

DISCRETIZATION ERROR OF STOCHASTIC INTEGRALS

MASAAKI FUKASAWA

ABSTRACT. Asymptotic error distribution for approximation of a stochastic integral with respect to continuous semimartingale by Riemann sum with general stochastic partition is studied. Effective discretization schemes of which asymptotic conditional mean-squared error attains a lower bound are constructed. Two applications are given; efficient delta hedging strategies with transaction costs and effective discretization schemes for the Euler-Maruyama approximation are constructed.

1. INTRODUCTION

The present article studies the asymptotic distribution of a sequence of continuous processes $Z^n = \{Z_t^n\}_{t \in [0, T]}$ defined as

$$(1) \quad Z_t^n = \int_0^t X_s dY_s - \sum_{j=0}^{\infty} X_{\tau_j^n} (Y_{\tau_{j+1}^n \wedge t} - Y_{\tau_j^n \wedge t})$$

for one-dimensional continuous semimartingales $X = \{X_t, \mathcal{F}_t\}$, $Y = \{Y_t, \mathcal{F}_t\}$ and sequences of $\{\mathcal{F}_t\}$ -stopping times $\tau^n = \{\tau_j^n\}$ with

$$(2) \quad 0 = \tau_0^n < \tau_1^n < \dots < \tau_j^n < \dots, \quad \lim_{j \rightarrow \infty} \tau_j^n = T \text{ a.s.},$$

where $T \in (0, \infty]$ is fixed and we suppose that the intervals $\tau_{j+1}^n - \tau_j^n$ converge to 0 as $n \rightarrow \infty$ in a sense specified later. Taking into mind the definition of stochastic integrals, the asymptotic behavior of Z^n is theoretically of interest. There may be no need to explain its practical importance. This problem was studied by Rootzén [21] in the case that Y is a Brownian motion and the asymptotic distribution was specified in the case that $\tau_j^n = j/n$ and $X_s = f(Y_s, s)$ for a smooth function f . Jacod [9] treated a related problem on the condition that each interval $\tau_{j+1}^n - \tau_j^n$ is $\mathcal{F}_{\tau_j^n}$ measurable. Jacod and Protter [10] considered the case $X = Y$ and $\tau_j^n = j/n$ and derived the asymptotic error distribution of the Euler-Maruyama scheme for stochastic differential equations. Hayashi and Mykland [7] discussed this problem again for the case $\tau_j^n = j/n$ in the context of discrete-time hedging error. Geiss and

Date: 2009-12-21.

Masaaki Fukasawa

Center for the Study of Finance and Insurance, Osaka University
Japan Science and Technology Agency, CREST
1-3 Machikaneyama, Toyonaka, Osaka, Japan.

Toivola [6] treated an irregular deterministic discretization scheme. The condition $\tau_j^n = j/n$, or more generally, that each interval $\tau_{j+1}^n - \tau_j^n$ is $\mathcal{F}_{\tau_j^n}$ measurable, played an important role in those preceding studies of central limit theorem. On the other hand, Karandikar [12] constructed a discretization scheme τ^n such that Z_t^n converges to 0 almost surely. Since the almost sure convergence is not usually attained by the time-equidistant scheme $\tau_j^n = j/n$, it means Karandikar's scheme is more effective in a sense. The scheme is defined using passage times of X , so that the measurability condition, that each interval $\tau_{j+1}^n - \tau_j^n$ is $\mathcal{F}_{\tau_j^n}$ measurable, is not satisfied. Recently, Fukasawa [4] gave the asymptotic distribution for such a scheme and Fukasawa [5] extended the results to a more general scheme in the case $X = Y$. The present article extends those limit theorems and constructs effective discretization schemes of which asymptotic conditional mean-squared error attains a lower bound. In particular, Karandikar's scheme is shown to be superior to the time-equidistant one also in terms of mean-squared error. An application to delta hedging with transaction costs is given, which can be directly used in practice. It remains for further research to extend the results to discontinuous semimartingales. In Section 2, we describe and prove main results. Effective discretization schemes are constructed in Section 3. The application to hedging is presented in Section 4. An application to the Euler-Maruyama approximation is given in Section 5.

2. CENTRAL LIMIT THEOREM

2.1. Notation and conditions. Here we give a rigorous formulation and describe several conditions on X , Y and τ^n . Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ be a filtered probability space. The filtration $\{\mathcal{F}_t\}$ is assumed to satisfy the usual conditions. We denote by $F_1 \cdot F_2$ the Stieltjes integral or the stochastic integral of F_1 with respect to F_2 .

Let us recall the definition of stable convergence. Let \mathbb{E} be a complete separable metric space and F^n be a sequence of \mathbb{E} -valued random variables defined on (Ω, \mathcal{F}, P) .

Definition 2.1. For a sub σ -field $\mathcal{G} \subset \mathcal{F}$, we say F^n converges \mathcal{G} -stably if for all \mathcal{G} -measurable random variable F_0 , the joint distribution (F^n, F_0) converges in law.

Our main results stated in the next subsection are stable convergences of Z^n defined as (1) with continuous semimartingales X and Y . Notice that a stable convergence is stable against, in particular, the usual localization procedure as well as the Girsanov-Maruyama transformation.

Denote by \mathcal{P} and \mathcal{P}_0 the set of the predictable processes and the set of locally bounded left-continuous adapted processes respectively. Let $T \in (0, \infty]$ be fixed. Given a continuous semimartingale M and $k \in \mathbb{N}$, put

$$\mathcal{P}_M^k = \left\{ H \in \mathcal{P}; |H|^k \cdot \langle M \rangle_t < \infty \text{ for all } t \in [0, T] \right\}.$$

Denote by \mathcal{S} the set of the continuous semimartingales (X, Y, M) satisfying the following Condition 2.2.

Condition 2.2. *There exist $\psi, \varphi, \kappa \in \mathcal{P}_M^2$ and a locally bounded predictable process $\gamma \in \mathcal{P}$ such that*

$$X = X_0 + \psi \cdot \langle M \rangle + \gamma \cdot M, \quad Y = Y_0 + \varphi \cdot \langle M \rangle + M^Y$$

on $[0, T)$, where M^Y is a continuous local martingale with

$$\langle M^Y \rangle = \kappa \cdot \langle M \rangle.$$

In addition, M is a continuous local martingale with $E[\langle M \rangle_T^6] < \infty$.

The integrability of $\langle M \rangle_T$ is not restrictive in light of the localization procedure. In order to describe conditions on τ^n , put

$$G_{j,n}^1 = E[|M_{\tau_{j+1}^n} - M_{\tau_j^n}| | \mathcal{F}_{\tau_j^n}], \quad G_{j,n}^k = E[(M_{\tau_{j+1}^n} - M_{\tau_j^n})^k | \mathcal{F}_{\tau_j^n}]$$

for a given continuous local martingale M with $E[\langle M \rangle_T^6] < \infty$ and $k \in \mathbb{N}$ with $2 \leq k \leq 12$. In addition, put

$$(3) \quad N[\tau^n]_\tau = \max\{j \geq 0; \tau_j^n \leq \tau\}$$

for a given stopping time τ . Denote by $\mathcal{T}(M)$ the set of the sequences of stopping times $\{\tau^n\}$ satisfying (2) and the following Condition 2.3.

Condition 2.3. *There exist a sequence ϵ_n with $\epsilon_n \rightarrow 0$ and $a, b \in \mathcal{P}_0$ such that*

$$G_{j,n}^4/G_{j,n}^2 = a_{\tau_j^n}^2 \epsilon_n^2 + o_p(\epsilon_n^2), \quad G_{j,n}^3/G_{j,n}^2 = b_{\tau_j^n} \epsilon_n + o_p(\epsilon_n)$$

and

$$G_{j,n}^6/G_{j,n}^2 = O_p(\epsilon_n^4), \quad G_{j,n}^{12}/G_{j,n}^2 = o_p(\epsilon_n^8),$$

uniformly in $j = 0, 1, \dots, N[\tau^n]_t$ for all $t \in [0, T)$.

Condition 2.3 is slightly stronger than Condition 1 of Fukasawa [5]; nevertheless, all examples given by Fukasawa [5] satisfy also this condition. Here, ϵ_n serves as a scale of increments of M . Note that $G_{j,n}^4/G_{j,n}^2 = O_p(\epsilon_n^2)$ implies $G_{j,n}^3/G_{j,n}^2 = O_p(\epsilon_n)$ by Lemma B.1. In usual cases, we have $G_{j,n}^{2k} = O_p(\epsilon_n^{2k})$, which in fact holds, for example, if $d\langle M \rangle_t/dt$ exists and is bounded and if $\tau_{j+1}^n - \tau_j^n$ is of $O(\epsilon_n^2)$ uniformly. Condition 2.3 is therefore a quite mild condition in the context of high-frequency asymptotics. It is often easily verified by using the Dambis-Dubins-Schwarz time-change technique for martingales when τ^n is a function of the path of M . Once it reduces to the Brownian motion case by the time-change, one can utilize various results on Brownian stopping times. See Fukasawa [5] for examples. In light of the Skorokhod stopping problem, the distribution of an increment can be any centered distribution with a suitable moment condition. The left continuity of a^2 and b corresponds to a local homogeneity property of the distributions of increments. It should be noted that $\sup_j |\tau_{j+1}^n \wedge t - \tau_j^n \wedge t| \rightarrow 0$ does not

follow from Condition 2.3 nor needed for our main results. See Lemma A.4 for what follows instead.

Denote by $\mathcal{T}_1(M)$, $\mathcal{T}_2(M)$ the subsets of $\mathcal{T}(M)$ satisfying the following Condition 2.4 and Condition 2.5 respectively.

Condition 2.4. *In addition to Condition 2.3, there exists $\zeta \in \mathcal{P}_0$ such that*

$$\zeta^{-1} \in \mathcal{P}_0, \quad \epsilon_n G_{j,n}^1 / G_{j,n}^2 = \zeta_{\tau_j^n} + o_p(1)$$

uniformly in $j = 0, 1, \dots, N[\tau^n]_t$ for all $t \in [0, T]$.

Condition 2.5. *In addition to Condition 2.3, there exists $q \in \mathcal{P}_0$ such that*

$$q^{-1} \in \mathcal{P}_0, \quad G_{j,n}^2 = q_{\tau_j^n}^2 \epsilon_n^2 + o_p(\epsilon_n^2)$$

uniformly in $j = 0, 1, \dots, N[\tau^n]_t$ for all $t \in [0, T]$.

Finally, for τ^n with (2) and $t \in [0, T]$, put

$$(4) \quad [M]_{j,n}^t = \langle M \rangle_{\tau_{j+1}^n \wedge t} - \langle M \rangle_{\tau_j^n \wedge t}.$$

2.2. Main results. Here we state general results on the asymptotic distribution of Z^n . The proofs are deferred to Section 2.3.

Theorem 2.6. *Let $(X, Y, M) \in \mathcal{S}$, $\tau^n \in \mathcal{T}(M)$ and Z^n be defined as (1). Assume one of the following two conditions to hold:*

- *M is the local martingale part of X , that is,*

$$(5) \quad \gamma \equiv 1.$$

- *for all $t \in [0, T]$,*

$$(6) \quad E \left[\sum_{j=0}^{\infty} \|[M]_{j,n}^t\|^k \right] = O(\epsilon_n^{2(k-1)})$$

for $k \in \{1, 2, 3, 4, 5\}$, where $[M]_{j,n}^t$ is defined as (4).

Then, Z^n / ϵ_n converges \mathcal{F} -stably to the law of

$$(7) \quad \frac{1}{3}(b\gamma) \cdot Y + \frac{1}{\sqrt{6}}(c\gamma) \cdot Y'$$

as a $C[0, T]$ -valued sequence, where

$$(8) \quad c^2 = a^2 - \frac{2}{3}b^2, \quad Y' = W_{\langle Y \rangle}$$

and W is a standard Brownian motion which is independent to \mathcal{F} .

Note that the asymptotic distribution (7) is an \mathcal{F} -conditionally Gaussian process, so that the marginal law is a mixed normal distribution. The following theorems give lower bounds for the conditional variance of the mixed normal distribution.

Theorem 2.7. Let $(X, Y, M) \in \mathcal{S}$, $\tau^n \in \mathcal{T}_1(M)$ and Z^n be defined as (1). Let $u \in \mathcal{P}_0$ and put

$$U_t^n = \sum_{j=0}^{\infty} |u_{\tau_j^n}| |M_{\tau_{j+1}^n \wedge t} - M_{\tau_j^n \wedge t}|.$$

Then, it holds

$$(9) \quad \epsilon_n U_t^n \rightarrow U_t := (|u|\zeta) \cdot \langle M \rangle_t$$

in probability for all $t \in [0, T)$. Moreover, if (5) or (6) holds, then $U^n Z^n$ converges \mathcal{F} -stably to the law of UZ as a $D[0, T)$ -valued sequence, where Z is defined as (7). The asymptotic conditional variance V_t of $U_t^n Z_t^n$ with $t \in [0, T)$ satisfies

$$(10) \quad V_t = \frac{1}{6} (c\gamma)^2 \cdot \langle Y \rangle_t (|u|\zeta) \cdot \langle M \rangle_t^2 \geq \frac{1}{6} \left((|u\gamma|^{2/3} \kappa^{1/3}) \cdot \langle M \rangle_t \right)^3 \text{ a.s..}$$

Theorem 2.8. Let $(X, Y, M) \in \mathcal{S}$, $\tau^n \in \mathcal{T}_2(M)$ and Z^n be defined as (1). Then, it holds that

$$(11) \quad \epsilon_n^2 N[\tau^n]_t \rightarrow N_t := q^{-2} \cdot \langle M \rangle_t$$

in probability for all $t \in [0, T)$. Moreover, if (5) or (6) holds, then $\sqrt{N[\tau^n]} Z^n$ converges \mathcal{F} -stably to the law of $\sqrt{N}Z$ as a $D[0, T)$ -valued sequence, where Z is defined as (7). The asymptotic conditional variance V_t of $\sqrt{N[\tau^n]} Z_t^n$ with $t \in [0, T)$ satisfies

$$(12) \quad V_t = \frac{1}{6} (c\gamma)^2 \cdot \langle Y \rangle_t q^{-2} \cdot \langle M \rangle_t \geq \frac{1}{6} \left\{ (|\gamma| \sqrt{\kappa}) \cdot \langle M \rangle_t \right\}^2 \text{ a.s..}$$

Note that the right hand sides of (10), (12) do not depend on the discretization scheme τ^n . In Section 3, we construct schemes $\tau^n \in \mathcal{T}_1(M) \cap \mathcal{T}_2(M)$ which attain the lower bounds (10), (12) respectively. Its practical meaning is discussed in Sections 4 and 5.

The condition (6) will be easily verified especially if M is a Brownian motion. Condition 2.3 is then also easily verified if, in addition, τ^n satisfies the condition that each interval $\tau_{j+1}^n - \tau_j^n$ is $\mathcal{F}_{\tau_j^n}$ measurable. It is therefore not difficult to recover the preceding results from Theorem 2.6. An irregular scheme treated in Geiss and Toivola [6] is an example.

2.3. Proof for Theorems. Here we give proofs for main results stated in the previous subsection.

Lemma 2.9. Let M be a continuous local martingale with $E[\langle M \rangle_T^6] < \infty$ and $\tau^n \in \mathcal{T}(M)$. Let H, γ be a locally bounded cag adapted process and a locally bounded predictable process respectively. Put $\bar{M} = \gamma \cdot M$ and define H^n, \bar{M}^n as

$$(13) \quad H_s^n = H_{\tau_j^n}, \quad \bar{M}_s^n = \bar{M}_{\tau_j^n}, \quad \text{for } j \geq 0 \text{ with } s \in [\tau_j^n, \tau_{j+1}^n).$$

Assume one of the following two conditions:

- $\gamma \equiv 1$,
- (6) holds.

Then, it holds

$$(14) \quad \begin{aligned} \epsilon_n^{-1}((\bar{M} - \bar{M}^n)H^n) \cdot \langle M \rangle_t &\rightarrow \frac{1}{3}(bH\gamma) \cdot \langle M \rangle_t, \\ \epsilon_n^{-2}((\bar{M} - \bar{M}^n)^2 H^n) \cdot \langle M \rangle_t &\rightarrow \frac{1}{6}(a^2 H \gamma^2) \cdot \langle M \rangle_t \end{aligned}$$

uniformly in t on compact sets of $[0, T]$ in probability. Moreover,

$$(15) \quad \epsilon_n^{-4}((\bar{M} - \bar{M}^n)^4 H^n) \cdot \langle M \rangle_t = O_p(1)$$

for all $t \in [0, T]$.

Proof: By the usual localization argument, we can assume H, γ, a, b, M and $\langle M \rangle$ are bounded without loss of generality. Let us suppose $\gamma \equiv 1$. Then, for any $l \in \mathbb{Z}$,

$$\begin{aligned} &\epsilon_n^{-l}((\bar{M} - \bar{M}^n)^l H^n) \cdot \langle M \rangle_t \\ &= \epsilon_n^{-l} \sum_{j=0}^{\infty} H_{\tau_j^n} \left\{ \alpha_l (M_{\tau_{j+1}^n \wedge t} - M_{\tau_j^n \wedge t})^{l+2} + \beta_j \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} (M_s - M_{\tau_j^n})^{l+1} dM_s \right\} \end{aligned}$$

by Itô's formula, where α_l, β_l are constants only depending on l and, in particular, $\alpha_1 = 1/3, \alpha_2 = 1/6$. Since

$$\begin{aligned} &\epsilon_n^{-2l} \sum_{j=0}^{N[\tau^n]_t} H_{\tau_j^n}^2 E \left[\int_{\tau_j^n}^{\tau_{j+1}^n} (M_s - M_{\tau_j^n})^{2l+2} d\langle M \rangle_s \middle| \mathcal{F}_{\tau_j^n} \right] \\ &= \epsilon_n^{-2l} \alpha_{2l+2} \sum_{j=0}^{N[\tau^n]_t} H_{\tau_j^n}^2 G_{j,n}^{2l+4} \rightarrow 0 \end{aligned}$$

for $l \in 1, 2, 4$ by Condition 2.3 and Lemma A.4, we have

$$\epsilon_n^{-l} \sum_{j=0}^{\infty} H_{\tau_j^n} \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} (M_s - M_{\tau_j^n})^{l+1} dM_s \rightarrow 0$$

as well as

$$\epsilon_n^{-l} \sum_{j=0}^{\infty} H_{\tau_j^n} (M_{\tau_{j+1}^n \wedge t} - M_{\tau_j^n \wedge t})^{l+2} - \epsilon_n^{-l} \sum_{j=0}^{N[\tau^n]_t} H_{\tau_j^n} G_{j,n}^{l+2} \rightarrow 0$$

by Lemma A.2. The result then follows from Lemma A.4.

Next, let us suppose (6) and $\gamma \neq 1$. Note that for all $\delta > 0$, there exists a bounded cag process γ^δ such that

$$E[|\gamma - \gamma^\delta|^k \cdot \langle M \rangle_t] < \delta$$

for $k = 4, 6, 12$ by Lemma A.3. Notice that for $p, q > 1$ with $1/p + 1/q = 1$,

$$\begin{aligned} & E \left[\epsilon_n^{-l} \sum_{j=0}^{\infty} \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} \left| \int_{\tau_j^n \wedge t}^s (\gamma_u - \gamma_u^\delta) dM_u \right|^l d\langle M \rangle_s \right] \\ & \leq E \left[\epsilon_n^{-l} \sum_{j=0}^{\infty} \sup_{s \in [\tau_j^n \wedge t, \tau_{j+1}^n \wedge t]} \left| \int_{\tau_j^n \wedge t}^s (\gamma_u - \gamma_u^\delta) dM_u \right|^l [M]_{j,n}^t \right] \\ & \leq C \epsilon_n^{-l} E \left[\sum_{j=0}^{\infty} \left| \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} (\gamma_u - \gamma_u^\delta)^2 d\langle M \rangle_u \right|^{pl/2} \right]^{1/p} E \left[\sum_{j=0}^{\infty} [M]_{j,n}^t \right]^{1/q} \end{aligned}$$

by the Hölder and the Burkholder-Davis-Gundy inequality. Furthermore,

$$\begin{aligned} & E \left[\sum_{j=0}^{\infty} \left| \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} (\gamma_u - \gamma_u^\delta)^2 d\langle M \rangle_u \right|^{pl/2} \right] \\ & \leq E \left[\sum_{j=0}^{\infty} \left| \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} |\gamma_u - \gamma_u^\delta|^{2pl} d\langle M \rangle_u [M]_{j,n}^t \right|^{1/2} \right] \\ & \leq E \left[\int_0^t |\gamma_u - \gamma_u^\delta|^{2pl} d\langle M \rangle_u \right]^{1/2} E \left[\sum_{j=0}^{\infty} [M]_{j,n}^t \right]^{1/2}. \end{aligned}$$

For $l = 1$, we take $p = 2$ and for $l = 2, 4$ we take $p = 3/2$ to have

$$E \left[\sum_{j=0}^{\infty} [M]_{j,n}^t \right]^{1/(2p)} E \left[\sum_{j=0}^{\infty} [M]_{j,n}^t \right]^{1/q} = O(\epsilon_n^l)$$

by the assumption (6). Since δ can be arbitrarily small, this estimate ensures that we can replace \bar{M} and γ by $\gamma^\delta \cdot M$ and γ^δ respectively in (14) and (15). Put

$$\gamma_s^{\delta,n} = \gamma_{\tau_j^n}^\delta \text{ for } s \in [\tau_j^n, \tau_{j+1}^n).$$

By the same argument, we can estimate

$$E \left[\epsilon_n^{-l} \sum_{j=0}^{\infty} \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} \left| \int_{\tau_j^n \wedge t}^s (\gamma_u^\delta - \gamma_u^{\delta,n}) dM_u \right|^l d\langle M \rangle_s \right]$$

to ensure that we can replace $(\bar{M} - \bar{M}^n)$ with $\gamma^{\delta,n}(M - M^n)$ in (14) and (15), where M^n is defined as

$$M_s^n = M_{\tau_j^n}, \text{ for } j \geq 0 \text{ with } s \in [\tau_j^n, \tau_{j+1}^n).$$

Then, repeat the proof for the case $\gamma \equiv 1$ with H replaced with $H\gamma^\delta$. // //

Proof of Theorem 2.6: Put $\bar{M} = \gamma \cdot M$, $A = \psi \cdot \langle M \rangle$ and define A^n, \bar{M}^n as (13) with $H = A$. Then we have

$$Z^n = (A - A^n) \cdot Y + ((\bar{M} - \bar{M}^n)\varphi) \cdot \langle M \rangle + (\bar{M} - \bar{M}^n) \cdot M^Y.$$

We shall prove that

$$\epsilon_n^{-1}(A - A^n) \cdot Y_t \rightarrow 0, \quad \epsilon_n^{-1}((\bar{M} - \bar{M}^n)\varphi) \cdot \langle M \rangle_t \rightarrow \frac{1}{3}(b\varphi) \cdot \langle M \rangle_t$$

in probability uniformly in t on compact sets of $[0, T)$ and that

$$(16) \quad D^n := \epsilon_n^{-1}(\bar{M} - \bar{M}^n) \cdot M^Y$$

converges \mathcal{F} -stably to

$$\frac{1}{3}(b\gamma) \cdot M^Y + \frac{1}{\sqrt{6}}(c\gamma) \cdot Y'.$$

Step a) Let us show

$$(17) \quad \epsilon_n^{-1}(A - A^n) \cdot Y_v \rightarrow 0$$

uniformly in $v \in [0, t]$ in probability. Fix $\delta_1, \delta_2 > 0$ arbitrarily and take a bounded cag process ψ^δ such that

$$P[|\psi - \psi^\delta|^2 \cdot \langle M \rangle_t > \delta_1] < \delta_2$$

by Lemma A.3. Observe that for any $v \in [0, t]$,

$$(18) \quad \begin{aligned} & \epsilon_n^{-1} \sum_{j=0}^{\infty} \int_{\tau_j^n \wedge v}^{\tau_{j+1}^n \wedge v} \int_{\tau_j^n \wedge v}^{\cdot v} \psi_u d\langle M \rangle_u \varphi_s d\langle M \rangle_s \\ &= \epsilon_n^{-1} \sum_{j=0}^{\infty} \int_{\tau_j^n \wedge v}^{\tau_{j+1}^n \wedge v} \int_{\tau_j^n \wedge v}^{\cdot s} \psi_u^\delta d\langle M \rangle_u \varphi_s d\langle M \rangle_s \\ &+ \epsilon_n^{-1} \sum_{j=0}^{\infty} \int_{\tau_j^n \wedge v}^{\tau_{j+1}^n \wedge v} \int_{\tau_j^n \wedge v}^{\cdot s} (\psi_u - \psi_u^\delta) d\langle M \rangle_u \varphi_s d\langle M \rangle_s \end{aligned}$$

and that

$$\begin{aligned} & \epsilon_n^{-1} \sum_{j=0}^{\infty} \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} \int_{\tau_j^n \wedge t}^{\cdot s} |\psi_u^\delta| d\langle M \rangle_u |\varphi_s| d\langle M \rangle_s \\ & \leq C^\delta \epsilon_n^{-1} \sum_{j=0}^{\infty} [M]_{j,n}^t \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} |\varphi_u| d\langle M \rangle_u \\ & \leq C^\delta \epsilon_n^{-1} \sum_{j=0}^{\infty} \left\{ |[M]_{j,n}^t|^3 \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} |\varphi_u|^2 d\langle M \rangle_u \right\}^{1/2} \\ & \leq C^\delta \left\{ \epsilon_n^{-2} \sum_{j=0}^{\infty} |[M]_{j,n}^t|^3 \right\}^{1/2} \left\{ \int_0^t |\varphi_u|^2 d\langle M \rangle_u \right\}^{1/2} \end{aligned}$$

for a constant C^δ . Using Lemmas A.2, A.4 and the Burkholder-Davis-Gundy inequality, we have

$$\epsilon_n^{-2} \sum_{j=0}^{\infty} |[M]_{j,n}^t|^3 \rightarrow 0$$

since

$$\epsilon_n^{-2} \sum_{j=0}^{N[\tau^n]_t} G_{j,n}^6 \rightarrow 0, \quad \epsilon_n^{-4} \sum_{j=0}^{N[\tau^n]_t} G_{j,n}^{12} \rightarrow 0$$

in probability by Condition 2.3. Hence the first term of the right hand side of (18) converges to 0 uniformly in $v \in [0, t]$ in probability. For the second term, we have

$$\begin{aligned} & \epsilon_n^{-1} \sum_{j=0}^{\infty} \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} \int_{\tau_j^n \wedge t}^s |\psi_u - \psi_u^\delta| d\langle M \rangle_u |\varphi_s| d\langle M \rangle_s \\ & \leq \epsilon_n^{-1} \sum_{j=0}^{\infty} \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} |\psi_u - \psi_u^\delta| d\langle M \rangle_u \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} |\varphi_s| d\langle M \rangle_s \\ & \leq \sqrt{|\psi - \psi^\delta|^2 \cdot \langle M \rangle_t} \epsilon_n^{-1} \sum_{j=0}^{\infty} \sqrt{[M]_{j,n}^t} \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} |\varphi_s| d\langle M \rangle_s \\ & \leq \sqrt{|\psi - \psi^\delta|^2 \cdot \langle M \rangle_t} \epsilon_n^{-1} \sum_{j=0}^{\infty} \left\{ |[M]_{j,n}^t|^2 \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} |\varphi_s|^2 d\langle M \rangle_s \right\}^{1/2} \\ & \leq \sqrt{|\psi - \psi^\delta|^2 \cdot \langle M \rangle_t} \left\{ \sum_{j=0}^{\infty} \epsilon_n^{-2} |[M]_{j,n}^t|^2 \right\}^{1/2} \left\{ \int_0^t |\varphi_s|^2 d\langle M \rangle_s \right\}^{1/2}. \end{aligned}$$

Using again Lemmas A.2, A.4 and the Burkholder-Davis-Gundy inequality, we have

$$\epsilon_n^{-2} \sum_{j=0}^{\infty} |[M]_{j,n}^t|^2 = O_p(1).$$

Since δ_1, δ_2 can be arbitrarily small, the left hand sum of (18) converges to 0 uniformly in $v \in [0, t]$ in probability. Next, observe that

$$\langle (A - A^n) \cdot M^Y \rangle = ((A - A^n)^2 \kappa) \cdot \langle M \rangle,$$

so that it suffices for (17) to prove

$$\epsilon_n^{-2} ((A - A^n)^2 \kappa) \cdot \langle M \rangle_t \rightarrow 0$$

in probability, in light of the Lenglart inequality. This follows from

$$\begin{aligned}
& \epsilon_n^{-2} \sum_{j=0}^{\infty} \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} \left\{ \int_{\tau_j^n \wedge t}^s \psi_u d\langle M \rangle_u \right\}^2 \kappa_s d\langle M \rangle_s \\
& \leq \epsilon_n^{-2} \sum_{j=0}^{\infty} \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} |\psi_u|^2 d\langle M \rangle_u |[M]_{j,n}^t| \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} \kappa_s d\langle M \rangle_s \\
& \leq \sup_{j \geq 0} \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} |\psi_u|^2 d\langle M \rangle_u \left\{ \epsilon_n^{-4} \sum_{j=0}^{\infty} |[M]_{j,n}^t|^3 \right\}^{1/2} \left\{ \int_0^t \kappa_s^2 d\langle M \rangle_s \right\}^{1/2},
\end{aligned}$$

and noticing that

$$\sup_{j \geq 0} \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} |\psi_u|^2 d\langle M \rangle_u \rightarrow 0, \quad \epsilon_n^{-4} \sum_{j=0}^{\infty} |[M]_{j,n}^t|^3 = o_p(1)$$

by Lemmas A.2, A.4 and Condition 2.3.

Step b) Let us show that

$$(19) \quad \epsilon_n^{-1} ((\bar{M} - \bar{M}^n) \varphi) \cdot \langle M \rangle_v \rightarrow \frac{1}{3} (b\varphi\gamma) \cdot \langle M \rangle_v$$

uniformly in $v \in [0, t]$ in probability. Fix $\delta_1, \delta_2 > 0$ arbitrarily and take a bounded cag adapted process φ^δ such that

$$P[|\varphi - \varphi^\delta|^2 \cdot \langle M \rangle_t > \delta_1] < \delta_2$$

by Lemma A.3. Notice that

$$(20) \quad |\epsilon_n^{-1} ((\bar{M} - \bar{M}^n)(\varphi - \varphi^\delta)) \cdot \langle M \rangle_v| \leq \sqrt{|\varphi - \varphi^\delta|^2 \cdot \langle M \rangle_t} \sqrt{\epsilon_n^{-2} (\bar{M} - \bar{M}^n)^2 \cdot \langle M \rangle_t}$$

and

$$\epsilon_n^{-2} (\bar{M} - \bar{M}^n)^2 \cdot \langle M \rangle_t = O_p(1)$$

by Lemma 2.9. Note also that

$$(21) \quad |(b\varphi) \cdot \langle M \rangle_v - (b\varphi^\delta) \cdot \langle M \rangle_v| \leq \sqrt{b^2 \cdot \langle M \rangle_t} \sqrt{|\varphi - \varphi^\delta|^2 \cdot \langle M \rangle_t}.$$

Since δ_1, δ_2 can be arbitrarily small, the estimates (20) and (21) ensures that we can suppose φ is a bounded cag adapted process without loss of generality. Then, putting

$$\varphi_s^n = \varphi_{\tau_j^n} \text{ for } j \geq 0 \text{ with } s \in [\tau_j^n, \tau_{j+1}^n),$$

we have also that

$$|\epsilon_n^{-1} ((\bar{M} - \bar{M}^n)(\varphi - \varphi^n)) \cdot \langle M \rangle_v| \leq \sqrt{|\varphi - \varphi^n|^2 \cdot \langle M \rangle_t} \sqrt{\epsilon_n^{-2} (\bar{M} - \bar{M}^n)^2 \cdot \langle M \rangle_t}.$$

Note that

$$|\varphi - \varphi^n|^2 \cdot \langle M \rangle_t \rightarrow 0$$

in probability as $n \rightarrow \infty$ because φ is now assumed to be bounded and left continuous. Applying Lemma 2.9, we have (19).

Step c) Let us study the asymptotic distribution of D^n defined as (16). Put

$$\hat{D}^n = D^n - \frac{1}{3}(b\gamma) \cdot M^Y.$$

In light of Theorem A.1, it suffices to show the following convergences in probability.

$$(1) \quad \langle \hat{D}^n, M^Y \rangle_t \rightarrow 0,$$

$$(2) \quad \langle \hat{D}^n \rangle_t \rightarrow \frac{1}{6}(c^2\gamma^2\kappa) \cdot \langle M \rangle_t,$$

$$(3) \quad \langle \hat{D}^n, \hat{M} \rangle_t \rightarrow 0,$$

for all $t \in [0, T]$ and for all bounded martingale \hat{M} orthogonal to M^Y . The last one is trivial. In order to see the first convergence, it suffices to see

$$\langle D^n, M^Y \rangle_t = \epsilon_n^{-1}((\bar{M} - \bar{M}^n)\kappa) \cdot \langle M \rangle_t \rightarrow \frac{1}{3}(b\gamma\kappa) \cdot \langle M \rangle_t$$

in probability. This is shown in the same manner as for (19). In order to see the second convergence, fix $\delta_1, \delta_2 > 0$ arbitrarily and take a bounded cag process κ^δ such that

$$P[|\kappa - \kappa^\delta|^2 \cdot \langle M \rangle_t > \delta_1] < \delta_2$$

by Lemma A.3. Notice that

$$(22) \quad \langle D^n \rangle = \epsilon_n^{-2}((\bar{M} - \bar{M}^n)^2\kappa^\delta) \cdot \langle M \rangle + \epsilon_n^{-2}((\bar{M} - \bar{M}^n)^2(\kappa - \kappa^\delta)) \cdot \langle M \rangle$$

and the second term is negligible since

$$\epsilon_n^{-2}((\bar{M} - \bar{M}^n)^2|\kappa - \kappa^\delta|) \cdot \langle M \rangle_t \leq \sqrt{\epsilon_n^{-4}(\bar{M} - \bar{M}^n)^4 \cdot \langle M \rangle_t} \sqrt{|\kappa - \kappa^\delta|^2 \cdot \langle M \rangle_t}$$

in light of Lemma 2.9 and the fact that δ_1, δ_2 can be arbitrarily small. Furthermore, putting

$$\kappa_s^{\delta,n} = \kappa_{\tau_j^n}^\delta \text{ for } j \geq 0 \text{ with } s \in [\tau_j^n, \tau_{j+1}^n),$$

we can replace κ^δ with $\kappa^{\delta,n}$ in the first term of (22) by the same argument. Then, we have from Lemma 2.9 that

$$\langle D^n \rangle_t \rightarrow \frac{1}{6}(a^2\gamma^2\kappa) \cdot \langle M \rangle_t$$

in probability. Since

$$\langle (b\gamma) \cdot M^Y \rangle = (b^2\gamma^2\kappa) \cdot \langle M \rangle,$$

it remains only to show

$$\langle D^n, (b\gamma) \cdot M^Y \rangle_t \rightarrow \frac{1}{3}(b^2\gamma^2\kappa) \cdot \langle M \rangle_t$$

in probability. Since the left hand side is

$$\epsilon_n^{-1}((\bar{M} - \bar{M}^n)b\gamma\kappa) \cdot \langle M \rangle_t,$$

the convergence follows from the same argument as for (19). ////

Proof of Theorem 2.7: The convergence (9) follows from Condition 2.4 and Lemma A.4. The convergence of $U^n Z^n$ in $D[0, T]$ is a consequence of the fact that the convergence of Z^n/ϵ_n is stable. To show (10), we first notice that

$$G_{j,n}^4/G_{j,n}^2 - \frac{3}{4}|G_{j,n}^3/G_{j,n}^2|^2 \geq |G_{j,n}^2/G_{j,n}^1|^2 \quad \text{a.s.},$$

which follows from Lemma B.2. In light of Lemma A.4 and Condition 2.4, this inequality implies

$$(Hc^2) \cdot \langle M \rangle \geq (H\zeta^{-2}) \cdot \langle M \rangle \quad \text{a.s.}$$

for any $H \in \mathcal{P}_M^1$. Thus we have

$$\begin{aligned} V_t &= \frac{1}{6}(c^2\gamma^2) \cdot \langle Y \rangle_t (|u|\zeta) \cdot \langle M \rangle_t^2 \\ &\geq \frac{1}{6}(\zeta^{-2}\gamma^2) \cdot \langle Y \rangle_t (|u|\zeta) \cdot \langle M \rangle_t^2 \geq \frac{1}{6}(|u\gamma|^{2/3}\kappa^{1/3}) \cdot \langle M \rangle_t^3 \end{aligned}$$

by Hölder's inequality. ////

Proof of Theorem 2.8: The convergence (11) follows from Condition 2.5 and Lemma A.4. The convergence of $\sqrt{N}[\tau^n]Z^n$ in $D[0, T]$ is a consequence of the fact that the convergence of Z^n/ϵ_n is stable. To show (12), we first notice that

$$G_{j,n}^4/G_{j,n}^2 - |G_{j,n}^3/G_{j,n}^2|^2 \geq G_{j,n}^2 \quad \text{a.s.},$$

which follows from Lemma B.1. In light of Lemma A.4 and Condition 2.5, this inequality implies

$$(Hc^2) \cdot \langle M \rangle \geq (Hq^2) \cdot \langle M \rangle \quad \text{a.s.}$$

for any $H \in \mathcal{P}_M^1$. Thus we have

$$\begin{aligned} V_t &= \frac{1}{6}(c^2\gamma^2) \cdot \langle Y \rangle_t q^{-2} \cdot \langle M \rangle_t \\ &\geq \frac{1}{6}(q^2\gamma^2) \cdot \langle Y \rangle_t q^{-2} \cdot \langle M \rangle_t \geq \frac{1}{6}(|\gamma|\sqrt{\kappa}) \cdot \langle M \rangle_t \end{aligned}$$

by the Cauchy-Schwarz inequality. ////

3. EFFECTIVE SCHEMES

Here we give effective discretization schemes. Let $(X, Y, M) \in \mathcal{S}$. For the sake of brevity, we suppose T is finite in this section. Then, by a localization argument, we can suppose without loss of generality that there exists $\delta > 0$ such that $\langle M \rangle$ is strictly increasing a.s. on $[T - \delta, T)$. In fact, we can consider a sequence M^K instead of M defined as, for example,

$$M_t^K = M_{t \wedge \sigma_K} + \hat{W}_t - \hat{W}_{t \wedge \sigma_K}, \quad \sigma_K = \inf\{t > 0; \langle M \rangle_t \geq K\} \wedge (T - 1/K)$$

with $K \rightarrow \infty$, where \hat{W} is a Brownian motion defined on an extension of (Ω, \mathcal{F}, P) . Recall that stable convergence is stable against such a localization procedure. Then, for any positive sequence ϵ_n with $\epsilon_n \rightarrow 0$ and for any $g \in \mathcal{P}_0$ with $g^{-1} \in \mathcal{P}_0$ the sequence of stopping times τ^n defined as

$$(23) \quad \tau_0^n = 0, \quad \tau_{j+1}^n = \inf \left\{ t > \tau_j^n; |M_t - M_{\tau_j^n}| = \epsilon_n g_{\tau_j^n} \right\} \wedge T$$

satisfies Conditions 2.4 and 2.5 with

$$b_s = 0, \quad a_s^2 = q_s^2 = \zeta_s^{-2} = g_s^2,$$

which follows from a famous property on the exit times of one-dimensional continuous local martingales.

Proposition 3.1. *Let $(X, Y, M) \in \mathcal{S}$ and $u \in \mathcal{P}_0$. The lower bound (10) is attained by τ^n defined as (23) with $g = |u|^{1/3} |\gamma^2 \kappa|^{-1/3}$ if $g, g^{-1} \in \mathcal{P}_0$.*

Proposition 3.2. *Let $(X, Y, M) \in \mathcal{S}$. The lower bound (12) is attained by τ^n defined as (23) with $g = |\gamma|^{-1/2} \kappa^{-1/4}$ if $g, g^{-1} \in \mathcal{P}_0$.*

Recall that the lower bound (12) was derived from a combined use of Lemma B.1 and the Cauchy-Schwarz inequality. Karandikar [12] studied a scheme which is defined as (23) with $g = 1$ and X instead of M to show the almost sure convergence of Z^n . In case that ψ appeared in Condition 2.2 is locally bounded and $\gamma \equiv 1$, we can suppose $X = M$ in light of the Girsanov-Maruyama theorem. Then, we can conclude that Karandikar's scheme is superior to the usual time-equidistant one in that it yields increments of the integrand which attain the equality in Lemma B.1. It is in fact optimal if $X = Y$.

Note that Lemma B.1 gives a more precise estimate

$$c^2 \cdot \langle Y \rangle_t = \left(a^2 - \frac{2}{3} b^2 \right) \cdot \langle Y \rangle_t \geq \left(\frac{1}{3} b^2 + q^2 \right) \langle Y \rangle_t.$$

The following proposition, for example, is easily shown by this estimate.

Proposition 3.3. *Let $(X, Y, M) \in \mathcal{S}$, Z^n be defined as (1) and $\beta, \delta \in \mathcal{P}_0$. Denote by $\mathcal{T}(\beta, \delta)$ the set of sequences of schemes τ^n which satisfies Condition 2.5 with $b = \beta$ and $q^2 = \delta$. Then, for all $t \in [0, T)$, Z_t^n / ϵ_n converges to a mixed normal distribution with the asymptotic conditional mean*

$$\frac{1}{3} (\beta \gamma) \cdot Y_t$$

and the asymptotic conditional variance V_t satisfying

$$V_t = \frac{1}{6} (c^2 \gamma^2) \cdot \langle Y \rangle_t \geq \frac{1}{6} \left\{ \left(\frac{1}{3} \beta^2 + \delta \right) \gamma^2 \right\} \cdot \langle Y \rangle_t \text{ a.s..}$$

The equality is attained by $\tau^n \in \mathcal{T}_1(M) \cap \mathcal{T}_2(M)$ defined as

$$\tau_{j+1}^n = \inf \left\{ t > \tau_j^n; M_t - M_{\tau_j^n} \geq \epsilon_n k_{\tau_j^n} \sqrt{\delta_{\tau_j^n}} \text{ or } M_t - M_{\tau_j^n} \leq -\epsilon_n k_{\tau_j^n}^{-1} \sqrt{\delta_{\tau_j^n}} \right\}$$

with $\tau_0^n = 0$, where

$$k_s = \frac{\beta_s \delta_s^{-1/2} + \sqrt{\beta_s^2 \delta_s^{-1} + 4}}{2}.$$

4. CONSERVATIVE DELTA HEDGING

This section treats conservative delta hedging of Mykland [18] as an example of financial applications. This framework includes the usual delta hedging for the Black-Scholes model; even for this classical model, results presented in this section give a new insight and a new practical technique for hedging derivatives. Let S be an asset price process and assume that

$$dS_t = S_t(\mu_t dt + \sigma_t dW_t)$$

is satisfied for predictable processes μ and σ and a standard Brownian motion W . Consider hedging an European contingent claim $f(S_T)$ for a convex function f of polynomial growth. Define a function p as

$$p(S, R, \Sigma) = e^{-R} \int_{\mathbb{R}} f\left(S \exp\{R - \Sigma/2 + \sqrt{\Sigma}z\}\right) \phi(z) dz$$

where ϕ is the standard normal density. Changing variable, it can be shown that

$$(24) \quad \frac{\partial p}{\partial \Sigma} = \frac{1}{2} S^2 \frac{\partial^2 p}{\partial S^2}, \quad \frac{\partial p}{\partial R} = S \frac{\partial p}{\partial S} - p.$$

Put

$$\eta_K = \inf\{t > 0; \langle \log(S) \rangle_t \geq K\}$$

for $K > 0$, $\tilde{V}_t = e^{-rt} V_t$, $\tilde{S}_t = e^{-rt} S_t$ and

$$\Sigma_t^K = K - \langle \log(S) \rangle_{t \wedge \eta_K}, \quad V_t = p(S_t, r(T-t), \Sigma_t^K), \quad \pi_t = \frac{\partial p}{\partial S}(S_t, r(T-t), \Sigma_t^K),$$

where $r > 0$ is a risk-free rate. Then, Itô's formula and (24) yield

$$\tilde{V}_{t \wedge \eta_K} = \int_0^{t \wedge \eta_K} \pi_u d\tilde{S}_u$$

for $t \in [0, T]$, that is, the portfolio strategy (π^0, π) with $\pi_t^0 = e^{-rt}(V_t - \pi_t S_t)$ is self-financing up to $\eta_K \wedge T$. Moreover, the convexity of f and (24) imply that p is increasing in Σ , so that

$$V_T \geq p(S_T, 0, 0) = f(S_T) \text{ on } \{\eta_K \geq T\}.$$

Note that p is the Black-Scholes price with cumulative volatility K and that π is the corresponding delta hedging strategy. The above inequality ensures that the delta hedging super-replicates any European contingent claim with convex payoff on the set $\{\eta_K \geq T\}$. As $K \rightarrow \infty$, $P[\eta_K \geq T] \rightarrow 1$, so that a hedge error due to the incompleteness of market converges to 0. Contracts such as variance swap serve as insurances against the event $\eta_K < T$ for predetermined K which is not so large. See Mykland [19] for an improvement of this conservative delta hedging. The purpose here is, however, not to treat such

a hedge error due to the incompleteness but to treat a hedge error due to the restriction that trades are executed finitely many times in practice. Note that the rebalancing of a portfolio is usually executed a few times per day while observation of S is almost continuous. Hence, the estimation error of $\langle \log(S) \rangle_t$ appeared in Σ_t is negligible compared to the discrete hedging error. Suppose for brevity that $\eta_K \geq T$ a.s.. An approximation of the strategy π is π^n defined as

$$\pi_s^n = \pi_{\tau_j^n} \text{ for } s \in [\tau_j^n, \tau_{j+1}^n)$$

for a discretization scheme τ^n . In this context, $N[\tau^n]_t$ is the number of transactions up to time $t < T$. The discounted replication error is given as

$$Z_t^n = e^{-rt}(V_t - V_t^n) = \int_0^t (\pi_u - \pi_u^n) d\tilde{S}_u.$$

Notice that after a Girsanov-Maruyama transformation, $X = \pi$ is a local martingale and $(X, Y, X) = (\pi, \tilde{S}, \pi) \in \mathcal{S}$. According to our results in the preceding section, the lower bound of the asymptotic variance of $\sqrt{N[\tau^n]_t} Z_t^n$ is attained by the scheme

$$\tau_0^n = 0, \quad \tau_{j+1}^n = \inf \left\{ t > \tau_j^n; |\pi_t - \pi_{\tau_j^n}|^2 = \epsilon_n^2 e^{r\tau_j^n} \Gamma_{\tau_j^n} \right\} \wedge T,$$

where

$$\Gamma_t = \frac{\partial^2 p}{\partial S^2}(S_t, r(T-t), \Sigma_t).$$

Note that Γ is what is called gamma in financial practice. In this case, $\tau^n \in \mathcal{T}_1(X)$ with

$$b_s = 0, \quad a_s^2 = c_s^2 = q_s^2 = e^{rs} \Gamma_s,$$

so that we have

$$Z \left\{ \frac{1}{6} \int_0^t e^{-ru} \Gamma_u d\langle S \rangle_u \right\}^{1/2}$$

as the asymptotic distribution of Z_t^n / ϵ_n , where Z is an independent standard normal variable. This scheme is efficient in that the conditional mean-squared replication error is asymptotically minimized for a conditionally given number of transactions. The number of transactions $N[\tau^n]_t$ is of course random; it is high if the path of Γ is of high level because $|\pi_t - \pi_{\tau_j^n}|^2 / \Gamma_{\tau_j^n} \approx |S_t - S_{\tau_j^n}|^2 \Gamma_{\tau_j^n}$. This property is intuitively expected in practice. Note that ϵ_n controls the expected number of transactions. The asymptotic distribution of $\sqrt{N[\tau^n]_t} Z_t^n$ is

$$\frac{Z}{\sqrt{6}} \int_0^t e^{-ru} \Gamma_u d\langle S \rangle_u.$$

In the equidistant case $\tau_j^n = j/n$, we can apply Theorem 2.6 to $\epsilon_n = 1/\sqrt{n}$, $M = W$, $\gamma = \Gamma\sigma S$, $Y = \tilde{S}$ with

$$b_s = 0, \quad q_s^2 = 1, \quad a_s^2 = c_s^2 = 3, \quad N[\tau^n]_t = [nt]$$

to have that $\sqrt{N[\tau^n]_t} Z_t^n$ converges \mathcal{F} -stably to

$$Z \left\{ \frac{t}{2} \int_0^t e^{-2ru} \Gamma_u^2 \sigma_u^2 S_u^2 d\langle S \rangle_u \right\}^{1/2}.$$

The inequality for the asymptotic conditional variance

$$\frac{1}{6} \left\{ \int_0^t e^{-ru} \Gamma_u d\langle S \rangle_u \right\}^2 \leq \frac{t}{2} \int_0^t e^{-2ru} \Gamma_u^2 \sigma_u^2 S_u^2 d\langle S \rangle_u$$

follows directly from the Cauchy-Schwarz inequality.

Karandikar's scheme is defined as

$$\tau_0^n = 0, \quad \tau_{j+1}^n = \inf \left\{ t > \tau_j^n; |\pi_t - \pi_{\tau_j^n}| = \epsilon_n \right\} \wedge T.$$

After the Girsanov-Maruyama transformation, we apply Theorem 2.6 to $X = M = \pi$ with

$$b_s = 0, \quad a_s^2 = c_s^2 = q_s^2 = 1$$

to have that $\{\sqrt{N[\tau^n]_t} Z_t^n\}$ converges \mathcal{F} -stably to

$$Z \left\{ \int_0^t \Gamma_u^2 d\langle S \rangle_u \right\}^{1/2} \left\{ \frac{1}{6} \int_0^t e^{-2ru} d\langle S \rangle_u \right\}^{1/2}.$$

The inequality for the asymptotic conditional variance

$$\frac{1}{6} \left\{ \int_0^t e^{-ru} \Gamma_u d\langle S \rangle_u \right\}^2 \leq \left\{ \int_0^t \Gamma_u^2 d\langle S \rangle_u \right\} \left\{ \frac{1}{6} \int_0^t e^{-2ru} d\langle S \rangle_u \right\}$$

follows again directly from the Cauchy-Schwarz inequality.

Taking the purpose of hedging into consideration, it might be preferable to use such a scheme τ^n that the asymptotic mean of Z_t^n/ϵ_n is negative. Proposition 3.3 presents an effective scheme for a given asymptotic conditional mean and a given asymptotic conditional number of transactions.

More importantly, we can incorporate linear transaction costs. Suppose that the total cost of the delta hedging with a discretization scheme τ^n up to time $t < T$ is proportional to

$$C_t^n = \sum_{j=0}^{\infty} |\pi_{\tau_{j+1}^n \wedge t} - \pi_{\tau_j^n \wedge t}| S_{\tau_{j+1}^n \wedge t}.$$

Let us study the asymptotic distribution of $C_t^n Z_t^n$. After the Girsanov-Maruyama transformation, π is a local martingale as before. Notice that

$$\epsilon_n C_t^n = \epsilon_n \sum_{j=0}^{\infty} S_{\tau_j^n} |\pi_{\tau_{j+1}^n \wedge t} - \pi_{\tau_j^n \wedge t}| + o_p(1)$$

if $\tau^n \in \mathcal{T}_1(M)$ with $M = \pi$. Apply Theorem 2.7 to $X = M = \pi$, $Y = \tilde{S}$ and $u = S$ to have that $\{C_t^n Z_t^n\}$ converges \mathcal{F} -stably to

$$(S\zeta) \cdot \langle S \rangle_t \left\{ \frac{1}{3} b \cdot \tilde{S}_t + Z \sqrt{\frac{1}{6} c^2 \cdot \langle \tilde{S} \rangle_t} \right\}$$

and that the asymptotic conditional variance of $\{C_t^n Z_t^n\}$ has a lower bound

$$\frac{1}{6} \left| \int_0^t |e^{-ru} S_u \Gamma_u^2|^{2/3} d\langle S \rangle_u \right|^3,$$

which is attained by τ^n defined as

$$\tau_0^n = 0, \quad \tau_{j+1}^n = \inf \left\{ t > \tau_j^n; |\pi_t - \pi_{\tau_j^n}|^3 = \epsilon_n^3 e^{2r\tau_j^n} S_{\tau_j^n} \Gamma_{\tau_j^n}^2 \right\} \wedge T.$$

This scheme is efficient in that the conditional mean-squared replication error is asymptotically minimized for a conditionally given amount of linear transaction costs.

5. EULER-MARUYAMA APPROXIMATION

Here we propose alternative discretization schemes for the Euler-Maruyama approximation as an application. Let us consider the stochastic differential equation

$$\begin{aligned} d\Xi_t &= \mu(\Xi_t, \eta_t)dt + \sigma(\Xi_t, \eta_t)dW_t, \\ d\eta_t &= \theta(\eta_t)dt, \end{aligned}$$

where W is a one-dimensional standard Brownian motion and μ, σ, θ are continuously differentiable functions. Since it is rarely possible to generate a path of Ξ fast and exactly, the Euler-Maruyama scheme is widely used to approximate to Ξ in simulation. For sequences $\tau^n = \{\tau_j^n\}$ with (2), the Euler-Maruyama approximation Ξ^n of Ξ is given as

$$\begin{aligned} d\Xi_t^n &= \mu(\bar{\Xi}_t^n, \bar{\eta}_t^n)dt + \sigma(\bar{\Xi}_t^n, \bar{\eta}_t^n)dW_t, \\ d\eta_t^n &= \theta(\bar{\eta}_t^n)dt, \end{aligned}$$

where $\bar{\Xi}_t^n = \Xi_{\tau_j^n}^n$, $\bar{\eta}_t^n = \eta_{\tau_j^n}^n$ for $j \geq 0$ with $t \in [\tau_j^n, \tau_{j+1}^n)$. Usually $\tau_j^n = j/n$ is taken. The convergence rate of the approximation has been extensively investigated; see e.g., Kloeden and Platen [14] for a well-known strong approximation theorem and Kohatsu-Higa [15], Bally and Talay [1, 2], Konakov and Mammen [16] for weak approximation theorems. Newton [20] treated passage times. Cambanis and Hu [3] studied efficiency of deterministic nonequidistant scheme. Hofmann, Müller-Gronbach and Ritter [8] treated a class of adaptive schemes. Here we exploit a result of Kurtz and Protter [17], Jacod and Protter [10] to deal with the asymptotic distribution of pathwise error. Our aim here is to construct discretization schemes which are more efficient than the usual equidistant sampling scheme.

With the aid of localization, we can suppose $\mu, \sigma, 1/\sigma, \theta$ and their derivatives to be bounded. Suppose that $\tau^n \in \mathcal{T}(W)$. By Theorem 2.6, there exists a conditionally Gaussian martingale Z such that

$$\epsilon_n^{-1}(W - W^n) \cdot W \rightarrow Z$$

\mathcal{F} -stably, where $W_t^n = W_{\tau_j^n}$ for $j \geq 0$ with $t \in [\tau_j^n, \tau_{j+1}^n)$. Put $L_t^n = \epsilon_n^{-1}(\Xi_t^n - \Xi)$. Then, applying Kurtz and Protter [17], we have that L^n converges to a process L which satisfies

$$dL_t = \partial_1 \mu(\Xi_t, \eta_t) L_t dt + \partial_1 \sigma(\Xi_t, \eta_t) [L_t dW_t - \sigma(\Xi_t, \eta_t) dZ_t],$$

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

$$L_t = -e_t \int_0^t e_s^{-1} \sigma(\Xi_s, \eta_s) \partial_1 \sigma(\Xi_s, \eta_s) [dZ_s - \partial_1 \sigma(\Xi_s, \eta_s) d\langle Z, W \rangle_s],$$

where

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

Therefore, in light of Theorem 2.6, the distribution of L_t is mixed normal with conditional mean

$$-\frac{1}{3} e_t \int_0^t e_s^{-1} \sigma(\Xi_s, \eta_s) \partial_1 \sigma(\Xi_s, \eta_s) b_s [dW_t - \partial_1 \sigma(\Xi_t, \eta_t) dt]$$

and conditional variance

$$(25) \quad \frac{1}{6} e_t^2 \int_0^t e_s^{-2} \sigma(\Xi_s, \eta_s)^2 \partial_1 \sigma(\Xi_s, \eta_s)^2 c_s^2 dt.$$

Proposition 5.1. *For any $T > 0$, the space-equidistant scheme τ_{sp}^n defined as (23) with $M = W$, $\epsilon_n = n^{-1/2}$, $g = 1$ is three times efficient than the usual time-equidistant scheme $\tau_{\text{tm}}^n = \{j/n\}$ in the following sense; for any $t \in [0, T)$,*

- $E[N_t^n] \leq nt$ and $N_t^n/n \rightarrow t$ in probability as $n \rightarrow \infty$ for both $N^n = N[\tau_{\text{sp}}^n]$ and $N^n = N[\tau_{\text{tm}}^n]$,
- the asymptotic conditional mean of L_t^n is 0 for the both schemes,
- the asymptotic conditional variance of L_t^n for τ_{sp}^n is one third of that for τ_{tm}^n .

Proof: For the space-equidistant case,

$$b_s \equiv 0, \quad a_s^2 \equiv c_s^2 \equiv 1, \quad P[N_t^n] = nP \left[\sum_{j=0}^{N_t^n-1} |W_{\tau_{j+1}^n} - W_{\tau_j^n}|^2 \right] \leq nt,$$

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. ////

Newton [20] studied this space-equidistant sampling scheme; the superiority of this scheme is more-or-less known. The above simple fact of asymptotic conditional variance, however, has not been recognized so far. The assumption that W is one-dimensional is a serious restriction. Nevertheless, also for a stochastic volatility model

$$\begin{aligned} d\Xi_t &= \hat{\mu}(t, \Xi_t)dt + \hat{\sigma}(t, V_t)[\rho(t, V_t)dW_t^1 + \sqrt{1 - \rho(t, V_t)^2}dW_t^2] \\ dV_t &= \mu(t, V_t)dt + \sigma(t, V_t)dW_t^1 \end{aligned}$$

with a two-dimensional standard Brownian motion (W^1, W^2) , a scheme defined as (23) with $M = W^1$ and $g = 1$ results in a one third conditional asymptotic variance of the Euler-Maruyama approximation error for Ξ . This is because in light of Theorem 2.6, discretization error is determined by only conditional moments of increments of integrand, which is a function of V independent of W^2 in this example.

Next, let us consider to minimize (25) in case that $\partial_1\sigma$ is nondegenerate. Define τ^n as

$$(26) \quad \tau_0^n = 0, \quad \tau_{j+1}^n = \inf\{t > \tau_j^n; |W_t - W_{\tau_j^n}|^2 = \epsilon(\tau_j^n)\},$$

where

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

and

$$\begin{aligned} \log(\hat{\epsilon}_{\tau_j^n}) &= \sum_{i=0}^{j-1} \left\{ \partial_1 \mu(\Xi_{\tau_i^n}^n, \eta_{\tau_i^n}^n) (\tau_{i+1}^n - \tau_i^n) + \partial_1 \sigma(\Xi_{\tau_i^n}^n, \eta_{\tau_i^n}^n) (W_{\tau_{i+1}^n} - W_{\tau_i^n}) \right. \\ &\quad \left. - \frac{1}{2} \partial_1 \sigma(\Xi_{\tau_i^n}^n, \eta_{\tau_i^n}^n)^2 (\tau_{i+1}^n - \tau_i^n) \right\}. \end{aligned}$$

Then, Condition 2.5 is satisfied with

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

Proposition 3.2 implies that this adaptive scheme attains a lower bound for (25) among $\tau^n \in \mathcal{T}(W)$. In this sense, this 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

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

for $\epsilon'_n > 0$ in order to ensure that 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 . It is sufficient that the distribution function F_ϵ of τ is available because

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

and τ, W_τ are independent. In fact, for a random variable U uniformly distributed on $(0, 1)$,

$$(\tau, W_\tau - W_0) \sim (F_\epsilon^{-1}(2U - [2U]), \epsilon(2[2U] - 1)).$$

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 [13], 2.8.11. Using the fact that

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

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

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

According to our numerical study, $G(x) \approx 1$ for $0 \leq x \leq 0.1$. This is not surprising because $G(0+) = 1$, $G'(0+) = G''(0+) = 0$. Note that if $x \geq 0.1$, the speed of convergence of the infinite series is very fast. We can therefore use

$$G(x) \approx \begin{cases} 4 \sum_{n=0}^{\lfloor N/x \rfloor} (\Phi((4n+3)x) - \Phi((4n+1)x)) & x \geq 0.1, \\ 1 & 0 \leq x < 0.1 \end{cases}$$

for, say, $N = 3$ 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 done very fast to generate τ repeatedly even if ϵ changes adaptively as in (26). Note also that

$$G(x) \leq 4(1 - \Phi(x)), \quad G^{-1}(y) \leq \Phi^{-1}(1 - y/4).$$

These inequalities will be useful in numerical calculation of G^{-1} for sufficiently small y (large x). Besides, if $x \geq 3$, $G(x) \approx 4(1 - \Phi(x))$ and $G^{-1}(y) \approx \Phi^{-1}(1 - y/4)$.

APPENDIX A. AUXILIARY RESULTS

Here we give auxiliary results for the proof of our theorems. The following limit theorem, which plays an essential role in this article, is a simplified version of a result of Jacod [9] and Jacod and Shiryaev [11], Theorem IX.7.3, which extends a result of Rootzén [21]. Let $M = \{M_t, \mathcal{F}_t, 0 \leq t < \infty\}$ be a continuous local martingale defined on (Ω, \mathcal{F}, P) and \mathcal{M}^\perp be the set of bounded $\{\mathcal{F}_t\}$ -martingales orthogonal to M .

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

$$\langle Z^n, \hat{M} \rangle_t \rightarrow 0, \quad \langle Z^n, M \rangle_t \rightarrow 0, \quad \langle Z^n \rangle_t \rightarrow V_t$$

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 following lemma is repeatedly used in our proofs.

Lemma A.2. *Consider a sequence of filtrations*

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

and random variables $\{U_j^n\}_{j \in \mathbb{N}}$ with U_j^n being \mathcal{H}_j^n -measurable. Let $N^n(\lambda)$ be a $\{\mathcal{H}_j^n\}$ -stopping time for each $n \in \mathbb{Z}_+$ and λ which is an element of a set Λ . Let $U(\lambda)$ be a random variable for each $\lambda \in \Lambda$. If it holds that there exists $\lambda_0 \in \Lambda$ such that

$$N^n(\lambda) \leq N^n(\lambda_0) \text{ a.s. for all } \lambda \in \Lambda \text{ and } \sum_{j=1}^{N^n(\lambda_0)} P[|U_j^n|^2 | \mathcal{H}_{j-1}^n] \rightarrow 0$$

as $n \rightarrow \infty$, then the following two are equivalent;

(1)

$$\sup_{\lambda \in \Lambda} \left| \sum_{j=1}^{N^n(\lambda)} U_j^n - U(\lambda) \right| \rightarrow 0 \text{ as } n \rightarrow \infty.$$

(2)

$$\sup_{\lambda \in \Lambda} \left| \sum_{j=1}^{N^n(\lambda)} P[U_j^n | \mathcal{H}_{j-1}^n] - U(\lambda) \right| \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Here the convergences are in probability.

Proof: Note that

$$\sup_{\lambda \in \Lambda} \left| \sum_{j=1}^{N^n(\lambda)} U_j^n - \sum_{j=1}^{N^n(\lambda)} E[U_j^n | \mathcal{H}_{j-1}^n] \right| \leq \sup_{k \in \mathbb{N}} \left| \sum_{j=1}^k V_j^n \right|,$$

where

$$V_j^n = (U_j^n - E[U_j^n | \mathcal{H}_{j-1}^n]) 1_{j \leq N^n(\lambda_0)}.$$

By the Lenglart inequality, we have

$$P \left[\sup_{k \in \mathbb{N}} \left| \sum_{j=1}^k V_j^n \right| \geq \epsilon \right] \leq \frac{\eta}{\epsilon^2} + P \left[\sum_{j=1}^{\infty} E[|V_j^n|^2 | \mathcal{H}_{j-1}^n] \geq \eta \right]$$

for any $\epsilon, \eta > 0$. The result then follows from the convergence

$$\sum_{j=1}^{\infty} E[|V_j^n|^2 | \mathcal{H}_{j-1}^n] \leq \sum_{j=1}^{N^n(\lambda_0)} E[|U_j^n|^2 | \mathcal{H}_{j-1}^n] \rightarrow 0$$

in probability. ////

The following lemma is well-known, so its proof is omitted.

Lemma A.3. *For all $H \in \mathcal{P}_M^k$, $t \in [0, T]$ and $\delta_1, \delta_2 > 0$, there exists a bounded adapted left-continuous process \hat{H} such that*

$$P[|H - \hat{H}|^k \cdot \langle M \rangle_t > \delta_1] < \delta_2.$$

The following lemma is taken from Fukasawa [5].

Lemma A.4. *Let M be a continuous local martingale with $E[\langle M \rangle_T^6] < \infty$ and suppose that $\tau^n \in \mathcal{T}(M)$. Then, for all $t \in [0, T]$,*

$$(27) \quad \sup_{j \geq 0, 0 \leq s \leq t} |M_{\tau_{j+1}^n \wedge s} - M_{\tau_j^n \wedge s}|^2 = o_p(\epsilon_n)$$

as well as

$$(28) \quad \sup_{0 \leq j \leq N[\tau^n]_t} |\langle M \rangle_{\tau_{j+1}^n} - \langle M \rangle_{\tau_j^n}| = o_p(\epsilon_n), \quad \sup_{0 \leq j \leq N[\tau^n]_t} |M_{\tau_{j+1}^n} - M_{\tau_j^n}|^2 = o_p(\epsilon_n),$$

where $N[\tau^n]_t$ is defined as (3). In particular,

$$(29) \quad N[\tau^n]_t \rightarrow \infty \text{ a.s.}$$

as $n \rightarrow \infty$. Moreover, for all locally bounded adapted cag process f , it holds

$$(30) \quad \sum_{j=0}^{N[\tau^n]_t} f_{\tau_j^n} G_{j,n}^2 \rightarrow \int_0^t f_s d\langle M \rangle_s$$

in probability, uniformly in t on compact sets of $[0, T]$.

Proof: Note that

$$\sum_{j=0}^{N[\tau^n]_t} G_{j,n}^2 = O_p(1)$$

since $E[\langle M \rangle_T] < \infty$. By the assumptions, it follows

$$\sum_{j=0}^{N[\tau^n]_t} E[(M_{\tau_{j+1}^n} - M_{\tau_j^n})^{2k} | \mathcal{F}_{\tau_j^n}] = o_p(\epsilon_n^k)$$

for $k = 3, 6$ and with the aid of Lemma A.2, we have

$$\sum_{j=0}^{N[\tau^n]_t} (M_{\tau_{j+1}^n} - M_{\tau_j^n})^6 = o_p(\epsilon_n^3).$$

On the other hand,

$$\sup_{0 \leq j \leq N[\tau^n]_t} |M_{\tau_{j+1}^n} - M_{\tau_j^n}|^2 \leq \left\{ \sum_{j=0}^{N[\tau^n]_t} (M_{\tau_{j+1}^n} - M_{\tau_j^n})^6 \right\}^{1/3},$$

so that the second of (28) follows.

To show (27), we use Doob's maximal inequality to have

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

for $k = 3, 6$. Using Lemma A.2 again, we obtain

$$\sum_{j=0}^{N[\tau^n]_t} \sup_{0 \leq s < \infty} |M_{\tau_{j+1}^n \wedge s} - M_{\tau_j^n \wedge s}|^6 = o_p(\epsilon_n^3),$$

which implies (27) since

$$\sup_{j \geq 0, 0 \leq s \leq t} |M_{\tau_{j+1}^n \wedge s} - M_{\tau_j^n \wedge s}|^2 \leq \left\{ \sum_{j=0}^{N[\tau^n]_t} \sup_{0 \leq s < \infty} |M_{\tau_{j+1}^n \wedge s} - M_{\tau_j^n \wedge s}|^6 \right\}^{1/3}.$$

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

$$\sum_{j=0}^{N[\tau^n]_t} E[|\langle M \rangle_{\tau_{j+1}^n} - \langle M \rangle_{\tau_j^n}|^k | \mathcal{F}_{\tau_j^n}] = o_p(\epsilon_n^k)$$

for $k = 3, 6$, which implies the first of (28) in the same manner. Note that (29) follows from

$$N[\tau^n]_t + 1 \geq \frac{\langle M \rangle_t}{\sup_{0 \leq j \leq N[\tau^n]_t} |\langle M \rangle_{\tau_{j+1}^n} - \langle M \rangle_{\tau_j^n}|}.$$

To see (30), again in light of Lemma A.2, it suffices to observe that

$$\sum_{j=0}^{N[\tau^n]_t} f_{\tau_j^n} (\langle M \rangle_{\tau_{j+1}^n} - \langle M \rangle_{\tau_j^n}) \rightarrow \int_0^t f_s d\langle M \rangle_s$$

and

$$\sum_{j=0}^{N[\tau^n]_t} f_{\tau_j^n}^2 G_{j,n}^4 = \epsilon_n^2 \sum_{j=0}^{N[\tau^n]_t} f_{\tau_j^n}^2 a_{\tau_j^n}^2 G_{j,n}^2 + o_p(\epsilon_n^2) \rightarrow 0.$$

////

APPENDIX B. KURTOSIS-SKEWNESS INEQUALITY

Lemma B.1. *Let X be a random variable with $E[X] = 0$ and $E[X^4] < \infty$. Then it holds*

$$(31) \quad \frac{E[X^4]}{E[X^2]^2} - \frac{E[X^3]^2}{E[X^2]^3} \geq 1.$$

The equality is attained if and only if X is a Bernoulli random variable.

Proof: This is called Pearson's inequality and shown easily as follows.

$$E[X^3]^2 = E[X(X^2 - E[X^2])]^2 \leq E[X^2]E[(X^2 - E[X^2])^2] = E[X^2](E[X^4] - E[X^2]^2).$$

If the equality holds, then X and $X^2 - E[X^2]$ must be linearly dependent, so that X must be Bernoulli. ////

The following lemma gives a similar inequality to the above. The proof is however rather different and the result itself is seemingly new.

Lemma B.2. *Let X be a random variable with $E[X] = 0$ and $E[X^4] < \infty$. Then it holds*

$$(32) \quad \frac{E[X^4]}{E[X^2]^2} - \frac{3 E[X^3]^2}{4 E[X^2]^3} \geq \frac{E[X^2]}{E[|X|]^2}.$$

The equality is attained if and only if X is a Bernoulli random variable.

Proof: We divide the proof into 4 steps.

Step a) It is straightforward to see that the equality holds if X is a Bernoulli random variable with $E[X] = 0$.

Step b) Let us show if $E[X] = 0$ and the support of X is a finite set, then the distribution P^X of X is a finite mixture of Bernoulli distributions with mean 0. First, consider the case $n = 3$. Suppose without loss of generality that

$$P[X = a] = p, P[X = b] = q, P[X = c] = r, p + q + r = 1$$

with $a > b \geq 0 > c$. Put

$$P_1(a) = \frac{-c}{a-c}, P_1(c) = \frac{a}{a-c}, P_2(b) = \frac{-c}{b-c}, P_2(c) = \frac{b}{b-c}.$$

Then P_1 and P_2 define Bernoulli distributions with mean 0 and supports $\{a, c\}$ and $\{b, c\}$ respectively. Putting $\lambda = (c - a)p/c = P^X(a)/P_1(a)$, we have

$$\lambda P_1(a) = p, (1 - \lambda)P_2(b) = q, \lambda P_1(c) + (1 - \lambda)P_2(c) = r,$$

which means

$$P^X = \lambda P_1 + (1 - \lambda)P_2.$$

Now, let us treat the general case by induction. Suppose that the claim holds for the case of n and consider the case of $n + 1$. Without loss of generality, we suppose

$$P[X = a_j] = p_j, j = 0, 1, \dots, n, p_0 + p_1 + \dots + p_n = 1$$

with

$$a_0 > a_1 > \dots > a_k \geq 0 > a_{k+1} > \dots > a_n$$

for an integer $k, 2 \leq k \leq n - 1$. Put

$$\tilde{a}_1 = \frac{a_0 p_0 + a_1 p_1}{p_0 + p_1}, \quad \tilde{p}_1 = p_0 + p_1, \quad \tilde{a}_j = a_j, \quad \tilde{p}_j = p_j, \quad 2 \leq j \leq n$$

and

$$\tilde{P}(\tilde{a}_j) = \tilde{p}_j, \quad j = 1, \dots, n.$$

Notice that \tilde{P} defines a distribution with mean 0 and support $\{\tilde{a}_1, a_2, \dots, a_n\}$. By the assumption of induction, there exists $\tilde{\lambda}_{ij} \geq 0, 1 \leq i \leq k, k < j \leq n$ such that

$$\sum_{i,j} \tilde{\lambda}_{ij} = 1, \quad \tilde{P} = \sum_{i,j} \tilde{\lambda}_{ij} \tilde{P}_{ij}, \quad \tilde{P}_{ij}(\tilde{a}_i) = \frac{-\tilde{a}_j}{\tilde{a}_i - \tilde{a}_j}, \quad \tilde{P}_{ij}(\tilde{a}_j) = \frac{\tilde{a}_i}{\tilde{a}_i - \tilde{a}_j}.$$

Here \tilde{P}_{ij} defines a Bernoulli distribution with mean 0 and support $\{\tilde{a}_i, \tilde{a}_j\}$. Now consider a distribution Q_j defined as

$$Q_j(a_0) = \frac{p_0}{p_0 + p_1} \tilde{P}_{1j}(\tilde{a}_1), \quad Q_j(a_1) = \frac{p_1}{p_0 + p_1} \tilde{P}_{1j}(\tilde{a}_1), \quad Q_j(a_j) = \tilde{P}_{1j}(\tilde{a}_j)$$

for $k < j \leq n$. Notice that Q_j is a distribution with mean 0 and support $\{a_0, a_1, a_j\}$. As seen above for the case $n = 3$, putting $\mu_j = Q_j(a_0)/P_0(a_0)$, we have

$$Q_j = \mu_j P_0 + (1 - \mu_j) P_{1j},$$

where we define

$$P_{ij}(a_i) = \frac{-a_j}{a_i - a_j}, \quad P_{ij}(a_j) = \frac{a_i}{a_i - a_j}, \quad 0 \leq i \leq k, \quad k < j \leq n.$$

Putting

$$\lambda_{0j} = \mu_j \tilde{\lambda}_{1j}, \quad \lambda_{1j} = (1 - \mu_j) \tilde{\lambda}_{1j}, \quad \lambda_{ij} = \tilde{\lambda}_{ij}, \quad 2 \leq i \leq k, \quad k < j \leq n,$$

we have

$$\sum_{i,j} \lambda_{ij} = 1, \quad P^X = \sum_{i,j} \lambda_{ij} P_{ij},$$

which completes the induction.

Step c) Let us show that the function $f(u, v, w, y)$ defined as

$$(33) \quad f(u, v, w, y) = u - \frac{3}{4} v^2 / w - w^3 / y^2$$

is a concave function. Note that the inequality (32) follows from Steps a, b and c since every distribution can be approximated arbitrarily close by a distribution supported by a finite set. By a straightforward calculation, the Hessian matrix of f is given by

$$(34) \quad H = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -\frac{3}{2w} & \frac{3v}{2w^2} & 0 \\ 0 & \frac{3v}{2w^2} & -\frac{3v^2}{2w^3} - \frac{6w}{y^2} & \frac{6w^2}{y^3} \\ 0 & 0 & \frac{6w^2}{y^3} & -\frac{6w^3}{y^4} \end{pmatrix}.$$

Again by a straightforward calculation, the determinant of $H - xI$ is of form $x^2(x + \alpha)(x + \beta)$ with $\alpha > 0$, $\beta > 0$, which means that H is negative semi-definite.

Step d) It remains to show that the equality holds only if X is a Bernoulli random variable. Suppose that there exists a random variable X with $E[X] = 0$ such that the equality holds in (32) which is not Bernoulli. Recall that the equality holds in (31) only if X is Bernoulli. It implies that the vector of the first four moments of X does not coincide with that of a Bernoulli random variable. Note that there exists a random variable \hat{X} of which the support is a finite set, such that

$$E[\hat{X}] = E[X], E[|\hat{X}|] = E[|X|], E[\hat{X}^2] = E[X^2], E[\hat{X}^3] = E[X^3], E[\hat{X}^4] = E[X^4].$$

This can be proved by the Hahn-Banach theorem. Hence, we assume the support of X is a finite set without loss of generality. Then, by Steps b and c, there exist Bernoulli distributions P_1 and P_2 and $\lambda \in (0, 1)$ such that $P_1 \neq P_2$ and

$$(35) \quad f(\lambda m_1 + (1 - \lambda)m_2) = 0 = \lambda f(m_1) + (1 - \lambda)f(m_2),$$

where f is defined as (33) and

$$m_i = \left(\int a^4 P_i(da), \int a^3 P_i(da), \int a^2 P_i(da), \int |a| P_i(da) \right)', \quad i = 1, 2.$$

Here \prime means the transpose of matrix. By (34), the eigenvectors of the Hessian matrix H associated to the eigenvalue 0 are

$$h_1 = (1, 0, 0, 0)', \quad h_2 = (0, v, w, y)'$$

Therefore, (35) implies that there exists a constant c such that $\bar{m}_2 = c\bar{m}_1$, where

$$\bar{m}_i = \left(\int a^3 P_i(da), \int a^2 P_i(da), \int |a| P_i(da) \right)', \quad i = 1, 2.$$

It suffices then to show that $\bar{m}_2 = c\bar{m}_1$ implies $c = 1$ and that \bar{m}_1 uniquely determines a Bernoulli distribution. Set

$$P_2(a) = p, \quad P_2(-b) = q, \quad p + q = 1, \quad ap = bq, \quad \bar{m}_1 = (v, w, y)'$$

and

$$a^3 p - b^3 q = cv, \quad a^2 p + b^2 q = cw, \quad ap + bq = cy.$$

Then we obtain that

$$a = \frac{2wq}{y}, \quad b = \frac{2wp}{y},$$

so that

$$\frac{2v}{y} = a^2 - b^2 = \frac{4w^2}{y^2}(1 - 2p).$$

Therefore, a, b, p, q are uniquely determined independently of c , which completes the proof. ////

Acknowledgement: This work is partially supported by JSPS, Grant-in-Aid for Young Scientists (B) and JST, CREST.

REFERENCES

- [1] Bally, V. and Talay, D. (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] Bally, V. and Talay, D. (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] Cambanis, S. and Hu, Y. (1996): Exact convergence rate of the Euler-Maruyama scheme, with application to sampling design, *Stoch. Stoch. rep.* 59, 211-240.
- [4] Fukasawa, M. (2007): Central limit theorem for the realized volatility based on tick time sampling, to appear in *Finance Stoch.*
- [5] Fukasawa, M. (2008): Realized volatility with stochastic sampling, Discussion Paper Series 2008-06, Center for the Study of Finance and Insurance, Osaka Univ.
- [6] Geiss, S. and Toivola, A. (2009): Weak convergence of error processes in discretizations of stochastic integrals and Besov spaces, to appear in *Bernoulli*.
- [7] Hayashi, T. and Mykland, P.A. (2005): Evaluating hedging errors: an asymptotic approach, *Math. Finance* 15, no. 2, 309-343.
- [8] Hofmann, N., Müller-Gronbach, T. and Ritter, K. (2001): The optimal discretization of stochastic differential equations, *J. Complexity* 17, no. 1, 117-153.
- [9] Jacod, J. (1994): Limit of random measures associated with the increments of a Brownian semimartingale, unpublished.
- [10] Jacod, J. and Protter, P. (1998): Asymptotic error distributions for the Euler method for stochastic differential equations, *Ann. Probab.* 26, no. 1, 267-307.
- [11] Jacod, J. and Shiryaev, A.N. (2002): *Limit theorems for stochastic processes*, 2nd ed., Springer-Verlag.
- [12] Karandikar, R.L. (1995): On pathwise stochastic integration, *Stochastic Process. Appl.* 57, no. 1, 11-18.
- [13] Karatzas, I. and Shreve, S.E. (1991): *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York.
- [14] Kloeden, P.E. and Platen, E. (1992): *Numerical solution of stochastic differential equations*, Applications of Mathematics (New York), 23, Springer-Verlag, Berlin.
- [15] Kohatsu-Higa, A. (2002): Weak approximations. A Malliavin calculus approach, *Math. Comp.* 70, no. 233, 135-172.
- [16] Konakov, V. and Mammen, E. (2002): Edgeworth type expansions for Euler schemes for stochastic differential equations, *Monte Carlo Methods Appl.* 8, no. 3, 271-285.
- [17] Kurtz, T. and Protter, P. (1991): Wong-Zakai corrections, random evolutions, and simulation schemes for SDEs. *Stochastic analysis*, 331-346, Academic Press, Boston, MA.
- [18] Mykland, P.A. (2000): Conservative delta hedging, *Ann. Appl. Probab.* 10, no. 2, 664-683.
- [19] Mykland, P.A. (2003): The interpolation of options, *Finance Stoch.* 7 (2003), no. 4, 417-432.
- [20] Newton, N.J. (1990): An efficient approximation for stochastic differential equations on the partition of symmetrical first passage times, *Stoch. Stoch. Rep.* 29, no. 2, 227-258.
- [21] Rootzén, H. (1980): Limit distributions for the error in approximations of stochastic integrals, *Ann. Probab.* 8, no. 2, 241-251.