

ASYMPTOTIC ANALYSIS OF HEDGING RISK UNDER ESTIMATED STRATEGY IN PRESENCE OF JUMPS

BY ZHI LIU BINGYI JING XINBING KONG

Hong Kong University of Science Technology

Lévy-type processes are considered as driven process of a financial security frequently in recent years. This paper is concerned with hedging risk between two securities driven by different Lévy-type processes with finite activity jumps. When jumps are present, we find the risk that hedge one security with other is larger than continuous case under high frequency sense. Using same scheme to estimate the residual sum of square, we get the rate is only $(\Delta t)^{1/4}$ which is different as the continuous case where the rate is $\sqrt{\Delta t}$.

1. Introduction. We consider a stochastic regression $d\Xi = \rho dS$ between two stochastic processes Ξ_t and S_t with the form:

$$d\Xi_t = \Xi_0 + \Xi^c + \Xi^J \star \mu^\Xi, \quad dS_t = S_0 + S_c + S^J \star \mu^S \quad (1.1)$$

Where, Ξ_t and S_t are Lévy-type processes with finite activity jumps which may represent the log price of two securities or derivatives in financial applications. The terms with superscript c means that it is continuous part which include finite variation drift and infinite variation martingale, i.e. Brownian integral, J and μ are jumps size and Lévy measure of which superscript, respectively. See section 2 for detail for Lévy type processes.

Hedging risk is most interesting research topic of financial management group. Under the assumption of completeness, self-financing property can guarantee that the derivatives can be hedged perfectly by underlying security, i.e., the hedging error is zero. But this imposed assumption is too naive for real financial world. Many evidences have been found that the market doesn't satisfy this assumption even in the interest rate model, see[10, 20], although classical Black-Scholes model established based this framework. The fail of complete assumption motivate the interesting to consider the hedging risk. When the market model is incomplete but still continuous, the problem are solved completely by the representative works of Föllmer and Martin Schweizer[11], they established the optimal strategy of hedging

AMS 2000 subject classifications: Primary 60G44, 62M09, 62M10, 62M10, 91B28, Quadratic hedging, central limit theorem, quadratic variation, threshold realized variation, Lévy-type processes, high frequency data.; secondary 60G42, 62G20, 62P20, 91B84

suppose underlying security is martingale, it is $\frac{\langle \Xi, S \rangle'}{\langle S, S \rangle'}$. Martin Schweizer[1] extend the result to general semimartingale case where the risk-minimizing strategy may not exist, he consider the so called local-risk minimizing strategy, the key tool they used is Kunita-Watanabe decomposition. However, Under the setting with jumps, the unhedgeable part include two parts, one dues to jumps and other dues to continuous component(for example, incomplete information), they are called market risk and operational risk (also known as process risk) in financial literature, respectively. As we will see in section 3, when the jump risk involved, then it play dominated role under sense of high frequency data. Some of the results here heavily rely upon the paper by Mykland and Zhang[14] in which similar problem have been solved when both Ξ and S are continuous.

The optimal strategy has been previously studied by many authors, but in the range of our knowledge, the only work concern the statistical estimation of minimize risk is Mykland and Zhang[14], they did it using a very straightforward method, obtain a estimation of minimum risk. we are interested in statistical estimation of minimum risk that the regression makes between two processes by using high frequency data. Consider the model

$$\Xi_t = \int_0^t \rho_t dS_t + Z_t \quad (1.2)$$

From pointview of quadratic hedging risk, we wish to estimate from the data: $\min_{\rho} [Z, Z]_T$, i.e., minimize the variance like that concerned in regression problem, and where the minimum is over all regression processes ρ , $[X, X]$ is quadratic variation of X defined as following:

$$[X, X]_t = \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})^2,$$

$0 = t_0 < t_1 < t_2 \cdots < t_n = T$ is partition of interval $[0, t]$ such that $\max_i (t_{i+1} - t_i) \rightarrow 0$.

There is two main purposes this paper, one is find the optimal hedging strategy ρ , then deduce the asymptotic behavior of the estimated $[Z, Z]_t$, as more and observations are available (or professionally high frequency data are handed) within a fixed time window. We can use different schemes to estimate $[\widehat{Z}, \widehat{Z}]_t^n$. However no matter which of our estimators is used, we start from following attempt:

$$[\widehat{Z}, \widehat{Z}]_t^n - [Z, Z]_t = \text{bias} + \text{variance}, \quad (1.3)$$

we consider the estimator of discrete and continuous parts separately, to first-order asymptotically, where $[Z, Z]_t^n$ is the sum of squares of the increments of the process Z at the sampling points, $[Z, Z]_t^n = \sum_{0 \leq t_i \leq t} (Z_{t_{i+1}} - Z_{t_i})^2$.

The paper are arranged as following, in section 2 we construct a framework of concerned model, section 3 is the main result, a central limit theory are considered. Last section, we propose a simple Monte Carlo simulation for fit our theory result. Proofs are given in appendix.

2. The Model: Description and Setup. Lévy-type processes, realized quadratic variation, threshold realized quadratic variation. We start from a Lévy-type processes defined on the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)_{0 \leq t \leq T}$ and we assume that $\{\mathcal{F}_t, t \geq 0\}$ satisfy the usual assumption and that $\mathcal{F} = \bigvee_{t \leq \infty} \mathcal{F}_t$. All the processes we will make use in following are assumed adapted with this filtration.

DEFINITION 1(Lévy-type processes). By saying that X is a Lévy-type process with finite activity jumps, we mean that X has the form:

$$X_t = X_0 + \int_0^t b(u)du + \int_0^t \sigma(u)dW_u + \int_0^t \int_R K(x, u)\mu(dx, du), \quad (2.1)$$

Where $\{W_t; t \geq 0\}$ is a standard Brownian motion and μ is an independent Poisson random measure with intensity measure ν satisfied following conditions

- (1) b, σ and K are predictable with respect to filtration $\{\mathcal{F}_t, t \geq 0\}$
- (2) $\int_0^T \int_R |b(u)|\nu(du, dx) < \infty, a.s.$
- (3) $\int_0^T \sigma(u)^2 du < \infty, a.s.$

This is a simple case of general Lévy-type integral, some authors call it jump diffusion, or the case of Blumenthal-Gatoor index is zero. See David Applebaum[16] and Cont Fama[17].

For this type of processes, the quadratic variation $[X, X]$ can be expressed in terms of the represents (2.1) by

$$\begin{aligned} [X, X]_t &= \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})^2 = [X^c, X^c]_t + [X^d, X^d]_t \\ &= \int_0^t \sigma^2(u)du + \int_0^t \int_R J^2(x, u)\mu(dx, du). \end{aligned}$$

We denote $[X^c, X^c]'_t$ the derivative of $[X^c, X^c]_t$ with respect to time t . Then $[X^c, X^c]'_t = \sigma^2(t)$, and this quantity is often known as square volatility in the finance literature.

We now suppose that we observe processes, in particular, S_t and Ξ_t , on a finite set $\mathcal{G} = \{0 = t_0 < t_1 < \dots < t_k = T\}$ of time points in the interval $[0, T]$.

Similarly, for two processes X and Y observed on a grid \mathcal{G} , realized quadratic covariation are defined by

$$[X, Y]_t^{\mathcal{G}} = \sum_{t_i \leq t} (\Delta_i X)(\Delta_i Y), \quad (2.3)$$

Since we consider the asymptotical property, we shall take limits with the help of a sequence of partitions $\mathcal{G}_n = \{0 = t_0 < t_1 < \dots < t_n = T\}$. As $n \rightarrow +\infty$, we let \mathcal{G}_n become dense in $[0, T]$, in the sense that the mesh tends to 0, i.e.,

$$\max_i |t_{i+1}^{(n)} - t_i^{(n)}| \rightarrow 0.$$

We will ignore the superscript (n) in where no confusion induced. In this case, $[X, Y]_t^{\mathcal{G}_n}$ converges to $[X, Y] = \frac{1}{4}([X + Y, X + Y]_t - [X - Y, X - Y]_t)$ uniformly in probability and L^2 . See Jacod and Shiryaev[12] or David Applebaum[16].

Under the assumption of jump diffusion, the one method separate the diffusion from jump is threshold realized quadratic variation by Mancini[18]

$$[X, X]^{TV} = \lim_{n \rightarrow \infty} \sum_{i=1}^n (X_i - X_{i-1})^2 \mathbf{1}_{((X_i - X_{i-1})^2 \leq \alpha(\Delta t_i)^\varpi)},$$

$$[X, Y]^{TV} = \lim_{n \rightarrow \infty} \sum_{i=1}^n (X_i - X_{i-1})(Y_i - Y_{i-1}) I_{((X_i - X_{i-1})^2 \leq \alpha(\Delta t_i)^\varpi, (Y_i - Y_{i-1})^2 \leq \alpha(\Delta t_i)^\varpi)},$$

Where $\varpi \in (0, 1/2)$, $\alpha > 0$.

We need some assumptions following

ASSUMPTION A: The process $\{\sigma_s^S, s \geq 0\}$ is càdlàg, locally bounded and positive almost surely.

ASSUMPTION 1: The observations is non-equidistance and satisfying

$$\frac{\sum_{t_i - h \leq t_i \leq t} (\Delta t_i)^2}{\Delta t h} \rightarrow H_2'(t).$$

ASSUMPTION 2: For each $n \in N$, we have a sequence of partitions $\{t_i^n\}$, $\Delta t_i^n = t_{i+1} - t_i$, the sequence $\{t_i^n\}$ is balance, We say that the sequence is *balanced* if $\max \Delta t_i / \min \Delta t_i$ is bounded above.

The assumption A assures that optimal ρ_t given by $\rho_t = \frac{[\Xi^c, S^c]_t}{[S^c, S^c]_t}$ is well defined. In the discontinuous world, where Ξ and S can only be observed

over grid times.

We shall use the following notations, where X is any semimartingales:

- $\Delta_i X$ is the increment $X_{t_{i+1}} - X_{t_i}$, $i = 1, 2, \dots$
- X_{J_t} is the size $X_t - X_{t-}$ of jump at time t (maybe zero), X_{J_j} is jump size of j th jump, N_t^X is jump process of X with arrival time T_j .
- $\overline{\Delta t} = t/n$, $t^* = \max\{t_i, t_i \leq t\}$.

3. Main result.

3.1. *Estimation Scheme.* Recall we consider the regression problem between Ξ and S :

$$d\Xi_t = \rho_t dS_t + dZ_t \quad (3.1)$$

Write (3.1) as $Z_t = \Xi_t - \int_0^t \rho_u dS_u$, simple calculation implies

$$\begin{aligned} [Z, Z]_t &= [\Xi, \Xi]_t + \int_0^t \rho_u^2 d[S, S]_u - 2 \int_0^t \rho_u d[\Xi, S]_u \\ &= [\Xi^c, \Xi^c]_t + [\Xi^d, \Xi^d]_t + \int_0^t \rho_u^2 d[S^c, S^c]_u \\ &\quad - 2 \int_0^t \rho_u d[\Xi^c, S^c]_u + \sum_{j=1}^{N_t^S} \rho_{T_j}^2 (S_{J_j})^2 - 2 \sum_{j=1}^{N_t^S} \rho_{T_j} \Xi_{J_j} S_{J_j}. \end{aligned} \quad (3.2)$$

Last term isn't zero when Ξ and S have common jumps, however this is reasonable and highly likely, for example, common jumps can be induced by jumps of index, in market efficiency and basis portfolio theory, jumps in the index can be generated by market-level news, while jumps in individual stocks be generated by either stock-specific or common news.

If S has continuous path, it is easy to get the Radon-Nikodym derivative $\rho_t = \frac{d[\Xi, S]_t}{d[S, S]_t}$ minimizes $[Z, Z]_T$, so the case reduced to the one considered by Mykland and Zhang[14] or Föllmer[11].

Otherwise, mathematically, the ρ which minimize $[Z, Z]_T$ is given by

$$\rho_t = \frac{[\Xi^c, S^c]_t'}{[S^c, S^c]_t'} \mathbf{1}_{(\Delta S_t=0)} + \frac{\Delta \Xi_t}{\Delta S_t} \mathbf{1}_{(\Delta S_t \neq 0)},$$

see figure 1 for a simple plot. From the form of ρ , $\rho_{T_{i+}} = \rho_{T_{i-}} \neq \rho_{T_i}$. However, a basic requirement for strategy should be predictable. Observe that if we let

$$\rho_t^c = \frac{[\Xi^c, S^c]_t'}{[S^c, S^c]_t'} \quad (3.3)$$

then $P(\rho_t^c \neq \rho_t) = P(\Delta S_t \neq 0) = 0$ for each $t \geq 0$, this implies ρ_t^c is a modification of ρ_t , it is predictable because of continuity of path.

The problem $\min_{\rho} [Z, Z]_T$ then connects to an ANOVA. Unfortunately, unlike the continuous case, the variance partly comes from residual and partly comes from jumps. Let Z_t be the residual in (1.1) for optimal ρ_t , so that the quantity $\min_{\rho} [Z, Z]_T$ can be written simple as $[Z, Z]_T$. In analogy with regular regression, substituting ρ_t^c into (3.2) gives rise to an ANOVA decomposition of the form

$$\underbrace{[\Xi, \Xi]}_{\text{total SS}} = \underbrace{\int_0^T \rho_u^2 d[S, S]_u}_{\text{SS explained}} + \underbrace{[Z, Z]_t}_{\text{RSS}} + 2 \underbrace{\sum_{j=1}^{N_t^S} \rho_{T_j} S_{J_j} (\Xi_{J_j} - \rho_{T_j} S_{J_j})}_{\text{JSS}}, \quad (3.4)$$

where "SS" is abbreviation for "sum of squares", "RSS" and "JSS" stand for "residual sum of squares" and "SS induced by Jump", respectively. If continuous observation are available, one can solve the problem (3.4) by using the ρ and Z defined above. Our target of inference, $[Z, Z]_t$ would then be observable. Discreteness of observation, however, creates the need for inference.

To estimate $[Z, Z]_t$, we have to estimate ρ_t first, i.e., $[S^c, S^c]'_t$ and $[S^c, \Xi^c]'_t$. Then one estimator of ρ given by

$$\hat{\rho}_t = \frac{\widehat{[\Xi^c, S^c]'_t}}{\widehat{[S^c, S^c]'_t}} = \frac{[\Xi, S]_t^{TV} - [\Xi, S]_{t-h_n}^{TV}}{[S, S]_t^{thp} - [S, S]_{t-h_n}^{TV}},$$

A detail description about threshold quadratic variation can be found in Mancini[18] and Jacod and Shirayayes[12]. Empirical simulation shows that optimal choice of ϖ is 0.48 or 0.49. Also here we have to use a smoothing bandwidth h_n . There will naturally be a tradeoff between h_n and $\overline{\Delta t}^{(n)}$. Here, we will use $h_n = O((\overline{\Delta t}^{(n)})^{1/2})$. Other way to estimate the integrated volatility from jump is so-called bipower or multipower method.

We now return to the estimation of the quadratic variation $[Z, Z]$ of residuals. Given the discrete data of (Ξ, S) , there are different methods to estimate the residual variation.

We also use the schemes given by Mykland and Zhang[12], first one is to start with model (3.1). For a fixed grid \mathcal{G} , one first estimates $\Delta_i Z$ through the relation $\Delta_i \hat{Z} = \Delta_i \Xi - \hat{\rho}_{t_i}(\Delta_i S)$, where $\Delta_i Z = Z_{t_{i+1}} - Z_{t_i}$, then the estimator of quadratic variation of Z given by

$$[\hat{Z}, \hat{Z}]_t^{(n)} = \sum_{t_{i+1} \leq t} (\Delta_i \hat{Z})^2 = \sum_{t_{i+1} \leq t} [\Delta_i \Xi - \hat{\rho}_{t_i}(\Delta_i S)]^2, \quad (3.5)$$

where the notation of square brackets is invoked, since the increments is taken over discrete times.

Alternatively, one can directly analyze the ANOVA version of (3.4) of the model, where

$$d[Z, Z]_t = d[\Xi, \Xi]_t - \rho_t^2 d[S, S]_t - 2\rho_t S_{J_t} (\Xi_{J_t} - \rho_t S_{J_t}).$$

This yields a second estimator of $[Z, Z]_t$,

$$\begin{aligned} [\tilde{Z}, \tilde{Z}]_t &= \sum_{t_{i+1} \leq t} [(\Delta_i \Xi)^2 - \hat{\rho}_{t_i}^2 (\Delta_i S)^2] \\ &- 2 \sum_{t_{i+1} \leq t} \hat{\rho}_{t_i} \Delta_i S (\Delta_i \Xi - \hat{\rho}_{t_i} \Delta_i S) \mathbf{1}_{((\Delta_i S)^2 \geq \alpha (\Delta t_i)^\varpi, (\Delta_i \Xi)^2 \geq \alpha (\Delta t_i)^\varpi)}. \end{aligned} \quad (3.6)$$

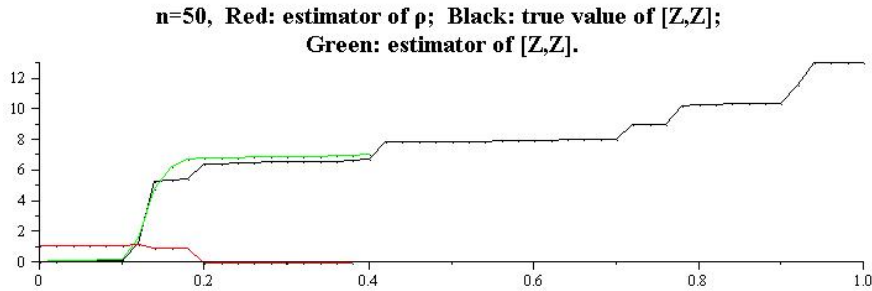
3.2. The asymptotical result: The distribution of risk. Recall that the square bracket with and without superscript n represent the variation of Z at discrete and continuous time scale, respectively.

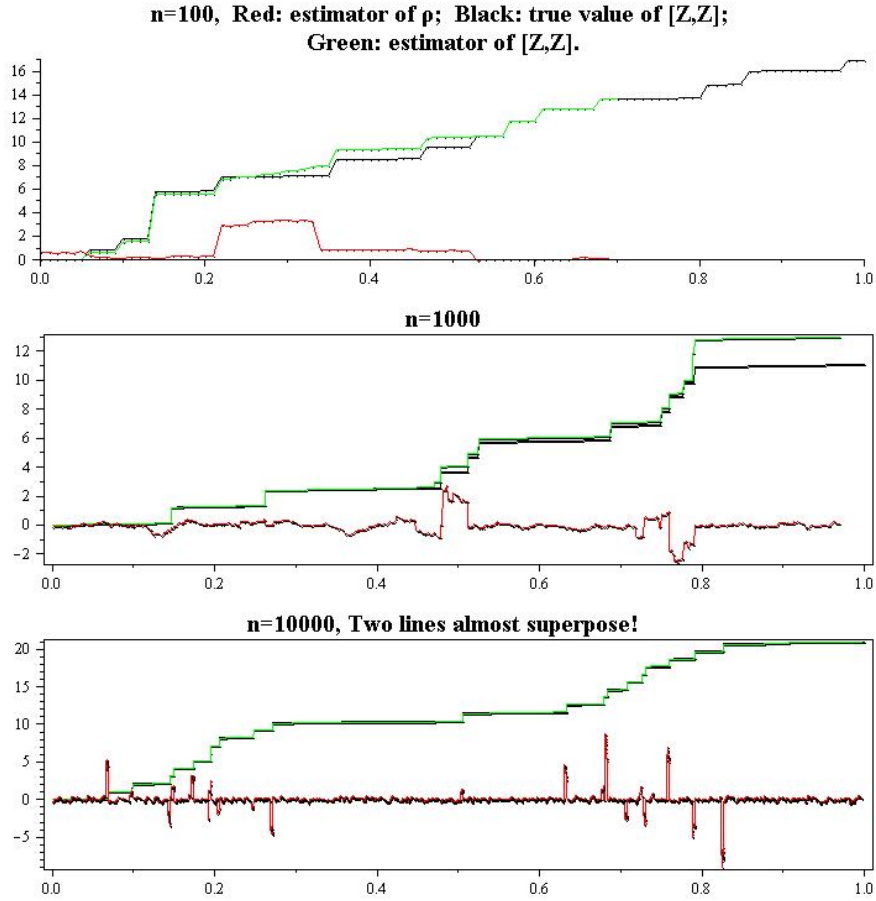
THEOREM 3.1. *Subject to some regular conditions, as $n \rightarrow \infty$,*

$$(\overline{\Delta t}^{(n)})^{-1/4} ([\hat{Z}, \hat{Z}]_t^{(n)} - [Z, Z]_t) \xrightarrow{\mathcal{L}} \xi_0 + 2 \sum_{j=1}^{N_t^S} S_{J_j} (\Xi_{J_j} - \rho_{T_j} S_{J_j}) \xi_j.$$

uniformly in $t \in [0, T]$. where ξ_0 is random variable $N(0, 4 \int_0^t V_{\hat{\rho}_s - \rho_s} ([S^c, S^c]_s')^2 ds)$, $V_{\hat{\rho}_t - \rho_t} = H_2'(t) (\frac{[\Xi^c, \Xi^c]_t'}{[S^c, S^c]_t'} - \rho_t^2)$, ξ_j is a sequence of standard normal variables independent of ξ_0 .

4. Simulation study. We simulate a simple case, where Ξ and S have same distribution, both have the form $W_t + J_t$, but they are generated independently, in this case, optimal ρ is 0, the simulation result is following:





REMARK 4.1. *In the first graph, we only use 20 data to estimate ρ , so the estimator haven't the graph after 0.4. Similar situation appears in next three graphs. For each estimate procedure, we regenerate the data, so the true variation is different in the four graphs.*

5. Appendix.

PROPOSTION 5.1. *Let Z be a semimartingale with finite activity jumps for which $\int_0^T ([Z^c, Z^c]_t)' dt < \infty$, a.s. and $\int_0^T Z_t^2 dt < \infty$ a.s.:*

$$\frac{1}{\sqrt{\Delta t^{(n)}}} ([Z, Z]_t^{(n)} - [Z, Z]_t^n) \xrightarrow{FD} \int_0^t \sqrt{2H'(t)} [Z^c, Z^c]_u' dW_u + \sum_{s \leq t} \sqrt{[Z^c, Z^c]_s} V_s (Z_s - Z_{s-}),$$

where FD is finite dimensionally and W is a standard Brownian motion, independent of underlying data process, V_s is a sequence of independent standard normal variables.

REMARK 5.1. *In this case, it is different with continuous case that we have not convergence stably for $\frac{1}{\sqrt{\Delta t^{(n)}}}([Z, Z]_t^{(n)} - [Z, Z]_t)$, or much more we even haven't convergence in law, the due to the untractable property of Skorohod topology. See Jacod[19].*

LEMMA 5.1. *Under our assumptions, we have*

1). *Consistency*

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

2). *Asymptotical normality*

$$\frac{1}{\sqrt{h}}(\hat{\rho}_t - \rho_t) \xrightarrow{\mathcal{L}} N(0, V_{\hat{\rho}_t - \rho_t}),$$

where $V_{\hat{\rho}_t - \rho_t} = H_2'(t) \left(\frac{[\Xi^c, \Xi^c]'_t}{[S^c, S^c]'_t} - \rho_t^2 \right)$.

PROOF: Note that

$$\begin{aligned} \hat{\rho}_{t_i} - \rho_{t_i} &= \frac{1}{[\widehat{S^c, S^c}]'_{t_i}} \left([\widehat{\Xi^c, S^c}]'_{t_i} - [\Xi^c, S^c]'_{t_i} - \rho_{t_i} ([\widehat{S^c, S^c}]'_{t_i} - [S^c, S^c]'_{t_i}) \right) \\ &= \frac{1}{[S^c, S^c]'_{t_i}} \left([\widehat{\Xi^c, S^c}]'_{t_i} - [\Xi^c, S^c]'_{t_i} - \rho_{t_i} ([\widehat{S^c, S^c}]'_{t_i} - [S^c, S^c]'_{t_i}) \right) + o_p([\widehat{S^c, S^c}]'_{t_i} - [S^c, S^c]'_{t_i}) \end{aligned}$$

it is enough to show that $[\widehat{S^c, S^c}]'_{t_i} - [S^c, S^c]'_{t_i} = O_p(\overline{\Delta t}^{1/4})$, while this follows from the limit property of threshold quadratic variation,

$$\begin{aligned} [\widehat{S^c, S^c}]'_{t_i} - [S^c, S^c]'_{t_i} &= \frac{1}{h} \sum_{t_i - h \leq t_j < t_{j+1} \leq t_i} (\Delta_j S)^2 \mathbf{1}_{\{(\Delta_j S)^2 \leq \alpha(\Delta t_j)^\varpi\}} - [S^c, S^c]'_{t_i} \\ &= \frac{1}{h} \sum_{t_i - h \leq t_j < t_{j+1} \leq t_i} \left((\Delta_j S)^2 \mathbf{1}_{\{(\Delta_j S)^2 \leq \alpha(\Delta t_j)^\varpi\}} - \Delta_j [S^c, S^c] \right) \\ &\quad + \frac{1}{h} \sum_{t_i - h \leq t_j < t_{j+1} \leq t_i} \Delta_j [S^c, S^c] - [S^c, S^c]'_{t_i} + O_p(\sqrt{\Delta t}) \\ &:= A^{SS} + B^{SS} + O_p(\sqrt{\Delta t}) \end{aligned}$$

$B^{SS} = O_p(h) = O_p(\sqrt{\Delta t})$ by Taylor expansion, now we show that $A^{SS} = O_p((\Delta t)^{1/4})$,

$$\begin{aligned} A^{SS} &= \frac{1}{h} \sum_{t_i - h \leq t_j < t_{j+1} \leq t_i} \left((\Delta_j S)^2 \mathbf{1}_{\{(\Delta_j S)^2 \leq \alpha(\Delta t_j)^\varpi\}} - \Delta_j [S^c, S^c] \right) \\ &= \sum_{t_i - h \leq t_j < t_{j+1} \leq t_i} \eta_j, \end{aligned}$$

where $\eta_j = \frac{1}{h} \left((\Delta_j S)^2 \mathbf{1}_{\{(\Delta_j S)^2 \leq \alpha(\Delta t_j)^\varpi\}} - \Delta_j[S^c, S^c] \right)$, a simple calculation can show that:

$$\begin{aligned} E[\eta_j^2 | \mathcal{F}_{i-1}] &= \overline{\Delta t} E \left[\left(\left(\frac{\Delta_j S}{\Delta t} \right)^2 \mathbf{1}_{\left\{ \left(\frac{\Delta_j S}{\sqrt{\Delta t_j}} \right)^2 \leq \alpha(\Delta t_j)^{\varpi-1/2} \right\}} - \frac{\Delta_j[S^c, S^c]}{\Delta t} \right)^2 \middle| \mathcal{F}_{i-1} \right] \\ &= \overline{\Delta t} E \left[\left(\left(\frac{\Delta t_i}{\Delta t} \right)^2 \sigma_j^2 \xi^2 \mathbf{1}_{\{\sigma_j^2 \xi^2 \leq \alpha(\Delta t_j)^{\varpi-1/2}\}} - \frac{\Delta_j[S^c, S^c]}{\Delta t} \right)^2 \middle| \mathcal{F}_{i-1} \right] \\ &= O_p(\overline{\Delta t}), \end{aligned}$$

ξ is standard normal variable, then $\sum_{t_i-h \leq t_j < t_{j+1} \leq t_i} E[\eta_j^2 | \mathcal{F}_{i-1}] = O_p(\sqrt{\Delta t})$, Hence by Lenglart inequality

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

Similarly, we know that $B^{\Xi S}, A^{\Xi S}$ have a order of $O_p((\Delta t)^{1/4})$, hence we proved 1).

2). It is suffice to prove that

$$\frac{1}{\sqrt{h}} B^{SS} \xrightarrow{P} 0, \quad \frac{1}{\sqrt{h}} B^{SS} \xrightarrow{P} 0,$$

and

$$\frac{1}{\sqrt{h}} \begin{pmatrix} A^{SS} \\ A^{\Xi S} \end{pmatrix} \xrightarrow{\mathcal{L}} N(0, M).$$

$$\text{Where } M = \begin{pmatrix} 2([S^c, S^c]'_t)^2 & 2[\Xi^c, \Xi^c]'_t [\Xi^c, \Xi^c]'_t \\ 2[\Xi^c, \Xi^c]'_t [\Xi^c, \Xi^c]'_t & [\Xi^c, \Xi^c]'_t [\Xi^c, \Xi^c]'_t + ([\Xi^c, S^c]'_t)^2 \end{pmatrix}.$$

Recall

$$A^{SS} = \frac{1}{h} \sum_{t_i-h \leq t_j < t_{j+1} \leq t_i} \left((\Delta S_j)^2 \mathbf{1}_{\{(\Delta S_j)^2 \leq \alpha(\Delta t_j)^\varpi\}} - \Delta_j[S^c, S^c] \right),$$

when jump intensity is finite, then by Mancini[18], for some large n , if $\omega < 1/2$,

$$\mathbf{1}_{\{(\Delta_j S)^2 \leq \alpha(\Delta t)^\omega\}} = \mathbf{1}_{\{N([t_j, t_{j+1}])=0\}}$$

almost surely. Therefore almost surely for large n ,

$$\begin{aligned} & \frac{1}{\sqrt{h}} \sum_{t_i-h \leq t_j \leq t_i} \left((\Delta_j S)^2 \mathbf{1}_{\{(\Delta_j S)^2 \leq \alpha(\Delta t)^\omega\}} - (\Delta_j S^c)^2 \right) \\ &= \frac{1}{h} \sum_{j=1}^{N_t^S} (\Delta_j S^c)^2 \\ &= O_P(h). \end{aligned}$$

That is

$$\frac{1}{\sqrt{h}}(A^{SS} - \frac{1}{h} \sum_{t_i-h \leq t_j \leq t_i} ((\Delta_j S^c)^2 - \Delta_j[S^c, S^c])) \xrightarrow{P} 0.$$

Direct application of Slusky's theorem, delta method and theorem 4.1 of Zhang[14] get required result.

PROOF of theorem 1: Now, we derive the limit distribution of $[\hat{Z}, \hat{Z}]_t^{(n)} - [Z, Z]_t$ or the limit for discretization error. The $[\hat{Z}, \hat{Z}]_t^{(n)} - [Z, Z]_t$ can be decomposed into bias and pure discretization error $[Z, Z]_t^{(n)} - [Z, Z]_t$. Start from first estimator, we may extend the definition of \hat{Z}_t when t is not a sampling point as following:

$$\hat{Z}_t = \hat{Z}_{t^*} + \Xi_t - \Xi_{t^*} - \int_{t^*}^t \hat{\rho}_{u^*} dS_u,$$

where, $t^* = \max\{t_i, t_i \leq t\}$. By definition of \hat{Z} ,

$$\begin{aligned} & [\hat{Z}, \hat{Z}]_t - [Z, Z]_t \\ &= \sum_{t_{i+1} \leq t} \Delta_i[\hat{Z}, \hat{Z}] - \sum_{t_{i+1} \leq t} \Delta_i[Z, Z] \\ &= \sum_{t_{i+1} \leq t} (\Delta_i[\Xi, \Xi] - 2 \int_{t_i}^{t_{i+1}} \hat{\rho}_u d[\Xi, S]_u + \int_{t_i}^{t_{i+1}} \hat{\rho}_u^2 d[S, S]_u) \\ &\quad - \sum_{t_{i+1} \leq t} (\Delta_i[\Xi, \Xi] - 2 \int_{t_i}^{t_{i+1}} \rho_u d[\Xi, S]_u + \int_{t_i}^{t_{i+1}} \rho_u^2 d[S, S]_u) \\ &= \int_0^t (\hat{\rho}_u - \rho_u)^2 d[S^c, S^c]_u + 2 \sum_{j=1}^{N_t^S} (\hat{\rho}_{T_j^*} - \rho_{T_j}) S_{J_j} (\Xi_{J_j} - \rho_{T_j} S_{J_j}). \end{aligned}$$

For prove the theorem, we have to know the order of $[\hat{Z}, \hat{Z}]_t^{(n)} - [Z, Z]_t$, i.e, following Lemma.

LEMMA 5.2. $[\hat{Z}, \hat{Z}]_t^{(n)} - [Z, Z]_t$ have an order of $(\Delta t)^{1/4}$.

Proof: Use the order of discretization error of $[Z^c, Z^c]_t$, $[Z^d, Z^d]_t$ and in-

crement of diffusion, we have

$$\begin{aligned}
& [\hat{Z}, \hat{Z}]_t^{(n)} - [\hat{Z}, \hat{Z}]_t \\
&= \sum_{t_i \leq t} (\Delta_i \Xi - \hat{\rho}_{t_i} \Delta_i S)^2 - [\hat{Z}, \hat{Z}]_t \\
&= O_P((\Delta t)^{1/2}) + 2 \sum_{t_i \leq t} (\rho_{t_i} - \hat{\rho}_{t_i}) \Delta_i S^c \Delta_i \Xi^c - 2 \sum_{t_i \leq t} (\rho_{t_i} - \hat{\rho}_{t_i}) (\Delta_i S^c)^2 \\
&\quad + 4 \sum_{j=1}^{N_t^S} (\rho_{T_j} - \hat{\rho}_{T_j^*}) (\Xi_{J_j} - \rho_{T_j} S_{J_i})
\end{aligned}$$

Therefore, it suffice to show *i)* and *ii)* following as $n \rightarrow \infty$.

- i).* $\frac{1}{\sqrt{h}} \sum_{t_i \leq t} (\rho_{t_i} - \hat{\rho}_{t_i}) (\Delta_i S^c)^2 - \frac{1}{\sqrt{h}} \sum_{t_i \leq t} (\rho_{t_i} - \hat{\rho}_{t_i}) \Delta_i [S^c, S^c] \xrightarrow{P} 0$,
ii). $\frac{1}{\sqrt{h}} \sum_{t_i \leq t} (\rho_{t_i} - \hat{\rho}_{t_i}) \Delta_i [S^c, S^c] \xrightarrow{\mathcal{L}} N(0, \int_0^t \text{Var}_{\hat{\rho}_s - \rho_s} ([S^c, S^c]'_s)^2 ds)$.

To prove *i)*, by proposition 1 and Lemma 1,

$$\begin{aligned}
& \frac{1}{\sqrt{h}} \sum_{t_i \leq t} (\rho_{t_i} - \hat{\rho}_{t_i}) (\Delta_i S^c)^2 - \frac{1}{\sqrt{h}} \sum_{t_i \leq t} (\rho_{t_i} - \hat{\rho}_{t_i}) \Delta_i [S^c, S^c] \\
&\leq \frac{\sup_i |\rho_{t_i} - \hat{\rho}_{t_i}|}{\sqrt{h}} \sum_{t_i \leq t} ((\Delta_i S^c)^2 - \Delta_i [S^c, S^c]) \\
&= O_p(1) O_P(\sqrt{\Delta t}) \\
&= o_p(1).
\end{aligned}$$

The left hand side of *ii)* just a sum of sequence of asymptotic independent normal random variables, therefore the limit still normal variable with mean zero, and variance is the sum of variance of all terms.

Following is some useful figures:

Figure 1 A plot of ρ_t

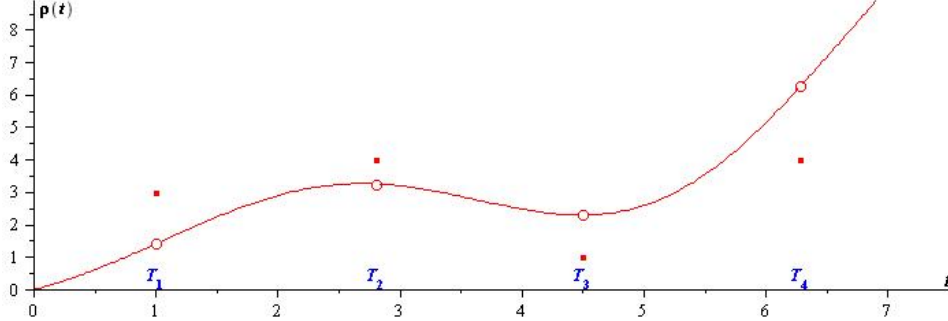


Figure 2 Graph of simulated Lévy-type process and its quadratic variation.

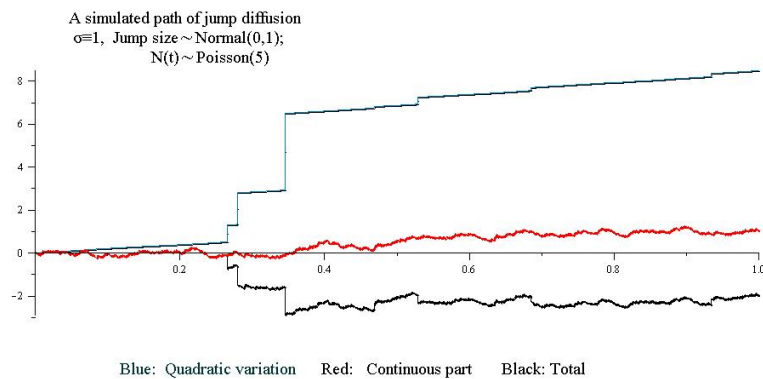


Figure 3 The result of threshold variation.

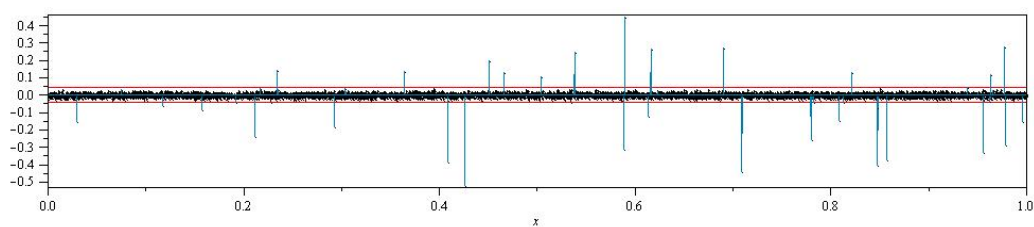
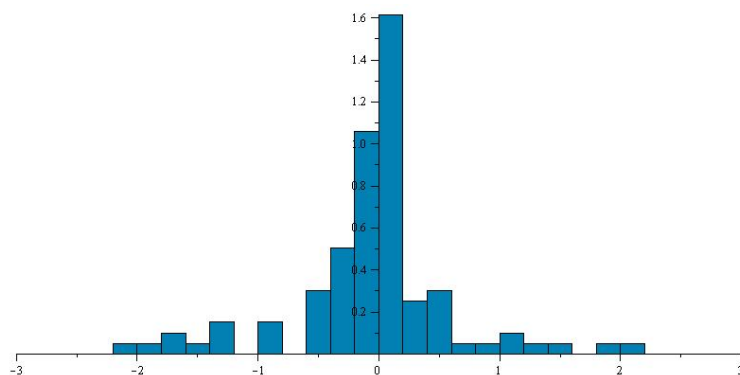


Figure 4 A histogram of $\hat{\rho} - \rho$



References.

- [1] M. SCHWERIZER (1991). Option Hedging for Semimartingales. *Stochastic Processes and their Applications*, **37**, 339-363.
- [2] AIT-SAHALIA, Y. (2000). Telling From Discrete Data Whether the Underlying Continuous-Time Model is a Diffusion. *Journal of Finance*, **57**, 2075-2112.
- [3] AIT-SAHALIA, Y. (2004). Distentangling Diffusion from Jumps. *Journal of Financial Econometrics*, **74**, 487-528.
- [4] AIT-SAHALIA, Y. P. MYKLAND AND L, ZHANG (2005). How often to Sample a Continuous Time Process in the presence of Market Microstructure Noise, *Review of Financial Studies*, **18**, 351-416.
- [5] ANDERSEN, T., T. BOLLERSLEV, AND F. DIEBOLD (2007). Roughing it Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility, *Review of Economics and Statistics*, **89**, forthcoming.
- [6] BARNDORFF-NIELSEN, O. AND N. SHEPHARD (2004). Power and Bipower Variation with Stochastic Volatility and Jumps, *Journal of Financial Econometrics*, **2**, 1-37.
- [7] BARNDORFF-NIELSEN, O. AND N. SHEPHARD (2006). Econometrics of Testing for Jumps in Financial Economics using Bipower Variation, *Journal of Financial Econometrics*, **4**, 1-30.
- [8] DACOROGNA, M., R. GENCAI, U. MULLER, R. OLSEN, AND O. PICTET (2001). *An Introduction to High Frequency Finance*, San Diego, Academic Press.
- [9] ERAKER, B. M. JOHANNES, AND N. POLSON (2003). The impact of Jumps in Volatility and Returns. *Journal of Finance*, **58**, 1269-1300
- [10] HUANG, X. AND G. TAUCHEN (2005). The Relative Contribution of Jumps to Total Price Variance. *Journal of Financial Econometrics*, **3**, 456-499.
- [11] FÖMMER, H AND SCHWEIZER, M (1991). Hedging of contingent claims under incomplete information. *Applied Stochastic Analysis*(M. H. A. Davis and R. J. Elliott, eds.) 389-414. Gordon and Breach, New York.
- [12] JACOD, J AND SHIRYAYEV, A. N. (2003). *Limit Theorems for Stochastic Processes*. Springer, New York.
- [13] KARATZAS. I. AND SHREVE, S. E. (1999). *Brownian Motion and Stochastic Calculus*. Springer, New York.
- [14] MYKLAND, P. (2006). *ANOVA for diffusions and Ito processes*. *Ann. stat.*, **34** 1931–1963.
- [15] TERENCE CHAN (1999). *Pricing contingent claims of stocks driven by Levy processes*. *The Annals of Applied Probability*, **2** 504-528.
- [16] DAVID APPLEBAUM (2004). *Levy processes and Stochastic Calculus*. Cambridge, New York.
- [17] RAMA CONT AND PETER TANKOV (2004). *Financial Modelling With Jump Processes*. CRC, London.
- [18] MANCINI, C. (2006). *Non parametric threshold estimation for models with stochastic diffusion coefficient and jumps*. *arxiv: math/0607378v1 [math. ST]*
- [19] JACOD, J. (2006). *Asymptotic properties of realized power variation and associated functions of semimartingale*. *arXiv*, 9 June 2006 n.0023146.
- [20] TIM BOLLERSLEV, TZU HANN LAW AND GEORGE TAUCHEN (2007). *Risk, Jumps, and Diversification*. Working paper.

DEPARTMENT OF MATHEMATICS
 HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY
 CLEAR WATER BAY, KOWLOON, HONG KONG
 E-MAIL: majing@ust.hk

DEPARTMENT OF MATHEMATICS
 HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY
 CLEAR WATER BAY, KOWLOON, HONGKONG
 E-MAIL: majing@ust.hk

DEPARTMENT OF MATHEMATICS
HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY
CLEAR WATER BAY, KOWLOON, HONGKONG
E-MAIL: kongxin@ust.hk