JUMP ACTIVITY ESTIMATION FOR PURE-JUMP SEMIMARTINGALES VIA SELF-NORMALIZED STATISTICS

By Viktor Todorov

Northwestern University

We derive a nonparametric estimator of the jump-activity index β of a "locally-stable" pure-jump Itô semimartingale from discrete observations of the process on a fixed time interval with mesh of the observation grid shrinking to zero. The estimator is based on the empirical characteristic function of the increments of the process scaled by local power variations formed from blocks of increments spanning shrinking time intervals preceding the increments to be scaled. The scaling serves two purposes: (1) it controls for the time variation in the jump compensator around zero and (2) it ensures self-normalization, i.e., that the limit of the characteristic function based estimator converges to a non-degenerate limit which depends only on β. The proposed estimator leads to nontrivial efficiency gains over existing estimators based on power variations. In the Lévy case, the asymptotic variance decreases multiple times for higher values of β. The limiting asymptotic variance of the proposed estimator, unlike that of the existing power variation based estimators, is constant. This leads to further efficiency gains in the case when the characteristics of the semimartingale are stochastic. Finally, in the limiting case of $\beta = 2$, which corresponds to jump-diffusion, our estimator of β can achieve a faster rate than existing estimators.

1. Introduction. In this paper we are interested in estimating the jump activity index of a process defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and given by

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

when L is locally stable pure-jump Lévy process (i.e., a pure-jump Lévy process whose Lévy measure around zero behaves like that of a stable process) and Y is a pure-jump process which is "dominated" at high-frequencies by L in a sense which is made precise below (see Assumption A). All formal conditions for X are given in Section 2. The jump activity index of X on a given fixed time interval is the infimum of the set of powers p for which the

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

Keywords and phrases: central limit theorem, high-frequency data, Itô semimartingale, jumps, jump activity index, stochastic volatility, power variation.

sum of p-th absolute moments of the jumps is finite. Provided σ does not vanish on the interval and has càdlàg paths, the jump activity index of X coincides with the Blumenthal-Getoor index of the driving Lévy process L (recall Y is dominated by L at high frequencies). The dominant role of L at high frequencies, together with its stable-like Lévy measure around zero, manifests into the following limiting behavior at high frequencies

$$h^{-1/\beta}(X_{t+sh} - X_t) \xrightarrow{\mathcal{L}} \sigma_t \times (S_{t+s} - S_t), \qquad \text{as } h \to 0 \text{ and } s \in [0,1], \tag{1.2}$$

for every t and where S is β-stable process, with the convergence being for the Skorokhod topology. (1.2) holds when $\beta > 1$ which is the case we consider in this paper (when $\beta < 1$ the drift will be the "dominant" component at high-frequencies and some of our results can be extended to this case as well). We study estimation of β from discrete equidistant observations of X on a fixed time interval with mesh of the observation grid shrinking to zero.

Estimation of the jump activity index has received a lot of attention recently. [20] consider estimation from low frequency observations in the setting of Lévy processes. [4] and [6] consider estimation from low-frequency data in the setting of time-changed Lévy processes with an independent time-change process. [2] consider estimation from low-frequency and options data. [3] and [5] consider estimation from low frequency data in certain stochastic volatility models. [27–29] propose estimation from high-frequency data using power variations in a pure-jump setting. [1] and [16] consider estimation in high-frequency setting when the underlying process can contain a continuous martingale via truncated power variations. [23] propose estimation of the jump activity index in pure-jump setting via power variations with adaptively chosen optimal power. [22] extend [23] via power variations of differenced increments which provide further robustness and efficiency gains. [15] consider jump activity estimation from noisy high-frequency data.

The estimation of β from high frequency data, thus far, makes use of the dependence of the scaling factor of the high-frequency increments in (1.2) on β. For example, consider the power variation

$$V(p, \Delta_n) = \sum_{i=1}^{n} |\Delta_i^n X|^p, \quad \Delta_i^n X = X_{\frac{i}{n}} - X_{\frac{i-1}{n}}, \quad \Delta_n = \frac{1}{n}, \quad p > 0. \tag{1.3}$$

Under certain technical conditions, (1.2) implies

$$\Delta_n^{1-p/\beta} V(p, \Delta_n) \xrightarrow{\mathbb{P}} \mu \int_0^1 |\sigma_s|^p ds, \quad (2\Delta_n)^{1-p/\beta} V(p, 2\Delta_n) \xrightarrow{\mathbb{P}} \mu \int_0^1 |\sigma_s|^p ds,$$

where μ is some constant. An estimate of β then can be simply formed as a nonlinear function of the ratio $\frac{V(p,\Delta_n)}{V(p,2\Delta_n)}$. This makes inference for β possible despite the unknown process σ.

The limit result in (1.2), however, contains much more information about β than previously used in estimation. In particular, (1.2) implies that over a short interval of time the increments of X, conditional on σ at the beginning of the interval, are approximately i.i.d. stable random variables. In this paper we propose a new estimator of β that utilizes this additional information in (1.2) and leads to significant efficiency gains over existent estimators based on high-frequency data.

The key obstacle in utilizing the result in (1.2) in inference for β is the fact that the process σ is unknown and time-varying. The idea of our method is to form a local estimator of σ using a block of high-frequency increments with asymptotically shrinking time span via a localized version of (1.3). We then divide the high-frequency increments of X by the local estimator of σ. The division achieves "self-normalization" in the following sense. First, the scale factor for the local estimator of σ and the high-frequency increment of X are the same and hence by taking the ratio they cancel. Second, both the high-frequency increment of X and the local estimator of σ are approximately proportional to the value of σ at the beginning of the high-frequency interval and hence taking their ratio cancels the effect of the unknown σ. The resulting scaled high-frequency increments are approximately i.i.d. stable random variables and we make inference for β via an analogue of the empirical characteristic function approach which has been used in various other contexts, see e.g., [8].

After removing an asymptotic bias, the limit behavior of the empirical characteristic function of the scaled high-frequency increments is determined by two correlated normal random variables. One of them is due to the limiting behavior of the empirical characteristic function of the high-frequency increments scaled by the limit of the local power variation. The other is due to the error in estimating the local scale by the local realized power variation. Importantly, because of the "self-normalization", the $\mathcal{F}$-conditional asymptotic variance of the empirical characteristic function of the scaled high-frequency increments is not random but rather a constant that depends only on β and the power p. This makes feasible inference very easy.

When comparing the new estimator with existing ones based on the power variation, we find nontrivial efficiency gains. There are two reasons for the efficiency gains. First, as we noted above, our estimator makes full use of the limiting result in (1.2) and not just the dependence of the scale of the high-frequency increments on β, which is the case for existing ones. Second, by locally removing the effect of the time-varying σ, we make the inference as if σ is constant, i.e., the limit variance is the same regardless whether X is Lévy or not. By contrast, the estimator based on the ratios of power

variations is asymptotically mixed normal with $\mathcal{F}$-conditional variance of the form $K(p,\beta)\frac{\int_0^1|\sigma_s|^{2p}ds}{\left(\int_0^1|\sigma_s|^p ds\right)^2}$, for some constant $K(p,\beta)$ and we note that $\frac{\int_0^1|\sigma_s|^{2p}ds}{\left(\int_0^1|\sigma_s|^p ds\right)^2}\geq 1$ with equality whenever the process $|\sigma|$ is almost everywhere constant on the interval $[0,1]$. That is, the presence of time-varying σ decreases the precision of the power-variation based estimator of β.

The efficiency gains of our estimator are bigger for higher values of β. In the limit case of $\beta=2$, which corresponds to L being a Brownian motion, we show that our estimator can achieve a faster rate of convergence than the standard $\sqrt{n}$ rate for existing estimators.

The rest of the paper is organized as follows. In Section 2 we introduce the setting. In Section 3 we construct our statistic and in Section 4 we derive its limit behavior. In Section 5 we build on the developed limit theory and construct new estimators of the jump activity and derive their limit behavior. This section also shows the efficiency gains of the proposed jump activity estimators over existing ones. Section 6 deals with the limiting case of jump-diffusion. Sections 7 and 8 contain a Monte Carlo study and an empirical application, respectively. Proofs are in Section 9.

2. Setting and Assumptions. We start with introducing the setting and stating the assumptions that we need for the results in the paper. We first recall that a Lévy process L with characteristic triplet (b,c,ν), with respect to truncation function κ (Definition II.2.3 in [14]), is a process with characteristic function given by

$$\mathbb{E}\left(e^{iuL_t}\right)=\exp\left[itub-tcu^2/2+t\int_{\mathbb{R}}\left(e^{iux}-1-iu\kappa(x)\right)\nu(dx)\right],\quad t\geq 0. \tag{2.1}$$

In what follows we will always assume for simplicity that $\kappa(-x)=-\kappa(x)$. Our assumption for the driving Lévy process in (1.1) as well as the "residual" jump component Y is given in assumption A.

Assumption A. *L in (1.1) is a Lévy process with characteristic triplet $(0,0,\nu)$ for ν a Lévy measure with density given by*

$$\nu(x)=\frac{A}{|x|^{1+\beta}}+\nu'(x),\quad \beta\in(0,2), \tag{2.2}$$

where $A>0$ and $\nu'(x)$ is such that there exists $x_0>0$ with $|\nu'(x)|\leq C/|x|^{1+\beta'}$ for $|x|\leq x_0$ and some $\beta'<\beta$.
Y is an Itô semimartingale with characteristic triplet ([14], Definition II.2.6) $\left(\int_0^t\int_{\mathbb{R}}\kappa(x)\nu_s^Y(dx)ds,0,dt\otimes\nu_t^Y(dx)\right)$ when $\beta'<1$ and $\left(0,0,dt\otimes\nu_t^Y(dx)\right)$

otherwise, with $\int_{\mathbb{R}}(|x|^{\beta'+\iota}\wedge 1)\nu_t^Y(dx)$ *being locally bounded predictable, for some arbitrary small* $\iota>0$.

Assumption A formalizes the sense in which Y is dominated at high frequencies by L: the activity index of Y is below that of L. We also stress that Y and L can have dependence. Therefore, as shown in [24], we can accommodate in our setup time-changed Lévy models, with absolute continuous time-change process, that have been extensively used in applied work. Finally, we note that (2.2) restricts only the behavior of ν around zero and ν' is a signed measure. Therefore many parametric jump specifications outside of the stable process are satisfied by Assumption A (e.g., the tempered stable process). We next state our assumption for the dynamics of α and σ.

Assumption B. *The processes* α *and* σ *are Itô semimartingales of the form*

(2.3)
$$\alpha_t=\alpha_0+\int_0^t b_s^\alpha ds+\int_0^t\int_E\kappa(\delta^\alpha(s,x))\widetilde{\underline{\mu}}(ds,dx)+\int_E\kappa'(\delta^\alpha(s,x))\underline{\mu}(ds,dx),$$
$$\sigma_t=\sigma_0+\int_0^t b_s^\sigma ds+\int_0^t\int_E\kappa(\delta^\sigma(s,x))\widetilde{\underline{\mu}}(ds,dx)+\int_E\kappa'(\delta^\sigma(s,x))\underline{\mu}(ds,dx),$$

where $\kappa'(x)=x-\kappa(x)$, *and*

(a) $|\sigma_t|^{-1}$ *and* $|\sigma_{t-}|^{-1}$ *are strictly positive;*
(b) $\underline{\mu}$ *is Poisson measure on* $\mathbb{R}_+\times E$, *having arbitrary dependence with the jump measure of* L, *with compensator* $dt\otimes\lambda(dx)$ *for some* σ*-finite measures* λ *on* E*;*
(c) $\delta^\alpha(t,x)$ *and* $\delta^\sigma(t,x)$ *are predictable, left-continuous with right limits in* t *with* $|\delta^\alpha(t,x)|+|\delta^\sigma(t,x)|\le\gamma_k(x)$ *for all* $t\le T_k$, *where* $\gamma_k(x)$ *is a deterministic function on* $\mathbb{R}$ *with* $\int_{\mathbb{R}}(|\gamma_k(x)|^{r+\iota}\wedge 1)\lambda(dx)<\infty$ *for arbitrary small* $\iota>0$ *and some* $0\le r\le\beta$, *and* T_k *is a sequence of stopping times increasing to* $+\infty$*;*
(d) b^α *and* b^σ *are Itô semimartingales having dynamics as in (2.3) with coefficients satisfying the analogues of conditions (b) and (c) above.*

We note that $\underline{\mu}$ does not need to coincide with the jump measure of L and hence it allows for dependence between the processes α, σ and L. This is of particular relevance for financial applications. For example, Assumption B is satisfied by the COGARCH model of [17] in which the jumps in σ are proportional to the squared jumps in X. More generally, Assumption B is satisfied if for example (X,α,σ) is modeled via a Lévy-driven SDE, with each of the elements of the driving Lévy process satisfying Assumption A.

3. Construction of the Self-Normalized Statistics. We continue next with the construction of our statistics. The estimation in the paper is based on observations of X at the equidistant grid times $0, \frac{1}{n}, ..., 1$ with $n \to \infty$, and we denote $\Delta_n = \frac{1}{n}$. To minimize the effect of the drift in our statistics, we follow [22] and work with the first difference of the increments, $\Delta_i^n X - \Delta_{i-1}^n X$, where $\Delta_i^n X = X_{\frac{i}{n}} - X_{\frac{i-1}{n}}$ for $i = 1, ..., n$. The above difference of increments is purged from the drift in the Lévy case, and in the general case the drift has a smaller asymptotic effect on it. For each $\Delta_i^n X - \Delta_{i-1}^n X$, we need a local power variation estimate for the scale. It is constructed from a block of k_n high-frequency increments, for some $1 < k_n < n - 2$, as follows

$$V_i^n(p) = \frac{1}{k_n} \sum_{j=i-k_n-1}^{i-2} |\Delta_j^n X - \Delta_{j-1}^n X|^p, \quad i = k_n + 3, ..., n. \tag{3.1}$$

Block-based local estimators of volatility have been also used in other contexts in a high-frequency setting, e.g., [13] and [25]. The empirical characteristic function of the scaled differenced increments is given by

$$\widehat{\mathcal{L}}^n(p, u) = \frac{1}{n - k_n - 2} \sum_{i=k_n+3}^{n} \cos\left(u \frac{\Delta_i^n X - \Delta_{i-1}^n X}{(V_i^n(p))^{1/p}}\right), \quad u \in \mathbb{R}_+. \tag{3.2}$$

We proceed with some notation needed for the limiting theory of $\widehat{\mathcal{L}}^n(p, u)$. Let S_1, S_2 and S_3 be random variables corresponding to the values of three independent Lévy processes at time 1, each of which with characteristic triplet $(0, 0, \nu)$, for any truncation function κ and where ν has the density $\frac{A}{|x|^{1+\beta}}$. Then we denote $\mu_{p,\beta} = (\mathbb{E}|S_1 - S_2|^p)^{\beta/p}$, which does not depend on κ, and we further use the shorthand notation $\mathbb{E}\left(e^{iu(S_1 - S_2)}\right) = e^{-A_\beta u^\beta}$ for any $u > 0$ with A_β being a (known) function of A and β. Using example 25.10 in [21] and references therein, we have

$$C_{p,\beta} = \frac{A_\beta}{\mu_{p,\beta}} = \left[\frac{2^p \Gamma\left(\frac{1+p}{2}\right) \Gamma\left(1 - \frac{p}{\beta}\right)}{\sqrt{\pi} \Gamma\left(1 - \frac{p}{2}\right)}\right]^{-\beta/p}, \tag{3.3}$$

which depends only on p and β but not on the scale parameter of the stable random variables S_1 and S_2. With this notation, we set

$$\mathcal{L}(p, u, \beta) = e^{-C_{p,\beta} u^\beta}, \quad u \in \mathbb{R}_+, \tag{3.4}$$

which will be the limit in probability of $\widehat{\mathcal{L}}^n(p,u)$. We finish with some more notation needed to describe the asymptotic variance of $\widehat{\mathcal{L}}^n(p,u)$. First, we denote for some $u \in \mathbb{R}_+$

$$\begin{aligned} \xi_1(p,u,\beta) &= \left(\cos\left(\frac{u(S_1-S_2)}{\mu_{p,\beta}^{1/\beta}}\right) - \mathcal{L}(p,u,\beta), \quad \frac{|S_1-S_2|^p}{\mu_{p,\beta}^{p/\beta}} - 1 \right)', \\ \xi_2(p,u,\beta) &= \left(\cos\left(\frac{u(S_2-S_3)}{\mu_{p,\beta}^{1/\beta}}\right) - \mathcal{L}(p,u,\beta), \quad \frac{|S_2-S_3|^p}{\mu_{p,\beta}^{p/\beta}} - 1 \right)'. \end{aligned} \tag{3.5}$$

We then set for $u, v \in \mathbb{R}_+$

$$\Xi_i(p,u,v,\beta) = \mathbb{E}\left(\xi_1(p,u,\beta)\xi_{1+i}'(p,v,\beta)\right), \qquad i = 0,1, \tag{3.6}$$

and

$$G(p,u,\beta) = \frac{\beta}{p} e^{-C_{p,\beta}u^\beta} C_{p,\beta} u^\beta, \; H(p,u,\beta) = G(p,u,\beta)\left(\frac{\beta}{p} C_{p,\beta} u^\beta - \frac{\beta}{p} - 1\right). \tag{3.7}$$

4. Limit Theory for $\widehat{\mathcal{L}}^n(p,u)$. We start with convergence in probability.

THEOREM 1. *Assume X satisfies Assumptions A and B for some $\beta \in (1,2)$ and $\beta' < \beta$. Let k_n be a deterministic sequence satisfying $k_n \asymp n^{\varpi}$ for some $\varpi \in (0,1)$. Then, for $0 < p < \beta$, we have*

$$\widehat{\mathcal{L}}^n(p,u) \xrightarrow{\mathbb{P}} \mathcal{L}(p,u,\beta), \quad \text{as } n \to \infty, \tag{4.1}$$

locally uniformly in $u \in \mathbb{R}_+$.

We note that we restrict $\beta > 1$, i.e., we focus on the infinite variation case. The above theorem will continue to hold for $\beta \le 1$, but for the subsequent results about the limiting distribution of $\widehat{\mathcal{L}}^n(p,u)$, we will need quite stringent additional restrictions in the case $\beta \le 1$. We do not pursue this here. The other conditions for the convergence in probability result are weak. The requirements for α and σ for Theorem 1 to hold are actually much weaker than what is assumed in Assumption B, but for simplicity of exposition we keep Assumption B throughout. We note that for consistency, we have a lot of flexibility about the block size k_n: (1) $k_n \to \infty$ so that we consistently estimate the scale via $V_i^n(p)$ and (2) $k_n/n \to 0$ so that the span of the block is asymptotically shrinking to zero and therefore no bias is generated due to the time variation of σ. In the case when X is a Lévy process, the second condition is obviously not needed.

To derive a central limit theorem (clt) for $\widehat{\mathcal{L}}^n(p,u)$, we will need to restrict the choice of k_n more. We will assume $k_n/\sqrt{n}\to 0$, so that biases due to the time variation in σ, which are hard to feasible estimate, are negligible. For such a choice of k_n, however, an asymptotic bias due to the sampling error of $V_i^n(p)$ appears and for stating a clt, we need to consider the following bias-corrected estimator

$$\begin{aligned}\widehat{\mathcal{L}}^n(p,u,\beta)' &= \widehat{\mathcal{L}}^n(p,u)\\ &\quad -\frac{1}{k_n}\frac{1}{2}H(p,u,\beta)\left(\Xi_0^{(2,2)}(p,u,u,\beta)+2\Xi_1^{(2,2)}(p,u,u,\beta)\right).\end{aligned} \tag{4.2}$$

We state the clt for $\widehat{\mathcal{L}}^n(p,u,\beta)'$ in the next theorem.

THEOREM 2. *Assume X satisfies Assumptions A and B with $\beta\in(1,2)$ and $\beta'<\frac{\beta}{2}$, and that the power p and block size k_n satisfy*

$$\frac{\beta\beta'}{2(\beta-\beta')}\bigvee\frac{\beta-1}{2}<p<\frac{\beta}{2}, \tag{4.3}$$

$$k_n\asymp n^{\varpi},\quad \frac{p}{\beta}\bigvee\frac{1}{3}<\varpi<\frac{1}{2}. \tag{4.4}$$

Then, as $n\to\infty$, we have

$$\sqrt{n}\left(\widehat{\mathcal{L}}^n(p,u,\beta)'-\mathcal{L}(p,u,\beta)\right)\ \xrightarrow{\mathcal{L}}\ Z_1(u)+G(p,u,\beta)Z_2(u), \tag{4.5}$$

locally uniformly in $u\in\mathbb{R}_+$. $Z_1(u)$ and $Z_2(u)$ are two Gaussian processes with the following covariance structure

$$\mathbb{E}\left(\mathbf{Z}(u)\mathbf{Z}(v)\right)=\Xi_0(p,u,v,\beta)+2\Xi_1(p,u,v,\beta),\quad u,v\in\mathbb{R}_+. \tag{4.6}$$

where $\mathbf{Z}(u)=(Z_1(u),Z_2(u))'$.

Let $\widehat{\beta}$ be an estimator of β with $\widehat{\beta}-\beta=o_p(k_n\sqrt{\Delta_n})$ as $n\to\infty$. Then

$$\sqrt{n}\left(\widehat{\mathcal{L}}^n(p,u,\widehat{\beta})'-\widehat{\mathcal{L}}^n(p,u,\beta)'\right)\ \xrightarrow{\mathbb{P}}\ 0, \tag{4.7}$$

locally uniformly in $u\in\mathbb{R}_+$.

The conditions for the power p in (4.3) are exactly the same as in [22] for the analysis of the realized power variation and they are relatively weak. For example, the condition $p>\frac{\beta-1}{2}$ will be always satisfied as soon as we pick power slightly above $\frac{1}{2}$. Moreover, this condition is not needed in the case

when X is a Lévy process. Further, the condition in (4.4) for k_n shows that we have more flexibility for the choice of k_n whenever p is not very close to its upper bound of $\beta/2$.

Due to the self-normalization in the construction of our statistic, the limiting distribution in (4.5) is Gaussian and not mixed Gaussian, which is the case for most limit results in high-frequency asymptotics (and in particular for the power variation based estimator of β), see [26] for another exception. This is very convenient as the estimation of the asymptotic variance is straightforward. The bias correction in (4.2) is infeasible, as it depends on β. However, (4.7) shows that a feasible version of the debiasing would work provided the initial estimator of β is $o_p(k_n\sqrt{\Delta_n})$. When one estimates β using $\widehat{\mathcal{L}}^n(p,u)$, with explicit estimators provided in the next section, $\widehat{\beta}-\beta$ will be $O_p(1/k_n)$. Hence, such a preliminary estimate of β will satisfy the required rate condition in Theorem 2.

5. Jump Activity Estimation. We now use the limit theory developed above to form estimators of β. The simplest one is based on $\widehat{\mathcal{L}}^n(p,u)$ and is given by

$$\widehat{\beta}^{fs}(p,u,v) = \frac{\log\Big(-\log(\widehat{\mathcal{L}}^n(p,u))\Big) - \log\Big(-\log(\widehat{\mathcal{L}}^n(p,v))\Big)}{\log(u/v)}, \tag{5.1}$$

for $u,v \in \mathbb{R}_+$ with $u \neq v$. Because of the asymptotic bias in $\widehat{\mathcal{L}}^n(p,u)$, $\widehat{\beta}^{fs}(p,u,v)-\beta$ will be only $O_p(1/k_n)$, with p and k_n satisfying (4.3)-(4.4). An explicit estimate of β using feasible debiasing is given by

$$\widehat{\beta}(p,u,v) = \frac{\log\Big(-\log(\widehat{\mathcal{L}}^n(p,u,\widehat{\beta}^{fs})')\Big) - \log\Big(-\log(\widehat{\mathcal{L}}^n(p,v,\widehat{\beta}^{fs})')\Big)}{\log(u/v)}, \tag{5.2}$$

for some $u,v \in \mathbb{R}_+$ with $u \neq v$, and where $\widehat{\beta}^{fs}$ is a suitable initial estimator of β (like the one in (5.1)). While convenient, the above estimators have two potential drawbacks. One, we do not take into account the information about β in the constant $C_{p,\beta}$. This is because in the asymptotic limit of the above estimators, $C_{p,\beta}$ gets canceled. Second, u and v are chosen arbitrary and one can include more moment conditions for the estimation of β using $\widehat{\mathcal{L}}^n(p,u,\widehat{\beta}^{fs})'$. In the next theorem we provide a general estimator of β which overcomes these drawbacks of the explicit estimators above.

THEOREM 3. *Assume X satisfies Assumptions A and B with $\beta \in (1,2)$ and $\beta' < \beta/2$, and that the conditions in (4.3) and (4.4) hold. Suppose $\widehat{\beta}^{fs}$ is a consistent estimator of β with $\widehat{\beta}^{fs}-\beta = o_p(k_n\sqrt{\Delta_n})$. Denote with $\widehat{\mathbf{u}}_l$*

and $\widehat{\mathbf{u}}_h$ *two sequences of* $K \times 1$*-dimensional vectors, for some finite* $K \geq 1$*, satisfying* $\widehat{\mathbf{u}}_l \xrightarrow{\mathbb{P}} \mathbf{u}_l$ *and* $\widehat{\mathbf{u}}_h \xrightarrow{\mathbb{P}} \mathbf{u}_h$ *as* $n \to \infty$*, for some* $\mathbf{u}_l, \mathbf{u}_h \in \mathbb{R}_+^K$ *with* $u_l^i < u_h^i$*,* $u_l^j < u_h^j$ *and* $(u_l^i, u_h^i) \cap (u_l^j, u_h^j) = \emptyset$ *for every* $i, j = 1, ..., K$ *with* $i \neq j$ *where* u_l^i *and* u_h^i *denote the* i*-th element of the vectors* $\mathbf{u}_l$ *and* $\mathbf{u}_h$ *respectively. Set further the shorthand* $\mathbf{u} = [\mathbf{u}_l;\ \mathbf{u}_h]$ *and* $\widehat{\mathbf{u}} = [\widehat{\mathbf{u}}_l;\ \widehat{\mathbf{u}}_h]$*.*

Let $\mathbf{W}(p, \mathbf{u}, \beta)$ *be* $K \times K$ *matrix with* (i, j) *element given by*

$$\mathbf{W}(p, \mathbf{u}, \beta)_{i,j} = \int_{u_l^i}^{u_h^i} \int_{u_l^j}^{u_h^j} w(p, u, v, \beta) du dv, \tag{5.3}$$

$$w(p, u, v, \beta) = \frac{1}{\mathcal{L}(p, u, \beta)\mathcal{L}(p, v, \beta)} \begin{pmatrix} 1 \\ G(p, u, \beta) \end{pmatrix}' \overline{\Xi}(p, u, v, \beta) \begin{pmatrix} 1 \\ G(p, v, \beta) \end{pmatrix},$$

where $\overline{\Xi}(p, u, v, \beta) = \Xi_0(p, u, v, \beta) + 2\Xi_1(p, u, v, \beta)$*.*

Define the $K \times 1$ *vector* $\widehat{\mathbf{m}}(p, \widehat{\mathbf{u}}, \widehat{\beta}^{fs}, \mathbf{u}, \beta)$ *by*

$$\widehat{\mathbf{m}}(p, \widehat{\mathbf{u}}, \widehat{\beta}^{fs}, \mathbf{u}, \beta)_i = \int_{\widehat{u}_l^i}^{\widehat{u}_h^i} \left(\log(\widehat{\mathcal{L}}^n(p, u, \widehat{\beta}^{fs})') - \log(\mathcal{L}(p, u, \beta)) \right) du, \tag{5.4}$$

for $i = 1, ..., K$ *and set*

$$\widehat{\beta}(p, \mathbf{u}) = \underset{\beta \in (1,2)}{argmin}\ \widehat{\mathbf{m}}(p, \widehat{\mathbf{u}}, \widehat{\beta}^{fs}, \mathbf{u}, \beta)' \mathbf{W}^{-1}(p, \widehat{\mathbf{u}}, \widehat{\beta}^{fs}) \widehat{\mathbf{m}}(p, \widehat{\mathbf{u}}, \widehat{\beta}^{fs}, \mathbf{u}, \beta). \tag{5.5}$$

Finally define the $K \times 1$ *vector* $\mathbf{M}(p, \mathbf{u}, \beta)$ *by*

$$\mathbf{M}(p, \mathbf{u}, \beta)_i = \int_{u_l^i}^{u_h^i} \nabla_\beta \log(\mathcal{L}(p, u, \beta)) du, \quad i = 1, ..., K. \tag{5.6}$$

Then for $\beta \in (1, 2)$*,* $p \in \left(\frac{\beta\beta'}{2(\beta - \beta')}, \frac{\beta}{2} \right)$*, and* $\beta' < \beta/2$*, we have*

$$\sqrt{n} \left(\widehat{\beta}(p, \mathbf{u}) - \beta \right) \xrightarrow{\mathcal{L}} \sqrt{\mathbf{M}(p, \mathbf{u}, \beta)' \mathbf{W}^{-1}(p, \mathbf{u}, \beta) \mathbf{M}(p, \mathbf{u}, \beta)} \times \mathcal{N}, \tag{5.7}$$

for $n \to \infty$ *with* $\mathcal{N}$ *being standard normal random variable.*

A consistent estimator for the asymptotic variance of $\widehat{\beta}(p, \mathbf{u})$ *is given by*

$$\mathbf{M}(p, \widehat{\mathbf{u}}, \widehat{\beta})' \mathbf{W}^{-1}(p, \widehat{\mathbf{u}}, \widehat{\beta}) \mathbf{M}(p, \widehat{\mathbf{u}}, \widehat{\beta}), \tag{5.8}$$

where $\mathbf{M}(p, \widehat{\mathbf{u}}, \widehat{\beta})$ *is defined as* $\mathbf{M}(p, \mathbf{u}, \beta)$ *with* $\mathbf{u}$ *and* β *replaced by* $\widehat{\mathbf{u}}$ *and* $\widehat{\beta}$*.*

Theorem 3 allows to adaptively choose the range of u over which to match $\widehat{\mathcal{L}}^n(p, u, \widehat{\beta}^{fs})'$ with its limit. This is convenient because the limiting variance

of $\widehat{\mathcal{L}}^n(p,u,\widehat{\beta}^{fs})'$ depends on β. For this reason also the weight function in (5.3) optimally weighs the moment conditions in the estimation. We discuss the practical issues regarding the construction of $\widehat{\mathbf{m}}(p,\widehat{\mathbf{u}},\widehat{\beta}^{fs},\mathbf{u},\beta)$ in Section 7.

We now illustrate the efficiency gains provided by the new method over existing power variation based estimators of β. The power variation estimator based on the differenced increments is given by (see [22])

$$\widetilde{\beta}(p) = \frac{p\log(2)}{\log\left[\widetilde{V}_2^n(p)/\widetilde{V}_1^n(p)\right]}1_{\{\widetilde{V}_1^n(p)\neq\widetilde{V}_2^n(p)\}}, \tag{5.9}$$

where

$$\widetilde{V}_1^n(p) = \sum_{i=2}^n |\Delta_i^n X - \Delta_{i-1}^n X|^p, \quad \widetilde{V}_2^n(p) = \sum_{i=4}^n |\Delta_i^n X - \Delta_{i-1}^n X + \Delta_{i-2}^n X - \Delta_{i-3}^n X|^p.$$

On Figure 1, we plot the limiting standard deviation of the estimators in (5.5) and (5.9) for different values of β (the estimator in (5.9) is derived under exactly the same assumptions for X as our estimator here). The asymptotic standard deviation of $\widetilde{\beta}(p)$ is computed from [22]. $\widehat{\beta}(p,u)$ is far less sensitive to the choice of p than $\widetilde{\beta}(p)$ is, with lower powers yielding marginally more efficient $\widehat{\beta}(p,u)$. The news estimator $\widehat{\beta}(p,u)$ provides nontrivial efficiency gains irrespective of the value of p and β. The gains are bigger for high values of the jump activity. For example, for $\beta = 1.75$, $\widehat{\beta}(p,u)$ is around two times more efficient (in terms of asymptotic standard deviation) than $\widetilde{\beta}(p)$.

6. The Limiting Case of Jump-Diffusion. So far our analysis has been for the pure-jump case of $\beta \in (1,2)$. We now look at the limiting case of $\beta = 2$, which corresponds to L in (1.1) being a Brownian motion. In this case the asymptotic behavior of the high-frequency increments in (1.2) holds with S being a Brownian motion. Thus, deciding $\beta = 2$ versus $\beta < 2$ amounts to testing pure-jump versus jump-diffusion specification for X. It turns out that when $\beta = 2$ our estimation method can lead to faster rate of convergence than the $\sqrt{n}$ rate we have seen for the case $\beta \in (1,2)$. This is unlike the power-variation based estimation methods for which the rate of convergence is $\sqrt{n}$, both for $\beta = 2$ and $\beta < 2$, see e.g., [23].

The faster rate of convergence in the case $\beta = 2$ can be achieved by letting the argument u of the empirical characteristic function $\widehat{\mathcal{L}}(p,u)$ drift towards zero as $n \to \infty$. In this case, $\frac{-\log(\widehat{\mathcal{L}}(p,u_n,2)')}{C_{p,2}u_n^2}$ and $\frac{-\log(\widehat{\mathcal{L}}(p,\rho u_n,2)')}{C_{p,2}\rho^2 u_n^2}$, for some $\rho > 0$, are asymptotically perfectly correlated and their difference converges at a faster rate. We note that this does not work in the pure-jump case of

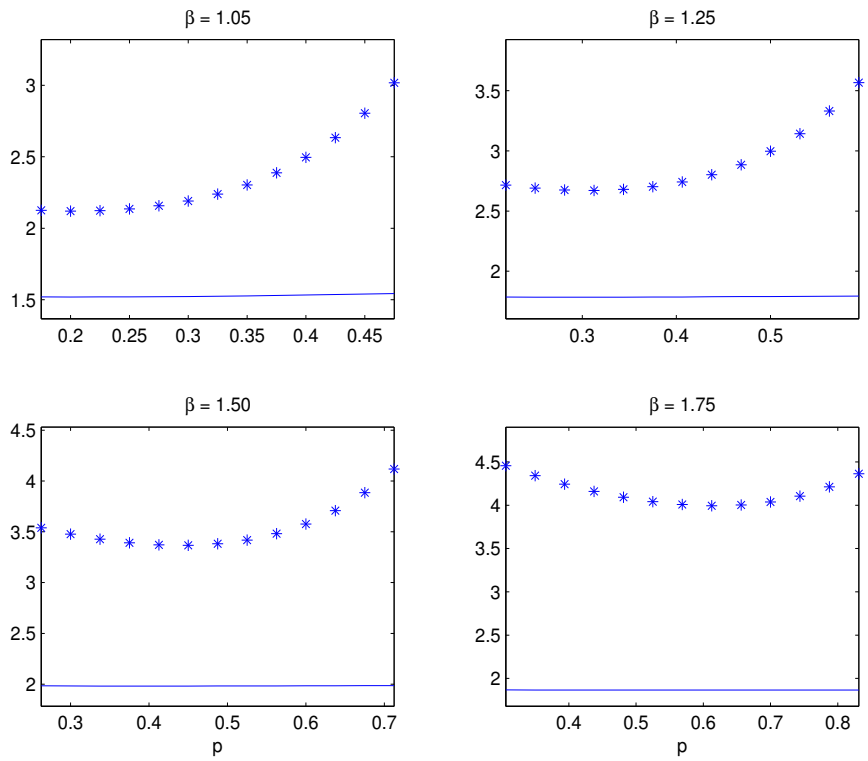

FIG 1. *Asymptotic standard deviation of jump activity estimators. The straight line corresponds to the asymptotic standard deviation of the characteristic function based estimator defined in (5.5) and the $*$ line to the power variation based estimator of [22] given in (5.9) (when σ is constant). For each cases of β, the power p ranges in the interval $p \in \left(\frac{7}{40}, \frac{19}{40}\right)\beta$. For the estimator in (5.5), the vector $\mathbf{u}_l = [0.1 : 0.05 : 5]$ and $\mathbf{u}_h = [0.15 : 0.05 : 5.05]$.*

$\beta < 2$. To state the formal result we first introduce some notation. For S_1, S_2 and S_3 being independent standard normal random variables we denote

$$\begin{aligned} \widetilde{\xi}_1(p) &= \left(\frac{|S_1 - S_2|^4}{\mu_{p,2}^2} - \frac{12}{\mu_{p,2}^2}, \quad \frac{|S_1 - S_2|^p}{\mu_{p,2}^{p/2}} - 1 \right)', \\ \widetilde{\xi}_2(p) &= \left(\frac{|S_2 - S_3|^4}{\mu_{p,2}^2} - \frac{12}{\mu_{p,2}^2}, \quad \frac{|S_2 - S_3|^p}{\mu_{p,2}^{p/2}} - 1 \right)', \end{aligned} \tag{6.1}$$

and then set $\widetilde{\Xi}_i(p) = \mathbb{E}\left(\widetilde{\xi}_1(p)\widetilde{\xi}'_{1+i}(p)\right)$ for $i = 0, 1$. The difference from the analogous expression for the case $\beta < 2$ is in the first terms of $\widetilde{\xi}_1(p)$ and $\widetilde{\xi}_2(p)$. Note that the expression for the bias-correction remains exactly the same as it involves only the variance and covariance of the second elements of $\widetilde{\xi}_1(p)$ and $\widetilde{\xi}_2(p)$ which remain the same as their pure-jump counterparts.

THEOREM 4. *Suppose X has dynamics given by (1.1) with L being a Brownian motion, Y satisfying the corresponding condition for it in Assumption A, and α and σ satisfying Assumption B for some $r < 2$. Suppose $p < 1$, $k_n\sqrt{\Delta_n} \to 0$ and $u_n \to 0$, and further*

$$\frac{\Delta_n^{\left(\frac{p}{\beta'}-\frac{p}{2}\right)\wedge\frac{p+1}{r\vee 1+1}-\iota} \vee k_n^{-\left(\frac{1}{p}\wedge\frac{3}{2}\right)+\iota} \vee (k_n\Delta_n)^{1-\iota}}{u_n^6\sqrt{\Delta_n}} \to 0, \quad \frac{(k_n\Delta_n)^{\frac{1}{r}\wedge\frac{2-p}{2}-\frac{1}{2}}}{u_n^6} \to 0. \tag{6.2}$$

Then for some $\rho > 0$

$$\widehat{\beta}^{fs}(p, u_n, \rho u_n) - 2 = O_p(k_n^{-1}u_n^2). \tag{6.3}$$

Further, if for some initial estimator $\widehat{\beta}^{fs} - 2 = o_p(k_n u_n^2\sqrt{\Delta_n})$, then

$$\frac{\sqrt{n}}{u_n^2(1-\rho^2)}\left(\widehat{\beta}(p, u_n, \rho u_n) - 2\right) \xrightarrow{\mathcal{L}} -\frac{1}{\log(\rho)}\left(\frac{1}{24C_{p,2}}Z_1 - \frac{2}{p}C_{p,2}Z_2\right), \tag{6.4}$$

where Z_1 and Z_2 are two zero-mean normal random variables with covariance given by $\widetilde{\Xi}_0(p) + 2\widetilde{\Xi}_1(p)$.

When X is a Lévy process, the requirement for k_n and u_n reduces to

$$u_n \to 0, \quad \frac{\Delta_n^{\left(\frac{p}{\beta'}\wedge 1-\frac{p}{2}\right)-\iota} \vee k_n^{-\left(\frac{1}{p}\wedge\frac{3}{2}\right)+\iota}}{u_n^6\sqrt{\Delta_n}} \to 0. \tag{6.5}$$

The rate of convergence of the estimator for β is now $\sqrt{n}u_n^{-2}$ and is faster, than the one in Theortem 3, when u_n converges to zero. The latter is determined by the restriction in (6.2), which in turn is governed by the presence of the "residual" term Y, the variation in σ and the sampling variation in measuring the scale via $V_i^n(p)$. For the condition to be satisfied we need $p \in (1/2, 1)$ and $\beta' < 1$, i.e., the jumps in X are of finite variation (for testing the null hypothesis of presence of diffusion when the process can contain infinite variation jumps under the null see the recent work of [19]). Without any prior knowledge on β' and r, we can set k_n according to (4.4), with $\beta = 2$, and then set $u_n \asymp \log(n)^{-1}$. The requirement on u_n can be further relaxed when X is a Lévy process as evident from (6.5). Finally, we can draw a parallel between our finding for faster rate of convergence of the estimator of β when $\beta = 2$ with the result in [9, 10] for faster rate of convergence for the maximum likelihood estimator of the stability index of i.i.d. β-stable random variables when $\beta = 2$.

7. Monte Carlo. We test the performance of the proposed method for jump activity estimation on simulated data from the following model

$$dX_t = \sigma_{t-}dL_t, \quad d\sigma_t = -0.03\sigma_t dt + dZ_t, \tag{7.1}$$

where L and Z are two Lévy processes independent of each other with Lévy densities given by $\nu_L(x) = e^{-\lambda|x|}\left(\frac{A_0}{|x|^{1+\beta}} + \frac{A_1}{|x|^{1+\frac{\beta}{3}}}\right)$ and $\nu_Z(x) = 0.0293\frac{e^{-3x}}{x^{1.5}}1_{\{x>0\}}$ respectively. σ is a Lévy-driven Ornstein-Uhlenbeck process with a tempered stable driving Lévy subordinator. The parameters governing the dynamics of σ imply $\mathbb{E}(\sigma_t) = 1$ and half-life of shock in σ of around one month (when unit of time is a day). L is a mixture of tempered stable processes with the parameter β coinciding with the jump activity index of X. We fix $\lambda = 0.25$ and consider four cases for β. In each of the cases we set A_0 and A_1 so that $A_0 \int_{\mathbb{R}} |x|^{1-\beta} e^{-\lambda|x|} dx = 1$ and $A_1 \int_{\mathbb{R}} |x|^{1-\frac{\beta}{3}} e^{-\lambda|x|} dx = 0.2$. The four cases are: (1) $\beta = 1.05$ and $A_0 = 0.1299$, $A_1 = 0.0113$; (2) $\beta = 1.25$ and $A_0 = 0.1443$, $A_1 = 0.0125$; (3) $\beta = 1.50$ and $A_0 = 0.1410$, $A_1 = 0.0141$ and (4) $\beta = 1.75$ and $A_0 = 0.0975$, $A_1 = 0.0158$.

In the Monte Carlo we set $T = 10$ and $n = 100$ which corresponds approximately to two weeks of 5-minute return data in a typical financial setting. We further set $k_n = 50$ and $p = 0.51$. The initial estimator to construct the moments and the optimal weight matrix is simply $\widehat{\beta}^{fs}(p, u, v)$ with $u = 0.1$ and $v = 1.1$. If $p \geq \widehat{\beta}^{fs}(p, u, v)/2$, then we reduce the power to $p = \widehat{\beta}^{fs}(p, u, v)/4$. Based on the initial beta estimator, we estimate the values of u for which $\mathcal{L}(p, u, \beta) = 0.95$ and $\mathcal{L}(p, u, \beta) = 0.25$, and then split this interval in five equidistant regions which are used in constructing the moment vector in (5.4).

Regarding the number of moment conditions, K, in the construction of our estimator, we should keep in mind the following. Larger K helps improve efficiency of the estimator as our equal weighting of the characteristic function within each moment condition is suboptimal. However, the feasible estimate of the optimal weight matrix is unstable in small samples when K is large (this is similar to "curse of dimensionality" problems occurring in related contexts, see e.g., [11] and [18]). Moreover, since the characteristic function is smooth, one typically does not need many moment conditions to gain efficiency. For example, we also experimented in the Monte Carlo with ten moment conditions (by splitting the region of u into ten equidistant regions). The performance of the estimator based on the ten moment conditions was very similar to the one based on the five moment conditions whose performance we summarize below.

The results from the Monte Carlo are reported in Table 1. For comparison, we also report results for $\widetilde{\beta}(p)$ where p is set to the level which minimizes the corresponding asymptotic standard deviation in Figure 1. We notice satisfactory finite sample performance of $\widehat{\beta}(p, \mathbf{u})$. In all cases for β, $\widehat{\beta}(p, \mathbf{u})$ contains relatively small upward biases. These biases, however, are

well below those of $\widetilde{\beta}(p)$. We note that the finite sample bias of $\widehat{\beta}(p, \mathbf{u})$ can be significantly reduced, if similar to $\widetilde{\beta}(p)$, one uses an adaptive choice of power in the range $(\beta/4, \beta/3)$. The superiority of $\widehat{\beta}(p, \mathbf{u})$ holds also in terms of precision in estimating β, with inter-quantile ranges of $\widehat{\beta}(p, \mathbf{u})$ typically well below those of $\widetilde{\beta}(p)$.

TABLE 1
Monte Carlo Results

Case	$\widehat{\beta}(p, \mathbf{u})$			$\widetilde{\beta}(p)$		
	median	IQR	MAD	median	IQR	MAD
$\beta = 1.05$	1.0801	0.0791	0.0518	1.1154	0.0925	0.0792
$\beta = 1.25$	1.3058	0.0817	0.0680	1.3229	0.1158	0.0932
$\beta = 1.50$	1.5398	0.0886	0.0622	1.5767	0.1405	0.1072
$\beta = 1.75$	1.7782	0.0806	0.0536	1.8196	0.1704	0.1183

Note: IQR is the inter-quartile range and MAD is the mean absolute deviation around the true value. The power p for $\widetilde{\beta}(p)$ is set to the value which minimizes the corresponding asymptotic standard deviation displayed on Figure 1.

8. Empirical Application. We now apply the developed inference procedures on high-frequency data for the VIX index. The VIX index is a option-based measure for volatility on the market (S&P 500 index). It serves as a popular indicator for investors uncertainty, and it is used as the underlying asset for many volatility-based derivative contracts traded on the financial exchanges. Earlier work, consistent with parametric models for volatility, has provided evidence that the VIX index is a pure-jump Itô semimartingale. Here, we estimate its jump activity index. The estimation is based on 5-minute sampled data during the trading hours for year 2010. Like in the Monte Carlo, we split the year in intervals of 10 days (two weeks) and estimate the jump activity over each of them. The moments, the power p and the block size k_n are selected in the same way as in the Monte Carlo. Estimation results are presented on Figure 2. The estimated jump activity index takes values around 1.6. Overall, our results support a pure-jump specification of the VIX index.

9. Proofs. In the proofs we use the shorthand notation $\mathbb{E}_i^n(\cdot) \equiv \mathbb{E}(\cdot|\mathcal{F}_{i\Delta_n})$ and $\mathbb{P}_i^n(\cdot) \equiv \mathbb{P}(\cdot|\mathcal{F}_{i\Delta_n})$. We also denote with K a positive constant that does not depend on n and u and might change from line to line in the inequalities that follow. When we want to highlight that the constant depends only on some parameters a and b, we write $K_{a,b}$.

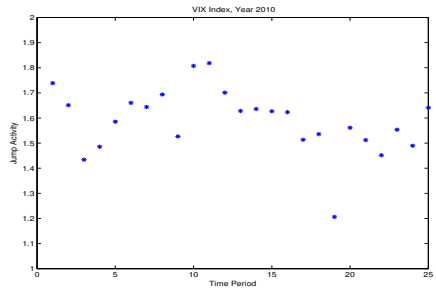

FIG 2. *Jump Activity for the VIX Index. Estimation is done over periods of* 10 *days in the year* 2010*. In the estimation, moments,* p*, and* k_n *are selected as in the Monte Carlo.*

9.1. *Decompositions and Additional Notation.* In what follows it is convenient to extend appropriately the probability space and then decompose the driving Lévy process L as follows

$$L_t + \widehat{S}_t = S_t + \widetilde{S}_t, \tag{9.1}$$

where S, $\widehat{S}$ and $\widetilde{S}$ are pure-jump Lévy processes with first two characteristics zero (with respect to the truncation function $\kappa(\cdot)$) and Lévy densities $\frac{A}{|x|^{1+\beta}}$, $2|\nu'(x)|1_{\{\nu'(x)<0\}}$, and $|\nu'(x)|$ respectively. We denote the associated counting jump measures with μ, μ_1 and μ_2 (note that there can be dependence between μ, μ_1 and μ_2).

S is β-stable process and $\widehat{S}$ and $\widetilde{S}$ are "residual" components whose effect on our statistic, as will be shown, is negligible (under suitable conditions). The proof of the decomposition in (9.1) as well as the explicit construction of S, $\widehat{S}$ and $\widetilde{S}$ can be found in Section 1 of the supplementary appendix of [24].

We now introduce some additional notation that will be used throughout the proofs. We denote for $i = k_n + 3, ..., n$

$$\widehat{V}_i^n(p) = \frac{1}{k_n} \sum_{j=i-k_n-1}^{i-2} |\sigma_{(j-2)\Delta_n-}|^p |\Delta_j^n S - \Delta_{j-1}^n S|^p,$$

$$\overline{V}_i^n(p) = \frac{1}{k_n} \sum_{j=i-k_n-1}^{i-2} \frac{|\Delta_j^n S - \Delta_{j-1}^n S|^p}{\mu_{p,\beta}^{p/\beta}},$$

$$\dot{V}_i^n(p) = \sum_{j=i-k_n-1}^{i-2} \left\{ \frac{[(i-j-4) \vee 0 + 1_{\{j<i-3\}}]}{k_n} \left(|\sigma_{j\Delta_n-}|^p - |\sigma_{(j-2)\Delta_n-}|^p \right) \right.$$
$$\left. + \frac{(|\sigma_{(j-1)\Delta_n-}|^p - |\sigma_{(j-2)\Delta_n-}|^p) 1_{\{j<i-2\}}}{k_n} \right\} |\Delta_j^n S - \Delta_{j-1}^n S|^p,$$

$$|\overline{\sigma}|_i^p = \frac{1}{k_n} \sum_{j=i-k_n-1}^{i-2} |\sigma_{(j-2)\Delta_n-}|^p.$$

We further denote the function

$$f_{i,u}(x) = \exp\left(-\frac{C_{p,\beta} u^\beta |\sigma_{(i-2)\Delta_n-}|^\beta}{x^{\beta/p}}\right),$$

and direct computation yields

$$\begin{cases} f'_{i,u}(x) = \frac{\beta}{p} f_{i,u}(x) \frac{C_{p,\beta} u^\beta |\sigma_{(i-2)\Delta_n-}|^\beta}{x^{\beta/p+1}}, \\ f''_{i,u}(x) = f_i(u,x) \left(\frac{\beta}{p} \frac{C_{p,\beta} u^\beta |\sigma_{(i-2)\Delta_n-}|^\beta}{x^{\beta/p+1}}\right)^2 - f_i(u,x) \frac{\beta}{p} \left(\frac{\beta}{p}+1\right) \frac{C_{p,\beta} u^\beta |\sigma_{(i-2)\Delta_n-}|^\beta}{x^{\beta/p+2}}. \end{cases}$$

We note

$$\sup_{x\in\mathbb{R}_+} |f_{i,u}(x) + f'_{i,u}(x) + f''_{i,u}(x) + f'''_{i,u}(x)| < K_u, \tag{9.2}$$

where the positive constant K_u depends only on u and is finite as soon as u is bounded away from zero.

With this notation, we make the following decomposition for any $u \in \mathbb{R}_+$

$$\widehat{\mathcal{L}}^n(p,u) - \mathcal{L}(p,u,\beta) = \frac{1}{n-k_n-2}\left[\widehat{Z}_1^n(u) + \widehat{Z}_2^n(u) + \sum_{j=1}^4 R_j^n(u)\right],$$

where $\widehat{Z}_j^n(u) = \sum_{i=k_n+3}^n z_i^j(u)$ for $j = 1, 2$ with

$$z_i^1(u) = \cos\left(u \frac{\sigma_{(i-2)\Delta_n-}(\Delta_i^n S - \Delta_{i-1}^n S)}{(V_i^n(p))^{1/p}}\right) - \exp\left(-\frac{A_\beta u^\beta |\sigma_{(i-2)\Delta_n-}|^\beta}{\Delta_n^{-1}(V_i^n(p))^{\beta/p}}\right),$$

$$z_i^2(u) = \exp\left(-\frac{C_{p,\beta} u^\beta |\sigma_{(i-2)\Delta_n-}|^\beta}{\Delta_n^{-1}\left(|\overline{\sigma}|_i^p \overline{V}_i^n(p)\right)^{\beta/p}}\right) - \exp\left(-\frac{C_{p,\beta} u^\beta |\sigma_{(i-2)\Delta_n-}|^\beta}{(|\overline{\sigma}|_i^p)^{\beta/p}}\right),$$

and $R_j^n(u) = \sum_{i=k_n+3}^n r_i^j(u)$ for $j = 1, 2, 3, 4$ with

$$r_i^1(u) = \cos\left(u \frac{\Delta_i^n X - \Delta_{i-1}^n X}{(V_i^n(p))^{1/p}}\right) - \cos\left(u \frac{\sigma_{(i-2)\Delta_n-}(\Delta_i^n S - \Delta_{i-1}^n S)}{(V_i^n(p))^{1/p}}\right),$$

$$r_i^2(u) = \exp\left(-\frac{A_\beta u^\beta |\sigma_{(i-2)\Delta_n-}|^\beta}{\Delta_n^{-1}(V_i^n(p))^{\beta/p}}\right) - \exp\left(-\frac{A_\beta u^\beta |\sigma_{(i-2)\Delta_n-}|^\beta}{\Delta_n^{-1}(\widehat{V}_i^n(p))^{\beta/p}}\right),$$

$$r_i^3(u) = \exp\left(-\frac{A_\beta u^\beta |\sigma_{(i-2)\Delta_n-}|^\beta}{\Delta_n^{-1}(\widehat{V}_i^n(p))^{\beta/p}}\right) - \exp\left(\frac{C_{p,\beta} u^\beta |\sigma_{(i-2)\Delta_n-}|^\beta}{\Delta_n^{-1}\left(|\overline{\sigma}|_i^p \overline{V}_i^n(p)\right)^{\beta/p}}\right),$$

$$r_i^4(u) = \exp\left(-\frac{C_{p,\beta} u^\beta |\sigma_{(i-2)\Delta_n-}|^\beta}{(|\overline{\sigma}|_i^p)^{\beta/p}}\right) - \exp\left(-C_{p,\beta} u^\beta\right).$$

We finally introduce the following: $\overline{Z}_1^n(u) = \sum_{i=k_n+3}^n \overline{z}_i^1(u)$, $\overline{Z}_2^{(a,n)}(u) = \sum_{i=k_n+3}^n \overline{z}_i^{(a,2)}(u)$ and $\overline{Z}_2^{(b,n)}(u) = \sum_{i=k_n+3}^n \overline{z}_i^{(b,2)}(u)$ where

$$\overline{z}_i^1(u) = \cos\left(u\Delta_n^{-1/\beta}\mu_{p,\beta}^{-1/\beta}(\Delta_i^n S - \Delta_{i-1}^n S)\right) - \mathcal{L}(p,u,\beta),$$

$$\overline{z}_i^{(a,2)}(u) = G(p,u,\beta)\left(\Delta_n^{-p/\beta}\overline{V}_i^n(p) - 1\right), \ \overline{z}_i^{(b,2)}(u) = \frac{1}{2}H(p,u,\beta)\left(\Delta_n^{-p/\beta}\overline{V}_i^n(p) - 1\right)^2.$$

9.2. *Localization.* We prove results under the following strengthened version of Assumption B:

Assumption SB. *We have assumption B and in addition*

(a) *the processes $|\sigma_t|$ and $|\sigma_t|^{-1}$ are uniformly bounded;*
(b) *the processes b^α and b^σ are uniformly bounded;*
(c) *$|\delta^\alpha(t,x)| + |\delta^\sigma(t,x)| \leq \gamma(x)$ for all t, where $\gamma(x)$ is a deterministic bounded function on $\mathbb{R}$ with $\int_{\mathbb{R}} |\gamma(x)|^{r+\iota}\lambda(dx) < \infty$ for arbitrary small $\iota > 0$ and some $0 \leq r \leq \beta$;*
(d) *the coefficients in the Itô semimartingale representation of b^α and b^σ satisfy the analogues of conditions (b) and (c) above;*
(e) *the process $\int_{\mathbb{R}}(|x|^{\beta'+\iota} \wedge 1)\nu_t^Y(dx)$ is bounded and the jumps of $\widehat{S}$, $\widetilde{S}$ and Y are bounded.*

Extending the results to the case of the more general Assumption B follows by standard localization arguments given in Section 4.4.1 of [12].

9.3. *Preliminary Results.* The strategy of the proofs is to bound the terms $R_j^n(u)$ for $j = 1,2,3,4$ as well as $\widehat{Z}_1^n(u) - \overline{Z}_1^n(u)$ and $\widehat{Z}_2^n(u) - \overline{Z}_2^{(a,n)}(u) - \overline{Z}_2^{(b,n)}(u)$, and to derive the asymptotic limits of $\overline{Z}_1^n(u)$, $\overline{Z}_2^{(a,n)}(u)$ and $\overline{Z}_2^{(b,n)}(u)$. We do this in a sequence of lemmas starting with one containing some preliminary bounds needed for the subsequent lemmas.

Lemma 1. *Under Assumptions A and SB and $k_n \asymp n^\varpi$ for $\varpi \in (0,1)$, we have for $0 < p < \beta$, $\iota > 0$ arbitrary small, and $1 \leq x < \frac{\beta}{p}$ and $y \geq 1$:*

$$\Delta_n^{-p/\beta}\mathbb{E}|V_i^n(p)) - \widehat{V}_i^n(p)| \leq K\alpha_n, \tag{9.3}$$

$$\alpha_n = \frac{\Delta_n^{(2-1/\beta)(1+(p-1/2)\wedge 0-\iota)}}{\sqrt{k_n}} \bigvee \Delta_n^{\frac{1}{\beta}-\iota} \bigvee \Delta_n^{\frac{p}{\beta'}\wedge 1-\frac{p}{\beta}-\iota} \bigvee \Delta_n^{\frac{p+1}{\beta+1}-\iota},$$

(9.4)

$$\mathbb{E}\left|\Delta_n^{-p/\beta}\widehat{V}_i^n(p) - \mu_{p,\beta}^{p/\beta}|\overline{\sigma}|_i^p\right|^x + \mathbb{E}\left|\Delta_n^{-p/\beta}\overline{V}_i^n(p) - 1\right|^x \leq K \begin{cases} k_n^{-x/2}, & \text{if } \beta/p > 2 \\ k_n^{1-x}, & \text{if } \beta/p \leq 2 \end{cases},$$

$$\left|\mathbb{E}_{i-k_n-3}^n(|\overline{\sigma}|_i^p - |\sigma_{(i-2)\Delta_n-}|^p)\right| \leq K k_n \Delta_n, \tag{9.5}$$

$$\mathbb{E}_{i-k_n-3}^n \left||\overline{\sigma}|_i^p - |\sigma_{(i-2)\Delta_n-}|^p\right|^y \leq K(k_n\Delta_n)^{\frac{y}{r}\wedge 1-\iota}, \tag{9.6}$$

$$\Delta_n^{-p/\beta}\left|\mathbb{E}_{i-k_n-3}^n\left(\widehat{V}_i^n(p) - \mu_{p,\beta}^{p/\beta}|\overline{\sigma}|_i^p\overline{V}_i^n(p) - \dot{V}_i^n(p)\right)\right| \leq K k_n \Delta_n, \tag{9.7}$$

$$\Delta_n^{-xp/\beta}\mathbb{E}\left|\widehat{V}_i^n(p) - \mu_{p,\beta}^{p/\beta}|\overline{\sigma}|_i^p\overline{V}_i^n(p) - \dot{V}_i^n(p)\right|^x \leq K(k_n\Delta_n)^{\frac{x}{r}\wedge 1-\iota}, \tag{9.8}$$

$$\Delta_n^{-xp/\beta}\mathbb{E}\left|\dot{V}_i^n(p)\right|^x \leq K\Delta_n^{\frac{\beta-xp}{\beta}\wedge\frac{x}{r}-\iota}. \tag{9.9}$$

Proof of Lemma 1. We start with (9.3). We apply exactly the same decomposition and bounds as for the term A_3 in Section 5.2.3 in [22] to get the result in (9.3). We continue with (9.4). Without loss of generality we assume $k_n \geq 2$ and we denote the two sets

$$\begin{cases} J_i^e = \left\{i - k_n - 1 + 2k : k = 0, ..., \lfloor\frac{k_n-1}{2}\rfloor\right\}, \\ J_i^o = \left\{i - k_n - 1 + 2k + 1 : k = 0, ..., \lfloor\frac{k_n-2}{2}\rfloor\right\}. \end{cases}$$

With this notation, we can decompose $\widehat{V}_i^n(p)$ into

$$\widehat{V}_i^{(e,n)}(p) = \frac{1}{k_n}\sum_{j\in J_i^e}|\sigma_{(j-2)\Delta_n-}|^p|\Delta_j^n S - \Delta_{j-1}^n S|^p, \quad \widehat{V}_i^{(o,n)}(p) = \widehat{V}_i^n(p) - \widehat{V}_i^{(e,n)}(p).$$

We further denote $|\overline{\sigma}|_{e,i}^p = \frac{1}{k_n}\sum_{j\in J_i^e}|\sigma_{(j-2)\Delta_n-}|^p$ and $|\overline{\sigma}|_{o,i}^p = \frac{1}{k_n}\sum_{j\in J_i^o}|\sigma_{(j-2)\Delta_n-}|^p$. Using triangular inequality, we then have

$$\begin{aligned} &\left|\Delta_n^{-p/\beta}\widehat{V}_i^n(p) - \mu_{p,\beta}^{p/\beta}|\overline{\sigma}|_i^p\right| \\ &\quad\leq \left|\Delta_n^{-p/\beta}\widehat{V}_i^{(e,n)}(p) - \mu_{p,\beta}^{p/\beta}|\overline{\sigma}|_{e,i}^p\right| + \left|\Delta_n^{-p/\beta}\widehat{V}_i^{(o,n)}(p) - \mu_{p,\beta}^{p/\beta}|\overline{\sigma}|_{o,i}^p\right|. \end{aligned}$$

Now, since $\mathbb{E}^n_{j-2}|\Delta^n_j S - \Delta^n_{j-1} S|^p = \Delta_n^{p/\beta}\mu^{p/\beta}_{p,\beta}$, the sums $\Delta_n^{-p/\beta}\widehat{V}_i^{(e,n)}(p) - \mu^{p/\beta}_{p,\beta}|\overline{\sigma}|^p_{e,i}$ and $\Delta_n^{-p/\beta}\widehat{V}_i^{(o,n)}(p) - \mu^{p/\beta}_{p,\beta}|\overline{\sigma}|^p_{o,i}$ are discrete martingales. From here, the result in (9.4) for the case $\beta/p \leq 2$ follows by a direct application of the Burkholder-Davis-Gundy inequality and the algebraic inequality

$$|\sum_i |a_i||^p \leq \sum_i |a_i|^p, \quad \forall p \in (0,1] \text{ and any real-valued } \{a_i\}_{i\geq 1}. \tag{9.10}$$

We are left with the case $\beta/p > 2$. We only show the bound involving the term $\widehat{V}_i^{(e,n)}(p)$, with the result for $\widehat{V}_i^{(o,n)}(p)$ being shown analogously. We first denote $\Delta_n^{-p/\beta}\widehat{V}_i^{(e,n)}(p) - \mu^{p/\beta}_{p,\beta}|\overline{\sigma}|^p_{e,i} = \frac{1}{k_n}\sum_{j\in J_i^e}\zeta_j^n$ where $\zeta_j^n = \Delta_n^{-p/\beta}|\sigma_{(j-2)\Delta_n-}|^p\left(|\Delta^n_j S - \Delta^n_{j-1}S|^p - \mu^{p/\beta}_{p,\beta}\right)$. Applying Burkholder-Davis-Gundy inequality, we have

$$\mathbb{E}|\sum_{j\in J_i^e}\zeta_j^n|^x \leq K\mathbb{E}\left(\sum_{j\in J_i^e}(\zeta_j^n)^2\right)^{x/2}.$$

If $x \leq 2$, the result in (9.4) then follows by Jensen's inequality. If $x > 2$, applying again Burkholder-Davis-Gundy, we have

$$\begin{aligned}
&\mathbb{E}\left(\sum_{j\in J_i^e}(\zeta_j^n)^2\right)^{x/2} \\
&\leq K\mathbb{E}\left(\sum_{j\in J_i^e}((\zeta_j^n)^2 - \mathbb{E}^n_{j-2}(\zeta_j^n)^2)\right)^{x/2} + K\mathbb{E}\left(\sum_{j\in J_i^e}\mathbb{E}^n_{j-2}(\zeta_j^n)^2\right)^{x/2} \\
&\leq K\mathbb{E}\left(\sum_{j\in J_i^e}((\zeta_j^n)^2 - \mathbb{E}^n_{j-2}(\zeta_j^n)^2)^2\right)^{x/4} + Kk_n^{x/2},
\end{aligned} \tag{9.11}$$

where we also made use of the fact that β-stable random variable has finite p-th absolute moment as soon as $p \in (0,\beta)$. If $x \leq 4$, the result will then follow from an application of (9.10). If $x > 4$, then we repeat (9.11) with x replaced by $x/2$ and ζ_j^n replaced $(\zeta_j^n)^2 - \mathbb{E}^n_{j-2}(\zeta_j^n)^2$. We continue this way applying $k = \sup\left\{i : 2^i < x\right\}$ times (9.11) and then (9.10). This shows (9.4).

We continue with (9.5) and (9.6). We make use of the following algebraic inequality

$$\left||a+b|^p - |a|^p - p\text{sign}\{a\}|a|^{p-1}b\right| \leq K_p|a|^{p-2}|b|^2,$$

for any $a, b \in R$ with $a \neq 0$, $0 < p < 1$ and K_p that depends only on p. Applying this inequality as well as triangular inequality, and using the fact that under assumption SB the process $|\sigma|$ is bounded from below, we have

$$|\mathbb{E}_s(|\sigma_t|^p - |\sigma_s|^p)| \leq K|t-s|, \quad 0 \leq s \leq t, \tag{9.12}$$

$$\mathbb{E}_s \left||\sigma_t|^p - |\sigma_s|^p\right|^q \leq K\mathbb{E}_s \left(|\sigma_t - \sigma_s|^q \vee |\sigma_t - \sigma_s|^{2q}\right), \quad 0 \leq s \leq t, \quad q \geq 1, \tag{9.13}$$

with some constant K that does not depend on s and t. From here (9.5) follows. Application of Corollary 2.1.9 of [12] further gives

$$\mathbb{E}_s|\sigma_t - \sigma_s|^q \leq K|t-s|^{\frac{q}{r} \bigwedge 1-\iota}, \quad 0 \leq s \leq t, \quad q \geq 1, \tag{9.14}$$

and applying this inequality with $q = y$ and $q = 2y$, for y the constant in (9.6), we have that result.

We proceed with showing the bounds in (9.7)-(9.9). We can decompose $|\overline{\sigma}|_i^p - |\sigma_{(k-2)\Delta_n-}|^p = \sum_{j=1}^4 a_k^j$ for $k = i - k_n - 1, ..., i-2$ and

$$\begin{cases} a_k^1 = \frac{1}{k_n}\sum_{j=k+3}^{i-2}\left(|\sigma_{(j-2)\Delta_n-}|^p - |\sigma_{k\Delta_n-}|^p\right), \\ a_k^2 = \frac{(i-k-4)\vee 0}{k_n}\left(|\sigma_{k\Delta_n-}|^p - |\sigma_{(k-2)\Delta_n-}|^p\right), \\ a_k^3 = \frac{(|\sigma_{k\Delta_n-}|^p - |\sigma_{(k-2)\Delta_n-}|^p)1_{\{k<i-3\}} + (|\sigma_{(k-1)\Delta_n-}|^p - |\sigma_{(k-2)\Delta_n-}|^p)1_{\{k<i-2\}}}{k_n}, \\ a_k^4 = \frac{1}{k_n}\sum_{j=i-k_n-1}^{k}\left(|\sigma_{(j-2)\Delta_n-}|^p - |\sigma_{(k-2)\Delta_n-}|^p\right), \end{cases}$$

with a_k^1 being zero for $k \geq i-4$. Using law of iterated expectations and the bound in (9.6), we have for $k = i - k_n - 1, ..., i-2$

$$\Delta_n^{-xp/\beta}\mathbb{E}\left(\left|a_k^1 + a_k^4\right| |\Delta_k^n S - \Delta_{k-1}^n S|^p\right)^x \leq K(k_n\Delta_n)^{\frac{x}{r} \bigwedge 1-\iota}. \tag{9.15}$$

Using Hölder inequality, the bound in (9.12), as well as the fact that a stable random variable has finite absolute moments for powers less than β, we have for $k = i - k_n - 1, ..., i-2$

$$\Delta_n^{-xp/\beta}\mathbb{E}\left(\left|a_k^2 + a_k^3\right| |\Delta_k^n S - \Delta_{k-1}^n S|^p\right)^x \leq K\Delta_n^{\left(\frac{\beta x/r}{\beta - xp} \bigwedge 1\right)\frac{\beta - xp}{\beta} - \iota}. \tag{9.16}$$

Combining (9.15) and (9.16), we get the results in (9.8) and (9.9).

Further, using (9.12), we get for $k = i - k_n - 1, ..., i-2$

$$\Delta_n^{-p/\beta}\left|\mathbb{E}_{i-k_n-3}^n\left((a_k^1 + a_k^4)|\Delta_k^n S - \Delta_{k-1}^n S|^p\right)\right| \leq Kk_n\Delta_n. \tag{9.17}$$

From here we get the result in (9.7). □

LEMMA 2. *Under Assumptions A and SB and $k_n \asymp n^{\varpi}$ for $\varpi \in (0,1)$, we have for $0 < p < \beta$, $\iota > 0$ arbitrary small, and every $0 < a < b < \infty$:*

$$\frac{1}{n-k_n-2}\mathbb{E}\left(\sup_{u\in[a,b]}|R_1^n(u)|\right) \leq K_{a,b}\left(\alpha_n \vee k_n^{-\left(\frac{\beta}{2p}\wedge\frac{\beta-p}{p}\right)+\iota}\right), \tag{9.18}$$

$$\frac{1}{n-k_n-2}\mathbb{E}\left(\sup_{u\in[a,b]}|R_2^n(u)|\right) \leq K_{a,b}\alpha_n, \tag{9.19}$$

$$\begin{aligned}&\frac{1}{n-k_n-2}\mathbb{E}\left(\sup_{u\in[a,b]}|R_3^n(u)|\right)\\ &\quad\leq \begin{cases} K_{a,b}\left((k_n\Delta_n)^{1-\iota} \vee k_n^{-1/2}(k_n\Delta_n)^{\frac{1}{r}\wedge\frac{\beta-p}{\beta}-\iota}\right), & \text{if } \beta/p > 2,\\ K_{a,b}\left((k_n\Delta_n)^{\frac{1}{r}\wedge 1-\iota} \vee \Delta_n^{\frac{\beta-p}{\beta}-\iota}\right), & \text{if } \beta/p \leq 2,\end{cases}\end{aligned} \tag{9.20}$$

$$\frac{1}{n-k_n-2}\mathbb{E}\left(\sup_{u\in[a,b]}|R_4^n(u)|\right) \leq K_{a,b}(k_n\Delta_n)^{1-\iota}, \tag{9.21}$$

where $K_{a,b}$ depends only on a and b and is finite-valued.

Proof of Lemma 2. We start with showing (9.18). We define the set

$$\mathcal{C}_i^n = \left\{\left|\Delta_n^{-p/\beta}V_i^n(p) - \mu_{p,\beta}^{p/\beta}|\overline{\sigma}|_i^p\right| > \frac{1}{2}\mu_{p,\beta}^{p/\beta}|\overline{\sigma}|_i^p\right), \quad i = k_n+3, ..., n,$$

and then we note that

$$\begin{aligned}1_{\{\mathcal{C}_i^n\}} \leq\ & 1\left(\Delta_n^{-p/\beta}|V_i^n(p) - \widehat{V}_i^n(p)| > \frac{1}{4}\mu_{p,\beta}^{p/\beta}|\overline{\sigma}|_i^p\right)\\ &+ 1\left(|\Delta_n^{-p/\beta}\widehat{V}_i^n(p) - \mu_{p,\beta}^{p/\beta}|\overline{\sigma}|_i^p| > \frac{1}{4}\mu_{p,\beta}^{p/\beta}|\overline{\sigma}|_i^p\right).\end{aligned}$$

Hence, we can apply (9.3) and (9.4) and conclude

$$\mathbb{E}\left[\sup_{u\in\mathbb{R}_+}(|r_i^1(u)|1_{\{\mathcal{C}_i^n\}})\right] \leq K\left(\alpha_n \vee k_n^{-\left(\frac{\beta}{2p}\wedge\frac{\beta-p}{p}\right)+\iota}\right). \tag{9.22}$$

We proceed with a sequence of inequalities. First, from assumption SB

$$\mathbb{E}_{i-2}^n\left|\int_{(i-1)\Delta_n}^{i\Delta_n}(\alpha_u - \alpha_{u-\Delta_n})du\right| \leq K\Delta_n^{1+\frac{1}{r\vee 1}-\iota}. \tag{9.23}$$

Next, if $\beta' < 1$ we can decompose

$$\widehat{S}_t = \int_0^t \int_{\mathbb{R}} x\mu_1(ds, dx) - t\int_{\mathbb{R}} \kappa(x) 2|\nu'(x)| 1_{\{\nu'(x)<0\}} dx, \tag{9.24}$$

and separate accordingly $\int_{(i-1)\Delta_n}^{i\Delta_n} \sigma_{u-} d\widehat{S}_u$ and $\int_{(i-2)\Delta_n}^{(i-1)\Delta_n} \sigma_{u-} d\widehat{S}_u$. For the difference of the integrals against time, we can proceed exactly as in (9.23). Further, using the the algebraic inequality in (9.10), as well as Assumption A for the measure ν', we have

$$\mathbb{E}_{i-1}^n \left| \int_{(i-1)\Delta_n}^{i\Delta_n} \int_{\mathbb{R}} \sigma_{u-} x \mu_1(du, dx) \right|^x \leq K\Delta_n^{x/\beta' - \iota}, \quad \text{for } x \leq \beta'. \tag{9.25}$$

When $\beta' \geq 1$, we can apply Burkholder-Davis-Gundy inequality and get

$$\mathbb{E}_{i-1}^n \left| \int_{(i-1)\Delta_n}^{i\Delta_n} \sigma_{u-} x d\widehat{S}_u \right|^x \leq K\Delta_n^{x/\beta' - \iota}, \quad \text{for } x \leq \beta'. \tag{9.26}$$

Exactly the same inequalities hold for the analogous integrals involving $\widetilde{S}$. Next, application of Burkholder-Davis-Gundy and Hölder inequalities, as well as Assumption SB yields

$$\mathbb{E}_{i-2}^n \left| \int_{(i-1)\Delta_n}^{i\Delta_n} (\sigma_{u-} - \sigma_{(i-2)\Delta_n -}) \kappa(x) \widetilde{\mu}(du, dx) \right| \leq K\Delta_n^{\frac{2}{\beta} - \iota}. \tag{9.27}$$

Finally, denoting $\kappa'(x) = x - \kappa(x)$ and upon noting that $\kappa'(x)$ is zero for x sufficiently close to zero, we have

$$\mathbb{E}_{i-2}^n \left| \int_{(i-1)\Delta_n}^{i\Delta_n} (\sigma_{u-} - \sigma_{(i-2)\Delta_n -}) \kappa'(x) \mu(du, dx) \right|^\iota \leq K\Delta_n, \quad \forall \iota > 0. \tag{9.28}$$

Combining the estimates in (9.23)-(9.28), as well as the inequality $|\cos(x) - \cos(y)| \leq 2|x-y|^p$ for every $x, y \in \mathbb{R}$ and $p \in (0, 1]$, we have

$$\mathbb{E}\left[\sup_{u \geq a} (|r_i^1(u)| 1_{\{(\mathcal{C}_i^n)^c\}})\right] \leq K_a \left(\Delta_n^{\frac{\beta - \beta'}{\beta(\beta' \vee 1)} - \iota} \vee \Delta_n^{\frac{1}{\beta} \wedge \frac{1}{r \vee 1} - \iota} \right). \tag{9.29}$$

(9.22) and (9.29) yield (9.18). We continue next with (9.19). This bound follows from a first-order Taylor expansion of $f_{i,u}(x)$ and the bounds in (9.2) and (9.3).

We proceed with showing the result for $R_4^n(u)$. Using second-order Taylor expansion and Cauchy-Schwarz inequality, as well as (9.6), we get

$$\mathbb{E}\left(\sup_{u\in[a,b]}\left|R_4^n(u)-\frac{\beta}{p}e^{-C_{p,\beta}u^\beta}C_{p,\beta}u^\beta\sum_{i=k_n+3}^{n}\widetilde{r}_i^4\right|\right)\leq Kk_n, \tag{9.30}$$

where

$$\widetilde{r}_i^4=\frac{|\sigma_{(i-2)\Delta_n-}|^p-|\overline{\sigma}|_i^p}{|\sigma_{(i-k_n-3)\Delta_n-}|^p}.$$

Using (9.5), we have

$$\mathbb{E}\left|\sum_{i=k_n+3}^{n}\mathbb{E}_{i-k_n-3}^n(\widetilde{r}_i^4)\right|\leq Kk_n. \tag{9.31}$$

Further, without loss of generality (because $k_n\Delta_n\to 0$) we assume $n\geq 2k_n+3$. Using the shorthand $\chi_i=\widetilde{r}_i^4-\mathbb{E}_{i-k_n-3}^n(\widetilde{r}_i^4)$, we then decompose

$$\sum_{i=k_n+3}^{n}\chi_i=\sum_{j=1}^{k_n+1}A_j+\sum_{i=2k_n+4+\left(\lfloor\frac{n-k_n-2}{k_n+1}\rfloor-1\right)(k_n+1)}^{n}\chi_i,$$

$$A_j=\sum_{i=1}^{\lfloor\frac{n-k_n-2}{k_n+1}\rfloor}\chi_{k_n+3+(j-1)+(i-1)(k_n+1)},\quad j=1,...,k_n+1.$$

Applying Burkholder-Davis-Gundy inequality for discrete martingales and making use of (9.6), we have

$$\mathbb{E}|A_j|\leq K(k_n\Delta_n)^{-\iota},\quad j=1,...,k_n+1. \tag{9.32}$$

Combining (9.30) and (9.32), we get the bound in (9.21).

We are left with (9.20). The case $\beta/p\leq 2$ follows from

$$\mathbb{E}|r_i^3(u)|\leq K_{a,b}\left|\Delta_n^{-p/\beta}\widehat{V}_i^n(p)-\mu_{p,\beta}^{p/\beta}\overline{V}_i^n(p)\right|,$$

and applying the bounds in (9.8)-(9.9). We now show (9.20) for the case $\beta/p>2$. We first decompose $r_i^3(u)=\sum_{j=1}^3\varrho_i^j(u)$, where

$$\varrho_i^1(u)=f'_{i,u}\left(\Delta_n^{-p/\beta}\overline{V}_i^n(p)|\overline{\sigma}|_i^p\right)\Delta_n^{-p/\beta}\left(\mu_{p,\beta}^{-p/\beta}\widehat{V}_i^n(p)-|\overline{\sigma}|_i^p\overline{V}_i^n(p)-\mu_{p,\beta}^{-p/\beta}\dot{V}_i^n(p)\right),$$

$$\varrho_i^2(u)=f'_{i,u}\left(\widetilde{x}\right)\Delta_n^{-p/\beta}\mu_{p,\beta}^{-p/\beta}\dot{V}_i^n(p),$$

$$\varrho_i^3(u) = \left(f'_{i,u}(\widetilde{x}) - f'_{i,u}\left(\Delta_n^{-p/\beta}\overline{V}_i^n(p)|\overline{\sigma}|_i^p\right)\right) \times \Delta_n^{-p/\beta}\left(\mu_{p,\beta}^{-p/\beta}\widehat{V}_i^n(p) - |\overline{\sigma}|_i^p\overline{V}_i^n(p) - \mu_{p,\beta}^{-p/\beta}\dot{V}_i^n(p)\right),$$

and $\widetilde{x}$ is a random number between $\Delta_n^{-p/\beta}\mu_{p,\beta}^{-p/\beta}\widehat{V}_i^n(p)$ and $\Delta_n^{-p/\beta}|\overline{\sigma}|_i^p\overline{V}_i^n(p)$. We further introduce

$$\widetilde{\varrho}_i^1(u) = \frac{G(p,u,\beta)}{|\sigma_{(i-k_n-3)\Delta_n-}|^p}\Delta_n^{-p/\beta}\left(\mu_{p,\beta}^{-p/\beta}\widehat{V}_i^n(p) - |\overline{\sigma}|_i^p\overline{V}_i^n(p) - \mu_{p,\beta}^{-p/\beta}\dot{V}_i^n(p)\right),$$

and note $G(p,u,\beta) = |\sigma_{(i-2)\Delta_n-}|^p f'_{i,u}\left(|\sigma_{(i-2)\Delta_n-}|^p\right)$. Then direct calculation for the function $xf'_{i,u}(x)$ and the boundedness of the process $|\sigma|$ yields

$$|\varrho_i^1(u) - \widetilde{\varrho}_i^1(u)| \le K_{a,b}(d_i^{(1)} + d_i^{(2)})e_i,$$

where

$$\begin{cases} d_i^{(1)} = |\Delta_n^{-p/\beta}\overline{V}_i^n(p) - 1|, \\ d_i^{(2)} = ||\overline{\sigma}|_i^p - |\sigma_{(i-2)\Delta_n-}|^p| + ||\sigma_{(i-2)\Delta_n-}|^p - |\sigma_{(i-k_n-3)\Delta_n-}|^p|, \\ e_i = \Delta_n^{-p/\beta}\left|\mu_{p,\beta}^{-p/\beta}\widehat{V}_i^n(p) - |\overline{\sigma}|_i^p\overline{V}_i^n(p) - \mu_{p,\beta}^{-p/\beta}\dot{V}_i^n(p)\right|. \end{cases}$$

From here, we use Hölder inequality and (9.4), (9.6) and (9.8), to get

$$\begin{cases} \mathbb{E}|d_i^{(1)}e_i| \le K\left(\mathbb{E}\left[\left(d_i^{(1)}\right)^{\frac{\beta}{p+\beta\iota}}\right]\right)^{\frac{p}{\beta}+\iota}\left(\mathbb{E}\left(e_i^{\frac{\beta}{\beta-p-\beta\iota}}\right)\right)^{\frac{\beta-p}{\beta}-\iota} \\ \qquad\qquad \le Kk_n^{-1/2}(k_n\Delta_n)^{\frac{1}{r}\wedge\frac{\beta-p}{\beta}-2\iota}, \\ \mathbb{E}|d_i^{(2)}e_i| \le \sqrt{\mathbb{E}(d_i^{(2)})^2\mathbb{E}(e_i)^2} \le K(k_n\Delta_n)^{1-\iota}. \end{cases} \tag{9.33}$$

For the sum $\sum_{i=k_n+3}^n \widetilde{\varrho}_i^1(u)$, using the bounds in (9.7) and (9.8), we can proceed exactly as for the analysis of $\sum_{i=k_n+3}^n \chi_i$ above and split it into k_n+1 terms which are the terminal values of discrete martingales. This altogether yields

$$\mathbb{E}\left(\sup_{u\in[a,b]}\left|\sum_{i=k_n+3}^n \widetilde{\varrho}_i^1(u)\right|\right) \le K_{a,b}k_n(k_n\Delta_n)^{-\iota}. \tag{9.34}$$

Next, using the bound in (9.9) as well as the boundedness of the derivative $f'_{i,u}(x)$ (for $u\in[a,b]$), we have

$$\mathbb{E}\left(\sup_{u\in[a,b]}|\varrho_i^2(u)|\right) \le K_{a,b}\Delta_n^{\frac{\beta-p}{\beta}\wedge\frac{1}{r}-\iota}. \tag{9.35}$$

We continue with the term $\varrho_i^3(u)$. We first introduce the set

$$\mathcal{E}_i^n = \left\{ \left| \mu_{p,\beta}^{-p/\beta} \widehat{V}_i^n(p) - |\overline{\sigma}|_i^p \overline{V}_i^n(p) - \mu_{p,\beta}^{-p/\beta} \dot{V}_i^n(p) \right| > 1 \right\}, \quad i = k_n + 3, ..., n.$$

With this notation, using (9.8) and the boundedness of the derivative $f'_{i,u}(x)$ (for $u \in [a,b]$), we have

$$\mathbb{E}\left(\sup_{u\in[a,b]} |\varrho_i^3(u)| 1_{\{\mathcal{E}_i^n\}} \right) \leq K_{a,b}(k_n\Delta_n)^{1-\iota}. \tag{9.36}$$

Next using the boundedness of the second derivative $f''_{i,u}(x)$, as well as the bounds in (9.8) and (9.9), we get

$$\mathbb{E}\left(\sup_{u\in[a,b]} |\varrho_i^3(u)| 1_{\{(\mathcal{E}_i^n)^c\}} \right) \leq K_{a,b}\left((k_n\Delta_n)^{1-\iota} \vee \Delta_n^{\frac{\beta-p}{\beta} \wedge \frac{1}{r} - \iota} \right). \tag{9.37}$$

Combining (9.33)-(9.37), we get the result in (9.20). □

Lemma 3. *Under Assumptions A and SB and $k_n \asymp n^{\varpi}$ for $\varpi \in (0,1)$, we have for $0 < p < \beta$, $\iota > 0$ arbitrary small, and every $0 < a < b < \infty$:*

$$\frac{1}{n-k_n-2} \sup_{u\in[a,b]} |\widehat{Z}_1^n(u) - \overline{Z}_1^n(u)| = o_p\left(\alpha_n \vee k_n^{-\left(\frac{\beta}{2p} \wedge \frac{\beta-p}{p}\right)+\iota} \vee \sqrt{\Delta_n} \right), \tag{9.38}$$

and further if $p < \beta/2$

$$\begin{aligned} &\frac{1}{n-k_n-2} \sup_{u\in[a,b]} |\widehat{Z}_2^n(u) - \overline{Z}_2^{(a,n)}(u) - \overline{Z}_2^{(b,n)}(u)| \\ &\qquad = o_p\left(k_n^{-\frac{1}{2}\left(\frac{\beta}{p}\wedge 3\right)+\iota} \vee k_n^{-1/2}(k_n\Delta_n)^{\frac{1}{r}\wedge\frac{\beta-p}{\beta}-\iota} \right). \end{aligned} \tag{9.39}$$

Proof of Lemma 3. We start with (9.38). We split $\widehat{Z}_1^n(u) - \overline{Z}_1^n(u) = E_1^n(u) + E_2^n(u)$ with $E_1^n(u) = \sum_{i=k_n+3}^n (z_i^1(u) - \overline{z}_i^1(u)) 1_{\{\mathcal{C}_i^n\}}$ and $E_2^n(u) = \sum_{i=k_n+3}^n (z_i^1(u) - \overline{z}_i^1(u)) 1_{\{(\mathcal{C}_i^n)^c\}}$. For $E_1^n(u)$, using Lemma 1, we easily have

$$\frac{1}{n-k_n-2}\mathbb{E}\left(\sup_{u\in[a,b]} |E_1^n(u)| \right) \leq K_{a,b}\left(\alpha_n \vee k_n^{-\left(\frac{\beta}{2p}\wedge\frac{\beta-p}{p}\right)+\iota} \right). \tag{9.40}$$

We proceed with $E_2^n(u)$. We first note that

$$\mathbb{E}_{i-2}^n\left[(z_i^1(u) - \overline{z}_i^1(u)) 1_{\{(\mathcal{C}_i^n)^c\}} \right] = 0. \tag{9.41}$$

Further, using the algebraic inequalities $|\cos(x)-\cos(y)|^2 \le 2|x-y|$ for $x,y\in\mathbb{R}$ and $|e^{-x}-e^{-y}|^2 \le 2|x-y|$ for $x,y\in\mathbb{R}_+$, as well as the definition of the set $\mathcal{C}_i^n$, we get

$$\mathbb{E}_{i-2}^n\left|(z_i^1(u)-\overline{z}_i^1(u))1_{\{(\mathcal{C}_i^n)^c\}}\right|^2 \le K_{a,b}\left|\Delta_n^{-p/\beta}V_i^n(p)-\mu_{p,\beta}^{p/\beta}|\sigma_{(i-2)\Delta_n-}|^p\right|.$$

Applying the above two inequalities, the bounds in (9.3), (9.4) and (9.6), as well as the algebraic inequality $2xy\le x^2+y^2$ for $x,y\in\mathbb{R}$, we have

$$\begin{aligned}\mathbb{E}\left(E_2^n(u)\right)^2 &= \mathbb{E}\left(\sum_{i=k_n+3}^n (z_i^1(u)-\overline{z}_i^1(u))^2 1_{\{(\mathcal{C}_i^n)^c\}}\right)\\ &\quad+\mathbb{E}\left(\sum_{i,j:|i-j|=1}(z_i^1(u)-\overline{z}_i^1(u))1_{\{(\mathcal{C}_i^n)^c\}}(z_j^1(u)-\overline{z}_j^1(u))1_{\{(\mathcal{C}_j^n)^c\}}\right)\\ &\le K_{a,b}\sum_{i=k_n+3}^n \mathbb{E}\left|\Delta_n^{-p/\beta}V_i^n(p)-\mu_{p,\beta}^{p/\beta}|\sigma_{(i-2)\Delta_n-}|^p\right|\\ &\le K_{a,b}\Delta_n^{-1}\left(\alpha_n\vee k_n^{-\left(\frac{\beta}{2p}\wedge\frac{\beta-p}{p}\right)+\iota}\vee(k_n\Delta_n)^{\frac{1}{r}\wedge 1-\iota}\right).\end{aligned}$$

As a result, $\frac{1}{\sqrt{n-k_n-2}}E_2^n(u)\xrightarrow{\mathbb{P}}0$ finite-dimensionally in u. We are left with showing that the convergence holds uniformly in $u\in[a,b]$. For this we apply a criteria for tightness on the space of continuous functions equipped with the uniform topology, see e.g., Theorem 12.3 of [7]. Using again (9.41), we have

$$\begin{aligned}&\mathbb{E}\left(E_2^n(u)-E_2^n(v)\right)^2\\ &\quad\le K\mathbb{E}\left(\sum_{i=k_n+3}^n (z_i^1(u)-\overline{z}_i^1(u)-z_i^1(v)+\overline{z}_i^1(v))^2 1_{\{(\mathcal{C}_i^n)^c\}}\right),\end{aligned}$$

hence for arbitrary small $\iota>0$

$$\begin{aligned}&\frac{1}{n-k_n-2}\mathbb{E}\left(\sum_{i=k_n+3}^n (z_i^1(u)-\overline{z}_i^1(u)-z_i^1(v)+\overline{z}_i^1(v))^2 1_{\{(\mathcal{C}_i^n)^c\}}\right)\\ &\qquad\le K\left\{|u^\beta-v^\beta|^2\vee|u-v|^{\beta-\iota}\right\},\end{aligned}$$

and since $\beta>1$, we have $\frac{1}{\sqrt{n-k_n-2}}\sup_{u\in[a,b]}|E_2^n(u)|\xrightarrow{\mathbb{P}}0$. We turn next to (9.39). We first introduce some additional notation. Based on a second-order Taylor expansion of the function $f_{i,u}(x)$, we can further decompose

$\widehat{Z}_2^n(u) = \widehat{Z}_2^{(a,n)}(u) + \widehat{Z}_2^{(b,n)}(u) + \widehat{Z}_2^{(c,n)}(u)$, with $\widehat{Z}_2^{(k,n)}(u) = \sum_{i=k_n+3}^n z_i^{(k,2)}(u)$ for $k = a, b, c$, where $z_i^{(c,2)}(u) = z_i^2(u) - z_i^{(a,2)}(u) - z_i^{(b,2)}(u)$ and

$$z_i^{(a,2)}(u) = f'_{i,u}\left(|\overline{\sigma}|_i^p\right)|\overline{\sigma}|_i^p\left(\Delta_n^{-p/\beta}\overline{V}_i^n(p) - 1\right),$$

$$z_i^{(b,2)}(u) = \frac{1}{2}f''_{i,u}\left(|\overline{\sigma}|_i^p\right)\left(|\overline{\sigma}|_i^p\right)^2\left(\Delta_n^{-p/\beta}\overline{V}_i^n(p) - 1\right)^2,$$

Note further that

$$\begin{cases} |\sigma_{(i-2)\Delta_n-}|^p f'_{i,u}\left(|\sigma_{(i-2)\Delta_n-}|^p\right) = G(p,u,\beta), \\ |\sigma_{(i-2)\Delta_n-}|^{2p} f''_{i,u}\left(|\sigma_{(i-2)\Delta_n-}|^p\right) = H(p,u,\beta). \end{cases}$$

Direct calculation, and using the boundedness of the process σ by Assumption SB, shows

$$\begin{aligned} &\left||\overline{\sigma}|_i^p f'_{i,u}\left(|\overline{\sigma}|_i^p\right) - G(p,u,\beta)\right| + \left|\left(|\overline{\sigma}|_i^p\right)^2 f''_{i,u}\left(|\overline{\sigma}|_i^p\right) - H(p,u,\beta)\right| \\ &\qquad \le K_{a,b}\left||\overline{\sigma}|_i^p - |\sigma_{(i-2)\Delta_n-}|^p\right|, \quad u \in [a,b], \quad i = k_n+3, ..., n, \end{aligned}$$

for some finite-valued constant $K_{a,b}$ which depends only a and b. From here, using the bounds in (9.4) and (9.6), we have

$$\mathbb{E}\left(\sup_{u\in[a,b]}|z_i^{(a,2)}(u) - \overline{z}_i^{(a,2)}(u)|\right) \le K_{a,b}\left(k_n^{-1/2}(k_n\Delta_n)^{\frac{1}{r}\wedge\frac{\beta-p}{\beta}-\iota}\right),$$

and similarly

$$\mathbb{E}\left(\sup_{u\in[a,b]}|z_i^{(b,2)}(u) - \overline{z}_i^{(b,2)}(u)|\right) \le K_{a,b}\left(k_n^{-\frac{\beta}{2p}+\iota} \vee k_n^{-1/2}(k_n\Delta_n)^{\frac{1}{r}\wedge\frac{\beta-p}{\beta}-\iota}\right).$$

Therefore,

$$\begin{aligned} &\frac{1}{n-k_n-2}\sup_{u\in[a,b]}\left|\widehat{Z}_2^{(a,n)}(u) - \overline{Z}_2^{(a,n)}(u) + \widehat{Z}_2^{(b,n)}(u) - \overline{Z}_2^{(b,n)}(u)\right| \\ &\qquad = o_p\left(k_n^{-\frac{\beta}{2p}+\iota} \vee k_n^{-1/2}(k_n\Delta_n)^{\frac{1}{r}\wedge\frac{\beta-p}{\beta}-\iota}\right). \end{aligned} \tag{9.42}$$

We are left with $\widehat{Z}_2^{(c,n)}(u)$. Using the boundedness of the derivatives in (9.2), we have

$$\left|z_i^{(c,2)}(u)\right| \le K_{a,b}\left|\Delta_n^{-p/\beta}\overline{V}_i^n(p) - 1\right|^x, \quad 2 < x < \beta/p \wedge 3.$$

From here, applying (9.19), we have

$$\frac{1}{n-k_n-2}\sup_{u\in[a,b]}\left|\widehat{Z}_2^{(c,n)}(u)\right| = o_p\left(k_n^{-\frac{1}{2}\left(\frac{\beta}{p}\wedge 3\right)+\iota}\right). \tag{9.43}$$

Combining the results in (9.42) and (9.43), we get (9.39). □

LEMMA 4. *Let $p\in(0,\beta/2)$. If $k_n \asymp n^{\varpi}$ for $\varpi\in(0,1)$, we have*

$$\frac{1}{\sqrt{n-k_n-2}}\begin{pmatrix}\overline{Z}_1^n(\mathbf{u})\\ \overline{Z}_2^{(a,n)}(\mathbf{u})\end{pmatrix} \xrightarrow{\mathcal{L}} \zeta(\mathbf{u}), \tag{9.44}$$

where $\zeta(\mathbf{u})$ is a Gaussian process with covariance function given by

$$\begin{pmatrix}1\\ G(p,u,\beta)\end{pmatrix}'\overline{\Xi}(p,u,v,\beta)\begin{pmatrix}1\\ G(p,v,\beta)\end{pmatrix},\quad u,v\in\mathbb{R}_+, \tag{9.45}$$

for $\overline{\Xi}(p,u,v,\beta)=\Xi_0(p,u,v,\beta)+2\Xi_1(p,u,v,\beta)$. The convergence in (9.44) is in the space of continuous functions $\mathbb{R}_+\to\mathbb{R}^2$ equipped with the local uniform topology. The convergence result for $\overline{Z}_1^n(\mathbf{u})$ in (9.44) continues to hold for $p\in[\beta/2,\beta)$.

Further, for some $\iota>0$,

$$\begin{aligned}\frac{k_n}{n-k_n-2}\overline{Z}_2^{(b,n)}(u)-\frac{1}{2}H(p,u,\beta)\left(\Xi_0^{(2,2)}(p,u,u,\beta)+2\Xi_1^{(2,2)}(p,u,u,\beta)\right)\\ = o_p\left((k_n\Delta_n)^{1-\frac{2p}{\beta}\vee\frac{1}{2}-\iota}\right),\end{aligned} \tag{9.46}$$

locally uniformly in $u\in\mathbb{R}_+$.

Proof of Lemma 4. We can write

$$\begin{pmatrix}\overline{Z}_1^n(u)\\ \overline{Z}_2^{(a,n)}(u)\end{pmatrix} = \sum_{i=k_n+1}^{n-k_n-1}\zeta_i(u)+E_l(u)+E_r(u), \tag{9.47}$$

where

$$\zeta_i(u)=\begin{pmatrix}\cos\left(u\Delta_n^{-1/\beta}\mu_{p,\beta}^{-1/\beta}(\Delta_i^nS-\Delta_{i-1}^nS)\right)-\mathcal{L}(p,u,\beta)\\ G(p,u,\beta)\left[\Delta_n^{-p/\beta}\mu_{p,\beta}^{-p/\beta}|\Delta_i^nS-\Delta_{i-1}^nS|^p-1\right]\end{pmatrix},$$

$$E_l(u)=\sum_{i=2}^{k_n}\frac{i-1}{k_n}\begin{pmatrix}0\\ \zeta_i^{(2)}(u)\end{pmatrix}-\sum_{i=k_n+1}^{k_n+2}\begin{pmatrix}\zeta_i^{(1)}(u)\\ 0\end{pmatrix},$$

$$E_r(u) = \sum_{i=n-k_n}^{n-2} \frac{n-1-i}{k_n} \begin{pmatrix} 0 \\ \zeta_i^{(2)}(u) \end{pmatrix} + \sum_{i=n-k_n}^{n} \begin{pmatrix} \zeta_i^{(1)}(u) \\ 0 \end{pmatrix}.$$

We note that for $u \in \mathbb{R}_+$

$$\mathbb{E}_{i-2}^n \left(\zeta_i(u)\right) = 0, \quad i = 2, ..., n. \tag{9.48}$$

Further, making using of the inequality $|\cos(x) - \cos(y)| \leq 2|x-y|^p$ for every $p \in (0,1]$ and $x, y \in \mathbb{R}$, we have for $u, v \in \mathbb{R}_+$

$$\mathbb{E}_{i-2}^n \left(\zeta_i^{(1)}(u) - \zeta_i^{(1)}(v)\right)^2 \leq K|u-v|^p \vee |u^\beta - v^\beta|^2, \quad 1 < p < \beta. \tag{9.49}$$

Making use of (9.48) and the fact that $\zeta_i^{(2)}(u)$ depends on u only through $H(p,u,\beta)$ and $\sup_{u\in\mathbb{R}^+} |H(p,u,\beta)|$ is a finite constant, we have

$$\frac{1}{k_n}\mathbb{E}\left(\sup_{u\in\mathbb{R}_+} |E_l(u)|^2\right) \leq K. \tag{9.50}$$

Making use of (9.49) and the differentiability of $G(p,u,\beta)$ in u, we also have

$$\frac{1}{k_n}\mathbb{E}\left(E_r(u) - E_r(v)\right)^2 \leq |F(u) - F(v)|^p,$$

for some increasing function $F(\cdot)$ and some $p > 1$. Applying then a criteria for tightness on the space of continuous functions equipped with the uniform topology, see e.g., Theorem 12.3 in [7], as well as making use of the fact that $k_n\Delta_n \to 0$, we have locally uniformly in u

$$\frac{1}{\sqrt{n-k_n-2}} E_r(u) \xrightarrow{\mathbb{P}} 0. \tag{9.51}$$

We are left with the first term on the right hand side of (9.47). First, we establish convergence for this term finite-dimensionally in u. We have the decomposition

$$\sum_{i=k_n+1}^{n-k_n-1} \zeta_i(u) = \sum_{i=k_n+1}^{n-k_n-1} (\zeta_i(u) - \mathbb{E}_{i-1}^n(\zeta_i(u))) + \sum_{i=k_n}^{n-k_n-2} \mathbb{E}_i^n(\zeta_{i+1}(u)).$$

From here, we can apply a clt for triangular arrays, see e.g., Theorem 2.2.13 of [12], to establish that $\frac{1}{\sqrt{n-2k_n-1}}\sum_{i=k_n+1}^{n-k_n-1}\zeta_i(u)$ converges finite-dimensionally in u to $\zeta(\mathbf{u})$. This convergence holds also locally uniformly in u using the bound in (9.49) and Theorem VI.4.1 in [14]. Combining the

latter with the asymptotic negligibility results in (9.50) and (9.51), together with the fact that $k_n/n \to 0$, we have the result in (9.44). Furthermore, since $\overline{Z}_1^n(\mathbf{u})$ depends on p only through $\mu_{p,\beta}$, the marginal convergence in (9.44) involving $\overline{Z}_1^n(\mathbf{u})$ holds for any $p \in (0, \beta)$.

We turn next to (9.46). We denote

$$\chi_i = k_n \left(\Delta_n^{-p/\beta} \overline{V}_i^n(p) - 1 \right)^2 - \left(\Xi_0^{(2,2)}(p, u, u, \beta) + 2\Xi_1^{(2,2)}(p, u, u, \beta) \right),$$

and we note that $\Xi_0^{(2,2)}(p, u, u, \beta)$ and $\Xi_1^{(2,2)}(p, u, u, \beta)$ do not depend on u.

Without loss of generality we can assume $n \geq 2k_n + 3$, and then we set

$$A_j = \sum_{i=1}^{\lfloor \frac{n-k_n-2}{k_n+1} \rfloor} \chi_{k_n+3+(j-1)+(i-1)(k_n+1)}, \quad j = 1, ..., k_n + 1.$$

Since $\mathbb{E}|\chi_i| < K$,

$$\left| \sum_{i=k_n+3}^{n} \chi_i - \sum_{j=1}^{k_n+1} A_j \right| = O_p(k_n). \tag{9.52}$$

Further, direct computation shows

$$\mathbb{E}_{i-k_n-3}^n (\chi_i) = 0, \quad i = k_n + 3, ..., n,$$

and applying Burkholder-Davis-Gundy inequality for discrete martingales, we have

$$\mathbb{E}|A_j|^x \leq K(k_n \Delta_n)^{-\left(\frac{x}{2} \vee 1\right)}, \quad 1 \leq x < \frac{\beta}{2p}. \tag{9.53}$$

Using inequality in means we further have

$$\left| \frac{1}{k_n + 1} \sum_{j=1}^{k_n+1} A_j \right|^x \leq \frac{1}{k_n + 1} \sum_{j=1}^{k_n+1} |A_j|^x, \quad 1 \leq x < \frac{\beta}{2p}.$$

Applying the above inequality with x sufficiently close to $\beta/(2p)$ and the bound in (9.53), we have $\Delta_n (k_n \Delta_n)^{\frac{2p}{\beta} \wedge \frac{1}{2} - 1 + \iota} \sum_{j=1}^{k_n+1} A_j \xrightarrow{\mathbb{P}} 0$, and together with the result in (9.52), this implies (9.46). □

9.4. *Proof of Theorems 1 and 2.* Theorem 1 and (4.5) of Theorem 2 follow readily by combining Lemmas 1-4 (and using (9.4) for bounding $\widehat{Z}_2^n(u)$ in the proof of Theorem 1). To show (4.7), we note first that $H(p,u,\beta)$ and $\Xi_i(p,u,u,\beta)$, for $i=0,1$, are continuously differentiable in β. For $H(p,u,\beta)$ this is directly verifiable and for $\Xi_i(p,u,u,\beta)$ with $i=0,1$, this follows from the continuous differentiability of the characteristic function $\beta \to e^{-A_\beta u^\beta}$ for $u \in \mathbb{R}_+$. Moreover, the derivative $\nabla_\beta H(p,u,\beta)$ is bounded in u. From here, (4.7) follows from an application of the continuous mapping theorem.

9.5. *Proof of Theorem 3.* We denote the true value of the parameter β with β_0. Then the claim in (5.7) will follow if we can show the following:

$$\widehat{\mathbf{m}}(p,\widehat{\mathbf{u}},\widehat{\beta}^{fs},\mathbf{u},\beta) \xrightarrow{\mathbb{P}} \mathbf{m}(p,\mathbf{u},\beta), \quad \text{uniformly in } \beta \in [1,2], \tag{9.54}$$

where $\mathbf{m}(p,\mathbf{u},\beta)$ is defined via

$$\mathbf{m}(p,\mathbf{u},\beta)_i = \int_{u_l^i}^{u_h^i} \left(\log(\mathcal{L}(p,u,\beta_0) - \log(\mathcal{L}(p,u,\beta))\right) du,$$

$$\sqrt{n}\, \widehat{\mathbf{m}}(p,\widehat{\mathbf{u}},\widehat{\beta}^{fs},\mathbf{u},\beta_0) \xrightarrow{\mathcal{L}} \mathbf{W}^{1/2}(p,\mathbf{u},\beta_0) \times \mathbf{N}, \tag{9.55}$$

where $\mathbf{N}$ is $K \times 1$ standard normal vector and

$$\mathbf{M}(p,\widehat{\mathbf{u}},\beta) \xrightarrow{\mathbb{P}} \mathbf{M}(p,\mathbf{u},\beta), \quad \text{uniformly in a neighborhood of } \beta_0. \tag{9.56}$$

This is because $\mathbf{m}(p,\mathbf{u},\beta) = \mathbf{0}$ iff $\beta = \beta_0$ and $W(p,\mathbf{u},\beta_0)$ is positive definite. We start with (9.54). We have $\int_{\widehat{u}_l^i}^{\widehat{u}_h^i} \log(\mathcal{L}(p,u,\beta))du \xrightarrow{\mathbb{P}} \int_{u_l^i}^{u_h^i} \log(\mathcal{L}(p,u,\beta))du$ uniformly in $\beta \in [1,2]$ for $i=1,...,K$ because of $\widehat{\mathbf{u}}_l \xrightarrow{\mathbb{P}} \mathbf{u}_l$ and $\widehat{\mathbf{u}}_h \xrightarrow{\mathbb{P}} \mathbf{u}_h$ as well as the continuity of the function u^β in β for every $u \in \mathbb{R}_+$. Exactly the same argument can be used to show (9.56). To show (9.54) it remains to show $\int_{\widehat{u}_l^i}^{\widehat{u}_h^i} \log(\widehat{\mathcal{L}}^n(p,u,\widehat{\beta}^{fs})')du \xrightarrow{\mathbb{P}} \int_{u_l^i}^{u_h^i} \log(\mathcal{L}(p,u,\beta_0))du$ for $i=1,...,K$. Due the continuous differentiability of the de-biasing term in β, $\widehat{\beta}^{fs} \xrightarrow{\mathbb{P}} \beta_0$ and the asymptotic boundedness of $\widehat{\mathbf{u}}_l$ and $\widehat{\mathbf{u}}_h$ and of $\widehat{\mathcal{L}}^n(p,u,\widehat{\beta}^{fs})'$ from below, we have $\int_{\widehat{u}_l^i}^{\widehat{u}_h^i} [\log(\widehat{\mathcal{L}}^n(p,u,\widehat{\beta}^{fs})') - \log(\widehat{\mathcal{L}}^n(p,u,\beta_0))]du \xrightarrow{\mathbb{P}} 0$. From here (9.54) follows by applying Theorem 1.

We are left with (9.55). This result follows from applying the uniform convergence of $\widehat{\mathcal{L}}^n(p,u,\widehat{\beta}^{fs})'$ in Theorem 2.

Finally, (5.8) follows from the continuity of $G(p,\mathbf{u},\beta)$ and $W^{-1}(p,\mathbf{u},\beta)$ in $\mathbf{u}$ and β.

9.6. *Proof of Theorem 4.* We will use the shorthand notation $v_n = \rho u_n$. We start with the following lemma.

LEMMA 5. *Under the conditions of Theorem 4 we have*

$$\widehat{\mathcal{L}}^n(p, u_n, \widehat{\beta}^{fs})' - \mathcal{L}(p, u_n, \widehat{\beta}^{fs}) = O_p(\sqrt{\Delta_n} u_n^2), \tag{9.57}$$

$$\frac{\sqrt{n}}{u_n^2 - v_n^2} \widehat{Z}_n \xrightarrow{\mathcal{L}} \frac{1}{24 C_{p,2}} Z_1 - \frac{2}{p} C_{p,2} Z_2, \tag{9.58}$$

where

$$\widehat{Z}_n = \frac{1}{C_{p,2} u_n^2} \left(\widehat{\mathcal{L}}^n(p, u_n, \widehat{\beta}^{fs})' - \mathcal{L}(p, u_n, \widehat{\beta}^{fs}) \right) - \frac{1}{C_{p,2} v_n^2} \left(\widehat{\mathcal{L}}^n(p, v_n, \widehat{\beta}^{fs})' - \mathcal{L}(p, v_n, \widehat{\beta}^{fs}) \right).$$

Proof of Lemma 5. We use exactly the same decomposition of $\widehat{\mathcal{L}}^n(p, u, \beta) - \mathcal{L}(p, u, \beta)$ as for the proofs of Theorems 1 and 2. We start with the leading terms $\overline{Z}_1^n(u_n)$, $\overline{Z}_2^{(a,n)}(u_n)$ and $\overline{Z}_2^{(b,n)}(u_n)$. Using Taylor series expansion, we have for any $u \in \mathbb{R}_+$ and $Z \in \mathbb{R}$

$$\cos(uZ) - 1 = -\frac{u^2 Z^2}{2} + \frac{u^4 Z^4}{24} + R(uZ), \quad |R(uZ)| \le K |uZ|^6,$$

$$1 - e^{-u^2} = u^2 - \frac{u^4}{2} + O(u^6), \quad \text{as } u \to 0.$$

Using this approximation we have (note that when L_t is a Brownian motion, then $A_\beta = 1$ and so $C_{p,\beta} = 1/\mu_{p,\beta}$)

$$\begin{aligned} &\frac{1}{C_{p,2} u_n^2} \overline{Z}_1^n(u_n) - \frac{1}{C_{p,2} v_n^2} \overline{Z}_1^n(v_n) \\ &\quad = \frac{u_n^2 - v_n^2}{24 C_{p,2}} \sum_{i=k_n+3}^{n} \left[\frac{n^2 (\Delta_i^n S - \Delta_{i-1}^n S)^4}{\mu_{p,2}^2} - \frac{12}{\mu_{p,2}^2} \right] + O_p(u_n^4 \sqrt{n}). \end{aligned} \tag{9.59}$$

We similarly get

$$\begin{aligned} &\frac{1}{C_{p,2} u_n^2} \overline{Z}_2^{(a,n)}(u_n) - \frac{1}{C_{p,2} v_n^2} \overline{Z}_2^{(a,n)}(v_n) \\ &\quad = (v_n^2 - u_n^2) \frac{2}{p} C_{p,2} \sum_{i=k_n+3}^{n} \left(\Delta_n^{-p/2} \overline{V}_i^n(p) - 1 \right) + O_p(u_n^4 \sqrt{n}), \end{aligned} \tag{9.60}$$

and also

(9.61)
$$\frac{k_n}{n-k_n-2}\sum_{i=k_n+3}^{n}\left(\Delta_n^{-p/2}\overline{V}_i^n(p)-1\right)^2-\left(\Xi_0^{(2,2)}(p,u_n,u_n,2)+2\Xi_1^{(2,2)}(p,u_n,u_n,2)\right)$$
$$=O_p\left(\sqrt{k_n\Delta_n}\right).$$

As in Lemma 4, it is easy to show

$$\frac{1}{\sqrt{n-k_n-2}}\sum_{i=k_n+3}^{n}\begin{pmatrix}\frac{n^2(\Delta_i^nS-\Delta_{i-1}^nS)^4}{\mu_{p,2}^2}-\frac{12}{\mu_{p,2}^2}\\ \Delta_n^{-p/2}\overline{V}_i^n(p)-1\end{pmatrix}\xrightarrow{\mathcal{L}}\begin{pmatrix}Z_1\\ Z_2\end{pmatrix}. \tag{9.62}$$

Next, using Taylor expansion as well as $\widehat{\beta}^{fs}-2=o_p(k_nu_n^2\sqrt{\Delta_n})$, we have

$$\frac{\sqrt{n}}{u_n^2k_n}\left(H(p,u_n,\widehat{\beta}^{fs})-H(p,u_n,2)\right)=o_p(1). \tag{9.63}$$

We proceed with the rest of the terms in the decomposition of $\widehat{\mathcal{L}}^n(p,u_n,\beta)-\mathcal{L}(p,u_n,\beta)$ and $\widehat{\mathcal{L}}^n(p,v_n,\beta)-\mathcal{L}(p,v_n\beta)$. We start with the term $R_1^n(u_n)$. It relies on the bound in (9.3), which in turn depends on the analysis of the term A_3 in Section 5.2.3 of [22]. When L is a Brownian motion, the bounds for this term get slightly changed. In particular, the bound in equation (41) of that paper becomes now $K\Delta_n^{1-\iota}$ for $q>r\vee 1$ (this follows from using integration by parts and Burkholder-Davis-Gundy inequality) and arbitrary small $\iota>0$. Using this, it is easy to show that, when L is a Brownian motion, the bound in (9.3) holds with α_n replaced by β_n, where

$$\beta_n=\frac{\Delta_n^{\frac{3}{2}(1+(p-1/2)\wedge 0-\iota)}}{\sqrt{k_n}}\bigvee\Delta_n^{\frac{1}{r\vee 1}-\iota}\bigvee\Delta_n^{\frac{p}{\beta'}\wedge 1-\frac{p}{2}-\iota}\bigvee\Delta_n^{\frac{p+1}{r\vee 1+1}-\iota}.$$

With this the bound for $R_1^n(u_n)$ becomes

$$\mathbb{E}\left|\frac{R_1^n(u_n)}{nu_n^2}\right|\leq K\left(\frac{\beta_n\vee k_n^{-\frac{1}{p}+\iota}}{u_n^2}\right), \tag{9.64}$$

Further, using exactly the same steps as in the proofs of Lemmas 1-3, as well as

$$\sup_{u,x\in\mathbb{R}_+}\left(|u|^p|f_{i,u}'(x)|+|u|^{2p}|f_{i,u}''(x)|\right)<\infty,$$

we get

$$\mathbb{E}\left|\frac{R_2^n(u_n)}{nu_n^2}\right|\leq K\beta_nu_n^{-2},\quad \mathbb{E}\left|\frac{R_4^n(u_n)}{nu_n^2}\right|\leq K(k_n\Delta_n)^{1-\iota}, \tag{9.65}$$

$$\mathbb{E}\left|\frac{R_3^n(u_n)}{nu_n^2}\right| \leq K u_n^{-2-2p}\left((k_n\Delta_n)^{1-\iota} \vee k_n^{-1/2}(k_n\Delta_n)^{\frac{1}{r}\wedge\frac{2-p}{2}-\iota}\right), \tag{9.66}$$

$$\frac{\mathbb{E}\left|\widehat{Z}_1^n(u_n) - \overline{Z}_1^n(u_n)\right|}{nu_n^2} \leq K\left(\frac{(\beta_n \vee k_n^{-\frac{1}{p}+\iota})}{u_n^2} \vee \sqrt{\Delta_n}(k_n\Delta_n)^{1/2-\iota}\right), \tag{9.67}$$

$$\begin{aligned}\frac{\mathbb{E}\left|\widehat{Z}_2^n(u_n) - \overline{Z}_2^{(a,n)}(u_n) - \overline{Z}_2^{(b,n)}(u_n)\right|}{nu_n^2} \\ \leq K\left(\frac{k_n^{-\frac{1}{p}+\iota}}{u_n^{2+2p}} \vee k_n^{-3/2+\iota} \vee k_n^{-1/2}(k_n\Delta_n)^{\frac{1}{r}\wedge\frac{2-p}{2}-\iota}\right).\end{aligned} \tag{9.68}$$

Combining the bounds in (9.64)-(9.68), together with (9.59)-(9.61), the result in (9.62) and (9.63), we establish Lemma 5. We further note that when X is a Lévy process $R_3^n(u)$ and $R_4^n(u)$ are identically zero. □

We proceed with the proof of Theorem 4. Using Taylor expansion and the result in (9.57), $\widehat{Z}_n$ defined in the statement of Lemma 5, is asymptotically equivalent to

$$\frac{1}{C_{p,2}u_n^2}\left(-\log(\widehat{\mathcal{L}}^n(p,u_n,\widehat{\beta}^{fs})') - C_{p,2}u_n^2\right) - \frac{1}{C_{p,2}v_n^2}\left(-\log(\widehat{\mathcal{L}}^n(p,v_n,\widehat{\beta}^{fs})') - C_{p,2}v_n^2\right).$$

Using again Taylor series expansion, the result in (9.57) and that $u_n^{-2}\sqrt{\Delta_n} \to 0$, we have that the above is asymptotically equivalent to

$$\begin{aligned}&\left(\log\left(-\log(\widehat{\mathcal{L}}^n(p,u_n,\widehat{\beta}^{fs})')\right) - \log(C_{p,2}u_n^2)\right) \\ &\qquad - \left(\log\left(-\log(\widehat{\mathcal{L}}^n(p,v_n,\widehat{\beta}^{fs})')\right) - \log(C_{p,2}v_n^2)\right).\end{aligned}$$

From here the result (6.4) in Theorem 4, both in the general and Lévy case, follows from Lemma 5.

Acknowledgements. Work partially supported by NSF grant SES-0957330. I would like to thank the Editor, the Associate Editor, and two anonymous referees for many useful suggestions and comments. I would also like to thank Denis Belomestny, Jose Manuel Corcuera, Valentine Genon-Catalot, Jean Jacod, Cecilia Mancini, Philip Protter, Markus Reiss, Peter Spreij, Mathias Vetter and seminar participants at the workshop on Statistical Inference for Lévy processes at the Lorentz Center, University of Leiden and the workshop on Statistics of High Frequency Data at Humboldt University.

References.

[1] 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.

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

[3] Belomestny, D. (2011a). Spectral Estimation of the Levy Density in Partially Observed Affine Models. *Stochastic Processes and Their Applications 121*, 1217–1244.

[4] Belomestny, D. (2011b). Statistical Inference for Time-Changed Levy Processes via Composite Characteristic Function Estimation. *Annals of Statistics 39*, 2205–2242.

[5] Belomestny, D. and V. Panov (2012). Abelian Theorems for Stochastic Volatility Models with Application to the Estimation of Jump Activity. *Stochastic Processes and Their Applications 123*, 15–44.

[6] Belomestny, D. and V. Panov (2013). Estimation of the Activity of Jumps in Time-Changed Levy Models. *Electronic Journal of Statistics 7*, 2970–3003.

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

[8] Chen, S., A. Delaigle, and P. Hall (2010). Noparametric Estimation for a Class of Levy Processes. *Journal of Econometrics 157*, 257–271.

[9] DuMouchel, W. (1973). On the Asymptotic Normality of the Maximum-Likelihood Estimate when Sampling from a Stable Distribution. *Annals of Statistics 1*, 948–957.

[10] DuMouchel, W. (1983). Estimating the Stable α Index in order to Measure Tail Thickness: A Critique. *Annals of Statistics 11*, 1019–1031.

[11] Fan, J., Y. Li, and Y. Ke (2012). Vast Volatility Matrix Estimation using High Frequency Data for Portfolio Selection. *Journal of the American Statistical Association 107*, 412–428.

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

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

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

[15] Jing, B., X. Kong, and Z. Liu (2011). Estimating the Jump Activity Index Under Noisy Observations Using High-Frequency Data. *Journal of the American Statistical Association 106*, 558–568.

[16] Jing, B., X. Kong, Z. Liu, and P. Mykland (2012). On the Jump Activity Index of Semimartingales. *Journal of Econometrics 166*, 213–223.

[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] Kong, X. (2013). A Direct Estimation Approach to Risk Approximation for Vast Portfolios under Gross-exposure Constraint. *TEST 22*, 647–669.

[19] Kong, X., Z. Liu, and B. Jing. Testing for Pure-Jump Processes for High-Frequency Data. *Annals of Statistics, forthcoming*.

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

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

[22] Todorov, V. (2013). Power Variation from Second Order Differences for Pure Jump Semimartingales. *Stochastic Processes and their Applications 123*, 2899–2920.

[23] Todorov, V. and G. Tauchen (2011). Limit Theorems for Power Variations of Pure-Jump Processes with Application to Activity Estimation. *Annals of Applied Probabil-*

ity 21, 546–588.

[24] Todorov, V. and G. Tauchen (2012). Realized Laplace Transforms for Pure-Jump Semimartingales. *Annals of Statistics 40*, 1233–1262.

[25] Todorov, V. and G. Tauchen (2014). Limit Theorems for the Empirical Distribution Function of Scaled Increments of Ito Semimartingales at High Frequencies. *Annals of Applied Probability 24*, 1850–1888.

[26] Vetter, M. (2012). Estimation of Correlation for Continuous Semimartingales. *Scandinavian Journal of Statistics 39*, 757–771.

[27] Woerner, J. (2003a). Purely Discontinuous Lévy Processes and Power Variation: Inference for the Integrated Volatility and the Scale Paramter. Working paper, University of Oxford.

[28] Woerner, J. (2003b). Variational Sums and Power Variation: A Unifying Approach to Model Selection and Estimation in Semimartingale Models. *Statistics and Decisions 21*, 47–68.

[29] Woerner, J. (2007). Inference in Lévy-type Stochastic Volatility Models. *Advances in Applied Probability 39*, 531–549.

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