Spot Volatility Estimation for High-Frequency Data*

Jianqing Fan

Princeton University

Yazhen Wang

University of Connecticut

To be published in *Statistics and Its Interface*

September 4, 2008

Abstract

The availability of high-frequency intraday data allows us to accurately estimate stock volatility. This paper employs a bivariate diffusion to model the price and volatility of an asset and investigates kernel type estimators of spot volatility based on high-frequency return data. We establish both pointwise and global asymptotic distributions

Jianqing Fan is Frederick Moore'18 Professor of Finance, Department of Operation Research and Financial Engineering, Princeton University, Princeton, NJ 08544. Yazhen Wang is professor, Department of Statistics, University of Connecticut, Storrs, CT 06269. Fan's research was partially supported by the NSF grant DMS-0532370 and Wang's research was partially supported by the NSF grant DMS-0504323. The authors thank Per Mykland for helpful comments and suggestions.

for the estimators.

Key words: Asymptotic normality, CIR model, constant elasticity of diffusion, extreme distribution, kernel estimator, long memory, stock price.

1 Introduction

Volatilities of asset returns are pivotal for many issues in financial economics. For example, market participants need to estimate volatility for the purpose of hedging, option pricing, risk analysis and portfolio management. With advance of computer technology, data availability is becoming less and less a problem. Nowadays it is relatively easy to obtain high frequency financial data such as complete records of quotes or transaction prices for stocks. The high-frequency financial data provide an incredible experiment for understanding market microstructure and more generally for analyzing financial markets. In particular we expect to estimate volatilities better using high-frequency returns directly. The field of high-frequency finance has evolved rapidly. Current main interests of volatility estimation are on instantaneous volatility (or spot volatility) and integrated volatility over a period of time, say, a day. Estimation methods for univariate integrated volatility include realized volatility (RV) [Andersen et. al. (2003)], bi-power realized variation (BPRV)[Barndorff-Nielsen and Shephard (2006)], two-time scale realized volatility (TSRV)[Zhang et. al. (2005)], multiple-time scale realized volatility (MSRV) [Zhang (2006)], wavelet realized volatility (WRV)[Fan and Wang (2007)], kernel realized volatility (KRV)[Barndorff-Nielsen et. al. (2004)], and Fourier realized volatility (FRV) [Mancino and Sanfelici (2008)]. For the case of multiple assets, estimation approaches of multivariate integrated volatility consist of realized co-volatility for synchronized high-frequency data [Barndorff-Nielsen and Shephard (2004)] and realized co-volatility based on overlap intervals and previous ticks for non-synchronized high-frequency data [Hayashi and Kusuoka (2005) and Zhang (2005)]. Wang, Yao, Li and Zou (2007) has proposed a matrix factor model to achieve dimension reduction and facilitate the estimation of integrated co-volatility in very high dimensions for non-synchronized high-frequency data. For spot volatility estimation, Foster and Nelson (1996) first showed that spot volatility can be estimated from high-frequency data by rolling and block sampling filters. For a general class of price and volatility processes, under a number of stringent conditions, they established pointwise asymptotic normality for rolling regression estimators of the spot volatility and establish the efficiency of different weighting schemes. The conditions and results are in quite abstract sense. For given examples, the conditions are hard to verify and asymptotic normality is difficult to evaluate as well. Andreou and Ghysels (2002) further investigated theoretical properties of rolling-sample volatility estimator and check its finite sample performance with simulation and empirical studies. In this paper, we assume price and volatility to follow a bivariate diffusion process and investigate asymptotic behaviors of the kernel type estimators of spot volatility for high-frequency data. Under the general but verifiable conditions, we derive explicit expressions for their pointwise and global asymptotic distributions. We show that these conditions are met by diffusion based volatility models often used in literature.

The paper is organized as follows. Section 2 presents the main results. Section 3 illustrates the common models and verifies the conditions for these models. Section 3 features key technical propositions about strong approximation for the spot volatility estimator.

2 Estimation of spot volatility

Consider d assets and let $X_t = (X_{1t}, \cdots, X_{dt})^T$ be the vector of the log prices of d assets. Assume that X_t follows a continuous-time diffusion model,

$$dX_t = \mu_t \, dt + \sigma_t \, dW_t, \qquad t \in [0, T], \tag{1}$$

where T is a positive constant, W_t is a d-dimensional Brownian motion, μ_t is a drift, and σ_t is a d by d matrix. We define instantaneous or spot volatility as

$$\Gamma_t = \sigma_t \, \sigma_t^\dagger = (\gamma_t^{ij})_{1 \le i,j \le d}.$$

The quadratic variation of X_t has expression

$$[X, X]_t = \int_0^t \Gamma_s \, ds, \qquad t \in [0, T].$$

Suppose that we observe X_t at n discrete time points $t_i = i \, T/n$, $i = 1, \cdots, n$. Our goal is to estimate

$$\Gamma_t = \frac{d[X, X]_t}{dt} = \sigma_t \, \sigma_t^\dagger = (\gamma_t^{kj})_{1 \le k,j \le d}.$$

Suppose K(x) is a kernel with support on $[-1, 1]$. We define the kernel type estimator

$$\begin{aligned} \hat{\Gamma}_t &= \frac{1}{b} \sum_{t_i = t-b}^{t+b} K\left(\frac{t_i - t}{b}\right) (X_{t_i} - X_{t_{i-1}}) \, (X_{t_i} - X_{t_{i-1}})^\dagger \\ &= \frac{1}{b} \sum_{t_i = t-b}^{t+b} K\left(\frac{s - t}{b}\right) \Delta_\delta X_{t_i} \, (\Delta_\delta X_{t_i})^\dagger \\ &= \frac{1}{b} \int_{t-b}^{t+b} K\left(\frac{s - t}{b}\right) d\widehat{[X, X]}_s, \end{aligned} \tag{2}$$

where b is bandwidth, $\delta = T/n$, $\Delta_\delta X_t$ is the increment of X_t over $[t - \delta, t]$ defined by

$$\Delta_\delta X_t = X_t - X_{t-\delta},$$

and $\widehat{[X,X]}_t$ is the realized volatility given by

$$\widehat{[X,X]}_t = \sum_{t_i \le t} \Delta_\delta X_{t_i} (\Delta_\delta X_{t_i})^\dagger.$$

For example, if $K = 1$, then the estimator results in a rolling average

$$\hat{\Gamma}_t \quad = \quad \frac{1}{b} \sum_{|t_i - t| \le b} \Delta_\delta X_{t_i} \; (\Delta_\delta X_{t_i})^\dagger \, .$$

One side kernel K with support on $[-1, 0]$ yields an estimator that uses the immediate past data, and one side exponential kernel $K(x) = e^x \, 1(x \le 0)$ results in an exponential smoothing in the RiskMetric [Fan et. al. (2003)].

Below we will establish asymptotic theory for $\hat{\Gamma}$. First we list some technical conditions. Let $\| \cdot \|$ denote the Euclidean norm for vectors and maximum norm for matrices.

A1 $sup \, \{ \|\sigma_s - \sigma_t\|, s, t \in [0, T], |s - t| \le a \} = O_P(a^{1/2} \, | \log \, a|^{1/2}), \quad \sup_{0 \le t \le T} \|\sigma_t^2\| = O_p(1),$

and for $j = 1, \cdots, d,$

$$\textbf{A2} \qquad sup \left\{ \left\| \int_{t_{i-1}}^{t_i} \{ \sigma(s) - \sigma(t_{i-1}) \} \, dW_s^j \right\|^2, i = 1, \cdots, n \right\} = O_P(n^{-2+\eta}),$$

where $\eta > 0$ is an arbitrarily small number. The drift μ_t in (1) satisfies

$$\textbf{A3} \qquad \sup \{ \|\mu_t - \mu_s\|, |t - s| \le a \} = O_P(a^{1/2} \, | \log \, a|^{1/2}).$$

Bandwidth b and kernel K satisfy

A4 $b \sim n^{-1/2} / \log \, n, \; K(\cdot)$ is twice differentiable with support $[-1, 1]$ and $\int_{-1}^1 K(x) \, dx = 1.$

We will show that Assumptions A1-A2 are very general and satisfied for common volatility processes in Section 3. Assumption A3 is about the mean drift in price processes and is often met by price models. We may select kernel and bandwidth to meet Assumption A4. Now we state the two main theorems whose proofs rely on technical propositions given in Section 4.

Theorem 1 *Under Assumptions A1-A4, we have that*

$$\sqrt{n\,b}\left\{\hat{\Gamma}_t - \Gamma_t\right\} \to \sigma^2(t)\,Z,$$

where the convergence is in distribution, and Z is a random matrix whose elements are independent and have normal distributions with mean zero and variance $2\,\lambda(K)$ for diagonal elements and $\lambda(K)$ for off-diagonal elements, where

$$\lambda(K) = \int_{-1}^{1} K^2(x)\,dx.$$

Proof. It is a direct consequence of Propositions 2 and 3 in Section 4.

Remark 1. Theorem 1 provides pointwise asymptotic distribution for $\hat{\Gamma}$. The limiting distribution is normal with explicit covariance matrix. The convergence rate in Theorem 1 matches up with the orders of convergence in Mykland and Zhang (2008) in terms of bandwidth and sample size.

Theorem 2 *Suppose that Assumptions A1-A4 are satisfied and that $\sigma_t, t \in [0,T]$, is stationary. Let*

$$M_n = \sup_{0 \le t \le T} \sqrt{n\,b}\left\|\hat{\Gamma}_t - \Gamma_t\right\|.$$

Then

$$(2\,\log\,n)^{1/2}\left(\frac{M_n}{\sqrt{\lambda(K)}} - d_n\right) \to exp(-2\,e^{-x}),$$

where the convergence is in distribution, and $\lambda(K)$ *is defined in Theorem 1,*

$$d_n = (2\,\log\,n)^{1/2} + \frac{1}{(2\,\log\,n)^{1/2}}\,\{\log\lambda_1(K) - 0.5\,\log\,\pi - 0.5\,\log\,\log\,n\}\,,$$

$$\lambda_1(K) = \frac{K^2(-1)+K^2(1)}{2\,\lambda(K)},$$

if $\lambda_1(K) > 0$*, and otherwise*

$$d_n = (2\,\log\,n)^{1/2} + \frac{1}{(2\,\log\,n)^{1/2}}\,\{\log\lambda_2(K) - \log\,(2\,\pi)\}$$

$$\lambda_2(K) = \frac{1}{2\,\lambda(K)}\int [K'(x)]^2\,dx.$$

Proof. M_n has the same asymptotic distribution as

$$\sup_{0\le t\le T} |V_n(t)|,$$

where $V_n(t)$ is defined by (7) in Section 4. The representation for $V_n(t)$ given by Propositions 2 and 3 in Section 4 allows us to establish the asymptotic distribution for the maximum of $|V_n(t)|$ by an application of Theorem A1 in Bickel and Rosenblatt (1973, section 5).

Remark 2. Theorem 2 gives the global asymptotic distribution for $\hat{\Gamma}$. The extreme limiting distribution may be used to construct confidence band for Γ_t over whole interval $t \in [0, T]$.

3 Common volatility models

Common volatility processes in literature include geometric Ornstein-Uhlenbeck(OU) process, Nelson GARCH diffusion process (Nelson, 1990), the CIR diffusion process (Cox, Ingersoll and Ross, 1985), and long-memory volatility process (Comte and Renault, 1998). We show below that Assumptions A1-A2 are satisfied for these volatility processes as well as their superpositions. Below we will examine the examples for which Assumptions A1-A2 are met.

Example 1. Geometric OU model,

$$d\log \sigma^2(t) = -\lambda \log \sigma^2(t)\, dt + dW_v(t), \tag{3}$$

where W_v is a standard Brownian motion, λ is a parameter, and the initial value $\sigma^2(0)$ is finite and independent of W_v.

Example 2. Nelson GARCH diffusion model,

$$d\sigma^2(t) = -\lambda\, \{\sigma^2(t) - \xi\}\, dt + \omega\, \sigma^2(t)\, dW_v(\lambda\, t), \tag{4}$$

where W_v is a standard Brownian motion, (λ, ξ, ω) are parameters, and the initial value $\sigma^2(0)$ is finite and independent of W_v.

Example 3. The CIR model,

$$d\sigma^2(t) = -\lambda\, \{\sigma^2(t) - \xi\}\, dt + \omega\, \sigma(t)\, dW_v(\lambda\, t), \tag{5}$$

where W_v is a standard Brownian motion, (λ, ξ, ω) are parameters, and the initial value $\sigma^2(0)$ is finite and independent of W_v.

Example 4. The Long-memory model,

$$d\log \sigma_t = -\kappa\, \log \sigma_t^2\, dt + \gamma\, dW_{v,\alpha}(t), \tag{6}$$

where $W_{v,\alpha}$ is a fractional Brownian motion with memory index $\alpha \in (1/2, 1)$, (κ, γ) are parameters, and the initial value $\sigma^2(0)$ is finite and independent of $W_{v,\alpha}$.

We first check Assumption A1 for each example. For Example 1, (3) has an explicit solution

$$\log \sigma^2(t) = e^{-\lambda t} \log \sigma^2(0) + \int_0^t e^{-\lambda (t-s)}\, dW_v(s).$$

From the sample path property of Brownian motion W_v, we immediately show that

$$\sup_t |\log\, \sigma^2(t)| = O_P(1),$$

which implies

$$\sup_t \sigma^2(t) = O_P(1).$$

This is the second condition in Assumption A1. For the first condition of Assumption A1, note that

$$\begin{aligned}
\log\, \sigma^2(t) - \log\, \sigma^2(s) &= \left(e^{-\lambda t} - e^{-\lambda s}\right) \log\, \sigma^2(0) + \int_0^t e^{-\lambda (t-u)}\, dW_v(u) - \int_0^s e^{-\lambda (s-u)}\, dW_v(u) \\
&= \left(e^{-\lambda t} - e^{-\lambda s}\right) \left\{\log\, \sigma^2(0) + \int_0^s e^{\lambda u}\, dW_v(u)\right\} + e^{-\lambda t} \int_s^t e^{\lambda u}\, dW_v(u).
\end{aligned}$$

Since $e^{-\lambda t} - e^{-\lambda s} = O(t-s)$, the first term in above equation is $O_P(t-s)$, and due to the increment property of Brownian motion W_v, the second term is $O_P(|(t-s)\, \log |t-s||^{1/2})$. Since

$$\sigma(t) - \sigma(s) = \sigma(s) \left\{\exp\left(\log\, \sigma(t) - \log\, \sigma(s)\right) - 1\right\},$$

so the first condition in Assumption A1 is satisfied for Example 1.

The equation (4) in Example 2 has solution

$$\sigma_t^2 = \exp\left\{\beta_1\, t + \beta_2\, W_v(t) - \beta_2^2\, t/2\right\} \left\{\sigma_0^2 + \beta_0 \int_0^t \exp\left(-\beta_1\, s - \beta_2\, W_v(s) + \beta_2^2\, s/2\right)\, ds\right\}.$$

where $\beta_0 = \lambda\, \xi$, $\beta_1 = -\lambda$, $\beta_2 = \sqrt{\lambda}\, \omega$. Again the sample path property of W_v shows that

$$\sup_t \sigma_t^2 = O_P(1),$$

which is the second condition of Assumption A1. For the first condition note that

$$
\begin{aligned}
&|\sigma_t - \sigma_s| = \left|\exp\left\{\beta_1\,(t-s)/2 + \beta_2\,(W_v(t) - W_v(s))/2 - \beta_2^2\,(t-s)/4\right\} - 1\right| \\
&\exp\left\{\beta_1\,s/2 + \beta_2\,W_v(s)/2 - \beta_2^2\,s/4\right\}\,\left(\sigma_0^2 + \beta_0 \int_0^s \exp\left(-\beta_1\,u - \beta_2\,W_v(u) + \beta_2^2\,u/2\right)\,du\right)^{1/2} \\
&+ \exp\left\{\beta_1\,t/2 + \beta_2\,W_v(t)/2 - \beta_2^2\,t/4\right\}\,|\beta_0/2| \int_s^t \exp\left(-\beta_1\,u - \beta_2\,W_v(u) + \beta_2^2\,u/2\right)\,du \\
&\left(\sigma_0^2 + \beta_0 \int_0^s \exp\left(-\beta_1\,u - \beta_2\,W_v(u) + \beta_2^2\,u/2\right)\,du\right)^{-1/2}.
\end{aligned}
$$

Due to the property for the order of increments of Brownian motion, the first term in above equation is $O_P(|(t-s)\,\log|t-s||^{1/2})$, and the second term is $O_P(|t-s|)$. Thus, the first condition in Assumption A1 is satisfied.

For Example 3, (5) has no explicit solution. However, it is well known that $\sigma^2(t)$ is a Gamma process with

$$
\sup_t \sigma^2(t) = O_P(1), \qquad \sup_t \sigma^{-2}(t) = O_P(1).
$$

So the second condition of Assumption A1 is met. For the first condition we have that

$$
d\sigma(t) = 0.5\,\lambda\,\{-\sigma(t) + (\xi - 0.25\,\omega)/\sigma(t)\}\,dt + 0.5\,\omega\,dW_v(\lambda\,t)
$$

$$
\sigma(t) - \sigma(s) = 0.5\,\lambda \int_s^t \{-\sigma(u) + (\xi - 0.25\,\omega)/\sigma(u)\}\,du + 0.5\,\omega\,[W_v(\lambda\,t) - W_v(\lambda\,s)]
$$

The first term is $O_P(t-s)$ and the second term has order $|(t-s)\,\log|t-s||^{1/2}$ in probability. Thus, the first condition in Assumption A1 is met.

The equation (6) in Example 4 has solution

$$
\log\,\sigma^2(t) = e^{-\kappa\,t}\,\log\,\sigma^2(0) + \gamma \int_0^t e^{-\kappa\,(t-s)}\,dW_{v,\alpha}(s).
$$

The maximum of sample paths of $W_{v,\alpha}$ in a bounded interval is $O_P(1)$, thus the

$$\max_t |\log \sigma^2(t)| = O_P(1),$$

which implies the second condition of Assumption 1. For the first condition, we have

$$\begin{aligned}
\log \sigma^2(t) - \log \sigma^2(s) &= \left(e^{-\kappa\, t} - e^{-\kappa\, s}\right) \log \sigma^2(0) + \gamma \int_0^t e^{-\kappa\,(t-u)}\, dW_{v,\alpha}(u) \\
&\quad - \gamma \int_0^s e^{-\kappa\,(s-u)}\, dW_{v,\alpha}(u) \\
&= \left(e^{-\kappa\, t} - e^{-\kappa\, s}\right) \left\{\log \sigma^2(0) + \gamma \int_0^s e^{\kappa\, u}\, dW_{v,\alpha}(u)\right\} + e^{-\kappa\, t} \gamma \int_s^t e^{\kappa\, u}\, dW_{v,\alpha}(u).
\end{aligned}$$

Again the first term in above equation is $O_P(t-s)$. The second term is $O_P(|(t-s)$ $log|t-s||^{\alpha})$, due to the increment property of fractional Brownian motion $W_{v,\alpha}$.

Remark 3. If volatility processes satisfy Assumption 1, their superpositions also meet Assumption 1. This shows in particular that a two factor volatility model, which is a superposition of two geometric OU processes, satisfies Assumption 1.

Now we consider Assumption 2. We have the following general result for models without leverage effect, where no leverage effect means that Brownian motion W in (1) driving price processes and Brownian motion W_v (or fractional Brownian motion $W_{v,\alpha}$) in (3)-(6) are independent.

Proposition 1 *Suppose that there is independence between Brownian motion in (1) for price process and Brownian motion (or fractional Brownian motion) in (3)-(6) for volatility processes. If Assumption A1 is satisfied, then Assumption A2 is automatically met.*

Proof. Conditional on whole paths of σ_t^2,

$$\int_{t_{i-1}}^{t_i} \{\sigma(s) - \sigma(t_{i-1})\}\, dW_s$$

are independent Gaussian random variables with mean zero and covariance

$$\int_{t_{i-1}}^{t_i} \{\sigma(s) - \sigma(t_{i-1})\} \{\sigma(s) - \sigma(t_{i-1})\}^\dagger \, ds.$$

Hence, with probability tending to one, the maximum of

$$\left\| \int_{t_{i-1}}^{t_i} \{\sigma(s) - \sigma(t_{i-1})\} \, dW_s \right\|^2, \qquad i = 1, \cdots, n,$$

is bounded by

$$2 \log n \; sup \left\{ \int_{t_{i-1}}^{t_i} \|\sigma(s) - \sigma(t_{i-1})\|^2 \, ds, i = 1, \cdots, n \right\},$$

which, by Assumption 1, has order $n^{-2} \log^2 n$. This gives Assumption A2.

For price and volatility models with leverage effect, that is, W and W_v are dependent, Assumption A2 needs to check case by case. Below we illustrate the check of Assumption 2 for the geometric OU model. Note that we have

$$\begin{aligned} \log \sigma^2(s) - \log \sigma^2(t_{i-1}) &= \left(e^{-\lambda s} - e^{-\lambda t_{i-1}} \right) \left\{ \log \sigma^2(0) + \int_0^{t_{i-1}} e^{\lambda u} \, dW_v(u) \right\} \\ &\quad + e^{-\lambda s} \int_{t_{i-1}}^{s} e^{\lambda u} \, dW_v(u), \end{aligned}$$

$$\int_{t_{i-1}}^{t_i} \{\sigma(s) - \sigma(t_{i-1})\} \, dW_s = \sigma(t_{i-1}) \int_{t_{i-1}}^{t_i} \{e^{\log \sigma(s) - \log \sigma(t_{i-1})} - 1\} \, dW_s,$$

and thus

$$\begin{aligned} \int_{t_{i-1}}^{t_i} \{e^{\log \sigma(s) - \log \sigma(t_{i-1})} - 1\} \, dW_s &= \int_{t_{i-1}}^{t_i} \left\{ e^{\log \sigma(s) - \log \sigma(t_{i-1})} - \exp\left(e^{-\lambda s} \int_{t_{i-1}}^{s} e^{\lambda u} \, dW_v(u) \right) \right\} dW_s \\ &\quad + \int_{t_{i-1}}^{t_i} \left\{ \exp\left(e^{-\lambda s} \int_{t_{i-1}}^{s} e^{\lambda u} \, dW_v(u) \right) - 1 \right\} dW_s \\ &\equiv I_i + J_i. \end{aligned}$$

We need to show that for both I_i and J_i, their maximum over $i = 1, \cdots, n$ are of order $n^{-1+\eta/2} \log n$.

I_i is a stochastic integral over $[t_{i-1}, t_i]$, and its integrand is of order n^{-1}. As $\sum_{\ell=1}^{i} I_i$ is a discrete martingale, and its quadratic variation $[I, I]$ is of order of the sum of squares of the integrand of I_i, which has order n^{-2}. Hence,

$$P\left(\max_{1\le i\le n} |I_i| \ge 2\,M\right) \le 2\,P\left(\max_{1\le i\le n} |\sum_{\ell=1}^{i} I_\ell| \ge M\right) \le \frac{2\,M_1^2}{M^2} + 2\,P([I,I] > M_1^2) \to 0,$$

where the last equality is due to Nuglart inequality (Jacod and Shiryaev, 2002), and $M_1 = n^{-1}\log^{1/2} n$ and $M = n^{-1}\log n$. We derive that the maximum of I_i is of order $n^{-1}\log n$.

Also J_i is a stochastic integral over $[t_{i-1}, t_i]$, but its integrand is of order $n^{-1/2}\log^{1/2} n$. However, J_i are independent. Applying BDG inequality (Jacod and Shiryaev, 2002) to each J_i, we obtain

$$\begin{aligned} E\left(|J_i|^{2\,p}\right) &\le C\,E\left(\int_{t_{i-1}}^{t_i}\left\{\exp\left(e^{-\lambda\,s}\int_{t_{i-1}}^{s} e^{\lambda\,u}\,dW_v(u)\right) - 1\right\}^2 ds\right)^p \\ &\le C\,n^{-p}\int_{t_{i-1}}^{t_i} E\left\{\exp\left(e^{-\lambda\,s}\int_{t_{i-1}}^{s} e^{\lambda\,u}\,dW_v(u)\right) - 1\right\}^{2\,p} ds \le C\,n^{-2\,p}, \end{aligned}$$

where C is a generic constant and $p > 0$ is a constant and will be chosen later. With $M = n^{1/(2\,p)-1}\log n$ we obtain

$$\begin{aligned} &P\left(\max_{1\le i\le n} |J_i| \le M\right) = \prod_i P\left(|J_i| \le M\right) \ge \prod_i (1 - C\,n^{-2\,p}/M^{2\,p}) \\ &= (1 - C\,n^{-2\,p}/M^{2\,p})^n \sim 1 - C\,n^{1-2\,p}/M^{2\,p} = 1 - C\,\log^{-2\,p} n \to 1, \end{aligned}$$

For large enough $p \ge 1/\eta$ we conclude that the maximum of J_i is of order $n^{-1+\eta/2}\log n$.

4 Strong approximation for spot volatility estimator

Define

$$V_n(t) = \sqrt{n\,b}\left\{\hat{\Gamma}_t - \Gamma_t^*\right\}, \tag{7}$$

where

$$\Gamma_t^* = \frac{1}{b} \sum_{t_i=t-b}^{t+b} K\left(\frac{t_i - t}{b}\right) \int_{t_{i-1}}^{t_i} \Gamma_s \, ds. \tag{8}$$

We establish the following strong approximation result for V_n. Strong approximation constructed on some probability spaces are held for versions of V_n, σ, Γ on the new probability spaces, which have identical distributions as V_n, σ, Γ, respectively. For simplicity, we use the same notations to denote their versions on the constructed probability spaces.

Proposition 2 *Suppose that Assumptions A1-A4 are satisfied. Then there exist matrix processes* $B_n(t)$ *on some probability spaces such that* $B_n(t) = B_n(t)^\dagger = \{(1+1(k=j))^{1/2} \, B_n^{kj}(t)\}_{d\times d}$ *with* $B_n^{kj}(t) = B_n^{jk}(t)$ *being independent standard Brownian motions, and independent of* (μ_t, σ_t, W_t)*, and*

$$\begin{aligned} V_n(t) &= \sigma(t) \frac{1}{\sqrt{b}} \int_{t-b}^{t+b} K\left(\frac{s-t}{b}\right) dB_n(s) \, \sigma(t)^\dagger + O_P(n^{-1/4+\eta/2} \log n) \\ &= \sigma(t) \int_{t/b-1}^{t/b+1} K\left(u - \frac{t}{b}\right) d\tilde{B}_n(u) \, \sigma(t)^\dagger + O_P(n^{-1/4+\eta/2} \log n), \end{aligned}$$

where $\tilde{B}_n(\cdot) = b^{-1/2} B_n(b\,\cdot)$ *are the rescaled of* B_n*, and the error order is uniformly over* $t \in [0, T]$.

Proof. Note that

$$X_{t_i} - X_{t_{i-1}} = \int_{t_{i-1}}^{t_i} \mu_s \, ds + \int_{t_{i-1}}^{t_i} \sigma_s \, dW_s,$$

Assumptions A1-A3 implies that $\int_{t_{i-1}}^{t_i} \mu_s \, ds$ is dominated by $\int_{t_{i-1}}^{t_i} \sigma_s \, dW_s$, so the drift term μ_t in (1) has no effect on asymptotic results (such as limiting distributions and convergence orders) for the estimator $\hat{\Gamma}_t$. Therefore, for simplicity we set $\mu_t = 0$ in the rest of proofs.

The second equality results from change variable and rescaling property of Brownian motion. We prove the first equality only. Let $\delta = t_i - t_{i-1} = T/n$. Then

$$\hat{\Gamma}_t - \Gamma_t^* = \frac{1}{b}\sum_{t_i=t-b}^{t+b} K\left(\frac{t_i - t}{b}\right)\left\{\left(\int_{t_{i-1}}^{t_i}\sigma(s)\,dW_s\right)\left(\int_{t_{i-1}}^{t_i}\sigma(s)\,dW_s\right)^\dagger - \int_{t_{i-1}}^{t_i}\Gamma_s\,ds\right\}$$

$$= \frac{1}{b}\sum_{t_i=t-b}^{t+b} K\left(\frac{t_i - t}{b}\right)\left\{\sigma(t_{i-1})\,(W_{t_i} - W_{t_{i-1}})\,(W_{t_i} - W_{t_{i-1}})^\dagger\,\sigma(t_{i-1})^\dagger - \Gamma_{t_{i-1}}\,(t_i - t_{i-1})\right\}$$

$$+ \frac{1}{b}\sum_{t_i=t-b}^{t+b} K\left(\frac{t_i - t}{b}\right)\left\{\sigma(t_{i-1})\,(W_{t_i} - W_{t_{i-1}})\left(\int_{t_{i-1}}^{t_i}\{\sigma(s) - \sigma(t_{i-1})\}\,dW_s\right)^\dagger\right.$$

$$+ \int_{t_{i-1}}^{t_i}\{\sigma(s) - \sigma(t_{i-1})\}\,dW_s\{\sigma(t_{i-1})\,(W_{t_i} - W_{t_{i-1}})\}^\dagger - \int_{t_{i-1}}^{t_i}\{\Gamma(s) - \Gamma(t_{i-1})\}\,ds$$

$$\left.+ \int_{t_{i-1}}^{t_i}\{\sigma(s) - \sigma(t_{i-1})\}\,dW_s\left(\int_{t_{i-1}}^{t_i}\{\sigma(s) - \sigma(t_{i-1})\}\,dW_s\right)^\dagger\right\}$$

$$= H_1 + H_2 + H_3 + H_4. \tag{9}$$

Lemmas 2-4 below will derive the orders for H_2, H_3 and H_4. Simple algebra shows

$$H_1 = \frac{1}{b}\sum_{t_i=t-b}^{t+b} K\left(\frac{s - t}{b}\right)\sigma(t_{i-1})\left\{\left(W_{t_i} - W_{t_{i-1}}\right)\left(W_{t_i} - W_{t_{i-1}}\right)^\dagger - \delta\,I_d\right\}\sigma(t_{i-1})^\dagger$$

$$= \frac{\delta\sqrt{n}}{b}\sum_{t_i=t-b}^{t+b} K\left(\frac{t_i - t}{b}\right)\sigma(t_{i-1})\,n^{-1/2}\,U_i\,\sigma(t_{i-1})^\dagger, \tag{10}$$

where I_d denotes the $d \times d$ identity matrix,

$$U_i = \left(W_{t_i} - W_{t_{i-1}}\right)\left(W_{t_i} - W_{t_{i-1}}\right)^\dagger/\delta - I_d.$$

As matrix random variables U_i are i.i.d., $E(U_i) = 0$, and the entries of U_i are uncorrelated and have variance 2 at diagonal and 1 off diagonal, then

$$n^{-1/2}\sum_{j=1}^{[i\,t]} U_j$$

weakly converges to $B(t) = B(t)^\dagger = \{(1+1(k=j))^{1/2}\,B^{kj}(t)\}_{d\times d}$ with $B^{kj}(t) = B^{jk}(t)$ being independent standard Brownian motions, and independent of (μ_t, σ_t, W_t). By KMT strong approximation (Komlós, Major, and Tusnády, 1975, 1976), there exists $B_n(t) = B_n(t)^\dagger$ on some probability spaces with $B_n(t)$ being versions of B such that

$$Cov(B_n, W) = 0, \qquad \max_{1\le i\le n} \left| n^{-1/2} \sum_{j=1}^{i} U_j - B_n(t_i) \right| = O_P(n^{-1/2} \log n). \tag{11}$$

Then from (10) we get

$$\begin{aligned}
& H_1 - \frac{\delta\sqrt{n}}{b} \sum_{t_i=t-b}^{t+b} K\left(\frac{t_i - t}{b}\right) \sigma(t_{i-1})\, \Delta B_n(t_i)\, \sigma(t_{i-1})^\dagger \\
&= \frac{\delta\sqrt{n}}{b} \sum_{t_i=t-b}^{t+b} K\left(\frac{t_i - t}{b}\right) \sigma(t_{i-1}) \left\{ n^{-1/2} U_i - \Delta B_n(t_i) \right\} \sigma(t_{i-1})^\dagger \\
&= \frac{\delta\sqrt{n}}{b} \sum_{t_i=t-b+\delta}^{t+b-\delta} \left\{ K\left(\frac{t_i - t}{b}\right) \sigma(t_{i-1}) \left(n^{-1/2} \sum_{j=1}^{i} U_j - B_n(t_i) \right) \sigma(t_{i-1})^\dagger \right. \\
& \left. - K\left(\frac{t_{i+1} - t}{b}\right) \sigma(t_i) \left(n^{-1/2} \sum_{j=1}^{i} U_j - B_n(t_i) \right) \sigma(t_i)^\dagger \right\} \\
& \pm \frac{\delta\sqrt{n}}{b} K\left(\frac{b \pm \delta}{b}\right) \sigma(t \pm (b-\delta)) \left(n^{-1/2} \sum_{j=1}^{n\,(t\pm b)} U_j - B_n(t \pm b) \right) \sigma(t \pm (b-\delta))^\dagger \\
& \equiv G_1 + G_2.
\end{aligned} \tag{12}$$

Because of (11) and order of b in Assumption A4, G_2 is of order

$$\frac{\delta\sqrt{n}}{b}\, n^{-1/2} \log n = n^{-1/2} \log^2 n.$$

The term in the bracket of G_1 is equal to

$$\begin{aligned}
& K\left(\frac{t_i - t}{b}\right) \sigma(t_{i-1}) \left(n^{-1/2} \sum_{j=1}^{i} U_j - B_n(t_i) \right) \{\sigma(t_{i-1}) - \sigma(t_i)\}^\dagger \\
& + K\left(\frac{t_i - t}{b}\right) \{\sigma(t_{i-1}) - \sigma(t_i)\} \left(n^{-1/2} \sum_{j=1}^{i} U_j - B_n(t_i) \right) \sigma(t_i)^\dagger \\
& + \left\{ K\left(\frac{t_i - t}{b}\right) - K\left(\frac{t_{i+1} - t}{b}\right) \right\} \sigma(t_i) \left(n^{-1/2} \sum_{j=1}^{i} U_j - B_n(t_i) \right) \sigma(t_i)^\dagger.
\end{aligned} \tag{13}$$

By Assumption A1, $\sigma(t_{i-1}) - \sigma(t_i)$ is of order $n^{-1/2} \log n$, and Assumption A4 implies

$$K\left(\frac{t_i - t}{b}\right) - K\left(\frac{t_{i+1} - t}{b}\right)$$

is of order $n^{-1/2} \log n$. These two results together with (11) show that each of the three terms in (13) is of order $n^{-1} \log^2 n$. Substituting above orders for (13) into G_1 given by (12) and using the order of b in Assumption A4, we derive the order for G_1

$$\frac{\delta\sqrt{n}}{b}\, n\, b\, n^{-1} \log^2 n = n^{-1/2} \log^2 n.$$

Using above obtained order $n^{-1/2} \log^2 n$ for both G_1 and G_2 and from (12) we have

$$\begin{aligned}
H_1 &= \frac{\delta\sqrt{n}}{b} \sum_{t_i=t-b}^{t+b} K\left(\frac{t_i - t}{b}\right) \sigma(t_{i-1})\, \Delta B_n(t_i)\, \sigma(t_{i-1})^\dagger + O(n^{-1/2} \log^2 n) \\
&= \frac{\delta\sqrt{n}}{b} \sum_{t_i=t-b}^{t+b} K\left(\frac{t_i - t}{b}\right) \sigma(t_{i-1}) \left\{B_n(t_i) - B_n(t_{i-1})\right\} \sigma(t_{i-1})^\dagger + O(n^{-1/2} \log^2 n) \\
&= \frac{\delta\sqrt{n}}{b} \sum_{t_i=t-b}^{t+b} \int_{t_{i-1}}^{t_i} K\left(\frac{t_i - t}{b}\right) \sigma(t_{i-1})\, dB_n(s)\, \sigma(t_{i-1})^\dagger + O(n^{-1/2} \log^2 n) \\
&= \frac{\delta\sqrt{n}}{b} \int_{t-b}^{t+b} K\left(\frac{s - t}{b}\right) \sigma(s)\, dB_n(s)\, \sigma(s)^\dagger + O(n^{-1/2} \log^2 n),
\end{aligned}$$

where the last equality is due to Lemma 1 below. Collecting together above result for H_1 and the orders for H_2, H_3 and H_4 given by Lemmas 2-4 below, and using equation (9) we arrive at

$$V_n = \sqrt{n\,b}\,(\hat{\Gamma}_t - \Gamma_t^*) = \frac{1}{\sqrt{b}} \int_{t-b}^{t+b} K\left(\frac{s - t}{b}\right) \sigma(s)\, dB_n(s)\, \sigma(s)^\dagger + O_p(n^{-1/4+\eta/2} \log n). \quad (14)$$

Finally we complete the proof by using the order of b in Assumption A4 and showing that $\sigma(s)$ in the stochastic integral on the right hand side of (14) can replaced by $\sigma(t)$ with an error of order $n^{-1/4} \log n$. In deed, note that

$$\frac{1}{\sqrt{b}} \int_{t-b}^{t+b} K\left(\frac{s - t}{b}\right) \sigma(s)\, dB_n(s)\, \sigma(s)^\dagger = \frac{1}{\sqrt{b}} \int_{t-b}^{t+b} K\left(\frac{s - t}{b}\right) \sigma(t)\, dB_n(s)\, \sigma(s)^\dagger$$

$$+\frac{1}{\sqrt{b}}\int_{t-b}^{t+b} K\left(\frac{s-t}{b}\right)[\sigma(s)-\sigma(t)]\,dB_n(s)\,\sigma(s)^\dagger. \tag{15}$$

The second stochastic integral on the right hand side of (15) has is of order $n^{-1/4}\log n$, because its quadratic variation is equal to

$$\frac{1}{b}\int_{t-b}^{t+b} K^2\left(\frac{s-t}{b}\right)[\sigma(s)-\sigma(t)]^2\,ds\,\sigma^2(s)^\dagger = \int_{-1}^{1} K^2(u)[\sigma(t+u\,b)-\sigma(t)]^2\,du\,\sigma^2(t+u\,b)^\dagger$$
$$= O_P(n^{-1/2}\log^2 n),$$

where the second equality is from the fact that by Assumption A1, the maximum of $|\sigma(t+u\,b)-\sigma(t)|^2$ over $u\in[-1,1]$ is of order $n^{-1/2}\log^2 n$. Similarly, $\sigma(s)$ in the first stochastic integral on the right hand side of (15) can be replaced by $\sigma(t)$ with a resulting error of order $n^{-1/4}\log n$.

Lemma 1 *Suppose that Assumptions A1-A4 are satisfied. Then*

$$\sum_{t_i=t-b}^{t+b}\int_{t_{i-1}}^{t_i} K\left(\frac{t_i-t}{b}\right)\sigma(t_{i-1})\,dB_n(s)\,\sigma(t_{i-1})^\dagger$$
$$=\int_{t-b}^{t+b} K\left(\frac{s-t}{b}\right)\sigma(s)\,dB_n(s)\,\sigma(s)^\dagger + O_P(n^{-1/2}\log n).$$

Proof. Define

$$\begin{aligned}
D_i &= \int_{t_{i-1}}^{t_i} K\left(\frac{t_i-t}{b}\right)\sigma(t_{i-1})\,dB_n(s)\,\sigma(t_{i-1})^\dagger - \int_{t_{i-1}}^{t_i} K\left(\frac{s-t}{b}\right)\sigma(s)\,dB_n(s)\,\sigma(s)^\dagger \\
&= \int_{t_{i-1}}^{t_i}\left[K\left(\frac{t_i-t}{b}\right)-K\left(\frac{s-t}{b}\right)\right]\sigma(t_{i-1})\,dB_n(s)\,\sigma(t_{i-1})^\dagger \\
&\quad + \int_{t_{i-1}}^{t_i} K\left(\frac{s-t}{b}\right)[\sigma(t_{i-1})-\sigma(s)]\,dB_n(s)\,\sigma(t_{i-1})^\dagger \\
&\quad + \int_{t_{i-1}}^{t_i} K\left(\frac{s-t}{b}\right)\sigma(s)\,dB_n(s)\,[\sigma(t_{i-1})-\sigma(s)]^\dagger.
\end{aligned} \tag{16}$$

B_n are independent of volatility process σ, and the entries of B_n are independent Brownian motions, then conditional on σ, the entries of $D_1,\cdots,D_n$ are independent normal random variables with mean zero. We work on each entry of matrix $\sigma\, dB_n\, \sigma^\dagger$. Denote by $\|D_i\|$ the maximum over all entries of D_i. Since D_i defined by (16) is equal to a sum of three stochastic integral with respect to Brownian motion, which have explicit quadratic variations, we have that conditional on σ,

$$\begin{aligned} E\left[\|D_i\|^2|\sigma\right] \;&\leq\; C_1 \sup_{0\leq s\leq T} \|\sigma(s)\|^4 \int_{t_{i-1}}^{t_i} \left|K\left(\frac{t_i-t}{b}\right) - K\left(\frac{s-t}{b}\right)\right|^2 ds \\ &\quad + C_2 \int_{t_{i-1}}^{t_i} \left(\|\sigma(s)\|^2 + \|\sigma(t_{i-1})\|^2\right) \|\sigma(s)-\sigma(t_{i-1})\|^2\, ds, \qquad (17) \end{aligned}$$

where C_1 and C_2 are generic constants. By Assumption A1, we have that for $s \in [t_{i-1}, t_i]$, $\sigma(s) - \sigma(t_{i-1})$ is of order $n^{-1/2} \log n$ uniformly over $1 \leq i \leq n$, and Assumption A4 implies that

$$K\left(\frac{t_i-t}{b}\right) - K\left(\frac{s-t}{b}\right)$$

is of order $n^{-1/2} \log n$. Hence, the right hand side of (17) is of order $n^{-2} \log^2 n$, and so is the conditional variance of D_i. Since the entries of all D_i are independent normal random variables with conditional variances uniformly bounded by a quantity of order $n^{-2} \log^2 n$. Conditional on σ, with probability tending to one,

$$\max_{1\leq i\leq n} \|D_i\| \leq \sqrt{\log\left(n\, d^2\right)\; \max_{1\leq i\leq n} E\left[\|D_i\|^2|\sigma\right]}.$$

Hence we have

$$\max_{1\leq i\leq n} \|D_i\| = O_P(n^{-1} \log^{3/2} n),$$

and

$$\sum_{t_i=t-b}^{t+b} D_i = O_P(n\, b\, n^{-1} \log^{3/2} n) = O_P(n^{-1/2} \log^{1/2} n).$$

This completes the proof of Lemma 1.

Lemma 2 *Suppose that Assumptions A1-A4 are satisfied. Then*

$$\frac{1}{b}\sum_{t_i=t-b}^{t+b} K\left(\frac{t_i-t}{b}\right)\int_{t_{i-1}}^{t_i}(\Gamma_s-\Gamma_{t_{i-1}})\,ds = O_P(n^{-1/2}\,\log^{1/2}\,n).$$

Proof. Note that

$$\begin{aligned}
&\left|\frac{1}{b}\sum_{t_i=t-b}^{t+b} K\left(\frac{t_i-t}{b}\right)\int_{t_{i-1}}^{t_i}(\Gamma_s-\Gamma_{t_{i-1}})\,ds\right| \\
&\le \frac{1}{b}\sum_{t_i=t-b}^{t+b} K\left(\frac{t_i-t}{b}\right)\int_{t_{i-1}}^{t_i}|\Gamma_s-\Gamma_{t_{i-1}}|\,ds \\
&\le \frac{C}{b}\sum_{t_i=t-b}^{t+b}\int_{t_{i-1}}^{t_i}|\Gamma_s-\Gamma_{i-1}|\,ds \\
&\le C\,sup\{n\int_{t_{i-1}}^{t_i}|\Gamma_s-\Gamma_{t_{i-1}}|\,ds, i=1,\cdots,n\} \\
&\le C\,sup\{|\Gamma_s-\Gamma_{t_{i-1}}|, s\in[t_{i-1},t_i], i=1,\cdots,n\}, \qquad (18)
\end{aligned}$$

where C is generic constant depending on kernel only. Since

$$|\Gamma_t-\Gamma_s| = |\sigma_t\,\sigma_t^\dagger-\sigma_s\,\sigma_s^\dagger| \le |(\sigma_t-\sigma_s)\,\sigma_t^\dagger| + |\sigma_s\,(\sigma_t-\sigma_s)^\dagger| \le |\sigma_t-\sigma_s|\,|\sigma_t| + |\sigma_s|\,|\sigma_t-\sigma_s|,$$

from Assumption A1, we immediately have

$$sup\{|\Gamma_s-\Gamma_{t_{i-1}}|, s\in[t_{i-1},t_i], i=1,\cdots,n\} = O_P(n^{-1/2}\,\log^{1/2}\,n).$$

Combining it with (18) we prove the lemma.

Lemma 3 *Suppose that Assumptions A1-A4 are satisfied. Then*

$$\frac{1}{b}\sum_{t_i=t-b}^{t+b} K\left(\frac{t_i-t}{b}\right)\int_{t_{i-1}}^{t_i}\{\sigma(s)-\sigma(t_{i-1})\}\,dW_s\left(\int_{t_{i-1}}^{t_i}\{\sigma(s)-\sigma(t_{i-1})\}\,dW_s\right)^\dagger = O_P(n^{-1+\eta}\,\log^2\,n).$$

Proof. Note that

$$\frac{1}{b}\sum_{t_i=t-b}^{t+b} K\left(\frac{t_i-t}{b}\right)\left\|\int_{t_{i-1}}^{t_i}\{\sigma(s)-\sigma(t_{i-1})\}\,dW_s\left(\int_{t_{i-1}}^{t_i}\{\sigma(s)-\sigma(t_{i-1})\}\,dW_s\right)^{\dagger}\right\|$$
$$\leq \frac{C}{b}\sum_{t_i=t-b}^{t+b}\sum_{j=1}^{d}\left(\int_{t_{i-1}}^{t_i}\{\sigma(s)-\sigma(t_{i-1})\}\,dW_s^j\right)^2$$
$$\leq C\,sup\left\{n\sum_{j=1}^{d}\left(\int_{t_{i-1}}^{t_i}\{\sigma(s)-\sigma(t_{i-1})\}\,dW_s^j\right)^2, i=1,\cdots,n\right\}.$$

Now the lemma is a consequence of Assumption A2.

Lemma 4 *Suppose that Assumptions A1-A4 are satisfied. Then both of*

$$\frac{1}{b}\sum_{t_i=t-b}^{t+b} K\left(\frac{t_i-t}{b}\right)\sigma(t_{i-1})\left(W_{t_i}-W_{t_{i-1}}\right)\left(\int_{t_{i-1}}^{t_i}\{\sigma(s)-\sigma(t_{i-1})\}\,dW_s\right)^{\dagger},$$

and

$$\frac{1}{b}\sum_{t_i=t-b}^{t+b} K\left(\frac{t_i-t}{b}\right)\int_{t_{i-1}}^{t_i}\{\sigma(s)-\sigma(t_{i-1})\}\,dW_s\{\sigma(t_{i-1})\left(W_{t_i}-W_{t_{i-1}}\right)\}^{\dagger}$$

are equal to $O_P(n^{-1/2+\eta/2}\log^{3/2} n)$.

Proof. Because of simplicity we need to prove the first one only. For Brownian motion W, we have

$$sup\left\{\|W_{t_i}-W_{t_{i-1}}\|, i=1,\cdots,n\right\}=O_P(n^{-1/2}\log^{1/2} n),$$

and Assumption A2 implies

$$sup\left\{\left\|\int_{t_{i-1}}^{t_i}\{\sigma(s)-\sigma(t_{i-1})\}\,dW_s\right\|, i=1,\cdots,n\right\}=O_p((n^{-1+\eta/2}\log n).$$

Thus,

$$\sup\left\{\left\|(W_{t_i}-W_{t_{i-1}})\left(\int_{t_{i-1}}^{t_i}\{\sigma(s)-\sigma(t_{i-1})\}\,dW_s\right)^{\dagger}\right\|, i=1,\cdots,n\right\}=O_p(n^{-3/2+\eta/2}\log^{3/2} n),$$

from which we conclude

$$
\begin{aligned}
&\frac{1}{b}\left\|\sum_{t_i=t-b}^{t+b} K\left(\frac{t_i-t}{b}\right)\sigma(t_{i-1})\,(W_{t_i}-W_{t_{i-1}})\,\left(\int_{t_{i-1}}^{t_i}\{\sigma(s)-\sigma(t_{i-1})\}\,dW_s\right)^{\dagger}\right\| \\
&\le n\left\|(W_{t_i}-W_{t_{i-1}})\,\left(\int_{t_{i-1}}^{t_i}\{\sigma(s)-\sigma(t_{i-1})\}\,dW_s\right)^{\dagger}\right\| \\
&= O_p(n^{-1/2+\eta/2}\,\log^{3/2}\,n).
\end{aligned}
$$

Proposition 3 *Suppose that Assumptions A1-A4 are satisfied. Then*

$$\sqrt{n\,b}\,\{\Gamma_t^*-\Gamma_t\}=o_P(1),$$

where Γ_t^* *is defined in (8).*

Proof. Note that

$$|\Gamma_t-\Gamma_s|=|\sigma_t\,\sigma_t^{\dagger}-\sigma_s\,\sigma_s^{\dagger}|\le|(\sigma_t-\sigma_s)\,\sigma_t^{\dagger}|+|\sigma_s\,(\sigma_t-\sigma_s)^{\dagger}|\le|\sigma_t-\sigma_s|\,|\sigma_t|+|\sigma_s|\,|\sigma_t-\sigma_s|.$$

From Assumption A1 and A4, we immediately have

$$sup\{|\Gamma_s-\Gamma_t|,s\in[t-b,t+b]\}=O_P(\sqrt{b\,|\log\,b|}).$$

Hence with $\delta=t_i-t_{i-1}=T/n$, we obtain

$$\int_{t_{i-1}}^{t_i}\Gamma_s\,ds=\Gamma_t\,\delta+\int_{t_{i-1}}^{t_i}(\Gamma_s-\Gamma_t)\,ds=\delta\left[\Gamma_t+O_P(\sqrt{b\,|\log\,b|})\right]. \tag{19}$$

On the other hand, Assumption A4 and simple calculus show that

$$\frac{\delta}{b}\sum_{t_i=t-b}^{t+b}K\left(\frac{t_i-t}{b}\right)=\int_{-1}^{1}K(u)\,du+O(n^{-1/2}\,\log\,n)=1+O(n^{-1/2}\,\log\,n). \tag{20}$$

Plugging (19) and (20) into (8), we have

$$E\left(\hat{\Gamma}_t|\Gamma_t\right)=\left[1+O(n^{-1/2}\,\log\,n)\right]\,\left[\Gamma_t+O_P(\sqrt{b\,|\log\,b|})\right]=\Gamma_t+O_P(\sqrt{b\,|\log\,b|}).$$

Thus,

$$\sqrt{n\,b}\;(\Gamma_t^* - \Gamma_t) = \sqrt{n\,b}\,O_P(\sqrt{b\,|\log\, b|}) = O_P(\sqrt{n\,b^2\,|\log\, b|}) = o_P(1),$$

where the last equality is from the order of b in Assumption 4.

REFERENCES

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2003). Modeling and forecasting realized volatility. Econometrica 71, 579-625,

Andreou, E. and Ghysels, E. (2002). Rolling-sample volatility estimators: some new theoretical, simulation, and empirical results. J. Business & Economic Statistics 20, 363-376.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2004). Regular and modified kernel-based estimators of integrated variance: the case with independent noise. Manuscript.

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Econometric analysis of realized covariance: high frequency based covariance, regression and correlation in financial economics. Econometrica, 72, 885-925.

Barndorff-Nielsen, O. E. and Shephard, N. (2006). Econometrics of testing for jumps in financial econometrics using bipower variation. Journal of Financial Econometrics 4, 1-30.

Bickel, P. J. and Rosenblatt, M. (1973). On some global measures of the deviations of density function estimates. Ann. Statist. 1, 1071-1095.

Comte, F. and Renault, E. (1998). Long memory in continuous time stochastic volatility models. Mathematical Finance 8, 291-323.

Cox, J. C., Ingersoll, J. E. J., and Ross, S. A. (1985). A theory of the term structure of interest rates. Econometrica 53, 385-407.

Dacorogna, M. M., Geçay, R., Müller, U. A., Pictet, O. V. and Olsen, R. B. (2001). An Introduction to High Frequency Finance. Academic Press. New York.

Fan, J.(2005). A selective review of financial econometrics. Statistical Science, 20, 317-357.

Fan, J., Jiang, J., Zhang, C. and Zhou, Z. (2003). Time-dependent diffusion models for term structure dynamics and the stock price volatility. Statistica Sinica 13, 965-992.

Fan, J. and Wang, Y. (2007). Multi-scale jump and volatility analysis for high-frequency financial data. Journal of the American Statistical Association 102, 1349-1362.

Foster, D. and Nelson, D. (1996). Continuous record asymptotics for rolling sample estimators. Econometrica 64, 139-174.

Fouque, J. P., Papanicolaou, G. and Sircar, K. R. (2000). Derivatives in Financial Markets with Stochastic Volatility. Cambridge University Press.

Hayashi, T. and Yoshida, N. (2005). On covariance estimation of non-synchronously observed diffusion processes. Bernoulli **11**, 359-379.

Jacod, J. and Shiryaev, A. N. (2002). Limit Theorems for Stochastic Processes. Springer. Second edition.

Komlós, J. Major, P. and Tusnády, G. (1975). An approximation of partial sums of independent RV's and the sample DF. I. Z. Wahrsch. Verw. Gebiete **32**, 111-131.

Komlós, J. Major, P. and Tusnády, G. (1976). An approximation of partial sums of independent RV's and the sample DF. II. Z. Wahrsch. Verw. Gebiete **34**, 33-58.

Mancino, M. E. and Sanfelici, S. (2008). Robustness of Fourier Estimator of Integrated Volatility in the Presence of Microstructure Noise. Computational Statistics & Data Analysis archive, 52, 2966-2989.

Mykland, P.A., and Zhang, L. (2006). ANOVA for diffusions and Ito processes. Ann. Statist. 34 1931-1963.

Mykland, P.A., and Zhang, L. (2007). Inference for continuous semimartingales observed at high frequency: A general approach.

Mykland, P.A., and Zhang, L. (2008). Inference for volatility-type objects and implications for hedging. To appear in SII.

Nelson, D. B. (1990). ARCH models as diffusion approximations. Journal of Econometrics **45**, 7-38.

Wang, Y., Yao, Q., Li, P. and Zou, J. (2007). High dimensional volatility modeling and analysis for high frequency financial data. Technical report.

Zhao, Z., and Wu, W.B. (2008). Confidence bands in nonparametric time series regression. To appear in Annals of Statistics.

Zhang, L., Mykland, P. A. and Aït-Sahalia, Y. (2005). A tale of two time scales: determining integrated volatility with noisy high-frequency data. Journal of the American Statistical Association, 100, 1394-1411.

Zhang, L. (2006). Estimating covariation: Epps effect, microstructure noise. Technical report.

Zhang, L. (2006). Efficient estimation of stochastic volatility using noisy observations: a multi-scale approach. Bernoulli, 12, 1019-1043.

Zhou, B. (1996). High-frequency data and volatility in foreign-exchange rates. Journal of Business and Economic Statistics, 14, 45-52.