Nonparametric Inference for the Spectral Measure of A Bivariate Pure-Jump Semimartingale*

Viktor Todorov†

February 4, 2018

Abstract

We develop a nonparametric estimator for the spectral density of a bivariate pure-jump Itô semimartingale from high-frequency observations of the process on a fixed time interval with asymptotically shrinking mesh of the observation grid. The process of interest is locally stable, i.e., its Lévy measure around zero is like that of a time-changed stable process. The spectral density function captures the dependence between the small jumps of the process and is time invariant. The estimation is based on the fact that the characteristic exponent of the high-frequency increments, up to a time-varying scale, is approximately a convolution of the spectral density and a known function depending on the jump activity. We solve the deconvolution problem in Fourier transform using the empirical characteristic function of locally studentized high-frequency increments and a jump activity estimator.

Keywords: deconvolution, Fourier transform, high-frequency data, Itô semimartingale, nonparametric inference, spectral density.

AMS 2000 subject classifications. 62G07, 62G20, 60H10, 60J75.

*Research partially supported by NSF Grant SES-1530748. I would like to thank the Associate Editor and a referee for many helpful comments and suggestions. Huidi Lin provided excellent research assistance.

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

1 Introduction

In this paper we are interested in studying the dependence between the small jumps of the following bivariate Itô semimartingale of pure-jump type:

$$d\mathbf{X}_t = \mathbf{a}_t dt + d\mathbf{J}_t,\ \mathbf{J}_t = \int_0^t \int_{\mathbb{R}^2} \kappa(\mathbf{x})\widetilde{\mu}(ds, d\mathbf{x}) + \int_0^t \int_{\mathbb{R}^2} \kappa'(\mathbf{x})\mu(ds, d\mathbf{x}), \tag{1.1}$$

where $\mathbf{a}_t$ takes values in $\mathbb{R}^2$; κ is a $\mathbb{R}^2$-valued symmetric C^3 function on $\mathbb{R}^2$ with $\kappa(\mathbf{x}) = \mathbf{x}$ in a neighborhood of the origin and $\kappa(\mathbf{x}) = \mathbf{0}$ when $||\mathbf{x}||$ is sufficiently large; $\kappa'(\mathbf{x}) = \mathbf{x} - \kappa(\mathbf{x})$; μ is an integer-valued measure on $\mathbb{R}_+ \times \mathbb{R}^2$ with compensator (Lévy measure) $dt \otimes \nu_t(d\mathbf{x})$ and for $\mathbf{x} = (r\cos\theta, r\sin\theta)$ with $r \in \mathbb{R}_+$ and $\theta \in [0, 2\pi)$, we have $\nu_t(d\mathbf{x}) = \nu_t(r,\theta)drd\theta$ where

$$\nu_t(r,\theta) = L_{t-}\frac{g(\theta)}{r^{1+\beta}} + \widetilde{\nu}_t(r,\theta),\ \beta \in (0,2), \tag{1.2}$$

with L_t being a positive-valued process with càdlàg paths, g is some bounded and nonnegative-valued function on $[0, 2\pi)$, and $\widetilde{\nu}_t(r,\theta)$ is some predictable and signed function on $\mathbb{R}_+ \times [0, 2\pi)$ with $\int_{\mathbb{R}_+}\int_0^{2\pi}(r^{\beta'} \wedge 1)|\widetilde{\nu}_t(r,\theta)|drd\theta$ being locally bounded for some $0 \le \beta' < \beta$. Finally, $\widetilde{\mu}(dt, d\mathbf{x}) = \mu(dt, d\mathbf{x}) - dt\nu_t(d\mathbf{x})$ is the compensated jump measure. Additional (more technical) assumptions for various processes in (1.1) and (1.2) are stated in the next section.

Around the origin, the first component of $\nu_t(r,\theta)$ on the right hand side of (1.2) dominates the second one. Hence, this piece of $\nu_t(r,\theta)$ determines the behavior of the "small" jumps of $\mathbf{X}$. In particular, β controls the activity of jumps (and it is equal to the so-called Blumenthal-Getoor index of the jumps, see e.g., Jacod and Protter (2012)) and the function g controls the dependence between the "small" jumps of the two components of the vector process $\mathbf{X}$. Our interest in this paper is the nonparametric estimation of g from discrete observations of $\mathbf{X}$ on a fixed time interval with asymptotically shrinking mesh of the observation grid.

The leading case of the above model is when $\mathbf{X}$ is a bivariate β-stable process, which corresponds to $\mathbf{a}_t = \mathbf{a}$, $L_t = L$ and $\widetilde{\nu}_t = 0$, see e.g., chapter 2 of Samorodnitsky and Taqqu (1994). In the stable case, g captures the dependence between the jumps of the two components of $\mathbf{X}$, regardless of their size, and is referred to as the spectral density. With a slight abuse of notation, henceforth we will continue to refer to g as the spectral density even in the general case when $\mathbf{X}$ is not necessarily bivariate stable but satisfies only (1.1)-(1.2).

More generally, the setup in (1.1)-(1.2) nests "locally" stable Lévy processes, i.e., Lévy processes whose measure around the origin behaves like that of a bivariate stable process. An example is the popular class of tempered stable Lévy processes (Rosiński (2007)). Further, (1.1)-(1.2) also holds for time-changed "locally" stable Lévy processes with absolutely continuous time-change, with L_t playing the role of the density of the time-change, see e.g., Carr et al. (2003) for their use in mathematical finance.

Models of pure-jump type have been used in various applications, see e.g., Cheng and Rachev (1995), Barndorff-Nielsen and Shephard (2001), Mikosch et al. (2002), Carr et al. (2002, 2003), Klüppelberg et al. (2004), Andrews et al. (2009), Klüppelberg et al. (2010) and Todorov and Tauchen (2011b). The small jumps replace the diffusion in capturing the small moves in $\mathbf{X}$. Unlike the case of diffusions, where the dependence between the elements of the vector process are captured by the (spot) covariance matrix (which is constant when the processes are Lévy), in the case of jumps we need a function (i.e., g above) to fully characterize the dependence between the small jumps. Most of the work to date on inference for pure-jump models has been either for the univariate case or in the special multivariate stable case. The goal of the current paper is to study dependence between small jumps in the general setting of (1.1)-(1.2) which allows for the Lévy measure to vary stochastically over time.

How can we estimate the spectral density function g from high-frequency observations of $\mathbf{X}$? To gain intuition, let us first consider the case when $\mathbf{X}$ is a symmetric bivariate stable process, i.e., a bivariate stable process with its spectral density satisfying $g(\theta) = g(\theta+\pi)$. That is, we consider the specification (1.1)-(1.2) with $\mathbf{a}_t = 0$, $L_t = 1$, $\widetilde{\nu}_t(r,\theta) = 0$, for $t \in [0,1]$, and $g(\theta) = g(\theta+\pi)$. In this case, the log-characteristic function $\log(\mathbb{E}(e^{i\langle \mathbf{u},\mathbf{X}_1\rangle}))$, for $\mathbf{u} = (u_1, u_2) \in \mathbb{R}^2$, is given by (Theorem 2.4.3 of Samorodnitsky and Taqqu (1994))

$$\Phi(\mathbf{u}) = -C_\beta \int_0^\pi |u_1 \cos\theta + u_2 \sin\theta|^\beta (g(\theta) + g(\theta+\pi))d\theta, \tag{1.3}$$

where C_β is some known function of β. If we restrict attention to $\mathbf{u} \in \mathbb{S}^1$, where $\mathbb{S}^1 = \left\{\mathbf{x} \in \mathbb{R}^2 : ||\mathbf{x}|| = 1\right\}$ is the unit circle, then the right-hand side of (1.3) can be expressed as a convolution of g with another function depending only on β, see e.g., Pivato and Seco (2003). Therefore, if β is known and $\Phi(\mathbf{u})$ can be estimated from the data, then we can solve the deconvolution problem in Fourier transform and estimate g nonparametrically, see e.g., Fan (1991) and the references therein.

In the general case of (1.1)-(1.2), $\Phi(\mathbf{u})$ is not the log-characteristic function of $\mathbf{X}_1$ but only a local approximation of it. Nevertheless, it turns out that we can still recover (up to a constant) $\Phi(\mathbf{u})$ from the high-frequency increments of $\mathbf{X}$. The recovery of $\Phi(\mathbf{u})$ is challenging as the increments of $\mathbf{X}$ do not have exact stable distribution and are heteroskedastic. Moreover, the law of the increments changes with the sampling frequency even in the stable case. To overcome these difficulties, we adopt a self-normalization approach, similar to the one proposed in Todorov (2015) for the estimation of the Blumenthal-Getoor index (i.e., β in (1.2)). First, we difference the high-frequency increments to mitigate the asymptotic effect of the drift term in $\mathbf{X}$ ($\mathbf{a}_t$ in (1.1)) on our statistics. Then, we normalize the differenced increments by local power variations formed from a local window of increments preceding the ones that are scaled. The law of the high-frequency increments of $\mathbf{X}$ is dominated by the first part of $\nu_t(r,\theta)$ on the right-hand side of (1.2) and with L_t kept fixed at its value at the beginning of the interval. Therefore, by the self-similarity property of the stable process, the differenced high-frequency increment of $\mathbf{X}$ behaves (conditionally) like a scaled bivariate symmetric stable random vector with a scale that depends on the sampling frequency and varies over time. The normalization of the increments purges this unknown random scale and thus the normalized increments behave approximately like uncorrelated and identically distributed stable random vectors. Therefore, our estimate of (a scaled version of) $\Phi(\mathbf{u})$ is simply the empirical characteristic function of the normalized differenced high-frequency increments.

Given the estimate of $\Phi(\mathbf{u})$ from the high-frequency data and an estimate of the jump activity index β as in Todorov (2015), the recovery of g (its symmetrized version $g(\theta) + g(\theta + \pi)$ to be precise) proceeds using the Fourier-based deconvolution technique discussed earlier. We derive the asymptotic order of the associated mean integrated squared error. As standard for deconvolution problems, the asymptotic order of the estimation error depends on the smoothness of the function g to be estimated as well as on the smoothness of the function it is convoluted with in the expression for $\Phi(\mathbf{u})$ in (1.3). The latter is determined by the value of β, with higher values of β corresponding to a slower rate of convergence of the estimator for a given level of smoothness of g. Since the error in recovering β is of the same order of magnitude as the one due to the estimation of $\Phi(\mathbf{u})$, the fact that the smoothness of the function that is convoluted with the object of interest is unknown does not impact the rate of convergence of the nonparametric estimator of g.

The results in the current paper relate to several strands of literature. First, Press (1972), Cheng and Rachev (1995), McCulloch et al. (2001) and Pivato and Seco (2003) consider estimation of the spectral density of a multivariate stable distribution from i.i.d. observations. By contrast, our asymptotic setup is the high-frequency one, i.e., we observe discretely a stochastic process on a fixed interval with asymptotically shrinking mesh. This is a non-trivial difference as in our setting the law of the increments depends on the sampling frequency. In addition, unlike the above cited papers, our setup is more general and covers processes which are only "locally" stable and can have time-varying jump intensity. Second, the current paper is also related to existing work on the recovery of the Blumenthal-Getoor index from high-frequency data, see e.g., Ait-Sahalia and Jacod (2009), Bull (2016), Jing et al. (2011), Jing et al. (2012), Kong et al. (2015), Todorov (2015), Todorov and Tauchen (2011a), Woerner (2003, 2007). Similar to that work, our interest here is also the behavior of the small jumps. However, unlike that work, the focus here is on the recovery of the spectral density function which governs the dependence between the small jumps of the observed vector. Third, the current paper is connected with existing studies on nonparametric deconvolution problems. These studies are primarily done in the i.i.d. random variables setup which as explained earlier is very different from our high-frequency setup. In addition, in the standard deconvolution problems the smoothness of the function that the object of interest is convoluted with is either known, see e.g., Fan (1991) and the many references therein, or is inferred from an independent source of information in the form of an additional i.i.d. data set, see e.g., Johannes (2009) and the many references therein. By contrast, in our case the smoothness of the function we are decoupling our object of interest (i.e., the spectral density) is determined by the unknown value of the jump activity index which we estimate from the same data set used for inferring $\Phi(\mathbf{u})$.

The rest of the paper is organized as follows. In Section 2 we introduce the setting and state our assumptions. In Section 3 we build our estimator of the spectral density from the high-frequency data and in Section 4 we characterize its asymptotic behavior. Section 5 presents numerical experiments on simulated data. Section 6 contains the proofs.

2 Setting and Assumptions

We start with the assumptions that we need for our results. The process $\mathbf{X}$ in (1.1) is defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. Throughout, we will make use of a reference stable process $\mathbf{S}$, which is defined exactly as $\mathbf{J}$ in (1.1)-(1.2) but with $L_t = 1$ and $\widetilde{\nu}_t(r, \theta) = 0$ for every $t \geq 0$. We will refer to the distributional properties of $\mathbf{S}$ only and hence its relation to all other quantities defined on the probability space is irrelevant. The real part of the log-characteristic function of $\mathbf{S}_1$, $\Re(\log(\mathbb{E}(e^{i\langle \mathbf{u}, \mathbf{S}_1\rangle})))$, is given by $\Phi(\mathbf{u})$ defined in (1.3). In what follows we will design a way to infer $\Phi(\mathbf{u})$ from a discrete record of $\mathbf{X}$.

We now state the assumptions. The first one is for the smoothness of g.

Assumption A. *For the bounded and nonnegative-valued function g, we have $|\int_0^{2\pi} g(x)e^{imx}dx| \leq C(1+m)^{-\alpha}$, for positive constants C and $\alpha > 1$ and $\forall m \in \mathbb{Z}$.*

As well known, the decay rate of the Fourier transform of g is connected with its smoothness properties, with higher values of α corresponding to smoother g. The constant α will play a central role in our analysis.

Our next assumption is for the drift, stochastic intensity and the "residual" jump compensator, i.e., $\mathbf{a}$, L and $\widetilde{\nu}_t(r, \theta)$. As in the introduction, we will denote with κ a truncation function.

Assumption B. *(a) The processes* $\mathbf{a}$ *and* L *have càdlàg paths and further there exists a sequence of stopping times* T_k *increasing to infinity and a sequence of positive numbers* Γ_k *such that* $\mathbb{E}||\mathbf{a}_{t\wedge T_k} - \mathbf{a}_{s\wedge T_k}||^2 \leq \Gamma_k|t-s|$ *and* $\mathbb{E}|L_{t\wedge T_k} - L_{s\wedge T_k}|^2 \leq \Gamma_k|t-s|$. *In addition,* L_t *and* L_{t-} *are strictly positive for every* $t \geq 0$.
(b) There exists a sequence of stopping times T_k *increasing to infinity and a sequence of positive numbers* Γ_k *such that for* $t < T_k$ *we have* $\int_{\mathbb{R}_+}(|r|^{\beta'} \wedge 1)|\widetilde{\nu}_t(r,\theta)|dr \leq \Gamma_k$, *for* $\theta \in [0, 2\pi)$ *and some nonnegative* $\beta' < \beta$.
(c) When $\beta < 1$, *we denote* $\widetilde{\mathbf{a}}_t = \mathbf{a}_t - \int_{\mathbb{R}^2}\kappa(\mathbf{x})\nu_t(d\mathbf{x})$ *and assume that* $\widetilde{\mathbf{a}}_t$ *is of the form* $\widetilde{\mathbf{a}}_t = \widetilde{\mathbf{a}}_0 + \int_0^t\int_{\mathbb{R}^2}\delta(s,\mathbf{x})\eta(ds,d\mathbf{x})$, *where* $\delta : \mathbb{R}_+ \times \mathbb{R}^2 \to \mathbb{R}^2$ *is some predictable function and* η *is a Poisson random measure on* $\mathbb{R}_+ \times \mathbb{R}^2$ *with compensator* $dt \times d\mathbf{x}$. *There is a localizing sequence* T_k *of stopping times and, for each* k, *a deterministic nonnegative function* $\Gamma_k(\mathbf{x})$ *on* $\mathbb{R}^2$ *satisfying* $\int_{\mathbb{R}^2}(\Gamma_k(\mathbf{x}) \wedge 1)^{\beta+\iota}d\mathbf{x} < \infty$, *for all* $\iota > 0$, *and such that* $||\delta(t,\mathbf{x})|| \leq \Gamma_k(\mathbf{x})$ *for all* $t < T_k$.

Part (a) of Assumption A is a weak "smoothness in squared expectation" condition for the processes $\mathbf{a}$ and L which is satisfied, for example, when they are Itô semimartingales. Part (b) of the assumption restricts the residual part $\widetilde{\nu}_t(r,\theta)$ of the jump compensator to be dominated near the origin by the stable part of $\nu_t(r,\theta)$. Importantly, we note that $\widetilde{\nu}_t$ is a signed function. Therefore, $\widetilde{\nu}_t$ can completely annihilate the stable part of $\nu_t(r,\theta)$ for the "big" jumps. Thus the above assumption restricts only the behavior of the "small" jumps of $\mathbf{X}$. Finally, in part (c), we impose additional assumptions for the "effective" drift term (i.e., the drift after controlling for the compensation of the small jumps) in the case when $\beta < 1$. This assumption is significantly more restrictive than the one imposed in the general case in part (a).

3 Estimation of the Spectral Density

Henceforth we assume that we observe $\mathbf{X}$ on a finite interval, which without loss of generality we set to $[0, 1]$. The observation times are given by the equidistant grid $0, \frac{1}{n}, ..., 1$, with $n \to \infty$. The distance between consecutive observation times is denoted with $\Delta_n = \frac{1}{n}$. Our estimation strategy is based on first recovering $\Phi(\mathbf{u})$ in (1.3) (up to a constant) and the jump activity index β from the high-frequency observations of $\mathbf{X}$, and then using deconvolution and Fourier inversion techniques to infer the spectral density from these estimates.

3.1 Estimation of $\Phi(\mathbf{u})$ and β

The estimation of $\Phi(\mathbf{u})$ and β is based on the fact that at high frequencies the "residual" component of the jump compensator, $\widetilde{\nu}_t(r,\theta)$, plays a negligible role and the same holds true for the drift term $\mathbf{a}_t$ in (1.1) (provided we difference the increments). Further, to account for the time-varying intensity process L, we will estimate it locally, and then we will use the estimate to standardize the high-frequency increments with it. The estimation of L will be achieved via local power variations constructed from blocks of k_n high-frequency increments preceding the increments to be standardized, where for the block size we have $k_n \to \infty$ and $k_n/n \to 0$, i.e., the block size increases but its time-span shrinks asymptotically. For a generic univariate process Y, we denote its p-th local power variation as

$$V(Y,p)_i^n = \frac{1}{k_n}\sum_{j\in I_n^i} |\Delta_j^n Y - \Delta_{j-1}^n Y|^p, \quad \Delta_j^n Y = Y_{\frac{j}{n}} - Y_{\frac{j-1}{n}}, \tag{3.4}$$

where $I_n^i = \{i-2k_n, i-2k_n+2, ..., i-2\}$. The differenced and standardized increments are then defined as:

$$\mathbf{X}^{(j)}(p)_i^n = \frac{\Delta_i^n \mathbf{X}^{(j)} - \Delta_{i-1}^n \mathbf{X}^{(j)}}{(V(\mathbf{X}^{(1)},p)_i^n + V(\mathbf{X}^{(2)},p)_i^n)^{1/p}}, \quad j = 1,2. \tag{3.5}$$

Our statistics will be based on $\mathbf{X}^{(j)}(p)_i^n$. The asymptotic results that will follow can be intuitively explained as follows. At high frequencies, we have $\Delta_i^n\mathbf{X} - \Delta_{i-1}^n\mathbf{X} \approx L_{(i-2)\Delta_n}^{1/\beta}(\Delta_i^n\mathbf{S} - \Delta_{i-1}^n\mathbf{S})$, where $\mathbf{S}$ is defined at the beginning of Section 2. Therefore, $V(\mathbf{X}^{(1)},p)_i^n + V(\mathbf{X}^{(2)},p)_i^n \approx C_{p,\beta} \times \Delta_n^{\frac{p}{\beta}} L_{(i-2)\Delta_n}^{p/\beta}$, for $0 < p < \beta$ and where $C_{p,\beta}$ is some constant that depends on p and β but importantly not on the intensity process L. Then, using the self-similarity property of the stable process, we have that $\mathbf{X}^{(j)}(p)_i^n$ is approximately the j-th component of a symmetric bivariate stable process with log-characteristic function which up to a constant is given by $\Phi(\mathbf{u})$. Moreover, for high frequencies, $\mathbf{X}^{(j)}(p)_k^n$ and $\mathbf{X}^{(j)}(p)_l^n$ become approximately independent whenever $|k-l| > 1$.

We note that the differencing of the increments in $\mathbf{X}^{(j)}(p)_i^n$ is essential for what follows and is done to minimize the impact of the drift term in $\mathbf{X}$ on our statistics. This is easiest to see when the drift process $\mathbf{a}$ in (1.1) is constant. In this case the drift component cancels in the first difference of the high-frequency increments. Otherwise, $\mathbf{a}$ determines the limit of the local power variation for $\beta < 1$ (i.e., the drift will "dominate" the finite variation jumps in the power variation) and it slows down its rate of convergence when $\beta \geq 1$. This carries over to the statistics we introduce below. Differencing of increments has been used in other related contexts, see e.g., Todorov (2013, 2015) and Kong et al. (2015). One consequence of the differencing of the increments is that we loose the information in the data about the asymmetry of the spectral density. That is, we can only recover the sum $g(\theta)+g(\theta+\pi)$. We stress, however, that the analysis that follows does not require symmetry, i.e., we do not assume $g(\theta)=g(\theta+\pi)$.

Our estimates of $\Phi(\mathbf{u})$ and β are based on the empirical characteristic function of the differenced and standardized increments which is defined as

$$\widehat{\mathcal{L}}_p^n(u_1,u_2)=\frac{1}{n-2k_n-1}\sum_{i=2k_n+2}^{n}\cos\left(u_1\mathbf{X}^{(1)}(p)_i^n+u_2\mathbf{X}^{(2)}(p)_i^n\right), \tag{3.6}$$

where $u_1,u_2\in\mathbb{R}$ and we further set

$$\overline{\mathcal{L}}_p^n(u_1,u_2)=\widehat{\mathcal{L}}_p^n(u_1,u_2)\bigvee\frac{1}{k_n}. \tag{3.7}$$

Under our assumptions and given the above discussion, we have

$$\overline{\mathcal{L}}_p^n(u_1,u_2)\ \stackrel{\mathbb{P}}{\longrightarrow}\ \mathcal{L}_p(u_1,u_2),\quad u_1,u_2\in\mathbb{R}, \tag{3.8}$$

where

$$\mathcal{L}_p(u_1,u_2)=\exp\left(-\int_0^\pi|u_1\cos(\theta)+u_2\sin(\theta)|^\beta\widetilde{g}(\theta)d\theta\right), \tag{3.9}$$

with

$$\widetilde{g}(\theta)=\frac{g(\theta)+g(\theta+\pi)}{K_{g,p}^{\beta/p}\mu_p^{\beta/p}},\quad \mu_p=\frac{\Gamma(\frac{1+p}{2})\Gamma(1-\frac{p}{\beta})}{\sqrt{\pi}\Gamma(1-\frac{p}{2})},$$

and

$$K_{g,p}=\left(\int_0^\pi|\cos(\theta)|^\beta(g(\theta)+g(\theta+\pi))d\theta\right)^{\frac{p}{\beta}}+\left(\int_0^\pi|\sin(\theta)|^\beta(g(\theta)+g(\theta+\pi))d\theta\right)^{\frac{p}{\beta}}.$$

The function $\widetilde{g}(\theta)$ is a scaled version of $g(\theta)+g(\theta+\pi)$. The scale factor $K_{g,p}^{\beta/p}\mu_p^{\beta/p}$ is due to the standardization of the increments of $\mathbf{X}$ by the local power variations. We note that the limit result in (3.8) is not sufficient for what follows as we will need the convergence of $\overline{\mathcal{L}}_p^n(u_1,u_2)$ as a function.

Next, we can construct easily an estimator of β from $\overline{\mathcal{L}}_p^n(u_1,u_2)$. Following Todorov (2015), our estimator is given by:

$$\overline{\beta}(p)=\frac{\log(-\log(\overline{\mathcal{L}}_p^n(u_1,u_1)))-\log(-\log(\overline{\mathcal{L}}_p^n(u_2,u_2)))}{\log(u_1/u_2)}, \tag{3.10}$$

for some fixed $u_1\neq u_2$ with $u_1,\,u_2\in\mathbb{R}_+$, and for simplicity we suppress the dependence on u_1 of u_2 in the notation of the above estimator of β. We further correct the above estimator by the following bias term that reflects the impact on the estimation of β which is due to the variation of the local power variation:

$$\widehat{\mathcal{B}}_{p,\beta}^n=\frac{1}{2k_n}\left(\frac{\overline{\beta}(p)}{p}\right)^2\frac{\log(\overline{\mathcal{L}}_p^n(u_2,u_2)/\overline{\mathcal{L}}_p^n(u_1,u_1))}{\log(u_1/u_2)}\left(\frac{\widehat{\Sigma}_p^n}{\widehat{\mathcal{K}}_p^n}-1\right). \tag{3.11}$$

Here $\widehat{\Sigma}_p^n$ and $\widehat{\mathcal{K}}_p^n$ are estimators of the asymptotic variance of the local power variation which are defined from

$$\xi_i^n = |\Delta_i^n \mathbf{X}^{(1)} - \Delta_{i-1}^n \mathbf{X}^{(1)}|^p + |\Delta_i^n \mathbf{X}^{(2)} - \Delta_{i-1}^n \mathbf{X}^{(2)}|^p, \tag{3.12}$$

as follows:

$$\widehat{\Sigma}_p^n = \sum_{i=1}^n (\xi_i^n)^2 \quad \text{and} \quad \widehat{\mathcal{K}}_p^n = \sum_{i=4}^n \xi_i^n \xi_{i-2}^n. \tag{3.13}$$

With this notation, the debiased jump activity estimator is given by

$$\widehat{\beta}(p) = \left[\left(\overline{\beta}(p) - \widehat{\mathcal{B}}_{p,\beta}^n\right) \bigvee \frac{1}{k_n}\right] \bigwedge \left(2 - \frac{1}{k_n}\right). \tag{3.14}$$

We note that the above jump activity estimator makes use of $\widehat{\mathcal{L}}_p^n(u,u)$ for two values of u only. Additional efficiency gains in the recovery of β can be achieved by incorporating information across a range of u-s, see e.g., Theorem 3 in Todorov (2017). To keep the analysis simple, we do not do this here.

Exactly as for the estimation of the jump activity, while $\log(\overline{\mathcal{L}}_p^n(u_1,u_2))$ is a valid estimator of $\log(\mathcal{L}_p(u_1,u_2))$, it contains an asymptotic bias due to the scaling by the local power variations. This bias can be removed in a feasible way, however, and this improves the rate of convergence of the debiased estimator. The bias correction term is given by

$$\begin{aligned}\widehat{\mathcal{B}}_p^n(u_1,u_2) = &\frac{1}{2k_n}\left(\frac{\widehat{\Sigma}_p^n}{\widehat{\mathcal{K}}_p^n} - 1\right) \\ &\times \left[\log(\overline{\mathcal{L}}_p^n(u_1,u_2))\frac{\widehat{\beta}(p)}{p}\left(\frac{\widehat{\beta}(p)}{p} + 1\right) + \log^2(\overline{\mathcal{L}}_p^n(u_1,u_2))\left(\frac{\widehat{\beta}(p)}{p}\right)^2\right],\end{aligned} \tag{3.15}$$

and our bias-corrected estimate of $\log(\mathcal{L}_p(u_1,u_2))$ is $\log(\overline{\mathcal{L}}_p^n(u_1,u_2)) - \widehat{\mathcal{B}}_p^n(u_1,u_2)$.

3.2 Deconvolution of the Spectral Density

We proceed with showing how to recover $\widetilde{g}(\theta)$ from the asymptotic limit of $\overline{\mathcal{L}}_p^n(u_1,u_2)$, i.e., $\mathcal{L}_p(u_1,u_2)$, and we will then develop the feasible counterpart of this deconvolution procedure. For the recovery of $\widetilde{g}(\theta)$, we will need only the log-characteristic function evaluated on points in $\mathbb{R}^2$ lying on the unit circle which we denote as

$$\gamma_p(y) = -\log(\mathcal{L}_p(\cos(y),\sin(y))) = \int_0^\pi |\cos(y-\theta)|^\beta \widetilde{g}(\theta)d\theta. \tag{3.16}$$

If we further denote the function $h(x) = |\cos x|^\beta$, then we can write

$$\gamma_p(y) = \int_0^\pi h(y-\theta)\widetilde{g}(\theta)d\theta, \quad y \in [0,\pi).$$

That is, the function γ is a convolution of the functions h and $\widetilde{g}$. Hence

$$\gamma_p^*(m) = \pi h^*(m)\widetilde{g}^*(m), \quad m \in \mathbb{Z}, \tag{3.17}$$

where for a generic function f on the interval $[0, \pi)$ belonging to $\mathbb{L}^1([0, \pi])$, we denote its Fourier transform via

$$f^*(m) = \frac{1}{\pi}\int_0^\pi f(x)e^{-2imx}dx, \quad m \in \mathbb{Z}.$$

Then, since under Assumption A the sequence $\{\widetilde{g}^*(m)\}_{m\in\mathbb{Z}}$ of Fourier coefficients is absolutely summable and as shown in the proofs $|h^*(m)| > 0$ for $m \in \mathbb{Z}$, we have by Fourier inversion

$$\widetilde{g}(\theta) = \frac{1}{\pi}\sum_{m=-\infty}^{\infty}\frac{\gamma_p^*(m)}{h^*(m)}e^{2im\theta}, \quad \theta \in [0, \pi). \tag{3.18}$$

Now an estimator of $\widetilde{g}(\theta)$ is easy to construct. We will replace in the above sum $\gamma_p^*(m)$ with an estimator based on $\widehat{\mathcal{L}}_p^n(u_1, u_2)$, $h^*(m)$ with an estimator based on $\widehat{\beta}(p)$, and we will further truncate the higher frequencies. The spectral density estimator is thus given by

$$\widehat{g}_p^n(\theta) = \frac{1}{\pi}\sum_{m=-m_n}^{m_n}\left(1 - \frac{|m|}{m_n}\right)\frac{\widehat{\gamma}_p^*(m)}{\widehat{h}_p^*(m)}e^{2im\theta}, \quad \theta \in [0, \pi), \tag{3.19}$$

where m_n is a deterministic sequence with $m_n \to \infty$ as $n \to \infty$, $\widehat{h}_p^*(m)$ is the counterpart of $h_p^*(m)$ in which β is replaced with $\widehat{\beta}(p)$, and further

$$\widehat{\gamma}_p^*(m) = \frac{1}{\pi}\int_0^\pi \frac{\log(\overline{\mathcal{L}}_p^n(\cos x, \sin x)) - \widehat{\mathcal{B}}_p^n(\cos x, \sin x)}{\log(\overline{\mathcal{L}}_p^n(1, 1)) - \widehat{\mathcal{B}}_p^n(1, 1)}e^{-2imx}dx. \tag{3.20}$$

We note that in $\widehat{\gamma}_p^*(m)$ we rescale the estimate of $\gamma_p^*(m)$ by $\log(\overline{\mathcal{L}}_p^n(1, 1)) - \widehat{\mathcal{B}}_p^n(1, 1)$. This implies that $\widehat{g}_p^n(\theta)$ will be an estimator for the following scaled version of $\widetilde{g}(\theta)$:

$$\begin{aligned} G(\theta) &= \frac{\widetilde{g}(\theta)}{\int_0^\pi |\cos(\theta) + \sin(\theta)|^\beta \widetilde{g}(\theta)d\theta} \\ &= \frac{g(\theta) + g(\theta + \pi)}{\int_0^\pi |\cos(\theta) + \sin(\theta)|^\beta (g(\theta) + g(\theta + \pi))d\theta}, \quad \theta \in [0, \pi), \end{aligned} \tag{3.21}$$

which does not depend on the power p used for the local power variation. Note that, as discussed already in Section 2, we can recover the spectral density up to a constant. The reason for the rescaling in (3.20) is to minimize the impact of the local power variation (which contains asymptotic biases) on our statistic. In $\widehat{\gamma}_p^*(m)$, due to the rescaling, the local power variation has no first-order asymptotic effect and this greatly improves the asymptotic behavior of the estimator.

3.3 Adapting to the Unknown β

The behavior of the power variation, used to scale the increments with, depends in a crucial way on the magnitude of p relative to the unknown β, and in particular we need $0 < p < \beta/2$. Hence, in order to improve the convergence rate of our estimator, we will use a preliminary estimator of β and we will set p as a function of it. In particular, we will use the first p_n increments, for some p_n with $p_n \to \infty$ and $p_n/n \to 0$, to estimate β as follows. We first set for some $q > 0$

$$\widetilde{V}_1^n(q) = \frac{1}{p_n-1}\sum_{j=1}^2\sum_{i=2}^{p_n}|\Delta_i^n\mathbf{X}^{(j)} - \Delta_{i-1}^n\mathbf{X}^{(j)}|^q,$$
$$\widetilde{V}_2^n(q) = \frac{1}{p_n-3}\sum_{j=1}^2\sum_{i=4}^{p_n}|\Delta_i^n\mathbf{X}^{(j)} - \Delta_{i-1}^n\mathbf{X}^{(j)} + \Delta_{i-2}^n\mathbf{X}^{(j)} - \Delta_{i-3}^n\mathbf{X}^{(j)}|^q.$$

With this notation, the preliminary estimate of β is given by

$$\widehat{\beta}_1 = \frac{q\log(2)}{\log\left(\frac{\widetilde{V}_2^n(q)}{\widetilde{V}_1^n(q)}\bigvee 2^{\frac{q}{2}}\frac{p_n}{p_n-1}\right)}. \tag{3.22}$$

We then set

$$\widehat{p} = \frac{\widehat{\beta}_1}{4}, \tag{3.23}$$

and use $\widehat{p}$ for the spectral density estimation, i.e., we work with $\widehat{g}^n_{\widehat{p}}(\theta)$. A more conservative choice for $\widehat{p}$ (which is asymptotically equivalent to the one above) is the following. On the basis of $\widehat{\beta}_1$ we can construct a confidence interval for β and then set $\widehat{p}$ to be 1/4 of the lower end of this interval.

We also note that we can use the preliminary activity estimator $\widehat{\beta}_1$ to optimally choose u_1 and u_2 for the construction of $\overline{\beta}(\widehat{p})$. Since this does not improve the rate of the convergence of the latter, we do not do this here.

Another extension of the above analysis is to consider the case $p_n = n$, i.e., to use all the data in the calculation of the preliminary estimator $\widehat{\beta}_1$. For such an extension, however, one would need uniform in p convergence results for $\widehat{\mathcal{L}}^n_p(u_1, u_2)$ and $\widehat{\beta}(p)$, similar to the analysis in Todorov and Tauchen (2011a) for the jump activity estimation. This is a nontrivial extension which does not provide asymptotic efficiency improvements, and for this reason we do not consider it here.

4 Asymptotic Behavior of the Spectral Density Estimator

We next state the asymptotic properties of $\widehat{g}^n_{\widehat{p}}(\theta)$ in the following theorem.

Theorem 1 *Let Assumptions A and B hold for the process* $\mathbf{X}$ *with* $\beta' < \beta/2$ *and* $\beta > q$ *for some* $q > 0$. *Set* $\widehat{p}$ *as in (3.22)-(3.23) with* $p_n \to \infty$ *and* $p_n \leq 2k_n$. *Let* $k_n \asymp n^{\varpi}$ *and* $m_n \asymp n^{\varrho}$, *where*

$\frac{3}{8} \leq \varpi < \frac{1}{2}$ *and* $\varrho > 0$. *Then, for* $\forall \iota > 0$, *we have:*

$$\int_0^\pi \left(\widehat{g}_{\widehat{p}}^n(\theta) - G(\theta)\right)^2 d\theta$$
$$= O_p\left(m_n^{-2\alpha+1} \bigvee m_n^{3+2\beta}\left(\Delta_n^{\beta\wedge 1-\iota} \vee \frac{\Delta_n^{\frac{1}{2}\left(\frac{\beta}{\beta'}-1\right)-\iota}}{k_n}\right)\right). \tag{4.24}$$

We note that in the above theorem we derive the order of magnitude in probability of the integrated squared error in recovering the spectral density rather than providing a bound for its expectation. This is because some of the error terms in estimating the spectral density can be bounded only in probability and not in expectation. The first term in the bound in (4.24) depends only on α and is due to the bias in the estimation coming from the use of a finite number of frequencies in the Fourier inversion. The smoother the spectral density G, the smaller the bias. The second term in the bound in (4.24) is due to the error in recovering the log-characteristic function Φ and the jump activity index β. It consists both of bias type terms and terms that are centered at zero and drive the limit distribution in estimating Φ and β. In particular, the term in (4.24) involving the block size k_n reflects a bias due to the scaling with local power variations. We note in this regard that we have significant flexibility in choosing k_n, i.e., ϖ can take values in the range $\left[\frac{3}{8}, \frac{1}{2}\right)$ (we can allow for even lower values of ϖ at the cost of more bias type terms in (4.24)). Of course, it is optimal to set ϖ as close as possible to 1/2, in which case the term involving k_n in (4.24) becomes of higher order.

The effect of the jump activity on the spectral density is twofold. On one hand, a higher β makes estimation more difficult as $h^*(m)$ becomes smaller and hence deconvolution is more difficult. On the other hand, when $\beta < 1$ the error in recovering the characteristic exponent Φ and the jump activity β from the high-frequency data becomes larger. In addition, since the order of magnitude of the errors in recovering Φ and β is the same, the fact that β has to be estimated from the data does not have an impact on the order of magnitude of the integrated squared error in recovering the spectral density.

We can alternatively write (4.24) as

$$\int_0^\pi \left(\widehat{g}_{\widehat{p}}^n(\theta) - G(\theta)\right)^2 d\theta = O_p(\Delta_n^{\Psi_n}), \tag{4.25}$$

where

$$\Psi_n = [(2\alpha-1)\varrho] \bigwedge [\beta \wedge 1 - \iota - (3+2\beta)\varrho] \bigwedge \left[\frac{1}{2}\left(\frac{\beta}{\beta'} - 1\right) - \iota + \varpi - (3+2\beta)\varrho\right].$$

When the residual term $\widetilde{\nu}_t(r,\theta)$ in (1.2) is absent, then the last term in Ψ_n disappears. In this case, the optimal choice of ϱ is

$$\varrho^* = \frac{\beta \wedge 1 - \iota}{2\alpha + 2\beta + 2},$$

for some $\iota > 0$ arbitrary small, and the corresponding Ψ_n is given by

$$\Psi_n^* = \frac{(2\alpha - 1)(\beta \wedge 1 - \iota)}{2\alpha + 2\beta + 2}.$$

Finally, we note that one important feature of the model (1.1)-(1.2) is that the dependence between the "small" jumps is time-invariant, i.e., the spectral density g does not depend on time. This mirrors our assumption for β, which is similarly assumed to be constant, consistent with the existing literature in univariate settings, see, e.g., Todorov (2017) and references therein. A generalization of the above setup will be to allow g to depend on time and be random. We conjecture that the same analysis we conducted here will go through for the estimation of g locally at a given point in time, i.e., from a block of increments with asymptotically shrinking time span, provided appropriate smoothness in expectation assumption is made for the time-variation in g. This, however, will be at the cost of much slower rates of convergence than the ones exhibited in Theorem 1.

Remark 1 *Our asymptotic setup is one of regular (equidistant) sampling. To extend the above analysis to the case of irregular sampling, we will need to appropriately rescale the increments when taking their difference in the construction of* $\widehat{\mathcal{L}}_p^n(u_1, u_2)$ *and* $\widehat{\beta}(p)$*, in order to account for the differences in the length of the intervals they are computed from. This will guarantee the cancelation of the drift when differencing in the case it is constant, and more generally it will reduce the impact of the drift on the statistic exactly as in the regular sampling case. Therefore, the result of Theorem 1 should be easy to extend to cover the situation when sampling times are irregular in the symmetric case (i.e., when* $g(\theta) = g(\theta + \pi)$*) at least when the sampling times are deterministic.*

5 Numerical Experiments

We now test the performance of our estimator on simulated data from a model for $\mathbf{X}$ in which the drift process $\mathbf{a}$ is some vector of constants (due to the differencing of the increments the values of these constants do not matter) and the Lévy density is given by

$$\nu_t(r,\theta) = A\frac{e^{-\lambda|r|}}{|r|^{\beta+1}} g(\theta)\sigma_t, \quad r \in \mathbb{R}_+, \quad \theta \in [0, 2\pi), \tag{5.26}$$

with the time-varying intensity process σ specified below. When σ_t is constant, then the Lévy density in (5.26) is that of a bivariate tempered stable process (Rosiński (2007)). The parameter β

coincides with the Blumenthal-Getoor index of the individual components of $\mathbf{X}$ and the parameter $\lambda > 0$ controls the tempering of the Lévy measure at infinity. The constant A is the following function of β

$$A = \frac{\beta(\beta-1)}{2\Gamma(2-\beta)|\cos(\beta\pi/2)|}, \quad \beta \in (1,2). \tag{5.27}$$

With this choice of A, the constant C_β in (1.3) equals 1/2. Finally, the stochastic intensity process σ is modeled as the square-root diffusion

$$d\sigma_t = 0.3 \times 252(252 - \sigma_t)dt + 0.1 \times 252\sqrt{\sigma_t}dW_t, \tag{5.28}$$

with W being a Brownian motion.

In all experiments we set the tempering parameter to $\lambda = 0.2$. We experiment with three values of the BG index: $\beta = 1.2$, 1.5 and 1.8. For the spectral density, $g(\theta)$, we consider the following model:

$$g(\theta) = \begin{cases} \frac{x}{2}, & \text{if } \theta \in (0, \frac{\pi}{4}], \\ \frac{\pi}{4} - \frac{x}{2}, & \text{if } \theta \in (\frac{\pi}{4}, \frac{\pi}{2}], \\ 0, & \text{if } \theta \in (\frac{\pi}{2}, \pi]. \end{cases} \tag{5.29}$$

The support of this function is $(0, \pi/2)$. $g(\theta)$ is continuous but its derivative $g'(\theta)$ has discontinuities at $\theta = 0$, $\theta = \frac{\pi}{4}$ and $\theta = \frac{\pi}{2}$. This parametric specification for the spectral density satisfies Assumption A with $\alpha = 2$.

The simulation from the bivariate tempered stable process is done via discretization of the dependence function $g(\theta)$ and simulation of univariate tempered stable processes. We simulate $\mathbf{X}$ on unit interval which corresponds to 1 year of 252 trading days. On each trading day we sample X at 100 equidistant points. This corresponds approximately to 5-minute sampling frequency for a typical financial application and results in $n = 25200$. [Alternatively, we can think of the unit interval representing approximately 50 trading days of 1 minute records of $\mathbf{X}$, etc.] We set $k_n = 100$ for the local power variation and $p_n = 2k_n$ for the initial jump activity estimator. We use $q = 0.25$ for $\widehat{\beta}_1$ and for the final jump activity estimator $\widehat{\beta}(\widehat{p})$ we use $u_1 = 1.5$ and $u_2 = 3$.

The results from 100 Monte Carlo replications are reported on Figure 1. As seen from the figure, we have a bias-variance tradeoff which is typical for nonparametric function estimation. Higher values of m_n lead to more precision but this is at the cost of noisier estimates of the spectral density function. We can also see from the figure the effect of the value of β on the precision of the estimation. For lower values of β, the accuracy of the estimation is higher and hence one can take higher values of m_n. For higher values of β, on the other hand, one has to be more conservative in the choice of m_n as the signal from the data becomes much weaker.

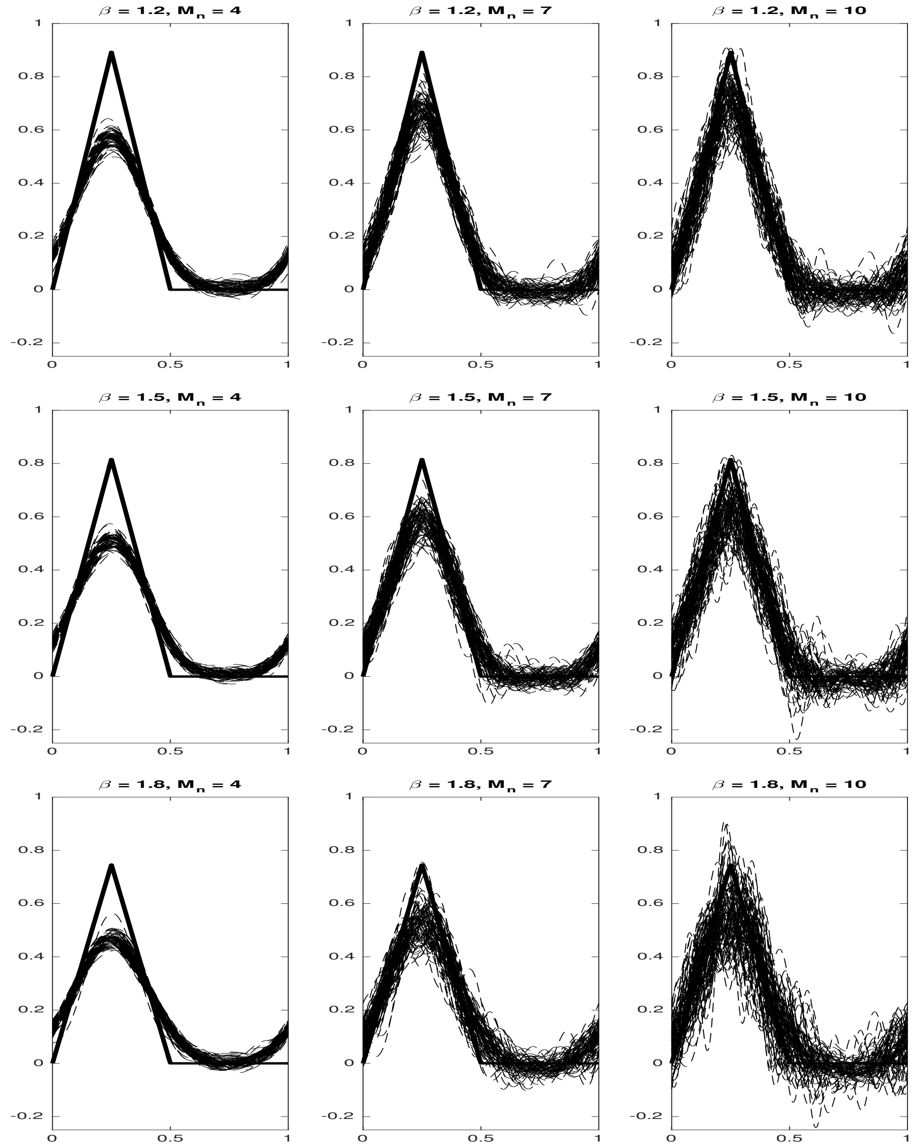

Figure 1: Estimated Spectral Densities. On each plot the solid thick line is the true spectral density and the thin dashed lines are estimates from 100 simulation draws with settings explained in the text. The x-axis is a multiple of π.

6 Proofs

Using Parseval's formula we have

$$\frac{1}{\pi}\int_0^{\pi}\left(\widehat{g}^n_{\widehat{p}}(\theta)-G(\theta)\right)^2 d\theta=\sum_{|m|\le m_n}(\widehat{g}^{n,*}_{\widehat{p}}(m)-G^*(m))^2+\sum_{|m|>m_n}|G^*(m)|^2, \tag{6.30}$$

where $\widehat{g}^{n,*}_{\widehat{p}}(m) = \frac{1}{\pi}\left(1-\frac{|m|}{m_n}\right)\frac{\widehat{\gamma}^*_{\widehat{p}}(m)}{\widehat{h}^*_{\widehat{p}}(m)}$. Therefore, we can bound the integrated squared error by a bias and a variance term:

$$\int_0^\pi \left(\widehat{g}^n_{\widehat{p}}(\theta) - G(\theta)\right)^2 d\theta \leq C(IB + IV), \tag{6.31}$$

for some positive constant C and where

$$IB = \frac{1}{m_n^2}\sum_{|m|\leq m_n}|mG^*(m)|^2 + \sum_{|m|>m_n}|G^*(m)|^2, \tag{6.32}$$

and $IV = \sum_{|m|\leq m_n}(A_1^n(m) + A_2^n(m) + A_3^n(m))$ with $A_1^n(m)$, $A_2^n(m)$ and $A_3^n(m)$ defined as:

$$A_1^n(m) = \frac{1}{h^*(m)^2}\left(\int_0^\pi |\widehat{a}^n(x) - a(x)|dx\right)^2,$$

$$A_2^n(m) = \left(\frac{1}{\widehat{h}^*_{\widehat{p}}(m)} - \frac{1}{h^*(m)}\right)^2\left(\int_0^\pi |\widehat{a}^n(x) - a(x)|dx\right)^2,$$

$$A_3^n(m) = \frac{G^*(m)^2}{\widehat{h}^*_{\widehat{p}}(m)^2}\left(\widehat{h}^*_{\widehat{p}}(m) - h^*(m)\right)^2,$$

with the additional notation

$$a(x) = \frac{\log(\mathcal{L}_{\widehat{p}}(\cos x, \sin x))}{\log(\mathcal{L}_{\widehat{p}}(1,1))}, \quad x \in [0,\pi],$$

$$\widehat{a}^n(x) = \frac{\log(\overline{\mathcal{L}}^n_{\widehat{p}}(\cos x, \sin x)) - \widehat{\mathcal{B}}^n_{\widehat{p}}(\cos x, \sin x)}{\log(\overline{\mathcal{L}}^n_{\widehat{p}}(1,1)) - \widehat{\mathcal{B}}^n_{\widehat{p}}(1,1)}, \quad x \in [0,\pi].$$

For the bias term, we have by Assumption A that IB $= O(m_n^{-2\alpha+1})$. For the variance term, we need to derive the order of magnitude of $h^*(m)$ (and related functionals) as $m \to \infty$ and analyze the asymptotic behavior of $\log(\overline{\mathcal{L}}^n_{\widehat{p}}(u_1,u_2)) - \widehat{\mathcal{B}}^n_{\widehat{p}}(u_1,u_2)$ and $\widehat{\beta}(p)$. We do this in the subsequent sections.

6.1 Alternative Representation of X and Localization

After appropriately extending the probability space and using Grigelionis decomposition (Theorem 2.1.2 in Jacod and Protter (2012)), we can represent $\mathbf{J}$ equivalently as

$$\mathbf{J}_t = \int_0^t \int_{\mathbb{R}^2} \kappa(\sigma_{s-}\mathbf{x})\widetilde{\mu}(ds, d\mathbf{x}) + \int_0^t \int_{\mathbb{R}^2} \kappa'(\sigma_{s-}\mathbf{x})\mu(ds, d\mathbf{x}) + Y_t, \tag{6.33}$$

where $\sigma_t = L_t^{1/\beta}$ and with slight abuse of notation, μ now is a Poisson measure with compensator $dt \otimes \nu(d\mathbf{x})$ with

$$\nu(d\mathbf{x}) = \frac{g(\theta)}{r^{1+\beta}}drd\theta, \quad \text{for } \mathbf{x} = (r\cos\theta, r\sin\theta), \tag{6.34}$$

and Y is a process of the form

$$
\begin{aligned}
Y_t = & \int_0^t \int_{\mathbb{R}^2} \kappa(\mathbf{x})\widetilde{\mu}_1(ds, d\mathbf{x}) + \int_0^t \int_{\mathbb{R}^2} \kappa'(\mathbf{x})\mu_1(ds, d\mathbf{x}) \\
& - \int_0^t \int_{\mathbb{R}^2} \kappa(\mathbf{x})\widetilde{\mu}_2(ds, d\mathbf{x}) - \int_0^t \int_{\mathbb{R}^2} \kappa'(\mathbf{x})\mu_2(ds, d\mathbf{x}),
\end{aligned}
\tag{6.35}
$$

where μ_1 and μ_2 are some integer-valued measures on $\mathbb{R}_+ \times \mathbb{R}^2$, having some dependence with μ, and whose compensators are of the form $dt \otimes \nu_{jt}(d\mathbf{x})$ with $\nu_{jt}(d\mathbf{x}) = \nu_{jt}(r, \theta) dr d\theta$, for $\mathbf{x} = (r\cos\theta, r\sin\theta)$ and $j = 1, 2$, and $\nu_{1t}(r, \theta) = |\widetilde{\nu}_t(r, \theta)|$ and $\nu_{2t}(r, \theta) = 2|\widetilde{\nu}_t(r, \theta)| 1_{\{\widetilde{\nu}_t(r,\theta) < 0\}}$.

We further denote (recall also Assumption B(c))

$$
\widetilde{\mathbf{a}}_t = \begin{cases} \mathbf{a}_t - \int_{\mathbb{R}^2} \kappa(\mathbf{x})\nu_t(d\mathbf{x}), & \text{if } \beta < 1, \\ \mathbf{a}_t - \int_{\mathbb{R}^2} \left(\kappa(\sigma_{t-}\mathbf{x}) - \sigma_{t-}\kappa(\mathbf{x})\right)\nu(d\mathbf{x}), & \text{if } \beta \geq 1, \end{cases}
\tag{6.36}
$$

$$
\widetilde{\mathbf{S}}_t = \begin{cases} \int_0^t \int_{\mathbb{R}^2} \mathbf{x}\mu(ds, d\mathbf{x}), & \text{if } \beta < 1, \\ \int_0^t \int_{\mathbb{R}^2} \kappa(\mathbf{x})\widetilde{\mu}(ds, d\mathbf{x}) + \int_0^t \int_{\mathbb{R}^2} \kappa'(\mathbf{x})\mu(ds, d\mathbf{x}), & \text{if } \beta \geq 1, \end{cases}
\tag{6.37}
$$

$$
\widetilde{Y}_t = \begin{cases} \int_0^t \int_{\mathbb{R}^2} \mathbf{x}\mu_1(ds, d\mathbf{x}) - \int_0^t \int_{\mathbb{R}^2} \mathbf{x}\mu_2(ds, d\mathbf{x}), & \text{if } \beta < 1, \\ Y_t, & \text{if } \beta \geq 1. \end{cases}
\tag{6.38}
$$

Note that $\widetilde{\mathbf{S}}_t$ has the same distribution as $\mathbf{S}_t$, where $\mathbf{S}_t$ is defined at the beginning of Section 2, and in particular, the real part of the log-characteristic function of $\widetilde{\mathbf{S}}_1$ is given by $\Phi(\mathbf{u})$ in (1.3). With this notation we can finally write:

$$
\mathbf{X}_t = \int_0^t \widetilde{\mathbf{a}}_u du + \int_0^t \sigma_{u-} d\widetilde{\mathbf{S}}_u + \widetilde{Y}_t.
\tag{6.39}
$$

We will prove the results under the following strengthened version of Assumption B:

Assumption SB. *Assume Assumption B holds and in addition:*

(a) The processes $\mathbf{a}$ *and* L *are bounded and further* $\mathbb{E}||\mathbf{a}_t - \mathbf{a}_s||^2 \leq \Gamma|t-s|$ *and* $\mathbb{E}|L_t - L_s|^2 \leq \Gamma|t-s|$ *for some finite constant* $\Gamma > 0$ *and every* $s, t \geq 0$*. In addition,* $|L_t|^{-1}$ *and* $|L_{t-}|^{-1}$ *are bounded by a finite positive constant.*

(b) The jumps of Y *are bounded and further* $\int_{\mathbb{R}_+} |r|^{\beta'} |\widetilde{\nu}_t(r, \theta)| dr \leq \Gamma$*, for* $\theta \in [0, 2\pi)$*, some nonnegative* $\beta' < \beta$ *and a finite constant* $\Gamma > 0$*.*

(c) When $\beta < 1$*, there is a positive constant* A *such that* $||\delta(t, \mathbf{x})|| \leq \Gamma(\mathbf{x})$*,* $\Gamma(\mathbf{x}) \leq A$ *and* $\int_{\mathbb{R}^2} (\Gamma(\mathbf{x}) \wedge 1)^{\beta+\iota} d\mathbf{x} < \infty$ *for all* $\iota > 0$*.*

Extending the proof to the weaker Assumption B follows by standard localization techniques, see e.g., Lemma 4.4.9 of Jacod and Protter (2012).

6.2 Auxiliary Results

Our first auxiliary result is a lower bound in absolute value and asymptotic size of the Fourier transform $h^*(m)$ as $m \to \infty$.

Lemma 1 *For the Fourier transform $h^*(m)$ of the function $h(x) = |\cos x|^\beta$ on the interval $[0, \pi]$ and for $\forall \epsilon > 0$, we have*

$$\inf_{\beta \in (\epsilon, 2-\epsilon)} |h^*(m)| > 0, \quad \text{for every } m \in \mathbb{Z}, \tag{6.40}$$

$$|h^*(m)| > C_\beta |m|^{-\beta-1}, \quad \text{for } |m| \text{ sufficiently large}, \tag{6.41}$$

where $C_\beta > 0$ is a constant that depends only on β and $\inf_{\beta \in \mathcal{B}} C_\beta > 0$ for any compact set $\mathcal{B}$ in $(0,1)$ or $[1,2)$. For all $\epsilon \in (0,1)$ and $\beta \in (\epsilon, 1]$, we have

$$|h^*(m)| > C_\epsilon |m|^{-2}, \quad \text{for } |m| \text{ sufficiently large}, \tag{6.42}$$

where $C_\epsilon > 0$ is a constant that depends only on ϵ. Further, we have

$$\left| \int_0^\pi h(x) \log^q |\cos x| e^{-2imx} dx \right| < C_\beta |m|^{-\beta-1} \log^q |m|, \ q \in \mathbb{N}, \tag{6.43}$$

for sufficiently large integer m and $C_\beta > 0$ is a constant that depends only on β and $\sup_{\beta \in \mathcal{B}} C_\beta < \infty$ for any compact set $\mathcal{B}$ in $(0,1]$ or $(1,2)$. For $\beta \in [1,2)$, we have

$$\left| \int_0^\pi h(x) \log^q |\cos x| e^{-2imx} dx \right| < C|m|^{-2} \log^q |m|, \ q \in \mathbb{N}, \tag{6.44}$$

where the constant C does not depend on β and m.

Proof of Lemma 1. We need to evaluate $\int_0^\pi \cos(mx)|\cos x|^\beta dx$ when m is an even integer and throughout the proof we will assume that this is the case. By splitting the region of integration and using standard trigonometric identities we have

$$h^*(m/2) = \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \cos(mx)|\cos(x)|^\beta dx, \tag{6.45}$$

and further

$$h^*(m/2) = \begin{cases} \frac{2}{\pi} \int_0^{\frac{\pi}{4}} \cos(mx)(|\cos x|^\beta + |\sin x|^\beta) dx, & \text{if } \frac{m}{2} \text{ is even}, \\ \frac{2}{\pi} \int_0^{\frac{\pi}{4}} \cos(mx)(|\cos x|^\beta - |\sin x|^\beta) dx, & \text{if } \frac{m}{2} \text{ is odd}. \end{cases} \tag{6.46}$$

The bound in (6.40) when $\beta \leq 1$. When $m/2$ is even, using (6.45), we have $h^*(m/2) = \frac{1}{\pi} \sum_{i=1}^{m/4} A_i$ where $A_i = \frac{2}{m} \int_{(i-1)2\pi}^{i2\pi} \cos x f\left(\frac{x}{m}\right) dx$ and $f(x) = |\cos x|^\beta$. By changing the variable of integration and using standard trigonometric identities we get

$$\begin{aligned} A_i = \frac{2}{m} \int_0^{\frac{\pi}{2}} \cos x \Bigg(f\left(\frac{x + (2i-2)\pi}{m}\right) - f\left(\frac{x + (2i-1)\pi}{m}\right) \\ - f\left(\frac{(2i-1)\pi - x}{m}\right) + f\left(\frac{2i\pi - x}{m}\right) \Bigg) dx. \end{aligned}$$

Since $f''(x) < 0$ for $x \in \left(0, \frac{\pi}{2}\right)$, we have $\sup_{\beta\in(\epsilon,1]} A_i < 0$ for every $0 < \epsilon < 1$ and every $i = 1, ..., \frac{m}{4}$.

When $m/2$ is odd we have $h^*(m/2) = \frac{2}{m\pi}\int_0^\pi \cos x f\left(\frac{x}{m}\right)dx + \frac{1}{\pi}\sum_{i=1}^{(m-2)/4} A_i$, where $A_i = \frac{2}{m}\int_{(i-1)2\pi+\pi}^{i2\pi+\pi} \cos x f\left(\frac{x}{m}\right)dx$ and $f(x) = |\cos x|^\beta$. By changing the variable of integration and using standard trigonometric identities, we get

$$A_i = -\frac{2}{m}\int_0^{\frac{\pi}{2}} \cos x\bigg(f\left(\frac{x+(2i-1)\pi}{m}\right) - f\left(\frac{x+i2\pi}{m}\right) - f\left(\frac{i2\pi - x}{m}\right) + f\left(\frac{(2i+1)\pi - x}{m}\right)\bigg)dx.$$

Since $f''(x) < 0$ for $x \in \left(0, \frac{\pi}{2}\right)$, we have $\inf_{\beta\in(\epsilon,1]} A_i > 0$ for every $0 < \epsilon < 1$ and every $i = 1, ..., \frac{m-2}{4}$. Further, since $f'(x) < 0$ for $x \in \left(0, \frac{\pi}{2}\right)$, we have

$$\frac{2}{m}\int_0^\pi \cos x f\left(\frac{x}{m}\right)dx = \frac{2}{m}\int_0^{\frac{\pi}{2}} \cos x\left(f\left(\frac{x}{m}\right) - f\left(\frac{\pi - x}{m}\right)\right)dx \geq 0.$$

The bound in (6.40) when $\beta > 1$. Using integration by parts twice, we can write

$$\int_0^{\frac{\pi}{4}} \cos(mx)|\cos x|^\beta dx = \frac{\beta}{m^2}\int_0^{\frac{\pi}{4}} \cos(mx)(|\cos x|^\beta - (\beta-1)|\cos x|^{\beta-2}\sin^2 x)dx + \frac{\sin(m\pi/4)}{m}|\cos(\pi/4)|^\beta.$$

Therefore

$$\left(1 - \frac{\beta}{m^2}\right)\int_0^{\frac{\pi}{4}} \cos(mx)|\cos x|^\beta dx = -\frac{\beta(\beta-1)}{m^2}\int_0^{\frac{\pi}{4}} \cos(mx)|\cos x|^{\beta-2}dx + \frac{\sin(m\pi/4)}{m}|\cos(\pi/4)|^\beta, \tag{6.47}$$

and similarly

$$\left(1 - \frac{\beta}{m^2}\right)\int_0^{\frac{\pi}{4}} \cos(mx)|\sin x|^\beta dx = -\frac{\beta(\beta-1)}{m^2}\int_0^{\frac{\pi}{4}} \cos(mx)|\sin x|^{\beta-2}dx + \frac{\sin(m\pi/4)}{m}|\sin(\pi/4)|^\beta. \tag{6.48}$$

Combining (6.46) with (6.47)-(6.48), we have

$$h^*(m/2) = \begin{cases} \frac{-2\beta(\beta-1)}{m^2-\beta}\frac{1}{\pi}\int_0^{\frac{\pi}{4}} \cos(mx)f(x)dx, & \text{if } \frac{m}{2} \text{ is even,} \\ \frac{-2\beta(\beta-1)}{m^2-\beta}\frac{1}{\pi}\int_0^{\frac{\pi}{4}} \cos(mx)g(x)dx, & \text{if } \frac{m}{2} \text{ is odd,} \end{cases} \tag{6.49}$$

where $f(x) = |\cos x|^{\beta-2} + |\sin x|^{\beta-2}$ and $g(x) = |\cos x|^{\beta-2} - |\sin x|^{\beta-2}$. Further, we can write $\int_0^{\frac{\pi}{4}} \cos(mx)f(x)dx = \frac{1}{m}\left(\sum_{i=1}^{\lfloor\frac{m}{8}\rfloor} A_{i,f} + R_f\right)$, where $A_{i,f} = \int_{(i-1)2\pi}^{i2\pi} \cos x f\left(\frac{x}{m}\right)dx$ and $R_f = \int_{\lfloor\frac{m}{8}\rfloor 2\pi}^{\frac{m\pi}{4}} \cos x f\left(\frac{x}{m}\right)dx$. Similarly, we have $\int_0^{\frac{\pi}{4}} \cos(mx)g(x)dx = \frac{1}{m}\left(\sum_{i=1}^{\lfloor\frac{m}{8}\rfloor} A_{i,g} + R_g\right)$, with $A_{i,g} = \int_{(i-1)2\pi}^{i2\pi} \cos x g\left(\frac{x}{m}\right)dx$ and $R_g = \int_{\lfloor\frac{m}{8}\rfloor 2\pi}^{\frac{m\pi}{4}} \cos x g\left(\frac{x}{m}\right)dx$. Therefore, it suffices to show that for $i = 2, ..., \lfloor\frac{m}{8}\rfloor$ and every $\epsilon > 0$, we have $\inf_{\beta\in(1,2-\epsilon)} A_{i,f} > 0$ and $\sup_{\beta\in(1,2-\epsilon)} A_{i,g} < 0$, $\inf_{\beta\in(1,2-\epsilon)} R_f \geq 0$ and also $\sup_{\beta\in(1,2-\epsilon)} R_g \leq 0$, as well as $\inf_{\beta\in(1,2-\epsilon)}(\beta-1)A_{1,f} > 0$ and $\sup_{\beta\in(1,2-\epsilon)}(\beta-1)A_{1,g} < 0$.

By splitting the integration over regions and using standard trigonometric identities, we have

$$A_{i,h} = \int_0^{\frac{\pi}{2}} \cos x \Bigg(h\left(\frac{x+(2i-2)\pi}{m}\right) - h\left(\frac{x+(2i-1)\pi}{m}\right) \\ - h\left(\frac{(2i-1)\pi - x}{m}\right) + h\left(\frac{2i\pi - x}{m}\right)\Bigg) dx, \quad h = f, g. \tag{6.50}$$

Therefore, since $f''(x) > 0$ and $g''(x) < 0$ for $x \in \left(0, \frac{\pi}{4}\right)$, we have $A_{i,f} > 0$ and $A_{i,g} < 0$ for every $\beta \in (1, 2)$. Note also that $\lim_{\beta\downarrow 1}(\beta - 1)A_{1,f} > 0$ and $\lim_{\beta\downarrow 1}(\beta - 1)A_{1,g} < 0$.

Next, we have

$$R_f = \int_0^{\frac{\pi}{2}} \cos x \left(f\left(\frac{x + \lfloor m/8 \rfloor 2\pi}{m}\right) - f\left(\frac{m\pi/4 - x}{m}\right)\right) dx.$$

From here, since $f'(x) < 0$ for $x \in \left(0, \frac{\pi}{4}\right)$, we have $R_f \geq 0$. We are left with R_g for which we have two cases depending on whether $(m-2)/4$ is even or odd. When $(m-2)/4$ is even, we have

$$R_g = \int_0^{\frac{\pi}{2}} \cos x g\left(\frac{x + \lfloor m/8 \rfloor 2\pi}{m}\right) dx,$$

and since $g(x) < 0$ for $x \in \left(0, \frac{\pi}{4}\right)$, we have $R_g \leq 0$ in this case. When $(m-2)/4$ is odd, we have

$$R_g = \int_0^{\frac{\pi}{2}} \cos x \left(g\left(\frac{x + \lfloor m/8 \rfloor 2\pi}{m}\right) - g\left(\frac{x + \lfloor m/8 \rfloor 2\pi + \pi}{m}\right) - g\left(\frac{\lfloor m/8 \rfloor 2\pi + \pi - x}{m}\right)\right) dx.$$

Since in this case $\lfloor m/8 \rfloor = \frac{m-6}{8}$, by applying integration by parts, we have

$$R_g = -\frac{1}{m}\int_0^{\frac{\pi}{2}} \sin x \Bigg(g'\left(\frac{x + \lfloor m/8 \rfloor 2\pi}{m}\right) - g'\left(\frac{x + \lfloor m/8 \rfloor 2\pi + \pi}{m}\right) \\ + g'\left(\frac{\lfloor m/8 \rfloor 2\pi + \pi - x}{m}\right)\Bigg) dx.$$

From here, making use of $g'(x) > 0$ and $g''(x) < 0$ for $x \in \left(0, \frac{\pi}{4}\right)$, we have $R_g \leq 0$ in the case when $(m-2)/4$ is odd.

We now turn to ((6.41)). First, the result can be shown by direct calculation for $\beta = 1$ and below we look separately at the cases $\beta > 1$ and $\beta < 1$.

The bound in ((6.41)) when $\beta > 1$. From (6.49) by applying integration by parts, we have

$$\left(1 - \frac{\beta}{m^2}\right) h^*(m/2) = -\frac{2\beta(\beta-1)}{\pi m^2} \int_0^{\lfloor \frac{m}{8} \rfloor \frac{2\pi}{m}} \cos(mx) |\sin x|^{\beta - 2} dx + \eta(\beta, m),$$

if $\frac{m}{2}$ is even and

$$\left(1 - \frac{\beta}{m^2}\right) h^*(m/2) = \frac{2\beta(\beta-1)}{\pi m^2} \int_0^{\lfloor \frac{m}{8} \rfloor \frac{2\pi}{m}} \cos(mx) |\sin x|^{\beta - 2} dx + \eta(\beta, m),$$

if $\frac{m}{2}$ is odd and where $\sup_{\beta\in[1,2]} |\eta(\beta,m)| = O\left(\frac{1}{m^3}\right)$. Next we analyze

$$\int_0^{\lfloor \frac{m}{8} \rfloor 2\pi} \cos x \left(\left| \sin\left(\frac{x}{m}\right) \right|^{\beta-2} - \left| \frac{x}{m} \right|^{\beta-2} \right) dx.$$

First, since $\beta - 2 \geq -1$ and using the inequality $|\sin x/x - 1| \leq \frac{1}{6}|x|^2$ for $x \neq 0$, we have for some positive constant C and m sufficiently high

$$\int_0^{2\pi} \left| \left| \sin\left(\frac{x}{m}\right) \right|^{\beta-2} - \left| \frac{x}{m} \right|^{\beta-2} \right| dx \leq C.$$

Then by integration by parts for any positive integer $k \leq \frac{m}{2}$ (note that $\sin(x/m) \geq 0$ for $x \in [0, 2k\pi]$), we have

$$\begin{aligned}
&\int_{2\pi}^{2k\pi} \cos x \left(\left| \sin\left(\frac{x}{m}\right) \right|^{\beta-2} - \left| \frac{x}{m} \right|^{\beta-2} \right) dx \\
&\quad = \frac{2-\beta}{m} \int_{2\pi}^{2k\pi} \sin x \left(\left| \sin\left(\frac{x}{m}\right) \right|^{\beta-3} \cos\left(\frac{x}{m}\right) - \left| \frac{x}{m} \right|^{\beta-3} \right) dx,
\end{aligned}$$

and using the trigonometric identity $1 - \cos x = 2\sin^2(x/2)$ and $\beta \geq 1$, we have that the last integral is equal to a term bounded in absolute value by a constant that does not depend on m and β and the following expression

$$\frac{2-\beta}{m} \int_{2\pi}^{2k\pi} |\sin x| \left(\left| \sin\left(\frac{x}{m}\right) \right|^{\beta-3} - \left| \frac{x}{m} \right|^{\beta-3} \right) dx,$$

which in turn is bounded in absolute value by

$$\frac{2-\beta}{m^{\beta-2}} \int_{2\pi}^{2k\pi} x^{\beta-3} \left| 1 - \left| \frac{\sin(x/m)}{x/m} \right|^{3-\beta} \right| \frac{1}{\left| \frac{\sin(x/m)}{x/m} \right|^{3-\beta}} dx,$$

and using the inequality $|\sin x/x - 1| \leq \frac{1}{6}|x|^2$ for $x \neq 0$, as well as the fact that $k \leq \lfloor \frac{m}{8} \rfloor$, we have that the last integral is bounded by a positive constant uniformly over $\beta \in [1,2]$ for m sufficiently high. Altogether:

$$h^*(m/2) = -\frac{2\beta(\beta-1)}{\pi m^2} \int_0^{\lfloor \frac{m}{8} \rfloor \frac{2\pi}{m}} \cos(mx)|x|^{\beta-2} dx + \eta(\beta,m), \quad \text{if } m/2 \text{ is even},$$

$$h^*(m/2) = \frac{2\beta(\beta-1)}{\pi m^2} \int_0^{\lfloor \frac{m}{8} \rfloor \frac{2\pi}{m}} \cos(mx)|x|^{\beta-2} dx + \eta(\beta,m), \quad \text{if } m/2 \text{ is odd},$$

where $\sup_{\beta\in[1,2]} |\eta(\beta,m)| = O\left(\frac{1}{m^3}\right)$.

So we are left with analyzing $\int_0^{2k\pi} \cos(x) x^{\beta-2} dx$ for some positive integer $k \leq \lfloor \frac{m}{8} \rfloor$. First, we have $\sup_{\beta\in(1,2]} (\beta-1) \int_0^{2\pi} |\cos x| x^{\beta-2} dx < C$ for some positive constant C that does not depend on m and β. Next, using integration by parts we have $\int_{2\pi}^{2k\pi} \cos(x) x^{\beta-2} dx = (2-\beta) \int_{2\pi}^{2k\pi} \sin(x) x^{\beta-3} dx$ and furthermore for $k \leq \lfloor \frac{m}{8} \rfloor$ we have $\sup_{\beta\in(1,2]} \int_{2\pi}^{2k\pi} |\sin x| x^{\beta-3} dx < C$ for some positive constant

C that does not depend on m and β. Therefore $\int_0^{2k\pi} \cos(x)x^{\beta-2}dx$ is finite and bounded by a constant that does not depend on m for $k \leq \lfloor \frac{m}{8} \rfloor$. We further have $\int_0^{2k\pi} \cos(x)x^{\beta-2}dx = \sum_{i=1}^k a_i$ where $a_i = \int_{2(i-1)\pi}^{2i\pi} \cos(x)x^{\beta-2}dx$. Then by using the properties of the cosine function we can write $a_i = \int_{2(i-1)\pi}^{2(i-1)\pi+\frac{\pi}{2}} \cos(x)f_\beta(x)dx$ for $f_\beta(x)$ being positive, and thus we have that all a_i have the same sign and are different from zero. Moreover, $|a_1| \geq \frac{1}{\sqrt{2}} \left| \int_0^{\frac{\pi}{4}} f_\beta(x)dx \right|$, and direct calculation shows $\inf_{\beta \in (1,2)}(\beta - 1) \left| \frac{1}{\sqrt{2}} \int_0^{\frac{\pi}{4}} f_\beta(x)dx \right| > 0$. Therefore, $|\sum_{i=1}^k a_i|$ is bounded from below by a positive constant uniformly in $\beta \in (1, 2 - \epsilon)$ for every $\epsilon > 0$.

The bound in (6.41) when $\beta < 1$. The proof is very similar to the case when $\beta > 1$ and therefore we provide only a sketch of it. Using integration by parts we can write for some $\eta(\beta, m)$ with $\sup_{\beta \in (0,1)} |\eta(\beta, m)| = O\left(\frac{1}{m^2}\right)$:

$$h^*(m/2) = \begin{cases} -\frac{2}{\pi}\frac{\beta}{m}\int_0^{\lfloor \frac{m}{8} \rfloor \frac{2\pi}{m}} \sin(mx)|\sin(x)|^{\beta-1}dx + \eta(\beta, m), & \text{if } \frac{m}{2} \text{ is even,} \\ \frac{2}{\pi}\frac{\beta}{m}\int_0^{\lfloor \frac{m}{8} \rfloor \frac{2\pi}{m}} \sin(mx)|\sin(x)|^{\beta-1}dx + \eta(\beta, m), & \text{if } \frac{m}{2} \text{ is odd.} \end{cases}$$

Then exactly the same analysis as for the case $\beta > 1$ leads to

$$h^*(m/2) = \begin{cases} -\frac{2}{\pi}\frac{\beta}{m}\int_0^{\lfloor \frac{m}{8} \rfloor \frac{2\pi}{m}} \sin(mx)|x|^{\beta-1}dx + \eta(\beta, m), & \text{if } \frac{m}{2} \text{ is even,} \\ \frac{2}{\pi}\frac{\beta}{m}\int_0^{\lfloor \frac{m}{8} \rfloor \frac{2\pi}{m}} \sin(mx)|x|^{\beta-1}dx + \eta(\beta, m), & \text{if } \frac{m}{2} \text{ is odd,} \end{cases}$$

with $\sup_{\beta \in (0,1)} |\eta(\beta, m)| = O\left(\frac{1}{m^2}\right)$.

From here, we can decompose $\int_0^{2k\pi} \sin(x)|x|^{\beta-1}dx = \sum_{i=1}^k a_i$ with $a_i = \int_{2(i-1)\pi}^{2i\pi} \sin(x)x^{\beta-1}dx = \int_{2(i-1)\pi}^{2i\pi-\pi} \sin(x)(x^{\beta-1} - (x+\pi)^{\beta-1})dx$. Therefore, $a_i > 0$ for $i = 1, ..., k$ as the integrand is always positive. Moreover, the sum $\sum_{i=1}^k a_i$ is bounded from above by a constant. Since a_i is a continuous function of β, the above results hold obviously uniformly over β lying in compact subsets of $(0, 1)$.

The bound in (6.42). We consider only the case when $m/2$ is even, with the case when $m/2$ is odd being proven in exactly the same way. By integration by parts we have

$$h^*(m/2) = -\frac{2\beta}{\pi m}\int_0^{\frac{\pi}{4}} \sin(mx)\left(|\sin x|^{\beta-1}\cos x - |\cos x|^{\beta-1}\sin x\right)dx.$$

We define $h_1^*(m/2) = -\frac{2\beta}{\pi m}\int_0^{\frac{\pi}{4}} \sin(mx)|\sin x|^{\beta-1}dx$ and $h_2^*(m/2) = h^*(m/2) - h_1^*(m/2)$. Using integration by parts again, for $\eta(\beta, m)$ with $\sup_{\beta \in (0,1)} |\eta(\beta, m)| = O\left(\frac{1}{m^3}\right)$, we have

$$h_2^*(m/2) = -\frac{2\beta}{\pi m^2}\cos\left(\frac{m\pi}{4}\right)\sin\left(\frac{\pi}{4}\right)^{\beta-1} + \eta(\beta, m).$$

$$\begin{aligned} h_1^*(m/2) = \frac{2\beta}{\pi m^2}\cos\left(\frac{m\pi}{4}\right)\sin\left(\frac{\pi}{4}\right)^{\beta-1} - \frac{2\beta}{\pi m^2}\sin\left(\lfloor \frac{m}{8} \rfloor \frac{2\pi}{m}\right)^{\beta-1} \\ - \frac{2\beta}{\pi m}\int_0^{\lfloor \frac{m}{8} \rfloor \frac{2\pi}{m}} \sin(mx)|\sin x|^{\beta-1}dx + \eta(\beta, m). \end{aligned}$$

Therefore, for $\eta(\beta,m)$ with $\sup_{\beta\in(0,1)}|\eta(\beta,m)| = O\left(\frac{1}{m^3}\right)$:

$$h^*(m/2) = -\frac{2\beta}{\pi m^2}\int_0^{\lfloor\frac{m}{8}\rfloor 2\pi}\sin x\left|\sin\left(\frac{x}{m}\right)\right|^{\beta-1}dx - \frac{2\beta}{\pi m^2}\sin\left(\lfloor\frac{m}{8}\rfloor\frac{2\pi}{m}\right)^{\beta-1} + \eta(\beta,m).$$

For $a_i = \int_{2(i-1)\pi}^{2i\pi}\sin x\left|\sin\left(\frac{x}{m}\right)\right|^{\beta-1}dx$ we can write $\int_0^{\lfloor\frac{m}{8}\rfloor 2\pi}\sin x\left|\sin\left(\frac{x}{m}\right)\right|^{\beta-1}dx = \sum_{i=1}^{\lfloor\frac{m}{8}\rfloor}a_i$. Now we have $a_i = \int_{2(i-1)\pi}^{(2i-1)\pi}\sin x\left(\left|\sin\left(\frac{x}{m}\right)\right|^{\beta-1} - \left|\sin\left(\frac{x+\pi}{m}\right)\right|^{\beta-1}\right)dx$. Since $\sin(x)$ is an increasing function for $x\in\left(0,\frac{\pi}{4}\right)$, we have that $a_i\geq 0$ for $i=1,...,\lfloor\frac{m}{8}\rfloor$ locally uniformly in β. Therefore, for m sufficiently high, we have $|h^*(m/2)| > (1-\iota)\frac{2\beta}{\pi m^2}\sin\left(\lfloor\frac{m}{8}\rfloor\frac{2\pi}{m}\right)^{\beta-1}$ for some small $\iota\in(0,1)$.

The bound in (6.43) when $\beta > 1$. By changing variable of integration, it suffices to look at the order of magnitude of $\int_0^{\frac{\pi}{2}}\cos(mx)|\cos x|^\beta\log^q|\cos x|dx$ and $\int_0^{\frac{\pi}{2}}\cos(mx)|\sin x|^\beta\log^q|\sin x|dx$ in m, for m being an even integer. Applying integration by parts twice, exactly as in the proof of (6.41), we have

$$\int_0^{\frac{\pi}{2}}\cos(mx)|\cos x|^\beta\log^q|\cos x|dx$$
$$= \frac{1}{m^2}\int_0^{\frac{\pi}{2}}\cos(mx)|\cos x|^{\beta-2}\left(\sum_{j=0}^{2}\alpha_j\log^{q-j}|\cos x|\right)dx + \eta(\beta,m),$$

where α_j are constants depending on q and β, $\alpha_j = 0$ if $q-j<0$, and $\sup_{\beta\in(1,2)}|\eta(\beta,m)| = O\left(\frac{1}{m^3}\right)$. We have similar expression for the integral in which $\cos x$ is replaced with $\sin x$. As in the proof of (6.41), we have

$$\left|\int_0^{\frac{\pi}{2}}\cos(mx)|\cos x|^{\beta-2}dx\right| < C_\beta m^{1-\beta} \text{ and } \left|\int_0^{\frac{\pi}{2}}\cos(mx)|\sin x|^{\beta-2}dx\right| < C_\beta m^{1-\beta},$$

with $\sup_{\beta\in\mathcal{B}}C_\beta<\infty$ for $\mathcal{B}$ being an arbitrary compact subset of $(1,2)$.

Hence we are left with deriving the order of magnitude of $\int_0^{\frac{\pi}{2}}\cos(mx)q(x)dx$ for $q(x) = |\cos x|^{\beta-2}\log^q|\cos x|$ or $q(x) = |\sin x|^{\beta-2}\log^q|\sin x|$ and q being a positive integer. Since we are interested in the above integrals only for even integers m, a change of variable of integration reduces the problem finally to evaluating the order of magnitude of the following integrals $\int_0^{\frac{\pi}{4}}\cos(mx)q(x)dx$, for $q(x)$ being $|\cos x|^{\beta-2}\log^q|\cos x|$ or $|\sin x|^{\beta-2}\log^q|\sin x|$. Integration by parts shows that $\sup_{\beta\in(1,2)}\left|\int_0^{\frac{\pi}{4}}\cos(mx)q(x)dx\right| = O(1/m)$, for $q(x) = |\cos x|^{\beta-2}\log^q|\cos x|$, and hence the case $q(x) = |\sin x|^{\beta-2}\log^q|\sin x|$ remains only. By changing the variable of integration we need to evaluate the integral $\frac{1}{m}\int_0^{m\frac{\pi}{4}}\cos xq(x/m)dx$ for m being an even integer. We split the interval of integration into three and bound each of the resulting integrals using the inequality $|\sin x/x - 1|\leq\frac{|x|}{2}$ and integration by parts for the integral over the middle region. Altogether we get for m sufficiently high

$$\int_0^{2\pi}|\cos xq(x/m)|dx + \left|\int_{2\pi}^{\lfloor\frac{m}{8}\rfloor 2\pi}\cos xq(x/m)dx\right| \leq C_\beta m^{2-\beta}\log^q(m),$$

and $\int_{\lfloor\frac{m}{8}\rfloor 2\pi}^{m\frac{\pi}{4}} |\cos x q(x/m)| dx \leq C_\beta$ for m sufficiently high. For C_β we have $\sup_{\beta\in\mathcal{B}} C_\beta < \infty$ for $\mathcal{B}$ being an arbitrary compact subset of $(1,2)$. Combining these results we have the bound in (6.43).

The bound in (6.43) when $\beta \leq 1$. Applying integration by parts we have

$$\begin{aligned}&\int_0^{\frac{\pi}{2}} \cos(mx)|\cos x|^\beta \log^q|\cos x| dx \\ &\quad = \frac{1}{m}\int_0^{\frac{\pi}{2}} \sin(mx)|\cos x|^{\beta-1}\left(\sum_{j=0}^{1} \alpha_j \log^{q-j}|\cos x|\right) dx + \eta(\beta, m),\end{aligned}$$

where α_j are constants depending on q and β, $\alpha_j = 0$ if $q - j < 0$, and $\sup_{\beta\in(0,1]} |\eta(\beta, m)| = O\left(\frac{1}{m^2}\right)$. Then exactly the same steps as for the case $\beta > 1$, reduce the problem to deducing the order of magnitude of the integrals $\frac{1}{m}\int_0^{m\frac{\pi}{4}} \sin(x) q(\frac{x}{m}) dx$ for $q(x) = |\sin x|^{\beta-1}\log^q|\sin x|$ and $q(x) = |\cos x|^{\beta-1}\log^q|\cos x|$. Deriving the order of magnitude of these integrals is done in exactly the same way as for their counterparts in the case $\beta > 1$ by splitting the region of integration into the intervals $(0, \frac{\pi}{2})$, $(\frac{\pi}{2}, \lfloor\frac{m-2}{4}\rfloor\pi + \frac{\pi}{2})$ and $(\lfloor\frac{m-2}{4}\rfloor\pi + \frac{\pi}{2}, \frac{m\pi}{4})$.

The bound in (6.44). This is shown in exactly the same way as the bound in (6.43) when $\beta \leq 1$. $\square$

Before stating the next lemma, we introduce the following additional notation:

$$\mathcal{B}_p^n(u_1, u_2) = \frac{1}{2k_n}\left(\frac{\Sigma_p}{(2C_\beta)^{\frac{2p}{\beta}}\mu_p^2 K_{g,p}^2} - 1\right)\overline{\mathcal{B}}_p(u_1, u_2),$$

where

$$\overline{\mathcal{B}}_p(u_1, u_2) = \left[\log(\mathcal{L}_p(u_1, u_2))\frac{\beta}{p}\left(\frac{\beta}{p} + 1\right) + \log^2(\mathcal{L}_p(u_1, u_2))\left(\frac{\beta}{p}\right)^2\right],$$

and

$$\Sigma_p = \mathbb{E}\left(\left|\widetilde{\mathbf{S}}_3^{(1)} - 2\widetilde{\mathbf{S}}_2^{(1)} + \widetilde{\mathbf{S}}_1^{(1)}\right|^p + \left|\widetilde{\mathbf{S}}_3^{(2)} - 2\widetilde{\mathbf{S}}_2^{(2)} + \widetilde{\mathbf{S}}_1^{(2)}\right|^p\right)^2,$$

with the bivariate stable process $\widetilde{\mathbf{S}}$ being defined in Section 6.1, and where $\widetilde{\mathbf{S}}^{(j)}$ denotes the j-th element of the vector process.

Similarly, we further denote with $\mathcal{B}_{p,\beta}^n$ the analogue of $\widehat{\mathcal{B}}_{p,\beta}^n$ in which $\overline{\beta}(p)$ is replaced with β, $\widehat{\Sigma}_p^n$ with Σ_p, and $\widehat{\mathcal{K}}_p^n$ with $\mathcal{K}_p = (2C_\beta)^{\frac{2p}{\beta}}\mu_p^2 K_{g,p}^2$.

Henceforth, we will denote with $C > 0$ a constant that does not depend on n, and when it depends in addition on some parameter k, we will denote it with $C_k > 0$, and in this case C_k will be locally bounded in k, i.e., $\sup_{k\in\overline{\mathcal{K}}} C_k < \infty$ for any compact set $\overline{\mathcal{K}}$. We also denote $\mathbf{1} = (1,1)$. Finally, we set for $\iota > 0$ arbitrary small constant:

$$\alpha_n = \Delta_n^{\frac{\beta}{2}\frac{p+1}{\beta+1}\wedge\left(\frac{p}{\beta'}\wedge 1-\frac{p}{\beta}\right)-\iota}.$$

Lemma 2 *Assume A and SB, $p \in \left(\frac{\beta}{4}-\varepsilon, \frac{\beta}{4}+\varepsilon\right)$ for some $0<\varepsilon<\left(\frac{1}{\beta}-\frac{\beta}{4}\right)\wedge\frac{\beta}{12}$, and $k_n \asymp n^{\varpi}$ and $m_n \asymp n^{\varrho}$, for $\varpi \in (0,\frac{1}{2})$ and $\varrho>0$. Denote $\mathbf{u}=(u_1,u_2)$, with $u_1,u_2\in\mathbb{R}_+$. We have for $\forall\iota>0$:*

$$\sup_{\mathbf{u}\in\mathbb{S}^1}|\widehat{\mathcal{L}}_p^n(u_1,u_2)-\mathcal{L}_p(u_1,u_2)| = O_p\left(\alpha_n \vee \Delta_n^{\frac{\beta\wedge 1}{2}-\iota}\vee\frac{1}{\sqrt{k_n}}\right), \tag{6.51}$$

$$\begin{aligned}\mathbb{E}&\left|\frac{\widehat{\mathcal{L}}_p^n(u_1,u_2)-\mathcal{L}_p(u_1,u_2)}{\mathcal{L}_p(u_1,u_2)}-\mathcal{B}_p^n(u_1,u_2)\right|^2\\ &\le C_{\mathbf{u}}\left[\frac{k_n}{n}\vee\alpha_n^2\vee\frac{1}{k_n}\left(\frac{k_n}{n}\right)^{1-\frac{2p}{\beta}-\iota}\vee\frac{1}{k_n^{\frac{\beta}{2p}\wedge 2-\iota}}\right],\end{aligned} \tag{6.52}$$

$$\begin{aligned}\mathbb{E}&\left|\frac{\widehat{\mathcal{L}}_p^n(u_1,u_2)}{\mathcal{L}_p(u_1,u_2)}-1-\mathcal{B}_p^n(u_1,u_2)-\frac{\Phi(\mathbf{u})}{\Phi(\mathbf{1})}\left(\frac{\widehat{\mathcal{L}}_p^n(1,1)}{\mathcal{L}_p(1,1)}-1-\mathcal{B}_p^n(1,1)\right)\right|\\ &\le C_{\mathbf{u}}\left[\Delta_n^{\frac{\beta\wedge 1}{2}-\iota}\vee\alpha_n^2\vee\frac{\alpha_n}{\sqrt{k_n}}\vee\frac{\Delta_n^{\frac{p}{2}-\iota}}{k_n}\vee k_n^{-\frac{3}{2}}\vee\frac{1}{k_n}\left(\frac{k_n}{n}\right)^{\frac{1}{2}-\frac{p}{\beta}-\iota}\right],\end{aligned} \tag{6.53}$$

$$\Delta_n^{1-\frac{2p}{\beta}}\widehat{\Sigma}_p^n-\Sigma_p\int_0^1|\sigma_s|^{2p}ds = O_p\left(\Delta_n^{\frac{p}{2}\wedge(\frac{p}{\beta'}-\frac{p}{\beta})-\iota}\right), \tag{6.54}$$

$$\Delta_n^{1-\frac{2p}{\beta}}\widehat{\mathcal{K}}_p^n-(2C_\beta)^{\frac{2p}{\beta}}\mu_p^2K_{g,p}^2\int_0^1|\sigma_s|^{2p}ds = O_p\left(\Delta_n^{\frac{p}{2}\wedge(\frac{p}{\beta'}-\frac{p}{\beta})-\iota}\right), \tag{6.55}$$

$$\widehat{\beta}(p)-\beta = O_p\left(\alpha_n^2\vee\Delta_n^{\frac{\beta\wedge 1}{2}-\iota}\vee\frac{\alpha_n}{\sqrt{k_n}}\vee\frac{\Delta_n^{\frac{p}{2}-\iota}}{k_n}\vee k_n^{-\frac{3}{2}}\vee\frac{1}{k_n}\left(\frac{k_n}{n}\right)^{\frac{1}{2}-\frac{p}{\beta}-\iota}\right). \tag{6.56}$$

Proof of Lemma 2. *Preliminary results.* In the proof we will denote with $\overline{\mathcal{U}}$ a compact subset of $\mathbb{R}^2$. We also use the shorthand notation

$$V(p)_i^n = V(\mathbf{X}^{(1)},p)_i^n + V(\mathbf{X}^{(2)},p)_i^n.$$

We will make use of several bounds which we now state and/or derive. First, using Assumption SB(a), the smoothness of the truncation function κ as well as the boundedness from below of the values of the process L_t, we have

$$\mathbb{E}_{i-2}^n\left|\left|\int_{(i-2)\Delta_n}^{i\Delta_n}(\widetilde{\mathbf{a}}_s-\widetilde{\mathbf{a}}_{s-\Delta_n})ds\right|\right|^q \le C\Delta_n^{\frac{3q}{2}},\quad q\in(0,2]. \tag{6.57}$$

and given the additional condition for $\widetilde{\mathbf{a}}$ in Assumption SB(c) we also have for some arbitrary small $\iota>0$:

$$\mathbb{E}_{i-2}^n\left|\left|\int_{(i-2)\Delta_n}^{i\Delta_n}(\widetilde{\mathbf{a}}_s-\widetilde{\mathbf{a}}_{s-\Delta_n})ds\right|\right|^q \le C\Delta_n^{q+\frac{q}{\beta}\wedge 1-\iota},\quad \text{if } \beta<1. \tag{6.58}$$

Using Assumption SB(c) as well as the algebraic inequality $|\sum_i a_i|^p \le \sum_i |a_i|^p$ for $p\in(0,1)$ and $a_i\in\mathbb{R}$ (and Burkholder-Davis-Gundy inequality if $q>1$), we have

$$\mathbb{E}_{i-2}^n||\Delta_i^n\widetilde{Y}-\Delta_{i-1}^n\widetilde{Y}||^q \le C\Delta_n^{\frac{q}{\beta'}\wedge 1-\iota},\quad q>0,\quad \forall\iota>0. \tag{6.59}$$

We next split $\widetilde{\mathbf{S}}_t=\widetilde{\mathbf{S}}_{t,1}+\widetilde{\mathbf{S}}_{t,2}$ where

$$\widetilde{\mathbf{S}}_{t,1}=\int_0^t\int_{\mathbb{R}^2}\kappa(\mathbf{x})\mu(ds,d\mathbf{x})\ \text{ and }\ \widetilde{\mathbf{S}}_{t,2}=\int_0^t\int_{\mathbb{R}^2}\kappa'(\mathbf{x})\mu(ds,d\mathbf{x}),\quad \text{when } \beta<1,$$

$$\widetilde{\mathbf{S}}_{t,1} = \int_0^t \int_{\mathbb{R}^2} \kappa(\mathbf{x})\widetilde{\mu}(ds, d\mathbf{x}) \text{ and } \widetilde{\mathbf{S}}_{t,2} = \int_0^t \int_{\mathbb{R}^2} \kappa'(\mathbf{x})\mu(ds, d\mathbf{x}), \quad \text{when } \beta \geq 1.$$

Using Burkholder-Davis-Gundy inequality, the algebraic inequality $|\sum_i a_i|^p \leq \sum_i |a_i|^p$ for $p \in (0,1)$ and $a_i \in \mathbb{R}$, and Jensen's inequality, we have for $j = 0, 1$:

$$\mathbb{E}_{i-2}^n \left|\left| \int_{(i-j-1)\Delta_n}^{(i-j)\Delta_n} (\sigma_{u-} - \sigma_{(i-2)\Delta_n}) d\widetilde{\mathbf{S}}_{u,1} \right|\right|^q \leq C\Delta_n^{\frac{q}{2}+\frac{q}{\beta}\wedge 1-\iota}, \ q \in (0,2], \ \forall \iota > 0. \tag{6.60}$$

Next, for $0 < q < \beta \wedge 1$, since the q-th power of the jumps of $\widetilde{\mathbf{S}}_{t,2}$ is absolutely summable, we have for $j = 0, 1$ and $q \in (0, \beta \wedge 1)$:

$$\mathbb{E}_{i-2}^n \left|\left| \int_{(i-j-1)\Delta_n}^{(i-j)\Delta_n} (\sigma_u - \sigma_{(i-2)\Delta_n}) d\widetilde{\mathbf{S}}_{u,2} \right|\right|^q \leq C\Delta_n^{1+\frac{q}{2}}. \tag{6.61}$$

Finally, using the above inequalities we can proceed exactly as in Lemma 1 of Todorov (2017) (note that given the restriction on p in the lemma we have $p < 1/\beta$ which is needed for applying Lemma 1 of Todorov (2017)), and show

$$\begin{aligned} &\Delta_n^{-p/\beta}\mathbb{E}_{i-2}^n||\Delta_i^n\mathbf{X}^{(j)} - \Delta_{i-1}^n\mathbf{X}^{(j)}|^p - |\sigma_{(i-2)\Delta_n}|^p|\Delta_i^n\widetilde{\mathbf{S}}^{(j)} - \Delta_{i-1}^n\widetilde{\mathbf{S}}^{(j)}|^p| \\ &\quad \leq C\alpha_n, \ j = 1, 2, \end{aligned} \tag{6.62}$$

$$\mathbb{P}\left((\mathcal{A}_i^n)^c\right) < Ck_n^{-\frac{\beta}{2p}+\iota}, \quad \mathcal{A}_i^n = \{\omega : \Delta_n^{-p/\beta}V(p)_i^n > \epsilon\}, \tag{6.63}$$

for some sufficiently small $\epsilon > 0$ (which depends on the lower bound on the process L, and the value of β). The bounds in (6.62) and (6.63) are derived in Todorov (2017) under the assumption $\beta \geq 1$ but they are easily extended to the case $\beta < 1$ given the results in (6.57)-(6.61) above (note that (6.41) in Todorov (2017) gets changed to $K\epsilon^{p-\beta}\mathbb{E}_{i-2}^n|\widetilde{\chi}_2|^\beta$ in the notation of that paper).

The results in (6.51)-(6.53). We denote the function

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

where $\Phi(\mathbf{u})$ is defined in (1.3) and we note that $f_i^n\left((2C_\beta)^{\frac{p}{\beta}}|\sigma_{(i-2)\Delta_n}|^p K_{g,p}\mu_p, \mathbf{u}\right) \equiv \mathcal{L}_p(u_1, u_2)$. With this notation, we can make the decomposition

$$\widehat{\mathcal{L}}_p^n(u_1, u_2) - \mathcal{L}_p(u_1, u_2) = Z^{(1)}(\mathbf{u}) + Z^{(2)}(\mathbf{u}) + Z^{(3)}(\mathbf{u}),$$

where $Z^{(j)}(\mathbf{u}) = \frac{1}{n-2k_n-1}\sum_{i=2k_n+2}^n z_i^{(j)}(\mathbf{u})$ for

$$z_i^{(1)}(\mathbf{u}) = \cos(u_1\mathbf{X}^{(1)}(p)_i^n + u_2\mathbf{X}^{(2)}(p)_i^n) - \mathbb{E}_{i-2}^n(\cos(u_1\mathbf{X}^{(1)}(p)_i^n + u_2\mathbf{X}^{(2)}(p)_i^n)),$$

$$z_i^{(2)}(\mathbf{u}) = \mathbb{E}_{i-2}^n(\cos(u_1\mathbf{X}^{(1)}(p)_i^n + u_2\mathbf{X}^{(2)}(p)_i^n)) - f_i^n(\Delta_n^{-p/\beta}V(p)_i^n, \mathbf{u}),$$

$$z_i^{(3)}(\mathbf{u}) = f_i^n(\Delta_n^{-p/\beta}V(p)_i^n, \mathbf{u}) - \mathcal{L}_p(u_1, u_2).$$

We further define $\widetilde{Z}^{(j)}(\mathbf{u}) = Z^{(j)}(\mathbf{u})/\mathcal{L}_p(u_1,u_2)$ and $\widetilde{z}^{(j)}(\mathbf{u}) = z^{(j)}(\mathbf{u})/\mathcal{L}_p(u_1,u_2)$.

We split the proof of (6.51)-(6.53) into several steps.

Step 1. Results for $Z^{(1)}(\mathbf{u})$. Using the bound in (6.63) we have

$$\mathbb{E}\left(\sup_{u\in\overline{\mathcal{U}}}(|z_i^{(1)}(\mathbf{u})|^q)1_{\{(\mathcal{A}_i^n)^c\}}\right) \le Ck_n^{-\frac{\beta}{2p}+\iota}, \quad \forall q>0. \tag{6.64}$$

Using successive conditioning and the boundedness of $z_i^{(1)}(\mathbf{u})$, we have

$$\mathbb{E}\left(\sum_{i=2k_n+2}^{n} z_i^{(1)}(\mathbf{u})1_{\{\mathcal{A}_i^n\}}\right)^2 \le C_{\mathbf{u}}n. \tag{6.65}$$

Combining the above two bounds, we have for some $\iota>0$ arbitrary small:

$$\mathbb{E}\big|\widetilde{Z}^{(1)}(\mathbf{u})\big|^2 \le C_{\mathbf{u}}\big(\Delta_n \vee k_n^{-\frac{\beta}{2p}+\iota}\big). \tag{6.66}$$

We next denote

$$\begin{aligned}\overline{z}_i^{(1)}(\mathbf{u}) = \cos&\left(\frac{u_1\sigma_{(i-2)\Delta_n}(\Delta_i^n\widetilde{\mathbf{S}}^{(1)}-\Delta_{i-1}^n\widetilde{\mathbf{S}}^{(1)})+u_2\sigma_{(i-2)\Delta_n}(\Delta_i^n\widetilde{\mathbf{S}}^{(2)}-\Delta_{i-1}^n\widetilde{\mathbf{S}}^{(2)})}{(V(p)_i^n)^{1/p}}\right)\\ &- f_i^n(\Delta_n^{-p/\beta}V(p)_i^n,\mathbf{u}).\end{aligned}$$

Using the bounds in (6.57)-(6.61), we have for $\forall\iota>0$:

$$\mathbb{E}\left(\sum_{i=2k_n+2}^{n}\sup_{\mathbf{u}\in\overline{\mathcal{U}}}|z_i^{(1)}(\mathbf{u})-\overline{z}_i^{(1)}(\mathbf{u})|^q1_{\{\mathcal{A}_i^n\}}\right) \le C_{\overline{\mathcal{U}}}\,n\Delta_n^{\frac{\beta\wedge 1}{2}-\iota}, \quad \forall q\ge 1. \tag{6.67}$$

If we further denote

$$\mathcal{B}_i^n = \left\{\omega : \Delta_n^{-\frac{1}{\beta}}|\Delta_i^n\widetilde{\mathbf{S}}^{(1)}-\Delta_{i-1}^n\widetilde{\mathbf{S}}^{(1)}| \le \sqrt{n}\ \cap\ \Delta_n^{-\frac{1}{\beta}}|\Delta_i^n\widetilde{\mathbf{S}}^{(2)}-\Delta_{i-1}^n\widetilde{\mathbf{S}}^{(2)}| \le \sqrt{n}\right\},$$

then with this notation we have for $\forall\iota>0$:

$$\mathbb{E}\left(\sum_{i=2k_n+2}^{n}\sup_{\mathbf{u}\in\overline{\mathcal{U}}}|\overline{z}_i^{(1)}(\mathbf{u})|^q1_{\{(\mathcal{B}_i^n)^c\}}\right) \le C_{\overline{\mathcal{U}}}n^{1-\frac{\beta}{2}+\iota}, \quad q>0. \tag{6.68}$$

Next, for $\mathbf{u},\mathbf{v}\in\mathbb{R}^2$, by splitting the sum $\sum_{i=2k_n+2}^n(\overline{z}_i^{(1)}(\mathbf{u})-\overline{z}_i^{(1)}(\mathbf{v}))1_{\{\mathcal{A}_i^n\cap\mathcal{B}_i^n\}})$ into sums over the odd and even terms, each of the latter becomes a sum of martingale increments, and then by applying Burkholder-Davis-Gundy inequality and inequality in means we have

$$\begin{aligned}\mathbb{E}\Bigg|\sum_{i=2k_n+2}^{n}&(\overline{z}_i^{(1)}(\mathbf{u})-\overline{z}_i^{(1)}(\mathbf{v}))1_{\{\mathcal{A}_i^n\cap\mathcal{B}_i^n\}}\Bigg|^q\\ &\le Cn^{\frac{q}{2}-1}\sum_{i=2k_n+2}^{n}\mathbb{E}(|(\overline{z}_i^{(1)}(\mathbf{u})-\overline{z}_i^{(1)}(\mathbf{v}))1_{\{\mathcal{A}_i^n\cap\mathcal{B}_i^n\}}|^q),\end{aligned}$$

for any $q\ge 2$. Now, using the asymptotic behavior of the tail probability of a stable random variable and Karamata's theorem (Proposition 1.5.8 in Bingham et al. (1987)), we have for $\iota>0$ sufficiently small and n sufficiently high:

$$\mathbb{E}|(\overline{z}_i^{(1)}(\mathbf{u})-\overline{z}_i^{(1)}(\mathbf{v}))1_{\{\mathcal{A}_i^n\cap\mathcal{B}_i^n\}}|^2 \le Cn^{\frac{1+\iota-\beta}{2}\vee 0}|\theta_u-\theta_v|^{1+\iota},$$

when $\mathbf{u} = (\cos(\theta_u), \sin(\theta_u))$ and $\mathbf{v} = (\cos(\theta_v), \sin(\theta_v))$, with $\theta_u,\ \theta_v \in [0, 2\pi)$. Therefore, applying standard criteria for tightness on a compact interval on the real line, see e.g., Theorem 12.3 in Billingsley (1968), we have $n^{\frac{\beta\wedge 1}{2}-\iota-1}\sum_{i=2k_n+2}^{n}\overline{z}_i^{(1)}(\mathbf{u})$, with $\mathbf{u}\in\mathbb{S}^1$ and $\iota>0$ arbitrary small, is tight. Combining this with the bounds in (6.64) and (6.67), we have

$$\frac{1}{n-2k_n-1}\sum_{i=2k_n+2}^{n} z_i^{(1)}(\mathbf{u}) = o_p\big(\Delta_n^{\frac{\beta\wedge 1}{2}-\iota} \vee k_n^{-\frac{\beta}{2p}+\iota}\big), \quad \text{uniformly for } \mathbf{u}\in\mathbb{S}^1. \tag{6.69}$$

Step 2. Results for $Z^{(2)}(\mathbf{u})$*.* As in the previous step we have

$$\mathbb{E}\left(\sup_{u\in\overline{\mathcal{U}}}(|z_i^{(2)}(\mathbf{u})|^q)1_{\{(\mathcal{A}_i^n)^c\}}\right) \leq Ck_n^{-\frac{\beta}{2p}+\iota}, \quad \forall q>0,$$

and

$$\mathbb{E}\left(\sup_{u\in\overline{\mathcal{U}}}|z_i^{(2)}(\mathbf{u})|^q 1_{\{\mathcal{A}_i^n\}}\right) \leq C_{\overline{\mathcal{U}}}\Delta_n^{q\frac{\beta\wedge 1}{2}-\iota}, \quad \forall q\geq 1.$$

Applying these bounds we easily get for $\iota>0$ arbitrary small

$$\mathbb{E}\left(\sup_{u\in\overline{\mathcal{U}}}|\widetilde{Z}^{(2)}(\mathbf{u})|\right)^q \leq C_{\overline{\mathcal{U}}}\,(k_n^{-\frac{\beta}{2p}+\iota} \vee \Delta_n^{q\frac{\beta\wedge 1}{2}-\iota}),\ q\in[1,2]. \tag{6.70}$$

Step 3. Results for $Z^{(3)}(\mathbf{u})$*.* First, we denote

$$\widetilde{V}(p)_i^n = \frac{1}{k_n}\sum_{j\in I_n^i}\mathbb{E}_{j-2}^n\{|\Delta_j^n\mathbf{X}^{(1)}-\Delta_{j-1}^n\mathbf{X}^{(1)}|^p + |\Delta_j^n\mathbf{X}^{(2)}-\Delta_{j-1}^n\mathbf{X}^{(2)}|^p\}.$$

Then using Burkholder-Davis-Gundy inequality and Hölder's inequality, we have

$$\mathbb{E}|\Delta_n^{-p/\beta}(V(p)_i^n - \widetilde{V}(p)_i^n)|^q \leq Ck_n^{-\frac{q}{2}},\ \ 0<q<\frac{\beta}{p}. \tag{6.71}$$

Next, using the smoothness in expectation assumption for σ in Assumption SB(a), we have

$$\mathbb{E}_s\big|\sigma_t-\sigma_s\big|^q \leq C|t-s|^{\frac{q}{2}\wedge 1}, \quad 0\leq s<t, \quad q>0. \tag{6.72}$$

Combining the above two results as well as using (6.62) and applying first-order Taylor expansion for $z_i^{(3)}(\mathbf{u})$, we have

$$\mathbb{E}\left(\sup_{u\in\overline{\mathcal{U}}}|z_i^{(3)}(\mathbf{u})|\right)^q \leq C_{\overline{\mathcal{U}}}(\alpha_n\vee k_n^{-1/2}\vee\sqrt{k_n\Delta_n})^q, \quad q\in[0,2]. \tag{6.73}$$

Next, we denote

$$\widetilde{z}_i^{(3,a)}(\mathbf{u}) = -\frac{\beta}{p}\log(\mathcal{L}_p(u_1,u_2))\frac{\Delta_n^{-p/\beta}V(p)_i^n - (2C_\beta)^{\frac{p}{\beta}}\mu_pK_{g,p}|\sigma_{(i-2)\Delta_n}|^p}{(2C_\beta)^{\frac{p}{\beta}}\mu_pK_{g,p}|\sigma_{(i-2)\Delta_n}|^p},$$

and the analogous expression, in which $\sigma_{(i-2)\Delta_n}$ in the denominator is replaced with $\sigma_{(i-2k_n-2)\Delta_n}$, with $\overline{z}_i^{(3,a)}(\mathbf{u})$. Then using Hölder's inequality, the bounds in (6.71)-(6.72) as well as the fact that $p < \beta/2$, we get

$$\mathbb{E}|\Delta_n^{-p/\beta}(V(p)_i^n - \widetilde{V}(p)_i^n)(\sigma_{(i-2)\Delta_n} - \sigma_{(i-2k_n-2)\Delta_n})|^2 \leq Ck_n^{-1}(k_n\Delta_n)^{1-\frac{2p}{\beta}-\iota},\ \forall\iota > 0,$$

and this inequality and applying again (6.62) and (6.72) leads to

$$\mathbb{E}|\widetilde{z}_i^{(3,a)}(\mathbf{u}) - \overline{z}_i^{(3,a)}(\mathbf{u})|^2 \leq C_{\mathbf{u}}\,(k_n^{-1}(k_n\Delta_n)^{1-\frac{2p}{\beta}-\iota} \vee k_n\Delta_n),\ \forall\iota > 0. \tag{6.74}$$

Next, using successive conditioning, the fact that $\mathbb{E}_{i-2k_n-2}^n\left(\frac{V(p)_i^n - \widetilde{V}(p)_i^n}{|\sigma_{(i-2k_n-2)\Delta_n}|^p}\right) = 0$ as well as Cauchy-Schwartz inequality and (6.71), we have

$$\Delta_n^{-2p/\beta}\mathbb{E}\left(\sum_{i=2k_n+2}^{n}\frac{V(p)_i^n - \widetilde{V}(p)_i^n}{|\sigma_{(i-2k_n-2)\Delta_n}|^p}\right)^2 \leq Cn. \tag{6.75}$$

Using first-order Taylor expansion, and the boundedness of the first and second derivatives of the function f in x, we have

$$|\widetilde{z}_i^3(\mathbf{u}) - \widetilde{z}_i^{(3,a)}(\mathbf{u})| \leq C_{\mathbf{u}}\,|\Delta_n^{-p/\beta}V(p)_i^n - (2C_\beta)^{p/\beta}\mu_p K_{g,p}|\sigma_{(i-2)\Delta_n}|^p|^{\frac{\beta}{2p}\bigwedge 2-\iota},\ \ \forall\iota > 0.$$

Therefore, using (6.62) and (6.71) as well as $p < \beta/3$ (note the restriction in ϵ in the statement of the lemma), we have

$$\mathbb{E}|\widetilde{z}_i^3(\mathbf{u}) - \widetilde{z}_i^{(3,a)}(\mathbf{u})|^2 \leq C_{\mathbf{u}}\,(k_n^{-\frac{\beta}{2p}\bigwedge 2+\iota} \vee \alpha_n^3 \vee k_n\Delta_n). \tag{6.76}$$

Combining (6.74)-(6.76) as well as (6.62) and (6.72), we finally get

$$\mathbb{E}(\widetilde{Z}^{(3)}(\mathbf{u}))^2 \leq C_{\mathbf{u}}\,(k_n^{-1}(k_n\Delta_n)^{1-\frac{2p}{\beta}-\iota} \vee \alpha_n^2 \vee k_n^{-\frac{\beta}{2p}\bigwedge 2+\iota} \vee k_n\Delta_n),\ \forall\iota > 0. \tag{6.77}$$

Using third-order Taylor series expansion, we can further decompose

$$\widetilde{z}_i^{(3)}(\mathbf{u}) = \widetilde{z}_i^{(3,a)}(\mathbf{u}) + \widetilde{z}_i^{(3,b)}(\mathbf{u}) + \widetilde{z}_i^{(3,c)}(\mathbf{u}),$$

where we denote

$$\widetilde{z}_i^{(3,b)}(\mathbf{u}) = \overline{\mathcal{B}}_p(u_1,u_2)\frac{(\Delta_n^{-p/\beta}V(p)_i^n - (2C_\beta)^{p/\beta}\mu_p K_{g,p}|\sigma_{(i-2)\Delta_n}|^p)^2}{(2C_\beta)^{2p/\beta}\mu_p^2 K_{g,p}^2|\sigma_{(i-2)\Delta_n}|^{2p}},$$

and $\widetilde{z}_i^{(3,c)}(\mathbf{u})$ is defined as the residual $\widetilde{z}_i^{(3,c)}(\mathbf{u}) = \widetilde{z}_i^{(3)}(\mathbf{u}) - \widetilde{z}_i^{(3,a)}(\mathbf{u}) - \widetilde{z}_i^{(3,b)}(\mathbf{u})$. We then denote $\overline{z}_i^{(3,b)}(\mathbf{u})$ from $\widetilde{z}_i^{(3,b)}(\mathbf{u})$ by replacing $\sigma_{(i-2)\Delta_n}$ in the denominator with $\sigma_{(i-2k_n-2)\Delta_n}$, and further $(2C_\beta)^{\frac{p}{\beta}}\mu_p K_{g,p}|\sigma_{(i-2)\Delta_n}|^p$ in the numerator with $\Delta_n^{-p/\beta}\widetilde{V}(p)_i^n$. Then, applying the bounds in (6.62) and (6.72) as well as Hölder's inequality, we have for $\iota > 0$ arbitrary small:

$$\mathbb{E}|\widetilde{z}_i^{(3,b)}(\mathbf{u}) - \overline{z}_i^{(3,b)}(\mathbf{u})| \leq C_{\mathbf{u}}\left[\alpha_n^2 \vee \frac{k_n}{n} \vee \frac{\alpha_n}{\sqrt{k_n}} \vee \sqrt{\Delta_n} \vee \frac{1}{k_n}\left(\frac{k_n}{n}\right)^{\frac{1}{2}-\frac{p}{\beta}-\iota}\right]. \tag{6.78}$$

We next set

$$\overline{\xi}_i^n = \sum_{j=1}^{2}\{|\Delta_i^n \mathbf{X}^{(j)} - \Delta_{i-1}^n \mathbf{X}^{(j)}|^p - \mathbb{E}_{i-2}^n(|\Delta_i^n \mathbf{X}^{(j)} - \Delta_{i-1}^n \mathbf{X}^{(j)}|^p)\},$$

$$\widehat{\xi}_i^n = |\sigma_{(i-2)\Delta_n}|^p \sum_{j=1}^{2}\{|\Delta_i^n \widetilde{\mathbf{S}}^{(j)} - \Delta_{i-1}^n \widetilde{\mathbf{S}}^{(j)}|^p - \mathbb{E}_{i-2}^n(|\Delta_i^n \widetilde{\mathbf{S}}^{(j)} - \Delta_{i-1}^n \widetilde{\mathbf{S}}^{(j)}|^p)\},$$

and $\widetilde{\xi}_i^n = \widehat{\xi}_i^n / |\sigma_{(i-2)\Delta_n}|^p$. With this notation we can split (recall the definition of I_n^i after (3.4))

$$(V(p)_i^n - \widetilde{V}(p)_i^n)^2 = \frac{1}{k_n^2}\sum_{k\in I_n^i}(\overline{\xi}_k^n)^2 + \frac{1}{k_n^2}\sum_{k,l\in I_n^i:|k-l|\geq 2}\overline{\xi}_k^n\overline{\xi}_l^n, \tag{6.79}$$

and analyze separately the sums on the right hand side of the above equality. We start with the first of the two. Applying Hölder's inequality and the bounds in (6.57)-(6.62), we get

$$\Delta_n^{-\frac{2p}{\beta}}\mathbb{E}||\overline{\xi}_i^n|^2 - |\widehat{\xi}_i^n|^2| \leq C(\Delta_n^{\frac{p}{2}-\iota} \vee \alpha_n),\ \forall \iota > 0. \tag{6.80}$$

Next, using Assumption SB(a) and successive conditioning, we have

$$\mathbb{E}\left|\frac{\sum_{k\in I_n^i}(\widehat{\xi}_k^n)^2}{|\sigma_{(i-2k_n-2)\Delta_n}|^p} - \sum_{k\in I_n^i}(\widetilde{\xi}_k^n)^2\right| \leq C k_n\sqrt{\frac{k_n}{n}}. \tag{6.81}$$

Finally, using Hölder's inequality and Burkholder-Davis-Gundy inequality and since $p < \beta/2$, we have (recall the notation of $\mathcal{K}_p$ before the statement of the lemma):

$$\mathbb{E}\left|\frac{1}{k_n(n-2k_n-1)}\sum_{i=2k_n+2}^{n}\left(\Delta_n^{-2p/\beta}\sum_{k\in I_n^i}(\widetilde{\xi}_k^n)^2\right) - (\Sigma_p - \mathcal{K}_p)\right| \leq C\Delta_n^{1-\left(\frac{1}{2}\vee\frac{2p}{\beta}\right)-\iota},\quad \forall\iota > 0. \tag{6.82}$$

Combining the bounds in (6.80)-(6.82) and using the fact that $\beta' < \beta/2$ as well as $p < \beta/3$ (so that $\frac{p}{2} < 1 - \left(\frac{1}{2}\vee\frac{2p}{\beta}\right)$), we have

$$\mathbb{E}\left|\frac{1}{k_n^2(n-2k_n-1)}\sum_{i=2k_n+2}^{n}\frac{\Delta_n^{-2p/\beta}\sum_{k\in I_n^i}(\overline{\xi}_k^n)^2}{|\sigma_{(i-2k_n-2)\Delta_n}|^p} - (\Sigma_p - \mathcal{K}_p)\right| \leq C\left(\frac{\Delta_n^{\frac{p}{2}-\iota}\vee\alpha_n}{k_n}\vee\frac{1}{\sqrt{k_n n}}\right). \tag{6.83}$$

We continue with the second term on the right hand side of (6.79). We denote $\chi_i^n = \Delta_n^{-\frac{2p}{\beta}}\frac{\sum_{k,l\in I_n^i:|k-l|\geq 2}\overline{\xi}_k^n\overline{\xi}_l^n}{|\sigma_{(i-2k_n-2)\Delta_n}|^p}$. With this notation we can split

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

Using law of iterated expectations we have $\mathbb{E}(\chi_i^n)^2 \leq Ck_n^2$ and further

$$\mathbb{E}(\overline{\chi}_j^n)^2 \leq \sum_{i=1}^{\lfloor \frac{n-2k_n-1}{2k_n} \rfloor} \mathbb{E}(\chi_{(i-1)2k_n+j}^n)^2 \leq Ck_n n,$$

and therefore

$$\mathbb{E}\left| \sum_{i=2k_n+2}^{n} \chi_i^n \right| \leq Ck_n\sqrt{k_n n}, \tag{6.84}$$

which leads to

$$\mathbb{E}\left| \frac{1}{k_n^2(n-2k_n-1)} \sum_{i=2k_n+2}^{n} \frac{\Delta_n^{-2p/\beta} \sum_{k,l\in I_n^i:|k-l|<2} \overline{\xi}_k^n \overline{\xi}_l^n}{|\sigma_{(i-2k_n-2)\Delta_n}|^p} \right| \leq \frac{C}{\sqrt{k_n n}}. \tag{6.85}$$

Finally, using the boundedness of the third derivative of the function f (in x) and the fact that $p < \beta/3$, we have by applying the bounds in (6.62), (6.71) and (6.72):

$$\mathbb{E}|\widetilde{z}_i^{(3,c)}(\mathbf{u})| \leq C_{\mathbf{u}} \left[k_n^{-\frac{3}{2}} \vee k_n\Delta_n \vee \alpha_n^3 \right]. \tag{6.86}$$

Step 4. Combining the bounds in (6.69), (6.70) and (6.73), we get the result in (6.51). Combining the bounds in (6.69), (6.70) and (6.77), and making use of $k_n/n \to 0$, we get the result in (6.52). Finally, combining the bounds in (6.69), (6.70), (6.78), (6.83), (6.85) and (6.86), and since $k_n/\sqrt{n} \to 0$, we get the result in (6.53).

The results in (6.54)-(6.55). The result (6.54) follows from using similar steps as for the derivation of (6.80) as well as using Burkholder-Davis-Gundy inequality and (6.72). For the result in (6.55), we first make use of the algebraic inequality $||a| - |b|| \leq C|a-b|$ for some positive constant C and $\forall\, a, b \in \mathbb{R}$. We then use successive conditioning, (6.57)-(6.61) and (6.72), as well as Burkholder-Davis-Gundy inequality.

The result in (6.56). Due to (6.51), on a set of probability approaching one, $\widehat{\beta}(p) = \overline{\beta}(p) - \widehat{\mathcal{B}}_{p,\beta}^n$. First, we can apply first-order Taylor expansion of $\overline{\beta}(p)$ as a function of $\widehat{\mathcal{L}}_p^n(u_1, u_1)$ and $\widehat{\mathcal{L}}_p^n(u_2, u_2)$. Then we can use the bounds in (6.69), (6.70), (6.78), (6.83), (6.85) and (6.86) to derive a result similar to that in (6.53) for $\widehat{\mathcal{L}}_p^n(u_1, u_1)$ and $\widehat{\mathcal{L}}_p^n(u_2, u_2)$, and from here by using in addition (6.51)-(6.52), to get for $\forall \iota > 0$:

$$\overline{\beta}(p) - \beta - \mathcal{B}_{p,\beta}^n = O_p\left(\alpha_n^2 \vee \Delta_n^{\frac{\beta\wedge 1}{2} - \iota} \vee \frac{\alpha_n}{\sqrt{k_n}} \vee \frac{\Delta_n^{\frac{p}{2}-\iota}}{k_n} \vee k_n^{-\frac{3}{2}} \vee \frac{1}{k_n}\left(\frac{k_n}{n}\right)^{\frac{1}{2} - \frac{p}{\beta} - \iota} \right).$$

Using the bounds in (6.51)-(6.52) and (6.54)-(6.55), we also have for $\forall\iota > 0$:

$$\widehat{\mathcal{B}}^n_{p,\beta} - \mathcal{B}^n_{p,\beta} = O_p\left(\frac{\Delta_n^{\frac{p}{2}-\iota}}{k_n} \vee \frac{1}{\sqrt{k_n n}} \vee \frac{1}{k_n^{1+1\wedge\frac{\beta}{4p}-\iota}} \vee \frac{1}{k_n\sqrt{k_n}}\left(\frac{k_n}{n}\right)^{\frac{1}{2}-\frac{p}{\beta}-\iota}\right).$$

Combining the above two results we get (6.56). □

In the following lemma, we set Σ_p, μ_p and $K_{g,p}$ to 1 if $p \geq \beta$.

Lemma 3 *Assume A and SB hold for* $\mathbf{X}$ *with* $\beta' < \beta/2$. *Then, if* $\beta > q$, $p_n \to \infty$ *and* $p_n \leq 2k_n$, *we have*

$$\widehat{\beta}_1 - \beta = O_p(\Delta_n^{\varepsilon}), \tag{6.87}$$

for some $\varepsilon > 0$. *In addition for* $k_n \asymp n^{\varpi}$ *with* $\varpi \in \left[\frac{3}{8}, \frac{1}{2}\right)$, *and with* $p = \beta/4$ *in* α_n, *we have for* $\forall\iota > 0$*:*

$$\sup_{\mathbf{u}\in\mathbb{S}^1} |\widehat{\mathcal{L}}^n_{\widehat{p}}(u_1,u_2) - \mathcal{L}_{\widehat{p}}(u_1,u_2)| = O_p\left(\alpha_n \vee \frac{1}{\sqrt{k_n}}\right), \tag{6.88}$$

$$\begin{aligned}&\int_0^{\pi}\left|\frac{\widehat{\mathcal{L}}^n_{\widehat{p}}(\cos x,\sin x) - \mathcal{L}_{\widehat{p}}(\cos x,\sin x)}{\mathcal{L}_{\widehat{p}}(\cos x,\sin x)} - \mathcal{B}^n_{\widehat{p}}(\cos x,\sin x)\right|^2 dx\\ &= O_p\left[\frac{k_n}{n} \vee \alpha_n^2 \vee \frac{\Delta_n^{-\iota}}{\sqrt{k_n n}}\right],\end{aligned} \tag{6.89}$$

$$\begin{aligned}&\int_0^{\pi}\left|\frac{\log(\mathcal{L}_{\widehat{p}}(\cos x,\sin x))}{\log(\mathcal{L}_{\widehat{p}}(1,1))}\left(\frac{\widehat{\mathcal{L}}^n_{\widehat{p}}(1,1)}{\mathcal{L}_{\widehat{p}}(1,1)} - 1 - \mathcal{B}^n_{\widehat{p}}(1,1)\right) -\right.\\ &\left.\left(\frac{\widehat{\mathcal{L}}^n_{\widehat{p}}(\cos x,\sin x)}{\mathcal{L}_{\widehat{p}}(\cos x,\sin x)} - 1 - \mathcal{B}^n_{\widehat{p}}(\cos x,\sin x)\right)\right|dx = O_p\left[\Delta_n^{\frac{\beta\wedge 1}{2}-\iota} \vee \frac{\Delta_n^{\frac{1}{4}\left(\frac{\beta}{\beta'}-1\right)-\iota}}{\sqrt{k_n}}\right],\end{aligned} \tag{6.90}$$

$$\Delta_n^{1-\frac{2\widehat{p}}{\beta}}\widehat{\Sigma}^n_{\widehat{p}} - \Sigma_{\widehat{p}}\int_0^1 |\sigma_s|^{2\widehat{p}}ds = O_p\left(\Delta_n^{\frac{\beta}{8}-\iota}\right), \tag{6.91}$$

$$\Delta_n^{1-\frac{2\widehat{p}}{\beta}}\widehat{\mathcal{K}}^n_{\widehat{p}} - (2C_\beta)^{\frac{2\widehat{p}}{\beta}}\mu^2_{\widehat{p}}K^2_{g,\widehat{p}}\int_0^1 |\sigma_s|^{2\widehat{p}}ds = O_p\left(\Delta_n^{\frac{\beta}{8}-\iota}\right), \tag{6.92}$$

$$\widehat{\beta}(\widehat{p}) - \beta = O_p\left[\Delta_n^{\frac{\beta\wedge 1}{2}-\iota} \vee \frac{\Delta_n^{\frac{1}{4}\left(\frac{\beta}{\beta'}-1\right)-\iota}}{\sqrt{k_n}}\right]. \tag{6.93}$$

Proof of Lemma 3. We start with (6.87). We have for all $\iota > 0$:

$$\left.\begin{array}{l}\Delta_n^{-q/\beta}\widetilde{V}^n_1(q) - (2C_\beta)^{\frac{q}{\beta}}\mu_q K_{g,q}|\sigma_0|^q\\ \Delta_n^{-q/\beta}\widetilde{V}^n_2(q) - (4C_\beta)^{\frac{q}{\beta}}\mu_q K_{g,q}|\sigma_0|^q\end{array}\right\} = O_p\left(\Delta_n^{\frac{q}{2}} \bigvee \sqrt{\frac{p_n}{n}} \bigvee \left(\frac{1}{p_n}\right)^{1-\left(\frac{1}{2}\vee\frac{q}{\beta}\right)-\iota}\right).$$

This result follows by combining the bounds in (6.57)-(6.61), using the "smoothness in expectation" assumption for L (and its boundedness from below) as well as a Burkholder-Davis-Gundy inequality. From here the result in (6.87) follows by making a Taylor series expansion of $\widehat{\beta}_1$ as a function of $\widetilde{V}^n_1(q)$ and $\widetilde{V}^n_2(q)$.

We continue with (6.88)-(6.93). We introduce the set

$$\Omega_n = \left\{ \omega : |\widehat{\beta}_1^n - \beta| > \Delta_n^{\varepsilon/2} \right\}. \tag{6.94}$$

Given (6.87), we have $\mathbb{P}(\Omega_n) \to 1$. Therefore, it is no restriction to restrict attention to the set Ω_n, and we note that for n sufficiently high on this set we have $\left|\widehat{p} - \frac{\beta}{4}\right| < \left(\frac{1}{\beta} - \frac{\beta}{4}\right) \wedge \frac{\beta}{12}$. Then, we can apply the bounds for the terms $\widetilde{Z}^{(1)}(\mathbf{u})$, $\widetilde{Z}^{(2)}(\mathbf{u})$ and $\widetilde{Z}^{(3)}(\mathbf{u})$ in the proof of Lemma 3. The only exception is for the terms $\{\widetilde{z}_i^{(j)}(\mathbf{u})\}_{i=2k_n+2,\ldots,4k_n}$, and $j = 1, 2, 3$. The reason is that for them there is an overlap between the increments of $\widehat{V}(\widehat{p})_i^n$ and those used in the calculation of $\widehat{\beta}_1^n$. However, given the boundedness of the cosine and exponential functions, we easily have $\frac{1}{n-2k_n-1}\sum_{j=1}^3 \sum_{i=2k_n+2}^{4k_n} \sup_{\mathbf{u}\in\mathbb{R}^2} |z_i^{(j)}(\mathbf{u})| = O_p\left(\frac{k_n}{n}\right)$. This implies immediately the results in (6.88)-(6.90) and (6.93). For the results in (6.91)-(6.92), we can use the following inequality on the set Ω_n

$$\left|\Delta_n^{-1/\beta}(\Delta_i^n \mathbf{X}^{(j)} - \Delta_{i-1}^n \mathbf{X}^{(j)})\right|^{\widehat{p}} \leq \left|\Delta_n^{-1/\beta}(\Delta_i^n \mathbf{X}^{(j)} - \Delta_{i-1}^n \mathbf{X}^{(j)})\right|^{\underline{p}^n} + \left|\Delta_n^{-1/\beta}(\Delta_i^n \mathbf{X}^{(j)} - \Delta_{i-1}^n \mathbf{X}^{(j)})\right|^{\overline{p}^n}, \quad j = 1, 2,$$

where $\underline{p}^n = \frac{\beta}{4} - \frac{\Delta_n^{\varepsilon/2}}{4}$ and $\overline{p}^n = \frac{\beta}{4} + \frac{\Delta_n^{\varepsilon/2}}{4}$. We can apply this inequality for the first $2k_n$ terms in $\widehat{\Sigma}_p^n$ and $\widehat{\mathcal{K}}_p^n$ to bound their effect on the results in (6.54)-(6.55) to $O_p(k_n/n)$ while for the rest of the summands the bounds in Lemma 2 apply directly. □

6.3 Proof of Theorem 1 continued

Now we are ready to complete the proof of Theorem 1. We will analyze separately the terms $\sum_{m=1}^{m_n}\left(\frac{1}{\widehat{h}_{\widehat{p}}^*(m)} - \frac{1}{h^*(m)}\right)^2$, $\sum_{m=1}^{m_n} G^*(m)^2 h^*(m)^2 \left(\frac{1}{\widehat{h}_{\widehat{p}}^*(m)} - \frac{1}{h^*(m)}\right)^2$ and $\int_0^\pi |\widehat{a}^n(x) - a(x)|dx$. We start with the first two. Since $\widehat{\beta}(\widehat{p})$ is a consistent estimate of β, we need only to look on a set on which $|\widehat{\beta}(\widehat{p}) - \beta| < \iota$ for some arbitrary small $0 < \iota < (q \wedge (2-\beta))$. Using Taylor expansion we have for some positive integer l:

$$\widehat{h}_{\widehat{p}}^*(m) - h^*(m) = \sum_{k=1}^{l} \frac{(\widehat{\beta}(\widehat{p}) - \beta)^k}{\pi k!} \int_0^\pi |\cos x|^\beta \log^k |\cos x| e^{-2imx} dx + \frac{(\widehat{\beta}(\widehat{p}) - \beta)^{l+1}}{\pi (l+1)!} \int_0^\pi |\cos x|^{\widetilde{\beta}(x)} \log^{l+1} |\cos x| e^{-2imx} dx,$$

where $\widetilde{\beta}(x)$ is an intermediate value between $\widehat{\beta}(\widehat{p})$ and β that depends on x. From here, using the bounds of Lemma 1, we have for some arbitrary small $\iota \in (0, q \wedge (2-\beta))$ on the set on which $|\widehat{\beta}(\widehat{p}) - \beta| < \iota$ (whose probability approaches one):

$$\sum_{m=1}^{m_n}\left(\frac{1}{\widehat{h}_{\widehat{p}}^*(m)} - \frac{1}{h^*(m)}\right)^2 \leq C \sum_{k=1}^{l} |\widehat{\beta}(\widehat{p}) - \beta|^{2k} m_n^{3+2\beta+2\iota} \log^{2k}(m_n) + C|\widehat{\beta}(\widehat{p}) - \beta|^{2l+2} m_n^{3+4\beta+2\iota},$$

for some constant $C > 0$ that does not depend on m_n (but depends on l). Now, $m_n \asymp n^{\varrho}$ for a finite positive ϱ, and from Lemma 2 and the conditions of the theorem, $\widehat{\beta}(\widehat{p}) - \beta = O_p(\Delta_n^{\alpha})$ for some $\alpha > 0$. Therefore, if we pick l sufficiently high (so that $\alpha l > \beta\varrho$), the last term is dominated asymptotically by the first term on the right-hand-side of the above inequality (and its leading term is its first summand). As a result,

$$\sum_{m=1}^{m_n} \left(\frac{1}{\widehat{h}^*_{\widehat{p}}(m)} - \frac{1}{h^*(m)} \right)^2 \asymp (\widehat{\beta}(\widehat{p}) - \beta)^2 m_n^{3+2\beta+\iota}, \quad \forall \iota > 0. \tag{6.95}$$

Similarly, and making use also of Assumption A, we have

$$\sum_{m=1}^{m_n} G^*(m)^2 h^*(m)^2 \left(\frac{1}{\widehat{h}^*_{\widehat{p}}(m)} - \frac{1}{h^*(m)} \right)^2 \asymp (\widehat{\beta}(\widehat{p}) - \beta)^2 (m_n^{1-2\alpha+\iota} \vee 1), \quad \forall \iota > 0. \tag{6.96}$$

We continue with $\int_0^{\pi} |\widehat{a}^n(x) - a(x)| dx$. We have the decomposition

$$\widehat{a}^n(x) - a(x) = \sum_{j=1}^{5} a_j^n(x),$$

where $a_j^n(x) = \widetilde{a}_j^n(x) / \left(\log(\overline{\mathcal{L}}^n_{\widehat{p}}(1,1)) - \widehat{\mathcal{B}}^n_{\widehat{p}}(1,1) \right)$ and

$$\widetilde{a}_1^n(x) = \left(\frac{\mathcal{L}_{\widehat{p}}(\cos x, \sin x)}{\widetilde{\mathcal{L}}^n_{\widehat{p}}(\cos x, \sin x)} - 1 \right) \left(\frac{\overline{\mathcal{L}}^n_{\widehat{p}}(\cos x, \sin x)}{\mathcal{L}_{\widehat{p}}(\cos x, \sin x)} - 1 \right),$$

$$\widetilde{a}_2^n(x) = -\left(\widehat{\mathcal{B}}^n_{\widehat{p}}(\cos x, \sin x) - \mathcal{B}^n_{\widehat{p}}(\cos x, \sin x) \right),$$

$$\widetilde{a}_3^n(x) = -\frac{\log(\mathcal{L}_{\widehat{p}}(\cos x, \sin x))}{\log(\mathcal{L}_{\widehat{p}}(1,1))} \left(\frac{\mathcal{L}_{\widehat{p}}(1,1)}{\widetilde{\mathcal{L}}^n_{\widehat{p}}(1,1)} - 1 \right) \left(\frac{\overline{\mathcal{L}}^n_{\widehat{p}}(1,1)}{\mathcal{L}_{\widehat{p}}(1,1)} - 1 \right),$$

$$\widetilde{a}_4^n(x) = \frac{\log(\mathcal{L}_{\widehat{p}}(\cos x, \sin x))}{\log(\mathcal{L}_{\widehat{p}}(1,1))} \left(\widehat{\mathcal{B}}^n_{\widehat{p}}(1,1) - \mathcal{B}^n_{\widehat{p}}(1,1) \right),$$

$$\begin{aligned} \widetilde{a}_5^n(x) = & \left[\frac{\overline{\mathcal{L}}^n_{\widehat{p}}(\cos x, \sin x) - \mathcal{L}_{\widehat{p}}(\cos x, \sin x)}{\mathcal{L}_{\widehat{p}}(\cos x, \sin x)} - \mathcal{B}^n_{\widehat{p}}(\cos x, \sin x) \right] \\ & - \frac{\log(\mathcal{L}_{\widehat{p}}(\cos x, \sin x))}{\log(\mathcal{L}_{\widehat{p}}(1,1))} \left[\frac{\overline{\mathcal{L}}^n_{\widehat{p}}(1,1) - \mathcal{L}_{\widehat{p}}(1,1)}{\mathcal{L}_{\widehat{p}}(1,1)} - \mathcal{B}^n_{\widehat{p}}(1,1) \right], \end{aligned}$$

with $\widetilde{\mathcal{L}}^n_{\widehat{p}}(\cos x, \sin x)$ being an intermediate value between $\overline{\mathcal{L}}^n_{\widehat{p}}(\cos x, \sin x)$ and $\mathcal{L}_{\widehat{p}}(\cos x, \sin x)$, and similarly $\widetilde{\mathcal{L}}^n_{\widehat{p}}(1,1)$ being an intermediate value between $\overline{\mathcal{L}}^n_{\widehat{p}}(1,1)$ and $\mathcal{L}_{\widehat{p}}(1,1)$. We now bound $\int_0^{\pi} \widetilde{a}_j^n(x) dx$, for $j = 1, ..., 5$, using Lemma 3.

We have

$$\begin{aligned} & \int_0^{\pi} |\widetilde{a}_1^n(x)| dx \\ & \quad \leq \frac{C}{\inf_{x \in [0,\pi]} |\widetilde{\mathcal{L}}^n_{\widehat{p}}(\cos x, \sin x)|} \int_0^{\pi} (\overline{\mathcal{L}}^n_{\widehat{p}}(\cos x, \sin x) - \mathcal{L}_{\widehat{p}}(\cos x, \sin x))^2 dx. \end{aligned}$$

Therefore, applying (6.88) (as well as the fact that $\inf_{\mathbf{u}\in\mathbb{S}^1}\mathcal{L}_{\widehat{p}}(u_1,u_2))>0$) and (6.89) (note that $\sup_{x\in[0,\pi]}|\mathcal{B}^n_{\widehat{p}}(\cos x,\sin x)|=O_p(k_n^{-2})$ and an analogue of (6.89) holds for $\widehat{\mathcal{L}}^n_{\widehat{p}}(1,1)$ as well), we have

$$\int_0^\pi|\widetilde{a}_1^n(x)|dx+\int_0^\pi|\widetilde{a}_3^n(x)|dx=O_p\left(\frac{k_n}{n}\vee\alpha_n^2\vee\frac{\Delta_n^{-\iota}}{\sqrt{k_n n}}\right). \tag{6.97}$$

To bound $\int_0^\pi\widetilde{a}_2^n(x)dx$, we fist use Taylor series expansion to bound

$$\begin{aligned}|\widehat{\mathcal{B}}^n_{\widehat{p}}(\cos x,\sin x)-\mathcal{B}^n_{\widehat{p}}(\cos x,\sin x)|\le\frac{\widehat{C}^n(x)}{k_n}\Bigg\{&\left|\Delta_n^{1-\frac{2p}{\beta}}\widehat{\Sigma}^n_{\widehat{p}}-\Sigma_{\widehat{p}}\int_0^1|\sigma_s|^{2p}ds\right|\\&+\left|\Delta_n^{1-\frac{2p}{\beta}}\widehat{\mathcal{K}}^n_{\widehat{p}}-(2C_\beta)^{\frac{\widehat{p}}{\beta}}\mu_{\widehat{p}}^2K^2_{g,\widehat{p}}\int_0^1|\sigma_s|^{2p}ds\right|\\&+|\widehat{\beta}(\widehat{p})-\beta|+|\widehat{\mathcal{L}}_{\widehat{p}}(\cos x,\sin x)-\mathcal{L}_{\widehat{p}}(\cos x,\sin x)|\Bigg\},\end{aligned}$$

where $\widehat{C}^n(x)$ is some nonnegative random function of x for which, using the bounds in (6.88) and (6.91)-(6.93) as well as $\inf_{\mathbf{u}\in\mathbb{S}^1}\mathcal{L}_{\widehat{p}}(u_1,u_2))>0$, we have $\sup_{x\in[0,\pi]}\widehat{C}^n(x)=O_p(1)$. Using the above bound (and an analogous one for $|\widehat{\mathcal{B}}^n_{\widehat{p}}(1,1)-\mathcal{B}^n_{\widehat{p}}(1,1)|$) as well as (6.88) and (6.91)-(6.93), we have

$$\int_0^\pi|\widetilde{a}_2^n(x)|dx+\int_0^\pi|\widetilde{a}_4^n(x)|dx=O_p\left(\frac{\Delta_n^{\frac{\beta}{8}-\iota}}{k_n}\vee\frac{1}{k_n\sqrt{k_n}}\right). \tag{6.98}$$

We are left with $\int_0^\pi\widetilde{a}_5^n(x)dx$. Application of (6.90) yields

$$\int_0^\pi|\widetilde{a}_5^n(x)|dx=O_p\left(\Delta_n^{\frac{\beta\wedge1}{2}-\iota}\vee\frac{\Delta_n^{\frac{1}{4}\left(\frac{\beta}{\beta'}-1\right)-\iota}}{\sqrt{k_n}}\right),\quad\forall\iota>0. \tag{6.99}$$

The above bounds continue to hold for the corresponding $\int_0^\pi|a_j^n(x)|dx$ terms as due to the bounds in Lemma 3, we have $\log(\overline{\mathcal{L}}^n_{\widehat{p}}(1,1))-\widehat{\mathcal{B}}^n_{\widehat{p}}(1,1)=O_p(1)$. From here, using the restriction on ϖ, we have

$$\int_0^\pi\sum_{j=1}^5|a_j^n(x)|dx=O_p\left(\Delta_n^{\frac{\beta\wedge1}{2}-\iota}\vee\frac{\Delta_n^{\frac{1}{4}\left(\frac{\beta}{\beta'}-1\right)-\iota}}{\sqrt{k_n}}\right)\quad\forall\iota>0. \tag{6.100}$$

To bound the order of magnitude of $A_1^n(m)$, $A_2^n(m)$ and $A_3^n(m)$, we then use the order of magnitude of $\widehat{\beta}(\widehat{p})-\beta$ in (6.93), as well as (6.95) and (6.96), and we get altogether

$$\sum_{|m|\le m_n}A_1^n(m)=O_p\left(m_n^{3+2\beta}\left(\Delta_n^{\beta\wedge1-\iota}\vee\frac{\Delta_n^{\frac{1}{2}\left(\frac{\beta}{\beta'}-1\right)-\iota}}{k_n}\right)\right),\quad\forall\iota>0,$$

$$\sum_{|m|\le m_n}A_2^n(m)=O_p\left(m_n^{3+2\beta+\iota}\left(\Delta_n^{2(\beta\wedge1)-\iota}\vee\frac{\Delta_n^{\left(\frac{\beta}{\beta'}-1\right)-\iota}}{k_n^2}\right)\right),\quad\forall\iota>0,$$

$$\sum_{|m|\le m_n}A_3^n(m)=O_p\left((m_n^{1-2\alpha+\iota}\vee1)\left(\Delta_n^{\beta\wedge1-\iota}\vee\frac{\Delta_n^{\frac{1}{2}\left(\frac{\beta}{\beta'}-1\right)-\iota}}{k_n}\right)\right),\quad\forall\iota>0.$$

From here, using the bound on the integrated squared bias, the result of the theorem follows.

References

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.

Andrews, B., M. Calder, and R. Davis (2009). Maximum Likelihood Estimation for Alpha Stable Autoregressive Processes. *Annals of Statistics 37*, 1946–1982.

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

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

Bingham, N., C. Goldie, and J. Teugels (1987). *Regular Variation.* Cambridge University Press.

Bull, A. (2016). Near-Optimal Estimation of Jump Activity in Semimartingales. *Annals of Statistics 44*, 58–86.

Carr, P., H. Geman, D. Madan, and M. Yor (2002). The Fine Structure of Asset Returns: An Empirical Investigation. *Journal of Business 75*, 305–332.

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

Cheng, B. and S. Rachev (1995). Multivariate Stable Futures Prices. *Mathematical Finance 5*, 133–153.

Fan, J. (1991). On the Optimal Rates of Convergence for Nonparametric Deconvolution Problems. *Annals of Statistics 19*, 1257–1272.

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

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.

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

Johannes, J. (2009). Deconvolution with Unknown Error Distribution. *Annals of Statistics 37*, 2301–2323.

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.

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

Kong, X., Z. Liu, and B. Jing (2015). Testing for Pure-Jump Processes for High-Frequency Data. *Annals of Statistics 43*, 847–877.

McCulloch, J., J. Nolan, and A. Panorska (2001). Estimation of Stable Spectral Measures. *Mathematical and Computer Modelling 34*, 1113–1122.

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

Pivato, M. and L. Seco (2003). Estimating the Spectral Measure of a Multivariate Stable Distribution via Spherical Harmonic Analysis. *Journal of Multivariate Analysis 87*, 219–240.

Press, J. (1972). Estimation in Univariate and Multivariate Stable Distributions. *Journal of the American Statistical Association 67*, 842–846.

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

Samorodnitsky, G. and M. Taqqu (1994). *Stable Non-Gaussian Random Processes*. New York: Chapman & Hall.

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

Todorov, V. (2015). Jump Activity Estimation for Pure-Jump Semimartingales via Self-Normalized Statistics. *Annals of Statistics 43*, 1831–1864.

Todorov, V. (2017). Testing for Time-Varying Jump Activity for Pure Jump Semimartingales. *Annals of Statistics 45*, 1284–1311.

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

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

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

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