# A New Stochastic Magnus Expansion For Linear Stochastic Differential Equations

Pei Yin, Xiaoling Wang

*Abstract*—**Based on the stochastic Magnus expansion, an explicit expression for the solution of the linear stochastic differential equations is proposed in this paper. By use of the Lie bracket and the rooted tree, the stochastic Magnus expansions, which can be used to compute the solutions directly, are analyzed in detail. Moreover, the global rate** $1.0$ **for the mean-square convergence is obtained in the numerical algorithm. Finally, some numerical experiments are given to show the advantages of this numerical algorithm.**

*Index Terms*—**Differential equations, Stochastic Magnus expansions, Lie bracket, Rooted tree, Mean-square convergence.**

## I. Introduction

FOR the matrix differential equation

$$\dot{Y} = A(t)Y, \tag{1}$$

where $A(t)$ is a $n \times n$ matrix. It was shown in [1] (also see [2]–[9]) that the solution of this equation is

$$Y(t) = \exp(\Omega(t))Y_0, \tag{2}$$

$\Omega$ is defined by

$$\dot{\Omega} = d\exp_{\Omega}^{-1}(A(t)),\ \Omega(0) = 0,$$

where

$$d\exp_{\Omega}^{-1}(H) = \sum_{k\geq 0} \frac{B_k}{k!}\mathbf{ad}[\Omega]^k[H].$$

$B_k$ are the Bernoulli numbers, $\mathbf{ad}[\Omega][A] = \Omega A - A\Omega$ is the adjoint operator. $\Omega$ satisfies the differential equation

$$\dot{\Omega} = A(t) - \frac{1}{2}[\Omega, A(t)] + \frac{1}{12}[\Omega.[\Omega, A(t)]] + \dots.$$

By Picard fixed point iteration, we have

$$\begin{aligned}\Omega(t) = &\int_0^t A(k)dk - \frac{1}{2}\int_0^t [\int_0^k A(\xi)d\xi, A(k)]dk \\ &+ \frac{1}{4}\int_0^t [\int_0^k [\int_0^\xi A(\eta)d\eta, A(\xi)]d\xi, A(k)]dk \\ &+ \frac{1}{12}\int_0^t [\int_0^k A(\eta)d\eta, [\int_0^k A(\xi)d\xi, A(k)]]dk + \dots,\end{aligned} \tag{3}$$

which is the so-called Magnus expansions.

If $A(t)$ commutes with $A(s)$, we have

$$\Omega(t) = \int_0^t A(\tau)d\tau. \tag{4}$$

The remainder in equation(3) is of size $O(t^5)$, the truncated series which are inserted into $Y(t) = \exp(\Omega(t))Y_0$ will produce a better approximation to the solution of the equation (1). The application of Magnus expansion to numerical computation of differential equations was firstly proposed by A. Iserles and developed by other authors (see [5]–[11] ). An important advantage of the Magnus expansion is that, even if equation (3) is truncated, it still preserves intrinsic geometric properties of the exact solution. For example, if equation (1) refers to the quantum mechanical evolution operator, the approximate solution obtained by the Magnus expansion is still unitary, no matter where the equation (3) is truncated. More generally, when equation (1) is considered on a Lie group $G$, $\exp(\Omega(t))$ will stay on $G$ for all $t$, provided $A(t)$ belongs to the Lie algebra associated with $G$. In the pioneering work, Iserles and Nørsett translated the advantage of the Magnus expansion into a powerful numerical algorithm. The methods produced better results than the classical numerical schemes in the different examples. The structure-preserving methods for both deterministic and stochastic differential equations have received much more attention in theory and application (see [8], [12]). In recent years, the stochastic differential equations have been widely used in the simulations of random phenomena appearing in physics, engineering, economics etc, (see [12], [13]). Some numerical methods for solving stochastic differential equations have been investigated and developed (see [12], [14], [15]). An interesting application of Magnus expansion to the stochastic case was given in [14], [16]. However, there has not been the general stochastic Magnus expansion. It is important to extend Magnus expansion in deterministic case to the stochastic counterpart. In this paper, we will pay attention to the linear stochastic differential equation in Stratonovich sense as follows

$$dy = a(t)ydt + \sum_{j=1}^{d} b_j(t)y \circ dW_j(t), y(t_0) = y_0, \quad y \in \mathbb{R}^n, \tag{5}$$

where $a(t)$ and $b_j(t)\,(j = 1, 2, \dots, d)$ are continuous matrix functions, and $W_j(t)\ (j = 1, 2, \dots, d)$ are the standard Winner processes. Although, there is a vast literature on the linear differential equation, (see [17], [18]), but lots of them can not be used to compute the solution directly. Therefore, the purpose of this paper is to present the formula of Magnus expansion, which can be immediately used in investigating the numerical solutions of the linear stochastic differential equations. In equation (5), there is no reason to expect

Manuscript received February 4, 2022; revised September 20, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 11301392, Shanghai Peak Discipline Program for Higher Education Institutions (Class I)-Civil Engineering, Fundamental Research Funds for the Central Universities under Grant 22120180529.

P.Yin is a lecturer of the School of Business, University of Shanghai for Science and Technology, Shanghai 200092, China, (e-mail:pyin@usst.edu.cn).

X.L.Wang is a lecturer of the School of Mathematics, Tianjin Chengjian University, Tianjin 300384, China, (corresponding author, e-mail: wangxiaolingcj@126.com) .

that the functions $a(t)$ and $b_j(t)$ associated with the Winner processes commute. This paper is organized as follows. In section 2, the formula of stochastic Magnus expansions is obtained by Lie brackets (see [19]–[21]). In section 3, based on the theory of rooted trees, we investigate the explicit form of stochastic Magnus expansions. In section 4, we truncate the explicit Magnus expansions and give an algorithm with the strong order 1. In the last section, we do some numerical experiments to show the advantages of our numerical algorithm.

## II. The Stochastic Magnus Expansion

In this section , we will use the definition of Lie bracket (see [11]) to present the formula of stochastic Magnus expansion. The binary operation is linear in each component, and subject to the Jacobi identity,

$$[a, b] = -[b, a], \quad \text{for } a, b \in g,$$

$$[a, [b, c]] + [b, [c, a]] + [c, [a, b]] = 0, \quad a, b, c \in g.$$

Equation (5) has the form, when $d = 1$,

$$dy(t) = a(t)y(t)dt + b(t)y(t) \circ dW(t), \quad y(t_0) = y_0, \ y(t) \in \mathbb{R}^n. \tag{6}$$

Suppose that the solution of (6) has the following form

$$y(t) = \exp(\sigma(t))y_0 \tag{7}$$

with

$$d\sigma(t) = \sigma_1(t)dt + \sigma_2(t) \circ dw(t). \tag{8}$$

**Theorem 2.1.** The functions $\sigma_1(t)$ and $\sigma_2(t)$ in equation (8) satisfy the following equations

$$\sum_{m=0}^{\infty} \frac{1}{(m+1)!}\mathbf{ad}[\sigma(t)]^m[\sigma_1(t)] = a(t), \tag{9}$$

$$\sum_{m=0}^{\infty} \frac{1}{(m+1)!}\mathbf{ad}[\sigma(t)]^m[\sigma_2(t)] = b(t), \tag{10}$$

where $\mathbf{ad}[p]^0[q] = q$, $\mathbf{ad}[p]^k[q] = [p, \mathbf{ad}[p]^{k-1}[q]]$, $k \in \mathbb{N}$.

**Proof.** Inserting equation(7) into (6), we have

$$\begin{aligned} dy &= d(\exp(\sigma(t))y_0) \\ &= a(t)\exp(\sigma(t))y_0 dt + b(t)\exp(\sigma(t))y_0 \circ dw(t), \end{aligned} \tag{11}$$

then

$$d\exp(\sigma(t)) = a(t)\exp(\sigma(t))dt + b(t)\exp(\sigma(t)) \circ dw(t). \tag{12}$$

From the fact that

$$\exp(\sigma(t)) = \sum_{k=0}^{\infty} \frac{1}{k!}(\sigma(t))^k, \tag{13}$$

we obtain

$$\begin{aligned} d(\exp(\sigma(t))) &= \sum_{k=1}^{\infty} \frac{1}{k!} d\sigma(t)^k = \sum_{k=1}^{\infty} \frac{1}{k!}(d\sigma(t)\sigma(t)^{k-1} \\ &+ \sigma(t)d\sigma(t)\sigma(t)^{k-2} + \ldots \sigma(t)^{k-1}d\sigma(t)). \end{aligned} \tag{14}$$

Combining equation (14) with equation (8), it implies that

$$\begin{aligned} d\exp(\sigma(t)) &= \sum_{k=1}^{\infty} \tfrac{1}{k!}(\sum_{j=1}^{k} \sigma(t)^{j-1}\sigma_1(t)\sigma(t)^{k-j})dt \\ &+ \sum_{k=1}^{\infty} \tfrac{1}{k!}(\sum_{j=1}^{k} \sigma(t)^{j-1}\sigma_2(t)\sigma(t)^{k-j}) \circ dw(t). \end{aligned} \tag{15}$$

From equation (12) and equation (15), it follows that

$$\sum_{k=1}^{\infty} \frac{1}{k!}(\sum_{j=1}^{k} \sigma(t)^{j-1}\sigma_1(t)\sigma(t)^{k-j}) = a(t)\exp(\sigma(t)), \tag{16}$$

$$\sum_{k=1}^{\infty} \frac{1}{k!}(\sum_{j=1}^{k} \sigma(t)^{j-1}\sigma_2(t)\sigma(t)^{k-j}) = b(t)\exp(\sigma(t)). \tag{17}$$

Therefore

$$\begin{aligned} a(t) &= \sum_{k=1}^{\infty} \frac{1}{k!}(\sum_{j=1}^{k} \sigma(t)^{j-1}\sigma_1(t)\sigma(t)^{k-j})\exp(-\sigma(t)) \\ &= \sum_{k=1}^{\infty} \frac{1}{k!}(\sum_{j=1}^{k} \sigma(t)^{j-1}\sigma_1(t)\sigma(t)^{k-j})[\sum_{l=0}^{\infty} \frac{(-1)^l}{l!}\sigma(t)^l] \\ &= \sum_{l=1}^{\infty} \frac{(-1)^l}{l!} \sum_{j=1}^{l}[\sum_{k=j}^{l}(-1)^k \begin{pmatrix} l \\ k \end{pmatrix}]\sigma(t)^{j-l}\sigma_1(t)\sigma(t)^{l-j}. \end{aligned} \tag{18}$$

It's not difficult to prove that

$$\sum_{k=j}^{l}(-1)^k \begin{pmatrix} l \\ k \end{pmatrix} = (-1)^j \begin{pmatrix} l-1 \\ j-1 \end{pmatrix}, \tag{19}$$

$$\begin{aligned} &\mathbf{ad}[\sigma(t)]^l[\sigma_1(t)] \\ &= \sum_{j=0}^{l}(-1)^{l-j} \begin{pmatrix} l \\ j \end{pmatrix} \sigma(t)^j\sigma_1(t)\sigma(t)^{l-j}, \quad l \in \mathbb{Z}^+. \end{aligned} \tag{20}$$

Based on the above discussion, we get equation (9) and equation (10). The proof is finished.

**Theorem 2.2.** If $\sigma(t)$ satisfies the equation

$$\begin{aligned} d\sigma(t) &= (\sum_{m=0}^{\infty} f_m\mathbf{ad}[\sigma(t)]^m[a(t)])dt \\ &+ \sum_{m=0}^{\infty} f_m\mathbf{ad}[\sigma(t)]^m[b(t)]) \circ dw(t), \end{aligned} \tag{21}$$

$f_m(m = 0, 1, ...)$ are coefficients of the power series

$$f(z) := \sum_{m=0}^{\infty} f_m z^m = \frac{1}{d(z)}$$

$$d(z) := \sum_{l=0}^{\infty} \frac{1}{(l+1)!} z^l,$$

then equation (9) and equation (10) remain true.

**Proof.** From equation (8) and equation (21), it follows that

$$\sigma_1(t) = \sum_{m=0}^{\infty} f_m\mathbf{ad}[\sigma(t)]^m[a(t)], \tag{22}$$

$$\sigma_2(t) = \sum_{m=0}^{\infty} f_m\mathbf{ad}[\sigma(t)]^m[b(t)]. \tag{23}$$

With the definition of $\mathbf{ad}[\ \cdot\ ][\ \cdot\ ]$, it is derived that

$$\mathbf{ad}[p]^{l-m}[\mathbf{ad}[p]^m[q]] = \mathbf{ad}[p]^l[q]. \quad (24)$$

Insteading $\sigma_1(t)$ in equation (9) and (22), we have

$$\begin{aligned}
&\sum_{l=0}^{\infty}\frac{1}{(l+1)!}\mathbf{ad}[\sigma(t)]^l[\sigma_1(t)]\\
&=\sum_{l=0}^{\infty}\frac{1}{(l+1)!}\mathbf{ad}[\sigma(t)]^l[\sum_{m=0}^{\infty}f_m\mathbf{ad}[\sigma(t)]^m[a(t)]]\\
&=\sum_{m=0}^{\infty}\sum_{l=m}^{\infty}f_m d_{l-m}\mathbf{ad}[\sigma(t)]^{l-m}[\mathbf{ad}[\sigma(t)]^m,[a(t)]]\\
&=\sum_{l=0}^{\infty}\sum_{m=0}^{l}f_m d_{l-m}\mathbf{ad}[\sigma(t)]^l[a(t)],
\end{aligned} \quad (25)$$

and

$$\sum_{l=0}^{\infty}\frac{1}{(l+1)!}\mathbf{ad}[\sigma(t)]^l[\sigma_1(t)] = a(t), \quad (26)$$

Similarly, we have

$$\begin{aligned}
&\sum_{l=0}^{\infty}\frac{1}{(l+1)!}\mathbf{ad}[\sigma(t)]^l[\sigma_2(t)]\\
&=\sum_{l=0}^{\infty}\frac{1}{(l+1)!}\mathbf{ad}[\sigma(t)]^l[\sum_{m=0}^{\infty}f_m\mathbf{ad}[\sigma(t)]^m[b(t)]]\\
&=\sum_{m=0}^{\infty}\sum_{l=m}^{\infty}f_m d_{l-m}\mathbf{ad}[\sigma(t)]^{l-m}[\mathbf{ad}[\sigma(t)]^m[b(t)]]\\
&=\sum_{l=0}^{\infty}\sum_{m=0}^{l}f_m d_{l-m}\mathbf{ad}[\sigma(t)]^l[b(t)]
\end{aligned} \quad (27)$$

and

$$\sum_{l=0}^{\infty}\frac{1}{(l+1)!}\mathbf{ad}[\sigma(t)]^l[\sigma_2(t)] = b(t). \quad (28)$$

The proof is finished.

If the coefficients of equation (21) satisfy the conditions of existence and uniqueness theorem for the stochastic differential equation, equation (21) can be written in the integral form

$$\sigma(t) = \sum_{m=0}^{\infty}f_m(\int_0^t \mathbf{ad}[\sigma(s)]^m[a(s)]ds + \mathbf{ad}[\sigma(s)]^m[b(s)]\circ dw(s). \quad (29)$$

Then we can get

$$\begin{aligned}
\sigma(t) &= \int_0^t a(k)dk + \int_0^t b(k)\circ dw(k)\\
&-\frac{1}{2}\int_0^t[\int_0^k a(\xi)d\xi, a(k)]dk\\
&-\frac{1}{2}\int_0^t[\int_0^k a(\xi)d\xi, b(k)]\circ dw(k)\\
&-\frac{1}{2}\int_0^t[\int_0^k b(\xi)dw(\xi), b(k)]\circ dw(k)\\
&-\frac{1}{2}\int_0^t[\int_0^k b(\xi)\circ dw(\xi), a(k)]dk + \dots\quad .
\end{aligned} \quad (30)$$

## III. The Analysis of Expansion via Rooted Trees

In this section, we will focus on the expansion of the general equation (21).

Let $\mathcal{E}$ be the set of all terms' derivatives in the expansion. We propose the following four composition rules, in which $\mathcal{E} := H^1 \bigcup H^2$ is defined recursively,

(1) $a(t) \in H^1,\ b(t) \in H^2$.

(2)

$$\int H_\tau = \begin{cases} \int H_\tau(t)dt, & H_\tau \in H^1,\\ \int H_\tau(t)\circ dw(t), & H_\tau \in H^2. \end{cases}$$

(3) $[\int w_1, w_2] \in \mathcal{E}$, if $w_1(t), w_2(t) \in \mathcal{E}$.

(4) $[\int_0^t w_1(t)dt, w_2(t)]$ belongs to $H^1$, if $w_2(t)$ belongs to $H^1$;

$[\int_0^t w_1(t)dt, w_2(t)]$ belongs to $H^2$, if $w_2(t)$ belongs to $H^2$.

In general, we look for an expansion in the following form

$$\sigma(t) = \sum_{k=1}^{\infty}\sum_{\tau\in\mathcal{T}_k}\alpha_\tau\int_0^t H_\tau^1(k)dk + \sum_{k=1}^{\infty}\sum_{\tau\in\mathcal{T}_k}\beta_\tau\int_0^t H_\tau^2(k)\circ dw(k), \quad (31)$$

where $\mathcal{T}_k$ is the set of all binary trees of order $k$. We can use the composition rules to get terms $H_\tau$. For example, $\mathcal{T}_1 = \{\omega_0^1, \omega_0^2\}, \mathcal{T}_4 = \{\omega_1^1, \omega_1^2, \omega_1^3, \omega_1^4\}$ and $\mathcal{T}_m = \emptyset$, if $m \neq 3n+1$, $n = 0, 1, \dots$ . The coefficients $\alpha_\tau$ and $\beta_\tau$ depend solely on the sequence $\{f_m\}$ $(m = 0, 1, \dots)$ .

From (31), it follows that

$$d\sigma(t) = \sum_{k=1}^{\infty}\sum_{\tau\in\mathcal{T}_k}\alpha_\tau H_\tau^1(t)dt + \sum_{k=1}^{\infty}\sum_{\tau\in\mathcal{T}_k}\beta_\tau H_\tau^2(t)\circ dw(t). \quad (32)$$

Let

$$U_m^1 = \mathbf{ad}[\sum_{k=1}^{\infty}\sum_{\tau\in\mathcal{T}_k}\alpha_\tau\int_0^t H_\tau^1 + \sum_{k=1}^{\infty}\sum_{\tau\in\mathcal{T}_k}\beta_\tau\int_0^t H_\tau^2]^m[a(t)], \quad (33)$$

$$U_m^2 = \mathbf{ad}[\sum_{k=1}^{\infty}\sum_{\tau\in\mathcal{T}_k}\alpha_\tau\int_0^t H_\tau^1 + \sum_{k=1}^{\infty}\sum_{\tau\in\mathcal{T}_k}\beta_\tau\int_0^t H_\tau^2]^m[b(t)], \quad (34)$$

According to (21), we have

$$\sigma_1(t) = \sum_{m=0}^{\infty}f_m U_m^1, \quad (35)$$

$$\sigma_2(t) = \sum_{m=0}^{\infty}f_m U_m^2. \quad (36)$$

Using the definition of $\mathbf{ad}[\cdot]^m[\cdot]$, it is concluded that

$$\begin{aligned}
U_m^1 &= [\sum_{k=1}^{\infty}\sum_{\tau\in\mathcal{T}_k}\alpha_\tau\int_0^t H_\tau^1 + \sum_{k=1}^{\infty}\sum_{\tau\in\mathcal{T}_k}\beta_\tau\int_0^t H_\tau^2,\\
&\quad \mathbf{ad}[\sum_{k=1}^{\infty}\sum_{\tau\in\mathcal{T}_k}\alpha_\tau\int_0^t H_\tau^1 + \sum_{k=1}^{\infty}\sum_{\tau\in\mathcal{T}_k}\beta_\tau\int_0^t H_\tau^2]^{m-1}[a]]\\
&= [\sum_{k=1}^{\infty}\sum_{\tau\in\mathcal{T}_k}\alpha_\tau\int_0^t H_\tau^1 + \sum_{k=1}^{\infty}\sum_{\tau\in\mathcal{T}_k}\beta_\tau\int_0^t H_\tau^2,\ U_{m-1}^1],
\end{aligned} \quad (37)$$

$$U_m^2 = [\sum_{k=1}^{\infty}\sum_{\tau\in\mathcal{T}_k}\alpha_\tau\int_0^t H_\tau^1 + \sum_{k=1}^{\infty}\sum_{\tau\in\mathcal{T}_k}\beta_\tau\int_0^t H_\tau^2,$$
$$\mathbf{ad}[\sum_{k=1}^{\infty}\sum_{\tau\in\mathcal{T}_k}\alpha_\tau\int_0^t H_\tau^1 + \sum_{k=1}^{\infty}\sum_{\tau\in\mathcal{T}_k}\beta_\tau\int_0^t H_\tau^2]^{m-1}[b]]$$
$$= [\sum_{k=1}^{\infty}\sum_{\tau\in\mathcal{T}_k}\alpha_\tau\int_0^t H_\tau^1 + \sum_{k=1}^{\infty}\sum_{\tau\in\mathcal{T}_k}\beta_\tau\int_0^t H_\tau^2,\ U_{m-1}^2]. \tag{38}$$

According to the composition rules, we can get

$$R^1(\tau_1,\tau_2,\ldots,\tau_r) := H_\tau^1 = [\int H_{\tau_1},[\int H_{\tau_2},\ldots,[\int H_{\tau_r},a]]], \tag{39}$$

$$R^2(\tau_1,\tau_2,\ldots,\tau_m) := H_\tau^2 = [\int H_{\tau_1},[\int H_{\tau_2},\ldots,[\int H_{\tau_m},b]]]. \tag{40}$$

**Proposition 3.1.** For any $m\in\mathbb{N}$, it is true that

$$U_m^1 = \sum_{k_1=1}^{\infty}\sum_{k_2=1}^{\infty}\cdots\sum_{k_3=1}^{\infty}\sum_{\tau_1\in\mathcal{T}_1}\sum_{\tau_2\in\mathcal{T}_2}\cdots \sum_{\tau_m\in\mathcal{T}_m} r_{\tau_1}^1 r_{\tau_2}^1\ldots r_{\tau_m}^1 R^1(\tau_1,\tau_2,\ldots,\tau_m), \tag{41}$$

$$r_{\tau_i}^1 = \begin{cases}\alpha_{\tau_i}, & H_{\tau_i}\in H^1,\\ \beta_{\tau_i}, & H_{\tau_i}\in H^2,\end{cases}$$

and

$$U_m^2 = \sum_{k_1=1}^{\infty}\sum_{k_2=1}^{\infty}\cdots\sum_{k_3=1}^{\infty}\sum_{\tau_1\in\mathcal{T}_1}\sum_{\tau_2\in\mathcal{T}_2}\cdots \sum_{\tau_m\in\mathcal{T}_m} r_{\tau_1}^2 r_{\tau_2}^2\ldots r_{\tau_m}^2 R^2(\tau_1,\tau_2,\ldots,\tau_m), \tag{42}$$

$$r_{\tau_i}^2 = \begin{cases}\alpha_{\tau_i}, & H_{\tau_i}\in H^1,\\ \beta_{\tau_i}, & H_{\tau_i}\in H^2.\end{cases}$$

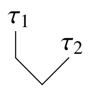
**Proof.** We use the following composition rules for the construction of rooted trees (see [6], [8], [22]). We associate the function $a(t)$ with the trivial tree of order one denoted by a black dot, and b(t) with the trivial tree of order one denoted by a black rectangle,

$$a(t) \rightsquigarrow \bullet, \qquad b(t) \rightsquigarrow \blacksquare.$$

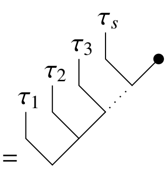
Then

$$\mathcal{T}_0 = \{\bullet,\ \blacksquare\}.$$

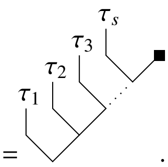
If $\mathcal{T}_k$ is defined for $k = 0,1,\ldots,m-1$, we can get

$\mathcal{T}_m = \Bigg\{$  $:\ \tau_1\in\mathcal{T}_{k_1},\ \tau_2\in\mathcal{T}_{k_2},\ k_1+k_2 = m-1 \Bigg\}.$

It's not difficult to deduce that every binary tree $\tau$ obtained by our composition rules can be uniquely written in the form

$\tau =$ 

or

$\tau =$  .

We will adopt the representations in the sequel.

**Proposition 3.2.** For any $r\in\mathbb{N}$ and $\tau_k\in\mathcal{T}_{m_k}$, $k = 1,2,\ldots,r$, it is true that

$$\mathbf{ord}\mathcal{R}(\tau_1,\tau_2,\ldots,\tau_r) = \sum_{k=1}^{r} m_k + 2r + 1, \tag{43}$$

where

$$\mathbf{ord}\mathcal{R}(\tau_1,\tau_2,\ldots,\tau_r) = \begin{cases}\mathbf{ord}\mathcal{R}^1(\tau_1,\tau_2,\ldots,\tau_r), & \mathcal{R}\in H^1,\\ \mathbf{ord}\mathcal{R}^2(\tau_1,\tau_2,\ldots,\tau_r), & \mathcal{R}\in H^2,\end{cases}$$

$\mathbf{ord}\mathcal{R}$ is the order of the tree $\mathcal{R}$.

**Proof.** we can get it by the induction method and the definition of $\mathcal{R}$.

**Corollary 3.1.** $\mathcal{T}_k = \emptyset$, when $k \neq 1 \bmod 3$, $k\in\mathbb{N}$.

**Proof.** According to equation (36), it yields

$$\sum_{\tau\in\mathcal{T}_k}\alpha_\tau H_\tau^1 = \sum_{l=1}^{\lfloor(k-1)/2\rfloor} f_l \sum_{n_1,n_2,\ldots,n_l\in\mathbb{N}}\ \sum_{\substack{\tau_i\in\mathcal{T}_i\\ i=1,2,\ldots,l}} r_{\tau_1}^1 r_{\tau_2}^1\ldots r_{\tau_l}^1 R^1(\tau_1,\tau_2,\ldots,\tau_l), \tag{44}$$

$$\sum_{\tau\in\mathcal{T}_k}\beta_\tau H_\tau^2 = \sum_{l=1}^{\lfloor(k-1)/2\rfloor} f_l \sum_{n_1,n_2,\ldots,n_l\in\mathbb{N}}\ \sum_{\substack{\tau_i\in\mathcal{T}_i\\ i=1,2,\ldots,l}} r_{\tau_1}^2 r_{\tau_2}^2\ldots r_{\tau_l}^2 R^2(\tau_1,\tau_2,\ldots,\tau_l). \tag{45}$$

Where $n_1+n_2+\ldots+n_l = k-2l-1$. This identity can be simplified in terms of corollary 3.1, since we just need to consider trees of order 1 mod 3.

$$\sum_{\tau\in\mathcal{T}_{3m+1}}\alpha_\tau H_\tau^1 = \sum_{l=1}^{\lfloor(3m)/2\rfloor} f_l \sum_{n_1,n_2,\ldots,n_l\in\mathbb{N}}\ \sum_{\substack{\tau_i\in\mathcal{T}_{3n_i+1}\\ i=1,2,\ldots,l}} r_{\tau_1}^1 r_{\tau_2}^1\ldots r_{\tau_l}^1 R^1(\tau_1,\tau_2,\ldots,\tau_l), \tag{46}$$

$$\sum_{\tau\in\mathcal{T}_{3m+1}}\beta_\tau H_\tau^2 = \sum_{l=1}^{\lfloor(3m)/2\rfloor} f_l \sum_{n_1,n_2,\ldots,n_l\in\mathbb{N}}\ \sum_{\substack{\tau_i\in\mathcal{T}_{3n_i+1}\\ i=1,2,\ldots,l}} r_{\tau_1}^2 r_{\tau_2}^2\ldots r_{\tau_l}^2 R^2(\tau_1,\tau_2,\ldots,\tau_l). \tag{47}$$

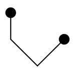
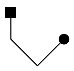
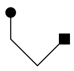
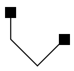
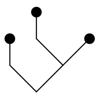
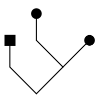
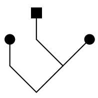
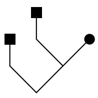
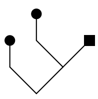
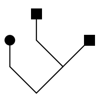
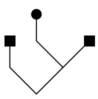
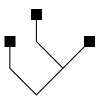
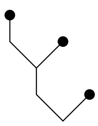
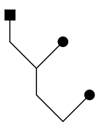
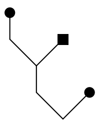
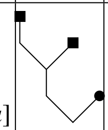
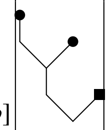
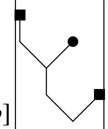
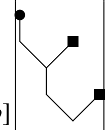
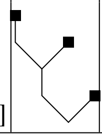
Where $n_1+n_2+\ldots+n_l = m-l$.

Comparing (46) with (47), we conclude that

$$\mathcal{T}_{3m+1} = \{\mathcal{R}(\tau_1,\tau_2,\ldots,\tau_l) : \tau_i\in\mathcal{T}_{3n_i+1},\ i = 1,2,\ldots,l,\ n_1+n_2+\ldots+n_l+l = m,\ l = 1,2,\ldots,\lfloor 3m/2\rfloor\}. \tag{48}$$

According to the equation (46) and (47), the coefficients $\alpha_\tau$ and $\beta_\tau$ can be evaluated. The values of $r^j_{\tau_i}$, $j = 1, 2$, $i = 1, 2, \ldots, l$, are the same as in equation (41) and (42).

TABLE I
THE STOCHASTIC MAGNUS EXPANSIONS OF TERMS $H^i_\tau$, $i = 1, 2$, WHEN THE ORDERS OF TREES $\tau$ ARE $\leq 7$.

| order | name | expression | tree | representation | coefficient |
|---|---|---|---|---|---|
| 1 | $\omega_0^1$ | a |  | - | $f_0$ |
| | $\omega_0^2$ | b |  | - | $f_0$ |
| 4 | $\omega_1^1$ | $[\int a, a]$ |  | $\mathcal{R}^1(\omega_0^1)$ | $f_0 f_1$ |
| | $\omega_1^2$ | $[\int b, a]$ |  | $\mathcal{R}^1(\omega_0^2)$ | $f_0 f_1$ |
| | $\omega_1^3$ | $[\int a, b]$ |  | $\mathcal{R}^2(\omega_0^1)$ | $f_0 f_1$ |
| | $\omega_1^3$ | $[\int b, b]$ |  | $\mathcal{R}^2(\omega_0^2)$ | $f_0 f_1$ |
| 7 | $\omega_2^1$ | $[\int a, [\int a, a]]$ |  | $\mathcal{R}^1(\omega_0^1, \omega_0^1)$ | $f_2 f_0^2$ |
| | $\omega_2^2$ | $[\int b, [\int a, a]]$ |  | $\mathcal{R}^1(\omega_0^2, \omega_0^1)$ | $f_2 f_0^2$ |
| | $\omega_2^3$ | $[\int a, [\int b, a]]$ |  | $\mathcal{R}^1(\omega_0^1, \omega_0^2)$ | $f_2 f_0^2$ |
| | $\omega_2^4$ | $[\int b, [\int b, a]]$ |  | $\mathcal{R}^1(\omega_0^2, \omega_0^2)$ | $f_2 f_0^2$ |
| | $\omega_2^5$ | $[\int a, [\int a, b]]$ |  | $\mathcal{R}^2(\omega_0^1, \omega_0^1)$ | $f_2 f_0^2$ |
| | $\omega_2^6$ | $[\int a, [\int b, b]]$ |  | $\mathcal{R}^2(\omega_0^1, \omega_0^2)$ | $f_2 f_0^2$ |
| | $\omega_2^7$ | $[\int b, [\int a, b]]$ |  | $\mathcal{R}^2(\omega_0^2, \omega_0^1)$ | $f_2 f_0^2$ |
| | $\omega_2^8$ | $[\int b, [\int b, b]]$ |  | $\mathcal{R}^2(\omega_0^2, \omega_0^2)$ | $f_2 f_0^2$ |
| | $\omega_2^9$ | $[\int [\int a, a], a]$ |  | $\mathcal{R}^1(\omega_1^1)$ | $f_1^2 f_0$ |
| | $\omega_2^{10}$ | $[\int [\int b, a], a]$ |  | $\mathcal{R}^1(\omega_1^2)$ | $f_1^2 f_0$ |
| | $\omega_2^{11}$ | $[\int [\int a, b], a]$ |  | $\mathcal{R}^1(\omega_1^3)$ | $f_1^2 f_0$ |
| 7 | $\omega_2^{12}$ | $[\int [\int b, b], a]$ |  | $\mathcal{R}^1(\omega_1^4)$ | $f_1^2 f_0$ |
| | $\omega_2^{13}$ | $[\int [\int a, a], b]$ |  | $\mathcal{R}^2(\omega_1^1)$ | $f_1^2 f_0$ |
| | $\omega_2^{14}$ | $[\int [\int b, a], b]$ |  | $\mathcal{R}^2(\omega_1^2)$ | $f_1^2 f_0$ |
| | $\omega_2^{15}$ | $[\int [\int a, b], b]$ |  | $\mathcal{R}^2(\omega_1^3)$ | $f_1^2 f_0$ |
| | $\omega_2^{16}$ | $[\int [\int b, b], b]$ |  | $\mathcal{R}^2(\omega_1^4)$ | $f_1^2 f_0$ |

$$\alpha_{\omega_0^1} = f_0, \; \beta_{\omega_0^2} = f_0, \tag{49}$$

$$\alpha_{\mathcal{R}^1(\tau_1, \tau_2, \ldots, \tau_l)} = f_l \prod_{i=1}^{l} r^1_{\tau_i}, \tag{50}$$

$$\beta_{\mathcal{R}^2(\tau_1, \tau_2, \ldots, \tau_l)} = f_l \prod_{i=1}^{l} r^2_{\tau_i}, \tag{51}$$

Table 1 displays all expansion terms, trees and coefficients of order seven. Assisted by Table 1, we present the terms of the stochastic Magnus expansion for the equation (21),

$$\begin{aligned} \sigma(t) =& f_0 \int_0^t a(s) ds + f_0 \int_0^t b(s) \circ dw(s) \\ &+ f_0 f_1 \int_0^t [\int_0^s a(k) dk, a(s)] ds \\ &+ f_0 f_1 \int_0^t [\int_0^s a(k) dk, b(s)] \circ dw(s) \\ &+ f_0 f_1 \int_0^t [\int_0^s b(k) dw(k), a(s)] ds \\ &+ f_0 f_1 \int_0^t [\int_0^s b(k) dw(k), b(s)] \circ dw(s) + \ldots . \end{aligned} \tag{52}$$

For the general stochastic differential equation in Stratonovich sense,we can get

$$dy(t) = a(t) y(t) dt + \sum_{j=1}^{d} b_j(t) y(t) \circ dW_j(t), \quad y(t_0) = y_0,$$

where $y(t) = \exp(\sigma(t)) y_0, y(t) \in \mathbb{R}^n$.

Our method can be easily extended to the case of $d > 2$ and we will give the general form of the expansions without proof.

**1**). Let $\mathcal{E}$ be the set of the derivatives of all terms in the expansion .We propose the following four composition rules. $\mathcal{E} := H^0 \bigcup H^1 \ldots \bigcup H^d$ is defined recursively, j=1,2...,d.

(1) $a(t) \in H^0,\ b_i(t) \in H^i,\ i = 1, 2 \ldots, d.$

(2)

$$\int H_\tau = \begin{cases} \int H_\tau(t)dt, & H_\tau \in H^0, \\ \int H_\tau(t) \circ dW_j(t), & H_\tau \in H^j, \end{cases}$$

(3) If $w_1(t), w_2(t) \in \mathcal{E}$, we can get $[\int_0^t w_1(t) \circ dW^l(t),\ w_2(t)] \in \mathcal{E}$,

$$dW^l(t) = \begin{cases} dt, & w_1(t) \in H^0, \\ dW_j(t), & w_1(t) \in H^j, \end{cases}$$

(4) $[\int_0^t w_1(t) \circ dW^l(t), w_2(t)]$ belongs to $H^i$, if $w_2(t)$ belongs to $H^i$, $i = 0, 1, 2, \ldots, d.$

**2)**. Analogously, we can get

$$\begin{aligned} d\sigma(t) &= (\sum_{m=0}^{\infty} f_m \mathbf{ad}[\sigma(t)]^m[a])dt \\ &+ \sum_{j=1}^{d}\sum_{m=0}^{\infty} f_m \mathbf{ad}[\sigma(t)]^m[b_j]) \circ dW_j(t), \end{aligned} \tag{53}$$

$$\sigma(t) = \sum_{i=0}^{d}\sum_{k=1}^{\infty}\sum_{\tau\in\mathcal{T}_k} \alpha_\tau^i \int_0^t H_\tau^i(t) \circ dW^i(t), \tag{54}$$

and

$$\begin{gathered} R^i(\tau_1, \tau_2, \ldots, \tau_r) := H_\tau^i \\ = [\int H_{\tau_1}, [\int H_{\tau_2}, \ldots, [\int H_{\tau_r}, b_i] \ldots]], \\ i = 1, 2 \ldots, d, \quad H_{\tau_i} \in \mathcal{E}, \end{gathered}$$

$$\begin{gathered} R^0(\tau_1, \tau_2, \ldots, \tau_r) := H_\tau^0 \\ = [\int H_{\tau_1}, [\int H_{\tau_2}, \ldots, [\int H_{\tau_r}, a]]], \\ H_{\tau_i} \in \mathcal{E}. \end{gathered}$$

**3)**. $\omega_0 := \{\omega_0^0 := a(t),\ \omega_0^1 := b_1(t), \ldots,\ \omega_0^d := b_d(t)\}$,

$$\alpha_{\omega_0^i}^i = f_0,\ i = 0, 1, \ldots, d,$$

$$\alpha_{\mathcal{R}^m(\tau_1,\tau_2,\ldots,\tau_l)} = f_l \prod_{i=1}^{l} r_{\tau_i}^m,$$

where

$$r_{\tau_i}^m = \begin{cases} \alpha_{\tau_i}^0, & H_{\tau_i} \in H^0, \\ \alpha_{\tau_i}^j, & H_{\tau_i} \in H^j. \end{cases}$$

Finally, we have

$$\begin{aligned} \sigma(t) &= f_0 \int_0^t a(s)ds + \sum_{j=1}^{d} f_0 \int_0^t b_j(s) \circ dW_j(s) \\ &+ f_0 f_1 \sum_{i=0}^{d}\sum_{j=0}^{d} \int_0^t\int_0^s [\omega_0^i, \omega_0^j] \circ dW^i(s) \circ dW^j(t) + \ldots . \end{aligned} \tag{55}$$

**Remark 1.**

Let's consider the stochastic differential equation in Stratonovich sense

$$\begin{aligned} dy(t) &= a(t)y(t)dt + a_1(t)dt + \sum_{j=1}^{d} b_j(t)y(t) \circ dW_j(t) \\ &+ \sum_{m=1}^{q} c_m(t) \circ dW_m(t), \quad y(t_0) = y_0, \quad y \in \mathbb{R}^n, \end{aligned}$$

where $a(t)$ is defined as a matrix or a vector, $b_j(t)\,(j = 1, 2, \ldots, d)$ are matrices, $c_m(t)\,(m = 1, 2, \ldots, q)$ are vectors, and $W_i(t)$ $(i = 1, 2 \ldots)$ are independent Winner processes. By the variation of constants formula, we have

$$\begin{aligned} y(t) &= \exp(\sigma(t))(y_0 + \int_0^t \exp(-\sigma(s))a_1(s)ds \\ &+ \sum_{m=1}^{q} \int_0^t \exp(-\sigma(s))c_m(s) \circ dW_m(s)), \end{aligned} \tag{56}$$

where $\sigma(t)$ is the same as the one in (46).

**Remark 2.**

If the matrices a(t) and b(t) are commutative, we have $\sigma(t) = \int_0^t a(s)ds + \int_0^t b(s) \circ dw(s)$. Then $y(t) = \exp(\sigma(t))y_0$, which coincides with the already-known result.

**Remark 3.**

When $a(t)$ and $b(t)$ are constant matrices and non-commutative, Bernoulli numbers are denoted by $f_m$ $(m = 0, 1, 2, \ldots)$. We have $f_0 = 1, f_1 = -\frac{1}{2}, f_2 = \frac{1}{12}$. Substituting $f_m$ into (51), we get

$$\begin{aligned} \sigma(t) &= \int_0^t a(s)ds + \int_0^t b(s) \circ dw(s) - \tfrac{1}{2}\int_0^t\int_0^k [a, b]ds \circ dw(k) \\ &- \tfrac{1}{2}\int_0^t\int_0^k [b, a] \circ dw(s)dk + \tfrac{1}{4}\int_0^t\int_0^k\int_0^\xi [[a, b], a]d(\eta) \circ dw(\xi)dk \\ &+ \tfrac{1}{4}\int_0^t\int_0^k\int_0^\xi [[a, b], b]d(\eta) \circ dw(\xi) \circ dw(k) \\ &+ \tfrac{1}{4}\int_0^t\int_0^k\int_0^\xi [[b, a], a] \circ dw(\eta)d(\xi)dk \\ &+ \tfrac{1}{4}\int_0^t\int_0^k\int_0^\xi [[b, a], b]dw(\eta)d(\xi) \circ dw(k) \\ &+ \tfrac{1}{12}\int_0^t\int_0^k\int_0^k [a, [b, a]]d(\eta) \circ dw(\xi)dk \\ &+ \tfrac{1}{12}\int_0^t\int_0^k\int_0^k [a, [a, b]]d(\eta)d(\xi) \circ dw(k) \\ &+ \tfrac{1}{12}\int_0^t\int_0^k\int_0^k [b, [b, a]] \circ dw(\eta) \circ dw(\xi)dk \\ &+ \tfrac{1}{12}\int_0^t\int_0^k\int_0^k [b, [a, b]] \circ dw(\eta)d(\xi) \circ dw(k) + \ldots . \end{aligned} \tag{57}$$

By the properties of the Lie bracket and multi-Stratonovich integral, we finally get

$$\begin{aligned} \sigma(t) =& at + bJ_{1,t} - \frac{1}{2}[a, b](J_{01,t} - J_{10,t}) \\ &+ [a, [b, a]](\frac{1}{2}J_{010,t} - \frac{1}{6}J_{1,t}J_{00,t}) \\ &+ [b, [a, b]](\frac{1}{2}J_{101,t} - \frac{1}{6}J_{0,t}J_{11,t}) + \ldots . \end{aligned} \tag{58}$$

## IV. Numerical Algorithms Based on The Linear Stochastic Magnus Expansion

In this section, we will investigate the numerical solution of the stochastic differential equation with the following form

$$dy = A(t)ydt + B(t)y \circ dW(t). \tag{59}$$

In order to construct efficient numerical algorithms based on the stochastic Magnus expansions, the multiple stochastic integrals should be easily computed. We will give several schemes with different strong orders (mean-square sense). And in all cases, we will choose the quadrature rules with equispaced points over the interval $[t_n, t_{n+1}]$. Our numerical algorithm errors contain two parts. One is the truncated error, the other is the error caused by the numerical schemes.

The $\sigma(t)$ is truncated in the following way

$$\sigma^m(t_n+h) = \sum_{i=0}^{1}\sum_{k=1}^{m}\sum_{\tau\in\mathcal{T}_{3k+1}} \alpha_\tau^i \int_{t_n}^{t_n+h} H_\tau^i(t)\circ dW^i(t)+I_{m+1}^J, m \geqslant 2. \tag{60}$$

$J = (1, 1, \cdots, 1)$ is an index of length $m + 1$,

$$I_{m+1}^J = \underbrace{\int_{t_n}^{t_n+h}\int_{t_n}^{s_1}\cdots\int_{t_n}^{s_{m+1}}}_{m+1}[B[B\cdots,[B,B]\cdots]] \tag{61}$$
$$\circ\, dW(s_{m+1})\circ dW(s_m)\cdots dW(s_2)\circ dW(s_1),$$

and

$$dW^i = \begin{cases} dt, & i=0,\\ dW(t), & i=1.\end{cases}$$

As discussed in Section 3, the order(the number of vertices) $k$ of the tree $\mathcal{T}_k$ is of the form $k = 3m+1, m = 0, 1, \cdots$, and the number $n(\mathcal{T}_k)$ of the tree with the order $k$ satisfies

$$n(\mathcal{T}_k) = \begin{cases} 2, & m=0,\\ 2^{2m}, & m>0.\end{cases}$$

Based on the multiple Stratonovich integrals and the multiple it$\hat{o}$ integrals, we can get the following lemma easily.

**Lemma 4.1.** If $A(t)$ and $B(t)$ are uniformly bounded in the finite intervals, the multiple Stratonovich integrals of the form

$$I_m^J = \underbrace{\int_{t_n}^{t_n+h}\int_{t_n}^{s_1}\cdots\int_{t_n}^{s_m}}_{m}[C^{j_1}[C^{j_2}\cdots,[C^{j_{m-1}},C^{j_m}]\cdots]] \tag{62}$$
$$\circ dW^{j_m}(s_m)\circ dW^{j_{m-1}}(s_{m-1})\cdots dW^{j_2}(s_2)\circ dW^{j_1}(s_1),$$

where

$$J = (j_1, j_2, \cdots, j_{m-1}, j_m), \tag{63}$$

$$C^{j_i} = \begin{cases} A, & j_i=0,\\ B, & j_i=1,\end{cases} \tag{64}$$

and $j_i \in \{0, 1\}$( $i = 1, 2, \cdots m$), satisfy

$$(\mathbb{E}|I_m^J|^2)^{1/2} \leqslant (8M^2)^{m/2}h^{\mathcal{L}(J)+\mathcal{N}(J)/2} \leqslant (8M^2)^{m/2}h^{m/2}. \tag{65}$$

$M$ is the bound of $A$ and $B$. $\mathcal{L}(J)$ is the number of components of $J$, and $\mathcal{L}(J) = 0$. $\mathcal{N}(J)$ is the number of components of $J$, and $\mathcal{N}(J) = 1$.

**Proof**:

Since $A$ and $B$ are uniformly bounded, we have $|A| \leqslant M$ and $|B| \leqslant M$. We will use the induction to prove this lemma.

When $m = 1$,

$$\mathbb{E}(|\int_{t_n}^{t_n+h} B(s)\circ dW(s)|)^2 = \mathbb{E}(|\int_{t_n}^{t_n+h} B(s)dW(s)|)^2$$
$$= \int_{t_n}^{t_n+h}\mathbb{E}(|B^2(s)|)dt \leqslant M^2h \leqslant 8M^2h,$$

According to the Cauchy-Schurz inequality, we can get

$$\mathbb{E}(|\int_{t_n}^{t_n+h} A(s)ds|)^2 \leqslant \int_{t_n}^{t_n+h} ds\mathbb{E}(\int_{t_n}^{t_n+h}|A(s)|^2ds$$
$$\leqslant M^2h^2 \leqslant 8M^2h^2.$$

For any $m \leqslant n$, the equation (65) holds, if we need to prove equation (65) holds for $m = n + 1$ and $J = (j_1, j_2, \cdots, j_n, j_{n+1})$.

Based on the Stratonovich integral and *Itô*-integral, we have

1) If $j_1 = 1$

$$I_{n+1}^J = \underbrace{\int_{t_n}^{t_n+h}\int_{t_n}^{s_1}\cdots\int_{t_n}^{s_{n+1}}}_{n+1}[B[C^{j_2}\cdots,[C^{j_n},C^{j_{n+1}}]\cdots]]$$
$$\circ\, dW^{j_{n+1}}(s_{n+1})\circ\cdots dW^{j_2}(s_2)\circ dW(s_1)$$
$$= \overbrace{\underbrace{\int_{t_n}^{t_n+h}\int_{t_n}^{s_1}\cdots\int_{t_n}^{s_{n+1}}}_{n+1}[B[C^{j_2}\cdots,[C^{j_n},C^{j_{n+1}}]\cdots]] \circ dW^{j_{n+1}}(s_{n+1})\circ\cdots\circ dW^{j_2}(s_2)dW(s_1)}^{I_{n+1}^J(1)}$$
$$+\frac{1}{2}\delta_{j_2j_1}\overbrace{\underbrace{\int_{t_n}^{t_n+h}\int_{t_n}^{s_1}\cdots\int_{t_n}^{s_n}}_{n}[B[C^{j_2}\cdots,[C^{j_n},C^{j_{n+1}}]\cdots]] \circ dW^{j_{n+1}}(s_n)\circ\cdots\circ dW^{j_3}(s_2)ds_1}^{I_{n+1}^J(2)}$$

where

$$\delta_{ij} = \begin{cases} 1, & i=j,\\ 0, & others.\end{cases}$$

Therefore, we get

$$\mathbb{E}|I_{n+1}^J(1)|^2$$
$$= \int_{t_n}^{t_n+h}\mathbb{E}(|\int_{t_n}^{s_1}\cdots\int_{t_n}^{s_{n+1}}[B[C^{j_2}\cdots,[C^{j_n},C^{j_{n+1}}]\cdots]]$$
$$\circ\, dW^{j_{n+1}}(s_{n+1})\circ\cdots\circ d^{j_2}W(s_2)|)^2ds_1$$
$$\leqslant 4M^2\int_{t_n}^{t_n+h}((8M^2)^n(s_1-t_n)^{2\mathcal{L}+\mathcal{N}})ds_1$$
$$= K_1h^{p_1} = \frac{(8M^2)^{n+1}}{2\mathcal{L}(J)+\mathcal{N}(J)}h^{2\mathcal{L}J+\mathcal{N}(J)}$$
$$\leqslant (8M^2)^{n+1}/(n+1)h^{2\mathcal{L}J+\mathcal{N}(J)},$$

where

$$p_1 = 2\mathcal{L}(j_2, j_3, \cdots, j_{n+1}) + \mathcal{N}(j_2, j_3, \cdots, j_{n+1}) + 1,$$

$$K_1 = \frac{(8M^2)^{n+1}}{2\mathcal{L}(j_2, j_3, \cdots, j_{n+1}) + \mathcal{N}(j_2, j_3, \cdots, j_{n+1}) + 1}.$$

Using the Cauchy-Schwarz inequality, we get

$$\begin{aligned}&\mathbb{E}|I_{n+1}^J(2)|^2\\ &\leqslant 4M^2h\int_{t_n}^{t_n+h}((8M^2)^n(s_1-t_n)^{2\mathcal{L}+\mathcal{N}})ds_1\\ &= K_2h^{p_2} \leqslant (8M^2)^{n+1}\frac{1}{n}h^{2\mathcal{L}(J)+\mathcal{N}(J)},\end{aligned} \tag{66}$$

where

$$p_2 = 2\mathcal{L}(j_2, j_3, \cdots, j_{n+1}) + \mathcal{N}(j_2, j_3, \cdots, j_{n+1}) + 2,$$

$$K_2 = \frac{(8M^2)^{n+1}}{2\mathcal{L}(j_2, j_3, \cdots, j_{n+1}) + \mathcal{N}(j_2, j_3, \cdots, j_{n+1}) + 1}.$$

Finally, we get

$$
\begin{aligned}
&(\mathbb{E}|I_{n+1}|^2)^{1/2} \leqslant (\mathbb{E}|I_{n+1}^1|^2)^{1/2} + \delta_{j_2j_1}(\mathbb{E}|I_{n+1}^2|^2)^{1/2} \\
&\leqslant \frac{(8M^2)^{n+1}}{n+1}h^{\mathcal{L}(J)+\mathcal{N}(J)/2} + \delta_{j_2j_1}(8M^2)^{n+1}\frac{1}{n}h^{\mathcal{L}(J)+\mathcal{N}(J)/2} \\
&\leqslant (8M^2)^{\frac{n+1}{2}}h^{\mathcal{L}(J)+\mathcal{N}(J)/2} \leqslant (8M^2)^{(n+1)/2}h^{(n+1)/2},
\end{aligned} \tag{67}
$$

2) If $j_1 = 0$,

$$
\begin{aligned}
I_{n+1}^J = &\underbrace{\int_{t_n}^{t_n+h}\int_{t_n}^{s_1}\cdots\int_{t_n}^{s_{n+1}}}_{n+1}[A[C^{j_2}\cdots,[C^{j_n},C^{j_{n+1}}]\cdots]] \\
&\circ dW^{j_{n+1}}(s_{n+1})\circ\cdots dW^{j_2}(s_2)ds_1,
\end{aligned} \tag{68}
$$

According to the Cauchy-Schwarz inequality, we have

$$
\begin{aligned}
&\mathbb{E}(|I_{n+1}^J|)^2 \\
&\leqslant \int_{t_n}^{t_n+h}ds\int_{t_n}^{t_n+h}\mathbb{E}(|\int_{t_n}^{s_1}\cdots\int_{t_n}^{s_{n+1}}[A[C^{j_2}\cdots,[C^{j_n}, \\
&C^{j_{n+1}}]]]\circ dW^{j_{n+1}}(s_{n+1})\cdots\circ d^{j_2}W(s_2)|)^2ds_1 \\
&\leqslant 4M^2h\int_{t_n}^{t_n+h}(8M^2)(n)(s_1-t_n)^{2\mathcal{L}+\mathcal{N}}ds_1 \\
&\leqslant K_3h^{p_3} \\
&\leqslant (8M^2)^{n+1}/(n+1)h^{2\mathcal{L}(J)+\mathcal{N}(J)},
\end{aligned} \tag{69}
$$

where

$p_3 = 2\mathcal{L}(j_2, j_3, \cdots, j_{n+1}) + \mathcal{N}(j_2, j_3, \cdots, j_{n+1}) + 2,$

$$K_3 = \frac{(8M^2)^{n+1}}{(2\mathcal{L}(j_2, j_3, \cdots, j_{n+1}) + \mathcal{N}(j_2, j_3, \cdots, j_{n+1}) + 1)}.$$

The proof is finished.

**Lemma 4.2.** Let $\sigma(t_n + h)$ be the exact solution of the equation (59) and $\sigma^m(t_n + h)$ be the truncated solution given by equation (60), it is true that

$$(\mathbb{E}|\sigma(t_n + h) - \sigma^m(t_n + h)|^2)^{1/2} = O(h^{(m+1)/2}). \tag{70}$$

**Proof**

$$\sigma(t_n+h)-\sigma^m(t_n+h) = \sum_{i=0}^{1}\sum_{k=m+1}^{\infty}\sum_{\tau\in\mathcal{T}_{3k+1}}\alpha_\tau^i\int_{t_n}^{t_n+h}H_\tau^i(t)\circ dW^i(t), \tag{71}$$

According to Cauchy-Schwarz inequality, we get

$$
\begin{aligned}
&(\mathbb{E}|\sigma(t_n + h) - \sigma^m(t_n + h)|^2)^{1/2} \\
&\leqslant \sum_{i=0}^{1}\sum_{k=m+1}^{\infty}\sum_{\tau\in\mathcal{T}_{3k+1}}|\alpha_\tau^i|(\mathbb{E}|\int_{t_n}^{t_n+h}H_\tau^i(t)\circ dW^i(t)|^2)^{1/2} \\
&\leqslant \sum_{k=m+1}^{\infty}2^{2k}(8M^2h)^{k/2} = \frac{(8\sqrt{2}Mh^{1/2})^{m+1}}{1-8\sqrt{2}Mh^{1/2}}.
\end{aligned} \tag{72}
$$

The second inequality is obtained by the lemma 4.1 and $|\alpha_\tau^i| \leqslant 1$. We complete the proof.

**Theorem 4.1.** Let $exp(\sigma(t_n + h))$ be the exact solution of equation (59) and $exp(\sigma^m(t_n + h))$ be the truncated solution of the system (59), we have

$$|\mathbb{E}(exp(\sigma(t_n + h)) - exp(\sigma^m(t_n + h)))| \leqslant O(h^{m/2+1}), \tag{73}$$

$$(\mathbb{E}|exp(\sigma(t_n+h)) - exp(\sigma^m(t_n+h)))|^2)^{1/2} \leqslant O(h^{m/2+1/2}). \tag{74}$$

**Proof**:

$exp(\sigma(t_n + h)) = exp(\sigma^m(t_n + h) + \bar{\delta})$

where

$$\bar{\delta} = \sum_{i=0}^{1}\sum_{k=m+1}^{\infty}\sum_{\tau\in\mathcal{T}_{3k+1}}\alpha_\tau^i\int_{t_n}^{t_n+h}H_\tau^i(t)\circ dW^i(t),$$

and $\bar{\delta}$ does not include the term (61).

The local remainder of the truncated Magnus expansion is

$$
\begin{aligned}
\mathcal{S}^{error} &= exp(\sigma(t_n + h)) - exp(\sigma^m(t_n + h)) \\
&= exp(\sigma^m(t_n + h) + \delta) - exp(\sigma^m(t_n + h)) \\
&= \delta + \frac{1}{2}(\sigma^m\delta + \delta\sigma^m) + O((\sigma^m)^2\delta).
\end{aligned}
$$

Combining Lemma 4.1with Lemma 4.2, we can prove the theorem easily.

In Theorem 4.1, the truncated error is $m/2$. If the numerical scheme which has the mean-square order at least $m/2$ is adopted, we will get the algorithm with strong order $m/2$.

**Remark**

- We can use any numerical algorithm with order $m/2$ to approximate the solution of formula, if we can exactly integrate the equation (60).
- To get a m-order algorithm, there is no need to include all the trees with order less than $2m$. We just include the trees which have strong order less than $m$.
- When $A$ and $B$ are constant matrices, we can reduce the explicit expansion to a more simple structure.

Next, we will give an algorithm of strong order 1.

In this case, $\sigma(t)$ corresponds to the following form

$$\sigma^1(t_n + h) = \int_{t_n}^{t_n+h}A(s)ds + \int_{t_n}^{t_n+h}B(s)\circ dW(s). \tag{75}$$

Using the Euler numerical algorithm, we can get

$$\bar{\sigma}^1(t_n + h) = A(t_n)h + B(t_n)\Delta W_n, \tag{76}$$

According to the Taylor expansion, we can get the stochastic expansion for the solution of the equation (59).

$$
\begin{aligned}
y(t_n + h) = &\, y(t_n) + A(t_n)y(t_n)h + B(t_n)y(t_n)\Delta W_n \\
&+ B(t_n)^2y(t_n)\Delta W_n^2/2 + O(h^{1.5}),
\end{aligned} \tag{77}
$$

The approximate solution can be expanded in the following form

$$
\begin{aligned}
\bar{y}(t_n + h) &= exp(\bar{\sigma}^1(t_n + h))y_{t_n} = exp(A(t_n)h + B(t_n)\Delta W_n) \\
&= y(t_n) + A(t_n)y(t_n)h + B(t_n)y(t_n)\Delta W_n + (A(t_n)h \\
&+ B(t_n)\Delta W_n)^2/2y(t_n) + O(h^{1.5}),
\end{aligned} \tag{78}
$$

It is easy to get

$$
\begin{aligned}
|\mathbb{E}y(t_n + h) - \bar{y}(t_n + h)| &= O(h^2), \\
(\mathbb{E}|y(t_n + h) - \bar{y}(t_n + h)|^2)^{1/2} &= O(h^{1.5}).
\end{aligned}
$$

According to the Milstein mean-square convergence theorem, the strong convergence order is 1.

## V. Numerical Experiment

### *A. The Lyapunov exponent of a linear system with small noise*

The negativeness of upper Lyapunov exponent is an important indication for the system stability. It is usually difficult to derive analytical formulas for Lyapunov exponents. An algorithm for the computation of Lyapunov exponent was proposed by Talay (see [23]) for the first time, and the algorithm was based on the weak schemes.

To test the numerical algorithm, we proposed in Section 4, we will use the two-dimensional linear *itô* system

$$dX = AXdt + \varepsilon BXdW(t). \tag{79}$$

$X$ is a two-dimensional vector, $A$ and $B$ are constant $2\times 2$ matrices, $W$ is the standard wiener process, and $\varepsilon > 0$.

In the ergodic case, there exists a unique Lyapunov exponent $\lambda$ of system (79),

$$\lambda = \lim_{t->\infty} \frac{1}{t}\mathbb{E}\rho(t) = \lim_{t->\infty} \frac{1}{t}\rho(t) a.s. \tag{80}$$

where $\rho(t) = ln|X(t)|$, $X(t)$ $(t \geqslant 0)$ is the non-trivial solution of the system (79).

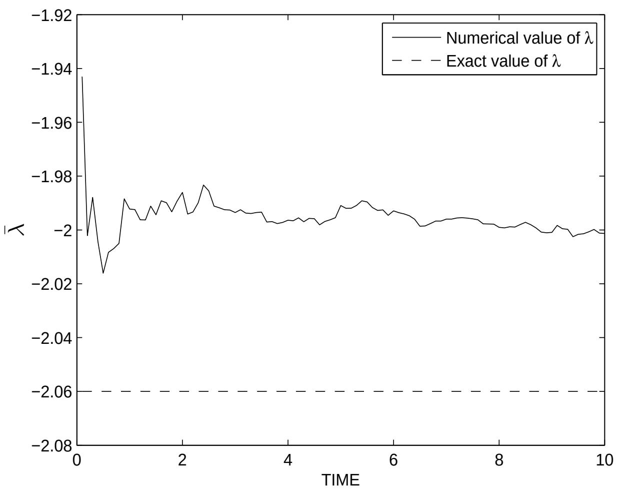



Fig. 1. The exact value and numerical value of the Lyapunov exponent over the time interval [0,10].

Herein, we consider system (79) with the matrices $A$ and $B$,

$$A = \begin{pmatrix} a & c \\ -c & a \end{pmatrix}, B = \begin{pmatrix} b & d \\ -d & b \end{pmatrix}, \tag{81}$$

the Lyapunov exponent is given by (3.4)(see [24]).

In order to study the system(79) conveniently, we transform it into its Stratonovich form

$$dX = (A - \frac{1}{2}B^2)Xdt + \varepsilon BX \circ dW(t). \tag{82}$$

We simulate the system (82) with the algorithm (76). The one-step algorithm is

$$Y_{n+1} = exp((A - \frac{1}{2}B^2)\Delta t_n + B\Delta W_n)Y_n, \tag{83}$$

In the following, we simulate $\lambda(T)$ with $a = -0.5$, $c = -3$, $b = 2$, $d = -1$, $\varepsilon = 0.2$, $X^1(0) = 0$, $X^2(0) = 1$. The

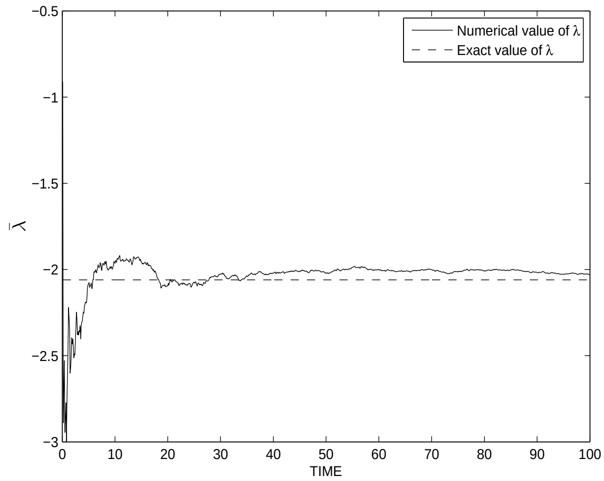



Fig. 2. The exact value and numerical value of the Lyapunov exponent over the time interval [0,100].

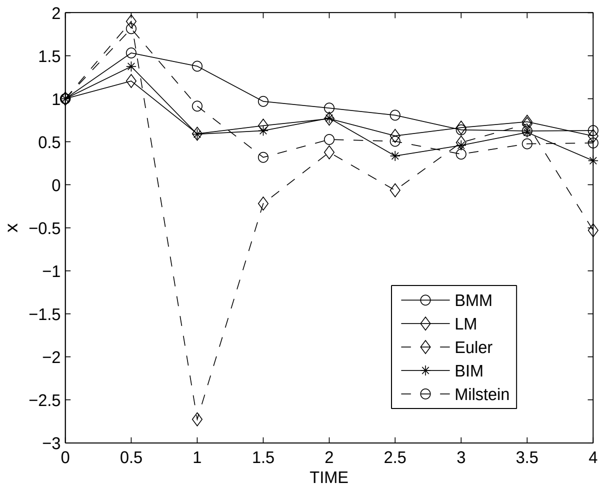



Fig. 3. The approximate solutions of the Magnus method, Balanced Milstein method, Euler method, Milstein method and Balanced Implicit method.

simulation results are shown in Figure 1 and Figure 2. In Figure 1, the time interval is [0,10], the stepsize of numerical approximation is denoted by $dT$, and $dT = 0.1$. The time-dependent function is $\bar{\lambda}(T) = \mathbb{E}\bar{\rho}(T)/T$. The number of realizations is denoted by $M$, and $M = 400$. In Figure 2, the time interval is [0,100]. The time-dependent function is $\bar{\rho}(T)/T$ which is computed along a single trajectory. The dashed line shows the exact value of the Lyapunov exponent, and $\lambda = -2.06$.

From Figure 1 and 2, we can see that our algorithm is strong convergent, and it is efficient to compute the Lyapunov exponent.

### *B. A linear system with nonnegative solution*

To test the numerical algorithm proposed in Section 4, we consider the other linear system

$$dX(t) = [a + bX(t)]dt + \sigma X(t)dw(t), \tag{84}$$

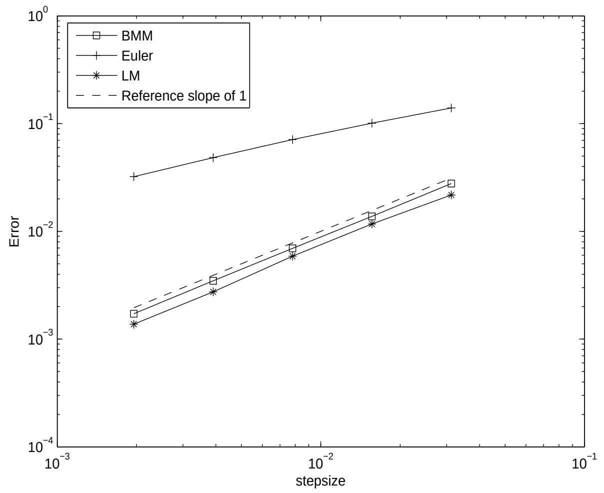



Fig. 4. $\mathcal{L}^2$-Error of the solutions for the Magnus method, Balanced Milstein method, Euler method, and Reference slope.

where $a \geqq 0, b \in R$ and $\sigma \geq 0$.

The strong solutions of this equation are nonnegative. To solve this problem, we split the original system into two subsystems

$$dX_1(t) = [b - \frac{1}{2}\sigma^2]X_1(t)dt + \sigma X(t) \circ dw(t), \quad (85)$$

and

$$dX_2(t) = adt. \quad (86)$$

For system (85), we use the magnus algorithm (76) to approximate the solution. For system (86), the solution can be computed exactly. Finally, we use the composite method to get one-step numerical scheme of the form

$$\bar{X}_{t_n+h} = exp(ah)\bar{X}_1(t_n + h). \quad (87)$$

where, $\bar{X}_1(t_n + h)$ is the approximate solution of the system (85).

In Figure 3, the time interval is [0,4], the stepsize of numerical approximations is $dT = 0.5$, and $a = 1, b = -1, \sigma = 1.4$. We compare our method with the Balanced Milstein method(see [25]), the Balanced implicit method(see [26]), the Euler method and the Milstein method. As one would expect, our method can preserve the positivity exactly and perform as well as the Balanced Milstein method.

Figure 4 shows that our method is of strong order 1 which verifies the theoretical analysis in Section 4. It can be seen from the image, the errors obtained by our method are less than other methods.

Table 2 shows the percentages of the negative paths in the simulation. Weight functions of BMM are defined by $d^0(x) = 1+0.5*1.4^2$, and $d^1(x) = 0$, the number of simulated paths is 1500. We can see that both Euler and Milstein methods have a certain percentage of negative paths. When the stepsize decreases, the number of negative paths also decreases. Our method is as good as the Balanced Milstein method in preserving the positivity of the solution. And the method is independent of the time interval and the stepsize of time.

TABLE II
The percentage of the negative paths for $dX(t) = (1 - X(t))dt + 1.4X(t)dW(t)$.

| Time | Stepsize | Euler | Milstein | BMM | LM |
|---|---|---|---|---|---|
| T=1 | dT=1/2 | 27.35% | 22.12% | 0% | 0% |
| | dT=1/4 | 26.35% | 8.21% | 0% | 0% |
| | dT=1/16 | 17.35% | 0.12% | 0% | 0% |
| T=4 | dT=1/2 | 69.35% | 53.45% | 0% | 0% |
| | dT=1/4 | 66.24% | 18.48% | 0% | 0% |
| | dT=1/16 | 57.89% | 2.45% | 0% | 0% |
| T=16 | dT=1/2 | 98.67% | 94.25% | 0% | 0% |
| | dT=1/4 | 96.56% | 58.48% | 0% | 0% |
| | dT=1/16 | 95.72% | 9.08% | 0% | 0% |

## VI. Conclusions

Based on the stochastic Magnus expansions, an explicit expression for the solutions of linear stochastic differential equations is proposed. In our numerical algorithm, the formula of Magnus expansion can be used in investigating the numerical solution immediately. Some numerical experiments are given to show the advantages of this numerical algorithm. At the same time, we also show that our method is efficient for preserving positivity of the models.

## References


[1] W. Magnus, "On the exponential solution of differential equations for a linear operator", *Communications on Pure and Applied Mathematics,* vol.7, no.2, pp 649-673, 1954.

[2] S. O. Edeki, I. Adinya, O. O. Ugbebor, "The Effect of Stochastic Capital Reserve on Actuarial Risk Analysis via an Integro-differential Equation," *IAENG International Journal of Applied Mathematics,* vol.44, no.2, pp 83-90, 2014.

[3] J. M. Yoon, S. Xie , V. Hrynkiv, "A Series Solution to a Partial Integro-Differential Equation Arising in Viscoelasticity", *IAENG International Journal of Applied Mathematics,* vol.43, no.4, pp 172-175, 2013.

[4] X. Huang, X. Liu, "Backward Stochastic Differential Equation with Monotone and Continuous Coefficient", *IAENG International Journal of Applied Mathematics,* vol.39, no.4, pp 231-235, 2009.

[5] S. Blanes, F. Casas, J. A. Oteo, J. Ros, "Magnus and Fer expansions for matrix differential equations:the convergence problem", *Journal of Physics A General Physics*, vol.31, no.1, pp 259-268, 1999.

[6] S. Blanes, F. Casas, J. A. Oteo, J. Ros, "Magnus expansion: mathematical study and physical applications", *Physics Reports* , vol.470, pp 151-238, 2009.

[7] S. Blanes, F. Casas, J. Ros, "Improved high order geometric integrators based on the Magnus expansion", *Bit Numerical Mathematics*, vol.40, no.3, pp 434-450, 2000.

[8] E. Hairer, C. Lubich, G. Wanner, "Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations", 2nd ed. Berlin : Springer, 2006.

[9] A. Iserles, "Solving linear ordinary differential equations by exponentials of iterated commutators", *Numerische Mathematik* vol.45, no.2, pp 183-199, 1984.

[10] A. Iserles, H. Z. Munthe-Kaas, S. P. Nørsett, A. Zanna, Lie group methods, *Acta Numerica*, vol.9, pp 215-365, 2000.

[11] A. Iserles, S. P. Nørsett, "On the solution of linear differential equation in Lie groups", *Philosophical Transactions Mathematical Physical and Engineering Sciences*, vol.357, no.1754 , pp 983-1019, 1997.

[12] G. N. Milstein, M. V. Tretyakov, "Solving the dirichlet problem for navier–stokes equations by probabilistic approach", *Bit Numerical Mathematics* , vol. 52 , no.1, pp 141-153, 2012.

[13] X. Li, Y. Wu, Q. Zhu, S. Hu, C. Qin, "A regression-based monte carlo method to solve two-dimensional forward backward stochastic differential equations", *Advances in Difference Equations*, vol.207, pp 1-13, 2021.

[14] K. Burrage, P. M. Burrage, "High strong order methods for non-commutative stochastic ordinary differential equation systems and the Magnus formula", *Physica D Nonlinear Phenomena*, vol.133, no.1, pp 34-48, 1999.

[15] S. Sheikhi, M. Matinfar, M. A. Firoozjaee, "Numerical Solution of Variable-Order Differential Equations via the Ritz-Approximation Method by Shifted Legendre Polynomials", *International Journal of Applied and Computational Mathematics*, vol.7, no.1, pp 259-268, 2021.

[16] G. Lord, S. J. A. Malham, A. Wiese, “Efficient integrators for linear stochastic systems”, *Siam Journal on Numerical Analysis*, vol.46, no.6, pp 2892-2919, 2008.

[17] F. Baudoin, *An introduction to the Geometry of stochastic Flows*, Imperial College Press, 2005.

[18] F. Castell, “Asymptotic expansion of stochastic flows”, *Probability Theory and Related Fields*, vol.96, no.2, pp 225-239, 1993.

[19] G. Yang, K. Burrage, Y. Komori, P. Burrage, X. Ding, “A class of new magnus-type methods for semi-linear non-commutative it stochastic differential equations”, *Numerical Algorithms*, vol.88, pp 1641-1665, 2021.

[20] H. Kunita, *Stochastic Flows and Stochastic Differential Equations, Cambridge Studies in Advanced Mathematics*, Cambridge: Cambridge University Press, 1990.

[21] T. Misawa, “A Lie algebraic approach to numerical integration of stochastic differential equations”, *SIAM Journal on Scientific Computing*, vol.23, no.3, pp 866-890, 2001.

[22] H. Hassani, J. Machado, M. S. Dahaghin, Z. Avazzadeh, “Relation between new rooted trees and derivatives of differential equations”, *Iranian Journal of Science and Technology, Transactions A: Science*, vol.45, no.3, pp 1025-1036, 2021.

[23] D. Talay, ”Approximation of upper Lyapunov exponents of bilinear stochastic differential equations”, *SIAM Journal on Numerical Analysis*, vol.28, no.4, pp 1141-1164, 1991.

[24] E. I. Auslender, G. N. Milstein, “Asymptotic expansions of the Lyapunov index for linear stochastic systems with small noises”, *Journal of Applied Mathematics and Mechanics*, vol.46, no.3, pp 358-365, 1982.

[25] C. Kahl, H. Schurz, “Balanced Milstein methods for ordinary SDEs”, *Monte Carlo Methods Applications*, vol.12, pp 143-170, 2006.

[26] G. N. Milstein, E. Platen, H. Schurz, “Balanced implicit methods for stiff stochastic systems”, *Siam Journal on Numerical Analysis*, vol.35, no.3, pp 1010-1019, 1998.