CARF Working Paper

CARF-F-335

A Weak Approximation with Asymptotic Expansion and Multidimensional Malliavin Weights

Akihiko Takahashi
The University of Tokyo
Toshihiro Yamada
The University of Tokyo & MTEC

First version: December 2013
Current version: January 2014

CARF is presently supported by Bank of Tokyo-Mitsubishi UFJ, Ltd., Dai-ichi Mutual Life Insurance Company, Meiji Yasuda Life Insurance Company, Nomura Holdings, Inc. and Sumitomo Mitsui Banking Corporation (in alphabetical order). This financial support enables us to issue CARF Working Papers.

A Weak Approximation with Asymptotic Expansion and Multidimensional Malliavin Weights

Akihiko Takahashi* and Toshihiro Yamada†

First Version: November, 2013, This Version: January 4, 2014

Abstract

This paper develops a new efficient scheme for approximations of expectations of the solutions to stochastic differential equations (SDEs). In particular, we present a method for connecting approximate operators based on an asymptotic expansion with multidimensional Malliavin weights to compute a target expectation value precisely. The mathematical validity is given based on Watanabe and Kusuoka theories in Malliavin calculus. Moreover, numerical experiments for option pricing under local and stochastic volatility models confirm the effectiveness of our scheme. Especially, our weak approximation substantially improve the accuracy at deep Out-of-The-Moneys (OTMs).
Keywords: Asymptotic expansion, Weak approximation, Malliavin calculus, Watanabe theory, Kusuoka Scheme, Option pricing

1 Introduction

Developing an approximation method for expectations of diffusion processes is an interesting topic in various research fields. In fact, it seems so useful that a precise approximation for the expectation would lead to substantial reduction of computational burden so that the subsequent analyses could be very easily implemented. Particularly, in finance it has drawn much attention for more than the past two decades since fast and precise computation is so important in terms of competition and risk management in practice such as in trading and investment.

An example among a large number of the related researches is an asymptotic expansion approach, which is mathematically justified by Watanabe theory (Watanabe (1987)) in Malliavin calculus (e.g. Malliavin (1997)). Especially, the asymptotic expansion have been applied to a broad class of problems in finance: for instance, see Takahashi and Yamada (2012a,b, 2013a,b) and references therein.

Although the asymptotic expansion up to the fifth order is known to be sufficiently accurate for option pricing (e.g. Takahashi et al. (2012)), the main criticism against the method would be that the approximate density function deviates from the true density at its tails that is, some region of the very deep Out-of-The-Money (OTM). However, there

*University of Tokyo

†University of Tokyo & MTEC

exist similar problems, at least implicitly in other well-known approximation methods such as Hagan et al. (2002).

On the other hand, the Monte Carlo simulation method is quite popular mainly due to the ease of its implementation. Nevertheless, in order to achieve accuracy sufficient enough in practice, there exists an unavoidable drawback in computational cost under the standard weak approximation schemes of SDEs such as the Euler-Maruyama scheme.

To overcome this problem, Kusuoka (2001, 2003b, 2004) developed a high order weak approximation scheme for SDEs based on Malliavin calculus and Lie algebra, which opened the door for the possibility that the computational speed and the accuracy in the Monte Carlo simulation satisfies stringent requirements in financial business. Independently, Lyons and Victoir (2004) developed a cubature method on the Wiener space. Since then, there have been a large number of researches for weak approximations and its applications to the computational finance inspired by those pioneering works. For instance see Crisan et al. (2013) for the Kusuoka's method and its related works (e.g. Bayer et al. (2013)).

This paper develops a new weak approximation scheme for expectations of functions of the solutions to SDEs. In particular, the scheme connects approximate operators constructed based on the asymptotic expansion. More concretely, a diffusion semigroup is defined as the expectation of an appropriate function of the solution to a certain SDE: for example, $P_t^{\varepsilon}f(x) = E[f(X_t^{x,\varepsilon})]$ with the solution $X_t^{x,\varepsilon}$ of a SDE with perturbation parameter ε and a function f. Then, we approximate P_t^{ε} by an operator $Q_t^{\varepsilon,m}$ which is constructed based on the asymptotic expansion up to a certain order m. Thus, given a partition of $[0,T]$, $\pi = \{(t_0, t_1, \cdots, t_n) : 0 = t_0 < t_1 < \cdots < t_n = T\}$, we are able to approximate $P_T^{\varepsilon}f(x)$ by connecting the expansion-based approximations sequentially: that is, with $s_k = t_k - t_{k-1}$, $k = 1, \cdots, n$,

$$P_T^{\varepsilon}f(x) \quad \simeq \quad Q_{s_n}^{\varepsilon,m}Q_{s_{n-1}}^{\varepsilon,m}\cdots Q_{s_1}^{\varepsilon,m}f(x).$$

This paper justifies this idea by applying Malliavin calculus, particularly, theories developed by Watanabe (1987), Kusuoka (2003a) and Kusuoka (2001, 2004).

Moreover, we show through numerical examples for option pricing that very few partition such as $n = 2$ is mostly enough to substantially improve the errors at deep OTMs of expansions with order $m = 1, 2$. For a related but different approach with similar motivation see Section 5 in Fujii (2013).

The organization of the paper is as follows. The next section introduces the setup and the basic results necessary for the subsequent analysis. Section 3 shows our main result for a new weak approximation of the expectation of diffusion processes. After Section 4 briefly describes an example for the implementation method of our scheme, Section 5 provides numerical experiments for option pricing under local and stochastic volatility models. Section 6 concludes. Appendix gives the proofs of the theorems 1 and 2 as well as the lemma 2 and its proof.

2 Preparation

Let $(\mathcal{W}, H, P)$ be the d-dimensional Wiener space and $\mathbf{D}^{k,p}(E)$, $k \geq 1$, $p \in [1, \infty)$ be the space of k-times Malliavin differentiable Wiener functionals $F \in L^p(\mathcal{W}, E)$, where E is a separable Hilbert space. See Watanabe (1987), Ikeda and Watanabe (1989) and Nualart (2006) for more details of the notations. Let $B_t = (B_t^1, \cdots, B_t^d)$ be a d-dimensional Brownian motion. In this paper, we consider the following general perturbed

N-dimensional stochastic differential equation with $\varepsilon \in (0,1]$:

$$X_t^{x,\varepsilon} = x + \int_0^t V_0(\varepsilon, X_s^{x,\varepsilon})ds + \varepsilon \sum_{j=1}^d \int_0^t V_j(X_s^{x,\varepsilon})dB_s^j, \tag{2.1}$$

where $V_0 \in C_b^\infty((0,1] \times \mathbf{R}^N; \mathbf{R}^N)$ and $V_j \in C_b^\infty(\mathbf{R}^N; \mathbf{R}^N)$, $j = 1, \cdots, d$ are bounded. Hereafter, we will use the notation $Vf(x) = \sum_{i=1}^N V^i(x)(\partial f/\partial x_i)(x)$ for $V \in C_b^\infty(\mathbf{R}^N; \mathbf{R}^N)$ and f a differentiable function $\mathbf{R}^N$ into $\mathbf{R}$. $X_t^{x,\varepsilon}$ can be written in the Stratonovich form:

$$X_t^{x,\varepsilon} = x + \int_0^t \tilde{V}_0(\varepsilon, X_s^{x,\varepsilon})ds + \varepsilon \sum_{j=1}^d \int_0^t \tilde{V}_j(X_s^{x,\varepsilon}) \circ dB_s^j, \tag{2.2}$$

where

$$\tilde{V}_0^i(\varepsilon, x) = V_0^i(\varepsilon, x) - \frac{\varepsilon^2}{2} \sum_{j=1}^d V_j V_j^i(x), \tag{2.3}$$

$$\varepsilon \tilde{V}_j^i(x) = \varepsilon V_j^i(x), \quad j = 1, \cdots, d. \tag{2.4}$$

Let Y_t^ε be a Wiener functional defined by

$$Y_t^\varepsilon = \frac{X_t^{x,\varepsilon} - X_t^{x,0}}{\varepsilon}. \tag{2.5}$$

Define the Malliavin covariance matrix of Y_t^ε:

$$\sigma_{i,j}^{Y_t^\varepsilon} = \sum_{k=1}^d \int_0^t D_{s,k} Y_t^{\varepsilon,i} D_{s,k} Y_t^{\varepsilon,j} ds, \quad 1 \le i, j \le N. \tag{2.6}$$

We assume the following condition **[H]** on the Malliavin covariance matrix in this paper.

[H] For any $t \in (0, T]$,

$$\limsup_{\varepsilon \downarrow 0} \left\| \det(\sigma^{Y_t^\varepsilon})^{-1} \right\|_{L^p} < \infty. \tag{2.7}$$

2.1 The space $\mathcal{K}_r$

This subsection introduces the space of Wiener functionals $\mathcal{K}_r$ developed by Kusuoka (2003a) and its properties. The element of $\mathcal{K}_r$ is called the *Kusuoka-Stroock function*. See Nee (2010, 2011)and Crisan et al. (2013) for more details of the notations and the proofs.

Definition 1. *Given $r \in \mathbf{R}$ and $n \in \mathbf{N}$, we denote by $\mathcal{K}_r(E, n)$ the set of functions $G : (0,1] \times \mathbf{R}^N \to \mathbf{D}^{n,\infty}(E)$ satisfying the following:*

1. *$G(t, \cdot)$ is n-times continuously differentiable and $[\partial^\alpha G/\partial x^\alpha]$ is continuous in $(t, x) \in (0, T] \times \mathbf{R}^N$ a.s. for any multi-index α of the elements of $\{1, \cdots, d\}$ with length $|\alpha| \le n$.*

2. For all $k \leq n - |\alpha|$, $p \in [1, \infty)$,

$$\sup_{t \in (0,1], x \in \mathbf{R}^N} t^{-r/2} \left\| \frac{\partial^\alpha G}{\partial x^\alpha}(t,x) \right\|_{\mathbf{D}^{k,p}} < \infty. \tag{2.8}$$

We write $\mathcal{K}_r$ for $\mathcal{K}_r(\mathbf{R}, \infty)$.

Next, we show the basic properties of the Kusuoka-Stroock functions. are the following.

Lemma 1. *[Properties of Kusuoka-Stroock functions]*

1. *The function $(t,x) \in (0,1] \times \mathbf{R}^N \mapsto X_t^{x,\varepsilon}$ belongs to $\mathcal{K}_0$.*
2. *Suppose $G \in \mathcal{K}_r(n)$ where $r \geq 0$. Then, for $i = 1, \cdots, d$,*

$$(a) \int_0^\cdot G(s,x) dB_s^i \in \mathcal{K}_{r+1}(n), \text{ and } (b) \int_0^\cdot G(s,x) ds \in \mathcal{K}_{r+2}(n). \tag{2.9}$$

3. *If $G_i \in \mathcal{K}_{r_i}(n_i)$, $i = 1, \cdots, N$, then*

$$(a) \prod_i^N G_i \in \mathcal{K}_{r_1 + \cdots + r_N}(\min_i n_i), \text{ and } (b) \sum_{i=1}^N G_i \in \mathcal{K}_{\min_i r_i}(\min_i n_i). \tag{2.10}$$

Then, we summarize the Malliavin's integration by parts formula using Kusuoka-Stroock functions. For any multi-index $\alpha^{(k)} := (\alpha_1, \cdots, \alpha_k) \in \{1, \cdots, N\}^k$, $k \geq 1$, we denote by $\partial_{\alpha^{(k)}}$ the partial derivative $\dfrac{\partial^k}{\partial x_{\alpha_1} \cdots \partial x_{\alpha_k}}$.

Proposition 1. *Let $G : (0,1] \times \mathbf{R}^N \to \mathbf{D}^\infty = \mathbf{D}^{\infty,\infty}(\mathbf{R})$ be an element of $\mathcal{K}_r$ and let f be a function that belongs to the space $C_b^\infty(\mathbf{R}^N; \mathbf{R})$. Then for any multi-index $\alpha^{(k)} \in \{1, \cdots, d\}^k$, $k \geq 1$, there exists $H_{\alpha^{(k)}}(X_t^{x,\varepsilon}, G(t,x)) \in \mathcal{K}_{r-|\alpha^{(k)}|}$ such that*

$$E\left[\partial_{\alpha^{(k)}} f(X_t^{x,\varepsilon}) G(t,x)\right] = E\left[f(X_t^{x,\varepsilon}) H_{\alpha^{(k)}}(X_t^{x,\varepsilon}, G(t,x))\right], \ t \in (0,1], \tag{2.11}$$

with

$$\sup_{x \in \mathbf{R}^N} \| H_{\alpha^{(k)}}(X_t^{x,\varepsilon}, G(t,x)) \|_{L^p} \leq t^{(r - |\alpha^{(k)}|)/2} C, \tag{2.12}$$

where $H_{\alpha^{(k)}}(X_t^{x,\varepsilon}, G(t,x))$ is recursively given by

$$H_{(i)}(X_t^{x,\varepsilon}, G(t,x)) = \delta \left(\sum_{j=1}^N G(t,x) \gamma_{ij}^{X_t^{x,\varepsilon}} D X_t^{x,\varepsilon,j} \right), \tag{2.13}$$

$$H_{\alpha^{(k)}}(X_t^{x,\varepsilon}, G(t,x)) = H_{(\alpha_k)}(X_t^{x,\varepsilon}, H_{\alpha^{(k-1)}}(X_t^{x,\varepsilon}, G(t,x))), \tag{2.14}$$

and a positive constant C. Here, δ is the Skorohod integral and $(\gamma_{ij}^{X_t^{x,\varepsilon}})_{1 \leq i,j \leq n}$ is the inverse matrix of the Malliavin covariance of $X_t^{x,\varepsilon}$.

Proof. Apply Corollary 3.7 of Kusuoka and Stroock (1984) and Lemma 8-(3) of Kusuoka (2003a) with Proposition 2.1.4 of Nualart (2006). □

Remark 1. *Kusuoka (2003a) shows that Proposition 1 holds under the UFG condition. See p. 262 of Kusuoka (2003a) for the definition of the UFG condition. We remark that if the vector fields V_0 and V_i, $i = 1, \cdots, d$ satisfy the uniform Hörmander condition, then they satisfy the UFG condition. We also remark that if the vector fields V_0 and V_i, $i = 1, \cdots, d$ satisfy the uniform ellipticity condition, then they satisfy the UFG condition.*

3 Weak Approximation with Asymptotic Expansion Method

Let $(P_t)_t$ be linear operators on $f \in C_b(\mathbf{R}^N;\mathbf{R})$ defined by

$$P_t f(x) = E[f(X_t^{x,\varepsilon})], \tag{3.1}$$

where

$$X_t^{x,\varepsilon} \quad = \quad x + \int_0^t V_0(\varepsilon, X_s^{x,\varepsilon})ds + \varepsilon \sum_{j=1}^d \int_0^t V_j(X_s^{x,\varepsilon})dB_s^j. \tag{3.2}$$

We remark that $(P_t)_t$ is a semigroup. Also let $(P_t^0)_t$ be linear operators on $f \in C_b(\mathbf{R}^N;\mathbf{R})$ defined by

$$P_t^0 f(x) = E[f(\bar{X}_t^{x,0})], \tag{3.3}$$

where

$$\bar{X}_t^{x,0} \quad = \quad X_t^{x,0} + \varepsilon \frac{\partial}{\partial \varepsilon} X_t^{x,\varepsilon}|_{\varepsilon=0}. \tag{3.4}$$

Note that $\frac{\partial}{\partial \varepsilon} X_t^{x,\varepsilon}|_{\varepsilon=0}$ is a Gaussian random variable.

Remark 2. *When $\tilde{V}_0(\varepsilon, x) = \varepsilon \tilde{V}_0(x)$, $\bar{X}_t^{x,0}$ is given by*

$$\bar{X}_t^{x,0} \quad = \quad x + \varepsilon \sum_{i=0}^d V_i(x) \int_0^t dB_s^i, \tag{3.5}$$

where $B_t^0 = t$.

In the remainder of the paper, we use the following norms and semi-norms:

$$\|f\|_\infty = \sup_{x\in\mathbf{R}^N} |f(x)|, \qquad \|\nabla f\|_\infty = \max_{i\in\{1,\cdots,N\}} \left\| \frac{\partial f}{\partial x_i} \right\|_\infty, \quad f \in C_b^\infty(\mathbf{R}^N;\mathbf{R}), \tag{3.6}$$

$$\|\nabla^i f\|_\infty = \max_{j_1,\cdots,j_i\in\{1,\cdots,N\}} \left\| \frac{\partial^i f}{\partial x_{j_1} \cdots \partial x_{j_i}} \right\|_\infty. \tag{3.7}$$

Next, as an approximation of P we introduce a linear operator Q^m below. Firstly, for $j \geq 1$, let $P_{\Phi^j}^0$ be a linear operator defined by the following expectation with Malliavin weight

$$P_{\Phi^j}^0(t)f(x) = E\left[f(\bar{X}_t^{x,0})\Phi_t^j\right], \tag{3.8}$$

where

$$\Phi_t^j \quad = \quad \sum_{k=1}^j \sum_{\beta_1+\cdots+\beta_k=j+k, \beta_l\geq 2} \sum_{\alpha^{(k)}\in\{1,\cdots,N\}^k} \frac{1}{k!} \tag{3.9}$$

$$H_{\alpha^{(k)}} \left(\frac{\partial}{\partial \varepsilon} X_t^{x,\varepsilon}|_{\varepsilon=0} \ , \ \prod_{l=1}^k \frac{1}{\beta_l!} \frac{\partial^{\beta_l}}{\partial \varepsilon^{\beta_l}}|_{\varepsilon=0} X_t^{x,\varepsilon,\alpha_l} \right). \tag{3.10}$$

Then, $\{Q^m_{(s)},\, s \in (0,1]\}$ is defined as a linear operator:

$$Q^m_{(s)} f(x) \quad = \quad P^0_s f(x) + \sum_{j=1}^{m} \varepsilon^j P_{\Phi^j}(s) f(x). \tag{3.11}$$

Remark 3. *When $\tilde{V}_0(\varepsilon, x) = \varepsilon \tilde{V}_0(x)$, $X^{x,\varepsilon}_t$ has the following expansion:*

$$\begin{aligned}
X^{x,\varepsilon}_t \quad = \quad & x + \varepsilon \sum_{j=0}^{d} \tilde{V}_j(x) \int_0^t \circ dB^j_s \\
& + \sum_{k=2}^{m} \varepsilon^k \sum_{(i_1,\cdots,i_k)\in\{0,1,\cdots,d\}^k, k+\#\{j|i_j=0\}\le m} \\
& \qquad (\tilde{V}_{i_1} \cdots \tilde{V}_{i_k})(x) \int_{0<t_1<\cdots<t_k<t} \circ dB^{i_1}_{t_1} \circ \cdots \circ dB^{i_k}_{t_k} \\
& + \varepsilon^{m+1} R_m(t, x, \varepsilon),
\end{aligned}$$

where $R_m(t,x,\varepsilon)$ is the residual. Here, we used the notation $B^0_t = t$.

Then, we have the following representations for $P^0_{\Phi^j}$, $j \ge 1$ and Q^m.

Proposition 2. *It holds*

$$P^0_{\Phi^j}(t) f(x) = E[f(\bar{X}^{x,0}_t) \mathcal{M}_{(j)}(t, x, \bar{X}^{x,0}_t)], \tag{3.12}$$

where $\mathcal{M}_{(j)}(t,x,y) = E[\Phi^j_t | \bar{X}^{x,0}_t = y]$ and

$$Q^m_{(s)} f(x) \quad = \quad E[f(\bar{X}^{x,0}_s) \mathcal{M}^m(s, x, \bar{X}^{x,0}_s)] \tag{3.13}$$

where $\mathcal{M}^m(s,x,y) = 1 + \sum_{j=1}^{m} \varepsilon^j \mathcal{M}_{(j)}(s,x,y)$.

Proof.
We have

$$\begin{aligned}
P^0_{\Phi^j}(t) f(x) \quad &= \quad \int_{\mathbf{R}^N} f(y) E\left[\Phi^j_t | \bar{X}^{x,0}_t = y\right] p(t,x,y) dy \\
&= \quad E\left[f(\bar{X}^{x,0}_t) \mathcal{M}_{(j)}(t, x, \bar{X}^{x,0}_t)\right]
\end{aligned}$$

and

$$\begin{aligned}
Q^m_{(s)} f(x) \quad &= \quad \int_{\mathbf{R}^N} f(y) \{1 + \sum_{j=1}^{m} \varepsilon^j E[\Phi^j_s | \bar{X}^{x,0}_s = y]\} p(s,x,y) dy \\
&= \quad E[f(\bar{X}^{x,0}_s) \mathcal{M}^m(s, x, \bar{X}^{x,0}_s)],
\end{aligned}$$

where $y \mapsto p(t,x,y)$ is the density of $\bar{X}^{x,0}_t$. $\square$

The next theorem shows the asymptotic approximation results for the expectations of the solution to the general perturbed SDEs $X^{x,\varepsilon}_s$.

Theorem 1. *We have the following estimates:*

1. *There exists $C > 0$ and $n \in \mathbf{N}$ such that*
$$\|P_s f - Q^m_{(s)} f\|_\infty \le \varepsilon^{m+1} C (\sum_{k=1}^{n} s^{(m+1+k)/2} \|\nabla^k f\|_\infty), \tag{3.14}$$
for any $s \in (0,1]$ and $f \in C^\infty_b(\mathbf{R}^N; \mathbf{R})$.
2. *There exists $C > 0$ such that*
$$\|P_s f - Q^m_{(s)} f\|_\infty \le \varepsilon^{m+1} C s^{(m+2)/2}, \tag{3.15}$$
for any $s \in (0,1]$ and Lipschitz continuous function $f : \mathbf{R}^N \to \mathbf{R}$.
3. *There exists $C > 0$ such that*
$$\|P_s f - Q^m_{(s)} f\|_\infty \le \varepsilon^{m+1} C s^{(m+1)/2}, \tag{3.16}$$
for any $s \in (0,1]$ and bounded Borel function $f : \mathbf{R}^N \to \mathbf{R}$.

Proof.
See Appendix A. □

Remark 4. *The above results are obtained based on the integration by parts argument for $G(s,x) \in \mathcal{K}_r$ with time $s \in (0,1]$. However, we are able to show that the same results hold for $s \in (0,T]$, $T > 0$, using the properties of the elements in the space $\mathcal{K}^T_r$ defined as in Crisan et al. (2013).*

Next, for $T > 0$, $\gamma > 0$, define a partition $\pi = \{(t_0, t_1, \cdots, t_n) : 0 = t_0 < t_1 < \cdots < t_n = T, t_k = k^\gamma T/n^\gamma, n \in \mathbf{N}\}$ and $s_k = t_k - t_{k-1}$, $k = 1, \cdots, n$. Using the asymptotic expansion operator Q^m of P, we can guess the following semigroup approximation.

$$E[f(X^{x,\varepsilon}_T)] = P_T f(x) = P_{s_n} P_{s_{n-1}} \cdots P_{s_1} f(x) \quad \simeq \quad Q^m_{(s_n)} Q^m_{(s_{n-1})} \cdots Q^m_{(s_1)} f(x).$$

The next theorem shows our main result on the approximation error for this scheme.

Theorem 2. *Let $T > 0$, $\gamma > 0$ and $n \in \mathbf{N}$.*

1. *For any $f \in C^\infty_b(\mathbf{R}^N; \mathbf{R})$, there exists $C > 0$ such that*
$$\|P_T f - Q^m_{(s_n)} Q^m_{(s_{n-1})} \cdots Q^m_{(s_1)} f\|_\infty \le \varepsilon^{m+1} \frac{C}{n^{\gamma(m+2)/2}}, \quad 0 < \gamma < m/(m+2), \tag{3.17}$$
$$\|P_T f - Q^m_{(s_n)} Q^m_{(s_{n-1})} \cdots Q^m_{(s_1)} f\|_\infty \le \varepsilon^{m+1} \frac{C}{n^{m/2}} (1 + \log n), \quad \gamma = m/(m+2), \tag{3.18}$$
$$\|P_T f - Q^m_{(s_n)} Q^m_{(s_{n-1})} \cdots Q^m_{(s_1)} f\|_\infty \le \varepsilon^{m+1} \frac{C}{n^{m/2}}, \quad \gamma > m/(m+2). \tag{3.19}$$

2. *For any Lipschitz continuous function $f : \mathbf{R}^N \to \mathbf{R}$, there exists $C > 0$ such that*

$$\|P_T f - Q^m_{(s_n)} Q^m_{(s_{n-1})} \cdots Q^m_{(s_1)} f\|_\infty \leq \varepsilon^{m+1} \frac{C}{n^{\gamma(m+2)/2}}, \quad 0 < \gamma < m/(m+2), \tag{3.20}$$

$$\|P_T f - Q^m_{(s_n)} Q^m_{(s_{n-1})} \cdots Q^m_{(s_1)} f\|_\infty \leq \varepsilon^{m+1} \frac{C}{n^{m/2}} (1 + \log n), \quad \gamma = m/(m+2), \tag{3.21}$$

$$\|P_T f - Q^m_{(s_n)} Q^m_{(s_{n-1})} \cdots Q^m_{(s_1)} f\|_\infty \leq \varepsilon^{m+1} \frac{C}{n^{m/2}}, \quad \gamma > m/(m+2). \tag{3.22}$$

3. *For any bounded Borel function $f : \mathbf{R}^N \to \mathbf{R}$, there exists $C > 0$ such that*

$$\|P_T f - Q^m_{(s_n)} Q^m_{(s_{n-1})} \cdots Q^m_{(s_1)} f\|_\infty \leq \varepsilon^{m+1} \frac{C}{n^{\gamma(m+1)/2}}, \quad 0 < \gamma < (m-1)/(m+1), \tag{3.23}$$

$$\|P_T f - Q^m_{(s_n)} Q^m_{(s_{n-1})} \cdots Q^m_{(s_1)} f\|_\infty \leq \varepsilon^{m+1} \frac{C}{n^{(m-1)/2}} (1 + \log n), \quad \gamma = (m-1)/(m+1), \tag{3.24}$$

$$\|P_T f - Q^m_{(s_n)} Q^m_{(s_{n-1})} \cdots Q^m_{(s_1)} f\|_\infty \leq \varepsilon^{m+1} \frac{C}{n^{(m-1)/2}}, \quad \gamma > (m-1)/(m+1). \tag{3.25}$$

Proof.
See Appendix C. □

Remark 5. *Due to the theorem above, the higher order asymptotic expansion provides the higher order weak approximation. In fact, we can mostly attain enough accuracy even when the expansion order m is low such as $m = 1, 2$. In Section 5 we confirm this fact through numerical examples.*

Remark 6. *When $\gamma = 1$, i.e. $s_k = T/n$ for all $k = 1, \cdots, n$, we have*

1. *For any $f \in C_b^\infty(\mathbf{R}^N; \mathbf{R})$, there exists $C > 0$ such that*

$$\|P_T f - (Q^m_{(T/n)})^n f\|_\infty \leq \varepsilon^{m+1} \frac{C}{n^{m/2}},$$

2. *For any Lipschitz continuous function $f : \mathbf{R}^N \to \mathbf{R}$, there exists $C > 0$ such that*

$$\|P_T f - (Q^m_{(T/n)})^n f\|_\infty \leq \varepsilon^{m+1} \frac{C}{n^{m/2}},$$

3. *For any bounded Borel function $f : \mathbf{R}^N \to \mathbf{R}$, there exists $C > 0$ such that*

$$\|P_T f - (Q^m_{(T/n)})^n f\|_\infty \leq \varepsilon^{m+1} \frac{C}{n^{(m-1)/2}}.$$

4 Computation with Malliavin Weights

This section illustrates computational scheme for implementation of our method.

4.1 Backward Discrete-time Approximation

Firstly, for preparation we describe a backward discrete time approximation of our method.

For $s \in (0,1]$ and $x, y \in \mathbf{R}^N$, define $p^m(s,x,y)$ as

$$Q^m_{(s)} f(x) = \int_{\mathbf{R}^N} f(y) p^m(s,x,y) dy. \tag{4.1}$$

We also recall that the m-th order Malliavin weight function $\mathcal{M}^m$ is given as

$$Q^m_{(s)} f(x) = E[f(\bar{X}^{x,0}_s) \mathcal{M}^m(s,x,\bar{X}^{x,0}_s)]. \tag{4.2}$$

Then,

$$p^m(s,x,y) = \mathcal{M}^m(s,x,y) p(s,x,y), \tag{4.3}$$

with

$$p(s,x,y) = \frac{1}{(2\pi\varepsilon^2)^{N/2} \det(\Sigma(s))^{1/2}} e^{-\frac{(y-x)\Sigma^{-1}(t)(y-x)^T}{2\varepsilon^2}}, \tag{4.4}$$

with the covariance matrix $\Sigma(s)$ of $\bar{X}^{x,0}_s$.

Then, we are able to calculate $(Q^m_{(T/n)})^n f(x)$ as follows:

$$(Q^m_{(T/n)})^n f(x) \tag{4.5}$$

$$= \int_{(\mathbf{R}^N)^n} f(y_n) \prod_{i=0}^{n-1} p^m(s_i, y_i, y_{i+1}) dy_n \cdots dy_1 \tag{4.6}$$

$$= \int_{(\mathbf{R}^N)^{n-1}} q_{n-1}(y_{n-1}) \prod_{i=0}^{n-2} p^m(s_i, y_i, y_{i+1}) dy_{n-1} \cdots dy_1 \tag{4.7}$$

$$= \int_{(\mathbf{R}^N)^{n-2}} q_{n-2}(y_{n-2}) \prod_{i=0}^{n-3} p^m(s_i, y_i, y_{i+1}) dy_{n-2} \cdots dy_1 \tag{4.8}$$

$$= \int_{\mathbf{R}^N} q_1(y_1) p^m(s_1, y_0, y_1) dy_1, \tag{4.9}$$

with $y_0 = x$.

4.2 Example of Computational Scheme

We are able to compute the expectation in the various ways such as numerical integration and Monte Carlo simulation. As an illustrative purpose and an example, this subsection briefly describes a scheme based on Monte Carlo simulation.

When we compute $(Q^m_{(T/n)})^n f(x)$ (i.e. $\gamma = 1$) with simulation, we store the j-th ($1 \leq j \leq M$) outcome of the simulation, $\bar{X}^{x_0,0,\pi,(j)}$ with the time grid π. Then, we calculate an approximate semigroup at each time grid. That is, $q_{n-1}(x)$, $q_{n-2}(x)$ are calculated as follows:

$$q_{n-1}(x) = \int_{\mathbf{R}^N} f(y) p^m(T/n, x, y) dy \tag{4.10}$$

$$= \int_{\mathbf{R}^N} f(y) \mathcal{M}^m(T/n, x, y) p(T/n, x, y) dy \tag{4.11}$$

$$\simeq \frac{1}{M} \sum_{j=1}^{M} f(\bar{X}^{x_0,0,\pi,(j)}_T) \mathcal{M}^m(T/n, x, \bar{X}^{x_0,0,\pi,(j)}_T) \tag{4.12}$$

$$
\begin{aligned}
q_{n-2}(x) &= \int_{\mathbf{R}^N} q_{n-1}(y) p^m(T/n, x, y) dy && (4.13)\\
&= \int_{\mathbf{R}^N} q_{n-1}(y) \mathcal{M}^m(T/n, x, y) p(T/n, x, y) dy && (4.14)\\
&\simeq \frac{1}{M} \sum_{j=1}^{M} q_{n-1}(\bar{X}_{t_{n-1}}^{x_0,0,\pi,(j)}) \mathcal{M}^m(T/n, x, \bar{X}_{t_{n-1}}^{x_0,0,\pi,(j)}). && (4.15)
\end{aligned}
$$

Therefore, in general,

$$
\begin{aligned}
q_{i-1}(x) &= \int_{\mathbf{R}^N} q_i(y) p^m(T/n, x, y) dy && (4.16)\\
&= \int_{\mathbf{R}^N} q_i(y) \mathcal{M}^m(T/n, x, y) p(T/n, x, y) dy && (4.17)\\
&\simeq \frac{1}{M} \sum_{j=1}^{M} q_i(\bar{X}_{t_i}^{x_0,0,\pi,(j)}) \mathcal{M}^m(T/n, x, \bar{X}_{t_i}^{x_0,0,\pi,(j)}). && (4.18)
\end{aligned}
$$

Finally, we obtain an approximation:

$$
\begin{aligned}
(Q_{(T/n)}^m)^n f(x) &= \int_{\mathbf{R}^N} q_1(y) p^m(T/n, x, y) dy && (4.19)\\
&= \int_{\mathbf{R}^N} q_1(y) \mathcal{M}^m(T/n, x, y) p(T/n, x, y) dy && (4.20)\\
&\simeq \frac{1}{M} \sum_{j=1}^{M} q_1(\bar{X}_{t_1}^{x_0,0,\pi,(j)}) \mathcal{M}^m(T/n, x, \bar{X}_{t_1}^{x_0,0,\pi,(j)}). && (4.21)
\end{aligned}
$$

We also remark that if the numerical integration method is applied, the scheme is based on the equations (4.17) and (4.20).

5 Numerical Example

This section demonstrates the effectiveness of our method through the numerical examples for option pricing under local and stochastic volatility models.

5.1 Local volatility model

The first example takes the following local volatility model:

$$
\begin{aligned}
dS_t &= \sigma(S_t) dB_t, && (5.1)\\
S_0 &= x.
\end{aligned}
$$

Then, let $(\bar{S}_t^0)_{t\geq 0}$ be the solution to the following SDE:

$$
\begin{aligned}
d\bar{S}_t^0 &= \sigma(x) dB_t, && (5.2)\\
\bar{S}_0^0 &= x.
\end{aligned}
$$

In this numerical example, for the payoff function $f(x) = \max\{x - K, 0\}$ or $f(x) = \max\{K - x, 0\}$ where K is a positive constant, we apply the first order asymptotic expansion operator, that is $m = 1$:

$$
Q_{(t)}^1 f(x) = E[f(\bar{S}_t^{x,0}) \mathcal{M}^1(t, x, \bar{S}_t^{x,0})] \tag{5.3}
$$

and the second order asymptotic expansion operator that is, $m = 2$:

$$Q^2_{(t)} f(x) = E[f(\bar{S}^{x,0}_t)\mathcal{M}^2(t, x, \bar{S}^{x,0}_t)]. \tag{5.4}$$

Here, the Malliavin weights $\mathcal{M}^1(t, x, y)$ and $\mathcal{M}^2(t, x, y)$ are given by

$$\mathcal{M}^1(t, x, y) \quad = \quad E\left[H_{(1)}\left(\bar{S}^{x,0}_t, \frac{1}{2}\frac{\partial^2}{\partial\varepsilon^2}|_{\varepsilon=0}S^{x,\varepsilon}_t\right)|\bar{S}^{x,0}_t = y\right],$$

and

$$\begin{aligned}\mathcal{M}^2(t, x, y) \quad = \quad & E\left[H_{(1)}\left(\bar{S}^{x,0}_t, \frac{1}{6}\frac{\partial^3}{\partial\varepsilon^3}|_{\varepsilon=0}S^{x,\varepsilon}_t\right)|\bar{S}^{x,0}_t = y\right] \\ & +\frac{1}{2}E\left[H_{(1,1)}\left(\bar{S}^{x,0}_t, \left(\frac{1}{2}\frac{\partial^2}{\partial\varepsilon^2}|_{\varepsilon=0}S^{x,\varepsilon}_t\right)^2\right)|\bar{S}^{x,0}_t = y\right].\end{aligned}$$

Also, we specify the local volatility function as $\sigma(x) = \sigma x^\beta$ with $\beta = 0.5$. The parameters are set to be $S_0 = 100$ and $\sigma = 4.0$ that corresponds to 40% under the log-normal model at initial value, $S_0 = 100$. The benchmark values are computed by Monte Carlo simulations with 10^7 trials and 1000 time steps for the 1 year maturity case or 2000 time steps for the 10 year maturity case.

Table 1 and Table 2 show the results. We observe that the increase in the number of the time steps improves the approximation. (See Error Rate AE 1order and Error Rate AE 1order WeakApprox $n = 2, 3$ in Table 1.) We also note that our scheme with the second order expansion and two time steps (Error Rate AE 2order WeakApprox $n = 2$) improves the base (analytical only) second order expansion (Error Rate AE 2order), and is able to provide an accurate approximation across all the strikes even for the long maturity case such as the 10 year maturity case in Table 2.

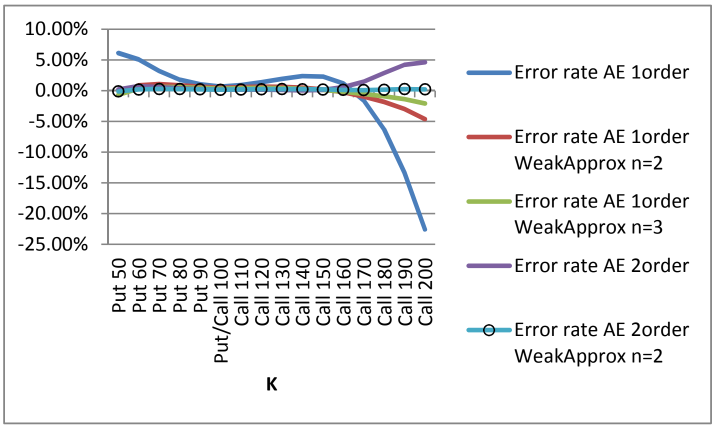

Figure 1: $T = 1$: Local volatility model, Error rates of the 1st and 2nd order asymptotic expansions and their weak approximations

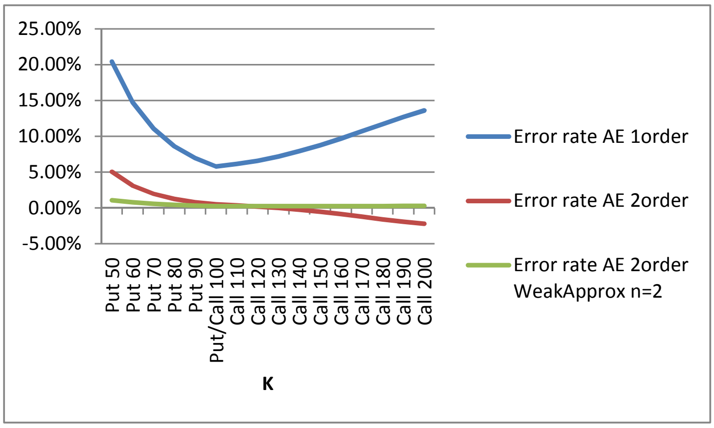

Figure 2: $T = 10$: Local volatility model, Error rates of the 1st and 2nd order asymptotic expansions and the weak approximations

5.2 Stochastic volatility model

The second example considers the following stochastic volatility model, which is also known as the log-normal SABR model.

$$dS_t = \sigma_t S_t dB_t^1, \ S_0 = z, \tag{5.5}$$
$$d\sigma_t = \nu\sigma_t(\rho dB_t^1 + \sqrt{1-\rho^2}dB_t^2), \ \sigma_0 = \sigma. \tag{5.6}$$

Let $X_{1,t} = \log S_t$ and $x = \log z$. Then, we have

$$dX_{1,t} = -\frac{\sigma_t^2}{2}dt + \sigma_t dB_t^1, \ X_{1,0} = x, \tag{5.7}$$
$$d\sigma_t = \nu\sigma_t(\rho dB_t^1 + \sqrt{1-\rho^2}dB_t^2), \ \sigma_0 = \sigma. \tag{5.8}$$

For some fixed $T > 0$ and $K > 0$, the target expectation is given by

$$E\left[f(X_{1,T},\sigma_T)\right] = E\left[\hat{f}(X_{1,T})\right] = E\left[\max\left\{e^{X_{1,T}} - K, 0\right\}\right] \text{ or } E\left[\max\left\{K - e^{X_{1,T}}, 0\right\}\right] \tag{5.9}$$

Next, let $(\bar{X}^0_{1,t}, \bar{\sigma}^0_t)_{t\geq 0}$ be the solution to the following SDE:

$$
\begin{aligned}
d\bar{X}^0_{1,t} &= -\frac{\sigma^2}{2}dt + \sigma dB^1_t,\ \bar{X}^0_{1,0} = x, &(5.10)\\
d\bar{\sigma}^0_t &= \nu\sigma(\rho dB^1_t + \sqrt{1-\rho^2}dB^2_t),\ \bar{\sigma}_0 = \sigma. &(5.11)
\end{aligned}
$$

The parameters are set to be $z = 100$, $\sigma = 0.3$, $\nu = 0.1$ and $\rho = -0.5$. The benchmark values are calculated by Monte Carlo simulations with 10^7 trials and 1000 time steps for the 1 year maturity case or 2000 times steps for the 2 year maturity case.

In this example, we use the first order two dimensional asymptotic expansion operator with two time steps, that is $m = 1$ and $n = 2$. Then, the calculation procedure corresponding to the one in the previous section is the following: Firstly, set $t_0 = 0$, $t_1 = T/2$, $t_2 = T$ and $s = t_k - t_{k-1} = T/2$, $(k = 1, 2)$.

- For $(\bar{X}^0_{1,t_1}, \bar{\sigma}^0_{1,t1}) = (x_1, \sigma_1)$ at $t = t_1$,

$$
q_1(x_1,\sigma_1) = E\left[\hat{f}\left(\bar{X}^{(x_1,\sigma_1),0}_{1,s}\right)\mathcal{M}^1\left(s,(x_1,\sigma_1),\left(\bar{X}^{(x_1,\sigma_1),0}_{1,s},\bar{\sigma}^{\sigma_1,0}_s\right)\right)\right]. \tag{5.12}
$$

- At $t = t_0 = 0$,

$$
q_0(x,\sigma) = E\left[q_1\left(\bar{X}^{(x,\sigma),0}_{1,s},\bar{\sigma}^{\sigma,0}_s\right)\mathcal{M}^1\left(s,(x,\sigma),\left(\bar{X}^{(x,\sigma),0}_{1,s},\bar{\sigma}^{\sigma,0}_s\right)\right)\right]. \tag{5.13}
$$

Here, $\mathcal{M}^1(t,(x,\sigma),(x',\sigma'))$ is the two dimensional Malliavin weight given by

$$
\begin{aligned}
&\mathcal{M}^1(t,(x,\sigma),(x',\sigma'))\\
=\ &E\left[H_{(1)}\left(\left(\bar{X}^{(x,\sigma),0}_{1,t},\bar{\sigma}^{\sigma,0}_t\right),\frac{1}{2}\frac{\partial^2}{\partial\varepsilon^2}|_{\varepsilon=0}X^{(x,\sigma),\varepsilon}_{1,t}\right)|\left(\bar{X}^{(x,\sigma),0}_{1,t},\bar{\sigma}^{\sigma,0}_t\right)=(x',\sigma')\right]\\
&+E\left[H_{(1)}\left(\left(\bar{X}^{(x,\sigma),0}_{1,t},\bar{\sigma}^{\sigma,0}_t\right),\frac{1}{2}\frac{\partial^2}{\partial\varepsilon^2}|_{\varepsilon=0}\sigma^{\sigma,\varepsilon}_t\right)|\left(\bar{X}^{(x,\sigma),0}_{1,t},\bar{\sigma}^{\sigma,0}_t\right)=(x',\sigma')\right].
\end{aligned}
$$

Actually, at t_1 we need not implement (5.12), but just compute the first order analytical asymptotic expansion for pricing options with the time-to-maturity $T/2$ and the initial value $(\bar{X}^0_{1,t_1}, \bar{\sigma}^0_{t1}) = (x_1, \sigma_1)$. That is,

$$
\hat{q}_1(x_1,\sigma_1) = E\left[\hat{f}\left(\bar{X}^{(x_1,\sigma_1),0}_{1,s}\right)\hat{\mathcal{M}}^1\left(s,(x_1,\sigma_1),\bar{X}^{(x_1,\sigma_1),0}_{1,s}\right)\right], \tag{5.14}
$$

where $\hat{\mathcal{M}}^1(s,(x_1,\sigma_1),y) = 1+\hat{\mathcal{M}}_{(1)}(s,(x_1,\sigma_1),y)$, and $\hat{\mathcal{M}}_{(1)}(s,(x_1,\sigma_1),y)$ stands for the first order one dimensional Malliavin weight:

$$
\hat{\mathcal{M}}_{(1)}(s,(x_1,\sigma_1),y) = E\left[H_{(1)}\left(\bar{X}^{(x_1,\sigma_1),0}_{1,s},\frac{1}{2}\frac{\partial^2}{\partial\varepsilon^2}|_{\varepsilon=0}X^{(x_1,\sigma_1),\varepsilon}_{1,s}\right)|\bar{X}^{(x_1,\sigma_1),0}_{1,s}=y\right]. \tag{5.15}
$$

Here, $\left(X^{(x_1,\sigma_1),\varepsilon}_t\right)_{t\geq 0}$ is the solution of the perturbed SDE:

$$
\begin{aligned}
dX^\varepsilon_{1,t} &= \varepsilon\left[-\frac{(\sigma^\varepsilon_t)^2}{2}dt + \sigma^\varepsilon_t S_t dB^1_t\right],\ X_{1,0} = x_1, &(5.16)\\
d\sigma^\varepsilon_t &= \varepsilon\left[\nu\sigma^\varepsilon_t(\rho dB^1_t + \sqrt{1-\rho^2}dB^2_t)\right],\ \sigma^\varepsilon_0 = \sigma_1. &(5.17)
\end{aligned}
$$

On the other hand, we apply a conditional expectation formula for multidimensional asymptotic expansions in Takahashi (1999) in order to evaluate the Malliavin weight $\mathcal{M}^1$ in (5.13).

Table 3 and Table 4 show the results. Again, our scheme with (5.14) and (5.13) (Error rate AE 1st order Weak Approx $n = 2$) improves the base first order expansion (Error rate AE 1st order) especially for the deep OTM calls and puts.

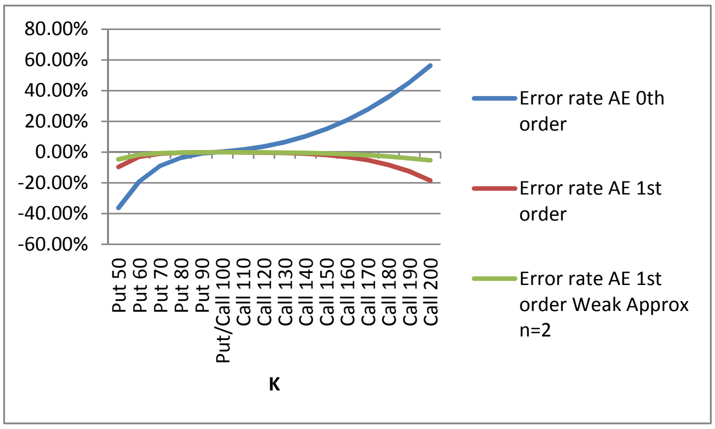

Figure 3: $T = 1$: Stochastic volatility model, Error rate of the 1st order 2 dimensional asymptotic expansion and the weak approximation

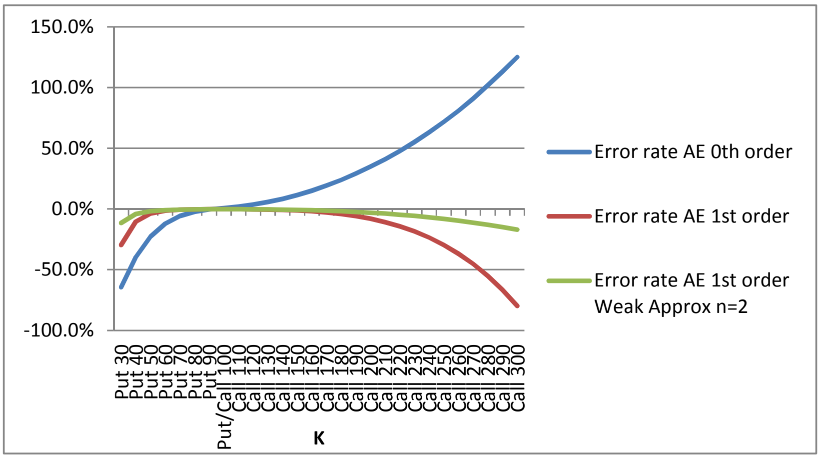

Figure 4: $T = 2$: Stochastic volatility model, Error rate of the 1st order 2 dimensional asymptotic expansion and the weak approximation

6 Concluding Remarks

In this paper, we have shown a new approximation method for the expectations of the functions of the solutions to SDEs applying an asymptotic expansion with Malliavin calculus. In particular, based on Kusuoka (2001, 2003 a, b, 2004) we have obtained an error estimate for our new weak approximation. Moreover, we have confirmed the validity of our method through the numerical examples for option pricing under local and stochastic volatility models. The further numerical examination under higher dimensional SDEs is the next our research topic, which will involve a higher order computational scheme for multidimensional expansions developed by Takahashi (1999) and Takahashi and Toda (2013).

References

[1] C. Bayer, P. Friz and R. Loeffen, Semi-Closed Form Cubature and Applications to Financial Diffusion Models, *Quantitative Finance*, (2013)

[2] D. Crisan, K. Manolarakis and C. Nee, Cubature Methods and Applications, preprint (2013).

[3] M. Fujii, Momentum-Space Approach to Asymptotic Expansion for Stochastic Filtering, forthcoming in *the Annals of Institute of Statistical Mathematics*, (2013).

[4] P.S. Hagan, D. Kumar, A.S. Lesniewski and D.E. Woodward, Managing smile risk, *Willmott Magazine*, 84-108, (2002).

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

[6] S. Kusuoka, Approximation of Expectation of Diffusion Process and Mathematical Finance, Taniguchi Conf. on Math. Nara. '98, *Advanced Studies in Pure Mathematics*, 31, 147-165, (2001).

[7] S. Kusuoka, Malliavin Calculus Revisited, *J. Math. Sci. Univ. Tokyo*, 261-277, (2003a).

[8] S. Kusuoka, Approximation of Expectation of Diffusion Process based on Lie Algebra and Malliavin Calculus, *Mathematical Economics, Kokyuroku* 1337, Research Institute for Mathematical Sciences (RIMS), Kyoto University, 205-209, (2003b).

[9] S. Kusuoka, Approximation of Expectation of Diffusion Process based on Lie Algebra and Malliavin Calculus, *Adv. Math. Econ.*, 6, 69-83, (2004).

[10] S. Kusuoka and D. Stroock, Applications of the Malliavin Calculus Part I, *Stochastic Analysis* (Katata/Kyoto 1982) 271-306 (1984)

[11] C. Li, Managing Volatility Risk: Innovation of Financial Derivatives, Stochastic Models and Their Analytical Implementation, Ph.D thesis in Columbia University, (2010).

[12] T. Lyons and N.Victor, Cubature on the Wiener Space, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 460, 168-198, (2004).

[13] C. Nee, Lecture notes on Gradient bounds for Solutions of stochastic Differential Equations, Applications to numerical schemes, (2010).

[14] C. Nee, Sharp Gradient Bounds for the Diffusion Semigroup, Ph.D thesis in Imperial College London, (2011).

[15] D. Nualart: *The Malliavin Calculus and Related Topics*, Springer, (2006).

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

[17] A. Takahashi, K. Takehara and M. Toda, A General Computation Scheme for a High-Order Asymptotic Expansion Method, *International Journal of Theoretical and Applied Finance*, 15(6), 2012.

[18] A. Takahashi and M. Toda, Note on an Extension of an Asymptotic Expansion Scheme, *International Journal of Theoretical and Applied Finance*, 16(5), 2013.

[19] A. Takahashi. and T. Yamada, An Asymptotic Expansion with Push-Down of Malliavin Weights, *SIAM Journal on Financial Mathematics*, 3, 95-136, (2012a).

[20] A. Takahashi. and T. Yamada, An Asymptotic Expansion for Forward-Backward SDEs: A Malliavin Caluclus Approach, preprint, (2012b).

[21] A. Takahashi. and T. Yamada, On Error Estimates for Asymptotic Expansions with Malliavin Weights -Application to Stochastic Volatility Model-, preprint, (2013a).

[22] A. Takahashi. and T. Yamada, An Asymptotic Expansion for Forward-Backward SDEs with a Perturbed Driver, preprint, (2013b).

[23] H. Tanaka and T. Yamada, Strong Convergence of Euler-Maruyama and Milstein Schemes with Asymptotic Method, Preprint, forthcoming in *International Journal of Theoretical and Applied Finance*, (2013).

[24] S. Violante, Asymptotics of Wiener Functionals and Applications to Mathematical Finance, Ph.D thesis in Imperial College London, (2012).

[25] S. Watanabe, Analysis of Wiener Functionals (Malliavin Calculus) and its Applications to Heat Kernels, *Ann. Probab.* **15**, no. 1, 1-39, (1987).

A Proof of Theorem 1

Using the Taylor formula for $E[f(Y_t^\varepsilon)]$ and the transform $X_t^{x,\varepsilon} = X_t^{x,0} + \varepsilon Y_t^\varepsilon$, we have

$$
\begin{aligned}
&E[f(X_t^{x,\varepsilon})] = E[f(\bar{X}_t^{x,0})] \\
&\quad + \sum_{i=1}^{m} \varepsilon^i \sum_{\alpha^{(k)},\beta^{(k)}}^{i} E\left[f(\bar{X}_t^{x,0}) H_{\alpha_k} \left(\frac{\partial}{\partial \varepsilon} X_t^{x,\varepsilon}|_{\varepsilon=0}, \prod_{l=1}^{k} \frac{1}{\beta_l!} \frac{\partial^{\beta_l}}{\partial \varepsilon^{\beta_l}} X_t^{x,\varepsilon,\alpha_l}|_{\varepsilon=0} \right) \right] \\
&\quad + \varepsilon^{m+1} \int_0^1 (1-u)^m \ (m+1) \sum_{\alpha^{(k)},\beta^{(k)}}^{m+1} E\left[f(X_t^{x,\varepsilon u}) H_{\alpha^{(k)}} \left(Y_t^{\varepsilon u}, \prod_{l=1}^{k} \frac{1}{\beta_l!} \frac{\partial^{\beta_l}}{\partial \eta^{\beta_l}} X_t^{x,\eta,\alpha_l}|_{\eta=\varepsilon u} \right) \right] du.
\end{aligned}
$$

Here,

$$
\sum_{\alpha^{(k)},\beta^{(k)}}^{i} = \sum_{k=1}^{i} \sum_{\sum_{j=1}^{k} \beta_j = i+k,\ \beta_j \geq 2} \ \sum_{\alpha^{(k)}=(\alpha_1,\cdots,\alpha_k) \in \{1,\cdots,N\}^k} \frac{1}{k!}. \tag{A.1}
$$

In order to estimate the residual term using Proposition 1, in the next lemma we characterize the differentiations of the solution to the general perturbed SDEs $X_t^{x,\varepsilon}$ with respect to ε as elements in the space $\mathcal{K}_r$.

Lemma 2.

$$
\frac{1}{j!} \frac{\partial^j}{\partial \varepsilon^j} X_t^{x,\varepsilon} \ \in \ \mathcal{K}_j, \quad j \geq 1,
$$

The proof is given in Appendix B.

Then, using the above lemma and Proposition 1, we obtain the following estimates: For $k \leq m+1$, $\sum_{j=1}^{k} \beta_j = m+1+k$, $\alpha^{(k)} = (\alpha_1, \cdots, \alpha_k) \in \{1, \cdots, N\}^k$,

1.
$$\sup_{x \in \mathbf{R}^N} \left| E\left[f(X_t^{x,\varepsilon}) H_{\alpha^{(k)}} \left(Y_t^{\varepsilon}, \prod_{l=1}^{k} \frac{1}{\beta_l !} \frac{\partial^{\beta_l}}{\partial \varepsilon^{\beta_l}} X_t^{x,\varepsilon,\alpha_l} \right) \right] \right| \leq C s^{(m+1+k)/2} \|\nabla^k f\|_\infty, \quad \text{(A.2)}$$
for any $f \in C_b^\infty(\mathbf{R}^N)$.

2.
$$\sup_{x \in \mathbf{R}^N} \left| E\left[f(X_t^{x,\varepsilon}) H_{\alpha^{(k)}} \left(Y_t^{\varepsilon}, \prod_{l=1}^{k} \frac{1}{\beta_l !} \frac{\partial^{\beta_l}}{\partial \varepsilon^{\beta_l}} X_t^{x,\varepsilon,\alpha_l} \right) \right] \right| \leq C s^{(m+2)/2} \|\nabla f\|_\infty, \quad \text{(A.3)}$$
for any $f \in C_b^1$.

3.
$$\sup_{x \in \mathbf{R}^N} \left| E\left[f(X_t^{x,\varepsilon}) H_{\alpha^{(k)}} \left(Y_t^{\varepsilon}, \prod_{l=1}^{k} \frac{1}{\beta_l !} \frac{\partial^{\beta_l}}{\partial \varepsilon^{\beta_l}} X_t^{x,\varepsilon,\alpha_l} \right) \right] \right| \leq C s^{(m+1)/2} \|f\|_\infty, \quad \text{(A.4)}$$
for an arbitrary bounded continuous function f.

Then, we have the assertion. □

B Proof of Lemma 2

We prove the assertion by induction using slightly modified argument in Takahashi and Yamada (2012, 2013). First, the differentiation of $\frac{\partial}{\partial \varepsilon} X_s^{x,\varepsilon}$ with respect to ε is given by

$$\begin{aligned} \frac{\partial}{\partial \varepsilon} X_s^{x,\varepsilon,l} &= \int_0^t \frac{\partial}{\partial \varepsilon} V_0^l(\varepsilon, X_s^{x,\varepsilon}) ds + \sum_{j=1}^{d} \int_0^t V_j^l(X_s^{x,\varepsilon}) dB_s^j \\ &+ \sum_{k=1}^{N} \int_0^t \partial_k V_0^l(\varepsilon, X_s^{x,\varepsilon}) \frac{\partial}{\partial \varepsilon} X_s^{x,\varepsilon,k} ds \\ &+ \varepsilon \sum_{k=1}^{N} \sum_{j=1}^{d} \int_0^t \partial_k V_j^l(X_s^{x,\varepsilon}) \frac{\partial}{\partial \varepsilon} X_s^{x,\varepsilon,k} dB_s^j, \quad l = 1, \cdots, N. \end{aligned} \quad \text{(B.1)}$$

The above SDE is linear and the order of the Kusuoka-Stroock function $\frac{\partial}{\partial \varepsilon} X_s^{x,\varepsilon}$ is determined by the following term

$$\sum_{j=1}^{d} \int_0^t J_t^{x,\varepsilon} (J_u^{x,\varepsilon})^{-1} V_j(X_u^{x,\varepsilon}) dB_u^j \in \mathcal{K}_1. \quad \text{(B.2)}$$

where $J_t^{x,\varepsilon} = \nabla_x X_t^{x,\varepsilon}$. Since this term gives the minimum order in the terms that consist of (B.1). Here, we use the properties $J_s^{x,\varepsilon}, (J_s^{x,\varepsilon})^{-1} \in \mathcal{K}_0$, $s \in (0,1]$ and the boundness of V_j, $j = 0, \cdots, d$. We have $\frac{\partial}{\partial\varepsilon} X_s^{t,x,\varepsilon} \in \mathcal{K}_1$ by using the properties 2 and 3 in Lemma 1.

For $i \geq 2$, $\frac{1}{i!}\frac{\partial^i}{\partial\varepsilon^i} X_s^{x,\varepsilon} = \left(\frac{1}{i!}\frac{\partial^i}{\partial\varepsilon^i} X_s^{x,\varepsilon,1}, \cdots, \frac{1}{i!}\frac{\partial^i}{\partial\varepsilon^i} X_s^{x,\varepsilon,N}\right)$ is recursively determined by the following:

$$
\begin{aligned}
& \frac{1}{i!}\frac{\partial^i}{\partial\varepsilon^i} X_s^{x,\varepsilon,n} \\
= \; & \frac{1}{i!}\int_0^t \frac{\partial^i}{\partial\varepsilon^i} V_0^n(\varepsilon, X_u^{x,\varepsilon}) du \\
& + \sum_{m=1}^{i} \sum_{i^{(k)},\alpha^{(k)}}^{(m)} \frac{1}{(i-1)!}\frac{1}{k!}\int_0^t \left(\prod_{l=1}^{k} \frac{1}{i_l!}\frac{\partial^{i_l}}{\partial\varepsilon^{i_l}} X_u^{x,\varepsilon,\alpha_l}\right) \sum_{j=1}^{d} \partial_{\alpha^{(k)}} \frac{\partial^{i-m}}{\partial\varepsilon^{i-m}} V_0^n(\varepsilon, X_u^{x,\varepsilon}) du \\
& + \sum_{i^{(k)},\alpha^{(k)}}^{(i-1)} \int_0^t \left(\prod_{l=1}^{k} \frac{1}{i_l!}\frac{\partial^{i_l}}{\partial\varepsilon^{i_l}} X_u^{x,\varepsilon,\alpha_l}\right) \sum_{j=1}^{d} \partial_{\alpha^{(k)}} V_j^n(X_u^{x,\varepsilon}) dB_u^j \\
& + \varepsilon \sum_{i^{(\beta)},\alpha^{(k)}}^{(i)} \int_0^t \left(\prod_{l=1}^{k} \frac{1}{i_l!}\frac{\partial^{i_l}}{\partial\varepsilon^{i_l}} X_u^{x,\varepsilon,\alpha_l}\right) \sum_{j=1}^{d} \partial_{\alpha^{(k)}} V_j^n(X_u^{x,\varepsilon}) dB_u^j, \; n = 1, \cdots, N,
\end{aligned}
\tag{B.3}
$$

where

$$
\sum_{i^{(k)},\alpha^{(k)}}^{(i)} := \sum_{k=1}^{i} \sum_{i_1+\cdots+i_k=i, i_l \geq 1} \sum_{\alpha^{(k)} \in \{1,\cdots,N\}^k} \frac{1}{k!}. \tag{B.4}
$$

The above SDE is linear and the order of the Kusuoka-Stroock function $\frac{1}{i!}\frac{\partial^i}{\partial\varepsilon^i} X_s^{x,\varepsilon}$ is determined inductively by the term

$$
\sum_{i^{(k)},\alpha^{(k)}}^{(i-1)} \int_0^t J_t^{x,\varepsilon} (J_u^{x,\varepsilon})^{-1} \left(\prod_{l=1}^{k} \frac{1}{i_l!}\frac{\partial^{i_l}}{\partial\varepsilon^{i_l}} X_u^{x,\varepsilon,\alpha_l}\right) \sum_{j=1}^{d} \partial_{\alpha^{(k)}} V_j(X_u^{x,\varepsilon}) dB_s^j \; \in \mathcal{K}_i, \tag{B.5}
$$

Since this term gives the minimum order in the terms that consist of (B.3). Then, $\frac{1}{i!}\frac{\partial^i}{\partial\varepsilon^i} X_s^{x,\varepsilon} \in \mathcal{K}_i$ by using the properties 2 and 3 in Lemma 1. $\square$

C Proof of Theorem 2

We follow the similar argument as in Kusuoka (2001,2003b,2004) and Chapter 3 of Crisan et al.(2013).

Note first that we have the following equality:

$$
\begin{aligned}
& P_T f(x) - Q^m_{(s_n)} Q^m_{(s_{n-1})} \cdots Q^m_{(s_1)} f(x) \\
= \;& P_{T-t_{n-1}} P_{t_{n-1}} f(x) - Q^m_{(s_n)} P_{t_{n-1}} f(x) \\
& + Q^m_{(s_n)} P_{t_{n-1}} f(x) - Q^m_{(s_n)} Q^m_{(s_{n-1})} P_{t_{n-2}} f(x) \\
& + \cdots \\
& + Q^m_{(s_n)} \cdots Q^m_{(s_2)} P_{t_1} f - Q^m_{(s_n)} Q^m_{(s_{n-1})} \cdots Q^m_{(s_1)} f \\
= \;& P_{T-t_{n-1}} P_{t_{n-1}} f(x) - Q^m_{(s_n)} P_{t_{n-1}} f(x) \\
& + Q^m_{(s_n)} (P_{s_{n-1}} P_{t_{n-2}} f(x) - Q^m_{(s_{n-1})} P_{t_{n-2}} f(x)) \\
& \cdots \\
& + Q^m_{(s_n)} \cdots Q^m_{(s_2)} (P_{t_1} f(x) - Q^m_{(s_1)} f(x)).
\end{aligned}
$$

Then, since Q^m is a Markov operator, we have

$$
\begin{aligned}
& \|P_T f - Q^m_{(s_n)} Q^m_{(s_{n-1})} \cdots Q^m_{(s_1)} f\|_\infty \\
\leq \;& \|P_{s_n} P_{t_{n-1}} f - Q^m_{(s_n)} P_{t_{n-1}} f\|_\infty \\
& + \|P_{s_{n-1}} P_{t_{n-2}} f - Q^m_{(s_{n-1})} P_{t_{n-2}} f\|_\infty \\
& \cdots \\
& + \|P_{t_1} f - Q^m_{(s_1)} f\|_\infty \\
= \;& \sum_{k=2}^{n} \|P_{s_k} P_{t_{k-1}} f - Q^m_{(s_k)} P_{t_{k-1}} f\|_\infty \\
& + \|P_{t_1} f - Q^m_{(s_1)} f\|_\infty .
\end{aligned}
$$

First, note that we can directly apply (3.14), (3.15) or (3.16) in Theorem1 to obtain an estimate of $\|P_{t_1} f - Q^m_{(s_1)} f\|_\infty$ for $f \in C_b^\infty(\mathbf{R}^N; \mathbf{R})$, a Lipschitz continuous function or a bounded Borel function, respectively. To obtain an estimate of $\sum_{k=2}^{n} \|P_{s_k} P_{t_{k-1}} f - Q^m_{(s_k)} P_{t_{k-1}} f\|_\infty$, we apply the results in Theorem 1 to $P_t f$ (in stead of f) as follows.

- By (3.14) in Theorem 1, for $s, t \in (0,1]$ and $f \in C_b^\infty(\mathbf{R}^N; \mathbf{R})$, there exists C and M such that

$$
\begin{aligned}
\|P_s P_t f - Q^m_{(s)} P_t f\|_\infty & \leq \sum_{l=1}^{M} s^{(m+1+l)/2} C \|\nabla^l P_t f\|_\infty & \text{(C.1)} \\
& \leq \sum_{l=1}^{M} s^{(m+1+l)/2} C \|\nabla^l f\|_\infty . & \text{(C.2)}
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\|P_T f - Q^m_{(s_n)} Q^m_{(s_{n-1})} \cdots Q^m_{(s_1)} f\|_\infty & \leq C \sum_{k=2}^{n} \sum_{l=1}^{M} s_k^{(m+1+l)/2} \|\nabla^l f\|_\infty & \text{(C.3)} \\
& + C \sum_{l=1}^{M} s_1^{(m+1+l)/2} \|\nabla^l f\|_\infty . & \text{(C.4)}
\end{aligned}
$$

- By (3.15) in Theorem 1, for $s,t \in (0,1]$ and $f \in C_b^1(\mathbf{R}^N;\mathbf{R})$, there exists C such that

$$\begin{aligned} \|P_sP_tf - Q^m_{(s)}P_tf\|_\infty &\leq s^{(m+2)/2}C\|\nabla P_tf\|_\infty && \text{(C.5)} \\ &\leq s^{(m+2)/2}C\|\nabla f\|_\infty. && \text{(C.6)} \end{aligned}$$

Hence,

$$\begin{aligned} \|P_Tf - Q^m_{(s_n)}Q^m_{(s_{n-1})}\cdots Q^m_{(s_1)}f\|_\infty &\leq C\sum_{k=2}^n s_k^{(m+2)/2}\|\nabla f\|_\infty && \text{(C.7)} \\ &\quad +Cs_1^{(m+2)/2}\|\nabla f\|_\infty. && \text{(C.8)} \end{aligned}$$

- By (3.16) in Theorem 1, for $s,t \in (0,1]$ and bounded Borel function f on $\mathbf{R}^N$, there exists C such that

$$\begin{aligned} \|P_sP_tf - Q^m_{(s)}P_tf\|_\infty &\leq s^{(m+1)/2}C\|P_tf\|_\infty && \text{(C.9)} \\ &\leq s^{(m+1)/2}C\|f\|_\infty. && \text{(C.10)} \end{aligned}$$

Hence,

$$\begin{aligned} \|P_Tf - Q^m_{(s_n)}Q^m_{(s_{n-1})}\cdots Q^m_{(s_1)}f\|_\infty &\leq C\sum_{k=2}^n s_k^{(m+1)/2}\|f\|_\infty && \text{(C.11)} \\ &\quad +Cs_1^{(m+1)/2}\|f\|_\infty. && \text{(C.12)} \end{aligned}$$

Next, we obtain more explicit and compact expressions with regard to n particularly for (C.3), (C.7) and (C.11).

Firstly, from the definition of s_k for $k \in \{2,\cdots,n\}$, we have

$$s_k = \frac{\gamma T(k-1)^{\gamma-1}}{n^\gamma}\int_{k-1}^k (u/(k-1))^{\gamma-1}du. \tag{C.13}$$

For $k \in \{2,\cdots,n\}$, $(u/(k-1))^{\gamma-1} \leq \max\{(k/(k-1))^{\gamma-1},1\} \leq \max\{2^{\gamma-1},1\}$. Then,

$$\begin{aligned} s_k^{l/2} &\leq \left(\frac{\gamma T(k-1)^{\gamma-1}}{n^\gamma}\max\{2^{\gamma-1},1\}\right)^{l/2} && \text{(C.14)} \\ &\leq C(1/n)^{\gamma l/2}(k-1)^{(\gamma-1)l/2} && \text{(C.15)} \end{aligned}$$

where $C = C(T,\gamma)$.

We consider the estimates for three different ranges of γ that are larger than, equal to and less than $(l-2)/l$, respectively. ($\gamma = (l-2)/l$ satisfies $(\gamma-1)l/2 = -1$.)

For $0 < \gamma < (l-2)/l$,

$$C(1/n)^{\gamma l/2}\sum_{k=2}^n (k-1)^{(\gamma-1)l/2} \leq C(1/n)^{\gamma l/2}. \tag{C.16}$$

For $\gamma = (l-2)/l$

$$
\begin{aligned}
& C(1/n)^{\gamma l/2} \sum_{k=2}^{n} (k-1)^{(\gamma-1)l/2} && \text{(C.17)} \\
= \;& C(1/n)^{(l-2)/2} \sum_{k=1}^{n} (k-1)^{-1} && \text{(C.18)} \\
\leq \;& C(1/n)^{(l-2)/2} \log n. && \text{(C.19)}
\end{aligned}
$$

For $\gamma > (l-2)/l$

$$
\begin{aligned}
& C(1/n)^{\gamma l/2} \sum_{k=2}^{n} (k-1)^{(\gamma-1)l/2} && \text{(C.20)} \\
= \;& C(1/n)^{(\gamma-1)l/2} (1/n)^{l/2} \sum_{k=2}^{n} (k-1)^{(\gamma-1)l/2} && \text{(C.21)} \\
= \;& C(1/n)^{(l-2)/2} \sum_{k=2}^{n} \left(\frac{k-1}{n} \right)^{(\gamma-1)l/2} \frac{1}{n} && \text{(C.22)} \\
\leq \;& C(1/n)^{(l-2)/2}. && \text{(C.23)}
\end{aligned}
$$

Then, by combining an estimate of $\|P_{t_1} f - Q_{(s_1)}^{m} f\|_\infty$ for $f \in C_b^\infty(\mathbf{R}^N; \mathbf{R})$, a Lipschitz continuous function or a bounded Borel function, we have the assertion. $\square$