



A class of new Magnus-type methods for semi-linear non-commutative Itô stochastic differential equations

Guoguo Yang^{1,2} · Kevin Burrage^{2,3} · Yoshio Komori⁴ · Pamela Burrage² · Xiaohua Ding¹

Received: 15 September 2020 / Accepted: 18 February 2021 / Published online: 30 March 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

In this paper, a class of new Magnus-type methods is proposed for non-commutative Itô stochastic differential equations (SDEs) with semi-linear drift term and semi-linear diffusion terms, based on Magnus expansion for non-commutative linear SDEs. We construct a Magnus-type Euler method, a Magnus-type Milstein method and a Magnus-type Derivative-free method, and give the mean-square convergence analysis of these methods. Numerical tests are carried out to present the efficiency of the proposed methods compared with the corresponding underlying methods and the specific performance of the simulation Itô integral algorithms is investigated.

Keywords Stochastic differential equations · Magnus-type methods · Magnus-type Euler method · Magnus-type Milstein method · Magnus-type Derivative-free method

✉ Guoguo Yang
17b912010@stu.hit.edu.cn

Kevin Burrage
kevin.burrage@qut.edu.au

Yoshio Komori
komori@ces.kyutech.ac.jp

Pamela Burrage
pamela.burrage@qut.edu.au

Xiaohua Ding
mathdxh@126.com

¹ Department of Mathematics, Harbin Institute of Technology, Harbin, People's Republic of China

² School of Mathematical Sciences, Queensland University of Technology, Brisbane, Australia

³ Department of Computer Science, University of Oxford, Oxford, UK

⁴ Department of Systems Design and Informatics, Kyushu Institute of Technology, Kitakyushu, Japan

1 Introduction

It is now generally accepted that stochastic differential equations (SDEs) can describe some problems more accurately than deterministic differential equations. For example, input data may be uncertain or systems may be subject to internal or external random fluctuations. It is often difficult to obtain explicit expressions of exact solutions for SDEs and so numerical methods become an important tool to provide good approximations.

In this paper, our main focus is the numerical approximations of the semilinear Itô SDEs

$$dy = (A_0y + g_0(y))dt + \sum_{i=1}^m (A_iy + g_i(y)) dW_i(t), \quad y(0) = y_0 \in \mathbb{R}^d, \quad (1)$$

where $W_j(t)$, $j = 1, \dots, m$, are independent Wiener processes on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a filtration $\{\mathcal{F}_t\}_{t>0}$ satisfying the usual conditions. Here $g_j: \mathbb{R}^d \rightarrow \mathbb{R}^d$, $j = 1, \dots, m$, are nonlinear functions of y and $A_j \in \mathbb{R}^{d \times d}$, $j = 0, \dots, m$, are constant matrices. The random initial value satisfies $\mathbb{E} \|y_0\|^2 < \infty$ and \mathbb{E} is the expectation.

Numerical integrators for (1) in the sense of Itô and Stratonovich are thoroughly studied in [1] based on the stochastic Runge–Kutta Lawson methods under the following commutative conditions

$$[A_l, A_k] = A_l A_k - A_k A_l = 0 \quad \text{for all } l, k = 0, 1, \dots, m, \quad (2)$$

where $[A, B] = AB - BA$ is called the Lie-product or matrix commutator of A and B . Under this commutative condition, Euler and Milstein versions of exponential methods are investigated in [2], and the strong convergence analysis of these two methods is given. Their numerical results show that these exponential methods are more effective than the corresponding underlying methods. Under similar commutative conditions, Yang, Burrage and Ding [3] investigate (1) in the sense of Stratonovich and construct structure-preserving stochastic exponential Runge–Kutta methods to the case of time independent matrix $A_0 = A(t)$, $A_k = 0$, $k = 1, \dots, m$. There is also some interesting work on the exact solution and the corresponding numerical solution under the commutative condition where A_0 is a constant matrix and $A_k = 0$, $k = 1, \dots, m$. For example, an exponential Euler method is constructed for the stiff problem and it is applied to solve an ion channel model in [4]. A class of weak second-order exponential Runge–Kutta methods is investigated for non-commutative SDEs in [5].

It is worth noting that when the commutative condition is not established, even the solution of the linear system cannot be expressed explicitly. We know that Magnus gave the deterministic Magnus expansion [6] in 1954, which expresses the solution as an exponential infinite matrix series. This topic has been further studied in [7–9]. In recent years, the Magnus-type expansion for SDEs has attracted an increasing number of researchers. The Magnus expansion for linear autonomous Stratonovich SDEs was derived by Burrage and Burrage [10]. They compare the truncated Magnus expansion with stochastic Runge–Kutta methods with the same convergence

order and concluded that the Magnus expansion enjoys a considerably smaller error coefficient. The authors of [11] consider the Magnus expansion for the linear and nonlinear Stratonovich SDEs and the convergence analysis is given based on binary rooted trees. In [12] a new explicit Magnus expansion is applied to solve a class of Stratonovich SDEs and the authors show that the Magnus expansion can preserve the positivity of the solution. A new version of Magnus expansion for linear Itô SDEs is derived in [13] and the authors have analyzed the convergence of the Magnus expansion and applied it to solve semi-discrete stochastic partial differential equations (SPDEs). In [14], for SPDEs driven by multiplicative noise, the authors first use the finite element method to discretize space, and then construct a Magnus-type method in the time direction.

We have to say that when the commutative condition is not established, the corresponding theoretical results on exponential numerical methods for the corresponding semi-linear problem are sparse. This is the motivation for us to consider a new class of Magnus-type integrators for the semi-linear problem (1).

The structure of this paper is organized as follows. In Section 2, we will give a brief review of the use of the Magnus formula in both linear non-autonomous ordinary differential equations (ODEs) and linear SDEs. In Section 3, we will derive a new class of Magnus-type integrators for the semi-linear problem (1) by the application of the Magnus expansion for linear non-commutative SDEs. In Section 4, we will give the results of the mean-square convergence analysis of the Magnus-type methods proposed in Section 3. In Section 5, we will compare three algorithms for simulating iterated Itô integrals and some details of numerical implementation will be presented on low-dimensional SDEs and a high-dimensional SDE obtained from a discretized SPDE, which illustrate the efficacy of Magnus-type methods.

2 Magnus formula

In this section we will briefly review the Magnus formula in both a linear deterministic setting and a linear stochastic setting.

2.1 Magnus formula for linear non-autonomous ODEs

Consider the non-autonomous linear initial value problem

$$Y'(t) = A(t)Y(t), t \geq 0, \quad Y(0) = Y_0 \in \mathbb{R}^{d \times d}. \tag{3}$$

This problem is investigated in [6] where the $d \times d$ matrix $A(t)$ is non-commutative. The solution is given in the form

$$Y(t) = \exp(\Omega(0, t)) Y_0, \tag{4}$$

where $\Omega(0, t)$ is the combination of integrals and nested Lie brackets of A , i.e.,

$$\begin{aligned} \Omega(0, t) = & \int_0^t A(s_1) ds_1 + \frac{1}{2} \int_0^t [A(s_1), \int_0^{s_1} A(s_2) ds_2] ds_1 \\ & + \frac{1}{4} \int_0^t [A(s_1), \int_0^{s_1} [A(s_2), \int_0^{s_2} A(s_3) ds_3] ds_2] ds_1 + \dots \end{aligned}$$

For this problem, the Magnus expansion has received some attention in the past few decades, see [15].

2.2 Magnus formula for linear SDEs

We consider autonomous linear SDE given by

$$dY(t) = \sum_{j=0}^m A_j Y(t) dW_j(t), \quad Y(0) = I \in \mathbb{R}^{d \times d}, \tag{5}$$

where $W_0(t) = t$. The $d \times d$ matrices A_j are constant and I is the identity $d \times d$ -matrix. When the commutative condition (2) is satisfied, the exact solution of (5) can be expressed explicitly by

$$Y(t) = \exp \left(\left(A_0 - \gamma^* \sum_{i=1}^m A_i^2 \right) t + \sum_{i=1}^m A_i W_i(t) \right),$$

where $\gamma^* = \frac{1}{2}$. When the commutative condition is not established, the solution of (5) cannot be expressed explicitly. The solution to (5) is given in the form $Y(t) = \exp(\Omega(0, t))I$ by Kamm and his coauthors [13] in the Itô setting. Here we apply the relationship between Itô and Stratonovich integrals [17] to obtain the expansion of the Itô case through $\Omega(0, t)$ given by Burrage and Burrage [10] in the Stratonovich setting,

$$\begin{aligned} \Omega(0, t) = & \sum_{j=0}^m \hat{A}_j \int_0^t dW_j \\ & + \frac{1}{2} \sum_{i=0}^m \sum_{j=i+1}^m [\hat{A}_i, \hat{A}_j] \left(\int_0^t \int_0^s dW_j dW_i - \int_0^t \int_0^s dW_i dW_j \right) \\ & + \sum_{i=0}^m \sum_{k=0}^m \sum_{j=k+1}^m [\hat{A}_i, [\hat{A}_j, \hat{A}_k]] \left\{ \frac{1}{3} \left(\int_0^t \int_0^s \int_0^{s_1} dW_k dW_j dW_i \right. \right. \tag{6} \\ & + \gamma^* I_{(j=i \neq 0)} \int_0^t \int_0^s dW_k dW_0 \left. \right) - \left(\int_0^t \int_0^s \int_0^{s_1} dW_j dW_k dW_i \right. \\ & + \gamma^* I_{(k=i \neq 0)} \int_0^t \int_0^s dW_j dW_0 \left. \right) \\ & \left. + \frac{1}{12} \int_0^t dW_i \left(\int_0^t \int_0^s dW_j dW_k - \int_0^t \int_0^s dW_k dW_j \right) \right\} + \dots, \end{aligned}$$

where $\hat{A}_0 = A_0 - \gamma^* \sum_{j=1}^m A_j^2$ and $\hat{A}_j = A_j, j \geq 1$. In the rest of this paper, we denote the iterated Itô integral as

$$I_{ij}(t_n, t_n + h) = \int_{t_n}^{t_n+h} \int_{t_n}^s dW_i(s_1) dW_j(s), \quad i, j \geq 1. \tag{7}$$

The relationship between Itô and Stratonovich integrals [17] is

$$\begin{aligned} J_\alpha &= I_\alpha, & l(\alpha) &= 0 \text{ or } 1, \\ J_\alpha &= I_\alpha + \frac{1}{2} I_{\{j_1=j_2 \neq 0\}} I_0, & l(\alpha) &= 2, \\ J_\alpha &= I_\alpha + \frac{1}{2} \left(I_{\{j_1=j_2 \neq 0\}} I_{(0, j_3)} + I_{\{j_2=j_3 \neq 0\}} I_{(j_1, 0)} \right), & l(\alpha) &= 3, \end{aligned}$$

where $l(\alpha)$ is the length of the index α and I_A is the indicator function, i.e., $I_A = 1$ if A is true, otherwise $I_A = 0$. The expansion (6) can also be obtained through the expansion rules in [13] and the convergence of the Magnus expansion for (5) is given by the following lemma.

Lemma 1 [13] *Let A_0, A_1, \dots, A_m be constant matrices. For $T > 0$ let $Y = (Y(t))_{t \in [0, T]}$ be the solution to (5) in the Itô case. There exists a strictly positive stopping time $\tau \leq T$ such that:*

(i) $Y(t)$ has a real logarithm $\Omega(0, t) \in \mathbb{R}^{d \times d}$ up to time τ , i.e.,

$$Y(t) = e^{\Omega(0, t)}, \quad 0 \leq t < \tau;$$

(ii) the following representation holds \mathbb{P} -almost surely:

$$\Omega(0, t) = \sum_{n=0}^{\infty} \Omega^{[n]}(0, t), \quad 0 \leq t < \tau,$$

where $\Omega^{[n]}(0, t)$ is the n th term in the stochastic Magnus expansion (6);

(iii) there exists a positive constant C , only dependent on $\|A_0\|, \dots, \|A_m\|, T$ and d , such that

$$\mathbb{P}(\tau \leq t) \leq Ct, \quad t \in [0, T].$$

3 A class of new Magnus-type methods for semi-linear SDEs

We shall now derive Magnus-type methods of mean-square order 1/2 and 1.0 for semi-linear SDEs (1). Throughout this paper, consider a partition $t_0 = 0 < t_1 < \dots < t_N = T$ of the interval $[0, T]$ with constant step size $h = t_j - t_{j-1}, j = 1, \dots, N$, and let y_n be the approximation of exact solution. To make sure of the existence of the unique solution of (1), we first give the following important result.

Theorem 1 [16] *There exists a constant $L > 0$ such that the global Lipschitz condition holds: for $y, z \in \mathbb{R}^d$,*

$$|g_0(y) - g_0(z)| + \sum_{i=1}^m |g_i(y) - g_i(z)| \leq L|y - z|.$$

Then there exists a unique solution $y(t)$ to (1).

Here we only require the Lipschitz condition as the linear growth condition

$$|g_0(y)|^2 + \sum_{i=1}^m |g_i(y)|^2 \leq L(1 + |y|^2) \text{ for } y \in \mathbb{R}^d,$$

is automatically satisfied from the Lipschitz condition in the autonomous case, see [18].

For the semi-linear Itô SDEs (1), we assume that the exact solution has the form

$$y(t) = Y(t)\tilde{y}(t), \tag{8}$$

where $Y(t)$ is the solution of the linear equation (5), and $\tilde{y}(t)$ is to be determined. Using the Itô chain rule to $y(t)$, we have

$$dy = (dY)\tilde{y} + Yd\tilde{y} + dYd\tilde{y} = \sum_{j=0}^m A_j Y \tilde{y} dW_j + Yd\tilde{y} + dYd\tilde{y}. \tag{9}$$

Comparing this with (1) shows that (8) is a solution of (1) if and only if

$$Yd\tilde{y} = g_0(y)dt - dYd\tilde{y} + \sum_{j=1}^m g_j(y)dW_j.$$

So, $d\tilde{y}$ has the form

$$d\tilde{y} = \exp(-\Omega(0, t))(g_0(y)dt - dYd\tilde{y}) + \sum_{j=1}^m \exp(-\Omega(0, t))g_j(y)dW_j.$$

Since

$$dYd\tilde{y} = \sum_{j=1}^m A_j g_j(y)dt,$$

then we get

$$\begin{aligned} y(t) &= \exp(\Omega(0, t))(y(0) + \int_0^t \exp(-\Omega(0, s))\tilde{g}_0(y(s))ds \\ &\quad + \sum_{j=1}^m \int_0^t \exp(-\Omega(0, s))g_j(y(s))dW_j(s)), \end{aligned} \tag{10}$$

where $\tilde{g}_0 = g_0 - \sum_{j=1}^m A_j g_j$.

It needs to be said that this transformation is also applicable to the case where the $A_j(t)$ depend on time t , that is, the non-autonomous case, but this paper focuses on the case of (1). Numerical methods will be derived by using this form. Different approximations to the integrals in the above equation will yield different numerical schemes, and we will consider Magnus-type Euler (ME) and Magnus-type Milstein (MM) methods.

3.1 Magnus-type Euler method

If the integrals in (10) are approximated as follows

$$\begin{aligned} &\exp(\Omega(t_n, t_{n+1})) \int_{t_n}^{t_{n+1}} \exp(-\Omega(t_n, s))g_j(y)dW_j(s) \\ &\approx \exp(\Omega^{[1]}(t_n, t_{n+1}))g_j(y_n)\Delta W_{jn}, \quad j = 1, \dots, m, \end{aligned}$$

where $\Omega^{[1]}(t_n, t_{n+1}) = \sum_{j=0}^m \hat{A}_j \int_{t_n}^{t_{n+1}} dW_j$, $\Delta W_{jn} = W_j(t_{n+1}) - W_j(t_n)$, $j \geq 1$, $\Delta W_{0n} = t_{n+1} - t_n$, the following ME method is obtained,

$$y_{n+1} = \exp(\Omega^{[1]}(t_n, t_{n+1})) \left(y_n + \tilde{g}_0(y_n)h + \sum_{j=1}^m g_j(y_n)\Delta W_{jn} \right). \tag{11}$$

In particular, when $g_j = 0$, $j \geq 1$, and if

$$\exp(\Omega^{[2]}(t_n, t_{n+1})) = \sum_{j=0}^m \hat{A}_j \int_{t_n}^{t_{n+1}} dW_j + \frac{1}{2} \sum_{i=1}^m \sum_{j=i+1}^m [\hat{A}_i, \hat{A}_j] (I_{ji} - I_{ij})$$

is selected instead of $\exp(\Omega^{[1]}(t_n, t_{n+1}))$, the resulting numerical scheme is mean-square 1 order convergent, which is actually a special case of the MM methods that we will give below. Next, we choose a higher order approximation for the diffusion terms, and then we obtain the MM method.

3.2 Magnus-type Milstein method

Note that $\hat{Y}(t) = \exp(-\Omega(t_n, t))$ is the solution of the following linear Itô SDE,

$$d\hat{Y}(t) = \left(-A_0 + \sum_{j=1}^m A_j^2 \right) \hat{Y}(t)dt - \sum_{j=1}^m A_j \hat{Y}(t)dW_j(t), \quad \hat{Y}(t_n) = I. \tag{12}$$

Applying the Itô–Taylor theorem to the stochastic integral

$$\begin{aligned} & \int_{t_n}^{t_{n+1}} \exp(-\Omega(0, s))g_j(y)dW_j(s) \\ &= \int_{t_n}^{t_{n+1}} \left(g_j(y_n) + \int_{t_n}^s \left((d\hat{Y})g_j(y) + \hat{Y}dg_j(y) + d\hat{Y}dg_j(y) \right) \right) dW_j(s) \\ &= \int_{t_n}^{t_{n+1}} \left(g_j(y_n) + \int_{t_n}^s \left((-A_0 + \sum_{l=1}^m A_l^2)\hat{Y}g_j(y)dt - \sum_{l=1}^m A_l \hat{Y}g_j(y)dW_l(s_1) \right. \right. \\ & \quad \left. \left. + \hat{Y}g'_j(y) \left((A_0y + g_0(y))dt + \sum_{l=1}^m (A_ly + g_l(y))dW_l(s_1) \right) \right. \right. \\ & \quad \left. \left. + \frac{1}{2}\hat{Y} \sum_{l=1}^m g''_j(A_ly + g_l(y), A_ly + g_l(y))dt + d\hat{Y}dg_j(y) \right) \right) dW_j(s) \\ &= \int_{t_n}^{t_{n+1}} \left(g_j(y_n) + \int_{t_n}^s \left(\sum_{l=1}^m (-A_lg_j(y_n) + g'_j(y_n)(A_ly_n + g_l(y_n))) \right) dW_l(s_1) \right) dW_j(s) \\ & \quad + h.o.t. \end{aligned}$$

Then we obtain the MM scheme

$$y_{n+1} = \exp(\mathcal{Q}^{[2]}(t_n, t_{n+1})) \left(y_n + \tilde{g}_0(y_n)h + \sum_{j=1}^m g_j(y_n)\Delta W_{jn} + \sum_{j,l=1}^m \mathbf{H}_{j,l}(y_n) I_{lj} \right), \tag{13}$$

where

$$\mathbf{H}_{jl}(y_n) = g'_j(y_n) (A_l y_n + g_l(y_n)) - A_l g_j(y_n).$$

As a particular example, we apply the MM method to solve the damped nonlinear Kubo oscillator (21) in Section 5 with $m = 2$, $\omega_0 = \omega_1 = \omega_2 = 1$, $\beta_0 = -1$, $\beta_1 = -1/2$, $\beta_2 = 0$ and $\alpha = 0$. It is

$$y_{n+1} = \exp(\mathcal{Q}^{[2]}(t_n, t_{n+1})) \left(y_n + \tilde{g}_0(y_n)h + \sum_{j=1}^2 g_j(y_n)\Delta W_{jn} + \sum_{j,l=1}^2 \mathbf{H}_{j,l}(y_n) I_{lj} \right),$$

where

$$\begin{aligned} \tilde{g}_0(y_n) &= g_0(y_n) - \begin{pmatrix} 0 & -1 \\ 1 & -1/2 \end{pmatrix} g_1(y_n) - \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} g_2(y_n), \\ H_{11}(y_n) &= g'_1(y_n) \left(\begin{pmatrix} 0 & -1 \\ 1 & -1/2 \end{pmatrix} y_n + g_1(y_n) \right) - \begin{pmatrix} 0 & -1 \\ 1 & -1/2 \end{pmatrix} g_1(y_n), \\ H_{12}(y_n) &= g'_1(y_n) \left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} y_n + g_2(y_n) \right) - \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} g_1(y_n), \\ H_{21}(y_n) &= g'_2(y_n) \left(\begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} y_n + g_1(y_n) \right) - \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} g_2(y_n), \\ H_{22}(y_n) &= g'_2(y_n) \left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} y_n + g_2(y_n) \right) - \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} g_2(y_n). \end{aligned}$$

As seen above, the disadvantage of the MM method is that a large number of derivatives and matrix operations need to be calculated for each step as the number of the nonlinear noise terms increases, which greatly reduces the efficiency of the method. From this point, it is natural to think of a Magnus-type Derivative-free (MDF) method, that is, use finite differences instead of these derivatives.

3.3 Magnus-type Derivative-free method

The MDF method can be derived from the MM method by replacing these derivatives with finite differences,

$$g'_j(y_n)(A_l y_n + g_l(y_n)) \approx \frac{g_j(Y_l) - g_j(y_n)}{\sqrt{h}},$$

where

$$Y_l = y_n + hg_0(y_n) + \sqrt{h}(A_l y_n + g_l(y_n)), \quad l = 1, 2, \dots, m.$$

We obtain the MDF method

$$\begin{aligned}
 y_{n+1} = \exp(\Omega^{[2]}(t_n, t_{n+1})) & \left(y_n + \tilde{g}_0(y_n)h + \sum_{j=1}^m g_j(y_n)\Delta W_{jn} \right. \\
 & \left. + \sum_{j,l=1}^m \hat{\mathbf{H}}_{jl} \frac{I_{lj}}{\sqrt{h}} \right), \tag{14}
 \end{aligned}$$

where $\hat{\mathbf{H}}_{jl} = g_j(Y_l) - g_j(y_n) + (\exp(-A_l\sqrt{h}) - I)g_j(y_n)$.

4 Convergence analysis

In this section we will give the mean-square convergence results for both the ME method and the MM method. From [19], we review the following fundamental convergence theorem of one-step numerical methods.

Lemma 2 *Suppose that the one-step approximation y_{t+h} has order of accuracy p_1 for the mathematical expectation of the deviation and order of accuracy p_2 for the mean-square deviation; more precisely, for arbitrary $t_0 \leq t \leq T - h$, $y(t) = y \in \mathbb{R}^d$ the following inequalities hold:*

$$\begin{aligned}
 |\mathbb{E}(y(t+h) - y_{t+h})| & \leq K\sqrt{1 + |y|^2}h^{p_1}, \\
 (\mathbb{E}|(y(t+h) - y_{t+h})|^2)^{1/2} & \leq K\sqrt{1 + |y|^2}h^{p_2},
 \end{aligned}$$

with $p_2 \geq 1/2$, $p_1 \geq p_2 + 1/2$, i.e., the approximation is consistent in the mean order p_1 and in the mean-square order p_2 . Then for any N and $k = 0, 1, \dots, N$ the following inequality holds:

$$\left[\mathbb{E} |y(t_k) - y_k|^2 \right]^{1/2} \leq K \left(1 + \mathbb{E} |y_0|^2 \right)^{1/2} h^{p_2 - 1/2},$$

i.e., the method is convergent of order $p_2 - 1/2$ in the sense of mean-square.

The following theorems show the convergence results of the Magnus-type method. We suppose that the coefficients of (1) satisfy the linear growth condition and the global Lipschitz condition. We also assume uniformly bounded derivatives up to order 2 for the MM method and MDF method. Let $y(t_n + h)$ be the exact evaluation of (1) at t_{n+1} starting from $y(t_n) = y_n$. We will estimate the p_1, p_2 for the Magnus-type method satisfying

$$\begin{aligned}
 |\mathbb{E}(y(t_n + h) - y_{n+1})| & = O(h^{p_1}), \\
 \left(\mathbb{E} |y(t_n + h) - y_{n+1}|^2 \right)^{1/2} & = O(h^{p_2}).
 \end{aligned}$$

Theorem 2 *Let y_n be an approximation to the solution of (1) using the ME method. Then for any N and $k = 0, 1, \dots, N$ the following inequality holds:*

$$\left[\mathbb{E} |y(t_k) - y_k|^2 \right]^{1/2} \leq K \left(1 + \mathbb{E} |y_0|^2 \right)^{1/2} h^{1/2},$$

i.e., the ME method is convergent of order 1/2 in the sense of mean-square.

Proof For the ME method, it can be readily shown that

$$y(t_n + h) - y_{n+1} = P_1 + P_2 + P_3,$$

where

$$\begin{aligned} P_1 &= \left(\exp(\Omega(t_n, t_{n+1})) - \exp(\Omega^{[1]}(t_n, t_{n+1})) \right) y_n, \\ P_2 &= \exp(\Omega(t_n, t_{n+1})) \int_{t_n}^{t_{n+1}} \exp(-\Omega(t_n, s)) \tilde{g}_0(y(s)) ds \\ &\quad - \exp(\Omega^{[1]}(t_n, t_{n+1})) \int_{t_n}^{t_{n+1}} \tilde{g}_0(y(t_n)) ds, \\ P_3 &= \exp(\Omega(t_n, t_{n+1})) \sum_{j=1}^m \int_{t_n}^{t_{n+1}} \exp(-\Omega(t_n, s)) g_j(y(s)) dW_j(s) \\ &\quad - \exp(\Omega^{[1]}(t_n, t_{n+1})) \sum_{j=1}^m \int_{t_n}^{t_{n+1}} g_j(y(t_n)) dW_j(s). \end{aligned}$$

For term P_1 , since $\exp(\Omega(t_n, t_{n+1}))y_n$ is the solution of (5), we can easily find

$$|\mathbb{E}(P_1)| = O(h^2), \quad (\mathbb{E}|P_1|^2)^{1/2} = O(h). \tag{15}$$

For P_2 , adding and subtracting $\exp(\Omega(t_n, t_{n+1})) \int_{t_n}^{t_{n+1}} \tilde{g}_0(y_n) ds$ give

$$\begin{aligned} P_2 &= \exp(\Omega(t_n, t_{n+1})) \int_{t_n}^{t_{n+1}} (\exp(-\Omega(t_n, s)) \tilde{g}_0(y(s)) - \tilde{g}_0(y_n)) ds \\ &\quad + \left(\exp(\Omega(t_n, t_{n+1})) - \exp(\Omega^{[1]}(t_n, t_{n+1})) \right) \int_{t_n}^{t_{n+1}} \tilde{g}_0(y(t_n)) ds \\ &= P_{21} + P_{22}. \end{aligned}$$

Utilizing (1) and (12),

$$\begin{aligned} & \exp(-\Omega(t_n, s)) \tilde{g}_0(y(s)) \\ &= \tilde{g}_0(y_n) + \int_{t_n}^s \left(-A_0 + \sum_{j=1}^m A_j^2 \right) Y(s_1) \tilde{g}_0(y(s_1)) ds_1 \\ & \quad - \int_{t_n}^s \sum_{j=1}^m A_j Y(s_1) \tilde{g}_0(y(s_1)) dW_j(s_1) + \int_{t_n}^s Y(s_1) d\tilde{g}_0(y(s_1)) \\ & \quad + \int_{t_n}^s \sum_{j=1}^m A_j Y(s_1) (A_j y(s_1) + g_j(y(s_1))) ds_1, \end{aligned}$$

we can easily find

$$\begin{aligned} |\mathbb{E}(P_{21})| &= O(h^2), \quad (\mathbb{E}|P_{21}|^2)^{1/2} = O(h^{3/2}), \\ |\mathbb{E}(P_{22})| &= O(h^3), \quad (\mathbb{E}|P_{22}|^2)^{1/2} = O(h^2). \end{aligned} \tag{16}$$

For P_3 , adding and subtracting $\exp(\Omega(t_n, t_{n+1})) \int_{t_n}^{t_{n+1}} g_j(y_n) dW_j(s)$ gives

$$\begin{aligned} P_3 &= \exp(\Omega(t_n, t_{n+1})) \sum_{j=1}^m \int_{t_n}^{t_{n+1}} (\exp(-\Omega(t_n, s)) g_j(y(s)) - g_j(y_n)) dW_j(s) \\ & \quad + \left(\exp(\Omega(t_n, t_{n+1})) - \exp(\Omega^{[1]}(t_n, t_{n+1})) \right) \sum_{j=1}^m \int_{t_n}^{t_{n+1}} g_j(y(t_n)) dW_j(s). \end{aligned}$$

Similar to term P_2 , we can get

$$|\mathbb{E}(P_3)| = O(h^2), \quad (\mathbb{E}|P_3|^2)^{1/2} = O(h). \tag{17}$$

With (15), (16) and (17), we have $p_1 = 2, p_2 = 1$. From Lemma 2, we know that the ME method is of mean-square order 0.5. □

For the MM methods, we can obtain the following convergence results.

Theorem 3 *Let y_n be an approximation to the solution of (1) using the MM method. Then for any N and $k = 0, 1, \dots, N$ the following inequality holds:*

$$\left[\mathbb{E} |y(t_k) - y_k|^2 \right]^{1/2} \leq K \left(1 + \mathbb{E} |y_0|^2 \right)^{1/2} h,$$

i.e., the MM method is convergent of order 1 in the sense of mean-square.

Proof For the MM method, through direct calculation, we find

$$y(t_n + h) - y_{n+1} = P_1 + P_2 + P_3,$$

where

$$\begin{aligned}
 P_1 &= \left(\exp(\Omega(t_n, t_{n+1})) - \exp(\Omega^{[2]}(t_n, t_{n+1})) \right) y_n, \\
 P_2 &= \exp(\Omega(t_n, t_{n+1})) \int_{t_n}^{t_{n+1}} \exp(-\Omega(t_n, s)) \tilde{g}_0(y(s)) ds \\
 &\quad - \exp(\Omega^{[2]}(t_n, t_{n+1})) \int_{t_n}^{t_{n+1}} \tilde{g}_0(y(t_n)) ds, \\
 P_3 &= \exp(\Omega(t_n, t_{n+1})) \sum_{j=1}^m \int_{t_n}^{t_{n+1}} \exp(-\Omega(t_n, s)) g_j(y(s)) dW_j(s) \\
 &\quad - \exp(\Omega^{[2]}(t_n, t_{n+1})) \sum_{j=1}^m \left(\int_{t_n}^{t_{n+1}} g_j(y(t_n)) dW_j(s) \right. \\
 &\quad \left. + \sum_{l=1}^m \mathbf{H}_{j,l}(y_n) \int_{t_n}^{t_{n+1}} \int_{t_n}^s dW_l(s_1) dW_j(s) \right).
 \end{aligned}$$

For term P_1 , since $\exp(\Omega(t_n, t_{n+1}))y_n$ is the solution of (5), we can easily see through an Itô-Taylor expansion

$$|\mathbb{E}(P_1)| = O(h^2), \quad (\mathbb{E}|P_1|^2)^{1/2} = O(h^{3/2}). \tag{18}$$

For P_2 , adding and subtracting $\exp(\Omega(t_n, t_{n+1})) \int_{t_n}^{t_{n+1}} \tilde{g}_0(y_n) ds$, we have

$$\begin{aligned}
 P_2 &= \exp(\Omega(t_n, t_{n+1})) \int_{t_n}^{t_{n+1}} (\exp(-\Omega(t_n, s)) \tilde{g}_0(y(s)) - \tilde{g}_0(y_n)) ds \\
 &\quad + \left(\exp(\Omega(t_n, t_{n+1})) - \exp(\Omega^{[2]}(t_n, t_{n+1})) \right) \int_{t_n}^{t_{n+1}} \tilde{g}_0(y(t_n)) ds \\
 &= P_{21} + P_{22},
 \end{aligned}$$

and

$$\begin{aligned}
 |\mathbb{E}(P_{21})| &= O(h^2), \quad (\mathbb{E}|P_{21}|^2)^{1/2} = O(h^{3/2}), \\
 |\mathbb{E}(P_{22})| &= O(h^3), \quad (\mathbb{E}|P_{22}|^2)^{1/2} = O(h^{5/2}).
 \end{aligned} \tag{19}$$

For term P_3 , since

$$\begin{aligned}
 &\exp(\Omega(t_n, t_{n+1})) \int_{t_n}^{t_{n+1}} \exp(-\Omega(t_n, s)) g_j(y(s)) dW_j(s) \\
 &= \exp(\Omega(t_n, t_{n+1})) \left(\int_{t_n}^{t_{n+1}} g_j(y_n) dW_j(s) + \int_{t_n}^{t_{n+1}} \int_{t_n}^s (g'_j(y_n) (A_l y_n + g_l(y_n)) \right. \\
 &\quad \left. - A_l g_j(y_n)) dW_l(s_1) dW_j(s) + R_{1j} \right),
 \end{aligned}$$

where

$$|\mathbb{E}(R_{1j})| = O(h^2), \quad (\mathbb{E}|R_{1j}|^2)^{1/2} = O(h^{3/2}),$$

we have

$$P_3 = \left(\exp(\Omega(t_n, t_{n+1})) - \exp(\Omega^{[2]}(t_n, t_{n+1})) \right) \sum_{j=1}^m \left(\int_{t_n}^{t_{n+1}} g_j(y(t_n)) dW_j(s) + \sum_{l=1}^m \mathbf{H}_{j,l}(y_n) \int_{t_n}^{t_{n+1}} \int_{t_n}^s dW_l(s_1) dW_j(s) \right) + \sum_{j=1}^m \exp(\Omega(t_n, t_{n+1})) R_{1j},$$

which gives

$$|\mathbb{E}(P_3)| = O(h^2), \quad (\mathbb{E}|P_3|^2)^{1/2} = O(h^{3/2}). \tag{20}$$

From (18), (19) and (20), $p_1 = 2, p_2 = 3/2$. Hence by Lemma 2, we see that the MM method is of mean-square order 1. \square

The following theorem describes the convergence of the MDF method. The proof is analogous to that of Theorem 3 and we omit it.

Theorem 4 *Let y_n be an approximation to the solution of (1) using the MDF method. Then for any N and $k = 0, 1, \dots, N$ the following inequality holds:*

$$\left[\mathbb{E} |y(t_k) - y_k|^2 \right]^{1/2} \leq K \left(1 + \mathbb{E} |y_0|^2 \right)^{1/2} h.$$

Thus the MDF method is convergent of order 1 in the sense of mean-square.

5 Implementation and numerical tests

Using the Milstein or MM method to generate numerical approximations, the iterated integrals (7) are included in the numerical scheme. For $i = j$, there is such a relationship $I_{ii}(t_n, t_n + h) = \frac{1}{2}((\Delta W_{in})^2 - h)$. For $i \neq j$, this issue has been fully described in [20, 21] and [22] by Kuznetsov and Wiktorsson from a different perspective, respectively. Both of them expand the iterated integrals (7) into infinite series, and then truncate the infinite series to approximate the iterated integrals. A brief overview is given in the Appendix.

To compare the convergence rate of the truncated series, for different step sizes and $m = 2$, we respectively give the minimum truncated terms indices q_w, q_t, q_p of the Wiktorsson’s algorithm, Kuznetsov’s algorithm with the orthonormal system of trigonometric functions and Legendre polynomials in Table 1. We can see that Wiktorsson’s algorithm requires smaller truncated indices to achieve the mean square error $O(h^3)$, especially for a smaller step size h , which also means that Wiktorsson’s algorithm needs to simulate fewer random variables.

In order to show the specific performance of the three algorithms, for the integral I_{21} , choosing $h = 2^{-6}$, the mean ($\mathbb{E}(I_{21}) = 0$) and standard deviation ($(\mathbb{E}(I_{21})^2)^{1/2} = h/\sqrt{2} \approx 0.01105$) are calculated in Table 2 with 2000 samples. Their mean and standard deviation are almost in the same range. At the same time, the numerical tests in the next will also show that using Kuznetsov and Wiktorsson algorithms in practical implementation can give almost the same simulation results for different step size.

Table 1 The minimum q_w, q_t, q_p that needs to be selected for different time steps

h	2^{-3}	2^{-4}	2^{-5}	2^{-6}	2^{-7}	2^{-8}	2^{-9}	2^{-10}	2^{-11}	2^{-12}	2^{-13}
q_w	1	2	2	3	4	5	7	10	14	19	27
q_t	2	3	5	10	20	39	78	156	312	623	1246
q_p	2	3	5	9	17	33	65	129	257	513	1025

For the rest of this section, the performances of the introduced Magnus-type methods are presented and then compared with classical stochastic methods, namely the Euler–Maruyama method and the Milstein method. For convenience, let M (W) denote the mean-square order 1.0 Milstein method, where the iterated Itô integrals are simulated by Wiktorsson’s algorithm. Let $M(K_t)$ and $M(K_p)$ denote the mean-square order 1.0 Milstein method, where the iterated Itô integrals are simulated by Kuznetsov’s algorithm with the orthonormal system of trigonometric functions and Legendre polynomials, respectively. This shorthand notation also applies to the MM method and the MDF method.

To present the performance of Wiktorsson’s and Kuznetsov’s algorithms for the iterated Itô integrals, we will compare the number of random variables required by Wiktorsson’s and Kuznetsov’s algorithms in each iteration as a measure of computational effort. The exponential function needs to be calculated during the implementation of the numerical algorithm, which is not the subject of this article, so we compare the time required by Wiktorsson’s and Kuznetsov’s algorithms to generate random variables as a measure of the efficiency. In all numerical simulations, we choose the minimum truncation number.

5.1 Damped nonlinear Kubo oscillator

As a first numerical test, we consider the damped nonlinear Kubo oscillator

$$dy(t) = \sum_{j=0}^m \left[\omega_j \begin{pmatrix} 0 & -1 \\ 1 & \beta_j \end{pmatrix} y(t) + \begin{pmatrix} 0 & -g_j(y(t)) \\ g_j(y(t)) & 0 \end{pmatrix} y(t) \right] dW_j(t) \quad (21)$$

Table 2 Iterated Itô integral I_{21} with truncating the first 3, 5, 7, 11 terms of the infinite series, respectively

	q = 3		q = 5		q = 7		q = 11	
	mean	stdev	mean	stdev	mean	stdev	mean	stdev
Wiktorsson	−0.0003	0.01113	0.0002	0.01096	−0.0003	0.01108	0.0001	0.01113
K_t	−0.0001	0.01059	−0.0002	0.01040	0.0002	0.01092	0.0001	0.01138
K_p	−0.0001	0.01056	−0.0002	0.01069	0.0002	0.01068	0.0001	0.01091

with $g_j : \mathbb{R}^2 \rightarrow \mathbb{R}$, where $t \in [0, T]$ and $\omega_j, \beta_j \in \mathbb{R}$. This problem is investigated in [24] with $\beta_j = 0$. We also let $m = 2$ and

$$g_0(y(t)) = \frac{1}{5} (y_1 + y_2)^5, g_1(y(t)) = 0, g_2(y(t)) = \frac{1}{3} (y_1 + y_2)^3.$$

Since we do not know the exact solution of (21), the reference solution is produced by the Milstein method with a small step $h = 2^{-19}$. Set the parameters $\omega_0 = 2, \omega_1 = 0.5, \omega_2 = 0.2, \beta_0 = -0.5, \beta_1 = -0.2, \beta_2 = -0.1$ and initial value $y_0 = (1, 1)$. The mean-square convergence order of ME method, MM method and MDF method is presented in Fig. 1. Here, the mean-square error

$$\|\text{ERROR}\|_{L2} = \sqrt{\frac{\sum_{j=1}^{2000} |y_N - y(T)|^2}{2000}}$$

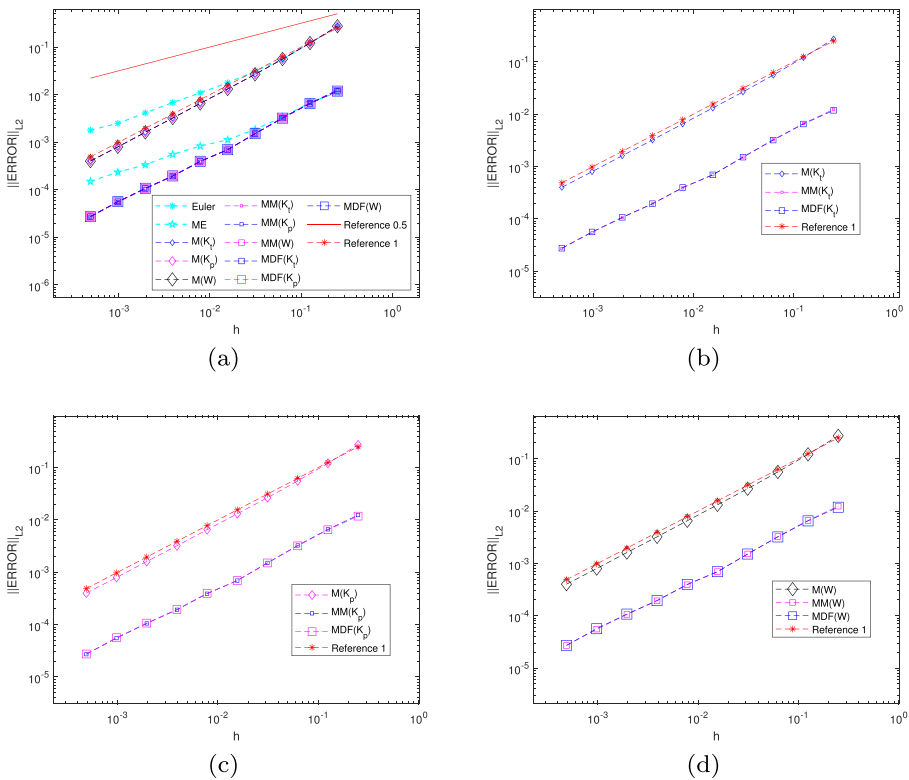


Fig. 1 **a** The convergence rate of the ME method, MM method and MDF method for solving (21) with $g_0(y(t)) = \frac{1}{5} (y_1 + y_2)^5, g_1(y(t)) = 0$ and $g_2(y(t)) = \frac{1}{3} (y_1 + y_2)^3$. **b** Kuznetsov's algorithms with the orthonormal system of trigonometric functions. **c** Kuznetsov's algorithms with the orthonormal system of Legendre polynomials. **d** Wiktorsson's algorithm

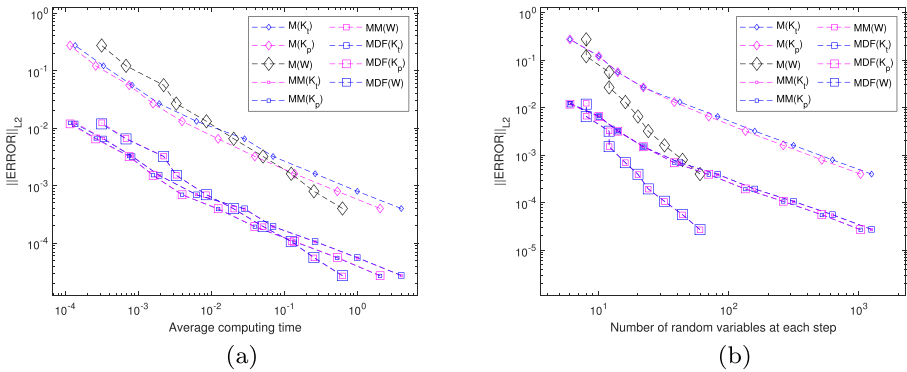


Fig. 2 Comparison of Wiktorsson’s and Kuznetsov’s algorithm in average computing time and number of random variables at each step for (21) with a set of increasing time steps $h = 2^i, i = -11, \dots, -2$. **a** Average computing time. **b** Number of random variables at each step

is calculated at the time $T = 1$ for a set of increasing time steps $h = 2^i, i = -11, \dots, -2$. As some of the lines are on the top of each other, we use separate figures to show the performance of Wiktorsson’s and Kuznetsov’s algorithms in Fig. 1. From Fig. 1a, we can see that the Magnus-type methods enjoy considerably smaller error coefficient compared to the Euler–Maruyama method and the Milstein method and the performance of the three simulations of the iterated Itô integral is similar. It can be seen from Fig. 2a that when the step size is large (i.e., the error is large), Wiktorsson’s and Kuznetsov’s algorithms perform similarly in terms of calculation time. As the step size becomes smaller (i.e., the error becomes smaller), Wiktorsson’s algorithm performs better, as it can be seen in (b) that the number of random variables that need to be simulated at each step of Wiktorsson’s algorithm is less. In addition, Kuznetsov’s algorithm with polynomial functions is slightly better than Kuznetsov’s algorithm with trigonometric functions in terms of calculation time and the number of random variables that need to be simulated at each step. The increment for different steps is given by

$$\begin{aligned}
 I_{i,j,t_n,t_n+ph} &= \int_{t_n}^{t_n+ph} \int_{t_n}^s dW_i(s_1)dW_j(s) \\
 &\approx I_{i,j,t_n,t_n+ph}^q \\
 &= \sum_{k=0}^{p-1} \left[I_{i,j,t_n+kh,t_n+(k+1)h}^q + \Delta W_{j,n+k} I_{k \geq 1} \sum_{l=0}^{k-1} \Delta W_{i,n+l} \right].
 \end{aligned}$$

5.2 SDE with linear non-commutative noise

We consider the following SDE [2] in \mathbb{R}^4 , with

$$dy = (rA_0y + \mathbf{F}(y))dt + G(y)d\mathbf{W}(t), \tag{22}$$

where $r = 4$, $\mathbf{F}(y) = (F_1, F_2, F_3, F_4)$, $F_j = \frac{y_j}{1+|y_j|}$, $j = 1, 2, 3, 4$, and initial value $y(0) = y_0$. $d\mathbf{W}(t) = (dW_1(t), dW_2(t), dW_3(t), dW_4(t))$. Here A_0 takes the form

$$A_0 = \begin{pmatrix} -2 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -2 \end{pmatrix},$$

which usually comes from a discrete Laplacian operator. Consider the following non-commutative noise

$$G(y) = \begin{pmatrix} \beta y_1 & 0 & 0 & 0 \\ 0 & \beta y_2 - \alpha y_1 & 0 & 0 \\ 0 & 0 & \beta y_3 - \alpha y_2 & 0 \\ 0 & 0 & 0 & \beta y_4 - \alpha y_3 \end{pmatrix}.$$

Then, we have

$$A_1 = \begin{pmatrix} \beta & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -\alpha & \beta & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, A_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -\alpha & \beta & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, A_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\alpha & \beta \end{pmatrix}.$$

Simulations with $\alpha = 0.5$, $\beta = 1$ and $y_0 = (1, 1, 1, 1)^T$ are carried out until the time $T = 1$ with a set of increasing time steps $h = 2^i$, $i = -11, \dots, -3$. The mean-square convergence order of the ME method and the MM method is presented in Fig. 3. In each case 2000 samples are generated as before. Here, the reference solution is produced by the Milstein method with the step $h = 2^{-16}$. Figure 3 compares the cases for $r = 4$, $\alpha = 0.8$, $\beta = 1$ in (a) and compares the cases for $r = 4$, $\alpha = 2$, $\beta = 0.1$ in (b). We see that the case $r = 4$, $\alpha = 0.8$, $\beta = 1$ enjoys a smaller error than the mildly stiff case $r = 4$, $\alpha = 2$, $\beta = 0.1$. Figure 4a shows that Wiktorsson’s and Kuznetsov’s algorithms perform similarly in terms of calculation time. This is because the truncation index of Wiktorsson’s algorithm is related to m , so when m and the step size are both large, Kuznetsov’s algorithm is

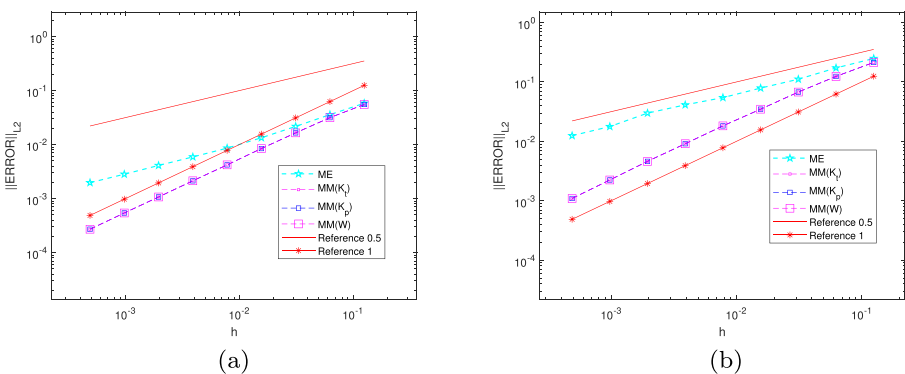


Fig. 3 The convergence rate of the ME method and the MM method for solving (22) with $m = 4$ in a $r = 4$, $\alpha = 0.8$, $\beta = 1$ and in b $r = 4$, $\alpha = 2$, $\beta = 0.1$

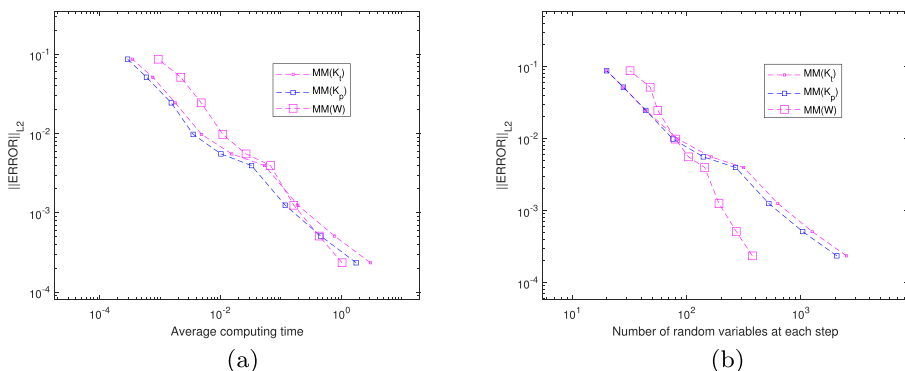


Fig. 4 Comparison of the algorithms in computing time and number of random variables at each step with a set of increasing time steps $h = 2^i, i = -11, \dots, -3$. Here $r = 4, \alpha = 0.8$ and $\beta = 1$. **a** Average computing time. **b** Number of random variables at each step

a sensible choice. In addition, Kuznetsov’s algorithm with polynomial functions is also slightly better than Kuznetsov’s algorithm with trigonometric functions in terms of calculation time and the number of random variables that need to be simulated at each step for $m = 4$. As the step size decreases, Wiktorsson’s algorithm has obvious advantages in the number of random variables at each step.

5.3 Stochastic Manakov equation

In order to confirm the performance of the methods for high-dimensional SDEs, we consider the stochastic Manakov system

$$idu + \left(\partial_x^2 u + |u|^2 u\right) dt + i\sqrt{\gamma} \sum_{k=1}^3 \sigma_k \partial_x u \circ dW_k = 0, \tag{23}$$

$$u(0, x) = (\cos(\pi/8) \sin(\pi/8)) \operatorname{sech}(x), \quad x \in [-a, a]$$

$$u(t, -a) = u(t, a) = 0, \quad t \in [0, T],$$

where $u = u(t, x) = (u_1, u_2) \in \mathbb{C}^2$ with $t \geq 0$ and $x \in \mathbb{R}$. The symbol \circ means that the stochastic integrals are established in the sense of Stratonovich and $\gamma \geq 0$ is the noise intensity. Here $|u|^2 u = (|u_1|^2 + |u_2|^2) u$ is the nonlinear term, and σ_1, σ_2 and σ_3 are the Pauli matrices taking the form

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The stochastic Manakov system (23) is usually used to describe pulse propagation in randomly birefringent optical fibers. The existence and uniqueness of the solution have been obtained in [25]. The equivalent Itô form is

$$du(t) = \left(C_\gamma \frac{\partial^2 u(t)}{\partial x^2} + i|u|^2 u \right) dt - \sqrt{\gamma} \sum_{k=1}^3 \sigma_k \frac{\partial u(t)}{\partial x} dW_k(t), \tag{24}$$

where $C_\gamma = i + \frac{3\gamma}{2}$. Applying the central finite-difference scheme to discretize the space interval by $N + 2$ uniform points

$$\partial_x u_t(x_i) \approx \frac{u_{t,i+1} - u_{t,i-1}}{2\Delta x} \text{ or } \frac{u_{t,i} - u_{t,i-1}}{\Delta x}, \quad \partial_{xx} u_t(x_i) \approx \frac{u_{t,i+1} - 2u_{t,i} + u_{t,i-1}}{\Delta x^2},$$

we get the non-commutative SDE

$$dy(t) = (C_\gamma A_0 y(t) + f(y(t)))dt - \sqrt{\gamma} \sum_{k=1}^3 \sigma_k A_k y(t) dW_k(t), \tag{25}$$

with

$$y(0) = \left[\cos\left(\frac{\pi}{8}\right)\text{sech}(x_1) \cdots \cos\left(\frac{\pi}{8}\right)\text{sech}(x_N) \quad \sin\left(\frac{\pi}{8}\right)\text{sech}(x_1) \cdots \sin\left(\frac{\pi}{8}\right)\text{sech}(x_N) \right]^\top,$$

where

$$y(t) \stackrel{\text{def}}{=} [u(t, x_1) \quad u(t, x_2) \quad \cdots \quad u(t, x_N)]^\top,$$

$$f(y) \stackrel{\text{def}}{=} i \left[y_1 |y_1|^2 \quad y_2 |y_2|^2 \quad \cdots \quad y_N |y_N|^2 \right]^\top,$$

and the matrices A_0, A_1, A_2 and A_3 are defined by

$$A_0 \stackrel{\text{def}}{=} N^2 \begin{bmatrix} -2 & 1 & & 0 \\ 1 & -2 & 1 & \\ & \ddots & \ddots & \ddots \\ & & 1 & -2 & 1 \\ 0 & & & 1 & -2 \end{bmatrix}, \quad A_k \stackrel{\text{def}}{=} N \begin{bmatrix} 1 & 0 & & 0 \\ -1 & 1 & 0 & \\ & \ddots & \ddots & \ddots \\ & & -1 & 1 & 0 \\ 0 & & & -1 & 1 \end{bmatrix}, \quad k = 1, 2, 3.$$

First, we set $a = 20, h = 0.001$ and $\Delta x = 2/5$ to simulate (23) with the MM method on the time interval $[0,3]$. The evolution of $|u_1|^2$ and $|u_2|^2$ is given in Fig. 5.

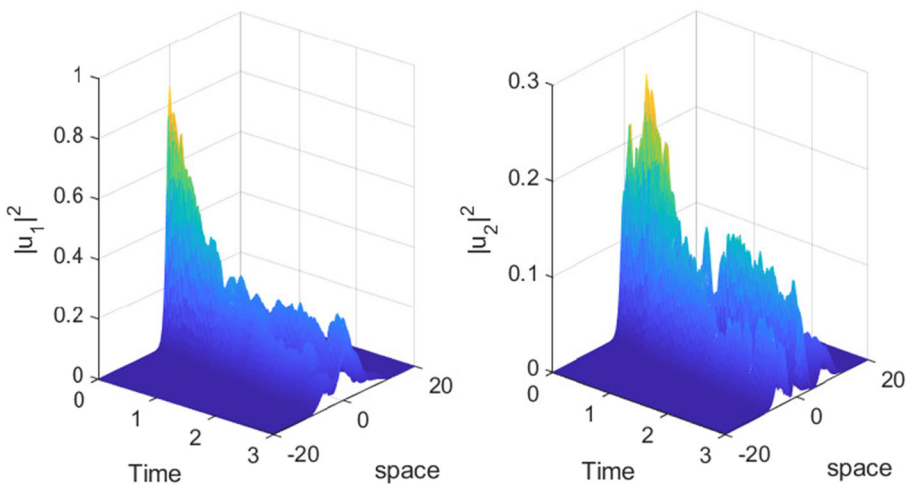


Fig. 5 Space-time evolution of the intensity of the first component (left) and the second component (right)

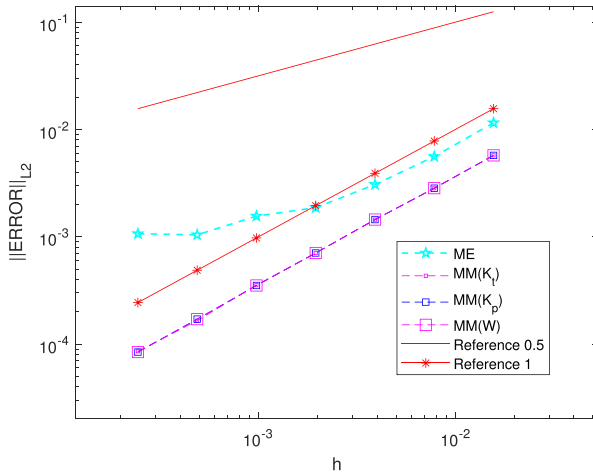


Fig. 6 The convergence rate of the ME and the MM method with $a = 50$ and $\Delta x = 2/5$, $T = 1/2$ and a set of increasing time steps $h = 2^i$, $i = -13, \dots, -7$

We can see that the results of the MM method show the energy exchange due to stochastic noise and nonlinearity. These results are similar to the structure-preserving method in [26].

Then, simulations with $a = 50$ and $\Delta x = 2/5$ are carried out until the time $T = 1/2$ with a set of increasing time steps $h = 2^i$, $i = -13, \dots, -7$. The mean-square convergence order of the ME method and the MM method is presented in Fig. 6. In each case 500 samples are generated. Here, the reference solution is produced by the Milstein method with the step $h = 2^{-16}$. In Fig. 7, (a) shows that Wiktorsson’s and Kuznetsov’s algorithms perform similarly in terms of calculation time. Wiktorsson’s algorithms is better than Kuznetsov’s algorithms in terms of calculation time and the number of random variables that need to be

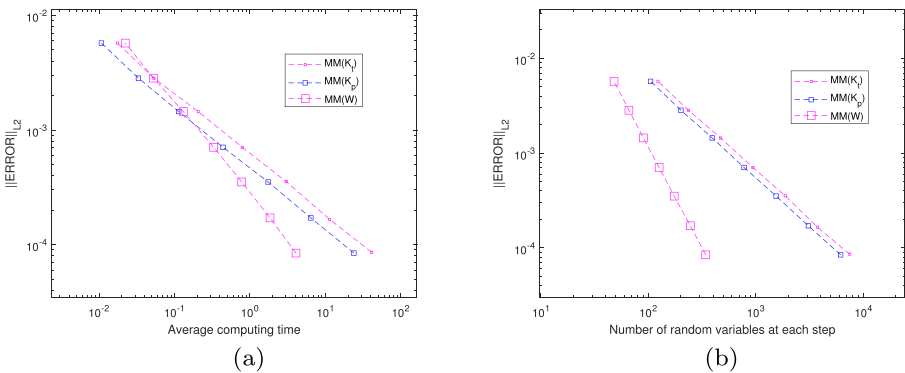


Fig. 7 Comparison of the algorithms in computing time and number of random variables at each step with a set of increasing time steps $h = 2^i$, $i = -13, \dots, -7$. Here $a = 50$, $\Delta x = 2/5$ and $T = 1/2$. **a** Average computing time. **b** Number of random variables at each step

simulated in each step. This is because as the step size decreases, the number of noise terms m has less impact on the truncation index of Wiktorsson's algorithm. In addition, as the step size decreases, the difference between Kuznetsov's algorithm with polynomial functions and Kuznetsov's algorithm with trigonometric functions is not significant.

Remark 1 Here we use the built-in `expdemo2` function in Matlab to calculate $\exp(\Omega^{[1]}(t_n, t_{n+1}))$ and $\exp(\Omega^{[2]}(t_n, t_{n+1}))$ using a Taylor series. For such high-dimensional problems, each path requires considerable computational time. Although the calculation of exponential function is not the subject of this article, the application of efficient techniques (such as Krylov subspace methods) will have a significant impact on the Magnus-type methods to address very high-dimensional semilinear SDEs.

Remark 2 In the high-dimensional problem (25), we choose mildly stiff matrices A_0 as opposed to the nonlinear term. When the stiffness is very strong, our methods require a small step size. For this type of stiff high-dimensional complex-valued SDEs, the combination of the SROCK methods [27, 28] and the Magnus method will be an interesting work that we will consider in future work.

Remark 3 With the specific performance in the above three test examples, we see that for the simulation of iterated Itô integrals, when the step size is large and the number of noise terms is large, Kuznetsov's algorithm with polynomial functions is superior, and when the step size is smaller, Wiktorsson's algorithm is better.

6 Conclusion

We have derived Magnus-type methods, based on Magnus expansion for non-commutative linear SDEs, for noncommutative Itô stochastic differential equations with semi-linear drift term and semi-linear diffusion terms. By truncating the Magnus series, the ME method, the MM method and the MDF method have been constructed. We have investigated the mean-square convergence of these methods and shown the same mean-square convergent order as the corresponding method, that is, the ME method is order 0.5, and the MM and MDF method is order 1.0. Then, we have compared two types of algorithms for simulating iterated Itô integrals. Numerical tests have been carried out to present the efficiency of the proposed methods for low-dimensional SDEs and a high-dimensional SDE from discretized SPDE.

Finally, we should make the following remarks. We can apply our methods to non-autonomous semi-linear SDEs, where we only need to truncate the non-autonomous Magnus expansion. However, the linear stability analysis of numerical methods for high-dimensional stochastic differential equations with non-commutative noises is quite complicated, especially for complex coefficient matrices. The mean-square stability analysis of Magnus-type methods for such problems is a motivation for future work. In addition, weak convergence Magnus methods will be a good choice to avoid simulating iterated stochastic integrals.

Appendix

A.1 The expansion of iterated Itô stochastic integrals based on the generalized multiple Fourier series

Let $\{\phi_j(x)\}_{j=0}^\infty$ be an orthonormal basis of the space $L_2([t_n, t_{n+1}])$. If the trigonometric series

$$\phi_j(s) = \frac{1}{\sqrt{h}} \begin{cases} 1, & j = 0, \\ \sqrt{2} \sin \frac{2\pi r(s-t_n)}{h}, & j = 2r - 1, \\ \sqrt{2} \cos \frac{2\pi r(s-t_n)}{h}, & j = 2r, \end{cases}$$

where $r = 1, 2, \dots$, are selected as an orthonormal basis, for the iterated Itô integral (7), using the expansion theorem in [20], the following representations are obtained,

$$I_{ij}(t_n, t_n + h) = \frac{1}{2}h \left[\zeta_0^{(j)} \zeta_0^{(i)} + \frac{1}{\pi} \sum_{r=1}^\infty \frac{1}{r} \left\{ \zeta_{2r}^{(j)} \zeta_{2r-1}^{(i)} - \zeta_{2r-1}^{(j)} \zeta_{2r}^{(i)} \right. \right. \tag{A.1} \\ \left. \left. + \sqrt{2} \left(\zeta_{2r-1}^{(j)} \zeta_0^{(i)} - \zeta_0^{(j)} \zeta_{2r-1}^{(i)} \right) \right\} \right], \quad i \neq j,$$

where $\zeta_0^{(i)} \stackrel{\text{def}}{=} W_i(h)/\sqrt{h}$ and $\zeta_j^{(i)} \stackrel{\text{def}}{=} \int_t^T \phi_j(s) dW_i(s)$, $i = 1, \dots, m$. The expansion (A.1) coincides with Kloeden, Platen and Wright’s (1992) algorithm [23] based on Kahunen–Loève expansion and it converges to the iterated Itô integral (7) in the mean-square sense.

If orthonormal Legendre polynomials are selected as an orthonormal basis,

$$\{\phi_j(s)\}_{j=0}^\infty, \phi_j(s) = \sqrt{\frac{2j+1}{h}} P_j \left(\left(s - \frac{h}{2} \right) \frac{2}{h} \right)$$

the expansion is

$$I_{ij}(t_n, t_n + h) = \frac{h}{2} \left[\zeta_0^{(i)} \zeta_0^{(j)} + \sum_{r=1}^\infty \frac{1}{\sqrt{4r^2 - 1}} \left\{ \zeta_{r-1}^{(i)} \zeta_r^{(j)} - \zeta_r^{(i)} \zeta_{r-1}^{(j)} \right\} \right]. \tag{A.2}$$

As the expansions (A.1) and (A.2) are infinite series, we need to truncate them for practical simulation. During the implementation of the Milstein and MM methods, the mean-square error of the approximated iterated integrals should not be larger than h^3 , which is to ensure mean-square convergence of 1.

Truncating the infinite series (A.1) and (A.2) to q terms, we have I_{ij}^q . From the truncated mean-square error in [21], we obtain the mean-square error for truncating (A.1)

$$E \left\{ \left(I_{ij}(t_n, t_n + h) - I_{ij}^q(t_n, t_n + h) \right)^2 \right\} \leq \frac{3h^2}{2\pi^2 q}.$$

Here we need $3h^2/(2\pi^2 q) \leq h^3$, so we should choose

$$q_t \geq \left\lceil \frac{3}{2\pi^2 h} \right\rceil.$$

Truncating the expansion (A.1) to q , we get the mean-square error

$$E \left\{ \left(I_{ij}(t_n, t_n + h) - I_{ij}^q(t_n, t_n + h) \right)^2 \right\} \leq -\frac{h^2}{8} \ln \left| 1 - \frac{2}{2q + 1} \right|.$$

Here we need the mean-square error less than h^3 , so we should choose

$$q_p \geq \left\lceil \left(\frac{2}{1 - e^{-8h}} - 1 \right) / 2 \right\rceil.$$

Since $q_p \approx 1/(8h)$, the truncated indices of A.1 and (A.2) both are $O(1/h)$, and it can be calculated that the convergence speed of (A.1) is about $q_t/q_p \approx 1.22$ times that of (A.2). This means that the approximation based on the multiple Legendre Fourier series is slightly more efficient than that based on the multiple trigonometric Fourier series.

A.2 Wiktorsson’s algorithm

If $\mathbf{I}(h)$ and $\mathbf{A}(h)$ are matrices with elements $I_{ij}(t_n, t_n + h)$, $A_{ii} = 0$ and

$$A_{ij}(h) = \frac{h}{2\pi} \sum_{r=1}^{\infty} \frac{1}{r} \left\{ \zeta_{2r}^{(j)} \zeta_{2r-1}^{(i)} - \zeta_{2r-1}^{(j)} \zeta_{2r}^{(i)} + \sqrt{2} \left(\zeta_{2r-1}^{(j)} \zeta_0^{(i)} - \zeta_0^{(j)} \zeta_{2r-1}^{(i)} \right) \right\}, i \neq j$$

respectively, the approximation based on the multiple trigonometric Fourier series can be written in matrix form

$$\begin{aligned} \mathbf{I}(h) &= \frac{\Delta \mathbf{W}(h) \Delta \mathbf{W}(h)^T - h I_m}{2} + \mathbf{A}(h), \\ \mathbf{A}(h) &= \frac{h}{2\pi} \sum_{k=1}^{\infty} \frac{1}{k} \left\{ \mathbf{X}_k \left(\mathbf{Y}_k + \sqrt{2/h} \Delta \mathbf{W}(h) \right)^T - \left(\mathbf{Y}_k + \sqrt{2/h} \Delta \mathbf{W}(h) \right) \mathbf{X}_k^T \right\}, \end{aligned} \tag{A.3}$$

where $\Delta \mathbf{W}(h) \sim N(0, h I_m)$, $\mathbf{X}_k \sim N(0_m, I_m)$ and $\mathbf{Y}_k \sim N(0_m, I_m)$, $k = 1, 2, \dots, q$ are all independent. As the Lévy stochastic area has the relationship $I_{ij} - I_{ji} = 2A_{ij}$, we only need to simulate I_{ij} for $i < j$ or $i > j$ in the simulation.

Wiktorsson’s algorithm is based on approximation of the tail-sum distribution of (A.3)

$$\varepsilon_q = \frac{h}{2\pi} \sum_{k=q+1}^{\infty} \frac{1}{k} \left\{ \mathbf{X}_k \left(\mathbf{Y}_k + \sqrt{2/h} \Delta \mathbf{W}(h) \right)^T - \left(\mathbf{Y}_k + \sqrt{2/h} \Delta \mathbf{W}(h) \right) \mathbf{X}_k^T \right\},$$

which improves the rate of convergence. Let $\text{vec}(\mathbf{I}(h)^T)$ be column vectors of (A.3), and the following is the procedure of Wiktorsson’s algorithm [22]:

1. First simulate $\Delta \mathbf{W}(h) \sim N(0_m, \sqrt{h} I_m)$.
2. Simulate the truncated first q terms

$$\tilde{A}^{(q)}(h) = \frac{h}{2\pi} \sum_{k=1}^q \frac{1}{k} K_m(P_m - I_{m^2}) \left\{ \left(\mathbf{Y}_k + \sqrt{\frac{2}{h}} \Delta \mathbf{W}(h) \right) \otimes \mathbf{X}_k \right\}$$

where $\mathbf{X}_k \sim N(0_m, I_m)$ and $\mathbf{Y}_k \sim N(0_m, I_m)$

3. Then simulate $\mathbf{G}_q \sim N(0_M, I_M)$ and add the approximation of tail-sum distribution:

$$\tilde{A}^{(q)'}(h) = \tilde{A}^{(q)}(h) + \frac{h}{2\pi} a_q^{1/2} \sqrt{\Sigma_\infty} \mathbf{G}_q$$

where $a_q = \sum_{k=q+1}^\infty 1/k^2$

4. Finally obtain the approximation $\text{vec}(\mathbf{I}(h)^T)^{(q)'}$ of $\text{vec}(\mathbf{I}(h)^T)$

$$\text{vec}(\mathbf{I}(h)^T)^{(q)'} = \frac{\Delta \mathbf{W}(h) \otimes \Delta \mathbf{W}(h) - \text{vec}(h I_{m^2})}{2} + (I_{m^2} - P_m) K_m^T \tilde{A}^{(q)'}(h)$$

Here P_m is the $m^2 \times m^2$ permutation matrix and for the specific expression of matrices P_m please refer to [22].

The mean-square error of Wiktorsson’s algorithm [22] is

$$\max_{i,j} E \left\{ \left(I_{ij}(h) - I_{ij}^{(q)'}(h) \right)^2 \right\} \leq \frac{5h^2}{24\pi^2 q^2} m^2(m-1).$$

Here we also need the mean-square error less than h^3 , so we should choose truncated indices as

$$q_w \geq \left\lceil \frac{\sqrt{5m^2(m-1)/(24\pi^2)}}{\sqrt{h}} \right\rceil,$$

and the truncated indices q_w is $O(1/\sqrt{h})$ that is much better than $O(1/h)$.

Acknowledgements The authors appreciate the valuable comments of the referee.

Funding Guoguo Yang was supported by China Scholarship Council (CSC) and the National Natural Science Foundation of China (No. 62073103 and 11701124) during his study at Queensland University of Technology. Yoshio Komori was partially supported for this work by JSPS Grant-in-Aid for Scientific Research 17K05369. Xiaohua Ding was partially supported by the National Key R&D Program of China (No. 2017YFC1405600).

References

1. Debrabant, K., Kværnø, A., Mattsson, N.C.: Runge–Kutta Lawson schemes for stochastic differential equations. arXiv:1909.11629 (2019)
2. Erdoğan, U., Lord, G.J.: A new class of exponential integrators for SDEs with multiplicative noise. IMA J. Numer. Anal. **39**, 820–846 (2019)
3. Yang, G., Burrage, K., Ding, X.: A new class of structure-preserving stochastic exponential Runge–Kutta integrators for stochastic differential equations. (Submitted)
4. Komori, Y., Burrage, K.: A stochastic exponential Euler scheme for simulation of stiff biochemical reaction systems. BIT Numer Math. **54**, 1067–1085 (2014)
5. Komori, Y., Cohen, D., Burrage, K.: Weak second order explicit exponential Runge–Kutta methods for stochastic differential equations. SIAM J. Sci. Comput. **39**, A2857–A2878 (2017)
6. Magnus, W.: On the exponential solution of differential equations for a linear operator. Commun. Pure Appl. Math. **7**, 649–673 (1954)

7. Khanamiryan, M.: Modified Magnus expansion in application to highly oscillatory differential equations. *BIT Numer. Math.* **52**, 383–405 (2012)
8. Iserles, A., Nørsett, S.P.: On the solution of linear differential equations in Lie groups. *Phil. Trans. R. Soc.* **357**, 983–1019 (1999)
9. Iserles, A., MacNamara, S.: Applications of Magnus expansions and pseudospectra to Markov processes. *Eur. J. Appl. Math.* **30**, 400–425 (2019)
10. Burrage, K., Burrage, P.M.: High strong order methods for non-commutative stochastic ordinary differential equation systems and the Magnus formula. *Phys. D Nonlinear Phenom.* **133**, 34–48 (1999)
11. Wang, Z., Ma, Q., Yao, Z., Ding, X.: The Magnus expansion for stochastic differential equations. *J. Nonlinear Sci.* **30**, 419–447 (2020)
12. Wang, X., Guan, X., Yin, P.: A new explicit Magnus expansion for nonlinear stochastic differential equations. *Mathematics* **8**, 183 (2020)
13. Kamm, K., Pagliarani, S., Pascucci, A.: On the stochastic Magnus expansion and its application to SPDEs. [arXiv:2001.01098](https://arxiv.org/abs/2001.01098) (2020)
14. Tambue, A., Mukam, J.D.: Magnus-type integrator for non-autonomous SPDEs driven by multiplicative noise. *Discrete Contin. Dyn. Syst. Ser. A.* **40**, 4597–4624 (2020)
15. Blanes, S., Casas, F., Oteo, J.A., Ros, J.: The Magnus expansion and some of its applications. *Phys. Rep. Rev. Sect. Phys. Lett.* **470**, 151–238 (2009)
16. Mao, X.: *Stochastic differential equations and applications*. Horwood, Chichester (2007)
17. Burrage, P.M.: Runge–Kutta methods for stochastic differential equations. Ph.D. Thesis, Dept. Maths., Univ. Queensland (1999)
18. Gard, T.C.: *Introduction to Stochastic Differential Equations*. Marcel Dekker Inc, New York-Basel (1988)
19. Milstein, G.N.: *Numerical integration of stochastic differential equations*. Kluwer Academic Publishers, Dordrecht (1995)
20. Kuznetsov, D.F.: Development and application of the Fourier method for the numerical solution of Itô stochastic differential equations. *Comput. Math. Math. Phys.* **58**, 1058–1070 (2018)
21. Kuznetsov, D.F.: A comparative analysis of efficiency of using the Legendre polynomials and trigonometric functions for the numerical solution of Itô stochastic differential equations. *Comput. Math. Math. Phys.* **59**, 1236–1250 (2019)
22. Wiktorsson, M.: Joint characteristic function and simultaneous simulation of iterated Itô integrals for multiple independent Brownian motions. *Ann. Appl. Probab.* **11**, 470–487 (2001)
23. Kloeden, P.E., Platen, E.: *Numerical solution of stochastic differential equations. Applications of Mathematics: Stochastic Modelling and Applied Probability*. Springer, Berlin (1995)
24. Debrabant, K., Kværnø, A., Mattsson, N.C.: Lawson schemes for highly oscillatory stochastic differential equations and conservation of invariants. [arXiv:1909.12287](https://arxiv.org/abs/1909.12287) (2019)
25. de Bouard, A., Gazeau, M.: A diffusion approximation theorem for a nonlinear PDE with application to random birefringent optical fibers. *Ann. Appl. Probab.* **22**, 2460–2504 (2012)
26. Berg, A., Cohen, D., Dujardin, G.: Exponential integrators for the stochastic Manakov equation. [arXiv:2005.04978v1](https://arxiv.org/abs/2005.04978v1) (2020)
27. Abdulle, A., Cirilli, S.: S-ROCK: Chebyshev methods for stiff stochastic differential equations. *SIAM J. Sci. Comput.* **30**, 997–1014 (2008)
28. Komori, Y., Burrage, K.: Weak second order S-ROCK methods for Stratonovich stochastic differential equations. *J. Comput. Appl. Math.* **236**, 2895–2908 (2012)

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.