Discussion Paper Series 2008-06

Center for the Study of Finance and Insurance, Osaka University

REALIZED VOLATILITY WITH STOCHASTIC SAMPLING

MASAAKI FUKASAWA

September 24, 2008

REALIZED VOLATILITY WITH STOCHASTIC SAMPLING

MASAAKI FUKASAWA

Abstract. A central limit theorem for the realized volatility based on a general stochastic sampling scheme is proved. The asymptotic distribution depends on the sampling scheme, which is written explicitly in terms of the asymptotic skewness and kurtosis of returns. The conditions for the central limit theorem to hold are examined for several concrete examples of schemes. A lower bound for mean-squared error is attained by a specific sampling scheme. More efficient sampling schemes for the Euler-Maruyama approximation than the usual equidistant scheme are given as an application.

1. Introduction

The realized volatility, which is defined as the sum of squared log-returns, is a popular statistic in the context of high-frequency data analysis. As is well-known, if we assume an asset log-price process is a continuous semimartingale, then the realized volatility is a consistent estimator of the quadratic variation of the semimartingale as the sampling frequency of price data goes to infinity. This consistency holds in a general nonparametric setting, even if sampling scheme is stochastic. Here, by a sampling scheme, we mean a sequence of increasing stopping times $\tau = \{\tau_j\}$ with

$$0 = \tau_0 < \tau_1 < \cdots < \tau_j < \cdots .$$

We suppose that available price data are given as (τ_j, X_{τ_j}) for $j = 0, 1, \dots N_T$ for a stopping time T, where $X = \{X_t\}$ is the log-price process and

$$N_T = \max\{j \geq 0; \tau_j \leq T\}$$

is the number of data which are obtained in the time interval $[0, T]$. For tick data, which was called ultra-high-frequency data by Engle [8], we can consider τ_j to be transaction times or quote-revision times. In such a case, as far as considering a stock with liquidity, the durations $\tau_{j+1} - \tau_j$ are very small, so that we can expect the realized volatility

$$\mathcal{R}v[\tau]_T = \sum_{j=0}^{N_T - 1} \left(X_{\tau_{j+1}} - X_{\tau_j}\right)^2$$

is a reliable estimator of the quadratic variation $\langle X \rangle_T$. The asymptotic theory of

2000 *Mathematics Subject Classification.* 62M05,60J60,91B28.
Key words and phrases. quadratic variation; realized volatility; stable convergence.
Center for the Study of Finance and Insurance, Osaka University.

the realized volatility and related statistics has been well-developed for the equidistant sampling case $\tau_j = j/n$ with $n \to \infty$; Jacod and Protter [19], Barndorff-Nielsen and Shephard [4], Barndorff-Nielsen et.al [3] among others. Deterministic but non-equidistant cases were treated in Mykland and Zhang [27], Barndorff-Nielsen and Shephard [5]. The author studied a specific random sampling scheme in Fukasawa [9, 10]. Very recently, a class of random sampling schemes is studied by Hayashi, Jacod and Yoshida [12]. Our aim in this article is to treat more general sampling schemes. We extend a result of the author's previous Japanese paper [11], which dealt with only the Black-Scholes model.

The importance to incorporate the information content of the durations into the estimation of the volatility has been recognized through empirical studies with GARCH type modeling such as Engle [8]. For example, variances of returns are found to be negatively influenced by long durations between trades. Nevertheless, there have been few studies on the asymptotic behavior of the realized volatility based on stochastic sampling schemes. We will show that the asymptotic distribution is determined by asymptotic skewness and kurtosis of returns; this simple fact was not recognized until Fukasawa [11]. In particular, we will see that negative relation between variance and duration reduces mean-squared error of the realized volatility.

Due to market microstructure noise, we encounter another problem; we have to design a subsampling scheme that is robust and efficient. Zhang, Mykland and Aït-Sahalia [32], Zhang [31] proposed two scales and multi-scale realized volatility in the case that the sampling times are deterministic. Oomen [29], Griffin and Oomen [14] found that tick time sampling, where prices are sampled with every price change, is superior to calendar time sampling, that is, the equidistant sampling scheme, in terms of mean-squared error. Although they exploited a simple pure jump process model, a similar assertion was validated in the usual semimartingale setting by Fukasawa [10]. The present article proves a central limit theorem for the realized volatility with a general stochastic sampling scheme. We do not assume any parametric form of sampling scheme nor such a measurability condition that was imposed in Jacod [17], Hayashi, Jacod and Yoshida [12]. Apart from microstructure noise, the realized volatility is found to be possibly biased if sampling scheme is path-dependent. We introduce models for tick time sampling, one of which is shown to be most efficient among all bias-free sampling schemes in terms of mean-squared error. Note that the scheme enjoys also robustness to market microstructure noise in a specific microstructure model which incorporates bid-ask bounce and price discreteness as shown by Fukasawa [10].

In Section 2, we recall stable convergence and describe a more or less known fundamental results on this topic. Section 3 presents the main result of this article. In Section 4, we consider several concrete examples of sampling schemes including calendar time sampling, business time sampling and tick time sampling. Section 5 treats efficiency problem. As an application, we propose alternative sampling schemes for the Euler-Maruyama approximation in Section 6.

2. Stable Convergence

In this section, we recall the definition of stable convergence and describe a well-known fundamental theorem which plays an essential role in the next section. Let

E be a complete separable metric space and $(\Omega, \mathcal{F}, P)$ be an probability space on which a sequence of E-valued random variables $\{Z^n\}$ is defined.

Definition 1. *For a sub σ-field $\mathcal{G} \subset \mathcal{F}$, we say $\{Z^n\}$ converges $\mathcal{G}$-stably if for all $\mathcal{G}$-measurable random variable Y, the joint distribution (Z^n, Y) converges in law.*

The following limit theorem is a simplified version of a result of Jacod [17, 18] and Jacod and Shiryaev [20], Theorem IX.7.3, which extends a result of Rootzén [30]. Let $M = \{M_t, \mathcal{F}_t, 0 \le t < \infty\}$ be a continuous local martingale defined on $(\Omega, \mathcal{F}, P)$ and $\mathcal{M}^\perp$ be the set of bounded $\{\mathcal{F}_t\}$-martingales orthogonal to M.

Theorem 1. *Let $\{Z^n\}$ be a sequence of continuous $\{\mathcal{F}_t\}$-local martingales. Suppose that there exist an $\{\mathcal{F}_t\}$-adapted process $V = \{V_t\}$ such that for all $N \in \mathcal{M}^\perp$, $t \in [0, \infty)$,*

$$\langle Z^n, N\rangle_t \to 0, \quad \langle Z^n, M\rangle_t \to 0, \quad \langle Z^n\rangle_t \to V_t$$

in probability. Then, the $C[0,\infty)$-valued sequence $\{Z^n\}$ converges $\mathcal{F}$-stably to the distribution of the time-changed process W'_V where W' is a standard Brownian motion independent of $\mathcal{F}$.

The above convergence implies that the marginal distribution Z^n_T converges $\mathcal{F}$-stably to a mixed normal distribution;

$$Z^n_T \Rightarrow N\sqrt{V_T}$$

where $N \sim \mathcal{N}(0,1)$ and is independent of $\mathcal{F}$.

The following lemma is also well-known and repeatedly used in the next section.

Lemma 1. *Consider a sequence of filtrations*

$$\mathcal{H}^n_j \subset \mathcal{H}^n_{j+1}, \quad j, n \in \mathbb{Z}_+ = \{0, 1, 2, \dots\}$$

and random variables $\{U^n_j\}_{j\in\mathbb{N}}$ with U^n_j being $\mathcal{H}^n_j$-measurable. If it holds

$$\sum_{j=0}^{\infty} P[|U^n_{j+1}|^2 | \mathcal{H}^n_j] \to 0 \quad as \quad n \to \infty,$$

then the following two are equivalent;

(1)

$$\sum_{j=0}^{\infty} U^n_{j+1} \to U \quad as\ n \to \infty.$$

(2)

$$\sum_{j=0}^{\infty} P[U^n_{j+1} | \mathcal{H}^n_j] \to U \quad as\ n \to \infty.$$

Here U is a common random variable and the convergences are in probability.

Proof. The proof is the same as in Genon-Catalot and Jacod [13], Lemma 9. □

3. Main result

Here we present a new central limit theorem for the realized volatility. As before, let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ be a filtered probability space and M be a continuous $\{\mathcal{F}_t\}$-local martingale. The filtration is supposed to satisfy the usual conditions. We consider a continuous semi-martingale $X = A + M$ with A being absolutely continuous with respect to $\langle M \rangle$. The Radon-Nikodym derivative $\psi = \{\psi_s\}$, say, is assumed to be locally bounded and left continuous. Let $\tau^n = \{\tau^n_j\}$ be a sampling scheme, that is, a sequence of $\{\mathcal{F}_t\}$-stopping times with

$$0 = \tau^n_0 < \tau^n_1 < \cdots < \tau^n_j < \tau^n_{j+1} < \cdots$$

Suppose that the number of data is finite, that is,

$$N^n_t := \max\{j \geq 0; \tau^n_j \leq t\} < \infty$$

for every $t \geq 0$, and assume for each j and n that

$$P[|\langle M \rangle_{\tau^n_{j+1}} - \langle M \rangle_{\tau^n_j}|^6] < \infty. \tag{1}$$

Note that τ^n_j is $\mathcal{F}_{\tau^n_j}$-measurable for each j and $N^n_t + 1$ is a stopping time with respect to the discrete-time filtration $\{\mathcal{F}_{\tau^n_j}\}$.

Now, put

$$\mathcal{G}^k_{j,n} := P[(M_{\tau^n_{j+1}} - M_{\tau^n_j})^k | \mathcal{F}_{\tau^n_j}]$$

and consider the following structure of sampling scheme;

Condition 1. For all $t \in [0, \infty)$, it holds

$$\sum_{j=0}^{N^n_t} \mathcal{G}^2_{j,n} = O_p(1)$$

as $n \to \infty$ and there exist a positive sequence $\{\epsilon_n\}$ with $\epsilon_n \to 0$ and $\{\mathcal{F}_t\}$-adapted locally bounded left continuous processes $\{a_s\}, \{b_s\}$ such that

$$\mathcal{G}^3_{j,n}/\mathcal{G}^2_{j,n} = b_{\tau^n_j}\epsilon_n + o_p(\epsilon_n), \quad \mathcal{G}^4_{j,n}/\mathcal{G}^2_{j,n} = a^2_{\tau^n_j}\epsilon^2_n + o_p(\epsilon^2_n),$$
$$\mathcal{G}^6_{j,n}/\mathcal{G}^2_{j,n} = o_p(\epsilon^3_n), \quad \mathcal{G}^8_{j,n}/\mathcal{G}^2_{j,n} = o_p(\epsilon^4_n), \quad \mathcal{G}^{12}_{j,n}/\mathcal{G}^2_{j,n} = o_p(\epsilon^6_n)$$

uniformly in $j = 0, 1, \ldots, N^n_t$ as $n \to \infty$.

Lemma 2. *Assume Condition 1 to be satisfied. Then, for all $t \geq 0$,*

$$\sup_{j \geq 0} |\langle M \rangle_{\tau^n_{j+1} \wedge t} - \langle M \rangle_{\tau^n_j \wedge t}| = o_p(\epsilon_n), \quad \sup_{j \geq 0} |M_{\tau^n_{j+1} \wedge t} - M_{\tau^n_j \wedge t}|^2 = o_p(\epsilon_n) \tag{2}$$

as well as

$$\sup_{0\le j\le N_t^n} |\langle M\rangle_{\tau_{j+1}^n} - \langle M\rangle_{\tau_j^n}| = o_p(\epsilon_n), \quad \sup_{0\le j\le N_t^n} |M_{\tau_{j+1}^n} - M_{\tau_j^n}|^2 = o_p(\epsilon_n). \tag{3}$$

Proof. By the assumptions, it follows

$$\sum_{j=0}^{N_t^n} P[(M_{\tau_{j+1}^n} - M_{\tau_j^n})^6|\mathcal{F}_{\tau_j^n}] = o_p(\epsilon_n^3), \quad \sum_{j=0}^{N_t^n} P[(M_{\tau_{j+1}^n} - M_{\tau_j^n})^{12}|\mathcal{F}_{\tau_j^n}] = o_p(\epsilon_n^6),$$

and with the aid of Lemma 1, we have

$$\sum_{j=0}^{N_t^n}(M_{\tau_{j+1}^n} - M_{\tau_j^n})^6 = o_p(\epsilon_n^3).$$

On the other hand,

$$\sup_{0\le j\le N_t^n} |M_{\tau_{j+1}^n} - M_{\tau_j^n}|^2 \le \left\{\sum_{j=0}^{N_t^n}(M_{\tau_{j+1}^n} - M_{\tau_j^n})^6\right\}^{1/3},$$

so that the second of (3) follows. To show the second of (2), use Doob's maximal inequality to have

$$P\left[\sup_{0\le t<\infty} |M_{\tau_{j+1}^n\wedge t} - M_{\tau_j^n\wedge t}|^{2k}|\mathcal{F}_{\tau_j^n}\right]/\mathcal{G}_{j,n}^2 = o_p(\epsilon_n^k)$$

for $k = 3, 6$. Using Lemma 1 again, we obtain

$$\sum_{j=0}^{N_t^n} \sup_{0\le s<\infty} |M_{\tau_{j+1}^n\wedge s} - M_{\tau_j^n\wedge s}|^6 = o_p(\epsilon_n^3),$$

so that the second of (2) follows since

$$\sup_{0\le j\le N_t^n}\sup_{0\le s<\infty} |M_{\tau_{j+1}^n\wedge s} - M_{\tau_j^n\wedge s}|^2 \le \left\{\sum_{j=0}^{N_t^n}\sup_{0\le s<\infty} |M_{\tau_{j+1}^n\wedge s} - M_{\tau_j^n\wedge s}|^6\right\}^{1/3}.$$

Using the Burkholder-Davis-Gundy inequality and Doob's maximal inequality, we have also

$$\sum_{j=0}^{N_t^n} P[|\langle M\rangle_{\tau_{j+1}^n} - \langle M\rangle_{\tau_j^n}|^3|\mathcal{F}_{\tau_j^n}] = o_p(\epsilon_n^3), \quad \sum_{j=0}^{N_t^n} P[|\langle M\rangle_{\tau_{j+1}^n} - \langle M\rangle_{\tau_j^n}|^6|\mathcal{F}_{\tau_j^n}] = o_p(\epsilon_n^6),$$

which implies the first of (3) in the same manner. It is then clear to see the first of (2). □

Now, put for a given process Y,

$$((Y))_p[\tau^n]_t := \sum_{j=0}^{\infty}(Y_{\tau^n_{j+1}\wedge t} - Y_{\tau^n_j\wedge t})^p$$

for $p \in \mathbb{N}$, $t \geq 0$ and consider a sequence of continuous local martingales $\{Z^n\}$ defined by

$$Z^n_t := \epsilon_n^{-1}\left(((M))_2[\tau^n]_t - \langle M\rangle_t\right)$$

By Lemma 2, we have

$$\mathcal{R}v[\tau^n]_t = ((M))_2[\tau^n]_t + 2\sum_{j=0}^{N^n_t}(A_{\tau^n_{j+1}} - A_{\tau^n_j})(M_{\tau^n_{j+1}} - M_{\tau^n_j}) + o_p(\epsilon_n).$$

Note that

$$Z^n_t = 2\epsilon_n^{-1}\sum_{j=0}^{\infty}\int_{\tau^n_j\wedge t}^{\tau^n_{j+1}\wedge t}(M_s - M_{\tau^n_j})dM_s$$

by Itô's formula, so that for all $N \in \mathcal{M}^{\perp}$, it holds $\langle Z^n, N\rangle = 0$.

Proposition 1. *Assume Condition 1 to be satisfied. Then for all $t \in [0,\infty)$,*

$$\begin{aligned}
&((M))_2[\tau^n]_t \to \langle M\rangle_t,\\
&\epsilon_n^{-1}((M))_3[\tau^n]_t \to \int_0^t b_s d\langle M\rangle_s,\\
&\epsilon_n^{-2}((M))_4[\tau^n]_t \to \int_0^t a_s^2 d\langle M\rangle_s,\\
&\langle Z^n, M\rangle_t \to \frac{2}{3}\int_0^t b_s d\langle M\rangle_s,\\
&\langle Z^n - \frac{2}{3}\sum_{j=0}^{\infty} b_{\tau^n_j}(M_{\tau^n_{j+1}\wedge\cdot} - M_{\tau^n_j\wedge\cdot})\rangle_t \to \frac{2}{3}\int_0^t c_s^2 d\langle M\rangle_s,
\end{aligned}$$

where

$$c_s^2 := a_s^2 - \frac{2}{3}b_s^2,$$

and the convergences are in probability.

Proof. By Lemma 2 and Lebesgue's convergence theorem, for all locally bounded left continuous process $\{g_s\}$, it holds that

$$\sum_{j=0}^{N_t^n} g_{\tau_j^n} |\langle M \rangle_{\tau_{j+1}^n} - \langle M \rangle_{\tau_j^n}| \to \int_0^t g_s d\langle M \rangle_s$$

in probability. Since

$$\sum_{j=0}^{N_t^n} g_{\tau_j^n}^2 \mathcal{G}_{j,n}^4 = O_p(\epsilon_n^2)$$

and

$$P[|\langle M \rangle_{\tau_{j+1}^n} - \langle M \rangle_{\tau_j^n}|^2 | \mathcal{F}_{\tau_j^n}] \leq C \mathcal{G}_{j,n}^4 \quad \text{a.s.}$$

for a constant C by the Burkholder-Davis-Gundy inequality and Doob's inequality, using Lemma 1, we have

$$\sum_{j=0}^{N_t^n} g_{\tau_j^n} \mathcal{G}_{j,n}^2 \to \int_0^t g_s d\langle M \rangle_s \tag{4}$$

in probability. The first three convergences follow from Lemmas 1 and 2 with the aid of (4). Now, note that Itô's formula gives

$$\langle Z^n, M \rangle_t = 2\epsilon_n^{-1} \sum_{j=0}^{N_t^n} \int_{\tau_j^n}^{\tau_{j+1}^n} (M_s - M_{\tau_j^n}) d\langle M \rangle_s + o_p(1) = \sum_{j=0}^{N_t^n} A_j + o_p(1),$$

where

$$A_j := \frac{2}{3} \epsilon_n^{-1} (M_{\tau_{j+1}^n} - M_{\tau_j^n})^3 - 2\epsilon_n^{-1} \int_{\tau_j^n}^{\tau_{j+1}^n} (M_s - M_{\tau_j^n})^2 dM_s.$$

Since

$$\sum_{j=0}^{M_t^n} P[A_j | \mathcal{F}_{\tau_j^n}] = \frac{2}{3} \sum_{j=0}^{M_t^n} b_{\tau_j^n} \mathcal{G}_{j,n}^2 + o_p(1), \quad \sum_{j=0}^{M_t^n} P[|A_j|^2 | \mathcal{F}_{\tau_j^n}] = o_p(\epsilon_n),$$

the fourth convergence also follows from Lemma 1 and (4). In order to see the last one, noting that

$$\begin{aligned}
&\langle Z^n - \frac{2}{3} \sum_{j=0}^{\infty} b_{\tau_j^n} (M_{\tau_{j+1}^n \wedge \cdot} - M_{\tau_j^n \wedge \cdot}) \rangle_t \\
&= \langle Z^n \rangle_t - 2 \langle Z^n, \frac{2}{3} \sum_{j=0}^{\infty} b_{\tau_j^n} (M_{\tau_{j+1}^n \wedge \cdot} - M_{\tau_j^n \wedge \cdot}) \rangle_t + \frac{4}{9} \sum_{j=0}^{\infty} b_{\tau_j^n}^2 (\langle M \rangle_{\tau_{j+1}^n \wedge t} - \langle M \rangle_{\tau_j^n \wedge t}),
\end{aligned}$$

it suffices to show

$$\langle Z^n \rangle_t \to \frac{2}{3}\int_0^t a_s^2 d\langle M\rangle_s$$
$$\langle Z^n, \frac{2}{3}\sum_{j=0}^{\infty} b_{\tau_j^n}(M_{\tau_{j+1}^n\wedge\cdot} - M_{\tau_j^n\wedge\cdot})\rangle_t \to \frac{4}{9}\int_0^t b_s^2 d\langle M\rangle_s$$

in probability. For the first of these two, observe that

$$\langle Z^n\rangle_t = 4\epsilon_n^{-2}\sum_{j=0}^{N_t^n}\int_{\tau_j^n}^{\tau_{j+1}^n}(M_s - M_{\tau_j^n})^2 d\langle M\rangle_s + o_p(1) = \sum_{j=0}^{N_t^n} B_j + o_p(1),$$

where

$$B_j := \frac{2}{3}\epsilon_n^{-2}(M_{\tau_{j+1}^n} - M_{\tau_j^n})^4 - \frac{8}{3}\epsilon_n^{-2}\int_{\tau_j^n}^{\tau_{j+1}^n}(M_s - M_{\tau_j^n})^3 dM_s,$$

and that

$$\sum_{j=0}^{N_t^n} P[B_j|\mathcal{F}_{\tau_j^n}] = \frac{2}{3}\sum_{j=0}^{N_t^n} a_{\tau_j^n}^2\mathcal{G}_{j,n}^2 + o_p(1),$$
$$\sum_{j=0}^{N_t^n} P[|B_j|^2|\mathcal{F}_{\tau_j^n}] = o_p(1).$$

For the second, observe that the left hand term can be written as

$$\frac{4}{3}\epsilon_n^{-1}\sum_{j=0}^{N_t^n} b_{\tau_j^n}\int_{\tau_j^n}^{\tau_{j+1}^n}(M_s - M_{\tau_j^n})d\langle M\rangle_s + o_p(1) = \sum_{j=0}^{N_t^n} C_j + o_p(1),$$

where

$$C_j := \frac{4}{3}\epsilon_n^{-1} b_{\tau_j^n}\left\{\frac{1}{3}(M_{\tau_{j+1}^n} - M_{\tau_j^n})^3 - \int_{\tau_j^n}^{\tau_{j+1}^n}(M_s - M_{\tau_j^n})^2 dM_s\right\},$$

and that

$$\sum_{j=0}^{N_t^n} P[C_j|\mathcal{F}_{\tau_j^n}] = \frac{4}{9}\sum_{j=0}^{N_t^n} b_{\tau_j^n}^2\mathcal{G}_{j,n}^2 + o_p(1),$$
$$\sum_{j=0}^{N_t^n} P[|C_j|^2|\mathcal{F}_{\tau_j^n}] = o_p(\epsilon_n).$$

Here we have repeatedly used Lemma 1. □

Theorem 2. *Assume Condition 1 to be satisfied. Then the $C[0,\infty)$-valued sequence $\{Z^n\}$ converges $\mathcal{F}$-stably to the distribution of*

$$\frac{2}{3}\int_0^{\cdot} b_s dM_s + \sqrt{\frac{2}{3}}\int_0^{\cdot} c_s dX'_s,$$

where $X' = W'_{\langle X\rangle}$ is a time-changed process of a standard Brownian motion W' which is independent of $\mathcal{F}$. Furthermore, the $C[0,\infty)$-valued sequence

$$\epsilon_n^{-1}\left(((X))_2[\tau^n] - \langle X\rangle\right)$$

converges $\mathcal{F}$-stably to the distribution of

$$\frac{2}{3}\int_0^{\cdot} b_s dX_s + \sqrt{\frac{2}{3}}\int_0^{\cdot} c_s dX'_s.$$

In particular, for every finite stopping time T,

$$\epsilon_n^{-1}\left(\mathcal{R}v[\tau^n]_T - \langle X\rangle_T\right)$$

converges in law to the mixed normal distribution

$$\mathcal{MN}\left(\frac{2}{3}\int_0^T b_s dX_s, \frac{2}{3}\int_0^T c_s^2 d\langle X\rangle_s\right).$$

Proof. By Proposition 1 and Theorem 1, we obtain the $\mathcal{F}$-stable convergence

$$Z^n - \frac{2}{3}\sum_{j=0}^{\infty} b_{\tau_j^n}(M_{\tau_{j+1}^n\wedge\cdot} - M_{\tau_j^n\wedge\cdot}) \Rightarrow \sqrt{\frac{2}{3}}\int_0^{\cdot} c_s dX'_s.$$

It is not difficult to see

$$\sum_{j=0}^{\infty} b_{\tau_j^n}(M_{\tau_{j+1}^n\wedge\cdot} - M_{\tau_j^n\wedge\cdot}) \to \int_0^{\cdot} b_s dM_s$$

in probability. It remains to prove

$$\epsilon_n^{-1}\sum_{j=0}^{N_t^n}(A_{\tau_{j+1}^n} - A_{\tau_j^n})(M_{\tau_{j+1}^n} - M_{\tau_j^n}) \to \frac{1}{3}\int_0^t b_s dA_s.$$

Since

$$
\begin{aligned}
&(A_{\tau^n_{j+1}} - A_{\tau^n_j})(M_{\tau^n_{j+1}} - M_{\tau^n_j}) \\
&= \int_{\tau^n_j}^{\tau^n_{j+1}} (M_s - M_{\tau^n_j}) dA_s + \int_{\tau^n_j}^{\tau^n_{j+1}} (A_s - A_{\tau^n_j}) dM_s \\
&= \psi_{\tau^n_j} \int_{\tau^n_j}^{\tau^n_{j+1}} (M_s - M_{\tau^n_j}) d\langle M\rangle_s + \int_{\tau^n_j}^{\tau^n_{j+1}} (M_s - M_{\tau^n_j})(\psi_s - \psi_{\tau^n_j}) d\langle M\rangle_s \\
&\qquad + \int_{\tau^n_j}^{\tau^n_{j+1}} \int_{\tau^n_j}^{s} \psi_u d\langle M\rangle_u dM_s,
\end{aligned}
$$

it suffices to see

$$
\begin{aligned}
&\epsilon_n^{-1} \sum_{j=0}^{N^n_t} \psi_{\tau^n_j} \int_{\tau^n_j}^{\tau^n_{j+1}} (M_s - M_{\tau^n_j}) d\langle M\rangle_s \to \frac{1}{3}\int_0^t b_s \psi_s d\langle M\rangle_s, \\
&\epsilon_n^{-1} \sum_{j=0}^{N^n_t} \int_{\tau^n_j}^{\tau^n_{j+1}} (M_s - M_{\tau^n_j})(\psi_s - \psi_{\tau^n_j}) ds \to 0, \\
&\epsilon_n^{-1} \sum_{j=0}^{N^n_t} \int_{\tau^n_j}^{\tau^n_{j+1}} \int_{\tau^n_j}^{s} \psi_u d\langle M\rangle_u dM_s \to 0.
\end{aligned}
$$

By localizing argument, we can assume ψ_s is uniformly bounded without loss of generality. Then, the first and third convergences follow from Lemma 1 by noting

$$
\begin{aligned}
&P\left[\psi_{\tau^n_j} \int_{\tau^n_j}^{\tau^n_{j+1}} (M_s - M_{\tau^n_j}) d\langle M\rangle_s \middle| \mathcal{F}_{\tau^n_j}\right] = \frac{1}{3}\psi_{\tau^n_j} \mathcal{G}^3_{j,n}, \\
&P\left[\left|\psi_{\tau^n_j} \int_{\tau^n_j}^{\tau^n_{j+1}} (M_s - M_{\tau^n_j}) d\langle M\rangle_s\right|^2 \middle| \mathcal{F}_{\tau^n_j}\right] = o_p(\epsilon_n^3)\mathcal{G}^2_{j,n}
\end{aligned}
$$

and

$$
\begin{aligned}
&P\left[\int_{\tau^n_j}^{\tau^n_{j+1}} \int_{\tau^n_j}^{s} \psi_u d\langle M\rangle_u dM_s \middle| \mathcal{F}_{\tau^n_j}\right] = 0, \\
&P\left[\left|\int_{\tau^n_j}^{\tau^n_{j+1}} \int_{\tau^n_j}^{s} \psi_u d\langle M\rangle_u dM_s\right|^2 \middle| \mathcal{F}_{\tau^n_j}\right] = o_p(\epsilon_n^3)\mathcal{G}^2_{j,n}
\end{aligned}
$$

respectively. The second convergence follows from the Cauchy-Schwarz inequality:

$$
\begin{aligned}
&\epsilon_n^{-1}\sum_{j=0}^{N_t^n}\int_{\tau_j^n}^{\tau_{j+1}^n}(M_s - M_{\tau_j^n})(\psi_s - \psi_{\tau_j^n})ds \\
&\leq \sqrt{\epsilon_n^{-2}\int_0^{\tau_{N_t^n+1}}(M_s - M_s^n)^2 d\langle M\rangle_s}\sqrt{\int_0^{\tau_{N_t^n+1}}(\psi_s - \psi_s^n)^2 d\langle M\rangle_s} \\
&\leq \sqrt{\langle Z^n\rangle_t/4 + o_p(1)}\sqrt{\int_0^{\tau_{N_t^n+1}}(\psi_s - \psi_s^n)^2 d\langle M\rangle_s} \to 0,
\end{aligned}
$$

where

$$M_s^n = M_{\tau_j^n}, \quad \psi_s^n = \psi_{\tau_j^n}, \quad \text{for } s \in [\tau_j^n, \tau_{j+1}^n).$$

Here we used the fact that ψ is left continuous. □

The asymptotic distribution is, therefore, determined by b and a, which are asymptotic skewness and kurtosis of returns in a sense, respectively. The larger the skewness, the more biased the realized volatility. The larger the kurtosis, the larger the mean-squared error. Remarkably, the drift term A affects the asymptotic distribution if $b \not\equiv 0$. Unfortunately, the randomness of the asymptotic mean and variance hampers practical use of the central limit theorem. Nevertheless, if we can assume $b \equiv 0$, the following corollary will be useful in constructing confidence intervals or testing hypothesis.

Corollary 1. *Let T be a finite stopping time. Assume Condition 1 to be satisfied with $b_s 1_{s\leq T} \equiv 0$ and $a_s 1_{s\leq T} \not\equiv 0$. Then*

$$\frac{\mathcal{R}v[\tau^n]_T - \langle X\rangle_T}{\sqrt{\mathcal{R}q[\tau^n]_T}} \Rightarrow \mathcal{N}(0, 2/3),$$

where

$$\mathcal{R}q[\tau^n]_T = \sum_{j=0}^{N_T^n-1}\left(X_{\tau_{j+1}^n} - X_{\tau_j^n}\right)^4.$$

This is a natural extension of a known result on studentizing for the equidistant sampling case. It should be noted that we do not need to specify the structure of the sampling scheme nor the normalizing sequence ϵ_n in constructing the above studentized statistic. Thus we obtain a statistic with the asymptotic normality in a nonparametric setting. It remains for further research to deal with the general case of $b \not\equiv 0$; however, assuming b to be constant, we obtain the following assertion.

Corollary 2. *Let T be a finite stopping time. Assume Condition 1 to be satisfied with $b_s 1_{s\leq T} \equiv \bar{b}$ for a constant $\bar{b}$ and $a_s 1_{s\leq T} \not\equiv 0$. Then,*

$$\frac{\mathcal{R}v[\tau^n]_T - \langle X\rangle_T - 2X_T\mathcal{R}t[\tau^n]_T/(3\mathcal{R}v[\tau^n]_T)}{\sqrt{\mathcal{R}q[\tau^n]_T - 2(\mathcal{R}t[\tau^n]_T)^2/(3\mathcal{R}v[\tau^n]_T)}} \Rightarrow \mathcal{N}(0, 2/3),$$

where

$$\mathcal{R}t[\tau^n]_T = \sum_{j=0}^{N_T^n - 1} \left(X_{\tau_{j+1}^n} - X_{\tau_j^n} \right)^3 .$$

4. Examples

4.1. **Conditionally independent sampling scheme.** Here we treat a sampling scheme $\tau^n = \{\tau_j^n\}$ with such a measurability condition that was imposed in Jacod [17] and Hayashi, Jacod and Yoshida [12]; we assume that the duration $\tau_{j+1}^n - \tau_j^n$ is conditionally independent of $M_{\cdot + \tau_j^n} - M_{\tau_j^n}$ given $\mathcal{F}_{\tau_j^n}$ for each $j \geq 0$. For example, suppose that

$$\tau_{j+1}^n - \tau_j^n = g_{\tau_j^n} \epsilon_n^2 + o_p(\epsilon_n^2)$$

with an $\{\mathcal{F}_t\}$-adapted positive left continuous process g which is locally bounded and bounded away from 0. Here the $o_p(\epsilon_n^2)$ term is supposed to be $\mathcal{F}_{\tau_j^n}$-measurable. In such a case, we obtain in a straightforward manner,

$$\mathcal{G}_{j,n}^{2k+1} = o_p(\epsilon_n^{2k+1}), \quad \mathcal{G}_{j,n}^{2k} = \frac{(2k)!}{2^k k!} \sigma_{\tau_j^n}^{2k} g_{\tau_j^n}^k \epsilon_n^{2k} + o_p(\epsilon_n^{2k})$$

provided that there exists an adapted locally bounded process σ such that

$$(5) \qquad M_t = \int_0^t \sigma_s dW_s, \quad \sup_{0 \leq j \leq N_t^n, s \geq 0} P\left[|\sigma_{\tau_{j+1}^n \wedge s} - \sigma_{\tau_j^n}|^{2k} | \mathcal{F}_{\tau_j^n} \right] \to 0$$

as $n \to \infty$ for each $t \geq 0$, $k \leq 6$, where W is a standard Brownian motion. Note that we can assume σ is bounded without loss of generality by localizing argument. Then, it is easy to see

$$\sum_{j=0}^{N_t^n} \mathcal{G}_{j,n}^2 = O_p(1),$$

so that we can apply Theorem 2 with

$$b_s \equiv 0, \quad c_s^2 = a_s^2 = 3 g_s \sigma_s^2.$$

This asymptotic distribution coincides, of course, with known results for deterministic sampling scheme.

Under the same assumption as (5), we can deal with the following Poisson sampling;

$$\tau_{j+1}^n - \tau_j^n \sim \epsilon_n^2 \lambda_j^n e,$$

conditionally to $\mathcal{F}_{\tau_j^n}$, where $e \sim \mathrm{Exp}(1)$ is an independent exponential variable and λ_j^n is $\mathcal{F}_{\tau_j^n}$-measurable positive random variable. More generally, if $\tau_{j+1}^n - \tau_j^n$ is conditionally independent of $M_{\cdot+\tau_j^n} - M_{\tau_j^n}$ given $\mathcal{F}_{\tau_j^n}$ for each $j \geq 0$ with

$$P[|\tau_{j+1}^n - \tau_j^n|^k | \mathcal{F}_{\tau_j^n}] = \epsilon_n^{2k} m_{\tau_j^n}^{(k)} + o_p(\epsilon^{2k}), \quad k = 1, 2, \dots, 6$$

for adapted locally bounded left continuous processes $m^{(k)}$, then we have

$$\mathcal{G}_{j,n}^{2k+1} = o_p(\epsilon_n^{2k+1}), \quad \mathcal{G}_{j,n}^{2k} = \frac{(2k)!}{2^k k!} \sigma_{\tau_j^n}^{2k} m_{\tau_j^n}^{(k)} \epsilon_n^{2k} + o_p(\epsilon_n^{2k})$$

and

$$b_s \equiv 0, \quad c_s^2 = a_s^2 = 3\sigma_s^2 m_s^{(2)} / m_s^{(1)}.$$

Here $m^{(1)}$ is assumed to be locally bounded away from 0.

4.2. **Calendar time and Business time sampling.** The so-called calendar time sampling is nothing but the equidistant sampling scheme $\tau_j^n = jT/n$ for a fixed interval $[0, T]$. This case has been extensively investigated and the corresponding asymptotic distribution of the realized volatility is specified as

$$b_s \equiv 0, \quad c_s^2 = a_s^2 = 3T\sigma_s^2$$

with $\epsilon_n = 1/\sqrt{n}$ in our terminology. The so-called business time sampling is given as

$$\tau_j^n = \inf\{t > 0; \langle X \rangle_t \geq \langle X \rangle_T j/n\},$$

or equivalently,

$$\langle X \rangle_{\tau_{j+1}^n} - \langle X \rangle_{\tau_j^n} = \langle X \rangle_T / n.$$

The sampling times are latent because they are defined from unobserved volatility path; however it is naively considered an ideal scheme in terms of estimation of volatility. See Hansen and Lunde [15] for detail. Since the above definition does not give a stopping times, let us consider a modified scheme

$$\tau_j^n = \inf\{t > 0; \langle X \rangle_t \geq P[\langle X \rangle_T] j/n\},$$

or equivalently,

$$\langle X \rangle_{\tau_{j+1}^n} - \langle X \rangle_{\tau_j^n} = P[\langle X \rangle_T] / n.$$

Here the integrability of $\langle X \rangle_T$ is assumed. Note that $P[N_T^n] \leq n$ since

$$N_T^n P[\langle X \rangle_T] / n = \sum_{j=0}^{N_T^n - 1} \langle X \rangle_{\tau_{j+1}^n} - \langle X \rangle_{\tau_j^n} = \langle X \rangle_{\tau_{N_T^n}^n} \leq \langle X \rangle_T.$$

If $\langle X \rangle$ is strictly increasing, then by the Dambis-Dubins-Schwarz theorem (see e.g., Karatzas and Shreve [22], 3.4.6, and Kallenberg [21], Proposition 7.9), there exists a standard Brownian motion $W = \{W_t, \mathcal{H}_t, 0 \leq t < \infty\}$ such that

$$M = W_{\langle X \rangle}, \quad \mathcal{F}_\tau = \mathcal{H}_{\langle X \rangle_\tau}$$

for every stopping time τ. Hence it holds that

$$\mathcal{G}^k_{j,n} = P\left[\left(W_{\langle X \rangle_{\tau^n_{j+1}}} - W_{\langle X \rangle_{\tau^n_j}}\right)^k \middle| \mathcal{H}_{\langle X \rangle_{\tau^n_j}}\right] \tag{6}$$

so that

$$\mathcal{G}^{2k+1}_{j,n} = 0, \quad \mathcal{G}^{2k}_{j,n} = \frac{(2k)!}{2^k k!} P[\langle X \rangle_T]^k / n^k$$

and putting $\epsilon_n = 1/\sqrt{n}$, we have

$$b_s \equiv 0, \quad a_s^2 \equiv 3P[\langle X \rangle_T].$$

The asymptotic variance of $\sqrt{n}(\mathcal{R}v[\tau^n]_T - \langle X \rangle_T)$ is therefore $2P[\langle X \rangle_T]^2$. Note that if $\langle X \rangle$ is absolutely continuous, that is, there exists an adapted process σ such that

$$\langle X \rangle_t = \int_0^t \sigma_s^2 ds,$$

then, by Jensen's inequality,

$$2P[\langle X \rangle_T]^2 \leq 2TP\left[\int_0^T \sigma_s^4 ds\right].$$

Since the right hand side is the asymptotic variance in calendar time sampling, we can conclude that business time sampling is more efficient than calendar time sampling as is expected. On the other hand, business time sampling is not optimal, as seen later.

4.3. **Tick time sampling.** Here we consider path-dependent sampling schemes which serve as models for the so-called tick time sampling. By tick time sampling, we mean a sampling scheme which is based on random times at which price changes. At first, define τ^n as

$$\tau^n_0 = 0, \quad \tau^n_{j+1} = \inf\{t > \tau^n_j; X_t \in \varphi(\epsilon_n \mathbb{D}) \setminus \{X_{\tau^n_j}\}\}, \tag{7}$$

where $\varphi : \mathbb{E} \to \mathbb{R}$ is a C^2 homeomorphism and $(\mathbb{E}, \mathbb{D}) = (\mathbb{R}_+, \mathbb{N})$ or $(\mathbb{R}, \mathbb{Z})$. This sampling scheme was introduced by Fukasawa [10] as a model for tick time sampling; if we assume bid price B_t is given as

$$B_t = \delta \lfloor (1 - \beta) S_t / \delta \rfloor \tag{8}$$

for latent price $S_t = \exp(X_t)$, tick size δ and discount factor $\beta \in [0,1)$ which represents a proportional cost for quote exposure, it holds

$$\tau^n_{j+1} = \inf\{t > \tau^n_j; B_t \geq B_{\tau^n_j} + \delta \text{ or } B_t \leq B_{\tau^n_j} - 2\delta\}$$

almost surely, by setting $\varphi = \log$, $(\mathbb{E}, \mathbb{D}) = (\mathbb{R}, \mathbb{N})$ and $\epsilon_n = \delta/(1-\beta)$. Hence τ^n_j corresponds to j-th continued price change of bid quote.

Now, let us assume $X = M$ for brevity; this restriction can be removed in the light of the Girsanov-Maruyama transformation. Besides, we assume X is a martingale without loss of generality by localizing argument. Then, by the optional sampling theorem,

$$P[X_{\tau^n_{j+1}} = \varphi(\varphi^{-1}(X_{\tau^n_j}) + \epsilon)|\mathcal{F}_{\tau_j}] = \frac{d_-(X_{\tau^n_j})}{d_+(X_{\tau^n_j}) + d_-(X_{\tau^n_j})},$$
$$P[X_{\tau^n_{j+1}} = \varphi(\varphi^{-1}(X_{\tau^n_j}) - \epsilon)|\mathcal{F}_{\tau_j}] = \frac{d_+(X_{\tau^n_j})}{d_+(X_{\tau^n_j}) + d_-(X_{\tau^n_j})},$$

where

$$d_+(x) = \varphi(\varphi^{-1}(x) + \epsilon_n) - x, \quad d_-(x) = x - \varphi(\varphi^{-1}(x) - \epsilon_n), \quad x \in \mathbb{R}.$$

It is then straightforward to see

$$\begin{aligned} &\mathcal{G}^2_{j,n} = d_+(X_{\tau^n_j})d_-(X_{\tau^n_j}) = \varphi'(\varphi^{-1}(X_{\tau^n_j}))^2\epsilon_n^2 + o_p(\epsilon_n^2), \\ &\mathcal{G}^3_{j,n}/\mathcal{G}^2_{j,n} = d_+(X_{\tau^n_j}) - d_-(X_{\tau^n_j}) = o_p(\epsilon_n), \\ &\mathcal{G}^4_{j,n}/\mathcal{G}^2_{j,n} = d_+(X_{\tau^n_j})^2 - d_+(X_{\tau^n_j})d_-(X_{\tau^n_j}) + d_-(X_{\tau^n_j})^2 = \mathcal{G}^2_{j,n} + o_p(\epsilon_n^2). \end{aligned} \tag{9}$$

It is also straightforward to see

$$\epsilon_n^2 N^n_t = \sum_{j=0}^{N^n_t - 1} \Big(\varphi^{-1}(X_{\tau^n_{j+1}}) - \varphi^{-1}(X_{\tau^n_j})\Big)^2 \to \langle \varphi^{-1}(X)\rangle_t$$

in probability, so that

$$\sum_{j=0}^{N^n_t} \mathcal{G}^2_{j,n} = O_p(1).$$

Consequently, we can apply Theorem 2 with

$$b_s \equiv 0, \quad a_s^2 = \varphi'(\varphi^{-1}(X_s))^2.$$

Here we have seen that the proof in Fukasawa [10] can be simplified using Theorem 2.

When considering (7) as a model for tick time sampling, we suppose that bid quotation is so high-frequently updated that (8) is satisfied for almost all $t \geq 0$.

In other words, we are neglecting time-discretization error and consider space-discretization effect which comes from price discreteness more significant. Now, let us consider another model which incorporates time-discretization effect. Define a sampling scheme τ^n as

$$\tau^n_{j+1} = \inf\{t > \sigma^n_j; X_t \in \varphi(\delta_n \mathbb{D})\}, \quad \sigma^n_j = \inf\{t > \tau^n_j; \langle X\rangle_t \geq \langle X\rangle_{\tau^n_j} + h_n\} \tag{10}$$

with $\tau^n_0 = 0$, where δ_n and h_n are positive numbers. Let $t_1, t_2, \dots$ are bid quote-revision times in tick data and $\hat{B}_{t_1}, \hat{B}_{t_2}, \dots$ are the corresponding bid prices. Let t_{m_1} be the first time at which the price changes, that is, $\hat{B}_{t_{m_1}} \neq \hat{B}_{t_{m_1-1}}$. Taking (8) into consideration, we identify $\tau^n_1 = t_{m_1}$

$$B_{\tau^n_1} = \begin{cases} \hat{B}_{t_{m_1}} & \text{if } \hat{B}_{t_{m_1}} > \hat{B}_{t_{m_1-1}}, \\ \hat{B}_{t_{m_1}} + \delta & \text{if } \hat{B}_{t_{m_1}} < \hat{B}_{t_{m_1-1}}, \end{cases}$$

where $\varphi = \log$, $\mathbb{D} = \mathbb{N}$ and $\delta_n = \delta/(1-\beta)$. Then, we identify $\sigma^n_1 = t_{m_1+1}$; we regard the duration $\sigma^n_j - \tau^n_j$ as a refractory period of quotation which represents time-discretization effect. Here a latent quantity h_n controls the length in business time scale. Let t_{m_2} be the first time after t_{m_1+1} at which price changes and identify $\tau^n_2 = t_{m_2}$ and $\sigma^n_2 = t_{m_2+1}$. Repeat this procedure to obtain a realization $(\tau^n_j, B_{\tau^n_j})$ as data. Note that $B_{\tau^n_j} = (1-\beta)S_{\tau^n_j}$, so that

$$\log(B_{\tau^n_{j+1}}) - \log(B_{\tau^n_j}) = X_{\tau^n_{j+1}} - X_{\tau^n_j},$$

which means we can construct $\mathcal{R}v[\tau^n]$ from these data.

Apart from this interpretation, the sampling scheme (10) has an interesting structure from mathematical point of view. By localizing argument, we can assume that φ' is bounded and bounded away from 0 without loss of generality. Assume also $X = M$ again and that $\langle X\rangle$ is strictly increasing. Then we have (6) by the Dambis-Dubins-Schwarz time-change method. Note that $\hat{\tau}^n_j = \langle X\rangle_{\tau^n_j}$ satisfies

$$\hat{\tau}^n_{j+1} = \inf\{t > \hat{\tau}^n_j + h_n; W_t \in \varphi(\delta_n \mathbb{D})\},$$

where W is a standard Brownian motion with $X = W_{\langle X\rangle}$. Put $x^n_k = \varphi(\delta_n k)$ for $k \in \mathbb{D}$. Then, using the strong Markov property of the Brownian motion, we have

(11)

$$\begin{aligned} \mathcal{G}^m_{j,n} =& \sum_{k\in\mathbb{D}} \int_{x^n_k}^{x^n_{k+1}} \left\{ (x^n_{k+1} - X_{\tau^n_j})^m \frac{x - x^n_k}{x^n_{k+1} - x^n_k} + (x^n_k - X_{\tau^n_j})^m \frac{x^n_{k+1} - x}{x^n_{k+1} - x^n_k} \right\} \\ & \phi(x, X_{\tau^n_j}, h_n) dx \\ =& h_n^{m/2} \sum_{k\in\mathbb{D}} (\phi(\varphi_n(k+1)) - \phi(\varphi_n(k))) \frac{\varphi_n(k)^m - \varphi_n(k+1)^m}{\varphi_n(k+1) - \varphi_n(k)} \\ & (\Phi(\varphi_n(k+1)) - \Phi(\varphi_n(k))) \frac{\varphi_n(k)^m \varphi_n(k+1) - \varphi_n(k+1)^m \varphi_n(k)}{\varphi_n(k+1) - \varphi_n(k)}, \end{aligned}$$

where $\varphi_n(k) = h_n^{-1/2}(\varphi(\delta_n k) - X_{\tau_j^n})$ and Φ, ϕ are the distribution function and the density of the standard normal distribution respectively. If $h_n, \delta_n \to 0$ with $\delta_n^2/h_n \to 0$, then $\varphi_n(k+1)-\varphi_n(k) \to 0$ uniformly, so that by Lebesgue's convergence theorem,

$$\mathcal{G}_{j,n}^m/h_n^{m/2} \to -m\int x^{m-1}\phi'(x)dx - (m-1)\int x^m\phi(x)dx = \int x^m \phi(x)dx.$$

Therefore Condition 1 holds with $\epsilon_n = h_n^{1/2}$, $b_s \equiv 0$ and $a_s \equiv 3$; the asymptotic distribution of the realized volatility coincides with that in business time sampling. On the other hand, if $h_n, \delta_n \to 0$ with $\delta_n^2/h_n \to \infty$, then $\varphi_n(k) \to \infty$ for $k \neq k_0$ uniformly, where k_0 is defined as $X_{\tau_j^n} = \varphi(\delta_n k_0)$. It follows then that

$$\mathcal{G}_{j,n}^m = h_n^{m/2}\left\{-\phi(0)\varphi_n(k_0-1)^{m-1} + \phi(0)\varphi_n(k_0+1)^{m-1} + o_p\left((h_n^{-1/2}\delta_n)^{m-1}\right)\right\},$$

so that if m is even, we obtain

$$\mathcal{G}_{j,n}^m = 2\varphi'(\varphi^{-1}(X_{\tau_j^n}))^{m-1}h_n^{1/2}\delta_n^{m-1} + o_p(h_n^{1/2}\delta_n^{m-1})$$

and if m is odd,

$$\mathcal{G}_{j,n}^m = o_p(h_n^{1/2}\delta_n^{m-1}).$$

Consequently, Condition 1 holds with $\epsilon_n = \delta_n$, $b_s \equiv 0$ and $a_s = \varphi'(\varphi^{-1}(X_{\tau_j^n}))^2$; the asymptotic distribution of the realized volatility coincides with that in the previous scheme (7). This is an example of such a sampling scheme that (15) below is not satisfied. These two cases, therefore, have totally different asymptotic behavior. In the former time-discretization effect is dominant and in the latter space-discretization effect is dominant. The intermediate case $\delta_n^2/h_n \to \alpha \in (0, \infty)$ bridges the gap between the two.

5. Efficient sampling scheme

We consider efficiency problem in this section. At first, let us deal with mean-squared error.

Lemma 3. *Assume $A = 0$ and that τ^n is a sampling scheme with (1). Then,*

$$P\left[|((X))_2[\tau^n]_T - \langle X\rangle_T|^2\right] \geq \frac{2}{3}\frac{|P[\langle X\rangle_T]|^2}{1 + P[N_T^n]} \tag{12}$$

for every finite stopping time T.

Proof. By Itô's formula, it holds

$$P\left[\left|(X_{\tau_{j+1}^n} - X_{\tau_j^n})^2 - (\langle X\rangle_{\tau_{j+1}^n} - \langle X\rangle_{\tau_j^n})\right|^2 |\mathcal{F}_{\tau_j^n}\right] = \frac{2}{3}P\left[(X_{\tau_{j+1}^n} - X_{\tau_j^n})^4|\mathcal{F}_{\tau_j^n}\right],$$

so that

$$P\left[|((X))_2[\tau^n]_T - \langle X\rangle_T|^2\right] = \frac{2}{3}P\left[((X))_4[\tau^n]_T\right]. \tag{13}$$

The assertion follows from Jensen's inequality and the Cauchy-Schwarz inequality;

$$\begin{aligned}P\left[\sum_{j=0}^{N_T^n}(X_{\tau_{j+1}^n\wedge T} - X_{\tau_j^n\wedge T})^4\right] &\geq P\left[\frac{1}{1+N_T^n}\left\{\sum_{j=0}^{N_T^n}(X_{\tau_{j+1}^n\wedge T} - X_{\tau_j^n\wedge T})^2\right\}^2\right] \\ &\geq \frac{1}{1+P[N_T^n]}P\left[\sum_{j=0}^{N_T^n}(X_{\tau_{j+1}^n\wedge T} - X_{\tau_j^n\wedge T})^2\right]^2.\end{aligned}$$

□

Proposition 2. *Assume $A = 0$. Then, tick time sampling defined as (7) with $\varphi = \mathrm{id}$, $(\mathbb{E}, \mathbb{D}) = (\mathbb{R}, \mathbb{Z})$ and $\epsilon_n^2 = P[\langle X\rangle_T]/n$ is asymptotically efficient in the sense that it asymptotically attains the lower bound (12) among sampling schemes with (1) and $\limsup_{n\to\infty} P[N_T^n]/n \leq 1$.*

Proof. Since $|X_{\tau_{j+1}^n} - X_{\tau_j^n}| = \epsilon_n$ for each $j \geq 1$,

$$\epsilon_n^2 P[N_T^n] = P\left[\mathcal{R}v[\tau^n]_T\right] + O(\epsilon_n^2) = P\left[\langle X\rangle_T\right] + O(\epsilon_n^2)$$

and

$$P\left[((X))_4[\tau^n]_T\right] = P\left[\mathcal{R}q[\tau^n]_T\right] + O(\epsilon_n^4) = \epsilon_n^4 P[N_T^n] + O(\epsilon_n^4). \tag{14}$$

In the light of (13), it suffices to observe

$$P\left[((X))_4[\tau^n]_T\right] = \frac{\epsilon_n^4 P[N_T^n]^2 + O(\epsilon_n^2)}{1+P[N_T^n]} = \frac{P[\langle X\rangle_T]^2}{1+P[N_T^n]} + O(\epsilon_n^4).$$

□

Next, let us treat the asymptotic conditional variance of the realized volatility.

Proposition 3. *In addition to Condition 1, assume the following condition to hold; there exists an $\{\mathcal{F}_t\}$-adapted left continuous process d_s which is locally bounded and bounded away from 0 such that*

$$\mathcal{G}_{j,n}^2 = d_{\tau_j^n}\epsilon_n^2 + o_p(\epsilon_n^2) \tag{15}$$

uniformly in $j = 0, 1, \ldots, N_t^n$ for all $t \in [0, \infty)$. Then, it holds

$$\epsilon_n^2 N_T^n \to \int_0^T d_s^{-1} d\langle X\rangle_s \tag{16}$$

in probability and

$$\left\{\int_0^T f_s \langle X\rangle_s\right\}^2 \leq \int_0^T f_s^2 a_s^2 d\langle X\rangle_s \int_0^T d_s^{-1} d\langle X\rangle_s$$

for every finite stopping time T and locally bounded left continuous adapted process f_s. In particular,

$$\int_0^T a_s^2 d\langle X\rangle \geq \langle X\rangle_T^2 \left\{\int_0^T d_s^{-1} d\langle X\rangle_s\right\}^{-1}. \tag{17}$$

Proof. The convergence (16) follows from (15) and (4). By Jensen's inequality, we have

$$\begin{aligned}\int_0^T f_s^2 a_s^2 d\langle X\rangle_s &= \lim_{n\to\infty} \epsilon_n^{-2} \sum_{j=0}^{N_T^n} f_{\tau_j^n}^2 \mathcal{G}_{j,n}^4 \\ &\geq \lim_{n\to\infty} \epsilon_n^{-2} \sum_{j=0}^{N_T^n} \{f_{\tau_j^n} \mathcal{G}_{j,n}^2\}^2 \\ &\geq \lim_{n\to\infty} \frac{1}{\epsilon_n^2 N_T^n} \left\{\sum_{j=0}^{N_T^n} f_{\tau_j^n} \mathcal{G}_{j,n}^2\right\}^2.\end{aligned}$$

The result then follows from (4) and (16).

□

Note that in tick time sampling (7) with $\varphi = \mathrm{id}$, we have $a_s \equiv 1$ and $d_s \equiv 1$, so that the identity holds in (17). Although d_s depends on sampling scheme by definition, the right hand side of (17) is approximately

$$\epsilon_n^{-2} \langle X\rangle_T^2 / N_T^n$$

so that it can be considered a lower bound for the left hand integral of (17) among sampling schemes of which number of data are N_T^n conditionally to $\mathcal{F}_T$. Consequently, as far as considering the restricted class of the sampling schemes with $b_s \equiv 0$, the asymptotic conditional variance of the realized volatility is minimized by tick time sampling (7) with $\varphi = \mathrm{id}$.

6. Application to the Euler-Maruyama approximation

The asymptotic distribution of the realized volatility is closely related to error distribution of the Euler-Maruyama approximation. In fact, Jacod and Protter [19] study the asymptotic distribution of the realized volatility with the equidistant sampling in the context of error calculus of the Euler-Maruyama scheme. Here we propose alternative sampling schemes for the Euler-Maruyama approximation as an application of the preceding sections. Let us consider the stochastic differential equation

$$\begin{aligned} d\xi_t &= \mu(\eta_t, \xi_t)dt + \sigma(\eta_t, \xi_t)dw_t, \\ d\eta_t &= \beta(\eta_t)dt, \end{aligned}$$

where w is a standard Brownian motion and μ, σ, β are continuously differentiable functions with linear growth. Note that in the case $\beta = 1$,

$$d\xi_t = \mu(t, \xi_t)dt + \sigma(t, \xi_t)dw_t.$$

Since it is rarely possible to generate a path of ξ fast and exactly, (see Beskos, Papaspiliopoulos and Roberts [6] for an exact simulation method), the Euler-Maruyama scheme is widely used to approximate to ξ in simulation. For a sampling scheme τ^n, the Euler-Maruyama approximation ξ^n of ξ is given as

$$\begin{aligned} d\xi_t^n &= \mu(\eta^n_{\psi_n(t)}, \xi^n_{\psi_n(t)})dt + \sigma(\eta^n_{\psi_n(t)}, \xi^n_{\psi_n(t)})dw_t, \\ d\eta_t^n &= \beta(\eta^n_{\psi_n(t)})dt, \end{aligned}$$

where $\psi_n(t) = \tau_j^n$ if $t \in [\tau_j^n, \tau_{j+1}^n)$. The convergence rate of the scheme has been extensively investigated; see e.g., Kloeden and Platen [23] for a well-known strong approximation theorem and Kohatsu-Higa [24], Bally and Talay [1, 2], Konakov and Mammen [25] for weak approximation theorems. Newton [28] treated passage times as sampling scheme. Cambanis and Hu [7] studied efficiency of deterministic nonequidistant sampling schemes. Hofmann, Müller-Gronbach and Ritter [16] treated a class of adaptive sampling schemes. Here we exploit a result of Kurtz and Protter [26], Jacod and Protter [19] to deal with the asymptotic distribution of pathwise error. Our aim here is to propose sampling schemes which are more efficient than the usual equidistant sampling scheme.

Assume that τ^n satisfies Condition 1 with $X = W$. By Theorem 2, there exists a conditionally Gaussian martingale Z such that

$$\epsilon_n^{-1}(((W))_2[\tau^n]_t - t) \Rightarrow Z_t$$

stably. Put $U_t^n = \epsilon_n^{-1}(\xi_t^n - \xi)$. Then, applying Kurtz and Protter [26], U^n converges to a process U which satisfies

$$dU_t = \partial_1\mu(\xi_t, \eta_t)U_t dt + \partial_1\sigma(\xi_t, \eta_t)\left[U_t dW_t - \frac{1}{2}\sigma(\xi_t, \eta_t)dZ_t\right],$$

where ∂_1 refers to the differential operator with respect to the first argument. Solving this stochastic differential equation, we obtain

$$U_T = -\frac{1}{2}e_T\int_0^T e_t^{-1}\sigma(\xi_t, \eta_t)\partial_1\sigma(\xi_t, \eta_t)\left[dZ_t - \partial_1\sigma(\xi_t, \eta_t)d\langle Z, W\rangle_t\right]$$

Therefore, applying Theorem 2, the distribution of U_T is mixed normal with conditional mean

$$-\frac{1}{3}e_T\int_0^T e_t^{-1}\sigma(\xi_t,\eta_t)\partial_1\sigma(\xi_t,\eta_t)b_t\left[dW_t-\partial_1\sigma(\xi_t,\eta_t)dt\right]$$

and conditional variance

$$\frac{1}{6}e_T^2\int_0^T e_t^{-2}\sigma(\xi_t,\eta_t)^2\partial_1\sigma(\xi_t,\eta_t)^2c_t^2dt, \tag{18}$$

where

$$e_t=\exp\left\{\int_0^t\partial_1\mu(\xi_s,\eta_s)ds+\int_0^t\partial_1\sigma(\xi_s,\eta_s)dW_s-\frac{1}{2}\int_0^t\partial_1\sigma(\xi_s,\eta_s)^2ds\right\}.$$

Now, let us first see that space-equidistant sampling scheme defined as (7) with $\varphi=\mathrm{id}$, $\epsilon_n=n^{-1/2}$, $X=W$ is three times efficient than the usual (time-)equidistant sampling scheme $\tau_j^n=j/n$. As already seen, for a deterministic time T,

$$b_s\equiv 0,\quad a_s^2\equiv c_s^2\equiv 1,\quad P[N_T^n]\leq nT$$

in the space-equidistant case, while

$$b_s\equiv 0,\quad a_s^2\equiv c_s^2\equiv 3,\quad N_T^n=[nT]$$

for the time-equidistant case. As mentioned above, Newton [28] studied this space-equidistant sampling scheme; the superiority of this scheme is more-or-less known. Nevertheless, the above simple fact of asymptotic conditional variance has not been recognized so far.

Next, let us consider to minimize (18) among sampling schemes with $b_s\equiv 0$. Define τ^n as

$$\tau_0^n=0,\quad \tau_{j+1}^n=\inf\{t>\tau_j^n;|W_{\tau_{j+1}^n}-W_{\tau_j^n}|^2=\epsilon(\tau_j^n)\}, \tag{19}$$

where

$$\epsilon(\tau_j^n)=\frac{\epsilon_n^2\hat{e}_{\tau_j^n}}{\sigma(\xi_{\tau_j^n}^n,\eta_{\tau_j^n}^n)\partial_1\sigma(\xi_{\tau_j^n}^n,\eta_{\tau_j^n}^n)}$$

and

$$\log(\hat{e}_{\tau_j^n})=\sum_{i=0}^{j-1}\big\{\partial_1\mu(\xi_{\tau_i^n}^n,\eta_{\tau_i^n}^n)(\tau_{i+1}^n-\tau_i^n)+\partial_1\sigma(\xi_{\tau_i^n}^n,\eta_{\tau_i^n}^n)(W_{\tau_{i+1}^n}-W_{\tau_i^n})$$
$$-\frac{1}{2}\partial_1\sigma(\xi_{\tau_i^n}^n,\eta_{\tau_i^n}^n)^2(\tau_{i+1}^n-\tau_i^n)\big\}.$$

Here σ and $\partial_1\sigma$ are assumed to be bounded away from 0. This is an adaptive sampling scheme and it is easy to see that Condition 1 is satisfied with

$$b_s \equiv 0, \quad a_s^2 = c_s^2 = \frac{e_s}{\sigma(\xi_s, \eta_s)\partial_1 \sigma(\xi_s, \eta_s)}.$$

Since (15) also is satisfied with $d_s = a_s^2$, we can apply Proposition 3 to see that this adaptive scheme attains a lower bound for (18) among sampling schemes with (15). In this sense, this sampling scheme is optimal. A disadvantage of this scheme is the difficulty to estimate the expected number of data. In other words, we cannot answer how to choose ϵ_n so that the expected number of data is less than n. In practice, it will be better to use

$$\tau_0^n = 0, \quad \tau_{j+1}^n = \inf\{t > \tau_j^n; |W_{\tau_{j+1}^n} - W_{\tau_j^n}|^2 = \epsilon(\tau_j^n) \vee \epsilon_n'\},$$

for $\epsilon_n' > 0$, or

$$\tau_0^n = 0, \quad \tau_{j+1}^n = \inf\{t > \tau_j^n + \epsilon_n'; |W_{\tau_{j+1}^n} - W_{\tau_j^n}|^2 = \epsilon(\tau_j^n)\}$$

in order to assure a simulation is done in a finite time.

We conclude this section by a remark on generating the random variable (τ, W_τ) satisfying

$$\tau = \inf\{t > 0; |W_t - W_0| = \epsilon\}$$

on a computer for a given ϵ. There is no difficulty in generating W_τ because

$$P[W_\tau = W_0 \pm \epsilon] = 1/2$$

and τ, W_τ are independent. To generate τ, it is sufficient that the distribution function F_ϵ of τ is available because

$$\tau \sim F_\epsilon^{-1}(U),$$

where U is a random variable uniformly distributed on $(0,1)$. It is known that the density of τ is given by

$$\frac{2}{\sqrt{2\pi t^3}} \sum_{n=-\infty}^{\infty} (4n+1)\epsilon \exp\left\{-\frac{(4n+1)^2\epsilon^2}{2t}\right\}$$

See Karatzas and Shreve [22], 2.8.11. Using the fact that

$$\int_0^t \frac{\alpha}{\sqrt{2\pi t^3}} e^{-\alpha^2/2t} dt = 2\int_{b/\sqrt{t}}^{\infty} \phi(x) dx$$

for $\alpha > 0$, we obtain $F_\epsilon(t) = G(\epsilon/\sqrt{t})$, where

$$G(x) = 4\left\{1 - \Phi(x) - \sum_{n=1}^{\infty} (\Phi((4n+1)x) - \Phi((4n-1)x))\right\}.$$

Since

$$\sum_{n=1}^{\infty}(\Phi((4n+1)x)-\Phi((4n-1)x)) \approx \sum_{n=0}^{\infty}(\Phi((4n+3)x)-\Phi((4n+1)x))$$

for small x, and

$$\sum_{n=1}^{\infty}(\Phi((n+1)x)-\Phi(nx)) \approx \frac{1}{2},$$

so that

$$G(x) \approx 4\left\{1-\Phi(x)-\frac{1}{4}\right\}$$

for sufficiently small $x \geq 0$. On the other hand, if $x \geq \delta > 0$, say, the speed of convergence of the infinite series is very fast. We can therefore use

$$G(x) \approx \begin{cases} 4\left\{1-\Phi(x)-\sum_{n=1}^{\lfloor N/x \rfloor}(\Phi((4n+1)x)-\Phi((4n-1)x))\right\} & x \geq \delta, \\ 4(1-\Phi(x))-1 & 0 \leq x < \delta \end{cases}$$

as a valid approximation of G. It is noteworthy that G is independent of ϵ, so that once we obtain the inverse function of G numerically, it is very fast to generate τ repeatedly even if ϵ changes adaptively as in (19).

References

[1] V. Bally and D. Talay (1996), The law of the Euler scheme for stochastic differential equations. I. Convergence rate of the distribution function, Probab. Theory Related Fields 104, no. 1, 43-60.

[2] V. Bally and D. Talay (1996), The law of the Euler scheme for stochastic differential equations. II. Convergence rate of the density, Monte Carlo Methods Appl. 2, no. 2, 93-128.

[3] O.E. Barndorff-Nielsen, S.E. Graverson, J. Jacod, M. Podolskij and N. Shephard (2006), A central limit theorem for realized power and bipower variations of continuous semimartingales, From stochastic calculus to mathematical finance, 33-68, Springer, Berlin.

[4] O.E. Barndorff-Nielsen and N. Shephard (2002), Econometric analysis of realized volatility and its use in estimating stochastic volatility models, J. R. Statist. Soc. Ser. B 64, (2), 253-280.

[5] O.E. Barndorff-Nielsen and N. Shephard (2006), Power variation and time change, Theory Probab. Appl., 50 (1), 1-15.

[6] A. Beskos, O. Papaspiliopoulos and G.O. Roberts (2006), Retrospective exact simulation of diffusion sample paths with applications, Bernoulli 12, no. 6, 1077-1098.

[7] S. Cambanis and Y. Hu (1996), Exact convergence rate of the Euler-Maruyama scheme, with application to sampling design, Stoch. Stoch. rep. 59, 211-240.

[8] R.F. Engle (2000), The econometrics of ultra-high-frequency data, Econometrica, 68, 1-22.

[9] M. Fukasawa (2006), Estimation of the integrated volatility from space-discretized data, preprint.

[10] M. Fukasawa (2007), Central limit theorem for the realized volatility based on tick time sampling, submitted.

[11] M. Fukasawa (2008), On the asymptotic distribution of the realized volatility (in Japanese), to appear in Tokei-Suri.

[12] T. Hayashi, J. Jacod and N. Yoshida (2008), Irregular sampling and central limit theorems for power variations: the continuous case, preprint

[13] V. Genon-Catalot and J. Jacod (1993), On the estimation of the diffusion coefficient for multi-dimensional diffusion processes, Ann. Inst. Henri Poincaré, Probab. Stat. 29, no. 1, 119-151.

[14] J.E. Griffin and R.C.A. Oomen (2006), Sampling returns for realized variance calculation: tick time or transaction time?, preprint.

[15] P.R., Hansen and A. Lunde (2006), Realized variance and market microstructure noise, J. Bus. Econom. Statisit. 24, no.2, 127-218.

[16] N. Hofmann, T. Müller-Gronbach and K. Ritter (2001), The optimal discretization of stochastic differential equations, J. Complexity 17, no. 1, 117-153.

[17] J. Jacod (1994), Limit of random measures associated with the increments of a Brownian semimartingale, unpublished.

[18] J. Jacod (1997), On continuous conditional Gaussian martingales and stable convergence in law, Séminaire Proba. XXXI, Lect. Notes in Math. 1655, 232-246, Springer-Verlag, Berlin.

[19] J. Jacod and P. Protter (1998), Asymptotic error distributions for the Euler method for stochastic differential equations, Ann. Probab. 26, no. 1, 267-307.

[20] J. Jacod and A. N. Shiryaev (2002), Limit theorems for stochastic processes, 2nd ed., Springer-Verlag.

[21] O. Kallenberg (2002), Foundations of modern probability. Second edition. Probability and its Applications (New York). Springer-Verlag, New York.

[22] I. Karatzas and S. Shreve (1991), Brownian motion and stochastic calculus, 2nd ed. Springer-Verlag.

[23] P.E. Kloeden and E. Platen (1992), Numerical solution of stochastic differential equations, Applications of Mathematics (New York), 23. Springer-Verlag, Berlin.

[24] A. Kohatsu-Higa (2001), Weak approximations. A Malliavin calculus approach, Math. Comp. 70, no. 233, 135-172.

[25] V. Konakov and E. Mammen (2002), Edgeworth type expansions for Euler schemes for stochastic differential equations, Monte Carlo Methods Appl. 8, no. 3, 271-285.

[26] T. Kurtz and P. Protter (1991), Wong-Zakai corrections, random evolutions, and simulation schemes for SDEs. Stochastic analysis, 331-346, Academic Press, Boston, MA.

[27] P.A. Mykland and L. Zhang (2006), ANOVA for diffusions, Ann. Statist. 34, no. 4, 1931-1963.

[28] N.J. Newton (1990), An efficient approximation for stochastic differential equations on the partition of symmetrical first passage times. Stoch. Stoch. Rep. 29, no. 2, 227-258.

[29] R.C.A. Oomen (2006), Properties of realized variance under alternative sampling schemes, J. Bus. Econom. Statist. 24, no. 2, 219-237.

[30] H. Rootzén (1980), Limit distributions for the error in approximations of stochastic integrals. Ann. Probab. 8, no. 2, 241-251.

[31] L. Zhang (2005), Efficient estimation of stochastic volatility using noisy observations: a multiscale approach, working paper, University of Illinois and Carnegie Mellon University.

[32] L. Zhang, P.A. Mykland and Y. Aït-Sahalia (2005), A tale of two time scales: Determining integrated volatility with noisy high-frequency data, J. Amer. Stat. Assoc. 100, No.472, 1394-1411.

1-3 Machikaneyama-cho, Toyonaka, Osaka, 560-8531, Japan
E-mail address: fukasawa@sigmath.es.osaka-u.ac.jp