Limit Theorems for the Empirical Distribution Function of Scaled Increments of Itô Semimartingales at high frequencies*

Viktor Todorov † and George Tauchen‡

February 14, 2013

Abstract

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.

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

JEL classification: C51, C52, G12.

*Research partially supported by NSF Grant SES-0957330.
†Department of Finance, Kellogg School of Management, Northwestern University, Evanston, IL 60208; e-mail: v-todorov@northwestern.edu.
‡Department of Economics, Duke University, Durham, NC 27708; e-mail: george.tauchen@duke.edu.
1 Introduction

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

\[ dX_t = \alpha_t dt + \sigma_t dW_t + dY_t, \quad (1.1) \]

where \( \alpha_t \) and \( \sigma_t \) 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 Jacod and Shiryaev (2003)).

At high-frequencies, provided \( \sigma_t \) 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

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

where \( B_t \) is a Brownian motion and the above convergence is for the Skorokhod topology. There are two distinctive features 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 \( \sigma_t \)). 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 Barndorff-Nielsen and Shephard (2004, 2006) and the many other alternative measures of powers of volatility summarized in the recent book of Jacod and Protter (2012). Another important example is the general approach of Mykland and Zhang (2009) (see also Mykland...
et al. (2012)) 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, 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 variance of the increments, $\sigma^2$, 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 \( \sigma_t \) 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., Barndorff-Nielsen and Shephard (2001); Andrews et al. (2009), Mikosch et al. (2002), Klüppelberg et al. (2010) and Woerner (2007). 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. Section 6 constructs a feasible test for local Gaussianity using our limit theory and in Sections 7 and 8 we apply the test on simulated and real financial data respectively. The proofs are given in Section 9.

## 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
\[ dX_t = \alpha_t dt + \sigma_t - dS_t + dY_t, \quad (2.1) \]
where \( \alpha_t, \sigma_t \) 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., Sato (1999), given by
\[
\log \left[ \mathbb{E}(e^{iuS_t}) \right] = -t|\beta (1 - i\gamma \text{sign}(u) \Phi), \quad \Phi = \begin{cases} 
\tan(\pi\beta/2) & \text{if } \beta \neq 1, \\
-2\pi \log |u| & \text{if } \beta = 1,
\end{cases}
\]
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 \( \beta < 2 \), \( X_t \) 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 \( \alpha_t \) and \( \sigma_t \)), 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 \( X_t \) 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
\[
\mathbb{E}\left( |\sigma_t^{(p)} - \sigma_s^{(p)}|^2 |\mathcal{F}_s \right) \leq K_p |t - s|, \text{ for every } 0 \leq s \leq t \leq 1. \quad (2.3)
\]

**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
\[
\mathbb{E}\left( |Y_t^p - Y_s^p|^q |\mathcal{F}_s \right) \leq K_p |t - s|, \text{ for every } 0 \leq s \leq t \leq 1. \quad (2.4)
\]
and for every $q \in (\beta', 2)$ where $\beta' < \beta$.

The assumption in (2.3) can be easily verified for Itô semimartingales which is the typical way of modeling $\sigma_t$, 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 Jacod and Protter (2012).

Under assumption A, we can extend the local Gaussianity result in (1.2) to (see e.g., Lemma 1 of Todorov and Tauchen (2012a))

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

(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. 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

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

(2.6)

where $\mu$ 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, $\sigma_t$ is an Itô semimartingale having the following representation

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

(2.7)

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

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

6
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 $|\delta^Y(t, x)| \wedge 1 + |\delta^\sigma(t, x)| \wedge 1 + |\delta^\tilde{\sigma}(t, x)| \wedge 1 \leq \gamma_p(x)$ for $t \leq T_p$.

The Itô semimartingale restriction on $\sigma_t$ (and its coefficients) is satisfied in most applications and from a practical point of view is not very restrictive. 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. This assumption can be further relaxed to allow for infinite activity jumps.

3 Empirical CDF of the “Devolatilized” High Frequency Increments

Throughout the paper we assume that $X$ is observed on the equidistant grid $0, \frac{1}{n}, \ldots, 1$ 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 $\Delta^n_i X = X_{i\frac{1}{n}} - X_{i+1\frac{1}{n}}$ are approximately Gaussian with variance given by the value of the process $\sigma^2_t$ at the beginning of the increment. Of course, the stochastic volatility $\sigma_t$ 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 $\sigma_t$ 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 $\sigma^2_t$ is given by

$$\hat{V}^n_j = \frac{\pi n}{2 k_n} \sum_{i=(j-1)k_n+2}^{jk_n} |\Delta^{n-1}_i X||\Delta^n_i X|, \quad j = 1, \ldots, \lfloor n/k_n \rfloor.$$ (3.1)

$\hat{V}^n_j$ is the Bipower Variation proposed by Barndorff-Nielsen and Shephard (2004, 2006) for measuring the quadratic variation of the diffusion component of $X$. 

We note that an alternative measure of $\sigma_t$ can be constructed using the so-called truncated variation. In 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 a truncated variation type estimator of $\sigma_t$ 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.

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 a feasible CLT later on. Finally, we need to remove the high-frequency increments that contain “big” jumps. The total number of increments used in our statistic is thus given by

$$N^n(\alpha, \varpi) = \sum_{j=1}^{\lfloor n/k_n \rfloor} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1 \left( |\Delta^n_i X| \leq \alpha \sqrt{\hat{V}^n_j n^{-\varpi}} \right),$$

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

$$\hat{F}^n(\tau) = \frac{1}{N^n(\alpha, \varpi)} \sum_{j=1}^{\lfloor n/k_n \rfloor} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1 \left( \frac{\sqrt{n} \Delta^n_i X}{\sqrt{\hat{V}^n_j}} \leq \alpha \sqrt{\hat{V}^n_j n^{-\varpi}} \right) \leq \tau,$$

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

4 Convergence in probability of $\hat{F}^n(\tau)$

We next derive the limiting behavior of $\hat{F}^n(\tau)$ 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

\[ k_n \sim n^q, \quad \text{for some } q \in (0, 1), \tag{4.4} \]

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

\[ \hat{F}_n(\tau) \xrightarrow{p} F_\beta(\tau), \quad \text{as } n \to \infty, \tag{4.5} \]

where the above convergence is uniform in \( \tau \) over compact subsets of \((-\infty, 0) \cup (0, +\infty)\); \( F_\beta(\tau) \) is the cdf of \( \sqrt{2\pi} \frac{S_1}{E|S_1|} \) (\( S_1 \) is the value of the \( \beta \)-stable process \( S_t \) at time 1) and \( F_2(\tau) \) equals the cdf of a standard normal variable \( \Phi(\tau) \).

We note that the convergence result in (4.5) is for values of \( \tau \) excluding a neighborhood of zero. The reason is that the cdf for those values of \( \tau \) are affected by the presence of the drift term as well as \( Y_t \). Similarly, due to the “big” jumps, we derive the behavior of the statistic only on compact sets of \( \tau \) as the estimate of the cdf for \( \tau \to \pm\infty \) is affected by the truncation.

Remark 1. The limit result in (4.5) shows that when \( S_t \) is stable with \( \beta < 2 \), \( \hat{F}_n(\tau) \) estimates the cdf of a \( \beta \)-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 \( \hat{V}_j^n \) by \( n^{1/\beta - 1/2} \) in order for the latter to converge to a non-degenerate limit (that is proportional to \( \sigma_t^2 \)). Hence the ratio

\[ \frac{\sqrt{n} \Delta_i^n X}{\sqrt{\hat{V}_j^n}} = \frac{n^{1/\beta} \Delta_i^n X}{n^{2/\beta - 1/2} \hat{V}_j^n}, \tag{4.6} \]

is appropriately scaled even in the case when \( \beta < 2 \) and importantly without knowing apriori the value of \( \beta \). 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 \( \beta \). That is, in all cases of \( \beta \), \( F_\beta(\tau) \) corresponds to the cdf of a random variable \( Z \) with \( E|Z| = \sqrt{\frac{2}{\pi}} \). Therefore, the difference between \( \beta < 2 \) and the null \( \beta = 2 \) will be in the relative probability assigned to “big” versus “small” values of \( \tau \).

We note further that in Theorem 1 we restrict \( \beta > 1 \). The reason is that for \( \beta \leq 1 \), the limit behavior of \( \hat{F}_n(\tau) \) 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 (4.5), we need to use $\Delta^n_i X - \Delta^n_{i-1} X$ in the construction of $\hat{F}_n(\tau)$ which essentially eliminates the drift term.

We next derive the limiting behavior of $\hat{F}_n(\tau)$ 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 $\hat{F}_n(\tau)$ be given by (3.3) with $\Delta^n_i X$ replaced with $\Delta^n_i X^*$ for $X^*_i = X_i + \epsilon_i$, and where $\{\epsilon_i\}_{i=1,...,n}$ are i.i.d. random variables defined on a product extension of the original probability space and independent from $\mathcal{F}$. Further, suppose $\mathbb{E}|\epsilon_i|^{1+i} < \infty$ for some $i > 0$. Finally, assume that the cdf of $\frac{1}{\sqrt{n}} (\epsilon_1 - \epsilon_i)$, $F_\epsilon(\tau)$, is continuous where we denote

$$
\mu = \sqrt{\frac{\pi}{2}} \sqrt{\mathbb{E} \left( |\epsilon_i - \epsilon_i| |\epsilon_i - \epsilon_{i-1} - \epsilon_{i-2}| \right)}.
\]

Then

$$
\hat{F}_n(\tau) \xrightarrow{p} F_\epsilon(\tau), \quad \text{as } n \to \infty,
$$

where the above convergence is uniform in $\tau$ over compact subsets of $(-\infty, 0) \cup (0, +\infty)$.

**Remark 2.** 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 $\Delta^n_i X^*$ but this is offset in the ratio in $\hat{F}_n(\tau)$ by a scaling factor for the local variance estimator $\hat{V}_j^n$ 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 $\hat{F}_n(\tau)$ is given by

$$
\frac{\sqrt{n} \Delta^n_i X}{\sqrt{\hat{V}_j^n}} = \frac{\Delta^n_i X}{\sqrt{n^{-1}\hat{V}_j^n}},
$$

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

The limit result in (4.7) provides an important insight into the noise by studying its distribution. We stress the fact that the presence of $\hat{V}_j^n$ in the
truncation is very important for the limit result in (4.7). 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 $\hat{V}_j^n$ (i.e., $\hat{V}_j^n$ 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 $\epsilon_n$ 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 $\hat{F}_n(\tau)$ 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 for $\hat{F}_n(\tau)$ 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

$$\frac{m_n}{k_n} \to 0, \quad k_n \sim n^q, \quad \text{for some } q \in (0,1/2), \quad \text{when } n \to \infty. \quad (5.9)$$

We then have locally uniformly in subsets of $(-\infty,0) \cup (0,\infty)$

$$\hat{F}_n(\tau) - \Phi(\tau) = \hat{Z}_1^n(\tau) + \hat{Z}_2^n(\tau) + \frac{1}{k_n} \tau^2 \Phi''(\tau) - \tau \Phi'(\tau) \left( \left( \frac{\pi}{2} \right)^2 + \pi - 3 \right) + o_p \left( \frac{1}{k_n} \right), \quad (5.10)$$

$$(\sqrt{[n/k_n]} m_n \hat{Z}_1^n(\tau) \quad \sqrt{[n/k_n]} k_n \hat{Z}_2^n(\tau)) \overset{d}{\to} (Z_1(\tau) \quad Z_2(\tau)), \quad (5.11)$$

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

\[
\begin{align*}
\text{Cov}(Z_1(\tau_1), Z_1(\tau_2)) &= \Phi(\tau_1 \wedge \tau_2) - \Phi(\tau_1)\Phi(\tau_2), \\
\text{Cov}(Z_2(\tau_1), Z_2(\tau_2)) &= \left[ \frac{\tau_1\Phi'(\tau_1) + \tau_2\Phi'(\tau_2)}{2} \right] \left( \left( \frac{\pi}{2} \right)^2 + \pi - 3 \right), \quad \tau_1, \tau_2 \in \mathbb{R} \setminus 0.
\end{align*}
\]

(5.12)

We make several observations regarding the limiting result in (5.10)-(5.12). The first term of \( \hat{F}_n(\tau) - \Phi(\tau) \) in (5.10), \( \hat{Z}_n^1(\tau) \), converges to \( Z_1(\tau) \) which is the standard Brownian bridge appearing in the Donsker theorem for empirical processes, see e.g., van der Vaart (1998). The second and third term on the right-hand side of (5.10) are due to the estimation error in recovering the local variance, i.e., the presence of \( \hat{V}_j^n \) in \( \hat{F}_n^n(\tau) \) instead of the true (unobserved) \( \sigma_t^2 \). \( \hat{Z}_n^2(\tau) \) converges to a centered Gaussian process, independent from \( Z_1(\tau) \), while the third term on the right-hand side of (5.10) is an asymptotic bias. Importantly, the asymptotic bias as well as the limiting variance of \( (Z_1(\tau) Z_2(\tau)) \) are all constants that depend only on \( \tau \) and not the stochastic volatility \( \sigma_t \). Therefore, feasible inference based on (5.10) 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 \( \hat{F}_n^n(\tau) \) arbitrary close to \( \sqrt{n} \). We should further point out that this is unlike the rate of estimating the spot \( \sigma_t^2 \) by \( \hat{V}_j^n \) (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 \( \hat{V}_j^n \).

The order of magnitude of the three components on the right-hand side of (5.10) 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 (5.10). The order of magnitude of the other two components can be made equal by setting \( m_n \propto \frac{k_n^3}{n} \), which provides the optimal rate of convergence for \( \hat{F}_n(\tau) - \Phi(\tau) \) with respect to the choice of \( m_n \).

Finally, we point out that a feasible CLT for \( \hat{F}_n(\tau) \) 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 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. The critical region of our proposed test is given by

$$C_n = \left\{ \sup_{\tau \in A} \sqrt{N_n(\alpha, \omega)} \left| \hat{F}_n(\tau) - \Phi(\tau) \right| > q_n(\alpha, A) \right\}$$

(6.13)

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

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

(6.14)

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

The test in (6.13) 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 $A$ over which the difference $\hat{F}_n(\tau) - \Phi(\tau)$ is evaluated. For reasons we already discussed, that are particular to our problem here, we need to exclude arbitrary high in magnitude values of $\tau$ as well as an arbitrary small interval excluding the zero.

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

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

(6.15)
where we make also use of the fact that the stable and standard normal variables have different cdf-s on compact subsets of $\mathbb{R} \setminus 0$ with positive Lebesgue measure. By Theorem 2, the above power result applies also to the case when we observe $X_{n+\epsilon_n}^i$, provided of course the limiting cdf of the noise in (4.7) differs from that of the standard normal on the set $A$.

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 (6.13), unlike earlier work, incorporates this distribution implication of (2.5) as well.

7 Monte Carlo

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

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

where $(W_t, B_t)$ is a vector of Brownian motions with $\text{Corr}(W_t, B_t) = -0.5$ and $\mu$ 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

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

where $S_t$ is a symmetric tempered stable martingale with Lévy measure $\frac{0.1089e^{-|x|}}{|x|^{1.8}}$ and $V_t$ is the square-root diffusion given in (7.1). The process in (7.2) 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 (7.1)). For this process the local Gaussianity does not hold and hence the behavior of the test on data from the model in (7.2) will allow us to investigate the power of the test.

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 1-3 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 $\hat{F}_n(\tau) - F(\tau)$ in our test is evaluated is set to

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

(7.3)

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

The results of the Monte Carlo are reported in Tables 1 and 2. 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 and, if anything, the test slightly favors the null, while it is seen to have excellent power in Table 2. 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 correct (and again the test slightly favors the null) while power is excellent. For larger values of $k_n$ relative to $n$ ($\lfloor n/k_n \rfloor = 2$ or $\lfloor n/k_n \rfloor = 1$) 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.
Table 1: Monte Carlo Results for Jump-Diffusion Model (7.1)

<table>
<thead>
<tr>
<th>Nominal Size</th>
<th>Rejection Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sampling Frequency</strong></td>
<td>$n = 100$</td>
</tr>
<tr>
<td>$\alpha = 1%$</td>
<td>4.1</td>
</tr>
<tr>
<td>$\alpha = 5%$</td>
<td>15.4</td>
</tr>
<tr>
<td><strong>Sampling Frequency</strong></td>
<td>$n = 200$</td>
</tr>
<tr>
<td>$\alpha = 1%$</td>
<td>1.2</td>
</tr>
<tr>
<td>$\alpha = 5%$</td>
<td>4.4</td>
</tr>
</tbody>
</table>

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 $\lfloor m_n/k_n \rfloor = 0.70$.

8 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 $A$ as in (7.3) and $\lfloor n/k_n \rfloor = 2$ for the five-minute sampling frequency and $\lfloor n/k_n \rfloor = 3$ 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-frequency returns by a time-of-day scale factor exactly as in Todorov and Tauchen (2012b).

The results from the test are shown on Figure 1. We can see from the figure that the local Gaussianity hypothesis works very 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 still fairly close to 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
Table 2: Monte Carlo Results for Pure-Jump Model (7.2)

<table>
<thead>
<tr>
<th>Nominal Size Rejection Rate</th>
<th>Sampling Frequency $n = 100$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$k_n = 33$</td>
<td>$k_n = 50$</td>
</tr>
<tr>
<td>$\alpha = 1%$</td>
<td>27.5</td>
<td>87.4</td>
</tr>
<tr>
<td>$\alpha = 5%$</td>
<td>66.6</td>
<td>97.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Nominal Size Rejection Rate</th>
<th>Sampling Frequency $n = 200$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$k_n = 67$</td>
<td>$k_n = 100$</td>
</tr>
<tr>
<td>$\alpha = 1%$</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>$\alpha = 5%$</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

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 $\lfloor m_n/k_n \rfloor = 0.70$.

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 $\beta < 2$.

9 Proofs

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

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

\[ \hat{V}_j^n = \frac{n \pi}{k_n} \sum_{i=(j-1)k_n+1}^{jk_n} |\Delta_{i-1} S||\Delta_i S|, \quad \hat{V}_j^n = \frac{n \pi}{k_n} \sum_{i=(j-1)k_n+1}^{jk_n} |\Delta_{i-1} A + \Delta_i B||\Delta_i A + \Delta_i B|, \]

\[ \tilde{V}_j^n = \frac{n \pi}{k_n} \sum_{i=(j-1)k_n+1}^{jk_n} |\Delta_{i-1} B||\Delta_i S|, \quad \tilde{V}_j^n = \frac{n \pi}{k_n} \sum_{i=(j-1)k_n+1}^{jk_n} |\Delta_{i-1} S||\Delta_i S|, \]

\[ \tilde{V}_j^n(i) = \tilde{V}_j^n - \frac{n \pi}{k_n} (|\Delta_{i-1} S||\Delta_i S| + |\Delta_i S||\Delta_{i+1} S|), \quad i = 1, \ldots, n, \]

where we set $\Delta_0 S = \Delta_{n+1} S = 0$. We also denote

\[ \tilde{F}^n(\tau) = \frac{N^n(\alpha, \varpi)}{[n/k_n] m_n} \tilde{F}^n(\tau). \]

17
Figure 1: Kolmogorov-Smirnov tests for local Gaussianity. The * corresponds to the value of the test \( \sup_{\tau \in A} \sqrt{N_n(\alpha, \varpi)} | \hat{F}_n(\tau) - F(\tau) | \) and the solid lines are the critical values \( q_n(\alpha, A) \) for \( \alpha = 5\% \) and \( \alpha = 1\% \).

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 \( \tau \). We will also use the shorthand notation \( E_n^i(\cdot) = E \left( \cdot | F_{i-1} \right) \).

9.1 Localization

We will prove Theorems 1-3 under the following stronger versions of assumption A and B:

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

**SB.** We have assumption B with all processes \( \alpha_t, \tilde{\alpha}_t, \sigma_t, \sigma_t^{-1}, \tilde{\sigma}_t, \tilde{\sigma}_t^t, Y_t \) and
the coefficients of the Itô semimartingale representations of $\tilde{\sigma}_t$ and $\tilde{\sigma}'_t$ being uniformly bounded on $[0, 1]$. Further $\left(\delta^Y(t, x)| + |\delta^\sigma(t, x)| + |\delta^\tilde{\sigma}(t, x)| + |\delta^\tilde{\sigma}'(t, x)|\right) \leq \gamma(x)$ 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 Jacod and Protter (2012).

9.2 Proof of Theorem 1

Without loss of generality, we will assume that $\tau < 0$, the case $\tau > 0$ being dealt with analogously (by working with $1 - \tilde{F}_n(\tau)$ instead). We first analyze the behavior of $\tilde{V}_j^n$. We denote with $\eta_n$ 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 $j = 1, \ldots, \lfloor n/k_n \rfloor$

$$P \left( n^{2/\beta - 1} |\tilde{V}_j^n - \dot{V}_j^n| \geq \eta_n \right) \leq K \frac{n^{1/\beta - (1/\beta') \wedge 1 + t}}{\eta_n}, \ \forall t > 0.$$ 

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

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

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

$$P \left( n^{2/\beta - 1} |\tilde{V}_j^n - V_j^n| \geq \eta_n \right) \leq K \frac{n^{1/\beta - 1 - \tau}}{\eta_n}, \ \forall t > 0.$$ 

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

$$P \left( n^{2/\beta - 1} V_j^n - \frac{\pi}{2} \sigma_2|^{1/2} (\mathbb{E}|S_1|)^2 | \geq \eta_n \right) \leq K \frac{1}{k_n^{\beta - 1 - \tau} \eta_n}, \ \forall t \in (0, 1 - \beta).$$
Combining these results, we get altogether for $\forall t \in (0, 1 - \beta)$
\[
\mathbb{P}\left( n^{2/\beta - 1} \hat{V}_j^n - \frac{\pi}{2} \sigma^2 \left( \frac{1}{2} \right) \left( \mathbb{E}[S_1] \right)^2 \geq \eta_n \right) \\
\leq K \left( \frac{n^{1/\beta - (1/\beta')} \wedge 1 + \tau}{\eta_n} \sqrt{\frac{k_n^{1/2 - \epsilon}}{n^{1/2 - \epsilon} \eta_n}} \sqrt{\frac{1}{k_n^{\beta - 1 - \epsilon} \eta_n^{\beta - \epsilon}}} \right). \quad (9.2)
\]

Next, for $i = (j - 1)k_n + 1, \ldots, (j - 1)k_n + m_n$ and $j = 1, \ldots, \lfloor n/k_n \rfloor$, we denote
\[
\left\{ \xi_{i,j}^n(1) = n^{1/\beta} \left( \Delta_t^n A + \Delta_t^n Y + \int_{(i-1)\Delta_n}^{i\Delta_n} (\sigma_u - \sigma_{(j-1)k_n}) dS_u \right) \right\}
\]
\[
\left\{ \xi_{i,j}^n(2) = n^{1/\beta} \sigma_{(j-1)k_n} \Delta_t^n S1 \left\{ |\Delta_t^n X| > \alpha \sqrt{\hat{V}_1^n} \right\} \right\}.
\]

With this notation, using similar inequalities as before, we get
\[
\mathbb{P}\left( |\xi_{i,j}^n(1)| \geq \eta_n \right) \leq K \left( \frac{n^{1/\beta - (1/\beta') \wedge 1 + \tau}}{\eta_n} \sqrt{\frac{k_n^{1/2 + \epsilon / 2}}{n^{3/2 + \epsilon / 2} \eta_n^{\beta + \epsilon / 2}}} \right). \quad (9.3)
\]

Next, using the result in (9.2) above as well as Hölder inequality, we get
\[
\mathbb{P}\left( |\xi_{i,j}^n(2)| \geq \eta_n \right) \leq K \frac{n^{-(1/2 - \epsilon) \wedge 1 + \tau} \wedge n^{1/\beta - (1/\beta') \wedge 1 + \tau} \wedge (k_n/n)^{1/2 - \epsilon} \wedge k_n^{1/\beta + 1 - \beta}}{\eta_n}, \quad \forall t > 0. \quad (9.4)
\]

We next denote the set (note that by assumption SA, $\sigma_t$ is strictly above zero on the time interval $[0, 1]$)
\[
A_{i,j}^n = \left\{ \omega : \frac{|\xi_{i,j}^n(1)| + |\xi_{i,j}^n(2)|}{\sqrt{\frac{2}{\sigma} \mathbb{E}[S_1]}} > \eta_n \cup \left| \frac{n^{1/\beta - 1/2} \sqrt{\hat{V}_j^n} - 1}{\sqrt{\frac{2}{\sigma} \sigma_{(j-1)k_n} \mathbb{E}[S_1]}} \right| > \eta_n \right\}, \quad (9.5)
\]
for $i = (j - 1)k_n + 1, \ldots, (j - 1)k_n + m_n$ and $j = 1, \ldots, \lfloor n/k_n \rfloor$.

We now can set (recall (4.4))
\[
\eta_n = n^{-x}, \quad 0 < x < \left[ \left( \frac{1}{\beta'} \wedge 1 - \frac{1}{\beta} \right) \wedge \frac{1 - q}{2} \wedge \frac{q(\beta - 1)}{\beta} \right], \quad (9.6)
\]
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 (4.4). With this choice of $\eta_n$, the results in (9.2), (9.3) and (9.4) imply
\[
\frac{1}{[n/k_n]m_n} \sum_{j=1}^{[n/k_n]} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} \mathbb{P}\left( A_{i,j}^n \right) = o(1). \quad (9.7)
\]
These two results altogether imply

\[
\sup_{\tau \in \mathcal{A}} |\hat{F}_n(\tau) - \hat{G}_n(\tau)| = o_p(1),
\]

(9.8)

where we denote

\[
\hat{G}_n(\tau) = \frac{1}{[n/k_n]m_n} \sum_{j=1}^{[n/k_n]} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1 \left\{ \frac{\sqrt{n} \Delta^n X}{\sqrt{V_j}} \left| \left| \Delta^n X \right| \leq \alpha \sqrt{V_j} \right\} \leq \tau \right\} \mathcal{A}_{n,c}^n.
\]

Taking into account the definition of the set \(\mathcal{A}_{i,j}^n\), we get

\[
\begin{cases}
\hat{G}_n(\tau) \geq \frac{1}{[n/k_n]m_n} \sum_{j=1}^{[n/k_n]} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1 \left\{ \frac{n^{1/2} \Delta^n S}{\sqrt{E[S_1]}} \leq \tau(1 - \eta_n) - \eta_n \right\}, \\
\hat{G}_n(\tau) \leq \frac{1}{[n/k_n]m_n} \sum_{j=1}^{[n/k_n]} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1 \left\{ \frac{n^{1/2} \Delta^n S}{\sqrt{E[S_1]}} \leq \tau(1 + \eta_n) + \eta_n \right\}.
\end{cases}
\]

Using Glivenko-Cantelli theorem, see e.g., Theorem 19.1 of van der Vaart (1998), we have

\[
\sup_{\tau} \left| \frac{1}{[n/k_n]m_n} \sum_{j=1}^{[n/k_n]} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1 \left\{ \frac{n^{1/2} \Delta^n S}{\sqrt{E[S_1]}} \leq \tau(1 - \eta_n) - \eta_n \right\} \right| - F_\beta(\tau(1 - \eta_n) - \eta_n) \xrightarrow{p} 0,
\]

\[
\sup_{\tau} \left| \frac{1}{[n/k_n]m_n} \sum_{j=1}^{[n/k_n]} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1 \left\{ \frac{n^{1/2} \Delta^n S}{\sqrt{E[S_1]}} \leq \tau(1 + \eta_n) + \eta_n \right\} \right| - F_\beta(\tau(1 + \eta_n) + \eta_n) \xrightarrow{p} 0,
\]

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

\[
\sup_{\tau} |F_\beta(\tau(1 - \eta_n) - \eta_n) - F_\beta(\tau)| \rightarrow 0, \quad \sup_{\tau} |F_\beta(\tau(1 + \eta_n) + \eta_n) - F_\beta(\tau)| \rightarrow 0.
\]

These two results altogether imply

\[
\sup_{\tau} |\hat{G}_n(\tau) - F_\beta(\tau)| \xrightarrow{p} 0,
\]

and from here, using (9.8), we have \(\sup_{\tau \in \mathcal{A}} |\hat{F}_n(\tau) - F_\beta(\tau)| = o_p(1)\) for any compact subset \(\mathcal{A}\) of \((-\infty, 0)\). Hence, to prove (4.5), we need only to show

\[
\frac{N^n(\alpha, \pi)}{[n/k_n]m_n} \xrightarrow{p} 1, \quad \text{as } n \to \infty.
\]

(9.9)
We have

\[ P\left(\left|\Delta_n^n X\right| > \alpha \sqrt{V_j^n n^{-\omega}}\right) \leq P\left(\left|\frac{n^{1/3-1/2}}{\sqrt{2} \sigma (\frac{j-1}{n})} E[S] - 1\right| > 0.5\right) \]

\[ + P\left(n^{1/3} \Delta_n^n X > 0.5 \alpha \sqrt{\frac{\pi}{2} \sigma (\frac{j-1}{n})} E[S_1 | n^{1/2-\omega}}\right). \]

From here we can use the bounds in (9.2) and (9.3) to conclude

\[ P\left(\left|\Delta_n^n X\right| > \alpha \sqrt{V_j^n n^{-\omega}}\right) \leq K \eta_n, \]

for some sufficiently small \( \epsilon > 0 \), (9.10)

hence the convergence in (9.9) holds which implies the result in (4.5).

9.3 Proof of Theorem 2

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

\[ P\left(\left|\frac{1}{n} \sqrt{V_j^n} - \mu\right| \geq \eta_n\right) \leq K \frac{n^{-1/2}}{\eta_n}, \]

(9.11)

\[ P\left(\left|\epsilon_n - \epsilon_{n-1}\right| 1\left\{\left|\Delta_n^n X\right| > \alpha \sqrt{V_j^n n^{-\omega}}\right\} \geq \eta_n\right) \leq K \frac{\eta_n^{-1/2}}{\eta_n}. \]

(9.12)

We denote

\[ B_{i,j} = \left\{ \omega : \left|\Delta_n^n X\right| 1\left\{\left|\Delta_n^n X\right| \leq \alpha \sqrt{V_j^n n^{-\omega}}\right\} \leq \tau \right\}, \]

for \( i = (j-1)k_n + 1, \ldots, (j-1)k_n + m_n \) and \( j = 1, \ldots, [n/k_n] \). We set \( \eta_n = n^{-x} \)

for \( 0 < x < \frac{1}{2} (1/2 - \omega) \land 1/2 \). With this choice

\[ \frac{1}{[n/k_n] m_n} \sum_{j=1}^{[n/k_n]} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} P(B_{i,j}) = o(1). \]

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

\[ \sup_{\tau \in A} \left|\tilde{F}_n(\tau) - \tilde{G}_n(\tau)\right| = o_p(1), \]

where we denote

\[ \tilde{G}_n(\tau) = \frac{1}{[n/k_n] m_n} \sum_{j=1}^{[n/k_n]} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1\left\{\frac{\sqrt{n} \Delta_n^n X}{\sqrt{V_j^n}} 1\left\{\left|\Delta_n^n X\right| \leq \alpha \sqrt{V_j^n n^{-\omega}}\right\} \leq \tau\right\} 1(B_{i,j}). \]
Taking into account the definition of the set $B_{i,j}^n$, we get
\[
\begin{align*}
\hat{G}_n(\tau) &\geq \frac{1}{(n/k_n)m_n} \sum_{j=1}^{[n/k_n]} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1_{\{\frac{1}{n} \left( \frac{\epsilon - \epsilon_{i-1}}{\mu} \right) \leq \tau(1 - \eta_n) - \frac{\eta_n}{\mu} \}} \\
\hat{G}_n(\tau) &\leq \frac{1}{(n/k_n)m_n} \sum_{j=1}^{[n/k_n]} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1_{\{\frac{1}{n} \left( \frac{\epsilon - \epsilon_{i-1}}{\mu} \right) \leq \tau(1 + \eta_n) + \frac{\eta_n}{\mu} \}}.
\end{align*}
\]

From here we can proceed exactly in the same way as in the proof of Theorem 1 to show that $\hat{G}_n(\tau) \xrightarrow{p} F_\tau(\tau)$ locally uniformly in $\tau$. Hence we need only show $\frac{N^{\mu}[\alpha, \pi]}{n/k_n} \xrightarrow{p} 1$ as $n \to \infty$. This follows from
\[
P \left(|\Delta^n X^*| > \alpha \sqrt{\frac{V^n_j}{n^{\pi}}} \right) \leq \frac{\left(\frac{\sqrt{V^n_j \nabla^n j}}{n^{\mu}} \right)^{1/2}}{0.5} + P \left(|\Delta^n X^*| > 0.5\alpha \mu n^{1/2 - \pi} \right)
\leq \frac{K}{n^\pi},
\]
for some sufficiently small $\epsilon > 0$,

which can be shown using (9.11), the fact that the noise term has a finite first moment and the Burkholder-Davis-Gundy inequality. \hfill $\square$

### 9.4 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 $\hat{V}_j^n - \hat{V}_j^n$. It is given by the following
\[
\hat{V}_j^n - \hat{V}_j^n = R_j^{(1)} + R_j^{(2)} + R_j^{(3)} + R_j^{(4)}, \quad j = 1, \ldots, \lfloor n/k_n \rfloor, \quad (9.13)
\]

\[
R_j^{(1)} = \frac{n \pi}{2 k_n} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} \left( (\Delta^n_{i-1} B||\Delta^n_{i-1} B) - \sigma_{i(j-1)k_n} \Delta^n_{i-1} W ||\Delta^n_{i-1} W) \\
+ (\sigma_{i(j-1)k_n} - \sigma_{(j-1)k_n})^2 |\Delta^n_{i-1} W||\Delta^n_{i-1} W| \right),
\]

\[
R_j^{(2)} = \frac{2 n \pi}{2 k_n} \sigma_{(j-1)k_n} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} \left[ \sigma_{i(j-1)k_n} - \sigma_{(j-1)k_n} - \int_{(j-1)k_n}^{i \pi/2} \bar{\sigma}_{(j-1)k_n} dW_u \\
- \int_{(j-1)k_n}^{i \pi/2} \bar{\sigma}_{(j-1)k_n} \bar{\sigma}_{(j-1)k_n} dW_u \right] |\Delta^n_{i-1} W||\Delta^n_{i-1} W|,
\]

\[
R_j^{(3)} = \frac{2 n \pi}{2 k_n} \sigma_{(j-1)k_n} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} \left[ \int_{(j-1)k_n}^{i \pi/2} \bar{\sigma}_{(j-1)k_n} dW_u + \int_{(j-1)k_n}^{i \pi/2} \bar{\sigma}_{(j-1)k_n} \bar{\sigma}_{(j-1)k_n} dW_u \right] \times \left( |\Delta^n_{i-1} W||\Delta^n_{i-1} W| - \frac{2}{\pi} \right),
\]

\[
23
\]
We further denote for $i = (j - 1)k_n + 1, \ldots, jk_n - 2$ we set

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

We further denote for $i = (j - 1)k_n + 1, \ldots, (j - 1)k_n + m_n$ and $j = 1, \ldots, \lfloor n/k_n \rfloor$,

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

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

$$\hat{\xi}^n_i(1) = \frac{\tilde{V}_j^n - \sigma_{\frac{(j-1)k_n}{n}}}{2 \sigma^2_{\frac{(j-1)k_n}{n}}}, \quad \hat{\xi}^n_i(2) = \frac{\left( \tilde{V}_j^n - \sigma_{\frac{(j-1)k_n}{n}} \right)^2}{8 \sigma^4_{\frac{(j-1)k_n}{n}}}.$$

$$\xi^n_i(3) = \sqrt{n} \Delta^n \frac{W}{\sigma_{\frac{(j-1)k_n}{n}}} \left[ \tilde{\sigma}_{\frac{(j-1)k_n}{n}} (W_{\frac{j-n}{k_n}} - W_{\frac{j-1}{k_n}}) + \tilde{\sigma}_{\frac{(j-1)k_n}{n}} (W_{\frac{j-n}{k_n}} - W_{\frac{j-1}{k_n}}) \right],$$

$$\xi^n_i(4) = 1 + \frac{1}{\sigma_{\frac{(j-1)k_n}{n}}} \left[ \tilde{\sigma}_{\frac{(j-1)k_n}{n}} (W_{\frac{j-n}{k_n}} - W_{\frac{j-1}{k_n}}) + \tilde{\sigma}_{\frac{(j-1)k_n}{n}} (W_{\frac{j-n}{k_n}} - W_{\frac{j-1}{k_n}}) \right].$$

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

$$\chi^n_{i,j}(1) = -\sqrt{n} \frac{1}{\sigma_{\frac{(j-1)k_n}{n}}} \left( \Delta^n A + \Delta^n Y + \int_{\frac{j-1}{n}}^{\frac{j}{n}} \left( \sigma_u - \sigma_{\frac{j-1}{n}} \right) dW_u \right) 1_{\{\Delta^n X \leq \alpha \sqrt{\tilde{V}_j^n n^{-\alpha}}\}}$$

$$+ \sqrt{n} \Delta^n W 1_{\{\Delta^n X > \alpha \sqrt{\tilde{V}_j^n n^{-\alpha}}\}} \left( \sigma_{\frac{j-1}{n}} - \sigma_{\frac{(j-1)k_n}{n}} - \xi^n_{i,j}(3) \right) 1_{\{\Delta^n X \leq \alpha \sqrt{\tilde{V}_j^n n^{-\alpha}}\}}$$

$$\chi^n_{i,j}(2) = \left( \frac{\sqrt{\tilde{V}_j^n}}{\sigma_{\frac{(j-1)k_n}{n}}} - 1 - \xi^n_{i,j}(1) + \xi^n_{i,j}(2) \right) + \left( \xi^n_{i,j}(1) - \xi^n_{i,j}(2) - \tilde{\xi}^n_{i,j}(1) + \hat{\xi}^n_{i,j}(2) \right).$$
Finally, we denote
\[ \tilde{G}_n(\tau) = \frac{1}{[n/k_n]m_n} \sum_{j=1}^{[n/k_n]} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1 \left( \frac{\sqrt{n}}{\sigma_{(j-1)k_n}} \frac{\Delta_i^n X}{\sqrt{n}} | \leq \alpha \sqrt{V_{n-w}} \right) \]
\[ \leq \frac{\sqrt{V_n}}{\sigma_{(j-1)k_n}} - \chi_{ij}^n (1) - \tau \chi_{ij}^n (2) \]
\[ = \frac{1}{[n/k_n]m_n} \sum_{j=1}^{[n/k_n]} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1 \left( \sqrt{n} \Delta_i^n W \leq \tau + \tau \xi_{ij}^n (1) - \tau \xi_{ij}^n (2) - \xi_{ij}^n (3) \right). \]

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

### 9.4.1 The difference \( \tilde{F}_n(\tau) - \tilde{G}_n(\tau) \)

We first collect some preliminary results that we then make use of in analyzing \( \tilde{F}_n(\tau) - \tilde{G}_n(\tau) \). We start with \( \max_{i=1,\ldots,m} |\Delta_i^n B| \). We have
\[ \max_{i=1,\ldots,m} |\Delta_i^n B| \leq \max_{i=1,\ldots,m} \left| \int_{(i-1)\Delta_n}^{i\Delta_n} (\sigma_u - \sigma_{(i-1)\Delta_n}) dW_u \right| + \sup_{s \in [0,1]} |\sigma_s| \max_{i=1,\ldots,m} |\Delta_i^n W|. \]

Using the the results for the maxima of a Brownian motion, see e.g., Embrechts et al. (2001), as well as the fact that \( \sigma_t \) is uniformly bounded on \([0,1]\), we have
\[ \sup_{s \in [0,1]} |\sigma_s| \max_{i=1,\ldots,m} |\Delta_i^n W| = o_p(n^{-1/2+\epsilon}), \quad \forall \epsilon > 0. \]

Next, we have
\[ \max_{i=1,\ldots,m} \left| \int_{(i-1)\Delta_n}^{i\Delta_n} (\sigma_u - \sigma_{(i-1)\Delta_n}) dW_u \right| \leq \sup_{s,t \in [0,1]} \left| \int_0^t \sigma_u^n dW_u \right|, \]
where \( \sigma_u^n = \sigma_u - \sigma_{(i-1)\Delta_n} \) for \( s \in [(i-1)\Delta_n, i\Delta_n] \). From here, using Burkholder-Davis-Gundy inequality and the fact that \( \sigma_t \) is an Itô semimartingale we have
\[ \mathbb{E} \left| \int_0^t \sigma_u^n dW_u \right|^{2+\epsilon} \leq \frac{K}{n} |t-s|^{1+\epsilon/2}, \quad \forall 0 \leq s \leq t \leq 1, \quad \forall \epsilon > 0. \]

Therefore, the process \( n^{1/2-\epsilon} \int_0^t \sigma_u^n dW_u \) is C-tight (see e.g., Theorem 12.3 in Billingsley (1968)). Then applying Theorem VI.3.26 in Jacod and Shiryaev.
(2003), we get
\[ P \left( n^{1/2-\epsilon} \max_{i=1,\ldots,n} \left| \int_{(i-1)\Delta_n}^{i\Delta_n} (\sigma_u - \sigma_{(i-1)\Delta_n}) dW_u \right| > \zeta \right) \rightarrow 0, \ \forall \zeta > 0, \ \forall \epsilon > 0. \]
Combining these results, we get altogether
\[ \max_{i=1,\ldots,n} |\Delta \nu_i B| = o_p \left( n^{-1/2+\epsilon} \right), \ \forall \epsilon > 0. \]  
(9.14)

Next, using assumption SB (in particular that jumps are of finite activity), we have
\[ P \left( \sqrt{n} \left| \int_{\frac{j}{n}}^{\frac{j+1}{n}} 1 \left( \delta^\phi (z, x) \neq 0 \right) \mu(ds, dx) \geq 1 \right| \right) \leq K \frac{k_n}{n}, \ \phi = Y, \sigma, \tilde{\sigma} \text{ and } \tilde{\sigma}'. \]  
(9.15)

We now provide bounds for the elements of \( \chi^{\nu} (1) \) and \( \chi^{\nu} (2) \). In what follows we denote with \( \eta_n \) some deterministic sequence of positive numbers that depends only on \( n \). We first have (recall the definition of \( \tilde{\sigma}_t \))
\[ P \left( \sqrt{n} \left| \int_{\frac{j}{n}}^{\frac{j+1}{n}} (\sigma_u - \sigma_{\frac{j-1}{n}}) dW_u \right| \geq \eta_n \right) \leq \frac{\eta_n}{P \left( \Delta \nu_i W \left| \sigma_{\frac{j-1}{n}} - \sigma_{\frac{j-1}{n}} \xi^{\nu}_{i,j} (3) \right| \geq \eta_n \right) \leq K \left( \frac{k_n}{n} \right) \left( \frac{1}{n^{p/2} \eta_n^p} \right)^{1/2} \].
\( p > 0 \)  
(9.16)

Similar calculations (using the fact that \( \tilde{\sigma}_t \) and \( \tilde{\sigma}'_t \) are Itô semimartingales), yields for \( p > 0 \)
\[ P \left( \sqrt{n} \left| \int_{\frac{j}{n}}^{\frac{j+1}{n}} (\sigma_u - \sigma_{\frac{j-1}{n}}) dW_u \right| \geq \eta_n \right) \leq K \left( \frac{k_n}{n} \right) \left( \frac{1}{n^{p/2} \eta_n^p} \right)^{1/2} \].
\( p > 0 \)  
(9.17)
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

$$P(\sqrt{n}|\Delta^n_i Y| \geq \eta_n) \leq \frac{n^{p/2}E^{\left(\int_{(i-1)\Delta_n \int_E |\delta^i_Y(s,x)|\mu(ds, dx)\right)}^t}{\eta_n^p} \leq \frac{n^{p/2}E^{\left(\int_{(i-1)\Delta_n \int_E |\delta^i_Y(s,x)|\mu(ds, dx)\right)}^t}{\eta_n^p} \leq K n^{-1+\epsilon/2}\eta_n^{-\epsilon}, \forall \epsilon \in (0, 1).$$

(9.18)

Further, Chebyshev inequality and the boundedness of $a_t$ easily implies

$$P(\sqrt{n}|\Delta^n_i A| \geq \eta_n) \leq \frac{n^{p/2}|\Delta^n_i A|^p}{\eta_n^p} \leq K \frac{1}{n^{p/2}\eta_n^p}. \quad (9.19)$$

We turn next to the difference $\hat{V}_n^n - \check{V}_n^n$. Using triangular inequality, we have

$$P\left(\left|\hat{V}_n^n - \check{V}_n^n\right| \geq \eta_n\right) \leq P\left(2 \frac{n}{k_n} \pi \max_{i=1,\ldots,n} |\Delta^n_i A + \Delta^n_i B| \int_{(i-1)\Delta_n \int_E |\delta^i_Y(s,x)| \wedge 1)\mu(ds, dx) \geq \eta_n\right).$$

Taking into account that $k_n/\sqrt{n} \to 0$ from (5.9), as well as (9.14), we have

$$P\left(2 \frac{n}{k_n} \pi \max_{i=1,\ldots,n} |\Delta^n_i A + \Delta^n_i B| \int_{(i-1)\Delta_n \int_E |\delta^i_Y(s,x)| \wedge 1)\mu(ds, dx) \geq \eta_n\right) \leq K \left(\frac{n^{1/2+\epsilon}}{\eta_n k_n}\right)^t \mathbb{E}\left[(n^{1/2-\epsilon}) \max_{i=1,\ldots,n} |\Delta^n_i A + \Delta^n_i B| \wedge 1\right]^t \times \mathbb{E}\left[\int_{(i-1)\Delta_n \int_E |\delta^i_Y(s,x)| \wedge 1)\mu(ds, dx) \right]^{\epsilon} \leq K \left(\frac{n^{1/2+\epsilon}}{\eta_n k_n}\right)^t \mathbb{E}\left[(n^{1/2-\epsilon}) \max_{i=1,\ldots,n} |\Delta^n_i A + \Delta^n_i B| \wedge 1\right]^t \times \mathbb{E}\left[\int_{(i-1)\Delta_n \int_E |\delta^i_Y(s,x)| \wedge 1)\mu(ds, dx) \right]^{\epsilon} \leq K \left(\frac{n^{1/2+\epsilon}}{\eta_n k_n}\right)^t \frac{k_n}{n}, \forall \epsilon > 0.$$

Thus altogether we get

$$P\left(\left|\hat{V}_n^n - \check{V}_n^n\right| \geq \eta_n\right) \leq K \left(\frac{n^{1/2+\epsilon}}{\eta_n k_n}\right)^t \frac{k_n}{n}. \quad (9.20)$$
We continue next with the difference $\hat{V}_j^n - \tilde{V}_j^n$. Application of triangular inequality gives

$$|\Delta_{i-1}^n A + \Delta_{i-1}^n B| |\Delta_i^n A + \Delta_i^n B| - |\Delta_{i-1}^n B| |\Delta_i^n B|$$

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

Using this inequality and applying Chebyshev inequality, we get

$$\mathbb{P} \left( |\hat{V}_j^n - \tilde{V}_j^n| \geq \eta_n \right) \leq K \left( \frac{1}{\sqrt{n}\eta_n} \right)^p, \quad \forall p \geq 1,$$

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 (9.15), we can easily get

$$\mathbb{P} \left( |R_j^{(1)}| \geq \eta_n \right) \leq \mathbb{P} \left( |R_j^{(1)}| \geq \eta_n, \int_{(j-1)k_n}^{jk_n} \int_E 1(\delta^\sigma(s, x) \neq 0) \mu(ds, dx) \geq 1 \right)$$

$$+ \mathbb{P} \left( |R_j^{(1)}| \geq \eta_n, \int_{(j-1)k_n}^{jk_n} \int_E 1(\delta^\sigma(s, x) \neq 0) \mu(ds, dx) = 0 \right)$$

$$\leq K \frac{k_n}{n} + K \left( \frac{1}{\sqrt{n}\eta_n} \right)^p + K \left( \frac{k_n}{m\eta_n} \right)^p, \quad \forall p \geq 1.$$ 

Similar calculations, and utilizing the fact that $\tilde{\sigma}, \tilde{\sigma}'$ are themselves Itô semi-martingales, yield

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

Next, by splitting

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

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

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

Finally, we trivially have

$$\left\{ \begin{array}{ll}
\mathbb{P} \left( |R_j^{(4)} - R_j^{(4)}(i)| \geq \eta_n \right) \leq K \left( \frac{1}{\sqrt{n}\eta_n} \right)^p, \\
\mathbb{E}|R_j^{(4)} - R_j^{(4)}(i)|^p \leq K \left( \frac{1}{\sqrt{n}} \right)^p, \quad \forall p > 0,
\end{array} \right.$$ 

(9.24)
the bounds in (9.20) and (9.21), we have
\[ \begin{align*}
\mathbb{P}(\mid \tilde{V}_j^n - \sigma \tilde{V}_j^n \mid \geq \eta_n) & \leq K \left( \frac{1}{k_n \eta_n} \right)^p, \\
\mathbb{E}[\mid \tilde{V}_j^n - \sigma \tilde{V}_j^n \mid] & \leq K \left( \frac{1}{k_n \eta_n} \right)^p, \quad \forall p > 0,
\end{align*} \tag{9.25} \]

Further, application of Burkholder-Davis-Gundy inequality gives
\[ \begin{align*}
\mathbb{P}(\mid \tilde{V}_j^n - \sigma \tilde{V}_j^n \mid \geq \eta_n) & \leq K \left( \frac{1}{k_n \eta_n^2} \right)^p, \\
\mathbb{P}(\mid \tilde{V}_j^n - \tilde{V}_j^n \mid \geq 0.5 \sigma \tilde{V}_j^n \mid \geq \eta_n) & \leq K \frac{k_n}{\eta_n}, \quad \forall p \geq 2.
\end{align*} \tag{9.26} \]

Now we can use the above results for the components of \( \tilde{V}_j^n - \sigma \tilde{V}_j^n \), to analyze the first term in \( \chi_{i,j}^n(2) \) involving \( \sqrt{\tilde{V}_j^n - \sigma \tilde{V}_j^n} \). 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 \( \tilde{V}_j^n \) and \( \sigma \tilde{V}_j^n \), respectively, as well the bounds in (9.20), (9.21) and (9.26), we get
\[ \begin{align*}
\mathbb{P} \left( \left| \frac{\tilde{V}_j^n}{\sigma \tilde{V}_j^n} - 1 - \xi_{i,j}^n(1) + \xi_{i,j}^n(2) \right| \geq \eta_n \right) \\
& \leq K \left[ \left( \frac{n^{1/2+\epsilon}}{\eta_n^{1/2+\epsilon} k_n} \right)^t \frac{k_n}{n} \sqrt{\frac{k_n}{n} \eta_n^{p/2} \wedge (n/k_n)^{p/2}} \right] \left( \frac{1}{\eta_n^{2p/3} k_n^{2/3}} \right) \left( \frac{1}{\eta_n^{p/3} k_n^{2/3}} \right), \tag{9.28} \end{align*} \]
for every \( p > 0 \). Similarly, using the following inequality
\[ \mathbb{P}(\mid x^2 - y^2 \mid \geq \epsilon) \leq \mathbb{P}(\mid x - y \mid \geq 0.5\epsilon) + \mathbb{P}(2\mid y \mid \geq K) + \mathbb{P}(\mid x - y \mid \geq 0.5\epsilon/K), \]
for any random variables \( x \) and \( y \) and constants \( \epsilon > 0 \) and \( K > 0 \), together with the bounds in (9.20) and (9.21), we have
\[ \begin{align*}
\mathbb{P} \left( \left| \xi_{i,j}^n(1) - \xi_{i,j}^n(1) - \tilde{\xi}_{i,j}^n(1) + \tilde{\xi}_{i,j}^n(2) \right| \geq \eta_n \right) \\
& \leq K \left[ \left( \frac{n^{1/2+\epsilon}}{\eta_n^{1/2+\epsilon} k_n} \right)^t \frac{k_n}{n} \sqrt{\frac{k_n}{n} \eta_n^{p/2} \wedge (n/k_n)^{p/2}} \right] \left( \frac{1}{\eta_n^{2p/3} k_n^{2/3}} \right), \tag{9.29} \end{align*} \]
for every \( p > 1 \) and arbitrary small \( \epsilon > 0 \).
We finally provide a bound for the last term in \( \chi_{i,j}^n(1) \). We can use Chebyshev inequality as well as Hölder inequality to get

\[
\mathbb{P} \left( \sqrt{n} \sigma \left( \frac{1}{n} \sum_{i,j} \Delta_n^\text{n} X \right) \geq 0.5 \sigma \left( \frac{1}{n} \sum_{i,j} \Delta_n^\text{n} X \right) \right) \leq K \left( \frac{k_n}{n} \right)^{1/(1+\text{i})} \frac{1}{\eta_n^\text{i}}, \quad \forall \text{i} > 0.
\]

We can further write

\[
\mathbb{P} \left( \left| \Delta_n^\text{n} X \right| > 0.5 \sigma \left( \frac{1}{n} \sum_{i,j} \Delta_n^\text{n} X \right) \right) \leq \mathbb{P} \left( \left| \Delta_n^\text{n} X \right| > 0.5 \sigma \left( \frac{1}{n} \sum_{i,j} \Delta_n^\text{n} X \right) \right).
\]

From here we can use the bounds in (9.20), (9.21) and (9.26) as well as (9.18) and conclude

\[
\mathbb{P} \left( \left| \Delta_n^\text{n} X \right| > 0.5 \sigma \left( \frac{1}{n} \sum_{i,j} \Delta_n^\text{n} X \right) \right) \leq \mathbb{P} \left( \left| \Delta_n^\text{n} X \right| > 0.5 \sigma \left( \frac{1}{n} \sum_{i,j} \Delta_n^\text{n} X \right) \right) \leq K \left( \frac{k_n}{n} \right)^{1/(1+\text{i})} \frac{1}{\eta_n^\text{i}}, \quad \forall \text{i} > 0.
\]

Combining the results in (9.16), (9.17), (9.18), (9.19), (9.28), (9.29) and (9.31), we get

\[
\mathbb{P} \left( \left| \chi_{i,j}^n(1) \right| + \left| \chi_{i,j}^n(2) \right| > \eta_n \right) \leq K \left( \frac{n^{1/2+\text{i}}}{\eta_n} \right)^{1/2} \frac{k_n}{n} \sqrt{\frac{k_n}{n}} \sqrt{\frac{1}{\eta_n^p}} \sqrt{\frac{1}{(n/k_n)^p}} \sqrt{\frac{1}{\eta_n^p}} \sqrt{\frac{1}{\eta_n^p}}.
\]

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

\[
\mathbb{E} \left[ \left( \sqrt{n} \Delta_n^\text{n} X \right)^1 \left| \Delta_n^\text{n} X \right| \leq \alpha \sqrt{\frac{1}{n}} \right] \leq K \left( \frac{n^{1/2+\text{i}}}{\eta_n} \right)^{1/2} \frac{k_n}{n} \sqrt{\frac{k_n}{n}} \sqrt{\frac{1}{\eta_n^p}} \sqrt{\frac{1}{(n/k_n)^p}} \sqrt{\frac{1}{\eta_n^p}} \sqrt{\frac{1}{\eta_n^p}}.
\]
Therefore, upon picking $\eta_n \propto n^{-q-i}$ for $i \in (0, 1/2 - q)$, we get finally for any compact subset $\mathcal{A}$ of $(-\infty, 0)$

\[
\sup_{\tau \in \mathcal{A}} |\hat{F}_n(\tau) - \hat{G}_n(\tau)| = o_p\left(\frac{1}{k_n}\right). \tag{9.32}
\]

### 9.4.2 The asymptotic behavior of $\hat{G}_n(\tau) - \Phi(\tau)$

We have

\[
\hat{G}_n(\tau) - \Phi(\tau) = \sum_{i=1}^{5} A_n^i, \quad A_1^n = \frac{1}{[n/k_n]m_n} \sum_{j=1}^{[n/k_n]} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} \left[1 \left(\sqrt{n} \Delta_i^n W \leq \tau\right) - \Phi(\tau)\right],
\]

\[
A_2^n = \frac{1}{[n/k_n]m_n} \sum_{j=1}^{[n/k_n]} \left(\Phi \left(\tau + \tau \tilde{\xi}_{i,j}^n(1) - \tau \tilde{\xi}_{i,j}^n(2)\right) - \Phi(\tau)\right), \quad A_3^n = \frac{1}{[n/k_n]m_n} \sum_{j=1}^{[n/k_n]} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} a_i^n,
\]

\[
a_i^n = 1 \left(\sqrt{n} \Delta_i^n W \leq \tau + \tau \tilde{\xi}_{i,j}^n(1) - \tau \tilde{\xi}_{i,j}^n(2)\right) - 1 \left(\sqrt{n} \Delta_i^n W \leq \tau\right) + \Phi(\tau) - \Phi \left(\tau + \tau \tilde{\xi}_{i,j}^n(1) - \tau \tilde{\xi}_{i,j}^n(2)\right),
\]

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

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

We first derive a bound for the order of magnitude of $A_2^n$, $A_3^n$ and $A_5^n$ and then analyze the limiting behavior of $A_1^n$ and $A_4^n$. Using the independence of $\Delta_i^n W$, $\Delta_i^n W$, $\Delta_i^n W'$, $\Delta_i^n W''$ from each other and $\mathcal{F}_{j-1}^{(k_n)}$, the fact that $\xi_{i,j}^n(4)$ is adapted to $\mathcal{F}_{j-1}^{(k_n)}$ as well as successive conditioning, we have for $i \neq h$ and $i, h = (j-1)k_n+1, \ldots, (j-1)k_n+m_n$

\[
E (a_i^n a_h^n) = 0, \quad E (a_i^n)^2 \leq K^2 \sqrt{\frac{k_n}{n}}, \quad i \neq h, \quad i, h = 1, \ldots, n.
\]

Therefore, we have

\[
A_3^n = \sqrt{\tau} \times O_p \left(\frac{1}{\sqrt{[n/k_n]m_n}} \sqrt{\frac{k_n}{n}}\right). \tag{9.34}
\]

For $A_1^n$, using a second-order Taylor expansion, the bounds in (9.24), (9.25) and (9.27), as well as the uniform boundedness of the probability density of the
standard normal distribution and its derivative, we get

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

(9.35)

Next, for $A_n^2$, 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

$$
\Phi\left(\frac{\tau + \tau \xi_{i,j}^n(1) - \tau \xi_{i,j}^n(2)}{\xi_{i,j}^n(4)}\right) - \Phi\left(\tau + \tau \xi_{i,j}^n(1) - \tau \xi_{i,j}^n(2)\right) = b^n_1(1) + b^n_2(2) + b^n_3(3),
$$

$$
b^n_1(1) = \left\{ \Phi\left(\frac{\tau + \tau \xi_{i,j}^n(1) - \tau \xi_{i,j}^n(2)}{\xi_{i,j}^n(4)}\right) - \Phi\left(\tau + \tau \xi_{i,j}^n(1) - \tau \xi_{i,j}^n(2)\right) \right\} 1_{\{\xi_{i,j}^n(4) - 1 \geq (\frac{\xi_{i,j}^n}{\sqrt{n}})^{1/2} - \iota\}},
$$

$$
b^n_2(2) = \Phi'(\tau + \tau \xi_{i,j}^n(1) - \tau \xi_{i,j}^n(2)) (\tau + \tau \xi_{i,j}^n(1) - \tau \xi_{i,j}^n(2)) (\xi_{i,j}^n(4) - 1) \{\xi_{i,j}^n(4) - 1 < (\frac{\xi_{i,j}^n}{\sqrt{n}})^{1/2} - \iota\},
$$

$$
|b^n_1(3)| \leq K \frac{|\tau + \tau \xi_{i,j}^n(1) - \tau \xi_{i,j}^n(2)|^2}{(1 - (k_n/n)^{1/2} - \iota)^2} |\xi_{i,j}^n(4) - 1|^2.
$$

For $b^n_1(1)$ and $b^n_3(3)$, we have

$$
\mathbb{E}(|b^n_1(1)| + |b^n_3(3)|) \leq K \frac{k_n}{n}.
$$

For $b^n_2(2)$, by an application of Hölder inequality, we first have

$$
\mathbb{E}\left|b^n_2(2) - \Phi'(\tau) \tau (\xi_{i,j}^n(4) - 1) 1_{\{\xi_{i,j}^n(4) - 1 < (\frac{\xi_{i,j}^n}{\sqrt{n}})^{1/2} - \iota\}}\right| \leq K \tau \frac{1}{\sqrt{n}}.
$$

Then,

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

Therefore, altogether we get

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

(9.36)

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

$$
A_2^n(1) = \frac{1}{[n/k_n]} \sum_{j=1}^{[n/k_n]} \Phi'(\tau) \tau \xi_j^n(1),
$$

$$
A_2^n(2) = \frac{1}{[n/k_n]} \sum_{j=1}^{[n/k_n]} (0.5 \Phi''(\tau) \tau^2 (\xi_j^n(1))^2 - \Phi'(\tau) \tau \xi_j^n(2)).
$$
With this notation, using the bounds in (9.27), as well as the boundedness of \( \Phi'' \), we have
\[
E|A_n^1 - A_1^1(1) - A_1^2(2)| \leq K \tau^3 \left[ \left( \frac{k_n}{n} \right)^{3/2} \sqrt{\left( \frac{1}{k_n} \right)^{3/2}} \right]. \tag{9.37}
\]

Further, upon denoting with \( \hat{A}_2^n(1) \) and \( \hat{A}_2^n(2) \) the counterparts of \( A_2^n(1) \) and \( A_2^n(2) \) with \( \tilde{\xi}_{ij}^n(1) \) and \( \tilde{\xi}_{ij}^n(2) \) replaced with \( \hat{\xi}_{ij}^n(1) \) and \( \hat{\xi}_{ij}^n(2) \) respectively, we have using the bounds in (9.27) (as well as the restriction on the rate of growth of \( k_n \) in (5.9))
\[
E|A_n^1(1) + A_2^n(2) - \hat{A}_2^n(1) - \hat{A}_2^n(2)| \leq K(\tau \vee \tau^2) \left( \frac{1}{\sqrt{n}} \sqrt{\frac{k_n}{n}} \right). \tag{9.38}
\]

Thus we are left with the terms \( A_1^n, \hat{A}_2^n(1) \) and \( \hat{A}_2^n(2) \). For \( \hat{A}_2^n(2) \), using
\[
E_n^\mathbb{C} \left( \hat{\xi}_{ij}^n(1) \right)^2 = 2E_n^\mathbb{C} \left( \hat{\xi}_{ij}^n(2) \right) = \frac{1}{4} k_n \left( \frac{\pi}{2} \right)^2 + \pi - 3,
\]
we have
\[
k_n \hat{A}_2^n(2) \xrightarrow{p} \frac{\tau^2 \Phi''(\tau) - \tau \Phi'(\tau)}{8} \left( \left( \frac{\pi}{2} \right)^2 + \pi - 3 \right), \tag{9.39}
\]
locally uniformly in \( \tau \). We finally will show that
\[
\left( \sqrt{\frac{n}{k_n}} m_n A_1^n \sqrt{\frac{n}{k_n}} k_n A_2^n(1) \right) \quad \xrightarrow{d} \quad \left( Z_1(\tau) \right. \left. Z_2(\tau) \right), \tag{9.40}
\]
locally uniformly in \( \tau \). We have
\[
\left( \sqrt{\frac{n}{k_n}} m_n A_1^n \sqrt{\frac{n}{k_n}} k_n A_2^n(1) \right) = \sum_{i=1}^{[n/k_n]} \left( \frac{\Phi'(\tau)}{2} \left( \frac{\zeta_i^n(1)}{\zeta_i^n(2)} + \zeta_i^n(3) \right) \right) + \left( \frac{0}{2} \frac{\zeta_i^n}{\tau} \right),
\]
with
\[
\zeta_i^n = \begin{cases} 
\frac{1}{\sqrt{[n/k_n] m_n}} \left[ 1 \left( \sqrt{n} \Delta_i^n \tau \leq \tau \right) \right] & \text{for } i \in \mathbb{I}^n, \\
\frac{1}{\sqrt{[n/k_n] k_n}} \left( \frac{1}{2} \sqrt{n} \Delta_i^n W \right) & \text{for } i \in \mathbb{I}^n.
\end{cases}
\]
where \( \mathbb{I}^n = \{ i = (j-1)k_n + 1, \ldots, (j-1)k_n + m_n, \quad j = 1, \ldots, [n/k_n] \} \), and for \( i, j = 1, \ldots, n \setminus \mathbb{I}^n, \zeta_i^n \) is exactly as above with only the first element being
replaced with zero, and finally
\[
\bar{\zeta}_n = -\left(\frac{\pi}{2}\right) \frac{\sqrt{n}}{\sqrt{n/k_n}} \sum_{j=1}^{\lfloor n/k_n \rfloor} \left| \sqrt{n}\Delta_{(j-1)k_n} W \right| \left( \left| \sqrt{n}\Delta_{(j-1)k_n+1} W \right| - \frac{\sqrt{2\pi}}{2} \right) \\
+ \sqrt{\frac{2\pi}{2}} \left( \left| \sqrt{n}\Delta_{jk_n} W \right| - \frac{\sqrt{2\pi}}{2} \right),
\]
where we set \(\Delta_0 W = 0\). With this notation, we have
\[
\mathbb{E}(\bar{\zeta}_n)^2 \leq \frac{K}{k_n}.
\]
Further,
\[
\mathbb{E}_{i=1}^{n} (\zeta_n^2) = 0,
\]
\[
\mathbb{E}_{i=1}^{n} ||\zeta_n||^2 \to 0, \quad \forall \tau > 0,
\]
\[
\sum_{i=1}^{\lfloor n/k_n \rfloor} \mathbb{E}_{i=1}^{n} \left[ \zeta_n^2 (\zeta_n')^2 \right] \to \left( \Phi(\tau) (1-\Phi(\tau)) \right)
\]
\[
\left( \begin{array}{ccc}
0 & \frac{1}{2} (1 - \frac{\pi}{2}) & 0 \\
0 & \frac{1}{2} (1 - \frac{\pi}{2}) & \frac{1}{2} (1 - \frac{\pi}{2}) \\
0 & \frac{1}{2} (1 - \frac{\pi}{2}) & \frac{1}{2} (1 - \frac{\pi}{2})
\end{array} \right).
\]
Combining the last two results, we have the convergence in (9.40), pointwise in \(\tau\), by an application of Theorem IX.7.28 in Jacod and Shiryaev (2003). Application of Theorem 12.3 in Billingsley (1968), extends the convergence to local uniform in \(\tau\).

Altogether, the limit behavior of \(\hat{G}_n(\tau) - \Phi(\tau)\) is completely characterized by the limits in (9.39)-(9.40) and
\[
\sup_{\tau \in \mathcal{A}} |\hat{G}_n(\tau) - \Phi(\tau) - A_1^n - \tilde{A}_2^n| = o_p \left( \frac{1}{k_n} \right), \quad (9.41)
\]
where \(\mathcal{A}\) is a compact subset of \((-\infty,0)\), with the result in (9.41) following from the bounds on the order of magnitude derived above.

9.4.3 The difference \(\hat{F}_n(\tau) - \tilde{F}_n(\tau)\)

To analyze the difference \(\hat{F}_n(\tau) - \tilde{F}_n(\tau)\), we use the following inequality
\[
\mathbb{P} \left( |\Delta^n_\tau X| > \alpha \sqrt{\frac{1}{\sqrt{j} n^{-\alpha}}} \right)
\leq \mathbb{P} \left( \left| \frac{\sqrt{V_j^n}}{\sigma_{(j-1)k_n}} - 1 \right| > 0.5 \right) + \mathbb{P} \left( |\Delta^n_\tau X| > 0.5 \alpha \sigma_{(j-1)k_n} n^{-\alpha} \right).
\]
For the first probability on the right-hand side of the above inequality we can use the bounds in (9.26), (9.27) and (9.28), while for the second one we can use the exponential inequality for continuous martingales with bounded variation, see e.g., Revuz and Yor (1999), as well as the algebraic inequality $|\sum_i a_i|^p \leq \sum_i |a_i|^p$ for $p \in (0, 1]$, to conclude

$$
P\left(|\Delta_i^n X| > \alpha \sqrt{V_j^n} n^{-\infty}\right) \leq K \left[\frac{k_n}{n} \sqrt{n^{-1+\varepsilon}}\right], \quad \forall \varepsilon > 0. \quad (9.42)$$

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

$$
\sup_{\tau \in A} |\hat{F}^n(\tau) - \hat{F}^n(\tau)| = o_p\left(\frac{1}{k_n}\right), \quad (9.43)
$$

for any compact subset $A$ of $(-\infty, 0)$. \hfill \Box

References


