\right\} 1_{\left\{|\Delta_i^n X| \leq \alpha\sqrt{\widehat{V}_j^n} n^{-\varpi}\right\}}, \tag{3.4}$$

which is simply the empirical cdf of the "devolatilized" increments that do not contain "big" jumps. In the jump-diffusion case of (1.1), $\widehat{F}_n(\tau)$ should be approximately the cdf of a standard normal random variable.

We note that all the results that follow for $\widehat{F}_n(\tau)$ will continue to hold if we do not truncate for the jumps in the construction of $\widehat{F}_n(\tau)$. The intuition for this is easiest to form in the case when X is a Lévy process without drift from the following $\mathbb{E}\left|1_{\{\sqrt{n}\Delta_i^n X \leq \tau\}} - 1_{\{\sqrt{n}\sigma\Delta_i^n W \leq \tau\}}\right| = O(n^{\beta'/2-1+\iota})$ for β' the constant of assumption A2 and $\iota > 0$ arbitrary small. Our rational for looking at the truncated increments only is that the order of magnitude of the above difference, i.e., the error due to the presence of jumps in X, can be slightly reduced by using truncation.

The construction of our statistic resembles the practice of standardizing increments of the process of fixed length by a measure for volatility constructed from high-frequency data within the interval (after correcting for jumps and leverage effect), see e.g. [2]. The main difference is that here the length of the increments that are standardized is shrinking and further the volatility estimator is local, i.e., over a shrinking time interval. Both these differences are crucial for deriving our feasible limit theory for $\widehat{F}_n(\tau)$.

4. Convergence in probability of $\widehat{F}_n(\tau)$. We next derive the limit behavior of $\widehat{F}_n(\tau)$ both under the null of model (1.1) as well as under a set of alternatives. We start with the case when X_t is given by (2.1).

THEOREM 1. *Suppose assumption A holds and assume the block size grows at the rate*

$$k_n \sim n^q, \quad \textit{for some } q \in (0,1), \tag{4.1}$$

and $m_n \to \infty$ *as* $n \to \infty$. *Then if* $\beta \in (1,2]$, *we have*

$$\widehat{F}_n(\tau) \xrightarrow{\mathbb{P}} F_\beta(\tau), \quad \textit{as } n \to \infty, \tag{4.2}$$

where the above convergence is uniform in τ over compact subsets of $\mathbb{R}$; $F_\beta(\tau)$ is the cdf of $\sqrt{\frac{2}{\pi}}\frac{S_1}{\mathbb{E}|S_1|}$ (S_1 is the value of the β-stable process S_t at time 1) and $F_2(\tau)$ equals the cdf of a standard normal variable $\Phi(\tau)$.

Since $\widehat{F}_n(\tau)$ and $F_\beta(\tau)$ are càdlàg and nondecreasing, the above result holds also uniformly on $\mathbb{R}$.

Remark 2. *The limit result in (4.2) shows that when S_t is stable with $\beta < 2$, $\widehat{F}_n(\tau)$ estimates the cdf of a β-stable random variable. We note that when $\beta < 2$, the correct scaling factor for the high-frequency increments is $n^{1/\beta}$. However, in this case we need also to scale $\widehat{V}_j^n$ by $n^{1/\beta-1/2}$ in order for the latter to converge to a non-degenerate limit (that is proportional to σ_t^2). Hence the ratio*

$$\frac{\sqrt{n}\Delta_i^n X}{\sqrt{\widehat{V}_j^n(i)}} = \frac{n^{1/\beta}\Delta_i^n X}{\sqrt{n^{2/\beta-1}\widehat{V}_j^n(i)}}, \tag{4.3}$$

is appropriately scaled even in the case when $\beta < 2$ and importantly without knowing apriori the value of β. We further note that the limiting cdf, $F_\beta(\tau)$, is of a random variable that has the same scale regardless of the value of β. That is, in all cases of β, $F_\beta(\tau)$ corresponds to the cdf of a random variable Z with $\mathbb{E}|Z| = \sqrt{\frac{2}{\pi}}$. Therefore, the difference between $\beta < 2$ and the null $\beta = 2$ will be in the relative probability assigned to "big" versus "small" values of τ.

We note further that in Theorem 1 we restrict $\beta > 1$. The reason is that for $\beta \leq 1$, the limit behavior of $\widehat{F}_n(\tau)$ is determined by the drift term in X (when present) and not S_t. To allow for $\beta \leq 1$ and still have a limit result of the type in (4.2), we need to use $\Delta_i^n X - \Delta_{i-1}^n X$ in the construction of $\widehat{F}_n(\tau)$ which essentially eliminates the drift term.

We next derive the limiting behavior of $\widehat{F}_n(\tau)$ in the situation when the Itô semimartingale X is "contaminated" by noise, which is of particular relevance in financial applications.

Theorem 2. *Suppose assumption A holds and $k_n \propto n^q$ for some $q \in (0,1)$ and $m_n \to \infty$ as $n \to \infty$. Let $\widehat{F}_n(\tau)$ be given by (3.4) with $\Delta_i^n X$ replaced with $\Delta_i^n X^*$ for $X^*_{\frac{i}{n}} = X_{\frac{i}{n}} + \epsilon_{\frac{i}{n}}$ and where $\left\{\epsilon_{\frac{i}{n}}\right\}_{i=1,\dots,n}$ are i.i.d. random variables defined on a product extension of the original probability space and independent from $\mathcal{F}$. Further, suppose $\mathbb{E}|\epsilon_{\frac{i}{n}}|^{1+\iota} < \infty$ for some $\iota > 0$. Finally, assume that the cdf of $\frac{1}{\mu}\left(\epsilon_{\frac{i}{n}} - \epsilon_{\frac{i-1}{n}}\right)$, $F_\epsilon(\tau)$, is continuous* *where we denote* $\mu = \sqrt{\frac{\pi}{2}}\sqrt{\mathbb{E}\left(|\epsilon_{\frac{i}{n}} - \epsilon_{\frac{i-1}{n}}||\epsilon_{\frac{i-1}{n}} - \epsilon_{\frac{i-2}{n}}|\right)}$. *Then*

$$\widehat{F}_n(\tau) \xrightarrow{\mathbb{P}} F_\epsilon(\tau), \qquad \text{as } n \to \infty, \tag{4.4}$$

where the above convergence is uniform in τ *over compact subsets of* $\mathbb{R}$.

Remark 3. *When* X *is observed with noise, the noise becomes the leading component at high-frequencies. Hence, our statistic recovers the cdf of the (appropriately scaled) noise component. Similar to the pure-jump alternative of* S_t *with* $\beta < 2$, *here* $\sqrt{n}$ *is not the right scaling for the increments* $\Delta_i^n X^*$ *but this is offset in the ratio in* $\widehat{F}_n(\tau)$ *by a scaling factor for the local variance estimator* $\widehat{V}_j^n$ *that makes it non-degenerate. Unlike the pure-jump alternative, in the presence of noise the correct scaling of the numerator and the denominator in the ratio in* $\widehat{F}_n(\tau)$ *is given by*

$$\frac{\sqrt{n}\Delta_i^n X^*}{\sqrt{\widehat{V}_j^n(i)}} = \frac{\Delta_i^n X^*}{\sqrt{n^{-1}\widehat{V}_j^n(i)}}, \tag{4.5}$$

that is, we need to scale down $\widehat{V}_j^n(i)$ *to ensure it converges to non-degenerate limit.*

The limit result in (4.4) provides an important insight into the noise by studying its distribution. We stress the fact that the presence of $\widehat{V}_j^n$ in the truncation is very important for the limit result in (4.4). This is because it ensures that the threshold is "sufficiently" big so that it does not matter in the asymptotic limit. If, on the other hand, the threshold did not contain $\widehat{V}_j^n$ (i.e., $\widehat{V}_j^n$ was replaced by 1 in the threshold), then in this case the limit will be determined by the behavior of the density of the noise around zero.

We finally note that when $\epsilon_{\frac{i}{n}}$ is normally distributed, a case that has received a lot of attention in the literature, the limiting cdf $F_\epsilon(\tau)$ is that of a centered normal but with variance that is below 1. Therefore, in this case $F_\epsilon(\tau)$ will be below the cdf of a standard normal variable, $\Phi(\tau)$, when $\tau < 0$ and the same relationship will apply to $1 - F_\epsilon(\tau)$ and $1 - \Phi(\tau)$ when $\tau > 0$.

On a more general level, the above results show that the empirical cdf estimator $\widehat{F}_n(\tau)$ can shed light on the potential sources of violation of the local Gaussianity of high-frequency data. It similarly can provide insights on the performance of various estimators that depend on this hypothesis.

5. CLT of $\widehat{F}_n(\tau)$ under Local Gaussianity.

THEOREM 3. *Let X_t satisfy (2.1) with S_t being a Brownian motion and assume that assumption B holds. Further, let the block size grow at the rate*

$$\frac{m_n}{k_n} \to 0, \quad k_n \sim n^q, \quad \text{for some } q \in (0, 1/2), \text{ when } n \to \infty, \tag{5.1}$$

such that $\frac{nm_n}{k_n^3} \to \lambda \geq 0$. We then have locally uniformly in subsets of $\mathbb{R}$

$$\begin{aligned} \widehat{F}_n(\tau) - \Phi(\tau) = \widehat{Z}_1^n(\tau) + \widehat{Z}_2^n(\tau) + \frac{1}{k_n} \frac{\tau^2 \Phi^{''}(\tau) - \tau\Phi'(\tau)}{8} \left(\left(\frac{\pi}{2}\right)^2 + \pi - 3 \right) \\ + o_p\left(\frac{1}{k_n}\right), \end{aligned} \tag{5.2}$$

$$\left(\sqrt{\lfloor n/k_n \rfloor m_n} \widehat{Z}_1^n(\tau) \quad \sqrt{\lfloor n/k_n \rfloor k_n} \widehat{Z}_2^n(\tau) \right) \xrightarrow{\mathcal{L}} \left(Z_1(\tau) \quad Z_2(\tau) \right), \tag{5.3}$$

where $\Phi(\tau)$ is the cdf of a standard normal variable and $Z_1(\tau)$ and $Z_2(\tau)$ are two independent Gaussian processes with covariance functions

$$\begin{aligned} &Cov(Z_1(\tau_1), Z_1(\tau_2)) = \Phi(\tau_1 \wedge \tau_2) - \Phi(\tau_1)\Phi(\tau_2), \\ &Cov(Z_2(\tau_1), Z_2(\tau_2)) = \left[\frac{\tau_1 \Phi'(\tau_1)}{2} \frac{\tau_2 \Phi'(\tau_2)}{2} \right] \left(\left(\frac{\pi}{2}\right)^2 + \pi - 3 \right), \quad \tau_1, \tau_2 \in \mathbb{R}. \end{aligned} \tag{5.4}$$

Due to the “big” jumps, we derive the CLT only on compact sets of τ since the error in the estimation of the cdf for $\tau \to \pm\infty$ is affected by the truncation.

We make several observations regarding the limiting result in (5.2)-(5.4). The first term of $\widehat{F}_n(\tau) - \Phi(\tau)$ in (5.2), $\widehat{Z}_1^n(\tau)$, converges to $Z_1(\tau)$ which is the standard Brownian bridge appearing in the Donsker theorem for empirical processes, see e.g., [21]. The second and third term on the right-hand side of (5.2) are due to the estimation error in recovering the local variance, i.e., the presence of $\widehat{V}_j^n$ in $\widehat{F}^n(\tau)$ instead of the true (unobserved) σ_t^2. $\widehat{Z}_2^n(\tau)$ converges to a centered Gaussian process, independent from $Z_1(\tau)$, while the third term on the right-hand side of (5.2) is an asymptotic bias. Importantly, the asymptotic bias as well as the variance of $(Z_1(\tau)\ Z_2(\tau))$ are all constants that depend only on τ and not the stochastic volatility σ_t. Therefore, feasible inference based on (5.2) is straightforward.

We note that by picking the rate of growth of m_n and k_n arbitrary close to $\sqrt{n}$, we can make the rate of convergence of $\widehat{F}_n(\tau)$ arbitrary close to $\sqrt{n}$. We should further point out that this is unlike the rate of estimating the spot σ_t^2 by $\widehat{V}_j^n$ (with the same choice of k_n) which is at most $n^{1/4}$. The reason for the better rate of convergence of our estimator is in the integration of the error due to the estimation $\widehat{V}_j^n$.

The order of magnitude of the three components on the right-hand side of (5.2) are different with the second term always dominated by the other two. Its presence should provide a better finite-sample performance of a test based on (5.2).

Finally, we point out that a feasible CLT for $\widehat{F}_n(\tau)$ is available with "only" arbitrarily close to $\sqrt{n}$ rate of convergence and not exactly $\sqrt{n}$. This is due to the presence of the drift term in X. The latter leads to asymptotic bias which is of order $1/\sqrt{n}$ and removing it via de-biasing is in general impossible as we cannot estimate the latter from high-frequency record of X.

6. Empirical CDF of "Devolatilized" High Frequency Increments with an Alternative Volatility Estimator. As mentioned in Section 3, an alternative estimator of the volatility is the Truncated Variation of [12] defined as

$$\widehat{C}_j^n = \frac{n}{k_n} \sum_{i=(j-1)k_n+1}^{jk_n} |\Delta_i^n X|^2 1\left(|\Delta_i^n X| \leq \alpha n^{-\varpi}\right), \quad j = 1, ..., \lfloor n/k_n \rfloor, \tag{6.1}$$

where $\alpha > 0$ and $\varpi \in (0, 1/2)$ and the corresponding one excluding the contribution of the i-th increment, for $i = (j-1)k_n + 1, ..., jk_n$, is

$$\widehat{C}_j^n(i) = \frac{k_n}{k_n - 1}\widehat{C}_j^n - \frac{n}{k_n - 1}|\Delta_i^n X|^2 1\left(|\Delta_i^n X| \leq \alpha n^{-\varpi}\right). \tag{6.2}$$

We define the corresponding empirical cdf of the "devolatilized" (and truncated) high-frequency increments as

$$\widehat{F}_n'(\tau) = \frac{1}{N'^n(\alpha,\varpi)} \sum_{j=1}^{\lfloor n/k_n \rfloor} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1\left\{\frac{\sqrt{n}\Delta_i^n X}{\sqrt{\widehat{C}_j^n(i)}} \leq \tau\right\} 1_{\{|\Delta_i^n X| \leq \alpha n^{-\varpi}\}}, \tag{6.3}$$

where for $\alpha > 0$ and $\varpi \in (0, 1/2)$

$$N'^n(\alpha,\varpi) = \sum_{j=1}^{\lfloor n/k_n \rfloor} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1\left(|\Delta_i^n X| \leq \alpha n^{-\varpi}\right). \tag{6.4}$$

In the next theorem we derive a CLT for $\widehat{F}_n'(\tau)$ when X is a jump-diffusion.

THEOREM 4. *Let X_t satisfy (2.1) with S_t being a Brownian motion and assume that assumption B holds. Let k_n and m_n satisfy (5.1). We then have locally uniformly in subsets of $\mathbb{R}$*

$$\widehat{F}_n'(\tau) - \Phi(\tau) = \widehat{Z}_1^n(\tau) + \widehat{Z}_2^n(\tau) + \frac{1}{k_n}\frac{\tau^2\Phi''(\tau) - \tau\Phi'(\tau)}{4} + o_p\left(\frac{1}{k_n}\right), \tag{6.5}$$

$$\left(\sqrt{\lfloor n/k_n \rfloor m_n}\widehat{Z}_1^n(\tau) \quad \sqrt{\lfloor n/k_n \rfloor k_n}\widehat{Z}_2^n(\tau)\right) \xrightarrow{\mathcal{L}} (Z_1(\tau) \quad Z_2(\tau)), \tag{6.6}$$

where $\Phi(\tau)$ is the cdf of a standard normal variable and $Z_1(\tau)$ and $Z_2(\tau)$ are two independent Gaussian processes with covariance functions

$$\begin{aligned} Cov(Z_1(\tau_1), Z_1(\tau_2)) &= \Phi(\tau_1 \wedge \tau_2) - \Phi(\tau_1)\Phi(\tau_2), \\ Cov(Z_2(\tau_1), Z_2(\tau_2)) &= [\tau_1\Phi'(\tau_1)\tau_2\Phi'(\tau_2)], \quad \tau_1, \tau_2 \in \mathbb{R}. \end{aligned} \tag{6.7}$$

Further, in the case when α_t, σ_t and $\delta^Y(t,x)$ do not depend on t, the above result continues to hold even when assumption B4 is replaced with the weaker condition $\int_E [|\delta^Y(x)|^{\beta'} \wedge 1]\nu(dx) < \infty$ for some $0 \le \beta' < 1$, provided for $\iota > 0$ arbitrary small, we have

$$k_n\left(n^{1-(4-\beta')\varpi} \vee n^{-\frac{1-\beta'/2}{1+\beta'}+\iota} \vee n^{-\frac{2}{3}(2-\beta')\varpi+\iota}\right) \to 0, \quad \frac{k_n^3}{m_n} n^{-(4-\beta')\varpi} \to 0. \tag{6.8}$$

The CLT for $\widehat{F}'_n(\tau)$ is similar to that for $\widehat{F}_n(\tau)$ with the only difference being that the asymptotic bias (the third term on the right side of (6.5)) and the limiting Gaussian process Z_2 are of smaller magnitude and with smaller variance respectively. This is not surprising as the Truncated Variation is known to be a more efficient estimator of volatility than the Bipower Variation.

The last part of the theorem shows that in the case when α_t, σ_t and $\delta^Y(t,x)$ do not depend on t, the CLT result continues to hold in presence of jumps of infinite activity (but finite variation) provided the growth condition (6.8) holds. This condition can be simplified when one uses a value for ϖ arbitrary close to $1/2$ (as is common) and m_n close to k_n.

7. Test for Local Gaussianity of High-Frequency Data. We proceed with a feasible test for a jump-diffusion model of the type given in (1.1) using the developed limit theory above. We will use $\widehat{F}_n(\tau)$ for this. The critical region of our proposed test is given by

$$C_n = \left\{\sup_{\tau \in \mathcal{A}} \sqrt{N^n(\alpha, \varpi)}|\widehat{F}_n(\tau) - \Phi(\tau)| > q_n(\alpha, \mathcal{A})\right\} \tag{7.1}$$

where recall $\Phi(\tau)$ denotes the cdf of a standard normal random variable, $\alpha \in (0,1)$, $\mathcal{A} \in \mathbb{R}$ is a finite union of compact sets with positive Lebesgue measure, and $q_n(\alpha, \mathcal{A})$ is the $(1-\alpha)$-quantile of

$$\sup_{\tau \in \mathcal{A}}\left| Z_1(\tau) + \sqrt{\frac{m_n}{k_n}} Z_2(\tau) + \sqrt{\frac{m_n}{k_n}} \frac{\sqrt{n}}{k_n} \frac{\tau^2\Phi''(\tau) - \tau\Phi'(\tau)}{8}\left(\left(\frac{\pi}{2}\right)^2 + \pi - 3\right)\right|, \tag{7.2}$$

with $Z_1(\tau)$ and $Z_2(\tau)$ being the Gaussian processes defined in Theorem 3. We can easily evaluate $q_n(\alpha, \mathcal{A})$ via simulation.

We note that in (7.1) we use $N^n(\alpha, \varpi)$ as a normalizing constant. This is justified because we have $\frac{N^n(\alpha,\varpi)}{\lfloor n/k_n \rfloor m_n} \xrightarrow{\mathbb{P}} 1$, both in the jump-diffusion case as well as in the two alternative scenarios considered in Section 4. The choice of k_n and m_n in general should be dictated by how much volatility of volatility in X we have. We illustrate this in the next section.

The test in (7.1) resembles a Kolmogorov-Smirnov type test for equality of continuous one-dimensional distributions. There are two differences between our test and the original Kolmogorov-Smirnov test. First, in our test we scale the high-frequency increments by a nonparametric local estimator of the volatility and this has an asymptotic effect on the test statistic, as evident from Theorem 3. The second difference is in the region $\mathcal{A}$ over which the difference $\widehat{F}_n(\tau) - \Phi(\tau)$ is evaluated. For reasons we already discussed, that are particular to our problem here, we need to exclude arbitrary high in magnitude values of τ.

Now, in terms of the size and power of the test, under assumptions A and B, using Theorem 1 and Theorem 3, we have

$$\lim_n \mathbb{P}\left(C_n\right) = \alpha, \quad \text{if } \beta = 2 \quad \text{and} \quad \liminf_n \mathbb{P}\left(C_n\right) = 1, \quad \text{if } \beta \in (1,2), \tag{7.3}$$

where we make also use of the fact that the stable and standard normal variables have different cdf-s on compact subsets of $\mathbb{R}$ with positive Lebesgue measure. By Theorem 2, the above power result applies also to the case when we observe $X_{\frac{i}{n}} + \epsilon_{\frac{i}{n}}$, provided of course the limiting cdf of the noise in (4.4) differs from that of the standard normal on the set $\mathcal{A}$.

We note that existing tests for presence of diffusive component in X are based only on the scaling factor of the high-frequency increments on the left side of (2.5). However, the limiting result in (2.5) implies much more. Mainly, the distribution of the "devolatilized" increments should be stable (and in particular normal in the jump-diffusion case). Our test in (7.1), unlike earlier work, incorporates this distribution implication of (2.5) as well.

We finally point out that using Theorem 3, one should be able to derive alternative tests for the presence of diffusive component in X, by adopting other measures of discrepancy between distributions like the Cramer-von Mises one.

8. Monte Carlo. We now evaluate the performance of our test on simulated data. We consider the following two models. The first is

$$dX_t = \sqrt{V_t} dW_t + \int_{\mathbb{R}} x\mu(ds, dx), \quad dV_t = 0.03(1.0 - V_t)dt + 0.1\sqrt{V_t} dB_t, \tag{8.1}$$

where (W_t, B_t) is a vector of Brownian motions with $\mathrm{Corr}(W_t, B_t) = -0.5$ and μ is a homogenous Poisson measure with compensator $\nu(dt, dx) = dt \otimes \frac{0.25e^{-|x|/0.4472}}{0.4472}dx$ which corresponds to double exponential jump process with intensity of 0.5 (i.e., a jump every second day on average). This model is calibrated to financial data by setting the means of continuous and jump variation similar to those found in earlier empirical work. Similarly, we allow for dependence between X_t and V_t, i.e., leverage effect. The second model is given by

$$X_t = S_{T_t}, \quad \text{with} \quad T_t = \int_0^t V_s ds, \tag{8.2}$$

where S_t is a symmetric tempered stable martingale with Lévy measure $\frac{0.1089e^{-|x|}}{|x|^{1+1.8}}$ and V_t is the square-root diffusion given in (8.1). The process in (8.2) is a time-changed tempered stable process. The parameters of S_t are chosen such that it behaves locally like 1.8-stable process and it has variance at time 1 equal to 1 (as the model in (8.1)). For this process the local Gaussianity does not hold and hence the behavior of the test on data from the model in (8.2) will allow us to investigate the power of the test. We also consider another alternative to the jump-diffusion, mainly the case when the process in (8.2) is contaminated with i.i.d. Gaussian noise. The variance of the noise is set to 0.01 consistent with empirical evidence in [8].

We turn next to the implementation of the test. We apply the test on one year worth of simulated data which consists of 252 days (our unit of time is one trading day). We consider two sampling frequencies: $n = 100$ and $n = 200$ which correspond to sampling every 5 and 2 minutes respectively in a typical trading day. We experiment with 1-4 blocks per day. In each block we use 75% or 70% of the increments in the formation of the test, i.e., we set $\lfloor m_n/k_n \rfloor = 0.75$ for $n = 100$ and $\lfloor m_n/k_n \rfloor = 0.70$ for $n = 200$. We found very little sensitivity of the test with respect to the choice of the ratio m_n/k_n. For the truncation of the increments, as typical in the literature, we set $\alpha = 3.0$ and $\varpi = 0.49$. Finally, the set $\mathcal{A}$ over which the difference $\widehat{F}_n(\tau) - F(\tau)$ in our test is evaluated is set to

$$\mathcal{A} = [Q(0.01) : Q(0.40)] \ \cup \ [Q(0.60) : Q(0.99)], \tag{8.3}$$

where $Q(\alpha)$ is the α-quantile of standard normal.

The results of the Monte Carlo are reported in Tables 1-3. For the smaller sample size, $n = 100$, and with no blocking at all ($k_n = n$) to account for volatility movements over the day, there are size distortions most noticeable at the conventional 5 percent level. With two blocks ($\lfloor n/k_n \rfloor = 2$), size is appropriate, while it is seen to have excellent power in Tables 2 and 3. But with three blocks on $n = 100$, there are size distortions because the noisy estimates of local volatility distort the test. Considering the larger sample size ($n = 200$), now with three blocks the test's size is approximately correct while power is excellent. For larger values of k_n relative to n ($\lfloor n/k_n \rfloor = 1$) the time variation in volatility over the day coupled with the relatively high precision of estimating a biased version of local volatility, leads to departures from Gaussianity of the (small) scaled increments and hence the over-rejections.

In Tables 1-3 we also report the performance on the simulated data of a test for presence of Brownian motion in high frequency data based on (truncated) power variations computed on two different frequencies, proposed in [1] (see also [18]). This test, unlike the test proposed here, does not exploit the distributional implication of the local Gaussianity result in (1.2). We can see from Table 1 that the test based on the power variations has reasonable behavior under the null of presence of a diffusion component in X. Table 2 further shows that for the optimal choice of the power ($p = 1$), the test has slightly lower power against the considered pure-jump alternative in (8.2) than the Kolmogorov-Smirnov test (when block size is chosen optimally).

When the pure-jump model is contaminated with noise, the scaling of the power variations is similar (for the considered frequencies) to that of a jump-diffusion model observed without noise. Hence, Table 3 reveals relatively low power of the test based on the power variations against the alternative of pure-jump process contaminated with noise. By contrast, the Kolmogorov-Smirnov test shows almost no change in performance compared with the alternative when the pure-jump process is observed without noise (Table 2). The reason is that the Kolmogorov-Smirnov test incorporates also the distributional implications of (1.2) and, under the pure-jump plus noise scenario, the scaled high-frequency increments have a distribution which is very different from standard normal.

9. Empirical Illustration. We now apply our test to two different financial assets, the IBM stock price and the VIX volatility index. The analyzed period is $2003 - 2008$ and like in the Monte Carlo we consider two and five minute sampling frequencies. The test is performed for each of the years in the sample. We set $\mathcal{A}$ as in (8.3) and $\lfloor n/k_n \rfloor = 3$ for the five-minute sampling frequency and $\lfloor n/k_n \rfloor = 4$ for the two-minute frequency. As in the Monte Carlo, the ratio $\lfloor m_n/k_n \rfloor$ is set to 0.75 and 0.70 for the five-minute and two-minute respectively sampling frequencies. Finally, to account for the well-known diurnal pattern in volatility we standardize the raw high-

Table 1
Monte Carlo Results for Jump-Diffusion Model (8.1)

	Kolmogorov-Smirnov Test			**Power Variation based Test**	
Nominal Size	**Rejection Rate**				
	Sampling Frequency $n = 100$				
	$k_n = 33$	$k_n = 50$	$k_n = 100$	$p = 1.0$	$p = 1.5$
$\alpha = 1\%$	0.0	0.8	5.6	1.5	0.7
$\alpha = 5\%$	0.4	4.3	16.8	7.7	5.4
	Sampling Frequency $n = 200$				
	$k_n = 50$	$k_n = 67$	$k_n = 200$	$p = 1.0$	$p = 1.5$
$\alpha = 1\%$	0.4	0.9	10.3	1.4	1.3
$\alpha = 5\%$	1.2	3.2	32.8	8.1	6.7

Note: For the cases with $n = 100$ we set $\lfloor m_n/k_n \rfloor = 0.75$ and for the cases with $n = 200$ we set $\lfloor m_n/k_n \rfloor = 0.70$. The power variation test is a one-sided test based on Theorem 2 in [1] with $k = 2$ and cutoff $u_n = 7\widehat{\sigma}\Delta_n^{0.49}$ with $\widehat{\sigma}$ being an estimate of volatility over the day using Bipower Variation.

frequency returns by a time-of-day scale factor exactly as in [20].

The results from the test are shown on Figure 1. We can see from the figure that the local Gaussianity hypothesis works relatively well for the 5-minute IBM returns. At 2-minute sampling frequency for the IBM stock price, however, our test rejects the local Gaussianity hypothesis at conventional significance levels. Nevertheless, the values of the test are not very far from the critical ones. The explanation of the different outcomes of the test on the two sampling frequencies is to be found in the presence of microstructure noise. The latter becomes more prominent at the higher frequency. Turning to the VIX index data, we see a markedly different outcome. For this data set, the local Gaussianity hypothesis is strongly rejected at both frequencies. The explanation for this is that the underlying model is of pure-jump type, i.e., the model (2.1) with $\beta < 2$.

10. Proofs. We start with introducing some notation that we will make use of in the proofs.

$$A_t = \int_0^t \alpha_s ds, \quad B_t = \int_0^t \sigma_s dS_s, \quad \overline{\sigma}_t = \sigma_t - \sum_{s \leq t} \Delta\sigma_s.$$

$$\dot{V}_j^n = \frac{n}{k_n - 1}\frac{\pi}{2} \sum_{i=(j-1)k_n+2}^{jk_n} |\Delta_{i-1}^n A + \Delta_{i-1}^n B||\Delta_i^n A + \Delta_i^n B|,$$

Table 2

Monte Carlo Results for Pure-Jump Model (8.2)

	Kolmogorov-Smirnov Test			Power Variation based Test	
Nominal Size	**Rejection Rate**				
	Sampling Frequency $n = 100$				
	$k_n = 33$	$k_n = 50$	$k_n = 100$	$p = 1.0$	$p = 1.5$
$\alpha = 1\%$	45.9	95.2	99.9	71.1	14.2
$\alpha = 5\%$	76.6	99.6	100.0	91.3	38.4
	Sampling Frequency $n = 200$				
	$k_n = 50$	$k_n = 67$	$k_n = 200$		
$\alpha = 1\%$	100.0	100.0	100.0	97.7	32.1
$\alpha = 5\%$	100.0	100.0	100.0	99.4	63.3

Note: Notation as in Table 1.

$$\widetilde{V}_j^n = \frac{n}{k_n - 1}\frac{\pi}{2}\sum_{i=(j-1)k_n+2}^{jk_n} |\Delta_{i-1}^n B||\Delta_i^n B|,\quad \overline{V}_j^n = \sigma^2_{\frac{(j-1)k_n}{n}}\frac{n}{k_n - 1}\frac{\pi}{2}\sum_{i=(j-1)k_n+2}^{jk_n} |\Delta_{i-1}^n S||\Delta_i^n S|,$$

and we define $\dot{V}_j^n(i)$, $\widetilde{V}_j^n(i)$ and $\overline{V}_j^n(i)$ from the above as in (3.3). We also denote

$$\widetilde{F}_n(\tau) = \frac{N^n(\alpha, \varpi)}{\lfloor n/k_n \rfloor m_n}\widehat{F}_n(\tau). \tag{10.1}$$

Finally, in the proofs we will denote with K a positive constant that might change from line to line but importantly does not depend on n and τ. We will also use the shorthand notation $\mathbb{E}_i^n(\cdot) = \mathbb{E}\left(\cdot|\mathcal{F}_{\frac{(i-1)}{n}}\right)$.

10.1. *Localization.* We will prove Theorems 1-4 under the following stronger versions of assumption A and B:

SA. *We have assumption A with α_t, σ_t and σ_t^{-1} being all uniformly bounded on $[0,1]$. Further, (2.3) and (2.4) hold for σ_t and Y_t respectively.*

SB. *We have assumption B with all processes α_t, $\widetilde{\alpha}_t$, σ_t, σ_t^{-1}, $\widetilde{\sigma}_t$, $\widetilde{\sigma}'_t$ and the coefficients of the Itô semimartingale representations of $\widetilde{\sigma}_t$ and $\widetilde{\sigma}'_t$ being uniformly bounded on $[0,1]$. Further $(|\delta^Y(t,x)| + |\delta^\sigma(t,x)| + |\delta^{\widetilde{\sigma}}(t,x)| + |\delta^{\widetilde{\sigma}'}(t,x)|) \leq \gamma(x)$ for some non-negative valued function $\gamma(x)$ on E satisfying $\int_E \nu(x : \gamma(x) = 0)dx < \infty$ and $\gamma(x) \leq K$ for some constant K.*

Extending the proofs to the weaker assumptions A and B follows by standard localization techniques exactly as Lemma 4.4.9 of [9].

TABLE 3
Monte Carlo Results for Pure-Jump Model (8.2) plus Noise

	Kolmogorov-Smirnov Test			Power Variation based Test	
Nominal Size	**Rejection Rate**				
	Sampling Frequency $n = 100$				
	$k_n = 33$	$k_n = 50$	$k_n = 100$	$p = 1.0$	$p = 1.5$
$\alpha = 1\%$	39.0	91.4	99.8	21.7	1.6
$\alpha = 5\%$	70.9	99.6	100.0	46.2	5.9
	Sampling Frequency $n = 200$				
	$k_n = 50$	$k_n = 67$	$k_n = 200$		
$\alpha = 1\%$	100.0	100.0	100.0	6.5	0.0
$\alpha = 5\%$	100.0	100.0	100.0	17.3	0.0

Note: Notation as in Table 1.

10.2. *Proof of Theorem 1.* Without loss of generality, we will assume that $\tau < 0$, the case $\tau \geq 0$ being dealt with analogously (by working with $1 - \widehat{F}_n(\tau)$ instead). We first analyze the behavior of $\widehat{V}_j^n$. We denote with η_n a deterministic sequence that depends only on n and vanishes as $n \to \infty$.

Using the triangular inequality, the Chebyshev inequality, successive conditioning, as well as Hölder inequality and assumption SA, we get for $j = 1, ..., \lfloor n/k_n \rfloor$

$$\mathbb{P}\left(n^{2/\beta-1}|\widehat{V}_j^n - \dot{V}_j^n| \geq \eta_n\right) \leq K\frac{n^{1/\beta-(1/\beta')\wedge 1+\iota}}{\eta_n}, \quad \forall \iota > 0.$$

Similarly, using triangular inequality, Chebyshev inequality as well as Hölder inequality, we get for $j = 1, ..., \lfloor n/k_n \rfloor$

$$\mathbb{P}\left(n^{2/\beta-1}|\dot{V}_j^n - \widetilde{V}_j^n| \geq \eta_n\right) \leq K\frac{n^{1/\beta-1+\iota}}{\eta_n}, \quad \forall \iota > 0.$$

Next, using, triangular inequality, Chebyshev inequality, Hölder inequality, Burkholder-Davis-Gundy inequality as well as assumption SA, we get for $j = 1, ..., \lfloor n/k_n \rfloor$

$$\mathbb{P}\left(n^{2/\beta-1}|\widetilde{V}_j^n - \overline{V}_j^n| \geq \eta_n\right) \leq K\frac{k_n^{1/2-\iota}}{n^{1/2-\iota}\eta_n}, \quad \forall \iota > 0.$$

Finally, using the self-similarity of the stable process and Burkholder-Davis-Gundy inequality (for discrete martingales), we get for $j = 1, ..., \lfloor n/k_n \rfloor$

$$\mathbb{P}\left(|n^{2/\beta-1}\overline{V}_j^n - \frac{\pi}{2}\sigma^2_{\frac{(j-1)k_n}{n}}(\mathbb{E}|S_1|)^2| \geq \eta_n\right) \leq K\frac{1}{k_n^{\beta-1-\iota}\eta_n^{\beta-\iota}}, \quad \forall \iota \in (0, 1-\beta).$$

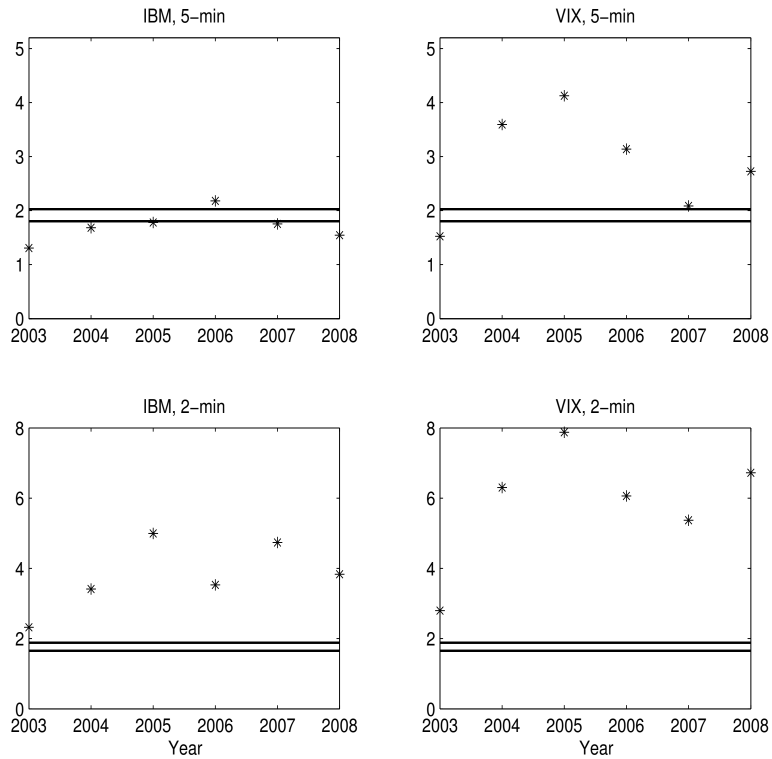

FIG 1. *Kolmogorov-Smirnov tests for local Gaussianity. The* $*$ *corresponds to the value of the test* $\sup_{\tau\in A}\sqrt{N^n(\alpha,\varpi)}|\widehat{F}_n(\tau)-F(\tau)|$ *and the solid lines are the critical values* $q_n(\alpha,\mathcal{A})$ *for* $\alpha=5\%$ *and* $\alpha=1\%$.

Combining these results, we get altogether for $\forall\iota\in(0,1-\beta)$

$$\tag{10.2}
\begin{aligned}
&\mathbb{P}\left(|n^{2/\beta-1}\widehat{V}_j^n-\frac{\pi}{2}\sigma^2_{\frac{(j-1)k_n}{n}}(\mathbb{E}|S_1|)^2|\geq\eta_n\right)\\
&\qquad\qquad\leq K\left(\frac{n^{1/\beta-(1/\beta')\wedge 1+\iota}}{\eta_n}\bigvee\frac{k_n^{1/2-\iota}}{n^{1/2-\iota}\eta_n}\bigvee\frac{1}{k_n^{\beta-1-\iota}\eta_n^{\beta-\iota}}\right).
\end{aligned}$$

Using the same proofs we can show that the result above continues to hold when $\widehat{V}_j^n$ is replaced with $\widehat{V}_j^n(i)$.

Next, for $i=(j-1)k_n+1...,(j-1)k_n+m_n$ and $j=1,...,\lfloor n/k_n\rfloor$, we denote

$$\begin{cases} \xi^n_{i,j}(1) = n^{1/\beta}\left(\Delta^n_i A + \Delta^n_i Y + \int_{(i-1)\Delta_n}^{i\Delta_n}(\sigma_{u-} - \sigma_{\frac{(j-1)k_n}{n}})dS_u\right) \\ \xi^n_{i,j}(2) = n^{1/\beta}\sigma_{\frac{(j-1)k_n}{n}}\Delta^n_i S 1_{\left\{|\Delta^n_i X| > \alpha\sqrt{\widehat{V}^n_j}n^{-\varpi}\right\}} \end{cases}.$$

With this notation, using similar inequalities as before, we get

$$\mathbb{P}\left(|\xi^n_{i,j}(1)| \geq \eta_n\right) \leq K\left(\frac{n^{1/\beta-(1/\beta')\wedge 1+\iota}}{\eta_n}\bigvee \frac{k_n^{\beta/2+\iota/2}}{n^{\beta/2+\iota/2}\eta_n^{\beta+\iota}}\right). \tag{10.3}$$

Next, using the result in (10.2) above as well as Hölder inequality, we get

$$\mathbb{P}\left(|\xi^n_{i,j}(2)| \geq \eta_n\right) \leq K\frac{n^{-(1/2-\varpi)\beta+\iota} \vee n^{1/\beta-(1/\beta')\wedge 1+\iota} \vee (k_n/n)^{1/2-\iota} \vee k_n^{1+\iota-\beta}}{\eta_n^{\iota}}, \quad \forall \iota > 0. \tag{10.4}$$

We next denote the set (note that by assumption SA, σ_t is strictly above zero on the time interval $[0,1]$)

$$\mathcal{A}^n_{i,j} = \left\{\omega : \frac{|\xi^n_{i,j}(1)| + |\xi^n_{i,j}(2)|}{\sqrt{\frac{\pi}{2}}\mathbb{E}|S_1|} > \eta_n \ \cup \ \left|\frac{n^{1/\beta-1/2}\sqrt{\widehat{V}^n_j(i)}}{\sqrt{\frac{\pi}{2}}\sigma_{\frac{(j-1)k_n}{n}}\mathbb{E}|S_1|} - 1\right| > \eta_n\right\}, \tag{10.5}$$

for $i = (j-1)k_n + 1, ..., (j-1)k_n + m_n$ and $j = 1, ..., \lfloor n/k_n \rfloor$.

We now can set (recall (4.1))

$$\eta_n = n^{-x}, \quad 0 < x < \left[\left(\frac{1}{\beta'} \wedge 1 - \frac{1}{\beta}\right) \bigwedge \frac{1-q}{2} \bigwedge \frac{q(\beta-1)}{\beta}\right], \tag{10.6}$$

and this choice is possible because of the restriction on the rate of increase of the block size k_n relative to n given in (4.1). With this choice of η_n, the results in (10.2), (10.3) and (10.4) imply

$$\frac{1}{\lfloor n/k_n \rfloor m_n}\sum_{j=1}^{\lfloor n/k_n \rfloor}\sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n}\mathbb{P}\left(\mathcal{A}^n_{i,j}\right) = o(1). \tag{10.7}$$

Therefore, for any compact subset $\mathcal{A}$ of $(-\infty, 0)$,

$$\sup_{\tau\in\mathcal{A}}|\widetilde{F}_n(\tau) - \widehat{G}_n(\tau)| = o_p(1), \tag{10.8}$$

where we denote

$$\widehat{G}_n(\tau) = \frac{1}{\lfloor n/k_n \rfloor m_n}\sum_{j=1}^{\lfloor n/k_n \rfloor}\sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1\left\{\frac{\sqrt{n}\Delta^n_i X}{\sqrt{\widehat{V}^n_j(i)}}1_{\left\{|\Delta^n_i X| \leq \alpha\sqrt{\widehat{V}^n_j}n^{-\varpi}\right\}} \leq \tau\right\}1_{\{(\mathcal{A}^n_{i,j})^c\}}.$$

Taking into account the definition of the set $\mathcal{A}^n_{i,j}$, we get

$$\begin{cases} \widehat{G}_n(\tau) \geq \frac{1}{\lfloor n/k_n \rfloor m_n} \sum_{j=1}^{\lfloor n/k_n \rfloor} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1\left\{ \frac{n^{1/\beta}\Delta_i^n S}{\sqrt{\frac{\pi}{2}}\mathbb{E}|S_1|} \leq \tau(1-\eta_n) - \eta_n \right\}, \\ \widehat{G}_n(\tau) \leq \frac{1}{\lfloor n/k_n \rfloor m_n} \sum_{j=1}^{\lfloor n/k_n \rfloor} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1\left\{ \frac{n^{1/\beta}\Delta_i^n S}{\sqrt{\frac{\pi}{2}}\mathbb{E}|S_1|} \leq \tau(1+\eta_n) + \eta_n \right\}. \end{cases}$$

Using Glivenko-Cantelli theorem, see e.g., Theorem 19.1 of [21], we have

$$\sup_\tau \left| \frac{1}{\lfloor n/k_n \rfloor m_n} \sum_{j=1}^{\lfloor n/k_n \rfloor} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1\left\{ \frac{n^{1/\beta}\Delta_i^n S}{\sqrt{\frac{\pi}{2}}\mathbb{E}|S_1|} \leq \tau(1-\eta_n) - \eta_n \right\} \right.$$
$$\left. - F_\beta(\tau(1-\eta_n)-\eta_n) \right| \stackrel{\mathbb{P}}{\longrightarrow} 0,$$

$$\sup_\tau \left| \frac{1}{\lfloor n/k_n \rfloor m_n} \sum_{j=1}^{\lfloor n/k_n \rfloor} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1\left\{ \frac{n^{1/\beta}\Delta_i^n S}{\sqrt{\frac{\pi}{2}}\mathbb{E}|S_1|} \leq \tau(1+\eta_n) + \eta_n \right\} \right.$$
$$\left. - F_\beta(\tau(1+\eta_n)+\eta_n \right| \stackrel{\mathbb{P}}{\longrightarrow} 0,$$

and further using the smoothness of cdf of the stable distribution we have

$$\sup_\tau |F_\beta(\tau(1-\eta_n)-\eta_n) - F_\beta(\tau)| \longrightarrow 0, \quad \sup_\tau |F_\beta(\tau(1+\eta_n)+\eta_n) - F_\beta(\tau)| \longrightarrow 0.$$

These two results altogether imply

$$\sup_\tau |\widehat{G}_n(\tau) - F_\beta(\tau)| \stackrel{\mathbb{P}}{\longrightarrow} 0,$$

and from here, using (10.8), we have $\sup_{\tau\in\mathcal{A}} |\widetilde{F}_n(\tau) - F_\beta(\tau)| = o_p(1)$ for any compact subset $\mathcal{A}$ of $(-\infty, 0)$. Hence, to prove (4.2), we need only to show

$$\frac{N^n(\alpha,\varpi)}{\lfloor n/k_n \rfloor m_n} \stackrel{\mathbb{P}}{\longrightarrow} 1, \quad \text{as } n\to\infty. \tag{10.9}$$

We have

$$\mathbb{P}\left(|\Delta_i^n X| > \alpha\sqrt{\widehat{V}_j^n} n^{-\varpi}\right) \leq \mathbb{P}\left(\left| \frac{n^{1/\beta-1/2}\sqrt{\widehat{V}_j^n}}{\sqrt{\frac{\pi}{2}}\sigma_{\frac{(j-1)k_n}{n}}\mathbb{E}|S_1|} - 1 \right| > 0.5 \right)$$
$$+ \mathbb{P}\left(n^{1/\beta}|\Delta_i^n X| > 0.5\alpha\sqrt{\frac{\pi}{2}}\sigma_{\frac{(j-1)k_n}{n}}\mathbb{E}|S_1| n^{1/2-\varpi} \right).$$

From here we can use the bounds in (10.2) and (10.3) to conclude

$$\mathbb{P}\left(|\Delta_i^n X| > \alpha\sqrt{\widehat{V}_j^n} n^{-\varpi}\right) \leq \frac{K}{n^\iota}, \quad \text{for some sufficiently small } \iota > 0, \tag{10.10}$$

hence the convergence in (10.9) holds which implies the result in (4.2). $\square$

10.3. *Proof of Theorem 2.* The proof follows the same steps as that of Theorem 1. We denote with η_n a deterministic sequence depending only on n and vanishing as $n \to \infty$. Then, using triangular inequality and successive conditioning, we have

$$\mathbb{P}\left(\left|\frac{1}{n}\widehat{V}_j^n - \mu^2\right| + \left|\frac{1}{n}\widehat{V}_j^n(i) - \mu^2\right| \geq \eta_n\right) \leq K\frac{n^{-1/2}}{\eta_n}, \tag{10.11}$$

$$\mathbb{P}\left((\epsilon_{\frac{i}{n}} - \epsilon_{\frac{i-1}{n}})1_{\left\{|\Delta_i^n X| > \alpha\sqrt{\widehat{V}_j^n}n^{-\varpi}\right\}} \geq \eta_n\right) \leq K\frac{n^{\varpi-1/2}}{\eta_n^{\iota}}. \tag{10.12}$$

We denote

$$\mathcal{B}_{i,j}^n = \left\{\omega : \left|\Delta_i^n X^* 1_{\left\{|\Delta_i^n X| \leq \alpha\sqrt{\widehat{V}_j^n}n^{-\varpi}\right\}} - (\epsilon_{\frac{i}{n}} - \epsilon_{\frac{i-1}{n}})\right| > \eta_n \ \cup \ \left|\frac{\sqrt{\widehat{V}_j^n(i)}}{\sqrt{n}\mu} - 1\right| > \eta_n\right\},$$

for $i = (j-1)k_n+1, ..., (j-1)k_n+m_n$ and $j = 1, ..., \lfloor n/k_n \rfloor$. We set $\eta_n = n^{-x}$ for $0 < x < \frac{1}{\iota}(1/2 - \varpi) \bigwedge 1/2$. With this choice

$$\frac{1}{\lfloor n/k_n \rfloor m_n}\sum_{j=1}^{\lfloor n/k_n \rfloor}\sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n}\mathbb{P}\left(\mathcal{B}_{i,j}^n\right) = o(1).$$

Therefore, for any compact subset $\mathcal{A}$ of $(-\infty, 0)$, we have

$$\sup_{\tau \in \mathcal{A}}|\widetilde{F}_n(\tau) - \widehat{G}_n(\tau)| = o_p(1),$$

where we denote

$$\widehat{G}_n(\tau) = \frac{1}{\lfloor n/k_n \rfloor m_n}\sum_{j=1}^{\lfloor n/k_n \rfloor}\sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1\left\{\frac{\sqrt{n}\Delta_i^n X}{\sqrt{\widehat{V}_j^n(i)}}1_{\left\{|\Delta_i^n X| \leq \alpha\sqrt{\widehat{V}_j^n}n^{-\varpi}\right\}} \leq \tau\right\} 1_{\{(\mathcal{B}_{i,j}^n)^c\}}.$$

Taking into account the definition of the set $\mathcal{B}_{i,j}^n$, we get

$$\left\{\begin{array}{l} \widehat{G}_n(\tau) \geq \frac{1}{\lfloor n/k_n \rfloor m_n}\sum_{j=1}^{\lfloor n/k_n \rfloor}\sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1\left\{\frac{1}{\mu}\left(\epsilon_{\frac{i}{n}} - \epsilon_{\frac{i-1}{n}}\right) \leq \tau(1-\eta_n) - \frac{\eta_n}{\mu}\right\}, \\ \widehat{G}_n(\tau) \leq \frac{1}{\lfloor n/k_n \rfloor m_n}\sum_{j=1}^{\lfloor n/k_n \rfloor}\sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1\left\{\frac{1}{\mu}\left(\epsilon_{\frac{i}{n}} - \epsilon_{\frac{i-1}{n}}\right) \leq \tau(1+\eta_n) + \frac{\eta_n}{\mu}\right\}. \end{array}\right.$$

From here we can proceed exactly in the same way as in the proof of Theorem 1 to show that $\widehat{G}_n(\tau) \xrightarrow{\mathbb{P}} F_{\epsilon}(\tau)$ locally uniformly in τ. Hence we need only show $\frac{N^n(\alpha,\varpi)}{\lfloor n/k_n \rfloor m_n} \xrightarrow{\mathbb{P}} 1$ as $n \to \infty$. This follows from

$$\mathbb{P}\left(|\Delta_i^n X^*| > \alpha\sqrt{\widehat{V}_j^n} n^{-\varpi}\right) \leq \mathbb{P}\left(\left|\frac{\sqrt{\widehat{V}_j^n}}{\sqrt{n}\mu}\right| > 0.5\right) + \mathbb{P}\left(|\Delta_i^n X^*| > 0.5\alpha\mu n^{1/2-\varpi}\right)$$

$$\leq \frac{K}{n^\iota}, \qquad \text{for some sufficiently small } \iota > 0,$$

which can be shown using (10.11), the fact that the noise term has a finite first moment and the Burkholder-Davis-Gundy inequality. □

10.4. *Proof of Theorem 3.* As in the proof of Theorem 1, without loss of generality we will assume $\tau < 0$. First, given the fact that $m_n/k_n \to 0$, it is no limitation to assume $k_n - m_n > 2$ and we will do so henceforth. Here we need to make some additional decomposition of the difference $\widetilde{V}_j^n - \overline{V}_j^n$. It is given by the following

$$\widetilde{V}_j^n - \overline{V}_j^n = R_j^{(1)} + R_j^{(2)} + R_j^{(3)} + R_j^{(4)}, \qquad j = 1, ..., \lfloor n/k_n \rfloor, \tag{10.13}$$

$$R_j^{(1)} = \frac{n}{k_n - 1}\frac{\pi}{2}\sum_{i=(j-1)k_n+2}^{jk_n}\Bigg[\left(|\Delta_{i-1}^n B||\Delta_i^n B| - \sigma_{(i-2)\Delta_n}^2|\Delta_{i-1}^n W||\Delta_i^n W|\right)$$
$$+ (\sigma_{(i-2)\Delta_n} - \sigma_{\frac{(j-1)k_n}{n}})^2|\Delta_{i-1}^n W||\Delta_i^n W|\Bigg],$$

$$R_j^{(2)} = 2\frac{n}{k_n - 1}\frac{\pi}{2}\sigma_{\frac{(j-1)k_n}{n}}\sum_{i=(j-1)k_n+2}^{jk_n}\Bigg[\sigma_{(i-2)\Delta_n} - \sigma_{\frac{(j-1)k_n}{n}} - \int_{\frac{(j-1)k_n}{n}}^{\frac{i-2}{n}}\widetilde{\sigma}_{\frac{(j-1)k_n}{n}}dW_u$$
$$- \int_{\frac{(j-1)k_n}{n}}^{\frac{i-2}{n}}\widetilde{\sigma}'_{\frac{(j-1)k_n}{n}}dW'_u\Bigg]|\Delta_{i-1}^n W||\Delta_i^n W|,$$

$$R_j^{(3)} = \frac{2}{k_n - 1}\frac{\pi}{2}\sigma_{\frac{(j-1)k_n}{n}}\sum_{i=(j-1)k_n+2}^{jk_n}\left[\int_{\frac{(j-1)k_n}{n}}^{\frac{i-2}{n}}\widetilde{\sigma}_{\frac{(j-1)k_n}{n}}dW_u + \int_{\frac{(j-1)k_n}{n}}^{\frac{i-2}{n}}\widetilde{\sigma}'_{\frac{(j-1)k_n}{n}}dW'_u\right]$$
$$\times\left(n|\Delta_{i-1}^n W||\Delta_i^n W| - \frac{2}{\pi}\right),$$

$$R_j^{(4)} = \frac{2}{k_n - 1}\sigma_{\frac{(j-1)k_n}{n}}\sum_{i=(j-1)k_n+2}^{jk_n}\left[\int_{\frac{(j-1)k_n}{n}}^{\frac{i-2}{n}}\widetilde{\sigma}_{\frac{(j-1)k_n}{n}}dW_u + \int_{\frac{(j-1)k_n}{n}}^{\frac{i-2}{n}}\widetilde{\sigma}'_{\frac{(j-1)k_n}{n}}dW'_u\right].$$

For $i = (j-1)k_n + 1, ..., jk_n - 2$ we denote the component of $R_j^{(4)}$ that does not contain the increments $\Delta_i^n W$ and $\Delta_i^n W'$ with

$$\widetilde{R}_{i,j}^{(4)} = R_j^{(4)} - \frac{2}{k_n - 1}\sigma_{\frac{(j-1)k_n}{n}}(jk_n - i - 1)\left[\int_{\frac{i-1}{n}}^{\frac{i}{n}}\widetilde{\sigma}_{\frac{(j-1)k_n}{n}}dW_u + \int_{\frac{i-1}{n}}^{\frac{i}{n}}\widetilde{\sigma}'_{\frac{(j-1)k_n}{n}}dW'_u\right].$$

We decompose analogously the difference $\widetilde{V}_j^n(i) - \overline{V}_j^n(i)$ into $R_j^{(k)}(i)$ for $k = 1, .., 4$ and $\widetilde{R}_{i,j}^{(4)}(i)$ is the component of $R_j^{(4)}(i)$ that does not contain the increments $\Delta_i^n W$ and $\Delta_i^n W'$. We further denote for $i = (j-1)k_n+1, ..., (j-1)k_n + m_n$ and $j = 1, ..., \lfloor n/k_n \rfloor$,

$$\xi_j^n(1) = \frac{\widehat{V}_j^n(i) - \sigma^2_{\frac{(j-1)k_n}{n}}}{2\sigma^2_{\frac{(j-1)k_n}{n}}}, \quad \xi_j^n(2) = \frac{\left(\widehat{V}_j^n(i) - \sigma^2_{\frac{(j-1)k_n}{n}}\right)^2}{8\sigma^4_{\frac{(j-1)k_n}{n}}},$$

$$\widetilde{\xi}_{i,j}^n(1) = \frac{\overline{V}_j^n(i) + \widetilde{R}_{i,j}^{(4)}(i) - \sigma^2_{\frac{(j-1)k_n}{n}}}{2\sigma^2_{\frac{(j-1)k_n}{n}}}, \quad \widetilde{\xi}_{i,j}^n(2) = \frac{\left(\overline{V}_j^n(i) + \widetilde{R}_{i,j}^{(4)}(i) - \sigma^2_{\frac{(j-1)k_n}{n}}\right)^2}{8\sigma^4_{\frac{(j-1)k_n}{n}}},$$

$$\overline{\xi}_j^n(1) = \frac{\overline{V}_j^n + R_j^{(4)} - \sigma^2_{\frac{(j-1)k_n}{n}}}{2\sigma^2_{\frac{(j-1)k_n}{n}}}, \quad \overline{\xi}_j^n(2) = \frac{\left(\overline{V}_j^n + R_j^{(4)} - \sigma^2_{\frac{(j-1)k_n}{n}}\right)^2}{8\sigma^4_{\frac{(j-1)k_n}{n}}},$$

$$\widehat{\xi}_j^n(1) = \frac{\overline{V}_j^n - \sigma^2_{\frac{(j-1)k_n}{n}}}{2\sigma^2_{\frac{(j-1)k_n}{n}}}, \quad \widehat{\xi}_j^n(2) = \frac{\left(\overline{V}_j^n - \sigma^2_{\frac{(j-1)k_n}{n}}\right)^2}{8\sigma^4_{\frac{(j-1)k_n}{n}}},$$

$$\xi_{i,j}^n(3) = \frac{\sqrt{n}\Delta_i^n W}{\sigma_{\frac{(j-1)k_n}{n}}}\left[\widetilde{\sigma}_{\frac{(j-1)k_n}{n}}(W_{\frac{i-1}{n}} - W_{\frac{(j-1)k_n}{n}}) + \widetilde{\sigma}'_{\frac{(j-1)k_n}{n}}(W'_{\frac{i-1}{n}} - W'_{\frac{(j-1)k_n}{n}})\right],$$

$$\xi_{i,j}^n(4) = 1 + \frac{1}{\sigma_{\frac{(j-1)k_n}{n}}}\left[\widetilde{\sigma}_{\frac{(j-1)k_n}{n}}(W_{\frac{i-1}{n}} - W_{\frac{(j-1)k_n}{n}}) + \widetilde{\sigma}'_{\frac{(j-1)k_n}{n}}(W'_{\frac{i-1}{n}} - W'_{\frac{(j-1)k_n}{n}})\right].$$

With this notation we set for $i = (j-1)k_n + 1, ..., (j-1)k_n + m_n$ and $j = 1, ..., \lfloor n/k_n \rfloor$

$$\begin{aligned}
\chi_{i,j}^n(1) = & -\sqrt{n}\frac{1}{\sigma_{\frac{(j-1)k_n}{n}}}\left(\Delta_i^n A + \Delta_i^n Y + \int_{\frac{i-1}{n}}^{\frac{i}{n}}\left(\sigma_u - \sigma_{\frac{i-1}{n}}\right)dW_u\right)1_{\left\{|\Delta_i^n X| \le \alpha\sqrt{\widehat{V}_j^n}n^{-\varpi}\right\}} \\
& + (\sqrt{n}\Delta_i^n W + \xi_{i,j}^n(3))1_{\left\{|\Delta_i^n X| > \alpha\sqrt{\widehat{V}_j^n}n^{-\varpi}\right\}} \\
& - \left(\frac{\sqrt{n}\Delta_i^n W}{\sigma_{\frac{(j-1)k_n}{n}}}(\sigma_{\frac{i-1}{n}} - \sigma_{\frac{(j-1)k_n}{n}}) - \xi_{i,j}^n(3)\right)1_{\left\{|\Delta_i^n X| \le \alpha\sqrt{\widehat{V}_j^n}n^{-\varpi}\right\}},
\end{aligned}$$

$$\chi_{i,j}^n(2) = \left(\frac{\sqrt{\widehat{V}_j^n(i)}}{\sigma_{\frac{(j-1)k_n}{n}}} - 1 - \xi_j^n(1) + \xi_j^n(2)\right) + (\xi_j^n(1) - \xi_j^n(2) - \widetilde{\xi}_{i,j}^n(1) + \widetilde{\xi}_{i,j}^n(2)).$$

Finally, we denote

$$\widehat{G}_n(\tau) = \frac{1}{\lfloor n/k_n \rfloor m_n} \sum_{j=1}^{\lfloor n/k_n \rfloor} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1\Bigg(\sqrt{n}\frac{\Delta_i^n X}{\sigma_{\frac{(j-1)k_n}{n}}} 1_{\left\{|\Delta_i^n X| \leq \alpha \sqrt{\widehat{V}_j^n} n^{-\varpi}\right\}} \leq \tau \frac{\sqrt{\widehat{V}_j^n(i)}}{\sigma_{\frac{(j-1)k_n}{n}}} - \chi_{i,j}^n(1) - \tau \chi_{i,j}^n(2)\Bigg)$$
$$= \frac{1}{\lfloor n/k_n \rfloor m_n} \sum_{j=1}^{\lfloor n/k_n \rfloor} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1\left(\sqrt{n}\Delta_i^n W \leq \tau + \tau \widetilde{\xi}_{i,j}^n(1) - \tau \widetilde{\xi}_{i,j}^n(2) - \xi_{i,j}^n(3)\right).$$

The proof consists of three parts: the first is showing the negligibility of $k_n(\widetilde{F}_n(\tau) - \widehat{G}_n(\tau))$, the second is deriving the limiting behavior of $\widehat{G}_n(\tau) - \Phi(\tau)$ and third part is showing negligibility of $k_n(\widehat{F}_n(\tau) - \widetilde{F}_n(\tau))$.

10.4.1. *The difference* $\widetilde{F}_n(\tau) - \widehat{G}_n(\tau)$. We first collect some preliminary results that we then make use of in analyzing $\widetilde{F}_n(\tau) - \widehat{G}_n(\tau)$. We start with $\max_{i=1,\dots,n} |\Delta_i^n B|$. Using maximal inequality we have

$$\mathbb{E}(\max_{i=1,\dots,n} |\Delta_i^n B|^p) \leq K n^{1-p/2}, \quad \forall p > 0. \tag{10.14}$$

Next, using assumption SB (in particular that jumps are of finite activity), we have

(10.15)
$$\mathbb{P}\left(\int_{\frac{(j-1)k_n}{n}}^{\frac{jk_n}{n}} \int_E 1\left(\delta^{\phi}(z,x) \neq 0\right) \mu(dz,dx) \geq 1\right) \leq K\frac{k_n}{n}, \quad \phi = Y,\ \sigma,\ \widetilde{\sigma} \text{ and } \widetilde{\sigma}'.$$

We now provide bounds for the elements of $\chi_{i,j}^n(1)$ and $\chi_{i,j}^n(2)$. In what follows we denote with η_n some deterministic sequence of positive numbers that depends only on n. We first have (recall the definition of $\overline{\sigma}_t$)

$$\mathbb{P}\left(\sqrt{n}\left|\int_{\frac{i-1}{n}}^{\frac{i}{n}} (\sigma_u - \sigma_{\frac{i-1}{n}}) dW_u\right| \geq \eta_n\right) \leq \mathbb{P}\left(\int_{\frac{(j-1)k_n}{n}}^{\frac{jk_n}{n}} \int_E 1\left(\delta^Y(s,x) \neq 0\right) \mu(ds,dx) \geq 1\right) + \mathbb{P}\left(\sqrt{n}\left|\int_{\frac{i-1}{n}}^{\frac{i}{n}} (\overline{\sigma}_u - \overline{\sigma}_{\frac{i-1}{n}}) dW_u\right| \geq \eta_n\right).$$

For the second term on the right hand side of the above inequality, we can use Chebyshev inequality as well as Burkholder-Davis-Gundy inequality, to get for $\forall p \geq 2$:

$$\mathbb{P}\left(\sqrt{n}\left|\int_{\frac{i-1}{n}}^{\frac{i}{n}} (\overline{\sigma}_u - \overline{\sigma}_{\frac{i-1}{n}}) dW_u\right| \geq \eta_n\right) \leq \frac{n^{p/2}\mathbb{E}\left|\int_{(i-1)\Delta_n}^{i\Delta_n} (\overline{\sigma}_u - \overline{\sigma}_{\frac{i-1}{n}})^2 du\right|^{p/2}}{\eta_n^p}.$$

Therefore, applying again Burkholder-Davis-Gundy inequality, we have altogether

(10.16)
$$\mathbb{P}\left(\sqrt{n}\left|\int_{\frac{i-1}{n}}^{\frac{i}{n}}(\sigma_u-\sigma_{\frac{i-1}{n}})dW_u\right|\geq\eta_n\right)\leq K\left[\left(\frac{k_n}{n}\right)\bigvee\left(\frac{1}{n^{p/2}\eta_n^p}\right)\right],\quad\forall p>0.$$

Similar calculations (using the fact that $\widetilde{\sigma}_t$ and $\widetilde{\sigma}'_t$ are Itô semimartingales), yields for $\forall p>0$

(10.17)
$$\mathbb{P}\left(\left|\frac{\sqrt{n}\Delta_i^nW}{\sigma_{\frac{(j-1)k_n}{n}}}\left(\sigma_{\frac{i-1}{n}}-\sigma_{\frac{(j-1)k_n}{n}}\right)-\xi_{i,j}^n(3)\right|\geq\eta_n\right)\leq K\left[\left(\frac{k_n}{n}\right)\bigvee\left(\frac{k_n}{n\eta_n}\right)^p\right].$$

Next, applying Chebyshev inequality and the elementary $|\sum_i|a_i||^p\leq\sum_i|a_i|^p$ for $p\in(0,1]$, we get

$$\begin{aligned}\mathbb{P}\left(\sqrt{n}|\Delta_i^nY|\geq\eta_n\right)&\leq\frac{n^{\iota/2}\mathbb{E}\left(\int_{(i-1)\Delta_n}^{i\Delta_n}\int_E|\delta^Y(s,x)|\mu(ds,dx)\right)^{\iota}}{\eta_n^{\iota}}\\&\leq\frac{n^{\iota/2}\mathbb{E}\left(\int_{(i-1)\Delta_n}^{i\Delta_n}\int_E|\delta^Y(s,x)|^{\iota}\mu(ds,dx)\right)}{\eta_n^{\iota}}\\&\leq Kn^{-1+\iota/2}\eta_n^{-\iota},\quad\forall\iota\in(0,1).\end{aligned}\tag{10.18}$$

Further, Chebyshev inequality and the boundedness of a_t easily implies

$$\mathbb{P}\left(\sqrt{n}|\Delta_i^nA|\geq\eta_n\right)\leq\frac{n^{p/2}\mathbb{E}(|\Delta_i^nA|^p)}{\eta_n^p}\leq K\frac{1}{n^{p/2}\eta_n^p}.\tag{10.19}$$

We turn next to the difference $\widehat{V}_j^n-\dot{V}_j^n$. Using triangular inequality and successive conditioning, we have

$$\begin{aligned}&\mathbb{P}\left(|\widehat{V}_j^n-\dot{V}_j^n|\geq\eta_n\right)\\&\quad\leq\mathbb{P}\left(2\frac{n}{k_n}\frac{\pi}{2}\max_{i=1,\dots,n}|\Delta_i^nA+\Delta_i^nB|\int_{\frac{(j-1)k_n}{n}}^{\frac{jk_n}{n}}\int_E(|\delta^Y(s,x)|\vee1)\mu(ds,dx)\geq\frac{\eta_n}{2}\right)+\frac{K}{n\eta_n}.\end{aligned}$$

From here we have

$$\begin{aligned}&\mathbb{P}\left(2\frac{n}{k_n}\frac{\pi}{2}\max_{i=1,\dots,n}|\Delta_i^nA+\Delta_i^nB|\int_{\frac{(j-1)k_n}{n}}^{\frac{jk_n}{n}}\int_E(|\delta^Y(s,x)|\vee1)\mu(ds,dx)\geq\frac{\eta_n}{2}\right)\\&\qquad\leq\mathbb{P}\left(\int_{\frac{(j-1)k_n}{n}}^{\frac{jk_n}{n}}\int_E\mu(ds,dx)\geq1\right)\leq K\frac{k_n}{n}.\end{aligned}$$

Thus altogether we get

$$\mathbb{P}\left(|\widehat{V}_j^n - \dot{V}_j^n| \geq \eta_n\right) \leq K\left(\frac{1}{n\eta_n}\bigvee\frac{k_n}{n}\right). \tag{10.20}$$

We continue next with the difference $\dot{V}_j^n - \widetilde{V}_j^n$. Application of triangular inequality gives

$$\begin{aligned}|\Delta_{i-1}^n A + \Delta_{i-1}^n B||\Delta_i^n A + \Delta_i^n B| - |\Delta_{i-1}^n B||\Delta_i^n B| \\ \leq |\Delta_{i-1}^n A + \Delta_{i-1}^n B||\Delta_i^n A| + |\Delta_{i-1}^n A||\Delta_i^n B|.\end{aligned}$$

Using this inequality and applying Chebyshev inequality, we get

$$\mathbb{P}\left(|\dot{V}_j^n - \widetilde{V}_j^n| \geq \eta_n\right) \leq K\left(\frac{1}{\sqrt{n}\eta_n}\right)^p, \qquad \forall p \geq 1, \tag{10.21}$$

and this inequality can be further strengthened but suffices for our analysis.

Turning next to $R_j^{(1)}$, using triangular inequality, Burkholder-Davis-Gundy inequality as well as (10.15), we can easily get

$$\begin{aligned}\mathbb{P}\left(|R_j^{(1)}| \geq \eta_n\right) &\leq \mathbb{P}\left(|R_j^{(1)}| \geq \eta_n, \quad \int_{\frac{(j-1)k_n}{n}}^{\frac{jk_n}{n}}\int_E 1(\delta^\sigma(s,x) \neq 0)\mu(ds,dx) \geq 1\right) \\ &\quad + \mathbb{P}\left(|R_j^{(1)}| \geq \eta_n, \quad \int_{\frac{(j-1)k_n}{n}}^{\frac{jk_n}{n}}\int_E 1(\delta^\sigma(s,x) \neq 0)\mu(ds,dx) = 0\right) \\ &\leq K\frac{k_n}{n} + K\left(\frac{1}{\sqrt{n}\eta_n}\right)^p + K\left(\frac{k_n}{n\eta_n}\right)^p, \quad \forall p \geq 1.\end{aligned}$$

Similar calculations, and utilizing the fact that $\widetilde{\sigma}_t$ $\widetilde{\sigma}'_t$ are themselves Itô semimartingales, yield

$$\mathbb{P}\left(|R_j^{(2)}| \geq \eta_n\right) \leq K\frac{k_n}{n} + K\left(\frac{k_n}{n\eta_n}\right)^p, \quad \forall p \geq 1. \tag{10.22}$$

Next, by splitting

$$n|\Delta_{i-1}^n W||\Delta_i^n W| - \frac{2}{\pi} = |\sqrt{n}\Delta_{i-1}^n W|\left(|\sqrt{n}\Delta_i^n W| - \sqrt{\frac{2}{\pi}}\right) + \sqrt{\frac{2}{\pi}}\left(|\sqrt{n}\Delta_{i-1}^n W| - \sqrt{\frac{2}{\pi}}\right),$$

we can decompose $R_j^{(3)}$ into two discrete martingales. Then applying Burkholder-Davis-Gundy inequality, we get

$$\mathbb{P}\left(|R_j^{(3)}| \geq \eta_n\right) \leq K\left(\frac{1}{\sqrt{n}\eta_n}\right)^p, \qquad \forall p \geq 2. \tag{10.23}$$

Next, we trivially have

$$\left\{\begin{array}{l}\mathbb{P}\left(|\overline{V}_j^n - \sigma^2_{\frac{(j-1)k_n}{n}}| \geq \eta_n\right) \leq K\left(\frac{1}{k_n\eta_n^2}\right)^p, \\ \mathbb{P}\left(|\widetilde{V}_j^n - \overline{V}_j^n| \geq 0.5\sigma^2_{\frac{(j-1)k_n}{n}}\right) \leq K\frac{k_n}{n}, \quad \forall p \geq 2.\end{array}\right. \tag{10.24}$$

Further, application of Burkholder-Davis-Gundy inequality gives

$$\left\{\begin{array}{l}\mathbb{E}|\overline{V}_j^n - \sigma^2_{\frac{(j-1)k_n}{n}}|^p \leq \frac{K}{k_n^{p/2}}, \\ \mathbb{E}^n_{(j-1)k_n}(R_j^{(4)}) = 0, \quad \mathbb{E}|R_j^{(4)}|^p \leq K\left(\frac{k_n}{n}\right)^{p/2}, \quad \forall p \geq 2.\end{array}\right. \tag{10.25}$$

The results in (10.20)-(10.25) continue to hold when $\widehat{V}_j^n$, $\dot{V}_j^n$, $\widetilde{V}_j^n$, $\overline{V}_j^n$, $R_j^{(1)}$, $R_j^{(2)}$ and $R_j^{(3)}$ are replaced with $\widehat{V}_j^n(i)$, $\dot{V}_j^n(i)$, $\widetilde{V}_j^n(i)$, $\overline{V}_j^n(i)$, $R_j^{(1)}(i)$, $R_j^{(2)}(i)$ and $R_j^{(3)}(i)$ respectively.

Further, using Burkholder-Davis-Gundy inequality for discrete martingales (note that $\overline{V}_j^n - \overline{V}_j^n(i)$ can be decomposed into discrete martingales and terms whose p-th moment is bounded by K/k_n^p), we have

$$\mathbb{E}\left(|R_j^{(4)} - \widetilde{R}_{i,j}^{(4)}|^p + |\widetilde{R}_{i,j}^{(4)} - \widetilde{R}_{i,j}^{(4)}(i)|^p\right) \leq K\left(\frac{1}{\sqrt{n}}\right)^p, \quad \forall p > 0, \tag{10.26}$$

$$\left|\mathbb{E}^n_{(j-1)k_n}\left(\overline{V}_j^n - \overline{V}_j^n(i)\right)\right| \leq \frac{K}{k_n^2}, \quad \mathbb{E}|\overline{V}_j^n - \overline{V}_j^n(i)|^p \leq K\left(\frac{1}{k_n}\right)^p, \quad \forall p \geq 1. \tag{10.27}$$

Now we can use the above results for the components of $\widehat{V}_j^n(i) - \sigma^2_{\frac{(j-1)k_n}{n}}$, to analyze the first term in $\chi_{i,j}^n(2)$ involving $\sqrt{\widehat{V}_j^n(i)} - \sigma_{\frac{(j-1)k_n}{n}}$. We make use of the following algebraic inequality

$$\left|\sqrt{x} - \sqrt{y} - \frac{x-y}{2\sqrt{y}} + \frac{(x-y)^2}{8y\sqrt{y}}\right| \leq \frac{(x-y)^4}{8y^{7/2}} + \frac{|x-y|^3}{2y^{5/2}},$$

for every $x \geq 0$ and $y > 0$. Using this inequality with x and y replaced with $\widehat{V}_j^n(i)$ and $\sigma^2_{\frac{(j-1)k_n}{n}}$ respectively, as well the bounds in (10.20)-(10.27), we get

$$\begin{aligned}&\mathbb{P}\left(\left|\frac{\sqrt{\widehat{V}_j^n(i)}}{\sigma_{\frac{(j-1)k_n}{n}}} - 1 - \xi_j^n(1) + \xi_j^n(2)\right| \geq \eta_n\right) \\ &\qquad \leq K\left[\frac{1}{n\eta_n^{1/3}} \bigvee \frac{k_n}{n} \bigvee \frac{1}{\eta_n^{p/3}[n^{p/2} \wedge (n/k_n)^{p/2}]} \bigvee \frac{1}{\eta_n^{2p/3}k_n^p}\right],\end{aligned} \tag{10.28}$$

for $\forall p \geq 1$ and $\forall \iota > 0$. Similarly, using the following inequality

$$\mathbb{P}\left(|x^2 - y^2| \geq \epsilon\right) \leq \mathbb{P}\left(|x-y|^2 \geq 0.5\epsilon\right) + \mathbb{P}\left(2|y| \geq K\right) + \mathbb{P}\left(|x-y| \geq 0.5\epsilon/K\right),$$

for any random variables x and y and constants $\epsilon > 0$ and $K > 0$, together with the bounds in (10.20)-(10.27), we have

$$\begin{aligned} &\mathbb{P}\left(|\xi_j^n(1) - \xi_j^n(2) - \widetilde{\xi}_{i,j}^n(1) + \widetilde{\xi}_{i,j}^n(2)| \geq \eta_n\right) \\ &\qquad \leq K\left[\frac{1}{n\eta_n} \bigvee \frac{k_n}{n} \bigvee \frac{1}{\eta_n^p[n^{p/2} \wedge (n/k_n)^p]}\right], \end{aligned} \tag{10.29}$$

for every $p \geq 1$ and arbitrary small $\iota > 0$.

We finally provide a bound for the second term in $\chi_{i,j}^n(1)$. We can use Chebyshev inequality as well as Hölder inequality to get

$$\begin{aligned} &\mathbb{P}\left((\sqrt{n}\sigma_{\frac{(j-1)k_n}{n}}|\Delta_i^n W| + |\xi_{i,j}^n(3)|)1\left(|\Delta_i^n X| > \alpha\sqrt{\widehat{V}_j^n}n^{-\varpi}\right) \geq \eta_n\right) \\ &\qquad \leq K\frac{\left[\mathbb{P}\left(|\Delta_i^n X| > \alpha\sqrt{\widehat{V}_j^n}n^{-\varpi}\right)\right]^{1/(1+\iota)}}{\eta_n^\iota}. \end{aligned} \tag{10.30}$$

We can further write

$$\begin{aligned} \mathbb{P}\left(|\Delta_i^n X| > \alpha\sqrt{\widehat{V}_j^n}n^{-\varpi}\right) &\leq \mathbb{P}\left(|\sqrt{\widehat{V}_j^n} - \sigma_{\frac{(j-1)k_n}{n}}| \geq 0.5\sigma_{\frac{(j-1)k_n}{n}}\right) \\ &\quad + \mathbb{P}\left(|\Delta_i^n X| > 0.5\alpha\sigma_{\frac{(j-1)k_n}{n}}n^{-\varpi}\right). \end{aligned}$$

From here we can use the bounds in (10.20)-(10.27) as well as (10.18) and conclude

$$\begin{aligned} &\mathbb{P}\left(\sqrt{n}\sigma_{\frac{(j-1)k_n}{n}}|\Delta_i^n W| 1\left(|\Delta_i^n X| > \alpha\sqrt{\widehat{V}_j^n}n^{-\varpi}\right) \geq \eta_n\right) \\ &\qquad \leq K\left(\frac{k_n}{n}\right)^{\frac{1}{1+\iota}}\frac{1}{\eta_n^\iota}, \quad \forall \iota > 0. \end{aligned} \tag{10.31}$$

Combining the results in (10.14), (10.16), (10.17), (10.18), (10.19), (10.28), (10.29) and (10.31), we get

$$\begin{aligned} &\mathbb{P}\left((|\chi_{i,j}^n(1)| + |\chi_{i,j}^n(2)|) > \eta_n\right) \\ &\qquad \leq K\left[\frac{1}{n\eta_n} \bigvee \frac{1}{\eta_n^p[n^{p/2} \wedge (n/k_n)^p \wedge k_n^{3p/2}]} \bigvee \left(\frac{k_n}{n}\right)^{\frac{1}{1+\iota}}\frac{1}{\eta_n^\iota}\right]. \end{aligned}$$

From here, using the fact that the probability density of a standard normal variable is uniformly bounded, we get

$$
\begin{aligned}
&\mathbb{E}\Bigg|1\left(\sqrt{n}\Delta_i^n X 1_{\left\{|\Delta_i^n X|\le\alpha\sqrt{\widehat{V}_j^n}n^{-\varpi}\right\}} \le \tau\sqrt{\widehat{V}_j^n(i)}\right) \\
&\qquad - 1\left(\sqrt{n}\frac{\Delta_i^n X}{\sigma_{\frac{(j-1)k_n}{n}}}1_{\left\{|\Delta_i^n X|\le\alpha\sqrt{\widehat{V}_j^n}n^{-\varpi}\right\}} \le \tau\frac{\sqrt{\widehat{V}_j^n(i)}}{\sigma_{\frac{(j-1)k_n}{n}}} - \chi_{i,j}^n(1) - \tau\chi_{i,j}^n(2)\right)\Bigg| \\
&\le \mathbb{P}\left((|\chi_{i,j}^n(1)| + |\chi_{i,j}^n(2)|) > \eta_n\right) \\
&\quad + \mathbb{E}\left|\Phi\left(\frac{\tau + \tau\widetilde{\xi}_{i,j}^n(1) - \tau\widetilde{\xi}_{i,j}^n(2) + \eta_n(1+|\tau|)}{\xi_{i,j}^n(4)}\right) - \Phi\left(\frac{\tau + \tau\widetilde{\xi}_{i,j}^n(1) - \tau\widetilde{\xi}_{i,j}^n(2) - \eta_n(1+|\tau|)}{\xi_{i,j}^n(4)}\right)\right| \\
&\le K\mathbb{P}\left((|\chi_{i,j}^n(1)| + |\chi_{i,j}^n(2)|) > \eta_n\right) + K\eta_n|\tau|.
\end{aligned}
$$

Therefore, upon picking $\eta_n \propto n^{-q-\iota}$ for $\iota \in (0, 1/2 - q)$ sufficiently small, we get finally for any compact subset $\mathcal{A}$ of $(-\infty, 0)$

$$
\sup_{\tau\in\mathcal{A}} |\widetilde{F}_n(\tau) - \widehat{G}_n(\tau)| = o_p\left(\frac{1}{k_n}\right). \tag{10.32}
$$

10.4.2. *The asymptotic behavior of $\widehat{G}_n(\tau) - \Phi(\tau)$.* We have (10.33)

$$
\widehat{G}_n(\tau) - \Phi(\tau) = \sum_{i=1}^5 A_i^n, \quad A_1^n = \frac{1}{\lfloor n/k_n\rfloor m_n}\sum_{j=1}^{\lfloor n/k_n\rfloor}\sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n}\left[1\left(\sqrt{n}\Delta_i^n W \le \tau\right) - \Phi(\tau)\right],
$$

$$
A_2^n = \frac{1}{\lfloor n/k_n\rfloor}\sum_{j=1}^{\lfloor n/k_n\rfloor}\left(\Phi\left(\tau + \tau\overline{\xi}_j^n(1) - \tau\overline{\xi}_j^n(2)\right) - \Phi(\tau)\right), \quad A_3^n = \frac{1}{\lfloor n/k_n\rfloor m_n}\sum_{j=1}^{\lfloor n/k_n\rfloor}\sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} a_i^n,
$$

$$
a_i^n = 1\left(\sqrt{n}\Delta_i^n W \le \frac{\tau + \tau\widetilde{\xi}_{i,j}^n(1) - \tau\widetilde{\xi}_{i,j}^n(2)}{\xi_{i,j}^n(4)}\right) - 1\left(\sqrt{n}\Delta_i^n W \le \tau\right) + \Phi(\tau) - \Phi\left(\frac{\tau + \tau\widetilde{\xi}_{i,j}^n(1) - \tau\widetilde{\xi}_{i,j}^n(2)}{\xi_{i,j}^n(4)}\right),
$$

$$
A_4^n = \frac{1}{\lfloor n/k_n\rfloor m_n}\sum_{j=1}^{\lfloor n/k_n\rfloor}\sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n}\left[\Phi\left(\tau + \tau\widetilde{\xi}_{i,j}^n(1) - \tau\widetilde{\xi}_{i,j}^n(2)\right) - \Phi\left(\tau + \tau\overline{\xi}_j^n(1) - \tau\overline{\xi}_j^n(2)\right)\right],
$$

$$
A_5^n = \frac{1}{\lfloor n/k_n\rfloor m_n}\sum_{j=1}^{\lfloor n/k_n\rfloor}\sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n}\left[\Phi\left(\frac{\tau + \tau\widetilde{\xi}_{i,j}^n(1) - \tau\widetilde{\xi}_{i,j}^n(2)}{\xi_{i,j}^n(4)}\right) - \Phi\left(\tau + \tau\widetilde{\xi}_{i,j}^n(1) - \tau\widetilde{\xi}_{i,j}^n(2)\right)\right].
$$

We first derive a bound for the order of magnitude of A_3^n, A_4^n and A_5^n and then analyze the limiting behavior of A_1^n and A_2^n. Using the independence of $\Delta_i^n W$, $\Delta_h^n W$, $\Delta_i^n W'$, $\Delta_h^n W'$ from each other (for $i \neq h$) and $\mathcal{F}_{\frac{(j-1)k_n}{n}}$, the fact that $\xi_{i,j}^n(4)$ is adapted to $\mathcal{F}_{i-1}^n$ as well as successive conditioning, we have $\mathbb{E}\left(a_i^n a_h^n\right)=0$ for $|i-h|>k_n$. For $0<i-h\le k_n$, we can first split a_h^n into a component in which the summand including the i-th increment $\Delta_i^n W$ is removed from $\widetilde{\xi}_{h,j}^n(1)$ and $\widetilde{\xi}_{h,j}^n(2)$. We denote this part of a_h^n with $\overline{a}_h^n$ and the residual with $\widetilde{a}_h^n = a_h^n - \overline{a}_h^n$. We further denote with $\widetilde{\xi}_{h,j}^{i,n}(1)$ and $\widetilde{\xi}_{h,j}^{i,n}(2)$ the terms $\widetilde{\xi}_{h,j}^n(1)$ and $\widetilde{\xi}_{h,j}^n(2)$ in which the summand corresponding to $\Delta_i^n W$ is removed. Then using successive conditioning, we have for $(j-1)k_n+1\le h<i\le (j-1)k_n+m_n$

$$\mathbb{E}\left(a_i^n \overline{a}_h^n\right)=0,\quad \mathbb{E}\left(a_i^n\right)^2\le K|\tau|\left(\sqrt{\frac{k_n}{n}}\bigvee\frac{1}{\sqrt{k_n}}\right).$$

Further, we can use triangular inequality for a_i^n and $\widetilde{a}_h^n$, the bounds in (10.25)-(10.27), and get for n sufficiently high

$$\begin{aligned}\mathbb{E}|a_i^n\widetilde{a}_h^n| &\le \mathbb{P}\left(|\widetilde{\xi}_{i,j}^n(4)-1|>\left(\frac{k_n}{n}\right)^{1/2-\iota}\ \cup\ |\widetilde{\xi}_{h,j}^n(4)-1|>\left(\frac{k_n}{n}\right)^{1/2-\iota}\right)\\
&+\mathbb{P}\left(|\widetilde{\xi}_{i,j}^n(1)-\widetilde{\xi}_{i,j}^n(2)|>k_n^{-1/2+\iota}\right)+\mathbb{P}\left(|\widetilde{\xi}_{h,j}^n(1)-\widetilde{\xi}_{h,j}^{i,n}(1)-\widetilde{\xi}_{h,j}^n(2)+\widetilde{\xi}_{h,j}^{i,n}(2)|>k_n^{-1+\iota}\right)\\
&+\mathbb{P}\left(\sqrt{n}\Delta_i^n W\in 2\tau(1-k_n^{-1/2+\iota},\ 1+k_n^{-1/2+\iota})\ \cap\ \sqrt{n}\Delta_h^n W\in 2\tau(1-k_n^{-1+\iota},\ 1+k_n^{-1+\iota})\right)\\
&+\mathbb{P}\left(|\widetilde{\xi}_{h,j}^n(1)-\widetilde{\xi}_{h,j}^n(2)|>k_n^{-1/2+\iota}\right)+K\frac{|\tau|\vee\tau^2}{k_n}.\end{aligned}$$

Therefore, using again (10.25)-(10.27), we have

$$A_3^n\le K(\sqrt{|\tau|}\vee\tau^2)\times\left(\frac{1}{\sqrt{\lfloor n/k_n\rfloor m_n}}\left(\frac{1}{k_n}\right)^{1/4}\bigvee\frac{1}{\sqrt{n}}\right). \tag{10.34}$$

For A_4^n, using a second-order Taylor expansion, the bounds in (10.24), (10.25) and (10.27), as well as the uniform boundedness of the probability density of the standard normal distribution and its derivative, we get

$$\mathbb{E}|A_4^n|\le K(|\tau|\vee\tau^2)\left(\frac{1}{k_n^{3/2}}\bigvee\frac{1}{\sqrt{\lfloor n/k_n\rfloor k_n}}\right). \tag{10.35}$$

Next, for A_5^n, we can use the boundedness of the probability density of the standard normal as well as a second-order Taylor expansion, to get for $\forall\iota>0$ and n sufficiently high

$$\Phi\left(\frac{\tau+\tau\widetilde{\xi}_{i,j}^n(1)-\tau\widetilde{\xi}_{i,j}^n(2)}{\xi_{i,j}^n(4)}\right)-\Phi\left(\tau+\tau\widetilde{\xi}_{i,j}^n(1)-\tau\widetilde{\xi}_{i,j}^n(2)\right)=b_i^n(1)+b_i^n(2)+b_i^n(3),$$

$$b_i^n(1)=\left\{\Phi\left(\frac{\tau+\tau\widetilde{\xi}_{i,j}^n(1)-\tau\widetilde{\xi}_{i,j}^n(2)}{\xi_{i,j}^n(4)}\right)-\Phi\left(\tau+\tau\widetilde{\xi}_{i,j}^n(1)-\tau\widetilde{\xi}_{i,j}^n(2)\right)\right\}1_{\left\{|\xi_{i,j}^n(4)-1|\ge\left(\frac{k_n}{n}\right)^{1/2-\iota}\right\}},$$

$$b_i^n(2) = \Phi'\left(\tau + \tau\widetilde{\xi}_{i,j}^n(1) - \tau\widetilde{\xi}_{i,j}^n(2)\right)(\tau+\tau\widetilde{\xi}_{i,j}^n(1)-\tau\widetilde{\xi}_{i,j}^n(2))(\xi_{i,j}^n(4)-1)1_{\left\{|\xi_{i,j}^n(4)-1|<(\frac{k_n}{n})^{1/2-\iota}\right\}},$$

$$|b_i^n(3)| \leq K\frac{|\tau+\tau\widetilde{\xi}_{i,j}^n(1)-\tau\widetilde{\xi}_{i,j}^n(2)|^2}{(1-(k_n/n)^{1/2-\iota})^3}|\xi_{i,j}^n(4)-1|^2.$$

For $b_i^n(1)$ and $b_i^n(3)$, we have

$$\mathbb{E}(|b_i^n(1)| + |b_i^n(3)|) \leq K(\tau^2 \vee 1)\frac{k_n}{n}.$$

For $b_i^n(2)$, by an application of Hölder inequality, we first have

$$\mathbb{E}\left|b_i^n(2) - \Phi'(\tau)\,\tau(\xi_{i,j}^n(4)-1)1_{\left\{|\xi_{i,j}^n(4)-1|<(\frac{k_n}{n})^{1/2-\iota}\right\}}\right| \leq K|\tau|\frac{1}{\sqrt{n}}.$$

Then,

$$\mathbb{E}\left(\frac{1}{\lfloor n/k_n\rfloor m_n}\sum_{j=1}^{\lfloor n/k_n\rfloor}\sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n}\Phi'(\tau)\,\tau(\xi_{i,j}^n(4)-1)1_{\left\{|\xi_{i,j}^n(4)-1|<(\frac{k_n}{n})^{1/2-\iota}\right\}}\right)^2 \leq K\frac{k_n m_n}{n^2}.$$

Therefore, altogether we get

$$\mathbb{E}|A_5^n| \leq K(|\tau| \vee \tau^2)\frac{k_n}{n}. \tag{10.36}$$

We turn now to A_1^n and A_2^n. Using secon-order Taylor expansion, we can extract the leading terms in A_n^2. In particular, we denote

$$\begin{cases} A_2^n(1) = \frac{1}{\lfloor n/k_n\rfloor}\sum_{j=1}^{\lfloor n/k_n\rfloor}\Phi'(\tau)\tau\overline{\xi}_j^n(1), \\ A_2^n(2) = \frac{1}{\lfloor n/k_n\rfloor}\sum_{j=1}^{\lfloor n/k_n\rfloor}(0.5\Phi''(\tau)\tau^2(\overline{\xi}_j^n(1))^2 - \Phi'(\tau)\tau\overline{\xi}_j^n(2)). \end{cases}$$

With this notation, using the bounds in (10.27), as well as the boundedness of Φ''', we have

$$\mathbb{E}|A_2^n - A_2^n(1) - A_2^n(2)| \leq K(|\tau|^3 \vee |\tau|^2)\left[\left(\frac{k_n}{n}\right)^{3/2}\bigvee\left(\frac{1}{k_n}\right)^{3/2}\right]. \tag{10.37}$$

Further, upon denoting with $\widehat{A}_2^n(1)$ and $\widehat{A}_2^n(2)$ the counterparts of $A_2^n(1)$ and $A_2^n(2)$ with $\overline{\xi}_j^n(1)$ and $\overline{\xi}_j^n(2)$ replaced with $\widehat{\xi}_j^n(1)$ and $\widehat{\xi}_j^n(2)$ respectively, we have using the bounds in (10.25) (as well as the restriction on the rate of growth of k_n in (5.1))

$$\mathbb{E}|A_2^n(1) + A_2^n(2) - \widehat{A}_2^n(1) - \widehat{A}_2^n(2)| \leq K(|\tau| \vee \tau^2)\left(\frac{1}{\sqrt{n}}\bigvee\frac{k_n}{n}\right). \tag{10.38}$$

Thus we are left with the terms A_1^n, $\widehat{A}_2^n(1)$ and $\widehat{A}_2^n(2)$. For $\widehat{A}_2^n(2)$, using

$$\mathbb{E}^n_{(j-1)k_n}\left(\widehat{\xi}_j^n(1)\right)^2 = 2\mathbb{E}^n_{(j-1)k_n}(\widehat{\xi}_j^n(2)) = \frac{1}{4}\frac{1}{k_n}\left(\left(\frac{\pi}{2}\right)^2+\pi-3\right)+o\left(\frac{1}{k_n}\right),$$

we have

$$k_n\widehat{A}_2^n(2) \xrightarrow{\mathbb{P}} \frac{\tau^2\Phi''(\tau)-\tau\Phi'(\tau)}{8}\left(\left(\frac{\pi}{2}\right)^2+\pi-3\right), \tag{10.39}$$

locally uniformly in τ. We finally will show that

$$\left(\sqrt{\lfloor n/k_n\rfloor m_n}A_1^n \quad \sqrt{\lfloor n/k_n\rfloor k_n}\widehat{A}_2^n(1)\right) \xrightarrow{\mathcal{L}} (Z_1(\tau) \quad Z_2(\tau)), \tag{10.40}$$

locally uniformly in τ. We have

$$\begin{pmatrix} \sqrt{\lfloor n/k_n\rfloor m_n}A_1^n \\ \sqrt{\lfloor n/k_n\rfloor k_n}\widehat{A}_2^n(1)\end{pmatrix} = \sum_{i=1}^{\lfloor n/k_n\rfloor k_n}\begin{pmatrix} \zeta_i^n(1) \\ \frac{\Phi'(\tau)\tau}{2}(\zeta_i^n(2)+\zeta_i^n(3))\end{pmatrix} + \begin{pmatrix} 0 \\ \frac{\Phi'(\tau)\tau}{2}\widetilde{\zeta}^n\end{pmatrix},$$

with

$$\zeta_i^n = \begin{pmatrix} \frac{1}{\sqrt{\lfloor n/k_n\rfloor m_n}}\left[1\left(\sqrt{n}\Delta_i^n W\le\tau\right)-\Phi(\tau)\right] \\ \frac{1}{\sqrt{\lfloor n/k_n\rfloor k_n}}\frac{\pi}{2}|\sqrt{n}\Delta_{i-1}^n W|\left(|\sqrt{n}\Delta_i^n W|-\sqrt{\frac{2}{\pi}}\right) \\ \frac{1}{\sqrt{\lfloor n/k_n\rfloor k_n}}\sqrt{\frac{\pi}{2}}\left(|\sqrt{n}\Delta_i^n W|-\sqrt{\frac{2}{\pi}}\right)\end{pmatrix}, \quad i\in I^n,$$

where $I^n=\{i=(j-1)k_n+1,...,(j-1)k_n+m_n, \quad j=1,...,\lfloor n/k_n\rfloor\}$, and for $i=1,....,n\setminus I^n$, ζ_i^n is exactly as above with only the first element being replaced with zero, and finally

$$\widetilde{\zeta}^n = -\frac{(\pi/2)}{\sqrt{\lfloor n/k_n\rfloor k_n}}\sum_{j=1}^{\lfloor n/k_n\rfloor}\Bigg[|\sqrt{n}\Delta^n_{(j-1)k_n}W|\left(|\sqrt{n}\Delta^n_{(j-1)k_n+1}W|-\sqrt{\frac{2}{\pi}}\right) \\ + \sqrt{\frac{2}{\pi}}\left(|\sqrt{n}\Delta^n_{jk_n}W|-\sqrt{\frac{2}{\pi}}\right)\Bigg],$$

where we set $\Delta_0^n W=0$. With this notation, we have

$$\mathbb{E}\left(\widetilde{\zeta}^n\right)^2 \le \frac{K}{k_n}.$$

Further,

$$\mathbb{E}^n_{i-1}(\zeta_i^n)=0, \qquad \sum_{i=1}^{\lfloor n/k_n\rfloor k_n}\mathbb{E}^n_{i-1}||\zeta_i^n||^{2+\iota} \longrightarrow 0, \quad \forall\iota>0,$$

$$\sum_{i=1}^{\lfloor n/k_n \rfloor k_n} \mathbb{E}_{i-1}^n \left[\zeta_i^n (\zeta_i^n)'\right] \longrightarrow \begin{pmatrix} \Phi(\tau)(1-\Phi(\tau)) & 0 & 0 \\ 0 & \left(\frac{\pi}{2}\right)^2 \left(1-\frac{2}{\pi}\right) & \frac{\pi}{2}\left(1-\frac{2}{\pi}\right) \\ 0 & \frac{\pi}{2}\left(1-\frac{2}{\pi}\right) & \frac{\pi}{2}\left(1-\frac{2}{\pi}\right) \end{pmatrix},$$

because recall $m_n/k_n \to 0$. Combining the last two results, we have the convergence in (10.40), pointwise in τ, by an application of Theorem VIII.3.32 in [10]. Application of Theorem 12.3 in [7], extends the convergence to local uniform in τ.

Altogether, the limit behavior of $\widehat{G}_n(\tau) - \Phi(\tau)$ is completely characterized by the limits in (10.39)-(10.40) and

$$\sup_{\tau \in \mathcal{A}} |\widehat{G}_n(\tau) - \Phi(\tau) - A_1^n - \widehat{A}_2^n(1) - \widehat{A}_2^n(2)| = o_p\left(\frac{1}{k_n}\right), \tag{10.41}$$

where $\mathcal{A}$ is a compact subset of $(-\infty, 0)$, with the result in (10.41) following from the bounds on the order of magnitude derived above.

10.4.3. *The difference* $\widehat{F}_n(\tau) - \widetilde{F}_n(\tau)$. To analyze the difference $\widehat{F}_n(\tau) - \widetilde{F}_n(\tau)$, we use the following inequality

$$\begin{aligned} &\mathbb{P}\left(|\Delta_i^n X| > \alpha \sqrt{\widehat{V}_j^n} n^{-\varpi}\right) \\ &\quad \leq \mathbb{P}\left(\left|\frac{\sqrt{\widehat{V}_j^n}}{\sigma_{\frac{(j-1)k_n}{n}}} - 1\right| > 0.5\right) + \mathbb{P}\left(|\Delta_i^n X| > 0.5\alpha\sigma_{\frac{(j-1)k_n}{n}} n^{-\varpi}\right). \end{aligned}$$

For the first probability on the right-hand side of the above inequality we can use the bounds in (10.24), (10.25) and (10.28), while for the second one we can use the exponential inequality for continuous martingales with bounded variation, see e.g., [16], as well as the algebraic inequality $|\sum_i a_i|^p \leq \sum_i |a_i|^p$ for $p \in (0, 1]$, to conclude

$$\mathbb{P}\left(|\Delta_i^n X| > \alpha \sqrt{\widehat{V}_j^n} n^{-\varpi}\right) \leq K\left[\frac{k_n}{n} \vee n^{-1+\iota\varpi}\right], \quad \forall \iota > 0. \tag{10.42}$$

Since $k_n/\sqrt{n} \to 0$ and from the result of the previous two subsections $\widetilde{F}_n(\tau) - \Phi(\tau) = O_p\left(\frac{1}{k_n}\right)$, we get from here

$$\sup_{\tau \in \mathcal{A}} |\widehat{F}_n(\tau) - \widetilde{F}_n(\tau)| = o_p\left(\frac{1}{k_n}\right), \tag{10.43}$$

for any compact subset $\mathcal{A}$ of $(-\infty, 0)$. □

10.5. *Proof of Theorem 4.* The proof follows exactly the same steps as the proof of Theorem 3 and we use analogous notation as in that proof. The only nontrivial difference in analyzing the term $\widetilde{F}'_n(\tau) - \widehat{G}'_n(\tau)$ regards the difference $|\widehat{C}^n_j - \dot{C}^n_j|$ (and $|\widehat{C}^n_j(i) - \dot{C}^n_j(i)|$). For it, we make use of the following algebraic inequality

$$\begin{aligned}|x^2 1_{\{|x|\le a\}} - y^2 1_{\{|y|\le a\}}| \le |x-y|^2 1_{\{|x-y|\le 2a\}} + 2a|x-y|1_{\{|x-y|\le 2a\}} \\ + 2|y|^2 1_{\{|y|>a/2\}} + a^2 1_{\{|x-y|>a/2\}}.\end{aligned}$$

Using the above inequality, the bound in (10.15), as well as the exponential inequality for continuous martingales with bounded variation, see e.g., [16], we have

$$\mathbb{P}\left(|\widehat{C}^n_j - \dot{C}^n_j| \ge \eta_n\right) \le K\frac{k_n}{n}. \tag{10.44}$$

Then, upon picking $\eta_n \propto n^{-q-\iota}$ for $\iota \in (0, 1/2 - q)$ sufficiently small, we get $\sup_{\tau\in\mathcal{A}} |\widetilde{F}'_n(\tau) - \widehat{G}'_n(\tau)| = o_p\left(\frac{1}{k_n}\right)$ for any compact subset $\mathcal{A}$ of $(-\infty, 0)$.

Further, for $\widehat{G}'_n(\tau) - \Phi(\tau)$ the only difference from the analysis of the corresponding term in the proof of Theorem 3 is that now we have

$$k_n \widehat{A}^n_2(2) \xrightarrow{\mathbb{P}} \frac{\tau^2\Phi''(\tau) - \tau\Phi'(\tau)}{4},$$

and further now

$$\begin{pmatrix} \sqrt{\lfloor n/k_n\rfloor m_n} A^n_1 \\ \sqrt{\lfloor n/k_n\rfloor k_n} \widehat{A}^n_2(1) \end{pmatrix} = \sum_{i=1}^{\lfloor n/k_n\rfloor k_n} \begin{pmatrix} \zeta^n_i(1) \\ \frac{\Phi'(\tau)\tau}{2}\zeta^n_i(2) \end{pmatrix},$$

with

$$(\zeta^n_i)' = \begin{pmatrix} \frac{1}{\sqrt{\lfloor n/k_n\rfloor m_n}}[1(\sqrt{n}\Delta^n_i W \le \tau) - \Phi(\tau)] & \frac{1}{\sqrt{\lfloor n/k_n\rfloor k_n}}((\sqrt{n}\Delta^n_i W)^2 - 1) \end{pmatrix}, \quad i \in I^n,$$

where $I^n = \{i = (j-1)k_n + 1, ..., (j-1)k_n + m_n, \quad j = 1, ..., \lfloor n/k_n\rfloor\}$, and for $i = 1, ..., n \setminus I^n$, ζ^n_i is exactly as above with only the first element being replaced with zero. From here the analysis of $\widehat{G}'_n(\tau) - \Phi(\tau)$ is done exactly as that of the corresponding term in the proof of Theorem 3.

We are left with showing the result in the case when jumps in X can be of infinite activity (under the conditions in the theorem). We again follow the steps of the proof of Theorem 3. We replace $R^{(4)}_j$ with $\widehat{C}^n_j - \dot{C}^n_j$ in $\overline{\xi}^n_j(1)$ and $\overline{\xi}^n_j(2)$ and similarly we replace $\widetilde{R}^{(4)}_{i,j}(i)$ with $\widehat{C}^n_j - \dot{C}^n_j - (\Delta^n_i X)^2 1_{\{|\Delta^n_i X| \le \alpha n^{-\varpi}\}} + |\Delta^n_i A + \Delta^n_i B|^2$ in $\widetilde{\xi}^n_{i,j}(1)$ and $\widetilde{\xi}^n_{i,j}(2)$.

Using the inequality in (10.44), and since $\int_E |\delta^Y(x)|^{\beta'}\nu(dx) < \infty$ (upon localization that bounds the size of the jumps), we have

$$\mathbb{E}\left|(\Delta_i^n X)^2 1_{\{|\Delta_i^n X| \le \alpha n^{-\varpi}\}} - (\Delta_i^n A + \Delta_i^n B)^2\right|^p \le K n^{-1-(2p-\beta')\varpi}, \quad \forall p \ge \beta'/2, \tag{10.45}$$

and from here

$$\mathbb{E}|\widehat{C}_j^n - \dot{C}_j^n|^p \le K n^{p-1-(2p-\beta')\varpi}, \quad \forall p \ge 1. \tag{10.46}$$

Using the bounds in (10.45) and (10.46), we can prove exactly as in the proof of Theorem 3 for some deterministic sequence of positive numbers η_n

$$\begin{aligned}&\mathbb{P}\left((|\chi_{i,j}^n(1)| + |\chi_{i,j}^n(2)|) > \eta_n\right)\\ &\le K\left[\frac{1}{\eta_n^p[n^{p/2} \wedge (n/k_n)^p \wedge k_n^{3p/2}]} \bigvee \left(\frac{k_n}{n}\right)^{\frac{1}{1+\iota}} \frac{1}{\eta_n^\iota} \bigvee \frac{n^{-1+\beta'/2}}{\eta_n^{\beta'}} \bigvee \frac{n^{-(2-\beta')\varpi}}{\eta_n^{1/2} \wedge \eta_n\sqrt{k_n}}\right].\end{aligned}$$

From here, using the rate of growth condition in (6.8), upon appropriately choosing η_n, we get

$$\sup_{\tau \in \mathcal{A}} |\widetilde{F}_n'(\tau) - \widehat{G}_n'(\tau)| = o_p\left(\frac{1}{k_n}\right), \tag{10.47}$$

for any compact subset $\mathcal{A}$ of $(-\infty, 0)$.

We turn next to $\widehat{G}_n'(\tau) - \Phi(\tau)$ and we derive the bounds of those terms in the decomposition of the latter which are different from the case of finite jump activity proved above (the term A_5 is identically zero since σ_t is constant). First, for A_3^n, using (10.46) as well as the independence of W_t and Y_t, we have

$$A_3^n \le K(\sqrt{|\tau|} \vee \tau^2) \times \left(\frac{1}{\sqrt{\lfloor n/k_n \rfloor m_n}} n^{1/2-(4-\beta')\varpi/2} \bigvee k_n^{1/4} n^{-1/4-(4-\beta')\varpi/4} \bigvee \frac{1}{\sqrt{n}}\right). \tag{10.48}$$

Next, if we exclude $\widehat{C}_j^n - \dot{C}_j^n$ from $\overline{\xi}_j^n(1)$ and $\overline{\xi}_j^n(2)$, we get for A_4, using (10.45) and (10.46), as well as applying Hölder inequality

$$\mathbb{E}|A_4^n| \le K(|\tau| \vee \tau^2)\left(\frac{1}{k_n^{3/2}} \bigvee \frac{1}{\sqrt{\lfloor n/k_n \rfloor} k_n} \bigvee \frac{n^{-(2-\beta')\varpi}}{\sqrt{k_n}} \bigvee n^{1-(4-\beta')\varpi}\right). \tag{10.49}$$

Combining the bounds in (10.48)-(10.49), and taking into account the growth condition in (6.8), we get

$$\sup_{\tau \in \mathcal{A}} |\widehat{G}_n'(\tau) - \Phi(\tau) - A_1^n - \widehat{A}_2^n(1) - \widehat{A}_2^n(2)| = o_p\left(\frac{1}{k_n}\right), \tag{10.50}$$

where $\mathcal{A}$ is a compact subset of $(-\infty, 0)$. The limit behavior of the triple $(A_1^n,\ \widehat{A}_2^n(1),\ \widehat{A}_2^n(2))$ is derived as in the finite jump activity case in the first part of the proof and this together with (10.47) and (10.50) yields the stated result in the case of infinite variation jumps. □

Acknowledgements. Research partially supported by NSF Grant SES-0957330. We would like to thank Dobrislav Dobrev, Jean Jacod, Per Mykland, Mark Podolskij, Markus Reiss, Mathieu Rosenbaum and many seminar participants for helpful comments and suggestions. We also thank an Associate Editor and a referee for careful read and many constructive comments.

References.

[1] Ait-Sahalia, Y. and J. Jacod (2010). Is Brownian Motion Necessary to Model High Frequency Data? *Annals of Statistics 38*, 3093–3128.

[2] Andersen, T., T. Bollerslev, and D. Dobrev (2007). No Arbitrage Semimartingale Restrictions for Continuous-time Volatility Models subject to Leverage Effects, Jumps and i.i.d. Noise: Theory and Testable Distributional Implications. *Journal of Econometrics 138*, 125–180.

[3] Andrews, B., M. Calder, and R. Davis (2009). Maximum Likelihood Estimation of α-stable Autoregressive Processes. *Annals of Statistics 37*, 1946–1982.

[4] Barndorff-Nielsen, O. and N. Shephard (2001). Non-Gaussian Ornstein–Uhlenbeck-Based Models and some of Their Uses in Financial Economics. *Journal of the Royal Statistical Society Series B, 63*, 167–241.

[5] Barndorff-Nielsen, O. and N. Shephard (2004). Power and Bipower Variation with Stochastic Volatility and Jumps. *Journal of Financial Econometrics 2*, 1–37.

[6] Barndorff-Nielsen, O. and N. Shephard (2006). Econometrics of Testing for Jumps in Financial Economics using Bipower Variation. *Journal of Financial Econometrics 4*, 1–30.

[7] Billingsley, P. (1968). *Convergence of Probability Measures*. New York: Wiley.

[8] Hansen, P. and A. Lunde (2006). Realized Variance and Market Microstructure Noise. *Journal of Business Economics and Statistics 24*, 127–161.

[9] Jacod, J. and P. Protter (2012). *Discretization of Processes*. Berlin: Springer-Verlag.

[10] Jacod, J. and A. Shiryaev (2003). *Limit Theorems For Stochastic Processes* (2nd ed.). Berlin: Springer-Verlag.

[11] Klüppelberg, C., T. Meyer-Brandis, and A. Schmidt (2010). Electricity Spot Price Modelling with a View Towards Extreme Spike Risk. *Quantitative Finance 10*, 963–974.

[12] Mancini, C. (2009). Non-parametric Threshold Estimation for Models with Stochastic Diffusion Coefficient and Jumps. *Scandinavian Journal of Statistics 36*, 270–296.

[13] Mikosch, T., S. Resnick, H. Rootzen, and A. Stegeman (2002). Is Network Traffic Approximated by Stable Levy Motion or Fractional Brownian Motion? *Annals of Applied Probability 12*, 23–68.

[14] Mykland, P., N. Shephard, and K. Sheppard (2012). Efficient and Feasible Inference for the Components of Financial Variation using Blocked Multipower Variation. Technical report.

[15] Mykland, P. and L. Zhang (2009). Inference for Continuous Semimartingales Observed at High Frequency. *Econometrica 77*, 1403–1445.

[16] Revuz, D. and M. Yor (1999). *Continuous Martingales and Brownian Motion*. Springer.

[17] Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge, UK: Cambridge University Press.

[18] Todorov, V. and G. Tauchen (2011). Limit Theorems for Power Variations of Pure-Jump Processes with Application to Activity Estimation. *Annals of Applied Probability 21*, 546–588.

[19] Todorov, V. and G. Tauchen (2012a). Realized Laplace Transforms for Pure-Jump Semimartingales. *Annals of Statistics 40*, 1233–1262.

[20] Todorov, V. and G. Tauchen (2012b). The Realized Laplace Transform of Volatility. *Econometrica 80*, 1105–1127.

[21] van der Vaart, A. (1998). *Asymptotic Statistics*. Cambridge: Cambridge University Press.

[22] Woerner, J. (2007). Inference in Lévy-type Stochastic Volatility Models. *Advances in Applied Probability 39*, 531–549.

Department of Finance
Northwestern University
Evanston, IL 60208-2001
E-mail: v-todorov@northwestern.edu

Department of Economics
Duke University
Durham, NC 27708-0097
E-mail: george.tauchen@duke.edu