Estimating the Volatility Occupation Time via Regularized Laplace Inversion*

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

April 30, 2015

Abstract

We propose a consistent functional estimator for the occupation time of the spot variance of an asset price observed at discrete times on a finite interval with the mesh of the observation grid shrinking to zero. The asset price is modeled nonparametrically as a continuous-time Itô semimartingale with non-vanishing diffusion coefficient. The estimation procedure contains two steps. In the first step we estimate the Laplace transform of the volatility occupation time and, in the second step, we conduct a regularized Laplace inversion. Monte Carlo evidence suggests that the proposed estimator has good small-sample performance and in particular it is far better at estimating lower volatility quantiles and the volatility median than a direct estimator formed from the empirical cumulative distribution function of local spot volatility estimates. An empirical application shows the use of the developed techniques for nonparametric analysis of variation of volatility.

*We wish to thank a co-editor and three anonymous referees for their detailed and thoughtful comments, which helped greatly improve the paper. We would also like to thank Tim Bollerslev, Marine Carrasco, Peter Carr, Nathalie Eisenbaum, Jean Jacod, Andrew Patton, Peter Phillips, Philip Protter, Markus Reiß, as well as seminar participants at various conferences for many useful suggestions. Li's and Todorov's work were partially supported by NSF grants SES-1326819 and SES-0957330 respectively.

[†]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 widely used to model financial prices. In its general form, an Itô semimartingale can be represented as

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sqrt{V_s} dW_s + J_t, \tag{1}$$

where b_t is the drift, V_t is the spot variance, W_t is a Brownian motion and J_t is a pure-jump process. Both the continuous and the jump components are known to be present in financial time series. From an economic point of view, volatility and jump risks are very different and this has spurred the recent interest in separately identifying these risks from high-frequency data on X; see, for example, Barndorff-Nielsen and Shephard (2006) and Mancini (2009). In this paper we focus attention on the diffusive volatility part of X while recognizing the presence of jumps in X.

Most of the existing literature has concentrated on estimating nonparametrically volatility functionals of the form $\int_0^T g(V_s)ds$ for some smooth function g, typically three times continuously differentiable (see, e.g., Andersen et al. (2013), Renault et al. (2014), Jacod and Protter (2012), Jacod and Rosenbaum (2013) and many references therein). The most important example is the integrated variance $\int_0^T V_s ds$, which is widely used in empirical work. These temporally integrated volatility functionals can be alternatively thought of as spatially integrated moments with respect to the occupation measure induced by the volatility process (Geman and Horowitz (1980)).

Motivated by this simple observation, Li et al. (2013) consider the estimation of the volatility occupation time (VOT), defined by

$$F_T(x) = \int_0^T 1_{\{V_s \le x\}} ds, \quad \forall x > 0, \tag{2}$$

which is the pathwise analogue of the cumulative distribution function (CDF).[1] Evidently, the VOT also takes the form $\int_0^T g(V_s)ds$ but with g discontinuous. The latency of V_t and the nonsmoothness of $g(\cdot)$ turn out to cause substantive complications in the estimation of the VOT.

To see the empirical relevance of the VOT, we note that the widely used integrated

[1] To make the analogy exact, one may normalize the expression in (2) by T^{-1}. Here, we follow the convention in the literature (see, e.g., Geman and Horowitz (1980)) without using this normalization.

variance $\int_0^T V_s ds$ is nothing but the mean of the occupation measure $\int_0^\infty x F_T(dx)$, where the equivalence is by the occupation formula[2]. Therefore, the relation between the VOT, the VOT quantiles and the integrated variance is exactly analogous to the relation between the CDF, its quantiles and the mean of a random variable. Needless to say, in classical econometrics and statistics, much can be learned from the CDF and quantiles beyond the mean. By the same logic, in the study of volatility risk, the VOT and its quantiles provide additional useful information (such as dispersion) of the volatility risk which has been well recognized as an important risk factor in modern finance.

Li et al. (2013) provide a two-step estimation method for estimating the VOT from a high-frequency record of X by first nonparametrically estimating the spot variance process over $[0, T]$ and then constructing a direct plug-in estimator corresponding to (2). Their estimation method is based on a thresholding technique (Mancini (2001)) to separate volatility from jumps and forming blocks of asymptotically decreasing length to account for the time variation of volatility (Foster and Nelson (1996), Comte and Renault (1998)).

In this paper we develop an alternative estimator for the VOT from a new perspective. The idea is to recognize that the informational content of the occupation time is the same as its pathwise Laplace transform, and the latter can be conveniently estimated as a sum of cosine-transformed logarithmic returns (Todorov and Tauchen (2012b)). Following this idea, our proposal is to first estimate the Laplace transform of the VOT and then conduct the Laplace inversion. The inversion is nontrivial because it is an ill-posed problem (Tikhonov and Arsenin (1977)). Indeed, Laplace inversion amounts to solving a Fredholm integral equation of the first kind, and the solution is not continuous in the Laplace transform. In order to obtain stable solutions, we regularize the inversion step by using the direct regularization method of Kryzhniy (2003a,b). The final estimator is known in closed form, up to a one-dimensional numerical integration, and can be easily computed using standard software.

The proposed inversion method and the plug-in method of Li et al. (2013) both involve some tuning parameters, but they play very different roles and reflect the different tradeoffs underlying these two methods. For the inversion method, the first-step estimation of the Laplace transform does not involve any tuning. In fact, the first-step is automatically robust to the presence of price jumps and achieves the parametric rate of convergence when jumps are not "too active;" see Todorov and Tauchen (2012b). In the second step, a

[2]See, for example, (6.4) in Geman and Horowitz (1980).

tuning parameter is introduced for stabilizing the Laplace inversion, at the cost of inducing a regularization bias. For the direct plug-in method of Li et al. (2013), the key is to recover the spot variance process, for which two types of tuning are needed. One is to select a threshold for eliminating jumps, for which the trade-off is to balance the pass-through of small jumps and the false elimination of large diffusive movements. The other is to select the block size of the local window by trading-off the bias induced by the time variation of the volatility and the sampling error induced by Brownian shocks. For both methods, the optimal choice of the tuning parameters remains an open, and likely very challenging, question. We provide some simulation results for assessing the finite-sample impact of these tuning parameters.

We can further compare our analysis here with Todorov and Tauchen (2012a), where somewhat analogous steps were followed to estimate the invariant probability density of the volatility process, but there are fundamental differences between the current paper and Todorov and Tauchen (2012a). First, 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. This is further illustrated by our empirical application which studies the randomness of the (occupational) interquartile range of various transforms of volatility. Thus, in this paper we impose neither the existence of invariant distribution of the volatility process nor mixing-type conditions. While such conditions may be reasonable for analyzing data from a long sample period, they are unlikely to "kick in" sufficiently fast in short samples in view of the high persistence of the volatility process (Comte and Renault (1998)). In our setup, we allow the volatility process to be nonstationary and strongly serially dependent. This asymptotic setting provides justification for estimating "distributional" or, to be more precise, occupational properties of the volatility process using data within relatively short (sub)sample periods. Second, and quite importantly from a technical point of view, the object of interest here (i.e., the VOT) is a random quantity with limited pathwise smoothness properties. It is well known that smoothness conditions are important in the analysis of ill-posed problems (Carrasco et al. (2007)). Indeed, our analysis of the stochastic regularization bias demands technical arguments that are very different from Todorov and Tauchen (2012a), where the invariant distribution is deterministic and smoother. As a technical by-product of our analysis, we provide primitive conditions for the smoothness of the volatility occupation density for a popular class of jump-diffusion stochastic volatility models. Overall, the current paper can be viewed as an extension of the results in Todorov and Tauchen (2012a) to the theoretically different setting of fixed time span and provides the theoretical justification for applying the method in Todorov and Tauchen (2012a) over different time horizons.

Finally, the current paper is also connected with the broad literature on ill-posed problems in econometrics; see Carrasco et al. (2007) for a comprehensive review. In the current paper, we adopt a direct regularization method for inverting transforms of the Mellin convolution type (Kryzhniy (2003a,b)), which is very different from spectral decomposition methods reviewed in Carrasco et al. (2007). In particular, we do not consider the Laplace transform as a compact operator for some properly designed Hilbert spaces. We prove the functional convergence for the VOT estimator under the local uniform topology, instead of under a (weighted) L_2 norm. The uniform convergence result is then used to prove consistency of estimators of the (random) VOT quantiles.

Our contribution is twofold. First, the proposed estimator is theoretically novel and has finite-sample performance that is generally better than the benchmark set by Li et al. (2013) in the presence of active jumps. To be specific, we provide Monte Carlo evidence that the regularized Laplace inversion estimates are more accurate than those of the direct plug-in method for estimating lower volatility quantiles as well as the volatility median in jump-diffusion models. This pattern appears in all Monte Carlo settings and is indeed quite intuitive: "small" jumps are un-truncated, and they induce a relatively large finite-sample bias for volatility estimation for lower quantiles. Moreover, this finding extends even to the estimation of higher volatility quantiles when (asymptotically valid) non-adaptive truncation thresholds are used for the direct plug-in method. That noted, we do observe a partial reversal of this pattern for estimating higher volatility quantiles when certain adaptive truncation thresholds are used for the direct plug-in method, so the proposed method does not always dominate that of Li et al. (2013). We further illustrate the empirical use of the proposed estimator by studying the dependence between the (occupational) interquartile range of various transforms of the volatility and the level of the volatility process. Such analysis sheds light on the modeling of volatility of volatility. Second, to the best of our knowledge, the ill-posed problem and the associated regularization is the first ever explored in a setting with discretely sampled semimartingales within a fixed time span. Other ill-posed problems within the high-frequency setting will naturally arise, for example, in nonparametric regressions involving elements of the diffusion matrix of a multivariate Itô semimartingale; see Härdle and Linton (1994) for a review in the classical long-span setting.

This 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 VOT, 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, followed by an empirical illustration in Section 5. 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 space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ with the jump component J_t given by

$$\begin{aligned} J_t \quad = \quad & \int_0^t \int_{\mathbb{R}} \left(\delta\left(s,z\right) 1_{\{|\delta(s,z)|\leq 1\}} \right) \tilde{\mu}\left(ds,dz\right) \\ & + \int_0^t \int_{\mathbb{R}} \left(\delta\left(s,z\right) 1_{\{|\delta(s,z)|>1\}} \right) \mu\left(ds,dz\right), \end{aligned} \tag{3}$$

where μ is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}$ with compensator ν of the form $\nu\left(dt,dz\right) = dt \otimes \lambda\left(dz\right)$ 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. Regularity conditions on X_t are collected below.

Assumption A: The following conditions hold for some constant $r \in (0,2)$ and a localizing sequence $(T_m)_{m\geq 1}$ of stopping times.[3]

A1. X is an Itô semimartingale given by (1) and (3), where the process b_t is locally bounded and the process V_t is strictly positive and càdlàg. Moreover, $|\delta\left(\omega,t,z\right)|^r \wedge 1 \leq \Gamma_m\left(z\right)$ for all $\omega \in \Omega$, $t \leq T_m$ and $z \in \mathbb{R}$, where $(\Gamma_m)_{m\geq 1}$ is a seqeuence of λ-integrable deterministic functions on $\mathbb{R}$.

A2. For a sequence $(K_m)_{m\geq 1}$ of real numbers, $\mathbb{E}|V_t - V_s|^2 \leq K_m|t-s|$ for all t,s in $[0,T_m]$ with $|t-s| \leq 1$.

Assumption A imposes very mild regularities on the process X and is standard in the literature on discretized processes; see Jacod and Protter (2012). The dominance condition in Assumption A is only required to hold locally in time up to the stopping time T_m, which often take forms of hitting times of adapted processes; this requirement is much weaker than a global dominance condition that corresponds to $T_m \equiv +\infty$. This more general setup,

[3]A localizing sequence of stopping times is a sequence of stopping times which increases to $+\infty$.

however, does not add any technical complexity into our proofs, thanks to the standard localization procedure in stochastic calculus; see Section 4.4.1 in Jacod and Protter (2012) for a review on the localization procedure.

We note that Assumption A imposes no parametric structure on the underlying process, allowing for jumps in X_t and V_t, and dependence between various components in an arbitrary manner. In particular, we allow the stochastic volatility process V_t to be dependent on the Brownian motion W_t, so as to accommodate the "leverage" effect (Black (1976)). The constant r in Assumption A1 controls the activity of small jumps, as it provides a bound for the generalized Blumenthal–Getoor index. The assumption is stronger when r is smaller.

Assumption A2 requires the spot variance process V_t to be (locally) 1/2-Hölder continuous under the L_2 norm. This assumption holds in the well-known case in which V_t is also an Itô semimartingale with locally bounded characteristics. It also holds for long-memory specifications that are driven by fractional Brownian motion; see Comte and Renault (1996). Assumption A2 coincides, albeit with a different norm, to the one maintained by Renault et al. (2014).

2.2 Occupation times

We next collect some assumptions on the VOT and the associated occupation density. In what follows we define $F_t(\cdot)$ as (2) with T replaced by t.

Assumption B: The following conditions hold for some localizing sequence $(T_m)_{m\geq 1}$ of stopping times and a constant sequence $(C_m)_{m\geq 1}$.

B1. Almost surely, the function $x \mapsto F_t(x)$ is piecewise differentiable with derivative $f_t(x)$ for all $t \in [0, T]$. For all $x, y \in (0, \infty)$, $\mathbb{P}(\{\text{the interval } (x, y) \text{ contains some nondifferentiable point of } F_T(\cdot)\} \cap \{T \leq T_m\}) \leq C_m |x - y|$.

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

Assumption B is used in our analysis on the estimation of $F_T(x)$ for fixed x. As in Assumption A, we only need the dominance conditions to hold locally up to the localizing sequence T_m. Assumption B1 holds if the occupation density of V_t exists, which is the case for general semimartingale processes with nondegenerate diffusive component and large classes of Gaussian processes; see, for example, Geman and Horowitz (1980), Protter (2004), Marcus and Rosen (2006), Eisenbaum and Kaspi (2007) and references therein. Assumption B1 holds more generally under settings where $F_t(\cdot)$ can be nondifferentiable (and even discontinuous) at random points, as long as these irregular points are located "diffusively" on the line, as formulated by the second part of Assumption B1. This generality accommodates certain pure-jump stochastic volatility processes, such as a compound Poisson process with bounded marginal probability density.[4] 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 stochastic volatility models.

To derive uniform convergence results, we need to strengthen Assumption B as follows.

Assumption C: The following conditions hold for some localizing sequence $(T_m)_{m\geq 1}$ of stopping times and constants $\tilde{\gamma} > \varepsilon > 0$.

C1. Almost surely, the function $x \mapsto F_t(x)$ is differentiable with derivative $f_t(x)$ for all $t \in [0, T]$.

C2. For any compact $\mathcal{K} \subset (0, \infty)$, $\sup_{x \in \mathcal{K}} \mathbb{E}\left[f_{T \wedge T_m}(x)^{1+\varepsilon}\right] < \infty$.

C3. 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 $\mathbb{E}\left[\sup_{t \leq T} \left|f_{t \wedge T_m}(x) - f_{t \wedge T_m}(y)\right|^{1+\varepsilon}\right] \leq C_m |x - y|^{\tilde{\gamma}}$.

Assumptions C1 and C2 are stronger than Assumption B. In addition, the Hölder-continuity condition in Assumption C3 is nontrivial to verify. We hence devote Section 2.3 to discussing primitive conditions for Assumption C that cover many volatility models used in financial applications, although this set of conditions is far from exhaustive. Finally, we note that C3 involves 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 that in C3.

2.3 Some primitive conditions for Assumption C

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

$$dV_t = a_t dt + s(V_t) dB_t + dJ_{V,t}, \tag{4}$$

[4] When V_t is a compound Poisson process, each nondifferentiable point of $F_T(\cdot)$ is a realized level of V_t. Therefore, the probability in Assumption B1 is bounded by $\mathbb{P}(V_t \in (x, y)$ for some $t \in [0, T])$. Since the expected number of jumps is finite and V_t has a bounded density, this probability is further bounded by $|x - y|$ up to a multiplicative constant.

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 $1/s(\cdot)$, that is, $g(x) = \int^x du/s(u)$ and the constant of integration is irrelevant. By Itô's formula, the continuous martingale part of $\widetilde{V}_t$ is B_t. Lemma 2.1(a) below shows that under some regularity conditions, the transformed process $\widetilde{V}_t$ satisfies Assumption C. To prove Lemma 2.1(a) we compute the occupation density of $\widetilde{V}_t$ explicitly in terms of stochastic integrals via the Meyer–Tanaka formula[5] and then we bound the corresponding spatial increments. Then Lemma 2.1(b) shows that V_t inherits the same property, that is, it satisfies Assumption C, provided that the transformation $g(\cdot)$ is smooth enough.

Lemma 2.1 *(a) Let $k > 1$. Consider a process $\widetilde{V}_t$ with the following form*

$$\widetilde{V}_t = \widetilde{V}_0 + \int_0^t \tilde{a}_s ds + B_t + \int_0^t \int_{\mathbb{R}} \tilde{\delta}(s,z)\, \mu(ds,dz), \tag{5}$$

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

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

(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

[5] This is possible because the continuous martingale part of $\widetilde{V}_t$ is a Brownian motion.

$\mathcal{K} \subset (0, \infty)$, there exists some constant $C > 0$, such that $|g'(x) - g'(y)| \le C|x-y|^{\bar{\gamma}}$ for all $x, y \in \mathcal{K}$. Then V_t satisfies Assumption C.

3 Estimating volatility occupation times

We now present our estimator for the VOT and its asymptotic properties. 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 VOT is to first estimate its Laplace transform and then to invert the latter.

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

$$\mathcal{L}_T(u) \equiv \int_0^T e^{-uV_s} ds, \quad \forall u > 0.$$

By the occupation density formula (see, e.g., (6.5) in Geman and Horowitz (1980)), the temporal integral above can be rewritten as a spatial integral under the occupation measure, that is,

$$\mathcal{L}_T(u) = \int_0^\infty e^{-u\,x} f_T(x) dx = \int_0^\infty e^{-u\,x} F_T(dx), \quad \forall u > 0.$$

The Laplace transform of the VOT can then be obtained by using Fubini's theorem and is given by

$$\frac{\mathcal{L}_T(u)}{u} = \int_0^\infty e^{-ux} F_T(x) dx. \tag{6}$$

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

$$\widehat{\mathcal{L}}_{T,n}(u) = \Delta_n \sum_{i=1}^{[T/\Delta_n]} \cos\left(\sqrt{2u}\Delta_i^n X / \Delta_n^{1/2}\right), \quad \forall u > 0, \tag{7}$$

where $[\cdot]$ denotes the largest smaller integer function and $\Delta_i^n X \equiv X_{i\Delta_n} - X_{(i-1)\Delta_n}$. Todorov and Tauchen (2012b) show that $\widehat{\mathcal{L}}_{T,n}(\cdot) \xrightarrow{\mathbb{P}} \mathcal{L}_T(\cdot)$ locally uniformly with an associated central limit theorem. These convergence results are robust to the presence of price jumps without appealing to the thresholding technique as in Mancini (2001) and Li et al. (2013). Consequently, $u^{-1}\widehat{\mathcal{L}}_{T,n}(u) \xrightarrow{\mathbb{P}} u^{-1}\mathcal{L}_T(u)$ for each $u \in (0, \infty)$.

Once the Laplace transform of the VOT is estimated from the data, in the next step we 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 $u^{-1}\mathcal{L}_T(u)$:

$$F_{T,R}(x) = \int_0^\infty \mathcal{L}_T(u)\Pi(R, ux)\frac{du}{u}, \quad \forall x > 0, \tag{8}$$

where $R > 0$ is a regularization parameter and the inversion kernel $\Pi(R, x)$ is defined as[6]

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

It can be shown that the regularized inversion $F_{T,R}(x)$ can be also written as (see (27))

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

That is, $F_{T,R}(x)$ is generated by smoothing the VOT via the kernel $2u^{1/2}\sin(R\ln u)/\pi(u^2 - 1)$, which approaches the Dirac mass at $u = 1$ as $R \to \infty$.[7]

Our estimator for the VOT is constructed by simply replacing $\mathcal{L}_T(u)$ in (8) with $\widehat{\mathcal{L}}_{T,n}(u)$, that is, it is given by

$$\widehat{F}_{T,n,R}(x) = \int_0^\infty \widehat{\mathcal{L}}_{T,n}(u)\Pi(R, ux)\frac{du}{u} = \int_{-\infty}^\infty \widehat{\mathcal{L}}_{T,n}(e^z)\Pi(R, xe^z)dz. \tag{9}$$

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

We now turn to the asymptotic properties of the estimator $\widehat{F}_{T,n,R_n}(x)$, where R_n is a

[6]This seemingly complicated inversion kernel corresponds to a simple stabilizing procedure for the Mellin transform of the integral equation (6); see Section 2 in Kryzhniy (2003a) for technical details.

[7]The fact that the regularized quantity can be considered as a type of convolution between the object of interest and a smoothing kernel that asymptotically collapses to a Dirac mass also arises in nonparametric kernel density estimation; see, for example, (9) in Härdle and Linton (1994).

sequence of (strictly positive) regularization parameters that grows to $+\infty$ asymptotically. Here, we allow R_n to be random so that it can be data-dependent, while the rate at which it grows is given by a deterministic sequence ρ_n. It is conceptually useful to decompose the estimation error $\widehat{F}_{T,n,R_n}(x) - F_T(x)$ into two components: the regularization bias $F_{T,R_n}(x) - F_T(x)$ and the sampling error $\widehat{F}_{T,n,R_n}(x) - F_{T,R_n}(x)$. Lemmas 3.1 and 3.2 below characterize the order of magnitude of each component.

Lemma 3.1 *Let $x > 0$ be a constant. Suppose that $R_n = O_p(\rho_n)$ and $R_n^{-1} = O_p(\rho_n^{-1})$ for some deterministic sequence ρ_n with $\rho_n \to \infty$. Under Assumptions A and B,*

$$F_{T,R_n}(x) - F_T(x) = O_p(\rho_n^{-1}\ln(\rho_n)).$$

Lemma 3.2 *Let $\eta \in (0, 1/2)$ be a constant and $\mathcal{K} \subset (0, \infty)$ be compact. Suppose that $R_n \leq \rho_n$ for some deterministic sequence ρ_n with $\rho_n \to \infty$. Under Assumption A,*

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

Lemma 3.1 describes the order of magnitude of the regularization bias. Lemma 3.2 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. Lemma 3.2 holds for any constant $\eta \in (0, 1/2)$. This constant arises as a technical device from the proof and should be taken close to $1/2$ so that the bound in Lemma 3.2 is sharper.

Combining Lemmas 3.1 and 3.2 and choosing the regularization parameter properly, we obtain the pointwise consistency of the VOT estimator.

Theorem 3.1 *Suppose (i) Assumptions A and B; (ii) $\rho_n = \delta\ln\left(\Delta_n^{-1}\right)$ for some $\delta \in \left(0, 2\bar{\delta}/\pi\right)$, where $\bar{\delta} \equiv \min\{(r\wedge 1)(1/r - 1/2), 1/2\}$; (iii) $R_n \leq \rho_n$ and $R_n^{-1} = O_p(\rho_n^{-1})$. Then for each $x > 0$,*

$$\widehat{F}_{T,n,R_n}(x) - F_T(x) \xrightarrow{\mathbb{P}} 0.$$

In Theorem 3.1, we set the regularization parameter R_n to grow slowly to infinity so that both the regularization bias and the sampling error vanish asymptotically. Condition (ii) specifies the admissible range of the tuning parameter, which depends on r when $r > 1$ and shrinks as r approaches 2 (the theoretical upper bound for jump activity of semimartingales). This phenomenon reflects the well-known difficulty of disentangling active jumps from the diffusive component. Estimators for jump activity (see, e.g., Aït-Sahalia and Jacod (2009)) may be used to assess the restrictiveness of this condition in a given sample.

More generally, the inversion method is not limited to the realized Laplace transform estimator $\widehat{\mathcal{L}}_{T,n}(\cdot)$. With a generic estimator $\widetilde{\mathcal{L}}_{T,n}(\cdot)$ for $\mathcal{L}_T(\cdot)$, we can associate an inversion estimator for the VOT as

$$\widetilde{F}_{T,n,R}(x) = \int_0^\infty \widetilde{\mathcal{L}}_{T,n}(u)\Pi(R,ux)\frac{du}{u}, \quad \forall x > 0.$$

Theorem 3.2 below shows that $\widetilde{F}_{T,n,R_n}(x)$ is a consistent estimator for $F_T(x)$ under a high-level condition concerning the estimation error of $\widetilde{\mathcal{L}}_{T,n}(\cdot)$ under the L_1 norm.

Theorem 3.2 *Suppose (i) there exist a localizing sequence $(T_m)_{m\geq 1}$ of stopping times and a sequence $(C_m)_{m\geq 1}$ of positive constants such that for some $\bar{c} \in (0,1/2)$, $\bar{\delta} > 0$ and all $u > 0$,*

$$\mathbb{E}\left|\widetilde{\mathcal{L}}_{T_m,n}(u) - \mathcal{L}_{T_m}(u)\right| \leq C_m\left(u^{-\bar{c}} + u^{1+\bar{c}}\right)\Delta_n^{\bar{\delta}}; \tag{10}$$

(ii) $\rho_n = \delta \ln\left(\Delta_n^{-1}\right)$ for some $\delta \in \left(0, 2\bar{\delta}/\pi\right)$; (iii) $R_n \leq \rho_n$ and $R_n^{-1} = O_p(\rho_n^{-1})$. Then for each $x > 0$,

$$\widetilde{F}_{T,n,R_n}(x) \xrightarrow{\mathbb{P}} F_T(x).$$

Theorem 3.1 can also be proved by using Theorem 3.2. Indeed, it can be seen from the proof of Lemma 3.2 that the estimator $\widehat{\mathcal{L}}_{T,n}(\cdot)$ verifies (10) for any $\bar{c} \in (0,1/2)$ with $\bar{\delta}$ as given in Theorem 3.1. In other settings, alternative estimators might be required to verify these conditions. The key to the proof of Theorem 3.2 is an extension of Lemma 3.2 under condition (10), but with a coarser bound.

A pessimistic theoretical bound on the rate of convergence for Theorem 3.1 is essentially $\ln\left(\Delta_n^{-1}\right)$, which is driven by the regularization bias. The plug-in estimator of Li et al. (2013), in contrast, can formally be bounded by a polynomial rate of convergence. However, the bounds might not be sharp. Efficiency issues in the estimation of integrated volatility functionals of the form $\int_0^T g(V_s)ds$ has recently been tackled by Jacod and Reiß (2013), Clément et al. (2013), Jacod and Rosenbaum (2013) and Renault et al. (2014) for smooth $g(\cdot)$. The VOT, on the other hand, corresponds to a discontinuous transform $g(\cdot) = 1_{\{\cdot \leq x\}}$. Assessing the efficiency of the VOT estimators remains to be an open question that is likely very challenging. That being said, at least intuitively, more efficient estimators of the integrated Laplace transform of volatility than the one in (7), like the ones considered in Renault et al. (2014), can help improve the efficiency of the VOT estimators based on regularized inversion. The theoretical results in Theorem 3.2 provide the foundations for doing this.

Another open question is how to optimally choose tuning parameters in order to minimize some loss criterion, such as the mean square error. Such analysis remains to be a technical challenge, not only for the current paper, but also for the in-fill analysis of high-frequency semimartingale data in general.[8] In Section 4, we provide simulation results in a realistically calibrated Monte Carlo setting for comparing the finite-sample performance of the two methods and for assessing the robustness of the proposed estimator with respect to the tuning parameter.

The pointwise convergence in Theorems 3.1 and 3.2 can be further strengthened to be uniform in the spatial variable, as shown below.

Theorem 3.3 *Suppose Assumption C. Then the following statements hold for any compact* $\mathcal{K} \subset (0, \infty)$.

(a) Under the conditions of Theorem 3.1,

$$\sup_{x \in \mathcal{K}} \left| \widehat{F}_{T,n,R_n}(x) - F_T(x) \right| \xrightarrow{\mathbb{P}} 0. \tag{11}$$

(b) Under the conditions of Theorem 3.2,

$$\sup_{x \in \mathcal{K}} \left| \widetilde{F}_{T,n,R_n}(x) - F_T(x) \right| \xrightarrow{\mathbb{P}} 0. \tag{12}$$

Next, we provide a refinement to the functional estimator $\widehat{F}_{T,n,R_n}(\cdot)$. The discussion below only requires the uniform convergence (11) to hold, so it also applies to the generic estimator $\widetilde{F}_{T,n,R_n}(\cdot)$ under (12). 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-

[8]The key difficulty lies in the characterization and estimation of asymptotic bias terms under a setting with asymptotically varying tuning parameters. In a recent paper, Kristensen (2010) considers the challenging question of the optimal choice of a bandwidth parameter in the estimation of spot volatility. In the setting without leverage effect and jumps, Kristensen (2010) remarks (see p. 77) that the optimal choice of tuning parameter is still difficult when the sample path of the volatility process is not differentiable with respect to time. Nondifferentiable paths, however, are common for the typical stochastic volatility models as the ones considered here.

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:

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

For any compact interval $\mathcal{K} \subset (0, \infty)$, 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 3.3 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 $\tau \in (0, T)$, $\widehat{Q}_{T,n,R_n}^{\mathcal{K}}(\tau) = \inf\{x : \widehat{F}_{T,n,R_n}^{\mathcal{K}}(x) \geq \tau\}$. 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 3.4 below.

Theorem 3.4 *Let $\mathcal{K} \subset (0, \infty)$ be a compact interval. If $F_T(\cdot)$ is continuous and*

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

then we have the following.

(a)

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

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

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

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)). Chernozhukov et al. (2010) shows that rearrangement leads to finite-sample improvement under very general settings; see Proposition 4 there.[9] Our asymptotic 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 $\sup_{x\in\mathcal{K}}|\widehat{F}_{T,n,R_n}(x)-F_T(x)|\xrightarrow{\mathbb{P}}0$ and, of course, our argument does not rely on the functional delta method.

4 Monte Carlo

We now examine the finite-sample performance of our estimator and compare it with the direct plug-in method proposed by Li et al. (2013). We consider the following jump-diffusion volatility model in which the log-volatility is a Lévy-driven Ornstein-Uhlenbeck (OU) process, that is,

$$dX_t=\sqrt{e^{V_t-1}}dW_t+dY_t,\quad dV_t=-0.03V_tdt+dL_t, \tag{13}$$

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. Our volatility specification is quite general as it allows for both diffusive and jump shocks in volatility, with the latter being of infinite

[9] Monotonization methods may also improve the rate of convergence; see Carrasco and Florens (2011) for such an example in the study of deconvolution problems. It may be interesting to explore this theoretical possibility in future research.

activity. The mean and the persistence of the volatility process are calibrated realistically to observed financial data. In particular, we set $\mathbb{E}[e^{V_t-1}] = 1$ (our unit of time is a trading day and we measure returns in percentage) and the persistence of a shock in V_t has a half-life of approximately 23 days. Finally, Y_t in (13) is a tempered stable Lévy process, i.e., a pure-jump Lévy process with Lévy measure $c\frac{e^{-\lambda|x|}}{|x|^{\beta+1}}$, which is independent from L_t and W_t. The tempered stable process is a flexible jump specification with separate parameters controlling small and big jumps: λ controls the jump tails and β coincides with the Blumenthal-Getoor index of Y_t (and hence controls the small jumps). We consider three cases in the Monte Carlo: (a) no price jumps, which corresponds to $c = 0$, (b) low-activity price jumps, with parameters $c = 6.2908$, $\lambda = 7$ and $\beta = 0.1$ and (c) high-activity price jumps, with parameters $c = 1.3408$, $\lambda = 7$ and $\beta = 0.9$. The value of λ in each case is set to produce jump tail behavior consistent with nonparametric evidence reported in Bollerslev and Todorov (2011). Further, in all considered cases for Y_t, we set the parameter c so that the second moment of the increment of Y on unit interval is equal to 0.3 which produces jump contribution in total quadratic variation of X similar to earlier nonparametric empirical evidence from high-frequency financial data.[10]

In the Monte Carlo we fix the time span to be $T = 22$ days, equivalent to one calendar month, and we consider $n = 80$ which corresponds to 5-minute sampling of intraday observations of X in a 6.5-hour trading day. For each realization we compute the 25-th, 50-th and 75-th volatility quantiles over the interval $[0, T]$ and assess the accuracy in measuring these random quantities by reporting bias and mean absolute deviation (MAD) around the true values for the considered estimators.

We first analyze the effect of the regularization parameter R_n on the volatility quantile estimation. For brevity, we conduct the analysis in the case when X_t does not contain price jumps, while noting that similar results hold in the other cases. In Table 1 we report results from the Monte Carlo for regularized Laplace inversion with values of the regularization parameter of $R_n = 2.5$, $R_n = 3.0$ and $R_n = 3.5$. Overall, the performance of our volatility quantile estimator is satisfactory with biases being small in relative terms. In general, the

[10]The price jumps specifications considered here are both of infinite activity, hence there are infinite numbers of jumps within a finite interval. However, "big" jumps are always of finite number. For example jumps of size bigger than 0.34%, which corresponds to an average three standard deviation move of the continuous price price increment at the 5-minute interval, occur on average 9.17 (case b) and 3.87 (case c) times on an interval of length 22 days. The low-activity jump specification generates more big jumps than the high-activity one, with the role reversed for the small jump sizes (recall that the quadratic variation of both jump specifications is constrained to be the same).

difference across the different values of the regularization parameter are relatively small. From Table 1 we can see the typical bias-variance tradeoff that arises in nonparametric estimation: for lower value of R_n (more smoothing) the biases are larger but the sampling variability is smaller, while for higher value of R_n (less smoothing) the opposite is true. The value of R_n that leads to the smallest MAD is $R_n = 3.0$ and henceforth we keep the regularization parameter at this value.

We next compare the performance of the regularized Laplace inversion approach for volatility quantile estimation with the direct plug-in method of Li et al. (2013). The latter is based on local estimators of the volatility process over blocks given by

$$\widehat{V}_{i\Delta_n} = \frac{1}{u_n}\sum_{j=1}^{k_n}\left(\Delta_{i+j}^n X\right)^2 1_{\{|\Delta_{i+j}^n X|\leq v_{n,i\Delta_n}\}}, \quad i = 0, ..., [T/\Delta_n] - k_n,$$

where $u_n = k_n\Delta_n$ and k_n denotes the number of high-frequency elements within a block (k_n satisfies $k_n \to \infty$ and $k_n\Delta_n \to 0$); $v_{n,t}$ is the threshold which takes the form $v_{n,t} = \alpha_{n,t}\Delta_n^{\varpi}$ for some strictly positive process $\alpha_{n,t}$ and $\varpi \in (0, 1/2)$. These local estimators are then used to approximate the volatility trajectory via

$$\widehat{V}_t = \widehat{V}_{iu_n}, \;\; t \in [iu_n, (i+1)\, u_n), \;\; \text{and} \;\; \widehat{V}_t = \widehat{V}_{([T/u_n]-1)u_n}, \;\; [T/u_n]u_n \leq t \leq T,$$

and from here the direct estimator of the volatility occupation time is given by

$$\widehat{F}^d_{T,n}(x) = \int_0^T 1_{\{\widehat{V}_s \leq x\}} ds, \quad x \in \mathbb{R}.$$

The direct estimator $\widehat{F}^d_{T,n}(x)$ has two tuning parameters. The first is the block size k_n which plays a similar role as the regularization parameter R_n in the regularized Laplace inversion method. We follow Li et al. (2013) and set $k_n = 4$ throughout. The second tuning parameter is the choice of the threshold $v_{n,t}$. There are various ways of setting this threshold which all lead to asymptotically valid results. One simple choice is a time-invariant threshold of the form $v_{n,t} = 3\overline{\sigma}\Delta_n^{0.49}$, where $\overline{\sigma}$ is an estimator of $\sqrt{\mathbb{E}(V_t)}$. Another is a time-varying threshold that takes into account the stochastic volatility. Here we follow Li et al. (2013) (and earlier work on threshold estimation) and experiment with $v_{n,t} = 3\sqrt{BV_j}\Delta_n^{0.49}$ and $v_{n,t} = 4\sqrt{BV_j}\Delta_n^{0.49}$ for $t \in [j-1, j)$, where $BV_j = \frac{\pi}{2}\sum_{i=[(j-1)/\Delta_n]+2}^{[j/\Delta_n]} |\Delta_{i-1}^n X||\Delta_i^n X|$ is

Table 1: Monte Carlo Results: Effect of R_n

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: Regularized Laplace Inversion with $R_n = 2.5$									
$V_0 = Q^V(0.25)$	0.1737	−0.0201	0.0282	0.2860	−0.0178	0.0322	0.4519	0.0008	0.0430
$V_0 = Q^V(0.50)$	0.3293	−0.0392	0.0527	0.5243	−0.0324	0.0583	0.8069	0.0078	0.0746
$V_0 = Q^V(0.75)$	0.6337	−0.0716	0.0975	0.9945	−0.0600	0.1104	1.5162	0.0178	0.1383
Panel B: Regularized Laplace Inversion with $R_n = 3.0$									
$V_0 = Q^V(0.25)$	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 C: Regularized Laplace Inversion with $R_n = 3.5$									
$V_0 = Q^V(0.25)$	0.1737	−0.0093	0.0262	0.2860	−0.0042	0.0335	0.4519	−0.0114	0.0556
$V_0 = Q^V(0.50)$	0.3293	−0.0180	0.0500	0.5243	−0.0093	0.0632	0.8069	−0.0080	0.0947
$V_0 = Q^V(0.75)$	0.6337	−0.0298	0.0981	0.9945	−0.0211	0.1201	1.5162	−0.0154	0.1736

Note: 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 the true value. The Monte Carlo replica is* 1000*.*

the Bipower Variation estimator of Barndorff-Nielsen and Shephard (2004).[11]

In Tables 2 and 3 we compare the precision of estimating the monthly volatility quantiles via regularized Laplace inversion (with $R_n = 3.0$) and via the direct method for the above discussed ways of setting the threshold parameter.[12] We consider only the empirically realistic scenarios in which X contains jumps in the comparison. An immediate observation from the tables is that the threshold parameter in $\widehat{F}^d_{T,n}$ plays a crucial role. Indeed, a constant threshold does a very poor job: it yields huge biases and it results also in very noisy estimates. The volatility quantile estimators based on $\widehat{F}^d_{T,n}$ work well only when a time-varying adaptive (to the current level of volatility) threshold is selected. Comparing these estimators with the one based on the regularized Laplace inversion, we see an interesting pattern. The estimation of the lower volatility quantiles is done significantly more precisely via the inversion method. For the lower volatility quantiles, the estimates based on $\widehat{F}^d_{T,n}$ contain nontrivial bias. This is due to small un-truncated jumps which play a relatively bigger role when estimating the lower volatility quantiles. The above observation continues to hold, albeit to a far less extent, for the volatility median. For the highest volatility quantile, we see a partial reverse. This volatility quantile is estimated more precisely via $\widehat{F}^d_{T,n}$ but mainly when the lower time-varying threshold $v_{n,t} = 3\sqrt{BV_j}\Delta_n^{0.49}$ is used. Overall, we find mixed results in this comparative analysis, but the evidence in Tables 2 and 3 clearly illustrates that the proposed volatility quantile estimator based on the regularized Laplace inversion provides an important alternative to the direct plug-in method.

5 Empirical Application

We illustrate the nonparametric quantile reconstruction technique with an 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

[11]In principle, the direct plug-in method of Li et al. (2013) can be applied to other jump-robust spot volatility estimators and may achieve better finite-sample performance. Improving the direct plug-in method in this direction is beyond the scope of the current paper.

[12]Comparing the results for $\widehat{F}_{T,n}$ in Tables 2 and 3 with those in Table 1, we notice that the negative biases for the first two quantiles in the case of no price jumps turn into positive biases in the two cases of price jumps. In the simulation scenarios with price jumps, the estimator $\widehat{F}_{T,n}$ contains biases both due to the regularization error and due to the separation of volatility from jumps. The bias due to the presence of price jumps is positive and dominates the bias due to the regularization error.

Table 2: Monte Carlo Results In Presence of Low Activity Jump Component

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: Regularized Laplace Inversion with $R_n = 3.0$									
$V_0 = Q^V(0.25)$	0.1703	0.0195	0.0318	0.2806	0.0705	0.0754	0.4564	0.1693	0.1858
$V_0 = Q^V(0.50)$	0.3234	0.0241	0.0505	0.5122	0.0961	0.1098	0.8139	0.2062	0.2408
$V_0 = Q^V(0.75)$	0.6235	0.0271	0.0854	0.9686	0.1308	0.1691	1.5180	0.2753	0.3216
Panel B: Direct Method with Constant Threshold $3\Delta_n^{0.49}$									
$V_0 = Q^V(0.25)$	0.1703	0.1028	0.1107	0.2806	0.1456	0.1880	0.4564	0.1839	0.2912
$V_0 = Q^V(0.50)$	0.3234	0.0905	0.1215	0.5122	0.0999	0.2142	0.8139	0.0608	0.3560
$V_0 = Q^V(0.75)$	0.6235	0.0272	0.1523	0.9686	-0.0466	0.2958	1.5180	-0.2538	0.5589
Panel C: Direct Method with Adaptive Threshold $3\sqrt{BV_j}\Delta_n^{0.49}$									
$V_0 = Q^V(0.25)$	0.1703	0.0576	0.0625	0.2806	0.0878	0.0960	0.4564	0.1294	0.1443
$V_0 = Q^V(0.50)$	0.3234	0.0595	0.0716	0.5122	0.0947	0.1148	0.8139	0.1421	0.1767
$V_0 = Q^V(0.75)$	0.6235	0.0486	0.0875	0.9686	0.0896	0.1414	1.5180	0.1554	0.2174
Panel D: Direct Method with Adaptive Threshold $4\sqrt{BV_j}\Delta_n^{0.49}$									
$V_0 = Q^V(0.25)$	0.1703	0.0852	0.0892	0.2806	0.1345	0.1394	0.4564	0.2018	0.2108
$V_0 = Q^V(0.50)$	0.3234	0.0935	0.1024	0.5122	0.1488	0.1622	0.8139	0.2226	0.2443
$V_0 = Q^V(0.75)$	0.6235	0.0893	0.1179	0.9686	0.1542	0.1870	1.5180	0.2441	0.2876

Note: Description as for Table 1.

Table 3: Monte Carlo Results In Presence of High Activity Jump Component

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: Regularized Laplace Inversion with $R_n = 3.0$									
$V_0 = Q^V(0.25)$	0.1619	0.0726	0.0833	0.2706	0.1526	0.1626	0.4579	0.2833	0.2992
$V_0 = Q^V(0.50)$	0.3084	0.0879	0.0997	0.4936	0.1708	0.1849	0.8169	0.3004	0.3371
$V_0 = Q^V(0.75)$	0.5985	0.0933	0.1278	0.9326	0.1888	0.2117	1.5245	0.3127	0.3944
Panel B: Direct Method with Constant Threshold $3\Delta_n^{0.49}$									
$V_0 = Q^V(0.25)$	0.1619	0.1666	0.1678	0.2706	0.2107	0.2410	0.4579	0.2317	0.3411
$V_0 = Q^V(0.50)$	0.3084	0.1556	0.1675	0.4936	0.1700	0.2562	0.8169	0.1062	0.4104
$V_0 = Q^V(0.75)$	0.5985	0.1062	0.1694	0.9326	0.0361	0.3048	1.5245	−0.2139	0.6120
Panel C: Direct Method with Adaptive Threshold $3\sqrt{BV_j}\Delta_n^{0.49}$									
$V_0 = Q^V(0.25)$	0.1619	0.1338	0.1347	0.2706	0.1706	0.1802	0.4579	0.2133	0.2302
$V_0 = Q^V(0.50)$	0.3084	0.1331	0.1375	0.4936	0.1726	0.1899	0.8169	0.2130	0.2541
$V_0 = Q^V(0.75)$	0.5985	0.1191	0.1361	0.9326	0.1621	0.2001	1.5245	0.2000	0.2951
Panel D: Direct Method with Adaptive Threshold $4\sqrt{BV_j}\Delta_n^{0.49}$									
$V_0 = Q^V(0.25)$	0.1619	0.1586	0.1590	0.2706	0.2123	0.2201	0.4579	0.2758	0.2897
$V_0 = Q^V(0.50)$	0.3084	0.1609	0.1642	0.4936	0.2173	0.2315	0.8169	0.2790	0.3106
$V_0 = Q^V(0.75)$	0.5985	0.1527	0.1643	0.9326	0.2137	0.2452	1.5245	0.2745	0.3508

Note: Description as for Table 1.

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 shows 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 $\ln(V_t)$. 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 variation of the (transformed) volatility process. The aim of these plots is to discover how the dispersion of volatility relates to the volatility level. We see that for both data sets, the IQRs of the spot variance and the spot volatility exhibit a clearly positive, and generally 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 is 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).[13] 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 $h(V_t)$ 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 variation 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 (13). This is consistent with earlier parametric evidence for superior performance of log-volatility models over affine models.[14]

[13]This also holds approximately true for two-factor models in which one of the factors is fast mean reverting and the other is very persistent (which is the case for most of the estimates of such models reported in empirical work). In such a setting, the fast mean reverting factor plays minimal role in the dependence of the interquantile range of various transforms of the spot variance over the interval on the level of volatility.

[14]Regarding log volatility, Chernov et al. (2003) present evidence from time series data while Cont and da Fonseca (2002) present evidence from the options-implied volatility surface.

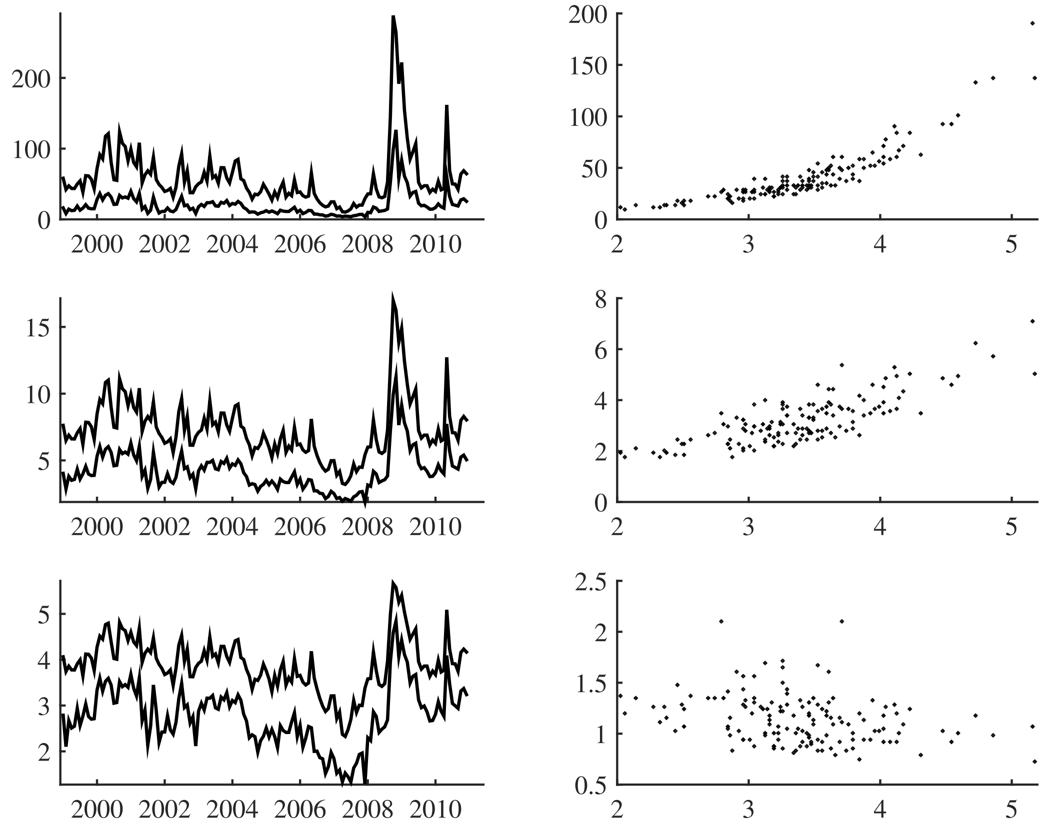

Figure 1: *Estimated Quantiles of the Monthly Occupation Measure of the Spot Volatility of the Euro/$ return, 1999–2010.* The three left-hand panels show the 25 and 75 percent quantiles of the monthly occupation measure of volatility expressed in terms of the local variance (left-top), the local standard 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 log-variance). Volatility is quoted annualized and in percentage terms.

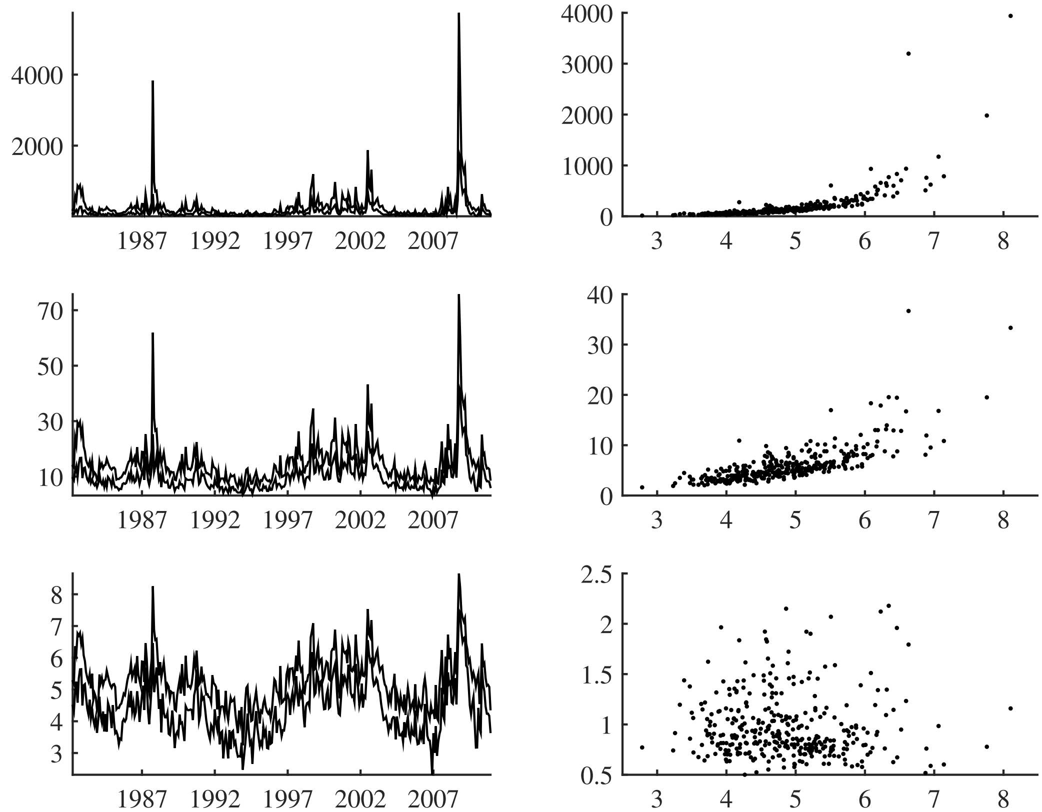

Figure 2: *Estimated Quantiles of the Monthly Occupation Measure of the Spot Volatility of the S&P500 index futures return, 1982–2010.* The organization is the same as Figure 1.

6 Conclusion

In this paper we use inverse Laplace transforms to generate a quick and easy nonparametric estimator of the volatility occupation time (VOT). The estimation is conducted based on discretely sampled Itô semimartingale increments over a fixed time interval with asymptotically shrinking mesh of the observation grid. We derive the asymptotic properties of the VOT estimator locally uniformly in the spatial argument and further invert it to estimate the corresponding quantiles of volatility over the time interval. Monte Carlo evidence shows good finite-sample performance that is significantly better than that of the benchmark estimator of Li et al. (2013) for estimating lower volatility quantiles. An empirical application illustrates the use of the estimator for studying the variation of volatility.

7 Appendix: Proofs

The appendix is organized as follows. We collect some preliminary estimates in Section 7.1. The rest of this 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 7.1 *Fix $c > 0$, $\eta_1 \in [0, 1/2)$ and $\eta_2 \in [0, 1/2)$. There exists some $K > 0$, such that for any $R \geq c$ and $x > 0$,*

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

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

$$|g_R(x)| \leq \int_0^\infty \frac{\sqrt{s}\,|\sin(xs)|}{s^2+1} ds \leq \int_0^\infty \frac{\sqrt{s}\,|\sin(xs)|^{\eta_1}}{s^2+1} ds \leq K x^{\eta_1}. \tag{14}$$

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

$$|g_R(x)| \leq K R x^{-1}. \tag{15}$$

Using integration by parts again, we get $g_R(x) = -x^{-2} \int_0^\infty h_R''(s) \sin(xs)\, ds$. By explicit computation, it is easy to see $|h_R''(s)| \leq K R^2 s^{-3/2}/(1+s^2)$. Hence, for $\eta_2 \in [0, 1/2)$,

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

Combining the inequality above with (14) and (15), we derive

$$|g_R(x)| \leq K \min\left\{x^{\eta_1}, R x^{-1}, R^2 x^{-1-\eta_2}\right\}.$$

Similarly, we can also show that

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

The assertion of the lemma then readily follows. ■

7.1.2 Estimates for the underlying process X

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

Assumption SA: We have Assumption A. Moreover, the processes b_t, V_t and V_t^{-1} and the sequence K_m are bounded, and for some bounded λ-integrable deterministic function Γ on $\mathbb{R}$, we have $|\delta(\omega, t, z)|^r \leq \Gamma(z)$.

For notational simplicity, we set

$$b'_t = \begin{cases} b_t & \text{if} \quad r > 1, \\ b_t - \int_{\mathbb{R}} \delta(t,z) \, 1_{\{|\delta(t,z)| \leq 1\}} \lambda(dz) & \text{if} \quad r \leq 1, \end{cases}$$

where r is the constant in Assumption A1. We also set $\sigma_t = \sqrt{V_t}$, $X'_t = X_0 + \int_0^t b'_s ds + \int_0^t \sigma_s dW_s$, $X''_t = X_t - X'_t$, and

$$\chi_i^n = \Delta_i^n X' / \Delta_n^{1/2}, \quad \beta_i^n = \sigma_{(i-1)\Delta_n} \Delta_i^n W / \Delta_n^{1/2}, \quad \lambda_i^n = \chi_i^n - \beta_i^n.$$

Lemma 7.2 *Under Assumption SA, there exists $K > 0$ such that for all $u \in \mathbb{R}_+$,*

$$\mathbb{E} \left| \Delta_n \sum_{i=1}^{[T/\Delta_n]} \left(\cos\left(\sqrt{2u}\chi_i^n\right) - \cos\left(\sqrt{2u}\beta_i^n\right) \right) \right| \leq K \min\left\{ u^{1/2} \Delta_n^{1/2}, 1 \right\}, \tag{16}$$

$$\mathbb{E} \left| \Delta_n \sum_{i=1}^{[T/\Delta_n]} \exp\left(-uV_{(i-1)\Delta_n}\right) - \int_0^T \exp\left(-uV_s\right) ds \right| \leq K \min\left\{ u \Delta_n^{1/2}, 1 \right\} + \Delta_n, \tag{17}$$

$$\mathbb{E} \left| \Delta_n \sum_{i=1}^{[T/\Delta_n]} \left(\cos\left(\sqrt{2u}\beta_i^n\right) - \exp\left(-uV_{(i-1)\Delta_n}\right) \right) \right| \leq K \Delta_n^{1/2} \min\left\{u, 1\right\}, \tag{18}$$

$$\mathbb{E} \left| \Delta_n \sum_{i=1}^{[T/\Delta_n]} \left(\cos\left(\sqrt{2u} \frac{\Delta_i^n X}{\sqrt{\Delta_n}}\right) - \cos\left(\sqrt{2u}\chi_i^n\right) \right) \right| \leq K \left(u^{1/2} \Delta_n^{1/r - 1/2} \right)^{r \wedge 1}. \tag{19}$$

Proof. By the Burkholder–Davis–Gundy inequality and Assumption SA, $\mathbb{E}\left|\lambda_i^n\right| \leq K \Delta_n^{1/2}$. Then (16) follows from a mean-value expansion and the triangle inequality. Turning to (17), we have, for $s \in [(i-1)\Delta_n, i\Delta_n]$, $\mathbb{E}\left|\exp\left(-uV_{(i-1)\Delta_n}\right) - \exp\left(-uV_s\right)\right| \leq K u \Delta_n^{1/2}$, by using a mean-value expansion and Assumption SA. By the triangle inequality, (17) readily follows. Now, consider (18). Denote $\zeta_i^n = \cos\left(\sqrt{2u}\beta_i^n\right) - \exp\left(-uV_{(i-1)\Delta_n}\right)$. It is easy to see that $(\zeta_i^n, \mathcal{F}_{i\Delta_n})_{i \geq 1}$ forms an array of martingale differences. Moreover,

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

Hence, $\mathbb{E}[(\Delta_n \sum_{i=1}^{[T/\Delta_n]} \zeta_i^n)^2] \leq K \Delta_n \min\{u^2, 1\}$. We then deduce (18) by using Jensen's inequality. Finally, we show (19). When $r \in (0,1]$, by Assumption SA and Lemma 2.1.7

in Jacod and Protter (2012),

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

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 $\mathbb{E}|\cos(\sqrt{2u}\Delta_i^n X/\sqrt{\Delta_n})-\cos(\sqrt{2u}\chi_i^n)|^r \leq Ku^{r/2}\Delta_n^{1-r/2}$, and then use Jensen's inequality to get $\mathbb{E}|\cos(\sqrt{2u}\Delta_i^n X/\sqrt{\Delta_n})-\cos(\sqrt{2u}\chi_i^n)| \leq Ku^{1/2}\Delta_n^{1/r-1/2}$. 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 (19) readily follows. ■

7.2 Proof of Lemma 2.1

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

$$\mathbb{E}\left[\sup_{t\leq T\wedge T_m}\left|\tilde{f}_t(x)-\tilde{f}_t(y)\right|^k\right] \leq K|x-y|^{(1-\beta)k\wedge(1/2)}. \tag{20}$$

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

$$\tilde{f}_t(y)-\tilde{f}_t(x) = 2\sum_{j=1}^{5} A_t^{(j)}, \tag{21}$$

where

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

Clearly, for any t, $|A_t^{(1)}| \leq 2\,|x-y|$. Hence, $\mathbb{E}[\sup_{t\leq T\wedge T_m} |A_t^{(1)}|^k] \leq K|x-y|^k$.

By (5), we have

$$A_t^{(2)} = \int_0^t 1_{\left\{x<\widetilde{V}_{s-}\leq y\right\}} \left(\tilde{a}_s ds + dB_s\right) + \int_0^t \int_{\mathbb{R}} 1_{\left\{x<\widetilde{V}_{s-}\leq y\right\}} \tilde{\delta}\left(s,z\right)\mu\left(ds,dz\right). \tag{22}$$

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

$$\mathbb{E}\left[\sup_{t\leq T\wedge T_m}\left|\int_0^t 1_{\left\{x<\widetilde{V}_{s-}\leq y\right\}} \tilde{a}_s ds\right|^k\right] \leq K\left|x-y\right|. \tag{23}$$

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

$$\mathbb{E}\left[\sup_{t\leq T\wedge T_m}\left|\int_0^t 1_{\left\{x<\widetilde{V}_{s-}\leq y\right\}} dB_s\right|^k\right] \leq K\mathbb{E}\left[\left(\int_0^T 1_{\left\{x<\widetilde{V}_{s-}\leq y\right\}} ds\right)^{k/2}\right] \leq K\left|x-y\right|^{1/2}. \tag{24}$$

Moreover, condition (i) implies that $\int_{\mathbb{R}}(\tilde{\Gamma}_m(z)^k + \tilde{\Gamma}_m(z))\lambda(dz) < \infty$. Then by Lemma 2.1.7

of Jacod and Protter (2012), we have

$$\begin{aligned}
&\mathbb{E}\left[\sup_{t\le T\wedge T_m}\left|\int_0^t\int_{\mathbb{R}}1_{\left\{x<\widetilde{V}_{s-}\le y\right\}}\tilde{\delta}\left(s,z\right)\mu\left(ds,dz\right)\right|^k\right]\\
&\le K\mathbb{E}\left[\int_0^T\int_{\mathbb{R}}1_{\left\{x<\widetilde{V}_{s-}\le y\right\}}\tilde{\Gamma}_m\left(z\right)^k\lambda\left(dz\right)ds\right]\\
&\quad+K\mathbb{E}\left[\left(\int_0^T\int_{\mathbb{R}}1_{\left\{x<\widetilde{V}_{s-}\le y\right\}}\tilde{\Gamma}_m\left(z\right)\lambda\left(dz\right)ds\right)^k\right]\\
&\le K\left|x-y\right|.
\end{aligned}\tag{25}$$

Combining (22)-(25), we derive $\mathbb{E}\left[\sup_{t\le T\wedge T_m}|A_t^{(2)}|^k\right]\le K\left|x-y\right|^{1/2}$.

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

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

From here, we readily obtain

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

where the second inequality is obtained by using Lemma 2.1.7 of Jacod and Protter (2012) and condition (i).

Finally, since $|A_t^{(4)}|\le\int_0^t\int_{\mathbb{R}}1_{\left\{x<\widetilde{V}_{s-}\le y\right\}}|\tilde{\delta}\left(s,z\right)|\mu\left(ds,dz\right)$, the same calculation as in (25) yields $\mathbb{E}[\sup_{t\le T\wedge T_m}|A_t^{(4)}|^k]\le K\left|x-y\right|$. We have shown that $\mathbb{E}[\sup_{t\le T\wedge T_m}|A_t^{(j)}|^k]\le K|x-y|^{(1-\beta)k\wedge(1/2)}$ for each $j\in\{1,\ldots,5\}$. In view of (21), we derive (20).

It remains to show $\mathbb{E}[\tilde{f}_{T\wedge T_m}(x)^k]\le K$ for all $x\in\tilde{\mathcal{K}}$. 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 (20) and the compactness of $\tilde{\mathcal{K}}$.

Part (b). Denote $\tilde{F}_t\left(y\right)=\int_0^t1_{\left\{\widetilde{V}_s\le y\right\}}ds$. Then $F_t\left(x\right)=\tilde{F}_t\left(g\left(x\right)\right)$. By the chain rule, $F_t\left(x\right)$ is differentiable with derivative $f_t\left(x\right)=\tilde{f}_t\left(g\left(x\right)\right)g'\left(x\right)$. Assumption C1 is thus verified. Let $\mathcal{K}\subset(0,\infty)$ be compact. Since g is continuously differentiable, $g'\left(\cdot\right)$ is bounded on $\mathcal{K}$. Moreover, the set $g\left(\mathcal{K}\right)$ is compact; hence by part (a), $\mathbb{E}[|\tilde{f}_{T\wedge T_m}(g(x))|^k]$ is bounded for

$x \in \mathcal{K}$, yielding $\mathbb{E}[|f_{T\wedge T_m}(x)|^k] = \mathbb{E}[|\tilde{f}_{T\wedge T_m}(g(x))|^k]|g'(x)|^k \leq K$. By Jensen's inequality, for any $\varepsilon \in (0, k-1)$, $\sup_{x\in\mathcal{K}} \mathbb{E}[|f_{T\wedge T_m}(x)|^{1+\varepsilon}] \leq K$. This verifies Assumption C2. Moreover, for $x, y \in \mathcal{K}$,

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

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

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

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 C3 for the process V_t. ∎

7.3 Proof of Theorem 3.1

We first prove Lemmas 3.1 and 3.2, and then prove Theorem 3.1.

Proof of Lemma 3.1. By localization, we can suppose Assumption SA and strengthen Assumptions B with the additional condition that $T \leq T_m$. Since $R_n = O_p(\rho_n)$ and $R_n^{-1} = O_p(\rho_n^{-1})$, we can also assume that

$$R_n \leq M\rho_n, \quad R_n^{-1} \leq M\rho_n^{-1} \tag{26}$$

for some fixed $M \geq 1$ in the proof without loss of generality. Otherwise, we can restrict calculations on the set for which (26) hold, while noting that the probability of the exception set can be made arbitrarily small by picking M large. The proof proceeds via several steps.

Step 1. Note that the inversion kernel $\Pi(R, x)$ differs from that in Kryzhniy (2003b) (see (3) there) by a factor of $1/\coth(\pi R)$. Hence, by (4) in Kryzhniy (2003b), we can rewrite (8) as

$$F_{T,R_n}(x) = \frac{2}{\pi}\int_0^\infty F_T(xu)\sqrt{u}\frac{\sin(R_n \ln u)}{u^2-1}du. \tag{27}$$

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

$$\begin{aligned} F_{T,R_n}(x) - F_T(x) &= F_T(x)\left(\frac{2}{\pi}\int_{-\infty}^\infty \frac{e^{3z/2}\sin(R_n z)}{e^{2z}-1}dz - 1\right) \\ &\quad + \frac{2}{\pi}\int_0^\infty G_T(z;x)\sin(R_n z)\,dz, \end{aligned} \tag{28}$$

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 (28) can be bounded as follows. By direct integration, we have

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

Hence,

$$\sup_{t\le T, x\ge 0}\left|F_t(x)\left(\frac{2}{\pi}\int_{-\infty}^\infty \frac{e^{3z/2}\sin(R_n z)}{e^{2z}-1}dz - 1\right)\right| \le Ke^{-2\pi R_n} = O_p(\rho_n^{-1}). \tag{29}$$

Below, we complete the proof by showing that the second term in (28) is $O_p(\rho_n^{-1}\ln(\rho_n))$.

Step 2. We denote $a_n = \pi/2R_n$ and, without loss of generality, we suppose that $R_n \ge 1$. In this step, we show that

$$\int_0^{a_n} G_T(z;x)\sin(R_n z)\,dz = O_p(\rho_n^{-1}). \tag{30}$$

Let $A_n = \{z \mapsto F_T(xe^z)$ is differentiable on $(-M\pi/2\rho_n, M\pi/2\rho_n)\}$. By Assumption B and

(26), we have $\mathbb{P}(A_n^c) \leq K\rho_n^{-1}$ and

$$\begin{aligned}
\mathbb{E}\left|1_{A_n}\int_0^{a_n} G_T(z;x)\sin(R_n z)\,dz\right| &= \mathbb{E}\left|1_{A_n}\int_{-a_n}^{a_n}\left(F_T(xe^z)-F_T(x)\right)\frac{e^{3z/2}\sin(R_n z)}{e^{2z}-1}dz\right| \\
&\leq \mathbb{E}\left|1_{A_n}\int_{-M\pi/2\rho_n}^{M\pi/2\rho_n}\left|F_T(xe^z)-F_T(x)\right|\frac{1}{|e^z-1|}dz\right| \\
&\leq K\rho_n^{-1}.
\end{aligned}$$

From here, (30) follows.

Step 3. For each $k \geq 0$, we denote $a_{n,k} = a_n + 2\pi k/R_n$. Let $N_n = \min\{k \in \mathbb{N} : a_{n,k} \geq 3\ln(\rho_n)\}$, where we assume that $\ln(\rho_n) \geq 2\pi$ without loss of generality. Note that for $0 \leq k \leq N_n$, we have $\pi/2M\rho_n \leq a_n \leq a_{n,k} \leq 4\ln(\rho_n)$. Moreover, $N_n \leq [M\rho_n \ln(\rho_n)]$. In this step, we show that

$$\mathbb{E}\left|\int_{a_n}^{a_n+2\pi N_n/R_n} G_T(z;x)\sin(R_n z)\,dz\right| \leq K\rho_n^{-1}\ln(\rho_n). \tag{31}$$

For $k \geq 1$, we define a binary random variable $I_{n,k}$ as follows: let $I_{n,k} = 1$ if the function $z \mapsto F_T(xe^z)$ is continuously differentiable on $(a_{n,k-1}, a_{n,k})$ and $I_{n,k} = 0$ otherwise. We first note that for each $k \geq 1$, if $I_{n,k} = 1$, then

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

where $G_{T,z}(z;x) \equiv \partial G_T(z;x)/\partial z$, the first equality is obtained by a change of variable and the second equality follows an integration by parts, using that $\cos(R_n z) = 0$ for $z = a_{n,k-1}$

and $z = a_{n,k} - \pi/R_n$. Therefore,

$$\begin{aligned}
&\mathbb{E}\left|\int_{a_{n,k-1}}^{a_{n,k}} I_{n,k} G_T(z;x)\sin(R_n z)\,dz\right| \\
&\quad\leq K\rho_n^{-1}\mathbb{E}\left[\int_{a_{n,k-1}}^{a_{n,k}-\pi/R_n} I_{n,k}\left|G_{T,z}(z;x) - G_{T,z}(z+\pi/R_n;x)\right|dz\right].
\end{aligned} \tag{32}$$

To bound the integrand on the majorant side of (32), we note that $G_{T,z}(z;x) = g_{T,z}(z;x) + g_{T,z}(-z;x)$, where $g_{T,z}(z;x) \equiv \partial g_T(z;x)/\partial z$. By setting $\phi_T(x) = x f_T(x)$, we can write $g_{T,z}(z;x) = \phi_T(xe^z)h(z) + (F_T(xe^z) - F_T(x))h'(z)$. Hence, for any $y, z \in \mathbb{R}$,

$$\begin{aligned}
&|g_{T,z}(z;x) - g_{T,z}(y;x)| \\
&\quad\leq\ |\phi_T(xe^z)h(z) - \phi_T(xe^y)h(y)| \\
&\qquad + \left|(F_T(xe^z) - F_T(x))h'(z) - (F_T(xe^y) - F_T(x))h'(y)\right| \\
&\quad\leq\ |\phi_T(xe^z) - \phi_T(xe^y)|\cdot h(z) + \phi_T(xe^y)\cdot|h(y) - h(z)| \\
&\qquad + |F_T(xe^z) - F_T(xe^y)|\cdot\left|h'(z)\right| + |F_T(xe^y) - F_T(x)|\cdot\left|h'(y) - h'(z)\right|.
\end{aligned}$$

We further note that under Assumption SA, $f_T(\cdot)$ and therefore $\phi_T(\cdot)$ are supported on a compact subset of $(0,\infty)$. Then, by Assumption B, (26) and (32),

$$\mathbb{E}\left|\int_{a_{n,k-1}}^{a_{n,k}} I_{n,k} G_T(z;x)\sin(R_n z)\,dz\right| \leq K\rho_n^{-1}k^{-1}. \tag{33}$$

Next, by Assumption B1,

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_{a_{n,k-1}}^{a_{n,k}}(1-I_{n,k})G_T(z;x)\sin(R_n z)\,dz\right|\right] \\
&\leq K\mathbb{E}\left[(1-I_{n,k})\int_{a_{n,k-1}}^{a_{n,k}}|h(z)|\,dz\right] \\
&\leq Kk^{-1}\mathbb{E}[1-I_{n,k}] \\
&\leq K\rho_n^{-1}k^{-1}.
\end{aligned} \tag{34}$$

Therefore,

$$\begin{aligned}\mathbb{E}\left|\int_{a_n}^{a_n+2\pi N_n/R_n} G_T(z;x)\sin(R_n z)\,dz\right| &\leq \mathbb{E}\left[\sum_{k=1}^{N_n}\left|\int_{a_{n,k-1}}^{a_{n,k}} G_T(z;x)\sin(R_n z)\,dz\right|\right] \\ &\leq K\rho_n^{-1}\sum_{k=1}^{[M\rho_n\ln(\rho_n)]} k^{-1},\end{aligned}$$

where the first inequality is by the triangle inequality; the second inequality is by (33), (34) and $N_n \leq [M\rho_n \ln(\rho_n)]$. From here, we readily derive (31).

Step 4. Now, note that

$$\begin{aligned}\left|\int_{a_n+2\pi N_n/R_n}^{\infty} g_T(z;x)\sin(R_n z)\,dz\right| &\leq \int_{3\ln(\rho_n)}^{\infty}\left|(F_T(xe^z)-F_T(x))\frac{e^{3z/2}}{e^{2z}-1}\right|dz \\ &\leq K\int_{3\ln(\rho_n)}^{\infty} e^{-z/2}dz,\end{aligned}$$

and

$$\begin{aligned}\left|\int_{a_n+2\pi N_n/R_n}^{\infty} g_T(-z;x)\sin(R_n z)\,dz\right| &\leq \int_{3\ln(\rho_n)}^{\infty}\left|\left(F_T\left(xe^{-z}\right)-F_T(x)\right)\frac{e^{-3z/2}}{e^{-2z}-1}\right|dz \\ &\leq K\int_{3\ln(\rho_n)}^{\infty} e^{-3z/2}dz.\end{aligned}$$

Recalling $G_T(z;x) = g_T(z;x) - g_T(-z;x)$, we derive

$$\left|\int_{a_n+2\pi N_n/R_n}^{\infty} G_T(z;x)\sin(R_n z)\,dz\right| \leq K\rho_n^{-3/2}. \tag{35}$$

Combining (30), (31) and (35), we have

$$\int_0^{\infty} G_T(z;x)\sin(R_n z)\,dz = O_p\left(\rho_n^{-1}\ln(\rho_n)\right). \tag{36}$$

The assertion of the lemma then follows from (28), (29) and (36). ■

Proof of Lemma 3.2. With a standard localization procedure, we can suppose that Assumption SA holds without loss of generality. We can further suppose that $R_n \geq 2$ and

$\Delta_n \leq 1$. It is easy to see that

$$\sup_{x\in\mathcal{K}} \left| \widehat{F}_{T,n,R_n}(x) - F_{T,R_n}(x) \right| \leq \sum_{j=1}^{4} \zeta_{j,n}, \tag{37}$$

where, with the notations of Lemma 7.2 and $\Pi^*(R,u) = \sup_{x\in\mathcal{K}} |\Pi(R,ux)|$, we set

$$\zeta_{1,n} = \int_0^\infty \left| \Delta_n \sum_{i=1}^{[T/\Delta_n]} \left(\cos\left(\sqrt{2u}\chi_i^n\right) - \cos\left(\sqrt{2u}\beta_i^n\right) \right) \right| \frac{\Pi^*(R_n,u)}{u} du,$$

$$\zeta_{2,n} = \int_0^\infty \left| \Delta_n \sum_{i=1}^{[T/\Delta_n]} \exp\left(-uV_{(i-1)\Delta_n}\right) - \int_0^T \exp\left(-uV_s\right) ds \right| \frac{\Pi^*(R_n,u)}{u} du,$$

$$\zeta_{3,n} = \int_0^\infty \left| \Delta_n \sum_{i=1}^{[T/\Delta_n]} \left(\cos\left(\sqrt{2u}\beta_i^n\right) - \exp\left(-uV_{(i-1)\Delta_n}\right) \right) \right| \frac{\Pi^*(R_n,u)}{u} du,$$

$$\zeta_{4,n} = \int_0^\infty \left| \Delta_n \sum_{i=1}^{[T/\Delta_n]} \left(\cos\left(\sqrt{2u}\frac{\Delta_i^n X}{\sqrt{\Delta_n}}\right) - \cos\left(\sqrt{2u}\chi_i^n\right) \right) \right| \frac{\Pi^*(R_n,u)}{u} du.$$

By Lemma 7.1 and $R_n \leq \rho_n$, given any $\eta_1, \eta_2 \in [0, 1/2)$, we have for all $u > 0$,

$$\Pi^*(R_n,u) \leq K \exp\left(\frac{\pi\rho_n}{2}\right) \min\left\{u^{\eta_1}, \rho_n u^{-1}, \rho_n^2 u^{-1-\eta_2}\right\}, \tag{38}$$

where we have used the fact that $\mathcal{K}$ is bounded above and away from zero.

By (16) and (38), where the latter is applied with $\eta_1 = 0$ and $\eta_2 = \eta$ for η being the constant in the statement of the lemma, we have

$$\begin{aligned} \mathbb{E}|\zeta_{1,n}| &\leq K \exp\left(\frac{\pi\rho_n}{2}\right) \int_0^\infty \min\left\{u^{1/2}\Delta_n^{1/2}, 1\right\} \min\left\{u^{-1}, \rho_n u^{-2}, \rho_n^2 u^{-2-\eta}\right\} du \\ &\leq K \exp\left(\frac{\pi\rho_n}{2}\right) \left(\rho_n^{1/2}\Delta_n^{1/2} + \rho_n^2 \Delta_n^{1+\eta}\right), \end{aligned} \tag{39}$$

where the second line is obtained by a direct (but somewhat tedious) calculation of the integral above.

By (17) and (38), for $\eta_1 \in (0, 1/2)$, we have

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

Moreover, by direction calculation, we have

$$\int_0^\infty \min\left\{u^{\eta_1-1}, \rho_n u^{-2}, \rho_n^2 u^{-2-\eta}\right\}du \leq K\rho_n^{\eta_1/(1+\eta_1)} + K\rho_n^{1-1/\eta},$$

and

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

Hence,

$$\mathbb{E}\left|\zeta_{2,n}\right| \leq K\exp\left(\frac{\pi\rho_n}{2}\right)\left(\rho_n \ln(\rho_n)\Delta_n^{1/2} + \rho_n^2\Delta_n^{(1+\eta)/2}\right). \tag{40}$$

By (18) and (38), we have

$$\begin{aligned}
\mathbb{E}\left|\zeta_{3,n}\right| &\leq K\Delta_n^{1/2}\exp\left(\frac{\pi\rho_n}{2}\right)\int_0^\infty \min\{u, 1\}\min\left\{u^{-1}, \rho_n u^{-2}, \rho_n^2 u^{-2-\eta}\right\}du \\
&\leq K\exp\left(\frac{\pi\rho_n}{2}\right)\ln(\rho_n)\Delta_n^{1/2}. \qquad (41)
\end{aligned}$$

Now, consider $\zeta_{4,n}$. We denote $\tilde{r} = r \wedge 1$. By (19) and (38), and a direct calculation of the integral below, we have

$$\begin{aligned}
\mathbb{E}\left|\zeta_{4,n}\right| &\leq K\Delta_n^{\tilde{r}(1/r-1/2)}\exp\left(\frac{\pi\rho_n}{2}\right)\int_0^\infty u^{\tilde{r}/2}\min\left\{u^{-1}, \rho_n u^{-2}, \rho_n^2 u^{-2-\eta}\right\}du \\
&\leq K\exp\left(\frac{\pi\rho_n}{2}\right)\rho_n^{\tilde{r}/2}\Delta_n^{\tilde{r}(1/r-1/2)}. \qquad (42)
\end{aligned}$$

Finally, we combine (37) and (39)–(42) to derive the assertion of the lemma. ■

Proof of Theorem 3.1. The assertion follows directly from Lemmas 3.1 and 3.2. ■

7.4 Proof of Theorem 3.2.

By localization, we can assume (10) holds for $T_m = T$ and $C_m = K$ without loss of generality. Since $\rho_n \to \infty$, we can also suppose $\rho_n \geq 1$. Let $\Pi^*(R_n, u)$ be defined as in the proof of Lemma 3.2. Recall from (38) that, given any $\eta_1, \eta_2 \in [0, 1/2)$, we have for all $u > 0$,

$$u^{-1}\Pi^*(R_n, u) \leq K \exp\left(\frac{\pi\rho_n}{2}\right) \min\left\{u^{-1+\eta_1}, \rho_n u^{-2}, \rho_n^2 u^{-2-\eta_2}\right\}. \tag{43}$$

Below, we fix η_1 and η_2 such that $\eta_1 \in (\bar{c}, 1/2)$ and $\eta_2 \in (\max\{\bar{c}, 1/3\}, 1/2)$.

We start with an estimate. Let $a \in (-\eta_1, 1 + \eta_2)$ be a constant. Straightforward algebra yields

$$\begin{aligned}
&\int_0^\infty u^a \min\left\{u^{-1+\eta_1}, \rho_n u^{-2}, \rho_n^2 u^{-2-\eta_2}\right\} du \\
&= \int_0^{\rho_n^{1/(1+\eta_1)}} u^{-1+\eta_1+a} du + \rho_n \int_{\rho_n^{1/(1+\eta_1)}}^{\rho_n^{1/\eta_2}} u^{a-2} du + \rho_n^2 \int_{\rho_n^{1/\eta_2}}^\infty u^{a-2-\eta_2} du \\
&\leq K\left(\rho_n^{\frac{a+\eta_1}{1+\eta_1}} + \rho_n^{1+\frac{a-2}{1+\eta_1}+\frac{1}{\eta_2}} + \rho_n^{\frac{a-1+\eta_2}{\eta_2}}\right) \\
&\leq K\rho_n^4.
\end{aligned} \tag{44}$$

Now, observe that for any compact $\mathcal{K} \subset (0, \infty)$,

$$\begin{aligned}
\mathbb{E}\left[\sup_{x\in\mathcal{K}}\left|\widetilde{F}_{T,n,R_n}(x) - F_{T,R_n}(x)\right|\right] &\leq \int_0^\infty \mathbb{E}\left|\widetilde{\mathcal{L}}_{T,n}(u) - \mathcal{L}_T(u)\right| \frac{\Pi^*(R_n, u)}{u} du \\
&\leq K\Delta_n^{\bar{\delta}} \int_0^\infty \left(u^{-\bar{c}} + u^{1+\bar{c}}\right) \frac{\Pi^*(R_n, u)}{u} du \\
&\leq K \exp\left(\frac{\pi\rho_n}{2}\right) \Delta_n^{\bar{\delta}} \rho_n^4,
\end{aligned} \tag{45}$$

where the first inequality is by the triangle inequality; the second inequality follows from (10); the third inequality follows from (43) and (44). Under condition (ii), the majorant side of the above display goes to zero. The assertion of the theorem then readily follows from Lemma 3.1 and (45). ■

7.5 Proof of Theorem 3.3

Part (a) Let $(T_m)_{m\geq1}$ be the localizing sequence of stopping times as in Assumption C. By localization, we can suppose that the stopped process $(V_{t\wedge T_m})_{t\geq0}$ 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 $x, y \in \mathcal{K}$ and $|z| \geq M_m$ imply $F_{T\wedge T_m}(xe^z) - F_{T\wedge T_m}(ye^z) = 0$. Hence, for any $x, y \in \mathcal{K}$,

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_{-\infty}^{\infty}\left(F_{T\wedge T_m}(xe^z) - F_{T\wedge T_m}(x)\right.\right.\right. \\
&\quad \left.\left.\left.- \left(F_{T\wedge T_m}(ye^z) - F_{T\wedge T_m}(y)\right)\right)\frac{e^{3z/2}\sin(R_n z)}{e^{2z}-1}dz\right|^{1+\varepsilon}\right] \\
&\leq K\mathbb{E}\left[|F_{T\wedge T_m}(x) - F_{T\wedge T_m}(y)|^{1+\varepsilon}\left|\int_{(-\infty,-M_m]\cup[M_m,\infty)}\frac{e^{3z/2}\sin(R_n z)}{e^{2z}-1}dz\right|^{1+\varepsilon}\right] \\
&\quad +K\mathbb{E}\left[\left|\int_{-M_m}^{M_m}\Big(F_{T\wedge T_m}(xe^z) - F_{T\wedge T_m}(x)\right.\right. \\
&\quad \left.\left.- \left(F_{T\wedge T_m}(ye^z) - F_{T\wedge T_m}(y)\right)\Big)\frac{e^{3z/2}\sin(R_n z)}{e^{2z}-1}dz\right|^{1+\varepsilon}\right].
\end{aligned} \tag{46}$$

By Assumption C2 and Hölder's inequality,

$$\begin{aligned}
&\mathbb{E}\left[|F_{T\wedge T_m}(x) - F_{T\wedge T_m}(y)|^{1+\varepsilon}\left|\int_{(-\infty,-M_m]\cup[M_m,\infty)}\frac{e^{3z/2}\sin(R_n z)}{e^{2z}-1}dz\right|^{1+\varepsilon}\right] \\
&\leq K_m\mathbb{E}\left[|F_{T\wedge T_m}(x) - F_{T\wedge T_m}(y)|^{1+\varepsilon}\right] \\
&\leq K_m|x-y|^{1+\varepsilon}.
\end{aligned} \tag{47}$$

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

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

Thus by Hölder's inequality, the second term on the majorant side of (46) can be bounded by

$$\begin{aligned}
&K_m\mathbb{E}\left[\int_{-M_m}^{M_m}\left(\int_x^y f_{T\wedge T_m}(ve^z)|e^z-1|\,dv\right)^{1+\varepsilon}\left|\frac{e^{3z/2}}{e^{2z}-1}\right|^{1+\varepsilon}dz\right] \\
&+K_m\mathbb{E}\left[\int_{-M_m}^{M_m}\left(\int_x^y |f_{T\wedge T_m}(ve^z) - f_{T\wedge T_m}(v)|\,dv\right)^{1+\varepsilon}\left|\frac{e^{3z/2}}{e^{2z}-1}\right|^{1+\varepsilon}dz\right].
\end{aligned} \tag{48}$$

The two terms in (48) can be further bounded as follows. By Hölder's inequality and Assumption C2, the first term in (48) can be bounded by

$$K_m \mathbb{E}\left[\int_{-M_m}^{M_m}\left(\int_x^y f_{T\wedge T_m}\left(ve^z\right)dv\right)^{1+\varepsilon}dz\right] \le K_m\left|x-y\right|^{1+\varepsilon}.$$

By Hölder's inequality and Assumption C3, the second term in (48) can be bounded by

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

Thus, (48) can be further bounded by $K_m\left|x-y\right|^{1+\varepsilon}$. Combining this estimate with (47), we deduce that

$$\text{The left-hand side of (46)} \le K_m\left|x-y\right|^{1+\varepsilon}. \tag{49}$$

Next, observe that by Assumption C2,

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

Equipped with (49) and the estimate displayed above, 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 (29), we see that the processes $F_{T\wedge T_m,R_n}(x) - F_{T\wedge T_m}(x)$, $x \in \mathcal{K}$, $n \geq 1$, are also stochastically equicontinuous. Combining this result with Lemma 3.2, 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 Theorem 3.1, $(\widehat{F}_{T,n,R_n}(x) - F_T(x))1_{\{T\leq T_m\}} = o_p(1)$ 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.

Part (b) The proof of part (b) is the same as part (a) except for the following difference: instead of using Lemma 3.2 and Theorem 3.1, we use (45) and Theorem 3.2 in the proof.■

7.6 Proof of Theorem 3.4

Part (a). Let $\mathbb{N}_1$ be an arbitrary subsequence of $\mathbb{N}$. Under the condition of the theorem, there exists a further subsequence $\mathbb{N}_2 \subseteq \mathbb{N}_1$, such that $\sup_{x\in\mathcal{K}} |\widehat{F}_{T,n,R_n}(x) - F_T(x)| \to 0$ 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 $Q_T^{\mathcal{K}}(\cdot)$. We have $1_{\{\widehat{F}_{T,n,R_n}(x)<\tau\}} \to 1_{\{F_T(x)<\tau\}}$ along $\mathbb{N}_2$ for $x \in \{x \in \mathcal{K} : F_T(x) \neq \tau\}$. For each $\tau \in \mathcal{T}$, the set $\{x \in \mathcal{K} : F_T(x) = \tau\}$ charges zero Lebesgue measure. By bounded convergence, along $\mathbb{N}_2$,

$$\widehat{Q}_{T,n,R_n}^{\mathcal{K}}(\tau) \to \inf\mathcal{K} + \int_{\inf\mathcal{K}}^{\sup\mathcal{K}} 1_{\{F_T(x)<\tau\}} dx = Q_T^{\mathcal{K}}(\tau), \quad \forall \tau \in \mathcal{T}. \tag{50}$$

Since $F_T(x)$ is continuous in x, by Lemma 21.2 of van der Vaart (1998) and (50), we have $\widehat{F}_{T,n,R_n}^{\mathcal{K}}(x) \to F_T(x)$ along $\mathbb{N}_2$ for all $x \in \mathcal{K}$. Since $\widehat{F}_{T,n,R_n}^{\mathcal{K}}(\cdot)$ is also increasing, we further have $\sup_{x\in\mathcal{K}} |\widehat{F}_{T,n,R_n}^{\mathcal{K}}(x) - F_T(x)| \to 0$ along $\mathbb{N}_2$.

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

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 $Q_T^{\mathcal{K}}(\cdot)$, by (50), we have $\widehat{Q}_{T,n,R_n}^{\mathcal{K}}(\tau^*) \to Q_T^{\mathcal{K}}(\tau^*)$ along $\mathbb{N}_2$ almost surely. The assertion of part (b) then follows a subsequence argument as in part (a). ■

References

Aït-Sahalia, Y. and J. Jacod (2009). Estimating the degree of activity of jumps in high frequency financial data. *Annals of Statistics 37*, 2202–2244.

Andersen, T. G., T. Bollerslev, P. F. Christoffersen, and F. X. Diebold (2013). Financial risk measurement for financial risk management. In G. Constanides, M. Harris, and R. Stulz (Eds.), *Handbook of the Economics of Finance, Vol.II*. Elsevier Science B.V.

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

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.

Black, F. (1976). Studies of stock price volatility changes. *Proceedings of the Business and Economics Section of the American Statistical Association*, 177–181.

Bollerslev, T. and V. Todorov (2011). Estimation of Jump Tails. *Economtrica 79*, 1727–17783.

Carrasco, M. and J.-P. Florens (2011). A Spectral Method for Deconvolving a Density. *Econometric Theory 27*(3), pp. 546–581.

Carrasco, M., J.-P. Florens, and E. Renault (2007). Linear inverse problems in structural econometrics estimation based on spectral decomposition and regularization. In *Handbook of Econometrics*, Volume 6B. Elsevier.

Chernov, M., A. Gallant, E. Ghysels, and G. Tauchen (2003). Alternative Models for Stock Price Dynamics. *Journal of Econometrics 116*, 225–257.

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

Clément, E., S. Delattre, and A. Gloter (2013). An infinite dimensional convolution theorem with applications to the efficient estimation of the integrated volatility. *Stochastic Processes and Their Applications 123*, 2500–2521.

Comte, F. and E. Renault (1996). Long memory continuous time models. *Journal of Econometrics 73*, 101–149.

Comte, F. and E. Renault (1998). Long memory in continuous-time stochastic volatility models. *Mathematical Finance 8*, 291–323.

Cont, R. and J. da Fonseca (2002). Dynamics of implied volatility surfaces. *Quantitative Finance 2*(1), 45–60.

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

Foster, D. and D. B. Nelson (1996). Continuous record asymptotics for rolling sample variance estimators. *Econometrica 64*, 139–174.

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

Härdle, W. and O. Linton (1994). Applied nonparametric methods. In R. F. Engle and D. L. McFadden (Eds.), *Handbook of Econometrics*, Volume 4, Chapter 38. Elsevier.

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.

Jacod, J. and M. Reiß (2013). A remark on the rates of convergence for integrated volatility estimation in the presence of jumps. *Annals of Statistics, Forthcoming*.

Jacod, J. and M. Rosenbaum (2013). Quarticity and Other Functionals of Volatility: Efficient Estimation. *Annals of Statistics 41*, 1462–1484.

Kristensen, D. (2010). Nonparametric filtering of the realized spot volatility: A kernel-based approach. *Econometric Theory 26*(1), pp. 60–93.

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.

Li, J., V. Todorov, and G. Tauchen (2013). Volatility occupation times. *Annals of Statistics 41*, 1865–1891.

Mancini, C. (2001). Disentangling the jumps of the diffusion in a geometric jumping Brownian motion. *Giornale dell'Istituto Italiano degli Attuari LXIV*, 19–47.

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.

Renault, E., C. Sarisoy, and B. J. Werker (2014). Efficient estimation of integrated volatility and related processes. Technical report, Brown University.

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 107*, 622–635.

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.