

A new higher-order weak approximation scheme for stochastic differential equations and the Runge–Kutta method

Mariko Ninomiya · Syoiti Ninomiya

Received: 12 December 2007 / Accepted: 2 January 2009 / Published online: 28 May 2009
© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract The authors report on the construction of a new algorithm for the weak approximation of stochastic differential equations. In this algorithm, an ODE-valued random variable whose average approximates the solution of the given stochastic differential equation is constructed by using the notion of free Lie algebras. It is proved that the classical Runge–Kutta method for ODEs is directly applicable to the ODE drawn from the random variable. In a numerical experiment, this is applied to the problem of pricing Asian options under the Heston stochastic volatility model. Compared with some other methods, this algorithm is significantly faster.

Keywords Free Lie algebra · Mathematical finance · Runge–Kutta method · Stochastic differential equations · Weak approximation

Mathematics Subject Classification (2000) 65C30 · 65C05 · 65L06 · 17B01 · 91B02

JEL Classification C63 · G12

This research was partly supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (C), 15540110, 2003 and 18540113, 2006, the 21st century COE program at Graduate School of Mathematical Sciences, the University of Tokyo, and JSPS Core-to-Core Program 18005.

M. Ninomiya (✉)
Graduate School of Mathematical Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku,
Tokyo 153-8914, Japan
e-mail: mariko.nnmy@gmail.com

S. Ninomiya
Center for Research in Advanced Financial Technology, Tokyo Institute of Technology, 2-12-1
Ookayama, Meguro-ku, Tokyo 152-8552, Japan
e-mail: ninomiya@craft.titech.ac.jp

1 Introduction

1.1 Background

In applied sciences, finding numerical solutions of stochastic differential equations (SDEs) is crucial. For example, the price of a financial derivative is obtained by the calculation of $E[f(X(T, x))]$ where $X(t, x)$ is the value at time t of the N -dimensional diffusion process which describes the asset price and f is the payoff. Therefore it is highly important to find fast and reliable algorithms for the numerical evaluation of $E[f(X(T, x))]$. A number of studies on this problem have been conducted [12, 15].

There are two approaches to the problem: the partial differential equation (PDE) approach and simulation. The former involves solving the partial differential equation (1.2) numerically. This method works only when the dimension is relatively small. We do not go into details on the subject here but refer to [21]. This condition is not necessarily satisfied in many practical problems; so we are forced to take the other approach in which we calculate $E[f(X(T, x))]$ directly by using the distribution of $X(T, x)$. When we take this approach, the problem can be classified into the following three cases:

- (1) The law of $X(T, x)$ is known explicitly.
- (2) The Fourier transform of $X(T, x)$ is known explicitly.
- (3) None of both.

In the last case, numerical schemes based on the stochastic asymptotic expansion are applied in order to construct a set of random variables $\{X^{(n)}(t_i, x)\}_{i=0}^n$ ($0 = t_0 < t_1 < \dots < t_n = T$) that approximates $X(t, x)$ in a precise sense. $\{X^{(n)}(t_i, x)\}_{i=0}^n$ is called the discretization of $X(t, x)$ or the discretized process of $X(t, x)$. Here n corresponds to the number of “ticks” included in $[0, T]$. This approach is called the probabilistic method or simulation. In this paper, we focus on that approach.

Now, $X^{(n)}(T, x)$ defines an \mathbb{R}^N -valued function defined over a finite-dimensional domain. We denote by $D(n)$ the number of dimensions of the domain. Then we have a map $X^{(n)}(T, x)(\cdot) : \mathbb{R}^{D(n)} \rightarrow \mathbb{R}^N$. $D(n)$ is determined by the discretization scheme, but at least it is an increasing function with respect to n , regardless of the scheme.

Using a simple volume transformation, we can reduce the problem to numerical integration over the unit cube $[0, 1]^{D(n)}$. One usually calculates such integrals by Monte Carlo or quasi-Monte Carlo methods because $D(n)$ becomes very large in practical problems.

Two types of approximation errors are thus involved in the calculation. One is the difference between $E[f(X(T, x))]$ and $E[f(X^{(n)}(T, x))]$ and the other is, if one uses the Monte Carlo method, the difference between $\text{MC}(f(X^{(n)}(T, x)), M)(\omega)$ and $E[f(X^{(n)}(T, x))]$, or if the quasi-Monte Carlo method is used, the difference between $\text{QMC}(f(X^{(n)}(T, x)), M)$ and $E[f(X^{(n)}(T, x))]$. Here, for a random variable W , $\text{MC}(W, M)$ denotes the random variable $(\sum_{i=1}^M W_i)/M$ where the W_i are independent random variables whose distributions are identical to that of W , and $\text{QMC}(W, M) = (\sum_{i=1}^M W(p_i))/M$ where $\{p_i\}_{i=1,2,\dots}$ is a deterministic sequence

generated by a low-discrepancy sequence [24]. In this paper, we call the former error discretization error and the latter integration error.

When the discretization error is $O(p)$, we say that the discretization scheme is of order p . When one simulates $X(t, x)$, usually the Euler–Maruyama scheme, which is of order 1, is used. By using higher-order schemes we can reduce $D(n)$, i.e., we can reduce the problem to a lower-dimensional integration problem. In the case of quasi-Monte Carlo methods, this means that the number of sample points needed for the numerical integration of $E[X^{(n)}(T, x)]$ is reduced. Detailed discussions are given in Sect. 6.

1.2 The problem and our results

1.2.1 The problem

Let (Ω, \mathcal{F}, P) be a probability space, $B^0(t) = t$, and $(B^1(t), \dots, B^d(t))$ a d -dimensional standard Brownian motion. $C_b^\infty(\mathbb{R}^N; \mathbb{R}^N)$ denotes the set of \mathbb{R}^N -valued infinitely differentiable functions defined on \mathbb{R}^N whose derivatives are all bounded. Our interest is in weak approximation, that is to say, approximations of the function $(P_t f)(x) = E[f(X(t, x))]$ where $f \in C_b^\infty(\mathbb{R}^N; \mathbb{R})$ and $X(t, x)$ is a solution to the stochastic differential equation written in Stratonovich form as

$$X(t, x) = x + \sum_{i=0}^d \int_0^t V_i(X(s, x)) \circ dB^i(s), \tag{1.1}$$

where $V_i \in C_b^\infty(\mathbb{R}^N; \mathbb{R}^N)$ for $i = 0, 1, \dots, d$. Here, $V_i \in C_b^\infty(\mathbb{R}^N; \mathbb{R}^N)$ is considered to be a vector field via

$$V_i f(x) = \sum_{j=1}^N V_i^j(x) \frac{\partial f}{\partial x_j}(x), \quad \text{for } f \in C_b^\infty(\mathbb{R}^N; \mathbb{R}).$$

It is well known (e.g. [14]) that $E[f(X(1, x))]$ is equal to $u(1, x)$ where u is the solution to the partial differential equation, with $L = V_0 + (1/2) \sum_{i=1}^d V_i^2$,

$$\frac{\partial u}{\partial t}(t, x) = Lu, \quad u(0, x) = f(x). \tag{1.2}$$

Usually, the Euler–Maruyama scheme is used to discretize $X(t, x)$ during simulations to weakly approximate $X(t, x)$. It is shown in [20, 25, 26, 30] that the new higher-order scheme introduced by Kusuoka in [17] calculates some finance problems much faster than the Euler–Maruyama scheme. Lyons and Victoir extensively developed the scheme in [23] using the notion of free Lie algebras.

Recent developments can be found in [2, 11], where weak, higher-order approximation schemes for SPDEs are introduced. There an additional feature appears. In the case of an SPDE, the size of n should be small since the larger n , the more PDEs have to be solved.

We shall discuss the reason why higher-order schemes greatly improve the speed of numerical weak approximation in the later part of this paper (Sect. 6).

1.2.2 Our results

In this paper, we describe how we successfully constructed in Theorem 1.3 and Corollary 1.4 a new higher-order weak approximation scheme for a broad class of stochastic differential equations. This scheme owes a great deal to the scheme shown in [17] and to the cubature method on Wiener space introduced in [23].

An intuitive explanation of our scheme is as follows. We construct an ODE (ordinary differential equation)-valued random variable whose average approximates the given stochastic differential equation. From this random variable, an ODE itself can be drawn at one time.

This scheme has the remarkable advantage that once an ODE is drawn, the conventional Runge–Kutta method can be applied so as to approximate the ODE. The approximating random variable is constructed using Theorems 1.3 and 1.6 and can be approximated by the Runge–Kutta method for ODEs via Theorem 4.15.

We should note that another higher-order weak approximation method based on the same approach [17, 23] is introduced in [27]. Although the algorithm in [27] and the new method presented in this paper share the same approach and have many common features, the algorithms themselves differ significantly.

1.3 Notation

Let $A = \{v_0, v_1, \dots, v_d\}$ be an alphabet where $d \in \mathbb{Z}_{\geq 1}$ and A^* denotes the set of all words consisting of the elements of A . The empty word 1 is the identity of A^* . For $u = v_{i_1} \cdots v_{i_n} \in A^*$, $|u|$ and $\|u\|$ are defined by $|u| = n$ and $\|u\| = |u| + \text{card}(\{k \mid i_k = 0\})$ where $\text{card}(S)$ denotes the cardinality of a set S . Here, $\|\cdot\|$ is related to the scaling property of the Brownian motion. A_m^* and $A_{\leq m}^*$ denote $\{w \in A^* \mid |w| = m\}$ and $\{w \in A^* \mid |w| \leq m\}$, respectively. Let $\mathbb{R}\langle A \rangle$ be the free \mathbb{R} -algebra with basis A^* and $\mathbb{R}\langle\langle A \rangle\rangle$ be the set of all \mathbb{R} -coefficient formal series with basis A^* . Then, $\mathbb{R}\langle A \rangle$ is a sub- \mathbb{R} -algebra of $\mathbb{R}\langle\langle A \rangle\rangle$. We call an element of $\mathbb{R}\langle A \rangle$ a non-commutative polynomial. $P \in \mathbb{R}\langle\langle A \rangle\rangle$ is written as

$$P = \sum_{w \in A^*} (P, w)w \quad \text{or} \quad \sum_{w \in A^*} a_w w,$$

where $(P, w) = a_w \in \mathbb{R}$ denotes the coefficient of w . Let

$$\mathbb{R}\langle A \rangle_m = \{P \in \mathbb{R}\langle A \rangle \mid (P, w) = 0, \text{ if } \|w\| \neq m\}.$$

The algebra structure is defined as usual, i.e.,

$$\left(\sum_{w \in A^*} a_w w \right) \left(\sum_{w \in A^*} b_w w \right) = \sum_{\substack{w=uv \\ w \in A^*}} a_u b_v w.$$

The Lie bracket is defined as $[x, y] = xy - yx$ for $x, y \in \mathbb{R}\langle\langle A \rangle\rangle$. For $w = v_{i_1} \cdots v_{i_n} \in A^*$, $\tau(w)$ denotes $[v_{i_1}, [v_{i_2}, [\dots, [v_{i_{n-1}}, v_{i_n}] \dots]]]$. We define $\mathcal{L}_{\mathbb{R}}(A)$ as the set of Lie polynomials in $\mathbb{R}\langle A \rangle$ and $\mathcal{L}_{\mathbb{R}}(\langle\langle A \rangle\rangle)$ as the set of Lie series. This means that $\mathcal{L}_{\mathbb{R}}(A)$ is the smallest \mathbb{R} -submodule of $\mathbb{R}\langle A \rangle$ including A and closed under the

Lie bracket, and that $\mathcal{L}_{\mathbb{R}}(\langle A \rangle)$ is the set of elements of $\mathbb{R}\langle\langle A \rangle\rangle$ whose homogeneous components belong to $\mathcal{L}_{\mathbb{R}}(A)$. We note that Lie polynomials correspond to vector fields while general polynomials do not necessarily. For $m \in \mathbb{Z}_{\geq 0}$, let j_m be the map defined by

$$j_m \left(\sum_{w \in A^*} a_w w \right) = \sum_{\|w\| \leq m} a_w w.$$

For arbitrary $P, Q \in \mathbb{R}\langle A \rangle$, the inner product $\langle P, Q \rangle$ is defined by

$$\langle P, Q \rangle = \sum_{w \in A^*} (P, w)(Q, w).$$

Moreover, we let $\|P\|_2 = (\langle P, P \rangle)^{1/2}$ for $P \in \mathbb{R}\langle A \rangle$. For $P \in \mathbb{R}\langle\langle A \rangle\rangle$ with $\langle P, 1 \rangle = 0$, we can define $\exp(P)$ as $1 + \sum_{k=1}^{\infty} P^k/k!$. In addition, $\log(Q)$ can be defined as $\sum_{k=1}^{\infty} (-1)^{k-1}(Q - 1)^k/k$ for $Q \in \mathbb{R}\langle\langle A \rangle\rangle$ with $\langle Q, 1 \rangle = 1$. Then we have the relations

$$\log(\exp(P)) = P \quad \text{and} \quad \exp(\log(Q)) = Q.$$

We can induce the direct product topology into $\mathbb{R}\langle\langle A \rangle\rangle$ by the identification $\mathbb{R}\langle\langle A \rangle\rangle \approx \mathbb{R}^{A^*} \approx \mathbb{R}^{\infty}$. We note that the first identification is not trivial ([9], Chap. 2, Sect. 4). Then $\mathbb{R}\langle\langle A \rangle\rangle$ becomes a Polish space with this topology. We can also consider its Borel σ -algebra $\mathcal{B}(\mathbb{R}\langle\langle A \rangle\rangle)$, $\mathbb{R}\langle\langle A \rangle\rangle$ -valued random variables, their expectations, and other notions as usual.

Let Φ be the homomorphism between $\mathbb{R}\langle A \rangle$ and the \mathbb{R} -algebra consisting of smooth differential operators over \mathbb{R}^N such that

$$\begin{aligned} \Phi(1) &= \text{Id}, \\ \Phi(v_{i_1} \cdots v_{i_n}) &= V_{i_1} \cdots V_{i_n} \quad \text{for } i_1, \dots, i_n \in \{0, 1, \dots, d\}. \end{aligned}$$

Considering the scaling property of the Brownian motion, we define the rescaling operator Ψ_s depending on $\|\cdot\|$. For $s \in \mathbb{R}_{>0}$, $\Psi_s : \mathbb{R}\langle\langle A \rangle\rangle \rightarrow \mathbb{R}\langle\langle A \rangle\rangle$ is defined by

$$\Psi_s \left(\sum_{m=0}^{\infty} P_m \right) = \sum_{m=0}^{\infty} s^{m/2} P_m \quad \text{where } P_m \in \mathbb{R}\langle A \rangle_m.$$

For a smooth vector field V , i.e., an element of $C_b^{\infty}(\mathbb{R}^N; \mathbb{R}^N)$, $\exp(V)(x)$ denotes the solution at time 1 of the ordinary differential equation

$$\frac{dz_t}{dt} = V(z_t), \quad z_0 = x.$$

We also define $\|V\|_{C^n}$ for $V \in C_b^{\infty}(\mathbb{R}^N; \mathbb{R}^N)$ by

$$\begin{aligned} \|V\| &= \sup \{ |V(x)|; x \in \mathbb{R}^N \}, \\ \|V^{(n)}\| &= \sup \{ |V_{(x)}^{(n)}(U_1, U_2, \dots, U_n)|; x \in \mathbb{R}^N \text{ and } |U_i| = 1, \text{ for } i = 1, \dots, n \}, \\ \|V\|_{C^n} &= \sum_{i=0}^n \|V^{(i)}\|. \end{aligned}$$

Here $V^{(k)}$ denotes the k th order total differential of V , i.e.,

$$V_{(x)}^{(n)}(U_1, U_2, \dots, U_n) = \sum_{i=1}^N \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{\partial^n V_i}{\partial x_{j_1} \cdots \partial x_{j_n}}(x) U_1^{j_1} \cdots U_n^{j_n} e_i$$

where each e_i denotes an N -dimensional unit vector, $\{e_1, \dots, e_N\}$ forms an orthonormal basis of \mathbb{R}^N , and U_k^j is the j th component of $U_k \in \mathbb{R}^N$.

1.4 Main results

Since in this paper we deal with operators that are not necessarily linear with respect to time t , we introduce the following definition.

Definition 1.1 A map g from $C_b^\infty(\mathbb{R}^N; \mathbb{R}^N)$ to the set of all maps from \mathbb{R}^N to \mathbb{R}^N is called an *integration scheme of order m* if there exists a positive constant C_m such that

$$\sup_{x \in \mathbb{R}^N} |g(W)(x) - \exp(W)(x)| \leq C_m \|W\|_{C^{m+1}}^{m+1}$$

for all $W \in C_b^\infty(\mathbb{R}^N; \mathbb{R}^N)$. Let $\mathcal{IS}(m)$ be the set of all integration schemes of order m .

This definition is a generalization of the usual order of approximation.

Definition 1.2 For $z_1, z_2 \in \mathcal{L}_{\mathbb{R}}((A))$, we define $z_2 \vdash z_1$ as $\log(\exp(z_2) \exp(z_1))$. Then from the definition, for $z_1, z_2, z_3 \in \mathcal{L}_{\mathbb{R}}((A))$,

$$(z_1 \vdash z_2) \vdash z_3 = \log(\exp(z_1) \exp(z_2) \exp(z_3)) = z_1 \vdash (z_2 \vdash z_3),$$

and so we can write for $z_1, \dots, z_n \in \mathcal{L}_{\mathbb{R}}((A))$

$$z_1 \vdash z_2 \vdash \cdots \vdash z_n = \log(\exp(z_1) \cdots \exp(z_n)).$$

We notice that $z_2 \vdash z_1 \in \mathcal{L}_{\mathbb{R}}((A))$ if $z_1, z_2 \in \mathcal{L}_{\mathbb{R}}((A))$ from the Baker–Campbell–Hausdorff formula ([4], Chap. II, 6.4).

The following are the main results.

Theorem 1.3 Let $m \geq 1$, $M \geq 2$, and Z_1, \dots, Z_M be $\mathcal{L}_{\mathbb{R}}((A))$ -valued random variables. Assume that Z_1, \dots, Z_M satisfy

$$\begin{aligned} Z_i &= j_m Z_i \quad \text{for } i = 1, \dots, M, \\ E[\|j_m Z_i\|_2] &< \infty \quad \text{for } i = 1, \dots, M, \\ E\left[\exp\left(a \sum_{j=1}^M \|\Phi(\Psi_s(Z_j))\|_{C^{m+1}}\right)\right] &< \infty \quad \text{for any } a > 0. \end{aligned} \tag{1.3}$$

Then for $p \in [1, \infty)$ and arbitrary $g_1, \dots, g_M \in \mathcal{IS}(m)$, there exists a positive constant $C_{m,M}$ such that

$$\left\| \sup_{x \in \mathbb{R}^N} \left| g_1(\Phi(\Psi_s(Z_1))) \circ \dots \circ g_M(\Phi(\Psi_s(Z_M)))(x) - \exp(\Phi(\Psi_s(j_m(Z_M \mathop{\text{H}} \dots \mathop{\text{H}} Z_1))))(x) \right\|_{L^p} \leq C_{m,M} s^{(m+1)/2} \tag{1.4}$$

for $s \in (0, 1]$ where $C_{m,M}$ depends only on m and M . Here for functions f and g , $f \circ g(x)$ denotes $f(g(x))$ as usual.

For $i = 1, \dots, d$, and $j = 1, \dots, M$, let S_j^i be \mathbb{R} -valued Gaussian random variables and for $j, j' = 1, \dots, M$, let c_j and $R_{jj'}$ be real numbers such that

$$\sum_{j=1}^M c_j = 1, \quad E[S_j^i] = 0, \quad \text{and} \quad E[S_j^i S_{j'}^{i'}] = R_{jj'} \delta_{ii'} \tag{1.5}$$

for $i, i' = 1, \dots, d$. We let $S_j^0 = c_j$ for convenience. Taking (1.2) into account, we let Z_1, \dots, Z_M be random variables such that $Z_j = c_j v_0 + \sum_{i=1}^d S_j^i v_i$ for $j = 1, \dots, M$ and that

$$E[j_m(\exp(Z_1) \cdots \exp(Z_M))] = j_m \left(\exp \left(v_0 + \frac{1}{2} \sum_{i=1}^d v_i^2 \right) \right). \tag{1.6}$$

In usual ODE cases, this type of approximation technique is known as a splitting method [13]. The stochastic versions of this technique are considered in [23, 27].

Recently, it is proved in a constructive way that for any $m \geq 2$ there exists a set of random variables Z_1, \dots, Z_M that satisfy (1.6) ([35]).

Corollary 1.4 *Suppose that the following UFG condition is satisfied:*

(UFG) *There exist an integer ℓ and $\varphi_{u,u'} \in C_b^\infty(\mathbb{R}^N; \mathbb{R})$ which satisfy*

$$\Phi(\tau(u)) = \sum_{u' \in A_{\leq \ell}^* \setminus \{1, v_0\}} \varphi_{u,u'} \Phi(\tau(u'))$$

for any $u \in A^* \setminus \{1, v_0\}$.

For $j = 1, \dots, M$ let Z_j be $\mathcal{L}_{\mathbb{R}}((A))$ -valued random variables constructed as above and define linear operators $Q_{(s)}$ for $s \in (0, 1]$ by

$$(Q_{(s)} f)(x) = E[f(g(\Phi(\Psi_s(Z_1))) \circ \dots \circ g(\Phi(\Psi_s(Z_M))))(x)]$$

where $f \in C_b^\infty(\mathbb{R}^N; \mathbb{R})$ and $g \in \mathcal{IS}(m)$. Then

$$\|P_s f - Q_{(s)} f\|_\infty \leq C s^{(m+1)/2} \|\text{grad}(f)\|_\infty$$

where C is a positive constant.

Remark 1.5 In [19], it is shown that for the operator $Q_{(s)}$ defined above, there exists a constant C such that

$$(P_s f)(x) - (Q_{(s)} f)(x) = Cs^{(m+1)/2} + O(s^{(m+3)/2})$$

holds. This means that the Romberg extrapolation can be applied to our new algorithm.

The intuitive understanding is that once we find the random variables Z_1, \dots, Z_M , we can numerically approximate $\exp(Z_i(\omega))$ by applying the integration scheme g_i repeatedly for each i as seen in (1.4) in Theorem 1.3. Therefore, our primary interest is in finding Z_1, \dots, Z_M .

Theorem 1.6 *Let $m = 5$ and $M = 2$. Then (1.6) holds if and only if*

$$\begin{aligned} c_1 &= \frac{\mp \sqrt{2(2u-1)}}{2}, & c_2 &= 1 \pm \frac{\sqrt{2(2u-1)}}{2}, & R_{11} &= u, \\ R_{22} &= 1 + u \pm \sqrt{2(2u-1)}, & R_{12} &= -u \mp \frac{\sqrt{2(2u-1)}}{2} \end{aligned} \tag{1.7}$$

for some $u \geq 1/2$.

Remark 1.7 We can show that in the case where $m = 7$ and $M = 3$ there is no solution to (1.6).

Now that we have obtained the random variables satisfying (1.6) for $m = 5$, we need a practical way of approximating these integration schemes g_1, \dots, g_M . We successfully extend the applicability of the general Runge–Kutta method to ODEs to find that it belongs to $\mathcal{IS}(m)$.

Let $A = (a_{ij})_{i,j=1,\dots,K}$ with $a_{ij} \in \mathbb{R}$ and $b = {}^t(b_1, \dots, b_K) \in \mathbb{R}^K$. If (A, b) satisfies the m th-order conditions defined as (4.2) in Sect. 4, the K -stage Runge–Kutta method of order m in the sense of [6] can be written as

$$\begin{aligned} Y_i(W, s) &= y_0 + s \sum_{j=1}^K a_{ij} W(Y_j(W, s)), \\ Y(y_0; W, s) &= y_0 + s \sum_{i=1}^K b_i W(Y_i(W, s)) \end{aligned} \tag{1.8}$$

for $W \in C_b^\infty(\mathbb{R}^N; \mathbb{R}^N)$, $s \in \mathbb{R}_{>0}$, and $y_0 \in \mathbb{R}^N$. Let $g(W)(y_0)$ be $Y(y_0; W, 1)$. We show that g belongs to $\mathcal{IS}(m)$ in Theorem 4.15.

Remark 1.8 Our scheme is fundamentally different from the class of numerical methods sometimes referred to as stochastic Runge–Kutta methods [5, 28, 29].

2 Proof of Theorem 1.3

For simplicity of notation, we let $\Phi_s(y)$ denote $\Phi(\Psi_s(y))$ for an element $y \in \mathcal{L}_{\mathbb{R}}(\langle A \rangle)$ in the following part. We split the left-hand side of (1.4) as

$$\begin{aligned} & \left\| \sup_{x \in \mathbb{R}^N} \left| g_1(\Phi_s(Z_1)) \circ \dots \circ g_M(\Phi_s(Z_M))(x) \right. \right. \\ & \quad \left. \left. - \exp(\Phi_s(j_m(Z_M \mathbb{H} \dots \mathbb{H} Z_1)))(x) \right\|_{L^p} \\ & \leq \left\| \sup_{x \in \mathbb{R}^N} \left| \exp(\Phi_s(Z_1)) \circ \dots \circ \exp(\Phi_s(Z_M))(x) \right. \right. \\ & \quad \left. \left. - \exp(\Phi_s(j_m(Z_M \mathbb{H} \dots \mathbb{H} Z_1)))(x) \right\|_{L^p} \\ & \quad + \left\| \sup_{x \in \mathbb{R}^N} \left| g_1(\Phi_s(Z_1)) \circ \dots \circ g_M(\Phi_s(Z_M))(x) \right. \right. \\ & \quad \left. \left. - \exp(\Phi_s(Z_1)) \circ \dots \circ \exp(\Phi_s(Z_M))(x) \right\|_{L^p}. \end{aligned} \tag{2.1}$$

The evaluation of each term of the right-hand side of (2.1) will be given by Lemma 2.6 or (2.9) in this section.

Proposition 2.1

(1) For any $V \in C_b^\infty(\mathbb{R}^N; \mathbb{R}^N)$, $f \in C^\infty(\mathbb{R}^N; \mathbb{R})$, $x \in \mathbb{R}^N$ and $n \geq 1$,

$$f(\exp(tV)(x)) = \sum_{k=0}^n \frac{t^k}{k!} (V^k f)(x) + \int_0^t \frac{(t-s)^n}{n!} (V^{n+1} f)(\exp(sV)(x)) ds. \tag{2.2}$$

(2) For all $z \in \mathcal{L}_{\mathbb{R}}(\langle A \rangle)$ and $n, m \geq 1$,

$$\begin{aligned} & \sup_{x \in \mathbb{R}^N} \left| f(\exp(\Phi(j_m z))(x)) - \sum_{k=0}^n \frac{1}{k!} (\Phi((j_m z)^k) f)(x) \right| \\ & \leq \frac{1}{(n+1)!} \|\Phi((j_m z)^{n+1}) f\|_\infty. \end{aligned} \tag{2.3}$$

Proof Since we have

$$\frac{d}{dt} f(\exp(tV)(x)) = (Vf)(\exp(tV)(x))$$

from the Taylor expansion, we obtain (2.2) by integration by parts and (2.3) can be derived from (2.2). □

Lemma 2.2 For all $n \geq 1$, there exists a constant $C_n > 0$ such that

$$\|\Phi(j_n z) f\|_\infty \leq C_n \|j_n z\|_2 \|\mathbf{grad}(f)\|_{C^{n-1}}$$

for all $z \in \mathcal{L}(\langle A \rangle)$ and $f \in C^\infty(\mathbb{R}^N; \mathbb{R})$.

Proof Let p_m be a map such that

$$p_m : \sum_{|\alpha|=0}^{\infty} a_{\alpha} D^{\alpha} \mapsto \sum_{|\alpha|=m} a_{\alpha} D^{\alpha}$$

where α is a multi-index, $a_{\alpha} \in C_b^{\infty}(\mathbb{R}^N; \mathbb{R})$, and $D^{\alpha} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_N^{\alpha_N}}$. Then we have

$$\Phi(w) = \sum_{i=1}^{|w|} p_i(\Phi(w)),$$

for $w \in A^* \setminus \{1\}$. Since there exists a constant $C_{w,i} > 0$ such that

$$\|p_i(\Phi(w))f\|_{\infty} \leq C_{w,i} \sup_{\substack{\alpha \in (\mathbb{Z}_{\geq 0})^N \\ |\alpha|=i-1}} \|D^{\alpha}(\text{grad}(f))\|_{\infty},$$

we see that there exists a constant $C'_n > 0$ such that

$$\begin{aligned} \|\Phi(j_n z)f\|_{\infty} &\leq \sum_{\substack{w \in A^* \\ 1 \leq \|w\| \leq n}} \|\Phi(w)f\|_{\infty} |\langle z, w \rangle| \\ &\leq \sum_{\substack{w \in A^* \\ 1 \leq \|w\| \leq n}} \sum_{i=1}^{|w|} C_{w,i} |\langle z, w \rangle| \sup_{\substack{\alpha \in (\mathbb{Z}_{\geq 0})^N \\ |\alpha|=i-1}} \|D^{\alpha}(\text{grad}(f))\|_{\infty} \\ &\leq C'_n \|j_n z\|_2 \sup_{\substack{\alpha \in (\mathbb{Z}_{\geq 0})^N \\ |\alpha| \leq n-1}} \|D^{\alpha}(\text{grad}(f))\|_{\infty} \\ &\leq C_n \|j_n z\|_2 \|\text{grad}(f)\|_{C^{n-1}}, \end{aligned}$$

where $C'_n = \text{card}(\{w \in A^* \mid 1 \leq \|w\| \leq n\}) \sup_{\substack{w \in A^* \\ 1 \leq \|w\| \leq n}} (\sum_{i=1}^{|w|} C_{w,i})$. □

Lemma 2.3

(1) *There exists a constant $C_{m,1} > 0$ such that*

$$\begin{aligned} \sup_{x \in \mathbb{R}^N} |f(\exp(\Phi_s(j_m z))(x)) - (\Phi_s(j_m \exp(j_m z))f)(x)| \\ \leq C_{m,1} s^{(m+1)/2} (1 + \|j_m z\|_2)^{m+1} \|\text{grad}(f)\|_{C^{m(m+1)-1}} \end{aligned} \tag{2.4}$$

for $z \in \mathcal{L}_{\mathbb{R}}((A))$.

(2) *There exists a constant $C_{m,M} > 0$ where $M \in \mathbb{Z}_{\geq 2}$ such that*

$$\begin{aligned} \sup_{x \in \mathbb{R}^N} |f(\exp(\Phi_s(j_m((j_m z_M) \text{H} \cdots \text{H}(j_m z_1))))(x)) \\ - (\Phi_s(j_m \exp(j_m((j_m z_M) \text{H} \cdots \text{H}(j_m z_1))))f)(x)| \\ \leq C_{m,M} s^{(m+1)/2} \left(1 + \sum_{i=1}^M \|j_m z_i\|_2\right)^{m+1} \|\text{grad}(f)\|_{C^{m(m+1)-1}} \end{aligned} \tag{2.5}$$

for $z_1, \dots, z_M \in \mathcal{L}_{\mathbb{R}}((A))$.

Proof From the fact that for $z \in \mathcal{L}_{\mathbb{R}}((A))$, we have

$$j_m(\exp(j_m z)) = \sum_{k=0}^m \frac{1}{k!} (j_m z)^k - \sum_{k=2}^m \frac{1}{k!} (j_m(m+1) - j_m)((j_m z)^k)$$

and from (2.3) in Proposition 2.1, we see that

$$\begin{aligned} & \left| f(\exp(\Phi(j_m z))(x)) - (\Phi(j_m(\exp(j_m z))))f(x) \right| \\ & \leq \frac{1}{(m+1)!} \|\Phi((j_m z)^{m+1})f\|_{\infty} + \left| \sum_{k=2}^m \frac{1}{k!} (\Phi((j_m(m+1) - j_m)((j_m z)^k))f)(x) \right|. \end{aligned}$$

Since we have

$$(j_m z)^{m+1} = (j_m(m+1) - j_m)(j_m z)^{m+1},$$

applying Lemma 2.2 yields

$$\begin{aligned} & \left| f(\exp(\Phi(j_m z))(x)) - (\Phi(j_m(\exp(j_m z))))f(x) \right| \\ & \leq \sum_{k=2}^{m+1} \frac{1}{k!} \|\Phi((j_m(m+1) - j_m)((j_m z)^k))f\|_{\infty} \\ & \leq C_m \sum_{k=2}^{m+1} \|(j_m(m+1) - j_m)(j_m z)^k\|_2 \|\text{grad}(f)\|_{C^{m(m+1)-1}} \\ & \leq C_{m,1} (1 + \|j_m z\|_2)^{m+1} \|\text{grad}(f)\|_{C^{m(m+1)-1}} \end{aligned}$$

where C_m and $C_{m,1}$ are positive constants. Thus (2.4) is proved.

Taking $(j_m z_M) \mathbb{H} \cdots \mathbb{H} (j_m z_1)$ as z above and evaluating by

$$\begin{aligned} & \sum_{k=2}^{m+1} \|(j_m(m+1) - j_m)(j_m((j_m z_M) \mathbb{H} \cdots \mathbb{H} (j_m z_1)))^k\|_2 \\ & \leq C_{m,M} \left(1 + \sum_{i=1}^M \|j_m z_i\|_2 \right)^{m+1}, \end{aligned}$$

we obtain (2.5). □

Lemma 2.4 *There exists a constant $C_{m,M} > 0$ such that*

$$\begin{aligned} & \sup_{x \in \mathbb{R}^N} \left| f(\exp(\Phi_s(j_m z_1)) \circ \cdots \circ \exp(\Phi_s(j_m z_M)))(x) \right. \\ & \quad \left. - (\Phi_s(j_m \exp(j_m((j_m z_M) \mathbb{H} \cdots \mathbb{H} (j_m z_1))))f)(x) \right| \\ & \leq C_{m,M} s^{(m+1)/2} \sum_{i=1}^M (1 + \|j_m z_i\|_2)^{m+1} \|\text{grad}(f)\|_{C^{m(m+M)-1}} \end{aligned} \tag{2.6}$$

for $z_1, \dots, z_M \in \mathcal{L}_{\mathbb{R}}((A))$. Here $C_{m,M}$ depends on m and M .

Proof We prove the lemma by induction on M . When $M = 1$, (2.4) and (2.6) are equivalent. Assume that (2.6) holds for M . Splitting the left-hand side of (2.6) for $M + 1$ as

$$\begin{aligned} & \left| f(\exp(\Phi_s(j_m z_1)) \circ \cdots \circ \exp(\Phi_s(j_m z_{M+1}))(x)) \right. \\ & \quad \left. - (\Phi_s(j_m \exp(j_m((j_m z_{M+1}) \mathbb{H} \cdots \mathbb{H}(j_m z_1))))f)(x) \right| \\ & \leq \left| f(\exp(\Phi_s(j_m z_1)) \circ \cdots \circ \exp(\Phi_s(j_m z_{M+1}))(x)) \right. \\ & \quad \left. - (\Phi_s(j_m \exp(j_m((j_m z_M) \mathbb{H} \cdots \mathbb{H}(j_m z_1))))f)(\exp(\Phi_s(j_m z_{M+1}))(x)) \right| \\ & \quad + \left| (\Phi_s(j_m \exp(j_m((j_m z_M) \mathbb{H} \cdots \mathbb{H}(j_m z_1))))f)(\exp(\Phi_s(j_m z_{M+1}))(x)) \right. \\ & \quad \left. - (\Phi_s(j_m \exp(j_m((j_m z_{M+1}) \mathbb{H} \cdots \mathbb{H}(j_m z_1))))f)(x) \right|, \end{aligned}$$

we can apply the induction hypothesis and (2.4) with

$$\Phi_s(j_m \exp((j_m z_M) \mathbb{H} \cdots \mathbb{H}(j_m z_1)))f$$

instead of f to obtain

$$\begin{aligned} & \sup_{x \in \mathbb{R}^N} \left| f(\exp(\Phi_s(j_m z_1)) \circ \cdots \circ \exp(\Phi_s(j_m z_{M+1}))(x)) \right. \\ & \quad \left. - (\Phi_s(j_m \exp(j_m((j_m z_{M+1}) \mathbb{H} \cdots \mathbb{H}(j_m z_1))))f)(x) \right| \\ & \leq C_1 s^{(m+1)/2} \left(\sum_{i=1}^M (1 + \|j_m z_i\|_2)^{m+1} + (1 + \|j_m z_{M+1}\|_2)^{m+1} \right) \\ & \quad \times \|\text{grad}(f)\|_{C^{m(m+M+1)-1}}, \end{aligned}$$

where $C_1 > 0$ is a constant depending on m and M . Hence, (2.6) is proved. □

From Lemmas 2.3 and 2.4, we have the following result.

Lemma 2.5 *For all $m \geq 1$, there exists a constant $C_{m,M} > 0$ such that*

$$\begin{aligned} & \sup_{x \in \mathbb{R}^N} \left| f(\exp(\Phi_s(j_m z_1)) \circ \cdots \circ \exp(\Phi_s(j_m z_M))(x)) \right. \\ & \quad \left. - f(\exp(\Phi_s(j_m((j_m z_M) \mathbb{H} \cdots \mathbb{H}(j_m z_1))))(x)) \right| \\ & \leq C_{m,M} s^{(m+1)/2} \sum_{i=1}^M (1 + \|j_m z_i\|_2)^{m+1} \|\text{grad}(f)\|_{C^{m(m+M)-1}} \end{aligned}$$

for all $s \in (0, 1]$, $z_1, \dots, z_M \in \mathcal{L}_{\mathbb{R}}((A))$, and $f \in C^\infty(\mathbb{R}^N; \mathbb{R})$.

Lemma 2.6 *Let Z_1, \dots, Z_M be $\mathcal{L}_{\mathbb{R}}((A))$ -valued random variables such that for $m \geq 1$, $E[\|j_m Z_i\|_2] < \infty$ for $i = 1, \dots, M$. Then, for $p \in [1, \infty)$ there exists a constant $C_{m,M} > 0$ such that*

$$\left\| \sup_{x \in \mathbb{R}^N} \left| \exp(\Phi_s(j_m Z_1)) \circ \dots \circ \exp(\Phi_s(j_m Z_M))(x) - \exp(\Phi_s(j_m((j_m Z_M) \text{H} \dots \text{H}(j_m Z_1))))(x) \right| \right\|_{L^p} \leq C_{m,M} s^{(m+1)/2} \tag{2.7}$$

for any $s \in (0, 1]$.

Proof If for $i \in \{1, \dots, N\}$, $f((x^1, \dots, x^N)) = x^i$, then $\|\text{grad}(f)\|_{C^{m(m+M)-1}} = 1$ for all $m \geq 1$. Therefore, applying Lemma 2.5 for this f , we obtain (2.7). \square

We note that in [31] a similar result to this lemma is obtained.

We now start the discussion about the latter term of the right-hand side of (2.1).

Proposition 2.7 *There exists a constant $C > 0$ such that*

$$|g(W)(x) - g(W)(y)| \leq C \|W\|_{C^{m+1}}^{m+1} + |x - y| \exp(\|W\|_{C^1}) \tag{2.8}$$

for $g \in \mathcal{IS}(m)$ and $W \in C_b^\infty(\mathbb{R}^N; \mathbb{R}^N)$.

Proof Since Gronwall’s inequality gives

$$|\exp(W)(x) - \exp(W)(y)| \leq |x - y| \exp(\|W\|_{C^1}),$$

(2.8) can be derived. \square

Since $g_i \in \mathcal{IS}(m)$ and each Z_i satisfies (1.3), we see that for some $C_1 > 0$,

$$\begin{aligned} & \left\| \sup_{x \in \mathbb{R}^N} |g_M(\Phi_s(Z_M))(x) - \exp(\Phi_s(Z_M))(x)| \right\|_{L^p} \\ & \leq \|C_m\| \|\Phi_s(Z_M)\|_{C^{m+1}}^{m+1} \|_{L^p} \leq C_1 s^{(m+1)/2}. \end{aligned}$$

From this fact and Proposition 2.7, there exists a constant $C_4 > 0$ such that

$$\begin{aligned} & \left\| \sup_{x \in \mathbb{R}^N} |g_{M-1}(\Phi_s(Z_{M-1})) \circ g_M(\Phi_s(Z_M))(x) - \exp(\Phi_s(Z_{M-1})) \circ \exp(\Phi_s(Z_M))(x)| \right\|_{L^p} \\ & \leq \left\| \sup_{x \in \mathbb{R}^N} |g_{M-1}(\Phi_s(Z_{M-1})) \circ \exp(\Phi_s(Z_M))(x) - \exp(\Phi_s(Z_{M-1})) \circ \exp(\Phi_s(Z_M))(x)| \right\|_{L^p} \\ & \quad + \left\| \sup_{x \in \mathbb{R}^N} |g_{M-1}(\Phi_s(Z_{M-1})) \circ g_M(\Phi_s(Z_M))(x) - g_{M-1}(\Phi_s(Z_{M-1})) \circ \exp(\Phi_s(Z_M))(x)| \right\|_{L^p} \end{aligned}$$

$$\begin{aligned} &\leq C_2 s^{(m+1)/2} + \left\| C_3 \|\Phi_s(Z_{M-1})\|_{C^{m+1}}^{m+1} \right. \\ &\quad \left. + \sup_{x \in \mathbb{R}^N} |g_M(\Phi_s(Z_M))(x) - \exp(\Phi_s(Z_M))(x)| \exp(\|\Phi_s(Z_{M-1})\|_{C^1}) \right\|_{L^p} \\ &\leq C_4 s^{(m+1)/2} \end{aligned}$$

where C_2 and C_3 are positive constants. Inductively,

$$\begin{aligned} &\left\| \sup_{x \in \mathbb{R}^N} |g_1(\Phi_s(Z_1)) \circ \dots \circ g_M(\Phi_s(Z_M))(x) \right. \\ &\quad \left. - \exp(\Phi_s(Z_1)) \circ \dots \circ \exp(\Phi_s(Z_M))(x) \right\|_{L^p} \leq C_5 s^{(m+1)/2} \end{aligned} \tag{2.9}$$

where $C_5 > 0$.

Lemma 2.6 and (2.9) complete the proof of Theorem 1.3.

3 Necessary conditions for the $\mathcal{L}_{\mathbb{R}}((A))$ -valued random variables Z_1, \dots, Z_M

Lemma 3.1 For $i = 1, \dots, M$, let Y_i be Gaussian random variables such that

$$E[Y_i] = 0 \quad \text{and} \quad E[Y_i Y_j] = R(i, j), \quad \text{for } i, j = 1, \dots, M$$

where $R(i, j) \in \mathbb{R}$. Moreover, for $i = 1, \dots, M$ let $m_i \in \mathbb{Z}_{\geq 0}$ be such that $\sum_{i=1}^M m_i$ is even. Then we have

$$\begin{aligned} &E[Y_1^{m_1} \dots Y_M^{m_M}] \\ &= \sum_{\{d_{ij}\}_{1 \leq i \leq j \leq M} \in e(m_1, \dots, m_M)} 2^{-\sum_{i=1}^M d_{ii}} \frac{\prod_{i=1}^M (m_i!)}{\prod_{1 \leq i \leq j \leq M} (d_{ij}!)} \prod_{1 \leq i \leq j \leq M} R(i, j)^{d_{ij}} \end{aligned} \tag{3.1}$$

where $e(m_1, \dots, m_M)$ is a set of $\{d_{ij}\}_{1 \leq i \leq j \leq M}$ satisfying that $d_{ij} \in \mathbb{Z}_{\geq 0}$ and

$$\sum_{1 \leq j < i} d_{ji} + 2d_{ii} + \sum_{i < j \leq M} d_{ij} = m_i$$

for $i = 1, \dots, M$.

Proof Let $\ell = \sum_{i=1}^M m_i$. We have

$$\begin{aligned} &E[Y_1^{m_1} \dots Y_M^{m_M}] \\ &= E \left[\frac{\partial^\ell}{\partial z_1^{m_1} \dots \partial z_M^{m_M}} \exp \left(\sum_{i=1}^M z_i Y_i \right) \right] \Bigg|_{z=0} \end{aligned}$$

$$\begin{aligned}
 &= \frac{\partial^\ell}{\partial z_1^{m_1} \dots \partial z_M^{m_M}} \left(\exp \left(\frac{1}{2} \sum_{1 \leq i, j \leq M} R(i, j) z_i z_j \right) \right) \Big|_{z=0} \\
 &= \frac{1}{2^{\ell/2} (\ell/2)!} \frac{\partial^\ell}{\partial z_1^{m_1} \dots \partial z_M^{m_M}} \left(\sum_{1 \leq i, j \leq M} R(i, j) z_i z_j \right)^{\ell/2} \Big|_{z=0} \tag{3.2}
 \end{aligned}$$

where $z = (z_1, \dots, z_M) \in \mathbb{R}^M$.

Let

$$e_\ell = \left\{ \{d_{ij}\}_{1 \leq i \leq j \leq M} \mid d_{ij} \in \mathbb{Z}_{\geq 0} \text{ and } \sum_{1 \leq i \leq j \leq M} d_{ij} = \frac{\ell}{2} \right\}.$$

Then

$$\begin{aligned}
 &\left(\sum_{1 \leq i, j \leq M} R(i, j) z_i z_j \right)^{\ell/2} \\
 &= \left(\sum_{i=1}^M R(i, i) z_i^2 + 2 \sum_{1 \leq i < j \leq M} R(i, j) z_i z_j \right)^{\ell/2} \\
 &= \sum_{\{d_{ij}\}_{1 \leq i \leq j \leq M} \in e_\ell} \frac{(\ell/2)!}{\prod_{1 \leq i \leq j \leq M} (d_{ij}!)} \prod_{i=1}^M (R(i, i) z_i^2)^{d_{ii}} \prod_{1 \leq i < j \leq M} (2R(i, j) z_i z_j)^{d_{ij}} \\
 &= \sum_{\{d_{ij}\}_{1 \leq i \leq j \leq M} \in e_\ell} \frac{(\ell/2)!}{\prod_{1 \leq i \leq j \leq M} (d_{ij}!)} 2^{\sum_{1 \leq i < j \leq M} d_{ij}} \\
 &\quad \times \left(\prod_{1 \leq i \leq j \leq M} R(i, j)^{d_{ij}} \right) \left(\prod_{i=1}^M z_i^{\sum_{1 \leq j < i} d_{ji} + 2d_{ii} + \sum_{i < j \leq M} d_{ij}} \right).
 \end{aligned}$$

Hence

$$\begin{aligned}
 &\frac{\partial^\ell}{\partial z_1^{m_1} \dots \partial z_M^{m_M}} \left(\sum_{1 \leq i, j \leq M} R(i, j) z_i z_j \right)^{\ell/2} \Big|_{z=0} \\
 &= (m_1! \dots m_M!) \left(\frac{\ell}{2} \right)! \\
 &\quad \times \sum_{\{d_{ij}\}_{1 \leq i \leq j \leq M} \in e(m_1, \dots, m_M)} \frac{2^{\sum_{1 \leq i \leq j \leq M} d_{ij}}}{\prod_{1 \leq i \leq j \leq M} (d_{ij}!)} \prod_{1 \leq i \leq j \leq M} R(i, j)^{d_{ij}}. \tag{3.3}
 \end{aligned}$$

Since we have from the definition of $e(m_1, \dots, m_M)$ that

$$\sum_{1 \leq i < j \leq M} d_{ij} = \frac{\ell}{2} - \sum_{i=1}^M d_{ii}$$

for $\{d_{ij}\} \in e(m_1, \dots, m_M)$, (3.1) is derived from (3.2) and (3.3). □

We need next a simple representation of the coefficient of each $v_{i_1} v_{i_2} \cdots v_{i_\ell}$ in $E[\exp(Z_1) \cdots \exp(Z_M)]$ where $(i_1, \dots, i_\ell) \in \{0, 1, \dots, d\}^\ell$ and Z_1, \dots, Z_M are $\mathcal{L}_{\mathbb{R}}((A))$ -valued random variables constructed with Gaussian random variables satisfying (1.5).

For $\ell, M \in \mathbb{Z}_{>0}$, let

$$\mathcal{K}_\ell(M) = \{\mathbf{k} = (k_1, \dots, k_M) \in (\mathbb{Z}_{\geq 0})^M \mid k_1 + \cdots + k_M = \ell\}.$$

For $w = v_{i_1} \cdots v_{i_\ell} \in A^*$, let $N^w : \{0, 1, \dots, d\} \times \{1, \dots, M\} \times \mathcal{K}_\ell(M) \rightarrow \mathbb{Z}_{\geq 0}$ be a function such that

$$N^w(i, j, \mathbf{k}) = \text{card}(\{r \mid i_r = i \text{ for } k_1 + \cdots + k_{j-1} + 1 \leq r \leq k_1 + \cdots + k_j\}).$$

Theorem 3.2 *Let $w = v_{i_1} v_{i_2} \cdots v_{i_\ell} \in A^*$ and*

$$n^w(i) = \text{card}(\{j \in \{1, \dots, \ell\} \mid i_j = i\})$$

for $i = 1, \dots, d$. Then the coefficient $C(w)$ of w in $E[\exp(Z_1) \cdots \exp(Z_M)]$ becomes as follows:

If $n^w(i)$ is odd for some $i \in \{1, \dots, d\}$, then

$$C(w) = 0. \tag{3.4}$$

If $n^w(i)$ is even for every $i \in \{1, \dots, d\}$, then

$$\begin{aligned} C(w) &= \sum_{\mathbf{k}=(k_1, \dots, k_M) \in \mathcal{K}_\ell(M)} \frac{1}{k_1! \cdots k_M!} \prod_{j=1}^M (c_j)^{N^w(0, j, \mathbf{k})} \\ &\quad \times \prod_{p=1}^d \left(\sum_{\substack{\{d_{ij}\}_{1 \leq i \leq j \leq M} \in \\ e(N^w(p, \mathbf{1}, \mathbf{k}), \dots, N^w(p, M, \mathbf{k}))}} 2^{-\sum_{i=1}^M d_{ij}} \right. \\ &\quad \left. \times \frac{\prod_{j=1}^M (N^w(p, j, \mathbf{k})!)}{\prod_{1 \leq i \leq j \leq M} (d_{ij}!)} \prod_{1 \leq i \leq j \leq M} R_{ij}^{d_{ij}} \right) \end{aligned} \tag{3.5}$$

where c_j and R_{ij} are real numbers defined in (1.5).

Proof In the case where $n^w(i)$ is odd for some $i \in \{1, \dots, d\}$, (3.4) is directly derived from (1.5). We therefore consider the other case. By the Taylor expansion of $\exp(Z_1) \cdots \exp(Z_M)$, we have that

$$\begin{aligned} &E[\exp(Z_1) \cdots \exp(Z_M)] \\ &= \sum_{k_1, \dots, k_M=0}^\infty \frac{1}{k_1! \cdots k_M!} E \left[\left(c_1 v_0 + \sum_{i=1}^d S_1^i v_i \right)^{k_1} \cdots \left(c_M v_0 + \sum_{i=1}^d S_M^i v_i \right)^{k_M} \right]. \end{aligned}$$

Hence

$$\begin{aligned}
 C(w) &= \langle E[\exp(Z_1) \cdots \exp(Z_M)], w \rangle \\
 &= \sum_{\mathbf{k}=(k_1, \dots, k_M) \in \mathcal{K}_\ell(M)} \frac{1}{k_1! \cdots k_M!} \\
 &\quad \times E \left[\underbrace{S_1^{i_1} \cdots S_1^{i_{k_1}}}_{k_1} \underbrace{S_2^{i_{k_1+1}} \cdots S_2^{i_{k_1+k_2}}}_{k_2} \cdots \underbrace{S_M^{i_{k_1+\cdots+k_{M-1}+1}} \cdots S_M^{i_{k_1+\cdots+k_M}}}_{k_M} \right] \\
 &= \sum_{\mathbf{k}=(k_1, \dots, k_M) \in \mathcal{K}_\ell(M)} \frac{1}{k_1! \cdots k_M!} \\
 &\quad \times E \left[(c_1)^{N^w(0,1,\mathbf{k})} \cdots (c_M)^{N^w(0,M,\mathbf{k})} (S_1^1)^{N^w(1,1,\mathbf{k})} \cdots \right. \\
 &\quad \left. \times (S_M^1)^{N^w(1,M,\mathbf{k})} \cdots (S_1^d)^{N^w(d,1,\mathbf{k})} \cdots (S_M^d)^{N^w(d,M,\mathbf{k})} \right].
 \end{aligned}$$

From the definition of S_j^i ,

$$\begin{aligned}
 C(w) &= \sum_{\substack{\mathbf{k}=(k_1, \dots, k_M) \\ \in \mathcal{K}_\ell(M)}} \frac{1}{k_1! \cdots k_M!} \\
 &\quad \times \prod_{j=1}^M (c_j)^{N^w(0,j,\mathbf{k})} \prod_{p=1}^d E[(S_1^p)^{N^w(p,1,\mathbf{k})} \cdots (S_M^p)^{N^w(p,M,\mathbf{k})}].
 \end{aligned}$$

Applying (3.1) from Lemma 3.1, we obtain (3.5). □

On the other hand, the value of the coefficient of each $v_{i_1} \cdots v_{i_\ell}$ in the expression $j_m(\exp(v_0 + (1/2) \sum_{i=1}^d v_i^2))$ can be obtained by the following result.

Proposition 3.3 *Let $A^0 = \{v_0, v_1 v_1, v_2 v_2, \dots, v_d v_d\} \subset A^*$. Then*

$$\exp\left(v_0 + \frac{1}{2} \sum_{i=1}^d v_i^2\right) = \sum_{\substack{w=w_1 \cdots w_\ell \\ w_1, \dots, w_\ell \in A^0}} \frac{1}{2^{|w|-\ell} \ell!} w. \tag{3.6}$$

Therefore, taking $\{S_j^i\}_{i=1, \dots, d, j=1, \dots, M}$ to equate (3.4) or (3.5) with (3.6) for $w = v_{i_1} v_{i_2} \cdots v_{i_\ell}$ with $\|w\| \leq m$, we can construct Z_1, \dots, Z_M .

For $m = 5$, we take $M = 2$ to obtain solvable simultaneous equations which in fact become the five equations

$$\begin{aligned}
 c_1 + c_2 &= 1, & \frac{1}{2}(c_1 R_{11} + c_2 R_{22}) + R_{12} &= \frac{1}{2}, \\
 \frac{1}{6}(c_1 R_{11} + c_2 R_{22}) + \frac{1}{2}c_1(R_{12} + R_{22}) &= \frac{1}{4}, \\
 \frac{1}{6}(c_1 R_{11} + c_2 R_{22}) + \frac{1}{2}c_2(R_{11} + R_{22}) &= \frac{1}{4}, \\
 \frac{1}{24}(R_{11}^2 + R_{22}^2) + \frac{1}{6}R_{12}(R_{11} + R_{22}) + \frac{1}{4}R_{11}R_{22} &= \frac{1}{8}.
 \end{aligned}$$

The solution is (1.7). Since we let $\{S_j^i\}_{i=1,\dots,d, j=1,\dots,M}$ be a Gaussian system, such random variables can be constructed.

Remark 3.4 If we let $m = 5$, then M must be at least two.

4 The Runge–Kutta method

We begin by briefly introducing the tree theory following [3, 6, 7]. For details of the Runge–Kutta method, see [6, 7, 28].

All trees introduced here are called directed or rooted trees in the literature listed above.

Definition 4.1 A labeled tree \mathbf{t} is a pair of finite sets $(V(\mathbf{t}), E(\mathbf{t}))$ that satisfies the conditions

- (1) $V(\mathbf{t}) \subset \mathbb{Z}$, $V(\mathbf{t}) \neq \emptyset$, and $E(\mathbf{t}) \subset \{(x, y) \in V(\mathbf{t}) \times V(\mathbf{t}) : x < y\}$.
- (2) For each $x \in V(\mathbf{t})$, if $(x, y) \in E(\mathbf{t})$ and $(x', y) \in E(\mathbf{t})$, then $x = x'$.
- (3) For two distinct elements $x, y \in V(\mathbf{t})$, one of the followings holds:
 - (i) There exists a path from x to y .
 - (ii) There exists a path from y to x .
 - (iii) For some $z \in V(\mathbf{t}) \setminus \{x, y\}$, there exist paths z to x and z to y .

Here a path from p_1 to p_ℓ is a sequence $(p_1, p_2), (p_2, p_3), \dots, (p_{\ell-1}, p_\ell)$ of elements of $E(\mathbf{t})$.

An element of $V(\mathbf{t})$ is called a *vertex* of \mathbf{t} and that of $E(\mathbf{t})$ is called an *edge* of \mathbf{t} .

A particular labeled tree τ_ℓ is that with $\text{card}(V(\tau_\ell)) = 1$ and $E(\tau_\ell) = \emptyset$.

For a labeled tree $\mathbf{t} = (V(\mathbf{t}), E(\mathbf{t}))$, let $\mathbf{r}(\mathbf{t})$ be $\text{card}(V(\mathbf{t}))$. We define \mathbf{T} as the set of all labeled trees.

Proposition 4.2 For each $\mathbf{t} = (V(\mathbf{t}), E(\mathbf{t}))$, there exists a unique vertex $r \in V(\mathbf{t})$ such that for any $x \in V(\mathbf{t}) \setminus \{r\}$, there is a path from r to x .

Such a vertex r is called the *root* of \mathbf{t} . Here, τ_ℓ consists of only the root.

Definition 4.3 For $i = 1, \dots, n$, let $\mathbf{t}_i = (V(\mathbf{t}_i), E(\mathbf{t}_i)) \in \mathbf{T}$ be such that $V(\mathbf{t}_i) \cap V(\mathbf{t}_j) = \emptyset$ if $i \neq j$. Then $[\mathbf{t}_1 \cdots \mathbf{t}_n]$ is defined as $\mathbf{t} = (V(\mathbf{t}), E(\mathbf{t})) \in \mathbf{T}$ such that

$$V(\mathbf{t}) = \{r\} \cup V(\mathbf{t}_1) \cup \cdots \cup V(\mathbf{t}_n),$$

$$E(\mathbf{t}) = \{(r, r_1), \dots, (r, r_n)\} \cup E(\mathbf{t}_1) \cup \cdots \cup E(\mathbf{t}_n),$$

where each r_i denotes the root of \mathbf{t}_i and $r = \min\{r_1, \dots, r_n\} - 1$.

Remark 4.4 For $\mathbf{t}_1, \dots, \mathbf{t}_n \in \mathbf{T}$, we have that

$$[\mathbf{t}_1 \cdots \mathbf{t}_n] = [\mathbf{t}_{\varpi(1)} \cdots \mathbf{t}_{\varpi(n)}]$$

for any permutation $\varpi \in \mathfrak{S}_n$.

Definition 4.5 Let $\mathbf{t}_i = (V(\mathbf{t}_i), E(\mathbf{t}_i)) \in \mathbf{T}$ for $i = 1, 2$. We say that \mathbf{t}_1 and \mathbf{t}_2 are *isomorphic*, written as $\mathbf{t}_1 \sim \mathbf{t}_2$, if there exists a bijection $\varpi : V(\mathbf{t}_1) \rightarrow V(\mathbf{t}_2)$ such that $(x, y) \in E(\mathbf{t}_1)$ if and only if $(\varpi(x), \varpi(y)) \in E(\mathbf{t}_2)$.

In particular, when $\mathbf{t}_1 \sim \mathbf{t}_2$ and $V(\mathbf{t}_1) = V(\mathbf{t}_2)$, that is, ϖ is a permutation, we say that \mathbf{t}_1 and \mathbf{t}_2 are *equivalent* and write $\mathbf{t}_1 \sim \mathbf{t}_2$.

Proposition 4.6 Both \sim and \sim are equivalence relations.

Proposition 4.7 Let $\mathbf{t}_i = (V(\mathbf{t}_i), E(\mathbf{t}_i)) \in \mathbf{T}$ and $\mathbf{u}_i = (V(\mathbf{u}_i), E(\mathbf{u}_i)) \in \mathbf{T}$ for $i = 1, \dots, n$. Suppose that $\mathbf{t}_i \sim \mathbf{u}_i$ for $i = 1, \dots, n$ and that

$$V(\mathbf{t}_i) \cap V(\mathbf{t}_j) = \emptyset \quad \text{and} \quad V(\mathbf{u}_i) \cap V(\mathbf{u}_j) = \emptyset$$

if $i \neq j$. Then

$$[\mathbf{t}_1 \cdots \mathbf{t}_n] \sim [\mathbf{u}_1 \cdots \mathbf{u}_n].$$

Definition 4.8 We define $T = \mathbf{T}/\sim$. An element $t \in T$ is called a *non-labeled tree*. For a labeled tree $\mathbf{t} \in \mathbf{T}$, $|\mathbf{t}|$ denotes the corresponding non-labeled tree $t \in T$.

Then, from Proposition 4.7, the following result can be derived.

Proposition 4.9 Under the same condition as Proposition 4.7,

$$|[\mathbf{t}_1 \cdots \mathbf{t}_n]| = |[\mathbf{u}_1 \cdots \mathbf{u}_n]|$$

holds.

By virtue of Proposition 4.9, we can define a non-labeled tree $t = [t_1 \cdots t_n]$ for $t_1, \dots, t_n \in T$ as $[[\mathbf{t}_1 \cdots \mathbf{t}_n]]$ where $\mathbf{t}_i \in \mathbf{T}$ is a representative labeled tree such that $|\mathbf{t}_i| = t_i$. In particular, we let $\tau = |\tau_\ell|$.

Proposition 4.10 For any $t \in T \setminus \{\tau\}$, there exist $t_1, \dots, t_n \in T$ such that $t = [t_1 \cdots t_n]$. Moreover

$$[t_1 \cdots t_n] = [t_{\varpi(1)} \cdots t_{\varpi(n)}]$$

for any permutation $\varpi \in \mathfrak{S}_n$.

Here, $[t_1^{m_1} \cdots t_n^{m_n}]$ denotes $[\underbrace{t_1 \cdots t_1}_{m_1} \cdots \underbrace{t_n \cdots t_n}_{m_n}]$ where $t_i \in T$ for $i = 1, \dots, n$.

Definition 4.11

(1) For $t = (V(t), E(t)) \in T$, we define $\alpha : T \rightarrow \mathbb{Z}_{\geq 1}$, $r : T \rightarrow \mathbb{Z}_{\geq 1}$, and $\sigma : T \rightarrow \mathbb{Z}_{\geq 1}$ by

$$\alpha(t) = \text{card}(\{\mathbf{u} \in \mathbf{T} \mid \mathbf{u} \sim \mathbf{t} \text{ where } \mathbf{t} \in \mathbf{T} \text{ is a representative element with } |\mathbf{t}| = t\}),$$

$$r(t) = \text{card}(V(t)),$$

$$\sigma(t) = \begin{cases} 1 & \text{if } t = \tau \\ \prod_{i=1}^{\ell} m_i! \sigma(t_i)^{m_i} & \text{if } t = [t_1^{m_1} \cdots t_{\ell}^{m_{\ell}}], \ell \geq 1. \end{cases}$$

(2) Let \mathcal{A} be the set of $K \times K$ real matrices. We inductively define derivative weights $\zeta_i : T \times \mathcal{A} \rightarrow \mathbb{R}$ for $i = 1, \dots, K$ by

$$\zeta_i(t; A) = \begin{cases} \sum_{j=1}^K a_{ij} & \text{if } t = \tau, \\ \sum_{j=1}^K a_{ij} \prod_{k=1}^{\ell} \zeta_j(t_k; A) & \text{if } t = [t_1 \cdots t_{\ell}], \ell \geq 1 \end{cases}$$

where $A = (a_{ij})_{i,j=1,\dots,K}$. We notice that α is well defined because α denotes the number of ways a tree may be labeled. In addition, we define the elementary differentials $D : C_b^{\infty}(\mathbb{R}^N; \mathbb{R}^N) \times T \rightarrow C_b^{\infty}(\mathbb{R}^N; \mathbb{R}^N)$ by

$$D(W, \tau)(x) = W(x)$$

and for $t = [t_1 t_2 \cdots t_{\ell}]$, $\ell \geq 1$,

$$D(W, t)(x) = W^{(\ell)}(x)(D(W, t_1)(x), D(W, t_2)(x), \dots, D(W, t_{\ell})(x)).$$

Let $y(W, s)$ be a solution to the ODE

$$\frac{d}{ds} y(W, s) = W(y(W, s)), \quad y(W, 0) = y_0 \tag{4.1}$$

where $W \in C_b^{\infty}(\mathbb{R}^N; \mathbb{R}^N)$ and $y_0 \in \mathbb{R}^N$.

Let $T_m = \{t \in T : r(t) = m\}$ and $T_{\leq m} = \bigsqcup_{n=0}^m T_n$ for $m \geq 0$ with $T_0 = \emptyset$. Then we have the following lemmas essentially proved in [7], pp. 139–145.

Lemma 4.12 For $m \in \mathbb{Z}_{\geq 1}$,

$$y^{(m)}(W, s) = \sum_{\mathbf{t} \in \mathbf{T}_m} D(W, |\mathbf{t}|)(y(W, s)).$$

Let $A_{ex} \in \mathcal{A}$ denote A in (1.8) for the explicit Runge–Kutta method. Then $Y_i(W, s)$ is definitely determined by A_{ex} with $a_{ij} = 0$ if $i \leq j$ and so $Y(y_0; W, s)$ can be constructed with b and $Y_i(W, s)$ as both seen in (1.8).

Lemma 4.13 *Let $m \geq 1$. If there exists a constant $C_m > 0$ such that*

$$\left| Y_i(W, s) - \left(y_0 + \sum_{t \in T_{\leq m-1}} s^{r(t)} \frac{\zeta_i(t)}{\sigma(t)} D(W, t)(y_0) \right) \right| \leq C_m s^m \|W\|_{C_m}^m$$

for $i = 1, \dots, K$, then there exists a constant C_{m+1} such that

$$\begin{aligned} & \left| sW(Y_i(W, s)) - \sum_{\substack{\ell=1, \dots, m-1 \\ t=[t_1 \dots t_\ell] \in T_{\leq m}}} s^{r(t)} \frac{\prod_{k=1}^\ell \zeta_i(t_k)}{\sigma(t)} D(W, t)(y_0) \right| \\ & \leq C_{m+1} s^{m+1} \|W\|_{C_{m+1}}^{m+1}. \end{aligned}$$

Applying these lemmas to evaluations of the solution to (4.1) and the Runge–Kutta method (1.8), we obtain the following result.

Theorem 4.14 *For y satisfying (4.1), there exists a constant C_{m+1} such that*

$$\begin{aligned} & \left| \exp(sW)(y_0) - \left(y_0 + \sum_{t \in T_{\leq m}} \frac{s^{r(t)}}{r(t)!} \alpha(t) D(W, t)(y_0) \right) \right| \\ & \leq C_{m+1} s^{m+1} \|W\|_{C_{m+1}}^{m+1}. \end{aligned}$$

On the other hand, for the Runge–Kutta method (1.8) there exists a constant C'_{m+1} such that

$$\begin{aligned} & \left| Y(y_0; W, s) - \left(y_0 + \sum_{\substack{\ell=1, \dots, m-1 \\ t=[t_1 \dots t_\ell] \in T_{\leq m}}} \frac{s^{r(t)}}{\sigma(t)} \sum_{i=1}^K b_i \prod_{k=1}^\ell \zeta_i(t_k; A_{ex}) D(W, t)(y_0) \right) \right| \\ & \leq C'_{m+1} s^{m+1} \|W\|_{C_{m+1}}^{m+1}. \end{aligned}$$

We say that (A, b) satisfies m th-order conditions if

$$\frac{\alpha(t)}{r(t)!} = \frac{\sum_{i=1}^K b_i \prod_{k=1}^\ell \zeta_i(t_k; A)}{\sigma(t)} \tag{4.2}$$

for all $t = [t_1 \dots t_\ell] \in T_{\leq m}$.

From Theorem 4.14, the following result can be directly derived.

Theorem 4.15 *Suppose that an explicit Runge–Kutta method (A_{ex}, b) satisfies the m th-order conditions (4.2). Let $g(W)(y_0) = Y(y_0; W, 1)$, where $Y(y_0; W, 1)$ is the Runge–Kutta method defined in (1.8). Then*

$$g \in \mathcal{IS}(m).$$

5 The new simulation scheme and Corollary 1.4

Corollary 1.4 indicates the new implementation method of the new higher-order scheme proposed by Kusuoka in [17, 18].

This implementation method seems to be distinct mainly because it has two advantages. One is that the approximation operator can be obtained by numerical calculations if the Runge–Kutta method is applicable to the calculation of each $\exp(Z_j)$, whereas the tediousness in symbolical calculations of the operator might be an obstacle for practical application, which can be observed in [20, 26, 30]. The other advantage is that the partial sampling problem discussed in [20, 26] can be resolved by using quasi-Monte Carlo methods. More precisely, the following two points make an effective use of low-discrepancy sequences, which are essential to quasi-Monte Carlo methods [24]:

- In this implementation, S_j^i can be taken to be a continuous random variable.
- The scheme itself is characterized by the need for a smaller number of discretization time steps, which leads to a reduction in the number of dimensions of the numerical integration.

6 Application

In this section we present a numerical example in order to illustrate the implementation method proposed in Corollary 1.4 and compare it with some existing schemes.

6.1 Simulation

Let $X(t, x)$ be the diffusion process defined by (1.1). The most popular scheme of first order is the Euler–Maruyama scheme, which is shown in [15, 34], for an arbitrary C^4 function f , to satisfy

$$\|E[f(X^{(EM),n}_1)] - E[f(X(1, x))]\| \leq C_f \frac{1}{n}$$

where $X^{(EM),n}_1$ denotes the Euler–Maruyama scheme approximating $X(t, x)$. We note that this inequality holds for measurable f if $\{V_i\}_{i=1,\dots,d}$ satisfies some more conditions [1, 16].

The construction of a higher-order scheme is based on the higher-order Itô–Taylor formula [8, 15]. When the vector fields $\{V_i\}_{i=0}^d$ commute, higher-order schemes can be simplified to a direct product of one-dimensional problems as seen in [15]. In contrast, for non-commutative $\{V_i\}_{i=0}^d$, the acquisition of all iterated integrals of Brownian motion is required, which is very demanding. This is done in [17, 20, 22, 32, 33] and generalized as the cubature method on Wiener space [23].

Once a p th-order scheme $\{X^{(\text{ord } p),n}_{k/n}\}_{k=0,\dots,n}$ is obtained and expanded with some constant K_f as

$$E[f(X^{(\text{ord } p),n}_1)] - E[f(X(1, x))] = K_f \frac{1}{n^p} + O\left(\frac{1}{n^{p+1}}\right),$$

the $(p + 1)$ th-order scheme can be derived as

$$\frac{2^p}{2^p - 1} E[f(X^{(\text{ord } p), 2n_1})] - \frac{1}{2^p - 1} E[f(X^{(\text{ord } p), n_1})].$$

This boosting method is called Romberg extrapolation and is shown to be applicable to the Euler–Maruyama scheme under certain conditions [34].

The simulation approach must be followed by the numerical calculation of the expectation $E[f(X^{(\text{ord } p), n_1})]$. However, when $n \times d$ is large, it is practically impossible to proceed with the integration by using the trapezoidal formula and so we fall back on the Monte Carlo or the quasi-Monte Carlo method [24]. Here we make only a few remarks on each method. For a more detailed analysis, see [27].

Remark 6.1 As long as we use the Monte Carlo method for numerical approximation of $E[f(X(1, x))]$, the number of sample points needed to attain a given accuracy is independent of the number of the dimensions of integration, namely both the number n of partitions and the order p of the approximation scheme.

Remark 6.2 In contrast to the Monte Carlo case, the number of sample points needed for the quasi-Monte Carlo method for numerical approximation of $E[f(X(1, x))]$ heavily depends on the number of the dimensions of integration. The fewer the dimensions, the fewer the samples that are needed.

6.2 The algorithm and competitors

6.2.1 The algorithm of the new method

We take the algorithm which is proposed in Theorem 1.6 and Corollary 1.4 with $u = 3/4$. From Corollary 1.4, we can implement the second-order algorithm with a numerical approximation of $\exp(Z_i)$ of at least fifth-order Runge–Kutta method because the order m for an integration scheme attained by Z_1 and Z_2 is five and so the order of the new implementation method becomes two. As a result of the same argument it can be shown that at least a seventh-order explicit Runge–Kutta method has to be applied to the approximation of $\exp(Z_i)$ when we boost the new method to the third order by Romberg extrapolation. Details of these Runge–Kutta algorithms used here are given in the Appendix.

6.2.2 Competing schemes

There are numerous studies on the acceleration of the Monte Carlo methods ([12]). We choose for the following reasons only the crude Euler–Maruyama scheme and the algorithm introduced in [27], which we refer to in the remainder of this paper as N–V method, both with and without Romberg extrapolation, as competitors:

- (i) Only these two schemes can be recognized as being comparable to the new method, since they are model-independent.
- (ii) Almost all variance reduction techniques and dimension reduction techniques applicable to the Euler–Maruyama scheme are also applicable to the new method.

6.3 Numerical results

We provide an example on financial option pricing in the following part of this paper.

6.3.1 Asian option under the Heston model

We consider an Asian call option written on an asset whose price process follows the Heston stochastic volatility model. Comparison with the N–V method will also be given as well from the result shown in [27].

The non-commutativity of this example should be noted here.

Let Y_1 be the price process of an asset following the Heston model

$$\begin{aligned}
 Y_1(t, x) &= x_1 + \int_0^t \mu Y_1(s, x) ds + \int_0^t Y_1(s, x) \sqrt{Y_2(s, x)} dB^1(s), \\
 Y_2(t, x) &= x_2 + \int_0^t \alpha(\theta - Y_2(s, x)) ds \\
 &\quad + \int_0^t \beta \sqrt{Y_2(s, x)} (\rho dB^1(s) + \sqrt{1 - \rho^2} dB^2(s)),
 \end{aligned}
 \tag{6.1}$$

where $x = (x_1, x_2) \in (\mathbb{R}_{>0})^2$, $(B^1(t), B^2(t))$ is a two-dimensional standard Brownian motion, $-1 \leq \rho \leq 1$, and α, θ, μ are some positive coefficients such that

$$2\alpha\theta - \beta^2 > 0 \tag{6.2}$$

to ensure the existence and uniqueness of a solution to the stochastic differential equation [10]. Then the payoff of an Asian call option on this asset with maturity T and strike K is $\max(Y_3(T, x)/T - K, 0)$ where

$$Y_3(t, x) = \int_0^t Y_1(s, x) ds. \tag{6.3}$$

Hence, the price of this option becomes $D \times E[\max(Y_3(T, x)/T - K, 0)]$ where D is an appropriate discount factor that we do not focus on here. We set $T = 1, K = 1.05, \mu = 0.05, \alpha = 2.0, \beta = 0.1, \theta = 0.09, \rho = 0$, and $(x_1, x_2) = (1.0, 0.09)$ and take

$$E[\max(Y_3(T, x)/T - K, 0)] = 6.0473534496 \times 10^{-2}$$

that is obtained by the new method with Romberg extrapolation and the quasi-Monte Carlo with $n = 96 + 48$, and $M = 8 \times 10^8$ where M denotes the number of sample points.

Let $Y(t, x) = {}^t(Y_1(t, x), Y_2(t, x), Y_3(t, x))$. Transformation of the stochastic differential equations (6.1) and (6.3) gives the Stratonovich-form stochastic differential equations

$$Y(t, x) = \sum_{i=0}^2 \int_0^t V_i(Y(s, x)) \circ dB^i(s),$$

Table 1 Number of dimensions involved in each method

Method	Number of dimensions
Euler–Maruyama	dn
N–V	$n + dn$ (n -Bernoulli and $(d \times n)$ -Gaussian)
New method	$2dn$

where

$$V_0(t(y_1, y_2, y_3)) = {}^t \left(y_1 \left(\mu - \frac{y_2}{2} - \frac{\rho\beta}{4} \right), \alpha(\theta - y_2) - \frac{\beta^2}{4}, y_1 \right),$$

$$V_1(t(y_1, y_2, y_3)) = {}^t \left(y_1 \sqrt{y_2}, \rho\beta \sqrt{y_2}, 0 \right),$$

$$V_2(t(y_1, y_2, y_3)) = {}^t \left(0, \beta \sqrt{(1 - \rho^2)y_2}, 0 \right).$$

We note that the vector fields in the Heston model violate the differentiable condition at the origin. We can, however, avoid the problem by approximating the vector fields by smooth vector fields because the process never touches the origin under the condition (6.2).

6.3.2 Dimensions of integrations

As mentioned in Remarks 6.1 and 6.2, the dimensions of integrations in these methods affect the quasi-Monte Carlo method. The relation among the number d of factors, the number n of partitions, and the dimensions of integration of each method can be summarized as in Table 1.

6.3.3 Discretization error

The relation between discretization error and the number of partitions of each algorithm is plotted in Fig. 1. We can observe from this figure that for 10^{-4} accuracy the new method with Romberg extrapolation takes the minimum number of partitions as $n = 1 + 2$ whereas $n = 16$ for the Euler–Maruyama scheme with the extrapolation. Even without the extrapolation, the new method attains that accuracy with $n = 10$ while the Euler–Maruyama scheme takes $n = 2000$. Moreover, it may be said that the N–V method shows slightly worse performance than the new method.

6.3.4 Integration error

Looking at Fig. 2, we can compare convergence errors of respective methods for each number M of sample points. For the Monte Carlo case, 2σ of 10 batches is taken as convergence error while for the quasi-Monte Carlo method, the absolute difference from the value to be convergent is considered. For 10^{-4} accuracy with 95% confidence level (2σ), $M = 10^8$ is taken for the Monte Carlo method. On the other hand, if we apply instead the quasi-Monte Carlo method, the new method and the N–V method require $M = 2 \times 10^5$ sample points, while $M = 5 \times 10^6$ has to be taken for the Euler–Maruyama scheme.

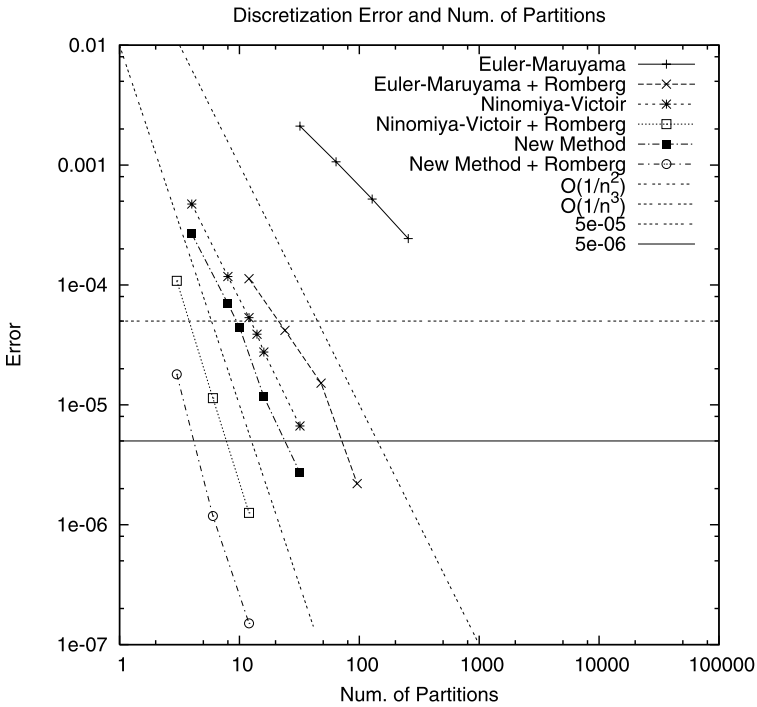


Fig. 1 Error coming from the discretization

6.3.5 Overall performance comparison

The number of partitions, the number of samples, and the amount of computation time required for 10^{-4} accuracy for each method are summarized in Table 2. The CPU used in this experiment is Athlon 64 3800+ by AMD.

Since the amount of time required to carry out the calculation for each sample point is proportional to the number of partitions, the total time spent on calculations is proportional both to the number of partitions and to the number of samples. We can see from Table 2 that the speed of the new method is approximately 100 times faster than that of the Euler–Maruyama scheme when Romberg extrapolation and the quasi-Monte Carlo are applied to each. Even when the extrapolation is not applied, the new method enables calculations some 37 times faster than the Euler–Maruyama scheme with Romberg extrapolation and the quasi-Monte Carlo method. This fact shows that the reduction in the number of partitions sufficiently compensates for the slowness of one step of the new method at least in the present study.

Lastly, Remarks 6.1 and 6.2 should be emphasized to reiterate that the advantage of the new method is deeply related to the properties of the quasi-Monte Carlo method.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

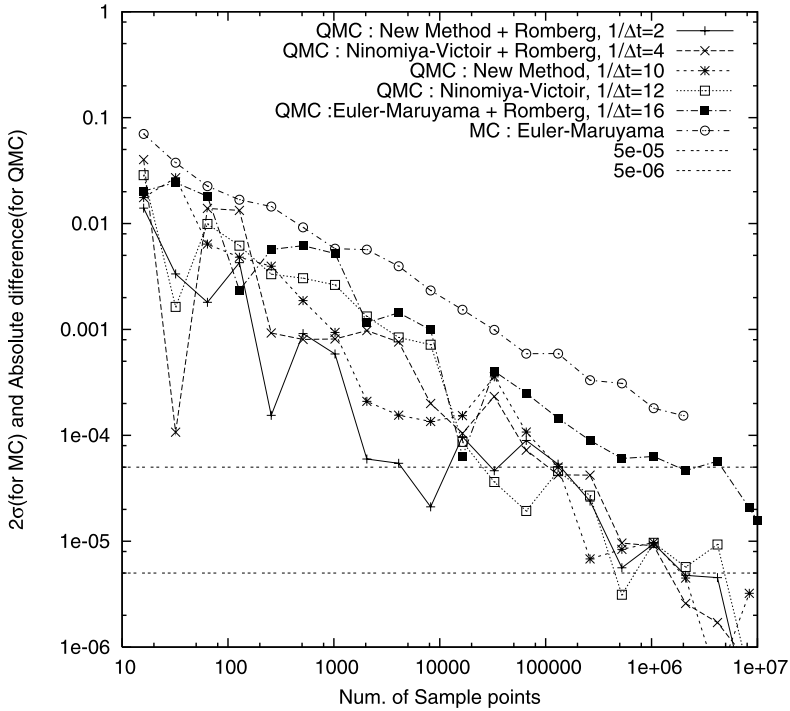


Fig. 2 Convergence error from quasi-Monte Carlo and Monte Carlo

Table 2 # partitions, # samples, dimension, and CPU time required for an accuracy of 10^{-4}

Method	# Part.	Dim.	# Samples	CPU time (s)
E-M + MC	2000	4000	10^8	1.72×10^5
E-M + Romb. + QMC	16 + 8	48	5×10^6	1.27×10^2
N-V + QMC	16	32 + 16	2×10^5	4.38
N-V + Romb. + QMC	4 + 2	12 + 6	2×10^5	1.76
New method + QMC	10	40	2×10^5	3.4
New method + Romb. + QMC	2 + 1	12	2×10^5	1.2

Appendix: The fifth-order and the seventh-order Runge–Kutta algorithms

We present here the concrete algorithms of the explicit fifth- and seventh-order Runge–Kutta methods applied in Sect. 6.2. The fifth-order method is taken from [6] as

$$a_{21} = \frac{2}{5}, \quad a_{31} = \frac{11}{64}, \quad a_{32} = \frac{5}{64}, \quad a_{43} = \frac{1}{2}, \quad a_{51} = \frac{3}{64},$$

$$\begin{aligned}
 a_{52} &= -\frac{15}{64}, & a_{53} &= \frac{3}{8}, & a_{54} &= \frac{9}{16}, & a_{62} &= \frac{5}{7}, & a_{63} &= \frac{6}{7}, \\
 a_{64} &= -\frac{12}{7}, & a_{65} &= \frac{8}{7}, \\
 a_{ij} &= 0 \quad \text{otherwise,} \\
 b &= \left(\frac{7}{90} \quad 0 \quad \frac{32}{90} \quad \frac{12}{90} \quad \frac{32}{90} \quad \frac{7}{90} \right).
 \end{aligned}$$

The seventh-order method is taken from [7] as

$$\begin{aligned}
 a_{21} &= \frac{1}{6}, & a_{32} &= \frac{1}{3}, & a_{41} &= \frac{1}{8}, & a_{43} &= \frac{3}{8}, & a_{51} &= \frac{148}{1331}, \\
 a_{53} &= \frac{150}{1331}, & a_{54} &= -\frac{56}{1331}, & a_{61} &= -\frac{404}{243}, & a_{63} &= -\frac{170}{27}, \\
 a_{64} &= \frac{4024}{1701}, & a_{65} &= \frac{10648}{1701}, & a_{71} &= \frac{2466}{2401}, & a_{73} &= \frac{1242}{343}, \\
 a_{74} &= -\frac{19176}{16807}, & a_{75} &= -\frac{51909}{16807}, & a_{76} &= \frac{1053}{2401}, \\
 a_{81} &= \frac{5}{154}, & a_{84} &= \frac{96}{539}, & a_{85} &= -\frac{1815}{20384}, & a_{86} &= -\frac{405}{2464}, \\
 a_{87} &= \frac{49}{1144}, & a_{91} &= -\frac{113}{32}, & a_{93} &= -\frac{195}{22}, & a_{94} &= \frac{32}{7}, \\
 a_{95} &= \frac{29403}{3584}, & a_{96} &= -\frac{729}{512}, & a_{97} &= \frac{1029}{1408}, & a_{98} &= \frac{21}{16}, \\
 a_{ij} &= 0 \quad \text{otherwise,} \\
 b &= \left(0 \quad 0 \quad 0 \quad \frac{32}{105} \quad \frac{1771561}{6289920} \quad \frac{243}{1560} \quad \frac{16807}{74880} \quad \frac{77}{1440} \quad \frac{11}{70} \right).
 \end{aligned}$$

References

1. Bally, V., Talay, D.: The law of the Euler scheme for stochastic differential equations, I: convergence rate of the distribution function. *Probab. Theory Relat. Fields* **104**, 43–60 (1996)
2. Bayer, C., Teichmann, J.: Cubature on Wiener space in infinite dimension. *Proc. R. Soc. A* **464**(2097), 2493–2516 (2008)
3. Bollobás, B.: *Graph Theory: An Introductory Course*. Springer, Berlin (1979)
4. Bourbaki, N.: *Éléments de Mathématique, Groupes et Algèbres de Lie*. Hermann, Paris (1972), Chaps. 2 and 3
5. Burrage, K., Burrage, P.M.: General order conditions for stochastic Runge–Kutta methods for both commuting and non-commuting stochastic ordinary differential equation systems. *Appl. Numer. Math.* **28**(2–4), 161–177 (1998)
6. Butcher, J.C.: *The Numerical Analysis of Ordinary Differential Equations*. Wiley, Chichester (1987)
7. Butcher, J.C.: *Numerical Methods for Ordinary Differential Equations*. Wiley, Chichester (2003)
8. Castell, F.: Asymptotic expansion of stochastic flows. *Probab. Theory Relat. Fields* **96**, 225–239 (1993)
9. Cohn, P.M.: *Skew Fields: Theory of General Division Rings*. Cambridge University Press, Cambridge (1995)

10. Feller, W.: Two singular diffusion problems. *Ann. Math.* **54**, 173–182 (1951)
11. Filipović, D., Tappe, S., Teichmann, J.: Jump-diffusions in Hilbert spaces: existence, stability and numerics. Preprint (2008). <http://arxiv.org/abs/0810.5023>
12. Glasserman, P.: *Monte Carlo Methods in Financial Engineering*. Springer, New York (2004)
13. Hairer, E., Lubich, C., Wanner, G.: *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, 2nd ed. Springer, Berlin (2006)
14. Ikeda, N., Watanabe, S.: *Stochastic Differential Equations and Diffusion Processes*. North Holland/Kodansha, Amsterdam (1981)
15. Kloeden, P.E., Platen, E.: *Numerical Solution of Stochastic Differential Equations*. Springer, Berlin (1999)
16. Kohatsu-Higa, A.: Weak approximations. A Malliavin calculus approach. *Math. Comput.* **70**, 135–172 (2000)
17. Kusuoka, S.: Approximation of expectation of diffusion process and mathematical finance. In: T. Sunada (ed.) *Proceedings of Final Taniguchi Symposium, Nara, 1998*. *Advanced Studies in Pure Mathematics*, vol. 31, pp. 147–165 (2001)
18. Kusuoka, S.: Approximation of expectation of diffusion processes based on Lie algebra and Malliavin calculus. *Adv. Math. Econ.* **6**, 69–83 (2004)
19. Kusuoka, S.: Kusuoka scheme and Gaussian type approximation. Presentation at Mathematical Finance Seminar in Graduate School of Mathematical Sciences the University of Tokyo (1/June/2005) (2005)
20. Kusuoka, S., Ninomiya, S.: A new simulation method of diffusion processes applied to finance. In: Akahori, J., Ogawa, S., Watanabe, S. (eds.) *Stochastic Processes and Application to Mathematical Finance, Proceedings of the Ritsumeikan International Symposium*, pp. 233–253. World Scientific, Singapore (2004)
21. Lapeyre, B., Pardoux, E., Sentis, R.: *Méthodes de Monte Carlo pour les Équations de Transport et de Diffusion*. *Mathematics and Applications*, vol. 29. Springer, Berlin (1998)
22. Liu, X.Q., Li, C.W.: Weak approximation and extrapolations of stochastic differential equations with jumps. *SIAM J. Numer. Anal.* **37**, 1747–1767 (2000)
23. Lyons, T., Victoir, N.: Cubature on Wiener space. *Proc. R. Soc. Lond. Ser. A Math. Phys. Sci.* **460**, 169–198 (2004)
24. Niederreiter, H.: *Random Number Generation and Quasi-Monte Carlo Methods*. SIAM, Philadelphia (1992)
25. Ninomiya, S.: A new simulation scheme of diffusion processes: application of the Kusuoka approximation to finance problems. *Math. Comput. Simul.* **62**(3–6), 479–486 (2003)
26. Ninomiya, S.: A partial sampling method applied to the Kusuoka approximation. *Monte Carlo Methods Appl.* **9**, 27–38 (2003)
27. Ninomiya, S., Victoir, N.: Weak approximation of stochastic differential equations and application to derivative pricing. *Appl. Math. Finance* **15**, 107–121 (2008)
28. Rössler, A.: *Runge–Kutta Methods for the Numerical Solution of Stochastic Differential Equations*. Shaker Verlag GmbH, Aachen (2003)
29. Rümelin, W.: Numerical treatment of stochastic differential equations. *SIAM J. Numer. Anal.* **19**, 604–613 (1982)
30. Shimizu, M.: Application of the Kusuoka approximation with tree based branching algorithm to pricing interest-rate derivatives with the HJM model. Master thesis, Imperial College of Science, Technology, and Medicine (2002)
31. Strichartz, R.E.: The Campbell–Baker–Hausdorff–Dynkin formula and solutions of differential equations. *J. Funct. Anal.* **72**, 320–345 (1987)
32. Talay, D.: Second-order discretization schemes of stochastic differential systems for the computation of the invariant law. *Stoch. Stoch. Rep.* **29**, 13–36 (1990)
33. Talay, D.: Simulation of stochastic differential systems. In: Kree, P., Wedig, W. (eds.) *Probabilistic Methods in Applied Physics*. LNP, vol. 451, pp. 54–96. Springer, Berlin (1995)
34. Talay, D., Tubaro, L.: Expansion of the global error for numerical schemes solving stochastic differential equations. *Stoch. Anal. Appl.* **8**, 483–509 (1990)
35. Teichmann, J.: Construction of cubature formulas for the KLV method. Preprint (2009)