NONPARAMETRIC SPOT VOLATILITY FROM OPTIONS

By Viktor Todorov

Northwestern University

We propose a nonparametric estimator of spot volatility from noisy short-dated option data. The estimator is based on forming portfolios of options with different strikes that replicate the (risk-neutral) conditional characteristic function of the underlying price in a model-free way. The separation of volatility from jumps is done by making use of the dominant role of the volatility in the conditional characteristic function over short time intervals and for large values of the characteristic exponent. The latter is chosen in an adaptive way in order to account for the time-varying volatility. We show that the volatility estimator is near rate-optimal in minimax sense. We further derive a feasible joint Central Limit Theorem for the proposed option-based volatility estimator and existing high-frequency return-based volatility estimators. The limit distribution is mixed-Gaussian reflecting the time-varying precision in the volatility recovery.

1. Introduction. Options provide a natural source of information for studying volatility. Indeed, following the seminal work of [14] and [33], any option written on an asset can be used to back out the unknown volatility of the asset. The resulting estimator of volatility is typically referred to as Black-Scholes Implied Volatility (BSIV). Unfortunately, the assumptions behind the model of [14] and [33], mainly constant volatility and no jump risk, are too simple for such volatility extraction to work in practice, see e.g., [19], [21] and [40]. Indeed, BSIV backed out from available options with strikes that are far from the current price level are typically too high when compared to historical averages based on returns data. These elevated implied volatility levels are a reflection of the importance of time-varying volatility and jump risk for investors. The goal of this paper is to develop nonparametric spot volatility estimator from options that works in general settings when jumps are present and volatility can vary over time.

Recent developments in financial markets make the construction of option-based nonparametric volatility estimates practically feasible. In particular, the availability and liquidity of very short-maturity options with a wide

AMS 2000 subject classifications: Primary 60F05, 60E10; secondary 60J60, 60J75.

Keywords and phrases: Itô Semimartingale, jumps, nonparametric inference, options, stable convergence, stochastic volatility.

range of strikes has significantly increased over the last few years, see e.g., [6].

We use these short-dated options in the construction of our estimator. A natural candidate for a spot volatility estimator is provided by the BSIV of options with strikes that are close to the current price level. The short time to expiration limits the effect of the time-varying volatility on these options. Similarly, the proximity of their strikes to the current price limits the effect of jumps on them. Nevertheless, we show that jumps cause an upward bias in the recovery of volatility from the BSIV of short-dated options with strikes close to the price level. This bias is nontrivial and cannot be ignored in practice.

In this paper, we take a different approach for estimating spot volatility from options, which allows for an efficient separation of volatility from jumps. Our approach makes use of the fact that the expected value of smooth functions of the price of the underlying asset at expiration can be replicated by portfolios of options with continuum of strike levels, see e.g., [15] as well as the earlier work of [26] and [38]. Using this insight, we construct portfolios of options which replicate the conditional risk-neutral characteristic function of the price at expiration. If the time to expiration is short, then the time variation in volatility has a negligible effect on the latter and can be ignored. The effect of the jumps on the characteristic function, on the other hand, is more subtle. If the value of the characteristic exponent is close to zero, then the jumps have a non-negligible effect. However, their effect diminishes for higher values of the characteristic exponent. We show that asymptotically (as the time to maturity shrinks) optimal separation of volatility from jumps can be achieved when the characteristic exponent is growing at a rate proportional to the square root of the time to expiration of the options. This leads to a volatility estimator which is significantly less biased in presence of jumps than the BSIV of options with strikes close to the current price.

We establish consistency of the proposed volatility estimator in an asymptotic setting in which options are observed with error, their maturity goes to zero together with shrinking mesh of the available strike grid. We further derive a Central Limit Theorem (CLT) for our volatility estimator. The limiting distribution is determined by the asymptotic behavior of the observation error in the available options. The convergence is stable and its asymptotic limit is mixed Gaussian. That is, the limit is centered Gaussian when conditioning on the sigma algebra on which the return and option data are defined. This allows for the asymptotic variance of the volatility estimator to depend, in particular, on the current level of volatility and more generally on any other variable that determines the quality of the option data. Hence, the precision in estimation will typically differ over different points in time. For feasible inference, we develop a simple estimator of the asymptotic variance which is based on an option portfolio that measures the sensitivity of the observed option prices to changes in their strikes.

There are many asymptotically valid choices for the characteristic exponent of the volatility estimator. However, for the successful performance of the estimator in practice, this choice matters a lot. Therefore, we develop an adaptive procedure for setting this tuning parameter by using an initial consistent estimator of volatility constructed from the option data. Our initial consistent estimator is the option analogue of the truncated high-frequency return volatility estimator of [31]. It is based on integrating the available options in a portfolio which spans a truncated second moment of the price at expiration (i.e., a function which behaves like the square function around zero and diminishes to zero for values of the argument diverging from zero).

We show that our estimator is near-rate optimal. In particular, in the specialized setting of Lévy jump-diffusion dynamics for the underlying price and Gaussian observation errors proportional to the true unobserved option prices, we show that the efficient rate (in a minimax sense) of recovering volatility from the noisy short-maturity option data coincides with the rate of convergence of our estimator up to a log term (the rate of convergence of our estimator is some power of the time to maturity). This is unlike a volatility estimator based on the average of close-to-money BSIV.

The nonparametric spot volatility estimator developed in this paper can be viewed as the option counterpart of the high-frequency return-based volatility estimators. In pioneering work, [9, 10] propose so-called multipower variation statistics as a way to separate volatility from jumps while [31, 32] develops truncated variance estimator that achieves the same goal. More recently, [30] propose the use of the empirical characteristic function of returns as a way to measure volatility in a jump-robust way, which allows also to deal with jumps of arbitrary high activity in an efficient way.

The high-frequency return-based volatility estimators use an asymptotically increasing number of increments in a local window of time to estimate volatility in a way similar to estimating volatility from a sequence of i.i.d. returns in classical settings. By contrast, the newly-proposed option-based estimator uses an asymptotically increasing number of short-dated options with different strikes to identify the expectation about the future volatility embodied in them. In turn, this conditional expectation of volatility converges to the spot volatility when the time to maturity of the options shrinks. We show that the convergence of the option-based and return-based volatility estimators holds jointly. This allows one to construct an optimal mixture of the two types of estimators which has the lowest asymptotic variance for measuring spot volatility from return and option data.

We evaluate the performance of the option-based volatility estimator in a Monte Carlo experiment whose setup mimics key features of available option data. The Monte Carlo shows satisfactory finite sample properties of the developed estimator and the inference about it. In an empirical application to short-dated S&P 500 index options, we find that the option-based volatility estimator is on average very close to an estimator based on high-frequency returns on the S&P 500 index but it is significantly more accurate.

The current paper is related to two strands of literature on volatility inference from option data. First, there is a large body of work that considers short-maturity expansions of at-the-money options or options which become at-the-money in the limit, see e.g., [12], [23], [24], [36], [35]. Unlike this body of work, we consider a portfolio of options across strikes which is important to reduce the asymptotic bias of the estimator. We further allow for observation error, derive a stable CLT for our estimator and establish its near rate-optimality. Second, [11], [20], [41], [42], [43, 44] consider nonparametric inference for the diffusive volatility in the class of exponential-Lévy models from options with fixed maturity. The major difference between the current paper and this strand of work is that our analysis applies to general Itô semimartingales and the asymptotic setup here is one with shrinking maturity of the options. The latter difference leads to a significantly faster rate of convergence of the volatility estimator in the current asymptotic setting as the shrinking maturity aids the separation of diffusive volatility from jumps.

The rest of the paper is organized as follows. In Section 2 we develop nonparametric methods for recovering volatility from options in the infeasible scenario where a continuum of short-maturity options with strikes spanning the positive real line are available. Section 3 adapts these procedures to the feasible setting where only a finite number of noisy option observations are available instead. In this section we further characterize the rate of convergence of the volatility estimator, derive a feasible CLT for it, and develop an adaptive method for selecting the tuning parameter used in its construction. Section 4 derives the minimax risk of recovering spot volatility from noisy short-dated options in the special Lévy case and Gaussian observation errors. Section 5 contains a Monte Carlo study and Section 6 an empirical application. The proofs are given in Section 7.

2. Option Portfolios and Volatility. We begin our analysis with showing how to identify volatility in the infeasible setting where short-dated options with arbitrary strikes are available and further when the options are observed without error. We will relax these assumptions about the option observation scheme in the next section.

The underlying asset price is denoted by X and is defined on the filtered probability space $\big(\Omega^{(0)}, \mathcal{F}^{(0)}, (\mathcal{F}_t^{(0)})_{t\geq 0}, \mathbb{P}^{(0)}\big)$. Since our focus in this paper is on extracting information from options, we will specify here only the behavior of X under the so-called risk-neutral measure $\mathbb{Q}$, which under no-arbitrage is locally equivalent to $\mathbb{P}^{(0)}$. For the return-based volatility estimates, which we use later on to compare the option-based volatility estimator with, we will need to impose some structure on the $\mathbb{P}^{(0)}$ dynamics of X as well (see assumption A6 in Section 3.2). The dynamics of the log-price $x = \ln(X)$ under $\mathbb{Q}$ is given by

$$x_t = \int_0^t a_s ds + \int_0^t \sigma_s dW_s + \int_0^t \int_{\mathbb{R}} x \widetilde{\mu}(ds, dx), \tag{2.1}$$

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 $\nu_t(x)dt \otimes dx$ and $\widetilde{\mu}$ is the martingale measure associated with μ (W and ν_t are defined with respect to $\mathbb{Q}$). The regularity conditions for the above quantities are given in Section 3.2.

Although equation (2.1) describes the dynamics of x under $\mathbb{Q}$, under no-arbitrage, σ_t continues to be the diffusive volatility of x under $\mathbb{P}^{(0)}$. Our goal here is to estimate the spot diffusive variance $V_t \equiv \sigma_t^2$ under general conditions, i.e., with minimal regularity assumptions about (a_t, σ_t, ν_t).

For the recovery of V_t, we will use options written on X at time t, which expire at $t + T$, for some $T > 0$. Since t will be fixed throughout, we will henceforth suppress the dependence on t in the notation of the option prices and other related quantities. For simplicity, we will further assume that the dividend yield associated with X and the risk-free interest rate are both equal to zero as their effect on short-dated options is known to be negligible. With these normalizations, the theoretical values of the options we will use in our analysis are given by

$$\kappa_T(k) = \begin{cases} \mathbb{E}_t^{\mathbb{Q}}(e^k - e^{x_{t+T}})^+, & \text{if } k \leq x_t, \\ \mathbb{E}_t^{\mathbb{Q}}(e^{x_{t+T}} - e^k)^+, & \text{if } k > x_t. \end{cases} \tag{2.2}$$

$\kappa_T(k)$ is the price of an out-of-the-money (OTM) option (i.e., an option which will be worth zero if it were to expire today). This is a call contract (an option to buy the asset) if $k > x_t$ and a put contract (an option to sell the asset) if $k \leq x_t$. In what follows, we will refer to $K \equiv e^k$ and k as the strike and log-strike, respectively, of the option.

To simplify analysis, in this section, we will assume that jumps are of finite activity, i.e., that

$$\int_t^{t+T}\int_{\mathbb{R}} \nu_s(dx)ds < \infty, \quad \text{a.s.} \tag{2.3}$$

Many jump models used in financial applications satisfy the above finite activity assumption and we will further relax it in the derivation of the formal results presented in the next section.

Henceforth, for a generic sequence of random variables Y_T and some deterministic sequence R_T, $Y_T = O_p(R_T)$ will mean that Y_T/R_T is bounded in probability and $Y_T = o_p(R_T)$ will mean that Y_T/R_T converges in probability to zero, with both statements being for $T \downarrow 0$, see e.g., Section 2.2 in [46].

Since the volatility accounts for the small moves in the asset price, a natural candidate for a spot volatility estimator is the at-the-money (ATM) Black-Scholes option implied volatility. Indeed, the ATM BSIV has often been used as a proxy for spot volatility in empirical work. When (2.3) holds and under some weak regularity type assumptions for (a_t, σ_t, ν_t), it is easy to show that

$$\kappa_T(0) = \frac{\sqrt{T}}{\sqrt{2\pi}}\sigma_t + O_p(T), \quad \text{as } T \downarrow 0. \tag{2.4}$$

This bound on the error for recovering σ_t from $\kappa_T(0)$ is sharp and a large component of it is due to the jumps in X. This can be illustrated using the seminal Merton jump-diffusion model ([34]) for which a higher-order expansion of $\kappa_T(0)$ can be derived. In the Merton model the volatility is constant and the jumps are compound Poisson with intensity λ and their size is drawn from a normal distribution with mean μ_j and variance σ_j^2. In this case, by directly expanding the option price by considering the leading cases of no jump or one jump in X until expiration, we get for the ATM option price, $\kappa_T^M(0)$, the following as $T \downarrow 0$:

$$\begin{aligned}\kappa_T^M(0) &= \frac{\sqrt{T}}{\sqrt{2\pi}}\sigma - \frac{T\sigma^2}{4} \\ &\quad + \lambda T\left(\Phi\left(-\frac{\mu_j}{\sigma_j}\right) - e^{\mu_j + \frac{\sigma_j^2}{2}}\Phi\left(-\frac{\mu_j}{\sigma_j} - \sigma_j\right)\right) + O_p(T^{3/2}),\end{aligned} \tag{2.5}$$

where Φ denotes the cdf of a standard normal random variable. The first two terms on the right-hand side of (2.5) are the leading terms of the option price when conditioning on no jumps in X until expiration. The third term is the leading component of the option price when conditioning on exactly one jump occurring during the life of the option.

The above parametric example shows that the bound in (2.4) is sharp. Using the ATM option price expansion in (2.4), we have

$$V_t = \frac{2\pi}{T}\kappa_T^2(0) + O_p(\sqrt{T}), \quad \text{as } T \downarrow 0, \tag{2.6}$$

and we can alternatively estimate V_t using the Black-Scholes implied volatility corresponding to $\kappa_T(0)$ (with obviously the same order of magnitude of the approximation error as above). In Figure 1 we illustrate the accuracy of the ATM BSIV for measuring spot volatility using volatility and option data generated from the parametric model used later in the Monte Carlo study. The maturity of the options in the experiment is set to $T = 2$ days. As seen from the figure, even for such short maturity, the bias due to the jumps in the ATM BSIV as a measure of V is rather nontrivial and increases as a function of the volatility. Moreover, in practice, we often do not have an option with k equal exactly to 0 (due to the discreteness of the available strike grid) and this will likely generate additional bias in the measurement of spot volatility.

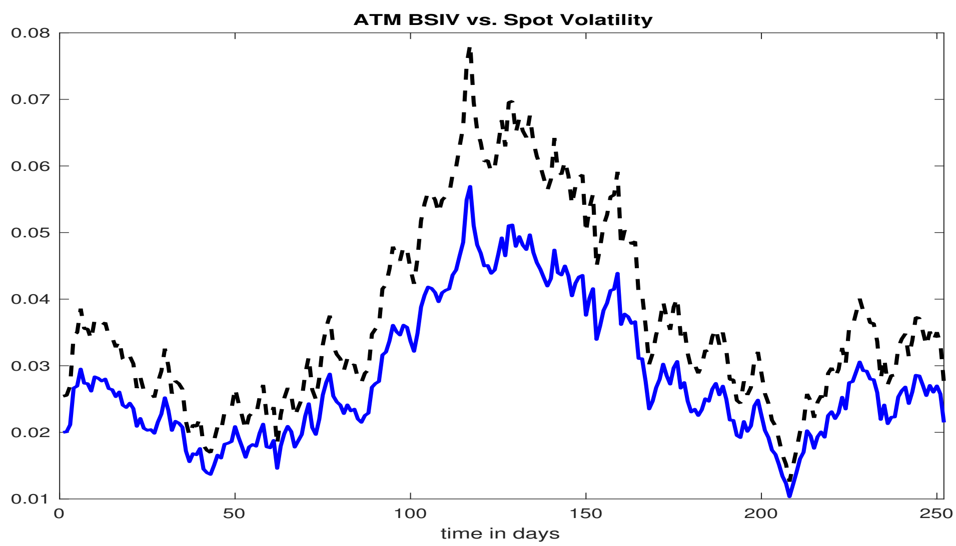

FIG 1. *ATM BSIV as a Measure of Volatility. The solid line corresponds to simulated path of V from the model in the Monte Carlo experiment in Section 5 (case H). The dashed line is the ATM BSIV with time to maturity of two business days.*

REMARK 1. *While Figure 1 shows that BSIV is a poor estimator of the true spot diffusive volatility of the underlying asset, we note nevertheless that BSIV is widely used in practice to quote options and also to generate option* *prices for strikes which are not available via interpolation in BSIV space. In addition, if one is interested solely in modeling the options, then (misspecified) diffusive stochastic volatility models which generate option prices free of arbitrage can be used. However, there is a large nonparametric evidence for presence of jumps in the underlying asset from return data (see e.g., [1]) as well as from option data (see e.g., [16]). Therefore, if one is interested in the joint modeling of the underlying asset and the derivatives written on it in a dynamically consistent way, then the nonparametric recovery of the spot diffusive volatility is important. Moreover, the volatility of the underlying asset is interesting in itself for addressing various practical risk management and theoretical asset pricing questions.*

We now develop an alternative strategy for recovering spot volatility from short-dated options which will have much smaller approximation error than the ATM BSIV. Our strategy builds on the fact that the conditional expectation (under $\mathbb{Q}$) of any sufficiently smooth functions of x_{t+T} can be spanned by a portfolio of options with continuum of strikes, $\{\kappa_T(k)\}_{k\in\mathbb{R}}$, see e.g., [15]. We note that this spanning result lies also behind the construction of the popular volatility VIX index computed by the CBOE options exchange.

The idea of our estimation strategy is to pick a function of the terminal price which will allow us to efficiently separate the volatility from the jumps. We will use the characteristic function to achieve this. Using $\{\kappa_T(k)\}_{k\in\mathbb{R}}$, we can recover $\mathbb{E}_t^{\mathbb{Q}}\left(e^{iu(x_{t+T}-x_t)}\right)$ (see the expression in (3.11) below for the explicit formula). For an appropriate choice of u, as we now show, we can disentangle volatility from jumps using $\mathbb{E}_t^{\mathbb{Q}}\left(e^{iu(x_{t+T}-x_t)}\right)$.

To help intuition, lets first assume that $x_{t+T}-x_t$ is, $\mathcal{F}_t$-conditionally, a Lévy process under $\mathbb{Q}$ (i.e., a process with i.i.d. increments). In this case, the Lévy-Khintchine formula ([39], Theorem 8.1) implies

$$\begin{aligned}(2.7)\quad & \mathbb{E}_t^{\mathbb{Q}}\left(e^{iu(x_{t+T}-x_t)/\sqrt{T}}\right) \\ & \quad = \exp\left(iu\sqrt{T}a_t - \frac{u^2}{2}V_t + T\int_{\mathbb{R}}(e^{iuT^{-1/2}x} - 1 - iuT^{-1/2}x)\nu_t(x)dx\right).\end{aligned}$$

Using our finite activity jump assumption in (2.3), we easily have that $\int_{\mathbb{R}}(\cos(uT^{-1/2}x)-1)\nu_t(x)dx = O_p(1)$, and therefore

$$V_t = -\frac{2}{u^2}\Re\left(\ln\left(\mathbb{E}_t^{\mathbb{Q}}\left(e^{iu(x_{t+T}-x_t)/\sqrt{T}}\right)\right)\right) + O_p(T), \quad \text{as } T\downarrow 0. \tag{2.8}$$

As we show in the appendix, the above approximation continues to hold even when $x_{t+T}-x_t$ is not $\mathcal{F}_t$-conditionally a Lévy process but it can instead have time-varying volatility and jump intensity. Comparing (2.6) and (2.8), we can see that the characteristic function based approach has asymptotically smaller error than the ATM BSIV for estimating the spot volatility.

In Figure 2, we illustrate the accuracy of the expression on the right-hand-side of (2.8) for measuring V_t in the context of the parametric model used in the Monte Carlo study below. We use two horizons of $T = 2$ and $T = 5$ days. As seen from the figure, when u is very small, the bias in estimating volatility due to the presence of jumps is rather nontrivial. Indeed, for the limit case of $u \downarrow 0$, $-\frac{2}{u^2}\Re\left(\ln\left(\mathbb{E}_t^{\mathbb{Q}}\left(e^{iu(x_{t+T}-x_t)/\sqrt{T}}\right)\right)\right)$ converges to the (expected) spot quadratic variation, $V_t + \int_{\mathbb{R}} x^2 \nu_t(x)dx$, which includes the (risk-neutral) second moment of the jump part. As u increases, the effect due to the jumps disappears and the characteristic function based volatility measures converge to V_t. This happens faster for the volatility measure based on the shorter of the two horizons and this volatility measure is also uniformly (across u) less biased.

Overall, consistent with our asymptotic analysis above, volatility estimation based on the conditional characteristic function can separate volatility from jumps far more efficiently than ATM BSIV (in the sense of smaller bias). For this to be of practical use, however, we need to be able to estimate reliably from the available options the conditional characteristic function of the returns for sufficiently high values of u for which the effect of the jumps is minimal. This is what we study next.

3. Nonparametric Option-Based Volatility Estimation.

We now develop the feasible counterpart of the volatility estimator based on the characteristic function proposed in the previous section and derive its asymptotic properties. We start with describing the observation scheme in Section 3.1 and stating our assumptions in Section 3.2, followed by a formal definition of the estimator in Section 3.3 and derivation of its asymptotic order. Section 3.4 proposes an option-based truncation volatility which we use in Section 3.5 to select the characteristic exponent of the volatility estimator in an adaptive way. This section further presents a feasible CLT.

3.1. *The Observation Scheme.*

Our data consists of OTM options at time t, expiring at $t + T$, and having log-strikes

$$\underline{k} \equiv k_1 < k_2 < \cdots k_N \equiv \overline{k}, \tag{3.1}$$

with the corresponding strikes given by

$$\underline{K} \equiv K_1 < K_2 < \cdots K_N \equiv \overline{K}. \tag{3.2}$$

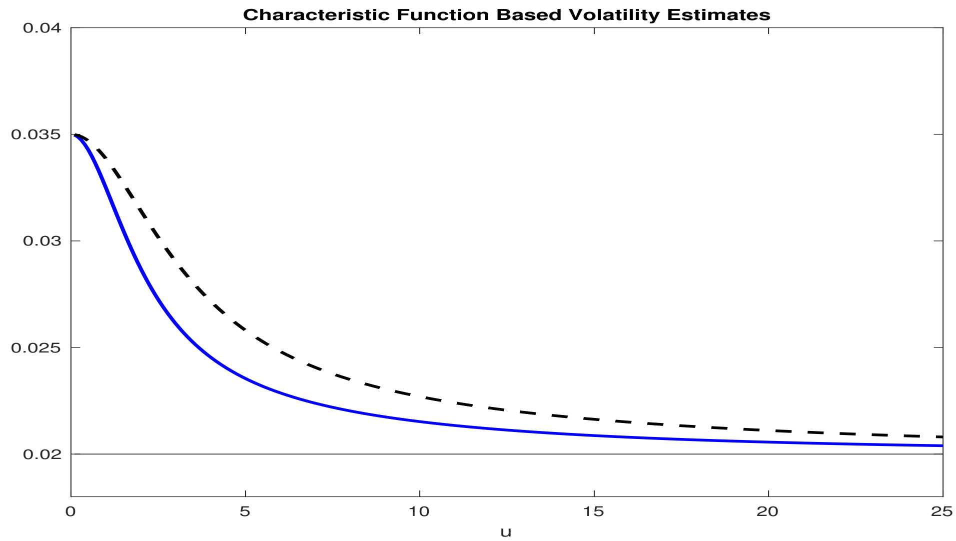

FIG 2. *Characteristic Function Based Volatility Estimates. The plot displays* $-\frac{2}{u^2}\Re\left(\ln\left(\mathbb{E}_t^{\mathbb{Q}}\left(e^{iu(x_{t+T}-x_t)/\sqrt{T}}\right)\right)\right)$ *as a function of* u *for the model in the Monte Carlo experiment in Section 5 (case H). The solid line corresponds to two business days and the dashed one to five business days. The horizontal solid line is at the true value of the spot variance.*

We denote the gaps between the log-strikes with $\Delta_i = k_i - k_{i-1}$, for $i = 2, ..., N$. We note that we do not assume an equidistant log-strike grid, i.e., we allow for Δ_i to differ across i-s. The asymptotic theory developed below is of joint type in which the time to maturity of the option T goes down to zero, the mesh of the log-strike grid $\sup_{i=2,\ldots,N} \Delta_i$ shrinks to zero and (in some cases) the log-strike limits $-\underline{k}$ and $\overline{k}$ increase to infinity.

Finally, as common in empirical derivatives pricing, we allow for observation error, i.e., instead of observing $\kappa_T(k_i)$ directly, we observe:

$$\widehat{\kappa}_T(k_i) = \kappa_T(k_i) + \epsilon_i, \tag{3.3}$$

where the sequence of observation errors $\{\epsilon_i\}_{i\geq 1}$ is defined on a space $\Omega^{(1)} = \mathbb{R}^{\mathbb{R}}$. This space is equipped with the product Borel σ-field $\mathcal{F}^{(1)}$ and with conditional probability $\mathbb{P}^{(1)}(\omega^{(0)}, d\omega^{(1)})$ from the original probability space $\Omega^{(0)}$ – on which X is defined – to $\Omega^{(1)}$. The reason for defining the errors on $\mathbb{R}^{\mathbb{R}}$ is that our asymptotic setup is of infill type and so we need to define an error for every log-strike which takes value in $\mathbb{R}$. We further define,

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

We will assume $\mathbb{E}\left(\epsilon_i|\mathcal{F}^{(0)}\right) = 0$ and that ϵ_i and ϵ_j are $\mathcal{F}^{(0)}$-conditionally independent for $i \neq j$. At the same time, we will allow for a general form of $\mathcal{F}^{(0)}$-conditional heteroskedasticity in the observation error.

3.2. *Assumptions.* We continue with our formal assumptions for the process x, the option observation scheme as well as the observation error.

A1. *$V_t > 0$ and the process σ has the following dynamics under $\mathbb{Q}$ for $s \geq t$:*

$$\sigma_s = \sigma_t + \int_t^s b_u du + \int_t^s \eta_u dW_u + \int_t^s \widetilde{\eta}_u d\widetilde{W}_u + \int_t^s \int_{\mathbb{R}} \delta^\sigma(u,z)\mu^\sigma(du,dz), \tag{3.4}$$

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 $\nu^\sigma(du,dz) = du \otimes dz$, having arbitrary dependence with the random measure μ; b, η and $\widetilde{\eta}$ are processes with càdlàg paths and $\delta^\sigma(u,z) : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ is left-continuous in its first argument.

A2-r. *With the notation of A1 and for some $r \in [0,1]$, there exists an $\mathcal{F}_t$-adapted random variable $\overline{t} > t$ such that for $s \in [t, \overline{t}]$:*

$$\mathbb{E}_t^{\mathbb{Q}}|a_s|^4 + \mathbb{E}_t^{\mathbb{Q}}|\sigma_s|^6 + \mathbb{E}_t^{\mathbb{Q}}(e^{4|x_s|}) + \mathbb{E}_t^{\mathbb{Q}}\left(\int_{\mathbb{R}}[(e^{3|z|}-1) \vee |z|^r]\nu_s(z)dz\right)^4 < C_t, \tag{3.5}$$

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

$$\mathbb{E}_t^{\mathbb{Q}}\left(\int_{\mathbb{R}}(|\delta^\sigma(s,z)|^4 \vee |\delta^\sigma(s,z)|)dz\right)^{1+\iota} \leq C_t. \tag{3.6}$$

A3. *With the notation of A1, there exists an $\mathcal{F}_t$-adapted random variable $\overline{t} > t$ such that for $s \in [t, \overline{t}]$:*

$$\mathbb{E}_t^{\mathbb{Q}}|a_s - a_t|^p + \mathbb{E}_t^{\mathbb{Q}}|\sigma_s - \sigma_t|^p + \mathbb{E}_t^{\mathbb{Q}}|\eta_s - \eta_t|^p + \mathbb{E}_t^{\mathbb{Q}}|\widetilde{\eta}_s - \widetilde{\eta}_t|^p \leq C_t|s-t|,\ p \in [2,4], \tag{3.7}$$

$$\mathbb{E}_t^{\mathbb{Q}}\left(\int_{\mathbb{R}}(e^{z\vee 0}|z| \vee |z|^2)|\nu_s(z) - \nu_t(z)|dz\right)^p \leq C_t|s-t|,\ p \in [2,3], \tag{3.8}$$

$$\mathbb{E}_t^{\mathbb{Q}}\left(\int_{\mathbb{R}}\left(\delta^\sigma(s,z) - \delta^\sigma(t,z)\right)^2 dz\right) \leq C_t|s-t|, \tag{3.9}$$

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

A4. *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*

$$\eta\overline{\Delta} \le \inf_{i=2,\dots,N} \Delta_i \le \sup_{i=2,\dots,N} \Delta_i \le \overline{\Delta}, \tag{3.10}$$

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

A5. *We have: (1)* $\mathbb{E}\left(\epsilon_i \middle| \mathcal{F}^{(0)}\right) = 0$*, (2)* $\mathbb{E}\left(\epsilon_i^2 \middle| \mathcal{F}^{(0)}\right) = \kappa_T(k_i)^2\sigma_{t,i}^2$ *where* $\sigma_{t,i} \equiv \sigma_t(k_i)$ *with* $\inf_{k\in\mathbb{R}} \sigma_t(k)$ *and* $\sup_{k\in\mathbb{R}} \sigma_t(k)$ *being finite-valued, positive and* $\mathcal{F}_t^{(0)}$*-adapted random variables, (3)* $\mathbb{E}\left(|\epsilon_i|^4 | \mathcal{F}^{(0)}\right) \le \zeta_t \kappa_T(k_i)^4$ *for some finite-valued* $\mathcal{F}_t^{(0)}$*-adapted random variable* ζ_t*, and (4)* ϵ_i *and* ϵ_j *are* $\mathcal{F}^{(0)}$*-conditionally independent whenever* $i \neq j$*.*

A6. *The dynamics of* X *and* σ *under* $\mathbb{P}$ *is as (2.1) and (3.4) but with* W*,* $\widetilde{W}$ *and* μ^σ *defined with respect to* $\mathbb{P}$*, and with* μ *having a compensator under* $\mathbb{P}$ *of the form* $\nu_t^{\mathbb{P}}(x)dt\otimes dx$*. The drift coefficient of* X *is locally bounded. Moreover, for a sequence of stopping times* (τ_n) *increasing to infinity and a sequence of functions* $\Gamma_n(z)$ *satisfying* $\int_{\mathbb{R}} \Gamma_n(z)dz < \infty$*, we have* $\int_{\mathbb{R}}(|z| \wedge 1)\nu_t^{\mathbb{P}}(dx) < \infty$ *and* $|\delta^\sigma(t,z)| \le \Gamma_n(z)$ *for* $t \le \tau_n$*.*

Assumption A1 imposes σ_t to be an Itô semimartingale under $\mathbb{Q}$, which is the case for many applications, e.g., for models in the popular affine class, see e.g., [22]. Assumption A2 imposes existence of conditional moments. This assumption also assumes that the so-called jump activity of X (see e.g., Section 3.2 of [27]) is bounded by $r \in [0,1]$. Assumption A3 imposes "smoothness in expectation" type conditions which are satisfied for example when the corresponding processes are Itô semimartingales. Assumption A4 is a weak condition on the strike grid and Assumption A5 is about the observation error. The latter is $\mathcal{F}^{(0)}$-conditionally centered at zero and it can have $\mathcal{F}^{(0)}$-conditional heteroskedasticity. The $\mathcal{F}^{(0)}$-conditional standard deviation of the observation error is proportional to the option price it is attached to and this determines the asymptotic order of the error as $T \downarrow 0$. Finally, assumption A6 is only needed for the high-frequency return-based volatility estimator and is taken from [27] (Assumption H in Section 9.1).

3.3. *Construction of the Volatility Estimator and its Rate of Convergence.* We proceed with formally defining our characteristic function based volatility estimator. Using Appendix 1 of [15], the conditional characteristic function of the log return, $\mathbb{E}_t^{\mathbb{Q}}(e^{iu(x_{t+T}-x_t)})$, can be spanned by the following portfolio of options

$$1-(u^2+iu)\int_{\mathbb{R}} e^{(iu-1)k-iux_t}\kappa_T(k)dk, \quad u\in\mathbb{R}. \tag{3.11}$$

The integral in the above expression is not computable, given available data, because we do not have option observations over a continuum of strikes and furthermore we do not observe directly $\kappa_T(k)$. The computable counterpart of the expression in (3.11) is formed by using a Riemann sum approximation of the integral in (3.11) constructed from the available noisy option observations:

$$\widehat{f}_{t,T}(u) = 1-(u^2+iu)\sum_{j=2}^{N} e^{(iu-1)k_{j-1}-iux_t}\widehat{\kappa}_T(k_{j-1})\Delta_j, \quad u\in\mathbb{R}. \tag{3.12}$$

While in general $x_{t+T}-x_t$ is not $\mathcal{F}_t^{(0)}$-conditionally the increment of a Lévy process, when T is small, the expression for the characteristic function in (2.7) nevertheless holds approximately true. This motivates the following estimator of the volatility:

$$\widehat{V}_{t,T}(u) = \frac{2}{Tu^2}\widehat{R}_{t,T}(u), \tag{3.13}$$

where $\widehat{R}_{t,T}(u)$ is given by

$$\widehat{R}_{t,T}(u) = -\Re\left(\ln\left(\widehat{f}_{t,T}(u)\vee T\right)\right). \tag{3.14}$$

For $\widehat{V}_{t,T}(u)$ to be a consistent estimator of V_t, we will need $\widehat{f}_{t,T}(u)$ to converge in probability to the expression in (3.11) and for this we will need the mesh of the discrete strike price grid in (3.1) to go to zero and the time to maturity T of the options to shrink. The formal result for the consistency and rate of convergence of $\widehat{V}_{t,T}(u)$ is given in the next theorem.

THEOREM 1. *Suppose assumptions A1-A5 in Section 3.2 hold for some $r\in[0,1]$ and in addition $\overline{\Delta}\asymp T^{\alpha}$, $\overline{K}\asymp T^{-\beta}$, $\underline{K}\asymp T^{\gamma}$ for some $\alpha>\frac{1}{2}$, $\beta\geq 0$ and $\gamma\geq 0$. Let (u_T) be an $\mathcal{F}_t^{(0)}$-adapted sequence such that*

$$u_T^2 T \xrightarrow{a.s.} \overline{u}, \quad \text{where } \overline{u} \text{ is a finite nonnegative random variable.} \tag{3.15}$$

Then, we have

$$\widehat{V}_{t,T}(u_T) - V_t = O_p\left(u_T^{r-2}\bigvee\frac{\sqrt{\overline{\Delta}}}{T^{1/4}}\bigvee u_T^{-1}e^{-2(|\underline{k}|\vee\overline{k})}\right). \tag{3.16}$$

Since the order of magnitude of the increment $x_{t+T} - x_t$ shrinks asymptotically as $T \downarrow 0$, it is intuitively clear that we need to consider sequences (u_T) which go to infinity. We look only at the case where u_T increases as fast as $1/\sqrt{T}$ because for sequences (u_T) going at a faster rate to infinity, the limit of $\widehat{f}_{t,T}(u)$ will be zero (recall (2.7)) and hence the signal about the volatility will be smaller.

The rate of convergence result in (3.16) reveals the role of the different sources of error in the volatility estimation. The first term on the right-hand side of (3.16) is a bias due to the presence of jumps in X. The parameter r controls the so-called jump activity (see assumption A2-r), with higher values of r implying more concentration of small jumps in X which in turn are harder to separate from the diffusive component. The case of finite activity jumps that we considered in the previous section corresponds to $r = 0$. Similar to earlier work on recovering volatility from high-frequency return data (e.g., [9, 10] and [31, 32]) and in line with earlier empirical evidence in [4] and [18], here we consider only the case of finite variation jumps, i.e., $r \leq 1$. For the infinite variation case, the bias due to the jumps becomes larger and a bias correction analogous to the one in [30] for the return-based estimator is probably needed for satisfactory performance of the volatility estimator in practice. Finally, from (3.16), it is clear that better separation of volatility from jumps is achieved for higher values of u_T.

The second term on the right-hand side of (3.16) is due to the observation error, i.e., due to the fact that we use $\widehat{\kappa}_T(k)$ in the estimation instead of $\kappa_T(k)$. The conditional volatility of the observation error is assumed to be of the same order of magnitude as the option price it is associated with (see assumption A5). This is intuitive and is motivated by the empirical evidence in [4] regarding the size of the relative bid-ask spread in available option data sets. The asymptotic order of magnitude of the option prices differ depending on the strike (and hence the same applies for the observation errors attached to the options). In particular, for log-strikes which are within a range from the current log-price of order $O_p(\sqrt{T})$, the option prices are of asymptotic order $O_p(\sqrt{T})$. On the other hand, for log-strikes which are of fixed size (different from the current log-price), the option prices are of asymptotic order $O_p(T)$ only. That is, for time-to-maturity T shrinking to zero, the option prices whose strikes are close to the current price level are of larger asymptotic order than the ones whose strikes are further away from it. Note that in (3.12) we use options with all available strikes (provided β and γ are strictly positive). The above discussion suggests that the effect of the observation error on the recovery of volatility will be determined by the option observations whose strikes are in the vicinity of the current price.

The third term on the right-hand side of (3.16) is due to the finite log-strike range of the available option data $(\underline{k}, \overline{k})$ used in the estimation. Intuitively, the order of magnitude of this error will depend on the probability mass in the tails of the risk-neutral $\mathcal{F}_t^{(0)}$-conditional distribution of $x_{t+T} - x_t$. With stronger assumptions for the latter, than what is currently assumed in assumption A2, the order of magnitude of this error can be further relaxed. From a practical point of view, this error is likely to have little impact on the estimation, as for the typical option data sets, the deepest available OTM option prices are very close to zero. This implies that the "effective" support of the conditional return distribution is covered by the available log-strike range $(\underline{k}, \overline{k})$. Indeed, earlier empirical work has documented that the effect of the finite strike range of the available options on the precision of the VIX index (which is another portfolio of options with different strikes) is typically small. We further note that since the argument of the characteristic function u_T is asymptotically drifting to infinity, we have that $\widehat{V}_{t,T}(u_T)$ is a consistent estimator of V_t even when the strike range of the options remains finite.

Finally, the recovery of the spot volatility from the short-dated options contains an error due to the time-variation in the volatility and the jump intensity over the interval $[t, t+T]$. The effect of this error on the volatility estimation is of order $O_p(T)$ and hence it is asymptotically dominated by the first term on the right-hand side of (3.16) (which recall is due to the separation of volatility from jumps and is present even if volatility is constant). We note in this regard that our interest here is in the effect of the error due to the time-variation in volatility and jump intensity on the recovery of the option portfolio in (3.11) and not on an individual option. The former is much smaller than what we can show for the latter. We also mention that it is only the stochastic changes in the volatility and the jump intensity which cause the above-mentioned bias in the estimation. Indeed, if conditional on $\mathcal{F}_t$ the process V has deterministic time-variation over the interval $[t, t+T]$, then $\widehat{V}_{t,T}(u_T)$ is an estimate of $\frac{1}{T}\int_t^{t+T} V_s ds$ without any bias due to the time-variation in V.

3.4. *Data-driven Choice of u_T and Option-Based Truncated Volatility.* From Theorem 1, it is clear that in order to minimize the impact of the jumps on the volatility recovery, it is optimal to set u_T to be of asymptotic order $O_p(1/\sqrt{T})$. This, of course, is an asymptotic statement and it does not give a specific guidance regarding the choice of u_T in finite samples. At the same time, from the expression for the log-characteristic function in (2.7), it is clear that its behavior is governed by the product $T \times u_T^2 \times V_t$. Therefore, one would like to set u_T such that $T \times u_T^2 \times V_t$ is some fixed constant. To do this, we will need a preliminary estimator of volatility and further we will have to show that our estimator $\widehat{V}_{t,T}(u_T)$ can be made adaptive, i.e., that u_T can be replaced with an estimate $\widehat{u}_T$ based on the data.

In this section we tackle the first problem, i.e., the construction of a preliminary volatility estimator from the option data, and in the next section we deal with making $\widehat{V}_{t,T}(u_T)$ adaptive. One natural choice of a preliminary volatility estimator is the ATM BSIV which has the additional advantage of being free of tuning parameters. However, given the documented large bias in the ATM BSIV, we propose an alternative one. Our initial consistent volatility estimator can be viewed as the option analogue of the truncated volatility estimator of [31] and is given by

$$\widehat{TV}_{t,T}(\eta) = \frac{1}{T}\sum_{j=2}^{N} h_\eta(k_{j-1})\widehat{\kappa}_T(k_{j-1})\Delta_j, \quad \eta \geq 0, \tag{3.17}$$

where we denote

$$h_\eta(k) = e^{-k-\eta(k-x_t)^2}\Big[4\eta^2(k-x_t)^4 + 2 - 10\eta(k-x_t)^2 + 2\eta(k-x_t)^3 - 2(k-x_t)\Big].$$

$\widehat{TV}_{t,T}(\eta)$ is a consistent estimator of $\mathbb{E}_t^{\mathbb{Q}}(e^{-\eta(x_{t+T}-x_t)^2}(x_{t+T}-x_t)^2)$ from the available options. In the special case when $\eta = 0$ we denote

$$\widehat{QV}_{t,T} \equiv \widehat{TV}_{t,T}(0), \tag{3.18}$$

and we note that $\widehat{QV}_{t,T}$ is an estimator of the expected risk-neutral spot quadratic variation

$$QV_{t,T} = V_t + \int_{\mathbb{R}} x^2\nu_t(x)dx. \tag{3.19}$$

Thus, $\widehat{QV}_{t,T}$ is the option counterpart of the realized variance computed from return data ([2] and [7, 8]). We note however a fundamental difference. The realized variance is an estimator of $\int_t^{t+\tau} V_s ds + \sum_{s\in[t,t+\tau]}(\Delta x_s)^2$ for some $\tau > 0$. By contrast, $QV_{t,T}$ can be viewed as the risk-neutral $\mathcal{F}_t^{(0)}$-conditional expectation of this quantity for τ small (and further standardized, i.e., divided, by τ). While for small τ we have $V_t \approx \mathbb{E}_t^{\mathbb{Q}}(V_{t+\tau})$, the same does not hold for the expected and realized jumps no matter how small τ is and regardless of whether the jump intensity varies over time or not (i.e., whether ν_t depends on t or not).

When η is a positive number, then $\widehat{TV}_{t,T}(\eta)$ estimates a truncated (conditional) second moment of the increment $x_{t+T} - x_t$ with the degree of truncation determined by η. When η is replaced with an increasing function depending on T, i.e., when the degree of truncation changes as we get more short-dated option data, then we can use $\widehat{TV}_{t,T}(\eta)$ to separate volatility from jumps. This is analogous to the truncated volatility estimator based on return data proposed by [31], with the difference being that, unlike [31], we use a smooth truncated square function here.

To implement the option-based truncated volatility estimator, we need to choose the truncation level. The tradeoff we face here is similar to the one for the return-based counterpart of our estimator. On one hand we would like to set the truncation as high as possible to minimize the impact of the jumps on the statistic. On the other hand, a more severe truncation will cause a downward bias in the recovery of volatility since such severe truncation will start eliminating even the contribution coming from the continuous part of the process in the second moment of the return. Therefore, an adaptive version of $\widehat{TV}_{t,T}(\eta)$ is necessary. We use the following data-driven choice for the cutoff parameter

$$\widehat{\eta}_T = \frac{\overline{\eta}_T}{T} \frac{1}{\widehat{QV}_{t,T} \vee T}, \tag{3.20}$$

for some deterministic sequence $\overline{\eta}_T$ that depends only on T and which goes to zero, but at a rate slower than the one at which T decreases. The reason for setting the truncation parameter this way is that the downward bias in $\widehat{TV}_{t,T}(\eta)$ caused by the truncation depends on the product $\eta \times T \times QV_{t,T}$. In the next theorem we present the consistency result for our truncation volatility estimator.

THEOREM 2. *Suppose assumptions A1-A5 in Section 3.2 hold for some* $r \in [0,1]$ *and in addition* $\overline{\Delta} \asymp T^{\alpha}$, $\overline{K} \asymp T^{-\beta}$, $\underline{K} \asymp T^{\gamma}$ *for some* $\alpha > \frac{1}{2}$, $\beta > 0$ *and* $\gamma > 0$. *We have*

$$\widehat{QV}_{t,T} \xrightarrow{\mathbb{P}} QV_{t,T}. \tag{3.21}$$

Suppose in addition that for $\overline{\eta}_T$ *in (3.20):*

$$\frac{\overline{\eta}_T}{\sqrt{T}} \to 0 \quad \textit{and} \quad \frac{\overline{\eta}_T}{T} \to \infty. \tag{3.22}$$

Then, we also have

$$\widehat{TV}_{t,T}(\widehat{\eta}_T) \xrightarrow{\mathbb{P}} V_t. \tag{3.23}$$

The result in (3.21) is of independent interest and can be used for making inference for the jump part of X. We note that we can further derive a CLT associated with the convergence in (3.21) which will allow us to assess the precision in the recovery of the jump part of the quadratic variation.

3.5. *Feasible CLT for the Characteristic Function Based Volatility Estimator.* Theorem 1 allows for the sequence (u_T) to be random. However, it restricts (u_T) to be $\mathcal{F}_t^{(0)}$-adapted and this rules out the case where (u_T) depends on the option data used in the construction of $\widehat{V}_{t,T}(u_T)$. The goal of this section is to make the volatility estimator adaptive by making u_T a function of our preliminary truncated volatility $\widehat{TV}_{t,T}(\widehat{\eta}_T)$. In particular, we set the characteristic exponent in the construction of $\widehat{V}_{t,T}(u)$ in the following data-driven way (recall our discussion at the beginning of Section 3.4)

$$\widehat{u}_T = \frac{\overline{u}}{\sqrt{T}} \frac{1}{\sqrt{\widehat{TV}_{t,T}(\widehat{\eta}_T)}}, \tag{3.24}$$

where $\overline{u}$ is some fixed positive number that does not depend on the data.

We will further derive a feasible CLT for $\widehat{V}_{t,T}(\widehat{u}_T)$ and for this we will need a consistent estimator for its conditional asymptotic variance. We now introduce the necessary notation for this. First, our estimates for the variance of the observation error are based on

$$\widehat{\epsilon}_j = \widehat{\kappa}_T(k_j) - \frac{1}{2}\left(\widehat{\kappa}_T(k_{j-1}) + \widehat{\kappa}_T(k_{j+1})\right), \ j = 2, ..., N-1 \text{ and } j \neq j^*, \tag{3.25}$$

where $j^* \in \{1, ..., N\}$ with $|k_{j^*} - x_t| \leq |k_j - x_t|$, for $j = 1, ..., N$, and

$$\widehat{\epsilon}_1 = \widehat{\epsilon}_2 \quad \text{and} \quad \widehat{\epsilon}_{N-1} = \widehat{\epsilon}_N,$$

$$\widehat{\epsilon}_{j^*} = \widehat{\kappa}_T(k_{j^*}) - \widehat{\kappa}_T(k_{j^*-1}) - (\widehat{\kappa}_T(k_{j^*-1}) - \widehat{\kappa}_T(k_{j^*-2}))\frac{K_{j^*} - K_{j^*-1}}{K_{j^*-1} - K_{j^*-2}}, \ \text{ if } k_{j^*} \leq x_t,$$

$$\widehat{\epsilon}_{j^*} = \widehat{\kappa}_T(k_{j^*}) - \widehat{\kappa}_T(k_{j^*+1}) - (\widehat{\kappa}_T(k_{j^*+1}) - \widehat{\kappa}_T(k_{j^*+2}))\frac{K_{j^*} - K_{j^*+1}}{K_{j^*+1} - K_{j^*+2}}, \ \text{ if } k_{j^*} > x_t.$$

Since the true option price is smooth in k, then for $j = 2, ..., N-1$ and $j \neq j^*$, $\widehat{\epsilon}_j$ is an estimate of $\epsilon_j - \frac{1}{2}(\epsilon_{j-1} + \epsilon_{j+1})$. We use a different estimate for the error associated with the available option with strike closest to the current price level. This is done so that we can incorporate the no-arbitrage restriction that the option price is a monotone function of its strike (decreasing for calls and increasing for puts).

Given $\{\widehat{\epsilon}_j\}_{j=2,\dots,N}$, we set

$$\widehat{C}_{t,T}(u) = \frac{2}{3}\sum_{j=2}^{N}\zeta_{j-1}(u)\zeta_{j-1}(u)^\top e^{-2k_{j-1}}\widehat{\epsilon}_{j-1}^2\Delta_j^2, \tag{3.26}$$

$$\zeta_j(u) = \begin{pmatrix} u^2\cos(uk_j - ux_t) - u\sin(uk_j - ux_t) \\ u\cos(uk_j - ux_t) + u^2\sin(uk_j - ux_t)\end{pmatrix}, \quad j = 1, ..., N,$$

and with it our estimate for the asymptotic variance is given by

$$\widehat{Avar}(\widehat{V}_{t,T}(u)) = \frac{4}{T^2u^4}\frac{1}{|\widehat{f}_{t,T}(u)|^4}\begin{pmatrix}\Re\widehat{f}_{t,T}(u) \\ \Im\widehat{f}_{t,T}(u)\end{pmatrix}^\top \widehat{C}_{t,T}(u)\begin{pmatrix}\Re\widehat{f}_{t,T}(u) \\ \Im\widehat{f}_{t,T}(u)\end{pmatrix}. \tag{3.27}$$

Theorem 3 gives a feasible CLT for $\widehat{V}_{t,T}(\widehat{u}_T)$. Below, $\mathcal{L} - s$ denotes stable convergence, i.e., convergence in law that holds jointly with any bounded positive random variable defined on the probability space, see e.g., [28] for further details.

Theorem 3. *Suppose assumptions A1-A5 in Section 3.2 hold for some $r \in [0,1]$ and in addition $\overline{\Delta} \asymp T^{\alpha}$, $\overline{K} \asymp T^{-\beta}$, $\underline{K} \asymp T^{\gamma}$ for some $\alpha > \frac{1}{2}$, $\beta > 0$ and $\gamma > 0$. If (3.22) holds and*

$$\alpha < \left(\frac{1}{2} + 2 - r\right)\bigwedge\left(\frac{1}{2} + 4(\beta\wedge\gamma)\right), \tag{3.28}$$

then

$$\frac{\widehat{V}_{t,T}(\widehat{u}_T) - V_t}{\sqrt{\widehat{Avar}(\widehat{V}_{t,T}(\widehat{u}_T))}} \xrightarrow{\mathcal{L}-s} N(0,1), \tag{3.29}$$

where the limit is defined on an extension of the original probability space and is independent of $\mathcal{F}$.

The condition in (3.28) ensures that the leading term in the difference $\widehat{V}_{t,T}(\widehat{u}_T) - V_t$ is due to the option observation error, and in particular that the biases in the estimation due to the separation of volatility from jumps and the finiteness of the strike range are of higher asymptotic order.

We note that the asymptotic limit of $\widehat{Avar}(\widehat{V}_{t,T}(\widehat{u}_T))$ (after appropriately rescaling it) is in general random. That is, the asymptotic limit of $\widehat{V}_{t,T}(\widehat{u}_T) - V_t$ is mixed Gaussian. This reflects the fact that the precision in the recovery of the random V_t is itself random. This mirrors the limit behavior of the return-based volatility estimators ([9, 10] and [31, 32]).

We further point out that the limit above is of self-normalizing type. That is, we do not establish the limit of appropriately scaled $\widehat{V}_{t,T}(\widehat{u}_T) - V_t$ and $\sqrt{\widehat{Avar}(\widehat{V}_{t,T}(\widehat{u}_T))}$ but only of their ratio. This allows, in particular, to incorporate general setups for the observed strike grid.

We finish this section by comparing the performance of our estimator with one constructed from high-frequency return data. We will use a local (in time) version of the truncated variance of [31, 32] in the comparison, but the results extend also to other return-based volatility estimators, e.g., the multipower variations of [9, 10]. The return-based truncated volatility estimator is given by

$$\widehat{V}_t^{hf} = \frac{n}{k_n}\sum_{i\in I_t^n}(\Delta_i^n x)^2 1_{\{|\Delta_i^n x|\leq \alpha n^{-\varpi}\}}, \quad \alpha > 0 \text{ and } \varpi\in(0,1/2), \tag{3.30}$$

where $I_t^n = \{i = 1, ..., k_n : \lfloor tn\rfloor - i\}$ denotes a local window used for the calculation of the volatility and $\Delta_i^n x = x_{\frac{i}{n}} - x_{\frac{i-1}{n}}$. An estimator for the asymptotic variance of $\widehat{V}_t^{hf}$ can be constructed as follows

$$\widehat{Avar}(\widehat{V}_{t,T}^{hf}) = \frac{2}{3}\frac{n^2}{k_n^2}\sum_{i\in I_t^n}(\Delta_i^n x)^4 1_{\{|\Delta_i^n x|\leq \alpha n^{-\varpi}\}}. \tag{3.31}$$

For the successful application of $\widehat{V}_t^{hf}$, it is important to set α in a data-driven way that accounts for the current level of volatility. In the next theorem, we show that the convergence of $\widehat{V}_{t,T}(\widehat{u}_T)$ holds jointly with that of $\widehat{V}_t^{hf}$.

THEOREM 4. *In addition to the conditions of Theorem 3 suppose also that assumption A6 in Section 3.2 holds and $k_n \asymp \sqrt{n}$. Then*

$$\frac{\widehat{V}_t^{hf} - V_t}{\sqrt{\widehat{Avar}(\widehat{V}_{t,T}^{hf})}} \xrightarrow{\mathcal{L}-s} N(0,1), \tag{3.32}$$

and this convergence holds jointly with the convergence in (3.29), with the limits defined on an extension of the original probability space and being independent of each other and of $\mathcal{F}$.

We note that $\widehat{V}_{t,T}(\widehat{u}_T)$ and $\widehat{V}_t^{hf}$ are only $\mathcal{F}$-conditionally independent of each other but, due to connections between their conditional asymptotic variances (which recall are random), they can have dependence unconditionally. The result of Theorem 4 suggests that we can benefit from combining the two volatility estimators. Indeed, we can optimally weight them (note that the weights given below are in general random variables and hence we need to use the fact that the convergence of the two estimators holds stably) according to their $\mathcal{F}$-conditional asymptotic variances, to get

(3.33)
$$\widehat{V}_t^{mix} = \widehat{w}_t \widehat{V}_{t,T}(\widehat{u}_T) + (1-\widehat{w}_t)\widehat{V}_t^{hf}, \quad \widehat{w}_t = \frac{\widehat{Avar}(\widehat{V}_{t,T}^{hf})}{\widehat{Avar}(\widehat{V}_{t,T}^{hf}) + \widehat{Avar}(\widehat{V}_{t,T}(\widehat{u}_T))}.$$

The noisier one of the two volatility estimators is, the less weight it receives in the combined estimator $\widehat{V}_t^{mix}$. In fact, if one of the two estimators converges at a faster rate, then asymptotically the weight it receives in $\widehat{V}_t^{mix}$ converges to one, i.e., it receives all the weight. A convenient feature of $\widehat{V}_t^{mix}$ is that the user does not need to take a stand on whether options or high-frequency returns are more efficient for recovering volatility at any point in time. The optimally weighted $\widehat{V}_t^{mix}$ automatically "adapts" to the situation at hand.

4. Minimax Risk for Recovering Volatility from Noisy Short-Dated Options.

We will now derive a lower bound for the minimax risk for recovering the spot volatility from noisy short-dated option data. This result will show that our nonparametric estimator $\widehat{V}_{t,T}(\widehat{u}_T)$ is near rate-efficient. We first introduce the necessary notation for stating the formal result. We will specialize attention to the case where x is a Lévy process (under the risk-neutral measure) with finite activity jumps, and hence we will drop the subscript t in the notation of the diffusive volatility and the jump compensator here. We will define the set $\mathcal{G}(R)$ of risk-neutral probability measures $\mathbb{Q}$ (under which the true option prices $\kappa_T(k)$ are computed according to (2.2)) for which x is a Lévy process with characteristics triplet with respect to the identity truncation function ([39], Definition 8.2) given by

$$\left(-\frac{1}{2}\sigma^2 - \int_{\mathbb{R}} (e^z - 1 - z)\nu(z)dz,\ \sigma^2,\ F(z)\right), \tag{4.1}$$

where $F(dz) = \nu(z)dz$ and we further have

$$\frac{1}{R} \le |\sigma| \le R, \quad \int_{\mathbb{R}} \left((e^{3|z|} - 1) \vee 1\right) \nu(z)dz \le R, \tag{4.2}$$

for some constant $R > 0$.

The option observations are given by

$$\widehat{\kappa}_T(k_i) = \kappa_T(k_i) + (\kappa_T(k_i) \vee T)\epsilon_i, \quad i = 1, ..., N, \tag{4.3}$$

where $\{\epsilon\}_{i\geq 1}$ is a sequence of i.i.d. $N(0,1)$ random variables defined on a product extension of $(\Omega^{(0)}, \mathcal{F}^{(0)}, (\mathcal{F}_t^{(0)})_{t\geq 0}, \mathbb{P}^{(0)})$ and independent of $\mathcal{F}^{(0)}$. One can show that the option prices for every strike are of order $O_p(T)$. Therefore, the truncation from below in the scale of the option error does not change its asymptotic order.

In what follows, we will denote with $\mathbb{E}_{\mathcal{T}}$ expectations under which the true (unobservable) option prices $\kappa_T(k)$ are computed according to the risk-neutral probability measure $\mathcal{T}$.

THEOREM 5. *In the setting of (4.1)-(4.3), assume further that* $\eta\overline{\Delta} \leq \Delta_i \leq \overline{\Delta}$, *for* $i = 1, ..., N$ *and some* $\eta \in (0,1]$ *and* $\overline{\Delta} > 0$. *Let* $\overline{\Delta} \asymp T^{\alpha}$, $\overline{K} \asymp T^{-\beta}$ *and* $\underline{K} \asymp T^{\gamma}$, *for* $\frac{1}{2} < \alpha < \frac{5}{2}$ *and* $\beta, \gamma > 0$ *as* $T \downarrow 0$.

We then have

$$\inf_{\widehat{\sigma}} \sup_{\mathcal{T}\in\mathcal{G}(R)} \mathbb{E}_{\mathcal{T}}\left(\frac{\sqrt{T}|\ln T|^{5/2}}{\overline{\Delta}}|\widehat{\sigma}-\sigma|^2\right) \geq c, \tag{4.4}$$

for some $c > 0$ *and where* $\widehat{\sigma}$ *is any estimator of* σ *based on the option data* $\{\widehat{\kappa}_T(k_i)\}_{i=1,...,N}$.

Under the conditions of Theorem 3, one can show that $\widehat{V}_{t,T}(\widehat{u}_T) - V_t = O_p\left(\frac{\sqrt{\overline{\Delta}}}{T^{1/4}}\right)$. Comparing this result with the efficient rate of convergence in Theorem 5 (in the special setting of that theorem), we see that $\widehat{V}_{t,T}(\widehat{u}_T)$ is near rate-optimal, i.e., its rate convergence is slower than the optimal one only by a log term.

5. Monte Carlo Study.

We now test the performance of the developed nonparametric techniques on simulated data. In order to generate option data, we need a parametric model for the risk-neutral dynamics of X. We use the following specification:

$$X_t = X_0 + \int_0^t \sqrt{V_s}dW_s + \int_0^t\int_{\mathbb{R}}(e^x-1)\widetilde{\mu}(ds,dx), \tag{5.5}$$

with W being a Brownian motion and V having the dynamics

$$dV_t = 3.6(0.02 - V_t)dt - 0.1\sqrt{V_t}dW_t + 0.2\sqrt{0.75}\sqrt{V_t}d\widetilde{W}_t, \tag{5.6}$$

where $\widetilde{W}$ is a Brownian motion orthogonal to W. The jump measure μ has a compensator $\nu_t(x)dt \otimes dx$ with ν_t given by

$$\nu_t(x) = c_- V_t \frac{e^{-20|x|}}{|x|^{0.5}}1_{\{x<0\}} + c_+ V_t \frac{e^{-100|x|}}{|x|^{0.5}}1_{\{x>0\}}. \tag{5.7}$$

The specification in (5.5)-(5.7) belongs to the affine class of models ([22]) commonly used in empirical option pricing work. Consistent with existing empirical evidence, the jumps have time-varying jump intensity. The jump size distribution is like the one of a tempered stable process which is found to provide good fit to observed option data. We set the model parameters in a way that results in option prices similar to observed equity index options. In particular, the parameter specification of V implies average annualized volatility of around 15% (our unit of time is one year) and negative correlation between the innovations in price and stochastic volatility (also known as leverage effect).

The parameters of the jump distribution are set in a way that produces jump tail behavior similar to that found in market index option data, see e.g., [5]. We consider three cases for $c_{\pm}$. In each of them, the ratio of expected negative to positive jump variation is 10 to 1, similar to what is found in the data. The different cases are of low, medium and high value of the jump variation, corresponding to total expected jump variation being $\frac{1}{4}$, $\frac{1}{2}$ and $\frac{3}{4}$, respectively, of the expected diffusive variation. The values of $c_{\pm}$ in the different cases are given in Table 1. Finally, we set $X_0 = 2000$ and draw V_0 from the stationary distribution of V_t under $\mathbb{Q}$ (which is Gamma distribution with shape and scale parameters of 3.6 and 0.02/3.6).

Case	c_-	c_+
L	0.3058×10^3	1.7097×10^3
M	0.6177×10^3	3.4194×10^3
H	0.9174×10^3	5.1291×10^3

TABLE 1
Monte Carlo Jump Parameter Settings.

We continue next with specifying our option observation scheme. The strike grid, strike range and the total number of options at a given point in time are calibrated to match roughly available S&P 500 index option data. In particular, we set $\underline{k} = -8 \times \sigma_{ATM}\sqrt{T}$, where we denote with σ_{ATM} the Black-Scholes implied volatility of the ATM option. We then set the strikes on an equidistant grid in increments of 5, exactly as for the available S&P 500 index option data. That is, we set $e^{k_i} = e^{k_{i-1}} + 5$ for $i = 2, ..., N$ and where $N = \inf\{i : k_i > 2.5 \times \sigma_{ATM}\sqrt{T}\}$. This way, we have approximately $\overline{k} = 2.5 \times \sigma_{ATM}\sqrt{T}$. We add observation error to the model-implied option prices equal to $\epsilon_i = 0.05 \times Z_i \times \kappa_T(k_i)$, where $\{Z_i\}_{i=1,...,N}$ is a sequence of i.i.d. standard normal random variables.

To implement the option-based volatility estimator $\widehat{V}_{t,T}(\widehat{u}_T)$ on the simulated data, we need to set $\overline{\eta}_T$ for the preliminary truncated volatility as well as $\overline{u}$ for the adaptive characteristic exponent $\widehat{u}_T$ in (3.24). Recall that $\overline{\eta}_T$ is a deterministic sequence converging to zero. We put $\overline{\eta}_T = T^{0.51}$, which for $T = 2/252$ takes value of approximately 0.085 and for $T = 5/252$ takes value of approximately 0.135. Our choice for $\overline{\eta}_T$ is motivated by the maximum downward bias in the measurement of volatility. By first-order Taylor expansion, in the case of no jumps, $3\overline{\eta}_T$ is equal approximately to the relative negative bias in the measurement of the spot variance by $\widehat{TV}_{t,T}(\widehat{\eta}_T)$.

Turning next to $\overline{u}$, we would like to pick this constant as low as possible to guard against the effect of the jumps while at the same time high enough so that the estimation is not too noisy. We experiment with two values, $\overline{u} = \sqrt{2\ln(1/0.1)}$ and $\overline{u} = \sqrt{2\ln(1/0.085)}$, which correspond to $|\mathbb{E}_t^{\mathbb{Q}}\big(e^{iu(x_{t+T}-x_t)}\big)|$ having values of 0.1 and 0.085, respectively (recall from the discussion in Section 2 that for high u we have $|\mathbb{E}_t^{\mathbb{Q}}\big(e^{iu(x_{t+T}-x_t)}\big)| \approx e^{-u^2\frac{V_tT}{2}}$). Finally, to guard against potential finite sample distortions in the data-driven choice of u, if the above choice of $\widehat{u}_T$ exceeds $\widehat{u}_{min} = \mathrm{argmin}_{u\in[0,400]}|\widehat{f}_T(u)|$, we set $\widehat{u}_T$ equal to the latter.

The results from the Monte Carlo are summarized in Tables 2 and 3. Overall, the performance of the option-based volatility estimator on the simulated data is consistent with theory. The estimator is nearly unbiased and volatility is recovered with good precision. The estimation tends to be a bit noisier for the shorted-dated options. This is because the number of available options decreases with T. Also, the precision of the estimator is lower for the higher of the two choices for $\overline{u}$. This is expected as the characteristic function is harder to estimate for higher values of its argument (in the sense of larger associated asymptotic variance).

Case	**Bias**	**MAE**	**Bias**	**MAE**	**Bias**	**MAE**
	$T = 2$ days		$T = 3$ days		$T = 5$ days	
			$\overline{u} = \sqrt{2\ln(1/0.085)}$			
L	0.0006	0.0034	0.0008	0.0034	0.0013	0.0035
M	0.0001	0.0034	0.0002	0.0031	0.0007	0.0030
H	−0.0001	0.0031	−0.0002	0.0028	0.0001	0.0024
			$\overline{u} = \sqrt{2\ln(1/0.1)}$			
L	0.0005	0.0029	0.0004	0.0029	0.0008	0.0030
M	0.0000	0.0028	−0.0001	0.0027	0.0001	0.0026
H	−0.0002	0.0026	−0.0005	0.0024	−0.0005	0.0022

TABLE 2
Monte Carlo Results: Bias and MAE. MAE stands for mean absolute error.

We turn next to the performance of the confidence intervals based on our inference theory which we summarize in Table 3. For the case of $\overline{u} = \sqrt{2\ln(1/0.085)}$, we can notice some under-coverage which is most severe for maturity $T = 5$ and the low jump variation scenario. As we mentioned above, for larger $\overline{u}$, the precision of the estimation is somewhat lower. On the other hand, for the case of $\overline{u} = \sqrt{2\ln(1/0.1)}$, the coverage rates are close to the nominal ones for the various Monte Carlo scenarios.

Case	**Coverage Rate**						**Coverage Rate**					
	0.90	0.95	0.90	0.95	0.90	0.95	0.90	0.95	0.90	0.95	0.90	0.95
	$T = 2$ days		$T = 3$ days		$T = 5$ days		$T = 2$ days		$T = 3$ days		$T = 5$ days	
	$\overline{u} = \sqrt{2\ln(1/0.085)}$						$\overline{u} = \sqrt{2\ln(1/0.1)}$					
L	0.90	0.92	0.85	0.90	0.82	0.86	0.90	0.93	0.88	0.91	0.84	0.89
M	0.91	0.95	0.90	0.92	0.86	0.90	0.92	0.95	0.91	0.94	0.90	0.93
H	0.93	0.95	0.91	0.93	0.89	0.91	0.94	0.96	0.92	0.96	0.91	0.95

TABLE 3
Monte Carlo Results: Coverage Probability.

Overall, the results from the Monte Carlo reveal satisfactory performance of the option-based nonparametric volatility estimation in empirically realistic settings.

6. Empirical Application. We next apply our nonparametric volatility procedure on real data. Our sample covers the period 2016-2017 and the underlying asset of the options in our analysis is the S&P 500 index. We use quotes on so-called weekly options traded on CBOE options exchange (which are settled at the end of the regular trading hours) recorded at market close on each Wednesday which is not a holiday. We take the shortest to maturity available options on each day with time to maturity between 2 and 5 business days. The median number of strikes per date in our option data set is 54 while on more than 84% of the days the time to maturity is 2 days. In addition to the options, we also make use of intraday data on the E-mini S&P 500 futures contract (with maturity closest to expiration) to construct return-based volatility estimates. The sampling frequency of the S&P 500 futures is five minutes which is sufficiently coarse to guard against the impact of microstructure related issues.

The tuning parameters of the option-based volatility estimator are chosen exactly as in the Monte Carlo (we use $\overline{u} = 0.1$). For the return-based estimator, following common practice, we set the truncation level in a data-driven way. More specifically, we set α and ϖ in (3.30) to

$$\alpha = 3\sqrt{RV_t \wedge BV_t} \quad \text{and} \quad \varpi = 0.49, \tag{6.8}$$

where

$$(6.9)\qquad RV_t = \sum_{i=\lfloor (t-1)n\rfloor+1}^{\lfloor tn\rfloor} (\Delta_i^n x)^2, \quad BV_t = \frac{\pi}{2}\sum_{i=\lfloor (t-1)n\rfloor+2}^{\lfloor tn\rfloor} |\Delta_{i-1}^n x||\Delta_i^n x|,$$

with RV_t being the realized volatility and BV_t being the Bipower Variation of [9, 10] over the trading day. The latter is a nonparametric jump-robust measure of integrated volatility that is free of tuning parameters. Finally, our local window for $\widehat{V}_t^{hf}$ consists of $k_n = 48$ five-minute returns before market close (which is the time when the option data is recorded).

In Figure 3, we plot the two spot volatility estimators. As seen from the figure, the two series are very close to each other on average. Indeed, the sample median of the option-based volatility estimate is within 1.5% of that of the return-based volatility estimator while the correlation between the two time series is 0.9. At the same time, though, we can note that the volatility estimate from the high-frequency data is significantly noisier. Indeed, the standard deviation of the return-based volatility estimate is over 32% higher than that of its option-based counterpart. In addition, the correlation between the first differences of the two estimators, for which the measurement error plays bigger role, is only 0.6 (first differences are used in the computation of measures of variation such as the quadratic variation). Also, the efficiency gains offered by the option-based estimator (based on the estimated asymptotic variances) are particularly pronounced at the beginning of the sample when the volatility was very high. This is to be expected as during such episodes the separation of volatility from the realization of jumps from return data is particularly challenging. Finally, $\widehat{V}_t^{mix}$ in (3.33) that combines optimally the option and return-based volatility estimators is very close to the former, with correlation between $\widehat{V}_{t,T}(\widehat{u}_T)$ and $\widehat{V}_t^{\mathrm{mix}}$ of over 0.99. This is due to the high weight assigned to $\widehat{V}_{t,T}(\widehat{u}_T)$ in forming $\widehat{V}_t^{\mathrm{mix}}$, particularly in the high volatility period.

Overall, the empirical analysis reveals nontrivial gains in measuring spot volatility by the use of short-dated options. The newly-proposed nonparametric volatility estimator should therefore greatly improve the precision in studying various features of the volatility process, e.g., the roughness of the volatility path (see e.g., [25]) as well as the presence of jumps in the volatility and their connection with those in the underlying price (see e.g., [29] and [3]). Answering these questions regarding the volatility trajectory from return data alone is known to be very difficult as the volatility is not directly observed and has to be filtered out from the data. This can be particularly challenging in the presence of persistent microstructure-related

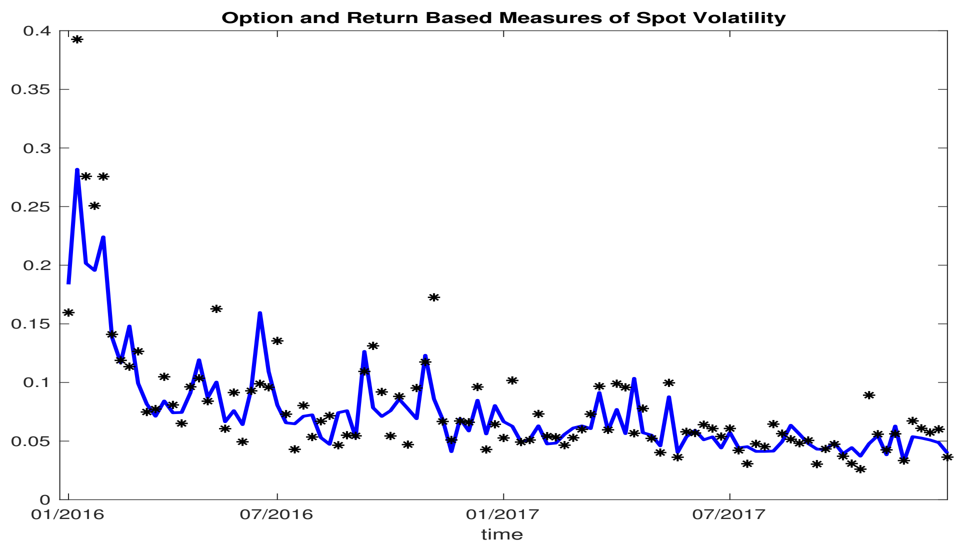

FIG 3. *Option and Return Based Measures of Volatility. The solid line corresponds to the option based one,* $\sqrt{\widehat{V}_{t,T}(\widehat{u}_T)}$*, and the stars to the one based on five-minute returns of the underlying asset,* $\sqrt{\widehat{V}_t^{hf}}$*. The x-axis ticks are at the beginning of the corresponding month.*

distortions in the high-frequency returns, see e.g., the recent work of [17]. In addition, the option-based spot volatility estimates should be of direct use for the purposes of volatility forecasting and risk management where a more precise volatility proxy is known to provide efficiency gains, see e.g., [2] in the case of volatility forecasting. In current work in progress I show this to be the case when using the newly-developed option-based volatility estimator for forecasting the future volatility of various assets.

7. Proofs.

In the proofs we will denote with C_t a finite-valued and $\mathcal{F}_t$-adapted random variable which might change from line to line. If the variable depends on some parameter q, then we will use the notation $C_t(q)$. Further, without loss of generality, in the proofs, we will set $X_t = 1$ or equivalently $x_t = 0$.

7.1. *Decomposition, Notation and Auxiliary Results.*

The jump part of the process x_t can be represented as an integral with respect to a Poisson random measure under $\mathbb{Q}$. In particular, using the so-called Grigelionis representation of the jump part of a semimartingale (Theorem 2.1.2 of [27]) and upon suitably extending the probability space, we can write

$$\int_0^t\int_{\mathbb{R}}\widetilde{\mu}(ds,dx)\equiv\int_0^t\int_E\delta^x(s,z)\widetilde{\mu}^x(ds,dz), \tag{7.1}$$

where $\mu^x(ds,dz)$ is a Poisson measure on $\mathbb{R}_+\times E$ with compensator $dt\otimes\lambda(dz)$ for some sigma-finite measure λ 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 $\lambda(dz)$ under the map $z\to\delta^x(t,z)$ on the set $\{z:\delta^x(\omega,t,z)\neq 0\}$.

There are different choices for E and the function δ^x. For the analysis here it will be convenient to use $\mathbb{E}=\mathbb{R}_+\times\mathbb{R}$ and $\delta^x(t,z)=z_2 1_{\{z_1\le\nu_t(z_2)\}}$ for $z=(z_1,z_2)$, with λ being the Lebesgue measure on E.

We proceed with introducing some notation that will be used throughout the proofs. By noting that $x_t=0$, we can split x_s into

$$x_s^c=\int_t^s a_u du+\int_t^s\sigma_u dW_u,\quad x_s^d=\int_t^s\int_E\delta^x(u,z)\widetilde{\mu}^x(du,dz),\quad s\ge t. \tag{7.2}$$

We now introduce two approximations for x_s. The first is $\widetilde{x}_s=\widetilde{x}_s^c+\widetilde{x}_s^d$, where for $s\ge t$

$$\widetilde{x}_s^c=a_t(s-t)+\sigma_t(W_s-W_t),\quad \widetilde{x}_s^d=\int_t^s\int_E\delta^x(t,z)\widetilde{\mu}^x(du,dz). \tag{7.3}$$

The second approximation is given by $\overline{x}_s=\overline{x}_s^c+\overline{x}_s^d$, where for $s\ge t$

$$\overline{x}_s^c=a_t(s-t)+\int_t^s\overline{\sigma}_u W_u,\quad \overline{x}_s^d=\widetilde{x}_s^d,\quad \overline{\sigma}_s=\sigma_t+\eta_t(W_s-W_t)+\overline{\eta}_t(\overline{W}_s-\overline{W}_t). \tag{7.4}$$

The OTM option prices at time t associated with log-terminal value $\widetilde{x}_{t+T}$ are denoted by $\widetilde{\kappa}_T(k)$, the ones with log-terminal value of $\overline{x}_{t+T}$ are denoted by $\overline{\kappa}_T(k)$, and the ones with log-terminal value of $\sigma_t(W_{t+T}-W_t)$ with $\widetilde{\kappa}_T^c(k)$.

Lemma 1. *Suppose assumptions A1-A3 hold. There exist $\mathcal{F}_t$-adapted random variables $\bar{t}>t$ and $C_t>0$ that do not depend on k, k_1, k_2 and T, such that for $T<\bar{t}$ we have*

$$\kappa_T(k)\le C_tT\begin{cases}\frac{e^{2k}}{e^{-k}-1}, & \text{if } k<0,\\ \frac{1}{e^k-1}, & \text{if } k>0,\end{cases} \tag{7.5}$$

$$|\kappa_T(k)-\overline{\kappa}_T(k)|\le C_t|\ln T|T^{3/2}, \tag{7.6}$$

$$|\overline{\kappa}_T(k) - \widetilde{\kappa}_T(k)| \le C_t \left(T^{\frac{3}{2}} \bigvee \left(\frac{T^{3/2}}{|e^k - 1|} \bigwedge T \right) \right), \tag{7.7}$$

$$\widetilde{\kappa}_T(k) \le C_t \left(\sqrt{T} \bigwedge \frac{T}{|e^k - 1|} \right), \tag{7.8}$$

$$\begin{cases} k < k_{l,t} \implies \widetilde{\kappa}_T(k) \le C_t \frac{e^{2k}}{\left(e^{-k-k_{l,t}}-1\right)^2} T, & k_{l,t} = -\sigma_t \sqrt{T} |\ln T|, \\ k > k_{h,t} \implies \widetilde{\kappa}_T(k) \le C_t \frac{1}{\left(e^{k-k_{h,t}}-1\right)^2} T, & k_{h,t} = \sigma_t \sqrt{T} |\ln T|, \end{cases} \tag{7.9}$$

$$|\widetilde{\kappa}_T(k_1) - \widetilde{\kappa}_T(k_2)| \le C_t \left[\left(\frac{T}{k_2^2} \bigwedge 1 \right) 1_{\{|k_2| \le 1\}} + \frac{T}{k_2^4} 1_{\{|k_2|>1\}} \right] \left| e^{k_1} - e^{k_2} \right|, \tag{7.10}$$

$$|\kappa_T(k_1) - \kappa_T(k_2)| \le C_t \left[\left(\frac{T}{k_2^2} \bigwedge 1 \right) 1_{\{|k_2| \le 1\}} + \frac{T}{k_2^4} 1_{\{|k_2|>1\}} \right] \left| e^{k_1} - e^{k_2} \right|, \tag{7.11}$$

where $k_1 < k_2 < 0$ or $k_1 > k_2 > 0$.

Proof of Lemma 1. All bounds but the last one are proved in Lemmas 2-7 of [37]. The bound in (7.11) can be proved exactly as Lemma 7 of [37] using the integrability assumptions in A2-r. ☐

Lemma 2. *Suppose assumptions A1-A3 hold. There exist $\mathcal{F}_t$-adapted random variables $\overline{t} > t$ and $C_t > 0$ that do not depend on k and T, such that for $T < \overline{t}$ we have*

$$|\widetilde{\kappa}_T(k) - \widetilde{\kappa}^c_T(k)| \le C_t T, \tag{7.12}$$

(7.13)

$$\left| \widetilde{\kappa}^c_T(k) - f\left(\frac{k - a_t T}{\sqrt{T}\sigma_t} \right) \sqrt{T}\sigma_t - (e^k - 1)\Phi\left(\frac{k - a_t T}{\sqrt{T}\sigma_t} \right) \right| \le C_t T, \quad \text{if } k \le 0,$$

$$\left| \widetilde{\kappa}^c_T(k) - f\left(\frac{k - a_t T}{\sqrt{T}\sigma_t} \right) \sqrt{T}\sigma_t + (e^k - 1)\left(1 - \Phi\left(\frac{k - a_t T}{\sqrt{T}\sigma_t} \right) \right) \right| \le C_t T, \quad \text{if } k > 0,$$

where f and Φ are the pdf and cdf, respectively, of a standard normal random variable.

Proof of Lemma 2. We look only at the case $k > 0$, with the proof for the case $k \le 0$ being done in analogous way. First, we have

$$\left| \left| e^{\widetilde{x}_{t+T}} - e^k \right|^+ - \left| e^{\sigma_t (W_{t+T} - W_t)} - e^k \right|^+ \right| \le e^{\sigma_t (W_{t+T} - W_t)} |e^{a_t T + \widetilde{x}^d_{t+T}} - 1|.$$

From here we can use the $\mathcal{F}_t$-conditional independence of $W_{t+T} - W_t$ and $\widetilde{x}^d_{t+T}$ and apply Lemma 1 of [37] to obtain the result in (7.12).

We continue with the bounds in (7.13). Direct calculation shows for $k > 0$

$$\widetilde{\kappa}^c_T(k) = e^{a_tT+\sigma_t^2T/2}\left(1-\Phi\left(\frac{k-a_tT}{\sqrt{T}\sigma_t}-\sqrt{T}\sigma_t\right)\right)-e^k\left(1-\Phi\left(\frac{k-a_tT}{\sqrt{T}\sigma_t}\right)\right).$$

From here the result of (7.13) follows by using Taylor expansion and the fact that the function f is bounded. □

7.2. *Proof of Theorem 1.* We introduce the following notation

$$f_{t,T}(u) = \mathbb{E}^{\mathbb{Q}}_t\left(e^{iu(x_{t+T}-x_t)}\right), \quad \widetilde{f}_{t,T}(u) = \mathbb{E}^{\mathbb{Q}}_t\left(e^{iu(\widetilde{x}_{t+T}-\widetilde{x}_t)}\right).$$

Using Appendix 1 of [15], $f_{t,T}(u)$ equals the expression in (3.11). We further note that by Lévy-Khintchine formula

$$\widetilde{f}_{t,T}(u) = \exp\left(iTua_t - T\frac{u^2}{2}V_t + T\int_{\mathbb{R}}(e^{iux}-1-iux)\nu_t(x)dx\right). \tag{7.14}$$

We start the proof with establishing a bound for the difference $f_{t,T}(u) - \widetilde{f}_{t,T}(u)$. In the proof we will denote with $\zeta_{t,T}(u)$ a random variable that depends on u and further satisfies

$$|\zeta_{t,T}(u)| \leq C_t(|u|T^{3/2} \vee |u|^4T^3),$$

where C_t is $\mathcal{F}_t$-adapted random variable that does not depend on u. This variable can change from one line to another.

We first study the real part of the difference $\Re(f_{t,T}(u)-\widetilde{f}_{t,T}(u))$. Applying Itô's lemma, using the normalization $x_t = 0$ and the integrability assumption A2, we have

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}_t(\cos(ux_{t+T})) - 1 = &-\mathbb{E}^{\mathbb{Q}}_t\left(\int_t^{t+T} u\sin(ux_s)a_sds + \frac{1}{2}\int_t^{t+T} u^2\cos(ux_s)\sigma_s^2ds\right)\\ &+\mathbb{E}^{\mathbb{Q}}_t\left(\int_t^{t+T}\int_{\mathbb{R}}\left(\cos(ux_s)(\cos(uz)-1)-\sin(ux_s)(\sin(uz)-u\sin(z))\right)\nu_s(z)dzds\right).\end{aligned}$$

We have an analogous expression for $\mathbb{E}^{\mathbb{Q}}_t(\cos(u\widetilde{x}_{t+T})) - 1$. Then, using assumption A3 as well as $\mathbb{E}^{\mathbb{Q}}_t|x_s - \widetilde{x}_s| \leq C_tT$ for $s \in [t, t+T]$ and C_t being $\mathcal{F}_t$-adapted random variable (which follows from using Doob's inequality and A3), we can write

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}_t(\cos(ux_{t+T})) - 1 = &-\mathbb{E}^{\mathbb{Q}}_t\left(\int_t^{t+T} u\sin(u\widetilde{x}_s)a_tds + \frac{1}{2}\int_t^{t+T} u^2\cos(ux_s)\sigma_s^2ds\right)\\ &+\mathbb{E}^{\mathbb{Q}}_t\left(\int_t^{t+T}\int_{\mathbb{R}}\left(\cos(u\widetilde{x}_s)(\cos(uz)-1)-\sin(u\widetilde{x}_s)(\sin(uz)-u\sin(z))\right)\nu_t(z)dzds\right)\\ &+\zeta_{t,T}(u).\end{aligned}$$

Therefore, we have

$$\begin{aligned}&\mathbb{E}_t^{\mathbb{Q}}(\cos(ux_{t+T})) - \mathbb{E}_t^{\mathbb{Q}}(\cos(u\widetilde{x}_{t+T}))\\&\quad = \frac{1}{2}\mathbb{E}_t^{\mathbb{Q}}\left(\int_t^{t+T} u^2(\cos(u\widetilde{x}_s)\sigma_t^2 - \cos(ux_s)\sigma_s^2)ds\right) + \zeta_{t,T}(u).\end{aligned}$$

Using the proof of Lemma 3 in [37] as well as (3.9) of A3, we have $\mathbb{E}_t^{\mathbb{Q}}|x_s - \overline{x}_s| \leq C_tT^{3/2}$ for $s \in [t, t+T]$ and C_t being $\mathcal{F}_t$-adapted random variable. In addition, using assumptions A1-A3 and Doob's inequality, we also have $\mathbb{E}_t^{\mathbb{Q}}|\sigma_s - \overline{\sigma}_s| \leq C_tT$ for $s \in [t, t+T]$ and C_t as before. Using these results and Cauchy-Schwartz inequality, we can write

$$\begin{aligned}&\mathbb{E}_t^{\mathbb{Q}}(\cos(ux_{t+T})) - \mathbb{E}_t^{\mathbb{Q}}(\cos(u\widetilde{x}_{t+T}))\\&\quad = \frac{1}{2}\mathbb{E}_t^{\mathbb{Q}}\left(\int_t^{t+T} u^2(\cos(u\widetilde{x}_s)\sigma_t^2 - \cos(u\overline{x}_s)\overline{\sigma}_s^2)ds\right) + \zeta_{t,T}(u).\end{aligned}\tag{7.15}$$

Next, we can decompose

$$\begin{aligned}\cos(u\overline{x}_s)\overline{\sigma}_s^2 - \cos(u\widetilde{x}_s)\sigma_t^2 &= (\cos(u\overline{x}_s) - \cos(u\widetilde{x}_s))\sigma_t^2 + \cos(u\overline{x}_s)(\overline{\sigma}_s - \sigma_t)^2\\&\quad + 2(\cos(u\overline{x}_s) - \cos(u\widetilde{x}_s))(\overline{\sigma}_s - \sigma_t)\sigma_t + 2\cos(u\widetilde{x}_s)(\overline{\sigma}_s - \sigma_t)\sigma_t.\end{aligned}\tag{7.16}$$

For the second and third terms on the right-hand side of the above equality we can use $\mathbb{E}_t^{\mathbb{Q}}(\overline{\sigma}_s - \sigma_t)^2 \leq C_tT$ and $\mathbb{E}_t^{\mathbb{Q}}(\overline{x}_s - \widetilde{x}_s)^2 = \mathbb{E}_t^{\mathbb{Q}}\left(\int_t^s(\overline{\sigma}_s - \sigma_t)dW_u\right)^2 \leq C_tT^2$, for $s \in [t, t+T]$ and C_t being $\mathcal{F}_t$-adapted random variable as well as Cauchy-Schwartz inequality, and conclude

$$u^2T\mathbb{E}_t^{\mathbb{Q}}(\overline{\sigma}_s - \sigma_t)^2 + u^2T\left|\mathbb{E}_t^{\mathbb{Q}}[(\cos(u\overline{x}_s) - \cos(u\widetilde{x}_s))(\overline{\sigma}_s - \sigma_t)]\right| = \zeta_{t,T}(u).$$

For the forth term on the right-hand side of (7.16), we can decompose $\cos(u\widetilde{x}_s) = \cos(u\widetilde{x}_s^c)\cos(u\widetilde{x}_s^d) - \sin(u\widetilde{x}_s^c)\sin(u\widetilde{x}_s^d)$. Then, using the symmetry of the density of the standard normal distribution as well as the independence of W and $\widetilde{W}$, we have

$$\begin{aligned}&\mathbb{E}_t^{\mathbb{Q}}(\cos(u\widetilde{x}_s^c)\cos(u\widetilde{x}_s^d)(\overline{\sigma}_s - \sigma_t))\\&\quad = -\sin(ua_t(s-t))\mathbb{E}_t^{\mathbb{Q}}(\sin(u\sigma_t(W_s - W_t))\cos(u\widetilde{x}_s^d)(\overline{\sigma}_s - \sigma_t)),\end{aligned}$$

which is $\zeta_{t,T}(u)$ because $\mathbb{E}_t^{\mathbb{Q}}|\overline{\sigma}_s - \sigma_t| \leq C_t\sqrt{T}$ for $s \in [t, t+T]$. Using the $\mathcal{F}_t$-conditional independence of $\widetilde{x}_s^d$ and $\overline{\sigma}_s - \sigma_t$, we also have

$$\begin{aligned}\mathbb{E}_t^{\mathbb{Q}}|\sin(u\widetilde{x}_s^c)\sin(u\widetilde{x}_s^d)(\overline{\sigma}_s - \sigma_t)| &\leq |u|\mathbb{E}_t^{\mathbb{Q}}(|\widetilde{x}_s^d||\overline{\sigma}_s - \sigma_t|)\\&= |u|\mathbb{E}_t^{\mathbb{Q}}|\widetilde{x}_s^d|\mathbb{E}_t^{\mathbb{Q}}|\overline{\sigma}_s - \sigma_t| = \zeta_{t,T}(u).\end{aligned}$$

Combining these bounds, we have

$$\mathbb{E}_t^{\mathbb{Q}}(\cos(u\widetilde{x}_s)(\overline{\sigma}_s - \sigma_t)) = \zeta_{t,T}(u).$$

Finally, for the first term on the right-hand side of (7.16) using the independence of W and $\widetilde{W}$ and integration by parts, we can write

$$\begin{aligned}
&\mathbb{E}_t^{\mathbb{Q}}(\cos(u\overline{x}_s) - \cos(u\widetilde{x}_s)) = -\mathbb{E}_t^{\mathbb{Q}}(\sin(u\widetilde{x}_s)\sin(u\overline{x}_s - u\widetilde{x}_s)) + \zeta_{t,T}(u) \\
&\quad = -u\mathbb{E}_t^{\mathbb{Q}}(\sin(u\widetilde{x}_s)(\overline{x}_s - \widetilde{x}_s)) + \zeta_{t,T}(u) = -\frac{u}{2}\eta_t\mathbb{E}_t^{\mathbb{Q}}(\sin(u\widetilde{x}_s)(W_s - W_t)^2) + \zeta_{t,T}(u) \\
&\quad = -\frac{u}{2}\eta_t\mathbb{E}_t^{\mathbb{Q}}(\sin(u\sigma_t(W_s - W_t))(W_s - W_t)^2) + \zeta_{t,T}(u) = \zeta_{t,T}(u),
\end{aligned}$$

where for the last two equalities we have made use of the $\mathcal{F}_t$-conditional independence of $\widetilde{x}_s^c$ and $\widetilde{x}_s^d$ as well as the symmetry of the density of the standard normal distribution. Altogether we get that $\mathbb{E}_t^{\mathbb{Q}}\left(\int_t^{t+T} u^2(\cos(u\overline{x}_s)\overline{\sigma}_s^2 - \cos(u\widetilde{x}_s)\sigma_t^2)ds\right)$ is $\zeta_{t,T}(u)$. Combining this result with (7.15), we have

$$\Re(f_{t,T}(u) - \widetilde{f}_{t,T}(u)) = \zeta_{t,T}(u). \tag{7.17}$$

Turning to $\Im(f_{t,T}(u) - \widetilde{f}_{t,T}(u))$, by making use of $\mathbb{E}_t^{\mathbb{Q}}|x_{t+T} - \widetilde{x}_{t+T}| \leq C_t T$, we have

$$|\mathbb{E}_t^{\mathbb{Q}}(\sin(ux_{t+T})) - \mathbb{E}_t^{\mathbb{Q}}(\sin(u\widetilde{x}_{t+T}))| \leq C_t|u|T. \tag{7.18}$$

The results in (7.17) and (7.18), together with the rate condition for the sequence u_T in (3.15), imply

$$\Re(f_{t,T}(u_T)) - \Re(\widetilde{f}_{t,T}(u_T)) = O_p(T),\ \Im(f_{t,T}(u_T)) - \Im(\widetilde{f}_{t,T}(u_T)) = O_p(\sqrt{T}). \tag{7.19}$$

From (7.14), we also have

$$\Re(\widetilde{f}_{t,T}(u_T)) = O_p(1), \quad \Im(\widetilde{f}_{t,T}(u_T)) = O_p(\sqrt{T}). \tag{7.20}$$

From here, using Taylor expansion, we have

$$\Re\left(\ln\left(f_{t,T}(u_T)\right)\right) - \Re\left(\ln\left(\widetilde{f}_{t,T}(u_T)\right)\right) = O_p(T). \tag{7.21}$$

Furthermore, for r being the constant in assumption A2-r, we have

$$\left|\Re\left(\ln\left(\widetilde{f}_{t,T}(u_T)\right)\right) + T\frac{u_T^2}{2}\sigma_t^2\right| \leq 2T|u_T|^r\int_{\mathbb{R}}|x|^r\nu_t(x)dx. \tag{7.22}$$

Altogether, we have

$$-\frac{2}{Tu_T^2}\Re\left(\ln\left(f_{t,T}(u_T)\right)\right)-\sigma_t^2=O_p(u_T^{r-2}). \tag{7.23}$$

We next decompose $\widehat{f}_{t,T}(u_T)-f_{t,T}(u_T)=\sum_{j=1}^{3}\widehat{f}_{t,T}^{(j)}$, where $\widehat{f}_{t,T}^{(j)}=-(u_T^2+iu_T)T\overline{f}_{t,T}^{(j)}$ and

$$\overline{f}_{t,T}^{(1)}=\frac{1}{T}\sum_{j=2}^{N}e^{(iu_T-1)k_{j-1}}\epsilon_{j-1}\Delta_j,$$

$$\overline{f}_{t,T}^{(2)}=\frac{1}{T}\sum_{j=2}^{N}\int_{k_{j-1}}^{k_j}\left(e^{(iu_T-1)k_{j-1}}\kappa_T(k_{j-1})-e^{(iu_T-1)k}\kappa_T(k)\right)dk,$$

$$\overline{f}_{t,T}^{(3)}=-\frac{1}{T}\int_{-\infty}^{\underline{k}}e^{(iu_T-1)k}\kappa_T(k)dk-\frac{1}{T}\int_{\overline{k}}^{\infty}e^{(iu_T-1)k}\kappa_T(k)dk.$$

Using the bounds of Lemma 1 and assumption A5 for the observation error, we have

$$\overline{f}_{t,T}^{(1)}=O_p\left(\frac{\sqrt{\overline{\Delta}}}{T^{1/4}}\right),\quad \overline{f}_{t,T}^{(2)}=O_p\left(\frac{\overline{\Delta}}{\sqrt{T}}|\ln T|\right). \tag{7.24}$$

Turning to $\overline{f}_{t,T}^{(3)}$, using integration by parts and Lemma 1, we get

$$\begin{aligned}\int_{-\infty}^{\underline{k}}e^{(iu_T-1)k}\kappa_T(k)dk&=\frac{i}{u_T}\int_{-\infty}^{\underline{k}}e^{(iu_T-1)k}(\kappa_T'(k)-\kappa_T(k))dk\\&\quad-\frac{i}{u_T}e^{(iu_T-1)\underline{k}}\kappa_T(\underline{k}).\end{aligned} \tag{7.25}$$

Using Lebesgue dominated convergence and the integrability conditions in Assumption A2, we have

$$\begin{aligned}\kappa_T'(k)=e^k\mathbb{Q}_t\left(x_{t+T}<k\right)&=e^k\mathbb{Q}_t\left(e^{-x_{t+T}}-1<e^{-k}-1\right)\\&\le C_te^k\frac{T}{(e^{-k}-1)^3},\quad k<0.\end{aligned} \tag{7.26}$$

From here by application of Lemma 1, we obtain $\int_{-\infty}^{\underline{k}}e^{(iu_T-1)k}\kappa_T(k)dk=O_p(u_T^{-1}Te^{-2|\underline{k}|})$. Exactly the same analysis can be done for the second integral in $\overline{f}_{t,T}^{(3)}$, and thus altogether we have

$$\overline{f}_{t,T}^{(3)}=O_p\left(u_T^{-1}e^{-2(|\underline{k}|\wedge|\overline{k}|)}\right). \tag{7.27}$$

Combining the bounds in (7.23), (7.24) and (7.27), using Taylor expansion and the rate condition in (3.15) as well as the conditions for the asymptotic behavior of T, $\overline{\Delta}$, $\underline{k}$ and $\overline{k}$ in the theorem, we have (3.16).

7.3. *Proof of Theorem 2.* We set $f_\eta(x) = e^{-\eta x^2}x^2$ for $\eta \geq 0$, and we denote

$$\eta_T = \frac{\overline{\eta}_T}{T}\frac{1}{QV_t}, \tag{7.28}$$

which is an $\mathcal{F}_t$-adapted random variable. Using Appendix 1 of [15], for every finite-valued, nonnegative and $\mathcal{F}_t$-adapted random variable η, we have

$$\mathbb{E}_t^{\mathbb{Q}}\left(f_\eta(\widetilde{x}_{t+T} - \widetilde{x}_t)\right) = \int_{-\infty}^{\infty} h_\eta(k)\widetilde{\kappa}_T(k)dk. \tag{7.29}$$

We will first show that $\mathbb{E}_t^{\mathbb{Q}}\left(f_0(\widetilde{x}_{t+T} - \widetilde{x}_t)\right)$ and $\mathbb{E}_t^{\mathbb{Q}}\left(f_{\eta_T}(\widetilde{x}_{t+T} - \widetilde{x}_t)\right)$ are close to QV_t and V_t, respectively. Applying Itô's lemma, taking expectations and using the integrability conditions of assumption A2-r, we have

(7.30)
$$\mathbb{E}_t^{\mathbb{Q}}\left(f_\eta(\widetilde{x}_{t+T} - \widetilde{x}_t)\right) = a_t\mathbb{E}_t^{\mathbb{Q}}\left(\int_0^T f_\eta'(\widetilde{x}_{t+s} - \widetilde{x}_t)ds\right) + \frac{V_t}{2}\mathbb{E}_t^{\mathbb{Q}}\left(\int_0^T f_\eta''(\widetilde{x}_{t+s} - \widetilde{x}_t)ds\right)$$
$$+ \mathbb{E}_t^{\mathbb{Q}}\left(\int_0^T\int_{\mathbb{R}}\left(f_\eta(\widetilde{x}_{t+s} - \widetilde{x}_t + z) - f_\eta(\widetilde{x}_{t+s} - \widetilde{x}_t) - f_\eta'(\widetilde{x}_{t+s} - \widetilde{x}_t)z\right)ds\nu_t(z)dz\right),$$

for any $\mathcal{F}_t$-adapted η. Using then the fact that (which follows by (7.3) and the integrability conditions in A2-r)

$$|\mathbb{E}_t^{\mathbb{Q}}(\widetilde{x}_{t+s} - \widetilde{x}_t)| \leq C_tT, \quad \text{for } s \in [t, t+T], \tag{7.31}$$

we have

$$\left|\frac{1}{T}\mathbb{E}_t^{\mathbb{Q}}\left(f_0(\widetilde{x}_{t+T} - \widetilde{x}_t)\right) - V_t - \int_{\mathbb{R}} z^2\nu_t(z)dz\right| = O_p\left(T\right). \tag{7.32}$$

Next, for some constant C that does not depend on η and x, we have for $\eta \in \mathbb{R}_+$ and $x \in \mathbb{R}$:

$$|f_\eta'(x)| \leq C|x| \quad \text{and} \quad |f_\eta''(x) - 2| \leq C\eta x^2, \tag{7.33}$$

$$\begin{aligned} &|f_\eta(x+z) - f_\eta(x) - f_\eta'(x)z - f_\eta(z)| \\ &\qquad \leq C\left(|x||z| + \eta|x|^2|z|^2 + \eta|x|^3|z| + e^{-\frac{\eta}{2}|z|^2}|z|^2\right). \end{aligned} \tag{7.34}$$

Using the above bounds and (7.30), the inequality $|x|e^{-|x|^2} \leq C$, the fact that $\int_{\mathbb{R}} |z|\nu_t(z)dz < \infty$ (due to A2-r), the bound $\mathbb{E}_t^{\mathbb{Q}}(|\widetilde{x}_s - \widetilde{x}_t|^2 \vee |\widetilde{x}_s - \widetilde{x}_t|^3) \leq C_t T$ for $s \in [t, t+T]$, as well as the integrability assumptions in A2-r, we have for η_T in (7.28):

$$\left|\frac{1}{T}\mathbb{E}_t^{\mathbb{Q}}\left(f_{\eta_T}(\widetilde{x}_{t+T} - \widetilde{x}_t)\right) - V_t\right| = O_p\left(\sqrt{T} \vee \eta_T T \bigvee \frac{1}{\sqrt{\eta_T}}\right). \tag{7.35}$$

Given the results in (7.32) and (7.35), to prove the claims of Theorem 2, we need to show the asymptotic negligibility of $\widehat{QV}_{t,T} - \frac{1}{T}\int_{-\infty}^{\infty} h_0(k)\widetilde{\kappa}_T(k)dk$ and $\widehat{TV}_{t,T}(\widehat{\eta}_T) - \frac{1}{T}\int_{-\infty}^{\infty} h_{\eta_T}(k)\widetilde{\kappa}_T(k)dk$.

First, using the bounds of Lemma 1 and assumption A5 for the observation error, we can easily conclude

$$\widehat{QV}_{t,T} - \frac{1}{T}\int_{-\infty}^{\infty} h_0(k)\kappa_T(k)dk = O_p\left(\frac{\sqrt{\overline{\Delta}}}{T^{1/4}} \bigvee e^{-2(|\underline{k}|\vee\overline{k})}(|\underline{k}| \vee \overline{k})\right). \tag{7.36}$$

In addition, using Itô's lemma as well as assumptions A1-A3, we have

$$\left|\mathbb{E}_t^{\mathbb{Q}}(f_0(x_{t+T} - x_t)) - \mathbb{E}_t^{\mathbb{Q}}(f_0(\widetilde{x}_{t+T} - \widetilde{x}_t))\right| = O_p(T\sqrt{T}). \tag{7.37}$$

The bounds in (7.36)-(7.37) together with (7.32) establish the consistency of $\widehat{QV}_{t,T}$. We continue with $\widehat{TV}_{t,T}(\widehat{\eta}_T)$. For its analysis, we first introduce the set

$$\Omega = \left\{\omega : |\widehat{QV}_{t,T} - QV_t| \leq \frac{1}{5}QV_t\right\}.$$

We will further make use of two algebraic inequalities. For $\eta \in \mathbb{R}_+$, $a \in \mathbb{R}_+$ with $|a - \eta| \leq \frac{\eta}{4}$, and $k \in \mathbb{R}$, we have

$$|h_\eta(k)| \leq Ce^{-k-\frac{\eta}{2}k^2}(|k| \vee 1), \tag{7.38}$$

$$|h_a(k) - h_\eta(k)| \leq C|a-\eta|k^2e^{-k-\frac{\eta}{2}k^2}\left(k^2 \vee 1 + \eta^2k^4\right), \tag{7.39}$$

for some C that does not depend on η, a and k (recall that $x_t = 0$).

Using the bounds of Lemma 1 and assumption A5 for the observation error as well as (7.32) and (7.36), we have for T being below some $\mathcal{F}_t^{(0)}$-adapted and positive random variable ζ_t (so that $\mathbb{P}(T > \zeta_t) \to 0$ as $T \to 0$):

$$\mathbb{P}\left(\Omega^c | T < \zeta_t\right) \leq C_t\frac{\overline{\Delta}}{\sqrt{T}}, \tag{7.40}$$

and therefore $1(\Omega^c)$ is $O_p\left(\frac{\overline{\Delta}}{\sqrt{T}}\right)$. Further, using the inequality in (7.38), the bounds of Lemma 1 as well as assumption A5 for the observation error, we have $\widehat{TV}_{t,T}(\widehat{\eta}_T) - \widehat{TV}_{t,T}(\eta_T) = O_p(1)$. Therefore, altogether

$$\left(\widehat{TV}_{t,T}(\widehat{\eta}_T) - \widehat{TV}_{t,T}(\eta_T)\right) 1_{\{\Omega^c\}} = O_p\left(\frac{\overline{\Delta}}{\sqrt{T}}\right). \tag{7.41}$$

Next, since on the set Ω we have $|\widehat{\eta}_T - \eta_T| \leq \frac{\eta_T}{4}$ for T small enough, we can apply (7.39) and bound

$$\begin{aligned}\left|\widehat{TV}_{t,T}(\widehat{\eta}_T) - \widehat{TV}_{t,T}(\eta_T)\right| 1_{\{\Omega\}} &\leq C_t \overline{\eta}_T |\widehat{QV}_{t,T} - QV_{t,T}| \\ &\times \frac{1}{T^2}\sum_{j=1}^N k_{j-1}^2 e^{-k_{j-1} - \frac{\eta_T}{2}k_{j-1}^2}\left(k_{j-1}^2 \vee 1 + \eta_T^2 k_{j-1}^4\right)\widehat{\kappa}_T(k_{j-1})\Delta_j,\end{aligned} \tag{7.42}$$

for some finite-valued $C_t > 0$ (note that because of A1 we have $QV_{t,T} > 0$). Using the bounds of Lemma 1 and assumption A5 for the observation error, we have

$$\frac{1}{T^2}\sum_{j=1}^N k_{j-1}^2 e^{-k_{j-1} - \frac{\eta_T}{2}k_{j-1}^2}\left(k_{j-1}^2 \vee 1 + \eta_T^2 k_{j-1}^4\right)\widehat{\kappa}_T(k_{j-1})\Delta_j = O_p(1), \tag{7.43}$$

and taking into account the bounds for $\widehat{QV}_{t,T}$ in (7.32) and (7.36), we have altogether

$$\left(\widehat{TV}_{t,T}(\widehat{\eta}_T) - \widehat{TV}_{t,T}(\eta_T)\right) 1_{\{\Omega\}} = O_p\left(\overline{\eta}_T \frac{\sqrt{\overline{\Delta}}}{T^{1/4}} \bigvee \overline{\eta}_T\sqrt{T} \bigvee \overline{\eta}_T e^{-2(|\underline{k}| \vee \overline{k})}(|\underline{k}| \vee \overline{k})\right). \tag{7.44}$$

We continue next with $\widehat{TV}_{t,T}(\eta_T) - \frac{1}{T}\int_{-\infty}^{\infty} h_{\eta_T}(k)\kappa_T(k)dk$ which we split into $\widehat{TV}_{t,T}^{(1)}$, $\widehat{TV}_{t,T}^{(2)}$ and $\widehat{TV}_{t,T}^{(3)}$ defined as

$$\widehat{TV}_{t,T}^{(1)} = \frac{1}{T}\sum_{j=2}^N h_{\eta_T}(k_{j-1})\epsilon_{j-1}\Delta_j,$$

$$\widehat{TV}_{t,T}^{(2)} = \frac{1}{T}\sum_{j=2}^N \int_{k_{j-1}}^{k_j} (h_{\eta_T}(k_{j-1})\kappa_T(k_{j-1}) - h_{\eta_T}(k)\kappa_T(k))dk,$$

$$\widehat{TV}_{t,T}^{(3)} = -\frac{1}{T}\int_{-\infty}^{k_1} h_{\eta_T}(k)\kappa_T(k)dk - \frac{1}{T}\int_{k_N}^{\infty} h_{\eta_T}(k)\kappa_T(k)dk.$$

Using assumption A5 and the bounds of Lemma 1, we have

$$\mathbb{E}\left(\widehat{TV}_{t,T}^{(1)} \middle| \mathcal{F}^{(0)}\right)^2 = O_p\left(\frac{\overline{\Delta}}{\sqrt{T}}\right), \quad \widehat{TV}_{t,T}^{(3)} = O_p\left(e^{-(\underline{k}^2 \wedge \overline{k}^2)\frac{1}{4QV_t}\frac{\overline{\eta}_T}{T}}\right). \tag{7.45}$$

For $\widehat{TV}^{(2)}_{t,T}$ we make use of the following algebraic inequality

$$|h_\eta(k_2) - h_\eta(k_1)| \leq C_t|k_2 - k_1|e^{|k_2| - \frac{\eta}{2}k_2^2}\left(1 + |k_2|\eta\right)\left(1 + k_2^2 + \eta^2 k_2^4\right), \tag{7.46}$$

for $|k_1| \leq \frac{1}{2}|k_2|$ and where C_t is a finite-valued $\mathcal{F}_t$-adapted random variable that does not depend on k_1, k_2 and η. Using this inequality, (7.38) as well as Lemma 1, we get

$$\widehat{TV}^{(2)}_{t,T} = O_p\left(\frac{\overline{\Delta}}{\sqrt{T}}\right). \tag{7.47}$$

Combining the bounds in (7.45) and (7.47), we have altogether (the bound below is not sharp and can be further relaxed)

$$\widehat{TV}_{t,T}(\eta_T) - \frac{1}{T}\int_{-\infty}^{\infty} h_{\eta_T}(k)\kappa_T(k)dk = O_p\left(\frac{\sqrt{\overline{\Delta}}}{T^{1/4}} \bigvee e^{-2(|\underline{k}|\wedge|\overline{k}|)}(|\underline{k}| \wedge |\overline{k}|)\right). \tag{7.48}$$

This result, together with (7.35), (7.41) and (7.44) as well as

$$\mathbb{E}^{\mathbb{Q}}_t|f_{\eta_T}(x_{t+T} - x_t) - f_{\eta_T}(\widetilde{x}_{t+T} - \widetilde{x}_t)| \leq C_t T^{3/2}, \tag{7.49}$$

implies the consistency of $\widehat{TV}_{t,T}(\widehat{\eta}_T)$.

7.4. *Proof of Theorem 3.* We use the notation in (7.14). Using the result in (7.23) in the proof of Theorem 1, the fact that $\widehat{u}_T$ is $\mathcal{F}_t$-adapted, the consistency of $\widehat{TV}_{t,T}(\widehat{\eta}_T)$ for V_t from Theorem 2 as well as the strict positivity of V_t from assumption A1, we have

$$-\frac{2}{T\widehat{u}_T^2}\Re\left(\ln\left(f_{t,T}(\widehat{u}_T)\right)\right) - \sigma_t^2 = O_p(T^{1-r/2}). \tag{7.50}$$

We denote

$$\breve{u}_T = \frac{\overline{u}}{\sqrt{T}}\frac{1}{\sqrt{\widehat{TV}_{t,T}(\eta_T)}}, \quad u_T = \frac{\overline{u}}{\sqrt{T}}\frac{1}{\sqrt{V_t}}.$$

Using the fact that $f_{t,T}(u)$ equals the expression in (3.11), we then decompose $\widehat{f}_{t,T}(\widehat{u}_T) - f_{t,T}(\widehat{u}_T) = \sum_{j=1}^4 \widehat{f}^{(j)}_{t,T}$, where $\widehat{f}^{(j)}_{t,T} = -(\widehat{u}_T^2 + i\widehat{u}_T)T\overline{f}^{(j)}_{t,T}$ and

$$\overline{f}^{(1)}_{t,T} = \frac{1}{T}\sum_{j=2}^{N} e^{(iu_T-1)k_{j-1}}\epsilon_{j-1}\Delta_j,$$

$$\overline{f}^{(2)}_{t,T} = \frac{1}{T}\sum_{j=2}^{N}(e^{(i\widehat{u}_T-1)k_{j-1}} - e^{(iu_T-1)k_{j-1}})\epsilon_{j-1}\Delta_j,$$

$$\overline{f}^{(3)}_{t,T} = \frac{1}{T}\sum_{j=2}^{N}\int_{k_{j-1}}^{k_j}\left(e^{(i\widehat{u}_T-1)k_{j-1}}\kappa_T(k_{j-1}) - e^{(i\widehat{u}_T-1)k}\kappa_T(k)\right)dk,$$

$$\overline{f}^{(4)}_{t,T} = -\frac{1}{T}\int_{-\infty}^{\underline{k}} e^{(i\widehat{u}_T-1)k}\kappa_T(k)dk - \frac{1}{T}\int_{\overline{k}}^{\infty} e^{(i\widehat{u}_T-1)k}\kappa_T(k)dk.$$

We analyze each of the terms in the above decomposition. We start with $\overline{f}^{(2)}_{t,T}$. We denote

$$\widehat{\xi}_{t,T}(u) = \frac{1}{T}\sum_{j=2}^{N} e^{\left(\frac{iu}{\sqrt{T}\sqrt{V_t}}-1\right)k_{j-1}}\epsilon_{j-1}\Delta_j, \quad u \in \mathbb{R}. \tag{7.51}$$

Taking into account the $\mathcal{F}^{(0)}$-conditional independence of the observation errors and applying Lemma 1, we have for some sufficiently small T:

$$\mathbb{E}\left(|\widehat{\xi}_{t,T}(u)|^2|\mathcal{F}^{(0)}\right) \leq C_t\frac{\overline{\Delta}}{\sqrt{T}}, \tag{7.52}$$

where C_t is $\mathcal{F}^{(0)}_t$-adapted random variable that does not depend on u. Further, by application of Burkholder-Gundy-Davis inequality and the algebraic inequality $|\sum_j a_i|^p \leq \sum_i |a_i|^p$ for $p \in (0,1]$ as well as the boundedness of the trigonometric functions, we have for some $\iota \in (0,1)$:

$$\begin{aligned}\mathbb{E}&\left(|\widehat{\xi}_{t,T}(u) - \widehat{\xi}_{t,T}(v)|^{2+\iota}\big|\mathcal{F}^{(0)}\right)\\ &\leq C_t|u-v|^{1+\iota}\sum_{j=2}^{N} e^{-(2+\iota)k_{j-1}}\frac{|k_{j-1}|^{1+\iota}}{T^{5/2+3\iota/2}}|\kappa_T(k_{j-1})|^{2+\iota}\Delta_j^{2+\iota}, \quad u,v\in\mathbb{R},\end{aligned} \tag{7.53}$$

where C_t is $\mathcal{F}^{(0)}_t$-adapted random variable that does not depend on u and v. From here, for T sufficiently small so that the bounds of Lemma 1 apply, we get

$$\mathbb{E}\left(|\widehat{\xi}_{t,T}(u) - \widehat{\xi}_{t,T}(v)|^{2+\iota}\big|\mathcal{F}^{(0)}\right) \leq C_t|u-v|^{1+\iota}\left(\frac{\overline{\Delta}}{\sqrt{T}}\right)^{1+\iota}|\ln T|, \tag{7.54}$$

and here again C_t is $\mathcal{F}^{(0)}_t$-adapted random variable that does not depend on u and v. Since by the rate conditions of the theorem, $\overline{\Delta}/\sqrt{T} \to 0$, we can apply Theorem 12.3 of [13] and conclude that $\frac{T^{1/4}}{\sqrt{\overline{\Delta}}}\widehat{\xi}_{t,T}(u)$ is $\mathcal{F}^{(0)}_t$-conditionally tight in the space of continuous functions of u for any arbitrary bounded interval on $\mathbb{R}$. Using the established tightness result, we can apply Theorem 8.2 of [13] and conclude that for arbitrary small $\epsilon > 0$ and $\eta > 0$, we have for T below some value and for some $\delta > 0$ (both being $\mathcal{F}_t^{(0)}$-adapted):

$$\mathbb{P}\left(\frac{T^{1/4}}{\sqrt{\Delta}} \sup_{|u-\overline{u}|\leq\delta} |\widehat{\xi}_{t,T}(u) - \widehat{\xi}_{t,T}(\overline{u})| \geq \epsilon \middle| \mathcal{F}^{(0)}\right) \leq \eta, \tag{7.55}$$

where $\overline{u}$ is the constant used in defining $\widehat{u}_T$ in (3.24). Now, if we have a sequence $\widehat{u}$ which $\mathcal{F}^{(0)}$-conditionally converges in probability to $\overline{u}$, then we can pick T sufficiently small, such that

$$\mathbb{P}\left(|\widehat{u} - \overline{u}| > \delta|\mathcal{F}^{(0)}\right) \leq \eta. \tag{7.56}$$

In turn, for T smaller than the values for which the above two results hold, we have

$$\mathbb{P}\left(\frac{T^{1/4}}{\sqrt{\Delta}} |\widehat{\xi}_{t,T}(\widehat{u}) - \widehat{\xi}_{t,T}(\overline{u})| \geq \epsilon \middle| \mathcal{F}^{(0)}\right) \leq 2\eta. \tag{7.57}$$

Therefore,

$$\widehat{\xi}_{t,T}(\widehat{u}) - \widehat{\xi}_{t,T}(\overline{u}) = o_p\left(\frac{\sqrt{\Delta}}{T^{1/4}}\right). \tag{7.58}$$

Applying this result with $\widehat{u} = \overline{u}\sqrt{\frac{V_t}{\widehat{TV}_{t,T}(\widehat{\eta}_T)}}$ and taking into account that from the proof of Theorem 2, we have that $\widehat{TV}_{t,T}(\widehat{\eta}_T)$ converges in probability $\mathcal{F}^{(0)}$-conditionally to V_t, we get

$$\overline{f}_{t,T}^{(2)} = o_p\left(\frac{\sqrt{\Delta}}{T^{1/4}}\right). \tag{7.59}$$

Next, using the bounds of Lemma 1 as well as the consistency result for $\widehat{TV}_{t,T}(\widehat{\eta}_T)$ of Theorem 2, we have

$$\overline{f}_{t,T}^{(4)} = O_p\left(\frac{\overline{\Delta}}{\sqrt{T}}|\ln T|\right), \quad \overline{f}_{t,T}^{(5)} = O_p\left(e^{-2(|\underline{k}|\wedge|\overline{k}|)}\right). \tag{7.60}$$

Overall, taking into account the rate condition in (3.28), we have

$$\sum_{j=2}^{4} \widehat{f}_{t,T}^{(j)} = o_p\left(\frac{\sqrt{\Delta}}{T^{1/4}}\right). \tag{7.61}$$

We are left with $\overline{f}_{t,T}^{(1)}$. First, using assumption A5 and the bounds of Lemma 1, we have

$$\mathbb{E}\left(\left(\Re\overline{f}_{t,T}^{(1)}\right)^2|\mathcal{F}^{(0)}\right)+\mathbb{E}\left(\left(\Im\overline{f}_{t,T}^{(1)}\right)^2|\mathcal{F}^{(0)}\right)\leq C_t\frac{\overline{\Delta}}{\sqrt{T}}. \tag{7.62}$$

Using again assumption A5 for the observation errors and the bounds of Lemma 1, we have

$$\mathbb{E}\left(\left(\begin{array}{cc}\Re\overline{f}_{t,T}^{(1)} & \Im\overline{f}_{t,T}^{(1)}\end{array}\right)^\top\left(\begin{array}{cc}\Re\overline{f}_{t,T}^{(1)} & \Im\overline{f}_{t,T}^{(1)}\end{array}\right)\Big|\mathcal{F}^{(0)}\right)=\mathbf{V}_{t,T}, \tag{7.63}$$

$$\frac{1}{T^4}\mathbb{E}\left(\sum_{j=2}^N e^{-4k_{j-1}}\epsilon_{j-1}^4\Delta_j^4\Big|\mathcal{F}^{(0)}\right)=O_p\left(\frac{\overline{\Delta}^3}{T^{3/2}}\right), \tag{7.64}$$

where $\mathbf{V}_{t,T}$ is 2×2 matrix with elements given as follows

$$\mathbf{V}_{t,T}^{lm}=\frac{1}{T^2}\sum_{j=2}^N\chi_{lm}(u_Tk_{j-1})e^{-2k_{j-1}}\sigma_{t,j-1}^2\kappa_T^2(k_{j-1})\Delta_j^2,$$

with $\chi_{lm}(x)=\cos^2(x)$ for $l=m=1$, $\chi_{lm}(x)=\sin^2(x)$ for $l=m=2$ and $\chi_{lm}(x)=\cos(x)\sin(x)$ for $l=1$ and $m=2$. Therefore,

$$\begin{aligned}\mathbb{E}&\left(\left(\begin{array}{c}Tu_T^2\Re\overline{f}_{t,T}^{(1)}-Tu_T\Im\overline{f}_{t,T}^{(1)}\\ Tu_T\Re\overline{f}_{t,T}^{(1)}+Tu_T^2\Im\overline{f}_{t,T}^{(1)}\end{array}\right)\left(\begin{array}{c}Tu_T^2\Re\overline{f}_{t,T}^{(1)}-Tu_T\Im\overline{f}_{t,T}^{(1)}\\ Tu_T\Re\overline{f}_{t,T}^{(1)}+Tu_T^2\Im\overline{f}_{t,T}^{(1)}\end{array}\right)^\top\Big|\mathcal{F}^{(0)}\right)\\ &=\mathbf{C}_{t,T}(u_T),\end{aligned} \tag{7.65}$$

where for $u\in\mathbb{R}$ we denote (recall the notation in (3.26))

$$\mathbf{C}_{t,T}(u)=\sum_{j=2}^N\zeta_{j-1}(u)\zeta_{j-1}(u)^\top e^{-2k_{j-1}}\sigma_{t,j-1}^2\kappa_T^2(k_{j-1})\Delta_j^2.$$

Further, since $f_{t,T}(u_T)$ is $\mathcal{F}^{(0)}$-adapted, if we set

$$\chi_{t,T}=\frac{\Re f_{t,T}(u_T)}{|f_{t,T}(u_T)|^2}\left(Tu_T^2\Re\overline{f}_{t,T}^{(1)}-Tu_T\Im\overline{f}_{t,T}^{(1)}\right)+\frac{\Im f_{t,T}(u_T)}{|f_{t,T}(u_T)|^2}\left(Tu_T\Re\overline{f}_{t,T}^{(1)}+Tu_T^2\Im\overline{f}_{t,T}^{(1)}\right),$$

we have for $V_\chi^T=\mathbb{E}\left(\chi_{t,T}^2|\mathcal{F}^{(0)}\right)$:

$$V_\chi^T=\frac{1}{|f_{t,T}(u_T)|^4}\left(\Re f_{t,T}(u_T)\ \Im f_{t,T}(u_T)\right)\mathbf{C}_{t,T}(u_T)\left(\Re f_{t,T}(u_T)\ \Im f_{t,T}(u_T)\right)^\top. \tag{7.66}$$

Using (7.19) and (7.62), we have

(7.67)
$$V_\chi^T = \frac{\overline{u}^4}{V_t^2}\frac{1}{|f_{t,T}(u_T)|^4}\big(\Re f_{t,T}(u_T)\ \Im f_{t,T}(u_T)\big)\mathbf{V}_{t,T}\big(\Re f_{t,T}(u_T)\ \Im f_{t,T}(u_T)\big)^\top + O_p\left(\overline{\Delta}\right)$$
$$= \frac{\overline{u}^4}{TV_t^2}\frac{1}{|f_{t,T}(u_T)|^2}\sum_{j=2}^{N}\cos^2\left(\frac{\overline{u}k_{j-1}}{\sqrt{T}\sigma_t} - \psi_{t,T}\right)e^{-2k_{j-1}}\sigma_{t,j-1}^2\kappa_T^2(k_{j-1})\Delta_j^2 + O_p\left(\overline{\Delta}\right),$$

where we use the notation

$$\psi_{t,T} = \frac{\overline{u}a_t\sqrt{T}}{\sigma_t} + T\int_{\mathbb{R}}\left(\sin\left(\frac{\overline{u}x}{\sqrt{T}\sigma_t}\right) - \frac{\overline{u}x}{\sqrt{T}\sigma_t}\right)\nu_t(x)dx.$$

Note that, given assumption A2-r, we have $|\psi_{t,T}| \leq C_t\sqrt{T}$. If we denote with $\overline{V}_\chi^T$ the expression in the second line of (7.67), then by using assumption A5 as well as the bounds of Lemmas 1 and 2, we have

$$\begin{aligned}
\overline{V}_\chi^T &\geq \frac{C_t\overline{\Delta}}{T^2}\sum_{j:|k_{j-1}|\leq\sqrt{T}}\cos^2(u_Tk_{j-1} - \psi_{t,T})\kappa_T^2(k_{j-1})\Delta_j \\
&\geq \frac{C_t\overline{\Delta}}{T^2}\sum_{j:|k_{j-1}|\leq\sqrt{T}}\cos^2(u_Tk_{j-1} - \psi_{t,T})\widetilde{\kappa}_T^c(k_{j-1})^2\Delta_j + O_p(\overline{\Delta}) \\
&\geq \frac{C_t\overline{\Delta}}{T^2}\int_{-\sqrt{T}}^{\sqrt{T}}\cos^2(u_Tk - \psi_{t,T})\widetilde{\kappa}_T^c(k)^2dk + O_p\left(\overline{\Delta}\vee\frac{\overline{\Delta}^2}{T}\right).
\end{aligned}$$

From here, using Lemma 2 and by a change of the variable of integration, we further have

(7.68)
$$\overline{V}_\chi^T \geq \frac{C_t\overline{\Delta}}{\sqrt{T}}\int_{-1}^{1}\cos^2\left(\frac{\overline{u}k}{\sigma_t} - \psi_{t,T}\right)\left(f\left(\frac{k}{\sigma_t}\right) - \frac{|k|}{\sigma_t}\Phi\left(-\frac{|k|}{\sigma_t}\right)\right)^2 dk + O_p\left(\overline{\Delta}\vee\frac{\overline{\Delta}^2}{T}\right).$$

We note that the function $f(k) - |k|\Phi(-|k|)$ is strictly positive obtaining its maximum at $k = 0$ and decaying to zero in the tails. Therefore, for T sufficiently small, $\frac{\sqrt{T}}{\overline{\Delta}}V_\chi^T$ is bounded from below by an $\mathcal{F}_t$-adapted and positive-valued random variable. This shows that $\frac{\overline{\Delta}}{\sqrt{T}}$ is the sharp order of magnitude of V_χ^T (and not an upper bound for it).

Using (7.64) and (7.66)-(7.68) as well as the $\mathcal{F}^{(0)}$-conditional independence of the observation errors from assumption A5, we can apply Theorem VIII.5.7 of [28] and get

$$\frac{\chi_{t,T}}{\sqrt{V_{\chi}^{T}}} \stackrel{\mathcal{L}-s}{\longrightarrow} \mathcal{N}(0,1), \tag{7.69}$$

where in the limit result above the notation $\mathcal{L}-s$ means convergence that is stable in law and further the limit is defined on an extension of the original probability space and independent of $\mathcal{F}$.

From the bounds on the terms $\{\overline{f}_{t,T}^{(j)}\}_{j=1,\ldots,5}$ as well as the asymptotic negligibility of $\widehat{TV}_{t,T}(\widehat{\eta}_T) - V_t$ by Theorem 2, we have

$$\Re(\widehat{f}_{t,T}(\widehat{u}_T))/\Re(f_{t,T}(u_T)) \stackrel{\mathbb{P}}{\longrightarrow} 1 \text{ and } \Im(\widehat{f}_{t,T}(\widehat{u}_T)) - \Im(f_{t,T}(u_T)) \stackrel{\mathbb{P}}{\longrightarrow} 0.$$

Therefore, by an application of delta method from the convergence result in (7.69) and upon taking into account (7.67) and (7.68) as well as (7.61), we get

$$\mathrm{Avar}(\widehat{V}_{t,T}(u_T))^{-1/2}\frac{2}{Tu_T^2}\left(\Re(\ln(\widehat{f}_{t,T}(\widehat{u}_T))) - \Re\left(\ln(f_{t,T}(\widehat{u}_T))\right)\right) \stackrel{\mathcal{L}-s}{\longrightarrow} \mathcal{N}\left(0,1\right), \tag{7.70}$$

where $\mathrm{Avar}(\widehat{V}_{t,T}(u_T))$ denotes the analogous expression as $\widehat{\mathrm{Avar}}(\widehat{V}_{t,T}(\widehat{u}_T))$ in which $\widehat{C}_{t,T}(\widehat{u}_T)$ is replaced with $C_{t,T}(u_T)$, $\widehat{f}_{t,T}(\widehat{u}_T)$ with $f_{t,T}(u_T)$ and $\widehat{u}_T$ with u_T. We would like to extend the above result to

$$\widehat{\mathrm{Avar}}(\widehat{V}_{t,T}(\widehat{u}_T))^{-1/2}\frac{2}{T\widehat{u}_T^2}\left(\Re(\ln(\widehat{f}_{t,T}(\widehat{u}_T))) - \Re\left(\ln(f_{t,T}(\widehat{u}_T))\right)\right) \stackrel{\mathcal{L}-s}{\longrightarrow} \mathcal{N}\left(0,1\right). \tag{7.71}$$

For this, we need to show that $\mathrm{Avar}(\widehat{V}_{t,T}(u_T))/\widehat{\mathrm{Avar}}(\widehat{V}_{t,T}(\widehat{u}_T)) \stackrel{\mathbb{P}}{\longrightarrow} 1$ and $\widehat{u}_T/u_T \stackrel{\mathbb{P}}{\longrightarrow} 1$. Given the asymptotic negligibility of $\widehat{TV}_{t,T}(\widehat{\eta}_T) - V_t$ and $\Re(\widehat{f}_{t,T}(\widehat{u}_T))/\Re(f_{t,T}(u_T)) \stackrel{\mathbb{P}}{\longrightarrow} 1$ and $\Im(\widehat{f}_{t,T}(\widehat{u}_T)) - \Im(f_{t,T}(u_T)) \stackrel{\mathbb{P}}{\longrightarrow} 0$ (established above), we only need to show that $\widehat{C}_{t,T}(\widehat{u}_T) - C_{t,T}(u_T) = o_p\left(\frac{\Delta}{\sqrt{T}}\right)$. First, using the bounds in Lemma 1 as well as assumption A5 for the observation error, we have $\widehat{C}_{t,T}(\widehat{u}_T) - \widehat{C}_{t,T}(u_T) = O_p\left(\alpha_T\frac{\Delta}{\sqrt{T}}\right)$ where $\widehat{TV}_{t,T}(\widehat{\eta}_T) - V_t = O_p(\alpha_T)$. Second, for $\widehat{C}_{t,T}(u_T) - C_{t,T}(u_T)$ we can use assumption A5 for the observation error and apply Burkholder-Davis-Gundy inequality and conclude $\widehat{C}_{t,T}(u_T) - C_{t,T}(u_T) = o_p\left(\frac{\Delta}{\sqrt{T}}\right)$. Thus, altogether, $\widehat{C}_{t,T}(\widehat{u}_T) - C_{t,T}(u_T) = o_p\left(\frac{\Delta}{\sqrt{T}}\right)$, and from here the result in (7.71) follows.

Combining this result with (7.50) and taking into account the rate condition in (3.28), we have the convergence result (3.29) of the theorem.

7.5. *Proof of Theorem 4.* Under assumption A6 and our rate condition for k_n, the conditions of Theorem 13.3.3 of [27] are satisfied. Therefore, we can apply this theorem and a counterpart to Theorem 9.3.2 of [27] for fourth power local truncated variations, and get

$$\frac{\widehat{V}_t^{hf} - V_t}{\sqrt{\widehat{Avar}(\widehat{V}_{t,T}^{hf})}} \xrightarrow{\mathcal{L}-s} Z, \tag{7.72}$$

with Z being a standard normal random variable defined on an extension of the original probability space and independent from $\mathcal{F}$. Lets denote $Z_1^n = \frac{\widehat{V}_{t,T}(u_T)-V_t}{\sqrt{\mathrm{Avar}(\widehat{V}_{t,T}(u_T))}}$ and $Z_2^n = \frac{\sqrt{k_n}\left(\widehat{V}_t^{hf}-V_t\right)}{\sqrt{2}\widehat{V}_t^{hf}}$, where $\mathrm{Avar}(\widehat{V}_{t,T}(u))$ is defined as in the proof of Theorem 3 above and as in that proof we set $u_T = \frac{\overline{u}}{\sqrt{T}}\frac{1}{\sqrt{V_t}}$. From the proof of Theorem 3 we have $\frac{\widehat{V}_{t,T}(\widehat{u}_T)-V_t}{\sqrt{\widehat{\mathrm{Avar}}(\widehat{V}_{t,T}(\widehat{u}_T))}} - Z_1^n = o_p(1)$ and hence to prove the result of Theorem 4 we need to establish the joint convergence of (Z_1^n, Z_2^n).

We further denote with Z_1 and Z_2 two independent standard normal variables, which are defined on an extension of the original probability space and independent of $\mathcal{F}$, and with g and h two bounded continuous functions on $\mathbb{R}$. Now, using our results from the proof of Theorem 3 for the term $\overline{f}_{t,T}^{(1)}$, we can apply Theorem VIII.5.25 of [28] (using assumption A5 for the observation errors and the separability of $\mathcal{F}^{(0)}$) and conclude that

$$\mathbb{E}\left(g\left(Z_1^n\right)\middle|\mathcal{F}^{(0)}\right) \longrightarrow \mathbb{E}(g(Z_1)), \quad \text{a.s.}, \tag{7.73}$$

and therefore since g and h are bounded functions

$$\mathbb{E}\left(\left(\mathbb{E}\left(g\left(Z_1^n\right)\middle|\mathcal{F}^{(0)}\right) - \mathbb{E}(g(Z_1))\right)h(Z_2^n)\right) \longrightarrow 0. \tag{7.74}$$

Therefore, using (7.72), we have for every bounded random variable Y on $\mathcal{F}$

$$\mathbb{E}\left(Yg(Z_1^n)h(Z_2^n)\right) \longrightarrow \mathbb{E}(Y)\mathbb{E}(g(Z_1))\mathbb{E}(h(Z_2)), \tag{7.75}$$

and this establishes $(Z_1^n, Z_2^n) \xrightarrow{\mathcal{L}-s} (Z_1, Z_2)$.

7.6. *Proof of Theorem 5.* Since $T \downarrow 0$, it is no restriction to assume $T < \bar{t}$, for $\bar{t}$ being the random variable of Lemmas 1 and 2, and we will do so in the proof without further mention. We will further define with C some positive constant which can change from line to line and depends on R in (4.2). Finally, we will use the notation $a_n \gtrsim b_n$ and $a_n \lesssim b_n$ to mean the respective inequality up to a constant independent of the parameter n.

The idea of the proof is to perturb locally σ and then derive the order of magnitude of the Kullback-Leibler divergence of the resulting two probability distributions of the observed option prices. Applying Theorem 2.2 in [45], we then have

$$\inf_{\widehat{\sigma}} \sup_{\mathcal{T}\in\mathcal{G}(R)} \mathbb{E}_{\mathcal{T}}\left(\frac{\sqrt{T}|\ln T|^{5/2}}{\overline{\Delta}}|\widehat{\sigma}-\sigma|^2\right) \geq V(\alpha), \tag{7.76}$$

where α is an upper bound on the Kullback-Leibler divergence of two probability distributions in $\mathcal{G}(R)$ with diffusive volatilities $\sigma^{(1)}$ and $\sigma^{(2)}$ such that $|\sigma^{(1)}-\sigma^{(2)}| = \frac{\sqrt{\overline{\Delta}}}{T^{1/4}|\ln T|^{5/4}}$ and $|a^{(1)}-a^{(2)}| \leq |\sigma^{(1)}-\sigma^{(2)}|$, and $V(\alpha)$ is some strictly positive function of α.

The KL divergence between the probability measures for the observed noisy option prices, corresponding to $\mathcal{T}$ with $\sigma^{(1)}$ and $\sigma^{(2)}$ and the same ν, both of which belong to $\mathcal{G}(R)$, is given by

$$\begin{aligned} KL(\sigma^{(1)},\sigma^{(2)}) &= \sum_{i=1}^{N}\frac{(\kappa_{T,1}(k_i)-\kappa_{T,2}(k_i))^2}{2(\kappa_{T,2}(k_i)\vee T)^2} \\ &\quad + \frac{1}{2}\sum_{i=1}^{N}\left(\left(\frac{\kappa_{T,1}(k_i)\vee T}{\kappa_{T,2}(k_i)\vee T}\right)^2 - 1 - \ln\left(\frac{\kappa_{T,1}(k_i)\vee T}{\kappa_{T,2}(k_i)\vee T}\right)^2\right), \end{aligned} \tag{7.77}$$

where the option prices corresponding to $\sigma^{(j)}$ are denoted by $O_{T,j}(k)$, for $j = 1, 2$.

In order to analyze the KL divergence, we will first establish lower and upper bounds on $\kappa_{T,1}(k)$ and $\kappa_{T,2}(k)$. In what follows we will use the notation of Section 7.2. For $\mathbb{Q} \in \mathcal{G}(R)$, we can decompose

$$\begin{aligned} \kappa_T(k) = \widetilde{\kappa}_T^c(k) &+ \mathbb{E}_t^{\mathbb{Q}}[(e^{x_{t+T}}-e^k)^+ 1_{\{\mu([t,t+T]\times\mathbb{R})\geq 1\}}] \\ &- \mathbb{E}_t^{\mathbb{Q}}[(e^{x^c_{t+T}}-e^k)^+ 1_{\{\mu([t,t+T]\times\mathbb{R})\geq 1\}}], \end{aligned} \tag{7.78}$$

for $k > 0$ and a similar decomposition holds for the case $k \leq 0$. Then using the independence of x^c_{t+T} and x^d_{t+T} (recall x is a Lévy process under $\mathbb{Q}$), the fact that $F(\mathbb{R}) < \infty$ (recall notation in (4.1)) as well as Hölder's inequality and the integrability assumptions for ν in (4.2), we have

$$|\kappa_T(k) - \widetilde{\kappa}_T^c(k)| \leq CT, \tag{7.79}$$

for some positive constant C which does not depend on k and T, and is a continuous function of σ.

We will henceforth concentrate on the case $k \leq 0$ with the other case $k > 0$ being treated analogously. Now, we can make use of Lemma 2, and hence it suffices to look at

$$f\left(\frac{k - aT}{\sqrt{T}\sigma}\right)\sqrt{T}\sigma + (e^k - 1)\Phi\left(\frac{k - aT}{\sqrt{T}\sigma}\right). \tag{7.80}$$

We have $-x\Phi(x) \sim f(x)$ for $x \downarrow -\infty$, and using this fact and Taylor expansion, we have

$$\left|\kappa_T(k) - f\left(\frac{k - aT}{\sqrt{T}\sigma}\right)\sqrt{T}\sigma - k\Phi\left(\frac{k - aT}{\sqrt{T}\sigma}\right)\right| \leq CT, \tag{7.81}$$

for some positive constant C which does not depend on k and T, and is a continuous function of σ. The function $f(x) + x\Phi(x)$ is increasing with value at zero of $f(0) > 0$ and $\lim_{x\downarrow -\infty}(f(x) + x\Phi(x)) = 0$. Furthermore, we have

$$\lim_{x\downarrow -\infty} \frac{f(x) + x\Phi(x)}{f(x)/x^2} = 1. \tag{7.82}$$

Therefore, for some $\epsilon \in (0, 1)$, there exists large in absolute value $x^* < 0$ such that for $x < x^*$, we have

$$(1 - \epsilon)\frac{f(x)}{x^2} \leq f(x) + x\Phi(x) \leq (1 + \epsilon)\frac{f(x)}{x^2}. \tag{7.83}$$

As a result, we have the following lower bounds for $\mathbb{Q} \in \mathcal{G}(R)$:

$$\begin{aligned} &\kappa_T(k) \vee T \geq \underline{C}_1 \left[f\left(\frac{k - aT}{\sqrt{T}\sigma}\right)\frac{T^{3/2}}{|k|^2} \bigvee T \right], \quad \text{for } k < x^*\sigma\sqrt{T} + aT, \\ &\kappa_T(k) \vee T \geq \underline{C}_2\sqrt{T}, \quad \text{for } k \in [x^*\sigma\sqrt{T} + aT, 0], \end{aligned} \tag{7.84}$$

and we note that x^* does not depend on σ while the constants $1 \geq \underline{C}_1 > 0$ and $\underline{C}_2 > 0$ do. We similarly have the upper bounds for $\mathbb{Q} \in \mathcal{G}(R)$:

$$\begin{aligned} &\kappa_T(k) \leq \overline{C}_1 \left[f\left(\frac{k - aT}{\sqrt{T}\sigma}\right)\frac{T^{3/2}}{|k|^2} \bigvee T \right], \quad \text{for } k < x^*\sigma\sqrt{T} + aT, \\ &\kappa_T(k) \leq \overline{C}_2\sqrt{T}, \quad \text{for } k \in [x^*\sigma\sqrt{T} + aT, 0], \end{aligned} \tag{7.85}$$

for some finite constants $\overline{C}_1 > 0$ and $\overline{C}_2 > 0$ which depend on σ. We note that $\underline{C}_1$, $\underline{C}_2$, $\overline{C}_1$ and $\overline{C}_2$ remain bounded both from below and above for $\mathbb{Q} \in \mathcal{G}(R)$. Further, for T sufficiently small, we have

$$\begin{cases} k < -\sqrt{2T}\sigma\sqrt{\ln(1/\sqrt{T})} \implies f\left(\frac{k-aT}{\sqrt{T}\sigma}\right)\frac{T^{3/2}}{|k|^2} < T, \\ k > -\sqrt{2T}\sigma \implies f\left(\frac{k-aT}{\sqrt{T}\sigma}\right)\frac{T^{3/2}}{|k|^2} > T. \end{cases} \tag{7.86}$$

As a result, for two risk-neutral probability laws $\mathbb{Q} \in \mathcal{G}(R)$ with the same ν and $|\sigma^{(1)} - \sigma^{(2)}| \leq CT^\eta$ (and $|a^{(1)} - a^{(2)}| \leq |\sigma^{(1)} - \sigma^{(2)}|$) with some positive $\eta > 0$, we have

$$\begin{aligned} k \in \left(-\sqrt{2T}\sigma^{(1)}\sqrt{\ln(1/\sqrt{T})}, -\sqrt{2T}\sigma^{(2)}\right) &\implies \\ 1 - \iota < f\left(\frac{k - a^{(1)}T}{\sqrt{T}\sigma^{(1)}}\right) \Big/ f\left(\frac{k - a^{(2)}T}{\sqrt{T}\sigma^{(2)}}\right) &< 1 + \iota, \end{aligned} \tag{7.87}$$

for some $\iota \in (0, 1)$ and where without loss of generality we have assumed $\sigma^{(1)} \leq \sigma^{(2)}$. Therefore,

$$0 < \underline{R} \leq \frac{\kappa_{T,1}(k) \vee T}{\kappa_{T,2}(k) \vee T} \leq \overline{R} < \infty, \tag{7.88}$$

for some $\underline{R}$ and $\overline{R}$ that depend on R.

Given this bound and using second-order Taylor expansion, we have

$$KL(\sigma^{(1)}, \sigma^{(2)}) \lesssim \sum_{i=1}^{N} \frac{(\kappa_{T,1}(k_i) - \kappa_{T,2}(k_i))^2}{\kappa_{T,2}(k_i)^2 \vee T^2}. \tag{7.89}$$

To proceed further, we need to analyze the difference $\kappa_{T,1}(k) - \kappa_{T,2}(k)$. By looking separately at the sets at which there is no jump in x on the interval $[t, t+T]$ and on which there is, we have

$$\begin{aligned} |\kappa_{T,1}(k) - \kappa_{T,2}(k)| &\leq |\widetilde{\kappa}^c_{T,1}(k) - \widetilde{\kappa}^c_{T,2}(k)| \\ &\quad + 2\mathbb{E}^{\mathbb{Q}}_t\left[(e^{x^d_{t+T}} \vee 1)|e^{x^{(c,1)}_{t+T}} - e^{x^{(c,2)}_{t+T}}|1_{\{\mu([t,t+T],\mathbb{R})\geq 1\}}\right], \end{aligned} \tag{7.90}$$

and therefore, taking into account the independence of x^c_{t+T} and x^d_{t+T} as well as $\nu(\mathbb{R}) < \infty$ and the tail decay of ν, we have

$$|\kappa_{T,1}(k) - \kappa_{T,2}(k)| \leq |\widetilde{\kappa}^c_{T,1}(k) - \widetilde{\kappa}^c_{T,2}(k)| + CT^{3/2}|\sigma^{(1)} - \sigma^{(2)}|. \tag{7.91}$$

In what follows, we consider the case $k \leq 0$ with $k > 0$ analyzed in a similar way. Direct calculation yields for $k \leq 0$:

$$\widetilde{\kappa}_T^c(k) = e^k \Phi\left(\frac{k - aT}{\sqrt{T}\sigma}\right) - e^{aT - \sigma^2 T/2} \Phi\left(\frac{k - aT}{\sqrt{T}\sigma} - \sqrt{T}\sigma\right), \tag{7.92}$$

and therefore using Taylor expansion, we can make the follow decomposition $\widetilde{\kappa}_T^c(k) = \sum_{j=1}^6 A_T^{(j)}(k)$ where with the shorthand $\overline{k}_T = \frac{k-aT}{\sqrt{T}\sigma}$, we denote

$$A_T^{(1)}(k) = \sigma\sqrt{T}\left[\overline{k}_T \Phi(\overline{k}_T) + f(\overline{k}_T)\right], \quad A_T^{(2)}(k) = aT\Phi(\overline{k}_T), \tag{7.93}$$

$$A_T^{(3)}(k) = \frac{1}{2} f'(\overline{k}_T)\sigma^2 T, \quad A_T^{(4)}(k) = (e^k - 1 - k)\Phi(\overline{k}_T), \tag{7.94}$$

$$A_T^{(5)}(k) = -\left(e^{aT - \sigma^2 T/2} - 1\right)\Phi(\overline{k}_T - \sigma\sqrt{T}), \tag{7.95}$$

and $A_T^{(6)}(k)$ is such that we have

$$\left|A_T^{(6)}(k)\right| \leq CT^{3/2}(|\overline{k}_T - \sigma\sqrt{T}|^2 \vee 1) f(\overline{k}_T). \tag{7.96}$$

For $l = 1, 2$, we denote with $\overline{k}_{T,l}$ and $A_{T,l}^{(j)}(k)$ the counterparts of $\overline{k}_T$ and $A_T^{(j)}(k)$ where a and σ are replaced with $a^{(l)}$ and $\sigma^{(l)}$. Using Taylor expansion and the monotonicity of f and Φ, we then have for $k \leq 0$

$$\begin{aligned} |A_{T,1}^{(1)}(k) - A_{T,2}^{(1)}(k)| \leq\ & C\sqrt{T}\big(\overline{k}_{T,1}\Phi(\overline{k}_{T,1}) + f(\overline{k}_{T,1})\big)|\sigma^{(1)} - \sigma^{(2)}| \\ & + C\sqrt{T}\left(\frac{|k|}{\sqrt{T}} \bigvee \sqrt{T}\right)\Phi(\overline{k}_{T,1} \vee \overline{k}_{T,2})|\sigma^{(1)} - \sigma^{(2)}|, \end{aligned} \tag{7.97}$$

$$|A_{T,1}^{(2)}(k) - A_{T,2}^{(2)}(k)| \leq CT\left(\frac{|k|}{\sqrt{T}} \bigvee \sqrt{T}\right) f(\overline{k}_{T,1} \vee \overline{k}_{T,2})|\sigma^{(1)} - \sigma^{(2)}|, \tag{7.98}$$

$$|A_{T,1}^{(3)}(k) - A_{T,2}^{(3)}(k)| \leq CT\left(\frac{|k|}{\sqrt{T}} \bigvee \frac{|k|^3}{T^{3/2}} \bigvee \sqrt{T}\right) f(\overline{k}_{T,1} \vee \overline{k}_{T,2})|\sigma^{(1)} - \sigma^{(2)}|, \tag{7.99}$$

$$|A_{T,1}^{(4)}(k) - A_{T,2}^{(4)}(k)| \leq C(|k|^2 \wedge |k|)\left(\frac{|k|}{\sqrt{T}} \bigvee \sqrt{T}\right) f(\overline{k}_{T,1} \vee \overline{k}_{T,2})|\sigma^{(1)} - \sigma^{(2)}|, \tag{7.100}$$

$$|A^{(5)}_{T,1}(k) - A^{(5)}_{T,2}(k)| \leq CT\left(\frac{|k|}{\sqrt{T}} \bigvee \sqrt{T}\right) f(\overline{k}_{T,1} \vee \overline{k}_{T,2})|\sigma^{(1)} - \sigma^{(2)}|, \tag{7.101}$$

and these bounds continue to hold for $k > 0$. Using them, the lower bounds for $\kappa_T(k)$ derived earlier, the fact that $-x \int_{-\infty}^{x} f^2(u)du \sim f^2(x)$ for $x \downarrow -\infty$, the asymptotic behavior of $\Phi(x)$ and $f(x)+x\Phi(x)$ for $x \downarrow -\infty$, and splitting the summations below into two sums according to whether $|k|$ is above or below $\sqrt{2T}\sigma^{(2)}\sqrt{\ln(1/\sqrt{T})}$, we have

$$\sum_{j=1}^{5}\sum_{i=1}^{N} \frac{|A^{(j)}_{T,1}(k_i) - A^{(j)}_{T,2}(k_i)|^2}{\kappa_{T,2}(k_i)^2 \vee T^2} \lesssim \frac{\sqrt{T}}{\overline{\Delta}}\eta_T^2|\ln T|^{5/2}, \tag{7.102}$$

and further

$$\sum_{i=1}^{N} \frac{|A^{(6)}_{T,1}(k_i) - A^{(6)}_{T,2}(k_i)|^2}{\kappa_{T,2}(k_i)^2 \vee T^2} \lesssim \frac{T^{5/2}}{\overline{\Delta}}|\ln T|^{9/2}. \tag{7.103}$$

Therefore, taking into account also the bound in (7.90), we have

$$\sum_{i=1}^{N} \frac{|\kappa_{T,1}(k_i) - \kappa_{T,2}(k_i)|^2}{\kappa_{T,2}(k_i)^2 \vee T^2} \lesssim \frac{\sqrt{T}}{\overline{\Delta}}\eta_T^2|\ln T|^{5/2} + \frac{T^{5/2}}{\overline{\Delta}}|\ln T|^{9/2}, \tag{7.104}$$

where we denote $\eta_T = \sigma^{(1)} - \sigma^{(2)}$. Evaluating the above bounds with $\eta_T = \frac{\sqrt{\overline{\Delta}}}{T^{1/4}|\ln T|^{5/4}}$ (and making use of the fact that $\alpha < 5/2$ by assumption), we get the result of the theorem.

Acknowledgements. The work is partially supported by NSF Grant SES-1530748. I would like to thank the Editor, Associate Editor, an anonymous referee, my discussants Dobrislav Dobrev, Marcel Rindisbacher and Dacheng Xiu, as well as seminar participants at various conferences and seminars for many comments and suggestions which have led to significant improvements.

References.

[1] Aït-Sahalia, Y. and J. Jacod (2014). *High-Frequency Financial Econometrics*. Princeton University Press.

[2] Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and Forecasting Realized Volatility. *Econometrica 71*, 579–625.

[3] Andersen, T. G., O. Bondarenko, and M. T. Gonzalez-Perez (2015). Exploring return dynamics via corridor implied volatility. *The Review of Financial Studies 28*(10), 2902–2945.

[4] Andersen, T. G., N. Fusari, and V. Todorov (2015a). Parametric Inference and Dynamic State Recovery from Option Panels. *Econometrica 83*, 1081–1145.
[5] Andersen, T. G., N. Fusari, and V. Todorov (2015b). The Risk Premia Embedded in Index Options. *Journal of Financial Economics 117*, 558–584.
[6] Andersen, T. G., N. Fusari, and V. Todorov (2017). The Pricing of Short-Term Market Risk: Evidence from Weekly Options. *Journal of Finance 72*, 1335–1386.
[7] Barndorff-Nielsen, O. and N. Shephard (2002). Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models. *Journal of the Royal Statistical Society Series B, 64*, 253280.
[8] Barndorff-Nielsen, O. and N. Shephard (2004a). Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica 72*(3), 885 – 925.
[9] Barndorff-Nielsen, O. and N. Shephard (2004b). Power and Bipower Variation with Stochastic Volatility and Jumps. *Journal of Financial Econometrics 2*, 1–37.
[10] Barndorff-Nielsen, O. and N. Shephard (2006). Econometrics of Testing for Jumps in Financial Economics using Bipower Variation. *Journal of Financial Econometrics 4*, 1–30.
[11] Belomestny, D. and M. Reiß (2006). Spectral Calibration of Exponential Lévy Models. *Finance and Stochastics 10*, 449–474.
[12] Bentata, A. and R. Cont (2012). Short-Time Asymptotics for Marginal Distributions of Semimartingales. Working paper.
[13] Billingsley, P. (1968). *Convergence of Probability Measures*. New York: Wiley.
[14] Black, F. and M. Scholes (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy 81*, 637–654.
[15] Carr, P. and D. Madan (2001). Optimal Positioning in Derivative Securities. *Quantitiative Finance 1*, 19–37.
[16] Carr, P. and L. Wu (2003). What Type of Process Underlies Options? A Simple Robust Test. *Journal of Finance 58*, 2581–2610.
[17] Christensen, K., R. Oomen, and R. Reno (2018). The Drift Burst Hypothesis. Working paper.
[18] Cont, R. and C. Mancini (2011). Nonparametric Tests for Pathwise Properties of Semimartingales. *Bernoulli 17*, 781–813.
[19] Cont, R. and P. Tankov (2004a). *Financial Modelling with Jump Processes*. Boca Raton, Florida, U.S.A.: Chapman and Hall.
[20] Cont, R. and P. Tankov (2004b). Nonparametric Calibration of Jump-Diffusion Option Pricing Models. *Journal of Computational Finance 7*, 1–49.
[21] Duffie, D. (2001). *Dynamic Asset Pricing Theory* (3rd ed.). Princeton University Press.
[22] Duffie, D., J. Pan, and K. Singleton (2000). Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica 68*, 1343–1376.
[23] Figueroa-Lopez, J., S. Gong, and C. Houdre (2016). High-Order Short-Time Expansions for ATM Option Prices of Exponential Lévy Models. *Mathematical Finance 26*, 516–557.
[24] Figueroa-Lopez, J. and S. Olafsson (2016). Short-time Expansions for Close-to-the-money Options under a Lévy Jump Model with Stochastic Volatility. *Finance and Stochastics 20*, 219–265.
[25] Gatheral, J., T. Jaisson, and M. Rosenbaum (2018). Volatility is Rough. *Quantitative Finance, forthcoming*.
[26] Green, R. and R. Jarrow (1987). Spanning and Completeness in Markets with Contingent Claims. *Journal of Economic Theory 41*, 202–210.

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

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

[29] Jacod, J. and V. Todorov (2010). Do Price and Volatility Jump Together? *Annals of Applied Probability 20*, 1425–1469.

[30] Jacod, J. and V. Todorov (2014). Efficient Estimation of Integrated Volatility in Presence of Infinite Variation Jumps. *Annals of Statistics 42*, 1029–1069.

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

[32] Mancini, C. (2009). Non-parametric Threshold Estimation for Models with Stochastic Diffusion Coefficient and Jumps. *Scandinavian Journal of Statistics 36*, 270–296.

[33] Merton, R. (1973). Theory of Rational Option Pricing. *Bell Journal of Economics and Management Science 4*, 141–183.

[34] Merton, R. (1976). Option Pricing when Underlying Asset Returns are Discontinuous. *Journal of Financial Economics 3*, 125–144.

[35] Mijatovic, A. and P. Tankov (2016). A New Look at Short-Term Implied Volatility in Asset Price Models with Jumps. *Mathematical Finance 26*, 149–183.

[36] Muhle-Karbe, J. and M. Nutz (2011). Small-Time Asymptotics of Option Prices and First Absolute Moments. *Journal of Applied Probability 48*, 1003–1020.

[37] Qin, L. and V. Todorov (2018). Nonparametric Implied Lévy Densities. *Annals of Statistics, forthcoming*.

[38] Ross, S. (1976). Options and Efficiency. *Quarterly Journal of Economics 90*, 75–89.

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

[40] Singleton, K. (2006). *Empirical Dynamic Asset Pricing*. Princeton University Press.

[41] Söhl, J. (2014). Confidence Sets in Nonparametric Calibration of Exponential Lévy Models. *Finance and Stochastics 18*, 617–649.

[42] Söhl, J. and M. Trabs (2014). Option Calibration of Exponential Lévy Models: Confidence Intervals and Empirical Results. *Journal of Computational Finance 18*, 91–119.

[43] Trabs, M. (2014). Calibration of Self-Decomposable Lévy Models. *Bernoulli 20*, 109–140.

[44] Trabs, M. (2015). Quantile Estimation for Lévy Measures. *Stochastic Processes and their Applications 125*, 3484–3521.

[45] Tsybakov, A. (2009). *Introduction to Nonparametric Estimation*. Springer-Verlag.

[46] van der Vaart, A. (1998). *Asymptotic Statistics*. Cambridge: Cambridge University Press.

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