REALIZED LAPLACE TRANSFORMS FOR PURE-JUMP SEMIMARTINGALES

By Viktor Todorov and George Tauchen

Northwestern University and Duke University

We consider specification and inference for the stochastic scale of discretely-observed pure-jump semimartingales with locally stable Lévy densities in the setting where both the time span of the data set increases and the mesh of the observation grid decreases. The estimation is based on constructing nonparametric estimate for the empirical Laplace transform of the stochastic scale over a given interval of time by aggregating high-frequency increments of the observed process on that time interval into a statistic we call realized Laplace transform. The realized Laplace transform depends on the activity of the driving pure-jump martingale and we consider both cases when the latter is known or has to be inferred from the data.

1. Introduction. Continuous-time semimartingales are used extensively for modeling many processes in various areas and in particular finance. Typically the model of interest is an Itô semimartingale (semimartingale with absolute continuous characteristics) given by

$$dX_t = \alpha_t dt + \sigma_{t-} dZ_t + dY_t, \tag{1.1}$$

where α_t and σ_t are some processes with càdlàg paths, Z_t is an infinite variation Lévy martingale, Y_t is a finite variation jump process satisfying certain regularity conditions (all technical conditions on the various processes will be given later). The martingale Z_t can be continuous (i.e., Brownian motion), jump-diffusion or of pure-jump type (i.e., without a continuous component). The presence of the last term in (1.1) might appear redundant as Z_t can already contain jumps, but its presence will allow us to encompass also the class of time-changed Lévy processes in our analysis. In any case, this last term in (1.1) is dominated over small scales by the term involving Z_t.

Our interest in this paper will be in inference about the process σ_t when Z_t is a pure-jump Lévy process. Pure-jump models have been used to study various processes of interest such as volatility and volume of financial prices [3, 5], traffic data [21], and electricity prices [18].

AMS 2000 subject classifications: Primary 62F12, 62M05; secondary 60H10, 60J75.

Keywords and phrases: Laplace transform, time-varying scale, high-frequency data, jumps, inference.

Parametric or nonparametric estimation of a model satisfying (1.1) in the pure-jump case is quite complicated for at least the following reasons. First, very often the transitional density of X_t is not known in closed-form. This holds true even in the relatively simple case when X_t is a pure-jump Lévy process. Second, in many situations the realistic specification of σ_t often implies that X_t is not a Markov process, with respect to its own filtration, and hence all developed methods for estimation of the latter will not apply. Third, the various parameters of the model (1.1) capture different statistical properties of the process X_t and hence will have various rates of convergence depending on the sampling scheme. For example, in general α_t and the tails of Z_t can be estimated consistently only when the span of the data increases, whereas the so-called activity of Z_t can be recovered even from a fixed-span data set, provided the mesh of the latter decreases (see [6] for estimation of the activity from low-frequency data set). Finally, the simulation of the process X_t can be in many cases difficult or time consuming.

In view of the above-mentioned difficulties, our goal here is specification analysis for only part of the model, mainly the process σ_t, in the case when X_t is a pure-jump process following (1.1). We conduct inference in the case when we have a high-frequency data set of X_t with increasing time span (see [6] and [22] for inference about jump processes based on low frequency). We refer to σ_t as the stochastic scale of the pure-jump process X_t in analogy with the scale parameter of a stable process, i.e., when X_t is a stable process then the constant σ is the scale of the process. σ_t is key in the specification of (1.1) and in particular it captures the time-variation of the process X_t over small intervals of time. Our goal here will be to make the inference about σ_t robust to the rest of the components of the model, i.e., the specification of α_t and Y_t as well as the dependence between σ_t and Z_t.

The inference in the paper is for processes for which the Lévy measure of the driving martingale Z_t in (1.1) behaves around zero like that of a stable process. This covers of course the stable process, but also many other Lévy processes of interest with details provided in Section 2 below. The idea of our proposed method of inference is to use the fact that when Z_t is locally stable, the leading component of the process X_t over small scales is governed by that of the "stable component" of Z_t. Moreover, when σ_t is an Itô semimartingale, then "locally" its changes are negligible and σ_t can be treated as constant. Intuitively then infill asymptotics can be conducted as if the increments of X_t are products of (a locally constant) stochastic scale and independent i.i.d. stable random variables. This in particular implies that the empirical characteristic function of the high-frequency increments over a small interval of time will estimate the characteristic function of a scaled stable process. The latter, however, is the Laplace transform for the locally constant stochastic scale. Therefore, aggregating over a fixed interval time the empirical characteristic function of the (appropriately re-scaled) high-frequency increments of X_t provides a nonparametric estimate for the empirical Laplace transform of the stochastic scale over that interval. We refer to this simple statistic as realized Laplace transform for pure-jump processes. The connection between the empirical characteristic function of the driving martingale and the Laplace transform of the stochastic scale in the context of time-changed Lévy processes, with time-change independent of the driving martingale, has been previously used for low-frequency estimation in [7].

The inference based on the realized Laplace transform is robust to the specification of α_t as well as the tail behavior of Z_t. Intuitively, this is due to our use of the high-frequency data whose marginal law is essentially determined by the small jumps of Z_t and the stochastic scale σ_t. Quite naturally, however, our inference depends on the activity of the small jumps of the driving martingale Z_t. The latter corresponds to the index of the stable part of Z and using the self-similarity of the stable process, it determines its scaling over different (high) frequencies. Therefore, the activity index enters directly in the calculation of the realized Laplace transform. We conduct inference both in the case where the activity is assumed known and when it needs to be estimated from the data. The estimation of the activity index however differs from the inference for the stochastic scale. While for the latter we need in general the time-span to increase to infinity (except for the degenerate case when σ_t is actually constant), for the former this is not the case. The activity index can be estimated only with a fixed span of high-frequency data and in general increasing time-span will not help for its nonparametric estimation. Therefore, we estimate the activity index of Z_t using initial part of the sample with a fixed span and then plug it in the construction of the realized Laplace transform. We further quantify the asymptotic effect from this plug-in approach on the inference for the Laplace transform of the stochastic scale.

The Laplace transform of the stochastic scale preserves the information for its marginal distribution. Therefore, it can be used for efficient estimation and specification testing. We illustrate this in a parametric setting by minimizing a distance between our nonparametric Laplace estimate and a model implied one, similar to estimation based on the empirical characteristic function as in [13].

Finally, the current paper studies the realized Laplace transform for the case when Z_t is pure-jump while [29] (and the empirical application of it in [30]) consider the case where Z_t is a Brownian motion. The pure-jump case is substantively different, starting from the very construction of the statistic as well as its asymptotic behavior. The leading component in the asymptotic expansions in the pure-jump case is a stable process with index less than 2 and this index is in general unknown and needs to be estimated, which further necessitates different statistical analysis from the continuous case. Also, the residual components in X_t, like α_t, play a more prominent role when X_t is of pure-jump type, and when the activity is low this requires modifying appropriately the realized Laplace transform to purge them.

The paper is organized as follows. Section 2 presents the formal setup and assumptions. Section 3 introduces the realized Laplace transform and derives its limit behavior. In Section 4 we conduct a Monte Carlo study and in Section 5 we present a parametric application of the developed limit theory. Section 6 concludes. The proofs are given in Section 7.

2. Setting and Assumptions.

Throughout the paper, the process of interest is denoted with X_t and is defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. Before stating our assumptions, we recall that a Lévy process L_t with characteristic triplet (b, c, ν) with respect to truncation function $\kappa(x) = x$ (we will always assume that the process has a finite first order moment) is a process with characteristic function given by

$$\mathbb{E}\left(e^{iuL_t}\right) = \exp\left(t\left(iub - u^2c/2 + \int_{\mathbb{R}}(e^{iux} - 1 - iux)\nu(dx)\right)\right). \tag{2.1}$$

With this notation, we assume that the Lévy process Z_t in (1.1) has a characteristic triplet $(0, 0, \nu)$ for ν some Lévy measure. Note that since the truncation function with respect to which the characteristics of the Lévy processes are presented is the identity, the above implies that Z_t is a pure-jump martingale. The first term in (1.1) is the drift term. It captures the persistence in the process and when X_t is used to model financial prices, the drift captures compensation for risk and time. The second term in (1.1) is defined in a stochastic sense since in assumption A below we will assume that Z_t is of infinite variation. The last term in (1.1) is a finite variation pure-jump process. Assumption A below will impose some restrictions on its properties, but we stress that there is no assumption of independence between the processes σ_t, Z_t and Y_t.

In the pure-jump model the jump martingale Z_t substitutes the Brownian motion used in jump-diffusions to model the "small" moves. We note that the "dominant" part of the increment of X_t over a short interval of time $(t, t+\Delta)$ is $\sigma_{t-} \times (Z_{t+\Delta} - Z_t)$. This term is of order $O_p(\Delta^\alpha)$ for $\alpha \in [1/2, 1)$, while the rest of the components of X_t are at most $O_p(\Delta)$ when $\Delta \downarrow 0$.

We recall from the introduction that our object of interest in this paper is the stochastic scale of the martingale component of X_t, i.e., σ_t. Of course we observe only X_t 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_t and the specification of σ_t itself (including the activity of the driving martingale). 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 scale with $Z_{t,t+\Delta}$. Our assumption A below characterizes the behavior of Z_t and Y_t over small scales.

Assumption A.The Lévy density of Z_t, ν, is given by

$$\nu(x) = \frac{A}{|x|^{\beta+1}} + \nu'(x), \quad \beta \in (1,2), \quad \int_{\mathbb{R}} |x|\nu(x)dx < \infty, \tag{2.2}$$

where

$$A = \left(\frac{4\Gamma(2-\beta)|\cos(\beta\pi/2)|}{\beta(\beta-1)}\right)^{-1}, \quad \beta \in (1,2), \tag{2.3}$$

and further there exists $x_0 > 0$ such that for $|x| \le x_0$ we have $|\nu'(x)| \le \frac{C}{|x|^{\beta'+1}}$ for some $\beta' < 1$ and a constant $C \ge 0$.

We further have Y_t absolutely integrable and $\mathbb{E}|Y_t - Y_s|^{\beta'} < C|t-s|| \log |t-s||$ for every $t,\ s \ge 0$ with $|t-s| \le 1$, some positive constant C, and $\beta' < 1$ being the constant above.

Assumption A implies that the small scale behavior of the driving martingale Z_t is like that of a stable process with index β. 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. Formally, β equals the Blumenthal-Getoor index of the Lévy process Z_t. The value of the index β is crucial for recovering σ_t from the discrete data on X_t, as intuitively it determines how big on average the increments $Z_{t,t+\Delta}$ should be for a given sampling frequency. The following lemma makes this formal.

LEMMA 1. *Let Z_t satisfy assumption A. Then for $h \to 0$ we have*

$$h^{-1/\beta} Z_{th} \xrightarrow{\mathcal{L}} S_t, \tag{2.4}$$

where the convergence is for the Skorokhod topology on the space of càdlàg functions: $\mathbb{R}_+ \to \mathbb{R}$, and S_t is a stable process with characteristic function $\mathbb{E}\left(e^{iuS_1}\right) = e^{-|u|^\beta/2}$.

The value of the constant A in (2.3) is a normalization that we impose. We are obviously free to do that since what we observe is X_t whose leading component over small scales is an integral of σ_t with respect to the jump martingale Z_t and we never observe the two separately. The above choice of A is a convenient one that ensures that when $\beta \to 2$, the jump process converges finite-dimensionally to Brownian motion. We note that in assumption A we rule out the case $\beta \leq 1$ but this is done for brevity of exposition as most processes of interest are of infinite variation, (although we rule out some important processes like the Generalized Hyperbolic).

In assumption A we restrict the "activity" of the "residual" jump components of X, i.e., we limit their effect in determining the small moves of X. The effect of the "residual" jump components on the small moves is controlled by the parameter β'. From (2.2), the leading component of $\nu(x)$ is the Lévy density of a stable and $\nu'(x)$ is the residual one. The restriction $\beta' < 1$, implies that the "residual" jump component is of finite variation. This restriction is not necessary for convergence in probability results (only $\beta' < \beta$ is needed for this) but is probably unavoidable if one needs also the asymptotic distribution of the statistics that we introduce in the paper. In most parametric models this restriction is satisfied.

We note that $\nu'(x)$ in (2.2) 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 [24], generated from the stable by tempering its tails, which has all its moments finite. Therefore, while assumption A ties the small scale behavior of the driving martingale Z_t with that of a stable process, it 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 by our assumption) and thus in particular unrelated with that of a stable process.

Remark 1. *Assumption A is analogous to the assumption used in [2]. It is also related with the so-called regular Lévy processes of exponential type studied in [9] with $\beta = \nu$ in the notation of that paper. Compared with the above mentioned processes of [9], we impose slightly more structure on the Lévy density around zero but no restriction outside of it. We note that if assumption A fails then the results that follow are not true. The degree of the violation depends on the sampling frequency and the deviation of the characteristic function of Z over small scales from that of a stable.*

Finally, the process Y_t also captures a "residual" jump component of X_t in terms of its small scale behavior. Assumption A limits its activity by β'. The component of Z corresponding to $\nu'(x)$ and Y_t control the jump measure of X_t away from zero. Unlike the former whose time variation is determined by σ_t, the latter has essentially unrestricted time variation. There is clearly some "redundancy" in the specification in (1.1) in terms of modeling the jumps of Z_t away from zero, but this is done to cover more general pure-jump models as clear from the following two remarks.

Remark 2. *Assumption A nests time-changed Lévy processes with absolute continuous time changes (see e.g., [11]), i.e., specifications of the form*

$$dX_t = \alpha_t dt + \int_0^t \int_{\mathbb{R}} x\widetilde{\mu}(ds, dx), \tag{2.5}$$

where μ is a random integer-valued measure with compensator $a_t dt \otimes \nu(dx)$ for some nonnegative process a_t and Lévy measure $\nu(dx)$ satisfying (2.2) of assumption A and $\widetilde{\mu} = \mu - \nu$. This can be shown using Theorem 14.68 of [14] or Theorem 2.1.2 of [15], linking integrals of random functions with respect to Poisson measure and random integer-valued measures, and implies that X_t in (2.5) can be equivalently represented as (1.1) with σ_t given by $a_t^{1/\beta}$.

On the other hand, if we start with X given by (1.1), with Z strictly stable and no Y, we can show using the definition of jump compensator and Theorem II.1.8 of [16] that the latter is a time-changed Lévy process with time change $|\sigma_{t-}|^\beta$. For more general "stable-like" Lévy processes, we need to introduce an additional term (this is Y_t in (1.1)) in addition to the above time-changed stable process.

We note that the connection of (1.1) with the time-changed Lévy processes does not depend on the presence of any dependence between σ_t and Z_t.

Remark 3. *Assumption A is also satisfied by the pure-jump Lévy-driven CARMA models (continuous-time autoregressive moving average) which have been used for modeling series exhibiting persistence, see e.g., [10] and the many references therein. For these processes σ_t in (1.1) is a constant.*

Our next assumption imposes minimal integrability conditions on α_t and σ_t and further limits the amount of variation 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 process σ_t is an Itô semimartingale given by

$$\sigma_t = \sigma_0 + \int_0^t \tilde{\alpha}_s ds + \int_0^t \tilde{\sigma}_s dW_s + \int_0^t \int_{\mathbb{R}} \underline{\delta}(s-, x)\underline{\tilde{\mu}}(ds, dx), \tag{2.6}$$

where W is a Brownian motion; $\underline{\mu}$ is a homogenous Poisson measure, with Lévy measure $\underline{\nu}(dx)$, having arbitrary dependence with μ, for μ being the jump measure of Z. We assume $|\underline{\delta}(t, x)| \leq \gamma_t \underline{\delta}(x)$ for some integrable process

γ_t and $\int_{\mathbb{R}}(\underline{\delta}(x)\wedge 1)^{\beta''}\underline{\nu}(dx)<\infty$ for some $\beta''<2$. Further for every t and s we have:

$$\left\{\begin{array}{l}\mathbb{E}\left(\alpha_t^2+\sigma_t^2+\tilde{\alpha}_t^2+\tilde{\sigma}_t^2+(\tilde{\sigma}'_t)^2+\int_{\mathbb{R}}\underline{\delta}^2(t,x)\underline{\nu}(dx)\right)<C,\\ \mathbb{E}\left(|\alpha_t-\alpha_s|^2+|\tilde{\sigma}_t-\tilde{\sigma}_s|^2+\int_{\mathbb{R}}(\underline{\delta}(t,x)-\underline{\delta}(s,x))^2\underline{\nu}(dx)\right)<C|t-s|,\end{array}\right. \tag{2.7}$$

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

Assumption B imposes σ_t to be an Itô semimartingale. This is a relatively mild assumption satisfied by the popular multifactor affine jump-diffusions [12] as well as the CARMA Lévy-driven models used to model persistent processes [10]. Assumption B rules out certain long-memory specifications for σ_t although we believe that at least for some of them the results in this paper will continue to hold.

Importantly, however, assumption B allows for jumps in the stochastic scale that can have arbitrary dependence with the jumps in X_t which is particularly relevant for modeling financial data, e.g., the parametric models of [17]. Finally, the second part of (2.7) will be satisfied when the corresponding processes are Itô semimartingales. The next assumption B′ restricts assumption B in a way that will allow us to strengthen some of the theoretical results in the next section.

Assumption B′. The process σ_t is an Itô semimartingale given by

$$\sigma_t=\sigma_0+\int_0^t\tilde{\alpha}_s ds+\int_0^t\tilde{\sigma}_s dW_s+\int_0^t\tilde{\sigma}'_s dZ_s+\int_0^t\int_{\mathbb{R}}\underline{\delta}(s-,x)\underline{\tilde{\mu}}(ds,dx), \tag{2.8}$$

with the same notation as assumption B with the only difference that $\underline{\mu}$ is now independent from μ. We also assume that corresponding condition (2.7) holds as well as that Z_t is square-integrable.

The strengthening in assumption B′ is in the modeling of the dependence between the jumps in X_t and σ_t. In assumption B′ this is done via the third integral in (2.8). This is similar to modeling dependence between continuous martingales using correlated Brownian motions. What is ruled out by assumption B′ is dependence between the jumps in X_t and σ_t that is different for the jumps of different size. Assumption B′ will be satisfied when the pair (X_t,σ_t) are modeled jointly via a Lévy-driven SDE.

Finally, in our estimation we make use of long-span asymptotics for 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 next.

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

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

3. Limit theory for RLT of Pure-Jump Semimartingales. Now we are ready to formally define the realized Laplace transform for the pure-jump model and derive its asymptotic properties. We assume the process X_t is observed 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 is then defined as:

(3.1)
$$V_T(X, \Delta_n, \beta, u) = \sum_{i=1}^{[T/\Delta_n]} \Delta_n \cos((2u)^{1/\beta}\Delta_n^{-1/\beta}\Delta_i^n X), \quad \Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n},$$

where β is the activity index of the driving martingale Z_t given in its Lévy density in (2.2). $V_T(X, \Delta_n, \beta, u)$ is the real part of the empirical characteristic function of the appropriately scaled increments of the process X_t. In the case of jump-diffusions, β in (3.1) is replaced with 2 as the activity of the Brownian motion has an index of 2 (i.e., for the Brownian motion Lemma 1 holds with β replaced by 2). We show in this section that $V_T(X, \Delta_n, \beta, u)/T$ is a consistent estimator for the empirical Laplace transform of $|\sigma_t|^\beta$ and further derive its asymptotic properties under various sampling schemes as well as assumptions regarding whether β is known or needs to be estimated.

3.1. *Fixed Span Asymptotics.* We start with the case when T is fixed and $\Delta_n \to 0$, i.e., the infill asymptotics, and we further assume we know β. Since the driving martingale over small scales behaves like β-stable (assumption A) and the stochastic scale changes over short intervals are not too big on average (assumption B), then the "dominant" part (in a infill asymptotic sense) of the increment $\Delta_i^n X$ (when Δ_n is small) is $\sigma_{(i-1)\Delta_n}\Delta_i^n Z$. $\Delta_i^n Z$ is approximately stable and from Lemma 1 we have approximately $\Delta_i^n Z \stackrel{d}{=} \Delta_n^{1/\beta} \times Z_1$ with the characteristic function of Z_1 given by $e^{-|u|^\beta/2}$. 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 $\int_0^T e^{-u|\sigma_t|^\beta} ds$, which is the empirical Laplace transform of $|\sigma_t|^\beta$ after dividing by T. The following theorem gives the precise infill asymptotic 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.

THEOREM 1. *For the process X_t, assume that assumptions A and B hold with $\beta' < \beta/2$ and let $\Delta_n \to 0$ with T fixed.*

(a) If $\beta > 4/3$, then we have

(3.2)
$$\frac{1}{\sqrt{\Delta_n}}\left(V_T(X,\Delta_n,\beta,u) - \int_0^T e^{-u|\sigma_s|^\beta}ds\right) \xrightarrow{\mathcal{L}-s} \sqrt{\int_0^T F_\beta(u^{1/\beta}|\sigma_s|)ds} \times E,$$

where E is a standard normal variable defined on an extension of the original probability space and independent from the σ-field $\mathcal{F}$; $F_\beta(x) = \frac{e^{-2^{\beta-1}x^\beta} - 2e^{-x^\beta}+1}{2}$ for $x > 0$.
A consistent estimator for the asymptotic variance is given by

$$\frac{V_T(X,\Delta_n,\beta,2^{\beta-1}u) - 2V_T(X,\Delta_n,\beta,u)+1}{2}. \tag{3.3}$$

(b) If $\beta \leq 4/3$, then

$$\left(V_T(X,\Delta_n,\beta,u) - \int_0^T e^{-u|\sigma_s|^\beta}ds\right) = O_p\left(\Delta_n^{2-2/\beta}\right). \tag{3.4}$$

In the case when X_t is a Lévy process, or more generally when only the scale σ_t is constant as is the case for the Lévy-driven CARMA models, the above theorem can be used to estimate the scale coefficient σ by either fixing some u or using a whole range of u-s as in the methods for estimation of stable processes based on the empirical characteristic function, see e.g., [19] and [1]. Furthermore, this can be done jointly with the nonparametric estimation of β by using for example the estimator we proposed in [27] that we define in (3.17) below.

The limit result in Theorem 1 is driven by the small jumps in X_t and this allows us to disentangle the stochastic scale (which drives their temporal variation) from the other components of the model, mainly the jumps away from zero. This is due to the fact that the cosine function is bounded and infinitely differentiable which limits the effect of the jumps of size away from zero on it. By contrast, for example, the infill asymptotic limit of the quadratic variation of the discretized process is the quadratic variation of X_t which is determined by all jumps not just the infinitely small ones.

Unfortunately when the activity of the driving martingale is relatively low, i.e., $\beta < 4/3$, we do not have a CLT for $V_T(X,\Delta_n,\beta,u)$. The reason is in the presence of the drift term, which for the purposes of our estimation starts behaving closer to the driving martingale Z_t and this slows the rate of convergence of our statistic. However, we can use the fact that over successive short intervals of time the contribution of the drift term in the increments of X_t is the same while sum or difference of i.i.d. stable random variables continues to have a stable distribution. Therefore, if we difference the increments of X_t, we will remove the drift term (up to the effect due to the time-variation in it which will be negligible) and the leading term will still be a product of the locally constant stochastic scale and a stable variable. Thus we consider the following alternative estimator

$$\widetilde{V}_T(X,\Delta_n,\beta,u) = \sum_{i=2}^{[T/\Delta_n]} \Delta_n \cos\left[u^{1/\beta}\Delta_n^{-1/\beta}(\Delta_i^n X - \Delta_{i-1}^n X)\right]. \tag{3.5}$$

THEOREM 2. *For the process X_t, assume that assumptions A and B hold with $\beta' < \beta/2$ and let $\Delta_n \to 0$ with T fixed. We have*

$$\frac{1}{\sqrt{\Delta_n}}\left(\widetilde{V}_T(X,\Delta_n,\beta,u) - \int_0^T e^{-u|\sigma_s|^\beta}ds\right) \xrightarrow{\mathcal{L}-s} \sqrt{\int_0^T \widetilde{F}_\beta(u^{1/\beta}|\sigma_s|)ds} \times \widetilde{E}, \tag{3.6}$$

where $\widetilde{E}$ is a standard normal variable defined on an extension of the original probability space and independent from the σ-field $\mathcal{F}$; $\widetilde{F}_\beta(x) = \left(\frac{e^{-2^{\beta-1}x^\beta}+1}{2}\right)^2 + 2e^{-x^\beta}\frac{e^{-2^{\beta-1}x^\beta}+1}{2} - 3e^{-2x^\beta} + \left(\frac{e^{-2^{\beta-1}x^\beta}-1}{2}\right)^2$ for $x > 0$.

A comparison of the standard errors in (1) and (2) shows that the latter can be up to 2.5 times higher than the former for values of $\beta > 4/3$. This is the cost of removing the effect of the drift term via the differencing of the increments. Therefore, $\widetilde{V}_T(X,\Delta_n,\beta,u)$ should be used only in the case when $\beta \leq 4/3$. For brevity, the results that follow will be presented only for $V_T(X,\Delta_n,\beta,u)$, but analogous results will hold for $\widetilde{V}_T(X,\Delta_n,\beta,u)$.

3.2. *Long Span Asymptotics: The Case of Known Activity.* We continue next with the case when the time span of the data increases together with the mesh of observation grid decreasing. The high-frequency data allows us to "integrate out" the increments $\Delta_i^n Z$, i.e. it essentially allows to "deconvolute" σ_t from the driving martingale of X_t in a robust way. After dividing by T, the infill asymptotic limit of (3.1) is the empirical Laplace transform of the stochastic scale and we henceforth denote it as $\widehat{\mathcal{L}}_\beta(u) = \frac{1}{T}V_T(X,\Delta_n,\beta,u)$. Then, by letting $T \to \infty$ we can eliminate the sampling variation due to the stochastic nature of σ_t, and thus recover its population properties, i.e., estimate $\mathcal{L}_\beta(u) = \mathbb{E}\left(e^{-u|\sigma_t|^\beta}\right)$ which is the Laplace transform of $|\sigma_t|^\beta$.

The next theorem gives the asymptotic behavior of $\widehat{\mathcal{L}}_\beta(u)$ when both $T \to \infty$ and $\Delta_n \to 0$. To state the result we first introduce some more notation. We henceforth use the shorthand

$$\widehat{Z}_{t,\beta}(u) = V_t(X,\Delta_n,\beta,u) - V_{t-1}(X,\Delta_n,\beta,u), \quad t = 1,...,T, \tag{3.7}$$

for the RLT over the time interval $(t-1,t)$. We further set

$$\widehat{C}_{k,\beta}(u,v) = \frac{1}{T}\sum_{t=k+1}^{T}\left(\widehat{Z}_{t,\beta}(u) - \widehat{\mathcal{L}}_\beta(u)\right)\left(\widehat{Z}_{t-k,\beta}(v) - \widehat{\mathcal{L}}_\beta(v)\right), \quad k \in \mathbb{N}. \tag{3.8}$$

THEOREM 3. *Suppose $T \to \infty$ and $\Delta_n \to 0$ and the process X_t satisfies assumptions A, B and C.*

(a) We have

$$\sqrt{T}\left(\widehat{\mathcal{L}}_\beta(u) - \mathcal{L}_\beta(u)\right) = Y_T^{(1)}(u) + Y_T^{(2)}(u), \tag{3.9}$$

$$\begin{cases} Y_T^{(1)}(u) \xrightarrow{\mathcal{L}} \Psi(u), \\ Y_T^{(2)}(u) = O_p\left(\sqrt{T}\left(|\log\Delta_n|\Delta_n^{1-\beta'/\beta} \vee \Delta_n^{(2-2/\beta)\wedge 1/2}\right) \vee \sqrt{\Delta_n}\right), \end{cases} \tag{3.10}$$

where the result for $Y_T^{(2)}(u)$ holds locally uniformly in u and the convergence of $Y_T^{(1)}(u)$ is on the space $\mathcal{C}(\mathbb{R}_+)$ of continuous functions indexed by u and equipped with the local uniform topology (i.e. uniformly over compact sets of $u \in \mathbb{R}_+$) and $\Psi(u)$ is a Gaussian process with variance-covariance for $u, v > 0$ given by

$$\begin{aligned}\Sigma_\beta(u,v) = \int_0^\infty \mathbb{E}\Big[&\left(e^{-u|\sigma_t|^\beta} - \mathcal{L}_\beta(u)\right)\left(e^{-v|\sigma_0|^\beta} - \mathcal{L}_\beta(v)\right) \\ &+ \left(e^{-v|\sigma_t|^\beta} - \mathcal{L}_\beta(v)\right)\left(e^{-u|\sigma_0|^\beta} - \mathcal{L}_\beta(u)\right)\Big]dt.\end{aligned} \tag{3.11}$$

If we strengthen assumption B to assumption B′, we get the stronger

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

(b) For arbitrary integer $k \geq 1$ and every $u, v > 0$ we have

$$\widehat{Z}_{1,\beta}(u)\widehat{Z}_{k,\beta}(v) \xrightarrow{\mathbb{P}} \int_0^1\int_{k-1}^k e^{-u|\sigma_t|^\beta}e^{-v|\sigma_s|^\beta}dsdt. \tag{3.12}$$

If further L_T is a deterministic sequence of integers satisfying $\frac{L_T}{\sqrt{T}} \to 0$ as $T \to \infty$ and $L_T \left(|\log \Delta_n| \Delta_n^{1-\beta'/\beta} \vee \Delta_n^{(2-2/\beta)\wedge 1/2}\right) \to 0$, we have

(3.13)
$$\widehat{\Sigma}_\beta(u,v) = \widehat{C}_{0,\beta}(u,v) + \sum_{i=1}^{L_T} \omega(i, L_T)(\widehat{C}_{i,\beta}(u,v) + \widehat{C}_{i,\beta}(v,u)) \xrightarrow{\mathbb{P}} \Sigma_\beta(u,v),$$

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

The result in (3.9) holds locally uniformly in u. This is important as in a typical application one needs the Laplace transform as a function of u. We illustrate in the next section an application of the above result to parametric estimation that makes use of the uniformity. We note also that $\Sigma_\beta(u,v)$ is well defined because of assumption C, see [16], Theorem VIII.3.79.

Under the conditions of Theorem 3, 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 $\sqrt{T}\left(\frac{1}{T}\int_0^T e^{-u|\sigma_t|^\beta} dt - \mathbb{E}\left(e^{-u|\sigma_t|^\beta}\right)\right)$, which is the empirical process corresponding to the case of continuous-record of X_t in which case $|\sigma_t|^\beta$ can be recovered exactly. 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_t in the estimation. For $Y_T^{(2)}(u)$ to be negligible we need a condition for the relative speed of $\Delta_n \to 0$ and $T \to \infty$ which in the general case of assumption B is given by $\sqrt{T}\left(|\log \Delta_n| \Delta_n^{1-\beta'/\beta} \vee \Delta_n^{(2-2/\beta)\wedge 1/2}\right) \to 0$.

The relative speed condition is driven by the biases that arise from using the discretized observations of X_t. The martingale term that determines the limit behavior of the statistic for a fixed span in Theorem 1 is dominated by the empirical process error $Y_T^{(1)}(u)$ when the time span increases. The leading biases due to the discretization are two: the drift term α_t and the presence of "residual" jump components in X_t in addition to its leading stable component at high frequencies. The bias in $\widehat{\mathcal{L}}_\beta(u)$ due to the "residual" jump components is $O_p(|\log \Delta_n| \Delta_n^{1-\beta'/\beta})$. The higher the activity of the "residual" jump components is, the stronger their effect is on measuring the Laplace transform of $|\sigma_t|^\beta$. Typically, β' will be determined from a Taylor expansion of the Lévy density of the driving martingale around zero. In this case $\beta' = \beta - 1$ and the bias will be bigger for the higher levels of activity β. The bias due to the drift term is $O_p(\Delta_n^{2-2/\beta})$ and it becomes bigger the lower the activity of the driving martingale is. This bias can be significantly reduced if we make use of $\widetilde{V}_T(X, \Delta_n, \beta, u)$ when estimating $\mathcal{L}_\beta(u)$. Finally, the orders of magnitude of the above biases can be shown to be optimal by deriving exactly the bias in the simple case (covered by our assumption A) in which X_t is Lévy and further Z_t is a sum of two independent stable processes with indexes β and β'.

The relative speed condition here can be compared with the corresponding one that arises in the problem of maximum likelihood estimation of Markov jump-diffusions, see e.g., [26]. The general condition in this problem is $T\Delta_n \to 0$ (also known as the rapidly increasing experimental design), i.e., the mesh of the grid should increase somewhat faster than the time span of the data. In our problem here we need weaker relative speed condition provided we use the stronger assumption B$'$ and the deviation of Z_t from a stable process at high frequencies is not too big, i.e., β' is relatively low.

Part b of Theorem 3 makes the limit result in (3.10) feasible, i.e., it provides estimates from the high-frequency data for the asymptotic variance of the leading term $Y_T^{(1)}(u)$. The first result in it, i.e., the limit in (3.12) is of independent interest. The sample average of the limit in (3.12) essentially identifies the integrated joint Laplace transform of $|\sigma_t|^\beta$. This is a natural extension of our results here for the marginal Laplace transform of $|\sigma_t|^\beta$ and can be used for estimation and testing of the transitional density specification of the stochastic scale. We do not pursue this any further here.

Finally, the proof of Theorem 3 implies also that $\widehat{\mathcal{L}}_\beta(u)$ converges to $\mathcal{L}_\beta(u)$ in $L_1(\mathbb{R}_+, \omega)$ where $\omega(u)$ is a bounded nonnegative-valued weight function with $\omega(u) = o(u^{-1-\iota})$ when $u \to \infty$ for arbitrary small $\iota > 0$. This can be used to invert $\widehat{\mathcal{L}}_\beta(u)$, using regularized kernels as those of [20], to estimate nonparametrically the density of the stochastic scale.

3.3. *Long Span Asymptotics: The Case of Estimated Activity.* The asymptotic results in Theorem 3 relied on the premise that β is known. The realized Laplace transform crucially relies on β, as the latter enters not only in its asymptotic limit and variance but also in its construction. If we put a wrong value of β in the calculation of the realized Laplace Transform, then it is easy to see that $\widehat{\mathcal{L}}_\beta(u)$ will converge either to 1 or 0 depending on whether the wrong value is above or below the true one respectively.

In this section we provide asymptotic results for the case where the activity β needs to be estimated from the data. Developing an estimate for β from the high-frequency data is relatively easy (we will give an example at the end of the section). Hence, here we investigate the effect of estimating β on our asymptotic results in Theorem 3.

THEOREM 4. *Suppose there exists an estimator of β, denoted with $\widehat{\beta}$ and assumptions A, B and C hold.*

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

$$\sqrt{T}\left(\widehat{\mathcal{L}}_{\widehat{\beta}}(u) - \widehat{\mathcal{L}}_{\beta}(u)\right) = o_p\left(\frac{1}{\sqrt{T}}\right). \tag{3.14}$$

(b) If $\widehat{\beta}$ 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 $\alpha > 1/(2\beta)$, $\beta' < \beta/2$, $\beta > 4/3$ and $T\Delta_n \to 0$, then we have (locally uniformly in u)

$$\sqrt{T}\left(\widehat{\mathcal{L}}_{\widehat{\beta}}(u) - \widehat{\mathcal{L}}_{\beta}(u)\right) - \frac{\sqrt{T}\log(2u/\Delta_n)\mathbb{E}(G_\beta(u^{1/\beta}|\sigma_t|))}{\beta^2}(\widehat{\beta}-\beta) \xrightarrow{\mathbb{P}} 0, \tag{3.15}$$

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

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

$$\widehat{G}_\beta = \frac{\Delta_n}{T}\sum_{i=1}^{[T/\Delta_n]}\left((2u)^{1/\widehat{\beta}}\Delta_n^{-1/\widehat{\beta}}\Delta_i^n X\right)\sin\left((2u)^{1/\widehat{\beta}}\Delta_n^{-1/\widehat{\beta}}\Delta_i^n X\right) \xrightarrow{\mathbb{P}} \mathbb{E}(G_\beta(u^{1/\beta}|\sigma_t|)). \tag{3.16}$$

Unlike the estimation of $\mathcal{L}_\beta(u)$, 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 will need the relative speed condition $T\Delta_n^\gamma \to 0$ for some $\gamma > 0$ to guarantee that the estimation of β does not have an asymptotic effect on the estimation of the Laplace transform of the stochastic scale. 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 scale. 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 $\mathcal{L}_\beta(u)$. For this, one needs to note that the errors in (3.9) and (3.15) in such case are asymptotically independent.

A more efficient estimator, in the sense of faster rate of convergence, will mean that the approximation error $\widehat{\mathcal{L}}_{\widehat{\beta}}(u)-\widehat{\mathcal{L}}_\beta(u)$ will be smaller asymptotically. Finally, the lower bound on α in part(b) of the above theorem would typically be satisfied when $\beta' < \beta/2$. We finish this section with providing an example of $\sqrt{\Delta_n}$-consistent nonparametric estimator of β (when $\beta' < \beta/2$) developed in [27]. The estimation is based on a ratio of power variations over two time scales for optimally chosen power. It is formally defined as

$$\widehat{\beta} = \frac{\ln(2)\, p^*}{\ln(2) + \ln\left[\Phi_T(X, p^*, 2\Delta_n)\right] - \ln\left[\Phi_T(X, p^*, \Delta_n)\right]}, \tag{3.17}$$

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

$$\Phi_T(X, p, \Delta_n) = \sum_{i=1}^{[T/\Delta_n]} |\Delta_i^n X|^p. \tag{3.18}$$

It is shown in [27] that $\widehat{\beta}$ in (3.17) is $\sqrt{\Delta_n}$-consistent for T fixed with an associated feasible Central Limit Theorem also available.

4. Monte Carlo Assessment. We now examine the properties of the estimators of the Laplace transform both in the case when activity of the driving martingale is known or needs to be estimated from the data, $\widehat{\mathcal{L}}_\beta(u)$ and $\widehat{\mathcal{L}}_{\widehat{\beta}}(u)$ respectively. The Monte Carlo setup is calibrated for a financial price series. In particular, we use $1,000$ Monte Carlo replications of $1,200$ "days" worth of 78 within-day price increments and this corresponds approximately to the span and the sampling frequency of our actual data set in the empirical application. The model used in the Monte Carlo is given by

$$dX_t = V_t^{1/\beta} dL_t, \quad dV_t = 0.02(1.0 - V_t)dt + 0.05\sqrt{V_t}dB_t, \tag{4.1}$$

where L_t is a Lévy process with characteristic triplet $(0, 0, \nu)$ for $\nu(x) = \frac{0.11}{|x|^{1+1.7}}$ or $\nu(x) = \frac{0.11e^{-0.25|x|}}{|x|^{1+1.7}}$. The first choice of the Lévy measure corresponds to that of a stable process with activity index of 1.7 while the second one is that of a tempered stable process with the same value of the activity index. For the second choice of $\nu(x)$, assumption A is satisfied with $\beta' = 0.7$, which indicates a rather active "residual" component in the driving martingale in addition to its stable part. Therefore, the second case represents a very stringent test for the small sample behavior of the RLT.

Table 1 summarizes the outcome of the Monte Carlo experiments. The first two columns of the table report the results for the case when the activity is known and fixed at its true value. In both cases, the estimate is very accurate and virtually unbiased. The third column presents the results for the case when the inference is done with $\beta = 2$ (with L_t being tempered stable) which corresponds to treating erroneously the process X_t as a jump-diffusion. As seen from Table 1 this results in a rather nontrivial upward bias. The reason is that in forming the realized Laplace transform the increments should be inflated by the factor $\Delta_n^{-1/1.7}$ but they are instead inflated by the much smaller $\Delta_n^{-1/2}$. Using the under-inflated increments in the computations induces a very large upward bias in the estimator.

The last two columns of Table 1 summarize the Monte Carlo results for the case where the index β is presumed unknown and estimated using (3.17) based on the first 252 "days" in the simulated data set. As to be be expected, the estimator of the Laplace transform is less accurate than when the activity is known. In the case when the driving martingale is tempered stable, our measure becomes slightly biased due to a small bias in the estimate of the activity level β. These biases however are relatively small when compared with the standard deviation of the estimator.

5. An Application to Parametric Estimation of the Stochastic Scale Law. We apply the preceding theoretical results to define a criterion for parametric estimation based on contrasting our nonparametric realized Laplace transform to that of a parametric model for the stochastic scale (or the time change).

THEOREM 5. *Suppose the conditions of Theorem 3 are satisfied. Let the Laplace transform of $|\sigma_t|^\beta$ be given by $\mathcal{L}_\beta(u;\theta)$ for some finite-dimensional parameter vector lying within a compact set $\theta \in \Theta$ with θ_0 denoting the true value and further assume that $\mathcal{L}_\beta(u;\theta)$ is twice continuously-differentiable in its second argument. If Θ^l is some local neighborhood of θ_0, assume $\sup_{\theta\in\Theta^l}\{|\nabla_\theta \mathcal{L}_\beta(u;\theta)| + \nabla_{\theta\theta'}\mathcal{L}_\beta(u;\theta)|\}$ bounded. Suppose for a kernel function with bounded support $\kappa : \mathbb{R}_+ \to \mathbb{R}_+$ we have that*

$$\int_{\mathbb{R}_+} \left(\mathcal{L}_\beta(u;\theta) - \mathcal{L}_\beta(u;\theta_0)\right)^2 \kappa(u)du > 0, \quad \theta \neq \theta_0.$$

Define the estimator

$$\widehat{\theta} = \underset{\theta\in\Theta}{\operatorname{argmin}} \int_{\mathbb{R}_+} \left(\widehat{\mathcal{L}}_\beta(u) - \mathcal{L}_\beta(u;\theta)\right)^2 \widehat{\kappa}(u)du, \tag{5.1}$$

TABLE 1
Monte Carlo Results

Activity	**S** **Fixed at true value**	**TS** **Fixed at true value**	**TS** **Fixed at** $\beta = 2$	**S** **Estimated**	**TS** **Estimated**
			$\mathcal{L}_\beta(0.10)$		
true value	0.9051	0.9051	0.9051	0.9051	0.9051
mean	0.9052	0.9085	0.9390	0.9057	0.9137
std	0.0063	0.0065	0.0045	0.0074	0.0072
			$\mathcal{L}_\beta(0.50)$		
true value	0.6112	0.6112	0.6112	0.6112	0.6112
mean	0.6111	0.6159	0.7771	0.6141	0.6434
std	0.0208	0.0218	0.0145	0.0292	0.0284
			$\mathcal{L}_\beta(1.25)$		
true value	0.3001	0.3001	0.3001	0.3001	0.3001
mean	0.2998	0.3035	0.5776	0.3050	0.3449
std	0.0249	0.0259	0.0231	0.0383	0.0393
			$\mathcal{L}_\beta(2.50)$		
true value	0.0980	0.0980	0.0980	0.0974	0.0980
mean	0.0977	0.0994	0.3753	0.1024	0.1312
std	0.0158	0.0164	0.0265	0.0261	0.0297
			$\mathcal{L}_\beta(3.75)$		
true value	0.0344	0.0344	0.0344	0.0344	0.0344
mean	0.0342	0.0350	0.2544	0.0374	0.0536
std	0.0084	0.0087	0.0249	0.0142	0.0179

Note: In all simulated scenarios $T = 1,200$ and $[1/\Delta_n] = 78$. The mean and the standard deviation (across the Monte Carlo replications) correspond to the estimator $\widehat{\mathcal{L}}_\beta(u)$ (the first three columns) or $\widehat{\mathcal{L}}_{\widehat{\beta}}(u)$ (the last two columns). The estimator $\widehat{\beta}$ is computed using (3.17) and the first 252 "days" of the sample. V_t has Gamma marginal law with corresponding Laplace transform of $\left(1 + u * 0.05^2/0.04\right)^{-0.04/0.05^2}$. The Monte Carlo replica is 1000.

where $\widehat{\kappa}$ is a nonnegative estimator of κ with $(u^{2+\iota} \vee 1) \sup_{u \in \mathbb{R}_+} |\widehat{\kappa}(u) - \kappa(u)| \xrightarrow{\mathbb{P}} 0$ for some $\iota > 0$. Then for $T \to \infty$ and $T\Delta_n \to 0$, we have

$$\sqrt{T}\left(\widehat{\theta} - \theta_0\right) \xrightarrow{\mathcal{L}-s} \left(\int_{\mathbb{R}_+} \nabla_\theta \mathcal{L}_\beta(u;\theta) \nabla_\theta \mathcal{L}_\beta(u;\theta)' \kappa(u) du\right)^{-1} \Xi^{1/2} E', \tag{5.2}$$

where E is a standard normal vector and

$$\Xi = \int_{\mathbb{R}_+} \int_{\mathbb{R}_+} \Sigma_\beta(u,v) \nabla_\theta \mathcal{L}_\beta(u;\theta) \nabla_\theta \mathcal{L}_\beta(v;\theta)' \kappa(u)\kappa(v) du dv, \tag{5.3}$$

for $\Sigma_\beta(u,v)$ the variance-covariance of Theorem 3.

Remark 4. *There are two types of pure-jump models used in practice. First are the time-changed Lévy processes, see e.g., [11] and [7]. As explained in*

Remark 2, the time-change a_t corresponds to $|\sigma_t|^\beta$ in (1.1). Therefore, $\widehat{\mathcal{L}}_\beta(u)$ provides an estimate of the Laplace transform of the time-change which is modeled directly in parametric settings. The second type of pure-jump models are the ones specified via Lévy-driven SDE. In this case we typically model σ_t and not $|\sigma_t|^\beta$. Therefore, to apply Theorem 5 in this case one will need to evaluate $\mathcal{L}_\beta(u;\theta)$ via simulation. In both cases, the use of RLT simplifies the estimation problem significantly as it preserves information about the stochastic scale and importantly is robust to any dependence between σ_t and Z_t, which particularly in financial applications is rather nontrivial.

The theorem was stated using $\widehat{\mathcal{L}}_\beta(u)$ but obviously the same result will apply if we replace it with $\widehat{\mathcal{L}}_{\widehat{\beta}}(u)$. By way of illustration, we apply the theory to the VIX index computed by the Chicago Board of Options Exchange; the VIX is an option-based measure of market volatility. The data set spans the period from September 22, 2003, until December 31, 2008, for a total of $1,212$ trading days. Within each day, we use 5-minute records of the VIX index corresponding to 78 price observations per day. [28] present nonparametric evidence indicating that the VIX is a pure-jump Itô semimartingale.

The underlying pure-jump model we consider for the log VIX index, denoted by v_t, is

$$dv_t = \alpha_t dt + \int_0^t \int_{\mathbb{R}} \tilde{\mu}(dx, dt),$$

where α_t is the drift term capturing the persistence of v_t, and $\tilde{\mu}$ is a random integer-valued measure that has been compensated by $a_t\, dt \otimes \nu(dx)$ for a_t a stochastic process capturing time varying intensity. The martingale component of v_t is a time-changed Lévy process as in [11]. Recall that the time-change a_t corresponds to $|\sigma_t|^\beta$ in the general model (1.1), and our interest here is in making inferences regarding its marginal distribution.

The parametric specification we use for the marginal distribution of the time-change is that of a tempered stable subordinator [24], which is a self-decomposable distribution, i.e., there is an autoregressive process of order one that generates it [25]. The Laplace transform of the tempered stable is

$$\mathcal{L}(u;\theta) = \begin{cases} \exp\left\{c\Gamma(-\alpha)\left[(\lambda+u)^\alpha - \lambda^\alpha\right]\right\}, & \text{if } \alpha \in (0,1), \\ \left(\frac{1}{1+\frac{u}{\lambda}}\right)^c & \text{if } \alpha = 0. \end{cases} \tag{5.4}$$

where $\theta = (\alpha\, c\, \lambda)$ is the parameter vector, $\Gamma(-\alpha) = -\frac{1}{\alpha}\Gamma(1-\alpha)$ for $\alpha \in (0,1)$, Γ denotes the standard Gamma function, $\alpha \in [0,1)$ can be interpreted as the activity index of the time-change a_t, c is the scale of the marginal distribution of a_t, and λ governs the tail.

To make the estimation feasible, we need an estimate of β and further specify the kernel κ of Theorem 5. For β we use the estimator defined in equation (3.17) over the first year of the sample exactly as in the Monte Carlo work; the point estimate is $\hat{\beta} = 1.862$ with standard error 0.034. We next follow [23] in using a Gaussian kernel $\kappa(u) = \exp(-2u^2/u_{max}^2)$ where u_{max} is defined via $\nabla_u \mathcal{L}_\beta(u_{max}, \theta_0) = -0.05$. The point u_{max} is set so that we collect most of the information available in the empirical Laplace transform. The feasible kernel $\widehat{\kappa}(u)$ is constructed from the infeasible by replacing u_{max} with a consistent estimator for it. It is easy to verify that this choice of the kernels satisfies the conditions of Theorem 5 above.

Table 2 shows the parameter estimates and asymptotic standard errors based on this feasible implementation of (5.1)–(5.3). Interestingly, α is estimated to be below that associated with the Inverse Gaussian ($\alpha = 1/2$), while the estimated tail parameter λ suggests relatively moderate dampening, but this parameter appears somewhat difficult to estimate with high precision given the time span of our data set. Figure 1 shows the fit of the model. Specifically, the heavy solid line is the model-implied Laplace transform evaluated at the estimated parameters. It plots on top of the (not visible) realized Laplace transform, (3.1), and thereby passes directly through the center of the (nonparametric) confidence bands. Overall, the fit of the tempered stable to the marginal law of the time-change is quite tight. From this point, one can follow the strategy of [5] and go further to develop a dynamic model for the time-change by coupling the fitted marginal law with a specification for the memory of the process.

TABLE 2
Estimation Results

Parameter	Estimate	Standard Error
α	0.2651	0.0453
c	1.2872	0.0469
λ	0.0377	0.0103

6. Conclusion. We derive the asymptotic properties of the realized Laplace transform for pure-jump processes computed from high-frequency data. The realized Laplace transform is shown to estimate the Laplace transform of the stochastic scale of the observed process. The results are (locally) uniform over the argument of the transform. We can thereby also derive the asymptotic properties of parameter estimates obtained by fitting parametric models for the marginal law of the stochastic scale to the realized Laplace transform. This estimation entails minimizing a measure of the discrepancy between between the model-implied and observed transforms.

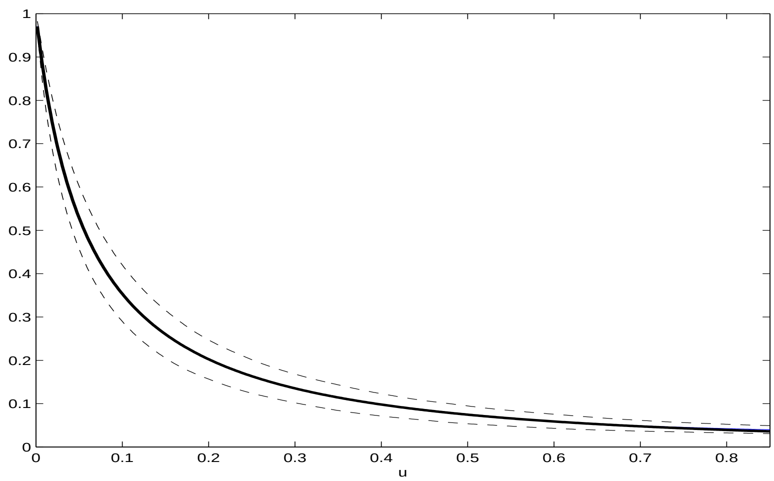

FIG 1. *The solid line is the fitted parametric Laplace transform of the marginal distribution of the scale of the log VIX index, which essentially plots on top of the estimated realized Laplace transform (not visible), and the dashed lines show 95 percent nonparametric confidence intervals about the estimated realized Laplace transform.*

7. Proofs. Here we give the proof of the main results in the paper: Lemma 1 and Theorems 1 and 3 with the rest shown in the supplementary appendix. 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 $\mathbb{E}\left(\cdot|\mathscr{F}_{(i-1)\Delta_n}\right)$. We start with stating some preliminary results, proof of which are in the supplement, and which we use in the proofs of the theorems.

7.1. *Preliminary results.* For a symmetric stable process with Lévy measure $\frac{c}{|x|^{\beta+1}}dx$ for some $c > 0$ and $\beta \in (1, 2)$, using Theorems 14.5 and 14.7 of [25], we can write its characteristic function at time 1 as

$$\exp\left(c\int_0^\infty (e^{iur} - 1 - iur)\frac{dr}{r^{1+\beta}} + c\int_0^\infty (e^{-iur} - 1 + iur)\frac{dr}{r^{1+\beta}}\right), \quad u \in \mathbb{R}.$$

Then using Lemma 14.11 of [25], we can simplify the above expression to

$$\exp\left(2c\Gamma(-\beta)\cos(\beta\pi/2)|u|^\beta\right),$$

where $\Gamma(-\beta) = \frac{\Gamma(2-\beta)}{\beta(\beta-1)}$ for $\beta \in (1, 2)$. Therefore, the Lévy measure of a β-stable process, L_t, with $\mathbb{E}\left(e^{-iuL_t}\right) = e^{-t|u|^\beta/2}$, is

$$A \times \frac{1}{|x|^{\beta+1}}dx,$$

for A defined in (2.3). Throughout, after appropriately extending the original probability space, we will use the following alternative representation of the process X_t (proof of which is given in the supplement)

$$\begin{aligned} X_t = X_0 + \int_0^t \overline{\alpha}_s ds + \int_0^t \int_{\mathbb{R}} \sigma_{s-} x \tilde{\mu}_1(ds,dx) \\ + \int_0^t \int_{\mathbb{R}} \sigma_{s-} x \mu_2(ds,dx) - \int_0^t \int_{\mathbb{R}} \sigma_{s-} x \mu_3(ds,dx) + Y_t, \end{aligned}$$

where μ_1, μ_2 and μ_3 are homogenous Poisson measures (the three measures are not mutually independent) with compensators respectively $\nu_1(dx) = \frac{A}{|x|^{\beta+1}}dx$, $\nu_2(dx) = |\nu'(x)|dx$ and $\nu_3(dx) = 2|\nu'(x)|1(\nu'(x) < 0)\,dx$ and $\overline{\alpha}_s = \alpha_s - \sigma_{s-}\int_{\mathbb{R}} x\nu'(x)dx$. Finally, to simplify notation we will also use the short-hand $L_t = \int_0^t\int_{\mathbb{R}} x\tilde{\mu}_1(ds,dx)$ and further for a symmetric bounded function κ with $\kappa(x) = x$ for x in a neighborhood of zero, we decompose

$$L_t = \overline{L}_t + \widetilde{L}_t, \quad \overline{L}_t = \int_0^t \int_{\mathbb{R}} (x - \kappa(x))\tilde{\mu}_1(ds,dx), \quad \widetilde{L}_t = \int_0^t \int_{\mathbb{R}} \kappa(x)\tilde{\mu}_1(ds,dx). \tag{7.1}$$

With this notation we make the following decomposition

$$V_T(X,\Delta_n,\beta,u) - \int_0^T e^{-u|\sigma_t|^\beta}dt = \sum_{i=1}^{[T/\Delta_n]}\sum_{j=1}^{3} \xi_{i,u}^{(j)} + \int_{[T/\Delta_n]\Delta_n}^T e^{-u|\sigma_t|^\beta}dt, \tag{7.2}$$

$$\begin{aligned} \xi_{i,u}^{(1)} &= \Delta_n \cos\left((2u)^{1/\beta}\sigma_{(i-1)\Delta_n-}\Delta_n^{-1/\beta}\Delta_i^n L\right) - \int_{(i-1)\Delta_n}^{i\Delta_n} e^{-u|\sigma_{(i-1)\Delta_n-}|^\beta}ds, \\ \xi_{i,u}^{(2)} &= \int_{(i-1)\Delta_n}^{i\Delta_n} \left(e^{-u|\sigma_{(i-1)\Delta_n-}|^\beta} - e^{-u|\sigma_s|^\beta}\right)ds, \\ \xi_{i,u}^{(3)} &= \Delta_n\left(\cos\left((2u)^{1/\beta}\Delta_n^{-1/\beta}\Delta_i^n X\right) - \cos\left((2u)^{1/\beta}\sigma_{(i-1)\Delta_n-}\Delta_n^{-1/\beta}\Delta_i^n L\right)\right). \end{aligned}$$

Starting with $\xi_{i,u}^{(1)}$, using the self-similarity of the stable process L_t, and the expression for its characteristic function, we have

$$\begin{cases} \mathbb{E}_{i-1}^n\left(\cos\left((2u)^{1/\beta}\sigma_{(i-1)\Delta_n-}\Delta_n^{-1/\beta}\Delta_i^n L\right) - e^{-u|\sigma_{(i-1)\Delta_n-}|^\beta}\right) = 0, \\ \mathbb{E}_{i-1}^n\left(\cos\left((2u)^{1/\beta}\sigma_{(i-1)\Delta_n-}\Delta_n^{-1/\beta}\Delta_i^n L\right) - e^{-u|\sigma_{(i-1)\Delta_n-}|^\beta}\right)^2 = F_\beta(u^{1/\beta}|\sigma_{(i-1)\Delta_n-}|), \\ \mathbb{E}_{i-1}^n\left(\cos\left((2u)^{1/\beta}\sigma_{(i-1)\Delta_n-}\Delta_n^{-1/\beta}\Delta_i^n L\right) - e^{-u|\sigma_{(i-1)\Delta_n-}|^\beta}\right)^4 \leq C. \end{cases} \tag{7.3}$$

Using first-order Taylor expansion we decompose $\xi_{i,u}^{(2)} = \sum_{j=1}^{3} \xi_{i,u}^{(2)}(j)$ where

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

$$\xi_{i,u}^{(2)}(2) = \int_{(i-1)\Delta_n}^{i\Delta_n} \left(\Upsilon(\sigma_s^*, u) - \Upsilon(\sigma_{(i-1)\Delta_n-,u}) \right) \Bigg(\int_{(i-1)\Delta_n}^{s} \tilde{\sigma}_u dW_u$$
$$+ \int_{(i-1)\Delta_n}^{s} \int_{\mathbb{R}^2} \underline{\delta}(u-, x) \underline{\tilde{\mu}}(du, dx) \Bigg) ds,$$

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

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

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

We derive the following bounds in the supplement for any finite $\overline{u} > 0$

$$\text{(7.4)} \quad \sum_{i=1}^{[T/\Delta_n]} \mathbb{E}_{i-1}^n \left(\xi_{i,u}^{(2)}(1) \right) = 0, \qquad \frac{\Delta_n^{-2}}{T} \mathbb{E} \left(\sum_{i=1}^{[T/\Delta_n]} \mathbb{E}_{i-1}^n \left(\xi_{i,u}^{(2)}(1) \right)^2 \right) \leq C,$$

$$\text{(7.5)} \quad (T\Delta_n^{\beta/2})^{-1} \sum_{i=1}^{[T/\Delta_n]} \mathbb{E} \left(\sup_{0 \leq u \leq \overline{u}} |\xi_{i,u}^{(2)}(2)| \right) \leq C,$$

$$\text{(7.6)} \quad (T\Delta_n)^{-1} \sum_{i=1}^{[T/\Delta_n]} \mathbb{E} \left(\sup_{0 \leq u \leq \overline{u}} |\xi_{i,u}^{(2)}(3)| \right) \leq C.$$

Turning to $\xi_{i,u}^{(3)}$, we can first make the following decomposition (recall the decomposition of L_t in (7.1))

$$\text{(7.7)} \qquad \cos(\chi_1) - \cos(\chi_5) = \sum_{j=1}^{4} [\cos(\chi_j) - \cos(\chi_{j+1})],$$

$$\chi_1 = (2u)^{1/\beta} \Delta_n^{-1/\beta} \Delta_i^n X, \quad \chi_2 = (2u)^{1/\beta} \Delta_n^{-1/\beta} \left(\int_{(i-1)\Delta_n}^{i\Delta_n} \overline{\alpha}_s ds + \int_{(i-1)\Delta_n}^{i\Delta_n} \sigma_{s-} dL_s \right),$$

$$\chi_3 = (2u)^{1/\beta} \Delta_n^{-1/\beta} \left(\Delta_n \overline{\alpha}_{(i-1)\Delta_n} + \int_{(i-1)\Delta_n}^{i\Delta_n} \sigma_{s-} d\widetilde{L}_s \right),$$

$$\chi_4 = (2u)^{1/\beta} \Delta_n^{-1/\beta} \sigma_{(i-1)\Delta_n-} \Delta_i^n \widetilde{L}, \quad \chi_5 = (2u)^{1/\beta} \Delta_n^{-1/\beta} \sigma_{(i-1)\Delta_n-} \Delta_i^n L.$$

Then, using the formula for $\cos(x) - \cos(y)$, $x, y \in \mathbb{R}$ for the first bracketed term on the right side of (7.7) and a second-order Taylor expansion for the third one, allows us to write $\xi^{(3)}_{i,u} = \sum_{j=1}^5 \xi^{(3)}_{i,u}(j)$ where

$$\xi^{(3)}_{i,u}(1) = -2\Delta_n \sin\left(0.5(2u)^{1/\beta}\Delta_n^{-1/\beta}\left(\Delta_i^n X + \int_{(i-1)\Delta_n}^{i\Delta_n}\overline{\alpha}_s ds + \int_{(i-1)\Delta_n}^{i\Delta_n}\sigma_{s-}dL_s\right)\right)$$
$$\times \sin\left(0.5(2u)^{1/\beta}\Delta_n^{-1/\beta}\left(\Delta_i^n X - \int_{(i-1)\Delta_n}^{i\Delta_n}\overline{\alpha}_s ds - \int_{(i-1)\Delta_n}^{i\Delta_n}\sigma_{s-}dL_s\right)\right),$$
$$\xi^{(3)}_{i,u}(2) = -(2u)^{1/\beta}\Delta_n^{2-1/\beta}\sin\left((2u)^{1/\beta}\sigma_{(i-1)\Delta_n-}\Delta_n^{-1/\beta}\Delta_i^n\widetilde{L}\right)\overline{\alpha}_{(i-1)\Delta_n},$$
$$\xi^{(3)}_{i,u}(3) = -(2u)^{1/\beta}\Delta_n^{1-1/\beta}\sin\left((2u)^{1/\beta}\sigma_{(i-1)\Delta_n-}\Delta_n^{-1/\beta}\Delta_i^n\widetilde{L}\right)\int_{(i-1)\Delta_n}^{i\Delta_n}(\sigma_{s-} - \sigma_{(i-1)\Delta_n-})d\widetilde{L}_s,$$
$$\xi^{(3)}_{i,u}(4) = \Delta_n[\cos(\chi_2) - \cos(\chi_3)] + \Delta_n[\cos(\chi_4) - \cos(\chi_5)],$$
$$\xi^{(3)}_{i,u}(5) = 0.5(2u)^{2/\beta}\Delta_n^{1-2/\beta}\cos(\tilde{\chi})\left(\Delta_n\overline{\alpha}_{(i-1)\Delta_n} + \int_{(i-1)\Delta_n}^{i\Delta_n}(\sigma_{s-} - \sigma_{(i-1)\Delta_n-})d\widetilde{L}_s\right)^2,$$

with $\tilde{\chi}$ denoting some value between χ_3 and χ_4.

We derive the following bounds in the supplement for any finite $\overline{u} > 0$

$$(7.8)\quad (T|\log(\Delta_n)|\Delta_n^{1-\beta'/\beta})^{-1}\sum_{i=1}^{[T/\Delta_n]}\mathbb{E}\left(\sup_{0\le u\le\overline{u}}|\xi^{(3)}_{i,u}(1)|\right) \le C.$$

$$(7.9)\quad \mathbb{E}^n_{i-1}\left(\xi^{(3)}_{i,u}(2)\right) = 0, \quad \frac{(\Delta_n^{3-2/\beta})^{-1}}{T}\mathbb{E}\left(\sum_{i=1}^{[T/\Delta_n]}\mathbb{E}^n_{i-1}\left(\xi^{(3)}_{i,u}(2)|\right)^2\right) \le C.$$

$$(7.10)\quad (T\Delta_n^{3/2-1/\beta})^{-1}\sum_{i=1}^{[T/\Delta_n]}\mathbb{E}\left(\sup_{0\le u\le\overline{u}}|\xi^{(3)}_{i,u}(4)|\right) \le C.$$

$$(7.11)\quad (T\Delta_n^{2-2/\beta})^{-1}\sum_{i=1}^{[T/\Delta_n]}\mathbb{E}\left(\sup_{0\le u\le\overline{u}}|\xi^{(3)}_{i,u}(5)|\right) \le C.$$

$$(7.12)\quad (T\Delta_n^{1-\iota})^{-1}\mathbb{E}\left(\sup_{0\le u\le\overline{u}}\left|\sum_{i=1}^{[T/\Delta_n]}\mathbb{E}^n_{i-1}\xi^{(3)}_{i,u}(3)\right|\right) \le C, \quad \text{under B}',$$

$$(7.13)\quad \left(T\Delta_n^{1/(\beta\vee\beta''+\iota)}\right)^{-1}\mathbb{E}\left(\sup_{0\le u\le\overline{u}}\left|\sum_{i=1}^{[T/\Delta_n]}\mathbb{E}^n_{i-1}\xi^{(3)}_{i,u}(3)\right|\right) \le C, \quad \text{under B}.$$

$$(7.14)\quad (T\Delta_n^{3-2/\beta})^{-1}\sum_{i=1}^{[T/\Delta_n]}\mathbb{E}\left(\sup_{0\le u\le\overline{u}}|\xi^{(3)}_{i,u}(3)|^2\right) \le C.$$

□

7.2. *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. [16], Corollary VII.3.6), i.e., we need to establish the following for $h \to 0$

$$\left\{ \begin{array}{l} h^{1-2/\beta}\int_{\mathbb{R}} \kappa^2(h^{-1/\beta}x)\nu(x)dx \longrightarrow \int_{\mathbb{R}} \kappa^2(x)\frac{A}{|x|^{\beta+1}}dx, \\ h\int_{\mathbb{R}} g(h^{-1/\beta}x)\nu(x)dx \longrightarrow \int_{\mathbb{R}} g(x)\frac{A}{|x|^{\beta+1}}dx, \end{array} \right. \tag{7.15}$$

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

The result in (7.15) follows by a change of variable in the integration and using the fact that by assumption A we have $|\nu'(x)| < \frac{C}{|x|^{\beta'+1}}$ for $|x| \le x_0$ where x_0 is fixed and $\beta' < \beta$. □

7.3. *Proof of Theorem 1.* Part (b) of the theorem holds from the bounds in (7.3)-(7.6) and (7.8)-(7.14), so we are left with showing part(a). First, we show that for $\Delta_n \to 0$, $\frac{1}{\sqrt{\Delta_n}}\sum_{i=1}^{[t/\Delta_n]}\xi_{i,u}^{(1)}$ converges stably as a process in t for the Skorokhod topology to the process $\int_0^t\sqrt{F_\beta(u^{1/\beta}|\sigma_s|)}dW'_s$ where W'_t is a Brownian motion defined on an extension of the original probability space and independent from the σ-field $\mathcal{F}$. Using the result in (7.3), we get for every $t > 0$

$$\left\{ \begin{array}{l} \frac{1}{\sqrt{\Delta_n}}\sum_{i=1}^{[t/\Delta_n]}\mathbb{E}_{i-1}^n(\xi_{i,u}^{(1)}) \stackrel{\mathbb{P}}{\longrightarrow} 0, \quad \frac{1}{\Delta_n}\sum_{i=1}^{[t/\Delta_n]}\mathbb{E}_{i-1}^n(\xi_{i,u}^{(1)})^2 \stackrel{\mathbb{P}}{\longrightarrow} \int_0^t F_\beta(u^{1/\beta}|\sigma_s|)ds, \\ \frac{1}{\Delta_n^2}\sum_{i=1}^{[t/\Delta_n]}\mathbb{E}_{i-1}^n(\xi_{i,u}^{(1)})^4 \stackrel{\mathbb{P}}{\longrightarrow} 0, \end{array} \right.$$

where for the second convergence above we made use of Riemann integrability. Thus to show the stable convergence, given the above result and upon using Theorem IX.7.28 of [16], we need to show only

$$\sum_{i=1}^{[t/\Delta_n]}\mathbb{E}_{i-1}^n\left(\sqrt{1/\Delta_n}\xi_{i,u}^{(1)}\Delta_i^nM\right) \stackrel{\mathbb{P}}{\longrightarrow} 0, \quad \forall t > 0, \tag{7.16}$$

where M is a bounded martingale defined on the original probability space.

When M is discontinuous martingale we can argue as follows. First, we can set $M_t^n = M_{[t/\Delta_n]\Delta_n}$ and $N_t^n = \sum_{i=1}^{[t/\Delta_n]}\sqrt{1/\Delta_n}\xi_{i,u}^{(1)}$ for any t. With this notation we have $[M^n, N^n]_t = \sum_{i=1}^{[t/\Delta_n]}\left(\sqrt{1/\Delta_n}\xi_{i,u}^{(1)}\Delta_i^nM\right)$ and $\langle M^n, N^n\rangle_t = \sum_{i=1}^{[t/\Delta_n]}\mathbb{E}_{i-1}^n\left(\sqrt{1/\Delta_n}\xi_{i,u}^{(1)}\Delta_i^nM\right)$. We trivially have that M^n converges (for the Skorokhod topology) to M, and furthermore from the results above the limit of N^n, which we denote here with N, is a continuous process. Therefore, using VI.3.33 (b) of [16], we have that (M^n, N^n) is tight. Then, using the fact that M is a bounded martingale (and hence it has bounded jumps), we can apply VI.6.29 of [16] and conclude that the limit of $[M^n, N^n]$ (up to taking a subsequence) is $[M, N]$. However, since continuous and pure-jump martingales are orthogonal, see e.g., I.4.11 of [16], we conclude that $[M, N] = 0$. Further, the difference $[M^n, N^n] - \langle M^n, N^n\rangle$ is a martingale and using Itô isometry, the fact that $\sqrt{1/\Delta_n}\xi^{(1)}_{i,u} \leq C\sqrt{\Delta_n}$, and the boundedness of M, we have

$$\begin{aligned}\mathbb{E}\left([M^n, N^n]_t - \langle M^n, N^n\rangle_t\right)^2 &= \mathbb{E}\left(\sum_{s\leq t}(\Delta M^n_s \Delta N^n_s)^2\right)\\ &\leq C\Delta_n \mathbb{E}\left(\sum_{s\leq t}(\Delta M^n_s)^2\right) \leq C\Delta_n.\end{aligned}$$

Therefore, $[M^n, N^n] - \langle M^n, N^n\rangle$ converges in probability to zero and hence so does $\langle M^n, N^n\rangle$.

When M is a continuous martingale, we can write $\mathbb{E}^n_{i-1}\left(\xi^{(1)}_{i,u}\Delta^n_i M\right) = \mathbb{E}^n_{i-1}\left(\Delta^n_i N \Delta^n_i M\right)$ where now we denote $N_t = \mathbb{E}(\xi^{(1)}_{i,u}|\mathscr{F}_t)$ for $t \in [(i-1)\Delta_n, i\Delta_n]$ (which is obviously a martingale with respect to the filtration $\mathscr{F}_t$). However, note that $\xi^{(1)}_{i,u}$ is uniquely determined by $\mathscr{F}_{(i-1)\Delta_n}$ and the homogenous Poisson measure μ_1. Therefore, N_t remains a martingale for the coarser filtration $\mathscr{F}^*_t = \mathscr{F}_{(i-1)\Delta_n} \cap \mathscr{F}^{\mu_1}_t$ for $\mathscr{F}^{\mu_1}_t$ being 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 $\mathscr{F}^*_t$ (note μ_1 is a homogenous Poisson measure), Theorem III.4.34 of [16], we can represent N_t as a sum of $\mathscr{F}_{(i-1)\Delta_n}$-adapted variable and an integral with respect to $\tilde{\mu}_1$. But then since pure-jump and continuous martingales are orthogonal, we have $\mathbb{E}^n_{i-1}\left(\xi^{(1)}_{i,u}\Delta^n_i M\right) = 0$.

This establishes the stable convergence of $\frac{1}{\sqrt{\Delta_n}}\sum_{i=1}^{[t/\Delta_n]}\xi^{(1)}_{i,u}$. Next, the bounds for $\xi^{(2)}_{i,u}$ and $\xi^{(3)}_{i,u}$ in (7.4)-(7.6) and (7.8)-(7.11) imply that $\frac{1}{\sqrt{\Delta_n}}\sum_{i=1}^{[T/\Delta_n]}(\xi^{(2)}_{i,u} + \xi^{(3)}_{i,u})$ is asymptotically negligible for $\Delta_n \to 0$ and T fixed. □

7.4. *Proof of Theorem 3.* **Part (a).** The proof consists of showing finite-dimensional convergence in u and tightness of the sequence.

(1) Finite-dimensional Convergence. First, given assumption C and using a CLT for stationary and ergodic process, see [16], Theorem VIII.3.79, we have for a finite-dimensional vector $\mathbf{u}$

$$\sqrt{T}\left(\frac{1}{T}\int_0^T e^{-\mathbf{u}|\sigma_t|^\beta}dt - \mathcal{L}_\beta(\mathbf{u})\right) \xrightarrow{\mathcal{L}} \Psi, \tag{7.17}$$

where Ψ is a zero-mean normal variable with elements of the variance-covariance matrix given by $\Sigma_\beta(u_i, u_j)$.

Next, the results in Section 7.1 imply for $T \to \infty$ and $\Delta_n \to 0$ under the weaker assumption B

$$\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{i=1}^{[T/\Delta_n]}\xi_{i,u}^{(1)} = o_p(\sqrt{\Delta_n}), \quad \frac{1}{\sqrt{T}}\sum_{i=1}^{[T/\Delta_n]}\xi_{i,u}^{(2)} = o_p(\sqrt{T}\Delta_n^{\beta/2} \vee \sqrt{\Delta_n}),\\
&\frac{1}{\sqrt{T}}\sum_{i=1}^{[T/\Delta_n]}\xi_{i,u}^{(3)} = o_p\left(\sqrt{T}(|\log(\Delta_n)|\Delta_n^{1-\beta'/\beta} \vee \Delta_n^{(2-2/\beta)\wedge 1/2}) \vee \sqrt{\Delta_n}\right),
\end{aligned} \tag{7.18}$$

with the last one replaced with the weaker

$$\frac{1}{\sqrt{T}}\sum_{i=1}^{[T/\Delta_n]}\xi_{i,u}^{(3)} = o_p\left(\sqrt{T}(|\log(\Delta_n)|\Delta_n^{1-\beta'/\beta} \vee \Delta_n^{2-2/\beta}) \vee \sqrt{\Delta_n}\right),$$

when the stronger assumption B$'$ holds.

(2) Tightness. Lets denote for arbitrary $u, v \geq 0$:

$$z_t = (e^{-u|\sigma_t|^\beta} - \mathcal{L}_\beta(u)) - (e^{-v|\sigma_t|^\beta} - \mathcal{L}_\beta(v)).$$

Then, using successive conditioning and Lemma VIII.3.102 in [16], together with the boundedness of z_t and assumption C, we get

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

where $\iota > 0$ is the constant of assumption C and $p > 3$. Using Theorem 12.3 of [8], the above bound implies the tightness of the sequence

$\frac{1}{\sqrt{T}}\int_0^T(e^{-u|\sigma_t|^\beta} - \mathcal{L}_\beta(u))dt$, and from here we have its convergence for the local uniform topology.

Turning now to $\frac{1}{\sqrt{\Delta_n}}\sum_{i=1}^{[T/\Delta_n]}\xi_{i,u}^{(1)}$, we can use the analogue of the result in (7.3) for $\xi_{i,u}^{(1)} - \xi_{i,v}^{(1)}$, to get

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

for some constant C. From here using Theorem 12.3 of [8], we get the tightness of $\frac{1}{\sqrt{T\Delta_n}}\sum_{i=1}^{[T/\Delta_n]}\xi_{i,u}^{(1)}$.

Similarly, using the analogue of (7.9) applied to $\xi_{i,u}^{(3)}(2) - \xi_{i,v}^{(3)}(2)$, we have

$$\frac{\Delta_n^{-(3-2/\beta)}}{T}\mathbb{E}\left(\sum_{i=1}^{[T/\Delta_n]}\left(\xi_{i,u}^{(3)}(2) - \xi_{i,v}^{(3)}(2)\right)\right)^2 \le C|u^{1/\beta} - v^{1/\beta}|^2. \tag{7.19}$$

This establishes tightness for $\Delta_n^{-(3/2-1/\beta)}\sum_{i=1}^{[T/\Delta_n]}\xi_{i,u}^{(3)}(2)$. We can do exactly the same for $\Delta_n^{-(3/2-1/\beta)}\sum_{i=1}^{[T/\Delta_n]}\left(\xi_{i,u}^{(3)}(3) - \mathbb{E}_{i-1}^n(\xi_{i,u}^{(3)}(3))\right)$ using the analogue of (7.14) applied to $\xi_{i,u}^{(3)}(3) - \xi_{i,v}^{(3)}(3) - \mathbb{E}_{i-1}^n(\xi_{i,u}^{(3)}(3) - \xi_{i,v}^{(3)}(3))$. Next,

$$\frac{\Delta_n^{-2}}{T}\mathbb{E}\left(\sum_{i=1}^{[T/\Delta_n]}\left(\xi_{i,u}^{(2)}(1) - \xi_{i,v}^{(2)}(1)\right)\right)^2 \le C\,(u-v)^2, \tag{7.20}$$

where we used successive conditioning and further made use of the inequality

$$|\Upsilon(x,u) - \Upsilon(x,v)| \le C|x|^{\beta-1}|u-v|, \quad x\in\mathbb{R},\ u,v\ge 0,$$

which follows from applying first-order Taylor expansion of $\Upsilon(x,u)$ in its second argument and using the fact that the derivative of $\Upsilon(x,u)$ in its second argument is bounded by $C|x|^{\beta-1}$. Therefore, $\sum_{i=1}^{[T/\Delta_n]}\frac{\Delta_n^{-1}}{\sqrt{T}}\xi_{i,u}^{(2)}(1)$ is tight on the space of continuous functions equipped with the local uniform topology.

Next, using the results in Section 7.1, it is easy to show that for any finite $\overline{u} > 0$ we have

$$\lim_{\Delta_n\downarrow 0, T\uparrow\infty}\mathbb{P}\left(\sup_{0\le u\le\overline{u}}\left|\sum_{i=1}^{[T/\Delta_n]}(T\Delta_n^{\beta/2})^{-1}\xi_{i,u}^{(2)}(2)\right| > \epsilon_n\right) = 0, \quad \forall\epsilon_n\uparrow\infty. \tag{7.21}$$

The same holds when in the above we replace $(T\Delta_n^{\beta/2})^{-1}\xi_{i,u}^{(2)}(2)$ with either of the following terms: $(T\Delta_n)^{-1}\xi_{i,u}^{(2)}(3)$, $(T|\log(\Delta_n)|\Delta_n^{1-\beta'/\beta})^{-1}\xi_{i,u}^{(3)}(1)$, $(T\Delta_n^{3/2-1/\beta})^{-1}\xi_{i,u}^{(3)}(4)$, $(T\Delta_n^{2-2/\beta})^{-1}\xi_{i,u}^{(3)}(5)$ as well as $(T\Delta_n^{1-\iota})^{-1}\mathbb{E}_{i-1}^n(\xi_{i,u}^{(3)}(3))$ under assumption B′ and $(T\Delta_n^{1/(\beta\vee\beta''+\iota)})^{-1}\mathbb{E}_{i-1}^n(\xi_{i,u}^{(3)}(3))$ under the weaker assumption B. This implies that those terms are uniformly in u bounded in probability.

Part (b). First, (3.12) follows directly from Theorem 1, so here we only show (3.13). If we denote for $k \geq 0$

$$C_{k,\beta}(u,v) = \frac{1}{T}\sum_{t=k+1}^{T}\int_{t-1}^{t}\left(e^{-u|\sigma_s|^\beta} - \mathcal{L}_\beta(u)\right)ds\int_{t-k-1}^{t-k}\left(e^{-v|\sigma_s|^\beta} - \mathcal{L}_\beta(v)\right)ds,$$

then under our assumptions, by standard arguments, see e.g., Proposition 1 in [4], we have

$$C_{0,\beta}(u,v) + \sum_{i=1}^{L_T}\omega(i,L_T)(C_{i,\beta}(u,v) + C_{i,\beta}(v,u)) \xrightarrow{\mathbb{P}} \Sigma_\beta(u,v). \tag{7.22}$$

Therefore, we are left showing

$$\begin{aligned}&(\widehat{C}_{0,\beta}(u,v) - C_{0,\beta}(u,v))\\ &\qquad + \sum_{i=1}^{L_T}\omega(i,L_T)(\widehat{C}_{i,\beta}(u,v) + \widehat{C}_{i,\beta}(v,u) - C_{i,\beta}(u,v) - C_{i,\beta}(v,u)) \xrightarrow{\mathbb{P}} 0.\end{aligned} \tag{7.23}$$

We note that for arbitrary $1 \leq k \leq T$ we have:

$$\Delta_n\sum_{i=[(k-1)/\Delta_n]+1}^{[k/\Delta_n]}\cos((2u)^{1/\beta}\Delta_n^{-1/\beta}\Delta_i^n X) \leq 1, \quad \text{and} \quad \int_{k-1}^{k} e^{-u|\sigma_s|^\beta}ds \leq 1.$$

Hence, for $k = 0, 1, ..., L_T$, we have

$$\begin{aligned}\left|\widehat{C}_{k,\beta}(u,v) - C_{k,\beta}(u,v)\right| &\leq \frac{1}{T}\sum_{t=1}^{T}\left|\widehat{Z}_{t,\beta}(u) - \int_{t-1}^{t}e^{-u|\sigma_s|^\beta}ds\right|\\ &+ \frac{1}{T}\sum_{t=1}^{T}\left|\widehat{Z}_{t,\beta}(v) - \int_{t-1}^{t}e^{-v|\sigma_s|^\beta}ds\right| + \left|\widehat{\mathcal{L}}_\beta(u) - \mathcal{L}_\beta(u)\right| + \left|\widehat{\mathcal{L}}_\beta(v) - \mathcal{L}_\beta(v)\right| + O\left(\frac{k}{T}\right).\end{aligned}$$

First, using the CLT result in (7.17), and since $L_T/\sqrt{T} \to 0$, we have

$$\sum_{i=1}^{L_T} |\omega(i, L_T)| \left| \widehat{\mathcal{L}}_\beta(u) - \mathcal{L}_\beta(u) \right| \stackrel{\mathbb{P}}{\longrightarrow} 0, \quad \forall u > 0. \tag{7.24}$$

Further, using the stationarity of the process σ_t and the bounds on the moments of the terms $\xi_{i,u}^{(j)}$ derived in Section 7.1, we have for every $t \geq 1$

$$\mathbb{E} \left| \widehat{Z}_{t,\beta}(u) - \int_{t-1}^{t} e^{-u|\sigma_s|^\beta} ds \right| \leq C \left(|\log \Delta_n| \Delta_n^{1-\beta'/\beta} \vee \Delta_n^{(2-2/\beta) \wedge 1/2} \right).$$

Therefore, using the relative speed condition between L_T and Δ_n in the theorem, we have

$$\frac{\sum_{i=1}^{L_T} |\omega(i, L_T)|}{T} \sum_{t=1}^{T} \mathbb{E} \left| \widehat{Z}_{t,\beta}(u) - \int_{t-1}^{t} e^{-u|\sigma_s|^\beta} ds \right| \longrightarrow 0, \quad \forall u > 0. \tag{7.25}$$

(7.24) and (7.25) imply (7.23) and this combined with (7.22) establishes the result in (3.13). □

Acknowledgements. Research partially supported by NSF Grant SES-0957330. We would like to thank the editor, an associate editor and two anonymous referees for many constructive comments which lead to significant improvements.

References.

[1] Ait-Sahalia, Y. and J. Jacod (2007). Volatility Estimators for Discretely Sampled Levy Processes. *Annals of Statistics*, 355–392.

[2] 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.

[3] Andrews, B., M. Calder, and R. Davis (2009). Maximum Likelihood Estimation of α-stable Autoregressive Processes. *Annals of Statistics 37*, 1946–1982.

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

[5] 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.

[6] Belomestny, D. (2010). Spectral Estimation of the Fractional Order of a Levy Process. *Annals of Statistics 38*, 317–351.

[7] Belomestny, D. (2011). Statistical Inference for Time-Changed Levy Processes via Composite Characteristic Function Estimation. *Annals of Statistics, forthcoming*.

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

[9] Boyarchenko, S. and S. Levendorskii (2002). Barrier Options and Touch-and-Out Options under Regular Levy Processes of Exponential Type. *Annals of Applied Probability 12*, 1261–1298.

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

[11] Carr, P., H. Geman, D. Madan, and M. Yor (2003). Stochastic Volatility for Lévy Processes. *Mathematical Finance 13*, 345–382.

[12] Duffie, D., D. Filipović, and W. Schachermayer (2003). Affine Processes and Applications in Finance. *Annals of Applied Probability 13(3)*, 984–1053.

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

[14] Jacod, J. (1979). *Calcul Stochastique et Problèmes de Martingales.* Lecture notes in Mathemtatics 714. Berlin Heidelberg New York: Springer-Verlag.

[15] Jacod, J. and P. Protter (2012). *Discretization of Processes.* Springer-Verlag.

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

[17] Klüppelberg, C., A. Lindner, and R. Maller (2004). A Continuous Time GARCH Process Driven by a Lévy Process: Stationarity and Second Order Behavior. *Journal of Applied Probability 41*, 601–622.

[18] Klüppelberg, C., T. Meyer-Brandis, and A. Schmidt (2010). Electricity Spot Price Modelling with a View Towards Extreme Spike Risk. *Quantitative Finance 10*, 963–974.

[19] Koutrouvelis, I. (1980). Regression-type Estimation of the Parameters of Stable Laws. *Journal of the American Statistical Association 75*, 918–928.

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

[21] Mikosch, T., S. Resnick, H. Rootzen, and A. Stegeman (2002). Is Network Traffic Approximated by Stable Levy Motion or Fractional Brownian Motion? *Annals of Applied Probability 12*, 23–68.

[22] Neumann, M. and M. Reiss (2009). Nonparametric Estimation for Lévy Processes from Low-Frequency Observations. *Bernoulli 15*, 223–248.

[23] Paulson, E. W., R. Holcomb, and A. Leitch (1975). The Estimation of the Parameters of the Stable Laws. *Biometrika 62*, 163–170.

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

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

[26] Shimizu, Y. and N. Yoshida (2006). Estimation of Parameters for Diffusion Processes with Jumps from Discrete Observations. *Statistical Inference for Stochastic Processes 9*, 227–277.

[27] Todorov, V. and G. Tauchen (2011a). Limit Theorems for Power Variations of Pure-Jump Processes with Application to Activity Estimation. *Annals of Applied Probability 21*, 546–588.

[28] Todorov, V. and G. Tauchen (2011b). Volatility Jumps. *Journal of Business and Economic Statistics 29*, 356–371.

[29] Todorov, V. and G. Tauchen (2012). The Realized Laplace Transform of Volatility. *Econometrica, forthcoming*.

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

Department of Finance
Northwestern University
Evanston, IL 60208-2001
E-mail: v-todorov@northwestern.edu

Department of Economics
Duke University
Durham, NC 27708-0097
E-mail: george.tauchen@duke.edu