Realized Range-based Threshold Estimation for Jump-diffusion Models

Jingwei Cai, Ping Chen, Xia Mei and Xiao Ji

Abstract—**We develop a framework for estimating the quadratic variation of discontinuous semi-martingales with intra-day high-low statistics. Restricting the realized range-based variance smaller than a suitably defined threshold, we propose an integrated volatility estimator and consider its consistency and asymptotic normality under a set of weak conditions. We find that the precision of our statistics is about five times greater than that of realized variance purely restricted by threshold. Simulation results illustrate the good finite sample properties of our estimator.**

Index Terms—**central limit theorem, realized range-based volatility, threshold, discontinuous semi-martingales, integrated volatility.**

I. Introduction

THE volatility of asset price plays a central role in both the theoretical and empirical finance literature (see [1], [2], [3], among others). It is one of the most important determinants of market decision, such as derivatives pricing, risk analysis, hedging, portfolio management. Lots of researchers considered estimation of volatility in continuous diffusion process ([4], [5], [6]). Christensen and Podolskij ([7]) used the high-low technique to discuss range-based estimation of integrated volatility, and showed its precision was five times greater than that of the realized variance under a set of weak conditions. To our knowledge, there are few works which are done on range-based estimation for jump-diffusion models.

With the increasing perfection of the financial market and rapid process of the computer technology, it is a more and more easy thing to obtain high frequency financial data. In practice, these high frequency data are often sparsely sampled due to the presence of microstructure noise([8]). Unfortunately, this sparse sampling method might neglect the important intra-day information of the price movement, and will lead to loss of information and efficiency. As stated in ([9]), the range-based estimation technique need not sample sparsely and can use the whole high frequency data.

It is undeniably that high frequency returns are often inevitably with the existence of jump ([10], [11], [12]). Let $X = (X_t)_{t\geq 0}$ be a logarithmic price process of a security. This jump-diffusion semi-martingale satisfies the following generic process

$$X_t = X_0 + \int_0^t a_u du + \int_0^t \sigma_u dB_u + \sum_{i=1}^{N_t} J_i, t \in]0,1] \quad (1)$$

where $a = (a_t)_{t\geq 0}$ (the drift) is locally bounded and predictable, $\sigma = (\sigma_t)_{t\geq 0}$ (the volatility) is càdlàg, and $B = (B_t)_{t\geq 0}$ is a standard Brownian motion. Also, $N = (N_t)_{t\geq 0}$ is a Poisson process with constant intensity λ, jumping at times denoted by $(\tau_i)_{i=1,2,\cdots,N_1}$, and each J_i is the size of jump occurred at τ_i. The random variables J_i are i.i.d. and independent of N.

The inference for jump-diffusion models is increasingly becoming one of foci of attention in financial mathematics, statistics and econometrics research. In view of the central role of integrated volatility IV (defined as $\int_0^1 \sigma_t^2 dt$) in financial economics, a large body of works are done on its estimation for jump-diffusion models for the moment, such as using nearest neighbor truncation ([10]), using blocked multi-power variation ([13]), using power and bi-power variation ([14]), the recent literatures ([15], [16]), and among others. In recent years, some researchers used threshold technique to consider the inference of jump-diffusion models ([17], [18], [19], [20]). Mancini ([21]) used threshold technique to provide a consistent estimate of integrated volatility with

$$\widehat{IV_\delta} = \sum_{i=1}^{n} (\Delta_i X)^2 I_{\{(\Delta_i X)^2 \leq r(\delta)\}}, \quad (2)$$

where $\Delta_i X$ is denoted by the increment $(X_{t_i} - X_{t_{i-1}})$ and $r(\delta)$ is a deterministic function of the lag δ between two adjacent observations $(X_{t_i}, X_{t_{i-1}})$, and prove the following central limit result

$$\frac{\widehat{IV_\delta} - IV}{\sqrt{\delta}} \xrightarrow{d} N\left(0, 2\widehat{IQ_\delta}\right), \quad (3)$$

where

$$\widehat{IQ_\delta} := \frac{1}{3\delta}\sum_{i=1}^{n} (\Delta_i X)^4 I_{\{(\Delta_i X)^2 \leq r(\delta)\}} \xrightarrow{P} \int_0^1 \sigma_t^4 dt.$$

The estimator (2) reflected quite a few good finite sample properties (robustness, consistency etc) when the sampling frequency was not very high. However, we should note that the definition of the estimator (2) is based on realized method, so it may not be a good estimator when the sampling frequency is high (due to microstructure noise ([5], [22])). In this situation, people have to resample available data sparsely, then information and efficiency will be lose inevitably.

As we have mentioned, the range-based method can remedy the weakness. Can these two methods be combined then? Motivated by the works of Christensen and Podolskij ([7]) and Mancini ([21]), given a discrete record of observations,

Manuscript received January 04, 2015; revised April 26, 2015. This work was supported in part by the National Natural Science Foundation of China under Grant 11271189 and 11201229.

J.W. Cai, P. Chen and X. Ji are with Department of Statistics and Financial Mathematics, Nanjing University of Science and Technology, Nanjing, 210094, China. E-mail: jingwei_cai@126.com and prob123@mail.njust.edu.cn.

J.W. Cai and X. Mei are with Foundation department, Jiangsu Polytechnic College of Agriculture and Forestry, Zhenjiang, 212400, China. E-mail: cxk_2008@yeah.net

we propose a realized range-based threshold estimator of integrated volatility. Our statistic has both the advantages existing in the range-based method and in the threshold technique, such as the former's estimation precision and the latter's efficiency of identifying jumps.

II. Preliminaries

In high-frequency volatility estimation, the quadratic variation plays a dominant role ([7]). The quadratic variation of the process $(X_t)_{t\geq 0}$ can be denoted by

$$[X]_t = \int_0^t \sigma_u^2 du + \sum_{i=1}^{N_t} J_i^2, \quad (4)$$

which is the integrated volatility plus the sum of squared jumps. The object of our interest is its continuous part (the integrated volatility IV). For the estimation of the total quadratic variation (including the contribution from the squared jumps), interested readers may reference [14] which used the popular realized volatility measure.

In the discussions of the asymptotic normality, we'll use the following regularity conditions of the volatility processσ:
Assumption(T). σ does not vanish ($T1$) and satisfies

$$\sigma_t = \sigma_0 + \int_0^t \mu'_u du + \int_0^t \sigma'_u dB_u + \int_0^t \nu'_u dB'_u, \quad (T2)$$

where $\mu' = (\mu'_t)_{t\geq 0}$ is locally bounded and predictable, $\sigma' = (\sigma'_t)_{t\geq 0}$ and $\nu' = (\nu'_t)_{t\geq 0}$ are càdlàg, and $B' = (B'_t)_{t\geq 0}$ is a Brownian motion independent of B.

The work is founded in a high-frequency record of X with n discretized observations which are supposed to be available at time $t_i, i = 0, 1, \cdots, n$.

Define the intra-day range at sampling times t_{i-1} and t_i as

$$s_{X_{t_i,\delta_i}} = \sup_{t_{i-1}\leq s,t\leq t_i} \{X_t - X_s\}\,. \quad (5)$$

In the same way, the range of a standard Brownian motion over $[t_{i-1}, t_i]$ may be defined as

$$s_{B_{t_i,\delta_i}} = \sup_{t_{i-1}\leq s,t\leq t_i} \{B_t - B_s\}\,. \quad (6)$$

For a scaled Brownian motion, $X_t = \sigma B_t$, the rth moment of its range can be denoted by [23]

$$E[s^r_{X_{t_i,\delta_i}}] = \lambda_r \delta_i^{r/2} \sigma^r (r \geq 1), \quad (7)$$

where $\lambda_r = E[s^r_{B_{1,1}}]$.

when the process X (B) is observed discretely at equidistant time points $\{t = t_1, t_2, \cdots, t_n\}$ with $\delta = 1/n = t_i - t_{i-1} (i = 1, 2, \cdots, n)$ which is a time distance between two consecutive observations, we abbreviate the range $s_{X_{t_i,\delta_i}}$ to $s_{X_{t_i}}$.

III. Main results

Theorem 1 Suppose that $\sum_{i=1}^{N_t} J_i$ is a finite activity jump process, where N is a non-explosive counting process and the random variable J_i satisfy, $\forall t \in [0, 1], P\{\Delta N_t \neq 0, J_{N_t} = 0\} = 0$. $r(\delta)$ is a deterministic function of the lag between two adjacent observations $(X_{t_i}, X_{t_{i-1}})$, such that $\lim_{\delta\to 0} r(\delta) = 0$ and $\lim_{\delta\to 0} (\delta \log \frac{1}{\delta})/r(\delta) = 0$. Then for P-almost all ω, $\exists \bar{\delta}(\omega) > 0$ such that $\forall \delta \leq \bar{\delta}(\omega)$ we have

$$I_{\{s^2_{X_{t_i}} \leq r(\delta)\}}(\omega) = I_{\{\Delta_i N = 0\}}(\omega) (\forall i = 1, ..., n). \quad (8)$$

Proof Without loss of generality, we suppose the drift function a and the diffusion function σ are both bounded and the partition is equidistant in the context ([7], [24]). In order to prove the theorem, it is enough to prove that both $\forall i \in \{1, ..., n\}$, $I_{\{s^2_{X_{t_i}} \leq r(\delta)\}}(\omega) \geq I_{\{\Delta_i N=0\}}(\omega)$ and $\forall i \in \{1, ..., n\}$, $I_{\{s^2_{X_{t_i}} \leq r(\delta)\}}(\omega) \leq I_{\{\Delta_i N=0\}}(\omega)$ hold.

(1) Let

$$J_{0,\delta} = \{i \in \{1, ..., n\} : \Delta_i N = 0\},$$

to prove that, for $\forall i \in \{1, ..., n\}$ and small δ,

$$I_{\{s^2_{X_{t_i}} \leq r(\delta)\}}(\omega) \geq I_{\{\Delta_i N=0\}}(\omega),$$

it is sufficient to show that, for $\forall i \in \{1, ..., n\}$ and small δ,

$$\sup_{J_{0,\delta}} s^2_{X_{t_i}} \leq r(\delta).$$

Notice that, as $\delta \to 0$,

$$\begin{aligned} &\sup_{J_{0,\delta}} \frac{s^2_{X_{t_i}}}{r(\delta)} \\ &= \sup_{J_{0,\delta}} \left(\frac{\sup\limits_{t_{i-1}\leq s,t\leq t_i} \left| \int_s^t a_u du + \int_s^t \sigma_u dB_u \right|}{\sqrt{\delta \log \frac{1}{\delta}}} \right)^2 \cdot \frac{\delta \log \frac{1}{\delta}}{r(\delta)} \end{aligned}$$

and

$$\begin{aligned} &\sup_{J_{0,\delta}} \frac{\sup\limits_{t_{i-1}\leq s,t\leq t_i} \left| \int_s^t a_u du + \int_s^t \sigma_u dB_u \right|}{\sqrt{\delta \log \frac{1}{\delta}}} \\ &\leq \sup_{J_{0,\delta}} \frac{\sup\limits_{t_{i-1}\leq s,t\leq t_i} \left| \int_s^t a_u du \right|}{\sqrt{\delta \log \frac{1}{\delta}}} \\ &+ \sup_{J_{0,\delta}} \frac{\sup\limits_{t_{i-1}\leq s,t\leq t_i} \left| \int_s^t \sigma_u dB_u \right|}{\sqrt{\delta \log \frac{1}{\delta}}} \end{aligned}$$

obviously,

$$\sup_{J_{0,\delta}} \frac{\sup\limits_{t_{i-1}\leq s,t\leq t_i} \left| \int_s^t a_u du \right|}{\sqrt{\delta \log \frac{1}{\delta}}} \to 0,$$

for the second term of the above inequality, using Burkholder-Davis-Gundy inequality ([25]), we have

$$\sup_{J_{0,\delta}} E\left[\frac{\sup\limits_{t_{i-1}\le s,t\le t_i}\left|\int\limits_s^t \sigma_u dB_u\right|}{\sqrt{2\delta\log\frac{1}{\delta}}}\right] \le C\cdot\sup_{J_{0,\delta}} E\left[\sqrt{\frac{\int\limits_{t_{i-1}}^{t_i}\sigma_u^2 du}{2\delta\log\frac{1}{\delta}}}\right] = o(1),$$

where C is a positive constant.
Obviously, for small δ,

$$\sup_{J_{0,\delta}} \frac{s^2_{X_{t_i}}}{r(\delta)} \to 0,$$

further,

$$\sup_{J_{0,\delta}} s^2_{X_{t_i}} \le r(\delta).$$

(2) Let

$$J_{1,\delta} = \{i \in \{1,...,n\} : \Delta_i N \neq 0\},$$

to prove that, for $\forall i \in \{1,...,n\}$ and small δ,

$$I_{\{s^2_{X_{t_i}} \le r(\delta)\}}(\omega) \le I_{\{\Delta_i N=0\}}(\omega),$$

it is sufficient to show that, for $\forall i \in \{1,...,n\}$ and small δ,

$$\inf_{i\in J_{1,\delta}} s^2_{X_{t_i}} \ge r(\delta).$$

We know form the proof of Theorem 1 in [21] that

$$\lim_{\delta}\inf_{i\in J_{1,\delta}} \frac{(\Delta_i X)^2}{r(\delta)} = +\infty,$$

so, for $\forall i \in J_{1,\delta}$,

$$\frac{s^2_{X_{t_i}}}{r(\delta)} \ge \frac{(\Delta_i X)^2}{r(\delta)} \to +\infty,$$

and then

$$I_{\{s^2_{X_{t_i}} \le r(\delta)\}}(\omega) \le I_{\{\Delta_i N=0\}}(\omega).$$

These complete the proof of the theorem.

Define

$$\widehat{RIV} = \frac{1}{\lambda_2}\sum_{i=1}^n s^2_{X_{t_i}} I_{\{s^2_{X_{t_i}} \le r(\delta)\}}. \tag{9}$$

Theorem 2 Under the assumptions in Theorem 1, we have, as $n\to\infty$,

$$P\lim_{\delta\to 0}\widehat{RIV} = IV. \tag{10}$$

Proof By the define of $\widehat{RIV}$,

$$\begin{aligned}
&P\lim_{\delta\to 0}\frac{1}{\lambda_2}\sum_{i=1}^n s^2_{X_{t_i}} I_{\{s^2_{X_{t_i}} \le r(\delta)\}}\\
&= \frac{1}{\lambda_2}P\lim_{\delta\to 0}\sum_{i=1}^n s^2_{X_{t_i}} I_{\{\Delta_i N=0\}}\\
&= \frac{1}{\lambda_2}P\lim_{\delta\to 0}\sum_{i=1}^n (\sup_{t_{i-1}\le s,t\le t_i}(\int_s^t a_u du + \int_s^t \sigma_u dB_u))^2\\
&- \frac{1}{\lambda_2}P\lim_{\delta\to 0}\sum_{i=1}^n (\sup_{t_{i-1}\le s,t\le t_i}(\int_s^t a_u du + \int_s^t \sigma_u dB_u))^2\\
&\cdot I_{\{\Delta_i N\neq 0\}}.
\end{aligned}$$

For the second item, as $\delta\to 0$,

$$\begin{aligned}
&\frac{1}{\lambda_2}\sum_{i=1}^n (\sup_{t_{i-1}\le s,t\le t_i}(\int_s^t a_u du + \int_s^t \sigma_u dB_u))^2 \cdot I_{\{\Delta_i N\neq 0\}}\\
&\le \frac{1}{\lambda_2}N_T\sup_i\left(\sup_{t_{i-1}\le s,t\le t_i}\left(\int_s^t a_u du + \int_s^t \sigma_u dB_u\right)\right)^2\\
&\to 0.
\end{aligned}$$

By Theorem 1 in [7], we have

$$\frac{1}{\lambda_2}\sum_{i=1}^n (\sup_{t_{i-1}\le s,t\le t_i}(\int_s^t a_u du + \int_s^t \sigma_u dB_u))^2 \to IV,$$

then the proof is complete.

Remark 1 For the case of unequally spaced observations, the lag $\Delta t_i := t_i - t_{i-1}$, between the observations $\{X_{t_0}, X_{t_1}, \cdots, X_{t_n}\}$ is not constant, then defining $\delta := \max\limits_i \Delta t_i$, Theorem 1 and Theorem 2 are still valid.

Theorem 3 Under the assumptions in Theorem 1, if a and $\sigma \neq 0$ are càdlàg process, as $\delta\to 0$, we have

$$\frac{\widehat{RIV} - IV}{\sqrt{\delta}} \xrightarrow{d} MN\left(0, \Lambda\int_0^1 \sigma_t^4 dt\right), \tag{11}$$

where $\Lambda = \frac{\lambda_4 - \lambda_2^2}{\lambda_2^2} \simeq 0.4073$.

Proof Using the result of Theorem 1, we know

$$\begin{aligned}
\frac{\widehat{RIV} - IV}{\sqrt{\delta}} &= \frac{\frac{1}{\lambda_2}\sum\limits_{i=1}^n s^2_{X_{t_i}} I_{\{\Delta_i N=0\}} - IV}{\sqrt{\delta}}\\
&= \frac{\frac{1}{\lambda_2}\sum\limits_{i=1}^n (\sup\limits_{t_{i-1}\le s,t\le t_i}(\int\limits_s^t a_u du + \int\limits_s^t \sigma_u dB_u))^2 - IV}{\sqrt{\delta}}\\
&- \frac{\frac{1}{\lambda_2}\sum\limits_{i=1}^n (\sup\limits_{t_{i-1}\le s,t\le t_i}(\int\limits_s^t a_u du + \int\limits_s^t \sigma_u dB_u))^2 I_{\{\Delta_i N\neq 0\}}}{\sqrt{\delta}}.
\end{aligned}$$

By the proof of Theorem 1, we know

$$
\begin{aligned}
&\frac{\frac{1}{\lambda_2}\sum\limits_{i=1}^{n}(\sup\limits_{t_{i-1}\le s,t\le t_i}(\int\limits_s^t a_u du+\int\limits_s^t \sigma_u dB_u))^2\cdot I_{\{\Delta_i N\neq 0\}}}{\sqrt{\delta}}\\
&\le \frac{\frac{1}{\lambda_2}N_T\sup_i\left(\sup\limits_{t_{i-1}\le s,t\le t_i}\left(\int\limits_s^t a_u du+\int\limits_s^t \sigma_u dB_u\right)\right)^2}{\sqrt{\delta}}\\
&\le \frac{2N_T}{\lambda_2\sqrt{\delta}}\sup_i\left(\sup_{t_{i-1}\le s,t\le t_i}\int\limits_s^t a_u du\right)^2\\
&+\frac{2N_T}{\lambda_2\sqrt{\delta}}\sup_i\left(\sup_{t_{i-1}\le s,t\le t_i}\int\limits_s^t \sigma_u dB_u\right)^2\\
&=R_1+R_2.
\end{aligned}
$$

Obviously, $R_1 = o(1)$. For the latter term, we exploit Burkholder-Davis-Gundy inequality:

$$
E[R_2]\le\frac{2CN_T}{\lambda_2\sqrt{\delta}}E[\sup_i\int\limits_{t_{i-1}}^{t_i}\sigma_u^2du]=o(1),
$$

where C is a positive constant.
Using the result of Theorem 2 in [7], we know

$$
\begin{aligned}
&\frac{\frac{1}{\lambda_2}\sum\limits_{i=1}^{n}(\sup\limits_{t_{i-1}\le s,t\le t_i}(\int\limits_s^t a_u du+\int\limits_s^t \sigma_u dB_u))^2-IV}{\sqrt{\delta}}\\
&\xrightarrow{d} MN\left(0,\Lambda\int\limits_0^T\sigma_t^4dt\right),
\end{aligned}
$$

and the proof is complete.

Theorem 4 Under the assumptions in Theorem 3, we have

$$
\widehat{RIQ}:=\frac{1}{\lambda_4\delta}\sum_{i=1}^{n}s_{X_t}^4 I_{\{s_{X_t}^2\le r(\delta)\}}\xrightarrow{P}\int\limits_0^1\sigma_t^4dt. \tag{12}
$$

Proof Using the result of Theorem 1, as $\delta\to 0$,

$$
\begin{aligned}
&P\lim_{\delta\to 0}\frac{1}{\lambda_4\delta}\sum_{i=1}^{n}s_{X_t}^4 I_{\{s_{X_t}^2\le r(\delta)\}}\\
&=P\lim_{\delta\to 0}\frac{1}{\lambda_4\delta}\sum_{i=1}^{n}s_{X_t}^4 I_{\{\Delta_i N=0\}}\\
&=P\lim_{\delta\to 0}\frac{1}{\lambda_4\delta}\sum_{i=1}^{n}(\sup_{t_{i-1}\le s,t\le t_i}(\int\limits_s^t a_u du+\int\limits_s^t \sigma_u dB_u))^4\\
&-P\lim_{\delta\to 0}\frac{1}{\lambda_4\delta}\sum_{i=1}^{n}(\sup_{t_{i-1}\le s,t\le t_i}(\int\limits_s^t a_u du+\int\limits_s^t \sigma_u dB_u))^4\\
&\cdot I_{\{\Delta_i N\neq 0\}}\\
&=R_3+R_4.
\end{aligned}
$$

The second term of the last equality above is dominated by

$$
\begin{aligned}
&P\lim_{\delta\to 0}\frac{N_1}{\lambda_4\delta}\sup_i\left(\sup_{t_{i-1}\le s,t\le t_i}\left(\int\limits_s^t a_u du+\int\limits_s^t \sigma_u dB_u\right)\right)^4\\
&=0.
\end{aligned}
$$

Using the result in [7], we have

$$
R_3=\int\limits_0^1\sigma_t^4dt.
$$

Corollary Assume that the conditions of Theorem 3 hold, then it holds that, as $\delta\to 0$,

$$
\frac{\widehat{RIV}-IV}{\sqrt{\delta\Lambda\widehat{RIQ}}}\xrightarrow{d}N(0,1), \tag{13}
$$

where $\Lambda=\frac{\lambda_4-\lambda_2^2}{\lambda_2^2}\simeq 0.4073$.

Proof It is a obvious result since Theorem 3 and Theorem 4 are valid.

Remark 2 The Λ scalar in front of $\widehat{RIQ}$ in Eq. (13) is roughly 0.4. In contrast, the number appearing in the CLT for $\widehat{IV}_\delta$ in Eq. (3) is 2.

In fact, continuous sampling is just an ideal consideration. In practice, it is impossible to sample continuously and to extract the true range, so the inference about IV will be from a finite sample. In order to scale properly, the number of high-frequency data used in forming the high-low should be accounted for. We follow the approach in ([7]) by using discretely sampled high-frequency data.

Just like the practice in ([7]), We now choose m observations in each segment of n intervals, then it brings about mn returns. The observed range over the ith interval is defined as

$$
S_{X_{t_i},m}=\max_{0\le s,t\le m}\{X_{(i-1)/n+t/mn}-X_{(i-1)/n+s/mn}\}, \tag{14}
$$

where s and t are integers.
We also let

$$
\mathrm{s}_{B,m}=\max_{0\le s,t\le m}\{B_{t/m}-B_{s/m}\}. \tag{15}
$$

And then we propose a new realized range-based estimator with discretely sampled high-frequency data by setting

$$
\widehat{RIV}_m=\frac{1}{\lambda_{2,m}}\sum_{i=1}^{n}s_{X_{t_i},m}^2 I_{\{s_{X_{t_i},m}^2\le r(\delta)\}}. \tag{16}
$$

where $\lambda_{r,m}=E[s_{B,m}^r]$, $\lambda_{r,m}$ is the rth moment of the range of a standard Brownian motion over an unit interval. It is obvious that $\lambda_{r,m}\to\lambda_r$ as $m\to\infty$.

Theorem 5 Under the assumptions of Theorem 1, we have, as $n\to\infty$,

$$
P\lim_{\delta\to 0}\widehat{RIV}_m=IV, \tag{17}
$$

where the convergence is uniform in m. Additionally, if a and $\sigma\neq 0$ are càdlàg processes and $m\to c\in N\cup\{\infty\}$, as

$\delta \to 0$,

$$\frac{\widehat{RIV}_m - IV}{\sqrt{\delta \Lambda_m \widehat{RIQ}_m}} \xrightarrow{d} N(0,1), \tag{18}$$

with $\Lambda_m = \left(\lambda_{4,m} - \lambda_{2,m}^2\right)/\lambda_{2,m}^2$ and

$$\widehat{RIQ}_m := \frac{1}{\lambda_{4,m}\delta} \sum_{i=1}^{n} s_{X_{t_i},m}^4 I_{\{s_{X_{t_i},m}^2 \le r(\delta)\}}. \tag{19}$$

Remark 3 When $m = 1$, Theorem 5 provides a CLT for $\widehat{IV}_\delta$ (derived in ([21])) since $\lambda_{2,1} = 1$ and $\lambda_{4,1} = 3$.

Remark 4 Our realized range-based threshold estimator $\widehat{RIV}_m$ may be used in any high frequency data since m can tend to infinite.

IV. Numerical simulation

In this section, we report simulation results documenting the finite sample performance of the realized range-based threshold estimator $\widehat{RIV}_m(m = 2, m = 5)$, and compare accuracy of the estimator to $\widehat{IV}_\delta$ (In fact, the estimator is a special case of our realized range-based threshold estimator $\widehat{RIV}_m$ with $m = 1$) in ([21]) obtained by using the pure threshold method.

The process we will consider is a jump-diffusion model

$$X_t = \sigma W_t + \sum_{j=1}^{N_t} Z_j (t \in [0,1]), \tag{20}$$

with Z_j $i.i.d.$ with law $N(0, \eta^2)$, where $\eta = 0.6, \sigma = 0.3$ and $\lambda = 5$, just as in ([21], [26]). We use an observation length of $T = 1$ day, consisting of 5 hours of trading (i.e., 18000 seconds). For the model, we generate $N = 5000$ sample paths. In each sample path, using the algorithms described in ([27]), we take $mn = 1500$ (6000) equally spaced observations with lag $\delta = 1/n$, namely, we will choose an observation every 12 (3) seconds.

From the equation (3) and (18), we know that the limit distribution of normalized bias term

$$\frac{\widehat{IV}_\delta - IV}{\sqrt{2\delta \widehat{IQ}_\delta}} \tag{21}$$

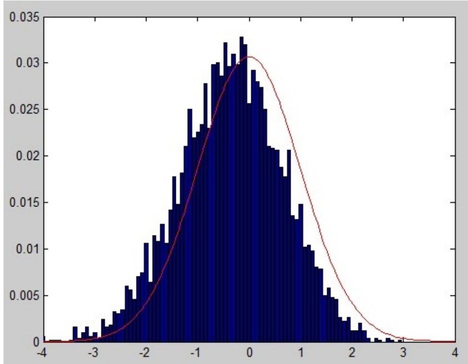
and normalized bias term

$$\frac{\widehat{RIV}_m - IV}{\sqrt{\delta \Lambda_m \widehat{RIQ}_m}} \tag{22}$$

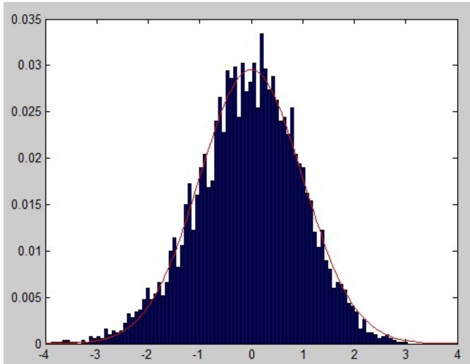
are both standard normal distribution. Next, we will use figure and table to illustrate accuracy of the two normalized bias terms.

In Figure 1 and 2, using the observations simulated for the model (20) in 5000 sample paths (in Figure 1 each sample path generates 1500 observations, while the number of observations in Figure 2 is 6000), we plot the histogram of the 5000 normalized bias terms for the threshold estimator $\widehat{IV}_\delta$ (remembering it is a special case of $\widehat{RIV}_m$ with $m = 1$) as in (21) and for the range-based threshold estimator $\widehat{RIV}_m(m = 2, 5)$ as in (22) respectively.

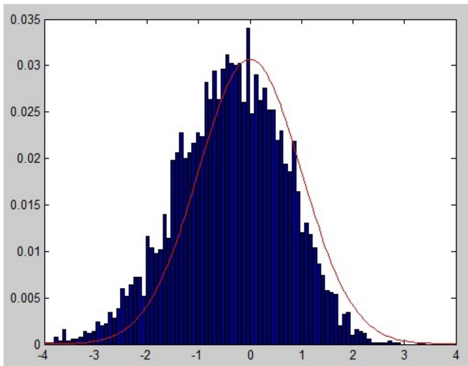
Figure1 and Table I illustrate that, in the three estimators, the estimator $\widehat{RIV}_2$ has highest accuracy when the number of the observations is 1500. Similarly, from Figure 2 and



(a)



(b)



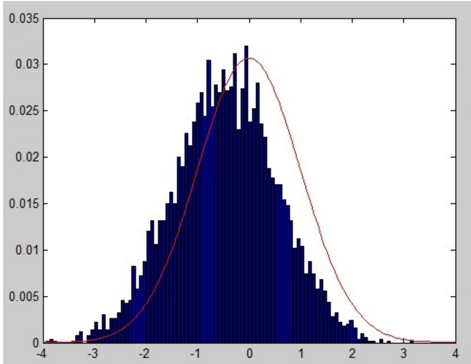
(c)

Fig. 1. Histograms of 5000 normalized bias terms with 1500 observations in each sample path. From up to down: it is for the realized threshold estimator $\widehat{IV}_\delta$, for the realized range-based threshold estimator $\widehat{RIV}_2$ and for the realized range-based threshold estimator $\widehat{RIV}_5$.

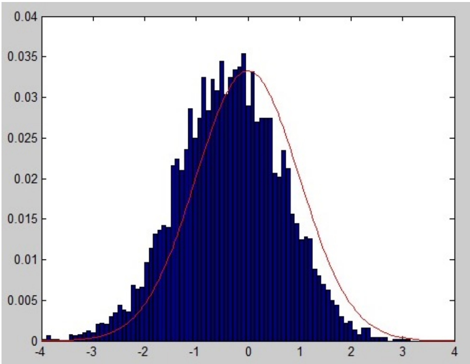
TABLE I

Statistics of the normalized bias terms under the model (20) with 1500 observations in each sample path. $m = 1$ denotes the normalized bias term (21), $m = 2$ and $m = 5$ denote the normalized bias term (22) with corresponding $m = 2$ and $m = 5$ respectively. Pct is the percentage of the 5000 realizations for which the normalized bias is in absolute value larger than 1.96 (asymptotically such a percentage has to be 0.05). Mean and SD are the mean and the standard deviation of the 5000 values assumed by each normalized bias term (asymptotically such Mean and SD have to be 0 and 1).

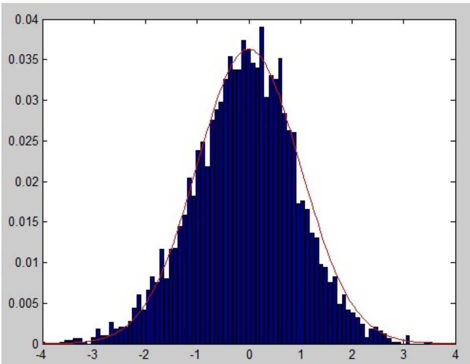
	Pct	Mean	SD
$m = 1$	0.0724	-0.3588	1.0382
$m = 2$	0.0566	-0.0541	1.0294
$m = 5$	0.0822	-0.3733	1.1487



(a)



(b)



(c)

Fig. 2. Histograms of 5000 normalized bias terms with 6000 observations in each sample path. See Fig. 1 for explanation.

TABLE II
STATISTICS OF THE NORMALIZED BIAS TERMS UNDER THE MODEL (20) WITH 6000 OBSERVATIONS IN EACH SAMPLE PATH. SEE TABLE I FOR EXPLANATION.

	Pct	Mean	SD
$m = 1$	0.0764	-0.4569	1.0267
$m = 2$	0.0664	-0.3563	1.0187
$m = 5$	0.0562	-0.0779	1.0171

Table II, we can obtain that the estimator $\widehat{RIV}_5$ is the best when 6000 observations are available in a trading day.

Remark 5 From the numerical simulation, we find that, for the estimator $\widehat{RIV}_m$, appropriate m value should be chosen for different frequency financial data. Exploring the relation between the choice of m value and the frequency of financial data will be one of issues of our future research.

V. CONCLUSIONS

In this paper, we consider an integrated volatility estimation procedure for jump-diffusion models. We are inspired by the precision of the range-based technique and the efficiency of the threshold method. Combining the advantages of the two methods, we propose a realized range-based threshold estimator for the volatility of jump-diffusion models. The simulation results illustrate that our estimator is more accurate than pure threshold estimator.

ACKNOWLEDGEMENT

The authors would like to thank the anonymous reviewers for their helpful comments and suggestions that largely improve the paper.

REFERENCES

[1] J. Jacod and M. Reiss, "A Remark on the Rates of Convergence for Integrated Volatility Estimation in the Presence of Jumps," *The Annas of Statistics*, Vol.42, no. 3, pp. 1131-1144, 2014.
[2] T. Busch, B.J. Christensen and M.Ø. Nielsen, "The Role of Implied Volatility in Forecasting Future Realized Volatility and Jumps in Foreign Exchange, Stock, and Bornd Markets," *Journal of Econometrics*, Vol. 160, pp. 48-57, 2011.
[3] G. Liu, Q. Zhao and G.D. Gu, "A Simple Control Variate Method for Options Pricing with Stochastic Volatility Models," *IAENG International Journal of Applied Mathematics*, vol. 45, no. 1, pp. 64-70, 2015.
[4] J.Q. Fan and Y.Z. Wang, "Spot Volatility Estimation for High-frequency Data," *Statistics and Its Interface*, Vol. 1, pp. 279-288, 2008.
[5] L. Zhang, P. Mykland and Y. Aït-Sahalia, "A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-frequency Data," *Journal of the American Statistical Association*, Vol. 100, no. 472, pp. 1394-1411, 2005.
[6] F.M. Bandi and P.C.B. Phillips, "Fully Nonparametric Estimation of Scalar Diffusion Models," *Econometrica*, Vol. 71, no. 1, pp. 241-283, 2003.
[7] K. Christensen and M. Podolskij, "Realized Range-based Estimation of Integrated Variance," *Journal of Econometrics*, Vol. 141, no. 2, pp. 323-349, 2007.
[8] K. Christensen and M. Podolskij, "Asymptotic Theory of Range-based Multipower Variation," *Journal of Financial Econometrics*, Vol. 10, no. 3, pp. 417-456, 2012.
[9] H.Q. Li and Y.M. Hong, "Financial Volatility Forecasting with Range-based Autoregressive Volatility Model," *Finance Research Letters*, Vol. 8, no. 2, pp. 69-76, 2011.
[10] T.G. Andersen, D. Dobrev and E. Schaumburg, "Jump-robust Volatility Estimation Using Nearest Neighbor Truncation," *Journal of Econometrics*, Vol. 169, pp. 75-93, 2012.
[11] S.P. Sidorov, A. Revutskiy, A. Faizliev, E. Korobov and V. Balash, "Stock Volatility Modelling with Augmented GARCH Model with Jumps," *IAENG International Journal of Applied Mathematics*, vol. 44, no. 4, pp. 212-220, 2014.
[12] K. Boudt, C. Croux and S. Laurent, "Robust Estimation of Intraweek Periodicity in Volatility and Jump Detection," *Journal of Empirical Finance*, Vol. 18, pp. 353-367, 2011.
[13] P.A. Mykland, N. Shephard and K. Sheppard, "Efficient and Feasible Inference for the Components of Financial Variation Using Blocked Multipower Variation," *Working Paper*, 2012.
[14] O.E. Barndorff-Nielsen and N. Shephard, "Power and Bipower Variation with Stochastic Volatility and Jumps," *Journal of Financial Econometrics*, Vol. 2, no. 1, pp. 1-37, 2004.
[15] K. Christensen, R.C.A. Oomen and M. Podolskij, "Fact or Friction: Jumps at Ultra High Frequency," *Journal of Financial Economics*, Vol. 114, no. 3, pp. 576-599, 2014.
[16] M. Bibinger and L. Winkelmann, "Econometrics of Co-jumps in High-frequency Data with Noise," *Journal of Econometrics*, Vol. 184, no. 2, pp. 361-378, 2015.
[17] C. Mancini and R. Renò, "Threshold Estimation of Jump-diffusion Models and Interest Rate Modeling," *Journal of Econometrics*, Vol. 160, pp. 77-92, 2011.
[18] C. Fulvio, P. Davide and R. Roberto, "Threshold Bipower Variation and the Impact of Jumps on Volatility Forecasting," *Journal of Econometrics*, Vol. 159, pp. 276-288, 2010.
[19] J.E. Figueroa-López and J. Nisen, "Optimally Thresholded Realized Power Variations for Lévy Jump Diffusion Models," *Stochastic Processes and their Applications*, Vol. 123, no. 7, pp. 2648-2677, 2013.

[20] I.R. Mitric, K.P. Sendovaa and C.C.L. Tsai, “On a Multi-threshold Compound Poisson Process Perturbed by Diffusion,” *Statistics and Probability Letters*, Vol. 80, pp. 366-375, 2010.
[21] C. Mancini, “Non-parametric Threshold Estimation for Models with Stochastic Diffusion Coefficient and Jumps,” *Scandinavian Journal of Statistics*, Vol. 36, pp. 270-296, 2009.
[22] J. Fleming and B.S. Paye, “High-frequency returns, jumps and the mixture of normals hypothesis,” *Journal of Econometrics*, Vol. 160, pp. 119-128, 2011.
[23] M. Parkinson, “The Extreme Value Method for Estimating the Variance of the Rate of Return,” *Journal of Business*, Vol. 53, no. 1, pp. 61-65, 1980.
[24] O.E. Barndorff-Nielsen, S.E. Graversen, J. Jacod, M. podolskij and N. Shephard, “A Central Limit Theorem for Realized Power and Bipower Variations of Contiuous Semimartingales: the Shiryaev Festschrift,” *From Stochastic Calculus to Mathematical Finance*, Eds. Y. Kabanov, R. Lipster and J. Stoyanov, Bernin: Springer-Verlag, pp. 33-68, 2006.
[25] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, 3rd ed. Bernin: Springer-Verlag, 1998.
[26] Y. Aït-Sahalia, “ Disentangling volatility from jumps,” *Journal of Financial Economics*, Vol. 74, pp. 487-528, 2004.
[27] P. Glasserman, *Monte Carlo methods in financial engineering*, New York: Springer-Verlag, 2004.