Nonparametric Option-Implied Volatility*

Viktor Todorov†

May 11, 2017

Abstract

We propose a nonparametric estimator of spot volatility from short-dated options. 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 later is chosen in an adaptive way in order to account for the time-varying volatility. We 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. Numerical experiments show the efficiency gains from the newly-developed option-based volatility extraction.

Keywords: Jumps, Large Data Sets, Nonparametric Inference, Options, Stable Convergence, Stochastic Volatility.

JEL classification: C51, C52, G12.

*The work is partially supported by NSF Grant SES-1530748. I would like to thank seminar participants at various conferences and seminars for useful comments and suggestions. I would also like to thank Nicola Fusari for the option pricing codes used in the Monte Carlo and Shumiao Ouyang for helping with the extraction of the option data used in the empirical application.

†Department of Finance, Kellogg School of Management, Northwestern University, Evanston, IL 60208; e-mail: v-todorov@northwestern.edu.
1 Introduction

Options provide a natural source of information for studying volatility. Indeed, following the seminal work of Black and Scholes (1973) and Merton (1973), 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 Black and Scholes (1973) and Merton (1973), mainly constant volatility and no jump risk, are too simple for such volatility extraction to work in practice, see e.g., Duffie (2001) and Singleton (2006). 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 range of strikes has significantly increased over the last few years, see e.g., Andersen et al. (2016). We use these short-dated options in the construction of our estimator. A natural candidate for a spot volatility estimator is provided by the Black-Scholes implied volatility 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., Carr and Madan (2001) as well as the earlier work of Ross (1976) and Green and Jarrow (1987).\footnote{This same principle has been used by the CBOE option exchange in the construction of their popular volatility VIX index which is model-free estimate of the risk-neutral expected future (one month) quadratic variation.} Using this insight, we construct portfolios of options, with different strikes and the same time to expiration, 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 more robust to jumps than the simple Black-Scholes implied volatility of options with strikes close to the current price.

We establish consistency of the proposed volatility estimator in an asymptotic setting in which the maturity of the options used in the estimation shrinks to zero together with the convergence of the available strike grid to the whole positive real line. We further derive a Central Limit Theorem (CLT) for our volatility estimator. The limiting distribution is determined by 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 Mancini (2001). 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 away from zero).

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, Barndorff-Nielsen and Shephard (2004b, 2006) propose so-called multipower variation statistics as a way to separate volatility from jumps while Mancini (2001, 2009) develops truncated variance estimator that achieves the same goal. More recently, Jacod and Todorov (2014) 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.\textsuperscript{2}

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.\textsuperscript{3} 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 test the option-based volatility estimator on simulated data, with characteristics that are similar to actual observed option data sets, and we find satisfactory performance of the developed procedure. We further apply our option-based method to study volatility of the S&P 500 market index over the period 2014-2015. We find option- and return-based volatility estimates that are on average very close. At the same time, the option-based volatility estimates in our application have significantly smaller asymptotic variances than the high-frequency ones and hence provide nontrivial efficiency improvements for measuring spot volatility.

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 illustrates the performance of the volatility estimator on simulated option data and Section 5 contains an empirical application. Section 6 concludes. The assumptions and proofs are given in Section 7.

\textsuperscript{2}The literature on nonparametric high-frequency volatility measurement is vast. A strand of it has also developed estimates for integrated functionals of volatility (other than the integrated volatility), see e.g., Barndorff-Nielsen et al. (2005), Jacod and Rosenbaum (2013), Mykland and Zhang (2009) and Todorov and Tauchen (2012).

\textsuperscript{3}The estimators of Barndorff-Nielsen and Shephard (2004b, 2006), Mancini (2001, 2009) and Jacod and Todorov (2014) are all measures of integrated over time volatility. However, by performing the estimation on local windows with asymptotically shrinking time span (as in e.g., Foster and Nelson (1996)), these estimators can be easily adopted to estimation of spot volatility.
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 with $X$ and is defined on the filtered probability space $(\Omega^{(0)}, \mathcal{F}^{(0)}, (\mathcal{F}_t^{(0)})_{t \geq 0}, \mathbb{P}^{(0)})$. 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 $Q$, which under no-arbitrage is locally equivalent to $\mathbb{P}^{(0)}$.\footnote{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 7.1).}

The dynamics of the log-price $x = \log(X)$ under $Q$ is given by

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

where $W$ is a Brownian motion, $\mu$ 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 $\tilde{\mu}$ is the martingale measure associated with $\mu$ ($W$ and $\nu_t$ are defined with respect to $Q$). The regularity conditions for the above quantities are given in Section 7.1.

Although equation (1) describes the dynamics of $x$ under $Q$, under no-arbitrage, $\sigma_t$ 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 $(a_t, \sigma_t, \nu_t)$.

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 trivial. With these normalizations, the theoretical values of the options we will use in our analysis are given by

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

$O_T(k)$ 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
\[
\int_t^{t+T} \int_{\mathbb{R}} \nu_s(dx)ds < \infty, \text{ a.s.}
\]  
(3)
Most 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.

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 (3) holds and under some weak regularity type assumptions for \((a_t, \sigma_t, \nu_t)\), it is easy to show that
\[
O_T(0) = \frac{\sqrt{T}}{\sqrt{2\pi}} \sigma_t + O_p(T), \text{ as } T \downarrow 0.
\]  
(4)
This bound on the error for recovering \(\sigma_t\) from \(O_T(0)\) is sharp and a large component of it is due to the jumps in \(X\). This can be illustrated using the seminal Merton (1976) jump-diffusion model for which a higher-order expansion of \(O_T(0)\) can be derived. In the Merton model the volatility is constant and the jumps are compound Poisson with intensity \(\lambda\) and their size is drawn from a normal distribution with mean \(\mu_j\) and variance \(\sigma_j^2\). 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, \(O^M_T(0)\), the following
\[
O^M_T(0) = \frac{\sqrt{T}}{\sqrt{2\pi}} \sigma - \frac{T\sigma^2}{4} + \lambda T \left( \Phi \left( -\frac{\mu_j}{\sigma_j} \right) - e^{\mu_j+\sigma_j^2/2} \Phi \left( -\frac{\mu_j+\sigma_j^2/2}{\sigma_j} \right) \right) + O(T^{3/2}), \text{ as } T \downarrow 0,
\]  
(5)
where \(\Phi\) denotes the cdf of a standard normal random variable. The first two terms on the right-hand side of (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 (4) is sharp. Using the ATM option price expansion in (4), we have
\[
V_t = \frac{2\pi}{T} O_T^2(0) + O_p(\sqrt{T}), \text{ as } T \downarrow 0,
\]  
(6)
and we can alternatively estimate \(V_t\) using the Black-Scholes implied volatility corresponding to \(O_T(0)\) (with obviously the same order of magnitude of the approximation error as above). On 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 generate additional bias in the measurement of spot volatility.

![Figure 1: ATM BSIV vs. Spot Volatility](image)

We now develop an alternative strategy for recovering spot volatility from short-dated options which will have much smaller approximation error. Our strategy builds on the fact that the conditional expectation (under $Q$) of any sufficiently smooth functions of $x_{t+T}$ can be spanned by a portfolio of options with continuum of strikes, $\{O_T(k)\}_{k \in \mathbb{R}}$, see e.g., Carr and Madan (2001). We note that this spanning result lies also behind the construction of the popular volatility VIX index.

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 $\{O_T(k)\}_{k \in \mathbb{R}}$, we can recover $\mathbb{E}_t^Q\left(e^{iu(x_{t+T} - x_t)}\right)$ (see the expression in (12) below for the explicit formula). For an appropriate choice of $u$, as we now show, we can disentangle volatility from jumps using $\mathbb{E}_t^Q\left(e^{iu(x_{t+T} - x_t)}\right)$.

To help intuition, let's first assume that $x_{t+T} - x_t$ is, $\mathcal{F}_t$-conditionally, a Lévy process under $Q$ (i.e., a process with i.i.d. increments). In this case, the Lévy-Khintchine formula implies

$$
\log \left( \mathbb{E}_t^Q \left( e^{iu(x_{t+T} - x_t)/\sqrt{T}} \right) \right) = iu \sqrt{T} a_t - \frac{u^2}{2} V_t + T \int_{\mathbb{R}} (e^{iuT^{-1/2}x} - 1 - iuT^{-1/2}x) \nu_t(dx). \quad (7)
$$

Using our finite activity jump assumption in (3), we easily have $\int_{\mathbb{R}} (\cos(uT^{-1/2}x) - 1) \nu_t(dx) = O_p(1)$,
and therefore
\[ V_t = -\frac{2}{u^2} \Re \left( \log \left( \mathbb{E}_t^Q \left( e^{iu(x_{t+T} - x_t)/\sqrt{T}} \right) \right) \right) + O_p(T), \quad \text{as } T \downarrow 0. \] (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 (6) and (8), we can see that the characteristic function based approach has asymptotically smaller error than the ATM BSIV for estimating the spot volatility.

On Figure 2, we illustrate the accuracy of the expression on the right-hand-side of (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 \( u \downarrow 0 \),
\[ -\frac{2}{u^2} \Re \left( \log \left( \mathbb{E}_t^Q \left( e^{iu(x_{t+T} - x_t)/\sqrt{T}} \right) \right) \right) \]
converges to the (expected) spot quadratic variation, \( V_t + \int_{\mathbb{R}} x^2 \nu_t(dx) \), 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. 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, followed by a formal definition of the estimator in Section 3.2 and derivation of its asymptotic order. Section 3.3 proposes an option-based truncation volatility which we use in Section 3.4 to select the characteristic exponent of the volatility estimator in a data-driven way. This section further presents a feasible CLT for the estimator.

3.1 The Observation Scheme

Our data consists of OTM options at time \( t \), expiring at \( t + T \), and having log-strikes
\[ \underline{k} \equiv k_1 < k_2 < \cdots k_N \equiv \overline{k}, \] (9)
Figure 2: Characteristic Function Based Volatility Estimates. The plot displays $-\frac{2}{u^2}\Re\left(\log\left(\mathbb{E}_t^Q\left(e^{iu(x_t-T-x_t)/\sqrt{T}}\right)\right)\right)$ as a function of $u$ for the model in the Monte Carlo experiment in Section 4 (case A-H). The solid line corresponds to two business days and the dashed one to five business days. The true value of the spot variance is marked with the dotted horizontal line.

with the corresponding strikes given by

$$K \equiv K_1 < K_2 < \cdots K_N \equiv \overline{K}. \quad (10)$$

We denote the gaps between the log-strikes with $\Delta_i = k_i - k_{i-1}$, for $i = 2, \ldots, N$. We note that we do not assume an equidistant log-strike grid, i.e., we allow for $\Delta_i$ 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 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 $O_T(k_i)$ directly, we observe:

$$\hat{O}_T(k_i) = O_T(k_i) + \epsilon_i, \quad (11)$$

where the sequence of observation errors $\{\epsilon_i\}_{i \geq 1}$ is defined on a space $\Omega^{(1)} = \bigtimes_{k \in \mathbb{R}} \mathcal{A}_k$, for $\mathcal{A}_k = \mathbb{R}$. This space is equipped with the product Borel $\sigma$-field $\mathcal{F}^{(1)}$ and with transition probability $\mathbb{P}^{(1)}(\omega^{(0)}, d\omega^{(1)})$ from the original probability space $\Omega^{(0)}$ – on which $X$ is defined – to $\Omega^{(1)}$. We further define,

$$\Omega = \Omega^{(0)} \times \Omega^{(1)}, \quad \mathcal{F} = \mathcal{F}^{(0)} \times \mathcal{F}^{(1)}, \quad \mathbb{P}(d\omega^{(0)}, d\omega^{(1)}) = \mathbb{P}^{(0)}(d\omega^{(0)}) \mathbb{P}^{(1)}(\omega^{(0)}, d\omega^{(1)}).$$
We will assume $\mathbb{E}(\epsilon_i | \mathcal{F}^{(0)}) = 0$ and that $\epsilon_i$ and $\epsilon_j$ 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 of the observation error.

Our formal assumptions for the process $x$, the option observation scheme as well as the observation error are stated in Section 7.1.

### 3.2 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 Carr and Madan (2001), the conditional characteristic function of the log return, $\mathbb{E}_t^Q(e^{iu(x_{t+T} - x_t)})$, can be spanned by the following portfolio of options

$$1 - (u^2 + iu) \int_{\mathbb{R}} e^{(iu-1)k-iux}O_T(k)dk, \quad u \in \mathbb{R}. \quad (12)$$

The integral in the above expression is infeasible, given available data, because we do not have option observations over a continuum of strikes and furthermore we do not observe directly $O_T(k)$.

The feasible counterpart of the expression in (12) is formed by using Riemann sum approximation of the integral in (12) constructed from the available noisy option observations:

$$\hat{f}_{t,T}(u) = 1 - (u^2 + iu) \sum_{j=2}^{N} e^{(iu-1)k_{j-1}-iux}\hat{O}_T(k_{j-1})\Delta_j, \quad u \in \mathbb{R}. \quad (13)$$

While in general $x_{t+T} - x_t$ is not $\mathcal{F}$-conditionally the increment of a Lévy process, when $T$ is small, the expression for the characteristic function in (7) nevertheless holds approximately true. This motivates the following estimator of the volatility:

$$\hat{V}_{t,T}(u) = \frac{2}{Tu^2} \hat{R}_{t,T}(u), \quad (14)$$

where $\hat{R}_{t,T}(u)$ is the real part of $\hat{f}_{t,T}(u)$

$$\hat{R}_{t,T}(u) = -\Re \left( \ln \left( \hat{f}_{t,T}(u) \vee T \right) \right). \quad (15)$$

For $\hat{V}_{t,T}(u)$ to be a consistent estimator of $V_t$, we will need $\hat{f}_{t,T}(u)$ to converge in probability to the expression in (12) and for this we will need the discrete strike grid in (9) to cover asymptotically the whole real line and the time to maturity $T$ of the options to shrink to zero. The formal result for the consistency and rate of convergence of $\hat{V}_{t,T}(u)$ is given in the next theorem.

**Theorem 1** Suppose assumptions A1-A5 in Section 7.1 hold for some $r \in [0,1]$ and in addition

$$\Delta \asymp T^\alpha, \quad K \asymp T^{-\beta}, \quad K \asymp T^\gamma$$

for some $\alpha > \frac{1}{2}$, $\beta > 0$ and $\gamma > 0$. Let $(u_T)$ be an $\mathcal{F}_t^{(0)}$-adapted sequence such that

$$u_T^2 \overset{a.s.}{\rightarrow} \overline{u}, \quad \text{where } \overline{u} \text{ is a finite nonnegative random variable.} \quad (16)$$
Then, we have

\[ \hat{V}_{t,T}(u_T) - V_t = O_p \left( \frac{\sqrt{\Delta T}}{T^{3/4}} \sqrt{\frac{\Delta T}{T^{3/4}}} \right). \]  

(17)

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 \( \hat{f}_{t,T}(u) \) will be zero (recall (7)) and hence the signal about the volatility diminishes.

The rate of convergence result in (17) reveals the role of the different sources of error in the volatility estimation. The first term on the right-hand side of (17) is a bias due to the presence of jumps in \( X \). The parameter \( r \) controls the so-called jump activity (see assumption A2-r in the appendix), 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 \( r = 0 \). From (17), 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 (17) is due to the observation error, i.e., due to the fact that we use \( \hat{O}_T(k) \) in the estimation instead of \( O_T(k) \). 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 in the appendix). This is intuitive and is motivated by the empirical evidence in Andersen et al. (2015a) 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 that of the observation errors attached to the options). In particular, for log-strikes which are within \( \sqrt{T} \) range away from the current log-price, 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 (13) we use options with all available strikes. The above discussion implies 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 (17) is due to the finite log-strike range of the available option data \((k, \bar{k})\) 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 \( \mathcal{F}_t \)-conditional distribution of \( x_{t+T} - x_t \). With stronger assumptions for the latter, than what is currently assumed in assumption A2 in the appendix, 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 \((k, \bar{k})\).

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 \(O_p(T)\) and hence it is asymptotically dominated by the first term on the right-hand side of (17) (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 (12) 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 \(\hat{V}_{t,T}(u_T)\) is an estimate of \(\frac{1}{T} \int_t^{t+T} V_s ds\) without any bias due to the time-variation in \(V\).

### 3.3 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 \(O(1/\sqrt{T})\). 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 (7), it is clear that its behavior is governed by the product \(T \times u_T^2 \times V_t\). Therefore, in practice, 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 need to show that our estimator \(\hat{V}_{t,T}(u_T)\) can be made adaptive, i.e., that \(u_T\) can be replaced with an estimate \(\hat{u}_T\) 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 \(\hat{V}_{t,T}(u_T)\) adaptive. Our initial consistent estimator can be viewed as the option analogue of the truncated volatility estimator of Mancini (2001) and is given by \(^5\)

\[
\overline{T}V_{t,T}(\eta) = \frac{1}{T} \sum_{j=2}^{N} h_\eta(k_{j-1}) \hat{O}_T(k_{j-1}) \Delta_j, \quad \eta \geq 0, \quad (18)
\]

\(^5\)An alternative choice for initial volatility estimator is the ATM BSIV which has the additional advantage of being free of tuning parameters.
where we denote
\[ h_\eta(k) = e^{-2k - \eta(k-x_t)^2} \left[ 4\eta^2(k-x_t)^4 + 2 - 10\eta(k-x_t)^2 + 2\eta(k-x_t)^3 - 2(k-x_t) \right]. \]

\( \hat{TV}_{t,T}(\eta) \) is a consistent estimator of \( \mathbb{E}^Q_t(e^{-\eta(x_{t+T}-x_t)^2}(x_{t+T}-x_t)^2) \) from the available options. In the special case when \( \eta = 0 \) we denote
\[ \hat{QV}_{t,T} \equiv \hat{TV}_{t,T}(0), \] (19)
and we note that \( \hat{QV}_{t,T} \) is an estimator of the expected risk-neutral spot quadratic variation
\[ QV_{t,T} = V_t + \int_{\mathbb{R}} x^2 \nu_t(dx). \] (20)

Thus, \( \hat{QV}_{t,T} \) is the option counterpart of the realized variance computed from return data (Andersen et al. (2003) and Barndorff-Nielsen and Shephard (2002, 2004a)).\(^6\) When \( \eta \) is a positive number, then \( \hat{TV}_{t,T}(\eta) \) estimates a truncated (conditional) second moment of the increment \( x_{t+T} - x_t \) with the degree of truncation determined by \( \eta \). When \( \eta \) 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 \( \hat{TV}_{t,T}(\eta) \) to separate volatility from jumps. This is analogous to the truncated volatility estimator based on return data proposed by Mancini (2001), with the difference being that, unlike Mancini (2001), 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 \( \hat{TV}_{t,T}(\eta) \) is necessary. We use the following data-driven choice for the cutoff parameter
\[ \hat{\eta}_T = \frac{\bar{\eta}_T}{T} \frac{1}{\hat{QV}_{t,T}}, \] (21)
for some deterministic sequence \( \bar{\eta}_T \) 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

\(^6\)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 \( \mathcal{F}_t \)-conditional expectation of this quantity for \( \tau \) small (and further standardized, i.e., divided, by \( \tau \)). While for small \( \tau \) we have \( V_t \approx \mathbb{E}^Q_t(V_{t+\tau}) \), the same does not hold for the expected and realized jumps no matter how small \( \tau \) is and regardless of whether the jump intensity varies over time or not (i.e., whether \( \nu_t \) depends on \( t \) or not).
this way is that the downward bias in \( \hat{T}V_{t,T}(\eta) \) 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 7.1 hold for some \( r \in [0, 1] \) and in addition \( \Delta \propto T^\alpha, K \propto T^{-\beta}, K \propto T^\gamma \) for some \( \alpha > \frac{1}{2}, \beta > 0 \) and \( \gamma > 0 \). We have

\[
\hat{QV}_{t,T} \xrightarrow{p} QV_{t,T}. \tag{22}
\]

Suppose in addition that for \( \bar{\eta}_T \) in (21):

\[
\bar{\eta}_T \to 0 \text{ and } \frac{\bar{\eta}_T}{T} \to \infty. \tag{23}
\]

Then, we also have

\[
\hat{T}V_{t,T}(\bar{\eta}_T) \xrightarrow{p} V_t. \tag{24}
\]

The result in (22) 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 (22) which will allow us to assess the precision in the recovery of the jump part of the quadratic variation. For brevity, as our main focus is volatility estimation, we do not pursue this any further here.

### 3.4 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 \( \mathcal{F}^{(0)}_t \)-adapted and this rules out the case where \((u_T)\) depends on the option data used in the construction of \( \hat{V}_{t,T}(u_T) \). The goal of this section is to make the volatility estimator adaptive by making \( u_T \) a function of our preliminary truncated volatility \( \hat{T}V_{t,T}(\bar{\eta}_T) \). In particular, we set the characteristic exponent in the construction of \( \hat{V}_{t,T}(u) \) in the following data-driven way (recall our discussion at the beginning of Section 3.3)

\[
\tilde{u}_T = \frac{\bar{u}}{\sqrt{T}} \frac{1}{\sqrt{\hat{T}V_{t,T}(\bar{\eta}_T)}}, \tag{25}
\]

where \( \bar{u} \) is some fixed positive number that does not depend on the option data.

We will further derive a feasible CLT for \( \hat{V}_{t,T}(\tilde{u}_T) \) 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

\[
\hat{\sigma}_j = \hat{O}_T(k_j) - \frac{1}{2} \left( \hat{O}_T(k_{j-1}) + \hat{O}_T(k_{j+1}) \right), \quad \text{for } j = 2, ..., N - 1 \text{ and } j \neq j^*, \tag{26}
\]
where \( j^* = \{ j = 1, \ldots, N : |k_j^* - x_t| \leq |k_j - x_t| \} \) and

\[
\hat{\epsilon}_1 = \hat{\epsilon}_2 \quad \text{and} \quad \hat{\epsilon}_{N-1} = \hat{\epsilon}_N,
\]

\[
\hat{\epsilon}_j^* = \begin{cases} \hat{\mathcal{O}}_T(k_j^*) - \hat{\mathcal{O}}_T(k_{j-1}^*) - (\hat{\mathcal{O}}_T(k_{j-1}^*) - \hat{\mathcal{O}}_T(k_{j-2}^*)) \frac{K_{j-1}^* - K_{j+1}^*}{K_{j-1}^* - K_{j+2}^*}, & \text{if } k_j^* \leq x_t, \\ \hat{\mathcal{O}}_T(k_j^*) - \hat{\mathcal{O}}_T(k_{j+1}^*) - (\hat{\mathcal{O}}_T(k_{j+1}^*) - \hat{\mathcal{O}}_T(k_{j+2}^*)) \frac{K_{j+1}^* - K_{j+2}^*}{K_{j+1}^* - K_{j+2}^*}, & \text{if } k_j^* > x_t. \end{cases}
\]

Since the true option price is smooth in \( k \) (this follows under standard weak assumptions for the \( Q \) dynamics of \( x \)), then for \( j = 2, \ldots, N-1 \) and \( j \neq j^* \), \( \hat{\epsilon}_j \) 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 \( \{\hat{\epsilon}_j\}_{j=2}^{N-1} \), we set

\[
\hat{\mathcal{C}}_{t,T}(u) = \frac{2}{3} \sum_{j=2}^{N} \left( u^2 \cos(uk_{j-1}) - u \sin(uk_{j-1}) \right) \left( u^2 \cos(uk_{j-1}) - u \sin(uk_{j-1}) \right)^\top 
\times e^{-2k_j^* \hat{\epsilon}_j^2 \Delta_j^2},
\]

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

\[
\hat{\text{Avar}}(\hat{V}_{t,T}(u)) = \frac{2}{Tu^2} \left| \hat{f}_{t,T}(u) \right|^4 \left( \Re \hat{f}_{t,T}(u) \right)^\top \hat{C}_{t,T}(u) \left( \Re \hat{f}_{t,T}(u) \right).
\]

Theorem 3 gives a feasible CLT for \( \hat{V}_{t,T}(\hat{u}_T) \). 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., Jacod and Shiryaev (2003) for further details.

**Theorem 3** Suppose assumptions A1-A5 in Section 7.1 hold for some \( r \in [0,1] \) and in addition \( \Delta \asymp T^\alpha, K \asymp T^{-\beta}, K \asymp T^\gamma \) for some \( \alpha > \frac{1}{2}, \beta > 0 \) and \( \gamma > 0 \). If

\[
\alpha < \left( \frac{1}{2} + 2 - r \right) \wedge \left( \frac{1}{2} + 4(\beta \wedge \gamma) \right),
\]

then

\[
\frac{\hat{V}_{t,T}(\hat{u}_T) - V_t}{\sqrt{\hat{\text{Avar}}(\hat{V}_{t,T}(\hat{u}_T))}} \stackrel{\mathcal{L} - s}{\Rightarrow} N(0,1),
\]

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

The condition in (29) ensures that the leading term in the difference \( \hat{V}_{t,T}(\hat{u}_T) - V_t \) 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.

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 Mancini (2001, 2009) in the comparison, but the results extend also to other return-based volatility estimators, e.g., the multipower variations of Barndorff-Nielsen and Shephard (2004b, 2006). The return-based truncated volatility estimator is given by

$$\widehat{V}_{hf}^t = \frac{n}{k_n} \sum_{i \in I^n_t} (\Delta^n_i x)^2 1_{\{|\Delta^n_i x| \leq \alpha n^{-\varpi}\}}, \quad \alpha > 0 \text{ and } \varpi \in (0, 1/2),$$

where $I^n_t = \{i = 1, ..., k_n : \lfloor tn \rfloor - i\}$ denotes a local window used for the calculation of the volatility and $\Delta^n_i x = x_i^n - x_{i-1}^n$. For the successful application of $\widehat{V}_{hf}^t$, it is important to set $\alpha$ in a data-driven way that accounts for the current level of volatility. We will use multipower variation as a preliminary estimator in setting $\alpha$, see the empirical section for further details. In the next theorem, we show that the convergence of $\widehat{V}_{t,T}(\widehat{u}_T)$ holds jointly with that of $\widehat{V}_{hf}^t$.

**Theorem 4** In addition to the conditions of Theorem 3 suppose also that assumption A6 in Section 7.1 holds and $k_n \asymp \sqrt{n}$. Then

$$\frac{\sqrt{k_n} (\widehat{V}_{hf}^t - V_t)}{\sqrt{2\widehat{V}_{hf}^t}} \xrightarrow{L^s} N(0, 1),$$

and this convergence holds jointly with the convergence in (30), with the limits defined on an extension of the original probability space and being independent of each other and $\mathcal{F}$.

We note that $\widehat{V}_{t,T}(\widehat{u}_T)$ and $\widehat{V}_{hf}^t$ 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

$$\widehat{V}_t^{mix} = \bar{w}_t \widehat{V}_{t,T}(\widehat{u}_T) + (1 - \bar{w}_t)\widehat{V}_{hf}^t, \quad \bar{w}_t = \frac{2 \left(\widehat{V}_{hf}^t\right)^2}{2 \left(\widehat{V}_{hf}^t\right)^2 + k_n \widehat{Avar}(\widehat{V}_{t,T}(\widehat{u}_T))}.$$

16
The noisier one of the two volatility estimators is, the less weight it receives in the combined estimator $\hat{V}^{mix}_t$. In fact, if one of the two estimators converges at a faster rate, then asymptotically the weight it receives in $\hat{V}^{mix}_t$ converges to one, i.e., it receives all the weight. A convenient feature of $\hat{V}^{mix}_t$ 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 $\hat{V}^{mix}_t$ automatically “adapts” to the situation at hand.

4 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:

$$X_t = X_0 + \int_0^t \sqrt{V_s} dW_s + \int_0^t \int_{\mathbb{R}} (e^x - 1) \tilde{\mu}(ds, dx),$$  \hspace{1cm} (34)

with $W$ being a Brownian motion and $V$ having the dynamics

$$dV_t = \kappa(\theta - V_t)dt + \sigma \sqrt{V_t} \rho dW_t + \sigma \sqrt{V_t} \sqrt{1 - \rho^2} d\tilde{W}_t,$$  \hspace{1cm} (35)

where $\tilde{W}$ is a Brownian motion orthogonal to $W$. The jump measure $\mu$ has a compensator $\nu_t(dx) \otimes dx$ for $\nu_t$ given by

$$\nu_t(dx) = c_+ V_t e^{-\lambda_+ |x|} \frac{|x|}{1 + \beta} dx 1_{\{x > 0\}} + c_- V_t e^{-\lambda_- |x|} \frac{|x|}{1 + \beta} dx 1_{\{x < 0\}}.$$  \hspace{1cm} (36)

The specification in (34)-(36) belongs to the affine class of models (Duffie et al. (2000)) 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 the tempered stable process which is found to provide good fit to observed option data. The parameter $\beta$ controls the behavior of the jump measure around zero, with $\beta < 0$ corresponding to finite jump activity and $\beta \geq 0$ to infinite activity jump specifications ($r$ in assumption A2-r equals $\beta \vee 0$ in this parametric model). The parameters $\lambda_{\pm}$, on the other hand, control the behavior of the jump measure in the tails while $c_{\pm}$ control the overall scale of the jump intensity. We set the model parameters in a way that results in option prices similar to observed equity index options. In particular, we set the parameters of the diffusive volatility to $\theta = 0.022$, $\kappa = 3.6124$, $\sigma = 0.2$ and $\rho = -0.5$ (our unit of time is one year). Further, in all considered cases we set the jump tail parameters to $\lambda_- = 20$ and $\lambda_+ = 100$, which are roughly in line with previous empirical studies, see e.g., Andersen et al. (2015b).
From our theoretical analysis in the previous section it is clear that the properties of the option-based volatility estimator depend in a critical way on the behavior of the jump measure around zero. Therefore, we experiment with different values of the parameters controlling the behavior of $\nu_t$ around zero in order to check the sensitivity of the volatility estimator to that feature of the model. In particular, we consider three cases for the jump activity parameter $\beta$: $\beta = -0.5$ (cases A), $\beta = 0.0$ (cases B) and $\beta = 0.25$ (cases C). In all cases the parameters $c_{\pm}$ are set so that the ratio of expected negative to positive jump variation is 10 to 1. For each case of jump activity, we further consider 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 parameter settings of all considered cases are given in Table 1. Finally, we set $X_0 = 2000$ and draw $V_0$ from the stationary distribution of $V_t$ under $Q$ (which is Gamma distribution with shape and scale parameters of $\frac{2\theta}{\sigma^2}$ and $\frac{\sigma^2}{2\kappa}$).

<table>
<thead>
<tr>
<th>Case</th>
<th>$\beta$</th>
<th>$c_-$</th>
<th>$c_+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-L</td>
<td>-0.50</td>
<td>0.3058 $\times$ 10$^3$</td>
<td>1.7097 $\times$ 10$^3$</td>
</tr>
<tr>
<td>A-M</td>
<td>-0.50</td>
<td>0.6177 $\times$ 10$^3$</td>
<td>3.4194 $\times$ 10$^3$</td>
</tr>
<tr>
<td>A-H</td>
<td>-0.50</td>
<td>0.9174 $\times$ 10$^3$</td>
<td>5.1291 $\times$ 10$^3$</td>
</tr>
<tr>
<td>B-L</td>
<td>0.00</td>
<td>0.9091 $\times$ 10$^2$</td>
<td>2.2727 $\times$ 10$^2$</td>
</tr>
<tr>
<td>B-M</td>
<td>0.00</td>
<td>1.8182 $\times$ 10$^2$</td>
<td>4.5454 $\times$ 10$^2$</td>
</tr>
<tr>
<td>B-H</td>
<td>0.00</td>
<td>2.7273 $\times$ 10$^2$</td>
<td>6.8181 $\times$ 10$^2$</td>
</tr>
<tr>
<td>C-L</td>
<td>0.25</td>
<td>0.4677 $\times$ 10$^2$</td>
<td>0.7820 $\times$ 10$^2$</td>
</tr>
<tr>
<td>C-M</td>
<td>0.25</td>
<td>0.9354 $\times$ 10$^2$</td>
<td>1.5640 $\times$ 10$^2$</td>
</tr>
<tr>
<td>C-H</td>
<td>0.25</td>
<td>1.4031 $\times$ 10$^2$</td>
<td>2.3460 $\times$ 10$^2$</td>
</tr>
</tbody>
</table>

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 $k = -8 \times \sigma_{ATM} \sqrt{T}$, where we denote with $\sigma_{ATM}$ 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 $k = 2.5 \times \sigma_{ATM} \sqrt{T}$.
We add observation error to the model-implied option prices equal to 
\[ \epsilon_i = \frac{1}{2} Z_i \times OT(k_i) \frac{\psi(k_i)}{Q_{0.995}}, \]
where \( \{Z_i\}_{i=1,...,N} \) is a sequence of i.i.d. standard normal random variables, \( Q_\alpha \) denotes the \( \alpha \)-quantile of the standard normal, and \( \psi(k) \) is a function of the log-strike determined by running a nonparametric kernel regression on the data used in the empirical application of the relative option bid-ask spread (i.e. the bid-ask spread divided by the mid-quote) as a function of the volatility-adjusted log-strike \( k \times \sigma_{ATM} \sqrt{T} \).

To implement the option-based volatility estimator \( \hat{V}_{t,T}(\hat{u}_T) \) on the simulated data, we need to set \( \eta_T \) for the preliminary truncated volatility as well as \( \overline{u} \) for the adaptive characteristic exponent \( \hat{u}_T \) in (25). Recall that \( \eta_T \) is a deterministic sequence converging to zero. We put \( \eta_T = 0.5/\log(1/T) \), which for \( T = 2/252 \) takes value of 0.1034 and for \( T = 5/252 \) takes value of 0.1276.\(^7\) For \( \overline{u} \), we would like to pick it 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, \( \overline{u} = \sqrt{2 \log(1/0.1)} \) and \( \overline{u} = \sqrt{2 \log(1/0.075)} \), which correspond to \( |E^Q_t(e^{iu(x_t - x_t + T - x_t)})| \) having values of 0.1 and 0.075, respectively (recall from the discussion in Section 2 that for high \( u \) we have \( |E^Q_t(e^{iu(x_t + T - x_t)})| \approx e^{-u^2 V_{t,T}^2} \)).

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. In most of the cases, the estimator is nearly unbiased and volatility is recovered with good precision. Recall from (7) that the volatility estimator should contain a positive (asymptotically negligible) bias due to the jumps which is equal to 
\[ \frac{2}{u_T} \int_{\mathbb{R}} (1 - \cos(u_T x)) \nu_t(dx). \]
This bias is larger for the higher jump activity cases, i.e., the cases where the jump distribution has more mass around zero. This can be clearly seen from our Monte Carlo evidence reported in Table 2. Indeed, the highest bias in the Monte Carlo exercise occurs for the simulation scenario C-H which corresponds to the case of highest jump activity and larger overall contribution of jumps to the total return variance. On the other hand, for simulation scenarios A-L, A-M and A-H, which correspond to finite activity jumps that are typically used in finance, the option-based volatility estimator is nearly unbiased. Consistent also with theory (recall Figure 2), the bias is smallest for estimation based on the shortest-dated options.

We turn next to the performance of the confidence intervals based on our inference theory which we summarize in Table 3. For most scenarios and choices of the tuning parameter \( \hat{u}_T \), the coverage probability of the volatility confidence interval is close to the nominal one. We notice some

\(^7\)Our choice for \( \eta_T \) is motivated by the maximum downward bias in the measurement of volatility. By first-order Taylor expansion, in the case of no jumps, \( 3\eta_T \) is equal approximately to the relative negative bias in the measurement of the spot variance by \( TV_{t,T}(\eta_T) \).
<table>
<thead>
<tr>
<th>Case</th>
<th>Bias</th>
<th>MAE</th>
<th>Bias</th>
<th>MAE</th>
<th>Bias</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$T = 2$ days</td>
<td>$T = 3$ days</td>
<td>$T = 5$ days</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-L</td>
<td>-0.0008</td>
<td>0.0035</td>
<td>-0.0008</td>
<td>0.0033</td>
<td>-0.0010</td>
<td>0.0031</td>
</tr>
<tr>
<td>A-M</td>
<td>-0.0001</td>
<td>0.0034</td>
<td>0.0000</td>
<td>0.0031</td>
<td>-0.0001</td>
<td>0.0029</td>
</tr>
<tr>
<td>A-H</td>
<td>0.0006</td>
<td>0.0033</td>
<td>0.0003</td>
<td>0.0028</td>
<td>0.0006</td>
<td>0.0026</td>
</tr>
<tr>
<td>B-L</td>
<td>-0.0007</td>
<td>0.0035</td>
<td>-0.0006</td>
<td>0.0034</td>
<td>-0.0006</td>
<td>0.0031</td>
</tr>
<tr>
<td>B-M</td>
<td>0.0003</td>
<td>0.0032</td>
<td>0.0005</td>
<td>0.0031</td>
<td>0.0006</td>
<td>0.0028</td>
</tr>
<tr>
<td>B-H</td>
<td>0.0010</td>
<td>0.0031</td>
<td>0.0013</td>
<td>0.0030</td>
<td>0.0018</td>
<td>0.0030</td>
</tr>
<tr>
<td>C-L</td>
<td>-0.0005</td>
<td>0.0034</td>
<td>-0.0004</td>
<td>0.0032</td>
<td>-0.0003</td>
<td>0.0031</td>
</tr>
<tr>
<td>C-M</td>
<td>0.0006</td>
<td>0.0033</td>
<td>0.0010</td>
<td>0.0032</td>
<td>0.0011</td>
<td>0.0029</td>
</tr>
<tr>
<td>C-H</td>
<td>0.0016</td>
<td>0.0032</td>
<td>0.0020</td>
<td>0.0033</td>
<td>0.0025</td>
<td>0.0033</td>
</tr>
</tbody>
</table>

\[ \bar{\mu} = \sqrt{2 \log(1/0.075)} \]

<table>
<thead>
<tr>
<th>Case</th>
<th>Bias</th>
<th>MAE</th>
<th>Bias</th>
<th>MAE</th>
<th>Bias</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$T = 2$ days</td>
<td>$T = 3$ days</td>
<td>$T = 5$ days</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-L</td>
<td>0.0001</td>
<td>0.0029</td>
<td>0.0000</td>
<td>0.0029</td>
<td>-0.0001</td>
<td>0.0026</td>
</tr>
<tr>
<td>A-M</td>
<td>0.0009</td>
<td>0.0029</td>
<td>0.0007</td>
<td>0.0026</td>
<td>0.0006</td>
<td>0.0023</td>
</tr>
<tr>
<td>A-H</td>
<td>0.0015</td>
<td>0.0028</td>
<td>0.0015</td>
<td>0.0026</td>
<td>0.0014</td>
<td>0.0023</td>
</tr>
<tr>
<td>B-L</td>
<td>0.0003</td>
<td>0.0029</td>
<td>0.0004</td>
<td>0.0028</td>
<td>0.0001</td>
<td>0.0025</td>
</tr>
<tr>
<td>B-M</td>
<td>0.0010</td>
<td>0.0027</td>
<td>0.0012</td>
<td>0.0026</td>
<td>0.0012</td>
<td>0.0024</td>
</tr>
<tr>
<td>B-H</td>
<td>0.0019</td>
<td>0.0029</td>
<td>0.0021</td>
<td>0.0028</td>
<td>0.0024</td>
<td>0.0028</td>
</tr>
<tr>
<td>C-L</td>
<td>0.0003</td>
<td>0.0030</td>
<td>0.0004</td>
<td>0.0027</td>
<td>0.0002</td>
<td>0.0026</td>
</tr>
<tr>
<td>C-M</td>
<td>0.0014</td>
<td>0.0029</td>
<td>0.0015</td>
<td>0.0027</td>
<td>0.0017</td>
<td>0.0026</td>
</tr>
<tr>
<td>C-H</td>
<td>0.0023</td>
<td>0.0031</td>
<td>0.0028</td>
<td>0.0032</td>
<td>0.0030</td>
<td>0.0033</td>
</tr>
</tbody>
</table>

Table 2: Monte Carlo Results: Bias and MAE. MAE stands for mean absolute error.
deviation for the simulation scenario C-H when the highest of the two considered choices for \( \hat{\sigma}_T \) is used and for the longest time to maturity \( T \) of the option data. The reason for the under-coverage of the constructed confidence interval in this case is the bias in the volatility estimation due to the jumps. Recall from the discussion in the previous paragraph that this bias is largest exactly in the above configuration. Nevertheless, we note that either a lower level for \( \hat{\sigma}_T \) or a smaller \( T \) will fix this problem, as evident from the coverage probability of the confidence interval in the corresponding settings reported in Table 3.

<table>
<thead>
<tr>
<th>Case</th>
<th>Coverage Rate</th>
<th>Coverage Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.90 0.95 0.90 0.95 0.95</td>
<td>0.90 0.95 0.95 0.95 0.95</td>
</tr>
<tr>
<td></td>
<td>( T = 2 ) days ( T = 3 ) days ( T = 5 ) days</td>
<td>( T = 2 ) days ( T = 3 ) days ( T = 5 ) days</td>
</tr>
<tr>
<td>A-L</td>
<td>0.86 0.90 0.83 0.88 0.82 0.86</td>
<td>0.91 0.94 0.89 0.93 0.88 0.91</td>
</tr>
<tr>
<td>A-M</td>
<td>0.89 0.92 0.86 0.90 0.85 0.89</td>
<td>0.94 0.96 0.92 0.95 0.91 0.94</td>
</tr>
<tr>
<td>A-H</td>
<td>0.90 0.93 0.89 0.92 0.88 0.92</td>
<td>0.95 0.97 0.92 0.96 0.91 0.96</td>
</tr>
<tr>
<td>B-L</td>
<td>0.86 0.90 0.85 0.89 0.84 0.88</td>
<td>0.91 0.94 0.91 0.94 0.88 0.92</td>
</tr>
<tr>
<td>B-M</td>
<td>0.91 0.94 0.89 0.92 0.90 0.93</td>
<td>0.93 0.96 0.93 0.95 0.93 0.96</td>
</tr>
<tr>
<td>B-H</td>
<td>0.92 0.94 0.92 0.95 0.91 0.95</td>
<td>0.94 0.96 0.91 0.95 0.87 0.93</td>
</tr>
<tr>
<td>C-L</td>
<td>0.89 0.92 0.87 0.90 0.85 0.89</td>
<td>0.91 0.94 0.90 0.94 0.89 0.92</td>
</tr>
<tr>
<td>C-M</td>
<td>0.91 0.94 0.91 0.93 0.91 0.94</td>
<td>0.94 0.96 0.92 0.96 0.92 0.96</td>
</tr>
<tr>
<td>C-H</td>
<td>0.93 0.96 0.93 0.95 0.91 0.95</td>
<td>0.92 0.96 0.88 0.93 0.81 0.89</td>
</tr>
</tbody>
</table>

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.

5 Empirical Application

We now apply our volatility estimation procedures to real data on short-dated options on the S&P 500 index traded on the CBOE options exchange. Our sample covers the period 2014-2015, and we collect quotes on OTM options recorded at the end of each trading day. We look only at the
shortest-dated options which have at least two trading days till expiration. We further remove observations with zero bid quotes. After applying the above filters, the median size of the option cross-section consists of 77 options per day.\textsuperscript{8}

We set $\bar{\eta}_T$ and $\bar{\pi} = \sqrt{2 \log(1/0.1)}$ as in the Monte Carlo. On Figure 3 we plot the option-based volatility estimates $\tilde{V}_{t,T}(\tilde{u}_T)$ together with estimates for the total (risk-neutral) quadratic variation $\tilde{QV}_t$. The period covered in the study is relatively quiet as revealed by the levels of the recovered spot volatility. There are a few episodes where volatility increases sharply, most notably at the end of August 2015 when there were concerns among investors for the global impact of a slowdown in the Chinese economy. The gap between the spot volatility and the total (risk-neutral) quadratic variation, which is due to the presence of jump risk, is substantial both during quiet as well as turbulent times. We note that the difference $\tilde{QV}_t - \tilde{V}_{t,T}(\tilde{u}_T)$ is an estimate of the risk-neutral expected jump variation which differs from the expected jump variation under the true probability by the (nontrivial) risk premium demanded by investors for bearing such risk.

![Figure 3: Option-Implied Measures. The solid line is $\tilde{V}_{t,T}(\tilde{u}_T)$ and the dashed one is $\tilde{QV}_t$.](image)

We next compare the option-based volatility estimator with a volatility estimate formed from high-frequency return data. We use the E-mini S&P 500 index futures contract traded on CME to

\textsuperscript{8}Most of the options in our analysis are SPX Weeklies which expire at the end of the trading day on Friday (in 2016 the CBOE introduced also Monday and Wednesday SPX Weeklies and this further shifts down the maturity of the shortest-dated available options). We also include in our analysis the standard SPX options whenever they get sufficiently close to expiration (so that they are the shortest available options on the S&P 500 index for that day). The standard SPX options expire at the opening of trading on Friday. Therefore, the time to expiration for these options is a fixed number of trading days plus the overnight period from market close on the Thursday preceding the Friday AM expiration. We use the ratio of variances of trading hour returns and overnight returns to determine the fraction of time in business days that is attributed to an overnight period.
compute the return-based volatility estimates. We sample every five minutes in order to minimize the impact of market microstructure noise in the high-frequency data. We further set $k_n = 24$ which corresponds to window length of 2 trading hours covering the period prior to market close. Finally, in order to annualize the high-frequency volatility estimates, we compute the ratio of the local window return variance over the close-to-close return variance. To increase precision, we compute this variance ratio using return data covering the period 2009-2015.

The tuning parameters of $\hat{V}^hf_t$ are set as follows. We use $\pi = 0.49$ and we set $\alpha$ to

$$\alpha = 3\sqrt{BV_t \wedge RV_t \wedge BV_{t-1} \wedge RV_{t-1}},$$

where

$$RV_t = \sum_{i=[tn]-n+1}^{[tn]} (\Delta^y_n x)^2, \quad BV_t = \frac{\pi}{2} \frac{n}{n-1} \sum_{i=[tn]-n+2}^{[tn]} |\Delta^y_{i-1} x| |\Delta^y_n x|.$$  

$RV_t$ is the daily realized volatility and $BV_t$ is the bipower variation of Barndorff-Nielsen and Shephard (2004b, 2006) which is a nonparametric and free of tuning parameters measure of the integrated daily volatility that is robust to jumps.

On Figure 4, we compare the two spot volatility estimators $\hat{V}_{t,T}(\hat{u}_T)$ and $\hat{V}^hf_t$. Overall the two series are quite close in terms of time series variation. However, the high-frequency volatility measure is considerably noisier than the option-based one. This is a manifest of the efficiency gains offered by the use of option data for the purposes of spot volatility recovery. We further note that the three highest peaks in the high-frequency volatility measure are associated with more modest increases in the corresponding option-based volatility estimates. There were large intraday swings in the market price for each of these three days in the data (one of them was the day of an FOMC announcement and the other two were days where the market experienced a severe drop followed by a sharp rebound during the trading day). This makes the separation of the jumps from volatility from high-frequency return data particularly difficult. This can at least partially explain the highly elevated levels of $\hat{V}^hf_t$ on these days.\footnote{\hat{V}^hf_t for these three days exceeds even $\hat{Q}V_{t,T}$ which includes also the risk-neutral second moment of the jumps.}

6 Conclusion

In this paper we develop a model-free measure of volatility from options that is robust to jumps in the underlying asset. The proposed volatility measure is based on nonparametric estimates of the conditional characteristic function of the asset return which in turn are constructed from portfolios of short-dated options. The volatility dominates asymptotically the real part of the
characteristic function of the asset return over short time intervals and for sufficiently high values of the characteristic exponent. We use this fact to build the volatility estimator from the conditional characteristic function of the returns spanned by the option portfolios. We further derive a feasible CLT for the option-based volatility estimator which holds jointly with the corresponding CLT for high-frequency return-based volatility estimators. The derived limit theory allows for optimally combining return and option data for the purposes of nonparametric spot volatility estimation.

7 Proofs

In the assumptions and 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 Assumptions

A1. $V_t > 0$ and the process $\sigma$ has the following dynamics under $Q$:

$$
\sigma_s = \sigma_t + \int_t^s b_u du + \int_t^s \eta_u dW_u + \int_t^s \tilde{\eta}_u d\tilde{W}_u + \int_t^s \int_E \delta^\sigma(u,z) \mu^\sigma(du,dz), \quad s \geq t,
$$

where $\tilde{W}$ is a Brownian motion independent of $W$; $\mu^\sigma$ is a Poisson random measure on $\mathbb{R} \times E$ with compensator $\nu^\sigma(du,dz) = du \otimes dz$, having arbitrary dependence with the random measure $\mu$; $b$, $\eta$
and $\tilde{\eta}$ are processes with cadlag paths and $\delta^\sigma(u,z)$ 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 $s \in [t,\overline{t}]$:

$$
\mathbb{E}_t^Q|a_s|^4 + \mathbb{E}_t^Q|\sigma_s|^6 + \mathbb{E}_t^Q(e^{4|x_s|} + \int_\mathbb{R} [(e^{\delta^\sigma(u,z)} - 1) \lor |z|^r] \nu_s(z)dz)^4 < C_t,
$$

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

$$
\mathbb{E}_t^Q\left(\int_\mathbb{E} (|\delta^\sigma(s,z)|^4 \lor |\delta^\sigma(s,z)|)dz\right)^{1+\epsilon} < C_t.
$$

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

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

$$
\mathbb{E}_t^Q\left(\int_\mathbb{R} (e^{\xi_0 |z|} \lor |z|^2)|\nu_s(z) - \nu_t(z)|dz\right)^p \leq C_t|s - t|, \quad p \in [2, 3],
$$

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

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

**A4.** On a set with probability approaching one, we have

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

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) $\mathbb{E}\left(\epsilon_i |\mathcal{F}(0)\right) = 0$, (2) $\mathbb{E}\left(\epsilon_i^2 |\mathcal{F}(0)\right) = O_T(k_i)\sigma_{t,i}^2$ 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) $\mathbb{E}\left(|\epsilon_i|^4 |\mathcal{F}(0)\right) \leq \zeta_t O_T(k_i)^4$ for some finite-valued $\mathcal{F}_t^{(0)}$-adapted random variable $\zeta_t$, and (4) $\epsilon_i$ and $\epsilon_j$ are $\mathcal{F}^{(0)}$-conditionally independent whenever $i \neq j$.

**A6.** The dynamics of $X$ and $\sigma$ under $\mathbb{P}$ is as (1) and (37) but with $W$, $\tilde{W}$ and $\mu^\sigma$ defined with respect to $\mathbb{P}$, and with $\mu$ having a compensator under $\mathbb{P}$ of the form $\nu^\mathbb{P}_t(x)dt \otimes dx$. Moreover, for a sequence of stopping times $(\tau_n)$ increasing to infinity and a sequence of functions $\Gamma_n(z)$ satisfying $\int_\mathbb{R} \Gamma_n(z)dz < \infty$, we have $\int_\mathbb{R} (|z| \lor 1)\nu^\mathbb{P}_t(dx) < \infty$ ad $|\delta^\sigma(t,z)| \leq \Gamma_n(z)$ for $t \leq \tau_n$. 

25
7.2 Decomposition, Notation and Auxiliary Results

The jump part of the process $x_t$ can be represented as an integral with respect to 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 Jacod and Protter (2012)) and upon suitably extending the probability space, we can write

$$
\int_0^t \int_{\mathbb{R}} \tilde{\mu}(ds,dx) \equiv \int_0^t \int_{E} \delta^x(s,z)\tilde{\mu}^x(ds,dz),
$$

where $\mu^x(ds,dz)$ is a Poisson measure on $\mathbb{R}_+ \times E$ with compensator $dt \otimes dz$, $\tilde{\mu}^x$ is the martingale counterpart of $\mu^x$, and $\delta^x$ 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 Lebesgue measure $dz$ under the map $z \rightarrow \delta^x(t, z)$ on the set $\{z : \delta^x(\omega, t, z) \neq 0\}$.

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

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

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

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

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

The second approximation is given by $\bar{x}_s = \bar{x}_s^c + \bar{x}_s^d$, where

$$
\bar{x}_s^c = a_t(s-t) + \int_t^s \sigma_u dW_u, \quad \bar{x}_s^d = \tilde{x}_s^d, \quad \bar{x}_s = \sigma_t + \eta_t(W_s - W_t) + \eta_t(\tilde{W}_s - \tilde{W}_t), \quad s \geq t.
$$

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

**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

$$
O_T(k) \leq C_t T \begin{cases} 
\frac{e^{2k}}{e^{-k} - 1}, & \text{if } k < 0, \\
\frac{1}{e^k - 1}, & \text{if } k > 0,
\end{cases}
$$

$$
|O_T(k) - \bar{O}_T(k)| \leq C_t |\ln T| T^{3/2},
$$

26
\[ |O_T(k) - \tilde{O}_T(k)| \leq C_t \left( T^{3/2} \sqrt{\frac{T^{3/2}}{|e^k - 1|} \wedge T} \right), \]
\[ \tilde{O}_T(k) \leq C_t \left( \sqrt{T} \wedge \frac{T}{|e^k - 1|} \right), \]
\[ \left\{ \begin{array}{ll}
    k < k_{l,t} & \Rightarrow \tilde{O}_T(k) \leq C_t \left( \frac{e^{2k}}{e^{k-k_{l,t}-1}} \right) \sqrt{T} \ln T, \\
k > k_{h,t} & \Rightarrow \tilde{O}_T(k) \leq C_t \left( \frac{1}{e^{k-k_{h,t}-1}} \right) \sqrt{T} \ln T, 
\end{array} \right. \]
\[ |O_T(k_1) - O_T(k_2)| \leq C_t \left[ \left( \frac{T}{k_2^2} \wedge 1 \right) 1_{\{|k_2| \leq 1\}} + \frac{T}{k_2^2} 1_{\{|k_2| > 1\}} \right] |e^{k_1} - e^{k_2}|, \]  
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 Qin and Todorov (2017). The bound in (54) can be proved exactly as Lemma 7 of Qin and Todorov (2017) using the integrability assumptions in A2-r.

**Lemma 2** Suppose assumptions A1-A3 hold. There exist \( \mathcal{F}_t \)-adapted random variables \( \tilde{t} > t \) and \( C_t > 0 \) that do not depend on \( k \) and \( T \), such that for \( T < \tilde{t} \) we have
\[ |\tilde{O}_T(k) - \tilde{O}_{\tilde{t}}(k)| \leq C_t T, \]
\[ \left| \tilde{O}_{\tilde{t}}(k) - f \left( \frac{k}{\sqrt{T} \sigma_t} \right) \sqrt{T} \sigma_t - (e^k - 1) \Phi \left( \frac{k}{\sqrt{T} \sigma_t} \right) \right| \leq C_t T, \quad \text{if} \ k \leq 0, \]
\[ \left| \tilde{O}_{\tilde{t}}(k) - f \left( \frac{k}{\sqrt{T} \sigma_t} \right) \sqrt{T} \sigma_t + (e^k - 1) \left( 1 - \Phi \left( \frac{k}{\sqrt{T} \sigma_t} \right) \right) \right| \leq C_t T, \quad \text{if} \ k > 0, \]
where \( f \) and \( \Phi \) 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 \( k \leq 0 \) being done in analogous way. First, we have
\[ |e^{\tilde{w}_{t+T}} - e^k| + |e^{\sigma_t(W_{t+T} - W_t)} - e^k| \leq e^{\sigma_t(W_{t+T} - W_t)} |e^{a_t T + \tilde{d}_{t+T}} - 1|, \]
From here we can use the \( \mathcal{F}_t \)-conditional independence of \( W_{t+T} - W_t \) and \( \tilde{w}_{t+T} \) and apply Lemma 1 of Qin and Todorov (2017) to obtain the result in (55).

We continue with the bounds in (56). Direct calculation shows for \( k > 0 \)
\[ \tilde{O}_T(k) = 1 - \Phi \left( \frac{k}{\sqrt{T} \sigma_t} - \sqrt{T} \sigma_t \right) - e^k \left( 1 - \Phi \left( \frac{k}{\sqrt{T} \sigma_t} \right) \right). \]
From here the result of (56) follows by using Taylor expansion and the fact that the function \( f \) is bounded.

\[ \square \]
7.3 Proof of Theorem 1

We introduce the following notation

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

Using Appendix 1 of Carr and Madan (2001), \( f_{t,T}(u) \) equals the expression in (12). We further note that by Lévy-Khintchine formula

\[ \tilde{f}_{t,T}(u) = \exp \left( iTu_x - T\frac{u^2}{2} V_t + T \int_{\mathbb{R}} (e^{iux} - 1 - iux) \nu_t(dx) \right). \]  \hspace{1cm} (59)

We start the proof with establishing a bound for the difference \( f_{t,T}(u) - \tilde{f}_{t,T}(u) \). 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 \( \Re(f_{t,T}(u) - \tilde{f}_{t,T}(u)) \). Applying Itô’s lemma, using the normalization \( x_t = 0 \) and the integrability assumption A2, we have

\[
\mathbb{E}^Q_t(\cos(u(x_{t+T} - x_t))) - 1 = -\mathbb{E}^Q_t\left( \int_t^{t+T} u \sin(u\bar{x}_s)a_s ds + \frac{1}{2} \int_t^{t+T} u^2 \cos(u\bar{x}_s)\sigma^2_s ds \right) \\
+ \mathbb{E}^Q_t\left( \int_t^{t+T} \int_{\mathbb{R}} (\cos(u\bar{x}_s) \cos(uz) - 1) - \sin(u\bar{x}_s)(\sin(uz) - u \sin(z)) \nu_t(dz) ds \right) .
\]

We have an analogous expression for \( \mathbb{E}^Q_t(\cos(u\bar{x}_{t+T})) - 1 \). Then, using assumption A3 as well as \( \mathbb{E}^Q_t|x_s - \bar{x}_s| \leq C_t T \) 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

\[
\mathbb{E}^Q_t(\cos(u(x_{t+T}))) - 1 = -\mathbb{E}^Q_t\left( \int_t^{t+T} u \sin(u\bar{x}_s)a_s ds + \frac{1}{2} \int_t^{t+T} u^2 \cos(u\bar{x}_s)\sigma^2_s ds \right) \\
+ \mathbb{E}^Q_t\left( \int_t^{t+T} \int_{\mathbb{R}} (\cos(u\bar{x}_s) \cos(uz) - 1) - \sin(u\bar{x}_s)(\sin(uz) - u \sin(z)) \nu_t(dz) ds \right) + \zeta_{t,T}(u) .
\]

Therefore, we have

\[
\mathbb{E}^Q_t(\cos(u(x_{t+T}))) - \mathbb{E}^Q_t(\cos(u\bar{x}_{t+T})) = \frac{1}{2} \mathbb{E}^Q_t\left( \int_t^{t+T} u^2(\cos(u\bar{x}_s)\sigma^2_t - \cos(u\bar{x}_s)\sigma^2_s) ds \right) + \zeta_{t,T}(u) .
\]

Using the proof of Lemma 3 in Qin and Todorov (2017) as well as (42) of A3, we have \( \mathbb{E}^Q_t|x_s - \bar{x}_s| \leq C_t T^{3/2} \) 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 \( \mathbb{E}^Q_t |\sigma_s - \sigma_t| \leq C_t T \) for \( s \in [t, t + T] \) and \( C_t \) as before. Using these results and Cauchy-Schwartz inequality, we can write

\[
\mathbb{E}^Q_t (\cos(u x_{t+T})) - \mathbb{E}^Q_t (\cos(u \tilde{x}_{t+T})) = \frac{1}{2} \mathbb{E}^Q_t \left( \int_t^{t+T} u^2 (\cos(u \tilde{x}_s) \sigma^2_t - \cos(u \overline{x}_s) \overline{\sigma}^2_s) ds \right) + \zeta_{t,T}(u). \tag{60}
\]

Next, we can decompose

\[
\cos(u \overline{x}_s) \overline{\sigma}^2_s - \cos(u \tilde{x}_s) \sigma^2_t = (\cos(u \overline{x}_s) - \cos(u \tilde{x}_s)) \sigma^2_t + \cos(u \overline{x}_s) (\overline{\sigma}_s - \sigma_t)^2 + 2 (\cos(u \overline{x}_s) - \cos(u \tilde{x}_s)) (\overline{\sigma}_s - \sigma_t) \sigma_t + 2 \cos(u \overline{x}_s) (\overline{\sigma}_s - \sigma_t) \sigma_t. \tag{61}
\]

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

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

For the forth term on the right-hand side of (61), we can decompose \( \cos(u \overline{x}_s) \cos(u \tilde{x}_s^d) - \sin(u \overline{x}_s^c) \sin(u \tilde{x}_s^d) \). Then, using the symmetry of the density of the standard normal distribution as well as the independence of \( W \) and \( \tilde{W} \), we have

\[
\mathbb{E}^Q_t (\cos(u \overline{x}_s^c) \cos(u \tilde{x}_s^d), (\sigma_s - \sigma_t)) = - \sin(u a x_t (s - t)) \mathbb{E}^Q_t (\sin(u \sigma_t (W_s - W_t)) \cos(u \tilde{x}_s^d) (\sigma_s - \sigma_t)),
\]

which is \( \zeta_{t,T}(u) \) because \( \mathbb{E}^Q_t |\sigma_s - \sigma_t| \leq C_t \sqrt{T} \) for \( s \in [t, t + T] \). Using the \( \mathcal{F}_t \)-conditional independence of \( \tilde{x}_s^d \) and \( \sigma_s - \sigma_t \), we also have \( \mathbb{E}^Q_t |\sin(u \overline{x}_s^c) \sin(u \tilde{x}_s^d) (\sigma_s - \sigma_t)| \leq |u| \mathbb{E}^Q_t (|\tilde{x}_s^d| |\sigma_s - \sigma_t|) = |u| \mathbb{E}^Q_t |\tilde{x}_s^d| \mathbb{E}^Q_t |\sigma_s - \sigma_t| = \zeta_{t,T}(u) \). Combining these bounds, we have

\[
\mathbb{E}^Q_t (\cos(u \overline{x}_s) (\sigma_s - \sigma_t)) = \zeta_{t,T}(u).
\]

Finally, for the first term on the right-hand side of (61) using the independence of \( W \) and \( \tilde{W} \) and integration by parts, we can write

\[
\mathbb{E}^Q_t (\cos(u \overline{x}_s) - \cos(u \tilde{x}_s)) = - \mathbb{E}^Q_t (\sin(u \overline{x}_s) \sin(u \overline{x}_s - u \tilde{x}_s)) + \zeta_{t,T}(u)
\]

\[
= - u \mathbb{E}^Q_t (\sin(u \overline{x}_s) (\overline{x}_s - \tilde{x}_s)) + \zeta_{t,T}(u) = - \frac{u}{2} \eta \mathbb{E}^Q_t (\sin(u \tilde{x}_s) (W_s - W_t)^2) + \zeta_{t,T}(u)
\]

\[
= - \frac{u}{2} \eta \mathbb{E}^Q_t (\sin(u \sigma_t (W_s - W_t)) (W_s - W_t)^2) = \zeta_{t,T}(u),
\]

where for the last two equalities we have made use of the \( \mathcal{F}_t \)-conditional independence of \( \overline{x}_s^c \) and \( \overline{x}_s^d \) as well as the symmetry of the density of the standard normal distribution. Altogether we get that

\[
\mathbb{E}^Q_t \left( \int_t^{t+T} u^2 (\cos(u \overline{x}_s) \sigma^2_t - \cos(u \tilde{x}_s) \sigma^2_t) ds \right) = \zeta_{t,T}(u) \). Combining this result with (60), we have

\[
\Re (\tilde{f}_{t,T}(u) - \tilde{f}_{t,T}(u)) = \zeta_{t,T}(u). \tag{62}
\]

29
Turning to $\Im(f_{t,T}(u) - \tilde{f}_{t,T}(u))$, by making use of $\mathbb{E}_t^Q|\tilde{x}_{t+T} - x_{t+T}| \leq C_t T$, we have

$$\mathbb{E}_t^Q(\sin(u\tilde{x}_{t+T})) - \mathbb{E}_t^Q(\sin(u\tilde{x}_{t+T})) \leq C_t |u| T.$$ (63)

The results in (62) and (63), together with the rate condition for the sequence $u_T$ in (16), imply

$$\Re(f_{t,T}(u_T)) - \Re(\tilde{f}_{t,T}(u_T)) = O_p(T), \quad \Im(f_{t,T}(u_T)) - \Im(\tilde{f}_{t,T}(u_T)) = O_p(\sqrt{T}).$$ (64)

From (59), we also have

$$\Re(\tilde{f}_{t,T}(u_T)) = O_p(1), \quad \Im(\tilde{f}_{t,T}(u_T)) = O_p(\sqrt{T}).$$ (65)

From here, using Taylor expansion, we have

$$\Re(\ln(f_{t,T}(u_T))) - \Re(\ln(\tilde{f}_{t,T}(u_T))) = O_p(T).$$ (66)

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

$$\left|\Re(\ln(\tilde{f}_{t,T}(u_T))) + Tu_T^2 \sigma_t^2\right| \leq 2T|u_T|^r \int_{\mathbb{R}} |x|^r \nu_t(dx).$$ (67)

Altogether, we have

$$-\frac{2}{T u_T^2} \Re(\ln(f_{t,T}(u_T))) - \sigma_t^2 = O_p(u_T^{-2}).$$ (68)

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

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

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

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

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

$$\hat{f}_{t,T}^{(1)} = O_p\left(\frac{\sqrt{\Delta}}{T^{1/4}}\right), \quad \hat{f}_{t,T}^{(2)} = O_p\left(\frac{\Delta}{\sqrt{T}}\ln T\right), \quad \hat{f}_{t,T}^{(3)} = O_p\left(e^{-2(|\epsilon|\wedge|\Delta|)}\right).$$ (69)

Combining the bounds in (68) and (69), using Taylor expansion and the rate condition in (16) as well as the conditions for the asymptotic behavior of $T$, $\Delta$, $\epsilon$ and $\Delta$ in the theorem, we have (17).
7.4 Proof of Theorem 2

We set \( f_\eta(x) = e^{-\eta x^2} x^2 \) for \( \eta \geq 0 \), and we denote

\[
\eta_T = \frac{\eta_T}{T} \frac{1}{QV_t},
\]

which is an \( \mathcal{F}_t \)-adapted random variable. Using Appendix 1 of Carr and Madan (2001), for every finite-valued, nonnegative and \( \mathcal{F}_t \)-adapted random variable \( \eta \), we have

\[
\mathbb{E}^Q_t (f_\eta(\bar{x}_{t+T} - \bar{x}_t)) = \int_{-\infty}^{\infty} h_\eta(k) \tilde{O}_T(k) dk.
\]

We will first show that \( \mathbb{E}^Q_t (f_0(\bar{x}_{t+T} - \bar{x}_t)) \) and \( \mathbb{E}^Q_t (f_\eta(\bar{x}_{t+T} - \bar{x}_t)) \) 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

\[
\mathbb{E}^Q_t (f_\eta(\bar{x}_{t+T} - \bar{x}_t)) = a_t \mathbb{E}^Q_t \left( \int_0^T f'_\eta(\bar{x}_{t+s} - \bar{x}_t) ds \right) + \frac{V_t}{2} \mathbb{E}^Q_t \left( \int_0^T f''_\eta(\bar{x}_{t+s} - \bar{x}_t) ds \right) \]

\[
+ \mathbb{E}^Q_t \left( \int_0^T \int_{\mathbb{R}} (f_\eta(\bar{x}_{t+s} - \bar{x}_t + z) - f_\eta(\bar{x}_{t+s} - \bar{x}_t) - f'_\eta(\bar{x}_{t+s} - \bar{x}_t) z) d\nu_t(dz) \right),
\]

for any \( \mathcal{F}_t \)-adapted \( \eta \). Using then the fact that (which follows by (46) and the integrability conditions in A2-r)

\[
|\mathbb{E}^Q_t (\bar{x}_{t+s} - \bar{x}_t)| \leq C_t T, \quad \text{for } s \in [t, t + T],
\]

we have

\[
\left| \frac{1}{T} \mathbb{E}^Q_t (f_0(\bar{x}_{t+T} - \bar{x}_t)) - V_t - \int_{\mathbb{R}} z^2 \nu_t(dz) \right| = O_p(T).
\]

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

\[
|f'_\eta(x)| \leq C|x| \quad \text{and} \quad |f''_\eta(x) - 2| \leq C \eta x^2,
\]

\[
|f_\eta(x + z) - f_\eta(x) - f'_\eta(x) z - f_\eta(z)| \leq C \left( |x| |z| + \eta |x|^2 |z|^2 + \eta |x|^3 |z| + e^{-\frac{2}{\eta} |z|^2} |z|^2 \right).
\]

Using the above bounds and (72), the inequality \( |x| e^{-|x|^2} \leq C \), the fact that \( \int_{\mathbb{R}} |z| \nu_t(dz) < \infty \) (due to A2-r), the bound \( \mathbb{E}^Q_t |\bar{x}_t - \bar{x}_t|^2 \leq C_t T \) for \( s \in [t, t + T] \), as well as the integrability assumptions in A2-r, we have for \( \eta_T \) in (70):

\[
\left| \frac{1}{T} \mathbb{E}^Q_t (f_\eta(\bar{x}_{t+T} - \bar{x}_t)) - V_t \right| = O_p \left( \sqrt{T} \vee \eta_T T \vee \frac{1}{\sqrt{\eta_T}} \right).
\]

Given the results in (74) and (77), to prove the claims of Theorem 2, we need to show the asymptotic negligibility of \( \sqrt{QV_t} - \frac{1}{T} \int_{-\infty}^{\infty} h_0(k) \tilde{O}_T(k) dk \) and \( \sqrt{TV_{t,T}}(\eta_T) - \frac{1}{T} \int_{-\infty}^{\infty} h_{\eta_T}(k) \tilde{O}_T(k) dk \).
First, using the bounds of Lemma 1 and assumption A5 for the observation error, we can easily conclude
\[
\hat{Q}V_{t,T} - \frac{1}{T} \int_{-\infty}^{\infty} h_0(k) \hat{O}_T(k) dk = \mathcal{O}_p \left( \frac{\sqrt{\Delta}}{T^{1/4}} \sqrt{e^{-2||k|\sqrt{T}}} \right),
\]  
(78)
and this together with (74) establishes the consistency of the \( \hat{Q}V_{t,T} \). We continue with \( \hat{T}V_{t,T}(\hat{\eta}_T) \).

For its analysis, we first introduce the set
\[
\Omega = \left\{ \omega : |\hat{Q}V_{t,T} - QV_t| \leq \frac{1}{5} QV_t \right\}.
\]
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
\[
|h_\eta(k)| \leq C e^{-2k - \frac{\eta}{2}} (|k| \vee 1),
\]
(79)
\[
|h_a(k) - h_\eta(k)| \leq C |a - \eta| k^2 e^{-2k - \frac{\eta}{2}} (k^2 \vee 1 + (\eta \vee \eta^2) k^4),
\]
(80)
for some \( C \) that does not depend on \( \eta, a \), and \( k \) (recall that \( x_t = 0 \)).

Using the bounds of Lemma 1 and assumption A5 for the observation error as well as (74) and (78), we have for \( T \) being below some \( \mathcal{F}_t^{(0)} \)-adapted and positive random variable \( \zeta_t \) (so that \( \mathbb{P}(T > \zeta_t) \to 0 \) as \( T \to 0 \)):
\[
\mathbb{P}(\Omega^c | T < \zeta_t) \leq C_t \frac{X}{\sqrt{T}},
\]
(81)
and therefore \( 1(\Omega^c) \) is \( \mathcal{O}_p \left( \frac{X}{\sqrt{T}} \right) \). Further, using the inequality in (79), the bounds of Lemma 1 as well as assumption A5 for the observation error, we have \( \hat{T}V_{t,T}(\hat{\eta}_T) - \hat{T}V_{t,T}(\eta_T) = \mathcal{O}_p(1) \).

Therefore, altogether
\[
\left( \hat{T}V_{t,T}(\hat{\eta}_T) - \hat{T}V_{t,T}(\eta_T) \right) 1_{\{\Omega^c\}} = \mathcal{O}_p \left( \frac{X}{\sqrt{T}} \right),
\]
(82)
Next, since on the set \( \Omega \) we have \( |\hat{\eta}_T - \eta_T| \leq \frac{\eta}{4} \), we can apply (80) and bound
\[
\left| \hat{T}V_{t,T}(\hat{\eta}_T) - \hat{T}V_{t,T}(\eta_T) \right| 1_{\{\Omega\}} \leq C_t \frac{\hat{\eta}_T}{QV_{t,T}} |QV_{t,T} - QV_{t,T}| \times \frac{1}{T^2} \sum_{j=1}^{N} k_{j-1}^2 e^{-2k_{j-1} - \frac{2X}{\Delta} k_{j-1}^2} (k_{j-1}^2 \vee 1 + (\eta_T \vee \eta_T^2) k_{j-1}^4) \hat{O}_T(k_{j-1}) \Delta_j,
\]
(83)
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
\[
\frac{1}{T^2} \sum_{j=1}^{N} k_{j-1}^2 e^{-2k_{j-1} - \frac{2X}{\Delta} k_{j-1}^2} (k_{j-1}^2 \vee 1 + (\eta_T \vee \eta_T^2) k_{j-1}^4) \hat{O}_T(k_{j-1}) \Delta_j = \mathcal{O}_p(1),
\]
(84)
and taking into account the bounds for $\tilde{QV}_{t,T}$ in (74) and (78), we have altogether

$$
(\tilde{TV}_{t,T}(\eta_T) - \tilde{TV}_{t,T}(\eta_T)) 1_{\Omega} = O_p \left( \frac{\Delta}{T^{1/4}} \sqrt{\eta_T} \sqrt{\eta_T T} \sqrt{\eta_T e^{-2(|k| \vee k)}} (|k| \vee k) \right). \tag{85}
$$

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

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

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

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

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

$$
E(\tilde{TV}_{t,T}^{(1)}|\mathcal{F}^{(0)})^2 = O_p \left( \frac{\Delta}{\sqrt{T}} \right), \quad \tilde{TV}_{t,T}^{(3)} = O_p \left( e^{-k^2/k^2} \frac{1}{\sqrt{\eta_T} T} \right). \tag{86}
$$

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

$$
|h_\eta(k_2) - h_\eta(k_1)| \leq C_1|k_2 - k_1|e^{2k_2 - \frac{1}{2}k_1^2} (1 + |k_2|\eta_1) (1 + k_2^2 + \eta^2k_2^4), \quad \text{for } |k_1| \leq \frac{1}{2}|k_2|, \tag{87}
$$

where $C_1$ is a finite-valued $\mathcal{F}_t$-adapted random variable that does not depend on $k_1, k_2$ and $\eta$. Using this inequality, (79) as well as Lemma 1, we get

$$
\tilde{TV}_{t,T}^{(2)} = O_p \left( \frac{\Delta}{\sqrt{T}} \right). \tag{88}
$$

Combining the bounds in (86) and (88), we have altogether (the bound below is not sharp and can be further relaxed)

$$
\tilde{TV}_{t,T}(\eta_T) - \frac{1}{T} \int_{-\infty}^{\infty} h_{\eta_T}(k)\tilde{O}_T(k)dk = O_p \left( \frac{\sqrt{\Delta}}{T^{1/4}} \sqrt{e^{-2(|k| \vee k)}} (|k| \vee k) \right). \tag{89}
$$

This result, together with (77), (82) and (85), implies the consistency of $\tilde{TV}_{t,T}(\eta_T)$.

### 7.5 Proof of Theorem 3

We use the notation in (59). Using the result in (68) in the proof of Theorem 1, the fact that $\tilde{u}_T$ is $\mathcal{F}_t$-adapted, the consistency of $\tilde{TV}_{t,T}(\eta_T)$ for $V_t$ from Theorem 2 as well as the strict positivity of $V_t$ from assumption A1, we have

$$
-\frac{2}{T\tilde{u}_T^2} \Re(\ln (f_{t,T}(\tilde{u}_T))) - \sigma_t^2 = O_p(T^{1-r/2}). \tag{90}
$$
Using the fact that $f_t(T(u))$ equals the expression in (12), we then decompose $\widehat{f}_{t,T}(\hat{u}_T) - f_{t,T}(u) = \sum_{j=1}^{5} \widehat{f}_{t,T}^{(j)}$, where $\widehat{f}_{t,T}^{(j)} = -(\hat{u}_T + i\hat{u}_T)T\overline{f}_{t,T}^{(j)}$ and

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

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

$$\overline{f}_{t,T}^{(3)} = \frac{1}{T} \sum_{j=2}^{N} (e^{(i\hat{u}_T-1)k_{j-1}} - e^{(i\hat{u}_T-1)k_{j-1}})\epsilon_{j-1}\Delta_j,$$

$$\overline{f}_{t,T}^{(4)} = \frac{1}{T} \sum_{j=2}^{N} \int_{k_{j-1}}^{k_j} \left( e^{(i\hat{u}_T-1)k_{j-1}}O_T(k_{j-1}) - e^{(i\hat{u}_T-1)k}O_T(k) \right) dk,$$

$$\overline{f}_{t,T}^{(5)} = -\frac{1}{T} \int_{-\infty}^{k} e^{(i\hat{u}_T-1)k}O_T(k)dk - \frac{1}{T} \int_{k}^{\infty} e^{(i\hat{u}_T-1)k}O_T(k)dk.$$

We start with $\overline{f}_{t,T}^{(3)}$. For its analysis we introduce the set

$$\Omega = \left\{ \omega : |\overline{TV}_{t,T}(\hat{\eta}_T) - \overline{TV}_{t,T}(\eta_T)| \leq \frac{1}{4} V_t \cap |\overline{TV}_{t,T}(\eta_T) - V_t| \leq \frac{1}{4} V_t \right\}.$$

Using the fact that $V_t > 0$ by assumption A1, we have for a finite-valued and $\mathcal{F}_t$-adapted $C_t$ and $T$ below some strictly positive $\mathcal{F}_t$-adapted random variable:

$$|\overline{f}_{t,T}^{(3)}\{\Omega\}| \leq C_t \left( |\overline{TV}_{t,T}(\hat{\eta}_T) - \overline{TV}_{t,T}(\eta_T)|^2 + |\overline{TV}_{t,T}(\eta_T) - \frac{1}{T} \int_{-\infty}^{\infty} h_{\eta_t}(k)\overline{O}_T(k)dk|^2 \right)$$

$$\times \frac{1}{T} \sum_{j=2}^{N} e^{-k_{j-1}}|\epsilon_{j-1}|\Delta_j,$$

and

$$|\overline{f}_{t,T}^{(3)}\{\Omega\}| \leq C_t |\overline{TV}_{t,T}(\hat{\eta}_T) - \overline{TV}_{t,T}(\eta_T)| \times \frac{1}{T\sqrt{T}} \sum_{j=2}^{N} e^{-k_{j-1}}|\epsilon_{j-1}||k_{j-1}|\Delta_j.$$

Using assumption A5 for the observation error as well as the bounds in Lemma 1, we have

$$\frac{1}{T} \sum_{j=2}^{N} e^{-k_{j-1}}|\epsilon_{j-1}|\Delta_j = O_p(|\ln T|), \quad \frac{1}{T\sqrt{T}} \sum_{j=2}^{N} e^{-k_{j-1}}|\epsilon_{j-1}||k_{j-1}|\Delta_j = O_p(1).$$

34
Combining this result with the bounds in (82), (85), (86) and (88), we have altogether
\[
\mathcal{T}_{t,T}^{(3)} = O_p\left( \eta_T \frac{\sqrt{\Delta}}{T^{1/4}} \sqrt{\eta_T T} \sqrt{\eta_T e^{-2(|\Delta|\sqrt{T})}(|\Delta| \vee \Delta)} \right). \tag{93}
\]
Next, using the bounds of Lemma 1 as well as the consistency result for \( \widehat{TV}_{t,T}(\eta_T) \) of Theorem 2, we have
\[
\mathcal{T}_{t,T}^{(4)} = O_p\left( \frac{\Delta}{\sqrt{T}} \right), \quad \mathcal{T}_{t,T}^{(5)} = O_p\left( e^{-2(|\Delta| \vee |\Delta|)} \right). \tag{94}
\]
We proceed with \( \mathcal{T}_{t,T}^{(2)} \). We first introduce some additional notation for its analysis. We denote
\[
\widehat{TV}_{t,T}^{ij}(\eta_T) = \sum_{t=2,\ldots,N, \ t \neq i,j} h_{\eta_T}(k_{i-1}) \tilde{O}_T(k_{i-1}) \Delta_i + \sum_{t=i,j} h_{\eta_T}(k_{l-1}) O_T(k_{l-1}) \Delta_i, \quad i,j = 2, \ldots, N,
\]
in which \( i \)-th and \( j \)-th noisy option observations have been replaced with the true option prices. We further set
\[
\tilde{u}_{ij}^T = \frac{\eta}{\sqrt{T}} \frac{1}{\sqrt{\widehat{TV}_{t,T}^{ij}(\eta_T)}},
\]
and denote the set
\[
\Omega_{ij} = \left\{ \omega : |\widehat{TV}_{t,T}^{ij}(\eta_T) - V_i| \leq \frac{1}{4} V_i \cap |\widehat{TV}_{t,T}^{ij}(\eta_T) - V_i| \leq \frac{1}{4} V_i \right\}.
\]
Our goal is to evaluate the \( \mathcal{F}^{(0)} \)-conditional expectation of \( \mathcal{T}_{t,T}^{(3)} \) and for this we will first derive a sequence of inequalities. Using assumption A5 for the observation error, successive conditioning, the proof of Theorem 2, the fact that \( h_{\eta_T}(k) \) is bounded by a finite-valued and \( \mathcal{F}_T \)-adapted \( C_T \) by (79), and denoting with \( \alpha_T \) a deterministic sequence that determines the rate of convergence of \( \widehat{TV}_{t,T}(\eta_T) - V_t \), we have for \( T \) below some \( \mathcal{F}_T \)-adapted and positive-valued random variable (so that \( \widehat{TV}_{t,T}^{(2)}, \widehat{TV}_{t,T}^{(3)} \) and \( \frac{1}{T} \mathbb{E}_T (f_{\eta_T}(\tilde{x}_{t+T} - \tilde{x}_t)) - V_t \) are sufficiently small):
\[
\begin{align*}
\mathbb{E} & \left[ \left( \cos(\tilde{u}_{ij}^T k_{i-1}) - \cos(u_{T} k_{i-1}) \right) \left( \cos(\tilde{u}_T k_{j-1}) - \cos(u_{T} k_{j-1}) \right) \epsilon_{i-1} \epsilon_{j-1} 1_{\{\Omega_{ij}\}} \right] \mathcal{F}^{(0)} \\
& \leq C_T \left( \frac{\Delta}{\sqrt{T}} \right)^{3/2} O_T(k_{i-1}) O_T(k_{j-1}) + C_T \frac{\Delta^2}{T^2} \tilde{O}_T^2(k_{i-1}) O_T(k_{j-1}) + C_T \frac{\Delta^2}{T^2} O_T(k_{i-1}) \tilde{O}_T^2(k_{j-1}), \tag{95}
\end{align*}
\]
\[
\begin{align*}
\mathbb{E} & \left[ \left( \cos(\tilde{u}_T k_{i-1}) - \cos(u_{T} k_{i-1}) \right) \left( \cos(\tilde{u}_T k_{j-1}) - \cos(u_{T} k_{j-1}) \right) \epsilon_{i-1} \epsilon_{j-1} 1_{\{\Omega_{ij}\}} \right] \mathcal{F}^{(0)} \\
& \leq C_T \left( \frac{\Delta}{\sqrt{T}} \right)^{3/2} O_T(k_{i-1}) O_T(k_{j-1}) + C_T \frac{\Delta^2}{T^2} \tilde{O}_T^2(k_{i-1}) O_T(k_{j-1}) + C_T \frac{\Delta^2}{T^2} O_T(k_{i-1}) \tilde{O}_T^2(k_{j-1}), \tag{96}
\end{align*}
\]
\[
\begin{align*}
\mathbb{E} & \left[ \left( \cos(\tilde{u}_{ij}^T k_{i-1}) - \cos(u_{T} k_{i-1}) \right) \left( \cos(\tilde{u}_T k_{j-1}) - \cos(u_{T} k_{j-1}) \right) \epsilon_{i-1} \epsilon_{j-1} 1_{\{\Omega_{ij}\}} \right] \mathcal{F}^{(0)} \\
& \leq C_T \alpha_T \frac{\Delta}{T^2} |k_{i-1}| |k_{j-1}| \left( O_T^2(k_{i-1}) O_T(k_{j-1}) + O_T(k_{i-1}) \tilde{O}_T^2(k_{j-1}) \right), \tag{97}
\end{align*}
\]
Using again assumption A5 for the observation errors and the bounds of Lemma 1, we have altogether

\[ \mathbb{E} \left[ \left( \cos(\tilde{u}_T k_{i-1}) - \cos(\tilde{u}_T^j k_{i-1}) \right) \left( \cos(\tilde{u}_T k_{j-1}) - \cos(\tilde{u}_T^j k_{j-1}) \right) \epsilon_{i-1} \epsilon_{j-1} \left| \mathcal{F}^{(0)} \right. \right] = 0, \ i \neq j, \]  

(99)

The inequalities (95)-(100) continue to hold if in any of the products the function \( \cos() \) is replaced with \( \sin() \). From here, using the bounds in Lemma 1, we have altogether

\[ \mathbb{E} (\tilde{J}_{t,T}^{(2)} | \mathcal{F}^{(0)}) \leq C_t \left( \left( \frac{\Delta}{T} \right)^{\frac{3}{2}} |\ln T|^{2} + \frac{\Delta^2}{T} \right). \]  

(101)

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

\[ \mathbb{E} \left( \left( \Re \tilde{J}_{t,T}^{(1)} \right)^2 | \mathcal{F}^{(0)} \right) + \mathbb{E} \left( \left( \Im \tilde{J}_{t,T}^{(1)} \right)^2 | \mathcal{F}^{(0)} \right) \leq C_t \frac{\Delta}{\sqrt{T}}. \]  

(102)

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

\[ \mathbb{E} \left( \left( \Re \tilde{J}_{t,T}^{(1)} \right)^{\top} \left( \Im \tilde{J}_{t,T}^{(1)} \right) \left| \mathcal{F}^{(0)} \right. \right) = \mathbf{V}_{t,T}, \]  

(103)

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

(104)

where \( \mathbf{V}_{t,T} \) is a \( 2 \times 2 \) matrix with elements given as follows

\[ \mathbf{V}_{t,T}^{ij} = \frac{1}{T} \sum_{j=2}^{N} \chi_{ij} (u_T k_{j-1}) e^{-2k_{j-1} \epsilon_{j-1}^4 \Delta_j} O_T^2 (k_{j-1}) \Delta_j, \]  

with \( \chi_{ij}(x) = \cos^2(x) \) for \( i = j = 1 \), \( \chi_{ij}(x) = \sin^2(x) \) for \( i = j = 2 \) and \( \chi_{ij}(x) = \cos(x) \sin(x) \) for \( i = 1 \) and \( j = 2 \). Therefore,

\[ \mathbb{E} \left( \left( \begin{array}{ll} T u_T^2 \Re \tilde{J}_{t,T}^{(1)} - T u_T \Im \tilde{J}_{t,T}^{(1)} \\ T u_T \Re \tilde{J}_{t,T}^{(1)} + T u_T^2 \Im \tilde{J}_{t,T}^{(1)} \end{array} \right) \left( \begin{array}{ll} T u_T^2 \Re \tilde{J}_{t,T}^{(1)} - T u_T \Im \tilde{J}_{t,T}^{(1)} \\ T u_T \Re \tilde{J}_{t,T}^{(1)} + T u_T^2 \Im \tilde{J}_{t,T}^{(1)} \end{array} \right)^{\top} | \mathcal{F}^{(0)} \right) = \mathbf{C}_{t,T} (u_T), \]  

(105)

where for \( u \in \mathbb{R} \) we denote

\[ \mathbf{C}_{t,T}(u) = \sum_{j=2}^{N} \left( \begin{array}{ll} u^2 \cos (uk_{j-1}) - u \sin (uk_{j-1}) \\ u \cos (uk_{j-1}) + u^2 \sin (uk_{j-1}) \end{array} \right) \left( \begin{array}{ll} u^2 \cos (uk_{j-1}) - u \sin (uk_{j-1}) \\ u \cos (uk_{j-1}) + u^2 \sin (uk_{j-1}) \end{array} \right)^{\top} \times e^{-2k_{j-1} \epsilon_{j-1}^4 \Delta_j} O_T^2 (k_{j-1}) \Delta_j, \]  

36
Further, since \( f_{t,T}(u_T) \) is \( \mathcal{F}^{(0)} \)-adapted, if we set
\[
\chi_{t,T} = \frac{\Re f_{t,T}(u_T)}{|f_{t,T}(u_T)|^2} \left( T u_T^2 \Re F^{(1)}_{t,T} - T u_T \Im F^{(1)}_{t,T} \right) + \frac{\Im f_{t,T}(u_T)}{|f_{t,T}(u_T)|^2} \left( T u_T \Re F^{(1)}_{t,T} + T u_T^2 \Im F^{(1)}_{t,T} \right),
\]
we have for \( V_X^T = \mathbb{E} \left( \chi_{t,T}^2 | \mathcal{F}^{(0)} \right) : \)
\[
V_X^T = \frac{1}{|f_{t,T}(u_T)|^4} \left( \Re f_{t,T}(u_T) \ \Im f_{t,T}(u_T) \right) C_{t,T}(u_T) \left( \Re f_{t,T}(u_T) \ \Im f_{t,T}(u_T) \right)^\top. \tag{106}
\]
Using (64) and (102), we have
\[
V_X^T = \frac{\pi^4}{V_t^2} \frac{1}{|f_{t,T}(u_T)|^4} \left( \Re f_{t,T}(u_T) \ \Im f_{t,T}(u_T) \right) V_{t,T} \left( \Re f_{t,T}(u_T) \ \Im f_{t,T}(u_T) \right)^\top + O_p(\Delta) \tag{107}
\]
where we use the notation
\[
\psi_{t,T} = \frac{\overline{u} \sqrt{T}}{\sigma_t} + T \int_{\mathbb{R}} \left( \sin \left( \frac{\overline{u} x}{\sqrt{T} \sigma_t} \right) - \frac{\overline{u} x}{\sqrt{T} \sigma_t} \right) \nu_t(dx).
\]
Note that, given assumption A2-r, we have \( |\psi_{t,T}| \leq C_t \sqrt{T} \). If we denote with \( \bar{V}_X^T \) the expression in the second line of (107), then by using assumption A5 as well as the bounds of Lemmas 1 and 2, we have
\[
\bar{V}_X^T \geq \frac{C_t \Delta}{T^2} \sum_{j:|k_{j-1}| \leq \sqrt{T}} \cos^2(u_T k_{j-1} - \psi_{t,T}) O_T^2(k_{j-1}) \Delta_j
\]
\[
\geq \frac{C_t \Delta}{T^2} \sum_{j:|k_{j-1}| \leq \sqrt{T}} \cos^2(u_T k_{j-1} - \psi_{t,T}) \bar{O}_T^c(k_{j-1})^2 \Delta_j + O_p(\Delta)
\]
\[
\geq \frac{C_t \Delta}{T^2} \int_{-\sqrt{T}}^{\sqrt{T}} \cos^2(u_T k - \psi_{t,T}) \bar{O}_T^c(k)^2 dk + O_p \left( \frac{\Delta \vee \Delta^2}{T} \right).
\]
From here, using Lemma 2 and by a change of the variable of integration, we further have
\[
\bar{V}_X^T \geq \frac{C_t \Delta}{\sqrt{T}} \int_{-1}^{1} \cos^2 \left( \frac{\overline{u} k}{\sigma_t} - \psi_{t,T} \right) \left( f \left( \frac{k}{\sigma_t} \right) - \frac{|k|}{\sigma_t} \Phi \left( -\frac{|k|}{\sigma_t} \right) \right)^2 \nu_t(dx). \tag{108}
\]
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, \( \frac{\sqrt{T}}{\Delta} \bar{V}_X^T \) is bounded from below by an \( \mathcal{F}_t \)-adapted and positive-valued random variable. This shows that \( \frac{\Delta}{\sqrt{T}} \) is the sharp order of magnitude (and not an upper bound for it) of \( V_X^T \).
Using (104) and (106)-(108) as well as the $\mathcal{F}^{(0)}$-conditional independence of the observation errors from assumption A5, we can apply Theorem VIII.5.7 of Jacod and Shiryaev (2003) and get

\[
\frac{\chi_{t,T}}{\sqrt{V_T}} \xrightarrow{L^s} \mathcal{N}(0,1), \quad (109)
\]

where in the limit result above the notation $L - s$ 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 $\{\tilde{T}^{(j)}_{t,T}\}_{j=1,...,5}$ as well as the asymptotic negligibility of $\tilde{T}V_{t,T}(\tilde{\eta}_T) - V_t$ by Theorem 2, we have $\mathbb{R}(\tilde{f}_{t,T}(\tilde{\eta}_T))/\mathbb{R}(f_{t,T}(u_T)) \xrightarrow{P} 1$ and $\mathbb{R}(\tilde{f}_{t,T}(\tilde{\eta}_T))/\mathbb{R}(f_{t,T}(u_T)) \xrightarrow{P} 1$. Therefore, by an application of delta method from the convergence result in (109) and upon taking into account (107) and (108), we get

\[
\text{Avar}(\tilde{V}_{t,T}(u_T))^{-1/2} \frac{2}{T u_T} \left( \mathbb{R}(\ln(\tilde{f}_{t,T}(\tilde{\eta}_T))) - \mathbb{R}(\ln(f_{t,T}(\tilde{\eta}_T))) \right) \xrightarrow{L^s} \mathcal{N}(0,1), \quad (110)
\]

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

\[
\widetilde{\text{Avar}}(\tilde{V}_{t,T}(\tilde{\eta}_T))^{-1/2} \frac{2}{T \alpha_T} \left( \mathbb{R}(\ln(\tilde{f}_{t,T}(\tilde{\eta}_T))) - \mathbb{R}(\ln(f_{t,T}(\tilde{\eta}_T))) \right) \xrightarrow{L^s} \mathcal{N}(0,1). \quad (111)
\]

For this, we need to show that $\text{Avar}(\tilde{V}_{t,T}(u_T))/\text{Avar}(\tilde{V}_{t,T}(\tilde{\eta}_T)) \xrightarrow{P} 1$ and $\tilde{\eta}_T/u_T \xrightarrow{P} 1$. Given the asymptotic negligibility of $\tilde{T}V_{t,T}(\tilde{\eta}_T) - V_t$ and $\mathbb{R}(\tilde{f}_{t,T}(\tilde{\eta}_T))/\mathbb{R}(f_{t,T}(u_T)) \xrightarrow{P} 1$ and $\mathbb{R}(\tilde{f}_{t,T}(\tilde{\eta}_T))/\mathbb{R}(f_{t,T}(u_T)) \xrightarrow{P} 1$ (established above), we only need to show that $\tilde{C}_{t,T}(\tilde{\eta}_T) - C_{t,T}(u_T) = O_p \left( \frac{\alpha_T}{\sqrt{T}} \right)$. First, using the bounds in Lemma 1 as well as assumption A5 for the observation error, we have $\tilde{C}_{t,T}(\tilde{\eta}_T) - C_{t,T}(u_T) = O_p \left( \alpha_T \frac{\sqrt{T}}{\sqrt{T}} \right)$ where $\tilde{T}V_{t,T}(\tilde{\eta}_T) - V_t = O_p(\alpha_T)$. Second, for $\tilde{C}_{t,T}(u_T) - C_{t,T}(u_T)$ we can use assumption A5 for the observation error and apply Burkholder-Davis-Gundy inequality and conclude $\tilde{C}_{t,T}(u_T) - C_{t,T}(u_T) = O_p \left( \frac{n}{\sqrt{T}} \right)$. Thus, altogether, $\tilde{C}_{t,T}(\tilde{\eta}_T) - C_{t,T}(u_T) = O_p \left( \frac{\alpha_T}{\sqrt{T}} \right)$, and from here the result in (111) follows. Combining this result with (90) and taking into account the rate condition in (29), we have the convergence result (30) of the theorem.

### 7.6 Proof of Theorem 4

Under assumption A6 and our rate condition for $k_n$, the conditions of Theorem 13.3.3 of Jacod and Protter (2012) are satisfied. Therefore, we have

\[
\sqrt{k_n} (\tilde{V}_{h,T}^{k} - V_t) \xrightarrow{L^s} Z, \quad (112)
\]
with $Z$ being a standard normal random variable defined on an extension of the original probability space and independent from $F$. Let's denote $Z^n_1 = \frac{\tilde{V}_{t,T}(u_T) - V_t}{\sqrt{\text{Avar}(\tilde{V}_{t,T}(u_T))}}$ and $Z^n_2 = \frac{\sqrt{T} \sqrt{n}}{\sqrt{2V_t}}$, where $\text{Avar}(\tilde{V}_{t,T}(u))$ is defined as in the proof of Theorem 3 above and as in that proof we set $u_T = \frac{\gamma}{\sqrt{T}}$. From the proof of Theorem 3 we have $Z^n_1 = o_p(1)$ and hence to prove the result of Theorem 4 we need to establish the joint convergence of $(Z^n_1, Z^n_2)$.

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 $F$, 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 $f^{(1)}_{t,T}$, we can apply Theorem VIII.5.25 (using assumption A5 for the observation errors and the separability of $F^{(0)}$) and conclude that

$$
\mathbb{E}\left( g(Z^n_1) \mid F^{(0)} \right) \to \mathbb{E}(g(Z_1)), \quad \text{a.s.,}
$$

and therefore since $g$ and $h$ are bounded functions

$$
\mathbb{E}\left( \left( \mathbb{E}\left( g(Z^n_1) \mid F^{(0)} \right) - \mathbb{E}(g(Z_1)) \right) h(Z^n_2) \right) \to 0.
$$

Therefore, using (112), we have for every bounded random variable $Y$ on $F$

$$
\mathbb{E}(Yg(Z^n_1)h(Z^n_2)) \to \mathbb{E}(Yg(Z_1))\mathbb{E}(h(Z_2)),
$$

and this establishes $(Z^n_1, Z^n_2) \xrightarrow{L} (Z_1, Z_2)$.

**References**


Y. Kabanov and R. Lipster (Eds.), *From Stochastic Analysis to Mathematical Finance, Festschrift for Albert Shiryaev*. Springer.


