Monte Carlo Simulation with Asymptotic Method

Akihiko TAKAHASHI and Nakahiro YOSHIDA

The University of Tokyo,

Frist Version:December 5, 2001,
Second Version:June 30, 2002,
Current Version:April 7, 2003

Abstract

We shall propose a new computational scheme with the asymptotic method to achieve variance reduction of Monte Carlo simulation for numerical analysis especially in 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 applying *Malliavin calculus*.

1 Introduction

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

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

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

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

where $[Z]_j$ $(j = 1, ..., N)$ denote independent copies of the random variable Z. We introduce a modified estimator $\mathbf{V}^*(n, N)$ defined by

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

where $\mathbf{E}[\hat{F}]$ is assumed to be analytically known. Intuitively, if we are able to find $\hat{F}$ such that the errors of $[F]_j$ and $[\hat{F}]_j$, that is, $[F]_j - \mathbf{V}$ and $[\hat{F}]_j - \mathbf{E}[\hat{F}]$ take close numerical values for each independent copy j, then $\mathbf{V}^*(n, N)$ becomes a better estimate since the error of each j in $\mathbf{V}^*(n, N)$ which is represented by the difference of the errors of $[F]_j$ and $[\hat{F}]_j$ becomes small. The asymptotic method provides such $\hat{F}$. That is, $\hat{F}$ obtained by the asymptotic method has a a strong correlation with F, and $\mathbf{E}[\hat{F}]$ is evaluated analytically. In this sense, the method is somewhat similar to *control variate technique*. (See chapter 3 of Robert and Casella(2000) on *control variate technique* for instance.) 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 $\mathbf{E}[\hat{F}]$ can be analytically obtained. Our method overcome this problem since the asymptotic method allow us to find such $\hat{F}$ in the unified way; the method can be applied to a broad class of Ito processes in the similar manner. (For details on this point, see Takahashi(1995,1999), Takahashi and Yoshida(2001), Kunitomo and Takahasi(2001, 2003), and others listed in their references.) We also note that our method may be used together with other accelation methods such as *antithetic variables technique* to pursue further variance reduction of Monte Carlo simulation. In the following sections, we will show this idea more rigorously and concretely.

In the next section, we will explain 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 show the proofs of theorems stated in section 2; we will first provide proofs for smooth cases in section 5, and extend it to non smooth cases in section 6 applying *Malliavin calculus*. In appendix, we will discuss mathematical validity of our asymptotic method with square-root processes used in the numerical examples.

2 Monte Carlo Simulation with the Asymptotic method

Let $(\Omega, \mathcal{F}, P)$ probability space and $T\in (0,\infty)$ denotes some fixed time horizon. $w(t) = \{(w^1(t), \cdots, w^r(t))^*; t \in [0,T]\}$ is $\mathbf{R}^r$-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, $\mathcal{F}_t^w = \sigma(w(s); 0 \le s \le t)$. Suppose that a $\mathbf{R}^D$-valued process $X_u(t,x)$ $(0 \le t \le u \le T, x \in \mathbf{R}^D)$ satisfy the stochastic integral equation:

$$X_u^\epsilon(t,x) \quad = \quad x + \int_t^u V_0(X_s^\epsilon(t,x), \epsilon)ds + \int_t^u V(X_s^\epsilon(t,x), \epsilon)dw_s. \qquad (1)$$

where ϵ is a parameter $\epsilon \in (0,1]$.

We first state the conditions on the process $X_u(t,x)$ $(0 \le t \le u \le T, x \in \mathbf{R}^D)$: We assume that (V_0, V) is *graded* according to $\mathbf{R}^D = \mathbf{R}^{d_1} \times \cdots \times \mathbf{R}^{d_q}$ by using the follwing definition of *grading*:

[Definition: Grading]
A *grading* of $\mathbf{R}^D$ is a decomposition $\mathbf{R}^{d_1} \times \cdots \times \mathbf{R}^{d_q}$ with $\sum_{i=1}^q d_i = D$. The coordinates of a point in $\mathbf{R}^D$ are always arranged in an increasing order along the subspaces $\mathbf{R}^{d_i}$. 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) are *graded* according to the grading $\mathbf{R}^D = \mathbf{R}^{d_1} \times \cdots \times \mathbf{R}^{d_q}$ if $V_0^i(x,\epsilon)$ and $V_\alpha^i(x,\epsilon)$, $\alpha = 1, 2, \cdots, r$ depend on x only through the coordinates $(x_m)_{1\le m\le M_l}$ when $M_{l-1} < i \le M_l$.

In this setting we also assume that $V_\alpha \in C_\uparrow^\infty(\mathbf{R}^D \times (0,1]; \mathbf{R}^D)$, $\alpha = 0, 1, \cdots, r$, where $C_\uparrow^\infty(R^D \times (0,1]; E)$ denotes a class of smooth mappings $f : \mathbf{R}^D \times (0,1] \to E$ whose derivatives $\partial_x^{\mathbf{n}} \partial_\epsilon^k f(x,\epsilon)$ are of polynomial growth orders for $\mathbf{n} \in \mathbf{Z}_+^D$ and $k \in \mathbf{Z}_+$. We further suppose that $\partial_{\hat{x}_l}^{\mathbf{n}} V_\alpha^i(x,\epsilon)$, $\alpha = 0, 1, \cdots, r$ are bounded for $\mathbf{n} \in \mathbf{Z}_+^{d_l}$ such that $|\mathbf{n}| \ge 1$ if $\hat{x}_l$ consists of coordinates from $(M_{l-1}+1)$-th to M_l-th and $M_{l-1} < i \le M_l$ for some $l \le q$.

Then, due to Chapter II-5 of Bichteler etal.(1987), $X_u(t,x)$ $(0 \le t \le u \le T, x \in \mathbf{R}^D)$ admits the unique solution and $\mathbf{E}[|X_u|^p] < \infty$ for all $p \ge 1$. We finally note that the Markovian system of SDEs (15) in section 3 is an example of this class.

2.1 Smooth Cases

Suppose that $f \in C_\uparrow^k(\mathbf{R}^D; \mathbf{R})$ for some large k where $C_\uparrow^k(\mathbf{R}^D; \mathbf{R})$ denotes a class of k times continuously differentiable functions $f : \mathbf{R}^D \to \mathbf{R}$ whose derivatives are of polynomial growth orders. For a stochastic approximation to $\mathbf{V} := \mathbf{E}[f(X_T^\epsilon(0,x))]$, an estimator by naive Monte Carlo simulation may be expressed as

$$\mathbf{V}(n,N) \quad = \quad \frac{1}{N}\sum_{j=1}^{N}\left[f(\bar{X}_T^\epsilon)\right]_j . \tag{2}$$

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

$$\bar{X}_u^\epsilon \quad = \quad x + \int_0^u V_0(\bar{X}_{\eta(s)}^\epsilon, \epsilon)ds + \int_0^u V(\bar{X}_{\eta(s)}^\epsilon, \epsilon)dw_s, \tag{3}$$

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

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

$$\mathbf{V}^*(\epsilon,n,N) \quad = \quad \mathbf{E}[f(X_T^{[0]}(0,x))] + \frac{1}{N}\sum_{j=1}^{N}\left[f(\bar{X}_T^\epsilon) - f(\bar{X}_T^{[0]})\right]_j \tag{4}$$

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

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

Our main objective is to state this intuition more rigorously. Starting with a naive estimator $\mathbf{V}(n,N)$, we have the following theorem:

Theorem 1 *Suppose that $f \in C_\uparrow^k(\mathbf{R}^D;\mathbf{R})$ for some large k. Then,*

(i) the bias of $\mathbf{V}(n,N)$ denoted by $\mathbf{Bias}[\mathbf{V}(n,N)]$ is

$$\mathbf{Bias}[\mathbf{V}(n,N)] = \mathbf{E}[\mathbf{V}(n,N)] - \mathbf{V} \quad = \quad O(\frac{1}{n}),$$

(ii) the variance of $\mathbf{V}(n,N)$ denoted by $\mathbf{Var}[\mathbf{V}(n,N)]$ is

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

and consequently,

(iii) the mean-square-error denoted by $\mathbf{MSE}[\mathbf{V}(n,N)]$ is

$$\mathbf{MSE}[\mathbf{V}(n,N)] := \mathbf{E}[(\mathbf{V}(n,N) - \mathbf{V})^2] \quad = \quad O(\frac{1}{n^2} + \frac{1}{N}).$$

where The bias of an estimator $\hat{Z}$ for Z denoted by $\mathbf{Bias}[\hat{Z}]$ *is defined as*

$$\mathbf{Bias}[\hat{Z}] := \mathbf{E}[\hat{Z}] - Z.$$

Proof: See section 5.1.

For our modified estimator, we can obtain a main theorem:

Theorem 2 *Suppose that $f \in C^k_{\uparrow}(\mathbf{R}^D;\mathbf{R})$ for some large k. Then,*

(i) the bias of $\mathbf{V}^*(\epsilon,n,N)$ is given by

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

(ii) the variance of $\mathbf{V}^*(\epsilon,n,N)$ is given by

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

and consequently,

(iii) the mean-square-error is given by

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

Proof: See section 5.2.

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

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

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

where

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

and

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

Then, we expect that $\theta_2(\epsilon,n,N)$ that is the order of ϵ^2, is small relative to $\frac{1}{N}\mathbf{Var}[f(\bar{X}^{\epsilon}_T)]$ and hence, that MSE of $\mathbf{V}^*(\epsilon,n,N)$ is the smaller than MSE of $\mathbf{V}(n,N)$.

2.2 Non Smooth Cases

If f is not smooth, in particular, f belongs to the class of Borel measurable functions with polynomial growth, we can still obtain the similar results as in the smooth cases under appropriate assumptions.

We will consider a stochastic approximation to $\mathbf{V} := \mathbf{E}[f(X_T^\epsilon(0,x))]$. An estimator may be obtained by a naive Monte Carlo simulation. However, in this case, Malliavin calculus is involved because of non smoothness of f. In order to apply Malliavin calculus effectively, we take a modified Euler-Maruyama scheme following Kohatsu-Higa(1996). That is, we compute

$$\mathbf{V}(n,N) \quad = \quad \frac{1}{N}\sum_{j=1}^{N}\left[f\left(\bar{X}_T^\epsilon + \frac{1}{n}\hat{w}_T\right)\right]_j, \tag{5}$$

instead of (2); $\frac{1}{N}\sum_{j=1}^N \left[f(\bar{X}_T^\epsilon)\right]_j$, where $\{\hat{w}_t; t \in [0,T]\}$ is a Wiener process independent of X^ϵ. Intuitively, randomness of $\hat{w}$ admits non-degeneracy of related Malliavin covariances under appropriate conditions. In fact, we use Malliavin calculus over product measure $P^w \otimes P^{\hat{w}}$.

Similarly, our new estimator (4) is modified as follows:

$$V^*(\epsilon,n,N) \quad = \quad \mathbf{E}[f(X_T^{[0]}(0,x))] + \frac{1}{N}\sum_{j=1}^{N}\left[f\left(\bar{X}_T^\epsilon + \frac{1}{n}\hat{w}_T\right) - f\left(\bar{X}_T^{[0]} + \frac{1}{n}\hat{w}_T\right)\right]. \tag{6}$$

To justifiy this scheme, we first make the following asuumption:

[A2]

$$\sup_\epsilon \mathbf{E}[|\sigma_{X_T^\epsilon(0,x)}|^{-p}] < \infty \text{ for all } p > 1.$$

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

Sometimes, it is difficult to check the condition [A2]. Then in stead of [A2], we may put the following condition [A3] which is practically more convenient.

[A3]

$$\mathbf{E}[|\sigma_{X_T^{[0]}(0,x)}|^{-p}] < \infty \text{ for all } p > 1$$

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

Then, we can obtain the similar results in the non smooth cases corresponding to theorems 1 and 2 in the smooth cases. In particular, the similar result to theorem 2 is obtained under condition [A2] or condition [A3]. The next theorem is a main result.

Theorem 3 *f belongs to the class of Borel measurable functions with polynomial growth. Then,* *(i) the bias of* $\mathbf{V}^*(\epsilon, n, N)$ *is given by*

$$\begin{aligned}
\mathbf{Bias}[\mathbf{V}^*(\epsilon, n, N)] &= O(\frac{\epsilon}{n}), \quad \textit{under [A2]} \\
&\textit{or} \\
\mathbf{Bias}[\mathbf{V}^*(\epsilon, n, N)] &= O\left(\frac{\epsilon}{n}\right) + O(\epsilon^K) \textit{ for all } K > 0 \textit{ under [A3]},
\end{aligned}$$

(ii) the variance of $\mathbf{V}^*(\epsilon, n, N)$ *is given by*

$$\mathbf{Var}[\mathbf{V}^*(\epsilon, n, N)] = \frac{1}{N}\mathbf{Var}[f(\bar{X}_T^\epsilon) - f(\bar{X}_T^0)] = O(\frac{\epsilon^2}{N}), \quad \textit{under [A2] and [A3]},$$

and consequently,

(iii) the mean-square-error is given by

$$\begin{aligned}
\mathbf{MSE}[\mathbf{V}^*(\epsilon, n, N)] &= O(\epsilon^2\left(\frac{1}{n^2} + \frac{1}{N}\right)) \quad \textit{under [A2]} \\
&\textit{or} \\
\mathbf{MSE}[\mathbf{V}^*(\epsilon, n, N)] &= O(\epsilon^2\left(\frac{1}{n^2} + \frac{1}{N}\right)) + O(\epsilon^K) \textit{ for all } K > 0 \textit{ under [A3]}.
\end{aligned}$$

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 the example belongs to *smooth cases* 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. $w(t) = \{(w^1(t), \cdots, w^r(t))^*; t \in [0, T]\}$ is $\mathbf{R}^r$-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 asstes and a locally riskless asset are described as follows.

$$\begin{aligned}
dS_i &= S_i(t)[b_i(t)dt + \sum_{j=1}^{r}\sigma_{ij}(t)dw_j(t)]; \; S_i(0) = s_i \; i = 1, \cdots, r \qquad (7) \\
dS_0 &= \gamma(t)S_0(t)dt; \; S_0(0) = 1
\end{aligned}$$

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 $|b(t)| := (\sum_{i=1}^r |b_i(t)|^2)^{\frac{1}{2}}$ 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 $\theta_i(t)$, $i = 1, 2, \cdots, r$ 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 $\pi(t) = \{\pi_i(t)\}^*_{i=1,\cdots,r}$ denotes the portfolio, which satisfy the integrability condition;

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

Let $\mathcal{A}(W)$ denote the set of stochastic processes (π, c) which generate $W(t) \geq 0$ 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;

$$\sup_{(\pi,c)\in\mathcal{A}(W)} E[U(W(T))] \tag{8}$$

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

$$\begin{aligned}
&U : (0,\infty) \to \mathbf{R}, \\
&\text{a strictly increasing, strictly concave function of class } \mathbf{C}^2 \\
&\text{with } U(0+) \equiv \lim_{c\downarrow 0} U(c) \in [-\infty,\infty),\ U^{'}(0+) \equiv \lim_{c\downarrow 0} U^{'}(c) = \infty \\
&\text{and } U^{'}(\infty) \equiv \lim_{c\to\infty} U^{'}(c) = 0.
\end{aligned} \tag{9}$$

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:

$$dX_u^\epsilon \;=\; V_0(X_u^\epsilon,\epsilon)du + V(X_u^\epsilon,\epsilon)dw_u, \qquad X_t^\epsilon = x \tag{10}$$

for $u \in [t,T]$ where $\epsilon \in (0,1]$, $V_0 \in C_b^\infty(\mathbf{R}^D \times (0,1]; \mathbf{R}^D)$, and $V = (V_\beta)_{\beta=1}^r \in C_b^\infty(\mathbf{R}^D \times (0,1]; \mathbf{R}^D \otimes \mathbf{R}^r)$. Here $C_b^\infty(R^d \times (0,1]; E)$ denotes a class of smooth mappings $f : \mathbf{R}^D \times (0,1] \to E$ whose derivatives $\partial_x^{\mathbf{n}}\partial_\epsilon^m f(x,\epsilon)$ are all bounded for $\mathbf{n} \in \mathbf{Z}_+^d$ such that $|\mathbf{n}| \geq 1$ and $m \in \mathbf{Z}_+$. 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:

$$\begin{cases} dY_{t,u}^{\epsilon} = \sum_{\alpha=0}^{r} \partial_x V_\alpha(X_u^\epsilon, \epsilon) Y_{t,u}^\epsilon dw_u^\alpha \\ \\ Y_{t,t}^\epsilon = \underline{I} \end{cases} \tag{11}$$

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 $\gamma \in C_b^\infty(\mathbf{R}^D; \mathbf{R}_+)$ and $\theta \in C_b^\infty(\mathbf{R}^D; \mathbf{R}^r)$. 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;

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

Then, due to Takahashi and Yoshida(2001) and Ocone and Karatzas(1991), the optimal proportions of risky assets in given wealth W at time t, are given by

$$\begin{aligned} \pi^*(t)/W \quad = \quad & \frac{1}{(1-\delta)}\theta(x)^*\sigma^{-1}(x) + \frac{\delta}{(1-\delta)}\frac{1}{\mathbf{E}\left[(H_{0,t,T})^{(\frac{-\delta}{1-\delta})}\right]} \times \\ & \mathbf{E}\left[(H_{0,t,T})^{(\frac{-\delta}{1-\delta})}\left(\int_t^T \partial\gamma(X_u^\epsilon)Y_{t,u}^\epsilon du + \sum_{\alpha=1}^{r}\int_t^T \partial\theta_\alpha(X_u^\epsilon)Y_{t,u}^\epsilon dw^\alpha(u)\right.\right. \\ & \left.\left. + \sum_{\alpha=1}^{r}\int_t^T \theta_\alpha(X_u^\epsilon)\partial\theta_\alpha(X_u^\epsilon)Y_{t,u}^\epsilon du\right)\right] V(x,\epsilon)\sigma^{-1}(x). \end{aligned} \tag{12}$$

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

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

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:

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

$$\left(\sum_{\alpha=1}^{r}\int_t^T \partial\theta_\alpha(X_u^\epsilon)Y_{t,u}^\epsilon dw^\alpha(u) + \sum_{\alpha=1}^{r}\int_t^T \theta_\alpha(X_u^\epsilon)\partial\theta_\alpha(X_u^\epsilon)Y_{t,u}^\epsilon du\right)\Bigg] V(x,\epsilon)\sigma^{-1}(x). \tag{13}$$

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

[**A1**] the deterministic limit condition: $V(\cdot, 0) \equiv 0$

Under the assumption [**A1**], 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(2001):

$$\text{mean variance} \equiv \frac{1}{(1-\delta)}\theta^*(x)\sigma^{-1}(x) \tag{14}$$

$$\text{IR-hedge} \equiv \epsilon\frac{\delta}{(1-\delta)}\left(\int_t^T \partial\gamma^{[0]}(u)Y_{t,u}du\right)\partial_\epsilon V(x,0)\sigma^{-1}(x)$$

$$\text{MPR-hedge} \equiv \epsilon\frac{\delta}{(1-\delta)^2}\left(\sum_{\alpha=1}^{r}\int_t^T \theta_\alpha^{[0]}(u)\partial\theta_\alpha^{[0]}(u)Y_{t,u}du\right)\partial_\epsilon V(x,0)\sigma^{-1}(x).$$

From now on, we illustrate our scheme by using MPR-hedge component (13). 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:

$$\begin{cases} dX_u^\epsilon = V_0(X_u^\epsilon,\epsilon)du + V(X_u^\epsilon,\epsilon)dw_u, \quad X_t^\epsilon = x \\ dY_{t,u}^\epsilon = \sum_{\alpha=0}^{r}\partial_x V_\alpha(X_u^\epsilon,\epsilon)Y_{t,u}^\epsilon dw_u^\alpha,\ Y_{t,t}^\epsilon = \underline{I} \\ dh_{0,t,u}^\epsilon = h_{0,t,u}^\epsilon[\{(\frac{\delta}{1-\delta})\gamma(X_u^\epsilon) + \frac{\delta}{2(1-\delta)^2}|\theta(X_u^\epsilon)|^2\}du + (\frac{\delta}{1-\delta})\theta(X_u^\epsilon)^* dw(u)], \\ h_{0,t,t}^\epsilon = 1 \\ d\eta_u^\epsilon = \sum_{\alpha=1}^{r}\theta_\alpha(X_u^\epsilon)\partial\theta_\alpha(X_u^\epsilon)Y_{t,u}^\epsilon du + \sum_{\alpha=1}^{r}\partial\theta_\alpha(X_u^\epsilon)Y_{t,u}^\epsilon dw^\alpha(u),\ \eta_t^\epsilon = 0 \end{cases} \tag{15}$$

We note that above system of equations (15) corresponds to the equation (1) in section 2. Then, the estimator for the denominator of MPR-hedge (13);

$$\mathbf{E}\left[(H_{0,t,T})^{(\frac{-\delta}{1-\delta})}\right] = \mathbf{E}\left[h_{0,t,T}^\epsilon\right] \tag{16}$$

by naive Monte Carlo simulation (2) may be expressed as

$$\frac{1}{N}\sum_{j=1}^{N}\left[\bar{h}_{0,t,T}^\epsilon\right]_j. \tag{17}$$

Similarly, the estimator for the numerator of MPR-hedge (13);

$$\mathbf{E}\left[(H_{0,t,T})^{(\frac{-\delta}{1-\delta})}\left(\sum_{\alpha=1}^{r}\int_t^T \partial\theta_\alpha(X_u^\epsilon)Y_{t,u}^\epsilon dw^\alpha(u)+\sum_{\alpha=1}^{r}\int_t^T \theta_\alpha(X_u^\epsilon)\partial\theta_\alpha(X_u^\epsilon)Y_{t,u}^\epsilon du\right)\right] \tag{18}$$

by naive Monte Carlo (2) may be expressed as

$$\frac{1}{N}\sum_{j=1}^{N}\left[\bar{h}_{0,t,T}^\epsilon \times \bar{\eta}_T^\epsilon\right]_j. \tag{19}$$

Next, we consider modified estimators for (16) and (18) in the following. First, we note that

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

where

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

and

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

A modified estimator for the denominator (16) is given by

$$\mathbf{E}[h_{0,t,T}^{[0]}]+\frac{1}{N}\sum_{j=1}^{N}\left\{\{[\bar{h}_{0,t,T}^\epsilon-\bar{h}_{0,t,T}^{[0]}]_j\right\} \tag{20}$$

where

$$\mathbf{E}[h_{0,t,T}^{[0]}] = C,$$

because clearly

$$\mathbf{E}[\xi_T^{[0]}] = 1.$$

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

$$\begin{cases} dh_{0,t,u}^{[0]} = h_{0,t,u}^{[0]}[\{(\frac{\delta}{1-\delta})\gamma_u^{[0]}+\frac{\delta}{2(1-\delta)^2}|\theta_u^{[0]}|^2\}du+(\frac{\delta}{1-\delta})\theta_u^{[0],*}dw(u)], \\ h_{0,t,t}^{[0]} = 1. \end{cases} \tag{21}$$

In the similar way, a modified estimator for the numerator (18) is given by

$$\mathbf{E}[h_{0,t,u}^{[0]}\eta_T^{[0]}]+\frac{1}{N}\sum_{j=1}^{N}\left\{\left[\bar{h}_{0,t,T}^\epsilon\times\bar{\eta}_T^\epsilon-\bar{h}_T^{[0]}\times\bar{\eta}_T^{[0]}\right]_j\right\} \tag{22}$$

where

$$\mathbf{E}[h_{0,t,u}^{[0]}\eta_T^{[0]}] = C\times\left(\frac{1}{1-\delta}\right)\left[\sum_{\alpha=1}^{r}\int_t^T\theta_\alpha^{[0]}(u)\partial\theta_\alpha^{[0]}(u)Y_{t,u}du\right],$$

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

$$d\eta_u^{[0]} = \sum_{\alpha=1}^{r}\theta_\alpha(X_u^{[0]})\partial\theta_\alpha(X_u^{[0]})Y_{t,u}^{[0]}du+\sum_{\alpha=1}^{r}\partial\theta_\alpha(X_u^{[0]})Y_{t,u}^{[0]}dw^\alpha(u),\quad \eta_t^{[0]} = 0 \tag{23}$$

3.2 Example 2:Pricing of an Average Call Option

The second example is pricing an average call option which belongs to *non smooth cases* 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:

$$dS_t^\epsilon = \gamma S_t^\epsilon dt + \epsilon\sigma(S_t^\epsilon, t)dw_t,\ S_0^\epsilon = S_0(>0) \tag{24}$$

where $\epsilon \in (0,1]$, $\sigma \in C_b^\infty(\mathbf{R}_+ \times [0,T]; \mathbf{R}_+)$, 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

$$V(T) = \left(\frac{1}{T}Z_T^\epsilon - K\right)_+. \tag{25}$$

Then, to obtain the price of an average call option at $t = 0$, we evaluate

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

given that

$$\begin{cases} dS_t^\epsilon = \gamma S_t^\epsilon dt + \epsilon\sigma(S_t^\epsilon, t)dw_t,\ S_0^\epsilon = S_0(>0) \\ \\ dZ_t^\epsilon = S_t^\epsilon dt,\ Z_0^\epsilon = 0. \end{cases} \tag{26}$$

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

$$V = e^{-\gamma T}\epsilon\mathbf{E}\left[\left(\frac{1}{T}X_{2T}^\epsilon + y\right)_+\right] \tag{27}$$

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:

$$\begin{cases} dX_{1t}^\epsilon = \gamma X_{1t}^\epsilon dt + \sigma(\epsilon X_{1t}^\epsilon + S_t^{[0]}, t)dw_t,\ X_{10}^\epsilon = 0 \\ \\ dX_{2t}^\epsilon = X_{1t}^\epsilon dt,\ X_{20}^\epsilon = 0 \end{cases} \tag{28}$$

Next, we assume the condition:

$$\Sigma \equiv \int_0^T \frac{1}{T^2}\left[\frac{e^{(T-u)}-1}{\gamma}\right]^2 \sigma^2(S_u^{[0]},u)du > 0. \tag{29}$$

Under the condition (29), 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 (27) is given by

$$e^{-\gamma T}\mathbf{E}\left[\left(\frac{1}{T}X_{2T}^{[0]}+y\right)_+\right]+\frac{1}{N}\sum_{j=1}^{N}\left\{\left[e^{-\gamma T}\left(\frac{1}{T}\bar{X}_{2T}^{\epsilon}+y+\frac{1}{n}\hat{w}_T\right)_+ - e^{-\gamma T}\left(\frac{1}{T}\bar{X}_{2T}^{[0]}+y+\frac{1}{n}\hat{w}_T\right)_+\right]_j\right\} \tag{30}$$

where

$$e^{-\gamma T}\mathbf{E}\left[\left(\frac{1}{T}X_{2T}^{[0]}+y\right)_+\right] = e^{-\gamma T}\left\{y\Phi\left(\frac{y}{\sqrt{\Sigma}}\right)+\Sigma\frac{1}{\sqrt{2\pi\Sigma}}\exp\left(\frac{-y^2}{2\Sigma}\right)\right\}. \tag{31}$$

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

$$\begin{cases} dX_{1t}^{[0]} = \gamma X_{1t}^{[0]}dt+\sigma(S_t^{[0]},t)dw_t,\ X_{10}^{[0]}=0 \\ dX_{2t}^{[0]} = X_{1t}^{[0]}dt,\ X_{20}^{[0]}=0. \end{cases} \tag{32}$$

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

4 Numerical Examples

4.1 Example 1:MPR-hedge

In this example, 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:

$$\begin{cases} dX_u^{\epsilon(1)} = \kappa_1(\bar{X}^{\epsilon(1)}-X_u^{\epsilon(1)})du-\epsilon(X_u^{\epsilon(1)})^{\frac{1}{2}}dw_u; \quad X_0^{\epsilon(1)}=\gamma_0 \\ dX_u^{\epsilon(2)} = \kappa_2(\bar{X}^{\epsilon(2)}-X_u^{\epsilon(2)})du+\epsilon\sigma_2 dw_u; \quad X_0^{\epsilon(2)}=\theta_0 \end{cases} \tag{33}$$

where w denotes one-dimensional Brownian motion.

Remark 2 *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 the smoothing does not make significant differences and the effects are negligible in the small disturbance asymptotic theory. This is also true for Example 2 in the next subsection. See appendix 5.4 for the rigorous argument on this point.*

We also suppose that there exist one risky asset and a locally riskless asset and assume that $\gamma(X_u^\epsilon)$ and $\theta(X_u^\epsilon)$ are smooth modifications of $\hat{\gamma}(X_u^\epsilon)$ and $\hat{\theta}(X_u^\epsilon)$ respectively where $\hat{\gamma}(X_u^\epsilon) = \min\{X_u^{\epsilon(1)}, M_1\}$, $\hat{\theta}(X_u^\epsilon) = \min\{X_u^{\epsilon(2)}, M_2\}$, and M_i,$i = 1, 2$ are some large positive integers. Then, the dynamics of both assets are described by

$$\begin{cases} dS_u^\epsilon = S_u^\epsilon(X_u^{\epsilon(1)} + \sigma X_u^{\epsilon(2)})du + S_u^\epsilon \sigma dw_u, \quad S^\epsilon(0) = s \\ \\ dS_{0u}^\epsilon = S_{0u}^\epsilon X_u^{\epsilon(1)} du, \quad S_0^\epsilon(0) = 1. \end{cases} \tag{34}$$

Further, we set the values of the parameters for X_u^ϵ following Detemple et al.(2000), 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 $\mathbf{n}$ is 365 and the number of trials N is 1,000,000 in each Monte Carlo simulation.

Mean variance, MPR-hedge and IR-hedge components and the sum of them denoted by *total demand* are listed in table 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 can observe that the results of asymptotic method and of Monte carlo simulation are so close for IR-hedge while there is some difference for MPR-hedge, but the difference is small relative to the total demand. We also notice that the second order scheme gives substantial improvement comparing with the first order scheme which is equivalent to the case that we ignore MPR-hedge and IR-hedge components. (Note that $O(1)$ for MPR-hedge and IR-hedge components are *zero*.)

To show that our new method to increase efficiency of Monte Carlo simulations is effective, we take the case that MPR-hedge with $T = 1$, and $R = 0.5$, in which the diviation of the value by the asymptotic method upto ϵ-order relative to 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 naive one for MPR-hedge (13): *hybrid* denotes the modified estimator expressed as the equation (22) divided by (20), that is (22)/(20) while *mc* denotes the naive estimator expressed as the equation (19) divided by (17), that is (19)/(17). In the figure 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 (22) that is, $\frac{1}{N}\sum_{j=1}^{N}\left[\bar{h}_T^{[0]} \times \bar{\eta}_T^{[0]}\right]_j$, and *num-mc* denotes the naive estimator (19). 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*.

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:

$$\begin{cases} dS_t^\epsilon = \gamma S_t^\epsilon dt + \epsilon\sqrt{S_t^\epsilon}dw_t, \ S_0^\epsilon = S_0 \\ \\ dZ_t^\epsilon = S_t^\epsilon dt, \ Z_0^\epsilon = 0 \end{cases} \tag{35}$$

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

$$\begin{cases} dX_{1t}^\epsilon = \gamma X_{1t}^\epsilon dt + \sqrt{\epsilon X_{1t}^\epsilon + e^{\gamma t}S_0}dw_t, \ X_{10}^\epsilon = 0 \\ \\ dX_{2t}^\epsilon = X_{1t}^\epsilon dt, \ X_{20}^\epsilon = 0, \end{cases} \tag{36}$$

and Σ is given by

$$\Sigma = \frac{S_0}{\gamma^3 T^2}(e^{2\gamma T} - 2\gamma e^{\gamma T} - 1). \tag{37}$$

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

$$\begin{cases} dX_{1t}^{[0]} = \gamma X_{1t}^{[0]} dt + e^{\frac{\gamma t}{2}}\sqrt{S_0}dw_t, \ X_{10}^{[0]} = 0 \\ \\ dX_{2t}^{[0]} = X_{1t}^{[0]} dt, \ X_{20}^{[0]} = 0. \end{cases} \tag{38}$$

Table 5 shows parameters' values and computational result in the 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,

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

$\gamma = 0.05(5\%)$, $T = 1.0$(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 (31), 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 (30), *mc* denotes the estimator by naive Monte Carlo for (27), that is

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

and *mc-asymp* denotes the estimator by naive Monte Carlo for (31), 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, figure 4 shows the comparison of the convergence of three estimators, and the same observation also holds in this case as in Figure 3.

5 Proofs of Theorems 1 and 2

5.1 Proof of Theorem 1

We only prove (i). (ii) and (iii) are easy.

Let

$$u_i^\epsilon(x) \quad = \quad \mathbf{E}\left[f(X_T^\epsilon(t_i,x))\right]. \tag{39}$$

where $t_i = iT/n$, $i = 0,1,2,\cdots,n$. Obviously, $u_n^\epsilon(x) = f(x)$, and

$$\begin{aligned}&u_n^\epsilon(\bar{X}^\epsilon_{t_n}) = u_n^\epsilon(\bar{X}^\epsilon_T) = f(\bar{X}^\epsilon_T),\\ &u_0^\epsilon(\bar{X}^\epsilon_{t_0}) = u_0^\epsilon(x) = \mathbf{E}[f(\bar{X}^\epsilon_T(0,x))].\end{aligned}$$

Define Δ_i^ϵ as

$$\Delta_i^\epsilon := \mathbf{E}[u_{i+1}^\epsilon(\bar{X}^\epsilon_{t_{i+1}})] - \mathbf{E}[u_i^\epsilon(\bar{X}^\epsilon_{t_i})]. \tag{40}$$

Then,

$$\mathbf{E}[f(\bar{X}^\epsilon_T)] - \mathbf{E}[f(X^\epsilon_T(0,x))] = \sum_{i=0}^{n-1}\Delta_i^\epsilon.$$

Next, define an operator L^ϵ which maps a function $u(x)$ to a function $L_y^\epsilon u(x)$ by

$$L_y^\epsilon u(x) \quad = \quad \sum_{i=1}^{D} V_0^{(i)}(y,\epsilon)\partial_i u(x) + \frac{1}{2}\sum_{i,j=1}^{D}\sum_{\alpha=1}^{r} V_\alpha^{(i)}V_\alpha^{(j)}(y,\epsilon)\partial_i\partial_j u(x).$$

Similarly, define $\mathcal{L}^\epsilon$ by

$$\mathcal{L}^\epsilon u(x) \quad = \quad \sum_{i=1}^{D} V_0^{(i)}(x,\epsilon)\partial_i u(x) + \frac{1}{2}\sum_{i,j=1}^{D}\sum_{\alpha=1}^{r} V_\alpha^{(i)}V_\alpha^{(j)}(x,\epsilon)\partial_i\partial_j u(x).$$

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

$$\begin{aligned}
\Delta^{\epsilon}_i &= \mathbf{E}\left[u^{\epsilon}_{i+1}\left(\bar{X}^{\epsilon}_{t_{i+1}}\right)\right] - \mathbf{E}\left[u^{\epsilon}_{i+1}\left(X^{\epsilon}_{t_{i+1}}\left(t_i, \bar{X}^{\epsilon}_{t_i}\right)\right)\right] \\
&= \mathbf{E}\Big[\int_{t_i}^{t_{i+1}} L^{\epsilon}_{\bar{X}^{\epsilon}_{t_i}} u^{\epsilon}_{i+1}(\bar{X}^{\epsilon}_t)dt - \int_{t_i}^{t_{i+1}} \mathcal{L}^{\epsilon} u^{\epsilon}_{i+1}(X^{\epsilon}_t(t_i, \bar{X}^{\epsilon}_{t_i}))dt\Big] \\
&= \mathbf{E}\Big[\int_{t_i}^{t_{i+1}} \{\mathcal{L}^{\epsilon} u_{i+1}(\bar{X}^{\epsilon}_{t_i}) - \mathcal{L}^{\epsilon} u^{\epsilon}_{i+1}(X^{\epsilon}_t(t_i, \bar{X}^{\epsilon}_{t_i}))\}dt\Big] \\
&\quad + \mathbf{E}\Big[\int_{t_i}^{t_{i+1}} \{L^{\epsilon}_{\bar{X}^{\epsilon}_{t_i}} u^{\epsilon}_{i+1}(\bar{X}^{\epsilon}_t) - L^{\epsilon}_{\bar{X}^{\epsilon}_{t_i}} u^{\epsilon}_{i+1}(\bar{X}^{\epsilon}_{t_i})\}dt\Big] \\
&= -\int_{t_i}^{t_{i+1}} \mathbf{E}[\mathcal{L}^{\epsilon} u^{\epsilon}_{i+1}(X^{\epsilon}_t(t_i, \bar{X}^{\epsilon}_{t_i})) - \mathcal{L}^{\epsilon} u_{i+1}(\bar{X}^{\epsilon}_{t_i})]dt \\
&\quad + \int_{t_i}^{t_{i+1}} \mathbf{E}[L^{\epsilon}_{\bar{X}^{\epsilon}_{t_i}} u^{\epsilon}_{i+1}(\bar{X}^{\epsilon}_t) - L^{\epsilon}_{\bar{X}^{\epsilon}_{t_i}} u^{\epsilon}_{i+1}(\bar{X}^{\epsilon}_{t_i})]dt
\end{aligned}$$

Hence,

$$\begin{aligned}
\Delta^{\epsilon}_i &= -\int_{t_i}^{t_{i+1}} \int_{t_i}^{t} \mathbf{E}[a^{\epsilon}_{i+1}(X^{\epsilon}_s(t_i, \bar{X}^{\epsilon}_{t_i}))]dsdt \\
&\quad + \int_{t_i}^{t_{i+1}} \int_{t_i}^{t} \mathbf{E}[b^{\epsilon}_{i+1}(\bar{X}_{t_i}; \bar{X}^{\epsilon}_s)]dsdt
\end{aligned} \tag{41}$$

where

$$a^{\epsilon}_{i+1}(x) := \mathcal{L}^{\epsilon}(\mathcal{L}^{\epsilon} u^{\epsilon}_{i+1}(x))$$

and

$$b^{\epsilon}_{i+1}(y; x) := L^{\epsilon}_y(L^{\epsilon}_y u^{\epsilon}_{i+1}(x))(x).$$

In fact, $a^{\epsilon}_{i+1}(x)$ is expressed as

$$\begin{aligned}
a^{\epsilon}_{i+1}(x) &= \sum_{k'=1}^{D} V_0^{(k')}(x, \epsilon)\partial_{k'}\{\sum_{k=1}^{D} V_0^{(k)}(x, \epsilon)\partial_k u^{\epsilon}_{i+1}(x) \\
&\quad + \frac{1}{2}\sum_{k,l=1}^{D}\sum_{\alpha=1}^{r} V_{\alpha}^{(k)}(x, \epsilon)V_{\alpha}^{(l)}(x, \epsilon)\partial_k\partial_l u^{\epsilon}_{i+1}(x)\} \\
&\quad + \frac{1}{2}\sum_{k',l'=1}^{D}\sum_{\alpha=1}^{r} V_{\alpha}^{(k')}(x, \epsilon)V_{\alpha}^{(l')}(x, \epsilon)\partial_{k'}\partial_{l'}\{\sum_{k=1}^{D} V_0^{(k)}(x)\partial_k u^{\epsilon}_{i+1}(x) \\
&\quad + \frac{1}{2}\sum_{k,l=1}^{D}\sum_{\alpha=1}^{r} V_{\alpha}^{(k)}(x, \epsilon)V_{\alpha}^{(l)}(x, \epsilon)\partial_k\partial_l u^{\epsilon}_{i+1}(x)\},
\end{aligned} \tag{42}$$

and $a^{\epsilon}_{i+1}(X^{\epsilon}_s(t_i, \bar{X}^{\epsilon}_{t_i}))$ means that $a^{\epsilon}_{i+1}(x)$ is evaluated at $x = X^{\epsilon}_s(t_i, \bar{X}^{\epsilon}_{t_i})$. $b^{\epsilon}_{i+1}(y; x)$ is expressed as

$$b^{\epsilon}_{i+1}(y; x) = \sum_{k'=1}^{D} V_0^{(k')}(y, \epsilon)\{\sum_{k=1}^{D} V_0^{(k)}(y, \epsilon)\partial_{k'}\partial_k u^{\epsilon}_{i+1}(x) \tag{43}$$

$$+\frac{1}{2}\sum_{k,l=1}^{D}\sum_{\alpha=1}^{r}V_\alpha^{(k)}(y,\epsilon)V_\alpha^{(l)}(y,\epsilon)\partial_{k'}\partial_k\partial_l u_{i+1}^\epsilon(x)\}$$
$$+\frac{1}{2}\sum_{k',l'=1}^{D}\sum_{\alpha=1}^{r}V_\alpha^{(k')}(x)V_\alpha^{(l')}(x)\{\sum_{k=1}^{D}V_0^{(k)}(y,\epsilon)\partial_{k'}\partial_{l'}\partial_k u_{i+1}^\epsilon(x)$$
$$+\frac{1}{2}\sum_{k,l=1}^{D}\sum_{\alpha=1}^{r}V_\alpha^{(k)}(y,\epsilon)V_\alpha^{(l)}(y,\epsilon)\partial_{k'}\partial_{l'}\partial_k\partial_l u_{i+1}^\epsilon(x)\},$$

and $b_{i+1}^\epsilon(\bar{X}_{t_{i+1}};\bar{X}_s^\epsilon)$ means that $b_{i+1}^\epsilon(y;x)$ is evaluated at $x=\bar{X}_s^\epsilon$ and $y=\bar{X}_{t_{i+1}}^\epsilon$.

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

$$\begin{array}{l} V_0^{(k_1)}, \partial_{k_2}V_0^{(k_1)}, \partial_{k_2}\partial_{l_2}V_0^{(k_1)}, \\ V_\alpha^{(k_2)}, \partial_{k_2}V_\alpha^{(k_1)}, \partial_{k_2}\partial_{l_2}V_\alpha^{(k_1)}, \\ u_{i+1}^\epsilon, \partial_{k_1}u_{i+1}^\epsilon, \partial_{k_1}\partial_{k_2}u_{i+1}^\epsilon, \partial_{k_1}\partial_{k_2}\partial_l u_{i+1}^\epsilon, \text{ and } \partial_{k_1}\partial_{k_2}\partial_{l_1}\partial_{l_2}u_{i+1}^\epsilon \end{array}$$

for $k_1,k_2,l_1,l_2=1,2,\cdots,D$ and $\alpha=1,2,\cdots,r$.

Note also that $V_\alpha(x)\in C_\uparrow^\infty(\mathbf{R}^D)$, $\alpha=0,1,\cdots,r$ and $f\in C_\uparrow^k(\mathbf{R}^D)$ for some large k.

Further, it is well known (see Chapter II-5 of Bichteler etal.(1987), for instance.) that

$$\begin{cases} \sup_\epsilon \sup_n \sup_{0\le s\le T}\mathbf{E}[|\bar{X}_s^\epsilon|^p]<\infty \\ \\ \sup_\epsilon \sup_n \sup_{t_i\le s\le t_{i+1}}\mathbf{E}[|X_s^\epsilon(t_i,\bar{X}_{t_i}^\epsilon)|^p]<\infty \end{cases} \tag{44}$$

for all $p\ge 1$.

Then, by using Hölder inequaility,

$$\sup_\epsilon\sup_n\sup_{i\in\{1,2,\cdots,n\}}\sup_{0\le s\le T}\mathbf{E}[|a_{i+1}^\epsilon(X_s^\epsilon(t_i,\bar{X}_{t_i}^\epsilon))|]<\infty. \tag{45}$$

Similarly,

$$\sup_\epsilon\sup_n\sup_{i\in\{1,2,\cdots,n\}}\sup_{0\le s\le T}\mathbf{E}[|b_{i+1}^\epsilon(\bar{X}_{t_{i+1}};\bar{X}_s^\epsilon)|]<\infty. \tag{46}$$

Thus, we conclude that

$$\begin{aligned} &\mathbf{E}[f(\bar{X}_T^\epsilon)]-\mathbf{E}[f(X_T^\epsilon(0,x))]=\sum_{i=0}^{n-1}\Delta_i^\epsilon \qquad (47)\\ &=\sum_{i=0}^{n-1}\{-\int_{t_i}^{t_{i+1}}\int_{t_i}^{t}\mathbf{E}[a_{i+1}^\epsilon(X_s^\epsilon(t_i,\bar{X}_{t_i}^\epsilon))]dsdt+\int_{t_i}^{t_{i+1}}\int_{t_i}^{t}\mathbf{E}[b_{i+1}^\epsilon(\bar{X}_{t_{i+1}};\bar{X}_s^\epsilon)]dsdt\}\\ &=O\left(\frac{1}{n}\right). \end{aligned}$$

In other words, we proved (i).

5.2 Proof of Theorem 2

We prove only (i), again.

First, we claim that

$$\sup_{s,i,n}\left|\mathbf{E}\left[a^{\epsilon}_{i+1}(X^{\epsilon}_s(t_i,\bar{X}^{\epsilon}_{t_i}))\right]-\mathbf{E}\left[a^{0}_{i+1}(X^{0}_s(t_i,\bar{X}^{0}_{t_i}))\right]\right| \quad = \quad O(\epsilon) \quad (\epsilon\downarrow 0) \tag{48}$$

and that

$$\sup_{s,i,n}\left|\mathbf{E}\left[b^{\epsilon}_{i+1}(\bar{X}^{\epsilon}_{t_i};\bar{X}^{\epsilon}_s)\right]-\mathbf{E}\left[b^{0}_{i+1}(\bar{X}^{0}_{t_i};\bar{X}^{0}_s)\right]\right| \quad = \quad O(\epsilon) \quad (\epsilon\downarrow 0). \tag{49}$$

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

We need to show that

$$\overline{\lim}_{\epsilon\downarrow 0}\frac{1}{\epsilon}\sup_{n}\sup_{i\in\{1,2,\cdots,n\}}\sup_{0\le s\le T}\left|\mathbf{E}\left[a^{\epsilon}_{i+1}(X^{\epsilon}_s(t_i,\bar{X}^{\epsilon}_{t_i}))\right]-\mathbf{E}\left[a^{0}_{i+1}(X^{0}_s(t_i,\bar{X}^{0}_{t_i}))\right]\right|<\infty. \tag{50}$$

Notice that

$$a^{\epsilon}_{i+1}(X^{\epsilon}_s(t_i,\bar{X}^{\epsilon}_{t_i}))=a^{0}_{i+1}(X^{0}_s(t_i,\bar{X}^{0}_{t_i}))+\epsilon\int_0^1\partial_{\epsilon}|_{\epsilon=u\epsilon}a^{\epsilon}_{i+1}(X^{\epsilon}_s(t_i,\bar{X}^{\epsilon}_{t_i}))du.$$

where

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

Then,

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

where $\|\cdot\|_1$ denotes $L_1(P)$-norm.

Note that $\partial_{\epsilon_1}a^{\epsilon_1}_{i+1}(X^{\epsilon_1}_s(t_i,\bar{X}^{\epsilon_1}_{t_i}))$, $0<\epsilon_1<\epsilon$ is a polynomial function of

$$\begin{aligned}
&\frac{\partial X^{\epsilon_1}_s(t_i,\bar{X}^{\epsilon_1}_{t_i})}{\partial\epsilon_1},\\
&V_0^{(k_1)},\partial_{k_2}V_0^{(k_1)},\partial_{k_2}\partial_{l_2}V_0^{(k_1)},\partial_{k_1}\partial_{k_2}\partial_{l_2}V_0^{(k_1)},\\
&V_\alpha^{(k_2)},\partial_{k_2}V_\alpha^{(k_1)},\partial_{k_2}\partial_{l_2}V_\alpha^{(k_1)},\partial_{k_1}\partial_{k_2}\partial_{l_2}V_\alpha^{(k_1)},\\
&u^{\epsilon_1}_{i+1},\partial_{k_1}u^{\epsilon_1}_{i+1},\partial_{k_1}\partial_{k_2}u^{\epsilon_1}_{i+1},\partial_{k_1}\partial_{k_2}\partial_{l}u^{\epsilon_1}_{i+1},\ \text{and}\ \partial_{k_1}\partial_{k_2}\partial_{l_1}\partial_{l_2}\partial_{m}u^{\epsilon_1}_{i+1}
\end{aligned}$$

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

Applying the similar argument in Chapter II-5 of Bichteler etal.(1987) to the system of the equations;

$$\begin{cases} \bar{X}_s^{\epsilon_1} = x + \int_0^s V_0(\bar{X}_{\eta(u)}^{\epsilon_1}, \epsilon_1) du + \int_0^s V(\bar{X}_{\eta(u)}^{\epsilon_1}, \epsilon_1) dw_u, \ s \in [0, T], \\ \\ X_s^{\epsilon_1}(t_i, \bar{X}_{t_i}^{\epsilon_1}) = \bar{X}_{t_i}^{\epsilon_1} + \int_{t_i}^s V_0(X_s^{\epsilon_1}(t_i, \bar{X}_{t_i}^{\epsilon_1}), \epsilon_1) du + \int_{t_i}^s V(X_s^{\epsilon_1}(t_i, \bar{X}_{t_i}^{\epsilon_1}), \epsilon_1) dw_u, \ s \in [t_i, t_{i+1}), \\ \\ \frac{\partial \bar{X}_s^{\epsilon_1}}{\partial \epsilon_1} = \left\{ \int_0^s \partial_{\epsilon_1} V_0(\bar{X}_{\eta(u)}^{\epsilon_1}, \epsilon_1) du + \int_0^s \partial_{\epsilon_1} V(\bar{X}_{\eta(u)}^{\epsilon_1}, \epsilon_1) dw_u \right\} \\ + \int_0^s \partial V_0(\bar{X}_{\eta(u)}, \epsilon_1) \left\{ \frac{\partial \bar{X}_{\eta(u)}^{\epsilon_1}}{\partial \epsilon_1} \right\} du + \sum_{\alpha=1}^r \int_0^s \partial V_\alpha(\bar{X}_{\eta(u)}^{\epsilon_1}, \epsilon_1) \left\{ \frac{\partial \bar{X}_{\eta(u)}^{\epsilon_1}}{\partial \epsilon_1} \right\} dw_u^\alpha, \\ \\ \frac{\partial X_s^{\epsilon_1}(t_i, \bar{X}_{t_i}^{\epsilon_1})}{\partial \epsilon_1} = \frac{\partial \bar{X}_{t_i}^{\epsilon_1}}{\partial \epsilon_1} + \left\{ \int_{t_i}^s \partial_{\epsilon_1} V_0(X_u^{\epsilon_1}(t_i, \bar{X}_{t_i}^{\epsilon_1})) du + \int_{t_i}^s \partial_{\epsilon_1} V(X_u^{\epsilon_1}(t_i, \bar{X}_{t_i}^{\epsilon_1}), \epsilon_1) dw_u \right\} \\ + \int_{t_i}^s \partial V_0(X_u^{\epsilon_1}(t_i, \bar{X}_{t_i}^{\epsilon_1})) \left\{ \frac{\partial X_u^{\epsilon_1}(t_i, \bar{X}_{t_i}^{\epsilon_1})}{\partial \epsilon_1} \right\} du \\ + \sum_{\alpha=1}^r \int_{t_i}^s \partial V_\alpha(X_u^{\epsilon_1}(t_i, \bar{X}_{t_i}^{\epsilon_1}), \epsilon_1) \left\{ \frac{\partial X_u^{\epsilon_1}(t_i, \bar{X}_{t_i}^{\epsilon_1})}{\partial \epsilon_1} \right\} dw_u^\alpha, \ s \in [t_i, t_{i+1}), \end{cases} \tag{51}$$

where ∂V_α, $\alpha = 0, 1, \cdots, r$ denotes the partial derivative with respect to the first argument, we can also show that

$$\begin{cases} \sup_n \sup_{0 \le s \le T} \sup_{0 < \epsilon_1 < \epsilon} \mathbf{E}[|\bar{X}_s^{\epsilon_1}|^p] < \infty, \\ \\ \sup_n \sup_{i \in \{1,2,\cdots,n\}} \sup_{t_i \le s \le t_{i+1}} \sup_{0 < \epsilon_1 < \epsilon} \mathbf{E}[|\bar{X}_s^{\epsilon_1}(t_i, \bar{X}_{t_i}^{\epsilon_1})|^p] < \infty \\ \\ \sup_n \sup_{0 \le s \le T} \sup_{0 < \epsilon_1 < \epsilon} \mathbf{E}\left[\left| \frac{\partial \bar{X}_s^{\epsilon_1}}{\partial \epsilon_1} \right|^p \right] < \infty \\ \\ \sup_n \sup_{i \in \{1,2,\cdots,n\}} \sup_{t_i \le s \le t_{i+1}} \sup_{0 < \epsilon_1 < \epsilon} \mathbf{E}\left[\left| \frac{\partial X_s^{\epsilon_1}(t_i, \bar{X}_{t_i}^{\epsilon_1})}{\partial \epsilon_1} \right|^p \right] < \infty \end{cases} \tag{52}$$

for all $p \ge 1$.

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

We return to prove (i).

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

Hence, using above, we conclude that

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

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 [A2] for a sufficiently large integer m, we have

$$\begin{aligned}&\mathbf{E}[f(X_T^\epsilon(0,x)+\frac{1}{n}\hat{w}_T)] - \mathbf{E}[f(X_T^\epsilon(0,x))] = \\ &\mathbf{E}[(A^{-m}f)(X_T^\epsilon(0,x)+\frac{1}{n}\hat{w}_T)\Psi_1^{(\epsilon)}] - \mathbf{E}[(A^{-m}f)(X_T^\epsilon(0,x))\Psi_2^{(\epsilon)}] = O(\frac{1}{n})\end{aligned} \tag{53}$$

for some Wiener functionals $\Psi_1^{(\epsilon)}$ and $\Psi_2^{(\epsilon)}$. The last equality holds because the differences between $(A^{-m}f)(X_T^\epsilon(0,x)+\frac{1}{n}\hat{w}_T)$ and $(A^{-m}f)(X_T^\epsilon(0,x))$, and between $\Psi_1^{(\epsilon)}$ and $\Psi_2^{(\epsilon)}$ are $O(\frac{1}{n})$.

Under [A2], we can also obtain

$$\begin{aligned}&\{\mathbf{E}[f(X_T^\epsilon(0,x)+\frac{1}{n}\hat{w}_T)] - \mathbf{E}[f(X_T^\epsilon(0,x)]\} \\ &-\{\mathbf{E}[f(X_T^{[0]}(0,x)+\frac{1}{n}\hat{w}_T)] - \mathbf{E}[f(X_T^{[0]}(0,x)]\} \\ &= O\left(\frac{\epsilon}{n}\right).\end{aligned} \tag{54}$$

Then, under [A3],

$$\begin{aligned}&\mathbf{E}[f(X_T^\epsilon(0,x)+\frac{1}{n}\hat{w}_T)] - \mathbf{E}[f(X_T^\epsilon(0,x))] \\ &= \left(\mathbf{E}\left[f(X_T^\epsilon(0,x)+\frac{1}{n}\hat{w}_T)\psi\left(|\sigma_{X_T^{[0]}(0,x)}|/|4\sigma_{X_T^\epsilon(0,x)}|\right)\right] - \mathbf{E}\left[f(X_T^\epsilon(0,x))\psi\left(|\sigma_{X_T^{[0]}(0,x)}|/|4\sigma_{X_T^\epsilon(0,x)}|\right)\right]\right) \\ &+\left(\mathbf{E}\left[f(X_T^\epsilon(0,x)+\frac{1}{n}\hat{w}_T)\left\{1-\psi\left(|\sigma_{X_T^{[0]}(0,x)}|/|4\sigma_{X_T^\epsilon(0,x)}|\right)\right\}\right]\right. \\ &\left.-\mathbf{E}\left[f(X_T^\epsilon(0,x))\left\{1-\psi\left(|\sigma_{X_T^{[0]}(0,x)}|/|4\sigma_{X_T^\epsilon(0,x)}|\right)\right\}\right]\right)\end{aligned} \tag{55}$$

where $\psi(x)$ is a $\psi : \mathbf{R} \to \mathbf{R}$ smooth function such that

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

The first parenthesis after the equality in the equation (55) is $O(\frac{1}{n})$ as in the equation (53). For the second parenthesis,

$$\mathbf{E}\left[f(X_T^\epsilon(0,x)+\frac{1}{n}\hat{w}_T)\left\{1-\psi\left(|\sigma_{X_T^{[0]}(0,x)}|/|4\sigma_{X_T^\epsilon(0,x)}|\right)\right\}\right]$$

$$
\begin{aligned}
&-\mathbf{E}\left[f(X_T^\epsilon(0,x))\left\{1-\psi\left(|\sigma_{X_T^{[0]}(0,x)}|/|4\sigma_{X_T^\epsilon(0,x)}|\right)\right\}\right] \\
&\leq C\left\|1-\psi\left(|\sigma_{X_T^{[0]}(0,x)}|/|4\sigma_{X_T^\epsilon(0,x)}|\right)\right\|_q \quad \text{(by Hölder inequality)} \\
&\leq C\times P(\{|\sigma_{X_T^{[0]}(0,x)}|/|\sigma_{X_T^\epsilon(0,x)}|>2\})^{\frac{1}{q}} \\
&\leq C\times 2^K\mathbf{E}\left[\left(\frac{|\sigma_{X_T^\epsilon(0,x)}-\sigma_{X_T^{[0]}(0,x)}|}{|\sigma_{X_T^{[0]}(0,x)}|}\right)^K\right] \quad \text{(by Markov inequality)} \\
&= O(\epsilon^K) \text{ for all } K>0 \qquad (56)
\end{aligned}
$$

for some positive costant C and $q>1$ where $\|\cdot\|_q$ denotes $L^q(P^w\otimes P^{\hat{w}})$-norm.

Hence under [A3], in stead of the equation (53) we have

$$
\mathbf{E}[f(X_T^\epsilon(0,x)+\frac{1}{n}\hat{w}_T)]-\mathbf{E}[f(X_T^\epsilon(0,x))]=O(\frac{1}{n})+O(\epsilon^K) \text{ for all } K>0. \qquad (57)
$$

Similarly, in stead of the equation (54), we have

$$
\begin{aligned}
&\{\mathbf{E}[f(X_T^\epsilon(0,x)+\frac{1}{n}w_T)]-\mathbf{E}[f(X_T^\epsilon(0,x)]\} \\
&-\{\mathbf{E}[f(X_T^{[0]}(0,x)+\frac{1}{n}w_T)]-\mathbf{E}[f(X_T^{[0]}(0,x)]\} \\
&=O(\frac{\epsilon}{n})+O(\epsilon^K) \text{ for all } K>0 \qquad (58)
\end{aligned}
$$

Next, we note that the *Bias* of $\mathbf{V}(n,N)$ is expressed as

$$
\begin{aligned}
\mathbf{Bias}[\mathbf{V}(n,N)] &= \mathbf{E}[\mathbf{V}(n,N)]-\mathbf{V} \\
&= \mathbf{E}\left[f\left(\bar{X}_T^\epsilon+\frac{1}{n}\hat{w}_T\right)\right]-\mathbf{E}[f(X_T^\epsilon(0,x)] \\
&= \left\{\mathbf{E}\left[f\left(\bar{X}_T^\epsilon+\frac{1}{n}\hat{w}_T\right)\right]-\mathbf{E}\left[f\left(X_T^\epsilon(0,x)+\frac{1}{n}\hat{w}_T\right)\right]\right\} \\
&+ \left\{\mathbf{E}\left[f\left(X_T^\epsilon(0,x)+\frac{1}{n}\hat{w}_T\right)\right]-\mathbf{E}[f(X_T^\epsilon(0,x)]\right\}. \qquad (59)
\end{aligned}
$$

Because after the last equality in this equation, the second term is $O\left(\frac{1}{n}\right)$ under [A2], or $O(\frac{1}{n})+O(\epsilon^K)$ for all $K>0$ under [A3], if the first term is $O\left(\frac{1}{n}\right)$, $\mathbf{Bias}[\mathbf{V}(n,N)]$ is $O\left(\frac{1}{n}\right)$ under [A2], or $O(\frac{1}{n})+O(\epsilon^K)$ for all $K>0$ under [A3].

Similarly, the *Bias* of $\mathbf{V}^*(\epsilon,n,N)$ is expressed as

$$
\begin{aligned}
\mathbf{Bias}[\mathbf{V}^*(\epsilon,n,N)] &= \mathbf{E}[\mathbf{V}^*(\epsilon,n,N)]-\mathbf{V} \\
&= \left\{\mathbf{E}\left[f\left(\bar{X}_T^\epsilon+\frac{1}{n}\hat{w}_T\right)\right]-\mathbf{E}\left[f\left(\bar{X}_T^{[0]}+\frac{1}{n}\hat{w}_T\right)\right]\right\} \\
&- \{\mathbf{E}[f(X_T^\epsilon(0,x)]-\mathbf{E}[f(X_T^{[0]}(0,x)]\}. \qquad (60)
\end{aligned}
$$

Note that the first term in the equation after the last equality is rewritten as follows:

$$
\begin{aligned}
&\left\{\mathbf{E}\left[f\left(\bar{X}^{\epsilon}_{T}+\frac{1}{n}\hat{w}_{T}\right)\right]-\mathbf{E}\left[f\left(\bar{X}^{[0]}_{T}+\frac{1}{n}\hat{w}_{T}\right)\right]\right\}= \\
&\left\{\mathbf{E}\left[f\left(\bar{X}^{\epsilon}_{T}+\frac{1}{n}\hat{w}_{T}\right)\right]-\mathbf{E}\left[f\left(X^{\epsilon}_{T}(0,x)+\frac{1}{n}\hat{w}_{T}\right)\right]\right\} \\
&-\left\{\mathbf{E}\left[f\left(\bar{X}^{[0]}_{T}+\frac{1}{n}\hat{w}_{T}\right)\right]-\mathbf{E}\left[f\left(X^{[0]}_{T}(0,x)+\frac{1}{n}\hat{w}_{T}\right)\right]\right\} \\
&+\mathbf{E}\left[f\left(X^{\epsilon}_{T}(0,x)+\frac{1}{n}\hat{w}_{T}\right)\right]-\mathbf{E}\left[f\left(X^{[0]}_{T}(0,x)+\frac{1}{n}\hat{w}_{T}\right)\right].
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\mathbf{Bias}[\mathbf{V}^{*}(\epsilon,n,N)] &= \left[\left\{\mathbf{E}\left[f\left(\bar{X}^{\epsilon}_{T}+\frac{1}{n}\hat{w}_{T}\right)\right]-\mathbf{E}\left[f\left(X^{\epsilon}_{T}(0,x)+\frac{1}{n}\hat{w}_{T}\right)\right]\right\}\right. \\
&- \left.\left\{\mathbf{E}\left[f\left(\bar{X}^{[0]}_{T}+\frac{1}{n}\hat{w}_{T}\right)\right]-\mathbf{E}\left[f\left(X^{[0]}_{T}(0,x)+\frac{1}{n}\hat{w}_{T}\right)\right]\right\}\right] \\
&+ \left[\left\{\mathbf{E}\left[f\left(X^{\epsilon}_{T}(0,x)+\frac{1}{n}\hat{w}_{T}\right)\right]-\mathbf{E}[f(X^{\epsilon}_{T}(0,x)]\right\}\right. \\
&- \left.\left\{\mathbf{E}\left[f\left(X^{[0]}_{T}(0,x)+\frac{1}{n}\hat{w}_{T}\right)\right]-\mathbf{E}[f(X^{[0]}_{T}(0,x)]\right\}\right]. \qquad (61)
\end{aligned}
$$

Because the second square bracket in the last equation is $O\left(\frac{\epsilon}{n}\right)$ under the condition [A2], if the first square bracket is $O\left(\frac{\epsilon}{n}\right)$, $\mathbf{Bias}[\mathbf{V}^{*}(\epsilon,n,N)]$ is $O\left(\frac{\epsilon}{n}\right)$ under [A2]. Similarly, because under the condition [A3], the second square bracket is $O\left(\frac{\epsilon}{n}\right)+O(\epsilon^{K})$ for all $K>0$, if the first square bracket is $O\left(\frac{\epsilon}{n}\right)$, $\mathbf{Bias}[\mathbf{V}^{*}(\epsilon,n,N)]$ is $O\left(\frac{\epsilon}{n}\right)+O(\epsilon^{K})$ for all $K>0$ under [A3].

In the sequel, we will evaluate

$$
\mathbf{E}\left[f\left(\bar{X}^{\epsilon}_{T}+\frac{1}{n}\hat{w}_{T}\right)\right]-\mathbf{E}\left[f\left(X^{\epsilon}_{T}(0,x)+\frac{1}{n}\hat{w}_{T}\right)\right] \qquad (62)
$$

in order to show that $\mathbf{Bias}[\mathbf{V}(n,N)]$ is $O\left(\frac{1}{n}\right)$.

First, define

$$
u^{\epsilon}_{i}(x):=\mathbf{E}\left[f\left(X^{\epsilon}_{T}(t_{i},x)+\frac{1}{n}\hat{w}_{T}\right)\right]. \qquad (63)
$$

Then, we can write

$$
\mathbf{E}[f(\bar{X}^{\epsilon}_{T}+\frac{1}{n}\hat{w}_{T})]-\mathbf{E}[f(X^{\epsilon}_{T}(0,x)+\frac{1}{n}\hat{w}_{T})]=\sum_{i=0}^{n-1}\Delta^{\epsilon}_{i}
$$

where

$$
\Delta^{\epsilon}_{i}:=\mathbf{E}[u^{\epsilon}_{i+1}(\bar{X}^{\epsilon}_{t_{i+1}})]-\mathbf{E}[u^{\epsilon}_{i}(\bar{X}^{\epsilon}_{t_{i}})] \qquad (64)
$$

Then,

$$
\begin{aligned}
&\mathbf{E}[u^{\epsilon}_{i}(\bar{X}^{\epsilon}_{t_{i}})]=\mathbf{E}[f(X^{\epsilon}_{T}(t_{i},\bar{X}^{\epsilon}_{t_{i}})+\frac{1}{n}\hat{w}_{T})] \\
&=\mathbf{E}\left[f\left(X^{\epsilon}_{T}(t_{i+1},X^{\epsilon}_{t_{i+1}}(t_{i},\bar{X}_{t_{i}}))+\frac{1}{n}\hat{w}_{T}\right)\right].
\end{aligned}
$$

Hence,

$$\mathbf{E}[u_i^\epsilon(\bar{X}_{t_i}^\epsilon)] = \mathbf{E}[u_{i+1}^\epsilon(X_{t_{i+1}}^\epsilon(t_i, \bar{X}_{t_i}^\epsilon))]. \tag{65}$$

Then, Δ_i^ϵ is formally expressed in the same manner as in the smooth cases(that is, $f \in C_\uparrow^k(\mathbf{R}^D)$): a_{i+1}^ϵ and b_{i+1}^ϵ are defined as the equations (42) and (43) respectively which includes partial derivatives of $u_{i+1}^\epsilon(x)$ with respect to x. These derivatives are justified in the sense of Malliavin based on the uniform non-degeneracies of related Malliavin covariances; for instance, the uniform non-degeneracy of the Malliavin covariance of $X_T^\epsilon(t_{i+1}, X_t^\epsilon(t_i, \bar{X}_{t_i}^\epsilon)) + \frac{1}{n}\hat{w}_T$, $t \in [t_i, t_{i+1})$ denoted by $\sigma\left(X_T^\epsilon(t_{i+1}, X_t^\epsilon(t_i, \bar{X}_{t_i}^\epsilon)) + \frac{1}{n}\hat{w}_T\right)$. That is,

$$\sup_{t,\epsilon,n} \mathbf{E}\left[\left|\sigma\left(X_T^\epsilon(t_{i+1}, X_t^\epsilon(t_i, \bar{X}_{t_i}^\epsilon)) + \frac{1}{n}\hat{w}_T\right)\right|^{-p}\right] < \infty, \tag{66}$$

and the uniform non-degeneracy of the Malliavin covariance of $X_T^\epsilon(t_{i+1}, \bar{X}_t^\epsilon + \frac{1}{n}\hat{w}_T)$, $t \in [t_i, t_{i+1})$ denoted by $\sigma\left(X_T^\epsilon(t_{i+1}, \bar{X}_t^\epsilon) + \frac{1}{n}\hat{w}_T\right)$. That is,

$$\sup_{t,\epsilon,n} \mathbf{E}\left[\left|\sigma\left(X_T^\epsilon(t_{i+1}, \bar{X}_t^\epsilon) + \frac{1}{n}\hat{w}_T\right)\right|^{-p}\right] < \infty, \tag{67}$$

Those uniform non-degeneracies are guranteed by the following lemma 1 of Kohatsu-Higa(1996):

Lemma 1 *(Kohatsu-Higa)*
Let $\{F_n^u\}_n$ and F^u random variables in D_∞ where u is a parameter, $D_\infty = \cap_{p>1} \cap_{s>0} D_{p,s}$, and $D_{p,s}$ denotes the Sobolev space of Wiener functionals. (See Ikeda and Watanabe(1989) for the details of the Sobolev space $D_{p,s}$.) Suppose also the followings:
i) there exists $\gamma > 0$ such that

$$\sup_u \|F_n^u - F^u\|_{1,p} = O(\frac{1}{n^\gamma}) \text{ for all } p > 1.$$

ii)

$$\sup_u \||\sigma_{F^u}|^{-1}\|_p < \infty \text{ for all } p > 1.$$

iii) for all $p > 1$, there exists $\nu(p) > 0$ such that

$$\sup_u \||\sigma_{F_n^u}|^{-1}\|_p = O(n^{\nu(p)}).$$

Then,

$$\sup_n \sup_u \||\sigma_{F_n^u}|^{-1}\|_p < \infty \text{ for all } p > 1.$$

Consequently, following the same procedure as in the smooth cases, we can obtain that

$$\mathbf{E}\left[f\left(\bar{X}_T^\epsilon + \frac{1}{n}\hat{w}_T\right)\right] - \mathbf{E}\left[f\left(X_T^\epsilon(0,x) + \frac{1}{n}\hat{w}_T\right)\right] = O\left(\frac{1}{n}\right). \tag{68}$$

Therefore,

$$\mathbf{Bias}[\mathbf{V}(n,N)] = O\left(\frac{1}{n}\right). \tag{69}$$

In the similar manner, we can show that

$$\begin{aligned}&\left[\left\{\mathbf{E}\left[f\left(\bar{X}^{\epsilon}_T + \frac{1}{n}\hat{w}_T\right)\right] - \mathbf{E}\left[f\left(X^{\epsilon}_T(0,x) + \frac{1}{n}\hat{w}_T\right)\right]\right\}\right. \\ &\left. - \left\{\mathbf{E}\left[f\left(\bar{X}^{[0]}_T + \frac{1}{n}\hat{w}_T\right)\right] - \mathbf{E}\left[f\left(X^{[0]}_T(0,x) + \frac{1}{n}\hat{w}_T\right)\right]\right\}\right] = O\left(\frac{\epsilon}{n}\right).\end{aligned} \tag{70}$$

Then, we have

$$\mathbf{Bias}[\mathbf{V}^*(\epsilon,n,N)] = O\left(\frac{\epsilon}{n}\right) \quad \text{under [A2]} \tag{71}$$

$$\mathbf{Bias}[\mathbf{V}^*(\epsilon,n,N)] = O\left(\frac{\epsilon}{n}\right) + O(\epsilon^K) \text{ for all } K > 0 \text{ under [A3].} \tag{72}$$

7 Appendix: On the Validity of Square-root Processes in the Asymptotic Method

Let processes $\{X^{\epsilon}_t; 0 \le t \le T\}$ and $\{\tilde{X}^{\epsilon}_t; 0 \le t \le T\}$ defined as follows:

$$\begin{cases} dX^{\epsilon}_t = (cX^{\epsilon}_t + d)dt + \epsilon\sqrt{X^{\epsilon}_t}dw_t, \; X^{\epsilon}_0 = x_0 \\ d\tilde{X}^{\epsilon}_t = (c\tilde{X}^{\epsilon}_t + d)dt + \epsilon g(\tilde{X}^{\epsilon}_t)dw_t, \; \tilde{X}^{\epsilon}_0 = x_0 \end{cases} \tag{73}$$

where $T < \infty$, c, d are some constants with $d \ge 0$, $x_0 > 0$, and $\epsilon \in (0,1]$. $g(x)$ is a smooth modification of $\sqrt{x}$ such that $g(x) = \sqrt{x}$ for $x \ge a^{'}$ 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 $\mathbf{R}$-valued functional F, $F(X^{\epsilon})$ and $F(\tilde{X}^{\epsilon})$ are $L_2(P)$-finite. Then, we have

$$\mathbf{E}[|F(X^{\epsilon}) - F(\tilde{X}^{\epsilon})|1_{\{X^{\epsilon}\neq\tilde{X}^{\epsilon}\}}] \le (\|F(X^{\epsilon})\|_2 + \|F(\tilde{X}^{\epsilon})\|_2)P(\{X^{\epsilon} \neq \tilde{X}^{\epsilon}\})^{\frac{1}{2}}$$

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

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

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

$$P(\{\sup_{0\leq s\leq T}|Z_s^\epsilon - Z_s^0| > a\}) \leq a_1 \exp(-a_2\epsilon^{-2}) \tag{74}$$

for all $a > 0$.

The lemma can be proved by slight modification of lemma5.3 in Yoshida(1992), or lemma 7.1 in Kunitomo and Takahashi(2003). Note also that X^ϵ and $\tilde{X}^\epsilon$ satisfy the condtions in lemma 2. Hence, if $\|F(X^\epsilon)\|_2 < \infty$ and $\|F(\tilde{X}^\epsilon)\|_2 < \infty$, then

$$\mathbf{E}[|F(X^\epsilon) - F(\tilde{X}^\epsilon)|] = o(\epsilon^n),\ n = 1, 2, \cdots. \tag{75}$$

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,

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

and

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

References

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

[2] Detemple, J. R., Garcia, J.R., and Rindisbacher, M. (2000), "A Monte Carlo Method for Optimal Portfolios," Working Paper, Columbia University.

[3] He, H. and Takahashi, A. (2000), "A Variable Reduction Technique for Pricing Average-Rate Options," *International Review of Finance*, Vol.1, 123-142.

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

[5] Kohatsu-Higa, A. (1996), "An approximation to the density of a diffusion," Preprint.

[6] Kunitomo, N. and Takahashi, A. (1992), "Pricing Average Options," *Japan Financial Review*, Vol.14, 1-20. (In Japanese.)

[7] Kunitomo, N. and Takahashi, A. (2001), "The Aysmptotic Expansion Approach to the Valuation of Interest Rate Contingent Claims," *Mathematical Finance*, Vol.11, 117-151.

[8] Kunitomo, N. and Takahashi, A. (2003), "On Validity of the Asymptotic Expansion Approach in Contingent Claim Analysis," to appear *Annals of Applied Probability*.

[9] Ocone, D. and Karatzas, I.(1991), "A Generalized Clark Representation Formula, with Application to Optimal Portfolios," *Stochastics and Stochastics Reports*, Vol.34, 187-220.

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

[11] Takahashi, A. (1995), "Essays on the Valuation Problems of Contingent Claims," Unpublished Ph.D. Dissertation, Haas School of Business, University of California, Berkeley.

[12] Takahashi, A. (1999), "An Asymptotic Expansion Approach to Pricing Contingent Claims," *Asia-Pacific Financial Markets*, Vol.6, 115-151.

[13] Takahashi, A. and Yoshida, N. (2001), "An Asymptotic Expansion Sheme for Optimal Portfolio for Investment," in *Kokyuroku* 1215, Research Institute for Mathematical Sciences, Kyoto University.

[14] Takahashi, A. and Yoshida, N. (2003), "An Asymptotic Expansion Scheme for Optimal Investment Problems," Preprint(submitted).

[15] Yoshida, N. (1992), "Asymptotic Expansions for Statistics Related to Small Diffusions," *Journal of Japan Statistical Society*, Vol.22, 139-159.