Evaluating the hedging error in price processes with jumps present*

BINGYI JING, XINBING KONG†, ZHI LIU AND BO ZHANG

In this draft, we consider a hedging strategy concerning only the continuous parts of two asset price processes which have jumps. Two consistent estimators of the hedging strategy, $\hat{\rho}$ and $\tilde{\rho}$, are presented in terms of realized bipower variation and threshold quadratic variation, respectively. Based on $\hat{\rho}$, estimators for operational risk, market risk (risk due to jumps) and total risk are investigated. It turns out that the variance of $\hat{\rho}$ enters into the bias of the operational risk estimator, whereas the variance is mainly due to jump influenced bipower estimation error. The convergence rate of the operational risk estimator (properly centralized) is $O_P((\overline{\Delta t})^{1/2})$. The convergence rate of the market risk is however $O_P((\overline{\Delta t})^{1/4})$. Based on $\tilde{\rho}$, the total risk is also studied, and it has the same convergence rate as that based on $\hat{\rho}$. Besides the interest in financial econometrics, it is also of significance in a statistical sense when we are interested in estimating the quadratic variation of the corresponding unhedgeable residual process.

1. INTRODUCTION

We consider the regression relationship between two stochastic processes Ξ_t and S_t,

$$(1.1) \qquad d\Xi_t = \rho_t dS_t + dZ_t, \quad 0 \le t \le T,$$

where Ξ_t and S_t are Lévy-type processes with finite activity jumps. This kind of process is widely used as an important

subclass of semimartingales. It also has a wide range of applications in the financial econometrics literature. Among them, [3, 8], and [13] are significant representatives. When Ξ_t and S_t are Itô processes, Mykland and Zhang [15] investigates intensively the structure of the estimation error of the quadratic variation of the residual process Z_t. One of their motivations comes from risk management in financial econometrics. Suppose that S_t and Ξ_t are the discounted values of two securities. At each time t, a financial institution is short one unit of Ξ and at the same time hold ρ_t units of security S. Z_t is then the gain/loss up to time t and $[Z,Z]_t$ defined by $\lim_{\max\{t_{i+1}-t_i\}\downarrow 0} \sum_i (Z_{t_{i+1}} - Z_{t_i})^2$ quantifies the unhedgeable part of the variation in asset Ξ in an incomplete market, when one adopts a feasible strategy ρ (regardless of the money market account units).

Under our setting of (1.1), the market is incomplete due to the presence of jumps. Then the unhedgeable risk includes two parts. One is due to jumps and the other due to possibly incomplete information of a continuous component. They are called market risk and operational risk (also known as process risk), respectively in the literature. Readers may find the motivations of risk-minimization in finance in [9] and [15] and references therein.

Since it is impossible to predict the jumps, we define the predictable hedging strategy as

$$(1.2) \qquad \rho_t = \frac{\langle \Xi, S \rangle'_t}{\langle S, S \rangle'_t},$$

where $\langle X, Y \rangle_t = \lim_{\max\{t_{i+1}-t_i\}\downarrow 0} \Delta X_{t_i}^c \Delta Y_{t_i}^c$ for two arbitrary semimartingales X and Y; X^c denotes the continuous part of X; $\Delta X_{t_i}^c$ defines the increment of X^c in interval $[t_i, t_{i+1}]$; and $\langle X, Y \rangle'_t$ denotes the derivative of $\langle X, Y \rangle_t$ with respect to time t. When Ξ and S are continuous semimartingales, ρ_t in (1.2) is the best choice of regression coefficient (or trading strategy) in the sense of minimizing $\langle Z, Z \rangle_T$ over all predictable processes, c.f. [15]. Under our circumstance, if Ξ and S are observed continuously,

$$(1.3) \qquad \underbrace{[\Xi, \Xi]_T}_{\text{total SS}} = \underbrace{\int_0^T \rho_u^2 d[S,S]_u}_{\text{Regression SS}} + \underbrace{[Z,Z]_T}_{\text{Residual SS}}$$
$$+ \underbrace{2 \sum_{0 \le u \le T} \rho_u J_u^S (J_u^\Xi - \rho_u J_u^S)}_{\text{Jump SC}},$$

where J_u^X is the jump size of an arbitrary process X at time u, i.e., $J_u^X = X(u) - X(u-)$, and

$$[X, X]_t = \lim_{\Delta_n \to 0} [X, X]_t^{(n)} =: \lim_{\Delta_n \to 0} \sum_{t_i \leq t} (X_{t_i} - X_{t_{i-1}})^2,$$

where $0 = t_0 < t_1 < t_2 < \cdots < t_n = T$ satisfying $\Delta_n :=$ $\max_i (t_{i+1} - t_i) \to 0$. $[X, X]_t^{(n)}$ is called realized volatility, c.f. [6]. In high frequency datasets, one can regard $t_0, t_1, \ldots, t_n$ as successive observation times.

The last term on the right-hand side of (1.3) stands for the sum of cross terms of regressed jumps and unregressed jumps (JSC). It is interesting to investigate the behavior and the influence of the JSC term in risk management. From JSC, to make risk more controllable, one seeks to choose a security S with jumps as small as possible. (1.3) differs from the standard ANOVA decomposition in regular regression because ρ_t concerns only the continuous part of two securities, Ξ and S. One easily argues that classical ANOVA decomposition exists if we set

$$\rho_t^* = \frac{[\Xi, S]_t'}{[S, S]_t'},$$

where $[X, X]'$ is interpreted as Radon-Nykodym derivative w.r.t time t. But unfortunately, ρ_t^* is not predictable right now and can neither be used as a trading strategy in finance nor an integrand in Itô calculus. This is also why instead of ρ_t^*, we present the hedging strategy as defined in (1.2). Actually, ρ_t is a continuous predictable modification of ρ_t^*.

The purpose of this paper is concerned with presenting a consistent estimator of the total risk $[Z, Z]_t$ and giving insights into the asymptotic behavior of the estimation error as more and more observations are available within a fixed time window. However, $[Z, Z]_t$ can be decomposed into two parts: the first part comes from the fluctuation of the continuous part of Z_t, i.e., $\langle Z, Z \rangle_t$, and the second part comes from the jumps, i.e., $\sum_{0 \leq u \leq t} (J_u^Z)^2$. The former is operational risk, whereas the latter market risk. Following natural questions arise:

Q_1 How to estimate the operational risk $\langle Z, Z \rangle_t$? What's the error distribution?

Q_2 How to estimate the market risk $\sum_{0 \leq T_j^z \leq t} (J_j^Z)^2$? What's the error distribution?

We tend to answer Q_1 and Q_2 respectively which are of individual interest and importance in finance.

To answer Q_1, we find that the estimator, $\widehat{\langle Z, Z \rangle}^{(n)}$ used in Section 2 in terms of bipower variation, has the first order approximation as

$$(1.4) \quad \widehat{\langle Z, Z \rangle}_t^{(n)} - \langle Z, Z \rangle_t \approx \text{bias}_t + (\langle Z, Z \rangle_t^{(n)} - \langle Z, Z \rangle_t).$$

The expression in (1.4) implies the separate effects of the two sources behind the asymptotics. The form of the bias

depends on the structure of the estimator. The variation component comes from the bipower estimation error common for arbitrary finite activity semimartingales. This is similar to that of Mykland and Zhang [15]. The main issue which differs from the ANOVA result in Mykland and Zhang [15] is to find a way to deal with jumps contained in discrete data and those jumps also influence the variance.

To answer Q_2, the convergence rate differs from that in Q_1. This demonstrates that the risk due to jumps in price process is unfortunately dominant in estimation error for the quadratic variation when we expect to require that $\widehat{\langle Z, Z \rangle}_t^{(n)} - \langle Z, Z \rangle_t$ and $\sum_{0 \leq T_j^z \leq t} \widehat{(J_j^Z)^2} - \sum_{0 \leq T_j^z \leq t} (J_j^Z)^2$ are asymptotically proportional to each other.

Putting the answers to Q_1 and Q_2 together, we get a full picture of the total estimation error for the quadratic variation of the residual process, $[Z, Z]_t$, by implementing the estimator:

$$\widehat{[Z, Z]}_t^{(n)} =: \widehat{\langle Z, Z \rangle}_t^{(n)} + \sum_{0 \leq T_j^z \leq t} \widehat{(J_j^Z)^2}.$$

After that, we also make use of the threshold variation method to present another consistent estimator of $[Z, Z]_t$, $\widetilde{[Z, Z]}_t^{(n)}$. The estimation error of $\widetilde{[Z, Z]}_t^{(n)}$ is studied as well. It has the same convergence rate as that of $\widehat{[Z, Z]}_t^{(n)}$. A comparison of these two estimators is then made.

This paper is organized as follows. In Section 2, the framework is set up and the estimators of ρ_t are introduced. Section 3 is devoted to answering the two questions we asked before and assessing the total risk estimators, $\widehat{[Z, Z]}_t^{(n)}$ and $\widetilde{[Z, Z]}_t^{(n)}$. Extensions to include stochastic volatilities are provided in Section 4. Proofs of the main results are deferred to Section 5.

2. THE SETUP

2.1 Jump diffusions

Throughout the paper, we restrict attention to the case of Ξ and S belonging to a subclass of Lévy type processes of the form:

$$\Xi_t = \int_0^t \mu_s^\Xi ds + \int_0^t \sigma_s^\Xi dB_s^\Xi + \sum_{0 \leq T_j^\Xi \leq t} J_j^\Xi,$$

$$S_t = \int_0^t \mu_s^S ds + \int_0^t \sigma_s^S dB_s^S + \sum_{0 \leq T_j^S \leq t} J_j^S,$$

where T_j^S and T_j^Ξ are respectively the jth jump times associated with the Poisson processes $\{N^\Xi(t); t \geq 0\}$ and $\{N^S(t); t \geq 0\}$. B_s^Ξ and B_s^S are two standard Brownian motions w.r.t the filtration $\mathscr{F}_s^\Xi$ and $\mathscr{F}_s^S$ respectively. Assume that $\langle B^\Xi, B^S \rangle' \in C[0, T]$ is bounded away from 0. Assume

that σ_t^Ξ and σ_t^S are deterministic positive functions away from zero belonging to $C^1[0,T]$ for simplicity. But when σ_t^Ξ and σ_t^S are stochastic, the main results can be extended naturally without much effort under some mild conditions; see Section 4. The drift coefficients μ^Ξ and μ^S are assumed to be bounded continuous processes (maybe random).

2.2 Bipower variation, threshold quadratic variation and estimators of ρ

If the processes Ξ_t and S_t can be continuously observed, the strategy can be fully figured out. However the discreteness of observations leads to the need of estimating ρ_t. Due to discrete observations, $\{\Xi_{t_i}, S_{t_i}; 0 \le t_i \le T\}$ is the only data set available to use. Because of jumps, a natural estimator of ρ_t is to use the realized bipower variation or covariation:

$$\langle X, X \rangle_t^* = \mu_1^{-2} \sum_{0 \le t_{j+1} \le t} |\Delta X_{t_j}||\Delta X_{t_{j+1}}|,$$

$$\langle X, Y \rangle_t^*$$
$$= \frac{1}{4}\mu_1^{-2}\Bigg[\sum_{0 \le t_{j+1} \le t} (|\Delta(X+Y)_{t_j}||\Delta(X+Y)_{t_{j+1}}|$$
$$- |\Delta(X-Y)_{t_j}||\Delta(X-Y)_{t_{j+1}}|)\Bigg],$$

where $\Delta X_{t_i} = X_{t_{i+1}} - X_{t_i}$ and $\mu_1 = (\frac{2}{\pi})^{1/2}$. Then a natural estimator of ρ_t is

$$(2.1) \qquad \hat{\rho}_t = \frac{\widehat{\langle \Xi, S \rangle'_t}}{\widehat{\langle S, S \rangle'_t}} = \frac{\langle \Xi, S \rangle_t^* - \langle \Xi, S \rangle_{t-h}^*}{\langle S, S \rangle_t^* - \langle S, S \rangle_{t-h}^*}.$$

For construction of bipower variation and its principle to get rid of jumps from observations, we refer to Barndorff and Neilson [7, 8]. One technical difficulty in dealing with bipower variation is the $|\cdot|$ operation and no Itô formula can be used for the artificial process $\langle X, X \rangle_t^*$ compared to $[X, X]_t^{(n)}$. $\hat{\rho}_t$ and $\hat{\rho}_s$ are asymptotically independent as $h \to 0$ when $t \ne s$ from the above definition.

Another way to separate the diffusion from jumps is to use realized threshold quadratic variation or covariation:

$$\langle X, X \rangle_t^\# = \sum_{t_{i+1} \le t} (X_{t_{i+1}} - X_{t_i})^2 I_{(|X_{t_{i+1}} - X_{t_i}| \le \alpha(\Delta t_i)^\varpi)},$$

$$\langle X, Y \rangle_t^\#$$
$$= \sum_{t_{i+1} \le t} (X_{t_{i+1}} - X_{t_i})(Y_{t_{i+1}} - Y_{t_i})$$
$$\times I_{(|X_{t_{i+1}} - X_{t_i}| \le \alpha_1(\Delta t_i)^\varpi, \ |Y_{t_{i+1}} - Y_{t_i}| \le \alpha_2(\Delta t_i)^\varpi)},$$

where $\varpi \in (0, 1/2), \alpha, \alpha_1, \alpha_2 > 0$. One then estimates ρ_t via

$$(2.2) \qquad \tilde{\rho}_t = \frac{\widetilde{\langle \Xi, S \rangle'_t}}{\widetilde{\langle S, S \rangle'_t}} = \frac{\langle \Xi, S \rangle_t^\# - \langle \Xi, S \rangle_{t-h}^\#}{\langle S, S \rangle_t^\# - \langle S, S \rangle_{t-h}^\#}.$$

The idea of this estimator actually comes from Lévy modulus characteristic. When $\max_i \Delta t_i$ is small enough, the increments of continuous part will be smaller than Δt_i^ϖ as long as $\varpi < 1/2$. A detailed description about threshold quadratic variation can be found in Mancini[13] and Jacod [11]. Empirical simulation shows that optimal choice of ϖ is 0.48 or 0.49, c.f. Jacod [10].

In (2.1) and (2.2), h plays the role of a smoothing bandwidth. The larger the h, the larger the bias and the less the variance. Hence h has to be chosen to tradeoff the bias and the variance. More discussion is given after Assumption 2.1.

From (1.1), we can estimate ΔZ_{t_i} through the difference

$$\widehat{\Delta Z}_{t_i} = \Delta \Xi_{t_i} - \hat{\rho}_{t_i}(\Delta S_{t_i}) \ \text{ or } \ \widetilde{\Delta Z}_{t_i} = \Delta \Xi_{t_i} - \tilde{\rho}_{t_i}(\Delta S_{t_i}).$$

Obviously at the sampling point t_i,

$$(2.3) \qquad \widehat{Z}_{t_i} = \sum_{t_{j+1} \le t_i} \widehat{\Delta Z}_{t_j} = \Xi_{t_i} - \int_0^{t_i} \hat{\rho}_u dS_u,$$

$$(2.4) \qquad \widetilde{Z}_{t_i} = \sum_{t_{j+1} \le t_i} \widetilde{\Delta Z}_{t_j} = \Xi_{t_i} - \int_0^{t_i} \tilde{\rho}_u dS_u,$$

where $\hat{\rho}_u = \hat{\rho}_{t_j}$ and $\tilde{\rho}_u = \tilde{\rho}_{t_j}$, if $t_j \le u < t_{j+1}$. (2.3) and (2.4) can be extended naturally to non-observation time as

$$(2.5) \qquad \widehat{Z}_t = \Xi_t - \int_0^t \hat{\rho}_u dS_u = Z_t + \int_0^t (\rho_u - \hat{\rho}_u)dS_u,$$

$$(2.6) \qquad \widetilde{Z}_t = \Xi_t - \int_0^t \tilde{\rho}_u dS_u = Z_t + \int_0^t (\rho_u - \tilde{\rho}_u)dS_u.$$

2.3 Variation of observation times

Consider a partition sequence $\varsigma_n = \{0 = t_1^{(n)} < \cdots < t_{n_k}^{(n)} = T\}$ which is assumed to be irregular and non-random in this paper. In the following, some regulations of variation of partition points are imposed to consider the asymptotical property of the estimation error.

Assumption 2.1. *Let* $\Delta t_i^{(n)} = t_{i+1}^{(n)} - t_i^{(n)}$, $\overline{\Delta t}^{(n)} = \frac{T}{n_k}$ *and* $\max_i \Delta t_i^{(n)} = \Delta^{(n)}$. *Suppose that:*

1. $\Delta^{(n)} \to 0$ *as* $n \to \infty$, *and* $\frac{\Delta^{(n)}}{\overline{\Delta t}^{(n)}} = O(1)$.
2. h *in (2.1) is chosen to variate with* n *satisfying* $\frac{\sqrt{\overline{\Delta t}^{(n)}}}{h} \to c^*$, *where* $0 < c^* < \infty$.
3. $H_{(n)}(t) = \frac{\sum_{t_{i+1}^{(n)} \le t} \Delta t_{i-1}^{(n)} \Delta t_i^{(n)}}{\overline{\Delta t}^{(n)}} \to H(t)$, *as* $n \to \infty$, *and* $H(t) \in C^1[0,T]$.
4. $\frac{H_{(n)}(t) - H_{(n)}(t-h)}{h} \to H'(t)$, *where the convergence is uniform in* t.
5. $G_{(n)}(t) =: \frac{1}{(\overline{\Delta t}^{(n)})^{1/4}}(\sum_{t_{i+1}^{(n)} \le t} \sqrt{\Delta t_{i-1}^{(n)} \Delta t_i^{(n)}} - t)$, $(G_{(n)}(t) - G_{(n)}(t-h))/h \to 0$, *uniformly in* t.

These assumptions are quite similar to those of Mykland and Zhang [15], except conditions 3 and 5. Actually, condition 3 is equivalent to (ii) of Assumption A in Mykland and Zhang [15], i.e.,

$$(2.7) \qquad \frac{\sum_{t_{i+1}^{(n)} \leq t}(\Delta t_i^{(n)})^2}{\overline{\Delta t}^{(n)}} \to H(t), \quad \text{as } n \to \infty.$$

In fact, condition 3 implies

$$\frac{\sum_{t_{i+1}^{(n)} \leq t}[(\Delta t_{i-1}^{(n)} + \Delta t_i^{(n)})^2 - (\Delta t_i^{(n)})^2 - (\Delta t_{i-1}^{(n)})^2]}{\overline{\Delta t}^{(n)}} \to H(t),$$

which is equivalent to (2.7).

Condition 5 is added technically to avoid too much bias due to bipower construction which will be seen clearly in proof of Lemma 6.1. Although we can use a smoothing bandwidth h, and there will naturally be a tradeoff between h and $\overline{\Delta t}^{(n)}$, in this paper, we will use $h = O((\overline{\Delta t}^{(n)})^{1/2})$.

One easily concludes Assumption 2.1 includes an equidistant partition as a particular example. For ease of notation, from now on, we suppress all the script (n), upper or lower, if there isn't confusion. So throughout the following sections, we interpret $\overline{\Delta t}$ as $\overline{\Delta t}^{(n)}$, Δ as $\Delta^{(n)}$, t_i as $t_i^{(n)}$, and so on.

3. MAIN RESULTS

3.1 Bipower variation, estimator and estimation error of the operational risk

In this subsection, the aim is to answer Q_1. It is of interest both in finance and statistics. In finance, we hope it can be used to control the unhedgeable risk while in statistics, it becomes the famous problem: estimating the integrated volatility (of the residual process). With respect to Q_1, a natural estimator is introduced via realized bipower variation,

$$(3.1) \qquad \widehat{\langle Z, Z \rangle}_t =: \mu_1^{-2} \sum_{t_{i+1} \leq t} |\widehat{\Delta Z_{t_i}}||\widehat{\Delta Z_{t_{i+1}}}|,$$

where $\mu_1 = \sqrt{\frac{2}{\pi}}$. Simple algebraic arrangement leads to the following decomposition.

$$\begin{aligned}
&\widehat{\langle Z, Z \rangle}_t - \langle Z, Z \rangle_t \\
&= \underbrace{(\widehat{\langle Z, Z \rangle}_t - \langle \widehat{Z}, \widehat{Z} \rangle_t)}_{\text{raw variance}} + \underbrace{(\langle \widehat{Z}, \widehat{Z} \rangle_t - \langle Z, Z \rangle_t)}_{\text{bias}} \\
&= \underbrace{(\langle \widehat{Z}, \widehat{Z} \rangle_t - \langle Z, Z \rangle_t)}_{\text{bias}} \\
&\quad + \underbrace{\left(\mu_1^{-2} \sum_{t_{i+1} \leq t} |\Delta Z_{t_i}||\Delta Z_{t_{i+1}}| - \langle Z, Z \rangle_t\right)}_{\text{vanriance due to distcretization}}
\end{aligned}$$

$$(3.2) \qquad \begin{aligned}
&+ \Bigg[\underbrace{(\widehat{\langle Z, Z \rangle}_t - \langle \widehat{Z}, \widehat{Z} \rangle_t)}_{} \\
&\quad - \underbrace{\left(\mu_1^{-2} \sum_{t_{i+1} \leq t} |\Delta Z_{t_i}||\Delta Z_{t_{i+1}}| - \langle Z, Z \rangle_t\right)}_{\text{correction}}\Bigg]
\end{aligned}$$

$$=: I + II + III.$$

It turns out that the correction term is $o(\sqrt{\overline{\Delta t}})$. So from (3.2), the bias is caused by the estimation method and the variance is mainly due to a kind of discretization (or jump influenced bipower estimation error). The main result of this section is

Theorem 3.1.

$$\begin{aligned}
&\frac{1}{\sqrt{\overline{\Delta t}}}(\widehat{\langle Z, Z \rangle}_t - \langle Z, Z \rangle_t) \\
&= D_t + \frac{1}{\sqrt{\overline{\Delta t}}\mu_1^2}\left(\sum_{t_{i+1} \leq t} |\Delta Z_{t_i}||\Delta Z_{t_{i+1}}| - \langle Z, Z \rangle_t\right) + o_p(1),
\end{aligned}$$

where $D_t = c'c^* \int_0^t (\sigma_u^Z)^2 H'(u)du$, and $c' = \frac{\pi^2}{4} + \pi - 3 \approx 2.6090$.

Remark 3.1. From the proof of Lemma 6.1 and Corollary 6.1, the variance of $\hat{\rho}_u - \rho_u$ is the main source of the bias D_t. Compared with the result of Mykland and Zhang [15], Theorem 3.1 shows another multiplier c' in D_t. This multiplier c' is caused by bipower construction which appears also in Barndorff and Neilson [7, 8]. When there are no jumps in price processes as studied in Mykland and Zhang [15], quadratic variation instead of bipower variation applies without appearance of c'.

Remark 3.2. The second term of the right-hand side converges to a mixture of a normal distribution and a compound Poisson distribution. When there are no jumps it is the same as obtained by Barndorff-Nielson et al. [8]. In the current paper jumps contained in price processes enter into the jumps of the limiting compound Poisson distribution as discussed in Proposition 6.2. This shows that to the second order the realized bipower variation cannot remove the influence of jumps as to the first order property (consistency) for jump diffusions.

3.2 Difference of quadratic variation and bipower variation, estimator and estimation error of the risk incurred from jumps

Since $\sum_{T_j^z \leq t}(J_j^Z)^2$ is the difference of the quadratic variation and the quadratic variation of the continuous part, i.e., $[Z, Z]_t - \langle Z, Z \rangle_t$, one easily has an estimator

$$(3.3) \qquad \sum_{t_i \leq t}(\widehat{\Delta Z_{t_i}})^2 - \widehat{\langle Z, Z \rangle}_t.$$

The main result with respect to this estimation error is

Theorem 3.2. *When n is large enough,*

$$\frac{1}{(\overline{\Delta t})^{1/4}}[\sum_{t_i}(\widehat{\Delta Z}_{t_i})^2 - \widehat{\langle Z,Z \rangle}_t - ([Z,Z]_t - \langle Z,Z \rangle_t)]$$

$$= 2\frac{1}{(\overline{\Delta t})^{1/4}}\sum_{T_j^S \leq t}(\rho_{T_j^S} - \widehat{\rho}_{T_j^S})J_j^S(J_j^\Xi - \rho_{T_j^S}J_j^S) + o_p(1),$$

and

$$\frac{1}{(\overline{\Delta t})^{1/4}}\sum_{T_j^S \leq t}(\rho_{T_j^S} - \widehat{\rho}_{T_j^S})J_j^S(J_j^\Xi - \rho_{T_j^S}J_j^S)$$

$$\xrightarrow{\mathscr{L}} \sum_{T_j^S \leq t}J_j^S(J_j^\Xi - \rho_{T_j^S}J_j^S)\frac{\sigma_{T_j}^Z}{\sigma_{T_j}^S}\sqrt{c'c^*H'(T_j)}\xi_j,$$

where ξ_j, for $j > 0$, are iid standard normal random variables which are independent of $\mathcal{F}^\Xi \bigvee \mathcal{F}^S$, on which (Ξ, S) are defined.

Remark 3.3. Theorem 3.2 shows that for market risk, the estimation error comes from the estimation error of the hedging strategy at jump times.

3.3 Assessment of estimators of the total risk $[Z, Z]_t$

3.3.1 $\widehat{[Z,Z]}_t$ as an estimator and its estimation error

By definition,

$$\widehat{[Z,Z]}_t = \sum_{t_{i+1} \leq t}(\widehat{\Delta Z}_{t_i})^2.$$

As shown by the following result with respect to estimation error, this is a consistent estimator.

Theorem 3.3. *As $n \to \infty$,*

$$\frac{1}{(\overline{\Delta t})^{1/4}}(\widehat{[Z,Z]}_t - [Z,Z]_t)$$

$$\xrightarrow{\mathscr{L}} 2\sum_{T_j^S \leq t}J_j^S(J_j^\Xi - \rho_{T_j^S}J_j^S)\frac{\sigma_{T_j}^Z}{\sigma_{T_j}^S}\sqrt{c'c^*H'(T_j)}\xi_j,$$

where $\xi_j's$ are a sequence of iid. standard normal variables which are independent of $\mathcal{F}^\Xi \bigvee \mathcal{F}^S$, on which (Ξ, S) are defined.

3.3.2 $\widetilde{[Z,Z]}_t$ as an estimator and its estimation error

In this subsection, we implement $\tilde{\rho}$ as the estimator of the hedging strategy, hence a natural estimator of the total risk $[Z, Z]_t$,

$$\widetilde{[Z,Z]}_t = \sum_{t_{i+1} \leq t}(\widetilde{\Delta Z}_{t_i}).$$

The main result in this subsection is

Theorem 3.4. *As $n \to \infty$,*

$$(3.4) \quad (\overline{\Delta t})^{-1/4}(\widetilde{[Z,Z]}_t - [Z,Z]_t)$$

$$\xrightarrow{\mathscr{L}} 2\sum_{T_j \leq t}J_j^S(J_j^\Xi - \rho_{T_j}J_j^S)\frac{\sigma_{T_j}^Z}{\sigma_{T_j}^S}\sqrt{c^*H'(T_j^S)}\xi_j.$$

where $\xi_j's$ are a sequence of iid. standard normal variables which are independent of $\mathcal{F}^\Xi \bigvee \mathcal{F}^S$, on which (Ξ, S) are defined.

Remark 3.4. This theorem also tells us that estimation error for the market risk dominates the estimation error for the operational risk. This is also an inspiration to explore the hedging risk due to a continuous component and jumps separately.

3.3.3 $\widehat{[Z,Z]}_t$ v.s. $\widetilde{[Z,Z]}_t$

As estimators of $[Z,Z]_t$, $\widehat{[Z,Z]}_t$ and $\widetilde{[Z,Z]}_t$ have the same convergence rate. When it comes to the finite sample performance, $\widetilde{[Z,Z]}_t$ is more efficient and the asymptotic relative efficiency of $\widehat{[Z,Z]}_t$ w.r.t. $\widetilde{[Z,Z]}_t$ is $\frac{1}{c'} \approx 0.38$. The difference comes from the difference in estimation scheme of ρ_t, see the comparisons on $\hat{\rho}_t$ and $\tilde{\rho}_t$. Simulations demonstrate that $\widetilde{[Z,Z]}_t$ is a bit better than $\widehat{[Z,Z]}_t$; see Section 5.

4. EXTENSIONS

The main results in Section 3 are obtained under the assumption that the volatilities of S and Ξ are deterministic functions. In this section, stochastic volatility models for S and Ξ are considered and Theorems 3.1–3.4 are extended. Obviously,

$$(\sigma_t^S)^2 = \langle S,S \rangle_t', \qquad (\sigma_t^\Xi)^2 = \langle \Xi,\Xi \rangle_t',$$

and

$$\int_0^t \sigma_u^S \sigma_u^\Xi d\langle B^\Xi, B^S \rangle_u = \langle \Xi,S \rangle_t.$$

First, we introduce some assumptions and notations used to deduce the results in this section. Assume that the spot volatility of Ξ and S are continuous Itô processes. Let

$$\langle \Xi,\Xi \rangle' = \int_0^t \mu_u^{V\Xi}du + \int_0^t \sigma_u^{V\Xi}dW_u^{V\Xi}$$

and

$$\langle S,S \rangle' = \int_0^t \mu_u^{VS}du + \int_0^t \sigma_u^{VS}dW_u^{VS},$$

where $W_s^{V\Xi}$ and W_s^{VS} are Winner processes w.r.t the filtration $\mathscr{F}_s^{V\Xi}$ and $\mathscr{F}_s^{VS}$ respectively.

Assumption 4.1. *1.1 $\langle \Xi,\Xi \rangle_u' \in \mathscr{F}_0^\Xi$ and $\langle S,S \rangle_u' \in \mathscr{F}_0^S$ for $0 \leq u \leq T$ are Itǒ processes, or*

1.2 $\langle \Xi, \Xi \rangle'$ and $\langle S, S \rangle'$ are independent of $\sigma\{\mathscr{F}_T^\Xi \vee \mathscr{F}_T^S\}$.

2. $\sigma^{V\Xi}$ and σ^{VS} are continuous processes bounded away from 0. $\mu^{V\Xi}$ and μ^{VS} are bounded continuous processes.

3. $\langle B^\Xi, B^S \rangle'_u \in C^1[0,T]$.

Conditions 1.1 and 1.2 trace back to [14] and [6] respectively. Now Theorem 3.1 is extended as follows.

Theorem 4.1. *Under Assumptions 2.1 and 4.1, Theorem 3.1 (including Remarks) holds with D_t replaced by*

$$c'c^* \int_0^t \langle Z, Z \rangle' H'(u) du + \frac{1}{3c^*} \int_0^t \langle \rho, \rho \rangle'_u d\langle S, S \rangle_u.$$

The bias changes when volatilities of Ξ and S are stochastic. This is because the effects of the continuous-estimation error of the instantaneous volatilities of Ξ and S in any interval like $[t-h, t]$ for $0 \le t \le T$ are not negligible compared to the normalization rate, as can be seen from comparison of (5.5) and Lemma 6.5.

Although volatilities become stochastic, it is still continuous. This should not affect much to the estimator of the jump risk. The following theorem shows what is expected and even the proof of it remains the same as that of Theorem 3.2 except replacing the r.h.s of corollary 6.1 by

$$\left(\frac{\sigma_u^Z}{\sigma_u^S} \sqrt{c'c^* H'(u)} + \sqrt{\frac{1}{3c^* \langle \rho, \rho \rangle'_u}} \right) \xi_u.$$

Theorem 4.2. *Under Assumptions 2.1 and 4.1, Theorem 3.2 holds with the limit replaced by*

$$\sum_{T_j^S \le t} J_j^S (J_j^\Xi - \rho_{T_j^S} J_j^S) \left(\frac{\sigma_{T_j}^Z}{\sigma_{T_j}^S} \sqrt{c'c^* H'(T_j^S)} + \sqrt{\frac{1}{3c^* \langle \rho, \rho \rangle'_{T_j}}} \right) \xi_j,$$

where ξ_j's are a sequence of iid. standard normal variables which are independent of $\mathcal{F}^\Xi \bigvee \mathcal{F}^S$, on which (Ξ, S) are defined.

Extensions to Theorems 3.3 and 3.4 are also listed below.

Theorem 4.3. *Under Assumptions 2.1 and 4.1, Theorem 3.3 holds with the limit replaced by*

$$2 \sum_{T_j^S \le t} J_j^S (J_j^\Xi - \rho_{T_j^S} J_j^S)$$
$$\times \left(\frac{\sigma_{T_j}^Z}{\sigma_{T_j}^S} \sqrt{c'c^* H'(T_j^S)} + \sqrt{\frac{1}{3c^* \langle \rho, \rho \rangle'_{T_j}}} \right) \xi_j,$$

where ξ_j's are a sequence of iid. standard normal variables which are independent of $\mathcal{F}^\Xi \bigvee \mathcal{F}^S$, on which (Ξ, S) are defined.

Theorem 4.4. *Under Assumptions 2.1 and 4.1, Theorem 3.4 holds with the limit replaced by*

$$2 \sum_{T_j^S \le t} J_j^S (J_j^\Xi - \rho_{T_j^S} J_j^S) \left(\frac{\sigma_{T_j}^Z}{\sigma_{T_j}^S} \sqrt{c^* H'(T_j^S)} + \sqrt{\frac{1}{3c^* \langle \rho, \rho \rangle'_{T_j}}} \right) \xi_j.$$

where ξ_j's are a sequence of iid. standard normal variables which are independent of $\mathcal{F}^\Xi \bigvee \mathcal{F}^S$, on which (Ξ, S) are defined.

5. SIMULATION STUDIES

In this section, we conduct some simulation studies to verify the consistency of the estimators of the hedge ratio ρ_t and the total variation of the hedging error in the model

$$d\Xi_t = \rho_t dS_t + dZ_t.$$

We generate n equally spaced discrete observations for both Ξ and S in the time interval $[0, 1]$ from two jump-diffusion processes,

$$(5.1) \qquad \Xi_t = \sigma^\Xi W_t^\Xi + J_t^\Xi, \qquad S_t = \sigma^S W_t^S + J_t^S,$$

where W_t is a standard Brownian motion and J_t is a Poisson process with intensity λ. For illustration, two cases are considered:

1. The two processes are independent, namely, W^Ξ and W^S are independent, and J_t^Ξ and J^S are independent. Hence the optimal hedge ratio is 0 according to our theoretical result.
2. The correlation between W^Ξ and W^S is ρ and two Poisson processes are independent.

Further, we choose the bandwidth $h = \sqrt{\Delta t}/2 = 1/(2\sqrt{n})$.

First we use a thresholding technique to remove the jumps effects, giving an estimator of hedge ratio ρ:

$$\tilde{\rho}_{t_i} = \frac{\widetilde{\langle \Xi, S \rangle}_{t_i+h} - \widetilde{\langle \Xi, S \rangle}_{t_i}}{\widetilde{\langle S, S \rangle}_{t_i+h} - \widetilde{\langle S, S \rangle}_{t_i}}$$
$$= \frac{\sum_{t_i \le t_j \le t_i+h} (\Delta_j \Xi)(\Delta_j S) \mathbf{1}_{\{\Delta_j \Xi < n^{-0.48}, \Delta_j S < n^{-0.48}\}}}{\sum_{t_i \le t_j \le t_i+h} (\Delta_j S)^2 \mathbf{1}_{\{\Delta_j S < n^{-0.48}\}}},$$

where $\Delta_i \Xi$ and $\Delta_i S$ are the differences of two consecutive observations of Ξ and S, respectively.

Secondly, we use $\Delta_i \widetilde{Z} = \Delta_i \Xi - \tilde{\rho}_{t_i} \Delta_i S$ to estimate the total variation of Z:

$$\widetilde{\langle Z, Z \rangle}_{t_i} = \sum_{t_j \le t_i} (\Delta_i \widetilde{Z})^2.$$

We also investigate the performance of Bipower estimators, $\hat{\rho}_{t_i}$ and $\widehat{\langle Z, Z \rangle}_{t_i}$ where we estimate the co-variation of Ξ and S by Bipower estimator:

$$\widehat{\langle \Xi, S \rangle}_t$$
$$= \frac{1}{4} \mu_1^{-2} \left[\sum_{0 \le t_{j+1} \le t} (|\Delta(\Xi + S)_{t_j}||\Delta(\Xi + S)_{t_{j+1}}| \right.$$
$$\left. - |\Delta(\Xi - S)_{t_j}||\Delta(\Xi - S)_{t_{j+1}}|) \right].$$

Figure 1. The path of $\widetilde{\langle Z, Z \rangle}_t$ (the thresholding estimator), $[Z, Z]_t$, and $\hat{\rho}_t, t \in (0, 1)$, with $n = 10,000$, $\rho = 0$ and $\lambda = 3$.

Figure 2. The path of $\widetilde{\langle Z, Z \rangle}_t$ (the thresholding estimator), $[Z, Z]_t$, and $\hat{\rho}_t, t \in (0, 1)$, with $n = 10,000$, $\rho = 0.5$ and $\lambda = 5$.

Table 1. The MADE values for two estimators, where we take $n = 10,000$

	Threshold estimator		Bipower estimator	
	$\widetilde{\langle Z, Z \rangle}$	$\tilde{\rho}$	$\widehat{\langle Z, Z \rangle}$	$\hat{\rho}$
$\rho = 0$	0.03	0.001	0.05	0.0015
$\rho = 0.5$	0.04	0.07	0.07	0.092

The thresholding estimators and the true values in interval [0,1] are displayed in Figures 1–2. Both figures show that the estimator is close to the true value, which is consistent with our theoretical results.

The mean absolute deviation errors (MADE) of two estimators are calculated as well, where, the MADE for $\widetilde{\langle Z, Z \rangle}_{t_i}$ is defined as

$$\text{MADE} = \frac{1}{m} \sum_{i=1}^{m} |\widetilde{\langle Z, Z \rangle}_{t_i} - \langle Z, Z \rangle_{t_i}|,$$

and MADE for $\tilde{\rho}$ is defined similarly, where $m = n - \lfloor \frac{h}{\Delta t} \rfloor$.

From Table 1, we can find both the MADE values are very small, this is in line with our theoretical result. We confirm the performance of our estimator with this numerical example.

6. PROOFS

6.1 Proof of Theorem 3.1

Since there are only finite jumps implied in the data and $h \to 0$ as $\Delta \to 0$, $\hat{\rho}_u = \frac{\widehat{\langle \Xi^c, S^c \rangle}_u}{\widehat{\langle S^c, S^c \rangle}_u}$ a.s. for h small enough. So without loss of generality, for calculating I, we suppress the script c and regard Ξ and S as continuous processes. From (2.3),

$$(6.1) \qquad Bias = \int_0^t (\hat{\rho}_u - \rho_u)^2 d\langle S, S \rangle_u.$$

To find the limit it is necessary to analyze the convergence rate of $\hat{\rho}_u$ to ρ_u. This is described by the following lemmas.

Lemma 6.1. $\hat{\rho}_u - \rho_u = O_P((\overline{\Delta t})^{1/4})$ uniformly for $0 \leq u \leq T$.

Proof. When $t_i \leq u \leq t_{i+1}$,

$$
\begin{aligned}
&\hat{\rho}_u - \rho_u \\
&= \hat{\rho}_{t_i} - \rho_{t_i} + \rho_{t_i} - \rho_u \\
&= \frac{1}{\langle S, S \rangle'_{t_i - h}} [\widehat{\langle \Xi, S \rangle}'_{t_i} - \langle \Xi, S \rangle'_{t_i} - \rho_{t_i - h}(\widehat{\langle S, S \rangle}'_{t_i}
\end{aligned}
$$

$$(6.2) \qquad - \langle S, S \rangle'_{t_i})] + O_P((\overline{\Delta t})^{1/2})$$

$$(6.3) \qquad = O_P((\overline{\Delta t})^{1/4}),$$

if we can show that $\widehat{\langle \Xi, S \rangle}'_{t_i} - \langle \Xi, S \rangle'_{t_i}$ and $\widehat{\langle S, S \rangle}'_{t_i} - \langle S, S \rangle'_{t_i}$ are $O_P((\overline{\Delta t})^{1/4})$, which will be clear later. From now on, we concentrate on $\widehat{\langle S, S \rangle}'_{t_i} - \langle S, S \rangle'_{t_i}$, and other terms can be deduced similarly.

$$
\begin{aligned}
&\widehat{\langle S, S \rangle}'_{t_i} - \langle S, S \rangle'_{t_i} \\
&= \mu_1^{-2} \frac{1}{h} \sum_{t_i - h \leq t_j \leq t_i} |\Delta S_{t_j}||\Delta S_{t_{j+1}}| \\
&\quad - \frac{1}{h} (\langle S, S \rangle_{t_i} - \langle S, S \rangle_{t_i - h}) \\
&\quad + \frac{1}{h} (\langle S, S \rangle_{t_i} - \langle S, S \rangle_{t_i - h}) - <S, S>'_{t_i}
\end{aligned}
$$

$$(6.4) \qquad =: A_i + B_i.$$

By Taylor's expansion,

$$(6.5) \qquad B_i = \frac{1}{h} \int_{t_i - h}^{t_i} [(\sigma_u^S)^2 - (\sigma_{t_i}^S)^2] du = O((\overline{\Delta t})^{1/2}),$$

uniformly in i. Now

$$
\begin{aligned}
&Bias \; of \; A_i \\
&= E \Big[\frac{\mu_1^{-2}}{h} \sum_{t_i - h \leq t_j \leq t_i} (\sigma_{t_j}^S)^2 |\Delta B_{t_{j-1}}||\Delta B_{t_j}|
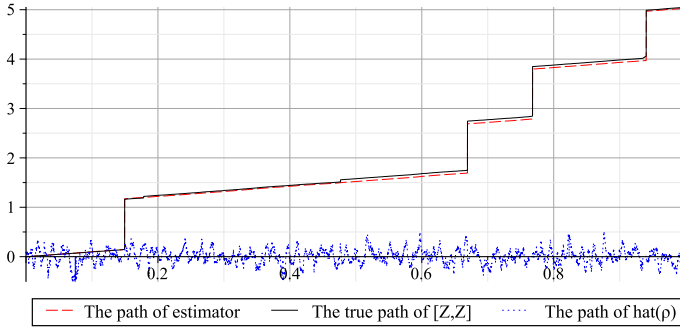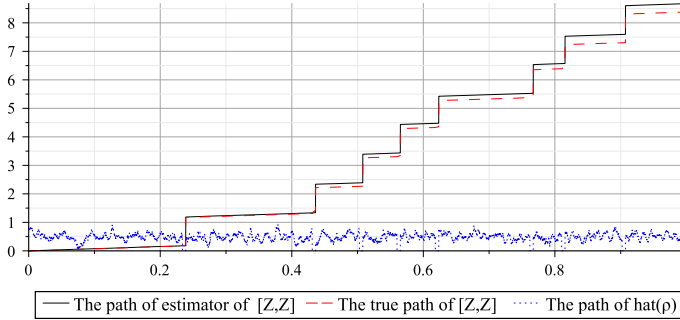\end{aligned}
$$

$$-\frac{1}{h}\sum_{t_i-h\leq t_j\leq t_i}(\sigma_{t_j}^S)^2\Delta t_j\Bigg] + O(\overline{\Delta t})$$

(6.6)
$$=\frac{1}{h}\sum_{t_i-h\leq t_j\leq t_i}(\sigma_{t_j}^S)^2(\sqrt{\Delta t_{j-1}\Delta t_j}-\Delta t_j)+O(\overline{\Delta t})$$

$$\leq \sup_{t_i}(\sigma_{t_i}^S)^2\frac{1}{h}\sum_{t_i-h\leq t_j\leq t_i}(\sqrt{\Delta t_{j-1}\Delta t_j}-\Delta t_j)$$
$$+O(\overline{\Delta t}).$$

If

(6.7)
$$\frac{1}{h}\sum_{t_i-h\leq t_j\leq t_i}(\sqrt{\Delta t_{j-1}\Delta t_j}-\Delta t_j)+o(\overline{\Delta t})^{1/4},$$

then $EA_i = o((\overline{\Delta t})^{1/4})$. Condition 5 of Assumption 2.1 ensures that (6.7) holds uniformly in i.

$$Var A_i =: E[\overline{A}_i]^2$$
$$=: E\Bigg[\mu_1^{-2}\frac{1}{h}\sum_{t_i-h\leq t_j\leq t_i}|\Delta S_{t_j}||\Delta S_{t_{j+1}}|$$
$$-E\Big[\mu_1^{-2}\frac{1}{h}\sum_{t_i-h\leq t_j\leq t_i}|\Delta S_{t_j}||\Delta S_{t_{j+1}}|\Big]\Bigg]^2$$
$$=\frac{1}{h^2}[(\mu_1^{-4}-1)+2(\mu_1^{-2}-1)]$$
$$\times\sum_{t_i-h\leq t_j\leq t_i}(\sigma^S)_{t_j}^4\Delta t_{j-1}\Delta t_j+O((\overline{\Delta t})^{3/4})$$

(6.8)
$$=:\frac{1}{h^2}c'\sum_j(\sigma^S)_{t_j}^4\Delta t_{j-1}\Delta t_j+O((\overline{\Delta t})^{3/4})$$

(6.9)
$$=O\Big(\frac{1}{h^2}\Big)\cdot\frac{h}{\overline{\Delta t}}\cdot(\overline{\Delta t})^2=O((\overline{\Delta t})^{1/2}),$$

uniformly in i. So, $\widehat{\langle S,S\rangle}'_{t_i}-\langle S,S\rangle'_{t_i}=O_P((\overline{\Delta t})^{1/4})$ by using Chebyshev's inequality. Similarly, $\widehat{\langle \Xi,S\rangle}'_{t_i}-\langle \Xi,S\rangle'_{t_i}=O_P((\overline{\Delta t})^{1/4})$. □

By (6.2), when $t_i\leq u\leq t_{i+1}$,

$$(\hat\rho_u-\rho_u)^2$$
$$=\frac{1}{(\langle S,S\rangle'_{t_i-h})^2}[(\widehat{\langle \Xi,S\rangle}'_{t_i}-\langle \Xi,S\rangle'_{t_i})^2$$
$$-2\rho_{t_i-h}(\widehat{\langle \Xi,S\rangle}'_{t_i}-\langle \Xi,S\rangle'_{t_i})(\widehat{\langle S,S\rangle}'_{t_i}-\langle S,S\rangle'_{t_i})$$
(6.10)
$$+(\rho_{t_i-h}(\widehat{\langle S,S\rangle}'_{t_i}-\langle S,S\rangle'_{t_i}))^2]+O_P((\overline{\Delta t})^{3/4})$$
$$=: IV_i+V_i+VI_i,$$

From (6.1), it suffices to show respectively the limit of

$$\frac{1}{\sqrt{\overline{\Delta t}}}\sum_{t_i}IV_i\langle S,S\rangle'_{t_i-h}\Delta t_i,\qquad \frac{1}{\sqrt{\overline{\Delta t}}}\sum_{t_i}V_i\langle S,S\rangle'_{t_i-h}\Delta t_i$$

and

$$\frac{1}{\sqrt{\overline{\Delta t}}}\sum_{t_i}VI_i\langle S,S\rangle'_{t_i-h}\Delta t_i.$$

It turns out that

Lemma 6.2.

$$\sum_{t_{i+1}\leq t}VI_i\langle S,S\rangle'_{t_i-h}\Delta t_i$$
$$=\sum_{t_{i+1}\leq t}\frac{\rho_{t_i-h}^2}{(\langle S,S\rangle'_{t_i-h})^2}E(\overline{A}_i)^2\langle S,S\rangle'_{t_i-h}\Delta t_i+O_P((\overline{\Delta t})^{3/4}).$$

Proof. Define $Y_i=(\overline{A}_i)^2-E[(\overline{A}_i)^2]$, since $A_i+B_i=\overline{A}_i+B_i+EA_i$, by (6.5) and the conclusion under (6.7) simple calculation results in

MSE
$$=E\Bigg(\sum_{t_{i+1}\leq t}\frac{\rho_{t_i-h}^2}{(\sigma_{t_i-h}^S)^2}Y_i\Delta t_i\Bigg)^2+o((\overline{\Delta t}))$$
$$=E\Bigg(\sum_{t_{i+1}\leq t}\sum_{t_{k+1}\leq t}\frac{\rho_{t_i-h}^2\rho_{t_k-h}^2}{(\sigma_{t_i-h}^S)^2(\sigma_{t_k-h}^S)^2}Y_iY_k\Delta t_i\Delta t_k\Bigg)+o(\overline{\Delta t})$$
$$\leq \frac{\sup_{0\leq u\leq T}(\sigma_u^\Xi)^4}{\inf_{0\leq u\leq T}(\sigma_u^S)^8}$$
$$\times\sum_{t_{i+1}\leq t}\sum_{t_{k+1}\leq t}\Delta t_i\Delta t_k I\{i\leq k;t_i>t_k-h\}E|Y_i||Y_k|$$
(6.11)
$$\leq C\Delta^2\sum_{t_{i+1}\leq t}\sum_{t_{k+1}\leq t}I\{i\leq k;t_i>t_k-h\}$$
$$\times\sqrt{E(Y_i^2)}\sqrt{E(Y_k)^2}$$
(6.12)
$$=O\Big((\overline{\Delta t})^2\cdot\frac{1}{\overline{\Delta t}}\frac{h}{\overline{\Delta t}}\cdot(\overline{\Delta t})^{1/2}(\overline{\Delta t})^{1/2}\Big)=O((\overline{\Delta t})^{3/2}).$$

From (6.11) to (6.12), we used the fact that $\max_i EY_i^2=\max_i Var(\overline{A}_i)^2=O((\overline{\Delta t}))$, which is indicated by (6.9) and can be calculated directly as well. □

Substitute (6.8) into Lemma 6.2, by interchange of summation,

$$\frac{1}{\sqrt{\overline{\Delta t}}}\sum_{t_{i+1}\leq t}\frac{\rho_{t_i-h}^2}{(\langle S,S\rangle'_{t_i-h})^2}E(\overline{A}_i)^2\langle S,S\rangle'_{t_i-h}\Delta t_i$$
$$=\frac{1}{\sqrt{\overline{\Delta t}}}\sum_{t_{i+1}\leq t}\frac{\rho_{t_i-h}^2}{(\langle S,S\rangle'_{t_i-h})^2}\Big(\frac{c'}{h^2}$$
$$\times\sum_{t_i-h\leq t_j\leq t_i}(\sigma_{t_j}^S)^4\Delta t_{j-1}\Delta t_j\Big)\langle S,S\rangle'_{t_i-h}\Delta t_i$$
$$=\frac{c'}{\sqrt{\overline{\Delta t}}}\frac{1}{h}\sum_{t_{j+2}\leq t}(\sigma_{t_j}^S)^4\Delta t_{j-1}\Delta t_j\Big(\frac{1}{h}$$
$$\times\sum_{t_{j+1}\leq t_i\leq t_j+h}\frac{\rho_{t_i-h}^2}{(\langle S,S\rangle'_{t_i-h})^2}\langle S,S\rangle'_{t_i-h}\Delta t_i\Big)$$

$$= \frac{c'}{\sqrt{\overline{\Delta t}}} \frac{1}{h} \sum_{t_{j+2} \le t} (\sigma_{t_j}^S)^4 \Delta t_{j-1} \Delta t_j \frac{\rho_{t_j}^2}{(\sigma_{t_j}^S)^2}$$

(6.13)
$$\xrightarrow{P} c'c^* \int_0^t \rho_u^2 (\sigma_u^S)^2 H'(u) du.$$

Similarly,
(6.14)
$$\frac{1}{\sqrt{\overline{\Delta t}}} \sum_{t_i} IV_i \langle S, S \rangle'_{t_i - h} \Delta t_i \xrightarrow{P} c'c^* \int_0^t (\sigma_u^\Xi)^2 H'(u) du.$$

and

(6.15)
$$\frac{1}{\sqrt{\overline{\Delta t}}} \sum_{t_i} V_i \langle S, S \rangle'_{t_i - h} \Delta t_i$$
$$\xrightarrow{P} -2c'c^* \int_0^t \rho_u \sigma_u^\Xi \sigma_u^S \langle B^\Xi, B^S \rangle_u H'(u) du.$$

Applying equations (6.1), (6.10), (6.13), (6.14) and (6.15), the following result can be obtained easily.

Propostion 6.1.

$$\frac{I}{\sqrt{\overline{\Delta t}}} \to c'c^* \int_0^t (\sigma_u^Z)^2 H'(u) du$$

in probability, where $c' \approx 2.6090$.

Corollary 6.1.

$$\frac{\rho_u - \hat{\rho}_u}{(\Delta t)^{1/4}} \to \frac{\sigma_u^Z}{\sigma_u^S} \sqrt{c'c^* H'(u)} \xi_u,$$

where ξ_u is a standard normal random variable which is independent of $\mathcal{F}^\Xi \bigvee \mathcal{F}^S$, on which (Ξ, S) are defined.

Next, we state a CLT result of the realized bipower variation related to Remark 3.1 after Theorem 3.1. Instead of applying the CLT result directly to the residual process Z, a general form is given in terms of an arbitrary jump diffusion process X, i.e.,

$$X_t = \mu^X t + \int_0^t \sigma_u^X dB_u + \sum_{j=1}^{N(t)} J_j^X,$$

where $N(t)$ is a Poisson process with intensity λ associated with the arriving points $0 < T_1 < T_2 < \cdots < \infty$ and σ_u^X and μ_u^X belongs to $C^1[0, T]$. This is of individual interest since, as far as we know, there are no such results on bipower variation of processes with jumps. Define

$$P_J = \{t_{T_j^- -}, t_{T_j^-}, t_{T_j^+}, t_{T_j^+ +}; j = 0, 1, 2, \ldots\},$$

where $t_{T_j^-} = \max\{t_i; t_i \le T_j\}$, $t_{T_j^+} = \min\{t_i; t_i \ge T_j\}$ are respectively the observation times before and after T_j, $t_{T_j^- -}$ is the observation time before $t_{T_j^-}$ and $t_{T_j^+ +}$ is the observation time after $t_{T_j^+}$.

Propostion 6.2.

$$\frac{1}{\sqrt{\overline{\Delta t}}} \left[\sum_{t_i} \frac{|\Delta X_{t_i}||\Delta X_{t_{i+1}}|}{\mu_1^2} - \langle X, X \rangle \right]$$
$$\to \Sigma_t \chi + \frac{1}{\mu_1^2} \sum_{T_j} \sigma_{T_j}^X (C_j + C_j') |\chi_j||J_j^X|$$

in distribution, where

$$\Sigma_t = \left(\frac{\pi^2}{4} + \pi - 3 \right) \int_0^t (\sigma_u^X)^2 H'(u) du$$
$$\approx 2.6090 \int_0^t (\sigma_u^X)^2 H'(u) du,$$
$$C_j = \lim_{\Delta \to 0} \frac{t_{T_j^-} - t_{T_j^- -}}{\Delta t}, \qquad C_j' = \lim_{\Delta \to 0} \frac{t_{T_j^+ +} - t_{T_j^+}}{\Delta t}$$

and χ and χ_j, $j = 1, \ldots, N(t)$, are independent standard normal random variables which are independent of $\mathcal{F}^\Xi \bigvee \mathcal{F}^S$, on which (Ξ, S) are defined.

Proof.

(6.16)
$$\sum_{t_i} \frac{|\Delta X_{t_i}||\Delta X_{t_{i+1}}|}{\mu_1^2} - \langle X, X \rangle_t$$
$$= \sum_{P_J} \frac{|\Delta X_{t_i}||\Delta X_{t_{i+1}}|}{\mu_1^2} + \sum_{P/P_J} \frac{|\Delta X_{t_i}||\Delta X_{t_{i+1}}|}{\mu_1^2} - \langle X, X \rangle_t.$$

Since in the subgrid P_J there are only finite terms almost surely, and X jumps only in the subgrid P_J,

$$\frac{1}{\sqrt{\overline{\Delta t}}} \left[\sum_{P/P_J} \frac{|\Delta X_{t_i}||\Delta X_{t_{i+1}}|}{\mu_1^2} - \langle X, X \rangle_t \right]$$

(6.17)
$$= \frac{1}{\sqrt{\overline{\Delta t}}} \left[\sum_P \frac{|\Delta X_{t_i}^c||\Delta X_{t_{i+1}}^c|}{\mu_1^2} - \langle X, X \rangle_t \right] + o_P(1).$$

Following the proof of Theorem 3 of Barndoff and Neilson [8], dominant term of (6.17) converges to a zero-mean normal distribution. The only gap we have to fill in is to calculate the limiting variance because in our case the observations are not equidistant.

(6.18) $Var\left(\sum_P \frac{|\Delta X_{t_i}^c||\Delta X_{t_{i+1}}^c|}{\mu_1^2} - \langle X, X \rangle_t \right)$

(6.19) $= \frac{1}{\mu_1^4} \sum_{t_i} [(1 - \mu_1^2)(\sigma_{t_i}^X)^2][(1 + 3\mu_1^2)(\sigma_{t_i}^X)^2](\Delta t_i)^2.$

So,

(6.20)
$$\frac{1}{\sqrt{\overline{\Delta t}}} Var\left(\sum_P \frac{|\Delta X_{t_i}^c||\Delta X_{t_{i+1}}^c|}{\mu_1^2} - \langle X, X \rangle_t \right)$$

$$\to \frac{1}{\mu_1^4} \int_0^t [(1-\mu_1^2)(\sigma_u^X)^2][(1+3\mu_1^2)(\sigma_u^X)^2]H'(u)du$$

(6.21)
$$= \left(\frac{\pi^2}{4} + \pi - 3\right) \int_0^t (\sigma_u^X)^2 H'(u)du.$$

On the other hand, by conditioning on the counting process $\{N(t); t \geq 0\}$ and using the continuity of volatility and Brownian process,

$$\frac{1}{\sqrt{\overline{\Delta t}}} \sum_{P_J} \frac{|\Delta X_{t_i}||\Delta X_{t_{i+1}}|}{\mu_1^2}$$

$$= \frac{1}{\sqrt{\overline{\Delta t}}} \left(\frac{1}{\mu_1^2} \sum_j (\sigma^X_{t_{T_j^-}} |B_{t_{T_j^-}} - B_{t_{T_j^-}}| \right.$$

$$\times |J_j^X + \sigma^X_{t_{T_j^-}} (B_{t_{T_j^+}} - B_{t_{T_j^-}})|$$

$$\left. + |J_j^X + \sigma^X_{t_{T_j^-}} (B_{t_{T_j^+}} - B_{t_{T_j^-}})||\sigma^X_{t_{T_j^+ +}} (B_{t_{T_j^+ +}} - B_{t_{T_j^+}})| \right)$$

$$+ o_P(1)$$

(6.22)
$$\xrightarrow{\mathscr{L}} \frac{1}{\mu_1^2} \sum_{T_j} \sigma^X_{T_j} (C_j + C_j')|\chi_j||J_j^X|. \qquad \Box$$

Now we prove Theorem 3.1.

Proof. By Propositions 6.1 and 6.2, and (3.2), it remains to prove that the correction term *III* is asymptotically negligible. Observe that from (2.5), up to the order $O_P((\overline{\Delta t})^{5/4})$,

$$\widehat{\Delta Z}_{t_i} = \Delta Z_{t_i} + \int_{t_i}^{t_{i+1}} (\rho_u - \hat{\rho}_u)dS_u^c$$

$$+ (\rho_{t_i} - \hat{\rho}_{t_i}) \sum_{T_j^S} J_j^S I\{t_i \leq T_j^S \leq t_{i+1}\})$$

and

$$\langle \hat{Z}, \hat{Z}\rangle_t = \langle Z, Z\rangle_t + \int_0^t (\rho_u - \hat{\rho}_u)^2 (\sigma_u^S)^2 du.$$

Observe also that, as Δ become small enough, $I\{t_i \leq T_j^S \leq t_{i+1}\}I\{t_{i+1} \leq T_k^S \leq t_{i+2}\} = 0$ a.s. for all j and k, by a similar trick in (6.22),

$$\sum_{0 \leq t_{i+1} \leq t} |\widehat{\Delta Z}_{t_i}||\widehat{\Delta Z}_{t_{i+1}}|$$

$$= \sum_{0 \leq t_{i+1} \leq t} \left| \Delta Z_{t_i} + \int_{t_i}^{t_{i+1}} (\rho_u - \hat{\rho}_u)dS_u^c \right|$$

$$\times \left| \Delta Z_{t_{i+1}} + \int_{t_{i+1}}^{t_{i+2}} (\rho_u - \hat{\rho}_u)dS_u^c \right| + o_P(\sqrt{\overline{\Delta t}}).$$

By definition of ρ, $\langle Z^c, B^S\rangle_u = 0$ for all $0 \leq u \leq T$. Therefore, regarding X_t in Proposition 6.2 as $Z_t + \int_0^t (\rho_u - \hat{\rho}_u)dS_u^c$

(depending on the partition sequence ς_n), following the proof of Proposition 6.2, it suffices to show that

$$\langle \widehat{X^c, X^c}\rangle_t - \langle \hat{Z}, \hat{Z}\rangle_t - (\langle \widehat{Z^c, Z^c}\rangle_t - \langle Z^c, Z^c\rangle_t) = o_P(\sqrt{\overline{\Delta t}}).$$

Simple calculation of the *MSE* of the left-hand side of the above equation is $o(\overline{\Delta t})$. Thus $III = o_P(\sqrt{\overline{\Delta t}})$. $\qquad \Box$

6.2 Proof of Theorem 3.2

Proof.

$$\frac{1}{(\overline{\Delta t})^{1/4}} \left[\sum_{t_i} (\widehat{\Delta Z}_{t_i})^2 - \frac{1}{\mu_1^2} \sum_{t_{i+1} \leq t} |\widehat{\Delta Z}_{t_i}||\widehat{\Delta Z}_{t_{i+1}}| \right.$$

$$- ([Z,Z]_t - \langle Z, Z\rangle_t) \Big]$$

$$= \frac{1}{(\overline{\Delta t})^{1/4}} \left[\left(\sum_{t_i} (\widehat{\Delta Z}_{t_i})^2 - [\hat{Z}, \hat{Z}] \right) + ([\hat{Z}, \hat{Z}] - [Z, Z]) \right.$$

(6.23)
$$\left. - \left(\frac{1}{\mu_1^2} \sum_{t_{i+1} \leq t} |\Delta \hat{Z}_{t_i}||\Delta \hat{Z}_{t_{i+1}}| - \langle Z, Z\rangle \right) \right].$$

Slight variation of Theorem 6.1 of Jacod and Protter [11] implies

(6.24)
$$\sum_{t_i} (\widehat{\Delta Z}_{t_i})^2 - [\hat{Z}, \hat{Z}] = O_P((\overline{\Delta t})^{1/2})).$$

Theorem 3.1 implies

(6.25)
$$\frac{1}{\mu_1^2} \sum_{t_{i+1} \leq t} |\widehat{\Delta Z}_{t_i}||\widehat{\Delta Z}_{t_{i+1}}| - \langle Z, Z\rangle = O_P((\overline{\Delta t})^{1/2})).$$

By (2.3), simple calculation yields

(6.26)
$$[\hat{Z}, \hat{Z}] - [Z, Z] = \int_0^t (\rho_u - \hat{\rho}_u)^2 d[S, S]_u$$

$$+ 2 \sum_{T_j^S \leq t} (\rho_{T_j^S} - \hat{\rho}_{T_j^S})J_j^S(J_j^\Xi - \rho_{T_j^S}J_j^S).$$

Since Lemma 6.1 holds and there are only finite jumps, by conditioning on the arriving times T_j^S, $j = 0, 1, \ldots$, and using Corollary 5.1,

$$\frac{1}{(\overline{\Delta t})^{1/4}} ([\hat{Z}, \hat{Z}] - [Z, Z])$$

$$= \frac{2}{(\overline{\Delta t})^{1/4}} \sum_{T_j^S \leq t} (\rho_{T_j^S} - \hat{\rho}_{T_j^S})J_j^S(J_j^\Xi - \rho_{T_j^S}J_j^S) + o_p(1)$$

(6.27)
$$\to 2 \sum_{T_j^S \leq t} J_j^S(J_j^\Xi - \rho_{T_j^S}J_j^S)\frac{(\sigma_{T_j}^Z)^2}{(\sigma_{T_j}^S)^2}\sqrt{c^*c^* H'(T_j)}\xi_j,$$

in distribution. $\qquad \Box$

6.3 Proof of Theorem 3.4

Lemma 6.3.

$$\sup_{t_i\in[0,T]}|\tilde{\rho}_{t_i}-\rho_{t_i}|=O_p(\overline{\Delta t}^{1/4}).$$

Proof. Note that

$$
\begin{aligned}
&\tilde{\rho}_{t_i}-\rho_{t_i}\\
&=\frac{1}{\widetilde{\langle S,S\rangle}'_{t_i}}\big(\langle\widetilde{\Xi,S}\rangle'_{t_i}-\langle\Xi,S\rangle'_{t_i}-\rho_{t_i}(\widetilde{\langle S,S\rangle}'_{t_i}-\langle S,S\rangle'_{t_i})\big)\\
&=\frac{1}{\langle S,S\rangle'_{t_i}}\big(\langle\widetilde{\Xi,S}\rangle'_{t_i}-\langle\Xi,S\rangle'_{t_i}-\rho_{t_i}(\widetilde{\langle S,S\rangle}'_{t_i}-\langle S,S\rangle'_{t_i})\big)\\
&\quad+o_p(\overline{\Delta t}^{1/4})
\end{aligned}
$$

if $\widetilde{\langle S,S\rangle}'_{t_i}-\langle S,S\rangle'_{t_i}=O_p(\overline{\Delta t}^{1/4})$, while this follows from the following estimate for realized threshold quadratic variation,

(6.28)
$$
\begin{aligned}
&\widetilde{\langle S,S\rangle}'_{t_i}-\langle S,S\rangle'_{t_i}\frac{1}{h}\\
&=\sum_{t_i-h\le t_j<t_{j+1}\le t_i}(\Delta S_j)^2\mathbf{1}_{\{(\Delta S_j)^2\le\alpha(\Delta t_j)^\varpi\}}-\langle S,S\rangle'_{t_i}
\end{aligned}
$$
(6.29)
$$=:\tilde{A}_i+\tilde{B}_i+O_p(\sqrt{\overline{\Delta t}})$$

$\tilde{B}_i=O_p(\sqrt{\overline{\Delta t}}\,)$ by Taylor expansion.

$$
\begin{aligned}
\tilde{A}_i&=\frac{1}{h}\sum_{t_i-h\le t_j<t_{j+1}\le t_i}((\Delta S_j)^2\mathbf{1}_{\{(\Delta S_j)^2\le\alpha(\Delta t_j)^\varpi\}}-\Delta\langle S,S\rangle_{t_j})\\
&=:\sum_{t_i-h\le t_j<t_{j+1}\le t_i}\eta_j.
\end{aligned}
$$

A simple calculation shows that,

$$
\begin{aligned}
&E[\eta_j^2]\\
&=\frac{(\Delta t_j)^2}{h^2}E\Big(\Big(\frac{\Delta S_j}{\sqrt{\Delta t_j}}\Big)^2\mathbf{1}_{\{(\frac{\Delta S_j}{\sqrt{\Delta t_j}})^2\le\alpha(\Delta t_j)^{\varpi-1/2}\}}\\
&\qquad-\frac{\Delta\langle S,S\rangle_{t_j}}{\Delta t_j}\Big)^2\\
&=\frac{(\Delta t_j)^2}{h^2}E\Big((\sigma_j^S)^2\mathscr{N}^2\mathbf{1}_{\{(\sigma_j^S)^2\mathscr{N}^2\le\alpha(\Delta t_j)^{\varpi-1/2}\}}\\
&\qquad-\frac{\Delta\langle S,S\rangle_{t_j}}{\Delta t_j}\Big)^2\\
&=O_p(\overline{\Delta t_j}).
\end{aligned}
$$

Then $\sum_{t_i-h\le t_j<t_{j+1}\le t_i}E\eta_j^2=O_p(\sqrt{\overline{\Delta t}}\,)$, hence by Chebyshev inequality

$$\tilde{A}_i=O_p((\overline{\Delta t})^{1/4}).$$

This concludes the proof. □

Lemma 6.4.

$$\frac{1}{(\overline{\Delta t})^{1/4}}(\tilde{\rho}_t-\rho_t)\xrightarrow{\mathscr{L}}\mathscr{N}(0,V_{\tilde{\rho}_t-\rho_t}),$$

where $\sqrt{V_{\tilde{\rho}_t-\rho_t}}=\frac{\sigma_t^Z}{\sigma_t^S}\sqrt{c^*H'(t)}$.

Proof. Select function $\varphi(t)$ satisfies

1. $\varphi(t)\in C^\infty$;
2. $\mathbf{1}_{(|\mathbf{x}|\le 1)}\le\varphi(\mathbf{x})\le\mathbf{1}_{(|\mathbf{x}|\le 2)}$;
3. $\lim_{y\to 0}\varphi(\frac{x}{y})=\varphi(\frac{x}{0})=:1$,

Then we have

$$
\begin{aligned}
&\sum_{t_i-h\le t_j<t_{j+1}\le t_i}(\Delta S_j)^2\varphi\Big(\frac{\Delta S_i}{\alpha\Delta t_j^\varpi}\Big)\\
&\le\sum_{t_i-h\le t_j<t_{j+1}\le t_i}(\Delta S_j)^2\mathbf{1}_{\{(\Delta S_j)^2\le\alpha(\Delta t_j)^\varpi\}}\\
&\le\sum_{t_i-h\le t_j<t_{j+1}\le t_i}(\Delta S_j)^2\varphi\Big(\frac{2\Delta S_i}{\alpha\Delta t_j^\varpi}\Big),
\end{aligned}
$$

hence also

$$
\begin{aligned}
&\Big|\sum_{t_i-h\le t_j<t_{j+1}\le t_i}((\Delta S_j)^2\mathbf{1}_{\{(\Delta S_j)^2\le\alpha(\Delta t_j)^\varpi\}}-\Delta\langle S,S\rangle_{t_j})\Big|\\
&\le\Big|\sum_{t_i-h\le t_j<t_{j+1}\le t_i}\Big((\Delta S_j)^2\varphi\Big(\frac{\Delta S_i}{\alpha\Delta t_j^\varpi}\Big)-\Delta\langle S,S\rangle_{t_j}\Big)\Big|\\
&\quad+\Big|\sum_{t_i-h\le t_j<t_{j+1}\le t_i}\Big((\Delta S_j)^2\varphi\Big(\frac{2\Delta S_i}{\alpha\Delta t_j^\varpi}\Big)-\Delta\langle S,S\rangle_{t_j}\Big)\Big|.
\end{aligned}
$$

It is clearly enough to prove the result for the first term and second term, and also that

$$\frac{1}{h}\sum_{t_i-h\le t_j<t_{j+1}\le t_i}\Big((\Delta S_j)^2\varphi\Big(\frac{\Delta S_i}{\alpha\Delta t_j^\varpi}\Big)-(\Delta S_j)^2\varphi\Big(\frac{2\Delta S_i}{\alpha\Delta t_j^\varpi}\Big)\Big)\xrightarrow{p}0.$$

This is a direct consequent of Theorem 2.11 of Jacod [10]. Therefore, there are no difference between our estimator and the one in Zhang [16], which completes the proof. □

Next, we turn to prove Theorem 3.4.

Proof. Simple algebraic arrangement yields,

$$
\begin{aligned}
&\widetilde{[Z,Z]}_t-[Z,Z]_t\\
&=:VII+IIX+IX\\
&=\big(\widetilde{[Z,Z]}_t-[\tilde{Z},\tilde{Z}]_t-([Z,Z]_t^{(n)}-[Z,Z]_t)\big)\\
&\quad+\big([\tilde{Z},\tilde{Z}]_t-[Z,Z]_t\big)+\big([Z,Z]_t^{(n)}-[Z,Z]_t\big).
\end{aligned}
$$

By (2.6),

(6.30) $$[\tilde{Z},\tilde{Z}]_t=[Z,Z]_t+\int_0^t(\rho_u-\tilde{\rho}_u)^2d[S,S]_u$$

Evaluating the hedging error in price processes with jumps present 423

$$+ 2 \sum_{0 \leq T_j \leq t} (\rho_{T_j} - \tilde{\rho}_{T_j}) J_j^S (J_j^\Xi - \rho_{T_j} J_j^S);$$

By Theorem 6.1 of Jacod and Protter [11],

$$(6.31) \qquad VII = O_P(\sqrt{\overline{\Delta t}}); \qquad IX = O_P(\sqrt{\overline{\Delta t}}).$$

Lemma 6.4, (6.30)–(6.31) imply the theorem. $\qquad \square$

6.4 Proof of Theorem 4.1

In the proof of Theorem 3.1, it proceeds by first getting rid of the negligible bias of A_i and the effect of B_i, and then take $\overline{A}_i$ as the dominant part of the estimation error of the estimator of the instantaneous volatility of the price process, S, which determines the convergence rate of $\hat{\rho}$ to ρ. The proof of Theorem 4.1 follows similar lines to that of theorem 3.1 with slight modification noting that B_i does not reduce in this case.

Proof. Firstly, replace all the expectations in the proof of Theorem 3.2 by E_σ, conditional expectation on both σ^Ξ and σ^S.

Secondly, instead of using Taylor's expansion of σ^S around observation times in (6.2) and (6.6) when the volatility is deterministic and continuously differentiable, the following two lemmas concerning the discrete error of $\langle S, S \rangle'$ and S, respectively, ensure that we can go through steps (6.2) and (6.6).

Lemma 6.5. *When $t_i \leq u \leq t_{i+1}$, under the Assumption 4.1, $\langle S, S \rangle'_u - \langle S, S \rangle'_{t_i} = O_P(\sqrt{\overline{\Delta t}})$.*

Proof.

$$\langle S, S \rangle'_u - \langle S, S \rangle'_{t_i} = \int_{t_i}^u \mu_t^{VS} dt + \int_{t_i}^u \sigma_t^{VS} dW_t^{VS},$$

$$\langle \langle S, S \rangle', \langle S, S \rangle' \rangle |_{t_i}^u = \int_{t_i}^u (\sigma_t^{VS})^2 dt$$
$$\leq \max_{0 \leq t \leq T} (\sigma_t^{VS})^2 \Delta = O_P(\overline{\Delta t}). \qquad \square$$

Lemma 6.6. *Under Assumption 4.1, for arbitrary $\epsilon > 0$*

$$E_\sigma[(|\Delta S_{t_i}| - \sigma_{t_i}^S |\Delta B_{t_i}|)|\Delta S_{t_{i+1}}|] = O_P(\overline{\Delta t})^{(2-\epsilon)},$$

uniformly in i.

Proof.

$$E_\sigma[(|\Delta S_{t_i}| - \sigma_{t_i}^S |\Delta B_{t_i}|)|\Delta S_{t_{i+1}}|]$$
$$\leq \sqrt{E_\sigma(|\Delta S_{t_i}| - \sigma_{t_i}^S |\Delta B_{t_i}|)^2} \sqrt{E_\sigma(\Delta S_{t_{i+1}})^2}$$
$$\leq \max_{0 \leq t \leq T} \sigma_t^S \max_{0 \leq t_i \leq T} \sup_{t_i \leq u \leq t_{i+1}} (\sigma_u - \sigma_{t_i})^2 \Delta,$$

by using Lévy continuity modulus theorem to time-changed Gaussian paths, c.f. proof of Theorem 3.1 in Mancini [12]. $\qquad \square$

Thirdly, (6.5) is invalid under this new circumstance, so we replace (6.5) by another estimate.

Lemma 6.7. *Under Assumption 4.1, $B_i = O_P((\overline{\Delta t})^{1/4})$.*

Proof. Integration by parts yields
$$(6.32)$$
$$B_i = \frac{1}{h} \left(\int_{t_i-h}^{t_i} \int_{t_i}^u \mu_t^{VS} dt du + \int_{t_i-h}^{t_i} (t_i - h - u) \sigma_u^{VS} dW_u^S \right).$$

Boundedness of μ^{VS} and σ^{VS} and simple calculation of the variation of B_i concludes the result. $\qquad \square$

Fourthly, taking Lemma 6.5 into consideration, Lemma 6.2 holds with $E_\sigma(\overline{A}_i)^2$ replaced by $E_\sigma(\overline{A}_i + B_i)^2$. Since $E_\sigma(\overline{A}_i B_i) = 0$, Lemma 6.2 follows with $E_\sigma(\overline{A}_i)^2$ replaced by $E_\sigma(\overline{A}_i)^2 + (B_i)^2$.

Next, we calculate

$$(6.33) \qquad \frac{1}{\sqrt{\overline{\Delta t}}} \sum_{t_{i+1} \leq t} \frac{\rho_{t_i-h}^2}{(\langle S, S \rangle'_{t_i-h})^2} E_\sigma(B_i)^2 \langle S, S \rangle'_{t_i-h} \Delta t_i.$$

Lemma 6.8. *$(B_i)^2 = \frac{h}{3} \langle \langle S, S \rangle', \langle S, S \rangle' \rangle'_{t_i} + O_P((\overline{\Delta t})^{3/4})$.*

Now, up to small term, (6.29) is

$$\frac{1}{\sqrt{\overline{\Delta t}}} \sum_{t_{i+1} \leq t} \frac{\rho_{t_i-h}^2}{(\langle S, S \rangle'_{t_i-h})^2} \frac{h}{3} \langle \langle S, S \rangle', \langle S, S \rangle' \rangle'_{t_i} \Delta t_i,$$

whose limit is

$$(6.34) \qquad \frac{1}{3} \frac{1}{c^*} \int_0^t \frac{\rho_u^2}{(\langle S, S \rangle'_u)^2} d\langle \langle S, S \rangle', \langle S, S \rangle' \rangle_u$$

in probability. Similarly, by decomposing $\widehat{\langle \Xi, S \rangle}'_{t_i} - \widehat{\langle \Xi, S \rangle}'_{t_i}$ and using Kunita-Watanabe inequality, we can get the same estimate as in Lemma 6.3–6.6 for $\widehat{\langle \Xi, S \rangle}'_{t_i} - \widehat{\langle \Xi, S \rangle}'_{t_i}$, hence

$$\frac{1}{\sqrt{\overline{\Delta t}}} \sum_{t_{i+1} \leq t} \frac{1}{(\langle S, S \rangle'_{t_i-h})^2} E_\sigma(B_i^{\Xi S})^2 \langle S, S \rangle'_{t_i-h} \Delta t_i$$

$$(6.35) \qquad \to^P \frac{1}{3} \frac{1}{c^*} \int_0^t \frac{1}{(\langle S, S \rangle'_u)} d\langle \langle \Xi, S \rangle', \langle \Xi, S \rangle' \rangle_u$$

and

$$\frac{1}{\sqrt{\overline{\Delta t}}} \sum_{t_{i+1} \leq t} \frac{-2\rho_{t_i-h}}{(\langle S, S \rangle'_{t_i-h})^2} E_\sigma[(B_i^{\Xi S})(B_i)] \langle S, S \rangle'_{t_i-h} \Delta t_i$$

$$(6.36)$$
$$\to^P \frac{1}{3} \frac{1}{c^*} \int_0^t \frac{-2\rho_u}{(\langle S, S \rangle'_u)} d\langle \langle \Xi, S \rangle', \langle S, S \rangle' \rangle_u,$$

where $B_i^{\Xi S}$ is $\frac{1}{h}(\langle \Xi, S \rangle_{t_i} - \langle \Xi, S \rangle_{t_i-h}) - \langle \Xi, S \rangle'_{t_i}$. Combining (6.13)–(6.15), (6.30)–(6.32) and using Itǒ's formula on ρ_u, the bias term of Theorem 4.1 is obtained.

Finally, Proposition 6.2 still holds without any changes to do with the proof in subsection 6.1. This completes the proof of Theorem 4.1. $\qquad \square$

6.5 Proof of Theorem 4.4

Proof. Replacing all the expectations in the proof of Theorem 3.4 by E_σ, conditional expectation on both σ^Ξ and σ^S and following the proof of Theorem 3.4 lead to completion of the proof. $\qquad\square$

REFERENCES

[1] AIT-SAHALIA, Y. (2000). Telling from discrete data whether the underlying continuous time model is a diffusion. *Journal of Finance* **57** 2075–2112.

[2] AIT-SAHALIA, Y. (2004). Distentangling diffusion from jumps. *Journal of Financial Econometrics* **74** 487–528.

[3] AIT-SAHALIA, Y. and JACOD, J. (2007). Volatility estimators for discretely sampled Levy processes. *Ann. Stat.* **35** 355–392. MR2332279

[4] AIT-SAHALIA, Y., MYKLAND, P. and ZHANG, L. (2005). How often to sample a continuous time process in the presence of market microstructrue noise. *Review of Financial Studies* **18** 351–416.

[5] ANDERSEN, T., BOLLERSLEV, T. and DIEBOLD, F. (2007). Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility. *Review of Economics and Statistics* **89** 701–720.

[6] BARNDORFF-NIELSEN, O. and SHEPHARD, N. (2001). Econometric analysis of realized volatility and its use in estimating. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **64** 253–280. MR1904704

[7] BARNDORFF-NIELSEN, O. and SHEPHARD, N. (2003a). Econometrics of testing for jumps in financial economics using bipower variation. Unpublished discussion paper: Nuffield College, Oxford.

[8] BARNDORFF-NIELSEN, O. and SHEPHARD, N. (2004b). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics.* **2** 1–48.

[9] FOLLMER, H. and SONDERMANN, D. (1986). Hedging of non-redundant contingent claims. In: *Contributions to Mathematical Economics* (W. Hildenbrand and A. Mas-Colell, eds). North Holland, Amesterdam, pp. 205–223. MR0902885

[10] JACOD, J. (2006). Asymptotic properties of realized power vaiation and associated functions of semimartingale. arXiv, 9 June 2006 n.0023146.

[11] JACOD, J. and PROTTER, P. (1998). Asymptotic error distributions for the Euler method for stochastic differential equations. *Ann. Probab.* **26** 267–307. MR1617049

[12] MANCINI, C. (2006). Non parametric threshold estimation for models with stochastic diffusion coefficient and jumps. arXiv: math/0607378vl [math. ST] MR2528985

[13] MANCINI, C. and RENO, R. (2006). Threshold estimation of jump-diffusion models and interest rate modelling. Working paper.

[14] MYKLAND, P. (2006). A Gaussian Calculus for Inference from High Frequency Data. TR 563, Dept of Statistics, The Univ. of Chicago.

[15] MYKLAND, P. and ZHANG, L. (2006). ANOVA for diffusions and Ito processes. *Ann. Stat.* **34** 1931–1963. MR2283722

[16] ZHANG, L. (2001). From martingales to ANOVA: Implied and realized volatility. Ph.D. dissertation, Dept. Statistics, Univ. Chicago. MR2716972

BingYi Jing
Department of Mathematics
Hong Kong Univ. of Science and Technology
Clear Water Bay
Kowloon, HongKong
People's Respublic of China
E-mail address: majing@ust.hk

XinBing Kong
Department of Statistics
School of Management
Fudan University
Guoshun Road
Shanghai
People's Respublic of China
E-mail address: kongxb@fudan.edu.cn

Zhi Liu
Department of Mathematics
Faculty of Science and Technology
University of Macau
People's Respublic of China
E-mail address: liuzhi@ust.hk

Bo Zhang
School of Statistics
Renmin University of China
#59 Zhongguancun Street
Beijing 100872
People's Respublic of China
E-mail address: mabzhang@ruc.edu.cn