The Annals of Applied Probability
2011, Vol. 21, No. 4, 1436–1465
DOI: 10.1214/10-AAP730

DISCRETIZATION ERROR OF STOCHASTIC INTEGRALS[1]

BY MASAAKI FUKASAWA

Osaka University

Limit distributions for the error in approximations of stochastic integrals by Riemann sums with stochastic partitions are studied. The integrands and integrators are supposed to be one-dimensional continuous semimartingales. Lower bounds for asymptotic conditional variance of the error are given and effective discretization schemes which attain the bounds are explicitly constructed. Two examples of their applications are given; efficient delta hedging strategies under fixed or linear 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 $Z^n = \{Z^n_t\}_{t\in[0,T)}$ defined as

$$Z^n_t = \int_0^t X_s \, \mathrm{d}Y_s - \sum_{j=0}^{\infty} X_{\tau^n_j}(Y_{\tau^n_{j+1}\wedge t} - Y_{\tau^n_j \wedge t}) \tag{1}$$

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 $\tau^n = \{\tau^n_j\}$ with

$$0 = \tau^n_0 < \tau^n_1 < \cdots < \tau^n_j < \cdots, \qquad \lim_{j\to\infty} \tau^n_j = T \qquad \text{a.s.}, \tag{2}$$

where $T \in (0,\infty]$ is fixed and the intervals $\tau^n_{j+1} - \tau^n_j$ are supposed to converge to 0 as $n \to \infty$ in a sense specified later. The stochastic integral is usually defined as a limit of Riemann sums and naturally approximated by them in practices, so the asymptotic behavior of Z^n is of interest. This problem was studied by Rootzén [23] in the case that Y is a Brownian motion and the asymptotic distribution of Z^n was specified in the case that $\tau^n_j = j/n$ and $X_s = f(Y_s, s)$ with a smooth function f. Jacod [10] treated a related problem on the condition that each interval $\tau^n_{j+1} - \tau^n_j$ is $\mathcal{F}_{\tau^n_j}$-measurable. Jacod and Protter [11] considered the case $X = Y$ and $\tau^n_j = j/n$ to derive the asymptotic error distribution of the Euler–Maruyama approximation for stochastic differential equations. Hayashi and Mykland [8] discussed this problem again for the case $\tau^n_j = j/n$ in a financial context of discrete-time hedging error. Geiss and Toivola [7] treated an irregular deterministic discretization scheme. The measurability condition that $\tau^n_{j+1} - \tau^n_j$ is $\mathcal{F}_{\tau^n_j}$-measurable

Received December 2009; revised July 2010.

[1]Supported by KAKENHI 21740074 (MEXT), Mirai Labo (Osaka Univ.), Cooperative Research Program (ISM) and CREST (JST).

MSC2010 subject classifications. 60F05, 60F17, 60H05.

Key words and phrases. Stable convergence, discrete hedging, Euler–Maruyama scheme.

for all j has played an indispensable role in those preceding studies. On the other hand, Karandikar [13] constructed a stochastic discretization scheme τ^n such that Z_t^n converges to 0 almost surely. Since the almost sure convergence does not hold in general for deterministic schemes, Karandikar's scheme is more effective in a sense. The scheme is constructed by using passage times of X and so does not satisfy the measurability condition. Recently, Fukasawa [5] proved a limit theorem for a class of discretization schemes including Karandikar's one in the case $X = Y$ and Fukasawa [6] extended it to general discretization schemes. This article extends those limit theorems to include general integrands X and presents lower bounds for the asymptotic conditional variance of Z^n. Effective discretization schemes which attain the bounds are explicitly constructed. Karandikar's scheme is shown to be superior to the deterministic scheme $\tau_j^n = j/n$ also in terms of mean squared error. An application to delta hedging under fixed or linear transaction costs is given which can be directly used in practice. Another application is to construct an alternative discretization scheme for the Euler–Maruyama approximation which results in a one third asymptotic mean squared error. It remains a matter for further research to extend the results to discontinuous semimartingales. The main results are given in Section 2. Effective discretization schemes are constructed in Section 3. The applications to hedging and the Euler–Maruyama approximation are presented in Sections 4 and 5, respectively.

2. Central limit theorem.

2.1. *Notation and conditions.* Here we give a rigorous formulation and 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 $\{\mathcal{F}_t\}_{t\geq 0}$ 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. For a positive sequence δ_n and sequences of random variables Ξ_n or $\Xi_{j,n}$, we write $\Xi_n = O_p(\delta_n)$ if $\delta_n^{-1}\Xi_n$ is tight. We say $\Xi_{j,n} = O_p(\delta_n)$ uniformly in j if $\sup_j \delta_n^{-1}|\Xi_{j,n}|$ is tight. We write $\Xi_n = o_p(\delta_n)$ if $\delta_n^{-1}\Xi_n$ converges to 0 in probability as $n \to \infty$. We say $\Xi_{j,n} = o_p(\delta_n)$ uniformly in j if $\sup_j |\delta_n^{-1}\Xi_{j,n}|$ converges to 0 in probability as $n \to \infty$.

Let us recall the definition of stable convergence.

DEFINITION 2.1. Let $\mathbb{E}$ be a complete separable metric space, F^n be a sequence of $\mathbb{E}$-valued random variables defined on $(\Omega, \mathcal{F}, P)$ and F be an $\mathbb{E}$-valued random variable defined on an extension of $(\Omega, \mathcal{F}, P)$. For a sub σ-field $\mathcal{G} \subset \mathcal{F}$, we say F^n converges $\mathcal{G}$-stably to F if for all $\mathcal{G}$-measurable random variable F_0, the joint distribution (F^n, F_0) converges to (F, F_0) in law.

Our main results stated in the next subsection are stable convergences of Z^n defined by (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

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

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

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

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 usual localization procedure. In order to describe conditions on τ^n, we put

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

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, we put

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

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 $\varepsilon_n \to 0$ and $a, b \in \mathcal{P}_0$ such that

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

and

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

uniformly in $j = 0, 1, \ldots, N[\tau^n]_t$ for all $t \in [0, T)$. Here $0/0$ is understood as 0.

Condition 2.3 is slightly stronger than [6], Condition 3.5; nevertheless, all examples given by Fukasawa [6] also satisfy this condition. Here, ε_n serves as a scale of increments of M. Note that $G^4_{j,n}/G^2_{j,n} = O_p(\varepsilon_n^2)$ implies $G^3_{j,n}/G^2_{j,n} = O_p(\varepsilon_n)$ by Lemma B.2. In usual cases, we have $G^{2k}_{j,n} = O_p(\varepsilon_n^{2k})$, which in fact holds, for example, if $\mathrm{d}\langle M\rangle_t/\mathrm{d}t$ exists and is bounded and if $\tau^n_{j+1} - \tau^n_j$ is of $O(\varepsilon_n^2)$ 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 then utilize many results on Brownian stopping times. See [6] 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 the increments. It should be noted that $\sup_j |\tau^n_{j+1} \wedge t - \tau^n_j \wedge t| \to 0$ does not follow from Condition 2.3 nor needed for our main results.

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

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

$$\zeta^{-1} \in \mathcal{P}_0, \qquad \varepsilon_n G^1_{j,n}/G^2_{j,n} = \zeta_{\tau^n_j} + o_p(1)$$

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

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

$$q^{-1} \in \mathcal{P}_0, \qquad G^2_{j,n} = q^2_{\tau^n_j}\varepsilon_n^2 + o_p(\varepsilon_n^2)$$

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

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

$$[M]^t_{j,n} = \langle M\rangle_{\tau^n_{j+1}\wedge t} - \langle M\rangle_{\tau^n_j \wedge t}. \tag{4}$$

2.2. *Main results.* Here we state main results on the limit 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 by* (1). *Assume one of the following two conditions hold*:

(i) *M is the local martingale part of X, that is,*

$$\gamma \equiv 1. \tag{5}$$

(ii) *For all* $t \in [0, T)$,

$$E\left[\sum_{j=0}^{\infty} |[M]^t_{j,n}|^k\right] = O\left(\varepsilon_n^{2(k-1)}\right) \tag{6}$$

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

Then, Z^n/ε_n *converges* $\mathcal{F}$-*stably to*

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

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

$$c^2 = a^2 - \tfrac{2}{3}b^2, \qquad Y' = W_{\langle Y \rangle} \tag{8}$$

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 by* (1). *Let* $u \in \mathcal{P}_0$ *and put*

$$U^n_t = \sum_{j=0}^{\infty} |u_{\tau^n_j}||M_{\tau^n_{j+1} \wedge t} - M_{\tau^n_j \wedge t}|.$$

Then, *it holds that*

$$\varepsilon_n U^n_t \to U_t := (|u|\zeta) \cdot \langle M \rangle_t \tag{9}$$

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

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

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

$$\varepsilon_n^2 N[\tau^n]_t \to N_t := q^{-2} \cdot \langle M \rangle_t \tag{11}$$

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* $\sqrt{N}Z$ *as a* $D[0, T)$-*valued sequence*, *where* Z *is defined by* (7). *The asymptotic conditional variance* V_t *of* $\sqrt{N[\tau^n]_t}Z^n_t$ *with* $t \in [0, T)$ *satisfies*

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

Note that the right-hand sides of (10) and (12) do not depend on 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) and (12). Its practical meaning is discussed in Sections 4 and 5.

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 $\tau^n_{j+1} - \tau^n_j$ is $\mathcal{F}_{\tau^n_j}$ measurable. It is, therefore, not difficult to recover the preceding results from Theorem 2.6. An irregular scheme treated in [7] 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 $E[\langle M\rangle_T^6] < \infty$. Let $\tau^n \in \mathcal{T}(M)$, $H \in \mathcal{P}_0$ and γ be a locally bounded predictable process. Put $\bar{M} = \gamma \cdot M$ and define H^n, $\bar{M}^n$ as*

$$H^n_s = H_{\tau^n_j}, \qquad \bar{M}^n_s = \bar{M}_{\tau^n_j} \qquad \text{for } j \geq 0 \text{ with } s \in [\tau^n_j, \tau^n_{j+1}). \tag{13}$$

If (5) or (6) holds, then it holds that

$$\begin{aligned} &\varepsilon_n^{-1}\big((\bar{M} - \bar{M}^n)H^n\big) \cdot \langle M\rangle_t \to \tfrac{1}{3}(bH\gamma) \cdot \langle M\rangle_t, \\ &\varepsilon_n^{-2}\big((\bar{M} - \bar{M}^n)^2 H^n\big) \cdot \langle M\rangle_t \to \tfrac{1}{6}(a^2 H\gamma^2) \cdot \langle M\rangle_t \end{aligned} \tag{14}$$

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

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

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 $l \in \mathbb{Z}_+$,

$$\begin{aligned} &\varepsilon_n^{-l}\big((\bar{M} - \bar{M}^n)^l H^n\big) \cdot \langle M\rangle_t \\ &\quad = \varepsilon_n^{-l} \sum_{j=0}^{\infty} H_{\tau^n_j} \left\{ \alpha_l (M_{\tau^n_{j+1}\wedge t} - M_{\tau^n_j \wedge t})^{l+2} + \beta_l \int_{\tau^n_j \wedge t}^{\tau^n_{j+1}\wedge t} (M_s - M_{\tau^n_j})^{l+1} \, \mathrm{d}M_s \right\} \end{aligned}$$

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

$$\begin{aligned} &\varepsilon_n^{-2l} \sum_{j=0}^{N[\tau^n]_t} H^2_{\tau^n_j} E\left[\int_{\tau^n_j}^{\tau^n_{j+1}} (M_s - M_{\tau^n_j})^{2l+2} \, \mathrm{d}\langle M\rangle_s \big| \mathcal{F}_{\tau^n_j} \right] \\ &\quad = \varepsilon_n^{-2l} \alpha_{2l+2} \sum_{j=0}^{N[\tau^n]_t} H^2_{\tau^n_j} G^{2l+4}_{j,n} \to 0 \end{aligned}$$

in probability for $l \in \{1, 2, 4\}$ by Condition 2.3, we have

$$\varepsilon_n^{-l} \sum_{j=0}^{\infty} H_{\tau_j^n} \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} (M_s - M_{\tau_j^n})^{l+1} \, \mathrm{d}M_s \to 0$$

in probability, as well as

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

in probability by Lemmas A.2 and A.4. The result then follows from the last assertion of Lemma A.4.

Next, let us suppose (6) and $\gamma \not\equiv 1$. Note that for all $\delta > 0$, there exists a bounded left-continuous process γ^δ such that

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

for any $k \in \{4, 6, 8, 10, 12\}$ by Lemma A.3. Notice that for $\xi \in \mathbb{N}$, $\eta \in \mathbb{Z}_+$ and $p, q, r > 1$ with $1/p + 1/q + 1/r = 1$,

$$\begin{aligned}
&E\left[\sum_{j=0}^{\infty} \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} \left| \int_{\tau_j^n \wedge t}^{s} (\gamma_u - \gamma_u^\delta) \, \mathrm{d}M_u \right|^\xi \left| \int_{\tau_j^n \wedge t}^{s} \gamma_n^\delta \, \mathrm{d}M_u \right|^\eta \mathrm{d}\langle M \rangle_s \right] \\
&\quad \le E\left[\sum_{j=0}^{\infty} \sup_{s \in [\tau_j^n \wedge t, \tau_{j+1}^n \wedge t]} \left| \int_{\tau_j^n \wedge t}^{s} (\gamma_u - \gamma_u^\delta) \, \mathrm{d}M_u \right|^\xi \right. \\
&\qquad \left. \times \sup_{s \in [\tau_j^n \wedge t, \tau_{j+1}^n \wedge t]} \left| \int_{\tau_j^n \wedge t}^{s} \gamma_u^\delta \, \mathrm{d}M_u \right|^\eta [M]_{j,n}^t \right] \\
&\quad \le C(\xi, \eta) \left\{ E\left[\sum_{j=0}^{\infty} \left| \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} (\gamma_u - \gamma_u^\delta)^2 \, \mathrm{d}\langle M \rangle_u \right|^{p\xi/2} \right] \right\}^{1/p} \\
&\qquad \times \left\{ E\left[\sum_{j=0}^{\infty} |[M]_{j,n}^t|^{q\eta/2} \right] \right\}^{1/q} \left\{ E\left[\sum_{j=0}^{\infty} |[M]_{j,n}^t|^{r} \right] \right\}^{1/r}
\end{aligned}$$

by the Hölder and the Burkholder–Davis–Gundy inequalities, where $C(\xi, \eta)$ is a constant. Furthermore,

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

For each $(\xi,\eta)\in\{(1,0),(1,1),(2,0),(1,3),(2,2),(3,1),(4,0)\}$, one can find suitable (p,q,r) such that it follows from the assumption (6) that

$$\left\{E\left[\sum_{j=0}^{\infty}|[M]^t_{j,n}|^{p\xi-1}\right]\right\}^{1/(2p)}\left\{E\left[\sum_{j=0}^{\infty}|[M]^t_{j,n}|^{q\eta/2}\right]\right\}^{1/q}\left\{E\left[\sum_{j=0}^{\infty}|[M]^t_{j,n}|^{r}\right]\right\}^{1/r}$$

is of $O(\varepsilon_n^{\xi+\eta})$. For example, take $(p,q,r)=(5/4,\infty,5)$ for $(\xi,\eta)=(4,0)$. Since δ can be arbitrarily small, this estimate ensures that one can replace $\bar{M}$ and γ by $\gamma^\delta\cdot M$ and γ^δ, respectively, in (14) and (15). Put

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

By the same argument, one can estimate

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

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

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

The rest of the proof is to repeat the argument for the case $\gamma\equiv 1$ by replacing H with $H\gamma^\delta$. $\square$

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

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

We shall prove that

$$\varepsilon_n^{-1}(A-A^n)\cdot Y_t\to 0,\qquad \varepsilon_n^{-1}\big((\bar{M}-\bar{M}^n)\varphi\big)\cdot\langle M\rangle_t\to\tfrac{1}{3}(b\varphi)\cdot\langle M\rangle_t$$

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

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

converges $\mathcal{F}$-stably to

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

Step (a). Let us show

$$\varepsilon_n^{-1}(A-A^n)\cdot Y_v\to 0 \tag{17}$$

uniformly in $v\in[0,t]$ in probability. We shall first see that

$$\varepsilon_n^{-1}\big((A-A^n)\varphi\big)\cdot\langle M\rangle_v\to 0 \tag{18}$$

uniformly in $v \in [0,t]$ in probability. Fix $\delta_1, \delta_2 > 0$ arbitrarily and take a bounded left-continuous 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 $v \in [0,t]$,

$$\begin{aligned}
&\varepsilon_n^{-1}\bigl((A - A^n)\varphi\bigr) \cdot \langle M\rangle_v \\
&\quad = \varepsilon_n^{-1} \sum_{j=0}^{\infty} \int_{\tau_j^n \wedge v}^{\tau_{j+1}^n \wedge v} \int_{\tau_j^n \wedge v}^{s} \psi_u \, \mathrm{d}\langle M\rangle_u \, \varphi_s \, \mathrm{d}\langle M\rangle_s \\
&\quad = \varepsilon_n^{-1} \sum_{j=0}^{\infty} \int_{\tau_j^n \wedge v}^{\tau_{j+1}^n \wedge v} \int_{\tau_j^n \wedge v}^{s} \psi_u^\delta \, \mathrm{d}\langle M\rangle_u \, \varphi_s \, \mathrm{d}\langle M\rangle_s \\
&\qquad + \varepsilon_n^{-1} \sum_{j=0}^{\infty} \int_{\tau_j^n \wedge v}^{\tau_{j+1}^n \wedge v} \int_{\tau_j^n \wedge v}^{s} (\psi_u - \psi_u^\delta) \, \mathrm{d}\langle M\rangle_u \, \varphi_s \, \mathrm{d}\langle M\rangle_s
\end{aligned} \tag{19}$$

and that

$$\begin{aligned}
&\varepsilon_n^{-1} \sum_{j=0}^{\infty} \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} \int_{\tau_j^n \wedge t}^{s} |\psi_u^\delta| \, \mathrm{d}\langle M\rangle_u |\varphi_s| \, \mathrm{d}\langle M\rangle_s \\
&\qquad \le C^\delta \varepsilon_n^{-1} \sum_{j=0}^{\infty} [M]_{j,n}^t \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} |\varphi_u| \, \mathrm{d}\langle M\rangle_u \\
&\qquad \le C^\delta \varepsilon_n^{-1} \sum_{j=0}^{\infty} \biggl\{ |[M]_{j,n}^t|^3 \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} |\varphi_u|^2 \, \mathrm{d}\langle M\rangle_u \biggr\}^{1/2} \\
&\qquad \le C^\delta \biggl\{ \varepsilon_n^{-2} \sum_{j=0}^{\infty} |[M]_{j,n}^t|^3 \biggr\}^{1/2} \biggl\{ \int_0^t |\varphi_u|^2 \, \mathrm{d}\langle M\rangle_u \biggr\}^{1/2}
\end{aligned}$$

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

$$\varepsilon_n^{-2} \sum_{j=0}^{\infty} |[M]_{j,n}^t|^3 = O_p(\varepsilon_n^2) \tag{20}$$

since

$$\varepsilon_n^{-2} \sum_{j=0}^{N[\tau^n]_t} G_{j,n}^6 = O_p(\varepsilon_n^2), \qquad \varepsilon_n^{-4} \sum_{j=0}^{N[\tau^n]_t} G_{j,n}^{12} \to 0$$

in probability by Condition 2.3. Hence, the first term of the right-hand side of (19) converges to 0 uniformly in $v \in [0,t]$ in probability. For the second term, we have

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

Again using Lemmas A.2, A.4 and the Burkholder–Davis–Gundy inequality, we have

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

Since δ_1, δ_2 can be arbitrarily small, this implies (18). Next, we shall prove

$$
\varepsilon_n^{-1} (A - A^n) \cdot M_v^Y \to 0
$$

uniformly in $v \in [0,t]$ in probability. By the Lenglart inequality

$$
P\Big[\sup_{v \in [0,t]} |\varepsilon_n^{-1} (A - A^n) \cdot M_v^Y| > \delta_1 \Big] \le \frac{\delta_2}{\delta_1^2} + P[\varepsilon^{-2} \langle (A - A^n) \cdot M^Y \rangle_t > \delta_2]
$$

for any $\delta_1, \delta_2 > 0$, it suffices to see

$$
\varepsilon^{-2} \langle (A - A^n) \cdot M^Y \rangle_t = \varepsilon_n^{-2} \big((A - A^n)^2 \kappa \big) \cdot \langle M \rangle_t \to 0 \tag{21}
$$

in probability. Observe that

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

and that

$$\sup_{j\geq 0}\int_{\tau_j^n\wedge t}^{\tau_{j+1}^n\wedge t} |\psi_u|^2\,\mathrm{d}\langle M\rangle_u \to 0$$

in probability by Lemma A.4. Then (21) follows from (20).

Step (b). Let us show that

$$\varepsilon_n^{-1}\big((\bar{M}-\bar{M}^n)\varphi\big)\cdot\langle M\rangle_v \to \tfrac{1}{3}(b\varphi\gamma)\cdot\langle M\rangle_v \tag{22}$$

uniformly in $v\in[0,t]$ in probability. Fix $\delta_1,\delta_2>0$ arbitrarily and take a bounded left-continuous 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

$$\begin{aligned}&\big|\varepsilon_n^{-1}\big((\bar{M}-\bar{M}^n)(\varphi-\varphi^\delta)\big)\cdot\langle M\rangle_v\big|\\ &\quad\leq\sqrt{|\varphi-\varphi^\delta|^2\cdot\langle M\rangle_t}\sqrt{\varepsilon_n^{-2}(\bar{M}-\bar{M}^n)^2\cdot\langle M\rangle_t}\end{aligned}\tag{23}$$

and

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

by Lemma 2.9. Also note that

$$|(b\varphi)\cdot\langle M\rangle_v-(b\varphi^\delta)\cdot\langle M\rangle_v|\leq\sqrt{b^2\cdot\langle M\rangle_t}\sqrt{|\varphi-\varphi^\delta|^2\cdot\langle M\rangle_t}. \tag{24}$$

Since δ_1,δ_2 can be arbitrarily small, the estimates (23) and (24) ensure that we can suppose φ is a bounded left-continuous adapted process without loss of generality. Then, putting

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

we also have that

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

Note that

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

in probability as $n\to\infty$ because φ is now assumed to be bounded and left-continuous. Then we obtain (22) by applying Lemma 2.9.

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

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

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

$$(1)\ \langle \hat{D}^n, M^Y\rangle_t \to 0,$$
$$(2)\ \langle \hat{D}^n\rangle_t \to \tfrac{1}{6}(c^2\gamma^2\kappa)\cdot\langle M\rangle_t,$$
$$(3)\ \langle \hat{D}^n, \hat{M}\rangle_t \to 0$$

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

$$\langle D^n, M^Y\rangle_t = \varepsilon_n^{-1}\big((\bar{M} - \bar{M}^n)\kappa\big)\cdot\langle M\rangle_t \to \tfrac{1}{3}(b\gamma\kappa)\cdot\langle M\rangle_t$$

in probability. This is shown in the same manner as (22). In order to see the second convergence, fix $\delta_1, \delta_2 > 0$ arbitrarily and take a bounded left-continuous 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

$$\langle D^n\rangle = \varepsilon_n^{-2}\big((\bar{M} - \bar{M}^n)^2\kappa^\delta\big)\cdot\langle M\rangle + \varepsilon_n^{-2}\big((\bar{M} - \bar{M}^n)^2(\kappa - \kappa^\delta)\big)\cdot\langle M\rangle \tag{25}$$

and the second term is negligible since

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

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

$$\kappa_s^{\delta,n} = \kappa^\delta_{\tau_j^n} \qquad \text{for } j \ge 0 \text{ with } s \in [\tau_j^n, \tau_{j+1}^n),$$

we can replace κ^δ with $\kappa^{\delta,n}$ in the first term of (25) by the same argument. Then, we have from Lemma 2.9 that

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

in probability. Since

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

it remains only to show

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

in probability. Since the left-hand side is

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

the convergence follows from the same argument as for (22). $\square$

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 $C[0, T)$ is a consequence of the fact that the convergence of Z^n/ε_n is stable. To show (10), we first notice that

$$G^4_{j,n}/G^2_{j,n} - \tfrac{3}{4}|G^3_{j,n}/G^2_{j,n}|^2 \geq |G^2_{j,n}/G^1_{j,n}|^2 \qquad \text{a.s.,}$$

which follows from Lemma B.3. In light of Lemma A.4 and Condition 2.4, this inequality implies

$$(Hc^2)\cdot\langle M\rangle \geq \tfrac{1}{12}(Hb^2)\cdot\langle M\rangle + (H\zeta^{-2})\cdot\langle M\rangle \geq (H\zeta^{-2})\cdot\langle M\rangle \qquad \text{a.s.}$$

for any $H \in \mathcal{P}^1_M$ with $H \geq 0$. Thus, we have

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

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

$$G^4_{j,n}/G^2_{j,n} - |G^3_{j,n}/G^2_{j,n}|^2 \geq G^2_{j,n} \qquad \text{a.s.,}$$

which follows from Lemma B.2. In light of Lemma A.4 and Condition 2.5, this inequality implies

$$(Hc^2)\cdot\langle M\rangle \geq \tfrac{1}{3}(Hb^2)\cdot\langle M\rangle + (Hq^2)\cdot\langle M\rangle \geq (Hq^2)\cdot\langle M\rangle \qquad \text{a.s.}$$

for any $H \in \mathcal{P}^1_M$ with $H \geq 0$. Thus, we have

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

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,

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

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

$$\tau^n_0 = 0, \qquad \tau^n_{j+1} = \hat{\tau}^n_{j+1} \wedge T, \qquad \hat{\tau}^n_{j+1} = \inf\{t > \tau^n_j; |M_t - M_{\tau^n_j}| = \varepsilon_n g_{\tau^n_j}\} \tag{26}$$

satisfies Conditions 2.4 and 2.5 with

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

This is because it holds uniformly in $j = 0, 1, \dots, N[\tau^n]_t$ with $t < T$ that

$$\begin{aligned} G_{j,n}^k &= E[(M_{\hat{\tau}_{j+1}^n} - M_{\tau_j^n})^k | \mathcal{F}_{\tau_j^n}] \\ &\quad - E\big[(M_{\hat{\tau}_{j+1}^n} - M_{\tau_j^n})^k 1_{\{\hat{\tau}_{j+1}^n > T\}} | \mathcal{F}_{\tau_j^n}\big] + E\big[(M_T - M_{\tau_j^n})^k 1_{\{\hat{\tau}_{j+1}^n > T\}} | \mathcal{F}_{\tau_j^n}\big] \\ &= \tfrac{1}{2}\big(1 + (-1)^k\big)\varepsilon_n^k g_{\tau_j^n}^k + O_p(\varepsilon_n^k) P[\hat{\tau}_{j+1}^n > T | \mathcal{F}_{\tau_j^n}] \\ &= \tfrac{1}{2}\big(1 + (-1)^k\big)\varepsilon_n^k g_{\tau_j^n}^k + o_p(\varepsilon_n^k). \end{aligned}$$

Here we have used the fact that

$$E[(M_{\hat{\tau}_{j+1}^n} - M_{\tau_j^n})^k | \mathcal{F}_{\tau_j^n}] = \tfrac{1}{2}\big(1 + (-1)^k\big)\varepsilon_n^k g_{\tau_j^n}^k,$$

which follows from a consequence of the optional sampling theorem

$$P[M_{\hat{\tau}_{j+1}^n} = M_{\tau_j^n} + \varepsilon_n g_{\tau_j^n} | \mathcal{F}_{\tau_j^n}] = P[M_{\hat{\tau}_{j+1}^n} = M_{\tau_j^n} - \varepsilon_n g_{\tau_j^n} | \mathcal{F}_{\tau_j^n}] = \tfrac{1}{2}.$$

Also, we have used that

$$\sup_{\tau_j^n \le t \le \hat{\tau}_{j+1}^n} |M_t - M_{\tau_j^n}| \le \varepsilon_n g_{\tau_j^n}$$

and that for $g \in \mathcal{P}_0$,

$$\sup_{0 \le s \le t} g_s < \infty \qquad \text{a.s.}$$

to see that

$$\frac{E[-(M_{\hat{\tau}_{j+1}^n} - M_{\tau_j^n})^k 1_{\{\hat{\tau}_{j+1}^n > T\}} + (M_T - M_{\tau_j^n})^k 1_{\{\hat{\tau}_{j+1}^n > T\}} | \mathcal{F}_{\tau_j^n}]}{P[\hat{\tau}_{j+1}^n > T | \mathcal{F}_{\tau_j^n}]} = O_p(\varepsilon_n^k).$$

To see $P[\hat{\tau}_{j+1}^n > T | \mathcal{F}_{\tau_j^n}] = o_p(1)$ uniformly in $j = 0, 1, \dots, N[\tau^n]_t$, recall that $\langle M \rangle$ is strictly increasing on $[T - \delta, T)$ so that $\lim_{n\to\infty} \sup_j \hat{\tau}_{j+1}^n \le (T - \delta) \vee t$.

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 by* (26) *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 by* (26) *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.2 and the Cauchy–Schwarz inequality. Karandikar [13] studied a scheme which is defined by (26) 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.2. It is in fact optimal if $X = Y$.

Note that Lemma B.2 gives a more precise estimate

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

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 by* (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^n_t/ε_n converges to a mixed normal distribution with the asymptotic conditional mean*

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

and the asymptotic conditional variance V_t satisfying

$$V_t = \tfrac{1}{6}(c^2\gamma^2) \cdot \langle Y\rangle_t \geq \tfrac{1}{6}\big\{\big(\tfrac{1}{3}\beta^2 + \delta\big)\gamma^2\big\} \cdot \langle Y\rangle_t \qquad a.s.$$

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

$$\tau^n_{j+1} = \inf\big\{t > \tau^n_j; M_t - M_{\tau^n_j} \geq \varepsilon_n k_{\tau^n_j}\sqrt{\delta_{\tau^n_j}} \text{ or } M_t - M_{\tau^n_j} \leq -\varepsilon_n k^{-1}_{\tau^n_j}\sqrt{\delta_{\tau^n_j}}\big\}$$

with $\tau^n_0 = 0$, 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 the conservative delta hedging of [20] 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 stand for an asset price process and assume that it is a positive continuous semimartingale satisfying

$$\mathrm{d}S_t = S_t(\mu_t\,\mathrm{d}t + \sigma_t\,\mathrm{d}W_t)$$

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

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

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

$$\frac{\partial p}{\partial \Sigma} = \frac{1}{2} S^2 \frac{\partial^2 p}{\partial S^2}, \qquad \frac{\partial p}{\partial R} = S \frac{\partial p}{\partial S} - p. \tag{27}$$

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

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

where $r > 0$ is a risk-free rate. Then, Itô's formula and (27) yield

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

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 (27) imply that p is increasing in Σ, so that

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

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 $K \to \infty$, $P[\eta_K \geq T] \to 1$, so that a hedge error due to the incompleteness of market converges to 0. Contracts such as variance swaps serve as insurance against the event $\eta_K < T$ for predetermined K which is not so large. See [21] 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. A natural approximation π^n of the strategy π is defined as

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

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

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

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

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

where

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

Note that Γ is what is called gamma in financial practice. In this case, $\tau^n \in \mathcal{T}_2(X)$ with

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

and so we have

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

as the asymptotic distribution of Z_t^n/ε_n, where Z is an independent standard normal variable. This scheme is efficient in that it asymptotically minimizes the mean squared replication error for a conditionally given number of transactions. The number of transactions $N[\tau^n]_t$ is, of course, random; it is large 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 given by

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

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

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

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

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

The inequality for the asymptotic conditional variance

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

follows directly from the Cauchy–Schwarz inequality.

Karandikar's scheme is defined as

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

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

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

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

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

The inequality for the asymptotic conditional variance

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

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

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

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

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

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

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

which is attained by τ^n defined as

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

This scheme is efficient in that it asymptotically minimizes the mean squared replication error for a conditionally given amount of linear transaction costs. In other words, denoting by L_t the mixed normal limit distribution of $C_t^n Z_t^n$, the asymptotic conditional variance of $Z_t^n \approx L_t / C_t^n$ is minimized by the above scheme subject to the same amount of the linear transaction costs in a conditional sense. For the Black–Scholes model, a hedging strategy under the linear transaction costs was given by Leland [19]. It is designed to absorb the costs by an adjustment of volatility parameter in the delta hedging strategy and is validated only for deterministic trading times under the particular model. An L^2 analysis of the Leland–Lott strategy with nonequidistant deterministic trading times is given by Denis and Kabanov [4]. Our strategy leaves the costs unabsorbed but is an optimal discretization scheme being valid in a model-free framework.

5. The Euler–Maruyama approximation. Here we propose alternative discretization schemes for the Euler–Maruyama approximation which outperform the usual time-equidistant scheme. Let us consider the stochastic differential equation

$$\begin{aligned} \mathrm{d}\Xi_t &= \mu(\Xi_t, \eta_t)\,\mathrm{d}t + \sigma(\Xi_t, \eta_t)\,\mathrm{d}W_t, \\ \mathrm{d}\eta_t &= \theta(\eta_t)\,\mathrm{d}t, \end{aligned}$$

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

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

where $\bar{\Xi}_t^n = \Xi_{\tau_j^n}^n$, $\bar{\eta}_t^n = \eta_{\tau_j^n}^n$ for $j \geq 0$ with $t \in [\tau_j^n, \tau_{j+1}^n)$. The standard choice is $\tau_j^n = j/n$. The convergence rate of the approximation has been extensively investigated; see, for example, [15] for a well-known strong approximation theorem and [1, 2, 16, 17] for weak approximation theorems. Newton [22] treated passage times. Cambanis and Hu [3] studied efficiency of deterministic nonequidistant scheme. Hofmann, Müller-Gronbach and Ritter [9] treated a class of adaptive schemes. Here we exploit a result of [11, 18] 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 that $\mu, \sigma, 1/\sigma, \theta$ and their derivatives are bounded without loss of generality. Suppose that $\tau^n \in \mathcal{T}(W)$. By applying Theorem 2.6 to $X = Y = W$, we have that $\varepsilon_n^{-1}(W - W^n) \cdot W$ converges $\mathcal{F}$-stably to

$$Z = \tfrac{1}{3} b \cdot W + \tfrac{1}{\sqrt{6}} c \cdot W',$$

where $W_t^n = W_{\tau_j^n}$ for $j \geq 0$ with $t \in [\tau_j^n, \tau_{j+1}^n)$ and W' is an independent standard Brownian motion. Put $L_t^n = \varepsilon_n^{-1}(\Xi_t^n - \Xi)$. Then, applying [18], we have that L^n converges to a process L which satisfies

$$\mathrm{d}L_t = \partial_1\mu(\Xi_t, \eta_t)L_t\,\mathrm{d}t + \partial_1\sigma(\Xi_t, \eta_t)[L_t\,\mathrm{d}W_t - \sigma(\Xi_t, \eta_t)\,\mathrm{d}Z_t],$$

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

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

where

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

Therefore, the distribution of L_t is mixed normal with conditional mean

$$-\frac{1}{3}e_t \int_0^t e_s^{-1}\sigma(\Xi_s, \eta_s)\,\partial_1\sigma(\Xi_s, \eta_s)b_s[\mathrm{d}W_s - \partial_1\sigma(\Xi_s, \eta_s)\,\mathrm{d}s]$$

and conditional variance

$$\frac{1}{6}e_t^2 \int_0^t e_s^{-2}\sigma(\Xi_s, \eta_s)^2\,\partial_1\sigma(\Xi_s, \eta_s)^2 c_s^2\,\mathrm{d}s. \tag{28}$$

PROPOSITION 5.1. *For any $T > 0$, the space-equidistant scheme τ_{sp}^n defined by* (26) *with $M = W$, $\varepsilon_n = n^{-1/2}$, $g = 1$ is three times more efficient than the usual time-equidistant scheme $\tau_{\mathrm{tm}}^n = \{j/n\}$ in the following sense. For any $t \in [0, T)$*:

- *$E[N_t^n] \leq nt$ and $N_t^n/n \to t$ in probability as $n \to \infty$ for both $N^n = N[\tau_{\mathrm{sp}}^n]$ and $N^n = N[\tau_{\mathrm{tm}}^n]$*;
- *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,

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

while

$$b_s \equiv 0, \qquad a_s^2 \equiv c_s^2 \equiv 3, \qquad N_t^n = [nt]$$

for the time-equidistant case. $\square$

Newton [22] 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, for a stochastic volatility model

$$\mathrm{d}\Xi_t = \hat{\mu}(t, \Xi_t)\,\mathrm{d}t + \hat{\sigma}(t, V_t)\big[\rho(t, V_t)\,\mathrm{d}W_t^1 + \sqrt{1-\rho(t, V_t)^2}\,\mathrm{d}W_t^2\big],$$

$$\mathrm{d}V_t = \mu(t, V_t)\,\mathrm{d}t + \sigma(t, V_t)\,\mathrm{d}W_t^1$$

with a two-dimensional standard Brownian motion (W^1, W^2), a scheme defined by (26) 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, the discretization error is determined by only the conditional moments of the increments of integrand, which is a function of V independent of W^2 in this example.

Next, let us consider to minimize (28) in case that $\partial_1\sigma$ is nondegenerate. Define τ^n as

$$\tau_0^n = 0, \qquad \tau_{j+1}^n = \inf\{t > \tau_j^n; |W_t - W_{\tau_j^n}|^2 = \varepsilon(\tau_j^n)\} \wedge T, \tag{29}$$

where

$$\varepsilon(\tau_j^n) = \frac{\varepsilon_n^2 \hat{e}_{\tau_j^n}}{\sigma(\Xi_{\tau_j^n}^n, \eta_{\tau_j^n}^n)\,\partial_1\sigma(\Xi_{\tau_j^n}^n, \eta_{\tau_j^n}^n)}$$

and

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

Then, Condition 2.5 is satisfied with

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

Proposition 3.2 implies that this adaptive scheme attains a lower bound for (28) among $\tau^n \in \mathcal{T}_2(W)$. In this sense, this scheme is optimal. A disadvantage of this scheme is the difficulty in estimating 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 would be better to use

$$\tau_0^n = 0, \qquad \tau_{j+1}^n = \inf\{t > \tau_j^n; |W_t - W_{\tau_j^n}|^2 = \varepsilon(\tau_j^n) \vee \varepsilon_n'\} \wedge T$$

for $\varepsilon_n' > 0$ 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

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

on a computer for a given ε. It is sufficient that the distribution function F_ε of τ is available because

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

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

$$(\tau, W_\tau - W_0) \sim \bigl(F_\varepsilon^{-1}(2U - \lfloor 2U \rfloor), \varepsilon(2\lfloor 2U \rfloor - 1)\bigr).$$

Here we use the fact that $2U - \lfloor 2U \rfloor$ is a uniform random variable on $(0, 1)$ independent to $\lfloor 2U \rfloor$ and $P[\lfloor 2U \rfloor = 0] = P[\lfloor 2U \rfloor = 1] = 1/2$. It is known that the density of τ is given by

$$\frac{2}{\sqrt{2\pi t^3}} \sum_{n=-\infty}^{\infty} (4n+1)\varepsilon \exp\biggl\{-\frac{(4n+1)^2\varepsilon^2}{2t}\biggr\}.$$

(See [14], Exercise 2.8.11.) Using the fact that

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

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

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

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

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

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 (29). Also note that

$$G(x) \le 4\bigl(1 - \Phi(x)\bigr), \qquad G^{-1}(y) \le \Phi^{-1}(1 - y/4).$$

These inequalities will be useful in numerical calculation of G^{-1} for sufficiently small y (large x). Besides, if $x \ge 3$, $G(x) \approx 4(1 - \Phi(x))$ 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 [10] and [12], Theorem IX.7.3, which extends a result of [23]. 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 A.1. *Let $\{Z^n\}$ be a sequence of continuous $\{\mathcal{F}_t\}$-local martingales. Suppose that there exists an $\{\mathcal{F}_t\}$-adapted continuous process $V = \{V_t\}$ such that for all $\hat{M} \in \mathcal{M}^\perp$, $t \in [0, \infty)$,*

$$\langle Z^n, \hat{M}\rangle_t \to 0, \qquad \langle Z^n, M\rangle_t \to 0, \qquad \langle Z^n\rangle_t \to V_t$$

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 following lemma is repeatedly used in our proofs.

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

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

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

$$N^n(\lambda) \le N^n(\lambda_0) \qquad \text{a.s. for all } \lambda \in \Lambda \text{ and } \sum_{j=1}^{N^n(\lambda_0)} P[|U^n_j|^2 | \mathcal{H}^n_{j-1}] \to 0$$

as $n \to \infty$, then

$$\sup_{\lambda\in\Lambda} \left| \sum_{j=1}^{N^n(\lambda)} U^n_j - \sum_{j=1}^{N^n(\lambda)} P[U^n_j | \mathcal{H}^n_{j-1}] \right| \to 0$$

in probability as $n \to \infty$.

PROOF. Note that

$$\sup_{\lambda\in\Lambda} \left| \sum_{j=1}^{N^n(\lambda)} U^n_j - \sum_{j=1}^{N^n(\lambda)} E[U^n_j | \mathcal{H}^n_{j-1}] \right| \le \sup_{k\in\mathbb{N}} \left| \sum_{j=1}^{k} V^n_j \right|,$$

where

$$V^n_j = (U^n_j - E[U^n_j | \mathcal{H}^n_{j-1}]) 1_{j \le N^n(\lambda_0)}.$$

By the Lenglart inequality, we have

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

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

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

in probability. $\square$

The following lemma is proved by a simple application of the monotone class theorem, so its proof is omitted.

LEMMA A.3. *Let $K \subset \mathbb{N}$ be a finite set. For all $H \in \bigcap_{k\in K} \mathcal{P}_M^k$, $t \in [0, T)$ and $\delta_1, \delta_2 > 0$, there exists a bounded adapted left-continuous process $\hat{H}$ with $|\hat{H}| \leq |H|$ such that*

$$P\left[\sum_{k\in K}\{|H - \hat{H}|^k \cdot \langle M\rangle_t\}^{1/k} > \delta_1\right] < \delta_2.$$

The following lemma is taken from [6].

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 $t \in [0, T)$,*

$$\sup_{0\leq j\leq N[\tau^n]_t, s\geq 0} |M_{\tau_{j+1}^n\wedge s} - M_{\tau_j^n\wedge s}|^2 = o_p(\varepsilon_n), \tag{30}$$

as well as

$$\sup_{0\leq j\leq N[\tau^n]_t} |\langle M\rangle_{\tau_{j+1}^n} - \langle M\rangle_{\tau_j^n}| = o_p(\varepsilon_n), \tag{31}$$

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

$$N[\tau^n]_t \to \infty \qquad a.s. \tag{32}$$

as $n \to \infty$. Moreover, for all locally bounded adapted left-continuous process f, it holds

$$\sum_{j=0}^{N[\tau^n]_t} f_{\tau_j^n} G_{j,n}^2 \to \int_0^t f_s \,\mathrm{d}\langle M\rangle_s \tag{33}$$

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 $E[\langle M\rangle_T] < \infty$. To show (30), we use Doob's maximal inequality to have

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

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

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

which implies (30) since

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

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

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

for $k = 3, 6$, which implies (31) in the same manner. Note that (32) follows from

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

To see (33), again in light of Lemma A.2, it suffices to observe that

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

and

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

□

APPENDIX B: KURTOSIS–SKEWNESS INEQUALITY

DEFINITION B.1. For a random variable X (resp., distribution μ), we say X (resp., μ) is Bernoulli if the support of X (resp., μ) consists of at most two points.

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

$$E[X^4]E[X^2]-|E[X^3]|^2 \geq |E[X^2]|^3. \tag{34}$$

The equality is attained if and only if X is Bernoulli.

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

$$\begin{aligned} E[X^3]^2 &= E\big[X(X^2-E[X^2])\big]^2 \leq E[X^2]E\big[|X^2-E[X^2]|^2\big] \\ &= E[X^2](E[X^4]-|E[X^2]|^2). \end{aligned}$$

This also implies that if the equality holds, then X and $X^2-E[X^2]$ must be linearly dependent, so that X must be Bernoulli. It can be directly checked that the equality is attained if X is Bernoulli. □

The following lemma gives a similar inequality to Lemma B.2. The proof is, however, rather different and the result itself is seemingly new.

LEMMA B.3. *Let X be a random variable with $E[X]=0$ and $E[X^4]<\infty$. Then*,

$$E[X^4]E[X^2]|E[|X|]|^2-\tfrac{3}{4}|E[X^3]|^2|E[|X|]|^2 \geq |E[X^2]|^4. \tag{35}$$

The equality is attained if and only if X is Bernoulli.

PROOF. We divide the proof into four 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

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

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

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

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

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

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

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

with

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

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

$$\tilde{a}_1=\frac{a_0p_0+a_1p_1}{p_0+p_1}, \qquad \tilde{p}_1=p_0+p_1,$$
$$\tilde{a}_j=a_j, \qquad \tilde{p}_j=p_j, \qquad 2\leq j\leq n$$

and

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

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

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

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

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

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

$$Q_j=\mu_jP_{0j}+(1-\mu_j)P_{1j},$$

where we define

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

Putting

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

we have

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

which completes the induction.

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

$$f(u, v, w, y) = u - \tfrac{3}{4}v^2/w - w^3/y^2 \tag{36}$$

is a concave function. Note that the inequality (35) follows from steps (a)–(c) since every distribution can be approximated arbitrarily close by a distribution supported by a finite set and by the concavity of f,

$$f\Bigl(\sum_j \lambda_j E_j[(X^4, X^3, X^2, |X|)]\Bigr) \geq \sum_j \lambda_j f(E_j[(X^4, X^3, X^2, |X|)])$$

for any mixture distribution $E = \sum_j \lambda_j E_j$. By a straightforward calculation, the Hessian matrix of f is given by

$$H = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -\dfrac{3}{2w} & \dfrac{3v}{2w^2} & 0 \\ 0 & \dfrac{3v}{2w^2} & -\dfrac{3v^2}{2w^3} - \dfrac{6w}{y^2} & \dfrac{6w^2}{y^3} \\ 0 & 0 & \dfrac{6w^2}{y^3} & -\dfrac{6w^3}{y^4} \end{pmatrix}. \tag{37}$$

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 semidefinite.

Step (d). It remains to show that the equality holds only if X is Bernoulli. Suppose that there exists a random variable X with $E[X] = 0$ such that the equality holds in (35) which is not Bernoulli. Recall that the equality holds in (34) 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

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

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

$$f\bigl(\lambda m_1 + (1-\lambda)m_2\bigr) = 0 = \lambda f(m_1) + (1-\lambda) f(m_2), \tag{38}$$

where f is defined by (36) and

$$m_i = \Bigl(\int a^4 P_i(\mathrm{d}a), \int a^3 P_i(\mathrm{d}a), \int a^2 P_i(\mathrm{d}a), \int |a| P_i(\mathrm{d}a)\Bigr)', \qquad i = 1, 2.$$

Here $'$ means the transpose of matrix. By the concavity of f, (38) holds for all $\lambda \in (0,1)$. By (37), the eigenvectors of the Hessian matrix H associated to the eigenvalue 0 are

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

Therefore, (38) implies that there exists a constant c such that $\bar{m}_2 = c\bar{m}_1$, where

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

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

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

and

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

Then we obtain that

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

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. The proof is complete. $\square$

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** 43–60. MR1367666

[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** 93–128. MR1401964

[3] CAMBANIS, S. and HU, Y. (1996). Exact convergence rate of the Euler–Maruyama scheme, with application to sampling design. *Stochastics Stochastics Rep.* **59** 211–240. MR1427739

[4] DENIS, E. and KABANOV, Y. (2010). Mean square error for the Leland–Lott hedging strategy: Convex pay-offs. *Finance Stoch.* **14** 625–667. MR2738026

[5] FUKASAWA, M. (2010). Central limit theorem for the realized volatility based on tick time sampling. *Finance Stoch.* **14** 209–233. MR2607763

[6] FUKASAWA, M. (2010). Realized volatility with stochastic sampling. *Stochastic Process. Appl.* **120** 829–852. MR2610328

[7] GEISS, S. and TOIVOLA, A. (2009). Weak convergence of error processes in discretizations of stochastic integrals and Besov spaces. *Bernoulli* **15** 925–954. MR2597578

[8] HAYASHI, T. and MYKLAND, P. A. (2005). Evaluating hedging errors: An asymptotic approach. *Math. Finance* **15** 309–343. MR2132193

[9] HOFMANN, N., MÜLLER-GRONBACH, T. and RITTER, K. (2001). The optimal discretization of stochastic differential equations. *J. Complexity* **17** 117–153. MR1817611

[10] JACOD, J. (1994). Limit of random measures associated with the increments of a Brownian semimartingale. Unpublished manuscript.

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

[12] JACOD, J. and SHIRYAEV, A. N. (2002). *Limit Theorems for Stochastic Processes*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **288**. Springer, Berlin. MR0959133

[13] KARANDIKAR, R. L. (1995). On pathwise stochastic integration. *Stochastic Process. Appl.* **57** 11–18. MR1327950

[14] KARATZAS, I. and SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. Springer, New York. MR1121940

[15] KLOEDEN, P. E. and PLATEN, E. (1992). *Numerical Solution of Stochastic Differential Equations*. *Applications of Mathematics* (*New York*) **23**. Springer, Berlin. MR1214374

[16] KOHATSU-HIGA, A. (2001). Weak approximations. A Malliavin calculus approach. *Math. Comp.* **70** 135–172. MR1680895

[17] KONAKOV, V. and MAMMEN, E. (2002). Edgeworth type expansions for Euler schemes for stochastic differential equations. *Monte Carlo Methods Appl.* **8** 271–285. MR1931967

[18] KURTZ, T. G. and PROTTER, P. (1991). Wong–Zakai corrections, random evolutions, and simulation schemes for SDEs. In *Stochastic Analysis* 331–346. Academic Press, Boston, MA. MR1119837

[19] LELAND, H. E. (1985). Option pricing and replication with transaction costs. *J. Finance* **40** 1283–1301.

[20] MYKLAND, P. A. (2000). Conservative delta hedging. *Ann. Appl. Probab.* **10** 664–683. MR1768218

[21] MYKLAND, P. A. (2003). The interpolation of options. *Finance Stoch.* **7** 417–432. MR2014243

[22] NEWTON, N. J. (1990). An efficient approximation for stochastic differential equations on the partition of symmetrical first passage times. *Stochastics Stochastics Rep.* **29** 227–258. MR1041038

[23] ROOTZÉN, H. (1980). Limit distributions for the error in approximations of stochastic integrals. *Ann. Probab.* **8** 241–251. MR0566591

CENTER FOR THE STUDY OF FINANCE
AND INSURANCE
JAPAN SCIENCE AND TECHNOLOGY
AGENCY, CREST
OSAKA UNIVERSITY
1-3 MACHIKANEYAMA, TOYONAKA
OSAKA
JAPAN
E-MAIL: fukasawa@sigmath.es.osaka-u.ac.jp