Volatility Occupation Times*

Jia Li[†] and Viktor Todorov[‡] and George Tauchen[§]

August 6, 2012

Abstract

We propose a nonparametric estimator of the occupation measure of the diffusion coefficient (stochastic volatility) of a discretely observed Itô semimrtingale on a fixed interval when the mesh of the observation grid shrinks to zero asymptotically. In a first step we recover the Laplace transform of the volatility occupation measure from the discrete observations of the process and then in a second step we invert the Laplace transform via a regularized kernel to estimate the (stochastic) volatility occupation measure. We derive the order of magnitude of the estimation error locally uniformly in space and we use the result to estimate nonparametrically the quantiles associated with the volatility occupation measure.

Keywords: Occupation time, Laplace transform, stochastic volatility, ill-posed problems, regularization, quantiles, nonparametric estimation, high-frequency data, stable convergence.

*We would like to thank Tim Bollerslev, Nathalie Eisenbaum, Jean Jacod, Andrew Patton and Philip Protter for helpful discussions. Todorov's work was partially supported by NSF grant SES-0957330.

[†]Department of Economics, Duke University, Durham, NC 27708; e-mail: jl410@duke.edu.

[‡]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

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

$$X_t = X_0 + \int_0^t \alpha_s ds + \int_0^t \sigma_s dW_s + J_t, \tag{1}$$

where α_t and σ_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 the model in (1) in the general case (either in a parametric or nonparametric context) is quite complicated because of the many "layers of latency", e.g., as typical in financial applications, σ_t and J_t can have randomness not captured by X_t.

When X is sampled discretely but with mesh of the observation grid shrinking to zero, i.e., high-frequency data on X are available, then the pathwise differences in the behavior of the different components in (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., Barndorff-Nielsen and Shephard (2006) and Mancini (2009), 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 much more smoothness requirements on $g(\cdot)$ needed to determine the rate of convergence); see Theorems 3.4.1 and 9.4.1 in Jacod and Protter (2012).

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

$$F_t(x) = \int_0^t 1_{\{V_s \leq x\}} ds, \quad \forall x > 0, \quad t \in [0, T]. \tag{2}$$

Evidently, the right-hand side of (2) is of the form $\int_0^t g(V_s)ds$ with $g(u) = 1_{\{u\leq x\}}$, which unlike earlier work is a discontinuous function.

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

$$f_t(x) = \lim_{\epsilon \downarrow 0} \frac{1}{2\epsilon} \left(F_t(x+\epsilon) - F_t(x-\epsilon) \right). \tag{3}$$

The occupation time of the volatility process "summarizes" in a convenient way information regarding the volatility behavior over the given interval of time. Indeed, for any bounded (or non-negative) Borel function $g(\cdot)$, see e.g., Theorem 6.4 of Geman and Horowitz (1980), we have

$$\int_0^t g(V_s)ds = \int_{\mathbb{R}_+} g(x)f_t(x)dx = \int_{\mathbb{R}_+} g(x)dF_t(x). \tag{4}$$

Thus, the occupation time can be considered as the pathwise analogue of the cumulative distribution function. We can further invert the volatility occupation time and compute estimates of the quantiles of the trajectory of volatility process over the interval of time.

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 e.g., the discussion in Bandi and Phillips (2003). 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 e.g., Dassios (1995) and Yor (1995) and references therein. Here, we show how to measure nonparametrically occupation times associated with the volatility risk of the price process.

The nonparametric estimation can be summarized as follows. First, we aggregate the high-frequency data on a fixed time interval into the realized Laplace transform, a function that provides a consistent (and asymptotically mixture normal) estimate of the empirical Laplace transform of the spot variance over the given interval. This statistic effectively deconvolutes the hidden volatility process from the Gaussian noise (the Brownian motion W_t in (1)) as well as the drift and jump components α_t and J_t. On a second step we use a regularized kernel to invert the realized Laplace transform, which in turn yields a nonparametric estimate of the volatility occupation time. As a by-product, we can compute estimates of the corresponding quantiles of the actual path of the volatility process over the fixed time interval.

The statistical estimation problem here differs significantly from the usual problem of integrated volatility estimation considered thus far in the literature, see e.g., Jacod and Protter (2012) and references therein. Firstly, for volatility occupation time estimation, unlike previous work, the function of volatility $g(\cdot)$ in (4) is discontinuous. As a result, we show that the precision of recovering the volatility occupation time depends on the smoothness of the volatility trajectories, which in turn depends on the type of volatility model, e.g., presence or not of volatility jumps as well as their activity. More generally, our analysis shows, constructively, that all spatial information about the volatility trajectory can be recovered nonparametrically from high-frequency observations.

We can further compare our analysis here with Todorov and Tauchen (2012a), where somewhat analogous steps are followed to estimate the invariant distribution of the volatility process, but there are fundamental differences between the current paper and Todorov and Tauchen (2012a). In the current paper, unlike Todorov and Tauchen (2012a), the time span of the data is fixed and hence we are interested in pathwise properties of the latent volatility process over the fixed time interval. Thus, we do not need and we do not impose here requirements on the existence of invariant distribution of the volatility process as well as mixing type conditions. In fact, the volatility process in our setup can be nonstationary. Second, and quite importantly, the object of interest here is a random quantity, mainly the occupation measure of the volatility process, and hence the regularization error is stochastic and not deterministic as is the case when we estimate the invariant density of the volatility.

Finally, our inference for the volatility occupation time can be compared with the estimation of occupation times (and densities) of recurrent Markov diffusion processes from discrete observations of the process, see e.g., Florens-Zmirou (1993) and Bandi and Phillips (2003). The main difference is that here the state vector, and therefore the stochastic volatility, is not fully observed. Hence we need to apply a very different estimation strategy that entails recovery of the empirical Laplace transform of volatility and its inverse, while the above-mentioned papers make use of standard nonparametric kernel-based estimators.

In the empirical application to financial data sets we document an interesting pattern: the interquartile range of log variance is unrelated to the level of the variance, unlike the interquantile ranges of other transforms of the variance (like the identity and the square root). This finding is consistent with volatility models in which the log-variance process has homoscedastic innovations. More generally, our estimates of the volatility occupation times provide an important tool for analyzing characteristics of the actual latent stochastic volatility process over a given interval of time.

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, derive its asymptotic properties, and use it to estimate the associated volatility quantiles. Section 4 reports results from a Monte Carlo study of our estimation technique. In Section 5 we use our estimator to study the volatility behavior of two financial data sets: Euro/$ exchange rate futures and S&P 500 index futures. Section 6 concludes. Section 7 contains all proofs.

2 Setup

2.1 The Underlying Process

We start with introducing the formal setup. The process X in (1) is defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ with α_t and σ_t being adapted to the filtration. Further, the jump component

J_t is defined as

$$J_t = \int_0^t \int_{\mathbb{R}} \left(\delta\left(s,z\right) 1_{\{|\delta(s,z)|\leq 1\}}\right) \left(\underline{\mu} - \underline{\nu}\right)\left(ds,dz\right) + \int_0^t \int_{\mathbb{R}} \left(\delta\left(s,z\right) 1_{\{|\delta(s,z)|>1\}}\right) \underline{\mu}\left(ds,dz\right), \quad (5)$$

where $\underline{\mu}$ is a Poisson measure on $\mathbb{R}_+ \times \mathbb{R}$ with compensator $\underline{\nu}$ of the form $\underline{\nu}\left(dt,dz\right) = dt \otimes \lambda\left(dz\right)$ for some σ-finite measure λ on $\mathbb{R}$ and $\delta : \Omega \times \mathbb{R}_+ \times \mathbb{R} \mapsto \mathbb{R}$ is a predictable function. Regularity conditions on X_t are collected below.

Assumption A. For some constant $r \in (0,2)$ and $C > 0$, we have

A1. X is an Itô semimartingale given by (1) and (5), and $|\delta\left(\omega,t,z\right)| \wedge 1 \leq \Gamma_m\left(z\right)$ for all (ω,t,z) with $t \leq T_m$, where $\left(T_m\right)_{m\geq 1}$ is a localizing sequence of stopping times and each Γ_m is a nonnegative function on $\mathbb{R}$ satisfying $\int \Gamma_m\left(z\right)^r \lambda\left(dz\right) < \infty$.

A2. We have $V_t > 0$ for all $t \in [0,T]$ a.s. and both V_t and V_t^{-1} are locally bounded.

A3. The process σ_t is also an Itô semimartingale with the form

$$\sigma_t = \sigma_0 + \int_0^t \tilde{\alpha}_s ds + \int_0^t v_s dW_s + \int_0^t v'_s dW'_s + \int_0^t \int_{\mathbb{R}} \delta'\left(s,z\right) \tilde{\mu}'\left(ds,dz\right),$$

where W' is a Brownian motion orthogonal to W, μ' is a Poisson measure with compensator $\nu'\left(dt,dz\right) = dt \otimes \lambda'\left(dz\right)$ for some σ-finite measure λ' on $\mathbb{R}$, $\tilde{\mu}' = \mu' - \nu'$ and $\delta' : \Omega \times \mathbb{R}_+ \times \mathbb{R} \mapsto \mathbb{R}$ is a predictable function. We have for every t,

$$\mathbb{E}\left(|\alpha_t|^2 + |\tilde{\alpha}_t|^2 + |\sigma_t|^2 + |v_t|^2 + \left|v'_t\right|^2 + \int_{\mathbb{R}} \left|\delta'\left(t,z\right)\right|^2 \lambda'\left(dz\right)\right) \leq C.$$

A4. For every t and s,

$$\mathbb{E}\left(|\alpha_t - \alpha_s|^2 + |v_t - v_s|^2 + \left|v'_t - v'_s\right|^2 + \int_{\mathbb{R}} \left(\delta'\left(t,z\right) - \delta'\left(s,z\right)\right)^2 \lambda'\left(dz\right)\right) \leq C\left|t-s\right|.$$

Assumptions A1-A3 impose very mild regularities on the process X and are standard in the literature on discretized processes; see Jacod and Protter (2012) (Assumption A2 can be also further relaxed). Assumption A4 imposes some additional smoothness on the coefficients; but this assumption is also mild. Importantly, Assumption A imposes no parametric structure on the underlying process, allowing for jumps in X_t and σ_t, and dependence between various components in an arbitrary manner. The constant r occurring in A1 describes the concentration of small jumps; the smaller r, the stronger the assumption. Most results in this paper hold as soon as $r < 2$, except for Theorem 1, which requires a stronger assumption with $r \leq 1$.

2.2 Occupation Times

We next collect some assumptions on the volatility occupation time and its occupation density.

Assumption B. Let $(T_m)_{m\geq 1}$ be a localizing sequence of stopping times and $\gamma \in [0,1]$ be a constant.

B1. Almost surely, the function $x \mapsto F_t(x)$ is differentiable with derivative $f_t(x)$ for all $t \in [0,T]$.

B2. For any compact $\mathcal{K} \subset (0,\infty)$, $\sup_{x\in\mathcal{K}} \mathbb{E}\left[f_{T\wedge T_m}(x)\right] < \infty$.

B3. For any compact $\mathcal{K} \subset (0,\infty)$, there exist constants $(C_m)_{m\geq 1}$ such that for all $x, y \in \mathcal{K}$, we have $\mathbb{E}\left[\sup_{t\leq T} |f_{t\wedge T_m}(x) - f_{t\wedge T_m}(y)|\right] \leq C_m |x-y|^\gamma$.

Assumption B1 imposes the existence of the occupation density of V_t. Assumption B2 imposes some mild integrability on the occupation density and is satisfied as soon as the probability density of V_t is uniformly bounded in the spatial variable and over $t \in [0,T]$ (which is the case for typical volatility models). Assumption B3 imposes Hölder continuity for the occupation density in expectation. Assumption B3 is implied by Assumption B2 if we take $\gamma = 0$; taking $\gamma > 0$ slightly improves the results in Theorem 2 below. Results on the existence and the continuity of occupation density can be studied using various methods, see e.g., Geman and Horowitz (1980), Protter (2004), Marcus and Rosen (2006) and Eisenbaum and Kaspi (2007) and the many references therein.

We sometimes need to strengthen Assumptions B2 and B3 as follows.

Assumption C. Let $(T_m)_{m\geq 1}$ be a localizing sequence of stopping times and $\tilde{\gamma} > \varepsilon > 0$ be constants. We have Assumption B1, as well as the following.

C1. For any compact $\mathcal{K} \subset (0,\infty)$, $\sup_{x\in\mathcal{K}} \mathbb{E}\left[f_{T\wedge T_m}(x)^{1+\varepsilon}\right] < \infty$.

C2. For any compact $\mathcal{K} \subset (0,\infty)$, there exist constants $(C_m)_{m\geq 1}$ such that for all $x, y \in \mathcal{K}$, we have $\mathbb{E}\left[\sup_{t\leq T} |f_{t\wedge T_m}(x) - f_{t\wedge T_m}(y)|^{1+\varepsilon}\right] \leq C_m |x-y|^{\tilde{\gamma}}$.

By Jensen's inequality, Assumption C is stronger than Assumption B. Below, we provide some primitive conditions for Assumption C that are easy to verify and cover many volatility models used in financial applications (although this set of conditions is far from exhaustive).

Remark. *We note that Assumptions B3 and C2 involve expectations and for establishing pathwise Hölder continuity in the spatial argument of the occupation density (via Kolmogorov's continuity theorem or some metric entropy condition, see e.g., Ledoux and Talagrand (1991)), one typically needs a stronger condition than those in B3 and C2.*

2.3 Some primitive conditions for Assumption C

We consider the following general class of jump-diffusion volatility models:

$$dV_t = a_t dt + s(V_t)\, dB_t + dJ_{V,t}, \tag{6}$$

where a_t is a locally bounded predictable process, B_t is a standard Brownian motion, $s(\cdot)$ is a deterministic function and $J_{V,t}$ is a pure jump process. This example includes many volatility models encountered in applications.

It is helpful to consider the Lamperti transform of V_t. More precisely, we set $\widetilde{V}_t = g(V_t)$, where $g(\cdot)$ is any primitive of the function $1/s(\cdot)$, i.e., $g(x) = \int^x du/s(u)$ and the constant of integration is irrelevant. By Itô's formula, the continuous martingale part of $\widetilde{V}_t$ is B_t. Lemma 1(a) below shows that under some regularity conditions, the transformed process $\widetilde{V}_t$ satisfies Assumption C. To prove Lemma 1(a) we compute the occupation density of $\widetilde{V}_t$ explicitly in terms of stochastic integrals via the Meyer-Tanaka formula (this is possible because the continuous martingale part of $\widetilde{V}_t$ is a Brownian motion) and we then bound the corresponding spatial increments. Then Lemma 1(b) shows that V_t inherits the same property, i.e., satisfies Assumption C, provided the transformation $g(\cdot)$ is smooth enough.

Lemma 1 *(a) Let $k > 1$. Consider a process $\tilde{V}_t$ with the following form*

$$\tilde{V}_t = \tilde{V}_0 + \int_0^t \tilde{a}_s ds + B_t + \int_0^t \int_{\mathbb{R}} \tilde{\delta}(s,z)\, \underline{\mu}(ds,dz) \tag{7}$$

where $\tilde{a}_t$ is a locally bounded predictable process, B_t is a Brownian motion, $\tilde{\delta}(\cdot)$ is a predictable function. Suppose that for some constant $C > 0$,

(i) $\left|\tilde{\delta}(\omega,t,z)\right| \leq \tilde{\Gamma}_m(z)$ for all (ω,t,z) with $t \leq S_m$, where $(S_m)_{m\geq 1}$ is a localizing sequence of stopping times and each $\tilde{\Gamma}_m$ is a nonnegative deterministic function satisfying

$$\int_{\mathbb{R}} \left(\tilde{\Gamma}_m(z)^{\beta} + \tilde{\Gamma}_m(z)^{k}\right) \lambda(dz) < \infty, \qquad \text{for some } \beta \in (0,1).$$

(ii) The probability density function of $\tilde{V}_t$ is bounded on compact subsets of $\mathbb{R}$ uniformly in $t \in [0,T]$.

(iii) The process $\widetilde{V}_t$ is locally bounded.

Then the occupation density of $\tilde{V}_t$, denoted by $\tilde{f}_t(\cdot)$, exists. Moreover, for any compact $\tilde{\mathcal{K}} \subset \mathbb{R}$, there exist a localizing sequence of stopping times $(T_m)_{m\geq 1}$, such that for any $x, y \in \tilde{\mathcal{K}}$, we have $\mathbb{E}\left[\tilde{f}_{T\wedge T_m}(x)^k\right] \leq K$ and $\mathbb{E}\left[\sup_{t\leq T\wedge T_m}\left|\tilde{f}_t(x) - \tilde{f}_t(y)\right|^k\right] \leq K\,|x-y|^{(1-\beta)k\wedge(1/2)}$ for some $K > 0$.

(b) Suppose, in addition, that $\widetilde{V}_t = g(V_t)$ for some continuously differentiable strictly increasing function $g : \mathbb{R}_+ \mapsto \mathbb{R}$. Also suppose that for some $\bar{\gamma} \in (0,1]$ and any compact $\mathcal{K} \subset (0,\infty)$, there exists some constant $C > 0$, such that $|g'(x) - g'(y)| \leq C|x-y|^{\bar{\gamma}}$ for all $x, y \in \mathcal{K}$. Then V_t satisfies Assumption C.

3 Estimating Volatility Occupation Times

We proceed next with our estimation results. We suppose that the process X_t is observed at discrete times $i\Delta_n$, $i = 0, 1, \ldots$, on $[0,T]$ for fixed $T > 0$, with the time lag $\Delta_n \to 0$ asymptotically when $n \to \infty$. Our strategy for estimating the occupation time $F_T(\cdot)$ is to first estimate its Laplace transform and then to invert the latter.

We define the empirical volatility Laplace transform over the interval $[0,T]$ as

$$\mathcal{L}_T(u) = \int_0^T e^{-uV_s} ds = \int_{\mathbb{R}_+} e^{-ux} f_T(x) dx, \quad u > 0,$$

where the second equality above follows from the occupation density formula given in (4). Then, by Fubini's theorem, the Laplace transform of the volatility occupation time is given by

$$\frac{\mathcal{L}_T(u)}{u} = \int_{\mathbb{R}_+} e^{-ux} F_T(x) dx.$$

Following Todorov and Tauchen (2012b), we estimate the empirical volatility Laplace transform $\mathcal{L}_T(u)$ using the realized Laplace transform of volatility defined as

$$\widehat{\mathcal{L}}_{T,n}(u) = \Delta_n \sum_{i=1}^{[T/\Delta_n]} \cos\left(\sqrt{2u}\Delta_i^n X / \Delta_n^{1/2}\right), \quad u \in \mathbb{R}_+. \tag{8}$$

Todorov and Tauchen (2012b) show that $\widehat{\mathcal{L}}_{T,n}(\cdot) \xrightarrow{\mathbb{P}} \mathcal{L}_T(\cdot)$ under the locally uniform topology on $\mathbb{R}_+$ with an associated CLT. Consequently, $u^{-1}\widehat{\mathcal{L}}_{T,n}(u) \xrightarrow{\mathbb{P}} u^{-1}\mathcal{L}_T(u)$ for each $u \in (0,\infty)$. This result however does not suffice for our purposes as we need the Laplace transform on the whole $\mathbb{R}_+$. Therefore, in Theorems 1 and 3 below we derive the limiting behavior of $\widehat{\mathcal{L}}_{T,n}(u)$ as an element in $\mathcal{L}_1(\mathbb{R}_+, w)$ for $w : \mathbb{R}_+ \to \mathbb{R}_+$ being some weight function (which in the case of Theorem 3 can depend on n).

Once the Laplace transform of the volatility occupation time is recovered from the data, in the next step we need to invert it in order to estimate $F_T(x)$. Inverting a Laplace transform, however, is an ill-posed problem and hence requires a regularization (Tikhonov and Arsenin (1977)). Here, we adopt an approach proposed by Kryzhniy (2003a,b) and implement the following regularized inversion of $u^{-1}\mathcal{L}_T(u)$:

$$F_{T,R}(x) = \int_0^\infty \mathcal{L}_T(u) \Pi(R, ux) \frac{du}{u}, \quad x > 0, \tag{9}$$

where $R > 0$ is a regularization parameter and the inversion kernel $\Pi(R,x)$ is defined as

$$\begin{aligned}\Pi(R,x) &= \frac{4}{\sqrt{2}\pi^2}\Big(\sinh(\pi R/2)\int_0^\infty \frac{\sqrt{s}\cos(R\ln(s))}{s^2+1}\sin(xs)\,ds \\ &\quad +\cosh(\pi R/2)\int_0^\infty \frac{\sqrt{s}\sin(R\ln(s))}{s^2+1}\sin(xs)\,ds\Big).\end{aligned}$$

Our estimator for the occupation time $F_T(x)$ is constructed by simply replacing $\mathcal{L}_T(u)$ in $F_{T,R}(x)$ with $\widehat{\mathcal{L}}_{T,n}(u)$, i.e., it is given by

$$\widehat{F}_{T,n,R}(x) = \int_0^\infty \widehat{\mathcal{L}}_{T,n}(u)\Pi(R,ux)\frac{du}{u} = \int_{-\infty}^\infty \widehat{\mathcal{L}}_{T,n}(e^z)\Pi(R,xe^z)dz, \quad x>0. \tag{10}$$

Todorov and Tauchen (2012a) use a similar strategy to estimate the invariant probability density of the spot volatility process (which also implies that in the above mentioned work, unlike here, $T\to\infty$). However, the problem here is more complicated, since the estimand $F_T(\cdot)$ itself is a random function, which in particular renders the regularization error random, whereas in Todorov and Tauchen (2012a) and Kryzhniy (2003a,b), the object of interest is deterministic.

We now discuss the asymptotic properties of $\widehat{F}_{T,n,R}(x)$ when $\Delta_n\to 0$. We first consider the case with the regularization parameter R fixed. In this case, a central limit theorem for $\widehat{F}_{T,n,R}(x)$ is available for fixed $x>0$. As a matter of fact, it is not much harder to prove a more general result which has interest of its own. The notation $\stackrel{\mathcal{L}\text{-}s}{\longrightarrow}$ indicates stable convergence in law.

Theorem 1 *Suppose that Assumption A holds with some $r\in(0,1]$. Let $w:\mathbb{R}_+\mapsto\mathbb{R}$ be a Borel function such that $\int_0^\infty (u\vee 1)\,|w(u)|\,du<\infty$. Then*

(a)

$$\Delta_n^{-1/2}\int_0^\infty \left(\widehat{\mathcal{L}}_{T,n}(u)-\mathcal{L}_T(u)\right)w(u)\,du \stackrel{\mathcal{L}\text{-}s}{\longrightarrow} \int_0^\infty \xi(u)\,w(u)\,du, \tag{11}$$

where $(\xi(u))_{u\in\mathbb{R}_+}$ is defined on an extension of the probability space $(\Omega,\mathcal{F},\mathbb{P})$ and, conditionally on $\mathcal{F}$, is a centered Gaussian process with covariance function

$$\Sigma_T(u,v) = \frac{1}{2}\int_0^T \left(e^{-\left(\sqrt{u}+\sqrt{v}\right)^2 V_s} + e^{-\left(\sqrt{u}-\sqrt{v}\right)^2 V_s} - 2e^{-(u+v)V_s}\right)ds, \quad u,v\in\mathbb{R}_+. \tag{12}$$

(b) The variance of the limit variable in (11) is given by

$$VAR_T = \int_0^\infty\int_0^\infty \Sigma_T(u,v)\,w(u)\,w(v)\,dudv, \tag{13}$$

and a consistent estimator for it is constructed from

$$\widehat{VAR}_{T,n} = \int_0^\infty\int_0^\infty \widehat{\Sigma}_{T,n}(u,v)\,w(u)\,w(v)\,dudv, \tag{14}$$

where, with $h_{u,v}(x) = \cos(\sqrt{2u}x)\cos(\sqrt{2v}x) - \cos(\sqrt{2(u+v)}x)$*,* $x \in \mathbb{R}$*, we set*

$$\widehat{\Sigma}_{T,n}(u,v) = \Delta_n \sum_{i=1}^{[T/\Delta_n]} h_{u,v}(\Delta_i^n X/\Delta_n^{1/2}). \tag{15}$$

Theorem 1(a) shows the stable convergence in law of a linear functional of the normalized empirical volatility Laplace transform. With $R > 0$ and $x > 0$ fixed, we set $w(u) = \Pi(R, xu)/u$ and verify that $\int_0^\infty (u \vee 1)|w(u)|\, du < \infty$ by using Lemma 2 in Section 7. Then (11) implies that $\Delta_n^{-1/2}\left(\widehat{F}_{T,n,R}(x) - F_{T,R}(x)\right)$ converges stably in law to $\int_0^\infty \xi(u) w(u)\, du$, which conditionally on $\mathcal{F}$, is centered Gaussian with variance VAR_T; the asymptotic variance can be consistently estimated by $\widehat{VAR}_{T,n}$ as shown in part (b) of the theorem.

We caution that the asymptotic distribution of $\widehat{F}_{T,n,R}(x)$ described in the previous paragraph is centered at the regularized version of the occupation time instead of the occupation time itself. Therefore, this result can not be applied to make inference for the occupation time. This said, the result provides a feasible quantification of the sampling variability of the estimator.

We now turn to the estimation of $F_T(x)$. It is conceptually useful to decompose the estimation error $\widehat{F}_{T,n,R}(x) - F_T(x)$ into two components: the regularization error $F_{T,R}(x) - F_T(x)$ and the sampling error $\widehat{F}_{T,n,R}(x) - F_{T,R}(x)$. Theorems 2 and 3 below characterize the order of magnitude of each component when $R \to \infty$ asymptotically.

Theorem 2 *Let* $x > 0$ *be a constant. Suppose that the processes* V_t *and* V_t^{-1} *are locally bounded. Under Assumption B, as* $R \to \infty$*,*

$$F_{T,R}(x) - F_T(x) = O_p\left(R^{-1} + R^{-1-\gamma}\ln(R)\right).$$

Theorem 3 *Let* $\eta \in (0, 1/2)$ *be a constant and* $\mathcal{K} \subset (0,\infty)$ *be compact. Suppose that as* $n \to \infty$*,* $\Delta_n \to 0$ *and* $R_n \to \infty$*. Under Assumptions A1-A3,*

$$\begin{aligned} & \sup_{x\in\mathcal{K}}\left|\widehat{F}_{T,n,R_n}(x) - F_{T,R_n}(x)\right| \\ = \; & O_p\left(\exp\left(\frac{\pi R_n}{2}\right)\left(R_n^{(r\wedge 1)/2}\Delta_n^{(r\wedge 1)(1/r-1/2)} + R_n \ln(R_n)\Delta_n^{1/2} + R_n^2 \Delta_n^{(1+\eta)/2}\right)\right). \end{aligned}$$

Theorem 2 describes the order of magnitude of the regularization error. In the absence of Assumption B3 (i.e. $\gamma = 0$), the regularization error is $O_p(R^{-1}\ln(R))$; when $\gamma > 0$, the rate can be improved to $O_p\left(R^{-1}\right)$. Theorem 3 describes the order of magnitude of the sampling error uniformly over $x \in \mathcal{K}$, where the set $\mathcal{K}$ is assumed bounded both above and away from zero. The parameter η should be set close to 1/2 in order to produce better estimates.

Combining Theorems 2 and 3 and choosing the regularization parameter properly, we arrive at the following estimate for the order of magnitude of the estimation error.

Corollary 1 *Suppose that Assumptions A1-A3 and B hold. We set* $R_n = \delta \ln\left(\Delta_n^{-1}\right)$ *for some* $\delta \in \left(0, 2\bar{\delta}/\pi\right)$*, where* $\bar{\delta} \equiv \min\left\{(r \wedge 1)\left(1/r - 1/2\right), 1/2\right\}$*. Then for each* $x > 0$*,*

$$\widehat{F}_{T,n,R_n}(x) - F_T(x) = O_p\left(R_n^{-1} + R_n^{-1-\gamma} \ln(R_n)\right).$$

Corollary 1 suggests that the rate of convergence of $\widehat{F}_{T,n,R_n}(x)$ towards $F_T(x)$ is essentially $\ln\left(\Delta_n^{-1}\right)$. This result is in sharp contrast with the fixed-R case (Theorem 1), where the rate of convergence of $\widehat{F}_{T,n,R}(x) - F_{T,R}(x)$ is $\Delta_n^{-1/2}$. The reason for the relatively slow rate of convergence is the regularization error. It, in turn, depends on the smoothness of the estimand, i.e., the volatility occupation time. As we saw in Lemma 1, the occupation density is typically Hölder (in expectation) of order up to 1/2 and this is determined by the presence of the martingale component in it. The relatively low level of smoothness of the estimand requires substantial smoothing in the inversion, by picking R relatively low, and this results in the somewhat slow rate of convergence of our estimator. By contrast, when the estimand is smoother, e.g., when estimating the invariant volatility density, one needs "less" regularization and this leads to faster rate of convergence.

Corollary 1 suggests that $\widehat{F}_{T,n,R_n}(x)$ consistently estimates $F_T(x)$. The convergence can be further strengthened to be uniform in x, as shown below.

Theorem 4 *Suppose the same conditions in Corollary 1 and Assumption C. Then for any compact* $\mathcal{K} \subset (0, \infty)$*,*

$$\sup_{x \in \mathcal{K}} \left| \widehat{F}_{T,n,R_n}(x) - F_T(x) \right| \xrightarrow{\mathbb{P}} 0.$$

Remark. *A trivial consequence of Theorem 4 is the following*

$$g(+\infty)T - \int_{\mathbb{R}_+} g'(x) \widehat{F}_{T,n,R_n}(x) dx \xrightarrow{\mathbb{P}} \int_0^T g(V_s) ds,$$

where $g(\cdot)$ *is* C^1 *function with* $g'(x) = 0$ *for* x *near zero or sufficiently large.*

Next, we provide a refinement to the functional estimator $\widehat{F}_{T,n,R_n}(\cdot)$. While the occupation time $x \mapsto F_T(x)$ is a pathwise increasing function by design, the proposed estimator $\widehat{F}_{T,n,R_n}(\cdot)$ is not guaranteed to be monotone. We propose a monotonization of $\widehat{F}_{T,n,R_n}(\cdot)$ via rearrangement, and as a by-product, consistent estimators of the quantiles of the occupation time. To be precise, for $\tau \in (0, T)$, we define the τ-quantile of the occupation time as its pathwise left-continuous inverse:

$$Q_T(\tau) = \inf\left\{x \in \mathbb{R}_+ : F_T(x) \geq \tau\right\}.$$

For any compact $\mathcal{K} \subset (0,\infty)$, we define the $\mathcal{K}$-constrained τ-quantile of $F_T(\cdot)$ as

$$Q_T^{\mathcal{K}}(\tau) = \inf\{x \in \mathcal{K} : F_T(x) \geq \tau\},$$

where the infimum over an empty set is given by $\sup \mathcal{K}$. While $Q_T(\tau)$ is of natural interest, we are only able to consistently estimate $Q_T^{\mathcal{K}}(\tau)$, although $\mathcal{K} \subset (0,\infty)$ can be arbitrarily large. This is due to the technical reason that the uniform convergence in Theorem 4 is only available over a nonrandom index set $\mathcal{K}$, which is bounded above and away from zero, but every quantile $Q_T(\tau)$ is itself a random variable and thus may take values outside $\mathcal{K}$ on some sample paths. Such a complication would not exist if $F_T(\cdot)$, and hence $Q_T(\tau)$, were deterministic—the standard case in econometrics and statistics. Of course, if the process V_t is known *a priori* to take values in some set $\mathcal{K} \subset (0,\infty)$, then $Q_T(\cdot)$ and $Q_T^{\mathcal{K}}(\cdot)$ coincide. In practice, the "$\mathcal{K}$-constraint" is typically unbinding as long as we do not attempt to estimate extreme (pathwise) quantiles of the process V_t.

We propose an estimator for $Q_T^{\mathcal{K}}(\tau)$ and a $\mathcal{K}$-constrained monotonized version $\widehat{F}_{T,n,R_n}^{\mathcal{K}}(\cdot)$ of the occupation time as follows:

$$\begin{aligned}\widehat{Q}_{T,n,R_n}^{\mathcal{K}}(\tau) &= \inf \mathcal{K} + \int_{\inf \mathcal{K}}^{\sup \mathcal{K}} 1\left\{\widehat{F}_{T,n,R_n}(y) < \tau\right\} dy, \quad \tau \in (0,T),\\ \widehat{F}_{T,n,R_n}^{\mathcal{K}}(x) &= \inf\left\{\tau \in (0,T) : \widehat{Q}_{T,n,R_n}^{\mathcal{K}}(\tau) > x\right\}, \quad x \in \mathbb{R},\end{aligned}$$

where on the second line, the infimum over an empty set is given by T. By construction, $\widehat{Q}_{T,n,R_n}^{\mathcal{K}} : (0,T) \mapsto \mathcal{K}$ is increasing and left continuous and $\widehat{F}_{T,n,R_n}^{\mathcal{K}} : \mathbb{R} \mapsto [0,T]$ is increasing and right continuous. Moreover, $\widehat{Q}_{T,n,R_n}^{\mathcal{K}}$ is the quantile function of $\widehat{F}_{T,n,R_n}^{\mathcal{K}}$, i.e., for $\tau \in (0,T)$, $\widehat{Q}_{T,n,R_n}^{\mathcal{K}}(\tau) = \inf\left\{x : \widehat{F}_{T,n,R_n}^{\mathcal{K}}(x) \geq \tau\right\}$. The asymptotic properties of $\widehat{F}_{T,n,R_n}^{\mathcal{K}}(\cdot)$ and $\widehat{Q}_{T,n,R_n}^{\mathcal{K}}(\tau)$ are given in Theorem 5 below.

Theorem 5 *Let $\mathcal{K} \subset (0,\infty)$ be compact. If $\sup_{x\in\mathcal{K}} \left|\widehat{F}_{T,n,R_n}(x) - F_T(x)\right| \xrightarrow{\mathbb{P}} 0$, then we have the following.*

(a)

$$\sup_{x\in\mathcal{K}} \left|\widehat{F}_{T,n,R_n}^{\mathcal{K}}(x) - F_T(x)\right| \xrightarrow{\mathbb{P}} 0.$$

(b) For every $\tau^* \in \{\tau \in (0,T) : Q_T(\cdot)$ is continuous at τ almost surely$\}$,

$$\widehat{Q}_{T,n,R_n}^{\mathcal{K}}(\tau^*) \xrightarrow{\mathbb{P}} Q_T^{\mathcal{K}}(\tau^*).$$

We note that the monotonization procedure here is similar to that in Chernozhukov et al. (2010), which in turn has a deep root in functional analysis (Hardy et al. (1952)). Our results are distinct from those of Chernozhukov et al. (2010) in two aspects. First, the estimand considered here, i.e. the occupation time, is a *random* function. Second, as we are interested in the convergence in probability, we only need to assume that $\sup_{x\in\mathcal{K}}\left|\widehat{F}_{T,n,R_n}(x)-F_T(x)\right|\xrightarrow{\mathbb{P}}0$ and, of course, our argument does not rely on the functional delta method.

4 Monte Carlo

We test the performance of our nonparametric procedures on two popular stochastic volatility models (both of which satisfy Lemma 1 and hence Assumption C). The first is the square-root diffusion volatility model, given by

$$dX_t=\sqrt{V_t}dW_t,\quad dV_t=0.03(1.0-V_t)dt+0.2\sqrt{V_t}dB_t, \tag{16}$$

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, i.e.,

$$dX_t=e^{V_t-1}dW_t,\quad dV_t=-0.03V_tdt+dL_t, \tag{17}$$

where L_t is a Lévy martingale uniquely defined by the marginal law of V_t which in turn has a self-decomposable distribution (see Theorem 17.4 of Sato (1999)) with characteristic triplet (Definition 8.2 of Sato (1999)) of $(0,1,\nu)$ for $\nu(dx)=\frac{2.33e^{-2.0|x|}}{|x|^{1+0.5}}1_{\{x>0\}}dx$ 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 (16), the transformation $\sqrt{V_t}$ is with constant diffusion coefficient while for (17) 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 the regularization parameter to $R_n=3.0$ for $n=80$ and we increase it to $R_n=3.5$ when $n=400$. For each realization we compute the 25-th, 50-th and 75-th volatility qunatiles 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 75-th quantile of its invariant distribution). Intuitively, in this case volatility drifts towards 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, i.e., the Brownian motion W_t in X_t. Consistent with our asymptotic results, the biases and the mean absolute deviation of all volatility quantiles and in all considered scenarios shrink as we increase the sampling frequency from $n = 80$ to $n = 400$.

5 Empirical Application

We illustrate the nonparametric quantile reconstruction technique with empirical application to two data sets: Euro/$ exchange rate futures (for the period 01/01/1999-12/31/2010) and S&P 500 index futures (for the period 04/22/1982-12/30/2010). Both series are sampled every 5 minutes during the trading hours. The time spans of the two data sets differ because of data availability but both data sets include some of the most quiescent and also the most volatile periods in modern financial history. These data sets thereby present a serious challenge for our method.

In the calculations of the volatility quantiles we use a time span of $T = 1$ month and as in the Monte Carlo we fix the regularization parameter at $R_n = 3$. Figure 1 shows the results for the Euro/$ rate and Figure 2 those for the S&P 500 index. The left panels show the time series of the 25-th and 75-th monthly quantiles of the spot variance V_t, the spot volatility $\sqrt{V_t}$ and the logarithm of the spot variance $\log(V_t)$. The estimated quantiles appear to track quite sensibly the behavior of volatility during times of either economic moderation or distress. The right panels show the associated interquartile range (IQR) versus the median of the logarithm of the spot variance; we use the IQR to measure the dispersion of the (transformed) volatility process. The aim of the plots is to discover how the dispersion of volatility relates to the level volatility. We see that for both data sets, the IQRs of the spot variance and the spot volatility exhibit a clear positive convex relationship with the median log-variance. In contrast, the IQR of the log-variance process shows no such pattern, suggesting that the log volatility process has homoscedastic, or at least independent from the level of volatility, innovations.

To guide intuition about our empirical findings, suppose we have $f(V_t) = f(V_0) + L_t$ on $[0, T]$, for L_t a Lévy process and $f(\cdot)$ some monotone function (this is approximately true for the typical volatility models like the ones in the Monte Carlo when T is relatively short and the volatility is very persistent as in the data). In this case, the interquartile range of the volatility occupation time of $f(V_t)$ on $[0, T]$ will be independent of the level V_0. On the other hand, for other functions $h(V_t)$, and in particular $h(V_t) = V_t$, the dispersion will depend in general on the level V_0. The IQR of the volatility occupation measure can be used, therefore, to study the important question of modeling the volatility of volatility. The evidence here points away from affine volatility models towards those models in which the log volatility has innovations that are independent from the level of volatility like the exponential OU model in (17).

Table 1: Monte Carlo Results

Start Value	$\widehat{Q}_{T,n}(0.25)$			$\widehat{Q}_{T,n}(0.50)$			$\widehat{Q}_{T,n}(0.75)$		
	True	Bias	MAD	True	Bias	MAD	True	Bias	MAD
Panel A: **Square-Root Volatility Model,** $n = 80$									
$V_0 = Q^V(0.25)$	0.3798	−0.0301	0.0477	0.5394	−0.0306	0.0518	0.7324	0.0043	0.0572
$V_0 = Q^V(0.50)$	0.6223	−0.0672	0.0824	0.8170	−0.0478	0.0712	1.0513	0.0300	0.0853
$V_0 = Q^V(0.75)$	0.9865	−0.1294	0.1447	1.2359	−0.0731	0.1014	1.5310	0.0774	0.1382
Panel B: **Square-Root Volatility Model,** $n = 400$									
$V_0 = Q^V(0.25)$	0.3798	−0.0147	0.0378	0.5394	−0.0177	0.0385	0.7324	−0.0026	0.0429
$V_0 = Q^V(0.50)$	0.6223	−0.0435	0.0629	0.8170	−0.0306	0.0535	1.0513	0.0186	0.0642
$V_0 = Q^V(0.75)$	0.9865	−0.0925	0.1109	1.2359	−0.0489	0.0766	1.5310	0.0574	0.1089
Panel C: **Log-Volatility Model,** $n = 80$									
$V_0 = Q^V(0.25)$	0.1737	−0.0128	0.0244	0.2860	−0.0082	0.0310	0.4519	−0.0010	0.0448
$V_0 = Q^V(0.50)$	0.3293	−0.0251	0.0465	0.5243	−0.0160	0.0558	0.8069	0.0038	0.0745
$V_0 = Q^V(0.75)$	0.6337	−0.0453	0.0860	0.9945	−0.0304	0.1088	1.5162	0.0094	0.1394
Panel D: **Log-Volatility Model,** $n = 400$									
$V_0 = Q^V(0.25)$	0.1737	−0.0050	0.0192	0.2860	−0.0003	0.0245	0.4519	−0.0114	0.0411
$V_0 = Q^V(0.50)$	0.3293	−0.0110	0.0354	0.5243	0.0005	0.0433	0.8069	−0.0151	0.0667
$V_0 = Q^V(0.75)$	0.6337	−0.0283	0.0645	0.9945	−0.0011	0.0792	1.5162	−0.0185	0.1190

Note: In all simulated scenarios $T = 22$ *and we set* $R_n = 3.0$ *for* $n = 80$ *and* $R_n = 3.5$ *for* $n = 400$. *In each of the cases, the volatility is started from a fixed point being the* 25*-th,* 50*-th and* 75*-th 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.

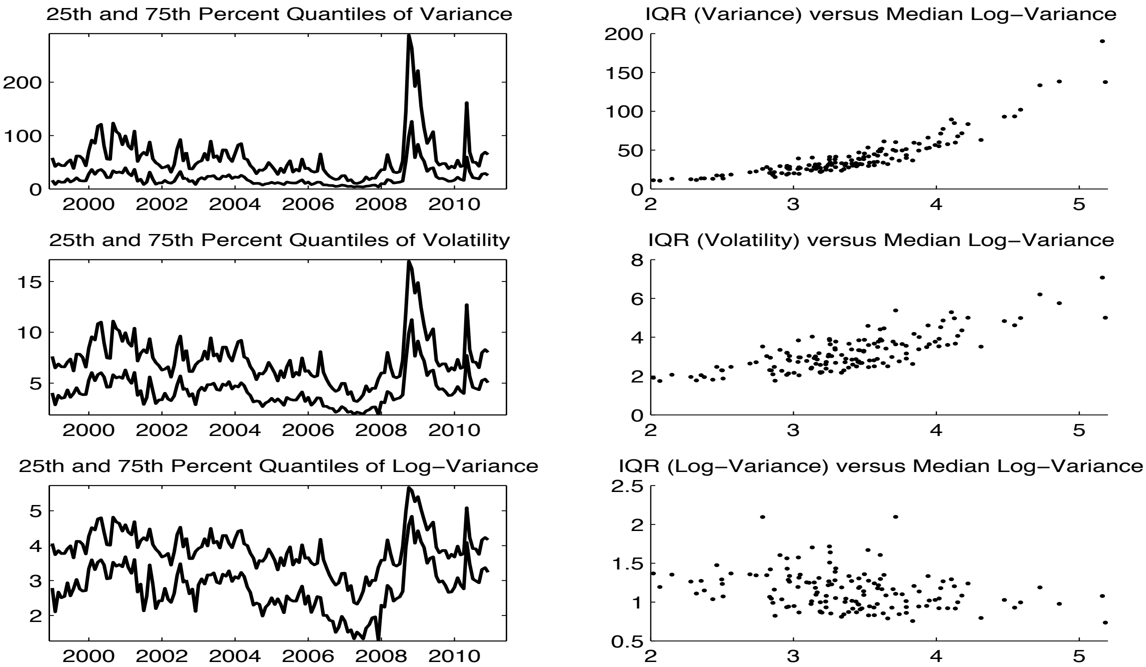

Figure 1: *Estimated Quantiles of the Monthly Occupation Measure of the Spot Volatility of the Euro/$ return, 1999–2010.* The three left-hand panels show the 25 and 75 percent quantiles of the monthly occupation measure of volatility expressed in terms of the local variance (left-top), the local standard deviation (left-middle), and the local log-variance (left-bottom). Each right-side panel is a scatter plot of the interquartile ranges of the associated monthly left-side distributions versus the medians of the distributions (in standard deviation). Volatility is quoted annualized and in percentage terms.

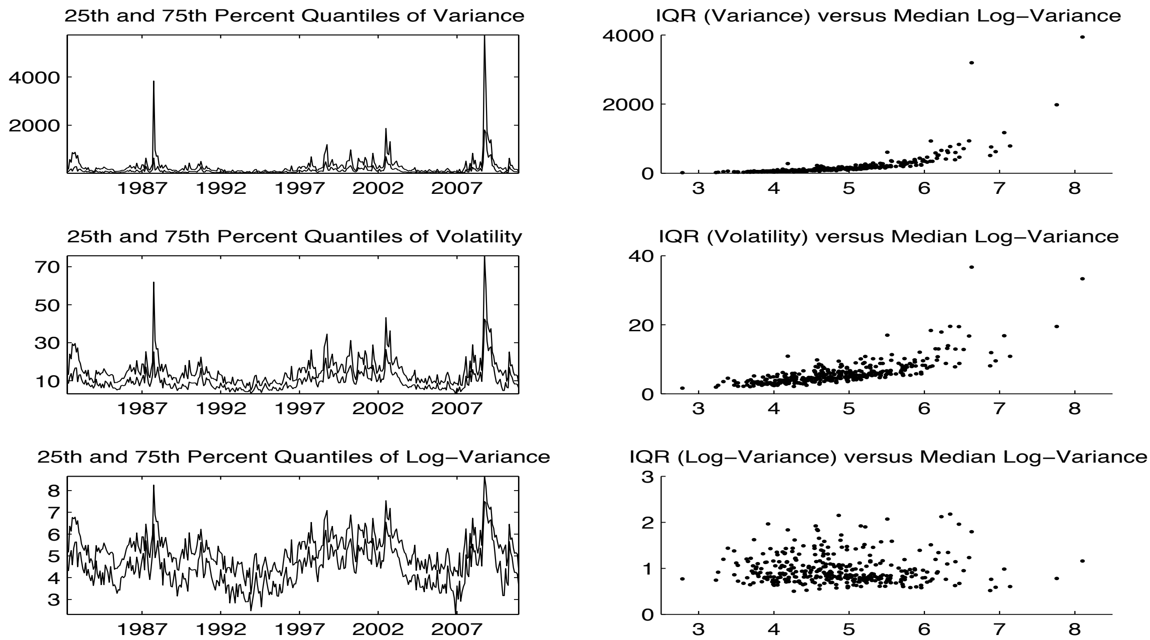

Figure 2: *Estimated Quantiles of the Monthly Occupation Measure of the Spot Volatility of the S&P500 index futures return, 1982–2010.* The organization is the same as Figure 1.

6 Conclusion

In this paper we propose a nonparametric estimator of the volatility occupation time from discrete observations of the process over a fixed time interval with asymptotically shrinking mesh of the observation grid. We derive the asymptotic properties of the volatility occupation time estimator locally uniformly in the spatial argument and further invert it to estimate the corresponding quantiles of volatility over the time interval. Monte Carlo and empirical applications illustrate the use of the statistic for studying volatility of volatility.

7 Proofs

This section is organized as follows. We collect some preliminary estimates in Section 7.1. The rest of the appendix is devoted to proving results in the main text. Throughout the proof, we use K to denote a generic positive constant that may change from line to line. We sometimes write K_m to emphasize the dependence of the constant on some parameter m.

7.1 Preliminary estimates

7.1.1 Estimates for the kernel $\Pi(R,x)$

Lemma 2 *Fix $x>0$, $c>0$, $\eta_1\in[0,1/2)$ and $\eta_2\in[0,1/2)$. There exists some $K>0$, such that for any $R\geq c$,*

$$|\Pi(R,x)|\leq K\exp\left(\frac{\pi R}{2}\right)\min\left\{x^{\eta_1},Rx^{-1},R^2x^{-1-\eta_2}\right\}.$$

Proof. To simplify notations, we denote

$$h_R(s)=\frac{\sqrt{s}\sin(R\ln(s))}{s^2+1},\quad g_R(x)=\int_0^\infty h_R(s)\sin(xs)\,ds.$$

Since $\eta_1\in[0,1/2)$, we have

$$|g_R(x)|\leq\int_0^\infty\frac{\sqrt{s}\,|\sin(xs)|}{s^2+1}ds\leq\int_0^\infty\frac{\sqrt{s}\,|\sin(xs)|^{\eta_1}}{s^2+1}ds\leq Kx^{\eta_1}. \tag{18}$$

Using integration by parts, we have $g_R(x)=x^{-1}\int_0^\infty h_R'(s)\cos(xs)\,ds$. With $h_R'(s)$ explicitly computed, we have

$$|g_R(x)|\leq x^{-1}\left|\int_0^\infty\left(\frac{R\cos(R\ln(s))}{\sqrt{s}\,(1+s^2)}+\frac{(1-3s^2)\sin(R\ln(s))}{2\sqrt{s}\,(1+s^2)^2}\right)ds\right|\leq KRx^{-1}. \tag{19}$$

Using integration by parts again, we have $g_R(x) = -x^{-2}\int_0^\infty h_R''(s)\sin(xs)\,ds$. Note that

$$\begin{aligned}\left|h_R''(s)\right| &= \left|\frac{4Rs^{1/2}\cos(R\ln(s))}{(1+s^2)^2} + \frac{\left(1+18s^2-15s^4+4R^2\left(1+s^2\right)^2\right)\sin(R\ln(s))}{4s^{3/2}(1+s^2)^3}\right| \\ &\le \frac{KR^2}{s^{3/2}(1+s^2)}.\end{aligned}$$

Hence, for $\eta_2 \in [0, 1/2)$,

$$\begin{aligned}|g_R(x)| &\le KR^2x^{-2}\int_0^\infty \frac{|\sin(xs)|}{s^{3/2}(1+s^2)}ds \\ &\le KR^2x^{-2}\int_0^\infty \frac{|\sin(xs)|^{1-\eta_2}}{s^{3/2}(1+s^2)}ds \\ &\le KR^2x^{-1-\eta_2}\int_0^\infty \frac{1}{s^{1/2+\eta_2}(1+s^2)}ds \\ &\le KR^2x^{-1-\eta_2}. \qquad (20)\end{aligned}$$

Combining (18), (19) and (20), we derive $|g_R(x)| \le K\min\left\{x^{\eta_1}, Rx^{-1}, R^2x^{-1-\eta_2}\right\}$. Similarly, we can also show that

$$\left|\int_0^\infty \frac{\sqrt{s}\cos(R\ln(s))}{s^2+1}\sin(xs)\right| \le K\min\left\{x^{\eta_1}, Rx^{-1}, R^2x^{-1-\eta_2}\right\}.$$

The assertion of the lemma then readily follows. *Q.E.D.*

7.1.2 Estimates for the underlying process X

As often in this kind of problem, it is convenient to strengthen Assumption A as follows.

Assumption SA. We have Assumptions A1-A3, and there are a constant $C > 0$ and a nonnegative function Γ on $\mathbb{R}$, such that $V_t \le C$, $|\delta(\omega, t, z)| \le \Gamma(z) \le C$ and $\int_{\mathbb{R}} \Gamma(z)^r \lambda(dz) \le C$.

For notational simplicity, we set

$$\begin{aligned}\alpha_t' &= \begin{cases}\alpha_t & \text{if } r > 1 \\ \alpha_t - \int_{\mathbb{R}} \delta(t,z)\,1_{\{|\delta(t,z)|\le 1\}}\lambda(dz) & \text{if } r \le 1,\end{cases} \\ J_t' &= \begin{cases} J_t & \text{if } r > 1 \\ \int_0^t\int_{\mathbb{R}} \delta(s,z)\,\underline{\mu}(ds,dz) & \text{if } r \le 1,\end{cases}\end{aligned}$$

where r is the constant in Assumption A1. We also set $X_t^c = X_0 + \int_0^t \alpha_s' ds + \int_0^t \sigma_s dW_s$, so that $X_t = X_t^c + J_t'$. We then define

$$\begin{aligned}\chi_i^n &= \Delta_i^n X^c/\Delta_n^{1/2}, \quad \beta_i^n = \sigma_{(i-1)\Delta_n}\Delta_i^n W/\Delta_n^{1/2}, \\ \lambda_i^n &= \chi_i^n - \beta_i^n = \Delta_n^{-1/2}\left(\int_{(i-1)\Delta_n}^{i\Delta_n} \alpha_s' ds + \int_{(i-1)\Delta_n}^{i\Delta_n}\left(\sigma_s - \sigma_{(i-1)\Delta_n}\right)dW_s\right).\end{aligned}$$

Lemma 3 *Under Assumption SA, there exists $K > 0$ such that for all $u \in \mathbb{R}_+$,*

$$\mathbb{E}\left|\Delta_n \sum_{i=1}^{[T/\Delta_n]} \left(\cos\left(\sqrt{2u}\chi_i^n\right) - \cos\left(\sqrt{2u}\beta_i^n\right)\right)\right| \leq K \min\left\{u^{1/2}\Delta_n^{1/2}, 1\right\} \tag{21}$$

$$\mathbb{E}\left|\Delta_n \sum_{i=1}^{[T/\Delta_n]} \exp\left(-uV_{(i-1)\Delta_n}\right) - \int_0^T \exp\left(-uV_s\right) ds\right| \leq K \min\left\{u\Delta_n^{1/2}, 1\right\} + \Delta_n \tag{22}$$

$$\mathbb{E}\left|\Delta_n \sum_{i=1}^{[T/\Delta_n]} \left(\cos\left(\sqrt{2u}\beta_i^n\right) - \exp\left(-uV_{(i-1)\Delta_n}\right)\right)\right| \leq K\Delta_n^{1/2} \min\left\{u, 1\right\} \tag{23}$$

$$\mathbb{E}\left|\Delta_n \sum_{i=1}^{[T/\Delta_n]} \left(\cos\left(\sqrt{2u}\frac{\Delta_i^n X}{\sqrt{\Delta_n}}\right) - \cos\left(\sqrt{2u}\chi_i^n\right)\right)\right| \leq K\left(u^{1/2}\Delta_n^{1/r-1/2}\right)^{r\wedge 1}. \tag{24}$$

Proof. Step 1. By the mean value theorem, $\mathbb{E}\left|\cos\left(\sqrt{2u}\chi_i^n\right) - \cos\left(\sqrt{2u}\beta_i^n\right)\right| \leq K \min\left\{\sqrt{u}\mathbb{E}\left|\lambda_i^n\right|, 1\right\}$. By the Burkholder-Davis-Gundy inequality,

$$\begin{aligned}
\mathbb{E}\left|\lambda_i^n\right| &\leq \Delta_n^{-1/2}\mathbb{E}\left|\int_{(i-1)\Delta_n}^{i\Delta_n} \alpha_s' ds + \int_{(i-1)\Delta_n}^{i\Delta_n} \left(\sigma_s - \sigma_{(i-1)\Delta_n}\right) dW_s\right| \\
&\leq \Delta_n^{-1/2}\left(K\Delta_n + \left(\int_{(i-1)\Delta_n}^{i\Delta_n} \mathbb{E}\left[\left(\sigma_s - \sigma_{(i-1)\Delta_n}\right)^2\right] ds\right)^{1/2}\right) \\
&\leq K\Delta_n^{1/2}.
\end{aligned}$$

Hence, $\mathbb{E}\left|\cos\left(\sqrt{2u}\chi_i^n\right) - \cos\left(\sqrt{2u}\beta_i^n\right)\right| \leq K\min\{u^{1/2}\Delta_n^{1/2}, 1\}$, which implies (21).

Step 2. We prove (22) as follows:

$$\begin{aligned}
&\mathbb{E}\left|\Delta_n \sum_{i=1}^{[T/\Delta_n]} \exp\left(-uV_{(i-1)\Delta_n}\right) - \int_0^T \exp\left(-uV_s\right) ds\right| \\
\leq& \sum_{i=1}^{[T/\Delta_n]} \int_{(i-1)\Delta_n}^{i\Delta_n} \mathbb{E}\left|\exp\left(-uV_s\right) - \exp\left(-uV_{(i-1)\Delta_n}\right)\right| ds + \Delta_n \\
\leq& K \sum_{i=1}^{[T/\Delta_n]} \int_{(i-1)\Delta_n}^{i\Delta_n} \mathbb{E}\left[\min\left\{u\left|V_s - V_{(i-1)\Delta_n}\right|, 1\right\}\right] ds + \Delta_n \\
\leq& K\min\left\{u\Delta_n^{1/2}, 1\right\} + \Delta_n.
\end{aligned}$$

where the first inequality follows the triangle inequality, the second is due to the mean value theorem and the last inequality follows $\mathbb{E}\left|V_t - V_s\right| \leq K\left|t - s\right|^{1/2}$, which in turn is implied by Assumption SA.

Step 3. Now, consider (23). Denote $\zeta_i^n = \cos\left(\sqrt{2u}\beta_i^n\right) - \exp\left(-uV_{(i-1)\Delta_n}\right)$. Note that conditionally on $\mathcal{F}_{(i-1)\Delta_n}$, β_i^n is $\mathcal{N}\left(0, V_{(i-1)\Delta_n}\right)$ distributed. It is then easy to see that $(\zeta_i^n, \mathcal{F}_{i\Delta_n})_{i\geq 1}$ is an array of martingale differences. Moreover,

$$\mathbb{E}\left[(\zeta_i^n)^2 \left|\mathcal{F}_{(i-1)\Delta_n}\right.\right] = \frac{1}{2}\left(1 - \exp\left(-2uV_{(i-1)\Delta_n}\right)\right)^2 \leq K\min\left\{u^2 V_{(i-1)\Delta_n}^2, 1\right\}.$$

Hence, $\mathbb{E}\left[(\zeta_i^n)^2\right] \leq K\min\left\{u^2, 1\right\}$ and $\mathbb{E}\left[\left(\Delta_n \sum_{i=1}^{[T/\Delta_n]} \zeta_i^n\right)^2\right] \leq K\Delta_n \min\left\{u^2, 1\right\}$. We then deduce (23) by using Jensen's inequality.

Step 4. Finally, we show (24). When $r \in (0,1]$, by Assumption SA and Lemma 2.1.7 in Jacod and Protter (2012),

$$\begin{aligned}
\mathbb{E}\left|\cos\left(\sqrt{2u}\frac{\Delta_i^n X}{\sqrt{\Delta_n}}\right) - \cos\left(\sqrt{2u}\chi_i^n\right)\right| &\leq K\mathbb{E}\left|\cos\left(\sqrt{2u}\frac{\Delta_i^n X}{\sqrt{\Delta_n}}\right) - \cos\left(\sqrt{2u}\chi_i^n\right)\right|^r \\
&\leq Ku^{r/2}\Delta_n^{-r/2}\mathbb{E}\left|\Delta_i^n J'\right|^r \\
&\leq Ku^{r/2}\Delta_n^{1-r/2}.
\end{aligned}$$

When $r \in [1,2)$, we use Assumption SA and Lemmas 2.1.5 and 2.1.7 in Jacod and Protter (2012) to derive

$$\begin{aligned}
\mathbb{E}\left|\cos\left(\sqrt{2u}\frac{\Delta_i^n X}{\sqrt{\Delta_n}}\right) - \cos\left(\sqrt{2u}\chi_i^n\right)\right|^r &\leq Ku^{r/2}\Delta_n^{-r/2}\mathbb{E}\left|\Delta_i^n J'\right|^r \\
&\leq Ku^{r/2}\Delta_n^{1-r/2},
\end{aligned}$$

and then use Jensen's inequality to get

$$\mathbb{E}\left|\cos\left(\sqrt{2u}\frac{\Delta_i^n X}{\sqrt{\Delta_n}}\right) - \cos\left(\sqrt{2u}\chi_i^n\right)\right| \leq Ku^{1/2}\Delta_n^{1/r-1/2}.$$

Combining the above estimates, we have for each $r \in (0,2)$,

$$\mathbb{E}\left|\cos\left(\sqrt{2u}\frac{\Delta_i^n X}{\sqrt{\Delta_n}}\right) - \cos\left(\sqrt{2u}\chi_i^n\right)\right| \leq K\left(u^{1/2}\Delta_n^{1/r-1/2}\right)^{r\wedge 1}.$$

Then (24) readily follows. *Q.E.D.*

Recall the notation $h_{u,v}(x)$ from Theorem 1. We set $\bar{h}_{u,v}(y) = \mathbb{E}\left[h_{u,v}(U)\right]$, for $y \geq 0$ and $U \sim \mathcal{N}(0, y)$, which is $\bar{h}_{u,v}(y) = \frac{1}{2}\left(e^{-\left(\sqrt{u}+\sqrt{v}\right)^2 y} + e^{-\left(\sqrt{u}-\sqrt{v}\right)^2 y} - 2e^{-(u+v)y}\right)$.

Lemma 4 *Under Assumption SA, there exists some $K > 0$ such that for all $u, v \in \mathbb{R}_+$,*

$$\mathbb{E}\left|\Delta_n \sum_{i=1}^{[T/\Delta_n]} \left(h_{u,v}\left(\chi_i^n\right) - h_{u,v}\left(\beta_i^n\right)\right)\right| \leq K \min\left\{\left(u^{1/2} + v^{1/2}\right)\Delta_n^{1/2}, 1\right\} \quad (25)$$

$$\mathbb{E}\left|\Delta_n \sum_{i=1}^{[T/\Delta_n]} \bar{h}_{u,v}\left(V_{(i-1)\Delta_n}\right) - \int_0^T \bar{h}_{u,v}\left(V_s\right) ds\right| \leq K \min\left\{(u+v)\Delta_n^{1/2}, 1\right\} + \Delta_n \quad (26)$$

$$\mathbb{E}\left|\Delta_n \sum_{i=1}^{[T/\Delta_n]} \left(h_{u,v}\left(\beta_i^n\right) - \bar{h}_{u,v}\left(V_{(i-1)\Delta_n}\right)\right)\right| \leq K\Delta_n^{1/2} \quad (27)$$

$$\mathbb{E}\left[\Delta_n \sum_{i=1}^{[T/\Delta_n]} \left|h_{u,v}\left(\frac{\Delta_i^n X}{\sqrt{\Delta_n}}\right) - h_{u,v}\left(\chi_i^n\right)\right|\right] \leq K\left(\left(u^{1/2} + v^{1/2}\right)\Delta_n^{1/r-1/2}\right)^{r\wedge 1}. \quad (28)$$

Proof. First, note that for any $x, y \in \mathbb{R}_+$,

$$\begin{aligned} \left|h_{u,v}(x) - h_{u,v}(y)\right| \leq & \left|\cos\left(\sqrt{2ux}\right) - \cos\left(\sqrt{2uy}\right)\right| + \left|\cos\left(\sqrt{2vx}\right) - \cos\left(\sqrt{2vy}\right)\right| \\ & + \left|\cos\left(\sqrt{2(u+v)x}\right) - \cos\left(\sqrt{2(u+v)y}\right)\right|. \end{aligned}$$

Hence, following the same calculations as in steps 1 and 4 of the proof of Lemma 3, we have (25) and (28).

Next, note that for any $x, y \in \mathbb{R}_+$,

$$\begin{aligned} & \left|\bar{h}_{u,v}(x) - \bar{h}_{u,v}(y)\right| \\ \leq & \left|e^{-\left(\sqrt{u}+\sqrt{v}\right)^2 x} - e^{-\left(\sqrt{u}+\sqrt{v}\right)^2 y}\right| + \left|e^{-\left(\sqrt{u}-\sqrt{v}\right)^2 x} - e^{-\left(\sqrt{u}-\sqrt{v}\right)^2 y}\right| + \left|e^{-(u+v)x} - e^{-(u+v)y}\right|. \end{aligned}$$

Similarly as in step 2 of the proof of Lemma 3, we have

$$\begin{aligned} & \mathbb{E}\left|\Delta_n \sum_{i=1}^{[T/\Delta_n]} \left(h_{u,v}\left(\beta_i^n\right) - \bar{h}_{u,v}\left(V_{(i-1)\Delta_n}\right)\right)\right| \\ \leq & K \min\left\{\left(\sqrt{u}+\sqrt{v}\right)^2 \Delta_n^{1/2}, 1\right\} + K \min\left\{\left(\sqrt{u}-\sqrt{v}\right)^2 \Delta_n^{1/2}, 1\right\} \\ & + K \min\left\{(u+v)\Delta_n^{1/2}, 1\right\} + \Delta_n, \end{aligned}$$

which implies (26).

Finally, we consider (27). To simplify notations, we denote $\zeta_i^n = h_{u,v}\left(\beta_i^n\right) - \bar{h}_{u,v}\left(V_{(i-1)\Delta_n}\right)$. It is easy to see that $|\zeta_i^n| \leq 4$. Since the $\mathcal{F}_{(i-1)\Delta_n}$-conditional distribution of β_i^n is $\mathcal{N}\left(0, V_{(i-1)\Delta_n}\right)$, $\mathbb{E}\left[h_{u,v}\left(\beta_i^n\right) \middle| \mathcal{F}_{(i-1)\Delta_n}\right] = \bar{h}_{u,v}\left(V_{(i-1)\Delta_n}\right)$. Hence, $\left(\zeta_i^n\right)_{i\geq 1}$ forms a triangular array of martingale differences. We then have

$$\mathbb{E}\left|\Delta_n \sum_{i=1}^{[T/\Delta_n]} \zeta_i^n\right|^2 = \Delta_n^2 \sum_{i=1}^{[T/\Delta_n]} \mathbb{E}\left|\zeta_i^n\right|^2 \leq K\Delta_n.$$

The claim then readily follows. *Q.E.D.*

7.2 Proof of Lemma 1

Part a. The existence of occupation density of $\widetilde{V}_t$ follows directly from Corollary 1 of Theorem IV.70 in Protter (2004). Since $\tilde{a}_t$ and $\widetilde{V}_t$ are locally bounded, we can find a localizing sequence of stopping times $(T_m)_{m\geq 1}$ such that $T_m \leq S_m$ and the stopped processes $\tilde{a}_{t\wedge T_m}$ and $\widetilde{V}_{t\wedge T_m}$ are bounded. We first show

$$\mathbb{E}\left[\sup_{t\leq T\wedge T_m}\left|\tilde{f}_t(x)-\tilde{f}_t(y)\right|^k\right]\leq K|x-y|^{(1-\beta)k\wedge(1/2)}. \tag{29}$$

By Theorem IV.68 of Protter (2004), we have for $x, y \in \tilde{\mathcal{K}}$, $x<y$,

$$\tilde{f}_t(y)-\tilde{f}_t(x)=2\sum_{j=1}^{5}A_t^{(j)}, \tag{30}$$

where

$$\begin{aligned}
A_t^{(1)} &= \left(\widetilde{V}_t-y\right)^+-\left(\widetilde{V}_t-x\right)^++\left(\widetilde{V}_0-x\right)^+-\left(\widetilde{V}_0-y\right)^+,\\
A_t^{(2)} &= \int_0^t 1_{\{x<\widetilde{V}_{s-}\leq y\}}d\widetilde{V}_s,\\
A_t^{(3)} &= \sum_{s\leq t}1_{\{\widetilde{V}_{s-}>y\}}\left[\left(\widetilde{V}_s-x\right)^--\left(\widetilde{V}_s-y\right)^-\right],\\
A_t^{(4)} &= \sum_{s\leq t}1_{\{x<\widetilde{V}_{s-}\leq y\}}\left[\left(\widetilde{V}_s-x\right)^--\left(\widetilde{V}_s-y\right)^+\right],\\
A_t^{(5)} &= \sum_{s\leq t}1_{\{\widetilde{V}_{s-}\leq x\}}\left[\left(\widetilde{V}_s-x\right)^+-\left(\widetilde{V}_s-y\right)^+\right].
\end{aligned}$$

Clearly, for any t,

$$|A_t^{(1)}|\leq 2|x-y|. \tag{31}$$

By (7), we have

$$A_t^{(2)}=\int_0^t 1_{\{x<\widetilde{V}_{s-}\leq y\}}(\tilde{a}_s ds+dB_s)+\int_0^t\int_{\mathbb{R}}1_{\{x<\widetilde{V}_{s-}\leq y\}}\tilde{\delta}(s,z)\,\underline{\mu}(ds,dz). \tag{32}$$

By Hölder's inequality, the boundedness of $\tilde{a}_{t\wedge T_m}$ and condition (ii), we have

$$\mathbb{E}\left[\sup_{t\leq T\wedge T_m}\left|\int_0^t 1_{\{x<\widetilde{V}_{s-}\leq y\}}\tilde{a}_s ds\right|^k\right]\leq K\mathbb{E}\left[\int_0^{T\wedge T_m}1_{\{x<\widetilde{V}_{s-}\leq y\}}|\tilde{a}_s|^k ds\right]\leq K|x-y|. \tag{33}$$

By the Burkholder-Davis-Gundy inequality and Jensen's inequality,

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\leq T\wedge T_m}\left|\int_0^t 1_{\left\{x<\widetilde{V}_{s-}\leq y\right\}}dB_s\right|^k\right] &\leq K\mathbb{E}\left[\left(\int_0^T 1_{\left\{x<\widetilde{V}_{s-}\leq y\right\}}ds\right)^{k/2}\right] \\
&\leq K\mathbb{E}\left[\left|\int_0^T 1_{\left\{x<\widetilde{V}_{s-}\leq y\right\}}ds\right|^{1/2}\right] \\
&\leq K\left(\mathbb{E}\left|\int_0^T 1_{\left\{x<\widetilde{V}_{s-}\leq y\right\}}ds\right|\right)^{1/2} \\
&\leq K\left|x-y\right|^{1/2}. \qquad (34)
\end{aligned}$$

Moreover, condition (i) implies that $\int_{\mathbb{R}}\left(\tilde{\Gamma}_m(z)^k+\tilde{\Gamma}_m(z)\right)\lambda(dz)<\infty$. Then by Lemma 2.1.7 (b) of Jacod and Protter (2012), we have

$$\begin{aligned}
&\mathbb{E}\left[\sup_{t\leq T\wedge T_m}\left|\int_0^t\int_{\mathbb{R}} 1_{\left\{x<\widetilde{V}_{s-}\leq y\right\}}\tilde{\delta}(s,z)\,\underline{\mu}(ds,dz)\right|^k\right] \\
\leq\;& \mathbb{E}\left[\left(\int_0^T\int_{\mathbb{R}} 1_{\left\{x<\widetilde{V}_{s-}\leq y\right\}}\tilde{\Gamma}_m(z)\,\underline{\mu}(ds,dz)\right)^k\right] \\
\leq\;& K\mathbb{E}\left[\int_0^T\int_{\mathbb{R}} 1_{\left\{x<\widetilde{V}_{s-}\leq y\right\}}\tilde{\Gamma}_m(z)^k\lambda(dz)\,ds\right]+K\mathbb{E}\left[\left(\int_0^T\int_{\mathbb{R}} 1_{\left\{x<\widetilde{V}_{s-}\leq y\right\}}\tilde{\Gamma}_m(z)\lambda(dz)\,ds\right)^k\right] \\
\leq\;& K\left|x-y\right|. \qquad (35)
\end{aligned}$$

Combining (32)-(35), we derive

$$\mathbb{E}\left[\sup_{t\leq T\wedge T_m}|A_t^{(2)}|^k\right]\leq K\left|x-y\right|^{1/2}. \qquad (36)$$

Turning to $A_t^{(3)}$ and $A_t^{(5)}$, we first can bound them as follows

$$\begin{aligned}
\sup_{t\leq T\wedge T_m}\left(|A_t^{(3)}|+|A_t^{(5)}|\right) &\leq \int_0^{T\wedge T_m}\int_{\mathbb{R}}(y-x)\wedge\left|\tilde{\delta}(s,z)\right|\underline{\mu}(ds,dz) \\
&\leq (y-x)^{1-\beta}\int_0^T\int_{\mathbb{R}}\left|\tilde{\Gamma}_m(z)\right|^{\beta}\underline{\mu}(ds,dz).
\end{aligned}$$

From here, we readily obtain

$$\begin{aligned}
&\mathbb{E}\left[\sup_{t\leq T\wedge T_m}\left(|A_t^{(3)}|+|A_t^{(5)}|\right)^k\right] \\
\leq\;& K(y-x)^{(1-\beta)k}\,\mathbb{E}\left[\left(\int_0^T\int_{\mathbb{R}}\left|\tilde{\Gamma}_m(z)\right|^{\beta}\underline{\mu}(ds,dz)\right)^k\right] \\
\leq\;& K(y-x)^{(1-\beta)k}\left(\int_{\mathbb{R}}\tilde{\Gamma}_m(z)^{\beta k}\lambda(dz)+K\left(\int_{\mathbb{R}}\tilde{\Gamma}_m(z)^{\beta}\lambda(dz)\right)^k\right) \\
\leq\;& K(y-x)^{(1-\beta)k}, \qquad (37)
\end{aligned}$$

where the second inequality is obtained by using Lemma 2.1.7 (b) of Jacod and Protter (2012), and the last inequality holds because $\int_{\mathbb{R}} \left(\tilde{\Gamma}_m(z)^{\beta k} + \tilde{\Gamma}_m(z)^{\beta}\right) \lambda(dz) < \infty$ under condition (i).

Finally, since $|A_t^{(4)}| \le \int_0^t \int_{\mathbb{R}} 1_{\{x<\widetilde{V}_{s-}\le y\}} |\tilde{\delta}(s,z)| \underline{\mu}(ds,dz)$, the same calculation as in (35) yields

$$\mathbb{E}\left[\sup_{t\le T\wedge T_m} |A_t^{(4)}|^k\right] \le K|x-y|. \tag{38}$$

Combining (30), (31), (36), (37) and (38), we derive (29).

It remains to show $\mathbb{E}\left[\tilde{f}_{T\wedge T_m}(x)^k\right] \le K$ for each $x \in \tilde{\mathcal{K}}$. Since $\widetilde{V}_{t\wedge T_m}$ is bounded, $\tilde{f}_{T\wedge T_m}(x^*) = 0$ for x^* large enough. The assertion then follows from (29) because $\mathbb{E}\left[\tilde{f}_{T\wedge T_m}(x)^k\right] \le K|x-x^*|^{(1-\beta)k\wedge(1/2)} \le K$.

Part b. Denote $\tilde{F}_t(y) = \int_0^t 1_{\{\tilde{V}_s\le y\}} ds$. Then $F_t(x) = \tilde{F}_t(g(x))$. By the chain rule, $F_t(x)$ is differentiable with derivative $f_t(x) = \tilde{f}_t(g(x))\, g'(x)$. Assumption B1 is thus verified. Let $\mathcal{K} \subset (0,\infty)$ be compact. Since g is continuously differentiable, $g'(\cdot)$ is bounded on $\mathcal{K}$. Moreover, the set $g(\mathcal{K})$ is compact; hence by part (a), $\mathbb{E}\left[\left|\tilde{f}_{T\wedge T_m}(g(x))\right|^k\right]$ is bounded for $x \in \mathcal{K}$, yielding $\mathbb{E}\left[|f_{T\wedge T_m}(x)|^k\right] = \mathbb{E}\left[\left|\tilde{f}_{T\wedge T_m}(g(x))\right|^k\right] |g'(x)|^k \le K$. By Jensen's inequality, for any $\varepsilon \in (0, k-1)$, $\sup_{x\in\mathcal{K}} \mathbb{E}\left[|f_{T\wedge T_m}(x)|^{1+\varepsilon}\right] \le K$. This verifies Assumption C1. Moreover, for $x, y \in \mathcal{K}$,

$$\begin{aligned}
& \mathbb{E}\left[|f_{T\wedge T_m}(x) - f_{T\wedge T_m}(y)|^k\right] \\
= {} & \mathbb{E}\left[\left|\tilde{f}_{T\wedge T_m}(g(x))\, g'(x) - \tilde{f}_{T\wedge T_m}(g(y))\, g'(y)\right|^k\right] \\
\le {} & K\mathbb{E}\left[\left|\tilde{f}_{T\wedge T_m}(g(x)) - \tilde{f}_{T\wedge T_m}(g(y))\right|^k\right] + K\mathbb{E}\left[\left|\tilde{f}_{T\wedge T_m}(g(y))\right|^k |g'(x) - g'(y)|^k\right] \\
\le {} & K|g(x) - g(y)|^{(1-\beta)k\wedge(1/2)} + K|x-y|^{\bar{\gamma}k} \\
\le {} & K|x-y|^{(1-\beta)k\wedge(1/2)} + K|x-y|^{\bar{\gamma}k}.
\end{aligned}$$

Hence, for any $\varepsilon \in (0, k-1)$, by Jensen's inequality,

$$\mathbb{E}\left[|f_{T\wedge T_m}(x) - f_{T\wedge T_m}(y)|^{1+\varepsilon}\right] \le K|x-y|^{(1-\beta)\wedge\frac{1}{2k}} + K|x-y|^{\bar{\gamma}}.$$

By setting $\tilde{\gamma} = (1-\beta) \wedge \frac{1}{2k} \wedge \bar{\gamma}$ and picking any $\varepsilon \in (0, \min\{\tilde{\gamma}, k-1\})$, we verify Assumption C2 for the process V_t. *Q.E.D.*

7.3 Proof of Theorem 1

Part a. By a standard localization procedure, we can suppose that Assumption SA holds without loss of generality. For any $M > 0$, we have $\int_0^M |w(u)|\, du < \infty$. It is then easy to see that the mapping $f \mapsto \int_0^M f(u)\, w(u)\, du$ is a continuous mapping on the space of continuous functions equipped with the uniform metric. By the continuous mapping theorem and Theorem 1 of Todorov and Tauchen (2012b),

$$\Delta_n^{-1/2} \int_0^M \left(\widehat{\mathcal{L}}_{T,n}(u) - \mathcal{L}_T(u) \right) w(u)\, du \xrightarrow{\mathcal{L}\text{-}s} \int_0^M \xi(u)\, w(u)\, du. \tag{39}$$

Observe that $\mathbb{E}\,|\xi(u)| \leq \sqrt{\mathbb{E}[\frac{1}{2}\int_0^T (1-e^{-2uV_s})^2\, ds]} \leq Ku$. Hence, $\mathbb{E}[\int_0^\infty |\xi(u)\, w(u)|\, du] \leq K\int_0^\infty u\,|w(u)|\, du < \infty$. By dominated convergence, as $M \to \infty$,

$$\int_0^M \xi(u)\, w(u)\, du \xrightarrow{\mathbb{L}^1} \int_0^\infty \xi(u)\, w(u)\, du. \tag{40}$$

Without loss of generality, suppose that $\Delta_n \leq 1$. By Lemma 3, we have for all $u \geq 1$, $\Delta_n^{-1/2}\mathbb{E}\left|\widehat{\mathcal{L}}_{T,n}(u) - \mathcal{L}_T(u)\right| \leq Ku$. Therefore, for $M \geq 1$,

$$\mathbb{E}\left|\Delta_n^{-1/2}\int_M^\infty \left(\widehat{\mathcal{L}}_{T,n}(u) - \mathcal{L}_T(u)\right) w(u)\, du\right| \leq K\int_M^\infty u\,|w(u)|\, du,$$

where the constant K does not depend on M. By the dominated convergence theorem and Markov's inequality, we have, for any $\varepsilon > 0$,

$$\lim_{M\to\infty}\limsup_{n\to\infty}\mathbb{P}\left(\left|\Delta_n^{-1/2}\int_M^\infty \left(\widehat{\mathcal{L}}_{T,n}(u) - \mathcal{L}_T(u)\right) w(u)\, du\right| > \varepsilon\right) = 0. \tag{41}$$

Combining (39)-(41), we readily derive the claim.

Part b. Recall the notations in Lemma 4 and observe that $\Sigma_T(u,v) = \int_0^T \bar{h}_{u,v}(V_s)\, ds$. By Lemma 4, we have $\mathbb{E}\left|\widehat{\Sigma}_{T,n}(u,v) - \Sigma_T(u,v)\right| \leq K(u\vee 1)(v \vee 1)\Delta_n^{1/2}$. Hence,

$$\begin{aligned} &\mathbb{E}\left[\int_0^\infty\int_0^\infty \left|\widehat{\Sigma}_{T,n}(u,v) - \Sigma_T(u,v)\right| |w(u)\, w(v)|\, dudv\right] \\ \leq\;& K\Delta_n^{1/2}\left(\int_0^\infty (u\vee 1)\,|w(u)|\, du\right)^2 \to 0. \end{aligned}$$

The assertion readily follows. *Q.E.D.*

7.4 Proof of Theorem 2, 3 and Corollary 1

Proof of Theorem 2. By localization, we can suppose that V_t is bounded above and away from zero. We can also strengthen Assumptions B2 and B3 as follows: for any compact $\mathcal{K} \subset (0,\infty)$ and any $x, y \in \mathcal{K}$,

$$\mathbb{E}\,|f_T(x)| \leq K, \quad \mathbb{E}\left[\sup_{t\leq T}|f_t(x) - f_t(y)|\right] \leq K\,|x-y|^\gamma;$$

this is without loss, because otherwise we can restrict our calculations in the event $\{T \leq T_m\}$, which charges an arbitrarily small probability mass for m sufficiently large. The proof proceeds via several steps.

Step 1. From Kryzhniy (2003b), we have

$$F_{t,R}(x) = \frac{2}{\pi} \int_0^\infty F_t(xu) \sqrt{u} \frac{\sin(R \ln u)}{u^2 - 1} du.$$

With a change of variable, we have the following decomposition:

$$F_{t,R}(x) - F_t(x) = F_t(x) \left(\frac{2}{\pi} \int_{-\infty}^\infty \frac{e^{3z/2} \sin(Rz)}{e^{2z} - 1} dz - 1 \right) + \frac{2}{\pi} \int_0^\infty G_t(z; x) \sin(Rz)\, dz, \qquad (42)$$

where we set

$$\begin{aligned} g_t(z; x) &= (F_t(xe^z) - F_t(x))\, h(z), \quad h(z) = \frac{e^{3z/2}}{e^{2z} - 1}, \\ G_t(z; x) &= g_t(z; x) - g_t(-z; x). \end{aligned}$$

The first term in (42) can be bounded as follows. By direct integration, we have

$$\frac{2}{\pi} \int_{-\infty}^\infty \frac{e^{3z/2} \sin(Rz)}{e^{2z} - 1} dz = \tanh(\pi R).$$

Hence,

$$\sup_{t \leq T, x \geq 0, \omega \in \Omega} \left| F_t(x) \left(\frac{2}{\pi} \int_{-\infty}^\infty \frac{e^{3z/2} \sin(Rz)}{e^{2z} - 1} dz - 1 \right) \right| = O\left(e^{-2\pi R}\right). \qquad (43)$$

Bounding the second term in (42) is the task below.

Step 2. Below, we denote $a = \pi/2R$ and, without loss of generality, we suppose that $R \geq 1$. In this step, we show that

$$\mathbb{E} \left| \int_0^a G_T(z; x) \sin(Rz)\, dz \right| \leq K R^{-1}. \qquad (44)$$

Indeed, by Assumption B2,

$$\begin{aligned} \mathbb{E} \left| \int_0^a G_T(z; x) \sin(Rz)\, dz \right| &= \mathbb{E} \left| \int_{-a}^a (F_T(xe^z) - F_T(x)) \frac{e^{3z/2}}{e^{2z} - 1} \sin(Rz)\, dz \right| \\ &\leq \mathbb{E} \left[\int_{-a}^a |F_T(xe^z) - F_T(x)| \frac{e^{3z/2}}{|e^{2z} - 1|} dz \right] \\ &\leq K \int_{-a}^a \frac{e^{3z/2} |e^z - 1|}{|e^{2z} - 1|} dz \\ &\leq K a, \end{aligned}$$

which implies (44).

Step 3. For each $k \geq 0$, we denote $a_{k,R} = a + 2\pi k/R$. Let $N_R = \min\{k \in \mathbb{N} : a_{k,R} \geq 1\}$. Note that for $0 \leq k \leq N_R$, we have $a \leq a_{k,R} \leq 3\pi$. In this step, we show that

$$\mathbb{E}\left|\int_a^{a+2\pi N_R/R} G_T(z;x)\sin(Rz)\,dz\right| \leq KR^{-1-\gamma}\ln(R) + KR^{-1}. \tag{45}$$

We denote $G_{t,z}(z;x) = \partial G_t(z;x)/\partial z$ and $g_{t,z}(z;x) = \partial g_t(z;x)/\partial z$. Note that $G_{t,z}(z;x) = g_{t,z}(z;x) + g_{t,z}(-z;x)$. To show (45), we first note that for each $k \geq 1$,

$$\begin{aligned}
&\int_{a_{k-1,R}}^{a_{k,R}} G_T(z;x)\sin(Rz)\,dz \\
=\;& \int_{a_{k-1,R}}^{a_{k,R}-\pi/R} \left(G_T(z;x) - G_T(z+\pi/R;x)\right)\sin(Rz)\,dz \\
=\;& R^{-1}\int_{a_{k-1,R}}^{a_{k,R}-\pi/R} \left(G_{T,z}(z;x) - G_{T,z}(z+\pi/R;x)\right)\cos(Rz)\,dz,
\end{aligned}$$

where the first equality is obtained by a change of variable and the second equality follows an integration by parts, using that $\cos(Rz) = 0$ for $z = a_{k-1,R}$ or $a_{k,R} - \pi/R$. Therefore,

$$\mathbb{E}\left|\int_{a_{k-1,R}}^{a_{k,R}} G_T(z;x)\sin(Rz)\,dz\right| \leq R^{-1}\int_{a_{k-1,R}}^{a_{k,R}-\pi/R} \mathbb{E}\left|G_{T,z}(z;x) - G_{T,z}(z+\pi/R;x)\right|dz. \tag{46}$$

To simplify notations, let $\phi_t(x) = xf_t(x)$. Then $g_{t,z}(z;x) = \phi_t(xe^z)h(z) + (F_t(xe^z) - F_t(x))h'(z)$. Hence, for any $y, z \in \mathbb{R}$,

$$\begin{aligned}
&\left|g_{t,z}(z;x) - g_{t,z}(y;x)\right| \\
\leq\;& \left|\phi_t(xe^z)h(z) - \phi_t(xe^y)h(y)\right| \\
&+ \left|(F_t(xe^z) - F_t(x))h'(z) - (F_t(xe^y) - F_t(x))h'(y)\right| \\
\leq\;& \left|\phi_t(xe^z) - \phi_t(xe^y)\right|\cdot h(z) + \phi_t(xe^y)\cdot\left|h(y) - h(z)\right| \\
&+ \left|F_t(xe^z) - F_t(xe^y)\right|\cdot\left|h'(z)\right| + \left|F_t(xe^y) - F_t(x)\right|\cdot\left|h'(y) - h'(z)\right|.
\end{aligned}$$

Moreover, under Assumptions B2 and B3, it is easy to see that if either (i) $z, y \in [-3\pi, -a]$ and $y = z - \pi/R$; or (ii) $z, y \in [a, 3\pi]$ and $y = z + \pi/R$, then

$$\begin{aligned}
&\mathbb{E}\left|\phi_T(xe^z) - \phi_T(xe^y)\right| \leq KR^{-\gamma}, && |h(z)| \leq K\left|e^z - 1\right|^{-1}, \\
&\mathbb{E}\left|\phi_T(xe^y)\right| \leq K, && |h(y) - h(z)| \leq KR^{-1}\left(e^z - 1\right)^{-2}, \\
&\mathbb{E}\left|F_T(xe^z) - F_T(xe^y)\right| \leq KR^{-1}, && |h'(z)| \leq K\left(e^z - 1\right)^{-2}, \\
&\mathbb{E}\left|F_T(xe^y) - F_T(x)\right| \leq K|z|, && |h'(y) - h'(z)| \leq KR^{-1}\left|e^z - 1\right|^{-3}.
\end{aligned}$$

Hence, for $z \in [a_{k-1,R}, a_{k,R}]$, $1 \leq k \leq N_R$, we have

$$\mathbb{E}\left|G_{T,z}(z;x) - G_{T,z}(z+\pi/R;x)\right| \leq KR^{-\gamma}(e^z-1)^{-1} + KR^{-1}(e^z-1)^{-2},$$

and by (46), we derive for each $1 \leq k \leq N_R$,

$$\mathbb{E}\left|\int_{a_{k-1,R}}^{a_{k,R}} G_T(z;x)\sin(Rz)\,dz\right| \leq KR^{-2-\gamma}\left(e^{a_{k-1,R}}-1\right)^{-1} + KR^{-3}\left(e^{a_{k-1,R}}-1\right)^{-2}. \tag{47}$$

Finally, observing $\mathbb{E}\left|\int_a^{a+2\pi N_R/R} G_T(z;x)\sin(Rz)\,dz\right| \leq \sum_{k=1}^{N_R}\mathbb{E}\left|\int_{a_{k-1,R}}^{a_{k,R}} G_T(z;x)\sin(Rz)\,dz\right|$, we readily derive (45) by using (47).

Step 4. Let M be a constant such that $M \geq a + 2\pi N_R/R$ and $\cos(MR) = 0$. Integration by parts yields

$$\int_{a+2\pi N_R/R}^{M} G_T(z;x)\sin(Rz)\,dz = R^{-1}\int_{a+2\pi N_R/R}^{M} G_{T,z}(z;x)\cos(Rz)\,dz. \tag{48}$$

Since V_t is bounded above and away from zero, $f_T(\cdot)$ is supported on some compact subset of $(0,\infty)$. Then under Assumption B2, for $z \geq 1$,

$$\begin{aligned}
\mathbb{E}\left|\phi_T(xe^z)\right| + \mathbb{E}\left|F_T(xe^z) - F_T(x)\right| &\leq K, \\
\mathbb{E}\left|\phi_T\left(xe^{-z}\right)\right| + \mathbb{E}\left|F_T\left(xe^{-z}\right) - F_T(x)\right| &\leq K.
\end{aligned}$$

Moreover, $|h(z)| + |h'(z)| \leq Ke^{-z/2}$ and $|h(-z)| + |h'(-z)| \leq Ke^{-3z/2}$. Hence, for $z \geq 1$,

$$\begin{aligned}
\mathbb{E}\left|g_{T,z}(z;x)\right| &= \mathbb{E}\left|\phi_T(xe^z)h(z) + (F_T(xe^z) - F_T(x))h'(z)\right| \leq Ke^{-z/2}, \\
\mathbb{E}\left|g_{T,z}(-z;x)\right| &= \mathbb{E}\left|\phi_T\left(xe^{-z}\right)h(-z) + \left(F_T\left(xe^{-z}\right) - F_T(x)\right)h'(-z)\right| \leq Ke^{-3z/2}.
\end{aligned}$$

Therefore, $\mathbb{E}\left|G_{T,z}(z;x)\right| \leq Ke^{-z/2}$. Since $a + 2\pi N_R/R \geq 1$, we use (48) to derive

$$\mathbb{E}\left|\int_{a+2\pi N_R/R}^{M} G_T(z;x)\sin(Rz)\,dz\right| \leq KR^{-1}\int_{a+2\pi N_R/R}^{M} e^{-z/2}dz \leq KR^{-1}, \tag{49}$$

where the constant K does not depend on M.

Step 5. Now, note that

$$\begin{aligned}
\left|\int_M^\infty g_T(z;x)\sin(Rz)\,dz\right| &\leq \int_M^\infty \left|(F_T(xe^z) - F_T(x))\frac{e^{3z/2}}{e^{2z}-1}\right| dz \\
&\leq K\int_M^\infty e^{-z/2}dz \\
&\leq Ke^{-M/2},
\end{aligned}$$

and

$$\begin{aligned}\left|\int_M^\infty g_T(-z;x)\sin(Rz)\,dz\right| &\leq \int_M^\infty \left|\left(F_T\left(xe^{-z}\right)-F_T(x)\right)\frac{e^{-3z/2}}{e^{-2z}-1}\right|dz \\ &\leq K\int_M^\infty e^{-3z/2}dz \\ &\leq Ke^{-3M/2}.\end{aligned}$$

Therefore, $\left|\int_M^\infty G_T(z;x)\sin(Rz)\,dz\right| \leq Ke^{-M/2}$. Taking $M \geq 3\ln(R)$, we derive

$$\left|\int_M^\infty G_T(z;x)\sin(Rz)\,dz\right| \leq KR^{-3/2}. \tag{50}$$

Step 6. Combining (44), (45), (49) and (50), we have

$$\mathbb{E}\left|\int_0^\infty G_T(z;x)\sin(Rz)\,dz\right| \leq KR^{-1-\gamma}\ln(R)+KR^{-1}. \tag{51}$$

The assertion of the theorem then follows from (42), (43) and (51). *Q.E.D.*

Proof of Theorem 3. With a standard localization procedure, we can suppose that Assumption SA holds without loss of generality, and further that $R_n \geq 2$ and $\Delta_n \leq 1$. It is easy to see that

$$\sup_{x\in\mathcal{K}}\left|\widehat{F}_{T,n,R_n}(x)-F_{T,R_n}(x)\right| \leq \sum_{j=1}^4 \zeta_{j,n}, \tag{52}$$

where, with the notations of Lemma 3 and $\Pi^*(R,u)=\sup_{x\in\mathcal{K}}|\Pi(R,ux)|$, we set

$$\begin{aligned}\zeta_{1,n} &= \int_0^\infty \left|\Delta_n\sum_{i=1}^{[T/\Delta_n]}\left(\cos\left(\sqrt{2u}\chi_i^n\right)-\cos\left(\sqrt{2u}\beta_i^n\right)\right)\right|\frac{\Pi^*(R_n,u)}{u}du, \\ \zeta_{2,n} &= \int_0^\infty \left|\Delta_n\sum_{i=1}^{[T/\Delta_n]}\exp\left(-uV_{(i-1)\Delta_n}\right)-\int_0^T\exp\left(-uV_s\right)ds\right|\frac{\Pi^*(R_n,u)}{u}du, \\ \zeta_{3,n} &= \int_0^\infty \left|\Delta_n\sum_{i=1}^{[T/\Delta_n]}\left(\cos\left(\sqrt{2u}\beta_i^n\right)-\exp\left(-uV_{(i-1)\Delta_n}\right)\right)\right|\frac{\Pi^*(R_n,u)}{u}du, \\ \zeta_{4,n} &= \int_0^\infty \left|\Delta_n\sum_{i=1}^{[T/\Delta_n]}\left(\cos\left(\sqrt{2u}\frac{\Delta_i^nX}{\sqrt{\Delta_n}}\right)-\cos\left(\sqrt{2u}\chi_i^n\right)\right)\right|\frac{\Pi^*(R_n,u)}{u}du.\end{aligned}$$

By Lemma 2, given any $\eta_1,\eta_2\in[0,1/2)$, we have for all $u>0$,

$$\begin{aligned}\Pi^*(R_n,u) &\leq K\exp\left(\frac{\pi R_n}{2}\right)\sup_{x\in\mathcal{K}}\min\left\{(xu)^{\eta_1},R_n(xu)^{-1},R_n^2(xu)^{-1-\eta_2}\right\} \\ &\leq K\exp\left(\frac{\pi R_n}{2}\right)\min\left\{u^{\eta_1},R_nu^{-1},R_n^2u^{-1-\eta_2}\right\},\end{aligned} \tag{53}$$

where the second inequality holds because $\mathcal{K}$ is bounded above and away from zero by assumption.

Below, for any Borel function $h:\mathbb{R}_{+}\mapsto\mathbb{R}_{+}$, and u', $u''\in\mathbb{R}_{+}$, we set the convention that $\int_{u'}^{u''}h(u)\,du=0$ whenever $u'\geq u''$. By (21) and (53) (the latter applied with $\eta_1=0$ and $\eta_2=\eta$, for η being the constant in the statement of the theorem), we have

$$\mathbb{E}\left|\zeta_{1,n}\right|\leq K\exp\left(\frac{\pi R_n}{2}\right)\int_0^\infty \min\left\{u^{1/2}\Delta_n^{1/2},1\right\}\min\left\{u^{-1},R_nu^{-2},R_n^2u^{-2-\eta}\right\}du.$$

Note that

$$\begin{aligned}
&\int_0^\infty \min\left\{u^{1/2}\Delta_n^{1/2},1\right\}\min\left\{u^{-1},R_nu^{-2},R_n^2u^{-2-\eta}\right\}du\\
\leq\;& \Delta_n^{1/2}\int_0^{R_n}u^{-1/2}du+\Delta_n^{1/2}R_n\int_{R_n}^{R_n^{1/\eta}}u^{-3/2}du\\
&+\Delta_n^{1/2}R_n^2\int_{R_n^{1/\eta}}^{\Delta_n^{-1}}u^{-3/2-\eta}du+R_n^2\int_{\Delta_n^{-1}}^\infty u^{-2-\eta}du\\
\leq\;& KR_n^{1/2}\Delta_n^{1/2}+R_n^{1-\frac{1}{2\eta}}\Delta_n^{1/2}+KR_n^2\Delta_n^{1+\eta}\\
\leq\;& KR_n^{1/2}\Delta_n^{1/2}+KR_n^2\Delta_n^{1+\eta}.
\end{aligned}$$

Hence,

$$\mathbb{E}\left|\zeta_{1,n}\right|\leq K\exp\left(\frac{\pi R_n}{2}\right)\left(R_n^{1/2}\Delta_n^{1/2}+R_n^2\Delta_n^{1+\eta}\right). \tag{54}$$

By (22) and (53), for $\eta_1\in(0,1/2)$, we have

$$\begin{aligned}
\mathbb{E}\left|\zeta_{2,n}\right|\;\leq\;& K\exp\left(\frac{\pi R_n}{2}\right)\int_0^\infty \min\left\{u\Delta_n^{1/2},1\right\}\min\left\{u^{-1},R_nu^{-2},R_n^2u^{-2-\eta}\right\}du\\
&+K\Delta_n\exp\left(\frac{\pi R_n}{2}\right)\int_0^\infty \min\left\{u^{\eta_1-1},R_nu^{-2},R_n^2u^{-2-\eta}\right\}du
\end{aligned}$$

Moreover, note that

$$\begin{aligned}
&\int_0^\infty \min\left\{u^{\eta_1-1},R_nu^{-2},R_n^2u^{-2-\eta}\right\}du\\
\leq\;& \int_0^{R_n^{1/(1+\eta_1)}}u^{\eta_1-1}du+R_n\int_{R_n^{1/(1+\eta_1)}}^{R_n^{1/\eta}}u^{-2}du+R_n^2\int_{R_n^{1/\eta}}^\infty u^{-2-\eta}du\\
\leq\;& KR_n^{\eta_1/(1+\eta_1)}+KR_n^{1-1/\eta},
\end{aligned}$$

and

$$\begin{aligned}
&\int_0^\infty \min\left\{u\Delta_n^{1/2},1\right\}\min\left\{u^{-1},R_nu^{-2},R_n^2u^{-2-\eta}\right\}du\\
\leq\;& R_n\Delta_n^{1/2}+R_n\Delta_n^{1/2}\int_{R_n}^{R_n^{1/\eta}}u^{-1}du+R_n^2\Delta_n^{1/2}\int_{R_n^{1/\eta}}^{\Delta_n^{-1/2}}u^{-1-\eta}du+R_n^2\int_{\Delta_n^{-1/2}}^\infty u^{-2-\eta}du\\
\leq\;& KR_n\ln\left(R_n\right)\Delta_n^{1/2}+KR_n^2\Delta_n^{(1+\eta)/2}.
\end{aligned}$$

Hence,

$$\mathbb{E}\left|\zeta_{2,n}\right| \leq K \exp\left(\frac{\pi R_n}{2}\right)\left(R_n \ln\left(R_n\right)\Delta_n^{1/2} + R_n^2 \Delta_n^{(1+\eta)/2}\right). \tag{55}$$

By (23) and (53), we have

$$\mathbb{E}\left|\zeta_{3,n}\right| \leq K\Delta_n^{1/2} \exp\left(\frac{\pi R_n}{2}\right)\int_0^\infty \min\{u,1\}\min\left\{u^{-1}, R_n u^{-2}, R_n^2 u^{-2-\eta}\right\}du.$$

Note that

$$\begin{aligned}
&\int_0^\infty \min\{u,1\}\min\left\{u^{-1}, R_n u^{-2}, R_n^2 u^{-2-\eta}\right\}du \\
\leq\ & 1 + \int_1^{R_n} u^{-1}du + R_n \int_{R_n}^{R_n^{1/\eta}} u^{-2}du + R_n^2 \int_{R_n^{1/\eta}}^\infty u^{-2-\eta}du \\
\leq\ & K + K\ln\left(R_n\right) + K R_n^{1-1/\eta} \\
\leq\ & K \ln\left(R_n\right).
\end{aligned}$$

Hence,

$$\mathbb{E}\left|\zeta_{3,n}\right| \leq K \exp\left(\frac{\pi R_n}{2}\right)\ln\left(R_n\right)\Delta_n^{1/2}. \tag{56}$$

Now, consider $\zeta_{4,n}$. We denote $\tilde{r} = r \wedge 1$. By (24) and (53),

$$\mathbb{E}\left|\zeta_{4,n}\right| \leq K\Delta_n^{\tilde{r}(1/r-1/2)} \exp\left(\frac{\pi R_n}{2}\right)\int_0^\infty u^{\tilde{r}/2}\min\left\{u^{-1}, R_n u^{-2}, R_n^2 u^{-2-\eta}\right\}du.$$

Note that

$$\begin{aligned}
&\int_0^\infty u^{\tilde{r}/2}\min\left\{u^{-1}, R_n u^{-2}, R_n^2 u^{-2-\eta}\right\}du \\
\leq\ & \int_0^{R_n} u^{\tilde{r}/2-1}du + R_n \int_{R_n}^{R_n^{1/\eta}} u^{\tilde{r}/2-2}du + R_n^2 \int_{R_n^{1/\eta}}^\infty u^{\tilde{r}/2-2-\eta}du \\
\leq\ & K R_n^{\tilde{r}/2} + K R_n^{1-(1-\tilde{r}/2)/\eta} \\
\leq\ & K R_n^{\tilde{r}/2}.
\end{aligned}$$

Hence,

$$\mathbb{E}\left|\zeta_{4,n}\right| \leq K \exp\left(\frac{\pi R_n}{2}\right) R_n^{\tilde{r}/2}\Delta_n^{\tilde{r}(1/r-1/2)}. \tag{57}$$

Finally, we combine (52) and (54)-(57) to derive the assertion of the theorem. *Q.E.D.*

Proof of Corollary 1. The assertion follows directly from Theorems 2 and 3. *Q.E.D.*

7.5 Proof of Theorem 4

Let $(T_m)_{m\geq 1}$ be the localizing sequence of stopping times as in Assumption C. By Assumption A2, we can suppose that the stopped process $(V_{t\wedge T_m})_{t\geq 0}$ takes values in $\mathcal{K}_m = [c_m, C_m]$ for some constants $C_m > c_m > 0$ without loss of generality. By enlarging $\mathcal{K}_m$ if necessary, we suppose that $\mathcal{K} \subseteq \mathcal{K}_m$ without loss. Let $M_m = \ln(C_m/c_m)$. Observe that

$$x, y \in \mathcal{K},\ |z| \geq M_m \Rightarrow F_{T\wedge T_m}(xe^z) - F_{T\wedge T_m}(ye^z) = 0. \tag{58}$$

Hence, for any $x, y \in \mathcal{K}$,

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_{-\infty}^{\infty}\left(F_{T\wedge T_m}(xe^z) - F_{T\wedge T_m}(x) - \left(F_{T\wedge T_m}(ye^z) - F_{T\wedge T_m}(y)\right)\right)\frac{e^{3z/2}\sin(R_n z)}{e^{2z}-1}dz\right|^{1+\varepsilon}\right] \\
\leq\ & K\mathbb{E}\left[\left|F_{T\wedge T_m}(x) - F_{T\wedge T_m}(y)\right|^{1+\varepsilon}\left|\int_{(-\infty,-M_m]\cup[M_m,\infty)}\frac{e^{3z/2}\sin(R_n z)}{e^{2z}-1}dz\right|^{1+\varepsilon}\right] \\
&+K\mathbb{E}\left[\left|\int_{-M_m}^{M_m}\Big(F_{T\wedge T_m}(xe^z) - F_{T\wedge T_m}(x)\right.\right. \\
&\left.\left.-\left(F_{T\wedge T_m}(ye^z) - F_{T\wedge T_m}(y)\right)\Big)\frac{e^{3z/2}\sin(R_n z)}{e^{2z}-1}dz\right|^{1+\varepsilon}\right]
\end{aligned} \tag{59}$$

By Assumption C1 and Hölder's inequality,

$$\begin{aligned}
&\mathbb{E}\left[\left|F_{T\wedge T_m}(x) - F_{T\wedge T_m}(y)\right|^{1+\varepsilon}\left(\int_{(-\infty,-M_m]\cup[M_m,\infty)}\frac{e^{3z/2}\sin(R_n z)}{e^{2z}-1}dz\right)^{1+\varepsilon}\right] \\
\leq\ & K_m\mathbb{E}\left[\left|F_{T\wedge T_m}(x) - F_{T\wedge T_m}(y)\right|^{1+\varepsilon}\right] \\
\leq\ & K_m\left|x-y\right|^{1+\varepsilon}.
\end{aligned} \tag{60}$$

Note that for any $x, y \in \mathcal{K}$ and $z \in \mathbb{R}$,

$$\begin{aligned}
&\left|F_{T\wedge T_m}(xe^z) - F_{T\wedge T_m}(x) - \left(F_{T\wedge T_m}(ye^z) - F_{T\wedge T_m}(y)\right)\right| \\
&= \left|\int_x^y f_{T\wedge T_m}(v)\,dv - \int_x^y e^z f_{T\wedge T_m}(ve^z)\,dv\right| \\
&\leq \int_x^y f_{T\wedge T_m}(ve^z)\left|e^z - 1\right|dv + \int_x^y \left|f_{T\wedge T_m}(ve^z) - f_{T\wedge T_m}(v)\right|dv.
\end{aligned} \tag{61}$$

By (61) and Hölder's inequality,

$$
\begin{aligned}
& \mathbb{E}\left[\left|\int_{-M_m}^{M_m} \left(F_{T\wedge T_m}\left(xe^{z}\right) - F_{T\wedge T_m}\left(x\right) - \left(F_{T\wedge T_m}\left(ye^{z}\right) - F_{T\wedge T_m}\left(y\right)\right)\right) \frac{e^{3z/2}\sin\left(R_n z\right)}{e^{2z}-1} dz\right|^{1+\varepsilon}\right] \\
\leq \;& K_m \mathbb{E}\left[\int_{-M_m}^{M_m} \left(\int_x^y f_{T\wedge T_m}\left(ve^{z}\right)\left|e^{z}-1\right| dv\right)^{1+\varepsilon} \left|\frac{e^{3z/2}}{e^{2z}-1}\right|^{1+\varepsilon} dz\right] \\
& + K_m \mathbb{E}\left[\int_{-M_m}^{M_m} \left(\int_x^y \left|f_{T\wedge T_m}\left(ve^{z}\right) - f_{T\wedge T_m}\left(v\right)\right| dv\right)^{1+\varepsilon} \left|\frac{e^{3z/2}}{e^{2z}-1}\right|^{1+\varepsilon} dz\right].
\end{aligned} \tag{62}
$$

The two terms on the majorant side of the above display can be further bounded as follows. By Hölder's inequality and Assumption C1

$$
\begin{aligned}
& \mathbb{E}\left[\int_{-M_m}^{M_m} \left(\int_x^y f_{T\wedge T_m}\left(ve^{z}\right)\left|e^{z}-1\right| dv\right)^{1+\varepsilon} \left|\frac{e^{3z/2}}{e^{2z}-1}\right|^{1+\varepsilon} dz\right] \\
\leq \;& K_m \mathbb{E}\left[\int_{-M_m}^{M_m} \left(\int_x^y f_{T\wedge T_m}\left(ve^{z}\right) dv\right)^{1+\varepsilon} dz\right] \\
\leq \;& K_m \left|x-y\right|^{1+\varepsilon}.
\end{aligned} \tag{63}
$$

Moreover, by Hölder's inequality and Assumption C2,

$$
\begin{aligned}
& \mathbb{E}\left[\int_{-M_m}^{M_m} \left(\int_x^y \left|f_{T\wedge T_m}\left(ve^{z}\right) - f_{T\wedge T_m}\left(v\right)\right| dv\right)^{1+\varepsilon} \left|\frac{e^{3z/2}}{e^{2z}-1}\right|^{1+\varepsilon} dz\right] \\
\leq \;& K_m \int_{-M_m}^{M_m} \mathbb{E}\left[\left(\int_x^y \left|f_{T\wedge T_m}\left(ve^{z}\right) - f_{T\wedge T_m}\left(v\right)\right| dv\right)^{1+\varepsilon}\right] \left|\frac{1}{e^{z}-1}\right|^{1+\varepsilon} dz \\
\leq \;& K_m \left|x-y\right|^{\varepsilon} \int_{-M_m}^{M_m} \mathbb{E}\left[\int_x^y \left|f_{T\wedge T_m}\left(ve^{z}\right) - f_{T\wedge T_m}\left(v\right)\right|^{1+\varepsilon} dv\right] \left|\frac{1}{e^{z}-1}\right|^{1+\varepsilon} dz \\
\leq \;& K_m \left|x-y\right|^{1+\varepsilon} \int_{-M_m}^{M_m} \frac{1}{\left|e^{z}-1\right|^{1+\varepsilon-\tilde{\gamma}}} dz \\
\leq \;& K_m \left|x-y\right|^{1+\varepsilon}.
\end{aligned} \tag{64}
$$

By (63) and (64), we see that the left-hand side of (62) can be further bounded by $K_m \left|x-y\right|^{1+\varepsilon}$. Combining this estimate with (59) and (60), we get

$$
\mathbb{E}\left[\left|\int_{-\infty}^{\infty} \left(F_{T\wedge T_m}\left(xe^{z}\right) - F_{T\wedge T_m}\left(x\right) - \left(F_{T\wedge T_m}\left(ye^{z}\right) - F_{T\wedge T_m}\left(y\right)\right)\right) \frac{e^{3z/2}\sin\left(R_n z\right)}{e^{2z}-1} dz\right|^{1+\varepsilon}\right] \leq K_m \left|x-y\right|^{1+\varepsilon}. \tag{65}
$$

Next, observe that by Assumption C1,

$$
\begin{aligned}
& \mathbb{E}\left[\left|\int_{-\infty}^{\infty}\left(F_{T\wedge T_m}\left(xe^z\right)-F_{T\wedge T_m}\left(x\right)\right)\frac{e^{3z/2}\sin\left(R_n z\right)}{e^{2z}-1}dz\right|^{1+\varepsilon}\right] \\
\leq\ & K\mathbb{E}\left[\left|\int_{-1}^{1}\left(F_{T\wedge T_m}\left(xe^z\right)-F_{T\wedge T_m}\left(x\right)\right)\frac{e^{3z/2}\sin\left(R_n z\right)}{e^{2z}-1}dz\right|^{1+\varepsilon}\right] \\
& +K\mathbb{E}\left[\left|\int_{\mathbb{R}\backslash[-1,1]}\left(F_{T\wedge T_m}\left(xe^z\right)-F_{T\wedge T_m}\left(x\right)\right)\frac{e^{3z/2}\sin\left(R_n z\right)}{e^{2z}-1}dz\right|^{1+\varepsilon}\right] \\
\leq\ & K\mathbb{E}\left[\int_{-1}^{1}\left|F_{T\wedge T_m}\left(xe^z\right)-F_{T\wedge T_m}\left(x\right)\right|^{1+\varepsilon}\left|\frac{1}{e^z-1}\right|^{1+\varepsilon}dz\right]+K \\
\leq\ & K\mathbb{E}\left[\int_{-1}^{1}\left|\int_{[x\wedge xe^z,x\vee xe^z]}f_{T\wedge T_m}\left(v\right)dv\right|^{1+\varepsilon}\left|\frac{1}{e^z-1}\right|^{1+\varepsilon}dz\right]+K \\
\leq\ & K. \qquad (66)
\end{aligned}
$$

Because of (65) and (66), we can apply Theorem 20 in Ibragimov and Has'minskii (1981) to show that the collection of processes

$$\int_{-\infty}^{\infty}\left(F_{T\wedge T_m}\left(xe^z\right)-F_{T\wedge T_m}\left(x\right)\right)\frac{e^{3z/2}\sin\left(R_n z\right)}{e^{2z}-1}dz,\quad x\in\mathcal{K},\ n\geq 1,$$

is stochastically equicontinuous. By (43), we see that $F_{T\wedge T_m,R_n}(x)-F_{T\wedge T_m}(x)$, $x\in\mathcal{K}$, $n\geq 1$, are also stochastically equicontinuous. Combining this result with Theorem 3, we deduce that $\widehat{F}_{T,n,R_n}(x)-F_T(x)$, $x\in\mathcal{K}$, $n\geq 1$, are stochastically equicontinuous in restriction to the event $\{T\leq T_m\}$. By Corollary 1, $(\widehat{F}_{T,n,R_n}(x)-F_T(x))1_{\{T\leq T_m\}}=o_p\left(1\right)$ for each $x\in\mathcal{K}$. Hence, for each $m\geq 1$,

$$\sup_{x\in\mathcal{K}}\left|\left(\widehat{F}_{T,n,R_n}\left(x\right)-F_T\left(x\right)\right)1_{\{T\leq T_m\}}\right|\xrightarrow{\mathbb{P}}0.$$

By using a standard localization argument, we readily derive the asserted convergence. *Q.E.D.*

7.6 Proof of Theorem 5

Part a. Let $\mathbb{N}_1$ be an arbitrary subsequence of $\mathbb{N}$. By assumption, $\sup_{x\in\mathcal{K}}|\widehat{F}_{T,n,R_n}\left(x\right)-F_T\left(x\right)|=o_p\left(1\right)$. Hence there exists a further subsequence $\mathbb{N}_2\subseteq\mathbb{N}_1$, such that $\sup_{x\in\mathcal{K}}|\widehat{F}_{T,n,R_n}\left(x\right)-F_T\left(x\right)|\to 0$ as $n\to\infty$ along $\mathbb{N}_2$ on some $\mathbb{P}$-full event Ω^*.

Now, fix a sample path in Ω^*. Let $\mathcal{T}$ be the collection of continuity points of $Q_T^{\mathcal{K}}\left(\cdot\right)$. We have $1\{\widehat{F}_{T,n,R_n}\left(x\right)<\tau\}\to 1\{F_T\left(x\right)<\tau\}$ along $\mathbb{N}_2$ for $x\in\{x\in\mathcal{K}:F_T\left(x\right)\neq\tau\}$. For each $\tau\in\mathcal{T}$, the set $\{x \in \mathcal{K} : F_T(x) = \tau\}$ charges zero Lebesgue measure. By bounded convergence, along $\mathbb{N}_2$,

$$\widehat{Q}^{\mathcal{K}}_{T,n,R_n}(\tau) \to \inf \mathcal{K} + \int_{\inf \mathcal{K}}^{\sup \mathcal{K}} 1\{F_T(x) < \tau\}\, dx = Q^{\mathcal{K}}_T(\tau), \quad \forall \tau \in \mathcal{T}. \tag{67}$$

Since $F_T(x)$ is continuous in x, by Lemma 21.2 of van der Vaart (1998) and (67), we have $\widehat{F}^{\mathcal{K}}_{T,n,R_n}(x) \to F_T(x)$ along $\mathbb{N}_2$ for all $x \in \mathcal{K}$. Since $\widehat{F}^{\mathcal{K}}_{T,n,R_n}(\cdot)$ is also increasing, we further have $\sup_{x\in\mathcal{K}} |\widehat{F}^{\mathcal{K}}_{T,n,R_n}(x) - F_T(x)| \to 0$ along $\mathbb{N}_2$.

We have shown that for any subsequence, we can extract a further subsequence along which $\sup_{x\in\mathcal{K}} |\widehat{F}^{\mathcal{K}}_{T,n,R_n}(x) - F_T(x)| \to 0$ almost surely. Hence, $\sup_{x\in\mathcal{K}} |\widehat{F}^{\mathcal{K}}_{T,n,R_n}(x) - F_T(x)| = o_p(1)$.

Part b. Let $\mathbb{N}_1$ and $\mathbb{N}_2$ be given as in part (a). Since the continuity points of $Q_T(\cdot)$ are also continuity points of $Q^{\mathcal{K}}_T(\cdot)$, by (67), we have $\widehat{Q}^{\mathcal{K}}_{T,n,R_n}(\tau^*) \to Q^{\mathcal{K}}_T(\tau^*)$ along $\mathbb{N}_2$ almost surely. The assertion of part (b) then follows a subsequence argument similar to part (a). *Q.E.D.*

References

Bandi, F. and P. Phillips (2003). Fully Nonparametric Estimation of Scalar Diffusion Models. *Econometrica 71*, 241–283.

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

Chernozhukov, V., I. Fernandez-Val, and A. Galichon (2010). Quantile and Probability Curves without Crossing. *Econometrica 78*, 1093–1125.

Dassios, A. (1995). The Distribution of the Quantiles of a Brownian with Drift. *Annals of Applied Probability 5*, 389–398.

Eisenbaum, N. and H. Kaspi (2007). On the Continuity of Local Times of Borel Right Markov Processes. *Annals of Probability 35*, 915–934.

Florens-Zmirou, D. (1993). On Estimating the Diffusion Coefficient from Discrete Observations. *Journal of Applied Probability 30*, 790–804.

Geman, D. and J. Horowitz (1980). Occupation Densities. *The Annals of Probability 8*, 1–67.

Hardy, G., J. Littlewood, and G. Polya (1952). *Inequalities.* Cambridge University Press.

Ibragimov, I. and R. Has'minskii (1981). *Statistical Estimation: Asymptotic Theory.* Berlin: Springer.

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

Kryzhniy, V. (2003a). Direct Regularization of the Inversion of Real-Valued Laplace Transforms. *Inverse Problems 19*, 573–583.

Kryzhniy, V. (2003b). Regularized Inversion of Integral Transformations of Mellin Convolution Type. *Inverse Problems 19*, 1227–1240.

Ledoux, M. and M. Talagrand (1991). *Probability in Banach Spaces.* Berlin: Springer.

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

Marcus, M. B. and J. Rosen (2006). *Markov Processes, Gaussian Processes, and Local Times.* Cambridge University Press.

Protter, P. (2004). *Stochastic Integration and Differential Equations* (2nd ed.). Berlin: Springer-Verlag.

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

Tikhonov, A. and V. Arsenin (1977). *Solutions of Ill-Posed Problems.* Washington: Winston.

Todorov, V. and G. Tauchen (2012a). Inverse Realized Laplace Transforms for Nonparametric Volatility Density Estimation in Jump-Diffusions. *Journal of the American Statistical Association, forthcoming*.

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

van der Vaart, A. W. (1998). *Asymptotic Statistics.* Cambridge University Press.

Yor, M. (1995). The Distribution of Brownian Quantiles. *Journal of Applied Probability 32*, 405–416.