VOLATILITY OCCUPATION TIMES

BY JIA LI1, VIKTOR TODOROV2 AND GEORGE TAUCHEN

Duke University, Northwestern University and Duke University

We propose nonparametric estimators of the occupation measure and the occupation density of the diffusion coefficient (stochastic volatility) of a discretely observed Itô semimartingale on a fixed interval when the mesh of the observation grid shrinks to zero asymptotically. In a first step we estimate the volatility locally over blocks of shrinking length, and then in a second step we use these estimates to construct a sample analogue of the volatility occupation time and a kernel-based estimator of its density. We prove the consistency of our estimators and further derive bounds for their rates of convergence. We use these results to estimate nonparametrically the quantiles associated with the volatility occupation measure.

1. Introduction. Continuous-time Itô semimartingales are used widely to model stochastic processes in various areas such as finance. The general Itô semimartingale process is given by

\[
X_t = X_0 + \int_0^t b_s \, ds + \int_0^t \sigma_s \, dW_s + J_t,
\]

where \(b_t\) and \(\sigma_t\) are processes with càdlàg paths, \(W_t\) is a Brownian motion and \(J_t\) is a jump process; formal conditions are given in the next section. Inference for model (1.1) in the general case (either in a parametric or a nonparametric context) is quite complicated because of the many “layers of latency,” for example, as typical in financial applications, \(\sigma_t\) and \(J_t\) can have randomness not captured by \(X_t\).

When \(X\) is sampled at discrete times but with the mesh of the observation grid shrinking to zero, that is, high-frequency data of \(X\) are available, the distinct pathwise behavior of the components in (1.1) can be used to nonparametrically separate them. Indeed, various techniques have been already proposed to estimate nonparametrically the integrated variance \(\int_0^T \sigma_s^2 \, ds\) over a specific interval \([0, T]\) (see, e.g., [3] and [13]), and more generally integrated variance measures of the form \(\int_0^T g(\sigma_s^2) \, ds\), where \(g(\cdot)\) is a continuous function with polynomial growth.
[and there are more smoothness requirements on \(g(\cdot)\) for determining the rate of convergence]; see Theorems 3.4.1 and 9.4.1 in [10] and the recent work of [12].

This paper extends the existing literature on high-frequency nonparametric volatility estimation by developing a nonparametric jump-robust estimator of the occupation time of the latent volatility process \((V_t)_{t \geq 0} \equiv (\sigma_t^2)_{t \geq 0}\) where the volatility occupation time is defined by

\[
F_t(x) = \int_0^t 1_{\{V_s \leq x\}} \, ds \quad \forall x > 0, \ t \in [0, T].
\]

Evidently, the right-hand side of (1.2) is of the form \(\int_0^t g(V_s) \, ds\) with \(g(v) = 1_{\{v \leq x\}}\), which unlike earlier work is a discontinuous function.

If \(F_t(\cdot)\) is absolutely continuous with respect to the Lebesgue measure, its derivative \(f_t(\cdot)\), that is, the volatility occupation density, is well-defined. By the Lebesgue differentiation theorem, the occupation density can be equivalently defined as

\[
f_t(x) = \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} (F_t(x + \varepsilon) - F_t(x - \varepsilon)).
\]

In addition to estimating \(F_t(x)\), in this paper we also develop a consistent estimator for the volatility occupation density \(f_t(x)\) using the high-frequency record of \(X\).

The occupation measure of the volatility process “summarizes” in a convenient way the information regarding the volatility behavior over the given time interval. Indeed, for any bounded (or nonnegative) Borel function \(g(\cdot)\) (see, e.g., Theorem 6.4 of [8]), we have

\[
\int_0^t g(V_s) \, ds = \int_{\mathbb{R}^+} g(x) f_t(x) \, dx = \int_{\mathbb{R}^+} g(x) \, dF_t(x).
\]

Thus, the occupation time and its density can be considered as the pathwise analogues of the cumulative distribution function and its density.

Our interest in occupation times stems from the fact that they are natural measures of risk, particularly in nonstationary settings where invariant distributions do not exist; see, for example, the discussion in [1]. Indeed, there has been a significant interest (both theoretically and in practice) in pricing options based on the occupation times of an underlying asset; see for example, [4] and [19] and references therein. Here, we show how to measure nonparametrically occupation times associated with the volatility risk of the price process. As a by-product, we also estimate the corresponding quantiles of the actual path of the volatility process over the fixed time interval. Since the pathwise volatility quantiles are preserved under monotone transformations, they provide a convenient way of studying the variability of the volatility and the relationship of the latter with the volatility process itself.

We summarize our estimation procedure as follows. We first split the fixed time interval into blocks of decreasing length and form local estimates of the unobserved stochastic variance over each of the blocks. The volatility estimates over
the blocks are truncated variations (see, e.g., [13] and [10]), and we further allow for adaptive choice of the truncation level that makes use of some preliminary estimates of the stochastic variance. Then, our estimator of the volatility occupation time is simply the empirical cumulative distribution function of the local volatility estimates over the blocks. Analogously, we estimate the volatility occupation density from the local volatility estimates using kernel smoothing.

Our estimation problem can be compared with the recent work of [12]. Jacod and Rosenbaum [12] show that an estimator of \( \int_0^T g(\sigma_s^2) \, ds \), for \( g(\cdot) \) a \( C^3 \) function, formed by plugging in local variance estimates formed over blocks of decreasing length, can achieve the efficient \( \Delta_n^{-1/2} \) rate of convergence (for \( \Delta_n \) being the length of the high-frequency intervals). Similar to [12], our estimator here is formed by plugging local variance estimates in our function of interest.

The main difference between the current work and [12] is that in our case the function \( g(\cdot) \) in (1.4) is discontinuous. As a result, the precision of estimating the volatility occupation time depends on the uniform rate of recovering the volatility process outside of the times of the “big” volatility jumps (with the size of the “big” jumps shrinking asymptotically to zero). Therefore, in the basic case when \( X \) and \( V \) are continuous, the rate of convergence of the volatility occupation time estimator is (almost) \( \Delta_n^{-1/4} \) which, as we show in the paper, is the optimal uniform rate for recovering the volatility trajectory from high-frequency observations. By contrast, [12] derive a central limit theorem for the convergence of their estimator to \( \int_0^T g(\sigma_s^2) \, ds \) by making use of the assumed smoothness of \( g \) and applying second-order Taylor expansion of the function \( g \) evaluated at the local volatility estimator in their bias-correction and asymptotic negligibility arguments.

Finally, our inference for the volatility occupation time and its density can be compared with the estimation of occupation time and density of a recurrent Markov diffusion process from discrete observations of the process; see, for example, [7] and [1]. The main difference is that here the state vector, and therefore the stochastic volatility, is not fully observed. Hence, we first need to recover non-parametrically the unobserved volatility trajectory, and the error associated with recovering the volatility trajectory determines the asymptotic behavior of our estimators.

The paper is organized as follows. In Section 2 we introduce the formal setup and state our assumptions. In Section 3 we develop our estimator of the volatility occupation measure and prove its consistency. In Section 4 we derive bounds for the rate of convergence of the volatility occupation time estimator. Section 5 derives a consistent estimator for the volatility occupation density. Section 6 reports results from a Monte Carlo study of our estimation technique. Section 7 concludes. Section 8 contains all proofs.

2. Setup and assumptions. We start with introducing the formal setup and stating our assumptions about \( X \). The process \( X \) in (1.1) is defined on a filtered
space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})\) with \(b_t\) and \(\sigma_t\) being adapted to the filtration. Further, the jump component \(J_t\) is defined as
\[
J_t = \int_0^t \int_{\mathbb{R}} \delta(s, z) 1_{\{|\delta(s, z)| \leq 1\}}(\mu - \nu)(ds, dz)
+ \int_0^t \int_{\mathbb{R}} \delta(s, z) 1_{\{|\delta(s, z)| > 1\}}\mu(ds, dz),
\]
where \(\mu\) is a Poisson measure on \(\mathbb{R}_+ \times \mathbb{R}\) with compensator \(\nu\) of the form
\[
\nu(dt, dz) = dt \otimes \lambda(dz)
\]
for some \(\sigma\)-finite measure \(\lambda\) on \(\mathbb{R}\) and \(\delta: \Omega \times \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}\) is a predictable function. Regularity conditions on \(X_t\) are collected below.

**Assumption A.** Let \(r \in [0, 2]\) be a constant. The process \(X_t\) is an Itô semimartingale given by (1.1) and (2.1), with \(b_t\) locally bounded and \(\sigma_t\) càdlàg. Moreover \(|\delta(\omega, t, z)| \wedge 1 \leq \Gamma_m(z)\) for all \((\omega, t, z)\) with \(t \leq \tau_m(\omega)\), where \((\tau_m)\) is a localizing sequence of stopping times, and each function \(\Gamma_m\) on \(\mathbb{R}\) satisfies \(\int_{\mathbb{R}} \Gamma_m(z)r\lambda(dz) < \infty\).

Assumption A can be viewed as a regularity-type condition. The coefficient \(r\) in Assumption A controls the degree of activity of the jump component \(J\) and will play an important role in the rate of convergence of our estimator. We note that \(r\) provides an upper bound for the (generalized) Blumenthal–Getoor index of \(J_t\); see, for example, Lemma 3.2.1 in [10]. We next state our assumption for the volatility occupation time.

**Assumption B.** Fix \(x \geq 0\). We have \(F_T(\cdot)\) a.s. differentiable with derivative \(f_T(\cdot)\) in a neighborhood \(\mathcal{N}_x\) containing \(x\). Moreover, \(\sup_{z \in \mathcal{N}_x} \mathbb{E}[f_T(z)] < \infty\).

Assumption B is mainly concerned with the pathwise smoothness of the occupation time. The differentiability condition amounts to the existence of occupation density \(f_T(\cdot)\), which is not a strong requirement; see, for example, [8] and [14]. The condition \(\sup_{z \in \mathcal{N}_x} \mathbb{E}[f_T(z)] < \infty\) only requires the temporal average (over \([0, T]\)) of the probability density of \(V_t\) uniformly bounded in the neighborhood \(\mathcal{N}_x\), which is satisfied by most stochastic volatility models. This condition is of course much weaker than requiring \(\mathbb{E}[\sup_{z \in \mathcal{N}_x} f_T(z)] < \infty\), as the latter would demand more on the pathwise regularity of the occupation density.

**Example.** Let \(V_t\) solve the following stochastic differential equation:
\[
dV_t = a_t^V dt + s(V_t) dW_t^V + dJ_t^V,
\]
where \(a_t^V\) is a locally bounded process, \(W_t^V\) is a Brownian motion, \(J_t^V\) is a finite-variational jump process and \(s(\cdot)\) has twice continuously differentiable reciprocal. Assume further that \(V\) has an invariant distribution which is \(C^1\) in a neighborhood
of \( x \). Then Assumption B holds. This follows from an application of Itô’s formula and Theorem IV.75 in [14]. This example includes many parametric models of interest like the square-root diffusion model and the more general constant elasticity of variance model.

For some of the results we will need a stronger condition on the volatility occupation density, mainly its continuity which we state formally in the next assumption.

**Assumption B’**. \( F_T (\cdot) \) is a.s. continuously differentiable on \( \mathbb{R} \) with derivative \( f_T (\cdot) \).

Assumption B’ is harder to verify than Assumption B. Necessary and sufficient conditions for the continuity of the occupation density (local time) of a Borel right Markov process are discussed in [6].

We finally state a slightly stronger condition on the volatility process that we will need for deriving the rate of convergence of our estimator.

**Assumption C**. The process \( \sigma_t \) is an Itô semimartingale with the form

\[
\sigma_t = \sigma_0 + \int_0^t \tilde{b}_s \, ds + \int_0^t \tilde{\sigma}_s \, dW_s + \int_0^t \tilde{\sigma}_s' \, dW'_s + \int_0^t \int_{\mathbb{R}} \tilde{\delta}(s, z)(\mu - \nu)(ds, dz),
\]

where the processes \( \tilde{b}, \tilde{\sigma}, \tilde{\sigma}' \) are locally bounded and adapted, \( W' \) is a Brownian motion orthogonal to \( W \), and \( \tilde{\delta}(\cdot) \) is a predictable function. Moreover \(|\tilde{\delta}(\omega, t, z)| \wedge 1 \leq \tilde{\Gamma}_m(z)\) for all \((\omega, t, z)\) with \( t \leq \tau_m(\omega) \), where \( (\tau_m) \) is a localizing sequence of stopping times, and for some \( \tilde{r} \in (0, 2] \), each function \( \tilde{\Gamma}_m \) on \( \mathbb{R} \) satisfies \( \int_{\mathbb{R}} \tilde{\Gamma}_m(z)^{\tilde{r}} \lambda(dz) < \infty \).

Assumption C assumes that \( \sigma_t \) is an Itô semimartingale, an assumption that is satisfied by most stochastic volatility models. We impose no restriction on the activity of the volatility jumps as well as the dependence between \( \sigma_t \) and \( X_t \), a generality that is important in practical applications (particularly in finance).

**3. The estimator and its consistency.** We next introduce our estimator of the volatility occupation time and derive its consistency. We suppose that the process \( X_t \) is observed at discrete times \( i \Delta_n, i = 0, 1, \ldots, \) on \([0, T]\) for a fixed \( T > 0 \) with the time lag \( \Delta_n \to 0 \) when \( n \to \infty \). The assumption for equidistant observations is merely for simplicity, and the theoretical results that follow (except Theorem 3) will continue to hold in the case of irregular (but nonrandom) sampling with \( \Delta_n \) replaced by the mesh of the irregular observation grid. In what follows the high-frequency increment of any process \( Y \) is denoted as \( \Delta_t^h Y = Y_{i \Delta_n} - Y_{(i-1) \Delta_n} \).

Our strategy of estimating \( F_T (\cdot) \) is to first form an approximation of the volatility trajectory and then use the latter to form a sample analogue of \( F_T (\cdot) \). To recover
the volatility trajectory we construct local approximations for the spot variance process $V$ over blocks of shrinking length. To this end, let $k_n$ be a sequence of integers with $k_n \to \infty$ and $k_n \Delta_n \to 0$. Henceforth we use the shorthand notation $u_n = k_n \Delta_n$. We also set a truncation process $v_{n,t}$ verifying the following assumption, which is maintained throughout the paper without further mention.

**ASSUMPTION D.** We have $v_{n,t} = \alpha_{n,t} \Delta_n^{\sigma}$, where $\sigma \in (0, 1/2)$ is constant, and $\alpha_{n,t}$ is a strictly positive real-valued process such that for some localizing sequence of stopping times $(\tau_m)$, $(\sup_{t \in [0,T]} (\alpha_{n,t \wedge \tau_m} \vee \alpha_{n,t \land \tau_m}))_{n \geq 1}$ is tight for each $m \geq 1$.

With this notation, for each $i = 0, \ldots, \lfloor T/\Delta_n \rfloor - k_n$, we set

$$
\begin{align*}
\hat{V}_{i,\Delta_n}^* &= \frac{1}{u_n} \sum_{j=1}^{k_n} (\Delta_{i+j}^n)^2, \\
\hat{V}_{i,\Delta_n} &= \frac{1}{u_n} \sum_{j=1}^{k_n} (\Delta_{i+j}^n)^2 1_{\{|\Delta_{i+j}^n| \leq v_{n,i,\Delta_n}\}}.
\end{align*}
$$

(3.1)

Here, $\hat{V}_{i,\Delta_n}^*$ is a local approximation of $V_{i,\Delta_n}$ when $X$ is continuous, while $\hat{V}_{i,\Delta_n}$ serves the same purpose but is robust to the presence of jumps in $X$. As a generalization to the standard truncation-based methods (see e.g., Chapter 9 of [10]), we allow explicitly the truncation parameter $\alpha_{n,t}$ to be time-varying and depend on $\{X_{i,\Delta_n}\}_{i=1,\ldots,\lfloor T/\Delta_n \rfloor}$. For example, one convenient and commonly used choice is to set $\alpha_{n,t} = c \hat{\sigma}_n$, where $c$ is a constant [typically in the range $(3, 5)$], and $\hat{\sigma}_n$ is a preliminary estimate of the average volatility over $[0, T]$. Assumption D is verified as soon as $\hat{\sigma}_n$ and $\hat{\sigma}_n^{-1}$ are tight. Another possibility is to make $\alpha_{n,t}$ adaptive by setting $\alpha_{n,t} = c \hat{\sigma}_t$ for $t \in [(i-1)u_n, iu_n)$, where $\hat{\sigma}_t$ is a preliminary estimate for the volatility in the local window $[(i-1)u_n, iu_n)$. For example, one may take $\hat{\sigma}_i$ to be a localized version of the Bipower variation estimator of [2]:

$$
(\hat{\sigma}_i = ((\pi/2)u_n^{-1} \sum_{j=1}^{k_n} |\Delta_{i+j}^n X||\Delta_{i+j+1}^n X|)^{1/2}.
$$

The tightness requirement in Assumption D can be easily fulfilled by replacing $\hat{\sigma}_i$ with $(\hat{\sigma}_i \vee (1/C)) \wedge C$ for some pre-specified regularization constant $C \geq 1$. Finally, in the above two examples for $\alpha_{n,t}$, we can further replace the constant $c$ with a deterministic sequence $c_n$ increasing at a logarithmic rate as $\Delta_n \to 0$.

We will use $\hat{V}_{i,\Delta_n}^*$ and $\hat{V}_{i,\Delta_n}$ to approximate for the volatility trajectory within the block. That is for $0 \leq i \leq \lfloor T/u_n \rfloor - 1$,

$$
\begin{align*}
\hat{V}_t^* &= \hat{V}_{iu_n}^* \quad \text{and} \quad \hat{V}_t = \hat{V}_{iu_n}, \\
\hat{V}_t &= \hat{V}_{(T/u_n - 1)u_n}^* \quad \text{and} \quad \hat{V}_t = \hat{V}_{(T/u_n - 1)u_n}, \quad T/u_n \leq t \leq T.
\end{align*}
$$

(3.2)

**REMARK 3.1.** We can alternatively define local estimators of volatility for each $i = 1, \ldots, \lfloor T/\Delta_n \rfloor$ by averaging the $k_n$ past squared increments below the threshold. All the results in the paper, except for Theorem 3 below, will hold for this alternative way of recovering the spot volatility.
Using $\hat{\nu}^*_t$ and $\hat{\nu}_t$, our proposed estimators of $F_T(\cdot)$ are defined as

$$\hat{F}^*_n,T(x) = \int_0^T 1_{[\hat{\nu}^*_s \leq x]} \, ds, \quad \hat{F}_{n,T}(x) = \int_0^T 1_{[\hat{\nu}_s \leq x]} \, ds, \quad x \in \mathbb{R}.$$ 

We first consider the pointwise consistency of $\hat{F}^*_n,T(x)$ and $\hat{F}_{n,T}(x)$. As a matter of fact, it is not much harder to prove a more general result as follows.

**Lemma 1.** Let $g : \mathbb{R}_+ \mapsto [0, 1]$ be a measurable function and $D_g$ be the collection of discontinuity points of $g$. Suppose:

(i) Assumption A holds for $r = 2$;

(ii) for Lebesgue a.e. $t \in [0, T]$, $\mathbb{P}(V_t \in D_g) = 0$.

Then we have

\[
\int_0^T g(\hat{\nu}_s) \, ds \xrightarrow{\mathbb{P}} \int_0^T g(V_s) \, ds;
\]

(3.3)

(b) if, in addition, $X$ is continuous, then (3.3) also holds when replacing $\hat{\nu}$ with $\hat{\nu}^*$.

Lemma 1 extends Theorem 9.4.1 in [10] by allowing for discontinuities in the test function $g(\cdot)$. For fixed $x \geq 0$, the pointwise consistency of $\hat{F}_{n,T}(x)$ and $\hat{F}^*_n,T(x)$ if $X$ is continuous, follows immediately [with $g(\cdot) = 1_{[\cdot \leq x]}$] provided that $\mathbb{P}(V_t = x) = 0$ for Lebesgue a.e. $t \in [0, T]$. The uniform consistency of $\hat{F}^*_n,T(\cdot)$ and $\hat{F}_{n,T}(\cdot)$ is available if the occupation time $F_T(\cdot)$ is a.s. continuous, as shown below.

**Theorem 1.** Suppose Assumption A holds for $r = 2$ and $F_T(\cdot)$ is a.s. continuous. We have:

(a) $\sup_{x \in \mathbb{R}} |\hat{F}_{n,T}(x) - F_T(x)| \xrightarrow{\mathbb{P}} 0$;

(b) if, in addition, $X$ is continuous, then (a) still holds when replacing $\hat{F}_{n,T}$ with $\hat{F}^*_n,T$.

Analogous to the classical notion of quantile for cumulative distribution functions, the quantile of the occupation time $F_T(\cdot)$ is naturally defined as the left-continuous functional inverse of $F_T(\cdot)$: for $\alpha \in (0, T)$, we set $Q_T(\alpha) = \inf\{x \in \mathbb{R} : F_T(x) \geq \alpha\}$. A natural estimator for $Q_T(\alpha)$ is $\hat{Q}_{n,T}(\alpha) = \inf\{x \in \mathbb{R} : \hat{F}_{n,T}(x) \geq \alpha\}$, and $\hat{Q}^*_n,T(\alpha)$ can be defined analogously for $\hat{F}^*_n,T(\cdot)$. The consistency of the quantile estimators is given by the next corollary.

**Corollary 1.** Suppose Assumption A holds for $r = 2$ and $F_T(\cdot)$ is a.s. continuous. Let $Q \equiv \{\alpha \in (0, T) : Q_T(\cdot) \text{ is continuous at } \alpha \text{ a.s.}\}$. We have for each $\alpha \in Q$:
1872 J. LI, V. TODOROV AND G. TAUCHEN

(a) \( \hat{Q}_{n,T}(\alpha) \xrightarrow{P} Q_T(\alpha) \);
(b) if, in addition, \( X \) is continuous, then \( \hat{Q}_{n,T}^*(\alpha) \xrightarrow{P} Q_T(\alpha) \).

4. Rate of convergence of \( \hat{F}_{n,T} \). We next study the rate of convergence of \( \hat{F}_{n,T}(x) \). We first consider in Section 4.1 the case when \( V \) is continuous and then the general case with discontinuous \( V \) is studied in Section 4.2, where the uniform rate of convergence of \( \hat{F}_{n,T}(\cdot) \) is also considered.

4.1. The continuous volatility case and the uniform approximation of \( V \). In the continuous volatility case we can link the rate of convergence of our volatility occupation time estimators with the rate of convergence of \( \hat{V}_t^* \) and \( \hat{V}_t \) toward \( V_t \) on the space of càdlàg functions equipped with the uniform norm. We denote the latter as

\[
\eta_n^* = \sup_{t \in [0,T]} |\hat{V}_t^* - V_t|, \quad \eta_n = \sup_{t \in [0,T]} |\hat{V}_t - V_t|.
\]

The rate of convergence of \( \hat{F}_{n,T} \) and \( \hat{Q}_{n,T} \) is then related with \( \eta_n \) through Lemma 2 below; an analogous result holds for \( \hat{F}_{n,T}^*, \hat{Q}_{n,T}^* \) and \( \eta_n^* \), but is omitted here for brevity.

**Lemma 2.** (a) For any \( x \geq 0 \) and \( \alpha \in (0, T) \), we have

\[
|\hat{F}_{n,T}(x) - F_T(x)| \leq F_T(x + \eta_n) - F_T(x - \eta_n) \quad \text{and} \quad |\hat{Q}_{n,T}(\alpha) - Q_T(\alpha)| \leq \eta_n.
\]

(b) Suppose Assumption B and \( \eta_n = O_p(a_n) \) for some nonrandom sequence \( a_n \rightarrow 0 \). Then \( \hat{F}_{n,T}(x) - F_T(x) = O_p(a_n) \).

In view of Lemma 2, bounding the rate of convergence of the occupation time estimators boils down to establishing the asymptotic order of magnitude of \( \eta_n \) and \( \eta_n^* \). Our main result concerning the uniform approximation of the \( V \) process is given by the following theorem.

**Theorem 2.** Suppose Assumptions A and C with \( \Delta V_s = 0 \) for \( s \in [0, T] \). Let \( k_n \asymp \Delta_n^{-\gamma} \) for some \( \gamma \in (r \sigma + (1 \lor r)(1 - 2\sigma), 1) \), \( \sigma \in ((1 \lor r - 1)/(2(1 \lor r) - r), 1/2) \) and \( \iota > 0 \) be arbitrarily small but fixed. We have:

(a) \( \eta_n = O_p(a_n) \), where

\[
a_n = \begin{cases} 
\Delta_n^{\gamma - 1 - (2-r)\sigma} / r - (1-\sigma) - \iota \lor \Delta_n^{\gamma / 2 - \iota} / \Delta_n^{(1-\gamma) / 2 - \iota}, & \text{if } r \leq 1, \\
\Delta_n^{\gamma / (r - (1-\sigma) - \iota)} / r - (1-\sigma) - \iota \lor \Delta_n^{\gamma / 2 - \iota} / \Delta_n^{(1-\gamma) / 2 - \iota}, & \text{if } r > 1, \\
\Delta_n^{\gamma / 2 - \iota} / \Delta_n^{(1-\gamma) / 2 - \iota}, & \text{if } X \text{ is continuous};
\end{cases}
\]

(b) if \( X \) is continuous, we also have \( \eta_n^* = O_p(a_n) \).
When $X$ is continuous, the terms $\Delta_n^{\gamma/2-\iota}$ and $\Delta_n^{(1-\gamma)/2-\iota}$ capture, respectively, the sampling variability and the discretization bias in the approximation of the spot variance. When $X$ is discontinuous, $a_n$ contains an additional term arising from the elimination of jumps, which of course depends on the concentration of “small” jumps through $r$ (recall Assumption A). The conditions on $\sigma$ and $\gamma$ imply $a_n \to 0$. In particular, when $r$ is close to 2, $\sigma$ and $\gamma$ need to be chosen close to $1/2$ and 1, respectively, to ensure that $a_n \to 0$, rendering the rate of convergence arbitrarily slow. Nonetheless, if $X$ is discontinuous, $\eta_n$ still has the same rate of convergence as in the continuous case, that is, $\Delta_n^{1/4-\iota}$, provided $r \in (0, 1/2)$. This rate can be achieved by setting $\sigma \in (3/(8-4r), 1/2)$ and $\gamma = 1/2$. Of course Theorem 2 provides only a bound for the rate of convergence of $\eta_n$ and $\eta_n^*$. The following theorem, however, establishes the exact asymptotic distributions of $\eta_n$ and $\eta_n^*$ in a simple model with constant volatility.

**Theorem 3.** Suppose:

(i) Assumption A holds with $V_t$ constant and $b_t = 0$ on $[0, T]$;

(ii) $r < 1/2$ and $\sigma \in (3/(8-4r), 1/2)$.

Then

\begin{equation}
\sqrt{\log([T/u_n])} (\sqrt{k_n} \eta_n - \sqrt{2} V m_n) \xrightarrow{\mathcal{L}} V \times \Lambda,
\end{equation}

provided $k_n \asymp \Delta_n^{-1/2}$ and where $\Lambda$ is a random variable with c.d.f. $\exp(-2 \exp(-x))$, and

\begin{equation}
m_n = \sqrt{2 \log([T/u_n])} - \frac{\log(\log([T/u_n])) + \log(4\pi)}{2\sqrt{2\log([T/u_n])}}.
\end{equation}

If we further assume (iii) $X_t$ is continuous, then (4.2) still holds with $\eta_n$ replaced by $\eta_n^*$.

**Remark 4.1.** Theorem 3 shows that the rates given in (4.1) are almost optimal when the jumps of $X_t$ are not very active ($r < 1/2$). To be precise, we observe that (4.1) suggests $\eta_n = O_p(\Delta_n^{1/4-\iota})$ for $\iota > 0$ fixed but arbitrarily small, while the optimal rate in Theorem 3 provides a slightly sharper bound $\eta_n = O_p(\Delta_n^{1/4} \log([T \Delta_n^{-1/2}])^{1/2})$.

The rate of convergence of our volatility occupation time estimators and their quantiles is a direct corollary of Lemma 2 and Theorem 2; the proof is omitted for brevity.

**Corollary 2.** Let $x \geq 0$ and $\alpha \in (0, T)$. Suppose Assumption B and the same setting as in Theorem 2. Then $\hat{F}_{n,T}(x) - F_T(x)$ and $\hat{Q}_{n,T}(\alpha) - Q_T(\alpha)$ are $O_p(a_n)$. If $X$ is continuous and $k_n \asymp \Delta_n^{-1/2}$, then $\hat{F}_{n,T}(x) - F_T(x)$, $\hat{F}_{n,T}(x) -$
\[ F_T(x), \hat{Q}_{n,T}(\alpha) - Q_T(\alpha) \] and \[ \hat{Q}_{n,T}^* (\alpha) - Q_T(\alpha) \] are \( O_p(\Delta_n^{1/4 - \iota}) \) for \( \iota > 0 \) arbitrarily small but fixed.

We should point out that in the trivial cases when \( x < \inf_{t \in [0,T]} V_t \) or \( x > \sup_{t \in [0,T]} V_t \) on a given path, the error in recovering the occupation time will become identically zero for \( n \) sufficiently high (up to taking a subsequence).

**Remark 4.2.** More generally, we can use Theorem 2 to show in the setting of the theorem that if \( L: \mathcal{D}([0,T]) \to \mathbb{R} \), where \( \mathcal{D}([0,T]) \) is the space of càdlàg functions on the interval \([0,T]\) equipped with the uniform topology, is a continuous function, we have \( \mathcal{L}(\hat{V}) \xrightarrow{p} \mathcal{L}(V) \). If further \( |\mathcal{L}(f) - \mathcal{L}(g)| \leq K \sup_{s \in [0,T]} |f_s - g_s| \) for any elements \( f, g \in \mathcal{D}([0,T]) \) and some positive constant \( K \), then the rate of convergence of \( \mathcal{L}(\hat{V}) \) to \( \mathcal{L}(V) \) is bounded by the order of magnitude of \( \eta_n \) (under the conditions of Theorem 2). An example of such a function is \( \mathcal{L}(f) = \sup_{s \in [0,T]} f_s \).

### 4.2. The discontinuous volatility case

We now turn to the general case when the volatility process contains jumps. When \( V \) is discontinuous, bounding the rate of convergence of the volatility occupation time estimators is much less straightforward. Lemma 2 is still valid, however, and the uniform approximation error \( \eta_n \) no longer vanishes asymptotically, due to the discontinuity in \( V \). This is true even if we consider the ideal case where \( \hat{V}_{iun} = V_{iun} \), that is, perfect pointwise approximation is available. Indeed, the lack of uniform approximation for a discontinuous process with its discretized version is well known in the study of convergence of processes.

Our strategy of bounding the rate of convergence of \( \hat{F}_{n,T}(x) \) and \( \hat{F}_{n,T}^*(x) \) is to pick out the “big” jumps in the \( V \) process and then consider the uniform rate of approximation to the “remainder” process. The idea is best illustrated in the basic case where the jumps in \( V \) are finitely active. In this case, the volatility jumps only occur within finitely many time blocks with the form \([iun, (i+1)un)\) and hence their total effect on the estimation is \( O_p(u_n) \). On time blocks not containing the jumps of \( V \), \( V \) is continuous, so Theorem 2 can be used to provide uniform bound.

The situation becomes considerably more complicated when \( V \) has infinitely active, or even infinite variational, jumps. In this case, one needs to compute the trade-off between picking out a smaller number of big jumps with less accurate uniform approximation to the remainder process, and picking out a larger number of big jumps with more accurate uniform approximation to the remainder process. The end result of this calculation is Theorem 4 below.

**Theorem 4.** Suppose Assumptions A, B and C. Let \( k_n \asymp \Delta_n^{-\gamma} \) for some \( \gamma \in (r\sigma + (1 \vee r)(1 - 2\sigma), 1), \sigma \in ((1 \vee r - 1)/(2(1 \vee r) - r), 1/2) \) and \( \iota > 0 \) be arbitrarily small but fixed. We have:
(a) $\widehat{F}_{n,T}(x) - F_T(x) = O_p(d_n)$, where $d_n = a_n \vee \Delta_n^{(1-\gamma)/(1+\tilde{r})^{-1}}$ and $a_n$ is given by (4.1); 
(b) if $X$ is continuous, we also have $\widehat{F}_{n,T}^*(x) - F_T(x) = O_p(d_n)$.

Theorem 4 establishes an upper bound for the pointwise rate of convergence of the occupation time estimators. We remind the reader that the rate $d_n$ depends on $r$ through $a_n$; recall (4.1) and the discussion following Theorem 2. Whether the rate is optimal or not is an open question. The rate optimality for jump-robust estimation of the integrated variance, that is, $\int_0^T V_s \, ds = \int_0^\infty x F_T(dx)$, is studied by [11].

In order to establish a uniform bound, we invoke the stronger Assumption $B'$ which assumes continuity of the volatility occupation density.

**THEOREM 5.** Consider the same setting as in Theorem 4 except with Assumption $B'$ replacing Assumption $B'$. Then (a) and (b) in Theorem 4 hold uniformly in $x \in \mathbb{R}$. Moreover, for $\alpha \in (0, T)$ with $f_T(Q_T(\alpha)) > 0$ a.s., we have $|\widehat{Q}_{n,T}(\alpha) - Q_T(\alpha)| = O_p(d_n)$ and, if $X$ is continuous, $|\widehat{Q}_{n,T}^*(\alpha) - Q_T(\alpha)| = O_p(d_n)$.

Theorem 5 establishes a bound for the uniform rate of convergence of the occupation time estimators; the rate of convergence of the quantiles follows as a consequence.

5. Estimation of the volatility occupation density. We now turn to estimating the volatility occupation density $f_T(x)$. Clearly, the uniform convergence of the occupation time (Theorem 1) does not directly lead to valid estimation for the occupation density. While the focus of the current paper is on the occupation time, we consider the estimation of $f_T(\cdot)$ theoretically complementary and empirically relevant. Our occupation density estimator is based on the local volatility estimates $\widehat{V}_t$ and $\widehat{V}_t^*$ and kernel smoothing. In particular, we propose the following kernel estimator of Nadaraya–Watson type:

$$\widehat{f}_{n,T}(x) \equiv \int_0^T \frac{1}{h_n} \kappa \left( \frac{\widehat{V}_s - x}{h_n} \right) \, ds,$$

where $h_n \to 0$ is a bandwidth sequence and the kernel function $\kappa : \mathbb{R} \mapsto \mathbb{R}_+$ is bounded $C^1$ with bounded derivative and $\int_{\mathbb{R}} \kappa(x) \, dx = 1$. We can define $\widehat{f}_{n,T}^*(x)$ similarly but with $\widehat{V}_s^*$ replacing $\widehat{V}_s$.

Below, we consider a weight function $w : \mathbb{R} \mapsto \mathbb{R}_+$ with $\int_{\mathbb{R}} w(x) \, dx < \infty$. For generic real-valued functions $g_1$ and $g_2$ on $\mathbb{R}$, we denote

$$\|g_1 - g_2\|_w \equiv \int_{\mathbb{R}} |g_1(x) - g_2(x)| \, w(x) \, dx.$$
THEOREM 6. Suppose:

(i) Assumptions A, B’ and C;
(ii) \( r \in (0, 2) \) and \( \sigma \in ((1 \lor r - 1)/(2(1 \lor r) - r), 1/2) \);
(iii) \( k_n \asymp \Delta_n^{-\gamma} \) for some \( \gamma \in (0, 1) \);
(iv) \( V_t^{-1} \) is locally bounded;
(v) for some \( \beta \in (0, 1) \) and any compact \( K \subset (0, \infty) \), there exists a constant \( C_K > 0 \), such that for all \( x, y \in K \), \( \mathbb{E}|f_T(x) - f_T(y)| \leq C_K|x - y|^{\beta} \);
(vi) \( \int_{\mathbb{R}} \kappa(z)|z|^{\beta} \, dz < \infty \).

We set
\[
\tilde{a}_n^* \equiv \frac{\gamma}{2} \Delta_n^{(1-\gamma)/2}, \quad \tilde{a}_n \equiv \tilde{a}_n^* \vee \Delta_n^{1-r\sigma-\theta}/(1\lor r)-\Delta_n^{1/(2+\beta)},
\]
where \( \theta = 0 \) when \( r \leq 1 \) and \( \theta > 0 \) is arbitrarily fixed when \( r > 1 \). Then for each \( x \geq 0 \), we have:

(a) \( \hat{f}_{n,T}(x) - f_T(x) \) and \( \| \hat{f}_{n,T} - f_T \|_w \) are \( O_p(h_n^{-2} \tilde{a}_n \vee h_n^{\beta}) \);
(b) if \( X \) is continuous, \( \hat{f}_{n,T}(x) - f_{n,T}(x) \) and \( \| \hat{f}_{n,T} - f_T \|_w \) are \( O_p(h_n^{-2} \tilde{a}_n^* \vee h_n^{\beta}) \) and moreover, the results still hold with \( \hat{f}_{n,T}(\cdot) \) replaced by \( \hat{f}_{n,T}^*(\cdot) \).

REMARK 5.1. Condition (v) in Theorem 6 requires the occupation density of \( V_t \) to be Hölder continuous on compacta with exponent \( \beta \) under the \( L_1 \)-norm. We preclude the analysis for cases in which \( f_T(\cdot) \) is differentiable, or in a Hölder class of higher order, because occupation densities of semimartingales in general do not enjoy such higher-order smoothness; recall from Assumption C that \( V_t \) is a semimartingale. For example, the occupation density of a one-dimensional Brownian motion is Hölder continuous in \( L_1 \) with exponent \( \beta = 1/2 \); see Exercise VI.1.32 in [15]. That being said, occupation densities of other processes, such as certain Gaussian processes (see, e.g., Table 2 in [8]), may enjoy higher-order smoothness. Such models have rarely been studied in the analysis of high-frequency financial data and are not directly compatible with Assumption C, so we do not pursue further results here. Notice that the rate \( \tilde{a}_n^* \) is optimized by setting \( \gamma = 1/2 \), resulting in \( \tilde{a}_n^* = \Delta_n^{1/4} \). Furthermore, when \( X \) is continuous, the estimation error of the occupation density is \( O_p(\Delta_n^{\beta/4(2+\beta)}) \), which is achieved by setting \( h_n \asymp \Delta_n^{1/(2+\beta)} \). Not surprisingly, the smoother the occupation density (larger \( \beta \)), the faster the rate of convergence.

REMARK 5.2. Theorem 6(a) implies \( \hat{f}_{n,T}(x) - f_T(x) \xrightarrow{\mathbb{P}} 0 \) and \( \| \hat{f}_{n,T} - f_T \|_w \xrightarrow{\mathbb{P}} 0 \), provided that \( h_n \to 0 \) and \( h_n^{-2} \tilde{a}_n \to 0 \). These results can be shown directly without conditions (iv)–(vi) using a very similar proof; the details are omitted for brevity. A similar comment applies to Theorem 6(b).
6. Monte Carlo. We test the performance of our nonparametric procedures on two popular stochastic volatility models. The first is the square-root diffusion volatility model, given by

\begin{equation}
\begin{aligned}
 dX_t &= \sqrt{V_t} \, dW_t, \\
 dV_t &= 0.03(1.0 - V_t) \, dt + 0.2\sqrt{V_t} \, dB_t,
\end{aligned}
\end{equation}

where \( W_t \) and \( B_t \) are two independent Brownian motions. Our second model is a jump-diffusion volatility model in which the log-volatility is a Lévy-driven Ornstein–Uhlenbeck (OU) process, that is,

\begin{equation}
\begin{aligned}
 dX_t &= e^{V_t} \, dW_t, \\
 dV_t &= -0.03V_t \, dt + dL_t,
\end{aligned}
\end{equation}

where \( L_t \) is a Lévy martingale uniquely defined by the marginal law of \( V_t \) which in turn has a selfdecomposable distribution (see Theorem 17.4 of [16]) with characteristic triplet (Definition 8.2 of [16]) of \((0, 1, \nu)\) for \( \nu(dx) = \frac{2.33e^{-2.0|x|}}{|x|^{1.465}} 1_{\{x > 0\}} \, dx \) with respect to the identity truncation function. The mean and persistence of both volatility specifications are calibrated realistically to observed financial data, and the two models differ in the presence of volatility jumps as well as in the modeling of the volatility of volatility: for model (6.1), the transformation \( \sqrt{V_t} \) is with constant diffusion coefficient while for (6.2) this is the case for the transformation \( \log V_t \).

In the Monte Carlo we fix the time span to \( T = 22 \) days (our unit of time is a day), equivalent to one calendar month, and we consider \( n = 80 \) and \( n = 400 \), which correspond to 5-minute and 1-minute, respectively, of intraday observations of \( X \) in a 6.5-hour trading day. We set \( k_n = 20 \) for \( n = 80 \) and we increase it to \( k_n = 40 \) when \( n = 400 \), which, respectively, correspond to 4 and 10 blocks per unit of time. We finally set the truncation process at \( v_{n,t} = 3\sqrt{BV_j \Delta_n^{0.49}} \) for \( t \in [j-1, j) \) and where \( BV_j = \frac{\pi}{2} \sum_{i=\lfloor (j-1)/\Delta_n \rfloor + 2}^{\lfloor j/\Delta_n \rfloor + 2} |\Delta_{i-1}^n X| \Delta_i^n X | \) is the Bipower Variation on the unit interval \([j-1, j)\). For each realization we compute the 25th, 50th and 75th volatility quantiles over the interval \([0, T]\). The results from the Monte Carlo are summarized in Table 1. Overall, the performance of our volatility quantile estimator is satisfactory. The highest bias arises for the square-root diffusion volatility model when volatility was started from a high value (the 75th quantile of its invariant distribution). Intuitively, in this case volatility drifts toward its unconditional mean, and this results in its larger variation over \([0, T]\), which in turn is more difficult to accurately disentangle from the Gaussian noise in the price process, that is, the Brownian motion \( W_t \) in \( X_t \). Consistently with our asymptotic results, the biases and the mean absolute deviations of all volatility quantiles shrink as we increase the sampling frequency from \( n = 80 \) to \( n = 400 \) in all considered scenarios.

7. Conclusion. In this paper we propose nonparametric estimators of the volatility occupation time and its density from discrete observations of the process over a fixed time interval with asymptotically shrinking mesh of the observation...
| Start value | \( \hat{Q}_{T,n}(0.25) \) | \( \hat{Q}_{T,n}(0.50) \) | \( \hat{Q}_{T,n}(0.75) \) |
|-------------|----------------|----------------|----------------|
|             | True  | Bias  | MAD  | True  | Bias  | MAD  | True  | Bias  | MAD  |
| \( V_0 = Q^V(0.25) \) | 0.3798 | -0.0536 | 0.0547 | 0.5394 | -0.0478 | 0.0514 | 0.7324 | -0.0190 | 0.0473 |
| \( V_0 = Q^V(0.50) \) | 0.6223 | -0.0916 | 0.0929 | 0.8170 | -0.0651 | 0.0703 | 1.0513 | -0.0081 | 0.0626 |
| \( V_0 = Q^V(0.75) \) | 0.9865 | -0.1516 | 0.1525 | 1.2359 | -0.0949 | 0.1027 | 1.5310 | 0.0110 | 0.0911 |

Panel A: Square-root volatility model, \( n = 80 \)

Panel B: Square-root volatility model, \( n = 400 \)

Panel C: Log-volatility model, \( n = 80 \)

Panel D: Log-volatility model, \( n = 400 \)

Notes: In all simulated scenarios \( T = 22 \), and we set \( k_n = 20 \) for \( n = 80 \) and \( k_n = 40 \) for \( n = 400 \). In each of the cases, the volatility is started from a fixed point being the 25th, 50th and 75th quantile of the invariant distribution of the volatility process, denoted correspondingly as \( Q^V(0.25) \), \( Q^V(0.50) \) and \( Q^V(0.75) \). The columns “True” report the average value (across the Monte Carlo simulations) of the true variance quantile that is estimated; MAD stands for mean absolute deviation around true value. The Monte Carlo replica is 1000.
grid. We derive the asymptotic properties of our volatility occupation time estimator and further invert it to estimate the corresponding quantiles of the volatility path over the fixed time interval. Monte Carlo shows satisfactory performance of the proposed estimation techniques.

8. Proofs. This section contains all proofs. Throughout the proof, we denote by \( K \) a generic constant that may change from line to line. We sometimes emphasize its dependence on some parameter \( p \) by writing \( K_p \). As is typical in this kind of problem, by a standard localization procedure, Assumptions A, C and D can be strengthened into the following stronger versions without loss of generality.

**Assumption SA.** We have Assumption A. The processes \( b_t \) and \( \sigma_t \) are bounded, and for some bounded nonnegative function \( \Gamma_1 \) on \( \mathbb{R} \),

\[
|\delta(\omega, t, z)| \leq \Gamma_1(z) \quad \text{and} \quad \int_{\mathbb{R}} \Gamma_1(z)^p \lambda(dz) < \infty.
\]

**Assumption SC.** We have Assumption C. The processes \( \tilde{b}_t \), \( \tilde{\sigma}_t \) and \( \tilde{\sigma}'_t \) are bounded, and for some bounded nonnegative function \( \Gamma_1\sigma \) on \( \mathbb{R} \),

\[
|\tilde{\delta}(\omega, t, z)| \leq \Gamma_1\sigma (z) \quad \text{and} \quad \int_{\mathbb{R}} \Gamma_1\sigma (z)^p \tilde{\lambda}(dz) < \infty.
\]

**Assumption SD.** We have Assumption D. Moreover, \( \alpha_{n,t} \) and \( \alpha^{-1}_{n,t} \) are uniformly bounded for all \( n, t \).

8.1. **Proofs in Section 3.**

**Proof of Lemma 1.** (a) We set \( \hat{V}_i^n = \hat{V}_{iu_n} \) for \( t \in ((i - 1)u_n, iu_n) \). Denote the left-hand side of (3.3) by \( S_n \) and \( T_n = \lfloor T/u_n \rfloor u_n \). We have

\[
S_n = \int_0^{([T/u_n] - 1)u_n} g(\hat{V}_s^+) ds + \int_0^{u_n} g(\hat{V}_s) ds + \int_{T_n}^T g(\hat{V}_s) ds.
\]

Since \( g \) is bounded,

\[
(8.1) \quad \mathbb{E}\left| S_n - \int_0^T g(V_s) ds \right| \leq Ku_n + \int_0^{([T/u_n] - 1)u_n} \mathbb{E}|g(\hat{V}_s^+) - g(V_s)| ds.
\]

Observe that for each \( s \in [0, ([T/u_n] - 1)u_n) \), \( \hat{V}_s^n \xrightarrow{p} V_s \). To see this, we recall from Assumption SD that \( \alpha_{n,t} \in [\alpha, \overline{\alpha}] \) for some constant \( \overline{\alpha} \geq \alpha > 0 \). Let \( \hat{V}_s^n(\overline{\alpha}) \) and \( \hat{V}_s^n(\alpha) \) be defined as \( \hat{V}_s^n \) except with \( \alpha_{n,t} \) replaced, respectively, by \( \overline{\alpha} \) and \( \alpha \). By Theorem 9.3.2 in [10], the right continuity of \( V \) and \( u_n \to 0 \), \( \hat{V}_s^n(\overline{\alpha}) \) and \( \hat{V}_s^n(\alpha) \) converge in probability to \( V_s \). The claim then follows \( \hat{V}_s^n(\alpha) \leq \hat{V}_s^n \leq \hat{V}_s^n(\overline{\alpha}) \).

Hence, by condition (ii) and bounded convergence, for Lebesgue a.e. \( s \in [0, T] \),

\[
\mathbb{E}|g(\hat{V}_s^n) - g(V_s)| = \mathbb{E}(g(\hat{V}_s^n) - g(V_s))1_{\{V_s \not\in D_s\}} \to 0.
\]

Applying bounded convergence on (8.1), we readily obtain (3.3). Part (b) can be shown similarly. \( \square \)
PROOF OF THEOREM 1. (a) For each $x \geq 0$, $F_T(x) = F_T(x-) \ a.s.$ by the continuity of $F_T(\cdot)$. Hence,
\[
\int_0^T \mathbb{P}(V_s = x) \, ds = \mathbb{E}[F_T(x) - F_T(x-)] = 0.
\]
Therefore, $\mathbb{P}(V_s = x) = 0$ for Lebesgue a.e. $s \in [0, T]$. By Lemma 1 with $g(\cdot) = 1_{\{\cdot \leq x\}}$, $\hat{F}_{n,T}(x) \xrightarrow{\mathbb{P}} F_T(x)$. Since $\hat{F}_{n,T}(\cdot)$ and $F_T(\cdot)$ are increasing, and $F_T(\cdot)$ is continuous, this convergence also holds locally uniformly. Since $V$ is càdlàg, $\bar{V} \equiv \sup_{t \in [0, T]} V_t = O_T(1)$. For any $\eta > 0$, there exists some $M > 0$, such that $\mathbb{P}(\bar{V} > M) < \eta$, yielding $\mathbb{P}(T \neq F_T(M)) < \eta$. Hence, for any $\varepsilon > 0$,
\[
\limsup_{n \to \infty} \mathbb{P}\left( \sup_{x \in \mathbb{R}} \left| \hat{F}_{n,T}(x) - F_T(x) \right| > \varepsilon \right) \leq \limsup_{n \to \infty} \mathbb{P}\left( \sup_{0 \leq x \leq M} \left| \hat{F}_{n,T}(x) - F_T(x) \right| > \varepsilon \right) \nonumber
\]
\[
+ \limsup_{n \to \infty} \mathbb{P}\left( \sup_{x \geq M} \left| \hat{F}_{n,T}(x) - F_T(x) \right| > \varepsilon \right) \leq \limsup_{n \to \infty} \mathbb{P}\left( \left| \hat{F}_{n,T}(M) - F_T(M) \right| > \varepsilon/2 \right) + \mathbb{P}(T - F_T(M) > \varepsilon/2) < \eta.
\]
Sending $\eta \to 0$, we readily derive the assertion in part (a). Part (b) can be proved similarly. □

PROOF OF COROLLARY 1. By Theorem 1, $\hat{F}_{n,T}(\cdot) \xrightarrow{\mathbb{P}} F_T(\cdot)$ uniformly. By a subsequence argument, we can then assume $\hat{F}_{n,T}(\cdot) \xrightarrow{a.s.} F_T(\cdot)$ uniformly without loss. The assertion in part (a) then follows Lemma 21.2 of [17]. The proof of part (b) is similar. □

8.2. Proofs in Section 4.1.

PROOF OF LEMMA 2. (a) Observe
\[
F_T(x - \eta_n) = \int_0^T 1_{\{V_s \leq x - \eta_n\}} \, ds = \hat{F}_{n,T}(x) \leq \int_0^T 1_{\{V_s \leq x + \eta_n\}} \, ds = F_T(x + \eta_n).
\]
Since $F_T(x - \eta_n) \leq F_T(x) \leq F_T(x + \eta_n)$, the first assertion in part (a) readily follows. Now consider the quantiles. By definition, $\hat{F}_{n,T}(\hat{Q}_{n,T}(\alpha)) \geq \alpha$. By (8.2), $F_T(\hat{Q}_{n,T}(\alpha) + \eta_n) \geq \alpha$. Therefore, $Q_T(\alpha) \leq \hat{Q}_{n,T}(\alpha) + \eta_n$. For any $\varepsilon > 0$, by (8.2), we have $F_T(\hat{Q}_{n,T}(\alpha) - \eta_n - \varepsilon) \leq \hat{F}_{n,T}(\hat{Q}_{n,T}(\alpha) - \varepsilon) < \alpha$. Hence,
\( \hat{Q}_{n,T}(\alpha) - \eta_n - \varepsilon < Q_T(\alpha) \). Since \( \varepsilon > 0 \) is arbitrary, \( \hat{Q}_{n,T}(\alpha) - \eta_n \leq Q_T(\alpha) \). The second assertion of part (a) is then obvious.

(b) Fix some \( \varepsilon > 0 \). There exists \( M > 0 \) such that \( \mathbb{P}(\eta_n \geq Ma_n) < \varepsilon / 2 \) for \( n \) sufficiently large. Since \( a_n \to 0 \), \( [x - \eta_n, x + \eta_n] \) is contained in \( \mathcal{N}_x \) with probability approaching one (w.p.a.1) and by part (a),

\[
\left| \hat{F}_{n,T}(x) - F_T(x) \right| \leq \int_{x - \eta_n}^{x + \eta_n} f_T(z) \, dz.
\]

Let \( M' = 4M \sup_{z \in \mathcal{N}_x} \mathbb{E}[f_T(z)] / \varepsilon \). We have for \( n \) sufficiently large,

\[
\mathbb{P} \left( \int_{x - \eta_n}^{x + \eta_n} f_T(z) \, dz > M'a_n \right) \\
\leq \mathbb{P} \left( \int_{x - Ma_n}^{x + Ma_n} f_T(z) \, dz > M'a_n \right) + \mathbb{P}(\eta_n \geq Ma_n) \\
\leq \frac{2M \sup_{z \in \mathcal{N}_x} \mathbb{E}[f_T(z)]}{M'} + \varepsilon / 2 \\
\leq \varepsilon.
\]

Hence, \( \int_{x - \eta_n}^{x + \eta_n} f_T(z) \, dz = O_p(a_n) \). The assertion in part (b) then readily follows. \( \square \)

We now prove Theorem 2, starting with two lemmas. Below, \( \| \cdot \|_p \) denotes the \( L_p \) norm.

**Lemma 3.** Let \( p \geq 1 \) be a constant and \( k_n \propto \Delta_n^{-\gamma} \) for some \( \gamma \in (0, 1) \). Suppose Assumption SA holds with \( X \) continuous and Assumption SD. Then for each \( 0 \leq i \leq \lfloor T / u_n \rfloor - 1 \),

\[
\left| \hat{V}_{i u_n}^* - V_{i u_n} \right| \leq \xi_{n,i} + \sup_{s \in [i u_n, (i+1) u_n)} |V_s - V_{i u_n}|,
\]

where the variable \( \xi_{n,i} \) satisfies \( \| \xi_{n,i} \|_p \leq K_p k_n^{-1/2} \). If we further have Assumption SC with \( \Delta V_s = 0 \) for \( s \in [0, T] \), then the majorant side of the above can be bounded by \( K_p (k_n^{-1/2} + u_n^{1/2}) \) in \( L_p \).

**Proof.** By Itô’s formula, \( \hat{V}_{i u_n}^* - V_{i u_n} = \zeta'_{n,i} + \zeta''_{n,i} \), where

\[
\zeta'_{n,i} = \frac{2}{u_n} \int_{i u_n}^{(i+1) u_n} (X_s - X_{n,s}) \, dX_s,
\]

\[
\zeta''_{n,i} = \frac{1}{u_n} \int_{i u_n}^{(i+1) u_n} (V_s - V_{i u_n}) \, ds.
\]
and $X_{n,s}$ is the discretized process given by $X_{n,s} = X_{iu_n + (j-1)\Delta_n}$ when $s \in [iu_n + (j-1)\Delta_n, iu_n + j\Delta_n)$. By classical estimates (note that $X$ is continuous),

$$
\mathbb{E} \left| \frac{2}{u_n} \int_{iu_n}^{(i+1)u_n} (X_s - X_{n,s}) b_s \, ds \right|^p \leq K_p \Delta_n^{p/2},
$$

$$
\mathbb{E} \left| \frac{2}{u_n} \int_{iu_n}^{(i+1)u_n} (X_s - X_{n,s}) \sigma_s \, dW_s \right|^p \leq K_p k_n^{-p/2}.
$$

Since $k_n = o(\Delta_n^{-1})$, we have $\|\xi_{n,i}''\|_p \leq K_p k_n^{-1/2}$ by Minkowski’s inequality. We also observe $|\xi_{n,i}''| \leq \sup_{s \in [iu_n, (i+1)u_n]} |V_s - V_{iu_n}|$.

Now note that $\hat{V}^*_{iun} - \hat{V}_{iun} = k_n^{-1} \sum_{j=1}^{k_n} (\Delta_{ik_n + j} \, X/\Delta_n^{1/2})^2 1_{[|\Delta_{ik_n + j} \, X| > v_{n,iun}]}$. Under Assumption SD, $\alpha_{n,t} \geq \alpha$ for some constant $\alpha > 0$. Hence, $v_{n,t} \geq v_n \equiv \alpha \Delta_n^{\sigma}$. Since $X$ is continuous, for any $q \geq 0$,

$$
\| (\Delta_{ik_n + j} \, X/\Delta_n^{1/2})^2 1_{[|\Delta_{ik_n + j} \, X| > v_{n,iun}]} \|_p \leq \left( \frac{\mathbb{E} |\Delta_{ik_n + j} \, X/\Delta_n^{1/2}|^{2p+q}}{(v_n/\Delta_n^{1/2})^q} \right)^{1/p} \leq K_p q \Delta_n^{q(1/2-\sigma)/p}.
$$

Since $\sigma \in (0, 1/2)$, when $q$ is taken sufficiently large, terms in the above display can be further bounded by $K_p k_n^{-1/2}$. Hence, $\|\hat{V}^*_{iun} - \hat{V}_{iun}\|_p \leq K_p k_n^{-1/2}$. The first assertion then readily follows by setting $\xi_{n,i} = |\xi_{n,i}'| + |\xi_{n,i}''| = |\hat{V}^*_{iun} - \hat{V}_{iun}|$.

Now, suppose Assumption SC together with $V$ being continuous. By standard estimates, for each $p \geq 1$, the second term on the right-hand side of (8.3) can be bounded by $K_p u_n^{-1/2}$ in $L_p$. The second assertion of the lemma is then obvious.

Under Assumption SA, we set

$$
X_t' = X_t - X_t'',
$$

$$
X_t'' = \begin{cases} 
\int_0^t \int_\mathbb{R} \delta(s,z) \mu(ds,dz), & \text{if } r \leq 1, \\
\int_0^t \int_\mathbb{R} \delta(s,z)(\mu - v)(ds,dz), & \text{if } r > 1.
\end{cases}
$$

We define $\hat{V}'$ as $\hat{V}^*$ in (3.1) but with $X'$ in place of $X$; in particular, $\hat{V}'_{iun} \equiv u_n^{-1} \sum_{j=1}^{k_n} (\Delta_{ik_n + j} X')^2$.

**Lemma 4.** Suppose that Assumption SA holds with some $r \in (0, 2)$ and Assumption SD. Let $p \geq r \lor 1$ and $\sigma \in \left( \frac{p-1}{2p-r}, \frac{1}{2} \right)$. Let $\theta \in (0, \infty)$ be arbitrarily fixed.
if \( r > 1 \) and \( \theta = 0 \) if \( r \leq 1 \). We have for each \( i \),

\[
\| \hat{V}_{iun} - \hat{V}'_{iun} \|_p \leq K_p \Delta_n^{(1-r \sigma - \theta)/p - (1-2 \sigma)}.
\]

**Proof.** Under Assumption SD, \( \alpha_{n,t} \in [\alpha, \bar{\alpha}] \) for constants \( \bar{\alpha} \geq \alpha > 0 \). We set \( \bar{v}_n = \bar{\alpha} \Delta_n^{p} \) and \( v_n = \alpha \Delta_n^{p} \). By applying Lemma 13.2.6 in \([10]\) [with \( s = 1, s' = 2, m = p, p' = 1, k = 1, F(x) = x^2 \)], we have

\[
\mathbb{E}[ (\Delta_{ik_n+j}^n X / \Delta_n^{1/2})^2 1_{|\Delta_{ik_n+j}^n X| \leq \bar{v}_n} - (\Delta_{ik_n+j}^n X' / \Delta_n^{1/2})^2 1_{|\Delta_{ik_n+j}^n X'| \leq \bar{v}_n} ]^p \leq K_p \Delta_n^{(2-r)/2 - \theta} + K_p \Delta_n^{1-r \sigma - p(1-2 \sigma) - \theta}
\]

By a similar argument as in the proof of Lemma 3,

\[
\| (\Delta_{ik_n+j}^n X' / \Delta_n^{1/2})^2 1_{|\Delta_{ik_n+j}^n X'| \leq \bar{v}_n} \|_p \leq K_p, q \Delta_n^{(1/2 - \sigma) / p}
\]

for any \( q \geq 0 \). Taking \( q \) sufficiently large, we then derive

\[
\| (\Delta_{ik_n+j}^n X' / \Delta_n^{1/2})^2 1_{|\Delta_{ik_n+j}^n X'| \leq \bar{v}_n} - (\Delta_{ik_n+j}^n X' / \Delta_n^{1/2})^2 \|_p \leq K_p \Delta_n^{(1-r \sigma)/p - (1-2 \sigma) - \theta / p}.
\]

By a similar argument, we can derive (8.4) when \( \bar{v}_n \) is replaced with \( \bar{v}_n \). Since \( \bar{v}_n \leq v_{n, iun} \leq \bar{v}_n \), (8.4) also holds when \( \bar{v}_n \) is replaced with \( v_{n, iun} \). The assertion of the lemma then follows from Minkowski’s inequality. \( \square \)

**Proof of Theorem 2.** Step 1. We first suppose \( X \) is continuous. Observe that

\[
\eta_n \leq \sup_{0 \leq i \leq [T/u_n] - 1} | \hat{V}_{iun} - V_{iun} | + 2 \sup_{0 \leq i \leq [T/u_n]} \sup_{s \in [iun, (i+1)un)} | V_{s} - V_{iun} |.
\]

Since \( V \) is continuous, \( \| \sup_{s \in [iun, (i+1)un)} | V_{s} - V_{iun} | \|_p \leq K_p u_n^{1/2} \) for any \( p \geq 1 \) by standard estimates. By Lemma 3 and the maximal inequality (e.g., Lemma 2.2.2 in \([18]\)), for any \( p \geq 1 \), \( \eta_n \|_p \leq K_p u_n^{1/2} \). Since \( k_n \propto \Delta_n^{-\gamma} \) by assumption, we derive \( \| \eta_n \|_p \leq K \Delta_n^{(\gamma \wedge (1-\gamma)) / 2 - t} \) by taking \( p \) sufficiently large.

The same argument yields \( \| \eta_n^* \|_p \leq K \Delta_n^{(\gamma \wedge (1-\gamma)) / 2 - t} \). This finishes the proof of part (a) with \( X \) continuous, as well as part (b).

Step 2. We now consider part (a) allowing \( X \) to be discontinuous. Let \( \hat{V}' \) be defined as in Lemma 4. Observe that

\[
\eta_n \leq \sup_{0 \leq i \leq [T/u_n] - 1} | \hat{V}_{iun} - \hat{V}'_{iun} | + \eta_n', \quad \text{where}
\]

\[
\eta_n' = \sup_{0 \leq i \leq [T/u_n] - 1} | \hat{V}'_{iun} - V_{iun} | + 2 \sup_{0 \leq i \leq [T/u_n]} \sup_{s \in [iun, (i+1)un)} | V_{s} - V_{iun} |.
\]
A similar argument as in part (a) yields $\eta_n^* = O_P(\Delta_n^{(1-\gamma)/2-\epsilon})$. By the maximal inequality and Lemma 4 for $p = 1 \lor r$,

$$\sup_{0 \leq i \leq \lfloor T/u_n \rfloor - 1} \| \hat{V}_{iun} - \hat{V}_{iun}' \|_p \leq K u_n^{-1/p} \Delta_n^{(1-r\omega-\theta)/p-(1-2\omega)}$$

$$\leq K \Delta_n^{(\gamma-r\omega-(1-2\omega)}$$

$$\leq \begin{cases} 
K \Delta_n^{(\gamma-r\omega-(1-2\omega)/(1-\omega)} & \text{if } r \leq 1, \\
K \Delta_n^{(\gamma-\theta)/r-(1-\omega)} & \text{if } r > 1.
\end{cases}$$

Taking $\theta$ sufficiently small in the $r > 1$ case, we readily derive the assertion in part (a). □

PROOF OF THEOREM 3. Step 1. We first prove the assertion on $\eta_n^*$, so condition (iii) is in force. In the constant volatility setting of the theorem, we have

$$\sqrt{k_n} \eta_n^* = \sqrt{2V} \times M_n,$$

where we denote

$$M_n = \sup_{i=0, \ldots, \lfloor T/u_n \rfloor - 1} |Z_i^n|,$$

(8.5)

$$Z_i^n = \sqrt{k_n} \left( \sqrt{2} V \left( \hat{V}_{iun}^* - V \right) \right), \quad i = 0, \ldots, \lfloor T/u_n \rfloor - 1.$$

Under our constant volatility assumption $\{Z_i^n\}_i$ are independent and identically distributed with distribution which is approximately standard normal. Therefore, we can use Edgeworth expansion of the c.d.f. together with extreme value theory to pin down the limit distribution of $M_n$. To this end, we set

$$\tau_n(x) = c_n x + m_n, \quad b_n = \lfloor T/u_n \rfloor, \quad x \in \mathbb{R}_+.$$

Note that $\tau_n(x) \asymp \sqrt{2 \log(b_n)}$ and hence increases to infinity as the number of blocks increases to infinity for every fixed $x$.

Using second-order Edgeworth expansion and denoting with $\Phi(\cdot)$, the c.d.f. of standard normal random variable (see Theorem 2.2 and Lemma 5.4 of [9]), we have

$$\mathbb{P}( |Z_i^n| \leq \tau_n(x) ) = \Phi(\tau_n(x)) - \Phi(-\tau_n(x)) + \frac{\log(b_n)^4}{b_n \sqrt{k_n}} K(x) + o\left( \frac{1}{k_n} \right)$$

(8.7)

for any $x$ where $K(x)$ is a polynomial of $x$. Then we have

$$\lim_{n \to \infty} \left[ \frac{\mathbb{P}( |Z_i^n| \leq \tau_n(x) )}{\Phi(\tau_n(x)) - \Phi(-\tau_n(x))} \right]^{b_n} = 1,$$

(8.8)

provided $k_n \asymp \Delta_n^{-1/2}$. This assumption on the rate of growth of $k_n$ guarantees that the distribution of $Z_i^n$ is “sufficiently close” to standard normal. Now we can
use (8.8) to get
\[
\mathbb{P}(c_n^{-1}(M_n - m_n) \leq x) = \mathbb{P}(|Z_n^x| \leq \tau_n(x))^{b_n}
\]
\[
\sim \left[ \Phi(\tau_n(x)) - \Phi(-\tau_n(x)) \right]^{b_n}
\]
as \(n \to \infty\). From here, using the results for the maximum domain of attraction of the Gumbel distribution (see e.g., Example 1.1.7 of [5]), we have
\[
\left[ \Phi(\tau_n(x)) - \Phi(-\tau_n(x)) \right]^{b_n} = \left[ 2\Phi(\tau_n(x)) - 1 \right]^{b_n}
\]
\[
\longrightarrow \exp(-2\exp(-x)) \quad \forall x,
\]
and hence
\[
c_n^{-1}(M_n - m_n) \xrightarrow{\mathcal{L}} \Lambda
\]
for \(\Lambda\) being a random variable with c.d.f. \(\exp(-2\exp(-x))\). From here the result in (4.2) (with \(\eta_n\) replaced by \(\eta_n^{*}\)) follows.

**Step 2.** We now prove (4.2) with condition (iii) relaxed. Let \(\eta_n^{*}\) be defined as \(\eta_n^{*}\) but with \(\hat{V}_t^*\) replaced by \(\hat{V}_t'\). By step 1, (4.2) holds with \(\eta_n\) replaced by \(\eta_n^{*}\). It remains to show that
\[
\log]\left(\left\lfloor T/un \right\rfloor^{1/2} \Delta_n^{-1/4}(\eta_n - \eta_n^{*}) \right) = o_p(1).
\]
Note that \(|\eta_n - \eta_n^{*}| \leq \sup_{0 \leq i \leq \lfloor T/un \rfloor - 1} |\hat{V}_{iun} - \hat{V}_{iun}'| = O_p(\Delta_n^{(2-r)\alpha - 1/2})\), where the stochastic order is shown in step 2 of the proof of Theorem 2. (8.12) then follows condition (ii). This completes the proof. \(\square\)

### 8.3. Proofs in Section 4.2

Under Assumption C, by Itô’s formula, we can represent \(V\) as
\[
V_t = V_0 + \int_0^t \tilde{b}_V,s \, ds + \int_0^t \tilde{\sigma}_V,s \, dW_s + \int_0^t \tilde{\sigma}'_V,s \, dW'_s + \int_0^t \int_\mathbb{R} \tilde{\delta}_V(s,z)(\mu - \nu)(ds,dz),
\]
where, by localization, we can assume without loss that the coefficients \(\tilde{b}_V\), \(\tilde{\sigma}_V\), \(\tilde{\sigma}'_V\) are bounded, and \(|\tilde{\delta}_V(\omega, s, z)| \leq \tilde{\Gamma}(z)\) for any \((\omega, s, z)\), where \(\tilde{\Gamma}(\cdot)\) is bounded and deterministic, and satisfies \(\int_\mathbb{R} \tilde{\Gamma}(z)\tilde{\lambda}(dz) < \infty\).

We consider the following decomposition: for \(q > 0\),
\[
V_t = V_t'(q) + V_t''(q), \quad \text{where}
\]
\[
V_t^c = V_0 + \int_0^t \tilde{b}_V,s \, ds + \int_0^t \tilde{\sigma}_V,s \, dW_s + \int_0^t \tilde{\sigma}'_V,s \, dW'_s,
\]
\[
V_t'(q) = \int_0^t \int_{\{z : \tilde{\Gamma}(z) \leq q\}} \tilde{\delta}_V(s,z)(\mu - \nu)(ds,dz)
\]
\[- \int_0^t \int \{ z : \Gamma(z) > q \} \delta_V(s, z) \nu(ds, dz), \]

\[ V_t''(q) = \int_0^t \int \{ z : \Gamma(z) > q \} \delta_V(s, z) \mu(ds, dz). \]

Denote \( I(n, i) = [iu_n, (i + 1)u_n) \). We also set \( \mathcal{I}_n(q) = \{ 0 \leq i \leq \lfloor T/u_n \rfloor - 1 : \mu(I(n, i) \times \{ z : \Gamma(z) > q \}) = 0 \} \) and \( \mathcal{I}_n(q) = \bigcup_{i \in \mathcal{I}_n(q)} I(n, i) \). Here, \( \mathcal{I}_n(q) \) collects indices of intervals not containing “big” jumps. We can decompose \( \hat{F}_{n, T}(x) = \hat{F}_{n, T}(x; q) + \hat{R}_{n, T}(x; q) \) where

\[ \hat{F}_{n, T}(x; q) = u_n \sum_{i \in \mathcal{I}_n(q)} 1_{\{ \hat{V}_{iu_n} \leq x \}} \]

\[ \hat{R}_{n, T}(x; q) = \int_{[0, T] \setminus \mathcal{I}_n(q)} 1_{\{ \hat{V}_s \leq x \}} ds. \]

Analogously, we have \( F_T(x) = F_{n, T}(x; q) + R_{n, T}(x; q) \), where

\[ F_{n, T}(x; q) = \int_{\mathcal{I}_n(q)} 1_{\{ V_s \leq x \}} ds, \]

\[ R_{n, T}(x; q) = \int_{[0, T] \setminus \mathcal{I}_n(q)} 1_{\{ V_s \leq x \}} ds. \]

Finally, we set

\[ \hat{\eta}_n(q) = \sup_{i \in \mathcal{I}_n(q)} |\hat{V}_{iu_n} - V_{iu_n}|, \]

\[ \eta'_n(q) = \sup_{0 \leq i \leq \lfloor T/u_n \rfloor} \sup_{s \in [iu_n, (i + 1)u_n)} |V'_s(q) - V'_{iu_n}(q)|, \]

\[ \eta_n(q) = \hat{\eta}_n(q) + \eta'_n(q). \]

We now generalize Lemma 2 as follows.

**Lemma 5.** Suppose \( \eta_n(q_n) = O_p(w_n) \) for some nonrandom sequences \( q_n \to 0 \) and \( w_n \to 0 \) and Assumption SA with \( r = 2 \). Then (a) under Assumption B, \( \hat{F}_{n, T}(x) - F_T(x) = O_p(u_nq_n^{-\frac{r}{2}}) + O_p(u_nq_n^{-\frac{r}{2}}) \); (b) under Assumption B’, the assertion in (a) holds uniformly in \( x \in \mathbb{R} \) and moreover (c) \( |\hat{Q}_{n, T}(\alpha) - Q_T(\alpha)| \leq \xi_n \sup_{x \in \mathbb{R}} |\hat{F}_{n, T}(x) - F_T(x)| \) for some tight sequence of variables \( \xi_n \), provided \( f_T(Q_T(\alpha)) > 0 \) a.s.

**Proof.** (a) Observe that \( \mathbb{E}[\int_{[0, T] \setminus \mathcal{I}_n(q_n)} ds] \leq Ku_nq_n^{-\frac{r}{2}} \), yielding

\[ \sup_{x \in \mathbb{R}} R_{n, T}(x; q_n) = O_p(u_nq_n^{-\frac{r}{2}}), \quad \sup_{x \in \mathbb{R}} \hat{R}_{n, T}(x; q_n) = O_p(u_nq_n^{-\frac{r}{2}}). \]

By definition, \( \hat{F}_{n, T}(x; q_n) = \int_{\mathcal{I}_n(q_n)} 1_{\{ V_s \leq x \}} ds \). Note that over \( \mathcal{I}_n(q_n) \), the process \( V''_t(q_n) \) is identically zero. Hence, \( \sup_{t \in \mathcal{I}_n(q_n)} |\hat{V}_t - V_t| \leq \eta_n(q_n) \). By a similar argument as in (8.2), we deduce

\[ |\hat{F}_{n, T}(x; q_n) - F_{n, T}(x; q_n)| \]

\[ \leq F_{n, T}(x + \eta_n(q_n); q_n) - F_{n, T}(x - \eta_n(q_n); q_n) \]

\[ \leq F_T(x + \eta_n(q_n)) - F_T(x - \eta_n(q_n)). \]
By an argument similar to part (b) of Lemma 2, we derive $F_T(x + \eta_n(q_n)) - F_T(x - \eta_n(q_n)) = O_p(w_n)$. The assertion of part (a) then follows (8.13) and (8.14).

(b) By localization, we can suppose that $V$ is bounded and thus $f_T(\cdot)$ is compactly supported. Since $f_T(\cdot)$ is continuous, $\sup_{x \in \mathbb{R}} \hat{f}_T(x) = O_p(1)$. By (8.14),

$$\sup_{x \in \mathbb{R}} \left| \hat{F}_{n,T}(x; q_n) - F_{n,T}(x; q_n) \right| \leq 2\eta_n(q_n) \sup_{z \in \mathbb{R}} f_T(z).$$

The assertion then readily follows (8.13) and (8.15).

(c) Observe that $\hat{F}_{n,T}(\hat{Q}_{n,T}(\alpha)) \geq \alpha = F_T(Q_T(\alpha))$, where the inequality follows the definition of quantiles and the equality is due to the continuity of $F_T(\cdot)$. Hence,

$$F_T(Q_T(\alpha)) - F_T(\hat{Q}_{n,T}(\alpha)) \leq \hat{F}_{n,T} (\hat{Q}_{n,T}(\alpha) - F_T(\hat{Q}_{n,T}(\alpha))$$

$$\leq \sup_{x \in \mathbb{R}} \left| \hat{F}_{n,T}(x) - F_T(x) \right|.$$ 

For any $\varepsilon > 0$, $\hat{F}_{n,T}(\hat{Q}_{n,T}(\alpha) - \varepsilon) < \alpha = F_T(Q_T(\alpha))$, yielding

$$F_T(\hat{Q}_{n,T}(\alpha) - \varepsilon) - F_T(\hat{Q}_{n,T}(\alpha)) < F_T(\hat{Q}_{n,T}(\alpha) - \varepsilon) - \hat{F}_{n,T}(\hat{Q}_{n,T}(\alpha) - \varepsilon)$$

$$\leq \sup_{x \in \mathbb{R}} \left| \hat{F}_{n,T}(x) - F_T(x) \right|.$$ 

Since $F_T(\cdot)$ is continuous, by sending $\varepsilon \downarrow 0$ we deduce

$$F_T(\hat{Q}_{n,T}(\alpha)) - F_T(Q_T(\alpha)) \leq \sup_{x \in \mathbb{R}} \left| \hat{F}_{n,T}(x) - F_T(x) \right|.$$ 

Let $K_{n,T}(\alpha)$ be the closed interval with endpoints $Q_T(\alpha)$ and $\hat{Q}_{n,T}(\alpha)$. Set $\xi_n \equiv \sup_{x \in K_{n,T}(\alpha)} f_T^{-1}(x)$. By a mean-value expansion,

$$F_T(\hat{Q}_{n,T}(\alpha) - \varepsilon) - F_T(Q_T(\alpha)) \leq \inf_{x \in K_{n,T}(\alpha)} f_T(x) \left| \hat{Q}_{n,T}(\alpha) - Q_T(\alpha) \right|.$$ 

Since $f_T(Q_T(\alpha)) > 0$ a.s., $\hat{Q}_{n,T}(\alpha) \overset{p}{\rightarrow} Q_T(\alpha)$ by Corollary 1. Since $f_T(\cdot)$ is continuous, $\inf_{x \in K_{n,T}(\alpha)} f_T(x) \overset{p}{\rightarrow} f_T(Q_T(\alpha)) > 0$; hence $\xi_n$ is tight. The assertion then follows (8.16), (8.17) and (8.18). \[\square\]

**Proof of Theorem 4.** Step 1. We first consider $\eta_n'(q_n)$. For each $i$,

$$\sup_{s \in [iu_n, (i+1)u_n)} \left| V_{s}^{'}(q_n) - V_{iu_n}^{'}(q_n) \right| \leq \xi_{n,i} + \zeta_{n,i} + \zeta_{n,i}^{''},$$

where

$$\xi_{n,i} = \sup_{s \in [iu_n, (i+1)u_n)} \left| \int_{iu_n}^{s} \int_{\{z : \tilde{\Gamma}(z) > q_n\}} \tilde{\delta}_V(s, z) \nu(ds, dz) \right|,$$

$$\zeta_{n,i} = \sup_{s \in [iu_n, (i+1)u_n)} \left| \int_{iu_n}^{s} \int_{\{z : \tilde{\Gamma}(z) \leq q_n\}} \tilde{\delta}_V(s, z)(\mu - v)(ds, dz) \right|,$$

$$\zeta_{n,i}^{''} = \sup_{s \in [iu_n, (i+1)u_n]} \left| V_{s}^{''} - V_{iu_n}^{''} \right|.$$
For $\zeta_{n,i}$, observe that
\[
\sup_{0 \leq i \leq \lfloor T/u_n \rfloor} \zeta_{n,i} \leq \sup_{0 \leq i \leq \lfloor T/u_n \rfloor} \int_{i u_n}^{(i+1)u_n} \int_{\{z : \tilde{\Gamma}(z) > q_n\}} \tilde{\Gamma}(z) \nu(ds, dz) \leq K u_n q_n^{(1-\tilde{r}) \wedge 0}.
\] (8.19)

Now turn to $\zeta'_{n,i}$. Let $p \geq 2$. For each $i \geq 0$, by Lemma 2.1.5 in [10],
\[
\mathbb{E}\|\zeta'_{n,i}\| \leq K p u_n \int_{\{z : \tilde{\Gamma}(z) \leq q_n\}} \tilde{\Gamma}(z)^p \lambda(dz) + K p u_n^{p/2} \left(\frac{1}{u_n} \int_{I(n,i)} ds \int_{\{z : \tilde{\Gamma}(z) \leq q_n\}} \tilde{\Gamma}(z)^2 \lambda(dz)\right)^{p/2} \leq K p u_n q_n^{1-\tilde{r}} + K p u_n^{p/2} q_n^{(2-\tilde{r}) p/2}.
\] Hence, $\|\zeta''_{n,i}\| \leq K p q_n^{1-\tilde{r}} + K p u_n^{1/2-1/p} q_n^{1-\tilde{r}/2}$. By the maximal inequality (Lemma 2.2.2 in [18]),
\[
\sup_{0 \leq i \leq \lfloor T/u_n \rfloor} \zeta''_{n,i} \leq K p u_n^{1/2-1/p} q_n^{1-\tilde{r}/2}.
\] (8.20)

Since $V^c$ is continuous, by a standard estimate for $\zeta''_{n,i}$ and the maximal inequality,
\[
\sup_{0 \leq i \leq \lfloor T/u_n \rfloor} \zeta''_{n,i} \leq K p u_n^{1/2-1/p}.
\] (8.21)

Combining (8.19), (8.20) and (8.21), we derive for $p \geq 2$, $\|\eta'_n(q_n)\| \leq K p a'_{n,p}$, where
\[
a'_{n,p} \equiv u_n q_n^{(1-\tilde{r}) \wedge 0} \wedge q_n^{1-\tilde{r}/p} \wedge u_n^{1/2-1/p}.
\]

Step 2. Observe that
\[
\hat{\eta}_n(q_n) \leq \sup_{i \in \mathcal{I}_n(q_n)} \left| \hat{V}_{i u_n} - \hat{V}'_{i u_n} \right| + \sup_{i \in \mathcal{I}_n(q_n)} \left| \hat{V}'_{i u_n} - V_{i u_n} \right|.
\] (8.22)

By Lemma 3, for $i \in \mathcal{I}_n(q_n)$, for some $\xi_{n,i}$ with $\|\xi_{n,i}\| \leq K p k_n^{-1/2}$,
\[
\left| \hat{V}'_{i u_n} - V_{i u_n} \right| \leq \xi_{n,i} + \sup_{s \in [i u_n, (i+1)u_n]} \left| V'_s(q_n) - V'_{i u_n}(q_n) \right|.
\]

Therefore, by a similar argument as in step 1,
\[
\sup_{i \in \mathcal{I}_n(q_n)} \left| \hat{V}'_{i u_n} - V_{i u_n} \right| \leq K p k_n^{-1/p} k_n^{-1/2} + K p a'_{n,p}.
\] (8.23)

Similarly as in step 2 of the proof of Theorem 2 [recall that $\theta = 0$ if $r \leq 1$ and $\theta \in (0, \infty)$ can be arbitrarily fixed when $r > 1$]
\[
\sup_{i \in \mathcal{I}_n(q_n)} \left| \hat{V}_{i u_n} - \hat{V}'_{i u_n} \right| = O_p(\Delta_n^{(\gamma_r - r\sigma)\wedge (1-r) - (1-2\sigma))})
\] (8.24)
Combining (8.22)–(8.24), we derive \( \hat{\eta}_n(q_n) = O_p(w_{n,p}) \) for \( p \geq 2 \), where

\[
w_{n,p} \equiv \Delta_n^{(\gamma - r - \sigma - \theta)/(1 + r) - (1 - 2\sigma)} \vee u_n^{-1/p} k_n^{-1/2} \vee a'_{n,p}.
\]

By step 1, we further derive \( \eta_n(q_n) = O_p(w_{n,p}) \).

By Lemma 5(a), we have

\[
\hat{F}_{n,T}(x) - F_T(x) = O_p(wn,p) + Op(\tilde{r} - \tilde{r}/p).
\]

Taking \( q_n = u_n^{1/(1 + \tilde{r} - \tilde{r}/p)} \) and recalling \( k_n \approx \Delta_n^{-\gamma} \), we have

\[
\hat{F}_{n,T}(x) - F_T(x) = O_p(\Delta_n^{(\gamma - r - \sigma - \theta)/(1 + r) - (1 - 2\sigma)} \vee u_n^{-1/p} k_n^{-1/2} \vee q_n^{1-\tilde{r}/p \vee u_n^{1/2-1/p}}
\]

and by taking \( p \) sufficiently large,

\[
\hat{F}_{n,T}(x) - F_T(x) = O_p(\Delta_n^{(\gamma - r - \sigma - \theta)/(1 + r) - (1 - 2\sigma)} \vee \Delta_n^{\gamma/2-1-\gamma} \vee \Delta_n^{(1-\gamma)(1/2-1/2)}/(1+\tilde{r}-\tilde{r}/p)).
\]

The discontinuous case in part (a) then readily follows the definitions of \( \theta, a_n \) [see (4.1)] and \( d_n \). The continuous case in part (a), as well as part (b), can be proved in a similar (but simpler) way. □

**Proof of Theorem 5.** We first show that \( \sup_{x \in \mathbb{R}} |\hat{F}_{n,T}(x) - F_T(x)| = O_p(d_n) \). The proof is similar as that of Theorem 4, except in step 2 of the proof, we use Lemma 5(b) instead of Lemma 5(a). The result for \( \hat{F}_{n,T}^* \) can be proved similarly. The assertion concerning \( \hat{Q}_{n,T}(\alpha) \) then follows from Lemma 5(c). The assertion on \( \hat{Q}_{n,T}^*(\alpha) \) can be proved in a similar (but simpler) way; the details are omitted for brevity. □

**8.4. Proofs in Section 5.**

**Proof of Theorem 6.** (a) By localization and condition (iv), we can assume that \( V_t \) takes value in some compact \( K \subset (0, \infty) \), and thus \( f_T(\cdot) \) is supported on \( K \). We set \( f_{n,T}(x) = \int_0^T h_n^{-1} \kappa(h_n^{-1}(V_s - x)) \, ds \). For each \( x \in \mathbb{R} \),

\[
\mathbb{E}|\hat{f}_{n,T}(x) - f_{n,T}(x)| \leq \mathbb{E} \left[ h_n^{-1} \int_0^T \left| \kappa \left( \frac{\hat{V}_s - x}{h_n} \right) - \kappa \left( \frac{V_s - x}{h_n} \right) \right| \, ds \right] \leq Kh_n^{-2} \mathbb{E} \left[ \int_0^T |\hat{V}_s - V_s| \, ds \right].
\]
By Lemmas 3 and 4, \( \mathbb{E} |\hat{V}_s - V_s| \leq K \tilde{a}_n \). Hence,

(8.25) \[ \mathbb{E} |\hat{f}_{n,T}(x) - f_{n,T}(x)| \leq K h_n^{-2} \tilde{a}_n, \quad \mathbb{E} [\|\hat{f}_{n,T} - f_{n,T}\|_w] \leq K h_n^{-2} \tilde{a}_n, \]

where \( K \) does not depend on \( x \). Now observe that \( f_{n,T}(x) = \int_{\mathbb{R}} h_n^{-1} \kappa(h_n^{-1}(y-x)) f_T(y) \, dy \). By a change of variable, \( f_{n,T}(x) = \int_{\mathbb{R}} \kappa(z) f_T(x + h_n z) \, dz \). Hence,

\[
\mathbb{E} |f_{n,T}(x) - f_T(x)| \leq \int_{\mathbb{R}} \kappa(z) \mathbb{E} |f_T(x + h_n z) - f_T(x)| \, dz \\
\leq K h_n^\beta \int_{\mathbb{R}} |z|^\beta \, dz \leq K h_n^\beta,
\]

which further implies \( \mathbb{E} [\|f_{n,T} - f_T\|_w] \leq K h_n^\beta \). Combining these estimates with (8.25) completes the proof of part (a). Part (b) can be proved in a similar way. \( \square \)

**Acknowledgments.** We would like to thank Tim Bollerslev, Nathalie Eisenbaum, Jean Jacod, Andrew Patton and Philip Protter for helpful discussions as well as an Associate Editor and two referees for very helpful suggestions. We are particularly grateful to Markus Reiss for suggesting the direct estimation approach adopted in the paper and the link with the uniform error in estimating the volatility path given in Lemma 2 of the paper.

**REFERENCES**

[1] Bandi, F. M. and Phillips, P. C. B. (2003). Fully nonparametric estimation of scalar diffusion models. *Econometrica* 71 241–283. MR1956859
[2] Barndorff-Nielsen, O. and Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics* 2 1–37.
[3] Barndorff-Nielsen, O. E. and Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics* 4 1–30.
[4] Dassios, A. (1995). The distribution of the quantile of a Brownian motion with drift and the pricing of related path-dependent options. *Ann. Appl. Probab.* 5 389–398. MR1336875
[5] De Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer, New York. MR2234156
[6] Eisenbaum, N. and Kaspi, H. (2007). On the continuity of local times of Borel right Markov processes. *Ann. Probab.* 35 915–934. MR2319711
[7] Florens-Zmirou, D. (1993). On estimating the diffusion coefficient from discrete observations. *J. Appl. Probab.* 30 790–804. MR1242012
[8] Geman, D. and Horowitz, J. (1980). Occupation densities. *Ann. Probab.* 8 1–67. MR0556414
[9] Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*. Springer, New York. MR1145237
[10] Jacod, J. and Protter, P. (2012). *Discretization of Processes. Stochastic Modelling and Applied Probability* 67. Springer, Heidelberg. MR2859096
[11] Jacod, J. and Reiss, M. (2012). A remark on the rates of convergence for integrated volatility estimation in the presence of jumps. Technical report. Available at arXiv:1209.4173v1.
[12] Jacod, J. and Rosenbaum, M. (2012). Quarticity and other functionals of volatility: Efficient estimation. Technical report. Available at arXiv:1207.3757.
[13] Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scand. J. Stat.* **36** 270–296. MR2528985

[14] Protter, P. E. (2004). *Stochastic Integration and Differential Equations: Stochastic Modelling and Applied Probability*, 2nd ed. *Applications of Mathematics* (New York) **21**. Springer, Berlin. MR2020294

[15] Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften* **293**. Springer, Berlin. MR1725357

[16] Sato, K.-i. (1999). *Lévy Processes and Infinitely Divisible Distributions. Cambridge Studies in Advanced Mathematics* **68**. Cambridge Univ. Press, Cambridge. MR1739520

[17] Van der Vaart, A. W. (1998). *Asymptotic Statistics. Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge Univ. Press, Cambridge. MR1652247

[18] Van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer, New York. MR1385671

[19] Yor, M. (1995). The distribution of Brownian quantiles. *J. Appl. Probab.* **32** 405–416. MR1334895

J. Li
G. Tauchen
DEPARTMENT OF ECONOMICS
DUKE UNIVERSITY
DURHAM, NORTH CAROLINA 27708-0097
USA
E-MAIL: jli410@duke.edu
george.tauchen@duke.edu

V. Todorov
DEPARTMENT OF FINANCE
NORTHWESTERN UNIVERSITY
Evanston, Illinois 60208-2001
USA
E-MAIL: v-todorov@northwestern.edu