J. Japan Statist. Soc.
Vol. 35 No. 2 2005 171–203

MONTE CARLO SIMULATION WITH ASYMPTOTIC METHOD

Akihiko Takahashi* and Nakahiro Yoshida**

We shall propose a new computational scheme with the asymptotic method to achieve variance reduction of Monte Carlo simulation for numerical analysis particularly for finance. We not only provide general scheme of our method, but also show its effectiveness through numerical examples such as computing optimal portfolio and pricing an average option. Finally, we show mathematical validity of our method.

Key words and phrases: Asymptotic method, average options, derivatives, finance, Malliavin calculus, Monte Carlo simulation, optimal portfolio.

1. Introduction

We propose a new method to increase efficiency of Monte Carlo simulation. We utilize the analytic approximation based on the asymptotic method to achieve variance reduction of the Monte Carlo simulation especially for numerical problems in finance. The idea of the method is as follows. Suppose that $F(w)$ is a Wiener functional and our objective is the evaluation of the expectation of $F(w)$. That is,

$$\boldsymbol{V} := \boldsymbol{E}[F(w)].$$

A typical estimate of $\boldsymbol{V}$ may be obtained by a naive Monte Carlo simulation based on Euler-Maruyama approximation. That is,

$$\boldsymbol{V}(n,N) = \frac{1}{N}\sum_{j=1}^{N}[F^{(n)}]_j,$$

where $[Z]_j$ $(j = 1,\dots,N)$ denote independent copies of the random variable Z, $Z^{(n)}$ represents a random variable obtained by discretization of Z depending on a continuous time parameter and n is the number of time points in discretization. We introduce a modified estimator $\boldsymbol{V}^*(n,N)$ defined by

$$\boldsymbol{V}^*(n,N) = \boldsymbol{E}[\hat{F}] + \frac{1}{N}\sum_{j=1}^{N}[F^{(n)} - \hat{F}^{(n)}]_j$$

where $\boldsymbol{E}[\hat{F}]$ is assumed to be analytically known. Intuitively, if we are able to find $\hat{F}$ such that the errors of $[F^{(n)}]_j$ and $[\hat{F}^{(n)}]_j$, that is, $[F^{(n)}]_j - \boldsymbol{V}$ and

Received May 22, 2004. Revised November 1, 2004. Accepted March 23, 2005.

*Graduate Scool of Economics, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan.

**Graduate Scool of Mathematical Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8914, Japan.

$[\hat{F}^{(n)}]_j - \boldsymbol{E}[\hat{F}]$ take close numerical values for each independent copy j, then $\boldsymbol{V}^*(n, N)$ becomes a better estimate since the error of each j in $\boldsymbol{V}^*(n, N)$ which is represented by the difference of the errors of $[F^{(n)}]_j$ and $[\hat{F}^{(n)}]_j$ becomes small. As seen below, the asymptotic method (or perturbation method) provides such $\hat{F}$. That is, $\hat{F}$ obtained by the asymptotic method has a strong correlation with F, and $\boldsymbol{E}[\hat{F}]$ is evaluated analytically.

Variance reduction methods in Monte Carlo simulations arising from finance has been examined by various authors. (See chapter 4 of Glasserman (2003) for the detail.) Among them, our method may be somewhat similar to the *control variate technique*. (For instance, see chapter 3 of Robert and Casella (2000) or section 4.1 of Glasserman (2003) on basics of *control variate technique*.) However, the main difficulty in the control variate technique is that it is generally difficult to find $\hat{F}$ strongly correlated with F whose expectation $\boldsymbol{E}[\hat{F}]$ can be analytically obtained. A well-known exception is pricing of an *arithmetic* average option under a log-normal price process where a *geometric* average option can be used as a control variate (Kemna and Vorst (1990)). However, this does not always work when the price process is not log-normal because the price of a *geometric* average option can not be analytically obtained in general. Newton (1994) derived theoretically optimal control variates, but this includes a term which is not easy to evaluate. He gave some approximations and claimed it was useful for some cases of numerical examples. Milshtein and Schoenmakers (2002) applied and extended Newton's idea to pricing of derivatives in finance without numerical examples.

Our method based on the perturbation of the stochastic differential equations overcomes the difficulty since the asymptotic method allows us to find such $\hat{F}$ in the unified way. In the following sections, we will show this idea more rigorously and concretely. We also note that our method may be used together with other acceleration methods such as the *antithetic variables technique* and an extrapolation method of Talay and Tubaro (1990) to pursue further variance reduction of Monte Carlo simulation. Moreover, an asymptotic expansion approach may be effectively applied with importance sampling technique developed by Newton (1994).

Asymptotic methods have been applied successfully to a broad class of Itô processes appearing in finance. Kunitomo and Takahashi (1992) proposed a normal approximation to evaluate average options in the Black-Scholes setting. Yoshida (1992b) applied the asymptotic expansion method to price path-dependent options for nonlinear price processes. This method was based on the Malliavin calculus and had been developed in statistics for stochastic processes (Yoshida (1992a, 1993)).

Takahashi (1999) presented a third-order pricing formula for plain options and second-order formulas for more complicated derivatives such as average options, basket options, and options with stochastic volatility in a general Markovian setting. Kunitomo and Takahashi (2001) derived expansions for interest rate models based on Heath-Jarrow-Morton (1992) which is not necessarily

Markovian, and provided pricing formulas for bond options (swap options), average options on interest rates. Takahashi (1995) also presented a second order scheme for average options on foreign exchange rates with stochastic interest rates in Heath-Jarrow-Morton framework.

Moreover, Takahashi and Yoshida (2004) extended the method to dynamic portfolio problems; starting with a result in Ocone and Karatzas (1991), they derived formulas for optimal portfolios associated with maximizing utility from terminal wealth in a general Markovian setting. Recently, Takahashi and Saito (2003) successfully applied the method to American options. For details of mathematical validity based on the Malliavin calculus and of further applications, see Kunitomo and Takahashi (2003a, 2003b, 2004).

The organization of the paper is as follows: In the next section, we will show our new scheme and state main theorems. In Section 3, we will give two examples to illustrate our method in finance; computing the market price of risk component in the optimal portfolio problem and pricing an average option. In Section 4, we will examine numerical examples for the problems discussed in Section 3. In Sections 5 and 6, we will give proofs of the main theorems. In Section 7, we will provide mathematical validity of the asymptotic method with square-root process used in the numerical examples.

2. Monte Carlo simulation with the asymptotic method

Let $(\Omega, \mathcal{F}, P)$ be probability space and $T \in (0, \infty)$ denotes some fixed time horizon. Process $w = \{(w^1(t), \dots, w^r(t))^*; t \in [0, T]\}$ is an $\boldsymbol{R}^r$-valued Brownian motion defined on $(\Omega, \mathcal{F}, P)$, and $\{\mathcal{F}_t\}$, $0 \le t \le T$ stands for P-augmentation of the natural filtration $\mathcal{F}_t^w = \sigma(w(s); 0 \le s \le t)$. Here we use the notation x^* as the transpose of x. Suppose that an $\boldsymbol{R}^D$-valued process $X_u(t, x)$ $(0 \le t \le u \le T, x \in \boldsymbol{R}^D)$ satisfy the stochastic integral equation:

$$X_u^\epsilon(t, x) = x + \int_t^u V_0(X_s^\epsilon(t, x), \epsilon) ds + \sum_{\alpha=1}^r \int_t^u V_\alpha(X_s^\epsilon(t, x), \epsilon) dw_s, \tag{2.1}$$

where ϵ is a parameter $\epsilon \in (0, 1]$ and $V_\alpha \in C_\uparrow^\infty(\boldsymbol{R}^D \times (0, 1]; \boldsymbol{R}^D)$, $\alpha = 0, 1, \dots, r$; $C_\uparrow^\infty(R^D \times (0, 1]; E)$ denotes the set of smooth mappings $f : \boldsymbol{R}^D \times (0, 1] \to E$ whose derivatives $\partial_x^{\boldsymbol{n}} \partial_\epsilon^k f(x, \epsilon)$ are of at most polynomial growth uniformly in ϵ for $\boldsymbol{n} \in \boldsymbol{Z}_+^D$ and $k \in \boldsymbol{Z}_+$. That is

$$\sup_{\epsilon \in (0,1]} |\partial_x^{\boldsymbol{n}} \partial_\epsilon^k f(x, \epsilon)| \le C_{n,k} (1 + |x|)^{C_{n,k}} \qquad \text{for some} \quad C_{n,k} > 0.$$

$\partial_x^{\boldsymbol{n}} \partial_\epsilon^k$ is defined by

$$\partial_x^{\boldsymbol{n}} \partial_\epsilon^k = \left(\frac{\partial}{\partial x_1}\right)^{n_1} \left(\frac{\partial}{\partial x_2}\right)^{n_2} \cdots \left(\frac{\partial}{\partial x_D}\right)^{n_D} \left(\frac{\partial}{\partial \epsilon}\right)^k$$

where $x = (x_m)_{1 \le m \le D}$ and $\boldsymbol{n} = (n_m)_{1 \le m \le D}$. We also assume that $(V_0, V_1, \dots, V_r)$ is *graded* according to $\boldsymbol{R}^D = \boldsymbol{R}^{d_1} \times \cdots \times \boldsymbol{R}^{d_q}$ in the sense of Definition 1 below.

DEFINITION 1. *A grading of $\boldsymbol{R}^D$ is a decomposition $\boldsymbol{R}^{d_1} \times \cdots \times \boldsymbol{R}^{d_q}$ with $\sum_{i=1}^{q} d_i = D$. The coordinates of a point in $\boldsymbol{R}^D$ are always arranged in an increasing order along the subspaces $\boldsymbol{R}^{d_i}$. We set $M_0 = 0$ and $M_l = \sum_{i=1}^{l} d_i$ for $1 \le l \le q$. We say that the coefficients $(V_0, V_1, \dots, V_r)$ are graded according to the grading $\boldsymbol{R}^D = \boldsymbol{R}^{d_1} \times \cdots \times \boldsymbol{R}^{d_q}$ if $V_\alpha^i(x, \epsilon)$, $\alpha = 0, 1, \dots, r$ depend on x only through the coordinates $(x_m)_{1 \le m \le M_l}$ when $M_{l-1} < i \le M_l$ where V_α^i denotes the i-th element of V_α.*

We further suppose that $\partial_{x^{(l)}}^{\boldsymbol{n}_l} V_\alpha^i(x, \epsilon)$, $\alpha = 0, 1, \dots, r$ are bounded for $\boldsymbol{n}_l \in \boldsymbol{Z}_+^{d_l}$ such that $|\boldsymbol{n}_l| \ge 1$ where $\boldsymbol{n}_l = (n_j)_{1 \le j \le d_l}$ and $|\boldsymbol{n}_l| = \sum_{j=1}^{d_l} n_j$; $x^{(l)} = (x_j^{(l)})_{1 \le j \le d_l}$ and $x_j^{(l)}$ denotes the $(M_{l-1} + j)$-th coordinate of $x \in \boldsymbol{R}^D$; $\partial_{x^{(l)}}^{\boldsymbol{n}_l}$ is defined by

$$\partial_{x^{(l)}}^{\boldsymbol{n}_l} = \left(\frac{\partial}{\partial x_1^{(l)}} \right)^{n_1} \left(\frac{\partial}{\partial x_2^{(l)}} \right)^{n_2} \cdots \left(\frac{\partial}{\partial x_{d_l}^{(l)}} \right)^{n_{d_l}}.$$

Due to Chapter II-5 of Bichteler *et al.* (1987), $X_u(t, x)$ admits a unique solution and $\sup_{0 \le u \le T} \boldsymbol{E}[|X_u(t, x)|^p] < \infty$ for all $p \ge 1$.

We finally note that the Markovian system (3.9) in Section 3 is an example of this class.

2.1. Smooth case

Suppose that $f \in C_\uparrow^\infty(\boldsymbol{R}^D; \boldsymbol{R})$, where $C_\uparrow^\infty(\boldsymbol{R}^D; \boldsymbol{R})$ denotes the set of smooth functions $f : \boldsymbol{R}^D \to \boldsymbol{R}$ whose derivatives are of at most polynomial growth. For stochastic approximation to $\boldsymbol{V} := \boldsymbol{E}[f(X_T^\epsilon(0, x))]$, an estimator by naive Monte Carlo simulation is given as

$$\boldsymbol{V}(\epsilon, n, N) = \frac{1}{N} \sum_{j=1}^{N} [f(\bar{X}_T^\epsilon)]_j. \tag{2.2}$$

Here $[Z]_j$ $(j = 1, \dots, N)$ denote independent copies of the random variable Z, and the Euler-Maruyama scheme $\bar{X}^\epsilon$ is defined by:

$$\bar{X}_u^\epsilon = x + \int_0^u V_0(\bar{X}_{\eta(s)}^\epsilon, \epsilon) ds + \sum_{\alpha=1}^{r} \int_0^u V_\alpha(\bar{X}_{\eta(s)}^\epsilon, \epsilon) dw_s \tag{2.3}$$

with $\eta(s) = [ns/T]T/n$.

In the sequel, we will consider a modified estimator for $\boldsymbol{V}$:

$$\boldsymbol{V}^*(\epsilon, n, N) = \boldsymbol{E}[f(X_T^{[0]}(0, x))] + \frac{1}{N} \sum_{j=1}^{N} [f(\bar{X}_T^\epsilon) - f(\bar{X}_T^{[0]})]_j, \tag{2.4}$$

where $X_T^{[0]}(0, x)$ and $\bar{X}_T^{[0]}$ denote $X_T^\epsilon(0, x)$ and $\bar{X}_T^\epsilon$ when $\epsilon = 0$ respectively. Intuitively, we expect that $\boldsymbol{V}^*(\epsilon, n, N)$ is a better estimate if $[f(\bar{X}_T^\epsilon)]_j - \boldsymbol{V}$ and $[f(\bar{X}_T^{[0]})]_j - \boldsymbol{E}[f(X_T^{[0]}(0, x))]$ take close values for each independent copy j since they are canceled with each other in each j of our estimator $\boldsymbol{V}^*(\epsilon, n, N)$. We can easily notice it by observing that the error of $\boldsymbol{V}^*(\epsilon, n, N)$ is given by the sample average of the difference between deviations of $[f(\bar{X}_T^\epsilon)]_j$ and $[f(\bar{X}_T^{[0]})]_j$ from their respective true values:

$$\begin{aligned}\boldsymbol{V}^*(\epsilon, n, N) - \boldsymbol{V} = \frac{1}{N}\sum_{j=1}^{N}[\{f(\bar{X}_T^\epsilon) - \boldsymbol{E}[f(X_T^\epsilon(0,x))]\} \\ - \{f(\bar{X}_T^{[0]}) - \boldsymbol{E}[f(X_T^{[0]}(0,x))]\}]_j.\end{aligned}$$

Our main objective is to state this intuition more rigorously. We shall start with a known error bound of the naive estimator $\boldsymbol{V}(\epsilon, n, N)$:

THEOREM 1. *Suppose that* $f \in C_\uparrow^\infty(\boldsymbol{R}^D; \boldsymbol{R})$. *Then*:

(i) *For the bias* $\mathbf{Bias}[\boldsymbol{V}(\epsilon, n, N)]$ *of* $\boldsymbol{V}(\epsilon, n, N)$,

$$\mathbf{Bias}[\boldsymbol{V}(\epsilon, n, N)] = \boldsymbol{E}[\boldsymbol{V}(\epsilon, n, N)] - \boldsymbol{V} = O\left(\frac{1}{n}\right).$$

(ii) *For the variance* $\mathbf{Var}[\boldsymbol{V}(\epsilon, n, N)]$ *of* $\boldsymbol{V}(\epsilon, n, N)$,

$$\mathbf{Var}[\boldsymbol{V}(\epsilon, n, N)] = \frac{1}{N}\mathbf{Var}[f(\bar{X}_T^\epsilon)] = O\left(\frac{1}{N}\right).$$

(iii) *For the mean-square-error* $\mathbf{MSE}[\boldsymbol{V}(\epsilon, n, N)] = \boldsymbol{E}[(\boldsymbol{V}(\epsilon, n, N) - \boldsymbol{V})^2]$,

$$\mathbf{MSE}[\boldsymbol{V}(\epsilon, n, N)] = O\left(\frac{1}{n^2} + \frac{1}{N}\right).$$

Theorem 1 is not a result we really want to show in this article. Presenting it here is just for comparison with our main results presented below. Since we will need the same procedure at the beginning of the proof of our main results, it is convenient to recall the proof of Theorem 1 in Section 5.1.

For our modified estimator $\boldsymbol{V}^*(\epsilon, n, N)$, we obtain a better error bound.

THEOREM 2. *Suppose that* $f \in C_\uparrow^\infty(\boldsymbol{R}^D; \boldsymbol{R})$. *Then*:

(i) *For the bias of* $\boldsymbol{V}^*(\epsilon, n, N)$, *it holds that*

$$\mathbf{Bias}[\boldsymbol{V}^*(\epsilon, n, N)] = O\left(\frac{\epsilon}{n}\right).$$

(ii) *For the variance of* $\boldsymbol{V}^*(\epsilon, n, N)$,

$$\mathbf{Var}[\boldsymbol{V}^*(\epsilon, n, N)] = \frac{1}{N}\mathbf{Var}[f(\bar{X}_T^\epsilon) - f(\bar{X}_T^0)] = O\left(\frac{\epsilon^2}{N}\right).$$

(iii) *The mean-square-error*

$$\mathbf{MSE}[\boldsymbol{V}^*(\epsilon, n, N)] = O\left(\epsilon^2\left(\frac{1}{n^2} + \frac{1}{N}\right)\right).$$

PROOF. See Section 5.2. □

Remark 1. We put the condition $f \in C_{\uparrow}^{\infty}(\boldsymbol{R}^D;\boldsymbol{R})$ for simplicity. This can be relaxed to a certain extent such as $f \in C_{\uparrow}^{k}(\boldsymbol{R}^D;\boldsymbol{R})$ for some positive k.

Remark 2. Though it is not so rigorous since $\boldsymbol{V}^*(\epsilon,n,N)$ is random, we may roughly regard $\boldsymbol{V}^*(\epsilon,n,N)$ approximating $\boldsymbol{V}$ with the same order of precision as the expansion of $\boldsymbol{V}$ up to the ϵ-order if $n \geq O(\epsilon^{-1})$ and $N \geq O(\epsilon^{-2})$.

Comparing $\boldsymbol{V}^*(\epsilon,n,N)$ with $\boldsymbol{V}(\epsilon,n,N)$ in mean-square-error, we see that

$$\begin{aligned}
&\mathbf{MSE}[\,\boldsymbol{V}(\epsilon,n,N)] - \mathbf{MSE}[\,\boldsymbol{V}^*(\epsilon,n,N)] \\
&\quad \geq \frac{1}{N}\{\mathbf{Var}[f(\bar{X}_T^{\epsilon})] - \mathbf{Var}[f(\bar{X}_T^{\epsilon}) - f(\bar{X}_T^{0})]\} - \theta_1(\epsilon,n) \\
&\quad \geq \frac{1}{N}\mathbf{Var}[f(\bar{X}_T^{\epsilon})] - \theta_2(\epsilon,n,N),
\end{aligned}$$

where

$$0 \leq \theta_1(\epsilon,n) = O\left(\left(\frac{\epsilon}{n}\right)^2\right)$$

and

$$0 \leq \theta_2(\epsilon,n,N) = O\left(\epsilon^2\left(\frac{1}{n^2} + \frac{1}{N}\right)\right).$$

We then expect that $\theta_2(\epsilon,n,N)$ is smaller than $N^{-1}\mathbf{Var}[f(\bar{X}_T^{\epsilon})]$, and hence that MSE of $\boldsymbol{V}^*(\epsilon,n,N)$ is smaller than MSE of $\boldsymbol{V}(\epsilon,n,N)$.

2.2. Non smooth case

If f is not smooth, in particular, if f is a Borel measurable function of at most polynomial growth, we can still obtain the similar results as in the smooth case under appropriate additional assumptions.

We consider a stochastic approximation to $\boldsymbol{V} := \boldsymbol{E}[f(X_T^{\epsilon}(0,x))]$. An estimator may be obtained by a naive Monte Carlo simulation. However, Malliavin calculus is used in order to give an error bound, because of non smoothness of f. To apply Malliavin calculus effectively, we will take a modified Euler-Maruyama scheme similar to Kohatsu-Higa (1997). That is, we compute

$$\boldsymbol{V}(\epsilon,n,N) = \frac{1}{N}\sum_{j=1}^{N}\left[f\left(\bar{X}_T^{\epsilon} + \frac{1}{n}\hat{w}_T\right)\right]_j, \tag{2.5}$$

instead of $\boldsymbol{V}(\epsilon,n,N)$ given in (2.2), where $\{\hat{w}_t; t \in [0,T]\}$ is a Wiener process independent of X^{ϵ}. Bally and Talay (1995) also applied the Malliavin calculus to derive an error bound when f is not smooth. We will use the Malliavin calculus over the product space of two Winer spaces equipped with the product measure $P^w \otimes P^{\hat{w}}$.

Similarly, our new estimator (2.4) is modified as follows:

$$\begin{aligned}(2.6)\qquad \mathbf{V}^*(\epsilon,n,N) &= \boldsymbol{E}[f(X_T^{[0]}(0,x))] \\ &\quad + \frac{1}{N}\sum_{j=1}^{N}\left[f\left(\bar{X}_T^{\epsilon}+\frac{1}{n}\hat{w}_T\right)-f\left(\bar{X}_T^{[0]}+\frac{1}{n}\hat{w}_T\right)\right]_j .\end{aligned}$$

To justify this scheme, we first make the following assumption:

[A1] For every $p>1$,

$$\sup_{\epsilon}\boldsymbol{E}[|\sigma_{X_T^{\epsilon}(0,x)}|^{-p}]<\infty,$$

where $\sigma_{X_T^{\epsilon}(0,x)}$ denotes the Malliavin covariance of $X_T^{\epsilon}(0,x)$.

It is sometimes difficult to check Condition [A1]. Then in stead of [A1], we may put the following condition [A2] which is practically more convenient.

[A2] For every $p>1$,

$$\boldsymbol{E}[|\sigma_{X_T^{[0]}(0,x)}|^{-p}]<\infty,$$

where $\sigma_{X_T^{[0]}(0,x)}$ denotes the Malliavin covariance of $X_T^{[0]}(0,x)$.

We can obtain similar results in the non smooth case corresponding to Theorems 1 and 2 in the smooth case. In particular, we have the following result similar to Theorem 2 under Condition [A1] or Condition [A2].

THEOREM 3. *Suppose that f is a Borel measurable function of at most polynomial growth. Suppose that for some positive constant ω, $\epsilon=o(n^{-\omega})$ as $n\to\infty$. Then under the Condition* [A1] *or* [A2], *the following properties hold*:

(i) *The bias of $\boldsymbol{V}^*(\epsilon,n,N)$ satisfies*

$$\mathbf{Bias}[\boldsymbol{V}^*(\epsilon,n,N)]=O\left(\frac{\epsilon}{n}\right).$$

(ii) *The variance of $\boldsymbol{V}^*(\epsilon,n,N)$ admits*

$$\mathbf{Var}[\boldsymbol{V}^*(\epsilon,n,N)]=\frac{1}{N}\mathbf{Var}[f(\bar{X}_T^{\epsilon})-f(\bar{X}_T^{0})]=O\left(\frac{\epsilon^2}{N}\right).$$

(iii) *The mean-square-error satisfies*

$$\mathbf{MSE}[\boldsymbol{V}^*(\epsilon,n,N)]=O\left(\epsilon^2\left(\frac{1}{n^2}+\frac{1}{N}\right)\right).$$

PROOF. See Section 6. □

3. Examples

In this section, we take two examples from finance to illustrate our method.

3.1. Example 1: Computation of optimal portfolio for investment

The first example is computation of the *Market Price of Risk* component of an optimal portfolio in multiperiod setting. (Hereafter, we call the component *MPR-hedge* following a convention in finance.) We note that this example belongs to *smooth case* in Section 2.1. We start with basic set up of the financial market.

Let $(\Omega, \mathcal{F}, P)$ probability space and $T \in (0, \infty)$ denotes some fixed time horizon of the economy. $w = \{(w^1(t), \ldots, w^r(t))^*; t \in [0, T]\}$ is $\boldsymbol{R}^r$-valued Brownian motion defined on $(\Omega, \mathcal{F}, P)$ and $\{\mathcal{F}_t\}$, $0 \le t \le T$ stands for P-augmentation of the natural filtration, $\mathcal{F}_t^w = \sigma(w(s); 0 \le s \le t)$. Here, we use the notation of x^* as the transpose of x.

For $t \in [0, T]$, the price processes of risky assets and a locally riskless asset are described as follows.

$$
\begin{aligned}
&(3.1)\quad dS_i = S_i(t)\left[b_i(t)dt + \sum_{j=1}^r \sigma_{ij}(t)dw_j(t)\right]; \qquad S_i(0) = s_i, \qquad i = 1, \ldots, r \\
&\qquad\quad dS_0 = \gamma(t)S_0(t)dt; \qquad S_0(0) = 1
\end{aligned}
$$

where $\gamma(t)$, $b_i(t)$, and $\sigma_{ij}(t)$ are progressively measurable with respect to $\{\mathcal{F}_t\}$. $b_i(t)$ and $\sigma_{ij}(t)$ satisfy the integrability conditions:

$$\int_0^T \{|b(t)| + |\sigma(t)|^2\}dt < \infty$$

where $|b(t)| := (\sum_{i=1}^r |b_i(t)|^2)^{1/2}$ and $|\sigma(t)|^2 := \sum_{i,j=1}^r |\sigma_{ij}(t)|^2$. $\sigma(t)$ is assumed to be non-singular Lebesgue-almost-every $t \in [0, T]$, a.s. Then, R^r-valued process $\theta(t)$, $t \in [0, T]$, the market price of risk process is well-defined as $\theta(t) := \sigma^{-1}(t)[b(t) - \gamma(t)\vec{1}]$. We further assume that $\gamma(t)$ and $\theta_i(t)$, $i = 1, 2, \ldots, r$ are bounded.

Next, we illustrate the problem of a (small) investor's optimal portfolio for investment in the multiperiod setting. Given the financial market described above, an investor's wealth $W(t)$ at time $t \in [0, T]$ is described as

$$dW(t) = [\gamma(t)W(t) - c(t)]dt + \pi(t)^*[(b(t) - \gamma(t)\mathbf{1})dt + \sigma(t)dw(t)];$$

where $W(0) = W > 0$ is the initial capital (wealth), $c(t)$ denotes the consumption rate and $\pi(t) = \{\pi_i(t)\}_{i=1,\ldots,r}^*$ denotes the portfolio, which satisfy the integrability condition;

$$\int_0^T \{|\pi(t)|^2 + c(t)\}dt < \infty \qquad \text{a.s.}$$

Let $\mathcal{A}(W)$ denote the set of stochastic processes (π, c) which generate $W(t) \ge 0$ for all $t \in [0, T]$ given $W(0) = W$. If $(\pi, c) \in \mathcal{A}(W)$, (π, c) is called admissible for W. Then, the problem of an investor's optimal portfolio for investment is formulated as follows;

$$\sup_{(\pi,c)\in\mathcal{A}(W)} E[U(W(T))] \tag{3.2}$$

where $E[\cdot]$ denotes the expectation operator under P, and U represents a utility function such that

$$\begin{aligned}(3.3)\qquad & U:(0,\infty)\to \boldsymbol{R},\\ & \text{a strictly increasing, strictly concave function of class } \boldsymbol{C}^2\\ & \text{with } U(0+)\equiv \lim_{c\downarrow 0}U(c)\in[-\infty,\infty),\ U'(0+)\equiv \lim_{c\downarrow 0}U'(c)=\infty\\ & \text{and } U'(\infty)\equiv \lim_{c\to\infty}U'(c)=0.\end{aligned}$$

From now on, we will concentrate on a *Markovian model*. We consider a Wiener space on $[t,T]$ for some fixed $t\in[0,T]$ and assume that all random variables will be defined on it. Let X_u^ϵ be a D-dimensional diffusion process defined by the stochastic differential equation:

$$dX_u^\epsilon = V_0(X_u^\epsilon,\epsilon)du + V(X_u^\epsilon,\epsilon)dw_u, \qquad X_t^\epsilon = x \tag{3.4}$$

for $u\in[t,T]$ where $\epsilon\in(0,1]$, $V_0\in C_b^\infty(\boldsymbol{R}^D\times(0,1];\boldsymbol{R}^D)$, and $V=(V_\beta)_{\beta=1}^r\in C_b^\infty(\boldsymbol{R}^D\times(0,1];\boldsymbol{R}^D\otimes\boldsymbol{R}^r)$. Here $C_b^\infty(R^d\times(0,1];E)$ denotes a class of smooth mappings $f:\boldsymbol{R}^D\times(0,1]\to E$ whose derivatives $\partial_x^{\boldsymbol{n}}\partial_\epsilon^m f(x,\epsilon)$ are all bounded for $\boldsymbol{n}\in\boldsymbol{Z}_+^d$ such that $|\boldsymbol{n}|\geq 1$ and $m\in\boldsymbol{Z}_+$. Note that time-dependent-coefficient diffusion processes are included in the above equation if we enlarge the process to a higer-dimensional one. Let $Y_{t,u}^\epsilon$ be a unique solution of the $D\times D$-matrix valued stochastic differential equation:

$$\begin{cases} dY_{t,u}^\epsilon = \sum_{\alpha=0}^r \partial_x V_\alpha(X_u^\epsilon,\epsilon)Y_{t,u}^\epsilon dw_u^\alpha \\ Y_{t,t}^\epsilon = \underline{I}\end{cases}. \tag{3.5}$$

We further assume the bounded processes $\gamma(u)$ (short rate) and $\theta(u)$ (the market price of risk) to be $\gamma(u)=\gamma(X_u^\epsilon)$ and $\theta(u)=\theta(X_u^\epsilon)$ where $\gamma\in C_b^\infty(\boldsymbol{R}^D;\boldsymbol{R}_+)$ and $\theta\in C_b^\infty(\boldsymbol{R}^D;\boldsymbol{R}^r)$. The case that $b(u)=b(X_u^\epsilon)$ and $\sigma(u)=\sigma(X_u^\epsilon)$ is an example in this formulation. Next, we suppose that a utility function is so called a power function;

$$U(x)=\frac{x^\delta}{\delta}, \qquad x\in(0,\infty), \quad \delta<1, \quad \delta\neq 0.$$

Then, due to Takahashi and Yoshida (2004), the optimal proportions of risky assets in given wealth W at time t, are provided by

$$\begin{aligned}\pi^*(t)/W = {} & \frac{1}{(1-\delta)}\theta(x)^*\sigma^{-1}(x) + \frac{\delta}{(1-\delta)}\frac{1}{\boldsymbol{E}[(H_{0,t,T})^{(-\delta/(1-\delta))}]}\\ & \times \boldsymbol{E}\Bigg[(H_{0,t,T})^{(-\delta/(1-\delta))}\Bigg(\int_t^T \partial\gamma(X_u^\epsilon)Y_{t,u}^\epsilon du\\ & \qquad + \sum_{\alpha=1}^r \int_t^T \partial\theta_\alpha(X_u^\epsilon)Y_{t,u}^\epsilon dw^\alpha(u)\\ & \qquad + \sum_{\alpha=1}^r \int_t^T \theta_\alpha(X_u^\epsilon)\partial\theta_\alpha(X_u^\epsilon)Y_{t,u}^\epsilon du\Bigg)\Bigg]V(x,\epsilon)\sigma^{-1}(x),\end{aligned} \tag{3.6}$$

where the state density process, $H_{0,t,T}$ is defined by

$$H_{0,t,T} \equiv \exp\left(-\int_t^T \theta(X_u^\epsilon)^* dw(u) - \frac{1}{2}\int_t^T |\theta(X_u^\epsilon)|^2 du - \int_t^T \gamma(X_u^\epsilon) du\right).$$

Next, we define *the mean variance*, *the interest rate hedge* (IR-hedge) and *the market price of risk hedge* (MPR-hedge) components of the optimal portfolio for a power utility function:

$$\begin{aligned}
(3.7)\quad \text{mean variance} &\equiv \frac{1}{(1-\delta)}\theta(x)^*\sigma^{-1}(x)\\
\text{IR-hedge} &\equiv \frac{\delta}{(1-\delta)}\frac{1}{\boldsymbol{E}[(H_{0,t,T})^{(-\delta/(1-\delta))}]}\\
&\quad \times \boldsymbol{E}\left[(H_{0,t,T})^{(-\delta/(1-\delta))}\int_t^T \partial\gamma(X_u^\epsilon)Y_{t,u}^\epsilon du\right]V(x,\epsilon)\sigma^{-1}(x)\\
\text{MPR-hedge} &\equiv \frac{\delta}{(1-\delta)}\frac{1}{\boldsymbol{E}[(H_{0,t,T})^{(-\delta/(1-\delta))}]}\\
&\quad \times \boldsymbol{E}\Bigg[(H_{0,t,T})^{(-\delta/(1-\delta))}\Bigg(\sum_{\alpha=1}^r \int_t^T \partial\theta_\alpha(X_u^\epsilon)Y_{t,u}^\epsilon dw^\alpha(u)\\
&\quad + \sum_{\alpha=1}^r \int_t^T \theta_\alpha(X_u^\epsilon)\partial\theta_\alpha(X_u^\epsilon)Y_{t,u}^\epsilon du\Bigg)\Bigg]V(x,\epsilon)\sigma^{-1}(x).
\end{aligned}$$

Then, we put a main assumption on the asymptotic method:

[**A3**] the deterministic limit condition: $V(\cdot,0) \equiv 0$.

Under the assumption [**A3**], each component of the optimal portfolio for a power utility function in the asymptotic method up to ϵ order is given due to Takahashi and Yoshida (2004):

$$\begin{aligned}
(3.8)\quad \text{mean variance} &\equiv \frac{1}{(1-\delta)}\theta^*(x)\sigma^{-1}(x)\\
\text{IR-hedge} &\equiv \epsilon\frac{\delta}{(1-\delta)}\left(\int_t^T \partial\gamma^{[0]}(u)Y_{t,u}du\right)\partial_\epsilon V(x,0)\sigma^{-1}(x)\\
\text{MPR-hedge}&\\
&\equiv \epsilon\frac{\delta}{(1-\delta)^2}\left(\sum_{\alpha=1}^r \int_t^T \theta_\alpha^{[0]}(u)\partial\theta_\alpha^{[0]}(u)Y_{t,u}du\right)\partial_\epsilon V(x,0)\sigma^{-1}(x).
\end{aligned}$$

From now on, we illustrate our scheme by using MPR-hedge component (3.7). Similar method can be applied to IR-hedge component. (Note that mean variance component is analytically obtained.)

Numerical computation of MPR-hedge

In computation of MPR-hedge, we first consider a naive estimator by Monte Carlo. Hereafter we set $t = 0$. A Markovian system of SDEs used in Monte Carlo simulation is given as follows:

$$\tag{3.9} \begin{cases} dX_u^\epsilon = V_0(X_u^\epsilon, \epsilon)du + V(X_u^\epsilon, \epsilon)dw_u, \qquad X_t^\epsilon = x \\ dY_{t,u}^\epsilon = \sum_{\alpha=0}^r \partial_x V_\alpha(X_u^\epsilon, \epsilon)Y_{t,u}^\epsilon dw_u^\alpha, \qquad Y_{t,t}^\epsilon = \underline{I} \\ dh_{0,t,u}^\epsilon = h_{0,t,u}^\epsilon[\{(\frac{\delta}{1-\delta})\gamma(X_u^\epsilon) + \frac{\delta}{2(1-\delta)^2}|\theta(X_u^\epsilon)|^2\}du \\ \qquad + (\frac{\delta}{1-\delta})\theta(X_u^\epsilon)^* dw(u)], \qquad h_{0,t,t}^\epsilon = 1 \\ d\eta_u^\epsilon = \sum_{\alpha=1}^r \theta_\alpha(X_u^\epsilon)\partial\theta_\alpha(X_u^\epsilon)Y_{t,u}^\epsilon du \\ \qquad + \sum_{\alpha=1}^r \partial\theta_\alpha(X_u^\epsilon)Y_{t,u}^\epsilon dw^\alpha(u), \qquad \eta_t^\epsilon = 0. \end{cases}$$

We note that above system of equations (3.9) corresponds to the equation (2.1) in Section 2. Then, the estimator based on naive Monte Carlo simulation (2.2) for the denominator of MPR-hedge (3.7);

$$\tag{3.10} \boldsymbol{E}[(H_{0,t,T})^{(-\delta/(1-\delta))}] = \boldsymbol{E}[h_{0,t,T}^\epsilon]$$

may be expressed as

$$\tag{3.11} \frac{1}{N}\sum_{j=1}^N [\bar{h}_{0,t,T}^\epsilon]_j.$$

Similarly, the estimator for the numerator of MPR-hedge (3.7);

$$\tag{3.12} \boldsymbol{E}\Bigg[(H_{0,t,T})^{(-\delta/(1-\delta))}\Bigg(\sum_{\alpha=1}^r \int_t^T \partial\theta_\alpha(X_u^\epsilon)Y_{t,u}^\epsilon dw^\alpha(u) + \sum_{\alpha=1}^r \int_t^T \theta_\alpha(X_u^\epsilon)\partial\theta_\alpha(X_u^\epsilon)Y_{t,u}^\epsilon du\Bigg)\Bigg]$$

may be expressed as

$$\tag{3.13} \frac{1}{N}\sum_{j=1}^N [\bar{h}_{0,t,T}^\epsilon \times \bar{\eta}_T^\epsilon]_j.$$

Next, we consider modified estimators for (3.10) and (3.12) in the following. First, we note that

$$(H_{0,t,T}^{[0]})^{(-\delta/(1-\delta))} = h_{0,t,T}^{[0]} = C \times \xi_T^{[0]}$$

where

$$\xi_T^{[0]} = e^{-1/2(\delta/(1-\delta))^2 \int_t^T |\theta^{[0]}(u)|^2 du + (\delta/(1-\delta))\int_t^T \theta^{[0]}(u)dw(u)}$$

and

$$C \equiv \exp\left\{\left(\frac{\delta}{1-\delta}\right)\int_t^T \gamma^{[0]}(u)du + \frac{\delta}{2(1-\delta)^2}\int_t^T |\theta^{[0]}(u)|^2 du\right\}.$$

A modified estimator for the denominator (3.10) is given by

$$\boldsymbol{E}[h^{[0]}_{0,t,T}] + \frac{1}{N}\sum_{j=1}^{N}\{\{[\bar{h}^{\epsilon}_{0,t,T} - \bar{h}^{[0]}_{0,t,T}]_j\} \tag{3.14}$$

where

$$\boldsymbol{E}[h^{[0]}_{0,t,T}] = C,$$

because clearly

$$\boldsymbol{E}[\xi^{[0]}_T] = 1.$$

Further, $\bar{h}^{[0]}_{0,t,u}$ denotes the Euler-Maruyama scheme of $h^{[0]}_{0,t,u}$:

$$\begin{cases} dh^{[0]}_{0,t,u} = h^{[0]}_{0,t,u}[\{(\frac{\delta}{1-\delta})\gamma^{[0]}_u + \frac{\delta}{2(1-\delta)^2}|\theta^{[0]}_u|^2\}du + (\frac{\delta}{1-\delta})\theta^{[0],*}_u dw(u)], \\ h^{[0]}_{0,t,t} = 1. \end{cases} \tag{3.15}$$

In the similar way, a modified estimator for the numerator (3.12) is given by

$$\boldsymbol{E}[h^{[0]}_{0,t,u}\eta^{[0]}_T] + \frac{1}{N}\sum_{j=1}^{N}\{[\bar{h}^{\epsilon}_{0,t,T} \times \bar{\eta}^{\epsilon}_T - \bar{h}^{[0]}_T \times \bar{\eta}^{[0]}_T]_j\} \tag{3.16}$$

where

$$\boldsymbol{E}[h^{[0]}_{0,t,u}\eta^{[0]}_T] = C \times \left(\frac{1}{1-\delta}\right)\left[\sum_{\alpha=1}^{r}\int_t^T \theta^{[0]}_\alpha(u)\partial\theta^{[0]}_\alpha(u)Y_{t,u}du\right],$$

and $\bar{\eta}^{[0]}_u$ denotes the Euler-Maruyama scheme of $\eta^{[0]}_u$:

$$\begin{aligned} d\eta^{[0]}_u &= \sum_{\alpha=1}^{r}\theta_\alpha(X^{[0]}_u)\partial\theta_\alpha(X^{[0]}_u)Y^{[0]}_{t,u}du \\ &\quad + \sum_{\alpha=1}^{r}\partial\theta_\alpha(X^{[0]}_u)Y^{[0]}_{t,u}dw^\alpha(u), \qquad \eta^{[0]}_t = 0. \end{aligned} \tag{3.17}$$

3.2. Example 2: Pricing of an average call option

The second example is pricing an average call option which belongs to *non smooth case* in Section 2.2. Given filtered probability space satisfying usual conditions $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{0\le t\le T})$ with one-dimensional Brownian motion $\{w_t; 0 \le t \le T\}$, where P represents a so called equivalent Martingale measure in finance. The underlying asset price process S_t, $0 \le t \le T$ is assumed to follow a one-dimensional diffusion process:

$$dS^\epsilon_t = \gamma S^\epsilon_t dt + \epsilon\sigma(S^\epsilon_t, t)dw_t, \qquad S^\epsilon_0 = S_0 (> 0) \tag{3.18}$$

where $\epsilon \in (0,1]$, $\sigma \in C_b^\infty(\boldsymbol{R}_+ \times [0,T]; \boldsymbol{R}_+)$, γ is a positive constant. The payoff of an average call option with strike price $K(>0)$ and with the maturity T is given by

$$V(T) = \left(\frac{1}{T}Z_T^\epsilon - K\right)_+, \tag{3.19}$$

where $(x)_+ = \max(x,0)$. Then, to obtain the price of an average call option at $t=0$, we evaluate

$$V = e^{-\gamma T}\boldsymbol{E}\left[\left(\frac{1}{T}Z_T^\epsilon - K\right)_+\right]$$

given that

$$\begin{cases} dS_t^\epsilon = \gamma S_t^\epsilon dt + \epsilon\sigma(S_t^\epsilon, t)dw_t, & S_0^\epsilon = S_0(>0) \\ dZ_t^\epsilon = S_t^\epsilon dt, & Z_0^\epsilon = 0. \end{cases} \tag{3.20}$$

(For details of *average options*, see Kunitomo and Takahashi (1992) and He and Takahashi (2000) for instance.) It is re-expressed by

$$V = e^{-\gamma T}\epsilon\boldsymbol{E}\left[\left(\frac{1}{T}X_{2T}^\epsilon + y\right)_+\right] \tag{3.21}$$

where

$$\begin{aligned} X_{1t}^\epsilon &\equiv \frac{S_t^\epsilon - S_t^{[0]}}{\epsilon}, \\ X_{2t}^\epsilon &\equiv \frac{Z_t^\epsilon - Z_t^{[0]}}{\epsilon}, \\ y &\equiv \frac{\frac{1}{T}Z_T^{[0]} - K}{\epsilon}, \\ S_t^{[0]} &= e^{\gamma t}S_0, \\ Z_t^{[0]} &= \frac{S_0}{\gamma}(e^{\gamma t} - 1). \end{aligned}$$

We also notice that X_{1t}^ϵ and X_{2t}^ϵ satisfy SDEs:

$$\begin{cases} dX_{1t}^\epsilon = \gamma X_{1t}^\epsilon dt + \sigma(\epsilon X_{1t}^\epsilon + S_t^{[0]}, t)dw_t, & X_{10}^\epsilon = 0 \\ dX_{2t}^\epsilon = X_{1t}^\epsilon dt, & X_{20}^\epsilon = 0. \end{cases} \tag{3.22}$$

Next, we assume the condition:

$$\Sigma \equiv \int_0^T \frac{1}{T^2}\left[\frac{e^{(T-u)} - 1}{\gamma}\right]^2 \sigma^2(S_u^{[0]}, u)du > 0. \tag{3.23}$$

Under condition (3.23), The asymptotic expansion of V upto ϵ-order is obtained by

$$V = e^{-\gamma T}\epsilon\left(y\Phi\left(\frac{y}{\sqrt{\Sigma}}\right) + \Sigma\frac{1}{\sqrt{2\pi\Sigma}}\exp\left(\frac{-y^2}{2\Sigma}\right)\right) + o(\epsilon).$$

Then, a modified estimator for (3.21) is given by

$$(3.24)\quad e^{-\gamma T}\boldsymbol{E}\left[\left(\frac{1}{T}X_{2T}^{[0]}+y\right)_+\right]+\frac{1}{N}\sum_{j=1}^{N}\Bigg\{\Bigg[e^{-\gamma T}\left(\frac{1}{T}\bar{X}_{2T}^{\epsilon}+y+\frac{1}{n}\hat{w}_T\right)_+ - e^{-\gamma T}\left(\frac{1}{T}\bar{X}_{2T}^{[0]}+y+\frac{1}{n}\hat{w}_T\right)_+\Bigg]_j\Bigg\}$$

where

$$(3.25)\quad e^{-\gamma T}\boldsymbol{E}\left[\left(\frac{1}{T}X_{2T}^{[0]}+y\right)_+\right] = e^{-\gamma T}\left\{y\Phi\left(\frac{y}{\sqrt{\Sigma}}\right)+\Sigma\frac{1}{\sqrt{2\pi\Sigma}}\exp\left(\frac{-y^2}{2\Sigma}\right)\right\}.$$

$\bar{X}_{it}^{[0]}$, $i=1,2$ denote the Euler-Maruyama scheme of $X_{it}^{[0]}$, $i=1,2$, which is given by

$$(3.26)\quad \begin{cases} dX_{1t}^{[0]} = \gamma X_{1t}^{[0]}dt+\sigma(S_t^{[0]},t)dw_t, & X_{10}^{[0]}=0 \\ dX_{2t}^{[0]} = X_{1t}^{[0]}dt, & X_{20}^{[0]}=0. \end{cases}$$

Here, $\Phi(x)$ denotes the standard normal distribution evaluated at x.

4. Numerical examples

4.1. Example 1: MPR-hedge

We take a numerical example in Takahashi and Yoshida (2004) where they computed the MPR-hedge component based on the analytic approximation (3.8). We will demonstrate our new scheme is effective in increasing the efficiency of Monte Carlo simulations as well as to aciheve further numerical accuracy for the case when the approximation error is relatively large. We start with brief explanation of the setup. (See Takahashi and Yoshida (2004) for the details.)

We suppose that $D=2$, that is $X_u^\epsilon=(X_u^{\epsilon(1)},X_u^{\epsilon(2)})^*$ and that they satisfy the following stochastic differential equations:

$$(4.1)\quad \begin{cases} dX_u^{\epsilon(1)} = \kappa_1(\bar{X}^{\epsilon(1)}-X_u^{\epsilon(1)})du-\epsilon(X_u^{\epsilon(1)})^{\frac{1}{2}}dw_u; & X_0^{\epsilon(1)}=\gamma_0 \\ dX_u^{\epsilon(2)} = \kappa_2(\bar{X}^{\epsilon(2)}-X_u^{\epsilon(2)})du+\epsilon\sigma_2 dw_u; & X_0^{\epsilon(2)}=\theta_0 \end{cases}$$

where w denotes one-dimensional Brownian motion.

Remark 3. The volatility function of $X^{\epsilon(1)}$ is not smooth at the origin and we need to use a smoothed version of the square root process at the origin in our framework. However, we can show that smoothing does not make significant differences and the effects are negligible for the *small disturbance asymptotic theory*. This is also true for Example 2 in the next subsection. See Section 7 for a rigorous argument on this point.

We also suppose that there exist one risky asset and a locally riskless asset, and assume that $\theta_u = X_u^{\epsilon(2)}$ and γ_u is a smooth modification of $\min\{X_u^{\epsilon(1)}, M\}$ where M is a positive large number. Then, the dynamics of both assets are described by

$$\begin{cases} dS_u^\epsilon = S_u^\epsilon(\gamma(X_u^{\epsilon(1)}) + \sigma\theta(X_u^{\epsilon(2)}))du + S_u^\epsilon \sigma dw_u, & S^\epsilon(0) = s \\ dS_{0u}^\epsilon = S_{0u}^\epsilon \gamma(X_u^{\epsilon(1)})du, & S_0^\epsilon(0) = 1. \end{cases} \tag{4.2}$$

Further, we set the values of the parameters for X_u^ϵ following Detemple *et al.* (2003), which were obtained by statistcal estimation; $\kappa_1 = 0.0824$, $\gamma_0 = \bar{X}^{\epsilon(1)} = 0.06$, $\epsilon = 0.03637$, $\kappa_2 = 0.6950$, $\bar{X}^{\epsilon(2)} = 0.0871$, $\sigma_2 = 0.21/0.03637$, $\theta_0 = 0.1$, $\sigma = 0.2$.

The benchmark value of each component in the optimal portfolios is obtained by naive Monte Carlo simulations based on the Euler-Maruyama approximation; the number of time steps $\boldsymbol{n}$ is 365 and the number of trials N is 1,000,000 in each Monte Carlo simulation.

The percentage-shares in total wealth of Mean variance, IR-hedge, MPR-hedge and the total demand which are sum of those three components are listed in Tables 1–4; the results for the asymptotic method are listed in Tables 1 and 3 while the results for the Monte Carlo simulation are listed in Tables 2 and 4. In addition, Tables 1 and 2 show the results for *investment horizons* $T = 1, 2, 3, 4, 5$ when *the Arrow-Pratt measure of relative risk aversion* $R(\equiv 1 - \delta)$ is fixed at 2, and Tables 3 and 4 show the results for $R = 0.5, 1, 1.5, 4, 5$ when $T = 1$.

We remark that *total demand* refers to the demand for risky assets and this may not be 100% because the remaining shares (100%-*total demand*) are invested into riskless assets. We also note that it may exceed 100% since selling (borrowing) riskless assets is admitted. We observe that the results of the asymptotic method and of Monte carlo simulations are very close for the IR-hedge while

Table 1. Asymptotic expansion ($R = 2.0$).

T (Investment horizon)	1	2	3	4	5
Total demand	25.31	26.41	27.80	29.26	30.70
Mean variance	25.00	25.00	25.00	25.00	25.00
IR-hedge	2.14	4.11	5.92	7.59	9.13
MPR-hedge	−1.83	−2.70	−3.12	−3.33	−3.43

Table 2. Monte Carlo simulation ($R = 2.0$).

T (Investment horizon)	1	2	3	4	5
Total demand	25.37	26.49	27.79	29.10	30.41
Mean variance	25.00	25.00	25.00	25.00	25.00
IR-hedge	2.14	4.12	5.95	7.63	9.19
MPR-hedge	−1.77	−2.63	−3.16	−3.53	−3.78

Table 3. Asymptotic expansion ($T = 1.0$).

$R(\equiv 1-\delta)$	0.5	1	1.5	4	5
Total demand	110.37	50	33.13	14.34	12.25
Mean variance	100.00	50.00	33.33	12.50	10.00
IR-hedge	−4.28	0	1.43	3.21	3.42
MPR-hedge	14.65	0	−1.63	−1.37	−1.17

Table 4. Monte Carlo simulation ($T = 1.0$).

$R(\equiv 1-\delta)$	0.5	1	1.5	4	5
Total demand	113.07	50.00	33.18	14.35	12.22
Mean variance	100.00	50.00	33.33	12.50	10.00
IR-hedge	−4.26	0.00	1.43	3.22	3.43
MPR-hedge	17.33	0.00	−1.58	−1.37	−1.22

there is some difference for the MPR-hedge, although small relative to the total demand. We also notice that the second order scheme gives substantial improvement compared with the first order scheme which is equivalent to the case that we ignore the MPR-hedge and IR-hedge components. (Note that the first orders of MPR-hedge and IR-hedge components are *zero*.)

To show that our new method to increase the efficiency of Monte Carlo simulations is effective, we take the case of MPR-hedge with $T = 1$, and $R = 0.5$, in which the diviation of the value based on the asymptotic method from the benchmark value is the largest. We follow the method illustrated in the previous section.

Figure 1 shows the comparison of the convergence between our modified estimator and a naive one for the MPR-hedge (3.7): *hybrid* denotes the modified estimator expressed as the equation (3.16) divided by (3.14), that is (3.16)/(3.14) while *mc* denotes the naive estimator expressed as the equation (3.13) divided by (3.11), that is (3.13)/(3.11). In Fig. 1, the horizontal axis is the number of trials N which varies from 1000 to 100,000, and the vertical axis is the errors (%) of *mc* and *hybrid* relative to their benchmark values. We observe that *hybrid* provides much faster convergence than *mc*. To examine our method more closely, we compare the covergence of three estimators for numerator of MPR-hedge; *num-hybrid* denotes the modified estimator, *num0-mc* denotes the estimator for $\epsilon = 0$ in (3.16) that is, $\frac{1}{N}\sum_{j=1}^{N}[\bar{h}_T^{[0]} \times \bar{\eta}_T^{[0]}]_j$, and *num-mc* denotes the naive estimator (3.13). Figure 2 clarifies that the errors of *num-mc* and *num0-mc* are canceled with each other, which results in the faster convergence of the modified estimator *num-hybrid*.

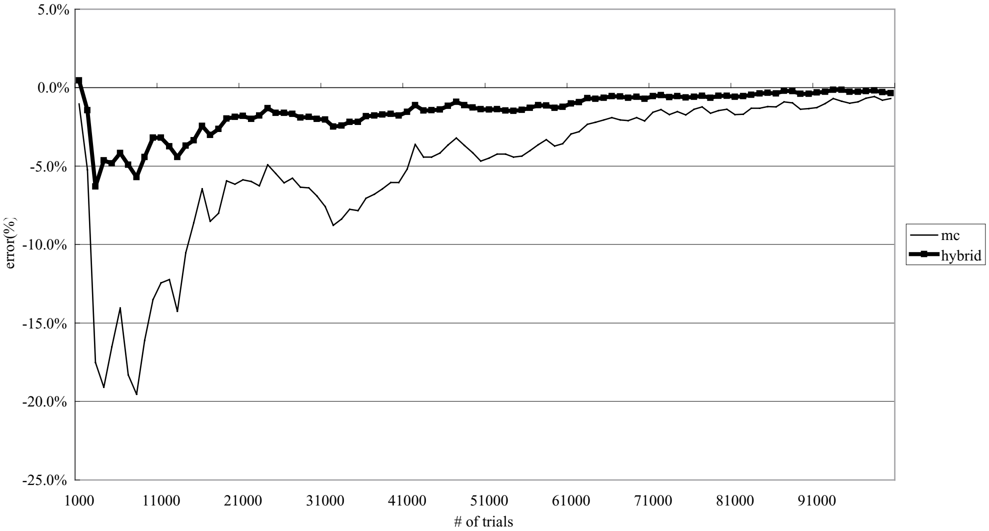

Figure 1. MPR-hedge convergence of Monte Carlo simulation.

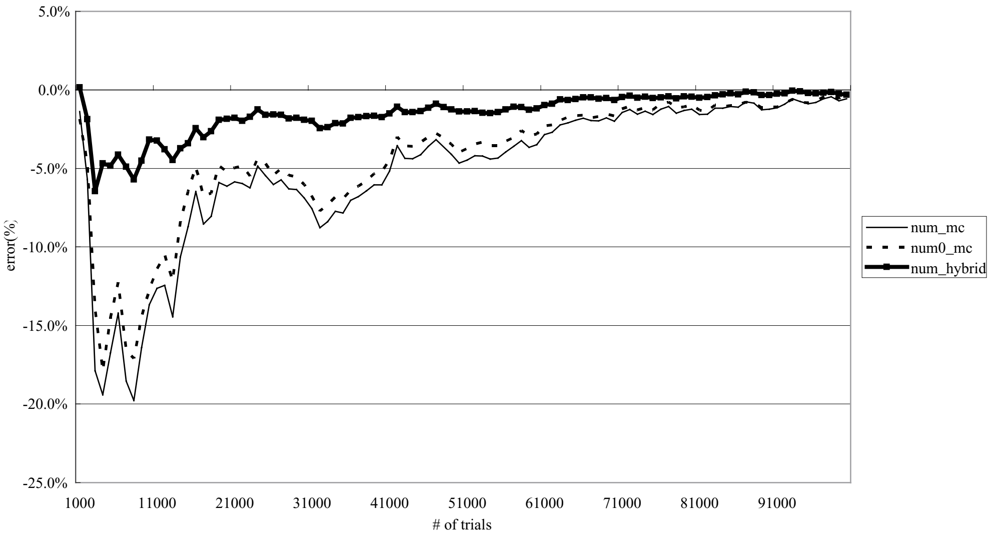

Figure 2. MPR-hedge (-Numerator-) convergence of Monte Carlo simulation.

4.2. Example 2: An average call option

On the second example, we take so called square-root process as the price process of the underlying asset:

$$\left\{\begin{array}{ll} dS_t^\epsilon = \gamma S_t^\epsilon dt + \epsilon\sqrt{S_t^\epsilon}dw_t, & S_0^\epsilon = S_0 \\ dZ_t^\epsilon = S_t^\epsilon dt, & Z_0^\epsilon = 0. \end{array}\right. \tag{4.3}$$

Then, the normalized price processes, X_{it}^ϵ, $i = 1, 2$ are expressed as

$$\left\{\begin{array}{ll} dX_{1t}^\epsilon = \gamma X_{1t}^\epsilon dt + \sqrt{\epsilon X_{1t}^\epsilon + e^{\gamma t}S_0}dw_t, & X_{10}^\epsilon = 0 \\ dX_{2t}^\epsilon = X_{1t}^\epsilon dt, & X_{20}^\epsilon = 0, \end{array}\right. \tag{4.4}$$

and Σ is given by

$$\Sigma = \frac{S_0}{\gamma^3 T^2}(e^{2\gamma T} - 2\gamma e^{\gamma T} - 1). \tag{4.5}$$

Finally, $X_{it}^{[0]}$, $i = 1, 2$ ($\epsilon = 0$) are obtained by

$$\left\{\begin{array}{ll} dX_{1t}^{[0]} = \gamma X_{1t}^{[0]} dt + e^{\frac{\gamma t}{2}}\sqrt{S_0}dw_t, & X_{10}^{[0]} = 0 \\ dX_{2t}^{[0]} = X_{1t}^{[0]} dt, & X_{20}^{[0]} = 0. \end{array}\right. \tag{4.6}$$

Table 5 shows the parameters' values and the computational results in a numerical example; $S_0 = 5.00$. $\epsilon = 0.671$ which is determined such that the coefficient of the diffusion term is equivalent to that of log-normal process at time 0 where the volatility is 30% that is,

$$\epsilon\sqrt{S_0} = \sigma S_0, \qquad \sigma = 0.3.$$

Table 5. Average call option (square-root process).

S_0	5	
ϵ	0.671	(the volatility is 30%)
γ	0.05	
T	1	
K	5.65	
$V^{[0]}$	0.145	(the error is −5.2%)
V	0.153	(a value obtained by 10,000,000 trials)

Table 6. % error (1000 trials, 100 cases).

	hybrid	mc	mc_asymp
avg	−0.1%	−0.9%	−0.9%
rmse	0.8%	6.7%	6.7%
max	1.6%	16.2%	16.2%
min	−1.6%	−14.3%	−14.3%

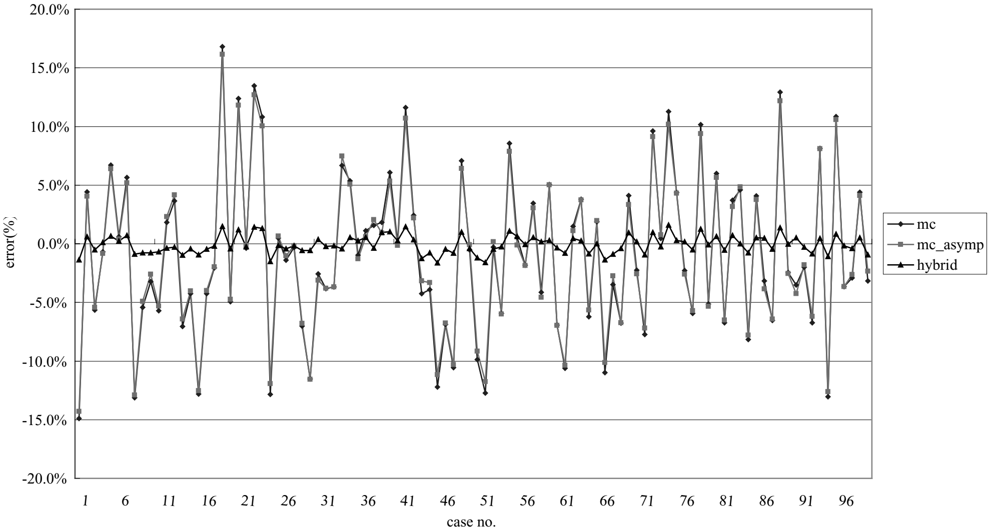

Figure 3. Average call options (square-root process) 1000 trials (100 cases).

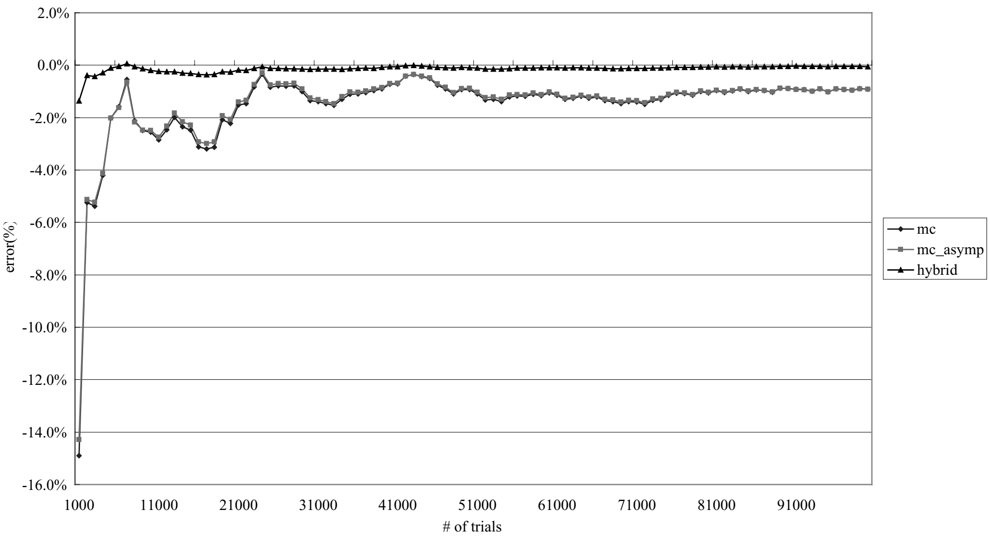

Figure 4. Average call options (square-root process) convergence of simulation.

$\gamma = 0.05$ (5%), $T = 1.0$ (1 year), and $K = 5.65$ (7.5% OTM). V denotes the benchmark value obtained by 10^7 trials of Monte Carlo simulation while $V^{[0]}$ denotes the value obtained by the asymptotics expansion upto ϵ-order, that is the equation (3.25), and it deviates from the benchmark value by -5.2%.

Table 6 shows average (avg), root-mean-square-error (rmse), maximum (max), and minimum (min) of error (%) of three estimators relative to their benchmark values for 100 cases; *hybrid* denotes the modified estimator given by the equation (3.24), *mc* denotes the estimator by naive Monte Carlo for (3.21), that is

$$e^{-\gamma T}\left\{\frac{1}{N}\sum_{j=1}^{N}\left[\left(\frac{1}{T}\bar{X}^{\epsilon}_{2T}+y+\frac{1}{n}\hat{w}_T\right)_+\right]_j\right\},$$

and *mc-asymp* denotes the estimator by naive Monte Carlo for (3.25), that is

$$e^{-\gamma T}\left\{\frac{1}{N}\sum_{j=1}^{N}\left[\left(\frac{1}{T}\bar{X}^{[0]}_{2T}+y+\frac{1}{n}\hat{w}_T\right)_+\right]_j\right\}.$$

Figure 3 shows the errors of three estimators for each 100 cases; the horizontal axis is the case number from 1 to 100 while the vertical axis is the error (%) of those estimators relative to their benchmark values. Clearly, we observe that our estimator is much better than the naive one for each case, and the figure clarifies that the errors of the estimators *mc* and *mc-asymp* are canceled with each other, which contributes to the better performance of our modified estimator *hybrid* for each case. Finally, Fig. 4 shows the comparison of the convergence of three estimators, and the same observation also holds in this case as in Fig. 3.

5. Proofs of Theorems 1 and 2

5.1. Proof of Theorem 1

Since we will need the same notations in the proof of our main results in later sections, we will present a proof of Theorem 1 for completeness. We only prove (i) because (ii) and (iii) are easy. Let

$$u^{\epsilon}_i(x) = \boldsymbol{E}\left[f(X^{\epsilon}_T(t_i,x))\right], \tag{5.1}$$

where $t_i = iT/n$, $i = 0, 1, 2, \ldots, n$. Obviously, $u^{\epsilon}_n(x) = f(x)$, and

$$\begin{aligned} u^{\epsilon}_n(\bar{X}^{\epsilon}_{t_n}) &= u^{\epsilon}_n(\bar{X}^{\epsilon}_T) = f(\bar{X}^{\epsilon}_T),\\ u^{\epsilon}_0(\bar{X}^{\epsilon}_{t_0}) &= u^{\epsilon}_0(x) = \boldsymbol{E}[f(\bar{X}^{\epsilon}_T(0,x))]. \end{aligned}$$

Define Δ^{ϵ}_i as

$$\Delta^{\epsilon}_i := \boldsymbol{E}[u^{\epsilon}_{i+1}(\bar{X}^{\epsilon}_{t_{i+1}})] - \boldsymbol{E}[u^{\epsilon}_i(\bar{X}^{\epsilon}_{t_i})]. \tag{5.2}$$

Then

$$\boldsymbol{E}[f(\bar{X}^{\epsilon}_T)] - \boldsymbol{E}[f(X^{\epsilon}_T(0,x))] = \sum_{i=0}^{n-1}\Delta^{\epsilon}_i.$$

Next, define an operator L_y^ϵ by

$$L_y^\epsilon u_i^\epsilon(x) = \sum_{k=1}^D V_0^{(k)}(y,\epsilon)\partial_k u_i^\epsilon(x) + \frac{1}{2}\sum_{k,j=1}^D \sum_{\alpha=1}^r V_\alpha^{(k)}(y,\epsilon)V_\alpha^{(j)}(y,\epsilon)\partial_k\partial_j u_i^\epsilon(x),$$

where $\partial_k u_i^\epsilon(x) = \frac{\partial u_i^\epsilon(x)}{\partial x_k}$, and $\partial_k\partial_j u_i^\epsilon(x) = \frac{\partial^2 u_i^\epsilon(x)}{\partial x_k \partial x_j}$. Here, $x_k(x_j)$ denotes the $k(j)$-th element of $x = (x_1, \ldots, x_D)$. Similarly, define $\mathcal{L}^\epsilon$ by

$$\begin{aligned}\mathcal{L}^\epsilon u_i^\epsilon(x) &= L_x^\epsilon u_i^\epsilon(x)\\ &= \sum_{k=1}^D V_0^{(k)}(x,\epsilon)\partial_k u_i^\epsilon(x) + \frac{1}{2}\sum_{k,j=1}^D \sum_{\alpha=1}^r V_\alpha^{(k)}(x,\epsilon)V_\alpha^{(j)}(x,\epsilon)\partial_k\partial_j u_i^\epsilon(x).\end{aligned}$$

We know the L^p estimates for the derivatives of $X_T^\epsilon(t,x)$: for any $p \geq 1$ and $l \in \boldsymbol{Z}_+$, there exsits a constant $C \in \boldsymbol{R}_+$ such that

$$\sup_{\substack{t\in[0,T]\\ \epsilon\in(0,1]}} \boldsymbol{E}\left[|\partial_x^l X_T^\epsilon(t,x)|^p\right] \leq C(1+|x|)^C \qquad (x \in \boldsymbol{R}^D)$$

because $\partial_x^l X_T^\epsilon(t,x)$ satisfies a graded stochastic differential equation; see Theorems 5–10 and 5–24 of Bichteler *et al.* (1987). Therefore $\mathcal{L}^\epsilon u_i^\epsilon(x)$ and $L_y^\epsilon u_i^\epsilon(x)$ are of at most polynomial growth in x and in (x,y), respectively. Since $\bar{X}_t^\epsilon$ is L^p-bounded uniformly in (t,ϵ), we have the L^p-boundedness of $\mathcal{L}^\epsilon u_{i+1}^\epsilon(X_t^\epsilon(t_i,\bar{X}_{t_i}^\epsilon))$ and $L_{\bar{X}_{t_i}^\epsilon}^\epsilon u_{i+1}^\epsilon(\bar{X}_t^\epsilon)$.

By the definition of the flow, applying Itô's formula and by the measurability of $\bar{X}_{t_i}^\epsilon$, we obtain:

$$\begin{aligned}\Delta_i^\epsilon &= \boldsymbol{E}[u_{i+1}^\epsilon(\bar{X}_{t_{i+1}}^\epsilon)] - \boldsymbol{E}[u_{i+1}^\epsilon(X_{t_{i+1}}^\epsilon(t_i,\bar{X}_{t_i}^\epsilon))]\\ &= \boldsymbol{E}\left[\int_{t_i}^{t_{i+1}} L_{\bar{X}_{t_i}^\epsilon}^\epsilon u_{i+1}^\epsilon(\bar{X}_t^\epsilon)dt - \int_{t_i}^{t_{i+1}} \mathcal{L}^\epsilon u_{i+1}^\epsilon(X_t^\epsilon(t_i,\bar{X}_{t_i}^\epsilon))dt\right]\\ &= \boldsymbol{E}\left[\int_{t_i}^{t_{i+1}} \{\mathcal{L}^\epsilon u_{i+1}^\epsilon(\bar{X}_{t_i}^\epsilon) - \mathcal{L}^\epsilon u_{i+1}^\epsilon(X_t^\epsilon(t_i,\bar{X}_{t_i}^\epsilon))\}dt\right]\\ &\quad + \boldsymbol{E}\left[\int_{t_i}^{t_{i+1}} \{L_{\bar{X}_{t_i}^\epsilon}^\epsilon u_{i+1}^\epsilon(\bar{X}_t^\epsilon) - L_{\bar{X}_{t_i}^\epsilon}^\epsilon u_{i+1}^\epsilon(\bar{X}_{t_i}^\epsilon)\}dt\right]\\ &= -\int_{t_i}^{t_{i+1}} \boldsymbol{E}[\mathcal{L}^\epsilon u_{i+1}^\epsilon(X_t^\epsilon(t_i,\bar{X}_{t_i}^\epsilon)) - \mathcal{L}^\epsilon u_{i+1}(\bar{X}_{t_i}^\epsilon)]dt\\ &\quad + \int_{t_i}^{t_{i+1}} \boldsymbol{E}[L_{\bar{X}_{t_i}^\epsilon}^\epsilon u_{i+1}^\epsilon(\bar{X}_t^\epsilon) - L_{\bar{X}_{t_i}^\epsilon}^\epsilon u_{i+1}^\epsilon(\bar{X}_{t_i}^\epsilon)]dt.\end{aligned}$$

Hence

$$\begin{aligned}\Delta_i^\epsilon &= -\int_{t_i}^{t_{i+1}}\int_{t_i}^t \boldsymbol{E}[a_{i+1}^\epsilon(X_s^\epsilon(t_i,\bar{X}_{t_i}^\epsilon))]dsdt\\ &\quad + \int_{t_i}^{t_{i+1}}\int_{t_i}^t \boldsymbol{E}[b_{i+1}^\epsilon(\bar{X}_{t_i};\bar{X}_s^\epsilon)]dsdt,\end{aligned} \tag{5.3}$$

where

$$a_{i+1}^{\epsilon}(x) := \mathcal{L}^{\epsilon}(\mathcal{L}^{\epsilon} u_{i+1}^{\epsilon}(x))$$

and

$$b_{i+1}^{\epsilon}(y;x) := L_y^{\epsilon}(L_y^{\epsilon} u_{i+1}^{\epsilon}(x))(x).$$

The function $a_{i+1}^{\epsilon}(x)$ is expressed as

$$\begin{aligned}(5.4)\qquad a_{i+1}^{\epsilon}(x) = {} & \sum_{k'=1}^{D} V_0^{(k')}(x,\epsilon)\partial_{k'} \\ & \times \Bigg\{ \sum_{k=1}^{D} V_0^{(k)}(x,\epsilon)\partial_k u_{i+1}^{\epsilon}(x) \\ & \quad + \frac{1}{2}\sum_{k,l=1}^{D}\sum_{\alpha=1}^{r} V_\alpha^{(k)}(x,\epsilon)V_\alpha^{(l)}(x,\epsilon)\partial_k\partial_l u_{i+1}^{\epsilon}(x)\Bigg\} \\ & + \frac{1}{2}\sum_{k',l'=1}^{D}\sum_{\alpha=1}^{r} V_\alpha^{(k')}(x,\epsilon)V_\alpha^{(l')}(x,\epsilon)\partial_{k'}\partial_{l'} \\ & \times \Bigg\{ \sum_{k=1}^{D} V_0^{(k)}(x)\partial_k u_{i+1}^{\epsilon}(x) \\ & \quad + \frac{1}{2}\sum_{k,l=1}^{D}\sum_{\alpha=1}^{r} V_\alpha^{(k)}(x,\epsilon)V_\alpha^{(l)}(x,\epsilon)\partial_k\partial_l u_{i+1}^{\epsilon}(x)\Bigg\}.\end{aligned}$$

Similarly, $b_{i+1}^{\epsilon}(y;x)$ is expressed as

$$\begin{aligned}(5.5)\qquad b_{i+1}^{\epsilon}(y;x) = {} & \sum_{k'=1}^{D} V_0^{(k')}(y,\epsilon) \\ & \times \Bigg\{ \sum_{k=1}^{D} V_0^{(k)}(y,\epsilon)\partial_{k'}\partial_k u_{i+1}^{\epsilon}(x) \\ & \quad + \frac{1}{2}\sum_{k,l=1}^{D}\sum_{\alpha=1}^{r} V_\alpha^{(k)}(y,\epsilon)V_\alpha^{(l)}(y,\epsilon)\partial_{k'}\partial_k\partial_l u_{i+1}^{\epsilon}(x)\Bigg\} \\ & + \frac{1}{2}\sum_{k',l'=1}^{D}\sum_{\alpha=1}^{r} V_\alpha^{(k')}(y)V_\alpha^{(l')}(y) \\ & \times \Bigg\{ \sum_{k=1}^{D} V_0^{(k)}(y,\epsilon)\partial_{k'}\partial_{l'}\partial_k u_{i+1}^{\epsilon}(x) \\ & \quad + \frac{1}{2}\sum_{k,l=1}^{D}\sum_{\alpha=1}^{r} V_\alpha^{(k)}(y,\epsilon)V_\alpha^{(l)}(y,\epsilon)\partial_{k'}\partial_{l'}\partial_k\partial_l u_{i+1}^{\epsilon}(x)\Bigg\}.\end{aligned}$$

Note that $a_{i+1}^{\epsilon}(x)$ is a ploynomial in

$$V_0^{(k_1)},\ \partial_{k_2}V_0^{(k_1)},\ \partial_{k_2}\partial_{l_2}V_0^{(k_1)},$$

$$V_\alpha^{(k_2)},\ \partial_{k_2}V_\alpha^{(k_1)},\ \partial_{k_2}\partial_{l_2}V_\alpha^{(k_1)},$$
$$\partial_{k_1}u_{i+1}^\epsilon,\ \partial_{k_1}\partial_{k_2}u_{i+1}^\epsilon,\ \partial_{k_1}\partial_{k_2}\partial_l u_{i+1}^\epsilon, \quad \text{and} \quad \partial_{k_1}\partial_{k_2}\partial_{l_1}\partial_{l_2}u_{i+1}^\epsilon$$

for $k_1, k_2, l_1, l_2 = 1, 2, \ldots, D$ and $\alpha = 1, 2, \ldots, r$. Note also that $V_\alpha(x) \in C_\uparrow^\infty(\boldsymbol{R}^D)$, $\alpha = 0, 1, \ldots, r$ and $f \in C_\uparrow^\infty(\boldsymbol{R}^D)$.

Further, it is well known (see e.g. Chapter II-5 of Bichteler *et al.* (1987)) that

$$\begin{cases} \sup_\epsilon \sup_n \sup_{0\le s\le T} \boldsymbol{E}[|\bar{X}_s^\epsilon|^p] < \infty \\ \sup_\epsilon \sup_n \sup_{t_i\le s\le t_{i+1}} \boldsymbol{E}[|X_s^\epsilon(t_i, \bar{X}_{t_i}^\epsilon)|^p] < \infty \end{cases} \tag{5.6}$$

for all $p \geq 1$. Then, by using the Hölder inequaility, we have

$$\sup_\epsilon \sup_n \sup_{i\in\{1,2,\cdots,n\}} \sup_{t_i\le s\le t_{i+1}} \boldsymbol{E}[|a_{i+1}^\epsilon(X_s^\epsilon(t_i, \bar{X}_{t_i}^\epsilon))|] < \infty. \tag{5.7}$$

Similarly,

$$\sup_\epsilon \sup_n \sup_{i\in\{1,2,\cdots,n\}} \sup_{t_i\le s\le t_{i+1}} \boldsymbol{E}[|b_{i+1}^\epsilon(\bar{X}_{t_{i+1}}; \bar{X}_s^\epsilon)|] < \infty. \tag{5.8}$$

Thus, we conclude that

$$\begin{aligned} &\boldsymbol{E}[f(\bar{X}_T^\epsilon)] - \boldsymbol{E}[f(X_T^\epsilon(0,x))] \\ &\quad = \sum_{i=0}^{n-1} \Delta_i^\epsilon = \sum_{i=0}^{n-1} \Bigg\{ -\int_{t_i}^{t_{i+1}} \int_{t_i}^{t} \boldsymbol{E}[a_{i+1}^\epsilon(X_s^\epsilon(t_i, \bar{X}_{t_i}^\epsilon))] ds dt \\ &\qquad\qquad\qquad + \int_{t_i}^{t_{i+1}} \int_{t_i}^{t} \boldsymbol{E}[b_{i+1}^\epsilon(\bar{X}_{t_{i+1}}; \bar{X}_s^\epsilon)] ds dt \Bigg\} \\ &\quad = O\left(\frac{1}{n}\right). \end{aligned} \tag{5.9}$$

5.2. Proof of Theorem 2

We follow a relatively standard argument in the proofs of Theorems 2 and 3. We only prove (i). Others are easy to show and we omit the proof.

First, we claim that

$$\sup_{s,i,n} |\boldsymbol{E}[a_{i+1}^\epsilon(X_s^\epsilon(t_i, \bar{X}_{t_i}^\epsilon))] - \boldsymbol{E}[a_{i+1}^0(X_s^0(t_i, \bar{X}_{t_i}^0))]| = O(\epsilon) \quad (\epsilon \downarrow 0) \tag{5.10}$$

and that

$$\sup_{s,i,n} |\boldsymbol{E}[b_{i+1}^\epsilon(\bar{X}_{t_i}^\epsilon; \bar{X}_s^\epsilon)] - \boldsymbol{E}[b_{i+1}^0(\bar{X}_{t_i}^0; \bar{X}_s^0)]| = O(\epsilon) \quad (\epsilon \downarrow 0). \tag{5.11}$$

We will show only the first one, and the second one can be obtained in a similar way.

We need to show that

$$\begin{aligned} \overline{\lim_{\epsilon\downarrow 0}} \frac{1}{\epsilon} \sup_n \sup_{i\in\{1,2,\ldots,n\}} \sup_{t_i\le s\le t_{i+1}} &|\boldsymbol{E}[a_{i+1}^\epsilon(X_s^\epsilon(t_i, \bar{X}_{t_i}^\epsilon))] \\ &- \boldsymbol{E}[a_{i+1}^0(X_s^0(t_i, \bar{X}_{t_i}^0))]| < \infty. \end{aligned} \tag{5.12}$$

Notice that

$$a_{i+1}^{\epsilon}(X_s^{\epsilon}(t_i,\bar{X}_{t_i}^{\epsilon})) = a_{i+1}^{0}(X_s^{0}(t_i,\bar{X}_{t_i}^{0})) + \epsilon\int_0^1 \partial_\epsilon\,|_{\epsilon=u\epsilon}\; a_{i+1}^{\epsilon}(X_s^{\epsilon}(t_i,\bar{X}_{t_i}^{\epsilon}))du,$$

where

$$\partial_\epsilon\,|_{\epsilon=u\epsilon}\; a_{i+1}^{\epsilon}(X_s^{\epsilon}(t_i,\bar{X}_{t_i}^{\epsilon})) \equiv \left.\frac{\partial a_{i+1}^{\epsilon}(X_s^{\epsilon}(t_i,\bar{X}_{t_i}^{\epsilon}))}{\partial\epsilon}\right|_{\epsilon=u\epsilon}.$$

Then

$$\begin{aligned}
&\frac{1}{\epsilon}\sup_{s,i,n}|\boldsymbol{E}[a_{i+1}^{\epsilon}(X_s^{\epsilon}(t_i,\bar{X}_{t_i}^{\epsilon})) - a_{i+1}^{0}(X_s^{0}(t_i,\bar{X}_{t_i}^{0}))]|\\
&\quad= \sup_{s,i,n}\left|\int_0^1 \boldsymbol{E}[\partial_\epsilon\,|_{\epsilon=u\epsilon}\; a_{i+1}^{\epsilon}(X_s^{\epsilon}(t_i,\bar{X}_{t_i}^{\epsilon}))]du\right|\\
&\quad\le \sup_{s,i,n}\int_0^1 \boldsymbol{E}[|\partial_\epsilon\,|_{\epsilon=u\epsilon}\; a_{i+1}^{\epsilon}(X_s^{\epsilon}(t_i,\bar{X}_{t_i}^{\epsilon}))|]du\\
&\quad\le \sup_{s,i,n}\sup_{0<\epsilon_1<\epsilon}\|\partial_{\epsilon_1}a_{i+1}^{\epsilon_1}(X_s^{\epsilon_1}(t_i,\bar{X}_{t_i}^{\epsilon_1}))\|_1,
\end{aligned}$$

where $\|\cdot\|_1$ denotes $L_1(P)$-norm. Note that $\partial_{\epsilon_1}a_{i+1}^{\epsilon_1}(X_s^{\epsilon_1}(t_i,\bar{X}_{t_i}^{\epsilon_1}))$, is a polynomial in partial derivatives of each term of (5.4) with respect to the parameter ϵ at $\epsilon=\epsilon_1$, and

$$\begin{aligned}
&\frac{\partial X_s^{\epsilon_1}(t_i,\bar{X}_{t_i}^{\epsilon_1})}{\partial\epsilon_1},\\
&V_0^{(k_1)},\ \partial_{k_2}V_0^{(k_1)},\ \partial_{k_2}\partial_{l_2}V_0^{(k_1)},\ \partial_{k_1}\partial_{k_2}\partial_{l_2}V_0^{(k_1)},\\
&V_\alpha^{(k_2)},\ \partial_{k_2}V_\alpha^{(k_1)},\ \partial_{k_2}\partial_{l_2}V_\alpha^{(k_1)},\ \partial_{k_1}\partial_{k_2}\partial_{l_2}V_\alpha^{(k_1)},\\
&\partial_{k_1}u_{i+1}^{\epsilon_1},\ \partial_{k_1}\partial_{k_2}u_{i+1}^{\epsilon_1},\ \partial_{k_1}\partial_{k_2}\partial_l u_{i+1}^{\epsilon_1},\quad \text{and}\quad \partial_{k_1}\partial_{k_2}\partial_{l_1}\partial_{l_2}\partial_m u_{i+1}^{\epsilon_1}
\end{aligned}$$

for $k_1,k_2,l_1,l_2,m = 1,2,\dots,D$ and $\alpha = 1,2,\dots,r$. Those are evaluated at $x = X_s^{\epsilon_1}(t_i,\bar{X}_{t_i}^{\epsilon_1})$, $0<\epsilon_1<\epsilon$.

We apply a similar argument in Chapter II-5 of Bichteler *et al.* (1987) to the system of equations:

$$(5.13)\quad \begin{cases}
\bar{X}_s^{\epsilon_1} = x + \int_0^s V_0(\bar{X}_{\eta(u)}^{\epsilon_1},\epsilon_1)du + \int_0^s V(\bar{X}_{\eta(u)}^{\epsilon_1},\epsilon_1)dw_u, \qquad s\in[0,T],\\
X_s^{\epsilon_1}(t_i,\bar{X}_{t_i}^{\epsilon_1}) = \bar{X}_{t_i}^{\epsilon_1} + \int_{t_i}^s V_0(X_u^{\epsilon_1}(t_i,\bar{X}_{t_i}^{\epsilon_1}),\epsilon_1)du\\
\qquad\qquad + \int_{t_i}^s V(X_u^{\epsilon_1}(t_i,\bar{X}_{t_i}^{\epsilon_1}),\epsilon_1)dw_u, \qquad s\in[t_i,t_{i+1}),\\
\frac{\partial\bar{X}_s^{\epsilon_1}}{\partial\epsilon_1} = \{\int_0^s \partial_{\epsilon_1}V_0(\bar{X}_{\eta(u)}^{\epsilon_1},\epsilon_1)du + \int_0^s \partial_{\epsilon_1}V(\bar{X}_{\eta(u)}^{\epsilon_1},\epsilon_1)dw_u\}\\
\qquad + \int_0^s \partial V_0(\bar{X}_{\eta(u)},\epsilon_1)\{\frac{\partial\bar{X}_{\eta(u)}^{\epsilon_1}}{\partial\epsilon_1}\}du\\
\qquad + \sum_{\alpha=1}^r \int_0^s \partial V_\alpha(\bar{X}_{\eta(u)}^{\epsilon_1},\epsilon_1)\{\frac{\partial\bar{X}_{\eta(u)}^{\epsilon_1}}{\partial\epsilon_1}\}dw_u^\alpha,\\
\frac{\partial X_s^{\epsilon_1}(t_i,\bar{X}_{t_i}^{\epsilon_1})}{\partial\epsilon_1} = \frac{\partial\bar{X}_{t_i}^{\epsilon_1}}{\partial\epsilon_1} + \{\int_{t_i}^s \partial_{\epsilon_1}V_0(X_u^{\epsilon_1}(t_i,\bar{X}_{t_i}^{\epsilon_1}))du\\
\qquad\qquad + \int_{t_i}^s \partial_{\epsilon_1}V(X_u^{\epsilon_1}(t_i,\bar{X}_{t_i}^{\epsilon_1}),\epsilon_1)dw_u\}\\
\qquad + \int_{t_i}^s \partial V_0(X_u^{\epsilon_1}(t_i,\bar{X}_{t_i}^{\epsilon_1}))\{\frac{\partial X_u^{\epsilon_1}(t_i,\bar{X}_{t_i}^{\epsilon_1})}{\partial\epsilon_1}\}du\\
\qquad + \sum_{\alpha=1}^r \int_{t_i}^s \partial V_\alpha(X_u^{\epsilon_1}(t_i,\bar{X}_{t_i}^{\epsilon_1}),\epsilon_1)\\
\qquad \times\{\frac{\partial X_u^{\epsilon_1}(t_i,\bar{X}_{t_i}^{\epsilon_1})}{\partial\epsilon_1}\}dw_u^\alpha, \qquad s\in[t_i,t_{i+1}),
\end{cases}$$

where ∂V_α, $\alpha = 0, 1, \dots, r$ denote the partial derivatives with respect to the first argument. Then, we can also show that

$$(5.14)\quad \begin{cases} \sup_n \sup_{0\le s\le T} \sup_{0<\epsilon_1<\epsilon} \boldsymbol{E}[|\bar{X}_s^{\epsilon_1}|^p] < \infty, \\ \sup_n \sup_{i\in\{1,2,\dots,n\}} \sup_{t_i\le s\le t_{i+1}} \sup_{0<\epsilon_1<\epsilon} \boldsymbol{E}[|\bar{X}_s^{\epsilon_1}(t_i, \bar{X}_{t_i}^{\epsilon_1})|^p] < \infty \\ \sup_n \sup_{0\le s\le T} \sup_{0<\epsilon_1<\epsilon} \boldsymbol{E}[|\frac{\partial \bar{X}_s^{\epsilon_1}}{\partial \epsilon_1}|^p] < \infty \\ \sup_n \sup_{i\in\{1,2,\dots,n\}} \sup_{t_i\le s\le t_{i+1}} \sup_{0<\epsilon_1<\epsilon} \boldsymbol{E}[|\frac{\partial X_s^{\epsilon_1}(t_i, \bar{X}_{t_i}^{\epsilon_1})}{\partial \epsilon_1}|^p] < \infty \end{cases}$$

for all $p \ge 1$.

Thus, $\partial_{\epsilon_1} a_{i+1}^{\epsilon_1}(X_s^{\epsilon_1}(t_i, \bar{X}_{t_i}^{\epsilon_1}))$ is L_p-bounded for any $p \ge 1$ uniformly in s, i, n and $0 < \epsilon_1 < \epsilon$.

We return to the proof of (i). We see

$$\begin{aligned} &\mathbf{Bias}[\boldsymbol{V}^*(\epsilon, n, N)] \\ &\quad = \boldsymbol{E}[V^*(\epsilon, n, N)] - V \\ &\quad = \{\boldsymbol{E}[f(\bar{X}_T^\epsilon)] - \boldsymbol{E}[f(X_T^\epsilon(0,x))]\} - \{\boldsymbol{E}[f(\bar{X}_T^{[0]})] - \boldsymbol{E}[f(X_T^{[0]}(0,x))]\} \\ &\quad = \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \int_{t_i}^{s} \{-\boldsymbol{E}[a_{i+1}^\epsilon(X_s^\epsilon(t_i, \bar{X}_{t_i}^\epsilon))] + \boldsymbol{E}[b_{i+1}^\epsilon(\bar{X}_{t_i}^\epsilon; \bar{X}_s^\epsilon)]\} dsdt \\ &\qquad - \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \int_{t_i}^{s} \{-\boldsymbol{E}[a_{i+1}^0(X_s^0(t_i, \bar{X}_{t_i}^0))] + \boldsymbol{E}[b_{i+1}^0(\bar{X}_{t_i}^0; \bar{X}_s^0)]\} dsdt \\ &\quad = \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \int_{t_i}^{s} -\{\boldsymbol{E}[a_{i+1}^\epsilon(X_s^\epsilon(t_i, \bar{X}_{t_i}^\epsilon))] - \boldsymbol{E}[a_{i+1}^0(X_s^0(t_i, \bar{X}_{t_i}^0))]\} dsdt \\ &\qquad + \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \int_{t_i}^{s} \{\boldsymbol{E}[b_{i+1}^\epsilon(\bar{X}_{t_i}^\epsilon; \bar{X}_s^\epsilon)]\}] - \boldsymbol{E}[b_{i+1}^0(\bar{X}_{t_i}^0; \bar{X}_s^0)]\} dsdt. \end{aligned}$$

Hence, using the estimate already obtained, we conclude that

$$\boldsymbol{E}[f(\bar{X}_T^\epsilon)] - \boldsymbol{E}[f(X_T^\epsilon(0,x))] - \boldsymbol{E}[f(\bar{X}_T^0)] + \boldsymbol{E}[f(X_T^0(0,x))] = O\left(\frac{\epsilon}{n}\right).$$

6. Proof of Theorems 3

We only prove (i) again. The others are easy. Let $A = 1 + |x|^2 - \frac{1}{2}\Delta$, and then A^{-1} is an integral operator. (See Ikeda and Watanabe (1989) or Sakamoto and Yoshida (1996) for the detail.) Then, under [A1] for a sufficiently large integer m depending on f, we obtain

$$(6.1)\quad \begin{aligned} &\boldsymbol{E}\left[f\left(X_T^\epsilon(0,x) + \frac{1}{n}\hat{w}_T\right)\right] - \boldsymbol{E}[f(X_T^\epsilon(0,x))] \\ &\quad = \boldsymbol{E}\left[(A^{-m}f)\left(X_T^\epsilon(0,x) + \frac{1}{n}\hat{w}_T\right)\Psi_1^{(\epsilon)}\right] \\ &\qquad - \boldsymbol{E}[(A^{-m}f)(X_T^\epsilon(0,x))\Psi_2^{(\epsilon)}] \end{aligned}$$

for some Wiener functionals $\Psi_1^{(\epsilon)}$ and $\Psi_2^{(\epsilon)}$ which correspond to the partial shifts only in the direction of w. Under [A1], the integration-by-parts formulas (6.1) (for ϵ and $\epsilon = 0$) and easy calculus with the Taylor formula yield

$$\begin{aligned}(6.2)\qquad &\left\{\boldsymbol{E}\left[f\left(X_T^\epsilon(0,x)+\frac{1}{n}\hat{w}_T\right)\right]-\boldsymbol{E}[f(X_T^\epsilon(0,x)]\right\}\\ &\qquad -\left\{\boldsymbol{E}\left[f\left(X_T^{[0]}(0,x)+\frac{1}{n}\hat{w}_T\right)\right]-\boldsymbol{E}[f(X_T^{[0]}(0,x)]\right\}\\ &= O\left(\frac{\epsilon}{n}\right).\end{aligned}$$

On the other hand, obviously,

$$\begin{aligned}(6.3)\quad &\boldsymbol{E}\left[f\left(X_T^\epsilon(0,x)+\frac{1}{n}\hat{w}_T\right)\right]-\boldsymbol{E}[f(X_T^\epsilon(0,x))]\\ &=\left(\boldsymbol{E}\left[f\left(X_T^\epsilon(0,x)+\frac{1}{n}\hat{w}_T\right)\psi(|\sigma_{X_T^{[0]}(0,x)}|/|4\sigma_{X_T^\epsilon(0,x)}|)\right]\right.\\ &\qquad\left.-\boldsymbol{E}[f(X_T^\epsilon(0,x))\psi(|\sigma_{X_T^{[0]}(0,x)}|/|4\sigma_{X_T^\epsilon(0,x)}|)]\right)\\ &\quad+\left(\boldsymbol{E}\left[f\left(X_T^\epsilon(0,x)+\frac{1}{n}\hat{w}_T\right)\{1-\psi(|\sigma_{X_T^{[0]}(0,x)}|/|4\sigma_{X_T^\epsilon(0,x)}|)\}\right]\right.\\ &\qquad\left.-\boldsymbol{E}[f(X_T^\epsilon(0,x))\{1-\psi(|\sigma_{X_T^{[0]}(0,x)}|/|4\sigma_{X_T^\epsilon(0,x)}|)\}]\right),\end{aligned}$$

where $\psi : \boldsymbol{R} \to [0,1]$ is a smooth function such that

$$\psi(x)=\begin{cases}1 & \text{if } |x|\le \frac{1}{2}\\ 0 & \text{if } |x|\ge 1.\end{cases}$$

For the second parenthesis,

$$\begin{aligned}(6.4)\quad &\boldsymbol{E}\left[f\left(X_T^\epsilon(0,x)+\frac{1}{n}\hat{w}_T\right)\{1-\psi(|\sigma_{X_T^{[0]}(0,x)}|/|4\sigma_{X_T^\epsilon(0,x)}|)\}\right]\\ &\qquad-\boldsymbol{E}[f(X_T^\epsilon(0,x))\{1-\psi(|\sigma_{X_T^{[0]}(0,x)}|/|4\sigma_{X_T^\epsilon(0,x)}|)\}]\\ &\le C\|1-\psi(|\sigma_{X_T^{[0]}(0,x)}|/|4\sigma_{X_T^\epsilon(0,x)}|)\|_q \qquad \text{(by the Hölder inequality)}\\ &\le C\times P(\{|\sigma_{X_T^{[0]}(0,x)}|/|\sigma_{X_T^\epsilon(0,x)}|>2\})^{1/q}\\ &\le C\times 2^K\boldsymbol{E}\left[\left(\frac{|\sigma_{X_T^\epsilon(0,x)}-\sigma_{X_T^{[0]}(0,x)}|}{|\sigma_{X_T^{[0]}(0,x)}|}\right)^K\right]\\ &\qquad\text{(by Markov's inequality)}\\ &=O(\epsilon^K)\end{aligned}$$

for any $K > 0$. Here C is some positive costant, $q > 1$, and $\|\cdot\|_q$ denotes the $L^q(P^w \otimes P^{\hat{w}})$-norm. It is also easy to obtain an estimate similar to (6.4) replacing $X_T^\epsilon(0,x)$ in f by $X_T^0(0,x)$. Hence under [A2], by the same argument as we obtained (6.2), we can estimate the gap

$$\begin{aligned}&\Big(\boldsymbol{E}\left[f\left(X_T^\epsilon(0,x)+\frac{1}{n}\hat{w}_T\right)\psi(|\sigma_{X_T^{[0]}(0,x)}|/|4\sigma_{X_T^\epsilon(0,x)}|)\right]\\&\qquad-\boldsymbol{E}[f(X_T^\epsilon(0,x))\psi(|\sigma_{X_T^{[0]}(0,x)}|/|4\sigma_{X_T^\epsilon(0,x)}|)]\Big)\\&\qquad-\Big(\boldsymbol{E}\left[f\left(X_T^0(0,x)+\frac{1}{n}\hat{w}_T\right)\psi(|\sigma_{X_T^{[0]}(0,x)}|/|4\sigma_{X_T^\epsilon(0,x)}|)\right]\\&\qquad\qquad-\boldsymbol{E}[f(X_T^0(0,x))\psi(|\sigma_{X_T^{[0]}(0,x)}|/|4\sigma_{X_T^\epsilon(0,x)}|)]\Big),\end{aligned}$$

and obtain

$$\begin{aligned}(6.5)\qquad&\left\{\boldsymbol{E}\left[f\left(X_T^\epsilon(0,x)+\frac{1}{n}w_T\right)\right]-\boldsymbol{E}[f(X_T^\epsilon(0,x)]\right\}\\&\qquad-\left\{\boldsymbol{E}\left[f\left(X_T^{[0]}(0,x)+\frac{1}{n}w_T\right)\right]-\boldsymbol{E}[f(X_T^{[0]}(0,x))]\right\}\\&\quad=O\left(\frac{\epsilon}{n}\right)+O(\epsilon^K)\end{aligned}$$

for every $K>0$.

The *Bias* of $\boldsymbol{V}^*(\epsilon,n,N)$ is expressed as

$$\begin{aligned}(6.6)\quad&\mathbf{Bias}[\boldsymbol{V}^*(\epsilon,n,N)]\\&=\Bigg[\left\{\boldsymbol{E}\left[f\left(\bar{X}_T^\epsilon+\frac{1}{n}\hat{w}_T\right)\right]-\boldsymbol{E}\left[f\left(X_T^\epsilon(0,x)+\frac{1}{n}\hat{w}_T\right)\right]\right\}\\&\qquad-\left\{\boldsymbol{E}\left[f\left(\bar{X}_T^{[0]}+\frac{1}{n}\hat{w}_T\right)\right]-\boldsymbol{E}\left[f\left(X_T^{[0]}(0,x)+\frac{1}{n}\hat{w}_T\right)\right]\right\}\Bigg]\\&\quad+\Bigg[\left\{\boldsymbol{E}\left[f\left(X_T^\epsilon(0,x)+\frac{1}{n}\hat{w}_T\right)\right]-\boldsymbol{E}[f(X_T^\epsilon(0,x)]\right\}\\&\qquad-\left\{\boldsymbol{E}\left[f\left(X_T^{[0]}(0,x)+\frac{1}{n}\hat{w}_T\right)\right]-\boldsymbol{E}[f(X_T^{[0]}(0,x)]\right\}\Bigg].\end{aligned}$$

From (6.2), the second square bracket on the right-hand side is $O(\frac{\epsilon}{n})$ under Condition [A1]. Hence if we show that the first square bracket is $O(\frac{\epsilon}{n})$, then $\mathbf{Bias}[\boldsymbol{V}^*(\epsilon,n,N)]$ turns out to be $O(\frac{\epsilon}{n})$ under [A1]. Similarly, because under Condition [A2] the second square bracket is $O(\frac{\epsilon}{n})+O(\epsilon^K)$ for every $K>0$ by (6.5), if we show that the first square bracket is $O(\frac{\epsilon}{n})$, then we can conclude that $\mathbf{Bias}[\boldsymbol{V}^*(\epsilon,n,N)]$ is $O(\frac{\epsilon}{n})+O(\epsilon^K)$ for every $K>0$ under [A2]. Because $O(\epsilon^K)$ is the smaller order than the order of $\frac{\epsilon}{n}$ for large K by the assumption that $\epsilon=o(n^{-\omega})$ for some positive constant ω as $n\to\infty$, $\mathbf{Bias}[\boldsymbol{V}^*(\epsilon,n,N)]$ is $O(\frac{\epsilon}{n})$ under [A2].

Hence, in order to complete the proof, we will evaluate the first square bracket on the right-hand side of (6.6). First, define u_i^ϵ by

$$(6.7)\qquad u_i^\epsilon(x)=\boldsymbol{E}\left[f\left(X_T^\epsilon(t_i,x)+\frac{1}{n}\hat{w}_T\right)\right].$$

We can write

$$\boldsymbol{E}\left[f\left(\bar{X}_T^\epsilon + \frac{1}{n}\hat{w}_T\right)\right] - \boldsymbol{E}\left[f\left(X_T^\epsilon(0,x) + \frac{1}{n}\hat{w}_T\right)\right] = \sum_{i=0}^{n-1} \Delta_i^\epsilon,$$

where

$$\Delta_i^\epsilon := \boldsymbol{E}[u_{i+1}^\epsilon(\bar{X}_{t_{i+1}}^\epsilon)] - \boldsymbol{E}[u_i^\epsilon(\bar{X}_{t_i}^\epsilon)], \tag{6.8}$$

and also

$$\begin{aligned}\boldsymbol{E}[u_i^\epsilon(\bar{X}_{t_i}^\epsilon)] &= \boldsymbol{E}\left[f\left(X_T^\epsilon(t_i, \bar{X}_{t_i}^\epsilon) + \frac{1}{n}\hat{w}_T\right)\right] \\ &= \boldsymbol{E}\left[f\left(X_T^\epsilon(t_{i+1}, X_{t_{i+1}}^\epsilon(t_i, \bar{X}_{t_i}^\epsilon)) + \frac{1}{n}\hat{w}_T\right)\right] \\ &= \boldsymbol{E}[u_{i+1}^\epsilon(X_{t_{i+1}}^\epsilon(t_i, \bar{X}_{t_i}^\epsilon))].\end{aligned}$$

The gaps Δ_i^ϵ are expressed in exactly the same form (5.3) as in the smooth case (i.e., $f \in C_\uparrow^\infty(\boldsymbol{R}^D)$). That is, a_{i+1}^ϵ and b_{i+1}^ϵ are defined as equations (5.4) and (5.5), respectively, and they include partial derivatives of $u_{i+1}^\epsilon(x)$ with respect to x. Even in the irregular case (i.e., where f is not necessarily differentiable nor continuous), these derivatives are justified by the (full) Malliavin calculus in which the shift operation is done in both directions of w and $\hat{w}$. (However, only for this purpose, the nondegeneracy of $\hat{w}$-terms is essential.)

In order to follow the same procedure as the proof of Theorem 2, we need to show the uniform boundedness

$$\overline{\lim_{n\to\infty,\epsilon\downarrow 0}} \sup_{i,s} |\boldsymbol{E}[\partial_\epsilon\{a_{i+1}^\epsilon(X_s^\epsilon(t_i, \bar{X}_{t_i}^\epsilon))\}]| < \infty, \tag{6.9}$$

for example. Here s moves over $[t_i, t_{i+1}]$. If we write out $\boldsymbol{E}[\partial_\epsilon\{a_{i+1}^\epsilon(X_s^\epsilon(t_i, \bar{X}_{t_i}^\epsilon))\}]$, there appear several terms. Among them, we have for example the following type of terms

$$I(f;i,s,\epsilon,n) := \boldsymbol{E}\left[\left\{B(x)\partial_\epsilon\partial_x^j \boldsymbol{E}\left[f\left(X_T^\epsilon(t_{i+1},x) + \frac{1}{n}\hat{w}_T\right)\right]\right\}\Big|_{x=X_s^\epsilon(t_i,\bar{X}_{t_i}^\epsilon)}\right],$$

where B is a smooth function of at most polynomial growth. Roughly speaking, it follows from the IBP-fomula that the functions $\partial_\epsilon\partial_x^j\boldsymbol{E}[f(X_T^\epsilon(t_{i+1},x) + \frac{1}{n}\hat{w}_T)]$ are nice functions of x, so that the functionals with $X_s^\epsilon(t_i, \bar{X}_{t_i}^\epsilon)$ substituted for x are also nice and have uniformly bounded norms. We will show this fact more rigorously.

Just for notational simplicity, we only consider one-dimensional X^ϵ. Let $\mathcal{S}$ denote the set of Schwartz test functions. For $f \in \mathcal{S}$,

(6.10) $I(f;i,s,\epsilon,n)$

$$= \boldsymbol{E}\Bigg[B(X_s^\epsilon(t_i,\bar{X}_{t_i}^\epsilon)) \sum_{k=1}^{j+1} \boldsymbol{E}\Bigg[(\partial^k f)\left(X_T^\epsilon(t_{i+1},x) + \frac{1}{n}\hat{w}_T\right)$$

$$\left.\left.\cdot P_k(\partial_x^\alpha \partial_\epsilon X_T^\epsilon(t_{i+1},x);\alpha=0,1,\ldots,j+1-k)\right]\right|_{x=X_s^\epsilon(t_i,\bar{X}_{t_i}^\epsilon)}\Bigg]$$

$$= \boldsymbol{E}\Bigg[B(X_s^\epsilon(t_i,\bar{X}_{t_i}^\epsilon))\sum_{k=1}^{j+1}(\partial^k f)\left(X_T^\epsilon(t_{i+1},X_s^\epsilon(t_i,\bar{X}_{t_i}^\epsilon))+\frac{1}{n}\hat{w}_T\right)$$

$$\cdot P_k(\partial_x^\alpha \partial_\epsilon X_T^\epsilon(t_{i+1},x)\mid_{x=X_s^\epsilon(t_i,\bar{X}_{t_i}^\epsilon)};\alpha=0,1,\ldots,j+1-k)\Bigg],$$

where P_k are polynomials, and we used independency.

Set

$$\check{X}(i,s,\epsilon,n)=X_T^\epsilon(t_{i+1},X_s^\epsilon(t_i,\bar{X}_{t_i}^\epsilon))+\frac{1}{n}\hat{w}_T.$$

We denote by $\sigma_{\check{X}(i,s,\epsilon,n)}$ the (full) Malliavin covariance of $\check{X}(i,s,\epsilon,n)$. We write the IBP-formula as

$$\boldsymbol{E}[(\partial^k f)(\check{X}(i,s,\epsilon,n))\psi]=\boldsymbol{E}[f(\check{X}(i,s,\epsilon,n))\Phi_k(\psi;\check{X}(i,s,\epsilon,n))]$$

for $f\in\mathcal{S}$ and smooth functional ψ. The functional $\Phi_1(\psi;\check{X}(i,s,\epsilon,n))$ is given by

$$\Phi_1(\psi;\check{X}(i,s,\epsilon,n))=D^*[\sigma^{-1}_{\check{X}(i,s,\epsilon,n)}\psi D\check{X}(i,s,\epsilon,n)]$$

with H-derivative D and its adjoint D^*, and $\Phi_k(\psi;\check{X}(i,s,\epsilon,n))$ are determined by repeated use of this expression. A similar formula exists for multi-dimensional case. Applying this IBP-formula, we obtain

$$I(f;i,s,\epsilon,n)=\sum_{k=1}^{j+1}\boldsymbol{E}[(A^{-m}f)(\check{X}(i,s,\epsilon,n))\check{\Psi}_{k+2m}] \tag{6.11}$$

for a sufficiently large integer m. Functional $\check{\Psi}_{k+2m}$ has an expression similar to that of $\Phi_k(\psi;\check{X}(i,s,\epsilon,n))$. The L^1-norm of $\check{\Psi}_{k+2m}$ is dominated by a polynomial of L^p-norms of $\sigma^{-1}_{\check{X}(i,s,\epsilon,n)}$ and $D_{p,s}$-norms of

$$(\partial_x^{\alpha_1}\partial_\epsilon^{\alpha_2}X_T^\epsilon(t_{i+1},x))\mid_{x=X_s^\epsilon(t_i,\bar{X}_{t_i}^\epsilon)}\qquad(\alpha_1,\alpha_2\in\boldsymbol{Z}_+) \tag{6.12}$$

as well as $\hat{w}_T$. The H-derivative of (6.12) is decomposed into the derivative component for $(w_t)_{t\in[0,t_{i+1}]}$ and that for $(w_t)_{t\in[t_{i+1},T]}$; therefore, estimation of its $D_{p,s}$-norms results in estimation of L^p-norms of $D_{p,s}$-norms of solutions of certain stochastic differential equations. It is just a routine job to show that those $D_{p,s}$-norms are bounded uniformly in i,s,ϵ,n.

Under Condition [A2] ([A1] in force), by Lemma 1 below, we know that $\check{X}(i,s,\epsilon,n)$ is uniformly nondegenerate:

$$\overline{\lim_{n\to\infty,\epsilon\downarrow 0}}\sup_i\sup_{s\in[t_i,t_{i+1}]}\boldsymbol{E}[\sigma^{-p}_{\check{X}(i,s,\epsilon,n)}]<\infty$$

for all $p>1$ (det should be put in multi-dimensional case). After all, we obtain

$$|I(f;i,s,\epsilon,n)|\le C\|f\|_{-2m}\qquad(f\in\mathcal{S}) \tag{6.13}$$

for any i, s and sufficiently large n and sufficiently small ϵ. Here C is a constant independent of i, s, ϵ, n, and $\|\cdot\|_{-2m}$ is the norm attached to the space C_{-2m} (see Ikeda and Watanabe (1989), Sakamoto and Yoshida (1996)).

Let ϕ_n be the density of the normal distribution $N(0, T/n^2)$. From (6.10), It is easy to see that $I(\cdot; i, s, \epsilon, n)$ is a signed-measure: for measurable functions f of at most polynomial growth,

$$I(f; i, s, \epsilon, n) = \int f(z)p(z)dz \tag{6.14}$$

with

$$p(z) = \sum_{k=1}^{j+1} \boldsymbol{E}[B(X_s^\epsilon(t_i, \bar{X}_{t_i}^\epsilon))(-\partial)^k \phi_n(z - \check{X}(i, s, \epsilon, n))$$
$$\cdot P_k(\partial_x^\alpha \partial_\epsilon X_T^\epsilon(t_{i+1}, x)\mid_{x=X_s^\epsilon(t_i, \bar{X}_{t_i}^\epsilon)}; \alpha = 0, 1, \ldots, j+1-k)].$$

Obviously, $p \in \mathcal{S}$. It follows from a slight modification of Lemma 4 of Sakamoto and Yoshida (1996) that for fixed measurable f, there exists a sequence $f_\nu \in \mathcal{S}$ such that for some large L, $f_\nu \to f$ in $L^1((1+|z|)^{-L}dz)$, and that for some large m, $f_\nu \to f$ in C_{-2m}. Therefore $I(f_\nu; i, s, \epsilon, n) \to I(f; i, s, \epsilon, n)$ as $\nu \to \infty$ due to (6.14), and hence Inequality (6.13) holds for that measurable function f. In this way, we can obtain (6.9). It is possible to obtain a similar estimate for terms involving b_{i+1}^ϵ. Consequently, following the same procedure as in the smooth case, the proof is finished. $\square$

$D_\infty(\boldsymbol{R}^d) = \cap_{p>1} \cap_{s>0} D_{p,s}(\boldsymbol{R}^d)$, and $D_{p,s}(\boldsymbol{R}^d)$ denotes the Sobolev space of $\boldsymbol{R}^d$-valued Wiener functionals. (See Ikeda and Watanabe (1989) for the details of the Sobolev space $D_{p,s}$.) Here is a simple but useful lemma originated by R. Leandre (cf. Kohatsu-Higa (1997)).

LEMMA 1. *Let $F_{n,\epsilon}^\theta$ and F^θ be in $D_\infty(\boldsymbol{R}^d)$ where θ is a parameter and $(n, \epsilon) \in \boldsymbol{N} \times (0, 1]$. Suppose that for some positive constant ω, $\epsilon = o(n^{-\omega})$ as $n \to \infty$. Suppose also the followings:*

(i) *There exists $\gamma > 0$ such that*

$$\sup_\theta \|F_{n,\epsilon}^\theta - F^\theta\|_{1,p} = O\left(\frac{1}{n^\gamma} + \epsilon^\gamma\right)$$

as $n \to \infty$ and $\epsilon \downarrow 0$ for every $p > 1$.

(ii) *For every $p > 1$,*

$$\sup_\theta \|\det \sigma_{F^\theta}^{-1}\|_p < \infty.$$

(iii) *For every $p > 1$, there exists $c_p > 0$ such that*

$$\sup_{\epsilon' \in (0,1], \theta} \|\det \sigma_{F_{n,\epsilon'}^\theta}^{-1}\|_p = O(n^{c_p}).$$

Then

$$\overline{\lim_{n\to\infty, \epsilon\downarrow 0}} \sup_\theta \|\det \sigma_{F_{n,\epsilon}^\theta}^{-1}\|_p < \infty$$

for every $p > 1$.

PROOF. Set $a = \det \sigma_{F^\theta}$ and $b = \det \sigma_{F^\theta_{n,\epsilon}}$. Then $\boldsymbol{E}[b^{-p}, b < 2^{-1}a] \leq \boldsymbol{E}[b^{-p}, |a-b| > 2^{-1}a] \leq 2^M \boldsymbol{E}[b^{-p}|a-b|^M a^{-M}] \leq$ const. $n^{pc_{3p}}(\frac{1}{n^\gamma} + \epsilon^\gamma)^M$, and take a sufficiently large M. □

Appendix: On the validity of square-root processes in the asymptotic method

Let processes $\{X^\epsilon_t; 0 \leq t \leq T\}$ and $\{\tilde{X}^\epsilon_t; 0 \leq t \leq T\}$ defined as follows:

$$\begin{cases} dX^\epsilon_t = (cX^\epsilon_t + d)dt + \epsilon\sqrt{X^\epsilon_t}dw_t, & X^\epsilon_0 = x_0 \\ d\tilde{X}^\epsilon_t = (c\tilde{X}^\epsilon_t + d)dt + \epsilon g(\tilde{X}^\epsilon_t)dw_t, & \tilde{X}^\epsilon_0 = x_0 \end{cases} \tag{A.1}$$

where $T < \infty$, c, d are some constants with $d \geq 0$, $x_0 > 0$, and $\epsilon \in (0,1]$. $g(x)$ is a smooth modification of $\sqrt{x}$ such that $g(x) = \sqrt{x}$ for $x \geq a'$ where $a' < a$, and $a \equiv \frac{1}{2}\min_{t\in[0,T]} X^0_t$. The process X^ϵ_t is a so called square-root process, and the process $\tilde{X}^\epsilon_t$ is a modified process of X^ϵ_t.

Suppose that for a $\boldsymbol{R}$-valued functional F, $F(X^\epsilon)$ and $F(\tilde{X}^\epsilon)$ are $L_2(P)$-finite. Then, we have

$$\boldsymbol{E}[|F(X^\epsilon) - F(\tilde{X}^\epsilon)|1_{\{X^\epsilon \neq \tilde{X}^\epsilon\}}] \leq (\|F(X^\epsilon)\|_2 + \|F(\tilde{X}^\epsilon)\|_2)P(\{X^\epsilon \neq \tilde{X}^\epsilon\})^{1/2}$$

where $\|\cdot\|_2$ denotes the $L_2(P)$-norm. It also holds that

$$\begin{aligned} &P(\{X^\epsilon \neq \tilde{X}^\epsilon\}) \\ &\quad = P(\{X^\epsilon_t \leq a' \text{ for some } t \in [0,T]\}) \\ &\quad \leq P\left(\left\{\sup_{0\leq t\leq T}|X^\epsilon_t - X^0_t| > a\right\}\right) \\ &\qquad + P\left(\{X^\epsilon_t \leq a' \text{ for some } t \in [0,T]\} \cap \left\{\sup_{0\leq t\leq T}|X^\epsilon_t - X^0_t| \leq a\right\}\right). \end{aligned}$$

We can easily see that the second term after the last inequality is 0. The first term is smaller than any ϵ^n for $n = 1, 2, \cdots$ by the following lemma of a large deviation inequality:

LEMMA 2. *Suppose that Z^ϵ_t, $t \in [0,T]$ follows a SDE*:

$$dZ^\epsilon_t = \mu(Z^\epsilon_t)dt + \epsilon\sigma(Z^\epsilon_t)dw_t,$$

where $\mu(z)$ satisfies Lipschitz and linear growth conditions, *and $\sigma(z)$ satisfies the linear growth condition. We assume that a unique strong solution exists. Then, there exists positive constants a_1 and a_2 independent of ϵ such that*

$$P\left(\left\{\sup_{0\leq s\leq T}|Z^\epsilon_s - Z^0_s| > a\right\}\right) \leq a_1 \exp(-a_2\epsilon^{-2}) \tag{A.2}$$

for all $a > 0$.

The lemma can be proved by slight modification of lemma 5.3 in Yoshida (1992b), or lemma 7.1 in Kunitomo and Takahashi (2003a). Note also that X^ϵ and $\tilde{X}^\epsilon$ satisfy the conditions in Lemma 2.

Hence, if $\|F(X^\epsilon)\|_2 < \infty$ and $\|F(\tilde{X}^\epsilon)\|_2 < \infty$, then

$$\boldsymbol{E}[|F(X^\epsilon) - F(\tilde{X}^\epsilon)|] = o(\epsilon^n), \qquad n = 1, 2, \ldots. \tag{A.3}$$

Therefore, the difference between $F(X^\epsilon)$ and $F(\tilde{X}^\epsilon)$ is negligible in the *small disturbance asymptotic theory*. Finally, we remark that functionals corresponding to F in the examples of Section 4 are $L_2(P)$ bounded, because $F(x) = \gamma(x)$ is bounded in example 1, and for $F(x) = (\frac{1}{T}\int_0^T x_t dt - K)_+$ with $K > 0$ in example 2,

$$\|F(X^\epsilon)\|_2 \le \left\| \frac{1}{T}\int_0^T X_t^\epsilon dt \right\|_2 \le \frac{1}{T}\int_0^T \|X_t^\epsilon\|_2 dt < \infty$$

and

$$\|F(\tilde{X}^\epsilon)\|_2 \le \left\| \frac{1}{T}\int_0^T \tilde{X}_t^\epsilon dt \right\|_2 \le \frac{1}{T}\int_0^T \|\tilde{X}_t^\epsilon\|_2 dt < \infty.$$

References

Bally, V. and Talay, D. (1995). The law of the Euler scheme for stochastic differential equations(I): convergence rate of the distribution function, *Probability Theory and Related Fields*, **102**, 43–60.

Bichteler, K., Gravereaux, J.-B. and Jacod, J. (1987). *Malliavin Calculus for Processes with Jumps*, Gordon and Breach Science Publishers.

Detemple, J. R., Garcia, J. R. and Rindisbacher, M. (2003). A Monte Carlo method for optimal portfolios, *Journal of Finance*, **58**, 401–446.

Glasserman, P. (2003). *Monte Carlo Methods in Financial Engineering*, Springer.

He, H. and Takahashi, A. (2000). A variable reduction technique for pricing average-rate options, *International Review of Finance*, **1**, 123–142.

Heath, D., Jarrow, R. and Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation, *Econometrica*, **60**, 77–105.

Ikeda, N. and Watanabe, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, Second Edition, North-Holland/Kodansha, Tokyo.

Kemna, A. G. Z. and Vorst, A. C. F. (1990). A pricing method for options based on average asset values, *Journal of Banking and Finance*, **14**, 113–129.

Kohatsu-Higa, A. (1997). High order Ito-Taylor approximations to heat kernels, *Journal of Mathematics of Kyoto University*, **37**, 129–151.

Kunitomo, N. and Takahashi, A. (1992). Pricing average options, *Japan Financial Review*, **14**, 1–20 (n Japanese).

Kunitomo, N. and Takahashi, A. (2001). The asymptotic expansion approach to the valuation of interest rate contingent claims, *Mathematical Finance*, **11**, 117–151.

Kunitomo, N. and Takahashi, A. (2003a). On validity of the asymptotic expansion approach in contingent claim analysis, *Annals of Applied Probability*, **13**, No. 3, Aug. 2003.

Kunitomo, N. and Takahashi, A. (2003b). Foundation of mathematical finance—application of Malliavin calculus and asymptotic expansion method—, *Toyo-Keizai* (in Japanese).

Kunitomo, N. and Takahashi, A. (2004). Applications of the asymptotic expansion approach based on Malliavin-Watanabe calculus in financial problems, *Stochastic Processes and Applications to Mathematical Finance*, 195–232, World Scientific.

Milshtein, G. and Schoenmakers, J. (2002). Numerical construction a hedging strategy against the multi-asset European claim, *Stochastics and Stochastics Reports*, **73**, 125–157.

Newton, N. J. (1994). Variance reduction for simulated diffusions, *SIAM Journal on Applied Mathematics*, **54**, 1780–1805.

Ocone, D. and Karatzas, I. (1991). A generalized Clark representation formula, with application to optimal portfolios, *Stochastics and Stochastics Reports*, **34**, 187–220.

Robert, C. P. and Casella, G. (2000). *Monte Carlo Statistical Methods*, Springer.

Sakamoto, Y. and Yoshida, N. (1996). Expansion of perturbed random variables based on generalized Wiener functionals, *Journal of Multivariate Analysis*, **59-1**, 34–59.

Takahashi, A. (1995). Essays on the valuation problems of contingent claims, Unpublished Ph.D. Dissertation, Haas School of Business, University of California, Berkeley.

Takahashi, A. (1999). An asymptotic expansion approach to pricing contingent claims, *Asia-Pacific Financial Markets*, **6**, 115–151.

Takahashi, A. and Saito, T. (2003). An asymptotic expansion approach to pricing American options, *Monetary and Economic Studies*, **22**, 35–87, Bank of Japan (in Japanese).

Takahashi, A. and Yoshida, N. (2004). An asymptotic expansion scheme for optimal investment problems, *Statistical Inference for Stochastic Processes*, **7-2**, 153–188.

Talay, D. and Tubaro, L. (1990). Expansion of the global error for numerical schemes solving stochastic differential equations, *Stochastic Analysis and Applications*, **8**, 483–509.

Yoshida, N. (1992a). Asymptotic expansion for small diffusions via the theory of Malliavin-Watanabe, *Prob. Theory Related Fields*, **92**, 275–311.

Yoshida, N. (1992b). Asymptotic expansions for statistics related to small diffusions, *J. Japan Statist. Soc.*, **22**, 139–159.

Yoshida, N. (1993). Asymptotic expansion of Bayes estimators for small diffusions, *Probab. Theory Relat. Fields*, **95**, 429–450.