The Realized Laplace Transform of Volatility *

Viktor Todorov † and George Tauchen ‡

September 15, 2011

Abstract

We introduce and derive the asymptotic behavior of a new measure constructed from high-frequency data which we call the Realized Laplace Transform of volatility. The statistic provides a nonparametric estimate for the empirical Laplace transform function of the latent stochastic volatility process over a given interval of time and is robust to presence of jumps in the price process. With a long span of data, i.e., under joint long-span and infill asymptotics, the statistic can be used to construct a nonparametric estimate of the volatility Laplace transform as well as of the integrated joint Laplace transform of volatility over different points of time. We derive feasible functional limit theorems for our statistic both under fixed span and infill asymptotics as well as under joint long span and infill asymptotics which allow to quantify the precision in estimation under both sampling schemes.

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

JEL classification: C51, C52, G12.

*We wish to thank the Editor (Harald Uhlig) and anonymous referees for their detailed comments that led to significant improvements of the paper. We also thank Torben Andersen, Tim Bollerslev, Ron Gallant, Jean Jacod, Nour Meddahi, Per Mykland, Andrew Patton, Mark Podolskij, Neil Shephard as well as many seminar participants. Todorov's work was partially supported by NSF grant SES-0957330.

†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

Time-varying volatility is a salient empirical feature of many economic and financial time series, and the importance of properly accounting for such dependencies in economic decision making is now widely recognized, see Engle (2004). The widespread use of continuous-time processes with stochastic volatility[1] in macroeconomics and finance also directly underscores this.

Inference for stochastic volatility models is complicated because the underlying volatility process is latent and not uniquely determined by observed variables.[2] The presence of the hidden volatility process presents statistical challenges far beyond those encountered in models with a fully observed state vector, for which there are a variety of tractable methods. The literature on statistical inference for stochastic volatility, either classical or Bayesian, is vast. As common throughout econometrics and statistics, most techniques involve integrating out the latent process(es) in some way or another.

The recent availability of high-frequency financial data provides an alternative. The leading example is the widely-studied realized variance, see e.g., Andersen et al. (2001, 2003) and Barndorff-Nielsen and Shephard (2002). The realized variance is the sum of squared returns over a given time period, usually a day, and it is a nonparametric measure of the unobserved quadratic variation over that period. Further, jump robust extensions of the measure (Barndorff-Nielsen and Shephard, 2004; Mancini, 2001) allow for nonparametric estimation of the integral of the spot variance over the time interval.

Some important issues regarding time aggregation and efficiency arise when using the realized variance and its jump robust extensions for making inference about the underlying volatility dynamics. Inferring distributional properties of the spot variance directly from those of integrated variance is difficult due to the time aggregation. The mapping between the probability distribution of the spot and integrated variances is not one-to-one in general. Also we typically have far more analytical tractability for the spot variance process rather than the integrated. For example, in the widely popular affine jump-diffusion class of Duffie et al. (2000), the conditional characteristic functions, and thereby Laplace transforms, of spot variables are known in closed form and easily computable.

[1] In what follows we adopt general financial econometrics usage of the term "volatility" generically, in reference to either scale or variance with the meaning evident from the context. Whenever the distinction is relevant, however, we use "spot volatility" to mean the local, or instantaneous scale, and "spot variance" for the local variance. Exact definitions are in the theoretical analysis below. Note that spot volatility always pertains to the diffusive component, apart from jumps, which are typically modeled separately. Finally, we refer to integrated over time spot volatility (or variance) simply as integrated volatility (or variance).

[2] This includes the prices of derivative contracts as the presence of a rather non-trivial risk premia in the latter makes the link with the unobserved volatility indirect.

In this paper we propose another way of aggregating the high-frequency returns data into a measure we call Realized Laplace Transform of volatility, hereafter abbreviated as RLT, which overcomes the above difficulties. A key distinction is that the realized variance (or its jump-robust extensions) is a mapping from the data to a random variable, while the RLT is a mapping from the data to a random function. The function estimates the empirical Laplace transform of the spot variance over an interval of time[3], and it preserves information about the characteristics of volatility (when the latter is a stationary process). The RLT is easy to compute, as it is simply an average of cosine transforms of the appropriately rescaled high-frequency increments.

The RLT measure is built on the idea that over small intervals of time the leading component of the price increment is (conditionally) a zero-mean Gaussian random variable with variance equal to the spot variance at the beginning of the interval.[4] Then, the characteristic function of a zero-mean normal random variable is a Laplace transform of its volatility and aggregating over the high-frequency increments, by Law of Large Numbers, our RLT measure estimates the empirical Laplace transform of the volatility over the time interval. We prove feasible functional Central Limit Theorems for the RLT measure when considered as a function of its dummy variable under both fixed and long span sampling scheme.

Our asymptotic analysis forms the basis for applying the measure in many parametric and non-parametric estimation contexts. Importantly, the measure has robustness with respect to jumps slightly better than the existing jump-robust realized measures. This robustness is achieved automatically without any need for explicit truncation, and hence the nontrivial issue of choosing tuning parameters is avoided.[5]

The empirical usefulness of the RLT becomes evident in a situation where a long span of data is available and stationarity-type conditions are reasonable: standard assumptions for economic applications, e.g., Hansen and Scheinkman (1995); Barndorff-Nielsen and Shephard (2002); Andersen et al. (2003) among many others. Then we are able to estimate the unconditional Laplace transform of the spot variance. As seen in our empirical illustration below, the approach provides evidence on the statistical significance of the error in treating the daily integrated variance as the spot variance, as is sometimes done. Also, we can discriminate across broad classes of volatility models and assess the magnitude of the distortion to the distribution of volatility induced by the

[3]The empirical Laplace transform of a continuous-time process X_t over an interval $[0, T]$ is the Laplace transform of X_t with respect to the empirical measure, i.e., it is $\frac{1}{T}\int_0^T e^{-uX_s}ds$ for $u \in \mathbb{R}_+$.

[4]The local Gaussianity of high-frequency returns has been used, either implicitly or explicitly, in constructing general volatility estimators in diffusion settings by Barndorff-Nielsen et al. (2006) and Mykland and Zhang (2009).

[5]The multipower variation measures of the integrated variance of Barndorff-Nielsen and Shephard (2004) similarly do not need a choice of a tuning parameter.

temporal aggregation associated with the integrated variance.

Much more generally, by considering products of the RLT measure over different time intervals, we can estimate nonparametrically the integrated joint Laplace transform of volatility over the time intervals. Matching moments of the latter with that implied by a model provides for efficient, robust, and often analytically convenient way of model determination and estimation of the volatility dynamics. This latter effort is far beyond the scope of this paper and is undertaken in a follow up paper (Todorov et al., 2011) that applies the limit theory developed here.

Finally, the analysis in this paper is for the case when the observed process is a jump-diffusion, which is the most typical in economic applications, but it can be easily extended for studying the stochastic volatility in a pure-jump setting. This "adaptivity" is another important advantage of the proposed measure.[6]

The paper is organized as follows. Section 2 introduces our setup and assumptions. In Section 3 we define formally the RLT measure and derive its asymptotic behavior. Section 4 presents a Monte Carlo study of our statistic and Section 5 contains the results from an empirical application. Section 6 concludes. Proofs are in Section 7 and a supplementary Appendix, which also contains details regarding all computations within the paper as well as some extensions of the results in the paper.

2 Setting and Assumptions

We start by introducing our 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 the following dynamics:

$$dX_t = \alpha_t dt + \sigma_t dW_t + \int_{\mathbb{R}} \delta(t-, x)\mu(dt, dx), \tag{1}$$

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) $dt \otimes \nu(dx)$; $\delta(t, x) : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ is càdlàg in t.[7]

Our goal in the paper will be to uncover the stochastic volatility σ_t, and its distribution and dynamics in particular, from observing only X while assuming as little as possible about the rest of the components of X and the volatility itself.

[6] On the other hand, in pure-jump setting the realized variance will measure the sum of the squared jumps which deviates from integrated variance (formally the integrated jump compensator) by a (local) martingale.

[7] In financial applications, the observed price differs from the model in (1) by the so-called market microstructure noise which can have nontrivial impact for very high frequencies. We will adopt a conservative approach of using coarser frequencies at which the impact of the noise is negligible, leaving an extension of the results to the case when noise is present for future research.

The first two components in (1) have continuous paths while the third one captures the discontinuous moves in X, i.e., jumps. Our first assumption restricts the behavior of the latter.

Assumption A. The Lévy measure of μ satisfies: $\mathbb{E}\left(\int_0^t \int_{\mathbb{R}}(|\delta(s,x)|^p \vee |\delta(s,x)|)ds\nu(dx)\right) < \infty$, for every $t > 0$ and every $p \in (\beta, 1)$, where $0 \leq \beta < 1$ is some constant.

Apart from the minor integrability condition, i.e., the first moment of the jump process exists, assumption A restricts the "activity" of the jump component of X. The "activity" of the jumps determines the "vibrancy" of their trajectory. We restrict $\beta < 1$, i.e., the jump component is of finite variation meaning that its trajectory is of finite length (and this is why we do not need a martingale measure to define it).[8] In most parametric continuous-time models used to date, e.g., the affine jump-diffusion models, the jump process is a compound Poisson process and assumption A is trivially satisfied in this case with $\beta = 0$.

Our next assumption imposes minimal integrability conditions on α_t and σ_t and further limits their variation 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. Assume that σ_t is an Itô semimartingale given by

$$\sigma_t = \sigma_0 + \int_0^t \tilde{\alpha}_s ds + \int_0^t v_s dW_s + \int_0^t v'_s dW'_s + \int_0^t \int_{\mathbb{R}} \delta'(s-,x)\tilde{\mu}'(ds,dx), \tag{2}$$

where W' is a Brownian motion independent from W; μ is a homogenous Poisson measure, with Lévy measure $dt \otimes \nu'(dx)$, having arbitrary dependence with μ and $\delta'(t,x) : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ is càdlàg in t. We have for every t and s:

$$\begin{cases} \mathbb{E}\left(|\alpha_t|^2 + |\tilde{\alpha}_t|^2 + |\sigma_t|^2 + |v_t|^2 + |v'_t|^2 + \int_{\mathbb{R}} |\delta'(t,x)|^2 \underline{\nu}(dx)\right) < C, \\ \mathbb{E}\left(|\alpha_t - \alpha_s|^2 + |v_t - v_s|^2 + |v'_t - v'_s|^2 + \int_{\mathbb{R}}(\delta'(t,x) - \delta'(s,x))^2 \underline{\nu}(dx)\right) < C|t-s|, \end{cases} \tag{3}$$

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

Assumption B is a very general assumption, which is satisfied by the multifactor stochastic volatility models that are widely used in financial econometrics, e.g., the popular affine jump-diffusion models. It allows for a completely arbitrary dependence between the increments in σ_t and X, i.e., so-called "leverage" effect (by linking either jumps or Brownian motions) is allowed.[9]

Finally, for some of our results we will make use of long-span asymptotics, for the process σ_t^2 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.

[8] This restriction is not necessary if one is interested only in convergence in probability results (only $\beta < 2$ is needed for this and the highest value of β is 2). However, if one needs also the asymptotic distribution of the statistics that we introduce in the paper, then this assumption is probably unavoidable.

[9] We can further relax this assumption but with the cost of weakening slightly some of our asymptotic results. Given the wide class of stochastic volatility models that is covered by assumption B we do not do this here.

Assumption C. The volatility σ_t is a stationary and α-mixing process with $\alpha_t^{mix} = O(t^{-3-\iota})$ for arbitrary small $\iota > 0$ when $t \to \infty$, where

$$\alpha_t^{\text{mix}} = \sup_{A \in \mathscr{F}_0,\ B \in \mathscr{F}^t} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|, \quad \mathscr{F}_0 = \sigma(\sigma_s, W_s, s \le 0) \text{ and } \mathscr{F}^t = \sigma(\sigma_s, W_s - W_t, s \ge t). \tag{4}$$

3 Limit theory for the Realized Laplace Transform

We next define the Realized Laplace Transform and derive its 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. The Realized Laplace Transform measure is formally defined as

$$V_T(X, \Delta_n, u) = \sum_{i=1}^{[T/\Delta_n]} \Delta_n \cos(\sqrt{2u}\Delta_n^{-1/2}\Delta_i^n X), \quad \Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}. \tag{5}$$

$V_T(X, \Delta_n, u)$ is simply the real part of the empirical characteristic function of the (appropriately scaled) increments of the process. When $\Delta_n \to 0$, we will show that $V_T(X, \Delta_n, u)/T$ is an estimate of the empirical Laplace transform function of the latent volatility $\int_0^T e^{-u\sigma_t^2} ds/T$ (regardless of whether T is fixed or not). This result in turn can be used to construct a feasible nonparametric estimator for the Laplace transform of volatility and the integrated joint Laplace transform over different points in time as we show in this section.

3.1 Infill asymptotics

We start our asymptotic analysis with the case of T fixed and $\Delta_n \downarrow 0$. Before presenting the formal results, we explain the intuition behind our RLT measure. Under assumptions A and B, the error due to replacing $\Delta_i^n X$ with $\sigma_{(i-1)\Delta_n}\Delta_i^n W$ in $V_T(X, \Delta_n, u)$ is asymptotically negligible. Then, note that $W_{i\Delta_n} - W_{(i-1)\Delta_n} \stackrel{d}{=} \sqrt{\Delta_n} \times N(0,1)$, and since the characteristic function of a standard normal variable is $e^{-u^2/2}$, we have $\mathbb{E}\left(\cos(\sqrt{2u}\Delta_n^{-1/2}\sigma_{(i-1)\Delta_n}\Delta_i^n W)\middle|\mathcal{F}_{(i-1)\Delta_n}\right) = e^{-u\sigma^2_{(i-1)\Delta_n}}$. Therefore, by a Law of Large Numbers for the sample average of a heteroscedastic data series, we have that $V_T(X, \Delta_n, u)$ will converge in probability to $\Delta_n \sum_{i=1}^{[T/\Delta_n]} e^{-u\sigma^2_{(i-1)\Delta_n}}$ and the latter in turn converges to $\int_0^T e^{-u\sigma_s^2} ds$. [10] The following theorem makes this result formal and it further gives an associated (feasible) CLT result. In it we denote with $\mathcal{L} - s$ convergence stable in law, which means that the

[10] The discussion here reveals the intrinsic link between the small-scale behavior of the price process and our RLT measure. Thus for example if the diffusion component of X is absent, i.e., in a pure-jump setting, the scaling of the high-frequency increments in the construction of the RLT measure in (5) should be corrected to reflect the small-scale behavior of the leading jump component of X.

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 and let $\Delta_n \to 0$ and T be fixed. Then, we have*

$$\frac{1}{\sqrt{\Delta_n}}\left(V_T(X,\Delta_n,u) - \int_0^T e^{-u\sigma_s^2}ds\right) \overset{\mathcal{L}-s}{\longrightarrow} \Psi_T(u), \tag{6}$$

where the convergence is on the space $\mathcal{C}(\mathbb{R}_+)$ of continuous functions indexed by u equipped with the local uniform topology (i.e. uniformly over compact sets of $u \in \mathbb{R}_+$). The process $\Psi_T(u)$ is defined on an extension of the original probability space and is $\mathcal{F}$-conditionally Gaussian process with zero mean function and covariance function of $\int_0^T F(\sqrt{u}\sigma_s, \sqrt{v}\sigma_s)ds$ for every $u, v \in \mathbb{R}_+$ with $F(x,y) = \frac{e^{-(x+y)^2} - 2e^{-x^2-y^2} + e^{-(x-y)^2}}{2}$ for $x, y \in \mathbb{R}_+$.

A consistent estimator for the covariance function of $\Psi_T(u)$ is given by

$$\Delta_n \sum_{i=1}^{[T/\Delta_n]} \left(\cos(\sqrt{2u}\Delta_n^{-1/2}\Delta_i^n X) - V_T(X,\Delta_n,u)\right)\left(\cos(\sqrt{2v}\Delta_n^{-1/2}\Delta_i^n X) - V_T(X,\Delta_n,v)\right), \quad u, v > 0.$$

The result of the theorem for the case of fixed u and X a diffusion follows from the general theory of Barndorff-Nielsen et al. (2006). The robustness to jumps of the CLT in (6) requires only that jumps are of finite variation (assumption A) and further does not require any explicit truncation of the increments and the associated choice of a tuning parameter. The robustness to jumps is easiest to see in the case when the jumps are of finite activity. In this case there is a finite number of increments $\Delta_i^n X$ that are affected by the jumps. Due to the boundedness of the cosine function, their impact on the RLT measure is limited by $K\Delta_n$ (where K is the number of jumps on the interval which of course depends on the realization) and hence they do not affect the result in (6). By contrast, in the case of the realized variance (which is the sum of the squared high-frequency data), jumps affect not only the limiting distribution but also the limit itself.

An important consequence of the proof of the above theorem is the following result about the bias of the Realized Laplace Transform as a measure of the Laplace transform of volatility (assuming in addition to assumptions A and B that σ_t is stationary)

$$\mathbb{E}\left(V_T(X,\Delta_n,u)\right) = \mathbb{E}\left(\int_0^T e^{-u\sigma_s^2}ds\right) + O\left(\Delta_n^{1-\beta/2-\iota}\right), \quad \forall \iota > 0. \tag{7}$$

To compare, we note that for the realized variance, as a measure of the integrated variance, the bias is of order $O(\Delta_n)$ when there are no price jumps (and is due to the drift term).

3.2 Joint Infill and Long Span Asymptotics

We continue next with the asymptotic results for the case when both the time span increases and the length between observations decreases. The preceding analysis shows that for any t, $\widehat{Z}_t(u) = V_t(X, \Delta_n, u) - V_{t-1}(X, \Delta_n, u)$, constructed from the high-frequency data in the interval $[t, t+1]$, is an estimate for $Z_t(u) = \int_{t-1}^{t} e^{-u\sigma_s^2} ds$. Taking sample averages of products of $Z_t(u)$ over different time intervals, i.e.,

$$\widehat{\mu}_k(u, v) = \frac{1}{T} \sum_{t=k+1}^{T} \widehat{Z}_t(u) \widehat{Z}_{t-k}(v), \tag{8}$$

for k integer and $u, v \geq 0$, and applying a standard Law of Large Numbers we can estimate consistently $\mu_k(u, v) = \mathbb{E}\left(Z_t(u) Z_{t-k}(v)\right)$ for $T \uparrow \infty$ and $\Delta_n \downarrow 0$. The latter (by stationarity) is equal to $\mathbb{E}\left(\int_k^{k+1} \int_0^1 e^{-v\sigma_{s_1}^2 - u\sigma_{s_2}^2} ds_1 ds_2\right)$ which is just integrated joint Laplace transform of volatility over different points in time. For $u = 0$ or $v = 0$, it reduces to the Laplace transform of the marginal distribution of the volatility process.

To make use of the above result, however, we need to know the precision with which we can recover the function $\mu_k(u, v)$ from the data. The estimation involves discretization error (from estimating $Z_t(u)$ by $\widehat{Z}_t(u)$) in addition to the empirical process $\frac{1}{T} \sum_{t=k+1}^{T} Z_t(u) Z_{t-k}(v) - \mu_k(u, v)$. Can we gauge the precision of $\widehat{\mu}_k(u, v)$ by a feasible estimate of the magnitude of the latter error? The answer to this depends on how big is the discretization error relative to the empirical process. In the next theorem we quantify the magnitudes of the two errors and provide a feasible CLT for the latter. We need some more notation for the asymptotic variance of $\widehat{\mu}_k(u, v)$ and its feasible estimate before we can present the theorem. In particular, we set

$$V_k([u_1, v_1], [u_2, v_2]) = \sum_{l=-\infty}^{\infty} \mathbb{E}\left[\left(Z_t(u_1) Z_{t-k}(v_1) - \mu_k(u_1, v_1)\right)\left(Z_{t-l}(u_2) Z_{t-l-k}(v_2) - \mu_k(u_2, v_2)\right)\right], \tag{9}$$

$$\widehat{C}_l([u_1, v_1], [u_2, v_2]) = \frac{1}{T} \sum_{t=k+l+1}^{T} \left(\widehat{Z}_t(u_1) \widehat{Z}_{t-k}(v_1) - \widehat{\mu}(u_1, v_1)\right)\left(\widehat{Z}_{t-l}(u_2) \widehat{Z}_{t-l-k}(v_2) - \widehat{\mu}(u_2, v_2)\right), \tag{10}$$

$$\widehat{V}_k([u_1, v_1], [u_2, v_2]) = \widehat{C}_0([u_1, v_1], [u_2, v_2]) + \sum_{i=1}^{L_T} \omega(i, L_T)(\widehat{C}_i([u_1, v_1], [u_2, v_2]) + \widehat{C}_i([u_2, v_2], [u_1, v_1])), \tag{11}$$

for $u_i, v_i \geq 0$ for $i = 1, 2$ and some nonnegative function ω. We note that $V_k([u_1, v_1], [u_2, v_2])$ is well defined when assumption C holds, see Jacod and Shiryaev (2003), Theorem VIII.3.79.

Theorem 2 *(a) Suppose $T \to \infty$ and $\Delta_n \to 0$. For the process X under assumptions A, B and C, for arbitrary integer $k \geq 0$ we have $\widehat{\mu}_k(u, v) \stackrel{\mathbb{P}}{\longrightarrow} \mu_k(u, v)$, and further*

$$\sqrt{T}\left(\widehat{\mu}_k(u, v) - \mu_k(u, v)\right) = Y_T^{(1)}(u, v) + Y_T^{(2)}(u, v), \quad Y_T^{(1)}(u, v) \stackrel{\mathcal{L}}{\longrightarrow} \Psi'(u, v), \tag{12}$$

$$
\begin{aligned}
Y_T^{(2)}(u,v) = \ & \frac{T-k}{\sqrt{T}} \sum_{i=[(t-k-1)/\Delta_n]+1}^{[(t-k)/\Delta_n]} \mathbb{E}\left[Z_t(u)\Delta_n\left(\cos\left(\sqrt{2v}\Delta_n^{-1/2}\sigma_{(i-1)\Delta_n}\Delta_i^n W\right) - e^{-v\sigma^2_{(i-1)\Delta_n}}\right)\right] \\
& + 1_{\{k=0\}}\sqrt{T} \sum_{i=[(t-1)/\Delta_n]+1}^{[t/\Delta_n]} \mathbb{E}\left[Z_t(v)\Delta_n\left(\cos\left(\sqrt{2u}\Delta_n^{-1/2}\sigma_{(i-1)\Delta_n}\Delta_i^n W\right) - e^{-u\sigma^2_{(i-1)\Delta_n}}\right)\right] \\
& + O_p\left(\sqrt{T}\Delta_n^{1-\beta/2-\iota}\right), \quad \forall \iota > 0,
\end{aligned} \tag{13}
$$

where the above limit results are for the space $\mathcal{C}(\mathbb{R}_+ \times \mathbb{R}_+)$ of continuous functions indexed by u and v and equipped with the local uniform topology. $\Psi'(u,v)$ is a Gaussian process with zero mean function and covariance function of $V_k([u_1,v_1],[u_2,v_2])$ which is defined in (9).

(b) If further L_T is a deterministic sequence of integers satisfying $\frac{L_T}{\sqrt{T}} \to 0$ as $T \to \infty$ and $L_T\Delta_n^{1-\beta/2-\iota} \to 0$, we have

$$
\widehat{V}_k([u_1,v_1],[u_2,v_2]) \stackrel{\mathbb{P}}{\longrightarrow} V_k([u_1,v_1],[u_2,v_2]), \tag{14}
$$

where $\omega(i, L_T)$ is either a Bartlett or a Parzen kernel.[11]

The two components of the estimation error, $Y_T^{(1)}(u,v)$ and $Y_T^{(2)}(u,v)$, are respectively the empirical process, $\sqrt{T}\left(\frac{1}{T}\sum_{t=k+1}^{T} Z_t(u)Z_{t-k}(v) - \mu_k(u,v)\right)$, and the discretization error. Naturally, $Y_T^{(1)}(u,v)$ is sole function of the time span T and does not depend on Δ_n unlike $Y_T^{(2)}(u,v)$. The first two terms in $Y_T^{(2)}(u,v)$ are due to the dependence between σ_t^2 and W_t. In general, they are $O(\sqrt{T\Delta_n})$. They are exactly 0 when $v=0$ or when σ_t^2 and W_t are independent. Even more generally, however, when σ_t^2 is a multifactor model with factors following Lévy-driven SDE-s (typical modeling assumption), and if in addition the conditional Laplace transform of σ_t^2 is twice differentiable with bounded second derivative (which is the case for example for the affine jump-diffusions), then these terms in $Y_T^{(2)}(u,v)$ are of order only $O(\sqrt{T}\Delta_n)$. When this is the case, the relative speed condition needed for $Y_T^{(2)}(u,v)$ to be negligible is $\sqrt{T}\Delta_n^{1-\beta/2-\iota} \to 0$ and is determined by the error due to the presence of jumps. This condition becomes more stringent for higher levels of β (recall assumption A) as for them the jumps are "closer" to the Brownian increments, for our estimation purposes, and this induces a more significant error in their disentangling. In the typical case of finite activity jumps, e.g., compound Poisson process, the relative speed condition reduces to $\sqrt{T}\Delta_n \to 0$ which allows the span of the data to increase much faster than the mesh of the observation grid. Compared with the standard requirement $T\Delta_n \to 0$ found in the related problem

[11] We refer to Andrews (1991) and the many references therein for the alternative kernels used in the construction of so-called heteroskedastic autocorrelation (HAC) estimators.

of maximum-likelihood estimation of diffusion processes with discrete data, see e.g., Prakasa Rao (1988), our relative speed condition is much weaker.

We note that the importance of the different components of the discretization error changes from fixed to long span asymptotics. The martingale part is the leading component in the fixed-span asymptotics and it determines the limiting distribution of the RLT measure in (6).[12] On the other hand, for the long-span asymptotics, the bias term due to the presence of jumps in the price increments dominates the martingale component of the discretization error and determines the order of magnitude of $Y_T^{(2)}(u,v)$.

While in (13) we give the order of magnitude of the discretization error, for an empirical application where we use fixed T and Δ_n, it is important to have an idea of the actual size of the bias it creates, in particular the one due to jumps. For simplicity we do this for the case when $v = 0$ and when the process has i.i.d. increments with compound Poisson jumps with symmetric distribution of the jump size. The bias due to the discretization error in this case is given by

$$\left|\frac{\mathbb{E}\left(V_T(X,\Delta_n,u)\right)}{Te^{-u\sigma^2}} - 1\right| \leq \Bigg[\left|\cos(\alpha\sqrt{2u}\sqrt{\Delta_n}) - 1\right| + \lambda\Delta_n\left|\cos(\alpha\sqrt{2u}\sqrt{\Delta_n})e^{\psi\left(\sqrt{2u/\Delta_n}\right)} - 1\right| + \frac{\lambda^2\Delta_n^2}{2}\Bigg],$$

where λ is the intensity of the jumps, i.e., $\lambda = \int_{\mathbb{R}} \nu(dx)$ with $\delta(t,x) = x$ in (1) and ψ is the characteristic function of the jump size distribution $\nu(dx)/\lambda$. The bias derived above is very small. For example, for intensity of 1 jump per day with $\Delta_n = 1/400$, the bias is less than 0.25% of the estimated value. In the Monte Carlo section we will further investigate the finite sample bias and variance of our estimator $\widehat{\mu}_k(u,v)$ for the time span and frequencies of the typical financial data sets that are available and will confirm our asymptotic analysis here.

The function $\widehat{\mu}_k(u,v)$ can be further generalized to products of more RLT measures over different time intervals. These functions essentially "summarize" the information for the latent volatility dynamics in the data. Indeed, if we assume that volatility stays constant over the intervals $[t,t+1]$ and is further Markov of finite order, it is well known (see e.g., Proposition 4.2 in Carrasco et al. (2007) and the references therein) that we can achieve the efficiency of the maximum likelihood estimator by minimizing the distance between these functions and their model implied analogues.

Of course, the volatility can change over the time interval $[t,t+1]$ so that $\mu_k(u,v)$ is not exactly the joint Laplace transform of volatility over arbitrary points of time, but rather an integrated version of it. Intuitively, this is the price to pay for the fact that we need to "recover" the latent

[12] The leading martingale component of the dicretization error, $\widehat{Z}_t(u) - Z_t(u)$, is given by $\Delta_n \sum_{i=[(t-1)/\Delta_n]+1}^{[t/\Delta_n]}\left(\cos\left(\sqrt{2u}\Delta_n^{-1/2}\sigma_{(i-1)\Delta_n}\Delta_i^n W\right) - e^{-u\sigma^2_{(i-1)\Delta_n}}\right)$.

volatility from the high-frequency data.[13] The loss of information compared with the infeasible case where the joint Laplace transform of volatility (and not an integrated version of it) is observed, is for the very short term moves in volatility which are hardest to pin down from discrete price data.

We can compare the use of $\widehat{\mu}_k(u,v)$ with that of the realized variance (or its jump-robust extensions) for the purposes of estimating continuous-time volatility models. In the latter case the inference has been based on matching the first few moments of the realized variance as those are known analytically for a wide range models (another alternative is Gaussian QMLE). This, however, typically leads to a significant loss of information about the volatility dynamics. For example, the volatility persistence, in such estimation, is inferred from the autocorrelation of the realized variance. The latter however is an "aggregated" measure of how volatility shocks on "average" propagate in the future. By contrast, using our measure $\widehat{\mu}_k(u,v)$ over different regions of (u,v), one can identify the "impulse response" to volatility shocks in different volatility regimes. This is often achieved in an analytically convenient way, since for wide classes of models, e.g., the affine jump-diffusions, the conditional Laplace transform is known in closed-form. Moreover, nonlinear transformations of the realized variance, needed to capture better the information in it, typically lead to a more prominent role of the discretization error, which is reflected in the stronger relative speed condition $T\Delta_n \to 0$ needed for this error to be negligible.

Finally, given the above-developed limit theory for the Laplace transform, it is natural to inquire about inverting the transform to generate a nonparametric estimator of the probability density. The inversion problem is well known to be ill posed, so some form of numerical regularization will be required (Kryzhniy, 2010). Furthermore, there is no imperative reason to invert, since almost all models, e.g. affine jump-diffusion models for the term structure and derivatives pricing, imply convenient forms for the Laplace transform, not the density.

4 Monte Carlo Assessment

We now examine the precision of estimating $\mu_k(u,v)$ via the RLT. We use the following two-factor stochastic volatility model

$$\begin{aligned} &dX_t = \sqrt{V_{1t}+V_{2t}}dW_t + dL_{1t}, \quad dV_{1t} = 0.02(0.5 - V_{1t})dt + 0.07\sqrt{V_{1t}}dB_t, \quad W_t \perp\!\!\!\perp B_t, \\ &dV_{2t} = -0.5V_{2t}dt + dL_{2t}, \quad L_{1t} \perp\!\!\!\perp L_{2t}, \end{aligned} \tag{15}$$

L_{1t} is pure-jump with Lévy density $\nu(x) = \frac{e^{-x^2/0.4}}{\sqrt{0.4\pi}}$ and L_{2t} is pure-jump with Lévy density $\nu(x) =$

[13] Alternatively we could have defined our RLT measures $\widehat{Z}_t(u)$ over intervals that shrink asymptotically. However, in this case the relative importance of the discretization error will increase. Since, we quantify the precision of $\widehat{\mu}_k(u,v)$ only by the associated empirical process, ignoring the discretization error, we do not consider such an extension.

$4e^{-4x}1_{\{x>0\}}$. The model is fairly general and captures most stylized features of asset returns data documented in empirical asset pricing.

Table 1: Monte Carlo Results

$k=0$		$v=0.00$							
u	0.50	1.25	2.50	3.75					
$\mu_k(u,v)$	0.6207	0.3256	0.1282	0.0563					
median	0.6182	0.3234	0.1266	0.0561					
MAD	0.0080	0.0100	0.0070	0.0042					
$k=1$		$v=0.50$			$v=1.25$			$v=2.50$	
u	0.50	1.25	2.50	0.50	1.25	2.50	0.50	1.25	2.50
$\mu_k(u,v)$	0.3967	0.2161	0.0895	0.2158	0.1232	0.0543	0.0889	0.0539	0.0258
median	0.3935	0.2136	0.0879	0.2132	0.1214	0.0533	0.0874	0.0529	0.0253
MAD	0.0099	0.0087	0.0056	0.0087	0.0067	0.0041	0.0056	0.0041	0.0024
$k=10$		$v=0.50$			$v=1.25$			$v=2.50$	
u	0.50	1.25	2.50	0.50	1.25	2.50	0.50	1.25	2.50
$\mu_k(u,v)$	0.3895	0.2074	0.0833	0.2074	0.1125	0.0464	0.0832	0.0464	0.0198
median	0.3865	0.2050	0.0819	0.2050	0.1109	0.0455	0.0819	0.0455	0.0194
MAD	0.0098	0.0086	0.0054	0.0086	0.0064	0.0036	0.0053	0.0036	0.0019

Note: The median and the median absolute deviation (MAD) correspond to the estimator $\widehat{\mu}_k(u,v)$. 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.

We simulate from the above model a data set of $4,000$ "days" worth of 400 within-day price increments which is similar to the data set we are going to use in the empirical application. Table 1 summarizes the results from the Monte Carlo experiment. As seen from the table, the estimator is very accurate and the finite sample biases of $\widehat{\mu}_k(u,v)$ are several times smaller than their sampling variation which further confirms the rather small effect of the discretization error implied by our theoretical analysis in the previous section.

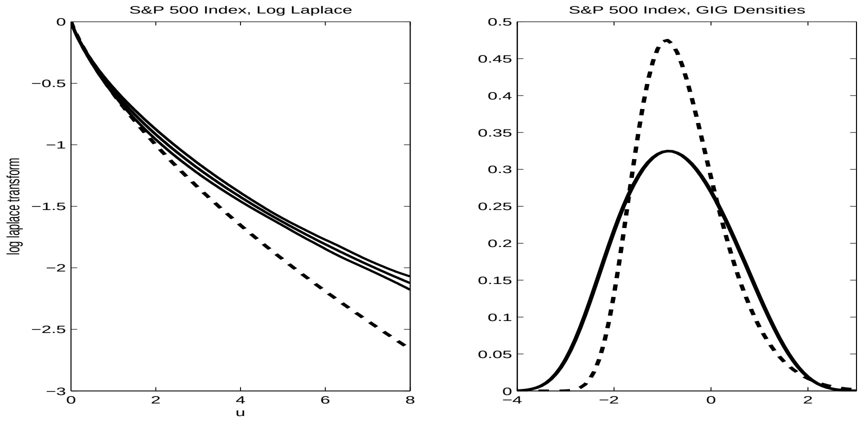
5 Empirical Application

The two questions addressed here are i) is there significant statistical error in treating the integrated variance as a spot variance, and ii) do we gain additional information from using our RLT measure?

The data are 1-minute observations on the S&P 500 futures index, January 1, 1990, to December 31, 2008. The strategy [14] is to juxtapose the estimate of the Laplace transform of the spot variance process, σ_t^2, as newly developed in this paper, to the transform of the widely studied daily integrated variance $\int_t^{t+1} \sigma_s^2 ds$. For ease of interpretation, we also juxtapose the implied probability densities.

The left panel of Figure 1 shows the estimate of the log-Laplace transform of the spot variance along with two-sigma confidence bands, obtained by using Theorem 2 with the function $\widehat{\mu}_0(u, 0)$. It also shows the empirical Laplace transform of the integrated variance, as estimated by the jump-robust truncated version of Mancini (2001), with the bipower variation (Barndorff-Nielsen and Shephard, 2004) used for the estimate of truncation — all details are in the supplementary appendix. There is a clear, statistically significant wedge between the two log-Laplace transforms. As a guide to interpreting the wedge, the right panel shows model-implied densities for logs of

Figure 1: Observed Log-Laplace Transforms and Implied Densities of the Log Variance



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 integrated variance, 1-minute S&P 500 index data, 1990–2008. Right: model-implied densities of the log of the log spot variance (solid) and the log of the realized variance (dashed), fitted to the Laplace transforms under the generalized inverse Gaussian specification.

the spot and the integrated variance under the Generalized-Inverse-Gaussian distribution. This three-parameter distribution nests many well-known positively supported distributions, and it is the marginal density for many important stochastic volatility models (Barndorff-Nielsen and Shephard, 2001). The parameter estimates were obtained by matching Laplace transforms at three widely dispersed points. The fitted Laplace transforms match observed with $R^2 \approx 1.00$ over the

[14]Extensive investigations indicated that microstructure noise is not a concern and the results below are robust across sampling frequencies. Further, the results are only marginally affected by the well-known deterministic within-day diurnal pattern in volatility. For a theoretical analysis when the latter is present, see the supplementary Appendix to the paper.

entire domain, and thereby the plotted densities are just alternative representations of the same information embedded in the Laplace transforms. On the other hand, as discussed in the supplementary appendix, the important special case of a gamma distribution is statistically rejected using criteria based on our limit theory, indicating that the theory is useful for discriminating across models. The Generalized-Inverse-Gaussian thereby appears to be the appropriate distribution of stochastic variance for these data. As to be expected, the density of the spot variance is more dispersed around the mode than is the density of the integrated variance. The integration used to accumulate the daily integrated variance smooths over sharp short-term movements, as would be induced by, say, volatility jumps. Our approach thereby provides empirical evidence on the magnitude to which the smoothing alters the distribution of the integrated relative to the spot variance, a difference to be kept in mind when modeling stochastic volatility.

6 Conclusions

In this paper we propose a new measure, we call Realized Laplace Transform of volatility, which estimates from high-frequency data over a given interval the empirical Laplace transform of the latent volatility process over that interval. We derive the asymptotic distribution of the statistic under settings of fixed and long span of data. Our asymptotic analysis and Monte Carlo work show the measure can be used to reliably estimate integrated joint Laplace transform of the volatility over different points in time. This provides an easy and efficient way to estimate and test performance of models with rich dynamics that are needed to capture the volatility risks evident in the data.

7 Proofs

In all the proofs C denotes a constant that does not depend on T and Δ_n, and further can change from line to line. We also use the short hand $\mathbb{E}_{i-1}^n$ for $\mathbb{E}\left(\cdot|\mathscr{F}_{(i-1)\Delta_n}\right)$. We start with some preliminary results that we use in the proofs of the theorems.

7.1 Preliminary Estimates

For every t and u we have $\widehat{Z}_t(u) - Z_t(u) = \sum_{i=[(t-1)/\Delta_n]+1}^{[t/\Delta_n]} \sum_{j=1}^{3} \xi_{i,u}^{(j)}$, with

$$\xi_{i,u}^{(1)} = \Delta_n \left[\cos\left(\sqrt{2u}\sigma_{(i-1)\Delta_n -}\Delta_n^{-1/2}\Delta_i^n W\right) - e^{-u\sigma^2_{(i-1)\Delta_n -}}\right], \quad \xi_{i,u}^{(2)} = \int_{(i-1)\Delta_n}^{i\Delta_n} \left(e^{-u\sigma^2_{(i-1)\Delta_n -}} - e^{-u\sigma_s^2}\right) ds,$$

$$\xi_{i,u}^{(3)} = \Delta_n \left(\cos\left(\sqrt{2u}\Delta_n^{-1/2}\Delta_i^n X\right) - \cos\left(\sqrt{2u}\sigma_{(i-1)\Delta_n -}\Delta_n^{-1/2}\Delta_i^n W\right)\right).$$

First, using $\mathbb{E}(e^{iuZ}) = e^{-u^2/2}$ for $u \in \mathbb{R}$ and $z \sim N(0,1)$, we have

$$\mathbb{E}_{i-1}^n\left(\xi_{i,u}^{(1)}\right) = 0, \quad \mathbb{E}_{i-1}^n\left(\xi_{i,u}^{(1)}\xi_{i,v}^{(1)}\right) = \Delta_n^2 F(\sqrt{u}\sigma_{(i-1)\Delta_n-}, \sqrt{v}\sigma_{(i-1)\Delta_n-}), \quad \mathbb{E}_{i-1}^n\left(\xi_{i,u}^{(1)}\right)^4 \leq C\Delta_n^4. \quad (16)$$

We move next to $\xi_{i,u}^{(2)}$. We decompose it using a first-order Taylor expansion as $\xi_{i,u}^{(2)} = \sum_{j=1}^3 \xi_{i,u}^{(2)}(j)$,

$$\xi_{i,u}^{(2)}(1) = K_1(\sigma_{(i-1)\Delta_n-}, u)\int_{(i-1)\Delta_n}^{i\Delta_n}\left(\sigma_{(i-1)\Delta_n-} - \widehat{\sigma}_s\right)ds, \quad \xi_{i,u}^{(2)}(2) = \int_{(i-1)\Delta_n}^{i\Delta_n}\left(e^{-u\widehat{\sigma}_s^2} - e^{-u\sigma_s^2}\right)ds,$$

$$\xi_{i,u}^{(2)}(3) = \int_{(i-1)\Delta_n}^{i\Delta_n}\left(K_1(\sigma_s^*, u) - K_1(\sigma_{(i-1)\Delta_n-}, u)\right)\left(\sigma_{(i-1)\Delta_n-} - \widehat{\sigma}_s\right)ds,$$

where $K_1(x,u) = -2uxe^{-ux^2}$, σ_s^* is a number between $\sigma_{(i-1)\Delta_n-}$ and $\widehat{\sigma}_s$, and

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

Then, using our integrability conditions in assumption B, successive conditioning and Itô isometry, as well as the the boundedness of the function $K_1(x,u)$, we have

$$\mathbb{E}_{i-1}^n\left(\xi_{i,u}^{(2)}(1)\right) = 0, \qquad \mathbb{E}\left|\xi_{i,u}^{(2)}(1)\right|^2 \leq C\Delta_n^3. \tag{17}$$

For $\xi_{i,u}^{(2)}(2)$, first-order Taylor expansion and the integrability conditions in assumption B imply

$$\mathbb{E}_{i-1}^n|\xi_{i,u}^{(2)}(2)| \leq Cu\mathbb{E}_{i-1}^n\left(\int_{(i-1)\Delta_n}^{i\Delta_n}\left|\widehat{\sigma}_s^2 - \sigma_s^2\right|ds\right) \implies \mathbb{E}|\xi_{i,u}^{(2)}(2)| \leq C\Delta_n^2. \tag{18}$$

Finally for $\xi_{i,u}^{(2)}(3)$, by using Cauchy-Schwarz inequality and Itô isometry, we can write

$$\mathbb{E}|\xi_{i,u}^{(2)}(3)| \leq C\Delta_n^{3/2}\sqrt{\mathbb{E}\left(\sup_{s\in[(i-1)\Delta_n, i\Delta_n]}(K_1(\sigma_s^*, u) - K_1(\sigma_{(i-1)\Delta_n-}, u))^2\right)}.$$

To continue further we make use of the bound $|K_1(x,u) - K_1(y,u)| \leq C|x-y|$ for $x, y \in \mathbb{R}$ and $u \geq 0$, where the constant C depends only on u. Plugging in the above inequality $x = \sigma_s^*$ and $y = \sigma_{(i-1)\Delta_n-}$ and using successive conditioning (first on the filtration $\mathscr{F}_{(i-1)\Delta_n}$) together with the Burkholder-Davis-Gundy inequality and the integrability conditions of assumption B, we get

$$\mathbb{E}|\xi_{i,u}^{(2)}(3)| \leq C\Delta_n^2. \tag{19}$$

Turning to $\xi_{i,u}^{(3)}$, we can decompose it as $\xi_{i,u}^{(3)} = \sum_{j=1}^5 \xi_{i,u}^{(3)}(j)$, where

$$\xi_{i,u}^{(3)}(1) = -2\Delta_n \sin\left(0.5\sqrt{2u}\Delta_n^{-1/2}\left(\Delta_i^n X + \int_{(i-1)\Delta_n}^{i\Delta_n} a_s ds + \int_{(i-1)\Delta_n}^{i\Delta_n}\sigma_s dW_s\right)\right)$$
$$\times \sin\left(0.5\sqrt{2u}\Delta_n^{-1/2}\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{\mathbb{R}}\delta(s-,x)\mu(ds,dx)\right),$$

$$\xi^{(3)}_{i,u}(2) = -\sqrt{2u}\Delta_n^{3/2}\sin\left(\sqrt{2}u\sigma_{(i-1)\Delta_n-}\Delta_n^{-1/2}\Delta_i^n W\right) a_{(i-1)\Delta_n},$$

$$\xi^{(3)}_{i,u}(3) = -u\cos(\tilde{x}_2)\left(\Delta_n a_{(i-1)\Delta_n} + \int_{(i-1)\Delta_n}^{i\Delta_n}(\sigma_s - \sigma_{(i-1)\Delta_n})dW_s\right)^2,$$

$$\xi^{(3)}_{i,u}(4) = -\sqrt{2u}\Delta_n^{1/2}\sin(\tilde{x}_1)\int_{(i-1)\Delta_n}^{i\Delta_n}(a_s - a_{(i-1)\Delta_n})ds,$$

$$\xi^{(3)}_{i,u}(5) = -\sqrt{2u}\Delta_n^{1/2}\sin\left(\sqrt{2u}\sigma_{(i-1)\Delta_n-}\Delta_n^{-1/2}\Delta_i^n W\right)\int_{(i-1)\Delta_n}^{i\Delta_n}(\sigma_s - \sigma_{(i-1)\Delta_n})dW_s,$$

where $\tilde{x}_1$ is between $\sqrt{2u}\Delta_n^{-1/2}\int_{(i-1)\Delta_n}^{i\Delta_n} a_s ds + \sqrt{2u}\Delta_n^{-1/2}\int_{(i-1)\Delta_n}^{i\Delta_n}\sigma_s dW_s$ and $\sqrt{2u}\Delta_n^{1/2}a_{(i-1)\Delta_n} + \sqrt{2u}\Delta_n^{-1/2}\int_{(i-1)\Delta_n}^{i\Delta_n}\sigma_s dW_s$ and $\tilde{x}_2$ is between the latter and $\sqrt{2u}\Delta_n^{-1/2}\sigma_{(i-1)\Delta_n-}\Delta_i^n W$.

Using the basic inequalities $|\sin(x)| \leq |x|$ and $|\sum_i |a_i||^p \leq \sum_i |a|_i^p$ for some $0 < p \leq 1$, we have

$$\mathbb{E}|\xi^{(3)}_{i,u}(1)| \leq C\Delta_n^{1-\beta/2-\iota/2}\mathbb{E}\left|\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{\mathbb{R}}\delta(s-,x)\mu(ds,dx)\right|^{\beta+\iota}$$
$$\leq C\Delta_n^{1-\beta/2-\iota/2}\mathbb{E}\int_{(i-1)\Delta_n}^{i\Delta_n}\int_{\mathbb{R}}|\delta(s-,x)|^{\beta+\iota}ds\nu(dx) \leq C\Delta_n^{2-\beta/2-\iota/2}, \quad \forall\iota \in (0, 1-\beta].$$

For $\xi^{(3)}_{i,u}(2)$, $\xi^{(3)}_{i,u}(3)$ and $\xi^{(3)}_{i,u}(4)$, using the boundedness of the functions $\sin(x)$ and $\cos(x)$, the symmetry of $\sin(x)$, the square integrability of a_s and σ_s (from assumption B), the second part of (3) in assumption B and applying Itô isometry, we trivially have

$$\mathbb{E}^n_{i-1}\left(\xi^{(3)}_{i,u}(2)\right) = 0, \quad \mathbb{E}\left|\xi^{(3)}_{i,u}(2)\right|^2 \leq C\Delta_n^3, \quad \mathbb{E}|\xi^{(3)}_{i,u}(3)| \leq C\Delta_n^2 \text{ and } \mathbb{E}|\xi^{(3)}_{i,u}(4)| \leq C\Delta_n^2. \tag{20}$$

We are left with $\xi^{(3)}_{i,u}(5)$. We denote $\xi^{(3,q)}_{i,u}(5) = -\sqrt{2u\Delta_n}\sin\left(\sqrt{2u}\sigma_{(i-1)\Delta_n-}\Delta_n^{-1/2}\Delta_i^n W\right)\int_{(i-1)\Delta_n}^{i\Delta_n}\zeta_s^{(q)}dW_s$ for $q = a, b$ and where

$$\zeta_s^{(a)} = \int_{(i-1)\Delta_n}^{s} v_{(i-1)\Delta_n}dW_u + \int_{(i-1)\Delta_n}^{s} v'_{(i-1)\Delta_n}dW'_u + \int_{(i-1)\Delta_n}^{s}\int_{\mathbb{R}}\delta'((i-1)\Delta_n-,x)\tilde{\mu}'(du,dx),$$
$$\zeta_s^{(b)} = \int_{(i-1)\Delta_n}^{s}\widetilde{\alpha}_u du + \int_{(i-1)\Delta_n}^{s}(v_u - v_{(i-1)\Delta_n})dW_u + \int_{(i-1)\Delta_n}^{s}(v'_u - v'_{(i-1)\Delta_n})dW'_u \tag{21}$$
$$+ \int_{(i-1)\Delta_n}^{s}\int_{\mathbb{R}}(\delta'(u-,x) - \delta'((i-1)\Delta_n-,x))\tilde{\mu}'(du,dx).$$

First, using Itô lemma and the fact that $\Delta_i^n W$ has symmetric distribution, we have

$$\mathbb{E}^n_{i-1}\left[\sin\left(\sqrt{2u}\Delta_n^{-1/2}\sigma_{(i-1)\Delta_n-}\Delta_i^n W\right)\int_{(i-1)\Delta_n}^{i\Delta_n}\left(\zeta_s^{(a)} - \int_{(i-1)\Delta_n}^{s} v_{(i-1)\Delta_n-}dW_u\right)dW_s\right] = 0,$$
$$\mathbb{E}^n_{i-1}\left[\sin\left(\sqrt{2u}\sigma_{(i-1)\Delta_n}\Delta_n^{-1/2}\Delta_i^n W\right)\int_{(i-1)\Delta_n}^{i\Delta_n}(W_s - W_{(i-1)\Delta_n})dW_s\right]$$
$$= 0.5\mathbb{E}^n_{i-1}\left[\sin\left(\sqrt{2u}\sigma_{(i-1)\Delta_n}\Delta_n^{-1/2}\Delta_i^n W\right)\left((\Delta_i^n W)^2 - \Delta_n\right)\right] = 0.$$

This result and application of Burkholder-Davis-Gundy inequality gives

$$\mathbb{E}^n_{i-1}\left(\xi^{(3,a)}_{i,u}(5)\right) = 0, \quad \mathbb{E}\left|\xi^{(3,a)}_{i,u}(5)\right|^2 \leq C\Delta_n^3, \quad \mathbb{E}|\xi^{(3,b)}_{i,u}(5)| \leq C\Delta_n^2. \tag{22}$$

Finally in the proofs of the theorems we will use the following shorthand notation

$$\widehat{Z}_{t,1}(u) = \sum_{[(t-1)/\Delta_n]+1}^{[t/\Delta_n]} \xi_{i,u}^{(1)}, \quad \widehat{Z}_{t,2}(u) = \sum_{[(t-1)/\Delta_n]+1}^{[t/\Delta_n]} (\xi_{i,u}^{(2)}(1) + \xi_{i,u}^{(3)}(2) + \xi_{i,u}^{(3,a)}(5)), \tag{23}$$

and $\widehat{Z}_{t,3}(u) = \widehat{Z}_t(u) - Z_t(u) - \widehat{Z}_{t,1}(u) - \widehat{Z}_{t,2}(u)$.

7.2 Proof of Theorem 1

The stable convergence result in (6) amounts to showing $\mathbb{E}\left(Yf(\Psi_T^n(u))\right) \to \mathbb{E}\left(Yf(\Psi_T(u))\right)$ for all f continuous bounded on $\mathcal{C}(\mathbb{R}_+)$ and Y bounded $\mathcal{F}$-measurable, where $\Psi_T^n(u) = \frac{1}{\sqrt{\Delta_n}}\sum_{t=1}^T(\widehat{Z}_t(u) - Z_t(u))$. For this we use similar argument as in the proof of Theorem VIII.5.8 of Jacod and Shiryaev (2003). By linearity, we can restrict attention to $Y \geq 0$ with $\mathbb{E}(Y) = 1$. We have $\mathbb{E}\left(Yf(\Psi_T^n(u))\right) = \tilde{\mathbb{E}}\left(f(\Psi_T^n(u))\right)$ where $\tilde{\mathbb{P}}$ is a new probability measure that has density Y with respect to the original one. Then, however, using the boundedness of Y, we have $\tilde{\mathbb{P}}(\Psi_T^n(u) \notin K) \leq a\mathbb{P}(\Psi_T^n(u) \notin K)$ for some constant $a > 0$ and K a compact subset of $\mathcal{C}(\mathbb{R}_+)$. Hence to show the convergence in (6) we need to show the result finite-dimensionally in u as well as tightness of the sequence $\Psi_T^n(u)$ (under the original probability measure).

We start with finite-dimensional convergence. Since $|F_i'(x,y)| \leq C$, for $F_i'(x,y)$, $i = 1,2$ denoting first derivatives, applying Cauchy-Schwarz inequality and using (2)-(3), we have

$$\begin{aligned} &\mathbb{E}\left(\sum_{i=1}^{[T/\Delta_n]}\int_{(i-1)\Delta_n}^{i\Delta_n} |F(\sqrt{u}\sigma_s, \sqrt{v}\sigma_s) - F(\sqrt{u}\sigma_{(i-1)\Delta_n-}, \sqrt{v}\sigma_{(i-1)\Delta_n-})|ds\right) \\ &\leq \frac{CT}{\Delta_n}\int_{(i-1)\Delta_n}^{i\Delta_n}\sqrt{\mathbb{E}(F_1'(\sqrt{u}\sigma_s^*, \sqrt{v}\sigma_s^{**}))^2 + F_2'(\sqrt{u}\sigma_s^*, \sqrt{v}\sigma_s^{**}))^2}\sqrt{\mathbb{E}(\sigma_s - \sigma_{(i-1)\Delta_n-})^2}ds \leq C\sqrt{\Delta_n}, \end{aligned} \tag{24}$$

where σ_s^* and σ_s^{**} are values between σ_s and $\sigma_{(i-1)\Delta_n-}$. Then, using the result in (16) and applying Theorem VIII.2.27 of Jacod and Shiryaev (2003), we get finite-dimensionally (in u)

$$\frac{1}{\sqrt{\Delta_n}}\sum_{i=1}^{[T/\Delta_n]}\xi_{i,u}^{(1)} \stackrel{\mathcal{L}}{\longrightarrow} \Psi_T(u). \tag{25}$$

Then, (25) and the bounds on the moments of the rest of the $\xi_{i,u}^{(j)}$ terms in Section 7.1 imply the convergence in (6) of Theorem 1 finite-dimensionally (in u). Further, Theorem 3 of Barndorff-Nielsen et al. (2006) imply that this (finite-dimensional) convergence holds stably.

Turning next to tightness, using the bounds in (16), (17), (20) and (22), we have for $u_1,\ u_2 \in \mathbb{R}_+$

$$\mathbb{E}\left(\frac{1}{\sqrt{\Delta_n}}\sum_{t=1}^T[\widehat{Z}_{t,1}(u_1) + \widehat{Z}_{t,2}(u_1) - \widehat{Z}_{t,1}(u_2) - \widehat{Z}_{t,2}(u_2)]\right)^2 \leq C|\sqrt{u_1} - \sqrt{u_2}|^2 \vee |u_1 - u_2|^2. \tag{26}$$

Using Theorem 20 of Ibragimov and Has'minskii (1981) we have $\frac{1}{\sqrt{\Delta_n}}\sum_{t=1}^T(\widehat{Z}_{t,1}(u)+\widehat{Z}_{t,2}(u))$ is tight. The tightness of $\Psi_T^n(u)$ then follows since for any $\overline{u}>0$ we have that $\left(\Delta_n^{\beta/2+\iota-1}\right)\sup_{0\le u\le\overline{u}}\left|\sum_{i=1}^{[T/\Delta_n]}\widehat{Z}_{t,3}(u)\right|$ is bounded in probability by using the bounds in Section 7.1.

The proof of the consistency of the estimator for the covariance function int the theorem follows trivially from the bounds in (16)-(22) and a Law of Large Numbers for $\frac{1}{T}\sum_{t=1}^T Z_t(u)$. □

7.3 Proof of Theorem 2

Part (a). The proof consists of showing finite dimensional convergence (in (u,v)) and tightness of the sequence. We first make the decomposition

$$\sqrt{T}\left(\frac{1}{T}\sum_{t=k+1}^T \widehat{Z}_t(u)\widehat{Z}_{t-k}(v)-\mu_k(u,v)\right)=\sqrt{T}\left(\frac{1}{T}\sum_{t=k+1}^T Z_t(u)Z_{t-k}(v)-\mu_k(u,v)\right)+\sum_{j=1}^3 R_T^{(j)}(u,v), \quad (27)$$

$$\begin{aligned}
R_T^{(1)}(u,v) &= \frac{1}{\sqrt{T}}\sum_{t=k+1}^T \mathbb{E}\left[Z_t(u)\widehat{Z}_{t-k,1}(v)+\widehat{Z}_{t,1}(u)Z_{t-k}(v)\right],\\
R_T^{(2)}(u,v) &= \frac{1}{\sqrt{T}}\sum_{t=k+1}^T \Big\{Z_t(u)\widehat{Z}_{t-k,1}(v)+\widehat{Z}_{t,1}(u)Z_{t-k}(v)+\widehat{Z}_{t,1}(u)\widehat{Z}_{t-k,1}(v)\\
&\qquad -\mathbb{E}\left[Z_t(u)\widehat{Z}_{t-k,1}(v)+\widehat{Z}_{t,1}(u)Z_{t-k}(v)+\widehat{Z}_{t,1}(u)\widehat{Z}_{t-k,1}(v)\right]\Big\},\\
R_T^{(3)}(u,v) &= \frac{1}{\sqrt{T}}\sum_{t=k+1}^T \Big\{[Z_t(u)+\widehat{Z}_{t,1}(u)][\widehat{Z}_{t-k,2}(v)+\widehat{Z}_{t-k,3}(v)]\\
&\qquad +[\widehat{Z}_{t,2}(u)+\widehat{Z}_{t,3}(u)]\widehat{Z}_{t-k}(v)+\mathbb{E}(\widehat{Z}_{t,1}(u)\widehat{Z}_{t-k,1}(v))\Big\}.
\end{aligned}$$

Note that $\mathbb{E}(\widehat{Z}_{t,1}(u)Z_{t-k}(v))=0$ for $k\ge1$ by an application of (16). Therefore, using stationarity, $R_T^{(1)}(u,v)$ equals the first two components of $Y_T^{(2)}(u,v)$ in (13).

Finite Dimensional Convergence. For the first term on the right-hand side of (27), given assumption C and using a CLT for stationary processes, see Jacod and Shiryaev (2003), Theorem VIII.3.79, we have finite-dimensionally (i.e., over a discrete grid of (u,v))

$$\sqrt{T}\left(\frac{1}{T}\sum_{t=k+1}^T Z_t(u)Z_{t-k}(v)-\mu_k(u,v)\right)\ \xrightarrow{\mathcal{L}}\ \Psi'(u,v). \quad (28)$$

For $R_T^{(2)}(u,v)$, applying successive conditioning and Lemma VIII.3.102 in Jacod and Shiryaev (2003) (which holds due to assumption C), and the bounds of (conditional) moments in Section 7.1, gives

$$\mathbb{E}\left(R_T^{(2)}(u,v)\right)^2\le C\Delta_n\int_0^\infty \alpha_t^{\mathrm{mix}}dt. \quad (29)$$

Finally, for $R_T^{(3)}(u,v)$, we apply the bounds of (conditional) moments in Section 7.1, and the boundedness of $Z_t(u)$ and $\widehat{Z}_t(u)$, to get $R_T^{(3)}(u,v) = O_p(\sqrt{T}\Delta_n^{1-\beta/2-\iota})$.

Tightness. We start with the first term on the right-hand side of (27). For any $[u_1, v_1]$ and $[u_2, v_2]$

$$Z_t(u_1)Z_{t-k}(v_1) - Z_t(u_2)Z_{t-k}(v_2) = \left(Z_t(u_1) - Z_t(u_2)\right) Z_{t-k}(v_1) + Z_t(u_2)\left(Z_{t-k}(v_1) - Z_{t-k}(v_2)\right),$$

and we can do the same for their means. Then, using successive conditioning and Lemma VIII.3.102 in Jacod and Shiryaev (2003), together with the boundedness of $Z_t(u)$ and assumption C, we get

$$T^2\mathbb{E}\left|\frac{1}{T}\sum_{t=k+1}^{T}\left[Z_t(u_1)Z_{t-k}(v_1) - Z_t(u_2)Z_{t-k}(v_2)\right] - \left(\mu_k(u_1,v_1) - \mu_k(u_2,v_2)\right)\right|^4 \leq C||\mathbf{x}_1 - \mathbf{x}_2||^{2+\iota},$$

where $\mathbf{x}_1 = (u_1, v_1)$, $\mathbf{x}_2 = (u_2, v_2)$ and some $\iota > 0$. Then, using Theorem 20 of Ibragimov and Has'minskii (1981), we can conclude the tightness of the sequence for the local uniform topology on $\mathcal{C}(\mathbb{R}_+ \times \mathbb{R}_+)$. Similarly, with the same notation as above

$$\mathbb{E}\left|R_T^{(2)}(u_1,v_1) - R_T^{(2)}(u_2,v_2)\right|^4 \leq C\Delta_n(||\sqrt{\mathbf{x}_1} - \sqrt{\mathbf{x}_2}||^{2+\iota} \vee ||\mathbf{x}_1 - \mathbf{x}_2||^{2+\iota}).$$

Finally, using the bounds in Section 7.1, it is easy to show that for any $\overline{u}, \overline{v} > 0$, we have that $\left(\sqrt{T}\Delta_n^{1-\beta/2-\iota}\right)^{-1} \sup_{0\leq u\leq\overline{u},\ 0\leq v\leq\overline{v}} \left|R_T^{(3)}(u,v)\right|$ is bounded in probability. This then implies the asymptotic negligibility of $R_T^{(3)}(u,v)$ on $\mathcal{C}(\mathbb{R}_+ \times \mathbb{R}_+)$ equipped with the local uniform topology.

Part b. We denote for $l \geq 0$,

$$C_l([u_1,v_1],[u_2,v_2]) = \frac{1}{T}\sum_{t=k+l+1}^{T}\left(Z_t(u_1)Z_{t-k}(v_1) - \mu_k(u_1,v_1)\right)\left(Z_{t-l}(u_2)Z_{t-l-k}(v_2) - \mu_k(u_2,v_2)\right).$$

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

$$C_0([u_1,v_1],[u_2,v_2]) + \sum_{i=1}^{L_T}\omega(i,L_T)\left(C_i([u_1,v_1],[u_2,v_2]) + C_i([u_2,v_2],[u_1,v_1])\right) \xrightarrow{\mathbb{P}} V([u_1,v_1],[u_2,v_2]). \quad (30)$$

Using the boundedness of $\widehat{Z}_t(u)$ and $Z_t(u)$, we next have

$$\left|\widehat{C}_i([u_1,v_1],[u_2,v_2]) - C_i([u_1,v_1],[u_2,v_2])\right| \leq C\sum_{j=1,2}\left(\frac{1}{T}\sum_{t=1}^{T}(|\widehat{Z}_t(u_j) - Z_t(u_j)| + |\widehat{Z}_t(v_j) - Z_t(v_j)|)\right.$$
$$\left. + \left|\frac{1}{T}\sum_{t=k+1}^{T} Z_t(u_j)Z_{t-k}(v_j) - \mu_k(u_j,v_j)\right|\right).$$

Using assumption C and Lemma VIII.3.102 in Jacod and Shiryaev (2003), as well as the bounds in Section 7.1, we get for arbitrary small $\iota > 0$ and $j = 1, 2$

$$\mathbb{E}\left|\frac{1}{T}\sum_{t=k+1}^{T} Z_t(u_j)Z_{t-k}(v_j) - \mu_k(u_j,v_j)\right| \leq \frac{C}{\sqrt{T}}, \quad \mathbb{E}\left|\widehat{Z}_t(u) - Z_t(u)\right| \leq C\Delta_n^{(1-\beta/2-\iota)\wedge 1/2}.$$

The result in (14) then follows from the relative speed condition between L_T, T and Δ_n. $\square$

References

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. 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*, 253–280.

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

Barndorff-Nielsen, O. E., S. E. Graversen, J. Jacod, and N. Shephard (2006). Limit theorems for realised bipower variation in econometrics. *Econometric Theory 22*, 677–719.

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.

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

Engle, R. (2004). Risk and volatility: Econometric models and financial practice. *American Economic Review 94*, 405–420.

Hansen, L. and J. Scheinkman (1995). Back to the Future: Generating Moment Implications for Continuous Time Markov-Processes. *Econometrica 63*, 767–804.

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

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

Kryzhniy, V. (2010). On Regularization Method for Numerical Inversion of The Laplace Transforms Computable At Any Point On The Real Axis. *Journal of Inverse Ill-Posed Problems 18*, 409–419.

Mancini, C. (2001). Disentangling the Jumps of the Diffusion in a Geometric Brownian Motion. *Giornale dell'Istituto Italiano degi Attuari LXIV*, 19–47.

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.

Todorov, V., G. Tauchen, and I. Grynkiv (2011). Realized Laplace Transforms for Estimation of Jump Diffusive Volatility Models. *Journal of Econometrics 164*, 367–381.