The Realized Laplace Transform of Volatility*

Viktor Todorov † and George Tauchen ‡

February 25, 2010

Abstract

We introduce a new measure constructed from high-frequency financial data which we call the Realized Laplace Transform of volatility. The statistic provides a nonparametric estimate for the empirical Laplace transform of the latent stochastic volatility process over a given interval of time. When a long span of data is used, i.e., under joint long-span and fill-in asymptotics, it is an estimate of the volatility Laplace transform. The asymptotic behavior of the statistic depends on the small scale behavior of the driving martingale. We derive the asymptotics both in the case when the latter is known and when it needs to be inferred from the data. When the underlying process is a jump-diffusion our statistic is robust to jumps and when the process is pure-jump it is robust to presence of less active jumps. We apply our results to simulated and real financial data.

Keywords: Laplace transform, stochastic volatility, Central Limit Theorem, activity index, jumps, high-frequency data.

JEL classification: C51, C52, G12.

*We wish to thank Torben Andersen, Tim Bollerslev, Ron Gallant, Nour Meddahi, Per Mykland, Andrew Patton, Mark Podolskij, Neil Shephard and seminar participants at Duke University, Oxford-Man Institute and IIT for helpful comments.

†Department of Finance, Kellogg School of Management, Northwestern University, Evanston, IL 60208; e-mail: v-todorov@kellogg.northwestern.edu

‡Department of Economics, Duke University, Durham, NC 27708; e-mail: george.tauchen@duke.edu.

1 Introduction

Day-to-day moves in asset prices are typically modeled with Gaussian shocks that have a slowly moving over time stochastic volatility plus occasional big jumps. The persistent market stochastic volatility typically captures the state of the economy and is a manifest of the time-varying opportunity set that investors face. This time-variation has nontrivial economic applications as investors seek to optimally allocate risks across time by hedging today against inverse changes in the investment opportunities tomorrow.

However, recent empirical evidence suggests that there are periods when the volatility can quickly increase over short time and then revert back to its pre-crisis level. This higher frequency moves in volatility are typically triggered by a jump in the stock market itself. Recent evidence from options markets indicate that the presence of this risk associated with the stochastic volatility can have a significant economic implications in helping explain the large magnitudes of various risk premia on the asset markets, see e.g., Bollerslev and Todorov (2009).

While recovering the spot volatility at a given point in time is theoretically possible by using a local window of high-frequency data, e.g., Foster and Nelson (1996), such estimation is very challenging as reflected by its slow rate of convergence (it is $\sqrt{n}$ where n is the number of observations in the local window but n increases at a slower rate (typically twice as slow) than the rate at which the mesh of the observation grid shrinks to zero, see Jacod and Todorov (2010)). An alternative, that has received a lot of attention recently (Andersen et al. (2001, 2003), Barndorff-Nielsen and Shephard (2002)), is to estimate instead integrals over time of the volatility process using the high-frequency data. This can be done by the Realized Variance, which is just the sum of high-frequency squared returns over a given period of time, and its jump-robust extensions (the Multipower Variations proposed by Barndorff-Nielsen and Shephard (2004) and the Truncated Variance of Mancini (2009)). By aggregating over time, a lot of the errors in estimating the spot volatility get "canceled" and this provides in most cases $\sqrt{n}$ rate of convergence for the integrated volatility estimators (for n denoting the total number of high-frequency observations on the time interval).

However, inferring distributional properties of the spot volatility directly from those of the integrated measures of volatility is difficult due to the time aggregation, e.g., the mapping between the probability distribution of the spot and integrated volatility is not one-to-one in general. In this paper we address this issue by proposing an alternative way of aggregating the high-frequency data into a realized measure which we call Realized Laplace Transform of volatility. Our measure estimates the empirical Laplace transform of the volatility over a given interval of time and thus preserves information about the characteristics of spot volatility. For example, when a long span of data is used, we are able to estimate the Laplace transform of the stochastic volatility which identifies uniquely its marginal distribution. The latter is of central importance for building stochastic volatility models of non-Gaussian type suggested by Barndorff-Nielsen and Shephard (2001) (see also references therein) as well as for the nonparametric estimation of the diffusion function in the context of Markov diffusion processes proposed by Ait-Sahalia (1996).

Similar to the (jump robust) analogues of the Realized Variance, the Realized Laplace Transform of volatility has typically $\sqrt{n}$ rate of convergence. Further, in the context of jump-diffusion models it posses robustness with respect to jumps as good as (and even slightly better than) that of the Truncated Variance.

We recall that for a generic non-negative random variable X with distribution function F and scalar $u \geq 0$, the real Laplace transform of X is given by[1]

$$\mathcal{L}_X(u) = \mathbb{E}\left(e^{-uX}\right) = \int_0^\infty e^{-ux}\, dF(x). \tag{1}$$

The family of functions $\{e^{ux}\}_{u\geq 0}$ is separating within the class of distribution functions supported on $[0,\infty)$ (Breiman, 1968, Proposition 8.51, p. 183), so the mapping from $F(x)$ to $\mathcal{L}_X(u)$ is one-to-one. Thus, any information in F is embodied in $\mathcal{L}_X$ as well and vice versa. Evidently $0 \leq \mathcal{L}_X(u) \leq 1$, and we further denote the log of the Laplace transform with $\ell_X(u) = \log\left[\mathcal{L}_X(u)\right]$ for $u \geq 0$. We have:

$$\begin{aligned} \text{location scale shift} \quad & Y = a + bX,\ b > 0 \Rightarrow \ell_Y(u) = -a + \ell_X(bu), \\ \text{mean preserving spread} \quad & Y = X + V,\ V \text{ independent of } X,\ \mathbb{E}\left(V\right) = 0, \Rightarrow \\ & \ell_Y(u) = \ell_X(u) + \ell_V(u), \\ \text{log convexity} \quad & \ell_X(u) \text{ is a convex function of } u \text{ (Boyd, 2004, Example 3.41, p. 106).} \end{aligned}$$

In general, the behavior of the log-Laplace transform near the origin conveys information about the tail behavior of X, and, vice versa, the tail behavior of the log-Laplace transform tells us about the behavior of the distribution of X near the origin; (Rozovsky, 2009) contains some interesting new results along these lines.

Given *observed* discrete-time data set $\{X\}_{t=1}^{T}$, the *empirical Laplace transform*, i.e., the Laplace transform of X_t with respect to the empirical measure, is simple to obtain,[2]

$$\hat{\mathcal{L}}_X(u) = \frac{1}{T}\sum_{t=1}^{T} e^{-uX_t}, \quad \hat{\ell}_X(u) = \log\left[\hat{\mathcal{L}}_X(u)\right]. \tag{2}$$

[1]Formally, this is the Lebesgue-Stieltjes unilateral Laplace transform.

[2]The continuous-time analogue, i.e., when we have a continuous record of X_t for $t \in [0,T]$, is $\hat{\mathcal{L}}_X(u) = \frac{1}{T}\int_0^T e^{-uX_t} dt$.

Under standard regularity conditions on the process X_t, the estimator $\hat{\mathcal{L}}_X(u)$ is consistent for $\mathcal{L}_X(u)$ and $\sqrt{T}$-asymptotically normal, as is $\hat{\ell}_X(u)$. However our situation here is much more complicated since the spot volatility $\{\sigma_t\}$ is unobserved. Using our Realized Laplace transform of volatility we are able to overcome this latency problem and make inference for the Laplace transform of the stochastic volatility feasible.

The intuition behind our estimator can be explained as follows. By decreasing the length of the interval over which we sample asset prices, we can get to frequencies where the stochastic volatility is approximately constant. Thus over such short intervals of time, most of the price increments (except for the occasional "big" jumps) are approximately draws[3] from the process $\sigma \times Z_i \times \sqrt{\Delta}$ where σ is the unknown level of (locally constant) volatility, Δ is the length of the high-frequency interval, and $\{Z_i\}$ is a sequence of independent standard normal variables.[4] Therefore, by using fill-in asymptotics we can estimate "locally" the characteristic function of the normal returns, i.e., $e^{-u^2\sigma^2/2}$. Recognizing that volatility changes over time and integrating to a unit interval of time, e.g., a day, the high-frequency data thus allows us to estimate nonparametrically $\int_t^{t+1} e^{-u^2\sigma_s^2/2}ds$. Then, using long-span asymptotics we are able to recover from the data $\mathbb{E}\left(e^{-u^2\sigma_t^2/2}\right)$. The latter, when viewed as function of $u^2/2$, is the Laplace transform of σ_t^2. Thus, by transforming appropriately the real part of the empirical characteristic function of the (scaled) high-frequency returns, we are able to recover the empirical Laplace transform of the unobserved spot volatility.

The above discussion was based on the premise that when we "zoom in" by sampling more frequently, most returns (apart from the ones containing the "big" jumps) look like coming from a Gaussian distribution (with different variance). However, there can be situations where local Gaussianity might be inappropriate even when looking at very fine intervals of time. A natural question then arises as to how our ability to identify nonparametrically the distribution of the stochastic volatility is connected with the type of "local" distribution of high-frequency returns. In terms of the underlying continuous-time semimartingale model for the price, this distributional distinction of "local" Gaussianity versus non-Gaussianity is equivalent to distinguishing whether the asset price is generated by a jump-diffusion or a pure-jump process. In the latter case the small jumps play similar role to the continuous martingale and account for the day-to-day moves in asset prices. Their so-called Blumnethal-Getoor index (see e.g., Ait-Sahalia and Jacod (2009)), taking values in $[0, 2]$, determines how "active" they are. The limit case when the index equals

[3]Formally, the error from such approximation in our estimation, when integrated over all high-frequency increments, is asymtotically negligible. Section 2 makes formal the discussion here about the local behavior of the high-frequency returns.

[4]The local Gaussianity of high-frequency returns has been used, either implicitly or explicitly, in constructing many estimators and tests, see e.g. Barndorff-Nielsen et al. (2005) and Mykland and Zhang (2009).

2 means essentially that the small asset price moves are generated by a diffusion, i.e. "locally" high-frequency returns are Gaussian.

Given the above discussion, we relax the local Gaussianity assumption by assuming more generally that "locally" the high-frequency returns behave approximately like generated from a stable distribution with index β ($\beta = 2$ corresponds to the locally Gaussian case discussed above) and this allows naturally to nest together the locally Gaussian and locally non-Gaussian cases.[5] In particular, in this more general case our realized measure builds on the observation that "locally" the high-frequency returns can be treated as draws from the process $\sigma \times Z_i \times \Delta^{1/\beta}$ where now $\{Z_i\}_i$ is a sequence of independent β-stable random variables. Our Realized Laplace Transform of volatility in this more general case then provides a nonparametric estimate for $\mathbb{E}e^{-|Z_\beta u \sigma_s|^\beta}$ where Z_β is some constant depending on β. The latter becomes a Laplace transform for $|\sigma_t|^\beta$ when viewed as a function of $|Z_\beta u|^\beta$.

Our analysis reveals the tight connection between the inference for the stochastic volatility and the small scale behavior of the price process. In particular, the activity of the small moves, i.e., the value of β, determines not only the limit behavior of our statistic but it also enters directly in its construction. The activity of the small price moves is of course unknown (although in many cases it is assumed to be 2 since the model that is used contains a continuous martingale) and in a final step of our theoretical analysis we show that we can perform the above analysis using estimated value for β and derive explicitly its asymptotic effect on the limit behavior of the Realized Laplace Transform of volatility.

We test the asymptotic theory for our estimator on a simulated data, and in an empirical application we apply it to two real financial data sets: S&P 500 futures index and the VIX index. For the S&P 500 index series we find that the Laplace transform of spot variance deviates from that of integrated variance even when looking at daily level, while the corresponding difference is relatively smaller for the VIX index series. We investigate model features that can generate this wedge and find that quickly mean-reverting jump-driven factors in volatility can account for it. The empirical analysis of the VIX index also illustrates the importance of treating the activity level of small moves in the estimation as unknown. For this data set we find that "local" Gaussianity is inappropriate and when one erroneously imposes it, it leads to underestimation of the level of volatility and an overestimation of the wedge in the Laplace transforms of spot and integrated volatility.

[5]The "locally" stable behavior of the high-frequency returns concerns only the scaling of the driving martingale over short intervals of time. Thus, it is satisfied by many processes whose increments, unlike those of the stable, have all their moments finite. The formal results and definitions are given in Section 2.

Finally, there are several antecedents in the literature of using trigonometric functions in problems related to the estimation of stochastic volatility. First, the use of the empirical characteristic function of the returns in estimation dates as early as Feuerverger and Mureika (1977) and more recent contributions include Singleton (2001) and Carrasco et al. (2007) (in which the conditional characteristic function is being used). The key difference with our approach is that in these papers the distance between observations is fixed. Therefore, while our method relies on the high-frequency data to allow us work directly with the volatility process (i.e., regardless of the presence in the price of lower activity jumps as well as the relationship between the driving martingale and the volatility, i.e. the so-called "leverage effect"), the above mentioned papers are applied in settings where the (conditional) characteristic function of the return process is available in closed-form.

Second, in the context of a pure-diffusion process (i.e., continuous semimartingale), Malliavin and Mancino (2009) propose estimator for the spot volatility using high-frequency data and estimates for the Fourier transform of the price process $\frac{1}{2\pi}\int_0^{2\pi} e^{-ius}dX_s$, ($X_s$ denoting the price process). By contrast here, our Realized Laplace transform of the volatility is the (real part of) the Fourier transform of the empirical measure of the return process. So, while Malliavin and Mancino (2009)'s estimator results in cross-products of the price increments over a given period of time, ours involves cosine transformation of the scaled price increments. Of course, this is also reflected in the different limits of the statistics, spot or weighted average of it (e.g., integrated variance) in Malliavin and Mancino (2009) versus the empirical Laplace transform of the volatility in our case. For the reasons discussed above, the contribution of this paper is to overcome some of the limitations of the integrated variance when making inference for stochastic volatility.

The rest of the paper is organized as follows. In Section 2 we introduce our setup and main assumptions. Section 3 defines formally the Realized Laplace Transform of volatility and derives its asymptotic behavior both in the case of fixed and estimated value of the activity of the small moves. In Section 4 we provide evidence from a Monte Carlo for the small sample behavior of our statistic and Section 5 contains the results from the empirical application. Section 6 concludes. All proofs are in Section 7.

2 Setting and Assumptions

Throughout the paper, the process of interest is denoted with X and is defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. We assume that X has one of the following dynamics:

(a) **Jump-Diffusion**

$$dX_t = \alpha_t dt + \sigma_t dW_t + \int_{\mathbb{R}} \delta(t-, x)\mu(dt, dx), \tag{3}$$

where α_t and σ_t are càdlàg processes; W_t is a Brownian motion; μ is a homogenous Poisson measure with compensator (Lévy measure) $\nu(x)dx$; $\delta(t,x) : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ is càdlàg in t;

(b) **Pure-Jump**

$$dX_t = \alpha_t dt + \int_{\mathbb{R}} \sigma_{t-}\kappa(x)\tilde{\mu}(dt,dx) + \int_{\mathbb{R}} \sigma_{t-}\kappa'(x)\mu(dt,dx), \tag{4}$$

where α_t and σ_t are càdlàg processes; κ is a symmetric function with bounded support with $\kappa(x) = x$ in a neighborhood of 0 and $\kappa'(x) = x - \kappa(x)$; $\tilde{\mu}(dt,dx) = \mu(dt,dx) - dt\nu(x)dx$.

The two specifications essentially encompass all continuous-time processes used for modeling financial prices. The drift term, α_t, captures compensation for time and risk associated with holding the asset and we will leave it unspecified. In the pure-jump model the jump martingale ($\int_0^t \int_{\mathbb{R}} \kappa(x)\tilde{\mu}(ds,dx)$) substitutes the Brownian motion in modeling the "small" moves in the asset prices. In both models the day to day variation in the price is determined by the martingales (the second components in (3) and (4)). To see this note that the "dominant" part of the price increment over a short interval of time $(t, t+\Delta)$ will be $\sigma_{t-} \times Z_{t,t+\Delta}$ where $Z_{t,t+\Delta}$ is either $W_{t+\Delta} - W_t$ or $\int_t^{t+\Delta} \int_{\mathbb{R}} \kappa(x)\tilde{\mu}(dt,dx)$. This term will be of order $O_p(\Delta^\alpha)$ for $\alpha \in [1/2, 1)$, while the rest of the components of X will be at most $O_p(\Delta)$ when $\Delta \downarrow 0$.

We will refer henceforth to W_t and $\int_0^t \int_{\mathbb{R}} \kappa(x)\tilde{\mu}(ds,dx)$, respectively in the jump-diffusion and pure-jump models, as the driving martingales of X. These are Lévy processes and hence are time-homogenous. In both models the stochastic volatility process, σ_t, is integrated with respect to these martingales and this results in martingale components of X that exhibit time-varying volatility (and other moments). Our object of interest in this paper is the stochastic volatility. Of course we observe only X and σ_t is hidden into it, so our goal in the paper will be to uncover σ_t, and its distribution in particular, with assuming as little as possible about the rest of the components of X and the volatility itself (including whether X is generated from a jump-diffusion or a pure-jump model). Given the preceding discussion, the scaling of the driving martingale components over short intervals of time will be of crucial importance for us, as at best we can observe only a product of the stochastic volatility with $Z_{t,t+\Delta}$. Our assumption A below characterizes the behavior of X over small scales.

Assumption A. The Lévy measure of μ satisfies:

(a) ***Jump-Diffusion***

$$\int_0^t \int_{\mathbb{R}} (|\delta(s,x)|^p \wedge 1) ds\nu(x)dx < \infty, \quad \int_{\mathbb{R}} |\delta(t,x)|\nu(x)dx < \infty, \tag{5}$$

for every $t > 0$ and every $p > \beta'$, where $0 \leq \beta' < 1$ is some constant.

(*b*) ***Pure-Jump***

$$\nu(x) = \frac{A}{|x|^{\beta+1}} + \nu'(x), \quad A > 0, \quad \beta \in (1,2), \quad \int_{\mathbb{R}} |x|\nu(x)dx < \infty, \tag{6}$$

where there exists $x_0 > 0$ *such that for* $|x| \leq x_0$ *we have* $|\nu'(x)| \leq \frac{K}{|x|^{\beta'+1}}$ *for some* $\beta' < 1$ *and a constant* $K \geq 0$.

The small scale behavior of the driving martingale in both models (W_t and $\int_0^t \int_{\mathbb{R}} \kappa(x)\tilde{\mu}(ds, dx)$ respectively) is like that of a stable process with index β. The case $\beta = 2$ corresponds to the jump-diffusion model and the case $\beta < 2$ to the pure-jump model. The index β determines the "activity" of the driving process, i.e., the vibrancy of its trajectories, and thus henceforth we will refer to it as the activity.[6] The higher β is the more active the process X is, i.e., the more vibrant its trajectories are. The value of the index β is crucial for recovering σ_t from the discrete data on X, as intuitively it determines how big on average the increments $Z_{t,t+\Delta}$ should be for a given sampling frequency. This will become clearer when we develop our estimators in the next section. The following lemma formalizes the above discussion on the small scale behavior of the driving martingale.

Lemma 1 *Let* $Z_t = W_t$ *and* $\beta = 2$ *in the case of jump-diffusion model (3) and* $Z_t = \int_0^t \int_{\mathbb{R}} \kappa(x)\tilde{\mu}(ds, dx)$ *and* β *be the parameter in (6) in the case of the pure-jump model (4). Then for* $h \to 0$ *we have*

$$h^{-1/\beta} Z_{th} \xrightarrow{\mathcal{L}} S_t, \tag{7}$$

where the convergence is for the local uniform in t *topology and* S_t *is a stable process with characteristic function* $\mathbb{E}\left(e^{iuS_1}\right) = e^{-|uZ_\beta|^\beta}$ *for*

$$Z_\beta = \begin{cases} \left(A \times \frac{2\Gamma(2-\beta)|\cos(\beta\pi/2)|}{\beta(\beta-1)}\right)^{1/\beta}, & \text{if } \beta \in (1,2), \\ \frac{1}{\sqrt{2}}, & \text{if } \beta = 2, \end{cases} \tag{8}$$

with A *being the constant in (6).*

In assumption A we also restrict the "activity" of the residual jump components of X, i.e., we limit their effect in determining the small moves of X. In both cases, jump-diffusion and pure-jump, this is conveniently captured by the parameter β'. In the case of the jump-diffusion model, the "leading" component is the diffusion and the "residual" is the jumps. In the pure-jump model the leading component is the "stable" part of the jump process and $\nu'(x)$ controls the "residual"

[6]Formally, β equals the generalized Blumenthal-Getoor index (Ait-Sahalia and Jacod (2009)) of the process (in the pure-jump model). We are not going to use this definition further in the paper and therefore to avoid unnecessary complications we do not provide a formal definition of the concept.

jump component. In both cases we restrict $\beta' < 1$, i.e., the "residual" jump component is of finite variation (and this is why we do not need a martingale measure to define it). This restriction is not necessary if one is interested only in convergence in probability results (only $\beta' < \beta$ in both cases is needed for this). However, if one needs also the asymptotic distribution of the statistics that we introduce in the paper, then this assumption is unavoidable. In most parametric continuous models used to date this restriction is satisfied.

It is important to note that $\nu'(x)$ in part (b) of the theorem is a signed measure and therefore assumption A restricts only the behavior of $\nu(x)$ for $x \sim 0$ to be like that of a stable process. However, for the big jumps, i.e., when $|x| > K$ for some arbitrary $K > 0$, the stable part of $\nu(x)$ can be completely eliminated or tempered by negative values of the "residual" $\nu'(x)$. An example of this, which is covered by our assumption A, is the tempered stable process of Rosiński (2007) generated from the stable by tempering its tails. The latter guarantees that unlike the stable, the tempered stable process has all its moments finite. We will use later this process in our Monte Carlo study. Therefore, while assumption A ties the small scale behavior of the driving martingale of X with that of a stable process, it importantly leaves its large scale behavior unrestricted (i.e., the limit of $h^{-\alpha}Z_{th}$ for some $\alpha > 0$ when $h \to \infty$ is unrestricted) and thus in particular unrelated with that of a stable process. Finally, part (b) of assumption A can be further weakened by requiring the "residual" term in X due to $\nu'(x)$ to have arbitrary time-dependence (and not necessarily tied with σ_t). This however would be only a mild extension that has not been used in the parametric pure-jump models to date and therefore we do not consider it here.

Our next assumption imposes minimal integrability conditions on α_t and σ_t and further limits the amount of variation we can have in these processes over short periods of time. Intuitively, we will need the latter to guarantee that by sampling frequently enough we can treat "locally" σ_t (and α_t) as constant.

Assumption B. The processes α_t and σ_t are square-integrable and satisfy for every $0 < s < t$

$$\mathbb{E}\left(\alpha_t - \alpha_s\right)^2 \leq C(t-s) \quad \text{and} \quad \mathbb{E}\left(\sigma_t - \sigma_s\right)^2 \leq C(t-s), \tag{9}$$

where C is a constant that does not depend on time.

Assumption B is very weak: it covers all processes used in continuous-time econometrics to date, including the most typical case of Itô semimartingales, but also Lévy-driven moving average processes (e.g. Brockwell (2001)) as well as long-memory models where the driving process is a fractional Brownian motion (e.g. Comte and Renault (1998)). We also state a slightly stronger version of assumption B, which allows to strengthen somewhat some of our asymptotic results.

Assumption B′. In addition to the requirement on the process α_t in assumption B, assume that σ_t is an Itô semimartingale given by

(a) ***Jump-Diffusion***

$$\sigma_t = \sigma_0 + \int_0^t \tilde{\alpha}_s ds + \int_0^t \tilde{\sigma}_s dW_s + \int_0^t \tilde{\sigma}'_s dW'_s + \int_0^t \int_{\mathbb{R}} \delta(s-, x)\underline{\tilde{\mu}}(ds, dx), \tag{10}$$

where W' is a Brownian motion independent from W; $\underline{\mu}$ is a homogenous Poisson measure, with Lévy measure $\underline{\nu}(dx)$, having arbitrary dependence with μ. We have for every t and s:

$$\begin{cases} \mathbb{E}\left(\sigma_t^6 + \alpha_t^6 + \tilde{\sigma}_t^6 + (\tilde{\sigma}'_t)^6 + \int_{\mathbb{R}} \delta^6(t,x)\underline{\nu}(dx)\right) < C, \\ \mathbb{E}\left(|\tilde{\sigma}_t - \tilde{\sigma}_s|^2 + |\tilde{\sigma}'_t - \tilde{\sigma}'_s|^2 + \int_{\mathbb{R}}(\delta(t,x) - \delta(s,x))^2\underline{\nu}(dx)\right) < C(t-s), \end{cases} \tag{11}$$

where $C > 0$ is some constant that does not depend on t and s.

(b) ***Pure-Jump***

$$\sigma_t = \sigma_0 + \int_0^t \tilde{\alpha}_s ds + \int_0^t \tilde{\sigma}_s dW_s + \int_0^t \int_{\mathbb{R}} \delta(s-, x)\underline{\tilde{\mu}}(ds, dx), \tag{12}$$

where $\underline{\mu}$ is a homogenous Poisson measure, with Lévy measure $\underline{\nu}(dx)$, having arbitrary dependence with μ. We have for every t and s:

$$\begin{cases} \mathbb{E}\left(|\sigma_t|^{4\beta-2} + |\alpha_t|^{4\beta-2} + |\tilde{\sigma}_t|^{4\beta-2} + \int_{\mathbb{R}} |\delta|^{4\beta-2}(t,x)\underline{\nu}(dx)\right) < C, \\ \mathbb{E}\left(|\tilde{\sigma}_t - \tilde{\sigma}_s|^2 + \int_{\mathbb{R}}(\delta(t,x) - \delta(s,x))^2\underline{\nu}(dx)\right) < C(t-s), \\ \int_0^t \int_{\mathbb{R}}(|\delta(s,x)|^p \wedge 1) ds \underline{\nu}(dx) < \infty, \quad \text{for } \forall p > \beta'' \text{ with } \beta'' < 2, \end{cases} \tag{13}$$

where $C > 0$ is some constant that does not depend on t and s.

Assumption B′ implies of course the weaker assumption B. It is still a very general assumption, which is satisfied by the multifactor stochastic volatility models that are widely used in financial econometrics (e.g. the affine class of Duffie et al. (2000)). In particular, note that assumption B′ allows for a completely arbitrary dependence between the increments in σ_t and those in X, i.e., so-called "leverage" effects in a most general form (by linking either jumps or Brownian motions) are allowed. Assumption B′ also strengthens the integrability assumption on σ_t by requiring existence of moments up to order 6, but this is relatively mild strengthening.

Finally, in our estimation we make use of long-span asymptotics, for (a transformation of) the process σ_t and the latter contains temporal dependence. Therefore, we need a condition on this dependence that guarantees that a Central Limit Theorem for the associated empirical process exists. This condition is given in our next assumption.

Assumption C-u. The volatility σ_t is a stationary and ergodic process and we further have for some $q \in [1,2]$

$$\int_0^\infty \left|\left|\mathbb{E}\left(e^{-u|\sigma_t|^\beta} - \mathbb{E}(e^{-u|\sigma_t|^\beta})|\mathscr{F}_0\right)\right|\right|_q dt < \infty, \tag{14}$$

where $\beta = 2$ if the process is a jump-diffusion and else is the constant appearing in (6).

The above assumption guarantees that the following integral, which will be the asymptotic variance of our estimator, is well defined (see Jacod and Shiryaev (2003), Theorem VIII.3.79):

$$V_\beta(u) = 2\int_0^\infty \mathbb{E}\left[\left(e^{-|Z_\beta u\sigma_t|^\beta} - \mathbb{E}\left(e^{-|Z_\beta u\sigma_t|^\beta}\right)\right)\left(e^{-|Z_\beta u\sigma_0|^\beta} - \mathbb{E}\left(e^{-|Z_\beta u\sigma_t|^\beta}\right)\right)\right] dt < \infty, \tag{15}$$

where the constant Z_β is given in (8). Assumption C-u can be easily verified when the conditional characteristic function of $|\sigma_t|^\beta$ is known in closed form, e.g., in the affine jump-diffusion models. We note also that more primitive sufficient conditions for (14) can be given in terms of mixing coefficients (if σ_t is mixing), an example is given by assumption C' below.

Assumption C'. The volatility σ_t is a stationary and α-mixing process with $\mathbb{E}|\sigma_t|^3 < \infty$ and $\alpha_t^{mix} = O(t^{-3-\iota})$ for arbitrary small $\iota > 0$ when $t \to \infty$, where

$$\alpha_t^{\text{mix}} = \sup_{A\in\mathscr{F}_0,\ B\in\mathscr{F}^t} |\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)|, \quad \mathscr{F}_0 = \sigma(\sigma_s, s\le 0) \text{ and } \mathscr{F}^t = \sigma(\sigma_s, s \ge t). \tag{16}$$

3 Limit theory for Realized Laplace Transform

3.1 The Case of Fixed Activity

Now we are ready to develop our estimators and derive their asymptotic properties. We will assume that we observe the process X at the equidistant times $0, \Delta_n, ..., i\Delta_n, ..., [T/\Delta_n]$ where Δ_n is the length of the high-frequency interval and T is the span of the data. Most of the asymptotics in this paper will be joint: fill-in ($\Delta_n \to 0$) and long-span ($T \to \infty$). This is appropriate for financial applications where one has long series of finely sampled asset prices. We start our analysis with the case when the activity level β is known to the econometrician.

We aggregate the high-frequency data into the following realized measure:

$$V_T(X, \Delta_n, \beta, u) = \sum_{i=1}^{[T/\Delta_n]} \Delta_n \cos(u\Delta_n^{-1/\beta}\Delta_i^n X), \quad \Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}. \tag{17}$$

$V_T(X, \Delta_n, \beta, u)$ is essentially the real part of the empirical characteristic function of the appropriately scaled increments of the process. Its connection with the Laplace transform of the volatility that we are after can be intuitively described as follows. Since the driving martingale over small scales behaves like β-stable (assumption A) and the volatility changes over short intervals are not too big on average (assumption B), then the "dominant" part (in a fill-in asymptotic sense) of the increment $\Delta_i^n X$ (when Δ_n is small) is $\sigma_{(i-1)\Delta_n} Z_{(i-1)\Delta_n, i\Delta_n}$ with $Z_{(i-1)\Delta_n, i\Delta_n}$ being approximately stable. Then, using the self-similarity of the stable process, $Z_{(i-1)\Delta_n, i\Delta_n}$ will behave like $\Delta_n^{1/\beta} \times Z_{0,1}$, and further the characteristic of a stable process is given by $e^{-|uZ_\beta|^\beta}$. Therefore, for a fixed T, $V_T(X, \Delta_n, \beta, u)$ is approximately a sample average of a heteroscedastic data series. Thus, by a law of large numbers (when $\Delta_n \to 0$), it will converge to $\int_0^T e^{-|u\sigma_t Z_\beta|^\beta} ds$, which is the empirical Laplace transform of $|\sigma_t|^\beta$ (when viewed as a function of $|uZ_\beta|^\beta$ and after dividing by T). For this reason we will refer to $V_T(X, \Delta_n, \beta, u)$ as the Realized Laplace Transform of volatility. The following theorem gives the precise fill-in asymptotics result. In it we denote with $\mathcal{L} - s$ convergence stable in law, which means that the convergence in law holds jointly with any random variable defined on the original probability space. We also use the standard notation $x \wedge y = \min\{x, y\}$ and $x \vee y = \max\{x, y\}$ for $x, y \in \mathbb{R}$.

Theorem 1 *For the process X, assume that assumptions A and B′ hold, fix T and let $\Delta_n \to 0$.*

(a) If $\beta > 4/3$ and $\beta' < \beta/2$ for β and β' the constants of assumption A, then we have

$$\frac{1}{\sqrt{\Delta_n}} \left(V_T(X, \Delta_n, \beta, u) - \int_0^T e^{-|u\sigma_s Z_\beta|^\beta} ds \right) \xrightarrow{\mathcal{L}-s} \sqrt{\int_0^T F_\beta(uZ_\beta \sigma_s) ds} \times E, \tag{18}$$

where E is a standard normal variable defined on extension of the original probability space and further $F_\beta(x) = \frac{e^{-2^\beta x^\beta} - 2e^{-2x^\beta} + 1}{2}$ for $x > 0$.

A consistent estimator for the asymptotic variance is given by

$$\frac{V_T(X, \Delta_n, \beta, 2u) - 2V_T(X, \Delta_n, \beta, 2^{1/\beta}u) + 1}{2}. \tag{19}$$

(b) If either $\beta \le 4/3$ or $\beta' \ge \beta/2$, then

$$\left(V_T(X, \Delta_n, \beta, u) - \int_0^T e^{-|u\sigma_s Z_\beta|^\beta} ds \right) = O_p\left(|\log \Delta_n| \Delta_n^{1-\beta'/\beta} \vee \Delta_n^{2-2/\beta} \right). \tag{20}$$

The high-frequency data allows us to "integrate out" the increments $Z_{(i-1)\Delta_n, i\Delta_n}$, i.e. it essentially allows to "deconvolute" σ_t from the driving martingale of X (W_t and $\int_0^t \int_\mathbb{R} \kappa(x)\tilde{\mu}(ds, dx)$ respectively). The fill-in asymptotic limit of (17) is the empirical Laplace transform of the stochastic volatility (after dividing by T). Then, by letting $T \to \infty$ (and dividing by T) we can eliminate the sampling variation due to the stochastic nature of σ_t, and thus recover its population properties, i.e., estimate $\mathbb{E}\left(e^{-|Z_\beta u \sigma_t|^\beta}\right)$. The latter is the Laplace transform of $|\sigma_t|^\beta$ that we are after (after an appropriate change of variable with respect to u). The next theorem gives the asymptotic behavior of $\frac{1}{T} V_T(X, \Delta_n, \beta, u)$ when both $T \to \infty$ and $\Delta_n \to 0$.

Theorem 2 *(a) Suppose $T \to \infty$ and $\Delta_n \to 0$. For the process X being either a jump-diffusion or pure-jump and under assumptions A, B and C-u for some $u > 0$, we have*

$$\sqrt{T}\left(\frac{1}{T}V_T(X,\Delta_n,\beta,u) - \mathbb{E}\left(e^{-|Z_\beta u\sigma_t|^\beta}\right)\right) = Y_T^{(1)}(u) + Y_T^{(2)}(u), \tag{21}$$

$$\begin{cases} Y_T^{(1)}(u) \overset{\mathcal{L}}{\longrightarrow} \sqrt{V_\beta(u)} \times Z, \\ Y_T^{(2)}(u) = O_p\left(\sqrt{T}\left(|\log\Delta_n|\Delta_n^{1-\beta'/\beta} \vee \Delta_n^{(2-2/\beta)\wedge 1/2}\right) \vee \sqrt{\Delta_n}\right), \end{cases} \tag{22}$$

where Z is standard normal random variable and $V_\beta(u)$ is defined in (15). $Y_T^{(2)}(u)$ is asymptotically negligible if

$$\sqrt{T}\left(|\log\Delta_n|\Delta_n^{1-\beta'/\beta} \vee \Delta_n^{(2-2/\beta)\wedge 1/2}\right) \to 0.$$

(b) If in part(a) we strengthen assumption B to assumption B′, we get the weaker

$$Y_T^{(2)}(u) = O_p\left(\sqrt{T}\left(|\log\Delta_n|\Delta_n^{1-\beta'/\beta} \vee \Delta_n^{2-2/\beta} \vee 1_{\{\beta\neq 2\}}\Delta_n^{1/(\beta\vee\beta'')-\iota}\right) \vee \sqrt{\Delta_n}\right),$$

for $\iota > 0$ arbitrary small. If further $\underline{\mu}$ in (12) is independent from μ, then we even have

$$Y_T^{(2)}(u) = O_p\left(\sqrt{T}\left(|\log\Delta_n|\Delta_n^{1-\beta'/\beta} \vee \Delta_n^{2-2/\beta}\right) \vee \sqrt{\Delta_n}\right).$$

(c) If in part (b) we further have $\sqrt{T}\left(|\log\Delta_n|\Delta_n^{1/2-\beta'/\beta} \bigvee \Delta_n^{3/2-2/\beta} \vee 1_{\{\beta\neq 2\}}\Delta_n^{1/(\beta\vee\beta'')-1/2-\iota}\right)$ $\to 0$, then the residual term $Y_T^{(2)}(u)$ satisfies

$$\frac{1}{\sqrt{\Delta_n}}Y_T^{(2)}(u) \overset{\mathcal{L}}{\longrightarrow} \sqrt{\mathbb{E}\left(F_\beta(uZ_\beta\sigma_t)\right)} \times U, \tag{23}$$

where U is standard normal random variable and the function F_β was defined in Theorem 1.

Under the condition of Theorem 2 (a), the leading component in the asymptotic expansion (for the joint asymptotics) of the scaled and centered realized Laplace transform can be split into two components, $Y_T^{(1)}(u)$ and $Y_T^{(2)}(u)$, that have different asymptotic behavior and capture different errors involved in the estimation. The first one, $Y_T^{(1)}(u)$, equals $\sqrt{T}\left(\frac{1}{T}\int_0^T e^{-|Z_\beta u\sigma_t|^\beta}dt - \mathbb{E}\left(e^{-|Z_\beta u\sigma_t|^\beta}\right)\right)$, which is the empirical process corresponding to the continuous-record case in which the driving martingale for X, i.e. the Brownian motion or the jump martingale, has been already "integrated out". Hence the magnitude of $Y_T^{(1)}(u)$ is sole function of the time span T. On the other hand, the term $Y_T^{(2)}(u)$ captures the effect from the discretization error, i.e., the fact that we use high-frequency data and not continuous record of X in the estimation. The magnitude of the discretization error is hence naturally a function of how big are the intervals between price observations, i.e., Δ_n.

Therefore, in order to determine the relative importance of the discretization error and the error associated with the empirical process, we need a condition for the relative speed of $\Delta_n \to 0$ and $T \to \infty$.

Part a and b of the above theorem provide such relative speed conditions. In analyzing them, it is helpful first to evaluate these conditions in the most typical case in finance, i.e., when X is a jump-diffusion model with finite activity jumps (e.g. compound Poisson). In this case $\beta = 2$ and $\beta' = 0$. If further assumption B$'$ holds (as in the multifactor stochastic volatility models), then the relative speed condition becomes $\sqrt{T}\Delta_n \to 0$. This is a very weak condition and allows in particular the span of the data to increase much faster than the sampling frequency. Compared with the standard requirement $T\Delta_n \to 0$ found in the related problem of maximum-likelihood estimation of diffusion processes with discrete data, see e.g., Prakasa Rao (1988), our relative speed condition is much weaker.

In the more general case for X, the relative speed condition is a function of the elements present in X. It is worth to keep in mind that the leading component determining the behavior of the realized Laplace transform $V_T(X, \Delta_n, \beta, u)$ is the "stable process part" of X ("convoluted" with the stochastic volatility σ_t). However, X contains other components, mainly less "active" jumps and a drift term, whose effect on $V_T(X, \Delta_n, \beta, u)$ although dominated by the "stable process part" might still affect its rate of convergence. In particular, the drift term in terms of its small scale behavior is similar to a stable process with activity level of $\beta = 1$. Hence, when the index β in (6) is lower, the leading component gets closer in small scale behavior to the drift term. Therefore, the relative speed condition deteriorates (this is captured by the term $\sqrt{T}\Delta_n^{2-2/\beta}$ in the order of magnitude of $Y_T^{(2)}(u)$ stated in Theorem 2). The effect of having additional jumps with activity β' (in addition to the "locally stable" part) is similar.

Finally, in part (c) of the theorem we give the limiting distribution of the discretization error term $Y_T^{(2)}(u)$ under the stronger relative speed condition $T\left(|\log\Delta_n|\Delta_n^{1-2\beta'/\beta} \bigvee \Delta_n^{3-4/\beta} \vee \Delta_n^{2/(\beta\wedge\beta'')-1-\iota}\right)$ $\to 0$. The processes Z and U will be independent when σ_t and the driving martingale in X are independent, i.e., the case of no "leverage". Otherwise there will be dependence between them determined by the "leverage" effect. Of course, the relative speed conditions discussed here give asymptotic limits and we will check in the empirical section if the asymptotic theory applies well for the time span and frequencies of the typical financial data sets that are available.

It is interesting to compare our Theorem 2 with the results of Van Es et al. (2003). In the context of a pure-diffusion model (i.e., the model in (5) but with no jumps), these authors propose to estimate the density of the stochastic volatility process by utilizing the fact that the squared high-frequency returns are approximately sum of the log-volatility (the signal) plus the log of a squared normal random variable (the noise). Then the estimation of the log-volatility density is done by using a deconvolution density estimator. However, since the density of the noise is very smooth, the deconvolution is very hard and results in mean squared errors decreasing at a logarithmic rate only.[7] By contrast, our estimator of the Laplace transform of volatility (which uniquely identifies its distribution) possesses the standard $\sqrt{T}$ rate of convergence and naturally extends without any modification in the case with jumps.

The limit results so far were derived for a fixed u. However, in a typical application one would like to know the Laplace transform as a function of u. Therefore, we next show in the following theorem that the asymptotic results for $\frac{1}{T}V_T(X,\Delta_n,\beta,u)$ in Theorem 2 hold also functionally, i.e., when one considers $\frac{1}{T}V_T(X,\Delta_n,\beta,u)$ as functions of u.

Theorem 3 *Suppose in Theorem 2, we replace assumption C-u with the stronger one C'. Then the limit results of Theorem 2 hold locally uniformly in u on the space $\mathcal{C}(\mathbb{R}_+)$ of continuous functions indexed by u (i.e. uniformly over compact sets of $u \in \mathbb{R}_+$):*

(i) The limit of $Y_T^{(1)}(u)$ is a Gaussian process with variance-covariance

$$2\int_0^\infty \mathbb{E}\left[\left(e^{-|Z_\beta u\sigma_t|^\beta} - \mathbb{E}\left(e^{-|Z_\beta u\sigma_t|^\beta}\right)\right)\left(e^{-|Z_\beta v\sigma_0|^\beta} - \mathbb{E}\left(e^{-|Z_\beta v\sigma_t|^\beta}\right)\right)\right]dt, \quad u,v>0. \quad (24)$$

(ii) The limit of $Y_T^{(2)}(u)$ in case (c) of Theorem 2 is a Gaussian process with variance-covariance matrix $\mathbb{E}\left(F_\beta(uZ_\beta\sigma_t, vZ_\beta\sigma_t)\right)$ for $u,v>0$ where

$$F_\beta(x,y) = \frac{e^{-|x+y|^\beta} - 2e^{-|x|^\beta-|y|^\beta} + e^{-|x-y|^\beta}}{2}, \quad x,y>0. \quad (25)$$

So far we have established the asymptotic behavior of the Realized Laplace Transform of the volatility. One might naturally ask whether this analysis can be extended to derivatives of our realized measure with respect to u, i.e., whether the latter converge to the corresponding derivatives of the limit $\mathbb{E}(e^{-|Z_\beta u\sigma_t|^\beta})$. Unfortunately this is not the case for two reasons. First, powers above 2 of a stable random variable with $\beta < 2$ do not exist, and therefore the derivatives of $\mathbb{E}(e^{-|Z_\beta u\sigma_t|^\beta})$ will not exist when evaluated at $u=0$. Second, and more importantly the derivatives of $\frac{1}{T}V_T(X,\Delta_n,\beta,u)$ will involve the terms $|\Delta_n^{-1/\beta}\Delta_i^n X|^p$ for $p\geq 2$, and the asymptotic behavior of these terms will be affected by the "big" jumps.

[7] Comte and Genon-Catalot (2006), building also on the deconvolution idea and still in a pure-diffusion setting (with an additional assumption of σ_t being independent from W_t), propose a penalized projection estimator for the volatility density. This estimator can improve the rates of convergence of the estimator of Van Es et al. (2003) when further conditions on the smoothness of the volatility density are imposed.

We turn now to making Theorem 2 feasible, i.e., deriving estimates from the high-frequency data for the asymptotic variance of the leading term $Y_T^{(1)}(u)$ (estimate for the asymptotic variance of $Y_T^{(2)}(u)$ is just a sample average of (19)). To state our next theorem we introduce some more notation. We denote for $t = 1, ..., T$

$$\widehat{z}_t(\beta, u) = \sum_{i=[(t-1)/\Delta_n]+1}^{[t/\Delta_n]} \Delta_n \cos(u\Delta_n^{-1/\beta}\Delta_i^n X) - \frac{1}{T}V_T(X, \Delta_n, \beta, u),$$

and we further set $\widehat{Z}_k(\beta, u) = \frac{1}{T}\sum_{t=k+1}^{T} \widehat{z}_t(\beta, u)\widehat{z}_{t-k}(\beta, u)$ for some $k \geq 0$.

Theorem 4 *Suppose assumptions A, B and C′ hold. Then, when $T \to \infty$ and $\Delta_n \to 0$, for arbitrary integer $k \geq 1$ and every $u, v > 0$ we have*

$$V_1(X, \Delta_n, \beta, u)\left(V_k(X, \Delta_n, \beta, v) - V_{k-1}(X, \Delta_n, \beta, v)\right) \stackrel{\mathbb{P}}{\longrightarrow} \int_0^1 \int_{k-1}^{k} e^{-|Z_\beta u\sigma_t|^\beta} e^{-|Z_\beta v\sigma_s|^\beta} ds dt. \quad (26)$$

If further L_T is a deterministic sequence of integers satisfying $\frac{L_T}{T} \to 0$ as $T \to \infty$ and

$$L_T\left(|\log \Delta_n|\Delta_n^{1-\beta'/\beta} \vee \Delta_n^{(2-2/\beta)\wedge 1/2}\right) \to 0,$$

we have

$$\widehat{V}_\beta(u) = \widehat{Z}_0(\beta, u) + 2\sum_{i=1}^{L_T} \omega(i, L_T)\widehat{Z}_i(\beta, u) \stackrel{\mathbb{P}}{\longrightarrow} V_\beta(u), \quad (27)$$

where $\omega(i, L_T)$ is either a Bartlett or a Parzen kernel.

The result in (26) is of independent interest. The sample average of the limit in (26) essentially identifies the joint Laplace transform of $|\sigma_t|^\beta$ and $|\sigma_s|^\beta$. This result can be used for estimation and testing of the transitional density specification of the volatility process. This is outside of the scope of the current paper and we do not pursue this any further here.

We finish this section with stating a corollary that summarizes the estimation of the Laplace transform of the volatility in the case when the true β is known. In this corollary we set A in (6) such that $Z_\beta = 2^{-1/\beta}$ for $\forall\beta \in (1, 2]$. From an econometric point of view we are free to set A in an arbitrary way since what we observe is X which is an integral of σ_t with respect to the jump martingale and we never observe the two separately (in the same spirit in the jump-diffusion model the integration is "normalized" to be with respect to the Brownian motion and not a multiple of it). The above choice of A ensures continuity across β (including the case $\beta = 2$).[8]

[8]An alternative would be to set A such that for all value of β the increment of the stable process has the same absolute value. Since this moment is not known in closed-form we do not adopt this alternative here.

Corollary 1 *Under the conditions of Theorem 2, we have that* $\widehat{\mathcal{L}}_{\sigma_t^\beta}(u) = \frac{1}{T}V_T(X, \Delta_n, \beta, Z_\beta^{-1}u^{1/\beta})$ *is a consistent and asymptotic normal estimate of the Laplace transform of* $|\sigma_t|^\beta$. *The estimated asymptotic standard error is given by* $\sqrt{\frac{1}{T}\widehat{V}_\beta(Z_\beta^{-1}u^{1/\beta})}$.

Finally, if we put a wrong value of β in the calculation of the Realized Laplace Transform of volatility, then it is easy to see that $\frac{1}{T}V_T(X, \Delta_n, \beta, u)$ will converge either to 1 or 0 depending on whether the wrong value is above or below the true one respectively.

3.2 The Case of Estimated Activity

The above asymptotic results relied on the premise that the value of β is known to the econometrician, which is somewhat unrealistic in practice. On the other hand, the realized Laplace transform crucially relies on the value of β, as the latter enters not only in its asymptotic limit and variance but also in its construction. At the same time developing an estimate for β from the high-frequency data is relatively easy (we will give an example in the next section). Hence, here we investigate the effect of estimating β on our asymptotic results from the previous section. The result is stated in the following theorem.

Theorem 5 *Suppose there exists an estimator of* β, *denoted with* $\widehat{\beta}$ *and assumptions A, B and C-u for some* $u > 0$ *hold.*

(a) If $\widehat{\beta} - \beta = o_p\left(\frac{\Delta_n^\alpha}{\sqrt{T}}\right)$ *for some* $\alpha > 0$, *then we have*

$$\sqrt{T}\left(\frac{1}{T}V_T(X, \Delta_n, \widehat{\beta}, u) - \frac{1}{T}V_T(X, \Delta_n, \beta, u)\right) = o_p\left(\frac{1}{\sqrt{T}}\right). \tag{28}$$

(b) If we have in addition assumption B′, $\widehat{\beta}$ *uses only information before the beginning of the sample or an initial part of the sample with a fixed time-span (i.e., one that does not grow with T), and further* $\widehat{\beta} - \beta = O_p(\Delta_n^\alpha)$ *for* $0 < \alpha < (1 - \beta'/\beta) \vee (2 - 2/\beta)$ *and* $\alpha \leq 1/2$, *then we have*

$$\sqrt{T}\left(\frac{1}{T}V_T(X, \Delta_n, \widehat{\beta}, u) - \frac{1}{T}V_T(X, \Delta_n, \beta, u)\right) - \frac{\sqrt{T}\log(\Delta_n)\mathbb{E}(G_\beta(uZ_\beta\sigma_t))}{\beta^2}(\widehat{\beta} - \beta) \overset{\mathbb{P}}{\longrightarrow} 0, \tag{29}$$

where $G_\beta(x) = \beta x^\beta e^{-x^\beta}$ *for* $x > 0$.

(c) Under the conditions of part (b), a consistent estimator for $\mathbb{E}(G_\beta(uZ_\beta\sigma_t))$ *is given by*

$$\widehat{G}_\beta = \frac{\Delta_n}{T}\sum_{i=1}^{[T/\Delta_n]}\left(u\Delta_n^{-1/\widehat{\beta}}\Delta_i^n X\right)\sin\left(u\Delta_n^{-1/\widehat{\beta}}\Delta_i^n X\right) \overset{\mathbb{P}}{\longrightarrow} \mathbb{E}(G_\beta(uZ_\beta\sigma_t)). \tag{30}$$

Unlike the estimation of $\mathbb{E}\left(e^{-|Z_\beta u \sigma_t|^\beta}\right)$, which requires both $\Delta_n \to 0$ and $T \to \infty$, the estimation of β can be performed with a fixed time span by only sampling more frequently. Therefore, typically the error $\widehat{\beta} - \beta$ will depend only on Δ_n. Thus, in the general case of part (a) of the theorem, we need the relative speed condition $T\Delta_n^{1+\alpha} \to 0$ for some arbitrary small $\alpha > 0$ to guarantee that the estimation of β does not have an asymptotic effect on the estimation of the Laplace transform of the volatility. By providing a bit more structure, mainly imposing the restriction that $\widehat{\beta}$ is estimated by previous part of the sample or an initial part of the current sample with a fixed time span, we can derive the leading component of the introduced error in our estimation. This is done in part (b) of the theorem, where it is shown that the latter is a linear function of $\widehat{\beta} - \beta$ (appropriately scaled). As mentioned earlier, $\widehat{\beta}$ does not need long span, just sampling more frequently, i.e., $\Delta_n \to 0$. Therefore, in a practical application one can estimate β from a short period of time at the beginning of the sample and use the estimated $\widehat{\beta}$ and the rest of the sample (or the whole sample) to estimate the Laplace transform of the stochastic volatility. In such a case, part (b) allows to incorporate the asymptotic effect of the error in estimating β into calculation of the standard errors for $\mathbb{E}\left(e^{-|Z_\beta u \sigma_t|^\beta}\right)$. For this, one needs to note that the errors in (21) and (29) in such case are asymptotically independent. Finally, the results of Theorem 5, suggest that a more efficient estimator, in the sense of faster rate of convergence, will mean that the approximation error $\frac{1}{T}V_T(X, \Delta_n, \widehat{\beta}, u) - \frac{1}{T}V_T(X, \Delta_n, \beta, u)$ will be smaller asymptotically. Hence, it is desirable to use estimators for $\widehat{\beta}$ that have good properties. And this is the case under fairly general conditions: $\sqrt{\Delta_n}$-consistent estimates for $\widehat{\beta}$ (when $\beta' < \beta/2$) can be constructed by using a ratio of power variations over two scales for appropriately chosen power, see Todorov and Tauchen (2009a) and equation (35) in the next section. We note finally that the upper bound on α in part(b) of the above theorem would typically be satisfied as it is due to the effect from the presence of less active (from the leading stable) components of X, and the latter will typically bound the rate of convergence of an estimator $\widehat{\beta}$ in a similar way.

Similar to Theorem 3 there is a functional analog of Theorem 5. We omit this (trivial) extension. We finish the section with stating the analogue of Corollary 1 in the case when β is estimated from the data. We make the same choice for the constant A as in that corollary.

Corollary 2 *Under the conditions of Theorem 5, we have that* $\widehat{\widehat{\mathcal{L}}}_{|\sigma_t|^\beta}(u) = \frac{1}{T}V_T(X, \Delta_n, \widehat{\beta}, (2u)^{1/\widehat{\beta}})$ *for some* $u > 0$ *is a consistent and asymptotic normal estimate of the Laplace transform of* $|\sigma_t|^\beta$. *The estimated asymptotic standard error is given by* $\sqrt{\frac{1}{T}\widehat{V}_{\widehat{\beta}}((2u)^{1/\widehat{\beta}}) + \frac{\log^2(0.5\Delta_n/u)\widehat{G}^2_{\widehat{\beta}}}{\widehat{\beta}^4}\widehat{var}(\widehat{\beta} - \beta)}$, *where* $\widehat{var}(\widehat{\beta} - \beta)$ *is a consistent estimate for the variance of the estimator* $\widehat{\beta}$.

4 Monte Carlo Assessment

We now examine the properties of the estimators of the Laplace transform given by Corollaries 1 and 2 of the preceding section. The Monte Carlo work is undertaken for the setups in **Cases A–F** described immediately below, which are based on five models calibrated for a financial price series evolving in continuous time. The results are then summarized in the two subsequent subsections. Since our estimators are intended for large dense data sets, we use $1,000$ Monte Carlo replications of $3,000$ "days" worth of 200 within-day price increments. In the descriptions of the cases below, W and B are generic Wiener processes and L is a generic pure-jump Lévy process with no drift.

Case A: Continuous affine stochastic volatility model

This model is the basic affine model without jumps and with persistent volatility:

$$dX_t = \sqrt{V_t}dW_t, \quad dV_t = 0.02(1.0 - V_t)dt + 0.05\sqrt{V_t}dB_t, \quad W_t \perp B_t. \tag{31}$$

Case B: Affine jump-diffusion

This model is (31) above but with added Merton-type jumps, i.e., compound Poisson jumps with normally distributed jump size:

$$\begin{aligned} &dX_t = \sqrt{V_t}dW_t + dL_t, \quad dV_t = 0.02(1.0 - V_t)dt + 0.05\sqrt{V_t}dB_t, \quad W_t \perp B_t, \\ &L_t \text{ is pure-jump with Lévy density } \nu(x) = 0.2 \times \frac{e^{-x^2}}{\sqrt{\pi}}. \end{aligned} \tag{32}$$

Case C: Jump-diffusion with pure-jump volatility

Here, a finitely active pure-jump volatility process replaces the continuous volatility process of the preceding case:

$$\begin{aligned} &dX_t = \sqrt{V_{t-}}dW_t + dL_{1t}, \; dV_t = -V_t dt + dL_{2t}, \quad L_{1t} \perp L_{2t}, \\ &L_{1t} \text{ is pure-jump with Lévy density as per } \textbf{Case B}, \\ &L_{2t} \text{ is pure-jump with Lévy density } \nu(x) = e^{-x}1_{\{x>0\}}. \end{aligned} \tag{33}$$

Case D: Stable pure-jump price process with continuous volatility

$$\begin{aligned} &dX_t = V_{t-}^{1/\beta}dL_t, \quad dV_t = 0.02(1.0 - V_t)dt + 0.05\sqrt{V_t}dB_t, \\ &L_t \text{ is pure-jump with Lévy density where } \nu(x) = \frac{0.11}{|x|^{1+1.7}} \; (\beta = 1.7). \end{aligned} \tag{34}$$

Case E: Tempered stable pure-jump price process with continuous volatility

This case is the same as **Case D** except L_t is a tempered stable also with activity $\beta = 1.7$; its Lévy density is

$$\nu(x) = \frac{0.11e^{-0.25|x|}}{|x|^{1+1.7}}.$$

We can represent the above tempered stable process as a sum of a stable process with index $\beta = 1.7$ and another Lévy process with activity 0.7 (these two processes will be dependent) and hence assumption A will be satisfied with $\beta' = 0.7$. This implies that X contains "residual" (to the stable) component which is quite active and therefore **Case E** represents a very stringent test for the small sample behavior of the Realized Laplace transform of volatility.

Case F: Incorrectly chosen value of the index β

The data generating process is the same as **Case E**, where the index is $\beta = 1.7$, but the computations are done under the presumption the index is 2.

4.1 Monte Carlo Results When Activity is Fixed

Table 1 summarizes the outcome of the Monte Carlo experiments pertinent to Corollary 1, where β is fixed in the computations. As seen from the table, in Cases **A**–**D** the estimator is very accurate and estimates the Laplace transform to within a precision of 0.01 or better. In **Case E** (tempered stable price jumps), however, there is a clear but relatively modest tendency to overestimate the Laplace transform across all values of the argument u.

The most interesting situation is **Case F**, where the true index is 1.7 but the computations are done on the assumption that it is 2.0. Now the estimator severely overestimates the true Laplace transform. The reason is that in forming the realized Laplace transform the increments should be inflated by the factor $(1/\Delta_n)^{1/1.7}$ but they are instead inflated by the much smaller $(1/\Delta_n)^{1/2.0}$. Using the under-inflated increments in the computations induces a very large upward bias in the estimator.

4.2 Monte Carlo Results When Activity is Estimated

Table 2 summarizes the Monte Carlo evidence pertinent to Corollary 2 where the index β is presumed unknown and estimated as using the method described in Todorov and Tauchen (2010, 2009a).[9] Only **Case A-Case E** are shown since **Case F** (an incorrectly chosen index) is not rele-

[9] The estimator is given by

$$\widehat{\beta} = \frac{\ln(2)\, p^*}{\ln(2) + \ln[\Phi_t(X, p^*, 2\Delta_n)] - \ln[\Phi_t(X, p^*, \Delta_n)]}, \tag{35}$$

Table 1: Monte Carlo Results when Activity is Fixed

	Case A	**Case B**	**Case C**	**Case D**	**Case E**	**Case F**
	$\mathcal{L}_{\lvert\sigma_t\rvert^\beta}(u)$, $u = 0.50$					
true value	0.6105	0.6105	0.6665	0.6105	0.6105	0.6105
median	0.6113	0.6106	0.6660	0.6105	0.6242	0.8065
MAD	0.0088	0.0088	0.0044	0.0089	0.0094	0.0056
	$\mathcal{L}_{\lvert\sigma_t\rvert^\beta}(u)$, $u = 1.25$					
true value	0.2992	0.2992	0.4443	0.2992	0.2992	0.2992
median	0.2991	0.2988	0.4442	0.3004	0.3172	0.6288
MAD	0.0114	0.0114	0.0056	0.0103	0.0116	0.0094
	$\mathcal{L}_{\lvert\sigma_t\rvert^\beta}(u)$, $u = 2.50$					
true value	0.0974	0.0974	0.2856	0.0974	0.0974	0.0974
median	0.0971	0.0970	0.2858	0.0978	0.1097	0.4385
MAD	0.0066	0.0066	0.0057	0.0065	0.0077	0.0114
	$\mathcal{L}_{\lvert\sigma_t\rvert^\beta}(u)$, $u = 3.75$					
true value	0.0341	0.0341	0.2104	0.0341	0.0341	0.0341
median	0.0340	0.0339	0.2107	0.0340	0.0404	0.3167
MAD	0.0034	0.0035	0.0052	0.0034	0.0040	0.0116

Note: In all simulated scenarios $T = 3000$ *and* $[1/\Delta_n] = 200$. *In all cases* $\mathbb{E}|\sigma_t|^\beta = 1$, *and the median and the median absolute deviation (MAD) correspond to the estimator* $\widehat{\mathcal{L}}_{|\sigma_t|^\beta}(u)$. *The true values of the volatility Laplace transform are computed using a sample average from a very long simulated series of the latent volatility process* σ_t. *The Monte Carlo replica is* 1000.

vant when the index is estimated. As to be be expected, the estimator of the Laplace transform is less accurate than when the activity is known as in **Case A**-**Case E** of Table 1 above. In the basic **Case A** the estimator remains quite precise in the presence of continuous stochastic volatility. On the other hand, the finitely active price or volatility jumps of **Case B** and **Case C** cause some mild downward bias in estimator while the infinitely active price jumps of **Case D** and **Case E** induce some upward bias. These biases however are small when compared with the median absolute deviation of the estimator.

5 Empirical Application

5.1 Initial Empirical Evidence

We use two data sets to illustrate the estimator and some potential uses. The first is 1-minute level data on the S&P 500 futures index, January 1, 1990, to December 31, 2008, yielding $1,900,000$ 1-minute price increments. Preliminary investigations indicated that the autocorrelations in this returns series are very small and insignificant, suggesting microstructure noise is not a serious concern for these returns[10]. The second data set consists of 5-minute observations on the S&P 500 volatility index, the VIX, from September 22, 2003, to December 31, 2008, for a total of $93,324$ price increments.[11] Similar preliminary investigations indicated the VIX returns are mildly autocorrelated at the 1-minute level, and thereby possibly affected by microstructure noise, so we use the 5-minute data where the autocorrelations are negligible.

The top two panels of Figure 1 show the estimated log-Laplace transform of the spot variance, σ_t^2, of the two series along with two-sigma confidence bands. The calculations were done using Corollary 1 with $\beta = 2$ for both series. We compare our estimates with the estimate of the log-Laplace transform of the Realized Variance, which is shown with a dashed line on the two top panels. There is a clear wedge between the two log-Laplace transforms for both data sets that becomes more evident from the LT$\times$LT plot (analogous to a Q-Q plot) in the lower panels of

where p^* is optimally chosen from a first-step estimation of the activity and the power variation $\Phi_T(X, p, \Delta_n)$ is defined as

$$\Phi_t(X, p, \Delta_n) = \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X|^p. \qquad (36)$$

[10]We also performed the estimation at 5-minute and found very little difference between the estimates at the two frequencies.

[11]The VIX index represents a traded security whose value is computed by the CBOE using a portfolio of out-of-the-money options written on the index over a wide range of strike prices. Its value is a very close approximation to that of a variance swap, a forward contract on the total quadratic variation of the log-price of the underlying asset over a fixed interval into the future. See Todorov and Tauchen (2009b) and the references therein for a much more extended discussion of the VIX index.

Table 2: Monte Carlo Results when Activity is Estimated

	Case A	**Case B**	**Case C**	**Case D**	**Case E**
		$\mathcal{L}_{\vert\sigma_t\vert^\beta}(u)$, $u = 0.50$			
true value	0.6105	0.6105	0.6665	0.6105	0.6105
median	0.6120	0.5947	0.6519	0.6291	0.6397
MAD	0.0153	0.0153	0.0103	0.0136	0.0134
		$\mathcal{L}_{\vert\sigma_t\vert^\beta}(u)$, $u = 1.25$			
true value	0.2992	0.2992	0.4443	0.2992	0.2992
median	0.3011	0.2779	0.4261	0.3252	0.3395
MAD	0.0199	0.0193	0.0128	0.0186	0.0191
		$\mathcal{L}_{\vert\sigma_t\vert^\beta}(u)$, $u = 2.50$			
true value	0.0974	0.0974	0.2856	0.0974	0.0974
median	0.0981	0.0836	0.2692	0.1158	0.1261
MAD	0.0134	0.0118	0.0124	0.0137	0.0142
		$\mathcal{L}_{\vert\sigma_t\vert^\beta}(u)$, $u = 3.75$			
true value	0.0341	0.0341	0.2104	0.0341	0.0341
median	0.0343	0.0272	0.1962	0.0440	0.0500
MAD	0.0069	0.0057	0.0106	0.0077	0.0082

Note: In all simulated scenarios $T = 3000$ *and* $[1/\Delta_n] = 200$. *In all cases* $\mathbb{E}|\sigma_t|^\beta = 1$, *and the median and the median absolute deviation (MAD) correspond to the estimator* $\widehat{\widehat{\mathcal{L}}}_{|\sigma_t|^\beta}(u)$. *In all cases* β *is estimated from the first* 252 *days of the simulated data. The true values of the volatility Laplace transform are computed using a sample average from a very long simulated series of the latent volatility process* σ_t. *The number of Monte Carlo replica is* 1000.

Figure 1.[12]

We recall that (when there are no jumps) the Realized Variance is a measure of the integrated variance $\int_t^{t+1} \sigma_s^2 ds$. Spot and integrated variance are distinct concepts and their distributions can be different, but nevertheless they are highly related as seen by a simple example. Suppose that the spot variance is a Lévy-driven process

$$\sigma_t^2 = \int_{-\infty}^{t} \kappa(t-s)\, dL(s),$$

where L is a nonnegative Lévy process and κ is a nonnegative function capturing the persistence in volatility. Then the integrated variance is

$$\int_{t-1}^{t} \sigma_u^2\, du = \int_{-\infty}^{t} \kappa^*(t-s)\, dL(s),$$

where the kernel κ^* of the integrated variance is readily determined from the kernel κ of the spot process by a simple interchange of order of integration. When volatility is persistent, κ and κ^* will be similar. Given this evident tight connection, it is interesting to explore the reasons for the wedges so apparent in Figure 1.

5.2 Model Features that Drive the Spot/Realized Variance Wedge

We undertake this investigation by computing the Laplace transforms on a single simulation for the four most pertinent cases, **B, C, D** and **F**, with the implied LT×LT plots displayed in Figure 2.[13] From **Case B** in the top left panel, it is seen that price level jumps contribute to the wedge in a manner consistent with the data, but the effect is small.[14] In **Case C** price jumps in conjunction with volatility jumps generate much more separation in the expected direction. In order to "net out" the effects of the price jumps, the top right panel also shows the LT×LT plot for the Truncated Variance of Mancini (2009) as well[15], because the relevant comparison is the spot to the integrated variation, exclusive of the contribution of squared price jumps. Using the Truncated Variance in

[12]An LT×LT plot is a plot of two log-Laplace transforms $(\log[\mathcal{L}_X(u)], \log[\mathcal{L}_Y(u)])$ as $u > 0$ varies. The LT×LT plot reveals differences just as a Q-Q plot in statistics, and if it lies everywhere on the 45-degree line then the two transforms are the same. Our convention is that the horizontal axis of the LT×LT plot corresponds to the log Laplace transform of the spot variance while the vertical axis corresponds to that of an alternative variance measure.

[13]As seen from the Monte Carlo, the amount of sampling variation is relatively small to have any impact on our comparisons here.

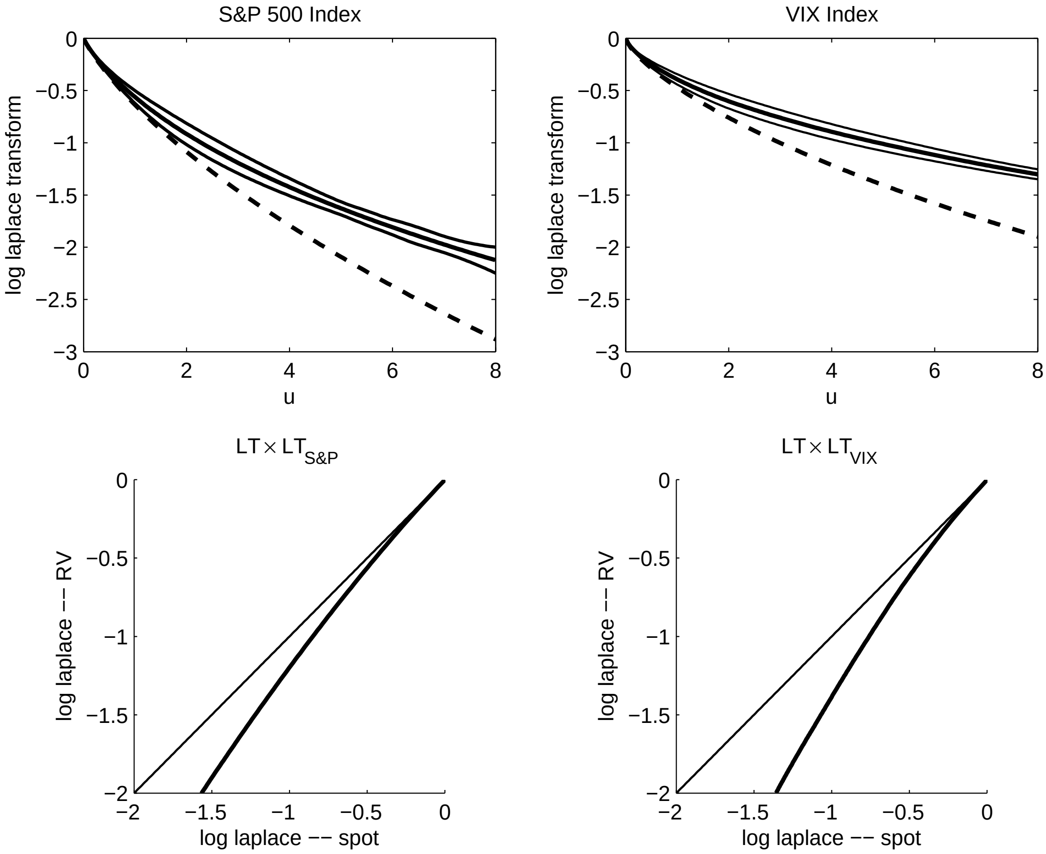
[14]The asymptotic limit of the Realized Variance in the jump-diffusion case is $\int_t^{t+1} \sigma_s^2 ds + \int_t^{t+1} \int_{\mathbb{R}} \delta^2(s,x)\mu(ds,dx)$.

[15]The truncated variance is

$$TV_T(\alpha, \varpi) = \sum_{i=1}^{[T/\Delta_n]} |\Delta_i^n X|^2 1_{\{|\Delta_i^n X| \le \alpha \Delta_n^{\varpi}\}}, \alpha > 0, \varpi \in (0, 1/2),$$

where here $\varpi = 0.49$, i.e. very close to 1/2 and α is $4 \times \sqrt{BV}$ for BV denoting the bipower variation over the day. These choices of the tuning parameters ensure reliable estimation of the integrated variance $\int_t^{t+1} \sigma_s^2\, ds$.

Figure 1: Observed Log-Laplace Transforms



Top left: Estimated log-Laplace transform and two-sigma pointwise confidence bands for spot volatility of the stock index along with the empirical Laplace transform of the daily Realized Variance, 1-minute S&P 500 index data, 1990–2008; the bottom left panel is the associated LT×LT. The two right-hand panels are same type of figures for 5-minute VIX data, 2003–2008.

place of the Realized Variance only entails a small reduction in a large wedge, suggesting that it is the volatility jumps that account for most of the separation seen in the top right panel of the figure.

In the bottom row, cases **D** and **F**, the data generation process is quite different and some care is needed in interpreting the two panels. For both cases the price process is pure-jump without a Brownian component. For the bottom left panel an appropriately adjusted Truncated Variance[16] is used to recover $\int_0^T |\sigma_s|^\beta ds$. Now the two LT×LT plots lie on top of each other, which is not surprising because the larger price jumps have been removed, volatility jumps are absent, volatility is persistent, and the *correct* value of $\beta = 1.7$ is used in computing. By contrast, for **Case F** in the lower right panel the data generation process is the same, but the incorrect value of 2.0 is used in the computations for the estimator of the Laplace transform of the spot volatility. Now there is a huge wedge which is due to the fact that the wrong scaling factor is used in the computations for the Realized Laplace transform.

The takeaway message from Figure 2 is twofold: First, both price and volatility jumps, and likely more so the latter, generate a wedge between the Laplace transforms of the spot and Realized Variance in a direction completely consistent with that seen in the data. Second, the use of an incorrect value for the activity index in computing the Realized Laplace transform can generate a very misleading estimate, because the price increments are incorrectly scaled in computing the cosine transformation (17) of the data.

5.3 Empirical Evidence Revisited

5.3.1 The Spot/Realized Variance Wedge

Guided by the findings of the preceding subsection, we re-estimated the plots of Figure 1. Now we use estimated values of the activity index β (using (35)), which are 2.002 for the S&P 500 Index and 1.733 for the VIX, in computing the spot volatility Laplace transform estimate. We also use

[16] In the pure-jump model the Truncated Variance would converge to zero. In order for it to trim out just the largest jumps, we need to scale it up. The following is easy to show for a re-scaled Truncated Variance, see e.g. Ait-Sahalia and Jacod (2010)

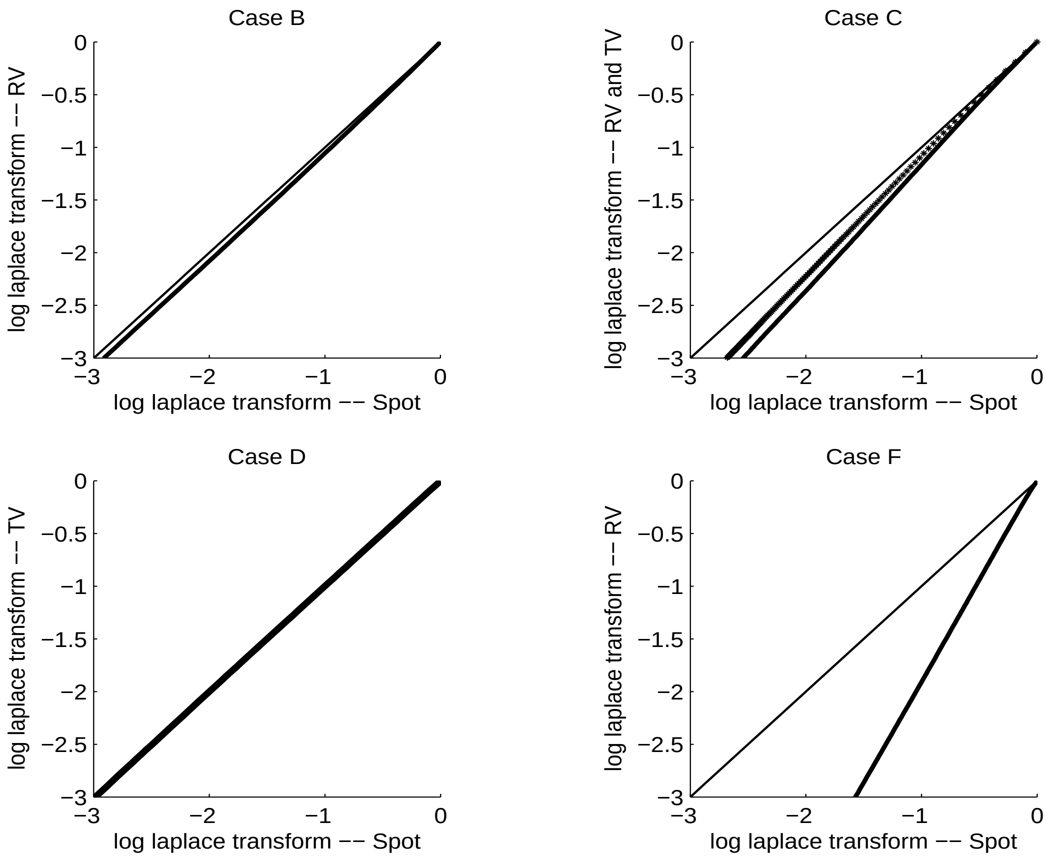
$$STV_T(\alpha, \varpi) = \Delta_n^{-(2-\beta)\varpi} \frac{\alpha^{\beta-2}(2-\beta)}{2A} TV_T(\alpha, \varpi) \xrightarrow{\mathbb{P}} \int_0^T |\sigma_s|^\beta ds.$$

Using our choice for A, this simplifies to

$$STV_T(\alpha, \varpi) = \Delta_n^{-(2-\beta)\varpi} \alpha^{\beta-2} \frac{2(2-\beta)\Gamma(2-\beta)|\cos(\beta\pi/2)|}{\beta(\beta-1)} TV_T(\alpha, \varpi),$$

from which it is easy to note that the scaling factor multiplying TV_T converges to 1 when β approaches 2 (i.e. the jump-diffusion case). In our application we set $\alpha = 15$ and $\varpi = 0.49$.

Figure 2: Model-Implied Log-Laplace Transforms Represented via LT×LT Plots



See equations (32)–(34) and the description of **Case F** in main text.

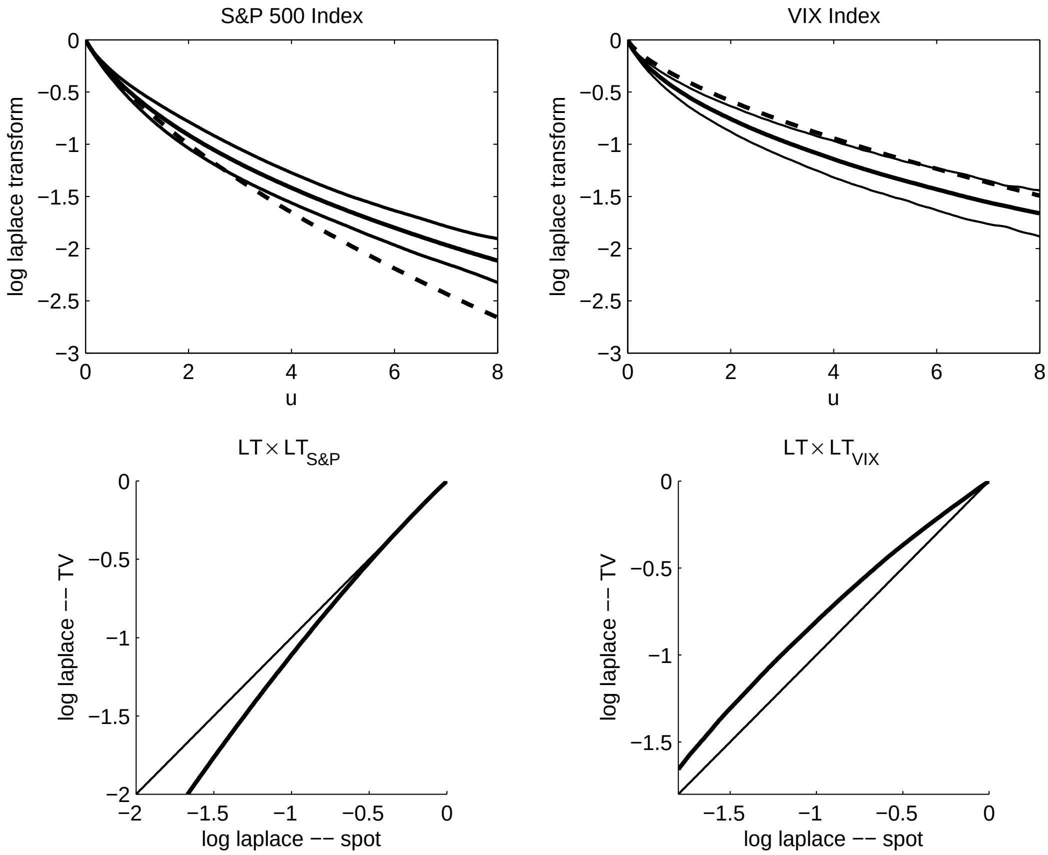
the Truncated Variance in place of the Realized Variance to compare. The results are shown in Figure 3. Interestingly, the left panels of Figure 1 and 3 for the S&P 500 Index are rather similar, except that the wedge is slightly smaller in Figure 3 due to the use of the Truncated Variance that nets out price jumps, leaving only volatility jumps to drive the wedge. On the other hand, the right panels for the VIX are quite different, with the empirical log-Laplace transform of the Truncated Variance lying just above the spot but still within the confidence band. The more modest wedge appears in the LT×LT plot in the lower right panel of Figure 3. A plausible explanation is that the VIX is a pure-jump process, consistent with the findings adduced in a much different manner by Todorov and Tauchen (2010). The use of the fixed value 2.0 for the activity index in Figure 1 improperly scaled the VIX price increments and left a misleading impression about the wedge, while the use of the estimated value of the activity index properly scales the increments and yields a more reliable depiction in Figure 3.

5.3.2 Unconditional Density of Volatility

As noted in the Introduction, an established strategy for modeling a stochastic volatility process entails specifying its unconditional density along with the dynamic dependence; our estimator is of direct use for the former and there are well established strategies for the latter. The left-hand panel of Figure 4 shows the implied log Laplace-transform of the unconditional density (a Gamma distribution) of the affine diffusion **Model A** in equation (31) along with the two-sigma confidence bands for the estimated spot variance based on Corollary 2. The model-implied transform is determined by calibrating the two parameters of the Gamma using numerical interpolation to two points at the outer range of the estimated transform of the S&P 500 index spot variance. Comparing the Gamma model-implied transform to the confidence bands suggests the empirical evidence discredits the Gamma as a candidate distribution. In contrast, the right-hand panel of Figure 4 shows the model-implied log-Laplace transform from a Generalized-Inverse-Gaussian. This three-parameter distribution nests the Gamma and is also the marginal density for many stochastic volatility models (Barndorff-Nielsen and Shephard, 2001). The parameters were calibrated by matching three points of the estimated spot volatility Laplace transform, two at outer range of the domain and one in the center. Unlike the Gamma, the Generalized-Inverse-Gaussian gives essentially a perfect fit ($R^2 \approx 1.00$) over the entire domain of the transform. The Generalized-Inverse-Gaussian thus appears to be an an excellent choice for the required marginal distribution of volatility.

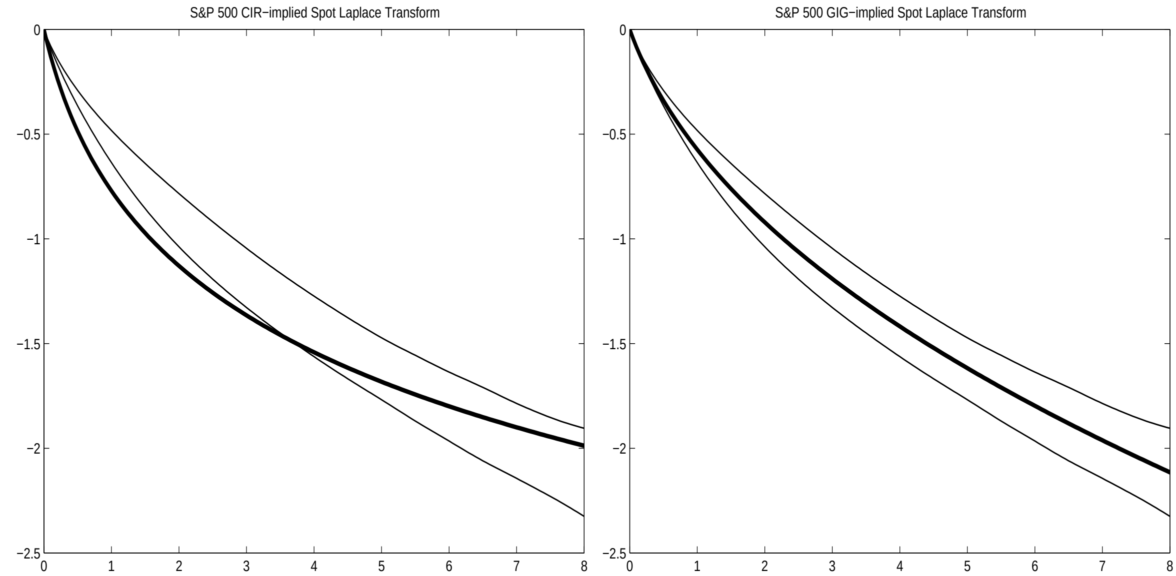
Since probability densities are easier to interpret, we show in Figure 5 the implied density under the Generalized-Inverse-Gaussian distribution for $\log |\sigma_t|^\beta$ along with that of the daily log

Figure 3: Laplace Transforms of Spot and Realized Variance with Estimated Activity Indexes



Log-Laplace transform of spot volatility and daily Realized Variance as per Figure 1 but using estimated indexes and Truncated Variance in place of the Realized Variance.

Figure 4: Model-Implied Log-Laplace Transforms of the S&P 500 Spot Variance



The left panel shows the implied log-Laplace transform for spot variance under a Gamma distribution along with the data-determined confidence interval for the nonparametric estimate of the log transform; the right panel is the same but for a Generalized-Inverse-Gaussian distribution fit.

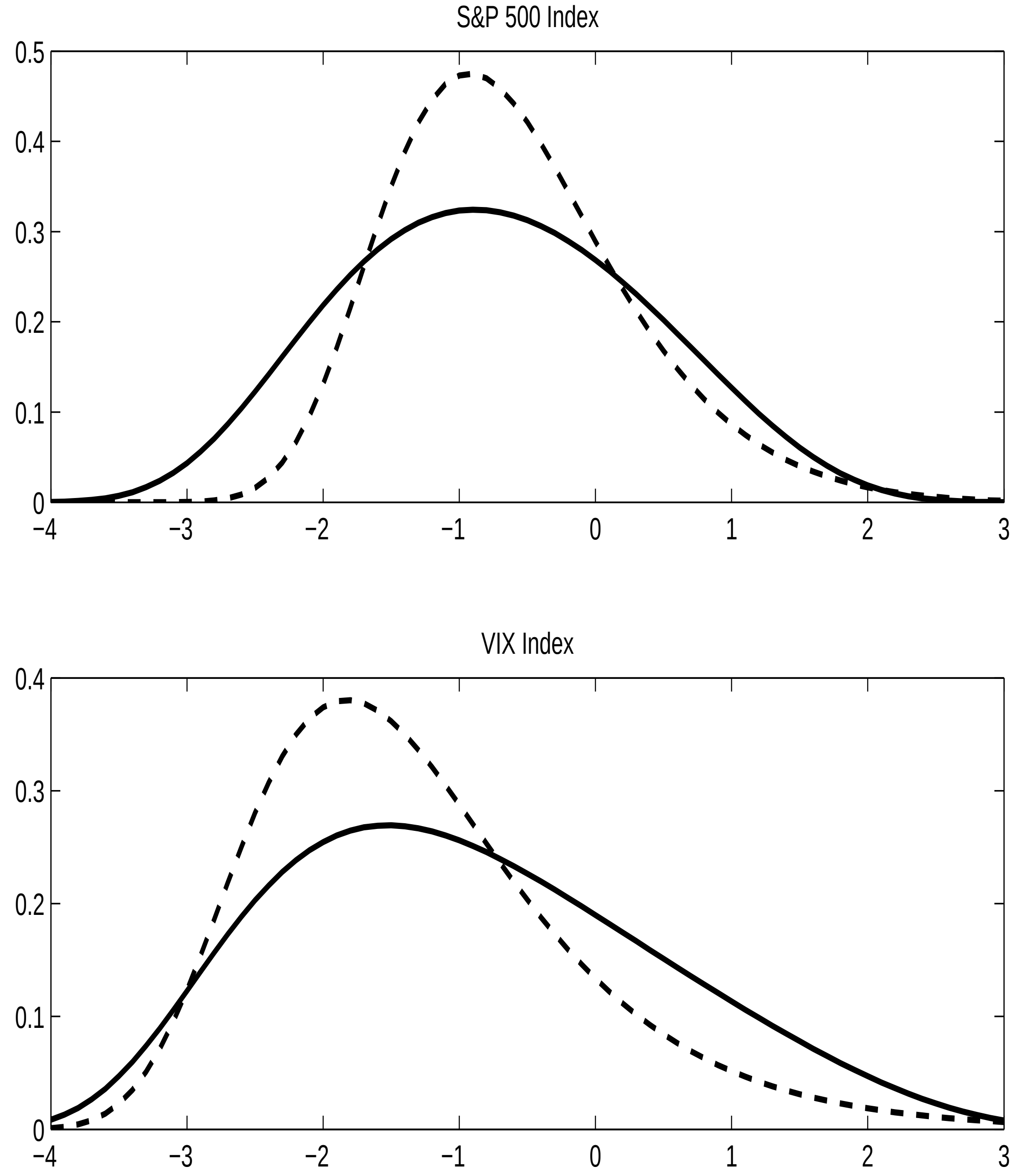
Truncated Variance (which measures the log of $\int_0^T |\sigma_s|^\beta ds$), both obtained by the interpolation scheme described above. Evidently, the density of the spot variance is more dispersed around the mode than is the the density of the integrated variance. The integration used to accumulate the daily integrated variance smoothed over sharp short-term within-day movements as would be induced by factors such as volatility jumps.[17] The density of the smoothed variance would therefore be misleading in the established two-pronged approach to modeling stochastic volatility.

6 Conclusions

We propose and derive the asymptotic properties of the Realized Laplace transform of volatility computed from high-frequency data. The results are sufficiently general to cover essentially all jump-diffusion and pure-jump processes that have been used in modeling financial prices. A crucial step in estimating reliably the Realized Laplace transform is to account correctly for the small scale behavior, i.e., the level of activity, of the driving martingale. We show that after properly accounting

[17]We checked also that the well-known deterministic within-day diurnal pattern in volatility is not a factor in this conclusion. We repeated the basic computations using diurnally adjusted data and the results were essentially the same. It is the *stochastic*, not deterministic, moves that are key here.

Figure 5: Implied Densities of Log Variance



Densities of $\log |\sigma_t|^\beta$ (solid line) and $\log \left(\int_0^T |\sigma_s|^\beta ds \right)$ (dashed line) implied by the Generalized-Inverse-Gaussian distribution fit.

for the activity, the estimator has very good robustness properties. This includes robustness against the presence of additional components in the discretely-observed process which are of lower activity like the drift term and jumps of finite variation. In a practical application we find that, after estimating and accounting for the activity levels, the Laplace transforms of the spot and integrated variance of S&P 500 index differ in an explicable way, while the transforms for the VIX index are closer.

Finally, our realized measure can be used in developing estimators for the volatility process. This can be done in two ways. One is to match the marginal density and the dependence structure with the ones observed in the data. Our realized measure is of direct use for the first part of such estimation. The second way is to match the model-implied joint Laplace transform of the volatility process over different points in time with that estimated from the data using our realized measure. This is particularly attractive for models in which the conditional Laplace transform of the volatility process is known in closed form. We leave these applications for future research.

7 Proofs

In all the proofs we will denote with C a constant that does not depend on T and Δ_n, and further it might change from line to line. We also use the short hand $\mathbb{E}_{i-1}^n$ for $\mathbb{E}\left(\cdot|\mathscr{F}_{(i-1)\Delta_n}\right)$. In all proofs we will restrict attention only to the case of pure-jump: the proofs for the jump-diffusion proceed in exactly the same way with only minor adjustments.

7.1 Proof of Lemma 1

Since $h^{-1/\beta}Z_{ht}$ is a Lévy process to prove the convergence of the sequence we need to show the convergence of its characteristics (see e.g. Jacod and Shiryaev (2003), Corollary VII.3.6), i.e., we need to establish the following for $h \to 0$

$$\begin{cases} h\int_{\mathbb{R}}\left(\kappa(h^{-1/\beta}x) - h^{-1/\beta}\kappa(x)\right)\nu(x)dx \longrightarrow 0, \\ h\int_{\mathbb{R}}\kappa^2(h^{-1/\beta}x)\nu(x)dx \longrightarrow \int_{\mathbb{R}}\kappa^2(x)\frac{A}{|x|^{\beta+1}}dx, \\ h\int_{\mathbb{R}}g(h^{-1/\beta}x)\nu(x)dx \longrightarrow \int_{\mathbb{R}}g(x)\frac{A}{|x|^{\beta+1}}dx, \end{cases} \tag{37}$$

where g is an arbitrary continuous and bounded function on $\mathbb{R}$, which is 0 around 0.

The first convergence result in (37) follows trivially as we have $\int_{\mathbb{R}}\kappa(x)\nu(x)dx < \infty$, and $\beta > 1$. The last two results of (37) follow by a change of variable in the integration and using the fact that by assumption A we have $|\nu'(x)| < \frac{K}{|x|^{\beta'+1}}$ for $|x| \leq x_0$ where x_0 is fixed and $\beta' < \beta$. □

7.2 Proof of Theorem 2

First, the Lévy measure of a standard β-stable process, i.e., a Lévy process L_t with $\mathbb{E}\left(e^{-iuL_t}\right) = e^{-t|u|^\beta}$, is

$$\frac{1}{2\int_0^\infty \frac{1-\cos(x)}{x^{\beta+1}}dx} \times \frac{1}{|x|^{\beta+1}}dx,$$

and further using Lemma 14.11 of Sato (1999), we can simplify $2\int_0^\infty \frac{1-\cos(x)}{x^{\beta+1}}dx = \frac{2\Gamma(2-\beta)|\cos(\beta\pi/2)|}{\beta(\beta-1)}$. Then the pure-jump, zero-mean stable Lévy process corresponding to the Lévy measure $A\frac{1}{|x|^{\beta+1}}dx$ is just $\left(A \times 2\int_0^\infty \frac{1-\cos(x)}{x^{\beta+1}}dx\right)^{1/\beta}$ times the pure-jump, zero-mean standard β-stable process (this follows from the self-similarity property of the stable process). Therefore it is no limitation to assume that the constant A in the theorem corresponds to that of a standard stable process and we will do so without further mention.

Throughout, after appropriately extending the original probability space, we will use an alternative representation of the process X, which is given by

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t \int_{\mathbb{R}} \sigma_{s-} x\tilde{\mu}_1(ds, dx) + \int_0^t \int_{\mathbb{R}} \sigma_{s-} x\mu_2(ds, dx) - \int_0^t \int_{\mathbb{R}} \sigma_{s-} x\mu_3(ds, dx),$$

where μ_1, μ_2 and μ_3 are homogenous Poisson measures with compensators respectively $\nu_1(dx) = \frac{A}{|x|^{\beta+1}}dx$, $\nu_2(dx) = |\nu'(x)|dx$ and $\nu_3(dx) = 2|\nu'(x)|1\left(\nu'(x) < 0\right)dx$ (the three measures are not mutually independent); $a_t = \alpha_t - \sigma_t \int_{\mathbb{R}} \kappa'(x)\frac{A}{|x|^{\beta+1}}dx - \sigma_t \int_{\mathbb{R}} \kappa(x)\nu'(x)dx$. Finally, to simplify notation we will also use $L_t = \int_0^t \int_{\mathbb{R}} x\tilde{\mu}_1(ds, dx)$ which is just a β-stable process.

With this notation we have the following decompostion

$$\sqrt{T}\left(\frac{1}{T}V_T(X, \Delta_n, \beta, u) - \mathbb{E}\left(e^{-|u\sigma_t|^\beta}\right)\right) = \sum_{i=1}^{[T/\Delta_n]} \sum_{j=1}^{3} \xi_{i,u}^{(j)} + \sqrt{T}\left(\frac{1}{T}\int_0^T e^{-|u\sigma_t|^\beta}dt - \mathbb{E}\left(e^{-|u\sigma_t|^\beta}\right)\right),$$

$$\xi_{i,u}^{(1)} = \frac{1}{\sqrt{T}}\left(\Delta_n \cos\left(u\sigma_{(i-1)\Delta_n -}\Delta_n^{-1/\beta}\Delta_i^n L\right) - \int_{(i-1)\Delta_n}^{i\Delta_n} e^{-|u\sigma_{(i-1)\Delta_n -}|^\beta}ds\right),$$

$$\xi_{i,u}^{(2)} = \frac{1}{\sqrt{T}}\int_{(i-1)\Delta_n}^{i\Delta_n}\left(e^{-|u\sigma_{(i-1)\Delta_n -}|^\beta} - e^{-|u\sigma_s|^\beta}\right)ds,$$

$$\xi_{i,u}^{(3)} = \frac{\Delta_n}{\sqrt{T}}\left(\cos\left(u\Delta_n^{-1/\beta}\Delta_i^n X\right) - \cos\left(u\sigma_{(i-1)\Delta_n -}\Delta_n^{-1/\beta}\Delta_i^n L\right)\right).$$

First, given assumption C-u and using a CLT for stationary and ergodic process, see Jacod and Shiryaev (2003), Theorem VIII.3.79, we have

$$\sqrt{T}\left(\frac{1}{T}\int_0^T e^{-|u\sigma_t|^\beta}dt - \mathbb{E}\left(e^{-|u\sigma_t|^\beta}\right)\right) \xrightarrow{\mathcal{L}} \sqrt{V_\beta(u)} \times Z,$$

where Z is the standard normal variable appearing in equation (22). Next, using the self-similarity of the stable process L, and the expression for its characteristic function, we have

$$\begin{cases} \mathbb{E}_{i-1}^n\left(\cos\left(u\sigma_{(i-1)\Delta_n -}\Delta_n^{-1/\beta}\Delta_i^n L\right) - e^{-|u\sigma_{(i-1)\Delta_n -}|^\beta}\right) = 0, \\ \mathbb{E}_{i-1}^n\left(\cos\left(u\sigma_{(i-1)\Delta_n -}\Delta_n^{-1/\beta}\Delta_i^n L\right) - e^{-|u\sigma_{(i-1)\Delta_n -}|^\beta}\right)^2 = F_\beta(u\sigma_{(i-1)\Delta_n -}), \\ \mathbb{E}_{i-1}^n\left(\cos\left(u\sigma_{(i-1)\Delta_n -}\Delta_n^{-1/\beta}\Delta_i^n L\right) - e^{-|u\sigma_{(i-1)\Delta_n -}|^\beta}\right)^4 \leq C. \end{cases}$$

Since $|F'_\beta(x)| \leq C|x|^{\beta-1}$ for $F'_\beta(x)$ denoting derivative with respect to x, the process σ_t is square integrable and stationary, and by an application of Cauchy-Schwartz inequality and assumption B, we can further write

$$\frac{1}{T}\mathbb{E}\left(\sum_{i=1}^{[T/\Delta_n]}\int_{(i-1)\Delta_n}^{i\Delta_n}|F_\beta(u\sigma_s) - F_\beta(u\sigma_{(i-1)\Delta_n -})|ds\right)$$
$$\leq \frac{C}{\Delta_n}\int_{(i-1)\Delta_n}^{i\Delta_n}\sqrt{\mathbb{E}(F'_\beta(u\sigma_s^*))^2}\sqrt{\mathbb{E}(\sigma_s - \sigma_{(i-1)\Delta_n -})^2}ds \leq C\sqrt{\Delta_n},$$

where σ_s^* is a value between σ_s and $\sigma_{(i-1)\Delta_n -}$. Then, applying Theorem VIII.2.27 of Jacod and Shiryaev (2003), we get

$$\frac{1}{\sqrt{\Delta_n}}\sum_{i=1}^{[T/\Delta_n]}\xi_{i,u}^{(1)} \xrightarrow{\mathcal{L}} \sqrt{\mathbb{E}\left(F_\beta(uZ_\beta\sigma_t)\right)} \times U, \tag{38}$$

for U the standard normal variable in equation (23).

We are left with the terms involving $\xi_{i,u}^{(2)}$ and $\xi_{i,u}^{(3)}$. We will determine their order of magnitude by bounding their first or second moment. We start with the term involving $\xi_{i,u}^{(2)}$. When only assumption B is made, we can use a first-order Taylor expansion and Cauchy-Schwartz inequality to get

$$\mathbb{E}|\xi_{i,u}^{(2)}| \leq C\frac{1}{\sqrt{T}}\mathbb{E}\left(\int_{(i-1)\Delta_n}^{i\Delta_n}|\sigma_s^*|^{\beta-1}|\sigma_s - \sigma_{(i-1)\Delta_n}|ds\right)$$
$$\leq C\frac{1}{\sqrt{T}}\int_{(i-1)\Delta_n}^{i\Delta_n}\sqrt{\mathbb{E}|\sigma_s^*|^{2\beta-2}\mathbb{E}|\sigma_s - \sigma_{(i-1)\Delta_n}|^2}ds \leq \frac{C\Delta_n^{3/2}}{\sqrt{T}},$$

where σ_s^* is some value between σ_s and $\sigma_{(i-1)\Delta_n}$. Therefore, under assumption B we get

$$(\sqrt{T\Delta_n})^{-1}\mathbb{E}\left(\sum_{i=1}^{[T/\Delta_n]}|\xi_{i,u}^{(2)}|\right) \leq C.$$

Now we derive the bound for $\xi_{i,u}^{(2)}$ when we assume the stronger assumption B$'$. In this case we first can decompose $\xi_{i,u}^{(2)}$ using first-order Taylor expansion as

$$\xi_{i,u}^{(2)} = \sum_{j=1}^{3}\xi_{i,u}^{(2)}(j),$$

where

$$\xi^{(2)}_{i,u}(1) = \frac{K_1(\sigma_{(i-1)\Delta_n-}, u)}{\sqrt{T}} \int_{(i-1)\Delta_n}^{i\Delta_n} \left(\int_{(i-1)\Delta_n}^{s} \tilde{\sigma}_u dW_u + \int_{(i-1)\Delta_n}^{s} \int_{\mathbb{R}} \delta(u-,x)\underline{\tilde{\mu}}(du,dx) \right) ds,$$

$$\begin{aligned} \xi^{(2)}_{i,u}(2) = \frac{1}{\sqrt{T}} \int_{(i-1)\Delta_n}^{i\Delta_n} \left(K_1(\sigma^*_s, u) - K_1(\sigma_{(i-1)\Delta_n-,u}) \right) \Big(\int_{(i-1)\Delta_n}^{s} \tilde{\sigma}_u dW_u \\ + \int_{(i-1)\Delta_n}^{s} \int_{\mathbb{R}^2} \delta(u-,x)\underline{\tilde{\mu}}(du,dx) \Big) ds, \end{aligned}$$

$$\xi^{(2)}_{i,u}(3) = \frac{1}{\sqrt{T}} \int_{(i-1)\Delta_n}^{i\Delta_n} \left(e^{-|u\widehat{\sigma}_s|^\beta} - e^{-|u\sigma_s|^\beta} \right) ds,$$

where $K_1(x,u) = -\beta \text{sign}\{x\}|u|^\beta |x|^{\beta-1} e^{-|ux|^\beta}$, σ^*_s is a number between $\sigma_{(i-1)\Delta_n-}$ and $\widehat{\sigma}_s$, and

$$\widehat{\sigma}_s = \sigma_{(i-1)\Delta_n-} + \int_{(i-1)\Delta_n}^{s} \tilde{\sigma}_u dW_u + \int_{(i-1)\Delta_n}^{s} \int_{\mathbb{R}} \delta(u-,x)\underline{\tilde{\mu}}(du,dx), \quad s \in [(i-1)\Delta_n, i\Delta_n].$$

Since $E^n_{i-1}\left(\int_{(i-1)\Delta_n}^{i\Delta_n} |\tilde{\sigma}_s|^2 ds + \int_{(i-1)\Delta_n}^{i\Delta_n} \int_{\mathbb{R}} |\delta(s,x)|^2 \underline{\nu}(ds,dx) \right) < \infty$, we have $\mathbb{E}^n_{i-1}\left(\xi^{(2)}_{i,u}(1) \right) = 0$. Using Itô 's isometry and square integrability we further have

$$\begin{aligned} \mathbb{E}^n_{i-1}\left(\xi^{(2)}_{i,u}(1) \right)^2 \leq & C\Delta_n \frac{K_1^2(\sigma_{(i-1)\Delta_n-}, u)}{T} \\ & \times \mathbb{E}^n_{i-1} \int_{(i-1)\Delta_n}^{i\Delta_n} \left(\int_{(i-1)\Delta_n}^{s} \tilde{\sigma}_u dW_u + \int_{(i-1)\Delta_n}^{s} \int_{\mathbb{R}} \delta(u-,x)\underline{\tilde{\mu}}(du,dx) \right)^2 ds \\ \leq & C\Delta_n \frac{K_1^2(\sigma_{(i-1)\Delta_n-}, u)}{T} \\ & \times \mathbb{E}^n_{i-1} \int_{(i-1)\Delta_n}^{i\Delta_n} \left(\int_{(i-1)\Delta_n}^{s} \tilde{\sigma}_u^2 du + \int_{(i-1)\Delta_n}^{s} \int_{\mathbb{R}} \delta^2(u-,x)\underline{\nu}(du,dx) \right) ds. \end{aligned}$$

Therefore using the integrability conditions in assumption B′, we have altogether

$$\sum_{i=1}^{[T/\Delta_n]} \mathbb{E}^n_{i-1}\left(\xi^{(2)}_{i,u}(1) \right) = 0, \qquad \Delta_n^{-2} \mathbb{E}\left(\sum_{i=1}^{[T/\Delta_n]} \mathbb{E}^n_{i-1}\left(\xi^{(2)}_{i,u}(1) \right)^2 \right) \leq C. \tag{39}$$

For $\xi^{(2)}_{i,u}(2)$, by using Cauchy-Schwartz inequality, Itô 's isometry and the integrability conditions of assumption B′, we can write

$$\begin{aligned} \mathbb{E}|\xi^{(2)}_{i,u}(2)| \leq & \frac{C}{\sqrt{T}} \int_{(i-1)\Delta_n}^{i\Delta_n} \sqrt{\mathbb{E}(K_1(\sigma^*_s,u) - K_1(\sigma_{(i-1)\Delta_n-},u))^2} \sqrt{s-(i-1)\Delta_n} ds \\ \leq & \frac{C\Delta_n^{3/2}}{\sqrt{T}} \sqrt{\mathbb{E}\left(\sup_{s \in [(i-1)\Delta_n, i\Delta_n]} (K_1(\sigma^*_s,u) - K_1(\sigma_{(i-1)\Delta_n-},u))^2 \right)}. \end{aligned}$$

To continue further we make use of the following algebraic inequality

$$\begin{aligned}|K_1(x,u)-K_1(y,u)| \leq & C|x-y|^{\beta-1}+C|y|^{\beta-1}1_{\{|x-y|\geq 0.5|y|\}}\\ & +C|y|^{2(\beta-1)}|x-y|+C|y|^{\beta-1}|x-y|^{\beta}, \quad x,y\in\mathbb{R},\ u>0,\end{aligned}$$

where the constant C depends only on the value of u. We can further simplify this inequality upon noticing $|y|^{\beta-1}1_{\{|x-y|\geq 0.5|y|\}} \leq C|x-y|^{\beta-1}$. Plugging in the above inequality $x=\sigma_s^*$ and $y=\sigma_{(i-1)\Delta_n-}$, using successive conditioning (first on the filtration $\mathscr{F}_{(i-1)\Delta_n}$), Burkholder-Davis-Gundy inequality, and finally the Holder inequality combined with the integrability conditions of the theorem, we get

$$\sqrt{\mathbb{E}\left(\sup_{s\in[(i-1)\Delta_n,i\Delta_n]}(K_1(\sigma_s^*,u)-K_1(\sigma_{(i-1)\Delta_n-},u))^2\right)} \leq C\Delta_n^{\beta/2-1/2}.$$

Therefore

$$(\sqrt{T}\Delta_n^{\beta/2})^{-1}\mathbb{E}\left(\sum_{i=1}^{[T/\Delta_n]}\mathbb{E}_{i-1}^n|\xi_{i,u}^{(2)}(2)|\right)\leq C. \tag{40}$$

Finally, first-order Taylor expansion implies

$$\mathbb{E}_{i-1}^n|\xi_{i,u}^{(2)}(3)| \leq \frac{C|u|^{\beta}}{\sqrt{T}}\mathbb{E}_{i-1}^n\int_{(i-1)\Delta_n}^{i\Delta_n}\left||\sigma_s|^{\beta}-|\widehat{\sigma}_s|^{\beta}\right|ds,$$

and using the integrability conditions in assumption B′, we can write

$$(\sqrt{T}\Delta_n)^{-1}\mathbb{E}\left(\sum_{i=1}^{[T/\Delta_n]}\mathbb{E}_{i-1}^n|\xi_{i,u}^{(2)}(3)|\right)\leq C. \tag{41}$$

Turning to $\xi_{i,u}^{(3)}$, we can decompose it as

$$\xi_{i,u}^{(3)}=\sum_{j=1}^{4}\xi_{i,u}^{(3)}(j),$$

where

$$\begin{aligned}\xi_{i,u}^{(3)}(1)=&\frac{-2\Delta_n}{\sqrt{T}}\sin\left(0.5u\Delta_n^{-1/\beta}\left(\Delta_i^nX+\int_{(i-1)\Delta_n}^{i\Delta_n}a_sds+\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{\mathbb{R}}\sigma_{s-}x\tilde{\mu}_1(ds,dx)\right)\right)\\ &\times\sin\left(0.5u\Delta_n^{-1/\beta}\int_{(i-1)\Delta_n}^{i\Delta_n}\sigma_{s-}x\mu_2(ds,dx)-0.5u\Delta_n^{-1/\beta}\int_{(i-1)\Delta_n}^{i\Delta_n}\sigma_{s-}x\mu_3(ds,dx)\right),\end{aligned}$$

$$\xi^{(3)}_{i,u}(2) = \frac{-u\Delta_n^{2-1/\beta}}{\sqrt{T}} \sin\left(u\sigma_{(i-1)\Delta_n-}\Delta_n^{-1/\beta}\Delta_i^n L\right) a_{(i-1)\Delta_n},\ \xi^{(3)}_{i,u}(3) = \frac{0.5u^2\Delta_n^{3-2/\beta}}{\sqrt{T}} \cos(\tilde{x}_2)\, a^2_{(i-1)\Delta_n},$$

$$\xi^{(3)}_{i,u}(4) = \frac{-u\Delta_n^{1-1/\beta}}{\sqrt{T}} \sin(\tilde{x}_1)\Big(\int_{(i-1)\Delta_n}^{i\Delta_n} (a_s - a_{(i-1)\Delta_n})ds + \int_{(i-1)\Delta_n}^{i\Delta_n}\int_{\mathbb{R}} \kappa'(x)(\sigma_{s-} - \sigma_{(i-1)\Delta_n-})\tilde{\mu}_1(ds,dx)\Big),$$

$$\xi^{(3)}_{i,u}(5) = \frac{-u\Delta_n^{1-1/\beta}}{\sqrt{T}} \sin\left(u\sigma_{(i-1)\Delta_n-}\Delta_n^{-1/\beta}\Delta_i^n L\right) \int_{(i-1)\Delta_n}^{i\Delta_n}\int_{\mathbb{R}} \kappa(x)(\sigma_{s-} - \sigma_{(i-1)\Delta_n-})\tilde{\mu}_1(ds,dx),$$

$$\xi^{(3)}_{i,u}(6) = \frac{0.5u^2\Delta_n^{1-2/\beta}}{\sqrt{T}} \cos(\tilde{x}_2) \left(\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{\mathbb{R}} \kappa(x)(\sigma_{s-} - \sigma_{(i-1)\Delta_n-})\tilde{\mu}_1(ds,dx)\right)^2,$$

where $\tilde{x}_1$ is between

$$u\Delta_n^{-1/\beta}\sigma_{(i-1)\Delta_n-}\Delta_i^n L + u\Delta_n^{1-1/\beta}a_{(i-1)\Delta_n} + u\Delta_n^{-1/\beta}\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{\mathbb{R}} \kappa(x)(\sigma_{s-} - \sigma_{(i-1)\Delta_n-})\tilde{\mu}_1(ds,dx)$$

and

$$u\Delta_n^{-1/\beta}\int_{(i-1)\Delta_n}^{i\Delta_n} a_s ds + u\Delta_n^{-1/\beta}\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{\mathbb{R}} \kappa(x)\sigma_{s-}\tilde{\mu}_1(ds,dx)$$

and $\tilde{x}_2$ is between

$$u\Delta_n^{-1/\beta}\sigma_{(i-1)\Delta_n-}\Delta_i^n L + u\Delta_n^{1-1/\beta}a_{(i-1)\Delta_n} + u\Delta_n^{-1/\beta}\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{\mathbb{R}} \kappa(x)(\sigma_{s-} - \sigma_{(i-1)\Delta_n-})\tilde{\mu}_1(ds,dx)$$

and $u\Delta_n^{-1/\beta}\sigma_{(i-1)\Delta_n-}\Delta_i^n L$.

Using the basic inequalities $|\sin(x)| \le |x|$ and $|\sum_i |a_i||^p \le \sum_i |a|_i^p$ for some $0 < p \le 1$ as well as the Burkholder-Davis-Gundy inequality, we have

$$\mathbb{E}|\xi^{(3)}_{i,u}(1)| \le \frac{C\Delta_n^{1-\beta'/\beta}}{\sqrt{T}}\mathbb{E}\left(\left|\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{\mathbb{R}} \sigma_{s-}x\mu_2(ds,dx)\right|^{\beta'} + \left|\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{\mathbb{R}} \sigma_{s-}x\mu_3(ds,dx)\right|^{\beta'}\right),$$

$$\begin{aligned}
\mathbb{E}\left|\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{\mathbb{R}} \sigma_{s-}x\mu_j(ds,dx)\right|^{\beta'} &\le C\Delta_n \\
&\quad + C\mathbb{E}\left(\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{|x|<\Delta_n^{1/\beta'}} x^2\sigma_{s-}^2\mu_j(ds,dx)\right)^{\beta'/2} + C\mathbb{E}\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{|x|\ge\Delta_n^{1/\beta'}} |x|^{\beta'}|\sigma_{s-}|^{\beta'}\mu_j(ds,dx) \\
&\le \left(\int_{|x|<\Delta_n^{1/\beta'}} x^2|\nu_2(x)|dx\right)^{\beta'/2}\left(\int_{(i-1)\Delta_n}^{i\Delta_n}\mathbb{E}\sigma_{s-}^2 ds\right)^{\beta'/2} + C\Delta_n\int_{|x|\ge\Delta_n^{1/\beta'}} |x|^{\beta'}|\nu_2(x)|dx \\
&\le C\Delta_n|\log\Delta_n|, \quad j = 2,3,
\end{aligned}$$

where we made use of the assumption that $|\nu_2(x)| \le \frac{K}{|x|^{\beta'+1}}$ for $|x|$ sufficiently small. Therefore

$$(\sqrt{T}|\log(\Delta_n)|\Delta_n^{1-\beta'/\beta})^{-1}\sum_{i=1}^{[T/\Delta_n]}\mathbb{E}|\xi^{(3)}_{i,u}(1)| \le C. \tag{42}$$

For $\xi_{i,u}^{(3)}(2)$, using the boundedness of the function $\sin(x)$ and the square integrability of a_s from assumption B, we trivially have

$$\mathbb{E}_{i-1}^{n}\left(\xi_{i,u}^{(3)}(2)\right) = 0, \quad (\Delta_n^{3-2/\beta})^{-1}\mathbb{E}\left(\sum_{i=1}^{[T/\Delta_n]} \mathbb{E}_{i-1}^{n}\left(\xi_{i,u}^{(3)}(2)|\right)^2\right) \leq C. \tag{43}$$

Similarly,

$$(\sqrt{T}\Delta_n^{2-2/\beta})^{-1}\mathbb{E}\left(\sum_{i=1}^{[T/\Delta_n]} |\xi_{i,u}^{(3)}(3)|\right) \leq C. \tag{44}$$

For $\xi_{i,u}^{(3)}(4)$ we first can use assumption B and apply Jensen's inequality to get

$$\mathbb{E}|a_s - a_{(i-1)\Delta_n}| \leq C\sqrt{s-(i-1)\Delta_n}, \quad \mathbb{E}|\sigma_s - \sigma_{(i-1)\Delta_n}| \leq C\sqrt{s-(i-1)\Delta_n},$$

where $s \in [(i-1)\Delta_n, i\Delta_n]$. Therefore, upon noticing that $\kappa'(x)$ is 0 around 0 (and hence the integral with respect to $\tilde{\mu}$ in $\xi_{i,u}^{(3)}(4)$ becomes the usual Riemann integral), we get from the above inequalities

$$(\sqrt{T}\Delta_n^{3/2-1/\beta})^{-1}\sum_{i=1}^{[T/\Delta_n]} \mathbb{E}|\xi_{i,u}^{(3)}(4)| \leq C. \tag{45}$$

Turning to $\xi_{i,u}^{(3)}(5)$, we have trivially when only assumption B holds:

$$(\sqrt{T\Delta_n})^{-1}\mathbb{E}\left|\sum_{i=1}^{[T/\Delta_n]} \xi_{i,u}^{(3)}(5)\right| \leq C. \tag{46}$$

We derive now a tighter bound when the stronger assumption B′ is assumed. The proof proceeds through splitting $\sigma_s - \sigma_{(i-1)\Delta_n}$ for $s \in [(i-1)\Delta_n, i\Delta_n]$, as follows

$$\begin{aligned}\sigma_s - \sigma_{(i-1)\Delta_n} = & \int_{(i-1)\Delta_n}^{s} \tilde{\alpha}_u du + \int_{(i-1)\Delta_n}^{s} \tilde{\sigma}_{(i-1)\Delta_n} dW_u + \int_{(i-1)\Delta_n}^{s}\int_{\mathbb{R}} \delta((i-1)\Delta_n-, x)\underline{\tilde{\mu}}(du, dx) \\ & + \int_{(i-1)\Delta_n}^{s} (\tilde{\sigma}_u - \tilde{\sigma}_{(i-1)\Delta_n}) dW_u + \int_{(i-1)\Delta_n}^{s}\int_{\mathbb{R}} (\delta(u-, x) - \delta((i-1)\Delta_n-, x))\underline{\tilde{\mu}}(du, dx).\end{aligned}$$

Then for each of the terms we can argues as follows. First we can split the range of integration:

$$\begin{aligned}&\mathbb{E}\left|\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{\mathbb{R}} \kappa(x) \int_{(i-1)\Delta_n}^{s} \left(\tilde{\sigma}_u - \tilde{\sigma}_{(i-1)\Delta_n}\right) dW_u \ \tilde{\mu}_1(ds, dx)\right| \\ &= \mathbb{E}\left|\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{|x|<\Delta_n^{1/\beta}} \kappa(x) \int_{(i-1)\Delta_n}^{s} \left(\tilde{\sigma}_u - \tilde{\sigma}_{(i-1)\Delta_n}\right) dW_u \ \tilde{\mu}_1(ds, dx)\right| \\ &\quad + \mathbb{E}\left|\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{|x|\geq\Delta_n^{1/\beta}} \kappa(x) \int_{(i-1)\Delta_n}^{s} \left(\tilde{\sigma}_u - \tilde{\sigma}_{(i-1)\Delta_n}\right) dW_u \ \tilde{\mu}_1(ds, dx)\right|.\end{aligned}$$

Then for the first integral on the right hand side of the above decomposition we can use Burkholder-Davis-Gundy inequality and get

$$\mathbb{E}\left|\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{|x|<\Delta_n^{1/\beta}}\kappa(x)\int_{(i-1)\Delta_n}^{s}\left(\tilde{\sigma}_u-\tilde{\sigma}_{(i-1)\Delta_n}\right)dW_u\ \tilde{\mu}_1(ds,dx)\right|$$
$$\leq C\Delta_n^{1/\beta-1/2}\sqrt{\int_{(i-1)\Delta_n}^{i\Delta_n}\mathbb{E}\left(\int_{(i-1)\Delta_n}^{s}\left(\tilde{\sigma}_u-\tilde{\sigma}_{(i-1)\Delta_n}\right)dW_u\right)^2 ds}$$
$$\leq C\Delta_n^{1/\beta-1/2}\sqrt{\mathbb{E}\left(\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{(i-1)\Delta_n}^{s}\left(\tilde{\sigma}_u-\tilde{\sigma}_{(i-1)\Delta_n}\right)^2 duds\right)}\leq C\Delta_n^{1/\beta+1}.$$

where we made use of the definition of $\nu_1(dx)$ and the fact that for x sufficiently close to 0, $\kappa(x)=x$. For the second integral we can decompose as integration with respect to μ and the compensated measure and then use again Burkholder-Davis-Gundy inequality to get

$$\mathbb{E}\left|\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{|x|\geq\Delta_n^{1/\beta}}\kappa(x)\int_{(i-1)\Delta_n}^{s}\left(\tilde{\sigma}_u-\tilde{\sigma}_{(i-1)\Delta_n}\right)dW_u\ \tilde{\mu}_1(ds,dx)\right|$$
$$\leq C\int_{|x|\geq\Delta_n^{1/\beta}}|\kappa(x)|\nu_1(dx)\mathbb{E}\left(\int_{(i-1)\Delta_n}^{i\Delta_n}\left|\int_{(i-1)\Delta_n}^{s}(\tilde{\sigma}_u-\tilde{\sigma}_{(i-1)\Delta_n})dW_u\right|ds\right)\leq C\Delta_n^{1/\beta+1}.$$

Similar analysis also gives the following bounds

$$\mathbb{E}\left|\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{\mathbb{R}}\kappa(x)\int_{(i-1)\Delta_n}^{s}\tilde{\alpha}_u du\ \tilde{\mu}_1(ds,dx)\right|\leq C\Delta_n^{1/\beta+1},$$
$$\mathbb{E}\left|\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{\mathbb{R}}\kappa(x)\int_{(i-1)\Delta_n}^{s}\int_{\mathbb{R}}\left(\delta(u-,x)-\delta((i-1)\Delta_n-,z)\right)\underline{\tilde{\mu}}(du,dz)\ \tilde{\mu}_1(ds,dx)\right|\leq C\Delta_n^{1/\beta+1}.$$

To continue further we denote $Y_s^n=\int_{(i-1)\Delta_n}^{s}\int_{\mathbb{R}}\kappa(x)\tilde{\mu}_1(du,dx)$ and $\widetilde{Y}_s^n=\int_{(i-1)\Delta_n}^{s}\int_{\mathbb{R}}\kappa'(x)\tilde{\mu}_1(du,dx)$ for $s\in[(i-1)\Delta_n,i\Delta_n]$. We also set $Z_s^n=\int_{(i-1)\Delta_n}^{s}\tilde{\sigma}_{(i-1)\Delta_n}dW_u+\int_{(i-1)\Delta_n}^{s}\int_{\mathbb{R}}\delta((i-1)\Delta_n-,x)\underline{\tilde{\mu}}(du,dx)$ when $\underline{\mu}$ is independent from μ and $Z_s^n=\int_{(i-1)\Delta_n}^{s}\tilde{\sigma}_{(i-1)\Delta_n}dW_u$ when this is not the case. Note that Z_s^n is time-homogenous martingale independent from the stable process L (and the measure μ). This follows from our assumption on $\underline{\mu}$ and the fact that the Brownian motion and a homogenous Poisson measure generate independent filtration. With this notation using integration by parts, we have

$$\mathbb{E}_{i-1}^n\left(\sin\left(\Delta_n^{-1/\beta}\sigma_{(i-1)\Delta_n-}(Y_{i\Delta_n}+\widetilde{Y}_{i\Delta_n})\right)\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{\mathbb{R}}\kappa(x)Z_{s-}^n\tilde{\mu}_1(ds,dx)\right)$$
$$=\mathbb{E}_{i-1}^n\left(\sin\left(\Delta_n^{-1/\beta}\sigma_{(i-1)\Delta_n-}(Y_{i\Delta_n}+\widetilde{Y}_{i\Delta_n})\right)\left(Y_{i\Delta_n}^n Z_{i\Delta_n}^n-\int_{(i-1)\Delta_n}^{i\Delta_n}Y_{s-}^n dZ_s^n\right)\right)=0.$$

where we made use of the independence of Z_s^n from Y_s^n and $\tilde{Y}_s^n$.[18] Finally for the case when μ and $\underline{\mu}$ are not necessarily independent, we have the following additional bound.

$$\begin{aligned}
&\mathbb{E}\left|\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{\mathbb{R}}\kappa(x)\int_{(i-1)\Delta_n}^{s}\int_{\mathbb{R}}\delta(u-,z)\underline{\tilde{\mu}}(du,dz)\tilde{\mu}_1(ds,dx)\right|^{\beta+\iota}\\
&\leq C\mathbb{E}\int_{(i-1)\Delta_n}^{i\Delta_n}\left|\int_{(i-1)\Delta_n}^{s}\int_{\mathbb{R}}\delta(u-,z)\underline{\tilde{\mu}}(du,dz)\right|^{\beta+\iota}ds\\
&\leq C\mathbb{E}\int_{(i-1)\Delta_n}^{i\Delta_n}\left|\int_{(i-1)\Delta_n}^{s}\int_{\mathbb{R}}|\delta(u-,z)|^{\beta\vee\beta''+\iota}\underline{\mu}(du,dz)\right|^{\frac{\beta+\iota}{\beta\vee\beta''+\iota}}ds\leq C\Delta_n^{1+\frac{\beta+\iota}{\beta\vee\beta''+\iota}},
\end{aligned}$$

for $\iota>0$ arbitrary small. Thus altogether under assumption B$'$ we have

$$\begin{cases}(\sqrt{T}\Delta_n)^{-1}\mathbb{E}\left|\sum_{i=1}^{[T/\Delta_n]}\mathbb{E}_{i-1}^n\xi_{i,u}^{(3)}(5)\right|\leq C, & \text{when } \mu \text{ and } \underline{\mu} \text{ are independent},\\ \left(\sqrt{T}\Delta_n^{1/(\beta\vee\beta''+\iota)}\right)^{-1}\mathbb{E}\left|\sum_{i=1}^{[T/\Delta_n]}\mathbb{E}_{i-1}^n\xi_{i,u}^{(3)}(5)\right|\leq C, & \text{where } \iota \text{ is arbitrary small.}\end{cases}\tag{47}$$

On the other hand using the boundedness of the $\sin(x)$ function, Itô 's isometry (note that $\kappa(x)$ has bounded support and therefore $\int_{\mathbb{R}}\kappa^2(x)\nu_1(dx)<\infty$), and the fact that $\mathbb{E}\int_{(i-1)\Delta_n}^{i\Delta_n}(\sigma_s-\sigma_{(i-1)\Delta_n})^2ds\leq C\Delta_n^2$ gives

$$(\Delta_n^{3-2/\beta})^{-1}\sum_{i=1}^{[T/\Delta_n]}\mathbb{E}\left(\xi_{i,u}^{(3)}(5)|\right)^2\leq C.\tag{48}$$

Similar transformations yield

$$(\sqrt{T}\Delta_n^{2-2/\beta})^{-1}\sum_{i=1}^{[T/\Delta_n]}\mathbb{E}|\xi_{i,u}^{(3)}(6)|\leq C.\tag{49}$$

□

7.3 Proof of Theorem 1

The results for $\sqrt{T}\xi_{i,u}^{(1)}$, $\sqrt{T}\xi_{i,u}^{(2)}$ and $\sqrt{T}\xi_{i,u}^{(3)}$ in the proof of Theorem 2 above can be applied directly in showing that $\frac{1}{\sqrt{\Delta_n}}\left(V_T(X,\Delta_n,\beta,u)-\int_0^T e^{-|u\sigma_t^\beta Z_\beta|^\beta}ds\right)$ converges in distribution to a continuous martingale with quadratic variation $\int_0^T F_\beta(uZ_\beta\sigma_s)ds$. Thus, we only need to show that the convergence holds stably on the original probability space. To prove the latter, using Theorem IX.7.28 of Jacod and Shiryaev (2003), we need to show only

$$\sum_{i=1}^{[T/\Delta_n]}\mathbb{E}_{i-1}^n\left(\sqrt{T}\xi_{i,u}^{(1)}\Delta_i^nM\right)\overset{\mathbb{P}}{\longrightarrow}0,\tag{50}$$

[18] In the case $\beta=2$, i.e., the jump-diffusion model, we will have an additional term $\sin\left(u\sigma_{(i-1)\Delta_n}\Delta_n^{-1/2}\Delta_i^nW\right)\int_{(i-1)\Delta_n}^{i\Delta_n}(W_s-W_{(i-1)\Delta_n})dW_s$. This term will be in expectation zero because of the symmetry of the Brownian motion and the fact that $\sin(x)$ is symmetric in x.

where M is a bounded martingale defined on the original probability space. When M is discontinuous martingale, the result follows from the fact that the limit of $\sum_{i=1}^{[T/\Delta_n]} \sqrt{T}\xi_{i,u}^{(1)}$ is a continuous process and the fact that continuous and pure-jump martingales are orthogonal, see e.g., I.4.11 of Jacod and Shiryaev (2003). When M is a continuous martingale, we can write $\mathbb{E}_{i-1}^n\left(\sqrt{T}\xi_{i,u}^{(1)}\Delta_i^n M\right) = \mathbb{E}_{i-1}^n\left(\Delta_i^n N \Delta_i^n M\right)$ where $N_t = \mathbb{E}(\sqrt{T}\xi_{i,u}^{(1)}|\mathscr{F}_t^*)$ for $t \in [(i-1)\Delta_n, i\Delta_n]$ and $\mathscr{F}_t^* = \mathscr{F}_{(i-1)\Delta_n} \cap \mathscr{F}_t^{\mu_1}$ for $\mathscr{F}_t^{\mu_1}$ denoting the filtration generated by the jump measure μ_1. Then using a martingale representation for the martingale $(N_t)_{t\geq(i-1)\Delta_n}$ with respect to the filtration $\mathscr{F}_t^{\mu_1}$ (note μ_1 is a homogenous Poisson measure), we can represent N_t as an integral with respect to $\tilde{\mu}_1$. But then since pure-jump and continuous martingales are orthogonal, we have $\mathbb{E}_{i-1}^n\left(\sqrt{T}\xi_{i,u}^{(1)}\Delta_i^n M\right) = 0$.
□

7.4 Proof of Theorem 3

In the proof we use the same decomposition of the difference $\sqrt{T}\left(\frac{1}{T}V_T(X,\Delta_n,\beta,u) - \mathbb{E}\left(e^{-|u\sigma_t|^\beta}\right)\right)$ as in the proof of Theorem 2, and as in that proof we choose A so that $Z_\beta = 1$. Since assumption C' implies C-u for any u, we have from Theorem 2 the finite dimensional convergence of $\frac{1}{\sqrt{T}}\int_0^T(e^{-|u\sigma_t|^\beta} - \mathbb{E}(e^{-|u\sigma_t|^\beta}))dt$ to a Gaussian process with variance-covariance matrix given in (24). Hence we are left with establishing the tightness of the sequence. For this lets denote for arbitrary $u, v \geq 0$:

$$z_t = (e^{-|u\sigma_t|^\beta} - \mathbb{E}(e^{-|u\sigma_t|^\beta})) - (e^{-|v\sigma_t|^\beta} - \mathbb{E}(e^{-|v\sigma_t|^\beta})).$$

Then, using successful conditioning, Holder's inequality and Lemma 3.102 in Jacod and Shiryaev (2003), together with the boundedness of z_t and assumption C$'$, we get

$$\begin{aligned}
\mathbb{E}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T z_t dt\right)^2 &= \frac{1}{T}\int_0^T\int_0^T \mathbb{E}\left(z_t z_s\right)dsdt \leq C|u^\beta - v^\beta|\frac{1}{T}\int_0^T\int_0^T \mathbb{E}\left(|\sigma_{s\wedge t}|^\beta \mathbb{E}(z_{s\vee t}|\mathscr{F}_{s\wedge t})\right)dsdt \\
&\leq C|u^\beta - v^\beta|^{1+3/2\iota}\frac{1}{T}\int_0^T\int_0^T\left(\alpha_{|t-s|}^{\text{mix}}\right)^{1/3-\iota}dtds \\
&\leq C|u^\beta - v^\beta|^{1+3/2\iota}\int_0^\infty\left(\alpha_s^{\text{mix}}\right)^{1-\iota}ds \leq C|u^\beta - v^\beta|^{1+3/2\iota}.
\end{aligned}$$

Using Theorem 12.3 of Billingsley (1968), the above bound implies the tightness of the sequence $\frac{1}{\sqrt{T}}\int_0^T(e^{-|u\sigma_t|^\beta} - \mathbb{E}(e^{-|u\sigma_t|^\beta}))dt$, and from here we have its convergence for the local uniform topology.

Turning now to $\frac{1}{\sqrt{\Delta_n}}\sum_{i=1}^{[T/\Delta_n]}\xi_{i,u}^{(1)}$, from the proof of Theorem 2, we have the finite-dimensional convergence to a Gaussian process with variance-covariance given by (25). Therefore, we only need to establish the tightness of the sequence. For this we use the proof of Theorem 2 to get

$\mathbb{E}\left(\frac{1}{\sqrt{\Delta_n}}\sum_{i=1}^{[T/\Delta_n]}(\xi_{i,u}^{(1)} - \xi_{i,v}^{(1)})\right)^2 \leq C|u-v|^p$ for some constant C and $1 < p < \beta$. From here using Theorem 12.3 of Billingsley (1968), we get the tightness of $\frac{1}{\sqrt{\Delta_n}}\sum_{i=1}^{[T/\Delta_n]}\xi_{i,u}^{(1)}$.

Similarly, using the proof of Theorem 2, we can show

$$\Delta_n^{-(3-2/\beta)}\mathbb{E}\left(\sum_{i=1}^{[T/\Delta_n]}\left(\xi_{i,u}^{(3)}(2) - \xi_{i,v}^{(3)}(2)\right)\right)^2 \leq C(|u-v|^2 + |u^{1/p+1} - v^{1/p+1}|^{2p}), \tag{51}$$

for $1/2 < p < \beta/2$ and ι arbitrary small. This establishes tightness for $\Delta_n^{-(3/2-1/\beta)}\sum_{i=1}^{[T/\Delta_n]}\xi_{i,u}^{(3)}(2)$ and together with the finite-dimensional result in the proof of Theorem 2 for it, we have convergence for the uniform topology. When assumption B$'$ holds, we can show in the same way

$$\begin{aligned}\Delta_n^{-2}\mathbb{E}\left(\sum_{i=1}^{[T/\Delta_n]}\left(\xi_{i,u}^{(2)}(1) - \xi_{i,v}^{(2)}(1)\right)\right)^2 &\leq C\left(K_1(\sigma_{(i-1)\Delta_n-}, u) - K_1(\sigma_{(i-1)\Delta_n-}, v)\right)^2 \\ &\leq C\left((u^\beta - v^\beta) + (u^{2\beta} - v^{2\beta})\right)^2,\end{aligned} \tag{52}$$

for $1/2 < p < \beta/2$ and ι arbitrary small. This together with the finite-dimensional result in the proof of Theorem 2 for this term implies that $\sum_{i=1}^{[T/\Delta_n]}\Delta_n^{-1}\xi_{i,u}^{(2)}(1)$ converges on the space of continuous functions equipped with the local uniform topology.

Next, using the proof of Theorem 2, it is easy to show that for any $\overline{u} > 0$ and when only assumption B holds, we have

$$\lim_{\Delta_n\downarrow 0, T\uparrow\infty}\mathbb{P}\left(\sup_{0\leq u\leq\overline{u}}\left|\sum_{i=1}^{[T/\Delta_n]}(\sqrt{T\Delta_n})^{-1}\xi_{i,u}^{(2)}\right| > \epsilon_n\right) = 0, \quad \forall\epsilon_n\uparrow\infty. \tag{53}$$

The same holds when in the above we replace $(\sqrt{T\Delta_n})^{-1}\xi_{i,u}^{(2)}$ with either of the following terms: $(\sqrt{T}\Delta_n^{\beta/2})^{-1}\xi_{i,u}^{(2)}(2)$ and $(\sqrt{T}\Delta_n)^{-1}\xi_{i,u}^{(2)}(3)$ (when the stronger assumption B$'$ holds), $(\sqrt{T}|\log(\Delta_n)|\Delta_n^{1-\beta'/\beta})^{-1}\xi_{i,u}^{(3)}(1)$, $(\sqrt{T}\Delta_n^{3/2-1/\beta})^{-1}\xi_{i,u}^{(3)}(3)$, $(\sqrt{T}\Delta_n^{3/2-1/\beta})^{-1}\xi_{i,u}^{(3)}(4)$, $(\sqrt{T\Delta_n})^{-1}\xi_{i,u}^{(5)}$ and $(\sqrt{T}\Delta_n^{2-2/\beta})^{-1}\xi_{i,u}^{(3)}(6)$. This implies that those terms are uniformly in u bounded in probability.

Finally, we are left with the term involving $\xi_{i,u}^{(5)}$ under the stronger assumption B$'$. We can argue for each of its subcomponents, according to the decomposition given in the proof of Theorem 2, using either the approach in (52) or the one in (53). □

7.5 Proof of Theorem 4

To simplify notation, as in the proof of Theorem 2, we will assume that A is such that $Z_\beta = 1$. If we denote for $k \geq 0$

$$Z_k(\beta, u) = \frac{1}{T}\sum_{t=k+1}^{T}\int_{t-1}^{t}\left(e^{-|u\sigma_s|^\beta} - \mathbb{E}(e^{-|u\sigma_s|^\beta})\right)ds\int_{t-k-1}^{t-k}\left(e^{-|u\sigma_s|^\beta} - \mathbb{E}(e^{-|u\sigma_s|^\beta})\right)ds, \tag{54}$$

then under our assumptions, by standard arguments, see e.g., Proposition 1 in Andrews (1991), we have

$$Z_0(\beta,u)+2\sum_{i=1}^{L_T}\omega(i,L_T)Z_i(\beta,u)\ \stackrel{\mathbb{P}}{\longrightarrow}\ V_\beta(u). \tag{55}$$

Therefore, we are left in showing that the error in estimating the integrals $\int_{t-1}^t e^{-|u\sigma_s|^\beta}ds$ for $t=1,...,T$ does not have any asymptotic effect. We note that for arbitrary $1\le k\le T$ we have: $\Delta_n\sum_{i=[(k-1)/\Delta_n]+1}^{[k/\Delta_n]}\cos(u\Delta_n^{-1/\beta}\Delta_i^nX)\le 1$ and $\int_{k-1}^k e^{-|u\sigma_s|^\beta}ds\le 1$. Further, using the stationarity of the process σ_t and the bounds on the moments of the terms $\xi_{i,u}^{(j)}$ derived in the proof of Theorem 2, we have for every t

$$\begin{aligned}\mathbb{E}\left|\sum_{i=[(t-1)/\Delta_n]+1}^{[t/\Delta_n]}\Delta_n\cos\left(u\Delta_n^{-1/\beta}\Delta_i^nX\right)-\int_{t-1}^t e^{-|u\sigma_s|^\beta}ds\right|\\ \le C\left(|\log\Delta_n|\Delta_n^{1-\beta'/\beta}\vee\Delta_n^{(2-2/\beta)\wedge 1/2}\right).\end{aligned} \tag{56}$$

From here, the result in (26) follows immediately. For (27), the proof follows by taking into account also (55) as well as the relative speed condition between L_T and Δ_n in the theorem. □

7.6 Proof of Theorem 5

Part a. Given Theorem 2, we need to prove that the difference $\frac{1}{\sqrt{T}}\left(V_T(X,\Delta_n,\widehat{\beta},u)-V_T(X,\Delta_n,\beta,u)\right)$ is asymptotically negligible. We have

$$\begin{aligned}\left|\frac{1}{\sqrt{T}}\left(V_T(X,\Delta_n,\widehat{\beta},u)-V_T(X,\Delta_n,\beta,u)\right)\right|\le&\frac{u}{(\beta^*)^2}\sqrt{T}|\log(\Delta_n)|\Delta_n^{-1/\beta^*+1/\beta}(\widehat{\beta}-\beta)\\ &\times\frac{\Delta_n}{T}\sum_{i=1}^{[T/\Delta_n]}|\sin\left(u\Delta_n^{-1/\beta^*}\Delta_i^nX\right)\Delta_n^{-1/\beta}\Delta_i^nX|,\end{aligned}$$

where β^* is between β and $\widehat{\beta}$. Then, using the integrability of the absolute values of the increments of X and also the fact that $\beta>1$, and upon applying Markov's inequality, we get

$$\mathbb{P}\left(\frac{\Delta_n}{T}\sum_{i=1}^{[T/\Delta_n]}|\sin\left(u\Delta_n^{-1/\beta^*}\Delta_i^nX\right)\Delta_n^{-1/\beta}\Delta_i^nX|>\log(\Delta_n)\right)\to 0.$$

Since $\widehat{\beta}-\beta=o_p(1)$, we have $\mathbb{P}\left(\left|\frac{1}{\widehat{\beta}}-\frac{1}{\beta}\right|>\alpha/2\right)\to 0$. Taking into account the assumed rate of convergence of $\widehat{\beta}$ the result follows.

Part b. In the case when $\widehat{\beta}$ uses an initial part of the sample (with fixed span) that is used in the construction of $V_T(X,\Delta_n,\beta,u)$, we can replace the latter with the same statistic but using only that part of the sample that is not used in the calculation of $\widehat{\beta}$. Since the time span of the sample used in the calculation of $\widehat{\beta}$ is fixed, this will have no asymptotic effect. Therefore, it is sufficient to consider only the case when $\widehat{\beta}$ uses only information before the beginning of the sample and we do so in the proof of the theorem.

First, since $\widehat{\beta} \xrightarrow{\mathbb{P}} \beta$, it is no limitation to restrict attention on the set for which $|1/\widehat{\beta} - 1/\beta| < \epsilon/2$ for some $\epsilon > 0$ such that $\epsilon < 1/(\beta \wedge \beta^{''}) - 1/2$ and $\epsilon < (1-\beta'/\beta) \wedge (2-2/\beta) - \alpha$. Then, using the proof of Theorem 2 and notation of that proof, we can write for arbitrary $\iota > 0$

$$\frac{1}{\sqrt{T}} V_T(X, \Delta_n, \beta, u) - \frac{1}{\sqrt{T}} \sum_{i=1}^{[T/\Delta_n]} \cos\left(u\Delta_n^{-1/\beta}\sigma_{(i-1)\Delta_n -}\Delta_i^n L\right) = o_p\left(\sqrt{T}\Delta_n^{(1-\beta'/\beta-\iota)\wedge(2-2/\beta-\iota)\wedge 1/2}\right).$$

Similar, using successive conditioning on the set of data used in the estimation of $\widehat{\beta}$, we can write

$$\frac{1}{\sqrt{T}} V_T(X, \Delta_n, \widehat{\beta}, u) - \frac{1}{\sqrt{T}} \sum_{i=1}^{[T/\Delta_n]} \cos\left(u\Delta_n^{-1/\widehat{\beta}}\sigma_{(i-1)\Delta_n -}\Delta_i^n L\right) = o_p\left(\sqrt{T}\Delta_n^{(1-\beta'/\beta-\iota-\epsilon)\wedge(2-2/\beta-\iota-\epsilon)\wedge 1/2}\right).$$

Thus, we need to prove asymptotic negligibility of

$$\begin{aligned}
&\frac{\Delta_n}{\sqrt{T}} \sum_{i=1}^{[T/\Delta_n]} \left(\cos\left(u\Delta_n^{-1/\widehat{\beta}}\sigma_{(i-1)\Delta_n -}\Delta_i^n L\right) - \cos\left(u\Delta_n^{-1/\beta}\sigma_{(i-1)\Delta_n -}\Delta_i^n L\right)\right) \\
&\quad = -\Delta_n^{1/\beta - 1/\beta^*} \frac{u\Delta_n \log(\Delta_n)}{\sqrt{T}(\beta^*)^2} \left(\widehat{\beta} - \beta\right) \sum_{i=1}^{[T/\Delta_n]} \sin\left(u\Delta_n^{-1/\beta^*}\sigma_{(i-1)\Delta_n -}\Delta_i^n L\right) \sigma_{(i-1)\Delta_n -}\Delta_n^{-1/\beta}\Delta_i^n L,
\end{aligned}$$

where we used a first-order Taylor expansion around the true value β and we further denoted with β^* some value between $\widehat{\beta}$ and β. As in the proof of Theorem 2, in what follows without loss of generality we will set A such that $Z_\beta = 1$. The proof consists of the following steps.

Step 1. We show $\frac{\Delta_n}{T}\sum_{i=1}^{[T/\Delta_n]} \sin\left(u\Delta_n^{-1/\beta}\sigma_{(i-1)\Delta_n -}\Delta_i^n L\right) u\sigma_{(i-1)\Delta_n -}\Delta_n^{-1/\beta}\Delta_i^n L \xrightarrow{\mathbb{P}} \mathbb{E}(G_\beta(u\sigma_t))$.

First upon differentiating in u both sides of the identity

$$\mathbb{E}(\cos(uL)) = e^{-|u|^\beta},$$

and using the self-similarity of the stable process, we have

$$\mathbb{E}_{i-1}^n\left(\sin\left(u\Delta_n^{-1/\beta}\sigma_{(i-1)\Delta_n -}\Delta_i^n L\right) u\Delta_n^{-1/\beta}\sigma_{(i-1)\Delta_n -}\Delta_i^n L\right) = G_\beta(u\sigma_{(i-1)\Delta_n -}).$$

From here using the fact that the function $G_\beta(x)$ is differentiable (in x), assumption B, the ergodicity of σ_t combined with a law of large numbers, we get

$$\frac{\Delta_n}{T} \sum_{i=1}^{[T/\Delta_n]} \mathbb{E}_{i-1}^n\left(\sin\left(u\Delta_n^{-1/\beta}\sigma_{(i-1)\Delta_n -}\Delta_i^n L\right) u\Delta_n^{-1/\beta}\sigma_{(i-1)\Delta_n -}\Delta_i^n L\right) \xrightarrow{\mathbb{P}} \mathbb{E}(G_\beta(u\sigma_t)).$$

The result then follows using Theorem VIII.2.29 of Jacod and Shiryaev (2003) and the fact that for some $1 < p < \beta$ we have

$$\mathbb{E}\left|\sin\left(u\Delta_n^{-1/\beta}\sigma_{(i-1)\Delta_n -}\Delta_i^n L\right)\sigma_{(i-1)\Delta_n -}\Delta_n^{-1/\beta}\Delta_i^n L\right|^p \leq C.$$

Step 2. We have

$$\frac{\Delta_n}{T}\sum_{i=1}^{[T/\Delta_n]}\Bigg(\sin\left(u\Delta_n^{-1/\beta^*}\sigma_{(i-1)\Delta_n -}\Delta_i^n L\right) - \sin\left(u\Delta_n^{-1/\beta}\sigma_{(i-1)\Delta_n -}\Delta_i^n L\right)\Bigg)\sigma_{(i-1)\Delta_n -}\Delta_n^{-1/\beta}\Delta_i^n L \xrightarrow{\mathbb{P}} 0.$$

First, we can write for some $1 < p < \beta$

$$\begin{aligned}|\sin\left(u\Delta_n^{-1/\beta^*}\sigma_{(i-1)\Delta_n -}\Delta_i^n L\right) - \sin\left(u\Delta_n^{-1/\beta}\sigma_{(i-1)\Delta_n -}\Delta_i^n L\right)| \\ \leq 2\sin\left(0.5u(\Delta_n^{-1/\beta^*} - \Delta_n^{-1/\beta})\sigma_{(i-1)\Delta_n -}\Delta_i^n L\right) \\ \leq C|\Delta_n^{-1/\beta^*+1/\beta} - 1|^{p-1}|\Delta_n^{-1/\beta}\Delta_i^n L|^{p-1},\end{aligned}$$

where we have made use of $|\cos(x)| \leq 1$ and the property $|\sin(x)| \leq |\sin(x)|^{p-1} \leq |x|^{p-1}$ since $0 < p-1 < 1$. Then we have

$$\begin{aligned}&\frac{\Delta_n}{T}\left|\sum_{i=1}^{[T/\Delta_n]}\Bigg(\sin\left(u\Delta_n^{-1/\beta^*}\sigma_{(i-1)\Delta_n -}\Delta_i^n L\right) - \sin\left(u\Delta_n^{-1/\beta}\sigma_{(i-1)\Delta_n -}\Delta_i^n L\right)\Bigg)\sigma_{(i-1)\Delta_n -}\Delta_n^{-1/\beta}\Delta_i^n L\right| \\ &\leq C|\Delta_n^{-1/\beta^*+1/\beta} - 1|^{p-1} \times \frac{\Delta_n}{T}\sum_{i=1}^{[T/\Delta_n]}|\sigma_{(i-1)\Delta_n -}\Delta_n^{-1/\beta}\Delta_i^n L|^p,\end{aligned}$$

from which the results follows by taking into account that the second term converges in L_1 and $\widehat{\beta} - \beta = o_p(\Delta_n^\alpha)$ for some $\alpha > 0$.

Step 3. The result of the theorem follows by taking into account that

$$\Delta_n^{1/\beta - 1/\beta^*} = 1 + (\beta^* - \beta)\frac{\Delta_n^{1/\beta - 1/\beta^{**}}}{(\beta^{**})^2}\log(\Delta_n),$$

where β^{**} is between β^* and β, and the fact that $\widehat{\beta} - \beta = o_p(\Delta_n^\alpha)$ for some $\alpha > 0$.

Part c. Given the result of Step 1 above, the only thing that remains to be proved is

$$\frac{\Delta_n}{T}\sum_{i=1}^{[T/\Delta_n]}\left(u\Delta_n^{-1/\beta}\Delta_i^n X\sin(u\Delta_n^{-1/\beta}\Delta_i^n X) - u\sigma_{(i-1)\Delta_n -}\Delta_n^{-1/\beta}\Delta_i^n L\sin(u\sigma_{(i-1)\Delta_n -}\Delta_n^{-1/\beta}\Delta_i^n L)\right) \xrightarrow{\mathbb{P}} 0.$$

For this we need only use the following algebraic inequality for $\forall x, y \in \mathbb{R}$

$$|x\sin(x) - y\sin(y)| \leq |x - y| + |y\sin((x-y)/2)|,$$

with $x = \Delta_n^{-1/\beta}\Delta_i^n X$ and $y = \Delta_n^{-1/\beta}\sigma_{(i-1)\Delta_n-}\Delta_i^n L$, Holder inequality, the fact that $\mathbb{E}|L|^p < \infty$ for $p < \beta$, and the following basic inequalities

$$\Delta_n^{-1/\beta}\mathbb{E}\left|\int_{(i-1)\Delta_n}^{i\Delta_n} a_s ds + \int_{(i-1)\Delta_n}^{i\Delta_n}\int_{\mathbb{R}} \sigma_{s-}x\mu_2(ds,dx) + \int_{(i-1)\Delta_n}^{i\Delta_n}\int_{\mathbb{R}} \sigma_{s-}x\mu_3(ds,dx)\right| \leq C\Delta_n^{1-1/\beta},$$

$$\Delta_n^{-1/\beta}\mathbb{E}\left|\int_{(i-1)\Delta_n}^{i\Delta_n} (\sigma_{s-} - \sigma_{(i-1)\Delta_n-})x\tilde{\mu}_1(ds,dx)\right| \leq C\Delta_n^{1+\beta/2-1/\beta-\epsilon},$$

where $\epsilon > 0$ is arbitrary small. □

References

Ait-Sahalia, Y. (1996). Nonparametric Pricing of Interest Rate Derivatives Securities. *Econeomtrica 64*, 527–560.

Ait-Sahalia, Y. and J. Jacod (2009). Estimating the Degree of Activity of Jumps in High Frequency Financial Data. *Annals of Statistics 37*, 2202–2244.

Ait-Sahalia, Y. and J. Jacod (2010). Is Brownian Motion Necessary to Model High Frequency Data? *Annals of Statistics, forthcoming*.

Andersen, T., T. Bollerslev, F. Diebold, and P. Labys (2001). The Distribution of Realized Exchange Rate Volatility. *Journal of the American Statistical Association 96*, 42–55.

Andersen, T., T. Bollerslev, F. Diebold, and P. Labys (2003). Modeling and Forecasting Realized Volatility. *Econometrica 71*, 579–625.

Andrews, D. (1991). Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econeomtrica 59*, 817–858.

Barndorff-Nielsen, O., S. Graversen, J. Jacod, M. Podolskij, and N. Shephard (2005). A Central Limit Theorem for Realised Power and Bipower Variations of Continuous Semimartingales. In Y. Kabanov and R. Lipster (Eds.), *From Stochastic Analysis to Mathematical Finance, Festschrift for Albert Shiryaev.* Springer.

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

Barndorff-Nielsen, O. and N. Shephard (2002). Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models. *Journal of the Royal Statistical Society Series B, 64*, 253280.

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

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

Bollerslev, T. and V. Todorov (2009). Tails, Fears and Risk Premia. Working paper, Northwestern University and Duke University.

Boyd, S. P. (2004). *Convex Optimization.* New York, NY: Cambridge University Press.

Breiman, L. (1968). *Probability* (1st ed.). Reading Massachussets: Addison Wesley.

Brockwell, P. (2001). Continuous-Time ARMA Processes. In D. Shanbhag and C. Rao (Eds.), *Handbook of Statistics*, Volume 19. North-Holland.

Carrasco, M., M. Chernov, J. Florens, and E. Ghysels (2007). Efficient Estimation of General Dynamic Models with a Continuum of Moment Conditions. *Journal of Econometrics 140*, 529–573.

Comte, F. and V. Genon-Catalot (2006). Penalized Projection Estimator for Volatility Density. *Scandinavian Journal of Statistics 33*, 875–893.

Comte, F. and E. Renault (1998). Long Memory in Continuous-Time Stochastic Volatility Models. *Mathematical Finance 8*, 291–323.

Duffie, D., J. Pan, and K. Singleton (2000). Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica 68*, 1343–1376.

Feuerverger, A. and R. Mureika (1977). The Empirical Characteristic Function and its Application. *Annals of Statistics 5*, 88–97.

Foster, D. and D. Nelson (1996). Continuous Record Asymptotics for Rolling Sample Variance Estimators. *Econometrica 64*, 139–174.

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

Jacod, J. and V. Todorov (2010). Do Price and Volatility Jump Together? *Annals of Applied Probability, forthcoming*.

Malliavin, P. and M. Mancino (2009). A Fourier Transform Method for Nonparametric Estimation of Multivariate Volatility. *Annals of Statistics 37*, 1983–2010.

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

Mykland, P. and L. Zhang (2009). Inference for Continuous Semimartingales Observed at High Frequency: A General Approach. *Econometrica 77*, 1403–1455.

Prakasa Rao, B. (1988). Statistical Inference from Sampled Data for Stochastic Processes. In *Contemporary Mathematics*, Volume 80. Amer. Math. Soc., Providence.

Rosiński, J. (2007). Tempering Stable Processes. *Stochastic Processes and their Applications 117*, 677–707.

Rozovsky, L. (2009). Remarks on a Link between the Laplace Transform and Distribution Function of a Nonnegative Random Variable. *Statistics and Probability Letters 79*, 1501–1508.

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

Singleton, K. (2001). Estimation of Affine Asset Pricing Models using the Empirical Characteristic Function. *Journal of Econometrics 102*, 111–141.

Todorov, V. and G. Tauchen (2009a). Limit Theorems for Power Variations of Pure-Jump Processes with Application to Activity Estimation. Working paper.

Todorov, V. and G. Tauchen (2009b). Volatility Jumps. Working paper.

Todorov, V. and G. Tauchen (2010). Activity Signature Functions for High-Frequency Data Analysis. *Journal of Econometrics 154*, 125–138.

Van Es, B., P. Spreij, and H. Van Zanten (2003). Nonparametric Volatility Density Estimation. *Bernoulli 9*, 451–465.