Nonparametric Jump Variation Measures from Options

Viktor Todorov[*]

April 27, 2021

Abstract

This paper proposes a novel nonparametric method for estimating tail jump variation measures from short-dated options, which can achieve rate-efficiency and works in a general infinite jump activity setting, avoiding parametric or semiparametric assumptions for the jump measure. The method is based on expressing the measures of interest as integrals of the Laplace transforms of the jump compensator and developing methods for recovering nonparametrically the latter from the available option data. The separation of volatility from jumps is done in a novel way by making use of the second derivative of the Laplace transform of the returns, de-biased using either the value of the Laplace transform or of its second derivative evaluated at high frequencies. A Monte Carlo study shows the superiority of the newly-developed method over existing ones in empirically realistic settings. In an empirical application to S&P 500 index options, we find risk-neutral negative market tail jump variation that is on average smaller than previous estimates of it, is generated by smaller-sized jumps, and has less dependence on the level of diffusive volatility.

Keywords: jump variation, Laplace transform, options, nonparametric estimation, tail risk.

JEL classification: C51, C52, G12.

[*]Department of Finance, Kellogg School of Management, Northwestern University; e-mail: v-todorov@kellogg.northwestern.edu.

1 Introduction

The availability of reliable data on deep out-of-the-money put options (i.e., puts with strikes that are significantly below the current price of the underlying asset) written on various assets, and in particular on market indices, allows for the nonparametric study of jump tail risk and its pricing. Bollerslev and Todorov (2011b, 2014) develop such techniques. Bollerslev et al. (2015) and Andersen et al. (2021), using the jump measures proposed in Bollerslev and Todorov (2011b, 2014), show that the pricing of downside jump tail risk varies significantly over time and is a good predictor of the equity risk premium.[1] In fact, these papers show that the predictive ability of the variance risk premium for future equity returns, documented in earlier work by Bollerslev et al. (2009) and Drechsler and Yaron (2011), is mostly due to the compensation for tail risk reflected in the option-based negative jump tail variation measures.

The existing option-based estimates of the jump variation measure, however, rely on strong semiparametric assumptions for the latter. The goal of the current paper, therefore, is to develop alternative fully nonparametric measures of jump variation from options which contain less approximation error than the existing ones and work under much weaker assumptions for the dynamics of the underlying asset than previously considered.

More specifically, the method of Bollerslev and Todorov (2011b, 2014) for measuring jump tail risk from options relies on several approximations. First, the increment of an Itô semimartingale over a short time interval is approximated with that of an Itô semimartingale whose characteristics are "frozen" at their values at the beginning of the interval. The approximate Itô semimartingale is thus conditionally a Lévy process, i.e., a process with i.i.d. increments. The second approximation used in the method of Bollerslev and Todorov (2011b, 2014) is to associate an increment of the process above a fixed threshold over a short time interval with the realization of a single jump exceeding that threshold.[2] These two approximations lead to the following asymptotic result for the prices of short-dated puts and calls:

$$\frac{P_{t,T}(k)}{TX_t} \xrightarrow{\mathbb{P}} \int_{\mathbb{R}} \left(e^{k-\log(X_t)} - e^x\right)^+ \nu_t(x)dx, \text{ as } T \downarrow 0 \text{ and for fixed } k < \log(X_t), \quad (1)$$

$$\frac{C_{t,T}(k)}{TX_t} \xrightarrow{\mathbb{P}} \int_{\mathbb{R}} \left(e^x - e^{k-\log(X_t)}\right)^+ \nu_t(x)dx, \text{ as } T \downarrow 0 \text{ and for fixed } k > \log(X_t), \quad (2)$$

[1] Related nonparametric option-based measures include the asymmetry measure of Du and Kapadia (2012) that is due to jumps, higher-order risk swaps of Schneider and Trojani (2015) and Orlowski et al. (2020) and option-based estimates of the Lévy densities in Lévy settings (Belomestny and Reiß (2006, 2015), Söhl (2014), Trabs (2014, 2015)) and in general Itô semimartingale settings (Qin and Todorov (2019)). These measures complement nonparametric tail measures extracted from high-frequency return data by Bollerslev and Todorov (2011a), from cross-section of returns by Kelly and Jiang (2014), from assets' bid-ask spreads by Weller (2019), and via a nonparametric bound on the stochastic discount factor by Almeida et al. (2017).

[2] Similar approximation is used in the construction of the truncated variance introduced by Mancini (2001, 2009), to separate jumps from diffusive volatility on the basis of high-frequency stock returns.

where $P_{t,T}(k)$ and $C_{t,T}(k)$ are the put and call price, respectively, at time t written on an asset whose value at that time is X_t, having a log-strike of k and expiring at $t+T$, and ν_t is the risk-neutral density of the jump compensator of the process $\log(X)$ at time t. The above approximation result is combined with a regular variation assumption (see e.g., Bingham et al. (1989)) for the tail integrals of $\frac{\nu_t(-\ln(y))}{y}$ and $\frac{\nu_t(\ln(y))}{y}$ as $y \to \infty$ by Bollerslev and Todorov (2011b, 2014). This leads to a simple (semiparametric) approximation of the tails of $\nu_t(x)$ by a function which is controlled by two parameters labeled as level and slope. These parameters can, in turn, be estimated from the prices of available deep out-of-the-money puts using the option approximations in (1)-(2) above.

While Bollerslev and Todorov (2011b, 2014) provide only consistency results for the above-described jump tail estimation in a setting with finite activity jumps, we can illustrate with an example the two approximation errors associated with the convergence result in (1)-(2). This is done for the parametric model that is used in our Monte Carlo experiment and the result is displayed in Figure 1. As seen from the figure, the approximation error due to "freezing" the semimartingale characteristics at their value at the beginning of the interval leads to a very small approximation error. In particular, the percentage error for the puts is very close to zero while that for the calls is somewhat bigger, which is probably due to the fact that the right tail of ν_t is much thinner than the left tail making the approximation error for the call bigger in relative terms. We also note that this type of Lévy approximation underlies most existing results in high-frequency financial econometrics, see e.g., Jacod and Protter (2012).

The second approximation in the method of Bollerslev and Todorov (2011b, 2014), which amounts to assuming that big absolute returns in the underlying asset are due to the occurrence of a single jump in the price, results in a much bigger approximation error as seen from Figure 1. Part of this approximation error is due to ignoring the contribution of the diffusive price component. Another is due to ignoring the possibility of multiple small jumps generating a large return over the option's horizon. This type of approximation error is going to be more severe for models with high jump activity, i.e., for models having higher probability of generating small-sized jumps. This fact can also explain, at least partially, why the approximation error for the calls is significantly larger than for the puts. Indeed, the positive jump compensator has a significantly thinner tail than its negative counterpart (the model jump specification used in the experiment is calibrated to match real index option data). This means that a higher percentage of the positive jump variation is generated by smaller-sized (positive) jumps and therefore for such jump specification the approximation based on the result in (2) regarding out-of-the-money calls is of limited practical value.

In this paper, we propose an alternative method for estimating jump variation measures, i.e., quantities of the type $\int_{x<-\vartheta} x^2\nu_t(x)dx$ and $\int_{x>\vartheta} x^2\nu_t(x)dx$ for some $\vartheta \geq 0$, from options which does not rely on the approximation in (1)-(2) and instead uses only Lévy approximation for the

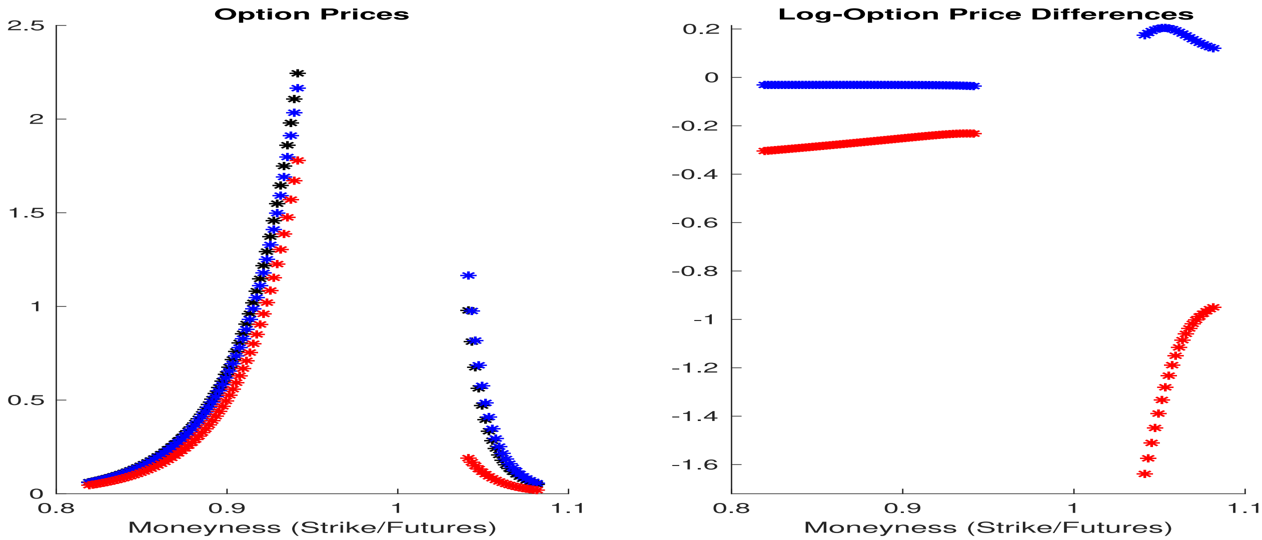

Figure 1: Approximations of Short-Dated Options. Left panel displays true option prices (in black), option prices based on a Lévy approximation (in blue) and option prices based on the distribution of one jump, given in equation (1) (in red). Right panel displays the differences in terms of log option prices between true model and the Lévy-based approximation (in blue) and the true model and approximation assuming one jump till expiration (in red). The options have 5 days till expiration and the model used for generating them is given in (53)-(54) with $\lambda_- = 20$ and $\lambda_+ = 50$.

underlying asset dynamics and smoothness of $\nu_t(x)$.[3] Unlike the method of Bollerslev and Todorov (2011b, 2014), which uses only deep out-of-the-money options in order to minimize the impact from the diffusion on the option prices, the method developed here takes advantage of the information regarding the jump variation measures contained in the options across all available strikes.

The newly-proposed method is based on several results. First, one can express the jump variation measures as integrals of the complex-valued Laplace transform of $x^2\nu_t(x)$, see e.g., Kawata (2014) and Kolassa (2006). Second, the complex-valued Laplace transform of $x^2\nu_t(x)$ can be recovered from $\nabla_{uu}\log(\mathbb{E}_t^{\mathbb{Q}}(e^{u\log(X_{t+T}/X_t)}))$ for complex u up to the unknown diffusive coefficient at time t and an (asymptotically negligible) approximation error due to the expected time variation in the semimartingale characteristics over the interval $[t, t+T]$. This is done using the Lévy-Khinchine formula for the moment generating function of a Lévy process (Theorem 25.17 in Sato (1999)). Third, the value of the unknown diffusion coefficient at t can be approximated using either $\nabla_{uu}\log(\mathbb{E}_t^{\mathbb{Q}}(e^{u\log(X_{t+T}/X_t)}))$ or $2\log(|\mathbb{E}_t^{\mathbb{Q}}(e^{u\log(X_{t+T}/X_t)})|)/u^2$, for some u with large (asymptotically increasing) imaginary part. Finally, $\mathbb{E}_t^{\mathbb{Q}}(e^{u\log(X_{t+T}/X_t)})$ can be recovered from the available

[3]Quantities like $\int_{x<-\vartheta} x^2\nu_t(x)dx$ and $\int_{x>\vartheta} x^2\nu_t(x)dx$ are of natural interest in applications as already discussed in the first paragraph of the introduction. For example, these quantities with $\vartheta = 0$ allow to study the dynamics of the risk-neutral jump intensity of positive and negative jumps, the behavior of which is of major interest in reduced-form and equilibrium asset pricing models, see e.g., Drechsler and Yaron (2011) and Wachter (2013), among others. Further, the difference $\int_{x<0} x^2\nu_t(x)dx - \int_{x>0} x^2\nu_t(x)dx$ is the spot risk-neutral semivariance which has also received a lot of attention in recent applied work, see e.g., Kilic and Shaliastovich (2019).

short-dated options using the option-spanning result in Carr and Madan (2001).

We derive the convergence of the above estimator under general conditions for the dynamics of the underlying process. In particular, unlike Bollerslev and Todorov (2011b, 2014), we allow for infinite activity jumps (but of finite variation) that can generate a lot of small jumps in the path of the underlying asset. The estimation error in recovering the jump variation measures depends on the smoothness of $\nu_t(x)$, the option observation error, the mesh of the strike grid and the time to maturity of the available options. The proposed jump variation estimator can achieve the optimal rate of convergence in a minimax sense which is faster than that of the option-based estimate of the Lévy density of Qin and Todorov (2019). This is not surprising as the jump variation measure is an integral transform of the Lévy density. Importantly, our approach for separating volatility from jumps is different from the one in Qin and Todorov (2019) and allows for more precise inference regarding the Lévy density of moderately-sized and small jumps. In particular, Qin and Todorov (2019) use the dominant role of jumps in the third derivative of $\log(\mathbb{E}_t^{\mathbb{Q}}(e^{u\log(X_{t+T}/X_t)}))$ while here we make use only of the second derivative of the latter and/or its level to achieve the separation.[4] This means that the current approach puts more weight (in relative terms) to near-the-money options, which are more liquid, than that of Qin and Todorov (2019). This also means that the current approach achieves higher precision for the behavior of the Lévy measure near the origin that that of Qin and Todorov (2019).

A Monte Carlo study shows significant bias reduction when using the new method for the estimation of the jump tail variation measures, particularly in settings with many small jumps and in periods of high volatility when the approximations in (1)-(2) do not work very well. Consistent with our theoretical results and the Monte Carlo evidence, in the empirical implementation of the newly-proposed techniques to S&P 500 index option data, we uncover left jump tail variation measures that are smaller on average than currently reported. These differences are bigger in high volatility periods and result in less correlation between the extracted left jump variation and the diffusive volatility. In addition, by comparing the tail jump variation measures with estimates of the total left jump variation, constructed using our method, we find that a significant part of the jump variation is generated by jumps of relatively small size.

The rest of the paper is organized as follows. In Section 2 we present our setup and introduce the option observation scheme. In Section 3 we develop the estimator and derive its rate of convergence and in Section 4 we describe the choice of the tuning parameters needed for its implementation.

[4] In their supplementary appendix, Qin and Todorov (2019) consider also inference procedure for the Lévy density based on the second derivative of $\log(\mathbb{E}_t^{\mathbb{Q}}(e^{u\log(X_{t+T}/X_t)}))$ with debiasing via a generic estimator of the diffusion coefficient. In its general form, as pointed out in Qin and Todorov (2019), this procedure does not lead to optimal rate of convergence because of the role of the near-the-money options in the inference. By contrast, here we use specifically designed debiasing methods that can suitably dampen the role of the near-the-money options in the estimator of the second derivative of $\log(\mathbb{E}_t^{\mathbb{Q}}(e^{u\log(X_{t+T}/X_t)}))$ and this leads to optimal rate of convergence of the resulting jump variation estimators.

Section 5 contains a Monte Carlo study and Section 6 an empirical application. Section 7 concludes. The proofs of the theoretical results are given in Section 8.

2 Setup and Option Observation Scheme

We start with introducing the notation and the option observation scheme. The log-price of the underlying asset, denoted with $x_t = \log(X_t)$, is defined on the filtered probability space $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\big)$. We assume the existence of a risk-neutral probability measure, denoted with $\mathbb{Q}$, under which the discounted at the risk-free rate prices become local martingales. Under regularity type conditions, see e.g., Duffie (2001), no-arbitrage implies the local equivalence of $\mathbb{P}$ and $\mathbb{Q}$, i.e., the equivalence of $\mathbb{P}$ and $\mathbb{Q}$ in restriction to $\mathcal{F}_t$, for any $t \geq 0$. We assume that x_t is an Itô semimartingale under $\mathbb{Q}$ with the following dynamics

$$x_t = \int_0^t a_s ds + \int_0^t \sigma_s dW_s + \int_0^t \int_{\mathbb{R}} x\widetilde{\mu}(ds, dx), \tag{3}$$

where W is a Brownian motion, μ is an integer-valued random measure on $\mathbb{R}_+ \times \mathbb{R}$, counting the jumps in x, with compensator $\nu_t(x)dt \otimes dx$, for some predictable function $\nu_t(x)$ satisfying $\int_{\mathbb{R}}(x^2 \wedge 1)\nu_t(x)dx < \infty$, and $\widetilde{\mu}(dt, dx) = \mu(dt, dx) - \nu_t(x)dt \otimes dx$ is the martingale measure associated with μ (W and ν_t are defined with respect to $\mathbb{Q}$).[5]

Our interest in this paper is measuring the tail jump variation measures defined as

$$RV_t(\vartheta) = \int_{x>\vartheta} x^2 \nu_t(x)dx \quad \text{and} \quad LV_t(\vartheta) = \int_{x<-\vartheta} x^2 \nu_t(x)dx, \quad \text{for some } \vartheta \geq 0. \tag{4}$$

The estimation will be based on European-style options written on the value of the asset X in the future, recorded at a fixed time point t and with the same time to maturity T. For each strike, we rely on the so-called out-of-the-money (OTM) option price–the cheaper of the call and the put for the given strike–which would be worth zero, if the option were to expire today. We denote OTM option price observed at time t by

$$O_{t,T}(k) = \min\{P_{t,T}(k), C_{t,T}(k)\} = \begin{cases} P_{t,T}(k) = e^{-rT}\mathbb{E}_t^{\mathbb{Q}}\left(e^k - X_{t+T}\right)^+ & \text{if } k \leq \log(F_{t,T}), \\ C_{t,T}(k) = e^{-rT}\mathbb{E}_t^{\mathbb{Q}}\left(X_{t+T} - e^k\right)^+ & \text{if } k > \log(F_{t,T}), \end{cases} \tag{5}$$

where $F_{t,T}$ is the time-t futures price of the asset with expiration date $t + T$ and r is the risk-free interest rate, which for simplicity of notation is assumed to be constant. The second part of the above equation follows from the definition of the option contracts and the risk-neutral probability measure.

[5] The jump compensator of μ is defined uniquely as the predictable random measure $\mu^{\mathbb{Q}}$ such that integrals of predictable functions with respect to $\mu - \mu^{\mathbb{Q}}$ are $\mathbb{Q}$-local martingales, see Theorem II.1.8 in Jacod and Shiryaev (2003). In defining the jump compensator, we have implicitly assumed that it is absolutely continuous in time, which is the case for Itô semimartingales.

The underlying asset price and true option prices are all defined on $\Omega^{(0)}$, with the associated σ-algebra $\mathcal{F}^{(0)}$ and filtration $(\mathcal{F}_t^{(0)})_{t\geq 0}$. The statistical (true) probability measure is denoted with $\mathbb{P}^{(0)}$.

Our data consists of options observed on the following log-strike grid,

$$\underline{k} \;\equiv\; k_1 \;<\; k_2 \;<\; \cdots k_N \;\equiv\; \overline{k}, \tag{6}$$

and we denote

$$\underline{K} = e^{\underline{k}} \text{ and } \overline{K} = e^{\overline{k}}. \tag{7}$$

Since in the estimation the option observation time t will be fixed, in order to keep notation simple, we do not add "t" in the notation of the strike grid. The gap between the log-strikes is denoted $\Delta_i = k_i - k_{i-1}$, for $i = 2, ..., N$. The log-strike grid need not be equidistant, i.e., Δ_i may differ across i's. The asymptotic theory developed below is of infill type, i.e., the mesh of the log-strike grid, $\sup_{i=2,...,N} \Delta_i$, shrinks towards zero.

Finally, we allow for observation error, i.e., instead of observing $O_{t,T}(k_j)$ directly, we observe,

$$\widehat{O}_{t,T}(k_j) = O_{t,T}(k_j) + \epsilon_{t,T}(k_j), \quad j = 1, ..., N, \tag{8}$$

where $\{\epsilon_{t,T}(k_j)\}_{j=1}^N$ is a sequence of observation errors. The probability space has the following product form

$$\Omega \;=\; \Omega^{(0)} \times \Omega^{(1)}, \quad \mathcal{F} \;=\; \mathcal{F}^{(0)} \times \mathcal{F}^{(1)}, \quad \mathbb{P}(d\omega^{(0)}, d\omega^{(1)}) \;=\; \mathbb{P}^{(0)}(d\omega^{(0)})\, \mathbb{P}^{(1)}(\omega^{(0)}, d\omega^{(1)}),$$

where the underlying price process and the true option prices are defined on $\Omega^{(0)}$ and the observation errors are defined on the space $\Omega^{(1)} = \mathbb{R}^{\mathbb{R}}$, which is equipped with the product Borel σ-field $\mathcal{F}^{(1)}$.

Throughout the paper when expectations are under the true probability measure $\mathbb{P}$, we will denote them with $\mathbb{E}$, i.e., without using superscript $\mathbb{P}$ in their notation. Henceforth, we will denote with $\Re(a)$ and $\Im(a)$, the real and imaginary part, respectively, of a complex number a. We will also use the notation $a \wedge b = \min\{a, b\}$ and $a \vee b = \max\{a, b\}$, for two real numbers a and b.

3 Jump Variation Estimators from Options

We now introduce our estimators of the jump variation measures, state the assumptions needed for characterizing their asymptotic behavior, and derive bounds for the asymptotic order of the estimation error in them.

3.1 Derivation of the Estimators

Our strategy for estimating $LV_t(\vartheta)$ and $RV_t(\vartheta)$ is to invert them from the Laplace transform of the measure $x^2\nu_t(x)$. More specifically, we fix $c > 0$, then for some $\vartheta \in [0, \infty)$ using Fubini's theorem, see e.g., Section 5 in Kolassa (2006) (and Theorem 7.4.3 in Kawata (2014)), we have[6]

$$RV_t(\vartheta) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{e^{-\vartheta(c+iz)}}{c+iz}\mathcal{F}\nu_t(c+iz)dz, \tag{9}$$

$$LV_t(\vartheta) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{e^{-\vartheta(c+iz)}}{c+iz}\mathcal{F}\nu_t(-c-iz)dz, \tag{10}$$

provided $\int_{\mathbb{R}} e^{cx}\nu_t(x)dx < \infty$ for $RV_t(\vartheta)$ and $\int_{\mathbb{R}} e^{-cx}\nu_t(x)dx < \infty$ for $LV_t(\vartheta)$ (these two conditions are implied by assumption A3 below provided $c < 3$), where we define the complex-valued Laplace transform of $x^2\nu_t(x)$ as

$$\mathcal{F}\nu_t(u) = \int_{\mathbb{R}} x^2 e^{ux}\nu_t(x)dx,\ u \in \mathbb{C}. \tag{11}$$

To get an estimate of $\mathcal{F}\nu_t(u)$, we will first recover from the option data the Laplace transform of the price increment $x_{t+T} - x_t$. Using the "spanning" result of Carr and Madan (2001), we have

$$\mathcal{L}_{t,T}(u) = \mathbb{E}_t^{\mathbb{Q}}\left(e^{u(x_{t+T}-x_t)}\right) = 1 + e^{-x_t}\int_{\mathbb{R}} f(u, k - x_t)O_{t,T}(k)dk,\ u \in \mathbb{C}, \tag{12}$$

where we denote

$$f(u,k) = \left(u^2 - u\right)e^{(u-1)k},\ u \in \mathbb{C},\ k \in \mathbb{R}, \tag{13}$$

provided the dividend yield of the underlying asset and the risk-free interest rate are both zero, an assumption that we will maintain here for simplicity given the fact that our asymptotics is for $T \downarrow 0$. Using this result, our estimate for the conditional Laplace transform of the price increment from the available options is given by

$$\widehat{\mathcal{L}}_{t,T}(u) = 1 + e^{-x_t}\sum_{j=2}^{N} f(u, k_{j-1} - x_t)\widehat{O}_{t,T}(k_{j-1})\Delta_j,\ u \in \mathbb{C}. \tag{14}$$

Next, to estimate $\mathcal{F}\nu_t(u)$ from $\widehat{\mathcal{L}}_{t,T}(u)$, we make use of the following approximation for T small

$$\widehat{\mathcal{L}}_{t,T}(u) \approx \exp\left(Tu\alpha_t + T\frac{u^2}{2}\sigma_t^2 + T\int_{\mathbb{R}}\left(e^{ux} - 1 - ux\right)\nu_t(x)dx\right), \tag{15}$$

and this approximation result is made formal in the proof. It follows from approximating $x_{t+T} - x_t$ with the increment of a $\mathcal{F}_t$-conditional Lévy process with spot characteristics equal to those of x "frozen" at their value at time t (the beginning of the interval) and making use of the expression for the moment generating function of a Lévy process (Theorem 25.17(iii) in Sato (1999)).

The approximation in (15) suggests that

$$\widehat{h}_{t,T}(u) = \left(\frac{1}{T}\frac{\nabla_{uu}\widehat{\mathcal{L}}_{t,T}(u)}{\widehat{\mathcal{L}}_{t,T}(u)} - \frac{1}{T}\left(\frac{\nabla_u\widehat{\mathcal{L}}_{t,T}(u)}{\widehat{\mathcal{L}}_{t,T}(u)}\right)^2\right)1\left(|\widehat{\mathcal{L}}_{t,T}(u)| \neq 0\right), \tag{16}$$

[6] Alternatively, we can invert $LV_t(\vartheta)$ using the formula $LV_t(\vartheta) = \frac{1}{2} - \frac{1}{\pi}\int_0^{\infty}\frac{1}{z}\Im\left(\mathcal{F}\nu_t(iz)e^{-iz\vartheta}\right)dz$, see e.g., Gil-Pelaez (1951) and more recently Shephard (1991). This inversion formula involves, however, the use of improper integration near the origin, and for this reason we proceed with inversion based on Laplace transform evaluated at complex numbers with strictly positive real part.

is an estimator of

$$h_t(u) = \sigma_t^2 + \mathcal{F}\nu_t(u). \tag{17}$$

To proceed further, we need a jump-robust estimator of the spot diffusive variance σ_t^2. A natural candidate is the option-based volatility estimator of Todorov (2019), which is defined as follows

$$\widehat{\sigma}_{t,T}^2(\overline{u}) = -\frac{2}{\overline{u}^2}\log\left(|\widehat{\mathcal{L}}_{t,T}(i\overline{u})|\right), \tag{18}$$

for some sequence $\overline{u} \to \infty$. As shown in Todorov (2019), to achieve optimal rate of convergence for the estimation of diffusive volatility, we need to set $\overline{u} \asymp 1/\sqrt{T}$. It turns out, however, that this choice will not yield in general a rate-efficient estimate of the tail variation measure. The reason for this is that its use will not lead to suitable cancelation in the weights assigned to options with strikes near the origin. The order of magnitude of the latter is bigger than that of the rest of the options, and the same applies to the observation errors attached to them. Therefore, if the role of the near-the-money options is not diminished in the estimation in a suitable way, this will result in a slower rate of convergence than what is feasible. For this reason, here we will use $\widehat{\sigma}_{t,T}^2(\overline{u})$ with an arbitrary sequence $\overline{u}$, and as seen later, its choice will have a nontrivial effect on the asymptotic properties of the resulting jump tail variation estimator:

$$\widehat{LV}_{t,T}(\vartheta, \overline{z}, \overline{u}) = \frac{1}{2\pi}\int_{-\overline{z}}^{+\overline{z}} \frac{e^{-\vartheta(c+iz)}}{c+iz}\left(\widehat{h}_{t,T}(-c-iz) - \widehat{\sigma}_{t,T}^2(\overline{u})\right)dz. \tag{19}$$

As we will show later, the above estimator can achieve the optimal rate of convergence for recovering the tail jump variation measure. However, for its use in practice, in addition to $\overline{z}$, we need to set optimally $\overline{u}$, which is not easy as this parameter needs to grow asymptotically at a certain rate that can depend on unknown features of the jump measure. For this reason, we consider also an alternative de-biasing procedure here based solely on the second derivative estimator $\widehat{h}_{t,T}(u)$. By the Riemann–Lebesgue lemma, $\mathcal{F}\nu_t(u) \to 0$ as $\Im(u) \to \pm\infty$. Therefore, an alternative estimator of σ_t^2 from $\widehat{h}_{t,T}(u)$ is given by $\frac{1}{2(\overline{\kappa}-\kappa)\overline{z}}\int_{|z|\in[\kappa\overline{z},\overline{\kappa}\overline{z}]}\widehat{h}_{t,T}(-c-iz)dz$, for some large $\overline{z}$ and some $\overline{\kappa} > \kappa \geq 1$. Combining these results, our alternative estimator of the negative jump variation measure becomes:

$$\begin{aligned}\widehat{LV}_{t,T}(\vartheta, \overline{z}, \kappa, \overline{\kappa}) &= \frac{1}{2\pi}\int_{-\overline{z}}^{+\overline{z}} \frac{e^{-\vartheta(c+iz)}}{c+iz}\widehat{h}_{t,T}(-c-iz)dz \\ &\quad - \frac{1}{4\pi(\overline{\kappa}-\kappa)\overline{z}}\int_{-\overline{z}}^{+\overline{z}} \frac{e^{-\vartheta(c+iz)}}{c+iz}dz\int_{|z|\in[\kappa\overline{z},\overline{\kappa}\overline{z}]}\widehat{h}_{t,T}(-c-iz)dz,\end{aligned} \tag{20}$$

for some $\overline{z} > 0$ and $\overline{\kappa} > \kappa \geq 1$.

We can similarly define estimators for the positive jump variation. The results that follow will also apply to such estimators as well. For brevity and given the interest in the negative market jumps, henceforth, we state only results for $\widehat{LV}_{t,T}(\vartheta, \overline{z}, \overline{u})$ and $\widehat{LV}_{t,T}(\vartheta, \overline{z}, \kappa, \overline{\kappa})$.

We can compare our estimators $\widehat{LV}_{t,T}(\vartheta, \overline{z}, \overline{u})$ and $\widehat{LV}_{t,T}(\vartheta, \overline{z}, \kappa, \overline{\kappa})$ with the nonparametric one for the Lévy density proposed by Qin and Todorov (2019). There are two fundamental differences between the current estimators and the one of Qin and Todorov (2019). First, the interest here is estimating integrals of the form $\int_{x<-\vartheta} x^2 \nu_t(x)dx$ while Qin and Todorov (2019) is about estimating the Lévy density $\nu_t(x)$. For the recovery of the former we need less smoothness of $\nu_t(x)$, i.e., slower tail decay of $\mathcal{F}\nu_t(u)$. That is, the rate of convergence of our estimator will be faster for the same smoothness of $\nu_t(x)$ than that of the Lévy density estimator of Qin and Todorov (2019). Intuitively, while the recovery of $x^2\nu_t(x)$ can be done via standard Fourier inversion from $\mathcal{F}\nu_t(u)$, in the integrands in (9)-(10), we have in addition division by $c + iz$, which helps reduce the bias from truncating the higher frequencies in the estimation.

Second, for separating volatility from jumps, Qin and Todorov (2019) consider the third derivative of $\log(\widehat{\mathcal{L}}_{t,T}(u))$ and estimate $x^3\nu_t(x)$. In this paper, on the other hand, we work with the second derivative of $\log(\widehat{\mathcal{L}}_{t,T}(u))$ and debias it via its value for arguments with the highest absolute value of their imaginary part or using directly the characteristic function of the returns for asymptotically increasing values of the exponent.[7] This method of debiasing is done in order to account for the diffusion in the return dynamics. An advantage of the current approach is that it puts more weight on options with closer to the money strikes which tend to be more liquid. Indeed, recovery of risk-neutral return moments above two is known to be significantly more difficult than the recovery of the risk-neutral variance. Moreover, the higher role of near-the-money options in the current method relative to Qin and Todorov (2019) does not come at a cost of slower rate of convergence even in the presence of higher activity jumps in X.

3.2 Assumptions

To derive the asymptotic behavior of our tail variation estimators, we will need assumptions for the smoothness of $\nu_t(x)$, the risk-neutral dynamics of x, the option sampling scheme and the option observation error. A smoothness assumption for $x^2\nu_t(x)$ translates into a rate decay of $\mathcal{F}\nu_t(u)$ as $\Im(u) \to \pm\infty$. Such an assumption is necessary because in $\widehat{LV}_{t,T}(\vartheta, \overline{z}, \overline{u})$ and $\widehat{LV}_{t,T}(\vartheta, \overline{z}, \kappa, \overline{\kappa})$, we integrate over the finite interval $[-\overline{z}, \overline{z}]$ instead of $(-\infty, \infty)$ used in $LV_t(\vartheta)$. The associated approximation thus relies on the tail decay of $\mathcal{F}\nu_t(u)$ for large in absolute value $\Im(u)$. Our assumption for this tail decay is given in the following:

A1. *For some constant $c > 0$, the function $x^2 e^{-cx}\nu_t(x)$ belongs to the class*

$$\mathcal{S}_r(C_t) = \left\{ f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}) : \int_{\mathbb{R}} |f^*(u)|^2 (1+u^2)^r du \leq C_t \right\}, \tag{21}$$

[7] As already mentioned in footnote 4, in the supplementary appendix of Qin and Todorov (2019) estimation based on the second derivative of $\log(\widehat{\mathcal{L}}_{t,T}(u))$ is also considered. This estimator does not achieve the optimal rate of convergence because Qin and Todorov (2019) use a generic estimate of the spot diffusive volatility for debiasing, which in general does not reduce the role of near-the-money options in the second derivative of $\log(\widehat{\mathcal{L}}_{t,T}(u))$.

for some positive constant r and some positive and $\mathcal{F}_t$-adapted random variable C_t, and where $f^*(u) = \int_{\mathbb{R}} e^{iux} f(x)dx$.

In the above assumption, the function $x^2 e^{-cx}\nu_t(x)$ is assumed to belong to a Sobolev class. A slightly stronger assumption will be to assume that $f^*(u)$ has a polynomial tail decay. A distribution with characteristic function satisfying such a tail decay is referred to as ordinary smooth, see e.g., Fan (1991).[8] An alternative stronger assumption is of super smooth distribution, like the normal distribution, for which the tail decay of its Fourier transform is exponential. We note that in the case of infinite activity jumps, the explosion of $\nu_t(x)$ around zero affects the smoothness of $x^2 e^{-cx}\nu_t(x)$ at the origin and hence the tail decay of the Fourier transform of $x^2 e^{-cx}\nu_t(x)$. Explosion of $\nu_t(x)$ near the origin will typically imply power law tail decay for the Fourier transform of $x^2 e^{-cx}\nu_t(x)$.

As an example, we can consider the tempered stable process with Lévy density of the form

$$\nu^{\text{ts}}(x) = c_- \frac{e^{-\lambda_-|x|}}{|x|^{\beta+1}} 1_{\{x<0\}} + c_+ \frac{e^{-\lambda_+|x|}}{|x|^{\beta+1}} 1_{\{x>0\}}, \; c_\pm \geq 0, \; \lambda_\pm \geq 0, \; \beta < 2. \tag{22}$$

The parameter β in this model for the jump compensator determines the jump activity, with $\beta < 0$ corresponding to the case of finite activity jumps and $\beta \geq 0$ to the case of infinite activity jumps. For this parametric model, it can be shown that the Fourier transform of $x^2 e^{-cx}\nu^{\text{ts}}(x)$ decays at the rate $u^{\beta-2}$ (for u being the argument of the Fourier transform), see e.g., Proposition 4.2 in Cont and Tankov (2003). This means that assumption A1 is satisfied with $r < 2 - \beta - \frac{1}{2}$. Therefore, the higher jump activity implies less smoothness of $x^2\nu^{\text{ts}}(x)$ and hence slower associated tail decay of its Fourier transform.

Apart from assumption A1 above, we will also need existence of certain conditional moments as well as assumptions about the dynamics of x and the option observation error. These assumptions are taken from Qin and Todorov (2019) with the exception of A6 which is slightly modified.

A2. *The process σ has the following dynamics under $\mathbb{Q}$:*

$$\sigma_t = \sigma_0 + \int_0^t b_s ds + \int_0^t \eta_s dW_s + \int_0^t \widetilde{\eta}_s d\widetilde{W}_s + \int_0^t \int_{\mathbb{R}} \delta^\sigma(s,u)\mu^\sigma(ds,du), \tag{23}$$

where $\widetilde{W}$ is a Brownian motion independent of W; μ^σ is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}$ with compensator $\nu^\sigma(ds,du) = ds \otimes du$, having arbitrary dependence with the random measure μ; b, η and $\widetilde{\eta}$ are processes with càdlàg paths and $\delta^\sigma(s,u) : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ is left-continuous in its first argument.

A3. *With the notation of A2, there exists an $\mathcal{F}_t$-adapted random variable $\bar{t} > 0$ such that for $s \in [t, \bar{t}]$:*

$$\mathbb{E}_t^{\mathbb{Q}}|a_s|^4 + \mathbb{E}_t^{\mathbb{Q}}|\sigma_s|^6 + \mathbb{E}_t^{\mathbb{Q}}(e^{4|x_s|}) + \mathbb{E}_t^{\mathbb{Q}}\left(\int_{\mathbb{R}} (e^{3|z|} - 1)\nu_s(z)dz\right)^4 < C_t, \tag{24}$$

[8] We note that up to a finite-valued $\mathcal{F}_{t-}$-adapted random variable, $x^2\nu_t(x)$ is the density of a probability distribution.

for some $\mathcal{F}_t$-adapted random variable C_t, and in addition for some $\iota > 0$

$$\mathbb{E}_t^{\mathbb{Q}}\left(\int_{\mathbb{R}}(|\delta^\sigma(s,z)|^4 \vee |\delta^\sigma(s,z)|)dz\right)^{1+\iota} \leq C_t. \tag{25}$$

A4. *With the notation of A2, there exists an $\mathcal{F}_t$-adapted random variable $\bar{t} > 0$ such that for $s, u \in [t, \bar{t}]$:*

$$\begin{aligned}\mathbb{E}_t^{\mathbb{Q}}|a_s - a_u|^p + \mathbb{E}_t^{\mathbb{Q}}|\sigma_s - \sigma_u|^p + \mathbb{E}_t^{\mathbb{Q}}|\eta_s - \eta_u|^p + \mathbb{E}_t^{\mathbb{Q}}|\widetilde{\eta}_s - \widetilde{\eta}_u|^p \\ \leq C_t|s-u|, \ \forall p \in [2,4],\end{aligned} \tag{26}$$

and

$$\mathbb{E}_t^{\mathbb{Q}}\left(\int_{\mathbb{R}}(e^{z\vee 0}|z| \vee |z|^2)|\nu_s(z) - \nu_u(z)|dz\right)^p \leq C_t|s-u|, \ \forall p \in [2,3], \tag{27}$$

for some $\mathcal{F}_t$-adapted random variable C_t.

A5. *The log-strike grid $\{k_i\}_{i=1}^N$ is $\mathcal{F}_t^{(0)}$-adapted and on a set with probability approaching one, we have*

$$\eta\Delta \leq \inf_{i=2,\ldots,N} \Delta_i \leq \sup_{i=2,\ldots,N} \Delta_i \leq \Delta, \tag{28}$$

where $\eta \in (0,1)$ is some positive constant and Δ is a deterministic sequence with $\Delta \to 0$.

A6. *We have $\epsilon_{t,T}(k_j) = O_{t,T}(k_j)\zeta_{t,T}(k_j)\bar{\epsilon}_j$, for some $\zeta_{t,T}(k_j)$ that is $\mathcal{F}^{(0)}$-adapted and is such that we have $\sup_{j=1,\ldots,N}|\zeta_{t,T}(k_j)| = O_p(1)$ as $N \to \infty$, and a sequence $\{\bar{\epsilon}_j\}_{j=1}^N$ of $\mathcal{F}^{(1)}$-adapted random variables that are independent of each other and of $\mathcal{F}^{(0)}$, have mean zero and finite absolute moments of all order.*

Assumption A2 specifies the diffusive volatility as an Itô semimartingale, which is the case for most parametric stochastic volatility models used in empirical work. Assumption A3 is a conditional integrability restriction. Note that x is the log-price and ν is the jump compensator of the log-jumps. Therefore, assumption A3 imposes, in particular, existence of conditional risk-neutral fourth moment of the price and similar conditional moment requirement for the price jumps.[9] Assumption A4 is a smoothness in expectation condition for various processes that appear in the dynamics of x. Such an assumption will be satisfied if these processes themselves are Itô semimartingales, which is the case for most parametric stochastic volatility models. Assumption A5 is a weak assumption regarding the strike grid and finally assumption A6 is about the observation errors. We note that the assumption for the $\mathcal{F}^{(0)}$-conditional independence of the observation errors in A6 can be trivially extended to allow for weak spatial dependence in the errors in the strike domain.

[9] For the existence of option prices, we need at least $\int_{\mathbb{R}} e^z \nu_t(z)dz < \infty$. The requirement in assumption A3 concerning the jump compensator, in particular, can accommodate regular variation condition in the tails (i.e., power law tail decay) for the jumps in "levels" that is given for example in assumption A2 of Bollerslev and Todorov (2011a) or in Bollerslev and Todorov (2014). The transformation of log-jumps to jumps in "levels" is done by $e^x - 1$ for $x > 0$ and by $e^{-x} - 1$ for $x < 0$. Our empirical estimates of the regular variation tail decay parameter in the empirical section are well above 3 required by our assumption A3 (in the case of regular variation in jump tails).

3.3 Asymptotic Bounds for the Estimators

We proceed with deriving bounds for the integrated squared error of the estimators. The next theorem contains the result for $\widehat{LV}_{t,T}(\vartheta, z_T, u_T)$.

Theorem 1. *Suppose assumption A1, with some $c > 0$, and assumptions A2-A6 hold. In addition,*

$$\int_{\mathbb{R}} (|x| \wedge 1)^{r'} \nu_t(x) dx < \infty, \tag{29}$$

for some $r' \in [0,1]$. Assume $\Delta \asymp T^{\alpha}$, $\underline{K} \asymp T^{\beta}$ and $\overline{K} \asymp T^{-\beta}$, for $\alpha > \frac{1}{2}$ and $\beta > 0$. Let $0 < c < \varsigma$, for some $\varsigma \in (0,2)$, and

$$z_T \to \infty,\ z_T^3 \Delta \to 0 \text{ and } z_T^3 T^{1-\frac{1}{3}(\varsigma+1)} |\log(T)|^8 \to 0, \tag{30}$$

as well as

$$u_T \to \infty \text{ and } u_T^2 T \to 0. \tag{31}$$

We then have:

$$\int_0^{\infty} \left(\widehat{LV}_{t,T}(\vartheta, z_T, u_T) - LV_t(\vartheta) \right)^2 d\vartheta = O_p\Big(z_T^{-2r-2} \bigvee \Delta z_T^3 \bigvee |\log(T)|^8 T^{1-\frac{1}{3}(\varsigma+1)} z_T^3 \bigvee |\log(T)|^4 T^{2(2-\varsigma)\beta} z_T^3 \bigvee u_T^{-2(2-r')} \bigvee \sqrt{T} u_T^2 \Delta \bigvee \frac{\Delta}{\sqrt{T} u_T^2} \Big). \tag{32}$$

The convergence of $\widehat{LV}_{t,T}(\vartheta, z_T, u_T)$ is in a joint asymptotic setting in which the time to maturity and the mesh of the strike grid shrink to zero, and the log-strike range increases to $(-\infty, +\infty)$. The third requirement for z_T in (30) ensures that $\widehat{\mathcal{L}}_{t,T}(u)$ is converging to 1 uniformly over the range of u used in the computation of the jump variation measure.

There are several sources of estimation error involved in the recovery of $LV_t(\vartheta)$ which determine the rate of convergence of $\widehat{LV}_{t,T}(\vartheta, z_T, u_T)$ given in (32). First, in the estimation we evaluate $\widehat{h}_{t,T}(u)$ over u with a finite range of its imaginary part. This generates bias in the estimation, the size of which depends on the smoothness of the measure $x^2 \nu_t(x)$, and is controlled by the parameter r in assumption A1. As noted above, the value of r will in general depend on the degree of jump activity, with higher activity implying smaller value of r, and thus slower rate of convergence of the estimator.

Second, the second term on the right-hand side of (32) is due to the option observation error. In this regard we note that the option error in assumption A5 is assumed to be proportional to the option price it is attached to. The option prices have different asymptotic order depending on their strike. In particular, options with $k - x_0$ in a range of zero that is of asymptotic size $\sqrt{T}$ are of asymptotic order $O_p(\sqrt{T})$ while options with $k - x_0$ away from zero are of the smaller asymptotic order $O_p(T)$. $\widehat{LV}_{t,T}(\vartheta, z_T, u_T)$ involves options of all strikes but they receive different weights. Due to the debiasing of $\widehat{h}_{t,T}(-c - iz_T)$ by $\widehat{\sigma}^2_{t,T}(u_T)$, the weight of the near-the-money options in the construction of $\widehat{LV}_{t,T}(\vartheta, z_T, u_T)$ can be minimized (provided that u_T is chosen optimally) so that they do not slow down the rate of convergence of the estimator. This is contrast to the estimation of the diffusive volatility coefficient for which the error in the near-the-money options determines its asymptotic distribution, see Todorov (2019).

The third term on the right-hand side of (32) is due to the time variation in the spot characteristics of x_t, i.e., due to the variation in its stochastic volatility, jump intensity and its drift term. If these are constant, then this term will be zero. We also conjecture that this condition might be possible to relax under some stronger conditions for the dynamics of x than those assumed in assumption A4.

The fourth term on the right-hand side of (32) is due to the finite range of the available options used in the estimation. Naturally, the size of this error depends on how fat are the tails of $\nu_t(x)$, with thinner tails corresponding to a smaller bias in the estimation. In practice, this error is expected to be small, at least for the case of index options, as the prices (mid-quotes) of the deepest available out-of-the-money options are typically no higher than two ticks.

Finally, the last three terms on the right-hand side of (32) are all due to the approximation error in the estimator $\widehat{\sigma}^2_{t,T}(u_T)$ of σ^2_t. The first of them is a bias due to the jumps while the other two are due to the measurement error in the options. The bias in $\widehat{\sigma}^2_{t,T}(u_T)$ due to the jumps is controlled by the parameter r' which is a bound on the jump activity and controls the rate of explosion of $\nu_t(x)$ near the origin. For the tempered stable example considered above, $r' < \beta$ and the two parameters r and r' are related via $r' = (2 - r - 1/2) \vee 0$. If u_T is set optimally, the last three terms on the right-hand side of (32) will be negligible in general. For example, if $r' < 2/3$, this will always be the case if we choose $u_T \asymp T^{-1/4}$.

In the case when the first two terms on the right-hand side of (32) dominate, which recall are due to the smoothness of the estimated density and the option observation error, then with z_T set to $z_T \asymp \Delta^{-\frac{1}{2r+5}}$, the convergence rate of the integrated squared error becomes of asymptotic order $O_p\left(\Delta^{\frac{2r+2}{2r+5}}\right)$. One can show that this is the optimal rate of convergence in a minimax sense of an estimator of $LV_t(\vartheta)$ from noisy short-dated option data for Lévy densities satisfying assumption A1.[10] This rate is faster than the rate of convergence of the option-based estimator of $x^3\nu_t(x)$ developed in Qin and Todorov (2019) whose integrated squared error is of size $O_p\left(\Delta^{\frac{2r+1}{2r+5}}\right)$. The situation here is analogous to comparing rates of convergence of estimators of the probability density (Fan (1991), Butucea and Comte (2009)) to those of the cumulative distribution function (Dattner et al. (2011), Butucea and Comte (2009) and Pensky (2017)) in nonparametric deconvolution problems.[11]

[10]This can be done by following the same steps as in the proof of Theorem 3 in the Appendix of Qin and Todorov (2019), with the only change being that the rate in equation (A.47) of that paper becomes now $2^{-2jr-2j}$.

[11]The estimation of the Lévy measure in the current setting can be naturally related to the standard nonparametric

We next state the counterpart of Theorem 1 for the alternative estimator $\widehat{LV}_{t,T}(\vartheta, z_T, \kappa_T, \overline{\kappa}_T)$.

Theorem 2. *Suppose assumption A1, with some $c > 0$, and assumptions A2-A6 hold. Assume $\Delta \asymp T^{\alpha}$, $\underline{K} \asymp T^{\beta}$ and $\overline{K} \asymp T^{-\beta}$, for $\alpha > \frac{1}{2}$ and $\beta > 0$. Let $0 < c < \varsigma$, for some $\varsigma \in (0, 2)$, and*

$$z_T \to \infty,\ z_T^3 \Delta \to 0 \text{ and } z_T^3 T^{1-\frac{1}{3}(\varsigma+1)} |\log(T)|^8 \to 0, \tag{33}$$

as well as

$$\overline{\kappa}_T > \kappa_T \geq 1,\ \overline{\kappa}_T \asymp \kappa_T,\ \ \kappa_T^8 z_T^2 \sqrt{T} |\log(T)| \to 0 \text{ and } \kappa_T^2 / z_T \to 0. \tag{34}$$

We then have:

$$\begin{aligned} \int_0^{\infty} \left(\widehat{LV}_{t,T}(\vartheta, z_T, \kappa_T, \overline{\kappa}_T) - LV_t(\vartheta)\right)^2 d\vartheta = O_p\Big(& z_T^{-2r-2} \bigvee \Delta z_T^3 \bigvee |\log(T)|^8 T^{1-\frac{1}{3}(\varsigma+1)} z_T^3 \\ & \bigvee |\log(T)|^4 T^{2(2-\varsigma)\beta} z_T^3 \bigvee \kappa_T^{-2r-1} z_T^{-2r-1} \Big). \end{aligned} \tag{35}$$

The difference between the results regarding the asymptotic orders of the approximation error in $\widehat{LV}_{t,T}(\vartheta, z_T, u_T)$ and $\widehat{LV}_{t,T}(\vartheta, z_T, \kappa_T, \overline{\kappa}_T)$ is that the last three terms in (32) are now replaced by the last term in (35). These differences reflect the different de-biasing terms used in the two estimators. If κ_T is chosen to be constant and the first two terms on the right-hand side of (35) are leading, then the best possible rate for the integrated squared error of $\widehat{LV}_{t,T}(\vartheta, z_T, \kappa_T, \overline{\kappa}_T)$ is $O_p\left(\Delta^{\frac{2r+1}{2r+4}}\right)$ which is slightly bigger than the optimal one of $O_p\left(\Delta^{\frac{2r+2}{2r+5}}\right)$ that can be achieved by $\widehat{LV}_{t,T}(\vartheta, z_T, u_T)$. Thus, with κ_T constant, $\widehat{LV}_{t,T}(\vartheta, z_T, \kappa_T, \overline{\kappa}_T)$ has a loss of efficiency but this loss is rather small. Indeed, in the typical case of finite activity jumps with C^1 Lévy density, which implies $r > 3/2$, the rate efficiency loss is less than 5%. On the other hand, if we choose κ_T to be increasing at the fastest rate such that condition (34) holds, then it is easy to show that the optimal rate $O_p\left(\Delta^{\frac{2r+2}{2r+5}}\right)$ can be achieved, provided $r > 3/2$.[12]

The above results are for the integrated squared error of the estimators. As is known, see e.g., Cai (2003), pointwise convergence rates in Sobolev spaces can differ. This turns out to be the case in our situation too, as we will now show. For simplicity, we will only consider $\widehat{LV}_{t,T}(\vartheta, z_T, u_T)$ and $\widehat{LV}_{t,T}(\vartheta, z_T, \kappa_T, \overline{\kappa}_T)$, for $\vartheta \neq 0$. The bounds on the rate of convergence of these estimators are provided in the following theorem.

Theorem 3. *Under the conditions of Theorem 1, we have*

$$\begin{aligned} \left(\widehat{LV}_{t,T}(\vartheta, z_T, u_T) - LV_t(\vartheta)\right)^2 = O_p\Big(& z_T^{-2r-1} \bigvee \Delta z_T^3 \bigvee |\log(T)|^8 T^{1-\frac{1}{3}(\varsigma+1)} z_T^3 \\ & \bigvee |\log(T)|^4 T^{2(2-\varsigma)\beta} z_T^3 \bigvee u_T^{-2(2-r')} \bigvee \sqrt{T} u_T^2 \Delta \bigvee \frac{\Delta}{\sqrt{T} u_T^2} \Big), \quad \text{for } \vartheta \neq 0, \end{aligned} \tag{36}$$

deconvolution problem in which one is interested in estimating the unknown density or cumulative distribution function of a random variable from its noisy observations. Indeed, like in the latter case, our estimation here is based on Fourier (or Laplace) inversions of a quantity $(\widehat{h}_{t,T}(u))$ constructed as a ratio of estimated from the data Laplace transforms.

[12] I am grateful to a referee for suggesting estimation with asymptotically increasing κ_T and $\overline{\kappa}_T$.

while under the conditions of Theorem 2, we have

$$\left(\widehat{LV}_{t,T}(\vartheta, z_T, \kappa_T, \overline{\kappa}_T) - LV_t(\vartheta)\right)^2 = O_p\bigg(z_T^{-2r-1} \bigvee \Delta z_T^3 \bigvee |\log(T)|^8 T^{1-\frac{1}{3}(\varsigma+1)} z_T^3 \bigvee |\log(T)|^4 T^{2(2-\varsigma)\beta} z_T^3\bigg), \quad \textit{for } \vartheta \neq 0. \tag{37}$$

The above statements continue to hold if we replace ϑ with $\widehat{\vartheta}_t$ that satisfies

$$\widehat{\vartheta}_t - \vartheta_t = O_p(\upsilon_T), \quad \textit{for some deterministic sequence } \upsilon_T \to 0, \tag{38}$$

and where ϑ_t is $\mathcal{F}_t^{(0)}$-adapted random variable which is almost surely positive, provided Δz_T^3 on the right-hand sides of (36) and (37) is replaced by $\Delta z_T^3(z_T^3 \upsilon_T^2 \vee 1)$.

We can make several observations about the results of the theorem. First, the first term on the right-hand side of (36) is bigger than its counterpart in (32). Recall that this term is due to the truncation of the limits of integration in $\widehat{LV}_{t,T}(\vartheta, z_T, u_T)$ and the size of it depends on the smoothness of $LV_t(\vartheta)$. Second, the rest of the terms on the right-hand side of (36) are of the same size as their counterparts in (32). Similar observations can be made when comparing the results in (37) and (35). Here, it is also interesting to note that choosing κ_T increasing provides no efficiency gains. If the first two terms on the right-hand sides of (36) and (37) are leading, then upon choosing z_T optimally, we get $O_p\left(\Delta_n^{\frac{2r+1}{2r+4}}\right)$ for the size of $\left(\widehat{LV}_{t,T}(\vartheta, z_T, u_T, \kappa_T, \overline{\kappa}_T) - LV_t(\vartheta)\right)^2$ and $\left(\widehat{LV}_{t,T}(\vartheta, z_T) - LV_t(\vartheta)\right)^2$. This bound on the pointwise squared error is slightly bigger than the optimal one for the integrated squared error.

The last part of Theorem 3 deals with the case when the cutoff parameter ϑ is replaced with an estimator $\widehat{\vartheta}_t$ which converges asymptotically to an $\mathcal{F}^{(0)}$-adapted positive random variable. In this case, if the rate of convergence υ_T of $\widehat{\vartheta}_t$ is not fast enough so that $z_T^3 \upsilon_T^2$ is negligible, then the rate of convergence of $\widehat{LV}_{t,T}(\widehat{\vartheta}_t, z_T, \kappa_T, \overline{\kappa}_T)$ and $\widehat{LV}_{t,T}(\widehat{\vartheta}_t, z_T, \kappa_T, \overline{\kappa}_T)$ will be slowed down compared to the case when using non-random cutoff ϑ. The result in Theorem 3 is for an abstract estimator $\widehat{\vartheta}_t$. This bound can be improved in more specific situations. For example, if $\widehat{\vartheta}_t$ is $\mathcal{F}^{(0)}$-conditionally independent from the option observation errors at time t, then (36) and (37) will continue to hold with ϑ replaced by $\widehat{\vartheta}_t$. This will be the case for example, if $\widehat{\vartheta}_t$ is formed from return data or from option data prior to time t. In addition, in cases when $\widehat{\vartheta}_t$ is formed from $\{\widehat{O}_T(k_j)\}_{j\geq 1}$, one can improve on the result of Theorem 3 with additional knowledge about $\widehat{\vartheta}_t$.

4 Choice of Tuning Parameters

We turn next to feasible implementation of the developed estimation method. Since implementing $\widehat{LV}_{t,T}(\vartheta, z_T, \kappa_T, \overline{\kappa}_T)$ is easier than implementing $\widehat{LV}_{t,T}(\vartheta, z_T, u_T)$, in what follows we will discuss its implementation only. We start first with estimating the diffusive spot volatility from the available option data using the estimator proposed in Todorov (2019). We will use this estimate to set the tail cutoff parameter ϑ in $\widehat{LV}_{t,T}(\vartheta, z_T, \kappa_T, \overline{\kappa}_T)$. Since on a given day we have a number of short maturities, in order to minimize the effect from the measurement error in the data, we integrate all maturities in the estimation. Towards this end, denote a set of (ordered) maturities with $\mathcal{T} = (T_1, T_2, ..., T_k)$. Then, we form the Laplace transform estimator from all available maturities as

$$\widehat{\mathcal{L}}_{t,\mathcal{T}}(u) = \prod_{i=1}^{k} \widehat{\mathcal{L}}_{t,T_i}(u). \tag{39}$$

We then set $\widehat{u}_{t,\mathcal{T}} = \widehat{u}_{t,\mathcal{T}}^{(1)} \wedge \widehat{u}_{t,\mathcal{T}}^{(2)}$ with,

$$\begin{aligned} \widehat{u}_{t,\mathcal{T}}^{(1)} &= \inf\left\{u \geq 0 : |\widehat{\mathcal{L}}_{t,\mathcal{T}}(iu)| \leq 0.1^k\right\}, \\ \widehat{u}_{t,\mathcal{T}}^{(2)} &= \operatorname{argmin}_{u \in [0,\overline{u}]} |\widehat{\mathcal{L}}_{t,\mathcal{T}}(iu)|, \quad \text{with} \quad \overline{u}_{t,\mathcal{T}} = \sqrt{\frac{2}{\sum_{i=1}^{k} T_i} \frac{\log(1/0.1)}{\widehat{\mathrm{BSIV}}_{t,T_1}^{2}}}, \end{aligned}$$

where $\widehat{\mathrm{BSIV}}_{t,T_1}$ is the Black-Scholes Implied Volatility for the shortest available maturity at time t and the strike closest to the forward level. The nearly rate-efficient spot volatility estimator proposed by Todorov (2019) is then given by:[13]

$$\widehat{\sigma}_{t,\mathcal{T}}^2 = -\frac{2}{\sum_{i=1}^{k} T_i} \frac{1}{\widehat{u}_{t,\mathcal{T}}^2} \log|\widehat{\mathcal{L}}_{t,\mathcal{T}}(i\widehat{u}_{t,\mathcal{T}})|. \tag{40}$$

Using the estimator of spot diffusive volatility, we follow Bollerslev et al. (2015) and Andersen et al. (2021) and set the cutoff parameter ϑ proportional to the current level of spot diffusive volatility:

$$\widehat{\vartheta}_t = 5 \times \widehat{\sigma}_{t,\mathcal{T}} \sqrt{5/252}. \tag{41}$$

We turn next to setting the tuning parameters for the jump variation estimation. As for the estimation of the spot volatility, we will utilize all short dated options, i.e., we will use

$$\begin{aligned} \widehat{LV}_{t,\mathcal{T}}(\vartheta, z_\mathcal{T}, \kappa_\mathcal{T}, \overline{\kappa}_\mathcal{T}) = \frac{1}{2\pi} \int_{-z_\mathcal{T}}^{+z_\mathcal{T}} \frac{e^{-\vartheta(c+iz)}}{c+iz} \Bigg(& \widehat{h}_{t,\mathcal{T}}(-c-iz) \\ & - \frac{1}{2(\overline{\kappa}_\mathcal{T} - \kappa_\mathcal{T}) z_\mathcal{T}} \int_{|v| \in [\kappa_\mathcal{T} z_\mathcal{T}, \overline{\kappa}_\mathcal{T} z_\mathcal{T}]} \widehat{h}_{t,\mathcal{T}}(-c-iv) dv \Bigg) dz, \end{aligned} \tag{42}$$

where

$$\widehat{h}_{t,\mathcal{T}}(u) = \frac{1}{k} \sum_{i=1}^{k} \widehat{h}_{t,T_i}(u). \tag{43}$$

Throughout the estimation we set c to some very small value, which in all of our implementations will be 0.01, $\kappa_\mathcal{T} = 1$ and $\overline{\kappa}_\mathcal{T} = 1.01$. Henceforth, we use the following simplifying notation:

$$\widehat{LV}_{t,\mathcal{T}}(\vartheta, z_\mathcal{T}) = \widehat{LV}_{t,\mathcal{T}}(\vartheta, z_\mathcal{T}, z_\mathcal{T}, 1.01 \times z_\mathcal{T}). \tag{44}$$

[13] Note that $\widehat{u}_{t,\mathcal{T}} \asymp 1/\sqrt{T_1}$.

The choice of $z_{\mathcal{T}}$ is in general difficult as it needs to balance several sources of error that are present in the estimation. We therefore set $z_{\mathcal{T}}$ to $\widehat{z}_{\mathcal{T}}$ which we determine in the following way. We choose $\widehat{z}_{\mathcal{T}}$ from the elements of the following set

$$\widehat{\mathcal{Z}}_{\mathcal{T}} = \left\{ \lfloor \widehat{u}_{t,\mathcal{T}}^{4/21}/0.5 \rfloor 0.5 : 0.5 : \lfloor \widehat{u}_{t,\mathcal{T}}^{4/21}/0.5 \rfloor 0.5 + 100 \log(1/T_1) \right\}, \tag{45}$$

and we remind the reader that $\widehat{u}_{t,\mathcal{T}} = O_p(1/\sqrt{T_1})$. This choice of $\widehat{z}_{\mathcal{T}}$ balances the asymptotic size of the first and third error terms on the right-hand side of (37) when $r = 2 - \frac{1}{2}$, which corresponds to the the case of finite activity jumps with sufficiently smooth density of the jump distribution. If the rest of the error terms in (37) are non-binding, this will mean that the squared error of $\widehat{LV}_{t,\mathcal{T}}(\vartheta, \widehat{z}_{\mathcal{T}})$ is of order $O_p\left(T_1^{\frac{8*0.99}{21}}\right)$.

Among the elements in $\widehat{\mathcal{Z}}_{\mathcal{T}}$, we select $\widehat{z}_{\mathcal{T}}$ in a way that balances the bias in the estimation due to using the finite interval $[-\widehat{z}_{\mathcal{T}}, \widehat{z}_{\mathcal{T}}]$ in the integration on one hand and the noisiness of the estimator on the other hand (first and second terms on the right-hand side of (37)). For the latter, we use an estimate of the asymptotic variance of $\widehat{LV}_{t,\mathcal{T}}(\vartheta, z_T)$ which we now describe. We denote

$$\begin{aligned} \widehat{\varkappa}_{t,T}(k,u) = & \frac{\nabla_{uu} f(u, k - x_t)}{\widehat{\mathcal{L}}_{t,T}(u)} - 2 \frac{\nabla_u \widehat{\mathcal{L}}_{t,T}(u)}{\widehat{\mathcal{L}}_{t,T}(u)^2} \nabla_u f(u, k - x_t) \\ & + \left(2 \frac{\left(\nabla_u \widehat{\mathcal{L}}_{t,T}(u)\right)^2}{\widehat{\mathcal{L}}_{t,T}(u)^3} - \frac{\nabla_u \widehat{\mathcal{L}}_{t,T}(u)}{\widehat{\mathcal{L}}_{t,T}(u)^2} \right) f(u, k - x_t), \end{aligned} \tag{46}$$

and

$$\widehat{\omega}_{t,T}^n(k, \overline{z}) = \frac{1}{2\pi T} \int_{-\overline{z}}^{\overline{z}} \frac{e^{-\vartheta(c+iz)}}{c+iz} \left(\widehat{\varkappa}_{t,T}(k, -c - iz) - \frac{1}{0.02\overline{z}} \int_{|v| \in [\overline{z}, 1.01\overline{z}]} \widehat{\varkappa}_{t,T}(k, -c - iv) dv \right) dz. \tag{47}$$

With this notation, our estimator of the asymptotic variance is given by

$$\widehat{\mathcal{V}}_{t,\mathcal{T}}(\overline{z}) = \frac{1}{k} \sum_{i=1}^{k} \widehat{\mathcal{V}}_{t,T_i}(\overline{z}), \tag{48}$$

where

$$\begin{aligned} \widehat{\mathcal{V}}_{t,T}(\overline{z}) = & \frac{2}{3} e^{-2x_t} \widehat{\omega}_{t,T}^n(k_1, \overline{z})^2 \widehat{O}_T(k_1)^2 \Delta_1^2 \\ & + \frac{2}{3} e^{-2x_t} \sum_{j=3}^{N-1} \widehat{\omega}_{t,T}^n(k_{j-1}, \overline{z})^2 \left(\widehat{O}_T(k_{j-1}) - \frac{1}{2}\widehat{O}_T(k_{j-2}) - \frac{1}{2}\widehat{O}_T(k_j) \right)^2 \Delta_j^2. \end{aligned} \tag{49}$$

Given $\widehat{\mathcal{V}}_{t,T}(\overline{z})$, we determine

$$\overline{z}_{\mathcal{T}} = \inf \left\{ z \in \widehat{\mathcal{Z}}_{\mathcal{T}} : \sqrt{\widehat{\mathcal{V}}_{t,\mathcal{T}}(z)} \geq 0.3 \times \widehat{LV}_{t,\mathcal{T}}(\vartheta, z) \right\}, \ \overline{\mathcal{Z}}_{\mathcal{T}} = \left\{ \lfloor \widehat{u}_{t,\mathcal{T}}^{4/21}/0.5 \rfloor 0.5 : 0.5 : \overline{z}_{\mathcal{T}} \right\}. \tag{50}$$

The point $\overline{z}_{\mathcal{T}}$ can be viewed as an upper limit beyond which the estimation is deemed very noisy. We then set

$$\widehat{z}_{\mathcal{T}} = \inf\left\{z \in \overline{\mathcal{Z}}_{\mathcal{T}} : \cap_{v\in\overline{\mathcal{Z}}_{\mathcal{T}}:v\geq z}\widehat{U}(v) \neq \emptyset\right\}, \tag{51}$$

where

$$\widehat{U}(z) = \left(\widehat{LV}_{t,\mathcal{T}}(\vartheta, z) - \frac{\log(1/T_1)}{16}\sqrt{\widehat{\mathcal{V}}_{t,\mathcal{T}}(z)},\, \widehat{LV}_{t,\mathcal{T}}(\vartheta, z) + \frac{\log(1/T_1)}{16}\sqrt{\widehat{\mathcal{V}}_{t,\mathcal{T}}(z)}\right). \tag{52}$$

The idea of this choice of $\widehat{z}_{\mathcal{T}}$ is that we pick the smallest value of z such that the change in a relatively small confidence interval around the estimate does not change by much when changing z (when starting from $\overline{z}_{\mathcal{T}}$), with the latter change being associated with a big bias due to the tail of $\mathcal{F}\nu_t(u)$ relative to the estimation uncertainty associated with recovering the jump variation.

Finally, if on a given day $\sqrt{\widehat{\mathcal{V}}_{t,\mathcal{T}}(z)} \geq 0.2 \times \widehat{LV}_{t,\mathcal{T}}(\vartheta, z)$ for a very small fixed value of $z = 0.5$, then we drop this day from the analysis as the data is deemed too noisy for reliable estimation. Indeed, asymptotically for a fixed z, this should not happen. In the real data this happens very rarely when there are occasional big strike gaps in the available options.

Remark 1. *Provided the tuning parameters are chosen in a way that the various biases in the estimation are negligible, one can derive a Central Limit Theorem for the left jump variation estimator with $\widehat{\mathcal{V}}_{t,\mathcal{T}}(z)$, given above, being an estimate for the $\mathcal{F}$-conditional limiting variance of $\widehat{LV}_{t,\mathcal{T}}(\vartheta, z)$. The derivations are similar to those for the Central Limit Theorem result for the volatility estimator in Todorov (2019), although the rate of convergence here will be very different than the one in that paper.*

5 Monte Carlo Study

We now evaluate the performance of our jump variation estimators in finite samples on simulated data. We start with introducing the option model used in the Monte Carlo as well as our option observation scheme, and we then present the simulation results.

5.1 Model Setup and Option Sampling Scheme

The dynamics of X under $\mathbb{Q}$ that we use in our Monte Carlo experiments is given by

$$\frac{dX_t}{X_{t-}} = \sqrt{V_t}dW_t + \int_{\mathbb{R}} (e^x - 1)\,\widetilde{\mu}(dt, dx),\; dV_t = 4(0.02 - V_t)dt + 0.2\sqrt{V_t}dB_t, \tag{53}$$

where (W_t, B_t) is a bivariate Brownian motion with $\text{cov}(dW_t, dB_t) = -0.5dt$, and $\widetilde{\mu} = \mu - \nu$, for μ being an integer-valued random measure on $\mathbb{R}_+ \times \mathbb{R}$ with the following compensator ν:

$$\nu(dt, dx) = \left(c_-\frac{e^{-\lambda_-|x|}}{|x|}1_{\{x<0\}} + c_+\frac{e^{-\lambda_+|x|}}{|x|}1_{\{x>0\}}\right)V_t dt \otimes dx, \tag{54}$$

$$c_- = 0.9\lambda_-^2 \quad \text{and} \quad c_- = 0.1\lambda_+^2. \tag{55}$$

We fix the value $\lambda_+ = 50$ in all Monte Carlo scenarios and consider three cases for the parameter λ_-: $\lambda_- = 10$, $\lambda_- = 20$ and $\lambda_- = 30$. The choice of these parameters matches key features of the option data set used in our empirical application. In particular, the three considered values of λ_- correspond roughly to the 10-th, 50-th and 90-th quantiles of the left tail index estimates from the data.

Turning next to the option sampling scheme, we set it in a way that mimics the one corresponding to the real option data used in the empirical application. In particular, on each day we consider options with four times to maturity, $T_1 = 3/252$, $T_2 = 5/252$, $T_3 = 7/252$ and $T_4 = 9/252$ (unit of time is a business day). For each maturity, the strikes grids are equidistant with strike gap of 5 and deepest out-of-the-money put and call corresponding to the lowest and highest, respectively, strikes on the grid for which the (true) option price exceeds in value 0.075. Finally, the observed option prices are given by $\widehat{O}_{t,T_i}(k_j) = O_{t,T_i}(k_j) \times (1 + 0.05 z_{i,j})$, for $\{z_{i,j}\}_{i\geq 1, j\geq 1}$ being a sequence of independent standard normal random variables. The initial value of the underlying asset is set to $X_t = 2500$.

The steps for computing the jump variation estimator are summarized as follows:

1. On each day, we aggregate the option data into an estimate of the Laplace transform according to (14) and (39).

2. We set the cutoff parameter ϑ to $\widehat{\vartheta}_t$ given in (41).

3. We set $z_{\mathcal{T}}$ to $\widehat{z}_{\mathcal{T}}$ given in (51).

4. We compute $\widehat{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t, \widehat{z}_{\mathcal{T}})$ according to (42) and (44).

5.2 Results

We compare the performance of the newly-proposed jump variation measure with the one based on the asymptotic result in (1), coupled with extreme value distribution approximation in the tails, developed in Bollerslev and Todorov (2011b, 2014). This approximation, in turn, leads to the following asymptotic result for the price of deep out-of-the-money puts

$$\frac{P_{t,T}(k)}{TX_t} \xrightarrow{\mathbb{P}} \phi_t \int_{\mathbb{R}} (e^{k-x_t} - e^u)^+ e^{\alpha_t u} du = \frac{\phi_t}{\alpha_t(\alpha_t+1)} \times e^{(\alpha_t+1)(k-x_t)}, \text{ as } T \downarrow 0 \text{ and } k \downarrow -\infty, \tag{56}$$

for some $\alpha_t > 0$ and $\phi_t > 0$. We note that the extreme value distribution approximation in the jump tails in the context of the parametric specification of the jump compensator in (54) boils down to replacing it with

$$\left(c_- e^{-\lambda_-|x|} 1_{\{x<0\}} + c_+ e^{-\lambda_+|x|} 1_{\{x>0\}}\right) V_t dt \otimes dx, \tag{57}$$

for large absolute value of the jump size x. The resulting jump variation estimator is given by

$$\widetilde{LV}_{t,\mathcal{T}}(\vartheta) = \frac{1}{k}\sum_{i=1}^{k}\widetilde{LV}_{t,T_i}(\vartheta), \quad \widetilde{LV}_{t,T_i}(\vartheta) = e^{-\vartheta\widehat{\alpha}_i}\widehat{\phi}_i\frac{\widehat{\alpha}_i\vartheta(\widehat{\alpha}_i\vartheta+2)+2}{\widehat{\alpha}_i^3}, \tag{58}$$

where

$$\widehat{\alpha}_i = \widehat{b}_{1,i} - 1 \quad \text{and} \quad \widehat{\phi}_i = e^{\widehat{b}_{0,i}}\widehat{\alpha}_i(\widehat{\alpha}_i+1), \tag{59}$$

and $(\widehat{b}_{0,i}, \widehat{b}_{1,i})$ are the nonlinear least squares coefficient estimates from the following nonlinear regression:

$$\frac{\widehat{O}_{t,T_i}(k_j)}{X_t} = \exp\left(b_0 + b_1 \times (k_j - x_t)\right) + \epsilon_j^{(i)}, \ j : k_j \in [x_t - 8\widehat{\sigma}_{t,\mathcal{T}}\sqrt{T_i}, x_t - 3\widehat{\sigma}_{t,\mathcal{T}}\sqrt{T_i}]. \tag{60}$$

We note that the above estimation differs from the one proposed in Bollerslev and Todorov (2014) which instead relies on an explicit estimate of α_t based on the log-differences in options with consecutive strikes. However, results which are not reported for brevity show that the current implementation of the method in Bollerslev and Todorov (2014) leads to significantly more efficient and less biased estimation.

The results from the Monte Carlo are summarized in Table 1 below. From the table we can see that, consistent with our theoretical results, the newly-proposed estimator provides a more accurate estimate of the risk-neutral jump variation. In all considered cases, $\widehat{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t, \widehat{z}_{\mathcal{T}})$ has significantly less bias than $\widetilde{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t)$. The bias reduction is biggest in relative terms for higher levels of the diffusive volatility. The intuition for this is that the separation of diffusive volatility from jumps is harder for higher levels of volatility when using only the prices of deep out-of-the-money puts. Recall that the estimator $\widetilde{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t)$ is based on approximating out-of-the-money option prices as if they are generated from a model without diffusive volatility. This type of approximation is worse in high volatility regimes. The estimator $\widehat{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t, \widehat{z}_{\mathcal{T}})$, which is based on the characteristic function of returns, provides a far more efficient way of separating volatility from jumps in such scenarios.

From the results in Table 1 we can also see that $\widehat{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t, \widehat{z}_{\mathcal{T}})$ provides significant bias reduction over $\widetilde{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t)$ for higher levels of the parameter λ_-. The reason for this is that the higher levels of λ_- correspond to thinner tails of the Lévy jump measure and most of the jump variation is generated by more frequent jumps of smaller size. This means that for these types of jump specifications, an approximation based on assuming that the time interval till the expiration of the option contains at most one "big" jump, on which $\widetilde{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t)$ is based, does not work very well even when the threshold $\widehat{\vartheta}_t$ is chosen relatively big. Again, $\widehat{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t, \widehat{z}_{\mathcal{T}})$ does not rely on such an approximation and is therefore far less sensitive to the value of λ_-.

Finally, $\widehat{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t, \widehat{z}_{\mathcal{T}})$ is somewhat noisier than $\widetilde{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t)$. This is not surprising given the fact that the latter estimator is semiparametric while the former is fully nonparametric. Nevertheless,

Table 1: Monte Carlo Results

Estimator	Median	IQR	Median	IQR	Median	IQR
$\lambda_- = 10,\ \lambda_+ = 50$	$V_t = 0.01,\ LV_t(\vartheta_t) = 0.0076$		$V_t = 0.02,\ LV_t(\vartheta_t) = 0.0133$		$V_t = 0.03,\ LV_t(\vartheta_t) = 0.0177$	
$\widehat{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t, \widehat{z}_{\mathcal{T}})$	0.0074	0.0003	0.0131	0.0003	0.0173	0.0004
$\widetilde{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t)$	0.0078	0.0001	0.0140	0.0001	0.0192	0.0001
$\lambda_- = 20,\ \lambda_+ = 50$	$V_t = 0.01,\ LV_t(\vartheta_t) = 0.0053$		$V_t = 0.02,\ LV_t(\vartheta_t) = 0.0073$		$V_t = 0.03,\ LV_t(\vartheta_t) = 0.0081$	
$\widehat{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t, \widehat{z}_{\mathcal{T}})$	0.0056	0.0001	0.0076	0.0002	0.0081	0.0002
$\widetilde{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t)$	0.0065	0.0001	0.0098	0.0001	0.0120	0.0001
$\lambda_- = 30,\ \lambda_+ = 50$	$V_t = 0.01,\ LV_t(\vartheta_t) = 0.0034$		$V_t = 0.02,\ LV_t(\vartheta_t) = 0.0036$		$V_t = 0.03,\ LV_t(\vartheta_t) = 0.0032$	
$\widehat{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t, \widehat{z}_{\mathcal{T}})$	0.0040	0.0001	0.0044	0.0002	0.0041	0.0002
$\widetilde{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t)$	0.0050	0.0001	0.0067	0.0001	0.0078	0.0001

the need to use sufficiently deep out-of-the-money options in the construction of $\widetilde{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t)$ limits the gains of $\widetilde{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t)$ in terms of reduction in noise in the estimation. This is particularly visible for the higher values of λ_-, which correspond to thinner jump tails and correspondingly fewer available deep out-of-the-money puts (recall that as in the real data we restrict our option sample by requiring that the true option prices exceed in value 0.075).

Overall, the results from the Monte Carlo show that $\widehat{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t, \widehat{z}_{\mathcal{T}})$ provides significant bias reduction in the estimation of the jump variation at the expense of a relatively small increase in the variance of the estimator.

6 Empirical Application

We apply the developed estimation methods to study the risk-neutral jump variation of the S&P 500 market index. The data is obtained from OptionMetrics and covers the period from the beginning of 2007 till the end of June of 2019. We apply standard filters to the data and exclude options with zero bids. We further exclude cross-sections of options with the same maturity on a given day if the cheapest out-of-the-money put exceeds 25 cents in value and/or if the gap between the strikes of the closest to the money options is above 20 dollars. From the remaining cross-sections of options on a given day with different tenors, we keep only those whose tenor is between three and twenty-two business days. If on a given day, we have more than two tenors in that range, we preserve only cross-sections of options with tenor not exceeding ten business days.

In Figure 2 we plot the two alternative tail variation measures $\widehat{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t, \widehat{z}_{\mathcal{T}})$ and $\widetilde{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t)$, and in Table 2, we report their time series quantiles. Consistent with the Monte Carlo results reported in the previous section, we find that $\widetilde{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t)$ is systematically higher than $\widehat{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t, \widehat{z}_{\mathcal{T}})$ throughout the sample. From Figure 2 we can see that the difference $\widehat{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t, \widehat{z}_{\mathcal{T}}) - \widetilde{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t)$ is highest during high volatility periods such as the financial crisis of 2008-2009, the sovereign debt crises of 2010 and 2011, etc. On the other hand, the difference $\widehat{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t, \widehat{z}_{\mathcal{T}}) - \widetilde{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t)$ is very small in low volatility regimes. The correlation between the new tail measure $\widehat{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t, \widehat{z}_{\mathcal{T}})$ and the spot volatility is less than that between the latter and $\widetilde{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t)$. This is likely due to the fact that $\widehat{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t, \widehat{z}_{\mathcal{T}})$ is far less upward biased than $\widetilde{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t)$ in high volatility regimes, as observed in our Monte Carlo experiment. Consistent with that evidence and the overall dynamics of the displayed two alternative tail variation measures, we find that $\widetilde{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t)$ is significantly more persistent than $\widehat{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t, \widehat{z}_{\mathcal{T}})$.

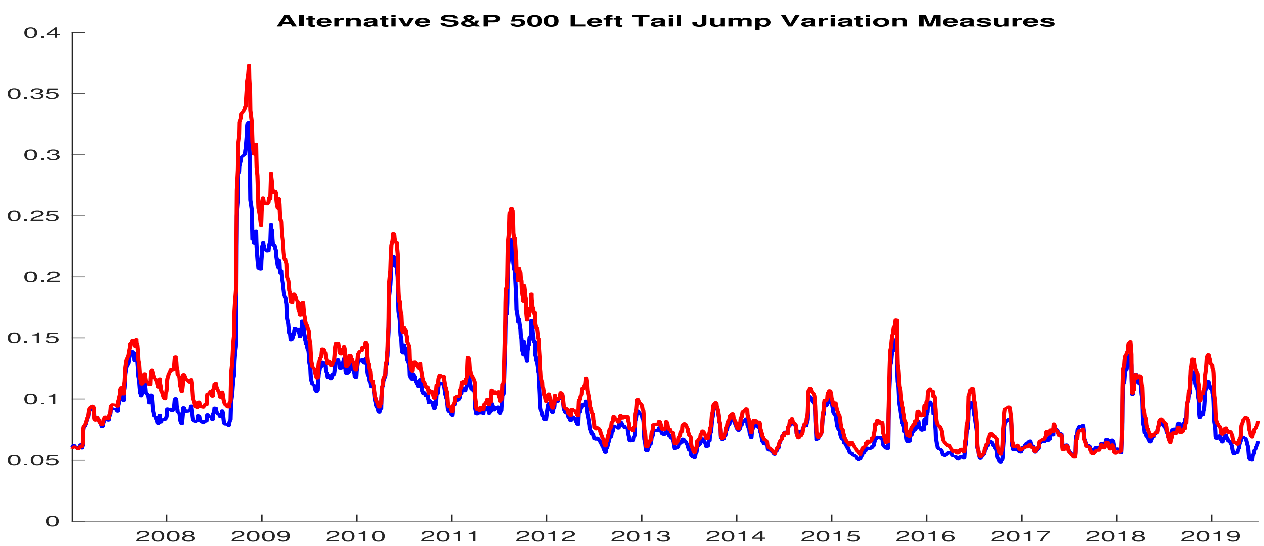

Figure 2: Alternative S&P 500 Left Tail Jump Variation Measures. Blue line corresponds to $\sqrt{\widehat{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t, \widehat{z}_{\mathcal{T}})}$ and red line to $\sqrt{\widetilde{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t)}$. The displayed series are 22-day moving averages of the corresponding daily estimates.

Table 2: Quantiles of S&P 500 Left Tail Jump Variation Measures

Estimator	Empirical Quantiles		
	25	50	75
$\widehat{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t, \widehat{z}_{\mathcal{T}})$	0.0037	0.0057	0.0104
$\widetilde{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t)$	0.0043	0.0070	0.0129

Unlike the method based on the approximation in (1)-(2) for deep out-of-the-money options, the current method allows us to study the variation of jumps of any size, and not only of the big ones. Therefore, we next compute the estimate for the total negative jump variation, using the new method, which we define formally as

$$\widehat{NJV}_{t,\mathcal{T}} \equiv \widehat{LV}_{t,\mathcal{T}}(0, \widehat{z}_{\mathcal{T}}). \tag{61}$$

$\widehat{NJV}_{t,\mathcal{T}}$ is an estimate of $\int_{z<0} z^2\nu_t(z)dz$. In Figure 3, we display this series along with the tail variation estimate $\widehat{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t, \widehat{z}_{\mathcal{T}})$ and the spot volatility $\sigma_{t,\mathcal{T}}$. All displayed series are reported in annualized volatility units. Comparing the total negative jump variation with the spot diffusive volatility, we can see that they have a similar level. We note that $\widehat{NJV}_{t,\mathcal{T}}$ is a risk-neutral measure. Its counterpart constructed from return data has a much lower level, with the difference reflecting risk premium demanded by investors for bearing negative market jump risk. We see also from the figure that the dynamics of $\widehat{NJV}_{t,\mathcal{T}}$ and $\sigma_{t,\mathcal{T}}$ differ. This is particularly notable in crisis periods where $\widehat{NJV}_{t,\mathcal{T}}$ tends to increase by more from its pre-crisis level than the diffusive volatility does. Examples of this behavior include the financial crisis in the Fall of 2008 and the two sovereign debt crises from 2010 and 2011. The latter two episodes were also characterized with more persistent elevation of $\widehat{NJV}_{t,\mathcal{T}}$, which is consistent with heightened fears in the aftermaths of these events. Similarly, the Brexit and the 2016 US elections had much more pronounced (but short-lived) impact on $\widehat{NJV}_{t,\mathcal{T}}$ than on $\sigma_{t,\mathcal{T}}$, which is again consistent with heightened fears during these episodes. On the other hand, the increased geopolitical uncertainty from the beginning of 2016 had a much bigger impact on the volatility than on the risk-neutral jump variation.

Comparing the total negative jump variation with its tail component, we can notice that there is a sizable gap between the two. This reflects the fact that the jump variation has a significant contribution that comes from smaller-sized market jumps. This type of dynamics cannot be generated by the traditional parametric jump-diffusion models used in prior empirical option pricing work, see e.g., Singleton (2009), where negative jumps are rare events and the jump variation is generated through jumps of large size (e.g., the estimate of the mean jump size in Pan (2002) is nearly -20%).

7 Conclusion

In this paper we develop a novel nonparametric method for estimating the risk-neutral tail jump variation from short-dated options, which works in general settings and utilizes all available option data. The method is based on inversion of an estimate of the Laplace transform of the jump variation measure. The latter is constructed using the second derivative of an option portfolio that recovers nonparametrically the (risk-neutral) Laplace transform of the underlying asset return. The separation of the volatility from jumps is achieved by making use of the leading role of the diffusion coefficient in the Laplace transform of returns and its second derivative at high frequencies. Simulation results show the superior performance of the new estimator over existing ones. In an empirical application to S&P 500 option data, we find smaller than previously reported left jump variation that has less correlation with diffusive volatility and is generated by smaller-sized jumps.

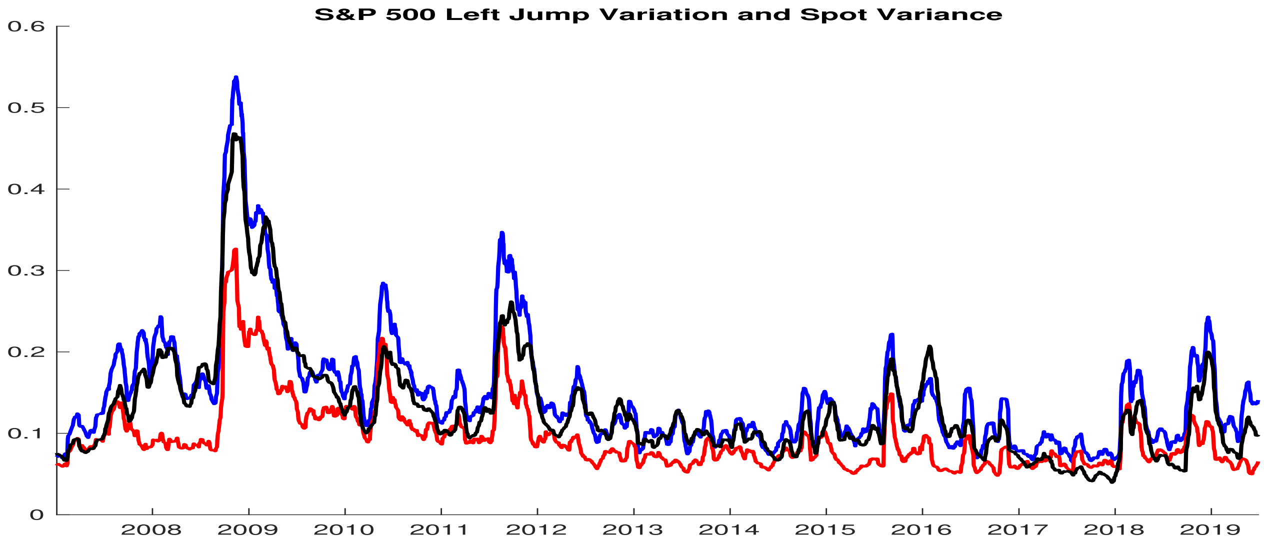

Figure 3: S&P 500 Left Jump Variation Measures and Spot Volatility. Blue line corresponds to $\sqrt{\widehat{NJV}_{t,\mathcal{T}}}$, red line to $\sqrt{\widehat{LV}_{t,\mathcal{T}}(\widehat{\vartheta}_t, \widehat{z}_{\mathcal{T}})}$, and black line to $\sigma_{t,\mathcal{T}}$. The displayed series are 22-day moving averages of the corresponding daily estimates.

8 Proofs

In the proofs, without loss of generality, we will set $t = 0$ and $x_0 = 0$. We will also drop the subscript t in the notation of various quantities, e.g., we will use the shorthand $\mathcal{L}_T(u)$ instead of $\mathcal{L}_{t,T}(u)$, etc.

8.1 Preliminary Estimates and Decompositions

We start with introducing an approximation for the process x in which its spot characteristics are frozen at their values at time zero. For this, we represent the jump part of the process x_t as an integral with respect to a Poisson random measure. In particular, using the so-called Grigelionis representation of the jump part of a semimartingale (Theorem 2.1.2 of Jacod and Protter (2012)), upon suitably extending the probability space, we can write

$$\int_0^t \int_{\mathbb{R}} \widetilde{\mu}(ds, dx) \equiv \int_0^t \int_E \delta^x(s, z) \widetilde{\mu}^x(ds, dz), \tag{62}$$

where $\mu^x(ds, dz)$ is a Poisson measure on $\mathbb{R}_+ \times E$ with compensator $dt \otimes \lambda(dz)$, for some σ-finite measure on E, $\widetilde{\mu}^x$ is the martingale counterpart of μ^x, and δ^x is a predictable and $\mathbb{R}$-valued function on $\Omega \times \mathbb{R}_+ \times E$ such that $\nu_t(z)dz$ is the image of the measure λ under the map $z \to \delta^x(t, z)$ on the set $\{z : \delta^x(\omega, t, z) \neq 0\}$.

With the above notation, we introduce the following approximation of x_t:

$$\widetilde{x}_t = a_0 t + \sigma_0 W_t + \int_0^t \int_E \delta^x(0, z) \widetilde{\mu}^x(du, dz), \quad t \geq 0. \tag{63}$$

The conditional characteristic function of $\widetilde{x}_t$ is given by:

$$\widetilde{\mathcal{L}}_t(u) = \mathbb{E}_0^{\mathbb{Q}}\left(e^{u(\widetilde{x}_t - x_0)}\right) = \exp\left(t\psi_0(u)\right), \tag{64}$$

where the characteristic exponent is given by

$$\psi_0(u) = a_0 u + \frac{u^2}{2}\sigma_0^2 + \int_{\mathbb{R}}(e^{ux} - 1)\nu_0(x)dx. \tag{65}$$

The above result holds provided the following exponential moment exist:

$$\int_{\mathbb{R}} e^{\Re(u)x}\nu_0(x)dx < \infty,$$

and this follows from Theorem 25.17(iii) of Sato (1999).

The out-of-the-money option price corresponding to payoff determined by $\widetilde{x}_T$ is denoted by $\widetilde{O}_T(k)$. That is, $\widetilde{O}_T(k)$ the counterpart of $O_T(k)$ in which the terminal price x_T is replaced with $\widetilde{x}_T$. In the next lemma, we provide bounds for the option prices $O_T(k)$ and $\widetilde{O}_T(k)$.

Lemma 1. *Suppose assumptions A2-A4 hold. For some $\mathcal{F}_0$-adapted positive random variables C_0 and $\bar{t}$, we have for $T < \bar{t}$:*

$$O_T(k) \le C_0 T\left(e^{3k} \wedge e^{-k}\right),\ \widetilde{O}_T(k) \le C_0 T\left(e^{4k} \wedge e^{-2k}\right),\ |k| \ge 1, \tag{66}$$

$$\widetilde{O}_T(k) \le C_0\left(\sqrt{T} \bigwedge \frac{T}{|e^k - 1|}\right), \tag{67}$$

$$\left|O_T(k) - \widetilde{O}_T(k)\right| \le C_0\left(|\log(T)|1_{\{|k|\ge 1\}} + 1\right)T^{3/2}\left(\frac{1}{|k| \vee \sqrt{T}} \bigvee 1\right),\ k \in \mathbb{R}, \tag{68}$$

$$\left|\widetilde{O}_T(k_1) - \widetilde{O}_T(k_2)\right| \le C_0|e^{k_1} - e^{k_2}|e^{-(3-\varepsilon)|k_2|}\left(\frac{T}{|k_2|^2 \vee |k_2|^4} \bigwedge 1\right), \tag{69}$$

for $k_1 < k_2 < 0$ or $k_1 > k_2 > 0$ and some arbitrary small $\varepsilon > 0$.

Proof of Lemma 1. The results follow from Lemmas 1-7 in Qin and Todorov (2019). The result of Lemma 7 in Qin and Todorov (2019) is strengthened by applying Lemma 1 of that paper. □

We next make the following decomposition:

$$\widehat{\mathcal{L}}_T(u) - \widetilde{\mathcal{L}}_T(u) = \widehat{A}_T^{(a)}(u) + \widehat{A}_T^{(b)}(u) + \widehat{A}_T^{(c)}(u) + \widehat{A}_T^{(d)}(u), \tag{70}$$

$$\widehat{A}_T^{(a)}(u) = \sum_{j=2}^{N} f(u, k_{j-1})\epsilon_T(k_{j-1})\Delta_j,\ \widehat{A}_T^{(b)}(u) = \sum_{j=2}^{N} f(u, k_{j-1})(O_T(k_{j-1}) - \widetilde{O}_T(k_{j-1}))\Delta_j, \tag{71}$$

$$\widehat{A}_T^{(c)}(u) = \sum_{j=2}^{N}\int_{k_{j-1}}^{k_j}\left(f(u, k_{j-1})\widetilde{O}_T(k_{j-1}) - f(u,k)\widetilde{O}_T(k)\right)dk, \tag{72}$$

$$\widehat{A}_T^{(d)}(u) = -\int_{-\infty}^{\underline{k}} f(u,k)\widetilde{O}_T(k)dk - \int_{\overline{k}}^{\infty} f(u,k)\widetilde{O}_T(k)dk. \tag{73}$$

In the next lemma, we derive bounds for $\widehat{A}_T^{(a)}(u)$, $\widehat{A}_T^{(b)}(u)$, $\widehat{A}_T^{(c)}(u)$ and $\widehat{A}_T^{(d)}(u)$ and their first two derivatives with respect to u.

Lemma 2. *Suppose assumptions A2-A6 hold. Assume $\Delta \asymp T^{\alpha}$, $\underline{K} \asymp T^{\beta}$ and $\overline{K} \asymp T^{-\beta}$, for $\alpha > \frac{1}{2}$ and $\beta > 0$. Let $-\varsigma < \Re(u) < 0$ for some $\varsigma \in (0, 2)$. For some $\mathcal{F}_0$-adapted positive random variables C_0 and $\bar{t}$ that do not depend on u, we have for $T < \bar{t}$:*

$$\mathbb{E}\left(|\widehat{A}_T^{(a)}(u)|^p \big| \mathcal{F}^{(0)}\right) \leq C_0|u|^{2p}(|\underline{k}| \vee |\overline{k}|)^{\frac{p}{2}-1}\Delta^{\frac{p}{2}}T^{\frac{1+p}{2}}, \ p \geq 2, \tag{74}$$

$$\mathbb{E}\left(|\nabla_u\widehat{A}_T^{(a)}(u)|^p \big| \mathcal{F}^{(0)}\right) \leq C_0|u|^{p}(|\underline{k}| \vee |\overline{k}|)^{\frac{p}{2}-1}\Delta^{\frac{p}{2}}T^{\frac{1+p}{2}}\left(|u|^pT^{\frac{p-1}{2}} \vee 1\right), \ p \geq 2, \tag{75}$$

$$\mathbb{E}\left(|\nabla_{uu}\widehat{A}_T^{(a)}(u)|^p \big| \mathcal{F}^{(0)}\right) \leq C_0(|\underline{k}| \vee |\overline{k}|)^{\frac{p}{2}-1}\Delta^{\frac{p}{2}}T^{\frac{1+p}{2}}\left(1 + (|u|^p + |u|^{2p})T^{\frac{p-1}{2}}\right), \ p \geq 2, \tag{76}$$

$$\begin{aligned} &|\widehat{A}_T^{(b)}(u)| + |\nabla_u\widehat{A}_T^{(b)}(u)| + |\nabla_{uu}\widehat{A}_T^{(b)}(u)| \\ &\qquad \leq C_0|u|^2T^{3/2-\frac{1}{6}(\varsigma+1)}|\log(T)|^4 + 1_{\{\beta \geq 1/6\}}C_0|u|^2T^{1+(2-\varsigma)\beta}|\log(T)|^2, \end{aligned} \tag{77}$$

$$|\widehat{A}_T^{(c)}(u)| \leq C_0|u|^2\Delta\sqrt{T}(|u|\sqrt{T} \vee 1), \ |\nabla_u\widehat{A}_T^{(c)}(u)| \leq C_0|u|\Delta\sqrt{T}(|u|\sqrt{T} \vee 1) + C_0|u|^3\Delta T|\log(T)|, \tag{78}$$

$$|\nabla_{uu}\widehat{A}_T^{(c)}(u)| \leq C_0\Delta\sqrt{T}(|u|\sqrt{T}|\log(T)| \vee 1) + C_0|u|^2\Delta T(|\log(T)| \vee |u|), \tag{79}$$

$$|\widehat{A}_T^{(d)}(u)| + |\nabla_u\widehat{A}_T^{(d)}(u)| + |\nabla_{uu}\widehat{A}_T^{(d)}(u)| \leq C_0|u|^2T^{1+(3-\varsigma)\beta}|\log(T)|^2. \tag{80}$$

Proof of Lemma 2. For the first three inequalities, we make use of the following bound

$$\mathbb{E}\left|\sum_{i=1}^{N}\epsilon_i\right|^p \leq CN^{\frac{p}{2}-1}\mathbb{E}\left(\sum_{i=1}^{N}|\epsilon_i|^p\right), \tag{81}$$

where $p \geq 2$, C is some positive constant, and $\{\epsilon_i\}_{i=1}^N$ is a sequence of independent random variables. The above inequality, follows from Burkholder-Davis-Gundy inequality and inequality in means. Applying this result and the bounds of Lemma 1 as well as our assumption for the observation error, we have the first three bounds of the lemma. The rest of the bounds to be proved follow by application of Lemma 1. □

We finish this section with decomposing $\widehat{h}_T(u)$. We will do so on the following set

$$\Omega_T = \left\{\omega : |\widehat{\mathcal{L}}_T(u)| > \alpha_T \ \cap \ |\mathcal{L}_T(u)| > \alpha_T, \text{ for } u = -c + iz \text{ and } |z| \leq z_T\right\}, \tag{82}$$

where c and z_T are the parameters of the estimator $\widehat{LV}_T(\vartheta, z_T)$, and α_T is a deterministic sequence satisfying the following condition

$$\alpha_T = O_p(|\log(T)|^k), \text{ for some } k > 0. \tag{83}$$

Using second-order Taylor expansion, on the set Ω_T, we have (recall notation in (17)):

$$\widehat{h}_T(u) - h_0(u) = \widehat{Z}_1(u) + \widehat{Z}_2(u) + \widehat{Z}_3(u) + \widehat{R}(u), \text{ for } u = -c + iz \text{ and } |z| \leq z_T, \tag{84}$$

where

$$\widehat{Z}_1(u) = \frac{1}{T}\frac{\nabla_{uu}\widehat{\mathcal{L}}_T(u) - \nabla_{uu}\widetilde{\mathcal{L}}_T(u)}{\widetilde{\mathcal{L}}_T(u)}, \ \widehat{Z}_2(u) = -\frac{2}{T}\frac{\nabla_u\widetilde{\mathcal{L}}_T(u)}{\widetilde{\mathcal{L}}_T(u)^2}\left(\nabla_u\widehat{\mathcal{L}}_T(u) - \nabla_u\widetilde{\mathcal{L}}_T(u)\right), \tag{85}$$

$$\widehat{Z}_3(u) = \frac{1}{T}\left(2\frac{\left(\nabla_u \widetilde{\mathcal{L}}_T(u)\right)^2}{\widetilde{\mathcal{L}}_T(u)^3} - \frac{\nabla_{uu}\widetilde{\mathcal{L}}_T(u)}{\widetilde{\mathcal{L}}_T(u)^2}\right)\left(\widehat{\mathcal{L}}_T(u) - \widetilde{\mathcal{L}}_T(u)\right), \tag{86}$$

and for C denoting a positive constant that does not depend on u, we have

$$|\widehat{R}(u)| \leq \frac{C}{T}(\widehat{R}_1(u) + \widehat{R}_2(u) + \widehat{R}_3(u)), \tag{87}$$

with

$$\widehat{R}_1(u) = \frac{1}{\alpha_T^2}|\nabla_{uu}\widehat{\mathcal{L}}_T(u) - \nabla_{uu}\widetilde{\mathcal{L}}_T(u)||\widehat{\mathcal{L}}_T(u) - \widetilde{\mathcal{L}}_T(u)| + \frac{1}{\alpha_T^2}|\nabla_u\widehat{\mathcal{L}}_T(u) - \nabla_u\widetilde{\mathcal{L}}_T(u)|^2, \tag{88}$$

$$\widehat{R}_2(u) = \frac{1}{\alpha_T^3}|\nabla_u\widehat{\mathcal{L}}_T(u) - \nabla_u\widetilde{\mathcal{L}}_T(u)||\widehat{\mathcal{L}}_T(u) - \widetilde{\mathcal{L}}_T(u)|(|\nabla_u\widehat{\mathcal{L}}_T(u)| + |\nabla_u\widetilde{\mathcal{L}}_T(u)|), \tag{89}$$

$$\widehat{R}_3(u) = \frac{1}{\alpha_T^4}|\widehat{\mathcal{L}}_T(u) - \widetilde{\mathcal{L}}_T(u)|^2\left(|\nabla_u\widehat{\mathcal{L}}_T(u)|^2 + |\nabla_u\widetilde{\mathcal{L}}_T(u)|^2 + |\nabla_{uu}\widehat{\mathcal{L}}_T(u)| + |\nabla_{uu}\widetilde{\mathcal{L}}_T(u)|\right). \tag{90}$$

We next set

$$\widehat{\mathcal{C}}_T(1) = \frac{2}{T}\sum_{j=2}^N(\widehat{O}_T(k_{j-1}) - \widetilde{O}_T(k_{j-1}))\Delta_j,\ \widehat{\mathcal{C}}_T(2) = \frac{2}{T}\sum_{j=2}^N\int_{k_{j-1}}^{k_j}(\widetilde{O}_T(k_{j-1}) - \widetilde{O}_T(k))dk, \tag{91}$$

and denote

$$\overline{Z}_1(u) = \widehat{Z}_1(u) - \widehat{\mathcal{C}}_T(1) - \widehat{\mathcal{C}}_T(2),\ \overline{h}_T(u) = \overline{Z}_1(u) + \widehat{Z}_2(u) + \widehat{Z}_3(u) + \widehat{R}(u). \tag{92}$$

We note that

$$\begin{aligned}\widehat{h}_T(-c-iz) - \mathcal{F}\nu_0(-c-iz) - \widehat{\sigma}_T^2(u_T) &= \overline{h}_T(-c-iz) - \widehat{\sigma}_T^2(u_T) \\ &+ \widehat{\mathcal{C}}_T(1) + \widehat{\mathcal{C}}_T(2) + \sigma_0^2, \quad \text{for } |z| \leq z_T.\end{aligned} \tag{93}$$

8.2 Proof of Theorem 1

In proving the theorem, we will make use of the following basic result:

Lemma 3. *For $a \geq 0$, $b > 0$ and $c > 0$, we have*

$$\left|\int_{-b}^{+b}\frac{e^{-iaz}}{c+iz}dz\right| \leq C\left(\left(ab\bigwedge\frac{1}{ab}\right)\bigvee 1\right), \tag{94}$$

for some constant $C > 0$ that does not depend on a, b and c.

Proof of Lemma 3. For $a \geq 0$, using the properties of the trigonometric functions, we have,

$$\int_{-b}^{+b}\frac{e^{-iaz}}{c+iz}dz = 2\int_0^b\frac{c\cos(az) - z\sin(az)}{c^2+z^2}dz. \tag{95}$$

Making use of $\int_0^\infty \frac{1}{c^2+z^2}dz \leq \int_0^c \frac{1}{c^2}dz + \int_c^\infty \frac{1}{z^2}dz = \frac{2}{c}$, we have

$$\left|\int_0^b\frac{c\cos(az)}{c^2+z^2}dz\right| \leq 2. \tag{96}$$

Using $|\sin(x)| \le |x|$, we have

$$\left|\int_0^b \frac{z\sin(az)}{c^2+z^2}dz\right| \le ab. \tag{97}$$

For $a > 0$, we have

$$\int_0^b \frac{z\sin(az)}{c^2+z^2}dz = \int_0^b \frac{\cos(az)}{a}\frac{c^2-z^2}{(c^2+z^2)^2}dz - \frac{b\cos(ab)}{a(c^2+b^2)}. \tag{98}$$

Therefore, altogether, we have

$$\left|\int_0^b \frac{z\sin(az)}{c^2+z^2}dz\right| \le C\left(ab\bigwedge\frac{1}{ab}\right). \tag{99}$$

Combining the above bounds, we get the result in (94). □

We proceed with proving the result of the theorem. First, we note that $LV_0(\vartheta)$ is well defined due to the integrability assumptions in A3 and the restriction of $c < \varsigma < 2$ imposed in the statement of the theorem. Next, using Lemma 2 (and the rate condition $\alpha > \frac{1}{2}$ imposed in the statement of the theorem), we have

$$\mathbb{P}\left(\Omega_T^c\right) \to 0. \tag{100}$$

Therefore, it suffices to look at the difference $\widehat{LV}_T(\vartheta, z_T, u_T) - LV_0(\vartheta)$ on the set Ω_T. On this set, using the algebraic inequality $(x+y)^2 \le 2x^2 + 2y^2$, for $x, y \in \mathbb{R}$, we have (note that $\vartheta \ge 0$ and $c > 0$):

$$\begin{aligned}&\int_0^\infty \left(\widehat{LV}_T(\vartheta, z_T, u_T) - LV_0(\vartheta)\right)^2 d\vartheta \\ &\quad\le \frac{1}{2\pi^2}\int_0^\infty\left(\int_{-\infty}^\infty e^{-i\vartheta z}\widehat{G}(z)dz\right)^2 d\vartheta + \frac{1}{2\pi^2}\int_0^\infty \widehat{H}^2(\vartheta)d\vartheta,\end{aligned} \tag{101}$$

where

$$\widehat{G}(z) = \begin{cases} \frac{1}{c+iz}\overline{h}_T(-c-iz), & \text{if } |z| \le z_T, \\ -\frac{\mathcal{F}\nu_0(-c-iz)}{c+iz}, & \text{if } |z| > z_T, \end{cases} \tag{102}$$

and

$$\widehat{H}(\vartheta) = \int_{-z_T}^{z_T} \frac{e^{-\vartheta(c+iv)}}{c+iv}dv\left(\widehat{\sigma}_T^2(u_T) - \widehat{C}_T(1) - \widehat{C}_T(2) - \sigma_0^2\right), \tag{103}$$

and we note that $\int_{-\infty}^\infty e^{-i\vartheta z}\widehat{G}(z)dz$ is real-valued, using the properties of holomorphic functions. Using Lemma 3, we have

$$\int_0^\infty \widehat{H}^2(\vartheta)d\vartheta \le C_0\left(\widehat{\sigma}_T^2(u_T) - \widehat{C}_T(1) - \widehat{C}_T(2) - \sigma_0^2\right)^2, \tag{104}$$

for some positive $\mathcal{F}_0$-adapted random variable C_0. By application of Parseval's equality, we have

$$\int_0^\infty\left(\int_{-\infty}^\infty e^{-i\vartheta z}\widehat{G}(z)dz\right)^2 d\vartheta \le \int_{-\infty}^\infty\left(\int_{-\infty}^\infty e^{-i\vartheta z}\widehat{G}(z)dz\right)^2 d\vartheta = 2\pi\int_{-\infty}^\infty |\widehat{G}(z)|^2 dz. \tag{105}$$

From here, using the notation of the previous section and (17), we have

$$\int_0^\infty \left(\widehat{LV}_T(\vartheta, z_T, u_T) - LV_0(\vartheta)\right)^2 d\vartheta \le C_0 \sum_{i=1}^6 \mathcal{W}_i, \tag{106}$$

where

$$\mathcal{W}_1 = \int_{|z|>z_T} \frac{1}{|c+iz|^2} |\mathcal{F}\nu_0(-c-iz)|^2 dz,\ \mathcal{W}_2 = \left(\widehat{\sigma}_T^2(u_T) - \widehat{\mathcal{C}}_T(1) - \widehat{\mathcal{C}}_T(2) - \sigma_0^2\right)^2, \tag{107}$$

$$\mathcal{W}_3 = \int_{-z_T}^{z_T} \frac{1}{|c+iz|^2} \left|\overline{Z}_1(-c-iz)\right|^2 dz,\ \mathcal{W}_j = \int_{-z_T}^{z_T} \frac{1}{|c+iz|^2} \left|\widehat{Z}_{j-2}(-c-iz)\right|^2 dz,\ j = 4, 5, \tag{108}$$

$$\mathcal{W}_6 = \int_{-z_T}^{z_T} \frac{1}{|c+iz|^2} \left|\widehat{R}(-c-iz)\right|^2 dz. \tag{109}$$

From assumption A, we have

$$\mathcal{W}_1 = O_p(z_T^{-2r-2}). \tag{110}$$

We proceed with a sequence of lemmas that provide bounds for the terms $\{\mathcal{W}_j\}_{j=2,3,4,5,6}$.

Lemma 4. *Suppose assumptions A2-A6 hold. Assume $\Delta \asymp T^\alpha$, $\underline{K} \asymp T^\beta$ and $\overline{K} \asymp T^{-\beta}$, for $\alpha > \frac{1}{2}$ and $\beta > 0$. Further, let the sequence z_T satisfy*

$$z_T \to \infty,\ \text{and}\ z_T^2 T |\log(T)|^2 \to 0. \tag{111}$$

We have

$$\begin{aligned} \alpha_T^8 \mathcal{W}_6 = O_p\big(& z_T^{15} \Delta^4 T^{\frac{5}{2}} |\log(T)|^3 \vee z_T^7 \Delta^2 \sqrt{T} |\log(T)| \vee z_T^7 T^{4-\frac{2}{3}(\varsigma+1)} |\log(T)|^{16} \\ & \vee z_T^7 T^{2+4(2-\varsigma)\beta} |\log(T)|^8\big). \end{aligned} \tag{112}$$

Proof of Lemma 4. Using the algebraic inequalities $2xy \le x^2 + y^2$ and $3x^2 y \le 4(x^3 + y^3)$ for two nonnegative real numbers x and y, we have

$$\begin{aligned} |\widehat{R}(u)| \le \frac{C}{T\alpha_T^4} \big[& h_{24}(|\widehat{\mathcal{L}}_T(u) - \widetilde{\mathcal{L}}_T(u)|) + h_{24}(|\nabla_u \widehat{\mathcal{L}}_T(u) - \nabla_u \widetilde{\mathcal{L}}_T(u)|) \\ & + h_{23}(|\nabla_{uu} \widehat{\mathcal{L}}_T(u) - \nabla_{uu} \widetilde{\mathcal{L}}_T(u)|)\big], \quad \text{for } |u|\sqrt{T} \le 1, \end{aligned} \tag{113}$$

where we denote the functions $h_{23}(x) = x^2 \vee x^3$ and $h_{24}(x) = x^2 \vee x^4$, and the constant $C > 0$ does not depend on u. From here, the result to be proved follows by an application of Lemma 2. □

Lemma 5. *Suppose assumptions A2-A6 hold. Assume $\Delta \asymp T^\alpha$, $\underline{K} \asymp T^\beta$ and $\overline{K} \asymp T^{-\beta}$, for $\alpha > \frac{1}{2}$ and $\beta > 0$. Further, let the sequence z_T satisfy the rate condition in (30). We then have*

$$\begin{cases} \mathcal{W}_3 = O_p\left(\Delta z_T^3 + \gamma(T, z_T)\right), \\ \mathcal{W}_4 = O_p\left(T^{3/2} \Delta z_T^3 + T^2 \Delta z_T^5 + T\Delta^2 z_T^3 + |\log(T)|^2 T \Delta^2 z_T^7 + T^2 z_T^2 \gamma(T, z_T)\right), \\ \mathcal{W}_5 = O_p\left(T^{3/2} \Delta z_T^3 + T^2 \gamma(T, z_T)\right), \end{cases} \tag{114}$$

where we denote

$$\gamma(T, z_T) = |\log(T)|^8 T^{1-\frac{1}{3}(\varsigma+1)} z_T^3 + |\log(T)|^4 T^{2(2-\varsigma)\beta} z_T^3. \tag{115}$$

Proof of Lemma 5. For the first term, we can make the decomposition

$$\left|\frac{\nabla_{uu}\widehat{A}_T^{(c)}(-c-iz)}{\widetilde{\mathcal{L}}_T(-c-iz)} - \mathfrak{a}_T(-c-iz) - \widehat{\mathcal{A}}_T^{(1)}(z)\right| \leq \widehat{\mathcal{A}}_T^{(2)}(z), \tag{116}$$

where

$$\mathfrak{a}_T(u) = \sum_{j=2}^{N}\left[\nabla_{uu}f(u,k_{j-1})\int_{k_{j-1}}^{k_j}(\widetilde{O}_T(k_{j-1}) - \widetilde{O}_T(k))dk\right], \tag{117}$$

and for T sufficiently small (so that $|\widetilde{\mathcal{L}}_T(-c-iz)|$ is bounded by a constant from below for $|z| \leq z_T$) and C_0 some $\mathcal{F}_0$-adapted random variable, we denote:

$$\widehat{\mathcal{A}}_T^{(1)}(z) = \sum_{j=2}^{N}\int_{k_{j-1}}^{k_j}(\nabla_{uu}f(-c-iz,k_{j-1}) - \nabla_{uu}f(-c-iz,k))\widetilde{O}_T(k)dk, \tag{118}$$

$$\widehat{\mathcal{A}}_T^{(2)}(z) = C_0 z_T^2 T|\mathfrak{a}_T(-c-iz) + \widehat{\mathcal{A}}_T^{(1)}(z)|. \tag{119}$$

We have

$$\nabla_{uu}f(u,k) = 2e^{(u-1)k} + 2(2u-1)ke^{(u-1)k} + (u^2-u)k^2e^{(u-1)k}, \tag{120}$$

and from here by Taylor expansion, for $u = -c-iz$, we have

$$\begin{aligned}|\nabla_{uu}f(u,k_2) - \nabla_{uu}f(u,k_1)| &\leq C|k_2-k_1|e^{-(1+c)(k_1\wedge k_2)} \\ &\quad\times\left(|z|\vee 1 + (z^2\vee 1)(|k_1|\vee|k_2|) + (|z|^3\vee 1)(|k_1|\vee|k_2|)^2\right),\end{aligned} \tag{121}$$

for two real k_1 and k_2 and some positive constant $C>0$ that depends on c only. Using this bound, Lemma 1 and the fact that $z_T^2 T \to 0$, we have

$$\int_{-z_T}^{z_T}\frac{1}{|c+iz|^2}\left|\widehat{\mathcal{A}}_T^{(1)}(z)\right|^2 dz = O_p(z_T^5 T^2\Delta^2), \tag{122}$$

$$\int_{-z_T}^{z_T}\frac{1}{|c+iz|^2}\left|\widehat{\mathcal{A}}_T^{(2)}(z)\right|^2 dz = O_p\left(z_T^3\Delta^2T^3 + z_T^7\Delta^2T^3\right). \tag{123}$$

Using Taylor expansion and the fact that $\Re(u) > -\varsigma$, we have

$$\left|\nabla_{uu}f(u,k) - 2 - 2k(2u-1)e^{(u-1)k} - (u^2-u)e^{(u-1)k}k^2\right| \leq C\left(e^{-(\varsigma+1)k}\vee 1\right)(|k| + |k||u|), \tag{124}$$

for some positive constant C that does not depend on k and u. This inequality and an application of Lemma 1 yield (note the cancelation in the loading of options with k near zero in the difference $\mathfrak{a}_T(-c-iz) - 2\sum_{j=2}^{N}\int_{k_{j-1}}^{k_j}(\widetilde{O}_T(k_{j-1}) - \widetilde{O}_T(k))dk$):

$$\int_{-z_T}^{z_T}\frac{\left|\mathfrak{a}_T(-c-iz) - 2\sum_{j=2}^{N}(\int_{k_{j-1}}^{k_j}\widetilde{O}_T(k_{j-1}) - \widetilde{O}_T(k))dk\right|^2}{|c+iz|^2}dz = O_p\left(z_T^3\Delta^2T^2\right), \tag{125}$$

Similarly, we have

$$\left| \frac{\nabla_{uu}\widehat{A}_T^{(a)}(-c-iz)}{\widetilde{\mathcal{L}}_T(-c-iz)} - \mathfrak{b}_T(-c-iz) \right| \leq \widehat{\mathcal{B}}_T(z), \tag{126}$$

where

$$\mathfrak{b}_T(u) = \sum_{j=2}^{N} \nabla_{uu} f(u, k_{j-1})\epsilon_{j-1}\Delta_j, \tag{127}$$

and

$$\widehat{\mathcal{B}}_T(z) = C_0 z_T^2 T \left| \sum_{j=2}^{N} \nabla_{uu} f(-c-iz, k_{j-1})\epsilon_{j-1}\Delta_j \right|. \tag{128}$$

Using Lemma 1 and $z_T^2 T \to 0$, we have

$$\int_{-z_T}^{z_T} \frac{1}{|c+iz|^2} \left| \widehat{\mathcal{B}}_T(z) \right|^2 dz = O_p(z_T^5 \Delta T^3), \tag{129}$$

Next, the bound in (124) and an application of Lemma 1 yield (note the cancelation in the loading of options with k near zero in the difference $\mathfrak{b}_T(-c-iz) - 2\sum_{j=2}^{N} \epsilon_{j-1}\Delta_j$):

$$\int_{-z_T}^{z_T} \frac{1}{|c+iz|^2} \left(\mathfrak{b}_T(-c-iz) - 2\sum_{j=2}^{N} \epsilon_{j-1}\Delta_j \right)^2 dz = O_p\left(z_T^3 \Delta T^2 \right). \tag{130}$$

From here, to derive the bound for $\mathcal{W}_3$, we can apply Lemma 2 for the bounds involving $\nabla_{uu}\widehat{A}_T^{(b)}(u)$ and $\nabla_{uu}\widehat{A}_T^{(d)}(u)$, and also use the fact that from the rate condition in (30), we have $z_T^2 T^{1/3} \to 0$ and since $\Delta/\sqrt{T} \to 0$ (assumed in the theorem), we have

$$\frac{z_T^5 \Delta^2 |\log(T)|^2}{z_T^3 \Delta} = z_n^2 \Delta |\log(T)|^2 = z_n^2 \sqrt{T} |\log(T)|^2 \times \frac{\Delta}{\sqrt{T}} = o(1). \tag{131}$$

The proof of the results for $\mathcal{W}_4$ and $\mathcal{W}_5$ makes use of the following bounds

$$|\nabla_u \widetilde{\mathcal{L}}_T(u)| \leq C_0 z_T T, \;\; |\nabla_{uu} \widetilde{\mathcal{L}}_T(u)| \leq C_0 T, \tag{132}$$

for $u = -c - iz$ and $|z| \leq z_T$, with $z_T^2 T \leq C_0$, and some positive constant $C_0 > 0$. From here, the results for $\mathcal{W}_4$ and $\mathcal{W}_5$ can be shown by making use of Lemma 2. □

Lemma 6. *Suppose assumptions A2-A6 hold. Assume $\Delta \asymp T^{\alpha}$, $\underline{K} \asymp T^{\beta}$ and $\overline{K} \asymp T^{-\beta}$, for $\alpha > \frac{1}{2}$ and $\beta > 0$. Further, let the sequence u_T satisfy the rate condition*

$$u_T \to \infty \text{ and } u_T^2 T \to 0. \tag{133}$$

We then have

$$\begin{aligned} &\widehat{\sigma}_T^2(u_T) - \widehat{\mathcal{C}}_T(1) - \widehat{\mathcal{C}}_T(2) - \sigma_0^2 \\ &\quad = O_p\left(u_T^{-(2-r')} \bigvee T^{1/4} u_T \sqrt{\Delta} \bigvee \frac{\sqrt{\Delta}}{T^{1/4} u_T} \bigvee |\log(T)|^4 T^{\frac{1}{2} - \frac{1}{6}(\varsigma+1)} \bigvee |\log(T)|^2 T^{(2-\varsigma)\beta} \right), \end{aligned} \tag{134}$$

where r' is the jump activity index appearing in the condition stated in equation (29).

Proof of Lemma 6. Using the bounds in Lemma 2, it is easy to check that $\widehat{\mathcal{L}}_T(iu_T)$ converges in probability to 1. Therefore, it suffices to work only on the set on which $|\widehat{\mathcal{L}}_T(iu_T)| > 1/2$. On this set, using Taylor expansion, we have

$$\left|\widehat{\sigma}_T^2(u_T) - \widehat{\mathcal{C}}_T(1) - \widehat{\mathcal{C}}_T(2) - \sigma_0^2\right| \leq \widehat{\Sigma}_T^{(1)} + \widehat{\Sigma}_T^{(2)} + \widehat{R}_T, \tag{135}$$

where

$$\widehat{\Sigma}_T^{(1)} = -\frac{2}{Tu_T^2}\psi_0(iu_T) - \sigma_0^2,\ \widehat{\Sigma}_T^{(2)} = -\frac{2}{Tu_T^2}(\widehat{\mathcal{L}}_T(iu_T) - \widetilde{\mathcal{L}}_T(iu_T)) - \widehat{\mathcal{C}}_T(1) - \widehat{\mathcal{C}}_T(2), \tag{136}$$

$$\widehat{R}_T = \frac{C_0}{Tu_T^2}\left|\widehat{\mathcal{L}}_T(iu_T) - \widetilde{\mathcal{L}}_T(iu_T)\right|^2 + C_0\left|\widehat{\mathcal{L}}_T(iu_T) - \widetilde{\mathcal{L}}_T(iu_T)\right|, \tag{137}$$

and C_0 is $\mathcal{F}_0$-adapted random variable. Using the finite variation assumption for the jumps, we have

$$\widehat{\Sigma}_T^{(1)} = O_p(u_T^{-(2-r')}). \tag{138}$$

Taking into account the fact that $\sqrt{\Delta}/T \to 0$ and using Lemma 2, we have

$$\widehat{R}_T = O_p\left(u_T^2\sqrt{\Delta}T^{\frac{3}{4}} \bigvee u_T^2|\log(T)|^4T^{\frac{3}{2}-\frac{1}{6}(\varsigma+1)} \bigvee u_T^2|\log(T)|^2T^{1+(2-\varsigma)\beta}\right). \tag{139}$$

Using Lemma 2, we have

$$\widehat{\Sigma}_T^{(2)} = \overline{A}_T^{(a)} + \overline{A}_T^{(c)} + O_p\left(|\log(T)|^4T^{\frac{1}{2}-\frac{1}{6}(\varsigma+1)} \bigvee |\log(T)|^2T^{(2-\varsigma)\beta}\right), \tag{140}$$

where we denote

$$\overline{A}_T^{(a)} = -\frac{2}{Tu_T^2}\widehat{A}_T^{(a)}(iu_T) - \frac{2}{T}\sum_{j=2}^{N}\epsilon_T(k_{j-1})\Delta_j, \tag{141}$$

$$\overline{A}_T^{(c)} = -\frac{2}{Tu_T^2}\widehat{A}_T^{(c)}(iu_T) - \frac{2}{T}\sum_{j=2}^{N}\int_{k_{j-1}}^{k_j}(\widetilde{O}_T(k_{j-1}) - \widetilde{O}_T(k))dk. \tag{142}$$

Using Lemma 1 and assumption A6 for the option observation error, we have

$$\overline{A}_T^{(a)} = O_p\left(T^{1/4}u_T\sqrt{\Delta} \bigvee \frac{\sqrt{\Delta}}{T^{1/4}u_T}\right). \tag{143}$$

Using Lemma 1 again, we have

$$\overline{A}_T^{(c)} = O_p\left(u_T\Delta \bigvee \frac{\Delta}{\sqrt{T}u_T}\right). \tag{144}$$

Note that $\overline{A}_T^{(c)}$ is of smaller order of magnitude than $\overline{A}_T^{(a)}$. Combining the above bounds, and using $\sqrt{\Delta}/T \to 0$, we get the result of the lemma. □

The proof the theorem now readily follows by combining the bound for $\mathcal{W}_1$ derived above, Lemmas 4, 5 and 6, and using the rate conditions on Δ, T, $\underline{k}$ and $\overline{k}$.

8.3 Proof of Theorem 2

In bounding $\int_0^\infty \left(\widehat{LV}_T(\vartheta, z_T, \kappa_T, \overline{\kappa}_T) - LV_0(\vartheta)\right)^2 d\vartheta$, we replace $\mathcal{W}_2$ term with the following ones:

$$\mathcal{W}_2' = \left| \frac{1}{\kappa_T z_T} \int_{|z|\in[\kappa_T z_T, \overline{\kappa}_T z_T]} \mathcal{F}\nu_0(-c-iz)dz \right|^2, \; \mathcal{W}_3' = \left| \frac{1}{\kappa_T z_T} \int_{|z|\in[\kappa_T z_T, \overline{\kappa}_T z_T]} \overline{Z}_1(-c-iz)dz \right|^2, \tag{145}$$

$$\mathcal{W}_j' = \frac{1}{\kappa_T z_T} \int_{|z|\in[\kappa_T z_T, \overline{\kappa}_T z_T]} \left|\widehat{Z}_{j-2}(-c-iz)\right|^2 dz \text{ and } \mathcal{W}_6' = \frac{1}{\kappa_T z_T} \int_{|z|\in[\kappa_T z_T, \overline{\kappa}_T z_T]} \left|\widehat{R}(-c-iz)\right|^2 dz, \tag{146}$$

for $j = 4, 5$. Using Cauchy–Schwarz inequality and assumption A1, we have

$$\mathcal{W}_2' = O_p(\kappa_T^{-2r-1} z_T^{-2r-1}). \tag{147}$$

For the term $\mathcal{W}_3'$, using Lemma 2, we first have

$$\begin{aligned} &\left| \frac{1}{T\kappa_T z_T} \int_{|z|\in[\kappa_T z_T, \overline{\kappa}_T z_T]} (\nabla_{uu}\widehat{\mathcal{L}}(-c-iz) - \nabla_{uu}\widetilde{\mathcal{L}}(-c-iz)) \left(\frac{1}{\widetilde{\mathcal{L}}(-c-iz)} - 1 \right) dz \right|^2 \\ &= O_p\Bigg(\Delta T^{3/2}(\kappa_T z_T)^4 + \Delta T^2 (\kappa_T z_T)^8 + \Delta^2 T (\kappa_T z_T)^4 \\ &\quad + (\kappa_T z_T)^8 \left(T^{3-\frac{1}{3}(\varsigma+1)} |\log(T)|^8 + T^{2+2(2-\varsigma)\beta} |\log(T)|^4 \right) \Bigg). \end{aligned} \tag{148}$$

From here, it is easy to check that this term is of smaller order than $\sum_{j=3}^6 |\mathcal{W}_j|$ when $z_T^3 \Delta \to 0$, $\kappa_T^8 z_T^2 \sqrt{T} |\log(T)| \to 0$ and $z_T^3 T^{1-\frac{1}{3}(\varsigma+1)} |\log(T)|^8 \to 0$. Therefore, to bound $\mathcal{W}_3'$, it suffices to derive bounds for

$$\left| \frac{1}{T\kappa_T z_T} \int_{|z|\in[\kappa_T z_T, \overline{\kappa}_T z_T]} \left(\sum_{j=2}^N (\nabla_{uu} f(-c-iz, k_{j-1}) - 2)(\widehat{O}_T(k_{j-1}) - \widetilde{O}_T(k_{j-1}))\Delta_j \right) dz \right|^2, \tag{149}$$

and

$$\left| \frac{1}{T\kappa_T z_T} \int_{|z|\in[\kappa_T z_T, \overline{\kappa}_T z_T]} \left(\sum_{j=b,c,d} \nabla_{uu} \widehat{A}_T^{(j)}(-c-iz) - \widehat{\mathcal{C}}_T(2) \right) dz \right|^2. \tag{150}$$

For deriving bounds for these terms, we make use of the following result which can be shown using integration by parts:

$$\int_{u_1}^{u_2} u^2 e^{iuk} du = \frac{u_2^2 e^{iu_2 k} - u_1^2 e^{iu_1 k}}{ik} - 2\int_{u_1}^{u_2} u \frac{e^{iuk}}{k} du, \; u_1, \, u_2 \in \mathbb{R}, \; k \neq 0. \tag{151}$$

Using this result, we also have

$$\begin{aligned} &\left| \int_{|z|\in[\kappa_T z_T, \overline{\kappa}_T z_T]} \left(z^2 k_2^2 e^{(-iz-c-1)k_2} - z^2 k_1^2 e^{(-iz-c-1)k_1} \right) dz \right| \\ &\leq C|k_2 - k_1| e^{-(1+c)(k_1 \vee k_2)} (\kappa_T^2 z_T^2 + (|k_2| \vee |k_1|)\kappa_T^3 z_T^3), \quad k_1, \, k_2 \in \mathbb{R}, \end{aligned} \tag{152}$$

for some constant C that does not depend on z_T, κ_T, $\overline{\kappa}_T$ and c. From here, to derive bounds for the terms in (149) and (150), we follow similar steps as for the analysis of $\mathcal{W}_3$ in the proof of Lemma 5. Upon using $z_T^3\Delta \to 0$, $\kappa_T^8 z_T^2\sqrt{T}|\log(T)| \to 0$ and $z_T^3 T^{1-\frac{1}{3}(\varsigma+1)}|\log(T)|^8 \to 0$, we then have the following bound:

$$\mathcal{W}_3' = O_p\left((\kappa_T^2 z_T^2 \vee z_T^3)\left(\Delta + |\log(T)|^8 T^{1-\frac{1}{3}(\varsigma+1)} + |\log(T)|^4 T^{2(2-\varsigma)\beta}\right)\right). \tag{153}$$

Next, exactly as Lemma 4, we have

$$\begin{aligned}\alpha_T^8\mathcal{W}_6' = O_p\big(&\kappa_T^{16} z_T^{16}\Delta^4 T^{\frac{5}{2}}|\log(T)|^3 \vee \kappa_T^8 z_T^8\Delta^2\sqrt{T}|\log(T)| \vee \kappa_T^8 z_T^8 T^{4-\frac{2}{3}(\varsigma+1)}|\log(T)|^{16} \\ &\vee \kappa_T^8 z_T^8 T^{2+4(2-\varsigma)\beta}|\log(T)|^8\big).\end{aligned} \tag{154}$$

Upon using the fact that $z_T^3\Delta \to 0$, $\kappa_T^8 z_T^2\sqrt{T}|\log(T)| \to 0$ and $z_T^3 T^{1-\frac{1}{3}(\varsigma+1)}|\log(T)|^8 \to 0$, we have

$$\frac{\kappa_T^{16} z_T^{16}\Delta^4 T^{\frac{5}{2}}|\log(T)|^3}{z_T^3\Delta} = O\left(\kappa_T^8 z_T^2\sqrt{T}\right), \; \frac{\kappa_T^8 z_T^8\Delta^2\sqrt{T}|\log(T)|}{z_T^3\Delta} = O\left(\kappa_T^8 z_T^2\sqrt{T}|\log(T)|\right), \tag{155}$$

$$\frac{\kappa_T^8 z_T^8 T^{4-\frac{2}{3}(\varsigma+1)}|\log(T)|^{16}}{z_T^3 T^{1-\frac{1}{3}(\varsigma+1)}|\log(T)|^8} = O\left(\kappa_T^8 z_T^2\sqrt{T}\right), \; \frac{\kappa_T^8 z_T^8 T^{2+4(2-\varsigma)\beta}|\log(T)|^8}{z_T^3 T^{2(2-\varsigma)\beta}|\log(T)|^4} = O\left(\kappa_T^8 z_T^2\sqrt{T}\right). \tag{156}$$

Therefore, given the definition of α_T, $\mathcal{W}_6'$ is dominated asymptotically by $\sum_{j=3}^6|\mathcal{W}_j|$. Next, exactly as in the proof of Lemma 5, we have

$$\begin{cases}\mathcal{W}_4' = O_p\left(T^{3/2}\Delta\kappa_T^4 z_T^4 + T^2\Delta\kappa_T^6 z_T^6 + T\Delta^2\kappa_T^4 z_T^4 + |\log(T)|^2 T\Delta^2\kappa_T^8 z_T^8 + T^2\kappa_T^2 z_T^2\widetilde{\gamma}(T,\kappa_T z_T)\right), \\ \mathcal{W}_5' = O_p\left(T^{3/2}\Delta\kappa_T^4 z_T^4 + T^2\widetilde{\gamma}(T,\kappa_T z_T)\right),\end{cases} \tag{157}$$

where we now denote

$$\widetilde{\gamma}(T,z) = |\log(T)|^8 T^{1-\frac{1}{3}(\varsigma+1)} z^4 + |\log(T)|^4 T^{2(2-\varsigma)\beta} z^4. \tag{158}$$

Using again $z_T^3\Delta \to 0$, $\kappa_T^8 z_T^2\sqrt{T}|\log(T)| \to 0$ and $z_T^3 T^{1-\frac{1}{3}(\varsigma+1)}|\log(T)|^8 \to 0$, we have

$$\frac{T^{3/2}\Delta\kappa_T^4 z_T^4 + T^2\Delta\kappa_T^6 z_T^6 + T\Delta^2\kappa_T^4 z_T^4 + |\log(T)|^2 T\Delta^2\kappa_T^8 z_T^8}{z_T^3\Delta} = O\left((\kappa_T^8 z_T^2\sqrt{T})^{\iota}\right), \tag{159}$$

for some $\iota > 0$ and we further have $(T^2\kappa_T^6 z_T^6)^{8/6} = o(\kappa_T^8 z_T^2\sqrt{T})$. This shows that $\mathcal{W}_4' + \mathcal{W}_5'$ is dominated asymptotically by $\sum_{j=3}^6|\mathcal{W}_j|$.

Combining the above bounds for $\{\mathcal{W}_j'\}_{j=2,\ldots,6}$ with those for $\mathcal{W}_1$ and $\{\mathcal{W}_j\}_{j=3,\ldots,6}$, we get the result of the theorem.

8.4 Proof of Theorem 3

Using Lemma 3, we have

$$\begin{aligned}|\widehat{LV}_T(\vartheta, z_T, u_T) - LV_0(\vartheta)| \leq C_0 \int_{|z|>z_T} \frac{1}{|c+iz|} |\mathcal{F}\nu_0(-c-iz)| dz \\ + C_0 \left| \widehat{\sigma}_T^2(u_T) - \widehat{\mathcal{C}}_T(1) - \widehat{\mathcal{C}}_T(2) - \sigma_0^2 \right| + C_0 \left| \int_{-z_T}^{z_T} \frac{e^{-i\vartheta z}}{c+iz} \overline{h}_T(-c-iz) dz \right|.\end{aligned} \tag{160}$$

Using assumption A1, we have

$$\left(\int_{|z|>z_T} \frac{1}{|c+iz|} |\mathcal{F}\nu_0(-c-iz)| dz \right)^2 = O_p(z_T^{-2r-1}). \tag{161}$$

Next, the bound for $\widehat{\sigma}_T^2(u_T) - \widehat{\mathcal{C}}_T(1) - \widehat{\mathcal{C}}_T(2) - \sigma_0^2$ follows from Lemma 6. We proceed with $\int_{|z|\leq z_T} \frac{e^{-i\vartheta z}}{c+iz} \overline{Z}_1(-c-iz)dz$. First, as in the proof of Theorem 2, we have

$$\begin{aligned}&\left| \frac{1}{T} \int_{-z_T}^{z_T} \frac{e^{-i\vartheta z}}{c+iz} (\nabla_{uu}\widehat{\mathcal{L}}(-c-iz) - \nabla_{uu}\widetilde{\mathcal{L}}(-c-iz)) \left(\frac{1}{\widetilde{\mathcal{L}}(-c-iz)} - 1 \right) dz \right|^2 \\ &= O_p\Bigg(\Delta T^{3/2} z_T^4 + \Delta T^2 z_T^8 + \Delta^2 T z_T^4 + z_T^8 \left(T^{3-\frac{1}{3}(\varsigma+1)} |\log(T)|^8 + T^{2+2(2-\varsigma)\beta} |\log(T)|^4 \right) \Bigg) \\ &= o_p\Bigg(\Delta z_T^2 + z_T^2 \left(T^{1-\frac{1}{3}(\varsigma+1)} |\log(T)|^8 + T^{2(2-\varsigma)\beta} |\log(T)|^4 \right) \Bigg).\end{aligned} \tag{162}$$

To proceed further, we make use of the following algebraic inequality, which can be shown using integration by parts and elementary trigonometric function inequalities:

$$\left| \int_{-b}^{b} e^{-iaz} z dz \right| \leq 8b \left(|a|b^2 \bigwedge \frac{1}{|a|} \right), \quad \text{for } b > 0 \text{ and } a \in \mathbb{R}. \tag{163}$$

Using this result, the bounds in Lemma 1 as well as the $\mathcal{F}^{(0)}$-conditional independence of the observation errors and the fact that $\vartheta \neq 0$ (note that $z_T\Delta \to 0$ and also that $|k+\vartheta| \leq 1/z_T$ implies $k < -\vartheta/2$ for Δ sufficiently small), we have:

$$\left| \frac{1}{T} \int_{-z_T}^{z_T} \frac{e^{-i\vartheta z}}{c+iz} \sum_{j=2}^{N} \left(\nabla_{uu} f(-c-iz, k_{j-1}) - 2 \right) \epsilon_{j-1} \Delta_j dz \right|^2 = O_p(z_T^3 \Delta). \tag{164}$$

Using inequality in means and the proof of Lemma 5 as well as (163) for $\nabla_{uu}\widehat{A}_T^{(b)}(-c-iz)$ and $\nabla_{uu}\widehat{A}_T^{(d)}(-c-iz)$ and the fact that $\vartheta \neq 0$, we have

$$\begin{aligned}&\left| \frac{1}{T} \int_{-z_T}^{z_T} \frac{e^{-i\vartheta z}}{c+iz} \left(\sum_{j=b,c,d} \nabla_{uu}\widehat{A}_T^{(j)}(-c-iz) - \widehat{\mathcal{C}}_T(2) \right) dz \right|^2 \\ &= O_p(z_T^6\Delta^2 + z_T^4\Delta^2 T + z_T^8\Delta^2 T + z_T^4\Delta^2 + z_T^3 \left(T^{1-\frac{1}{3}(\varsigma+1)} |\log(T)|^8 + T^{2(2-\varsigma)\beta} |\log(T)|^4 \right) \\ &= o_p\left(z_T^3\Delta + z_T^3 \left(T^{1-\frac{1}{3}(\varsigma+1)} |\log(T)|^8 + T^{2(2-\varsigma)\beta} |\log(T)|^4 \right) \right),\end{aligned} \tag{165}$$

where for the last equality we made use of the fact that $z_T^3\Delta \to 0$. Altogether, we have

$$\left|\int_{-z_T}^{z_T} \frac{e^{-i\vartheta z}}{c+iz}\overline{Z}_1(-c-iz)dz\right|^2 = O_p\left(z_T^3\left(\Delta + |\log(T)|^8 T^{1-\frac{1}{3}(\varsigma+1)} + |\log(T)|^4 T^{2(2-\varsigma)\beta}\right)\right). \quad (166)$$

Inequality in means also implies

$$\left|\int_{-z_T}^{z_T} \frac{e^{-i\vartheta z}}{c+iz}(\overline{h}_T(-c-iz) - \overline{Z}_1(-c-iz))dz\right|^2 \leq C_0 z_T \sum_{j=4}^{6} |\mathcal{W}_j|. \quad (167)$$

Combining the above bounds, the result for $\widehat{LV}_T(\vartheta, z_T, u_T)$ follows. These bounds as well as the bounds for $\{\mathcal{W}'_j\}_{j=2,\ldots,6}$ derived in the proof of Theorem 2 imply the result for $\widehat{LV}_T(\vartheta, z_T, \kappa_T, \overline{\kappa}_T)$.

We are left with showing the results in the case when ϑ is replaced with $\widehat{\vartheta}_0$. First note that all of the above results hold trivially if we replace ϑ with ϑ_0, for ϑ_0 being almost surely positive and $\mathcal{F}^{(0)}$-adapted. This is because of the structure of the option observation errors. Next, we note that from Lemma 3, we have

$$\left|\int_{-z_T}^{z_T} \frac{e^{-\widehat{\vartheta}_0(c+iz)} - e^{-\vartheta_0(c+iz)}}{c+iz}(\widehat{\sigma}_T^2(u_T) - \widehat{\mathcal{C}}_T(1) - \widehat{\mathcal{C}}_T(2) - \sigma_0^2)dz\right| \leq C|\widehat{\sigma}_T^2(u_T) - \widehat{\mathcal{C}}_T(1) - \widehat{\mathcal{C}}_T(2) - \sigma_0^2|. \quad (168)$$

Next, by Cauchy–Schwarz inequality and the rate conditions for z_T, Δ and T, we have

$$\begin{aligned}&\left|\int_{-z_T}^{z_T} \frac{e^{-\widehat{\vartheta}_0(c+iz)} - e^{-\vartheta_0(c+iz)}}{c+iz}(\widehat{Z}_2(-c-iz) + \widehat{Z}_3(-c-iz) + \widehat{R}(-c-z))dz\right|^2 \\ &\leq \int_{-z_T}^{z_T} |e^{-\widehat{\vartheta}_0(c+iz)} - e^{-\vartheta_0(c+iz)}|^2 dz \times \sum_{j=4}^{6}\mathcal{W}_j = O_p\left(z_T^3\Delta + \gamma(T, z_T)\right),\end{aligned} \quad (169)$$

and similarly

$$\left|\int_{-z_T}^{z_T} \frac{e^{-\widehat{\vartheta}_0(c+iz)} - e^{-\vartheta_0(c+iz)}}{c+iz}\left(\frac{\nabla_{uu}\widehat{A}_T^{(a)}(-c-iz)}{T\widetilde{L}_T(-c-iz)} - \widehat{\mathcal{C}}_T(1)\right)dz\right|^2 = O_p\left(z_T^3\Delta \times z_T^3 \upsilon_T^2\right). \quad (170)$$

On the other hand, the term

$$\int_{-z_T}^{z_T} \frac{e^{-\widehat{\vartheta}_0(c+iz)} - e^{-\vartheta_0(c+iz)}}{c+iz}\left(\frac{1}{T\widetilde{L}_T(-c-iz)}\sum_{j=b,c,d}\nabla_{uu}\widehat{A}_T^{(j)}(-c-iz) - \widehat{\mathcal{C}}_T(2)\right)dz,$$

can be analyzed exactly in the deterministic case for ϑ above as the stochasticity here plays no role (note that because of the consistency of $\widehat{\vartheta}_t$ for ϑ_t and the strict positivity of the latter, on a set of probability approaching one $\widehat{\vartheta}_t$ is above $\vartheta_t/2 > 0$).

Altogether, we have

$$\begin{aligned}&\left|\int_{-z_T}^{z_T} \frac{e^{-\widehat{\vartheta}_0(c+iz)} - e^{-\vartheta_0(c+iz)}}{c+iz}\left(\widehat{h}_T(-c-iz) - h_0(-c-iz) - \widehat{\sigma}_T^2(u_T) + \sigma_0^2\right)dz\right| \\ &= O_p\left(z_T^3\Delta(z_T^3\upsilon_T^2 \vee 1) + \gamma(T, z_T)\right).\end{aligned} \quad (171)$$

Finally,

$$\int_{-z_T}^{z_T} \frac{e^{-\widehat{\vartheta}_0(c+iz)}}{c+iz}(h_0(-c-iz)-\sigma_0^2)dz = LV_0(\widehat{\vartheta}_0) - \int_{|z|>z_T} \frac{e^{-\widehat{\vartheta}_0(c+iz)}}{c+iz}\mathcal{F}\nu_0(-c-iz)dz, \tag{172}$$

and similar expression holds with $\widehat{\vartheta}_0$ replaced with ϑ_0. From assumption A1, we have

$$\left|\int_{|z|>z_T} \frac{e^{-\widehat{\vartheta}_0(c+iz)}}{c+iz}\mathcal{F}\nu_0(-c-iz)dz\right| + \left|\int_{|z|>z_T} \frac{e^{-\vartheta_0(c+iz)}}{c+iz}\mathcal{F}\nu_0(-c-iz)dz\right| = O_p\left(z_T^{-r-1/2}\right). \tag{173}$$

Combining the above bounds, we have the result of the theorem.

Acknowledgements

I would like to thank Serena Ng (the Editor), the Associate Editor and two anonymous referees for many useful comments and suggestions.

References

Almeida, C., K. Ardison, R. Garcia, and J. Vicente (2017). Nonparametric tail risk, stock returns, and the macroeconomy. *Journal of Financial Econometrics 15*(3), 333–376.

Andersen, T. G., V. Todorov, and M. Ubukata (2021). Tail risk and return predictability for the japanese equity market. *Journal of Econometrics 222*, 344–363.

Belomestny, D. and M. Reiß (2006). Spectral calibration of exponential Lévy models. *Finance and Stochastics 10*, 449–474.

Belomestny, D. and M. Reiß (2015). Lévy Matters IV: Estimation for discretely observed Lévy processes. In *Lecture Notes in Mathematics 2128*. Springer.

Bingham, N. H., C. M. Goldie, and J. L. Teugels (1989). *Regular variation*, Volume 27. Cambridge university press.

Bollerslev, T., G. Tauchen, and H. Zhou (2009). Expected stock returns and variance risk premia. *The Review of Financial Studies 22*(11), 4463–4492.

Bollerslev, T. and V. Todorov (2011a). Estimation of jump tails. *Econometrica 79*(6), 1727–1783.

Bollerslev, T. and V. Todorov (2011b). Tails, fears, and risk premia. *The Journal of Finance 66*(6), 2165–2211.

Bollerslev, T. and V. Todorov (2014). Time-varying jump tails. *Journal of Econometrics 183*(2), 168–180.

Bollerslev, T., V. Todorov, and L. Xu (2015). Tail risk premia and return predictability. *Journal of Financial Economics 118*(1), 113–134.

Butucea, C. and F. Comte (2009). Adaptive estimation of linear functionals in the convolution model and applications. *Bernoulli 15*(1), 69–98.

Cai, T. T. (2003). Rates of convergence and adaptation over Besov spaces under pointwise risk. *Statistica Sinica 13*, 881–902.

Carr, P. and D. Madan (2001). Optimal positioning in derivative securities. *Quantitative Finance 1*, 19–37.

Cont, R. and P. Tankov (2003). *Financial modelling with jump processes.* Chapman and Hall/CRC.

Dattner, I., A. Goldenshluger, and A. Juditsky (2011). On deconvolution of distribution functions. *The Annals of Statistics 39*(5), 2477–2501.

Drechsler, I. and A. Yaron (2011). What's vol got to do with it. *The Review of Financial Studies 24*(1), 1–45.

Du, J. and N. Kapadia (2012). The tail in the volatility index. *U. Massachusetts, Amherst Working paper*.

Duffie, D. (2001). *Dynamic asset pricing theory* (3rd ed.). Princeton University Press.

Fan, J. (1991). On the optimal rates of convergence for nonparametric deconvolution problems. *The Annals of Statistics 19*(3), 1257–1272.

Gil-Pelaez, J. (1951). Note on the inversion theorem. *Biometrika 38*(3-4), 481–482.

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

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

Kawata, T. (2014). *Fourier analysis in probability theory*, Volume 15. Academic Press.

Kelly, B. and H. Jiang (2014). Tail risk and asset prices. *The Review of Financial Studies 27*(10), 2841–2871.

Kilic, M. and I. Shaliastovich (2019). Good and bad variance premia and expected returns. *Management Science 65*(6), 2522–2544.

Kolassa, J. E. (2006). *Series approximation methods in statistics*, Volume 88. Springer Science & Business Media.

Mancini, C. (2001). Disentangling the jumps of the diffusion in a geometric jumping brownian motion. *Giornale dell'Istituto Italiano degli Attuari 64*(19-47), 44.

Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics 36*(2), 270–296.

Orlowski, P., A. Sali, P. Schneider, and F. Trojani (2020). On the nature of jump risk premia. Technical report.

Pan, J. (2002). The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of financial economics 63*(1), 3–50.

Pensky, M. (2017). Minimax theory of estimation of linear functionals of the deconvolution density with or without sparsity. *The Annals of Statistics 45*(4), 1516–1541.

Qin, L. and V. Todorov (2019). Nonparametric implied Lévy densities. *Annals of Statistics 47*, 1025–1060.

Sato, K. (1999). *Lévy processes and infinitely divisible distributions.* Cambridge, UK: Cambridge University Press.

Schneider, P. and F. Trojani (2015). Fear trading. *Swiss Finance Institute Research Paper* (15-03).

Shephard, N. G. (1991). From characteristic function to distribution function: a simple framework for the theory. *Econometric theory 7*, 519–529.

Singleton, K. J. (2009). *Empirical dynamic asset pricing: model specification and econometric assessment*. Princeton University Press.

Söhl, J. (2014). Confidence sets in nonparametric calibration of exponential Lévy models. *Finance and Stochastics 18*, 617–649.

Todorov, V. (2019). Nonparametric spot volatility from options. *The Annals of Applied Probability 29*(6), 3590–3636.

Trabs, M. (2014). Calibration of self-decomposable Lévy models. *Bernoulli 20*, 109–140.

Trabs, M. (2015). Quantile estimation for Lévy measures. *Stochastic Processes and their Applications 125*, 3484–3521.

Wachter, J. A. (2013). Can time-varying risk of rare disasters explain aggregate stock market volatility? *The Journal of Finance 68*(3), 987–1035.

Weller, B. M. (2019). Measuring tail risks at high frequency. *The Review of Financial Studies 32*(9), 3571–3616.