



# Higher strong order methods for linear Itô SDEs on matrix Lie groups

Michelle Muniz<sup>1</sup> · Matthias Ehrhardt<sup>1</sup> · Michael Günther<sup>1</sup> ·  
Renate Winkler<sup>1</sup>

Received: 25 January 2021 / Accepted: 17 December 2021 / Published online: 4 January 2022  
© The Author(s) 2022, corrected publication 2022

## Abstract

In this paper we present a general procedure for designing higher strong order methods for linear Itô stochastic differential equations on matrix Lie groups and illustrate this strategy with two novel schemes that have a strong convergence order of 1.5. Based on the Runge–Kutta–Munthe–Kaas (RKMK) method for ordinary differential equations on Lie groups, we present a stochastic version of this scheme and derive a condition such that the stochastic RKMK has the same strong convergence order as the underlying stochastic Runge–Kutta method. Further, we show how our higher order schemes can be applied in a mechanical engineering as well as in a financial mathematics setting.

**Keywords** Lie group methods · Stochastic Runge–Kutta methods · Geometric integration

**Mathematics Subject Classification** 60H10 · 70G65 · 91G80

---

Communicated by Antonella Zanna Munthe-Kaas.

---

✉ Michelle Muniz  
muniz@uni-wuppertal.de  
Matthias Ehrhardt  
ehrhartd@uni-wuppertal.de  
Michael Günther  
guenther@uni-wuppertal.de  
Renate Winkler  
winkler@uni-wuppertal.de

<sup>1</sup> Chair of Applied Mathematics/Numerical Analysis, Faculty of Mathematics and Natural Sciences, Bergische Universität Wuppertal, Gaußstrasse 20, 42119 Wuppertal, Germany

## 1 Introduction

In the last decades, more and more interrelations in mechanics and finance have been modeled by stochastic differential equations (SDEs) on Lie groups. A trend can be observed that shows that kinematic models, which were previously expressed by ordinary differential equations (ODEs), are now extended by terms that include stochastic processes in order to include possible stochastic perturbations. Examples can be found in the modeling of rigid bodies such as satellites, vehicles and robots [5,6,16,31]. Furthermore, SDEs on Lie groups are also considered in the estimation of object motion from a sequence of projections [28] and in the representation of the precessional motion of magnetization in a solid [1].

In financial mathematics, the consideration of stochastic processes is essential, and the solution of SDEs has been performed for many years, but usually not on Lie groups. However, the use of Lie groups to solve existing or to create new financial models could be of central importance for dealing with geometric constraints. We are confronted with geometric constraints, e.g. in the form of a positivity constraint on interest rates [14,25,30] or a symmetry and positivity constraint on covariance and correlation matrices [22], which are important in, for example, risk management and portfolio optimization.

Despite these diversified applications, the available literature on analysis and numerical methods for SDEs on Lie groups is limited, in contrast to the available literature on ODEs on Lie groups (for example [4,7,9,11,23,24]). Furthermore, the available literature on Lie group SDEs mainly concerns Stratonovich SDEs [1,3,8,16,21,30]. Readers interested in Itô SDEs on Lie groups will only find the *geometric Euler–Maruyama* scheme with strong order  $\gamma = 1$  appearing in [18,19,26] and more recently the existence and convergence proof of the stochastic Magnus expansion in [12]. However, the consideration of Itô SDEs is crucial for its application in finance.

Our contribution to this field of research is a general procedure on how to set up structure-preserving schemes of higher strong order for Itô SDEs on matrix Lie groups. Based on the *Magnus expansion* we apply Itô–Taylor schemes or stochastic Runge–Kutta (SRK) schemes to solve a corresponding SDE in the Lie algebra. Using a SRK method can be interpreted as a stochastic version of Runge–Kutta–Munthe–Kaas (RKMK) methods and if our considered Itô SDEs were transformed to Stratonovich SDEs, this approach would be equivalent to the *stochastic Munthe–Kaas* schemes presented in [16]. Nevertheless, a proof of convergence for this method is still missing. Under these circumstances, we derive a condition such that the stochastic RKMK scheme inherits the strong convergence order  $\gamma$  of the SRK method applied in the Lie algebra.

The remainder of the paper is organized as follows. We start with an introduction to matrix Lie groups, their corresponding Lie algebras and the linear Itô matrix SDE which we consider in this geometric setting in Sect. 2. In Sect. 3 we take a closer look on how SDEs on Lie groups can be solved numerically and present our higher strong order methods. Then we provide some numerical and application examples in Sect. 4. A conclusion of our results is given in Sect. 5.

## 2 SDEs on matrix Lie groups

A Lie group is a differentiable manifold, which is also a group  $G$  with a differentiable product that maps  $G \times G \rightarrow G$ . Matrix Lie groups are Lie groups, which are also subgroups of  $GL(n)$  for  $n \in \mathbb{N}$ . The tangent space at the identity of a matrix Lie group  $G$  is called Lie algebra  $\mathfrak{g}$ . The Lie algebra is closed under forming Lie brackets  $[\cdot, \cdot]$  (also called commutators) of its elements. For further details on Lie groups and Lie algebras we refer the interested reader to [10].

The matrix exponential,  $\exp(\Omega) = \sum_{k=0}^{\infty} \frac{1}{k!} \Omega^k$ , serves as a map from the Lie algebra to the corresponding Lie group,  $\exp: \mathfrak{g} \rightarrow G$ , and is a local diffeomorphism near  $\Omega = 0_{n \times n}$ . Its directional derivative in the direction of an arbitrary matrix  $H \in \mathfrak{g}$  is given by

$$\left( \frac{d}{d\Omega} \exp(\Omega) \right) H = (d \exp_{\Omega}(H)) \exp(\Omega) = \exp(\Omega) (d \exp_{-\Omega}(H))$$

where

$$d \exp_{-\Omega}(H) = \sum_{k=0}^{\infty} \frac{1}{(k+1)!} \text{ad}_{-\Omega}^k(H), \tag{2.1}$$

(see [9, p. 83]). By  $\text{ad}_{\Omega}: \mathfrak{g} \rightarrow \mathfrak{g}$ ,  $\text{ad}_{\Omega}(H) = [\Omega, H] = \Omega H - H \Omega$  we express the adjoint operator which is used iteratively

$$\text{ad}_{\Omega}^0(H) = H, \quad \text{ad}_{\Omega}^k(H) = [\Omega, \text{ad}_{\Omega}^{k-1}(H)] = \text{ad}_{\Omega}(\text{ad}_{\Omega}^{k-1}(H)), \quad k \geq 1.$$

The inverse of  $d \exp_{-\Omega}$  is given in the following Lemma [9, p. 84].

**Lemma 2.1** (Baker, 1905) *If the eigenvalues of the linear operator  $\text{ad}_{\Omega}$  are different from  $2\ell\pi i$  with  $\ell \in \{\pm 1, \pm 2, \dots\}$ , then  $d \exp_{-\Omega}$  is invertible. Furthermore, we have for  $\|\Omega\| < \pi$  that*

$$d \exp_{-\Omega}^{-1}(H) = \sum_{k=0}^{\infty} \frac{B_k}{k!} \text{ad}_{-\Omega}^k(H), \tag{2.2}$$

where  $B_k$  are the Bernoulli numbers, defined by  $\sum_{k=0}^{\infty} (B_k/k!)x^k = x/(e^x - 1)$ .

We recall that the first three Bernoulli numbers are given by  $B_0 = 1$ ,  $B_1 = -\frac{1}{2}$ ,  $B_2 = \frac{1}{6}$  and that  $B_{2m+1} = 0$  holds for  $m \in \mathbb{N}$ .

Let  $(\Theta, \mathcal{F}, \mathbb{P})$  be a complete probability space and let  $W_t = (W_t^1, W_t^2, \dots, W_t^d)$  be a  $d$ -dimensional standard Brownian motion w.r.t. a filtration  $\mathcal{F}_t$  for  $t \geq 0$  which satisfies the usual conditions. On a matrix Lie group  $G$  we consider the linear matrix-valued Itô SDE

$$dQ_t = Q_t K_t dt + Q_t \sum_{i=1}^d V_{i,t} dW_t^i, \quad Q_0 = I_{n \times n}, \tag{2.3}$$

where  $K_t, V_{1,t}, \dots, V_{d,t} \in \mathbb{R}^{n \times n}$  are given coefficient matrices and  $I_{n \times n}$  is the  $n$ -dimensional identity matrix. In general, there exists no closed form solution to (2.3). However, a solution can be defined via a *Magnus expansion*  $Q_t = Q_0 \exp(\Omega_t)$  (see [15,19,24]), where  $\Omega_t \in \mathfrak{g}$  obeys the following matrix SDE

$$d\Omega_t = A(\Omega_t) dt + \sum_{i=1}^d \Gamma_i(\Omega_t) dW_t^i, \quad \Omega_0 = 0_{n \times n}. \tag{2.4}$$

The drift and diffusion coefficients,  $A, \Gamma_i : \mathfrak{g} \rightarrow \mathfrak{g}$ , are given by

$$A(\Omega_t) = d \exp_{-\Omega_t}^{-1} \left( K_t - \frac{1}{2} \sum_{i=1}^d \left( V_{i,t}^2 + C_i(\Omega_t) \right) \right), \quad \Gamma_i(\Omega_t) = d \exp_{-\Omega_t}^{-1} (V_{i,t}) \tag{2.5}$$

with  $C_i : \mathfrak{g} \rightarrow \mathfrak{g}$ ,

$$C_i(\Omega_t) = \left( \frac{d}{d\Omega_t} d \exp_{-\Omega_t} (\Gamma_i(\Omega_t)) \right) \Gamma_i(\Omega_t), \tag{2.6}$$

which can be specified by

$$C_i(\Omega_t) = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{1}{(p+q+2)} \frac{(-1)^p}{p!(q+1)!} \text{ad}_{\Omega_t}^p \left( \text{ad}_{\Gamma_i(\Omega_t)} (\text{ad}_{\Omega_t}^q (\Gamma_i(\Omega_t))) \right), \tag{2.7}$$

for  $i = 1, \dots, d$ . We refer to [19] for the proof.

### 2.1 The Cayley map as local parameterization

With all these series related to the matrix exponential, the question arises whether there is another mapping  $\psi : \mathfrak{g} \rightarrow G$ , which is also a local diffeomorphism near 0 but not based on the evaluation of an infinite number of summands. In case of a *quadratic Lie group* the answer is yes, there is a mapping, namely the *Cayley transformation*

$$\text{cay}(\Omega) = (I - \Omega)^{-1} (I + \Omega).$$

A quadratic Lie group  $G$  is a set of matrices  $Q$  that fulfill the equation  $Q^T P Q = P$  for a given constant matrix  $P$ . For the derivative of  $\text{cay}(\Omega)$  we have

$$\left( \frac{d}{d\Omega} \text{cay}(\Omega) \right) H = (d\text{cay}_{\Omega}(H)) \text{cay}(\Omega) = \text{cay}(\Omega) (d\text{cay}_{-\Omega}(H)). \tag{2.8}$$

The analogue expression to (2.1) is

$$d\text{cay}_{-\Omega}(H) = 2(I + \Omega)^{-1} H (I - \Omega)^{-1}$$

with the inverse given by

$$d\text{cay}_{-\Omega}^{-1}(H) = \frac{1}{2}(I + \Omega)H(I - \Omega), \quad (2.9)$$

see [9, p. 128].

Using the Cayley map instead of the matrix exponential as a local parameterization, the coefficients of (2.4) are

$$A(\Omega_t) = d\text{cay}_{-\Omega_t}^{-1} \left( K_t - \frac{1}{2} \sum_{i=1}^d (V_{i,t}^2 + C_i(\Omega_t)) \right), \quad \Gamma_i(\Omega_t) = d\text{cay}_{-\Omega_t}^{-1}(V_{i,t}),$$

$$C_i(\Omega_t) = \left( \frac{d}{d\Omega_t} d\text{cay}_{-\Omega_t}(\Gamma_i(\Omega_t)) \right) \Gamma_i(\Omega_t) = V_{i,t} \Omega_t V_{i,t} \quad (2.10)$$

for  $i = 1, \dots, d$  (see Appendix for proof).

## 2.2 Example: SDEs on $SO(n)$

As an example of a matrix Lie group, we take a closer look at the special orthogonal group

$$SO(n) = \{X \in GL(n) : X^\top X = I, \det(X) = 1\}$$

which is a quadratic Lie group, such that the Cayley map is also applicable as a local parameterization. The corresponding Lie algebra consists of skew-symmetric matrices:

$$\mathfrak{so}(n) = \{Y \in GL(n) : Y + Y^\top = 0\}.$$

Since we are interested in structure preservation we need conditions that tell us when the solution of an SDE on  $SO(n)$  is kept on the manifold.

**Theorem 2.1** *For the solution  $Q_t$  of (2.3), then  $Q_t \in SO(n)$  if and only if the coefficient matrices satisfy  $V_{i,t} \in \mathfrak{so}(n)$  for  $i = 1, \dots, d$  and  $K_t + K_t^\top = \sum_{i=1}^d V_{i,t}^2$ .*

For the proof of this theorem we refer to [19].

## 3 Numerical methods for SDEs on Lie groups

Applying standard numerical methods for SDEs directly to the linear matrix-valued Itô SDE (2.3) will result in a *drift off*, i.e. the numerical approximations do not stay on the manifold. Consequently, one needs to consider special numerical methods that preserve the geometric properties of the Lie group  $G$ .

As the Lie algebra  $\mathfrak{g}$  represents a linear space with Euclidean-like geometry, it appears reasonable to compute the numerical approximations of the matrix SDE (2.4) and to project the solution back onto the Lie group  $G$ .

A simple scheme based on the Runge–Kutta–Munthe–Kaas schemes for ODEs [23] that puts the described approach into practice can be found in [18] and is presented in the following algorithm.

**Algorithm 3.1** Divide the time interval  $[0, T]$  uniformly into  $J$  subintervals  $[t_j, t_{j+1}]$ ,  $j = 0, 1, \dots, J - 1$  and define the time step  $\Delta = t_{j+1} - t_j$ . Let  $Q_t = Q_0\psi(\Omega_t)$  with  $\psi: \mathfrak{g} \rightarrow G$  be a local parameterization of the Lie group  $G$ . Starting with  $t_0 = 0$ ,  $Q_0 = I_{n \times n}$  and  $\Omega_0 = 0_{n \times n}$  the following steps are repeated over successive intervals  $[t_j, t_{j+1}]$  until  $t_{j+1} = T$ .

1. **Initialization step:** Let  $Q_j$  be the approximation of  $Q_t$  at time  $t = t_j$ . Similarly, let  $K_j$  and  $V_{i,j}$  be the coefficient matrices  $K_t$  and  $V_{i,t}$  evaluated at  $t = t_j$  for  $i = 1, \dots, d$ .
2. **Numerical method step:** Let  $\Omega_\Delta$  be the exact solution of (2.4) after one time step, i.e. at  $t = t_1 = \Delta$ . Compute an approximation  $\Omega_1 \approx \Omega_\Delta$  by applying a stochastic Itô–Taylor or stochastic Runge–Kutta method to the matrix SDE (2.4).
3. **Projection step:** Set  $Q_{j+1} = Q_j \psi(\Omega_1)$ .

Stochastic Runge–Kutta schemes consist of linear operations which would lead to approximations drifting off the Lie group  $G$  if they were applied directly to (2.3). In Algorithm 3.1 these linear operations are all carried out in the vector space  $\mathfrak{g}$  such that the approximation  $\Omega_1$  of (2.4) is also an element of the Lie algebra  $\mathfrak{g}$ . This approximation is then mapped to the Lie group  $G$  by  $\Omega_1 \mapsto Q_{j+1} = Q_j \psi(\Omega_1)$ , i.e. an approximation of (2.3) is obtained. Thus, all schemes obtained by following Algorithm 3.1 produce approximations in  $G$  and the structure of the Lie group is preserved.

The order of convergence of these Lie group structure-preserving schemes clearly depends on the numerical method used in the second step of the algorithm since  $\Omega_1 \mapsto Q_{j+1} = Q_j \psi(\Omega_1)$  is a smooth mapping, i.e. the order on  $G$  is as high as the order of the scheme used for the corresponding equation in  $\mathfrak{g}$ . In order to analyze the accuracy of our geometric numerical methods we recall that an approximating process  $X_t^\Delta$  is said to converge in a strong sense with order  $\gamma > 0$  to the Itô process  $X_t$  if there exists a finite constant  $K$  and a  $\Delta' > 0$  such that

$$\mathbb{E}[|X_T - X_T^\Delta|] \leq K \Delta^\gamma \tag{3.1}$$

for any time discretization with maximum step size  $\Delta \in (0, \Delta')$  [13].

### 3.1 Geometric schemes of strong order 1

Using the Euler–Maruyama scheme in the numerical method step of Algorithm 3.1 results in

$$\begin{aligned}
\Omega_1 &= \Omega_0 + A(\Omega_0)\Delta + \sum_{i=1}^d \Gamma_i(\Omega_0)\Delta W^i \\
&= d\psi_{-\Omega_0}^{-1} \left( K_j - \frac{1}{2} \sum_{i=1}^d V_{i,j}^2 \right) \Delta + \sum_{i=1}^d d\psi_{-\Omega_0}^{-1}(V_{i,j})\Delta W^i, \\
Q_{j+1} &= Q_j \psi(\Omega_1),
\end{aligned} \tag{3.2}$$

where  $\Delta W \sim \mathcal{N}(0, \Delta)$ . Note that  $C_i(\Omega_0) = 0_{n \times n}$ ,  $i = 1, \dots, d$  for both mappings  $\psi = \exp$  (see (2.7)) and  $\psi = \text{cay}$  (see (2.10)) which is why we neglect this coefficient from here on.

Since this scheme (3.2) preserves the geometry of the Lie group  $G$  it was called the *geometric Euler–Maruyama scheme* [19]. It can be specified according to the mapping.

For  $\psi = \exp$ , we get

$$\begin{aligned}
\Omega_1 &= d \exp_{-\Omega_0}^{-1} \left( K_j - \frac{1}{2} \sum_{i=1}^d V_{i,j}^2 \right) \Delta + \sum_{i=1}^d d \exp_{-\Omega_0}^{-1}(V_{i,j})\Delta W^i \\
&= \left( K_j - \frac{1}{2} \sum_{i=1}^d V_{i,j}^2 \right) \Delta + \sum_{i=1}^d V_{i,j} \Delta W^i, \\
Q_{j+1} &= Q_j \exp(\Omega_1),
\end{aligned} \tag{3.3}$$

where inserting  $\Omega_0 = 0_{n \times n}$  is equivalent to truncating the infinite series (2.2) after the first summand, right before any dependence on  $\Omega$  appears.

Using  $\psi = \text{cay}$  instead, we obtain

$$\begin{aligned}
\Omega_1 &= d \text{cay}_{-\Omega_0}^{-1} \left( K_j - \frac{1}{2} \sum_{i=1}^d V_{i,j}^2 \right) \Delta + \sum_{i=1}^d d \text{cay}_{-\Omega_0}^{-1}(V_{i,j}) \Delta W^i \\
&= \frac{1}{2} \left( K_j - \frac{1}{2} \sum_{i=1}^d V_{i,j}^2 \right) \Delta + \frac{1}{2} \sum_{i=1}^d V_{i,j} \Delta W^i, \\
Q_{j+1} &= Q_j \text{cay}(\Omega_1) = Q_j (I - \Omega_1)^{-1} (I + \Omega_1).
\end{aligned}$$

In both cases we see that the diffusion term is only dependent on time and not on the solution itself. This is called *additive noise* [13] and it is the reason why these schemes have strong order  $\gamma = 1$  instead of  $\gamma = 0.5$  as expected for the traditional Euler–Maruyama method. A general proof of the geometric Euler–Maruyama method converging with strong order  $\gamma = 1$  can be found in [26].

### 3.2 Geometric schemes of higher order

A higher strong order than  $\gamma = 1$  can be achieved by applying the strong Itô–Taylor approximation of order  $\gamma = 1.5$  (see [13, p. 351]) in the second step of Algorithm 3.1.

By doing so, we obtain for  $d = 1$

$$\begin{aligned} \Omega_1 &= A(\Omega_0)\Delta + \Gamma(\Omega_0) \Delta W + \frac{1}{2}\Gamma' \Gamma(\Omega_0)((\Delta W)^2 - \Delta) \\ &\quad + A' \Gamma(\Omega_0)\Delta Z \\ &\quad + \frac{1}{2} \left( A' A(\Omega_0) + \frac{1}{2} A'' \Gamma^2(\Omega_0) \right) \Delta^2 \\ &\quad + \left( \Gamma' A(\Omega_0) + \frac{1}{2} \Gamma'' \Gamma^2(\Omega_0) \right) (\Delta W \Delta - \Delta Z) \\ &\quad + \frac{1}{2} (\Gamma' \Gamma(\Omega_0))' \Gamma(\Omega_0) \left( \frac{1}{3} (\Delta W)^2 - \Delta \right) \Delta W, \\ \mathcal{Q}_{j+1} &= \mathcal{Q}_j \psi(\Omega_1). \end{aligned} \tag{3.4}$$

As seen in the previous section  $\Omega_1$  depends on  $t_j$  since

$$A(\Omega_0) = d\psi_{-\Omega_0}^{-1} \left( K_j - \frac{1}{2} V_j^2 \right), \quad \Gamma(\Omega_0) = d\psi_{-\Omega_0}^{-1} (V_j).$$

Representing the double integral  $\int_{\tau_\ell}^{\tau_{\ell+1}} \int_{\tau_\ell}^{s_2} dW_{s_1} ds_2$ , the random variable  $\Delta Z$  is normally distributed with mean  $\mathbb{E}[\Delta Z] = 0$ , variance  $\mathbb{E}[(\Delta Z)^2] = \frac{1}{3} \Delta^3$  and covariance  $\mathbb{E}[\Delta Z \Delta W] = \frac{1}{2} \Delta^2$ . We consider the matrix derivatives as directional derivatives, e.g.

$$A' H = \left( \frac{d}{d\Omega} A(\Omega) \right) H = \frac{d}{d\epsilon} A(\Omega + \epsilon H)|_{\epsilon=0}$$

which we then evaluate at  $\Omega_0$ . The computation of the needed matrix derivatives for  $\psi = \exp$  and  $\psi = \csc$  is provided in the Appendix.

In the case where  $d > 1$  one could apply the corresponding Itô–Taylor scheme for  $d > 1$  (see [13, p. 353]) to (2.4) followed by a projection step.

A strong order of  $\gamma = 1.5$  can also be achieved by applying a stochastic Runge–Kutta method of that order to the SDE (2.4). By using the stochastic Runge–Kutta scheme of order  $\gamma = 1.5$  of Rößler [27], we can avoid computing the derivatives in (3.4) and we obtain for  $d = 1$

$$\begin{aligned} \Omega_1 &= + \left( \frac{1}{3} A(H_1) + \frac{2}{3} A(H_2) \right) \Delta \\ &\quad + \left( \frac{13}{4} \Gamma(\tilde{H}_1) - \frac{9}{4} \Gamma(\tilde{H}_2) - \frac{9}{4} \Gamma(\tilde{H}_3) + \frac{9}{4} \Gamma(\tilde{H}_4) \right) \Delta W \\ &\quad + \left( -\frac{15}{4} \Gamma(\tilde{H}_1) + \frac{15}{4} \Gamma(\tilde{H}_2) + \frac{3}{4} \Gamma(\tilde{H}_3) - \frac{3}{4} \Gamma(\tilde{H}_4) \right) \frac{1}{2\sqrt{\Delta}} ((\Delta W)^2 - \Delta) \\ &\quad + \left( -\frac{9}{4} \Gamma(\tilde{H}_1) + \frac{9}{4} \Gamma(\tilde{H}_2) + \frac{9}{4} \Gamma(\tilde{H}_3) - \frac{9}{4} \Gamma(\tilde{H}_4) \right) \frac{\Delta Z}{\Delta} \end{aligned}$$



$$\begin{aligned}
 &+ \left(6\Gamma(\tilde{H}_1) - 9\Gamma(\tilde{H}_2) + 3\Gamma(\tilde{H}_4)\right) \frac{1}{3!\Delta} \left((\Delta W)^2 - 3\Delta\right) \Delta W, \\
 Q_{j+1} &= Q_j \psi(\Omega_1),
 \end{aligned}
 \tag{3.5}$$

with the stage values

$$\begin{aligned}
 H_1 &= H_3 = \tilde{H}_1 = \Omega_0, \quad H_2 = \frac{3}{4}A(H_1)\Delta + \frac{3}{2}\Gamma(\tilde{H}_1)\frac{\Delta Z}{\Delta}, \\
 \tilde{H}_2 &= \frac{1}{9}A(H_1)\Delta + \frac{1}{3}\Gamma(\tilde{H}_1)\sqrt{\Delta}, \\
 \tilde{H}_3 &= -\frac{5}{9}A(H_1)\Delta + \frac{1}{3}A(H_2)\Delta - \frac{1}{3}\Gamma(\tilde{H}_1)\sqrt{\Delta} + \Gamma(\tilde{H}_2)\sqrt{\Delta}, \\
 \tilde{H}_4 &= -A(H_1)\Delta + \frac{1}{3}A(H_2)\Delta + A(H_3)\Delta \\
 &\quad + \Gamma(\tilde{H}_1)\sqrt{\Delta} - \Gamma(\tilde{H}_2)\sqrt{\Delta} + \Gamma(\tilde{H}_3)\sqrt{\Delta}.
 \end{aligned}$$

The exploitation of stochastic Runge–Kutta methods gives us the benefit of a derivative-free scheme. However, using the mapping  $\psi = \exp$  raises the question of how large the truncation index  $q$  must be chosen in the truncated approximation for (2.2),

$$\sum_{k=0}^q \frac{B_k}{k!} \text{ad}_{-\Omega}^k(H) = H - \frac{1}{2}[-\Omega, H] + \frac{1}{12}[-\Omega, [-\Omega, H]] + \dots, \tag{3.6}$$

in order to maintain a strong order of  $\gamma = 1.5$ . More generally, a condition is needed that connects the truncation index  $q$  with the aimed strong convergence order  $\gamma$ .

Inspired by [9, Theorem IV.8.5.] for Runge–Kutta–Munthe–Kaas methods to solve deterministic matrix ODEs we formulate the following theorem.

**Theorem 3.2** *Consider Algorithm 3.1 with  $\psi = \exp$ . Let the applied stochastic Runge–Kutta method in the second step of Algorithm 3.1 be of strong order  $\gamma$ . If the truncation index  $q$  in (3.6) satisfies  $q \geq 2\gamma - 2$ , then the method of Algorithm 3.1 is of strong order  $\gamma$ .*

**Proof** According to the definition of strong convergence (3.1) we have to show that

$$\mathbb{E}[\|\Omega_\Delta - \Omega_1\|] \leq K \Delta^{(q+2)/2}$$

where  $\Omega_\Delta$  is the exact solution of (2.4) with  $\psi = \exp$  at  $t = \Delta$ ,  $\Omega_1$  is the numerical approximation obtained in the second step of Algorithm 3.1 and  $K$  is a finite constant.

Let  $\Omega_\Delta^q$  be the exact solution of the truncated version of (2.4) with  $\psi = \exp$  at  $t = \Delta$ , namely

$$\Omega_\Delta^q = \int_0^\Delta \sum_{k=0}^q \frac{B_k}{k!} \text{ad}_{-\Omega_t}^k \left( K_s - \frac{1}{2} \sum_{i=1}^d V_{i,s}^2 \right) ds + \sum_{i=1}^d \int_0^\Delta \sum_{k=0}^q \frac{B_k}{k!} \text{ad}_{-\Omega_t}^k (V_{i,s}) dW_s^i.$$

Our proof is divided into six steps.

Step 1: Numerical error

We consider the Frobenius norm of matrices in the Lie algebra  $\mathfrak{g}$  and estimate the error in the  $L^1$ -norm by the  $L^2$ -norm. Then, we use the Minkowski inequality by introducing  $\Omega_\Delta^q$ ,

$$\begin{aligned} \mathbb{E}[\|\Omega_\Delta - \Omega_1\|] &\leq \left(\mathbb{E}\left[\|\Omega_\Delta - \Omega_1\|^2\right]\right)^{1/2} \\ &\leq \left(\mathbb{E}\left[\|\Omega_\Delta - \Omega_\Delta^q\|^2\right]\right)^{1/2} + \left(\mathbb{E}\left[\|\Omega_\Delta^q - \Omega_1\|^2\right]\right)^{1/2}. \end{aligned}$$

We are left with the modelling error, which corresponds to the first summand, and the numerical error, the second summand. The numerical error can be estimated by

$$\left(\mathbb{E}\left[\|\Omega_\Delta^q - \Omega_1\|^2\right]\right)^{1/2} \leq \tilde{K} \Delta^\gamma \quad \text{for } \tilde{K} < \infty,$$

as we assume that we are applying a SRK method of strong order  $\gamma$ .

In other words, it remains to be shown that

$$\left(\mathbb{E}\left[\|\Omega_\Delta - \Omega_\Delta^q\|^2\right]\right)^{1/2} \leq K \Delta^{(q+2)/2}$$

holds for the modelling error.

Step 2: Itô isometry

Inserting the integral equation of (2.4) and its truncated version, we get

$$\begin{aligned} &\left(\mathbb{E}\left[\|\Omega_\Delta - \Omega_\Delta^q\|^2\right]\right)^{1/2} \\ &= \left(\mathbb{E}\left[\left\|\int_0^\Delta \sum_{k=q+1}^\infty \frac{B_k}{k!} \text{ad}_{-\Omega_s}^k \left(K_s - \frac{1}{2} \sum_{i=1}^d V_{i,s}^2\right) ds \right.\right.\right. \\ &\quad \left.\left.\left. + \sum_{i=1}^d \int_0^\Delta \sum_{k=q+1}^\infty \frac{B_k}{k!} \text{ad}_{-\Omega_s}^k (V_{i,s}) dW_s \right\|^2\right]\right)^{1/2} \\ &\leq \left(\mathbb{E}\left[\left\|\int_0^\Delta \sum_{k=q+1}^\infty \frac{B_k}{k!} \text{ad}_{-\Omega_s}^k \left(K_s - \frac{1}{2} \sum_{i=1}^d V_{i,s}^2\right) ds \right\|^2\right]\right)^{1/2} \\ &\quad + \sum_{i=1}^d \left(\mathbb{E}\left[\left\|\int_0^\Delta \sum_{k=q+1}^\infty \frac{B_k}{k!} \text{ad}_{-\Omega_s}^k (V_{i,s}) dW_s \right\|^2\right]\right)^{1/2} \\ &\leq \left(\int_0^\Delta \mathbb{E}\left[\left\|\sum_{k=q+1}^\infty \frac{B_k}{k!} \text{ad}_{-\Omega_s}^k \left(K_s - \frac{1}{2} \sum_{i=1}^d V_{i,s}^2\right)\right\|^2\right] ds\right)^{1/2} \end{aligned}$$

$$\begin{aligned}
 & + \sum_{i=1}^d \left( \int_0^\Delta \mathbb{E} \left[ \left\| \sum_{k=q+1}^\infty \frac{B_k}{k!} \text{ad}_{-\Omega_s}^k(V_{i,s}) \right\|^2 \right] ds \right)^{1/2} \\
 & \leq \left( \int_0^\Delta \mathbb{E} \left[ \left( \sum_{k=q+1}^\infty \frac{|B_k|}{k!} \left\| \text{ad}_{-\Omega_s}^k \left( K_s - \frac{1}{2} \sum_{i=1}^d V_{i,s}^2 \right) \right\| \right)^2 \right] ds \right)^{1/2} \\
 & + \sum_{i=1}^d \left( \int_0^\Delta \mathbb{E} \left[ \left( \sum_{k=q+1}^\infty \frac{|B_k|}{k!} \left\| \text{ad}_{-\Omega_s}^k(V_{i,s}) \right\| \right)^2 \right] ds \right)^{1/2},
 \end{aligned}$$

where we also used the Minkowski inequality, the Itô isometry and the properties of a matrix norm. Now, the summands in the last line differ only in the input matrix of the adjoint operator.

**Step 3: Adjoint operator**

We estimate the Frobenius norm of the adjoint operator of  $V_{i,s}$  for a fixed  $s \in [0, \Delta]$  and keep in mind that analogous estimates hold for the adjoint operator of  $K_s - \frac{1}{2} \sum_{i=1}^d V_{i,s}^2$ . Since the Frobenius norm is submultiplicative, we have

$$\|\text{ad}_{-\Omega_s}(V_{i,s})\| = \|[-\Omega_s, V_{i,s}]\| \leq \|\Omega_s V_{i,s}\| + \|V_{i,s} \Omega_s\| \leq 2\|\Omega_s\| \|V_{i,s}\|.$$

As a direct consequence, it holds

$$\|\text{ad}_{-\Omega_s}^k(V_{i,s})\| \leq 2^k \|\Omega_s\|^k \|V_{i,s}\|,$$

which can also be shown via induction. Inserting this result in the expected value considered in the last line of the previous step, we get

$$\begin{aligned}
 \mathbb{E} \left[ \left( \sum_{k=q+1}^\infty \frac{|B_k|}{k!} \left\| \text{ad}_{-\Omega_s}^k(V_{i,s}) \right\| \right)^2 \right] & \leq \mathbb{E} \left[ \left( \sum_{k=q+1}^\infty \frac{|B_k|}{k!} 2^k \|\Omega_s\|^k \|V_{i,s}\| \right)^2 \right] \\
 & = \|V_{i,s}\|^2 \mathbb{E} \left[ \left( \sum_{k=q+1}^\infty \frac{|B_k|}{k!} 2^k \|\Omega_s\|^k \right)^2 \right].
 \end{aligned}$$

**Step 4: Estimate for the remainder**

It is known that the Bernoulli numbers are implicitly defined by  $\sum_{k=0}^\infty (B_k/k!)x^k = x/(e^x - 1)$ . Inserting the absolute values of the Bernoulli numbers instead, then

$$\sum_{k=0}^\infty \frac{|B_k|}{k!} x^k = \frac{x}{2} \left( 1 + \cot\left(\frac{x}{2}\right) \right) + 2$$

(see [11, p. 48]). Let  $f : I \rightarrow \mathbb{R}, x \mapsto \frac{x}{2} (1 + \cot(\frac{x}{2})) + 2$  with  $I = \{x \in \mathbb{R} : \frac{x}{2\pi} \notin \mathbb{Z}\}$ . Applying Taylor’s theorem to the function  $f$  at the point 0 gives

$$f(x) = \sum_{k=0}^q \frac{f^{(k)}(0)}{k!} x^k + R_q(x), \quad R_q(x) = \frac{f^{(q+1)}(\xi)}{(q+1)!} x^{q+1},$$

where we consider the Lagrange form of the remainder for some real number  $\xi$  between 0 and  $x$ .

Setting  $x = 2\|\Omega_s\|$  and recalling that the expression (2.2) only converges for  $\|\Omega\| < \pi$ , we now consider  $f|_{\tilde{I}} : \tilde{I} \rightarrow \mathbb{R}, x \mapsto \frac{x}{2} (1 + \cot(\frac{x}{2})) + 2$  with  $\tilde{I} = \{x \in \mathbb{R} : |x| < 2\pi\}$ . The restriction of  $f$  to  $\tilde{I}$  is bounded, in particular there exists an upper bound  $M_q$  such that  $|f|_{\tilde{I}}^{(q+1)}(\xi)| \leq M_q$  for all  $\xi$  between 0 and  $x$ . Moreover, the following estimate for the remainder holds

$$|R_q(x)| = \left| \frac{f|_{\tilde{I}}^{(q+1)}(\xi)}{(q+1)!} x^{q+1} \right| \leq \frac{M_q}{(q+1)!} |x|^{q+1} \leq \frac{M_q}{(q+1)!} (2\|\Omega_s\|)^{q+1}.$$

Using this estimate in the expected value of the last line of the previous step results in

$$\begin{aligned} \mathbb{E} \left[ \left( \sum_{k=q+1}^{\infty} \frac{|B_k|}{k!} (2\|\Omega_s\|)^k \right)^2 \right] &\leq \mathbb{E} \left[ \left( \frac{M_q}{(q+1)!} (2\|\Omega_s\|)^{q+1} \right)^2 \right] \\ &= \left( \frac{2^{q+1} M_q}{(q+1)!} \right)^2 \mathbb{E} \left[ \|\Omega_s\|^{2q+2} \right]. \end{aligned}$$

**Step 5: Itô–Taylor expansion**

The goal of this step is to find an estimate for  $\mathbb{E} [\|\Omega_s\|^{2q+2}]$ . For this purpose, we examine the following Itô–Taylor expansion

$$\Omega_s = \Omega_0 + R_s = R_s, \quad \mathbb{E} [\|R_s\|^2] \leq C_1 s$$

where  $C_1$  is a finite constant, for details see [13, Proposition 5.9.1]. This leads to the estimate

$$\mathbb{E} [\|\Omega_s\|^{2(q+1)}] = \mathbb{E} [\|R_s\|^{2(q+1)}] \leq C_1 s^{q+1}.$$

**Step 6: Overall estimate**

Gathering the results of the previous steps and inserting a Taylor expansion for  $V_{i,s}$  where  $C_2 < \infty$  gives

$$\begin{aligned} & \mathbb{E} \left[ \left( \sum_{k=q+1}^{\infty} \frac{|B_k|}{k!} \|\text{ad}_{-\Omega_s}^k(V_{i,s})\| \right)^2 \right] \\ & \leq (\|V_{i,0}\| + C_2s)^2 \left( \frac{2^{q+1}M_q}{(q+1)!} \right)^2 C_1s^{q+1} = \mathcal{O}(s^{q+1}). \end{aligned}$$

Thus,

$$\left( \int_0^\Delta \mathbb{E} \left[ \left( \sum_{k=q+1}^{\infty} \frac{|B_k|}{k!} \|\text{ad}_{-\Omega_s}^k(V_{i,s})\| \right)^2 \right] ds \right)^{1/2} = \mathcal{O}(\Delta^{(q+2)/2}).$$

Analogously, one can show that

$$\left( \int_0^\Delta \mathbb{E} \left[ \left( \sum_{k=q+1}^{\infty} \frac{|B_k|}{k!} \|\text{ad}_{-\Omega_s}^k(K_s - \frac{1}{2}V_{i,s}^2)\| \right)^2 \right] ds \right)^{1/2} = \mathcal{O}(\Delta^{(q+2)/2}),$$

which concludes the proof. □

**Remark 3.1** 1. Assuming that multiple Wiener integrals can be simulated efficiently, a general procedure for designing methods of strong order  $\gamma$  for linear Itô SDEs on matrix Lie groups can be achieved by applying a SRK scheme of order  $\gamma$  in the second step of Algorithm 3.1 while evaluating the sum (3.6) involved in the coefficients of equation (2.4) only up to index  $2\gamma - 2$ . We refer the reader to [17] and the references therein for more details on the approximation of iterated stochastic integrals.

2. Theorem 3.2 is in accordance with our results of Sect. 3.1, where the geometric Euler–Maruyama scheme (3.3) can be interpreted as a stochastic RKMK method with  $\gamma = 1$  and  $q = 0$ .
3. Due to the definition of the Cayley map as a finite product of matrices, there is no modelling error and therefore, no such theorem is needed if  $\psi = \text{cay}$  is chosen as the local parameterization in Algorithm 3.1.

### 4 Numerical examples

In the following we provide numerical examples that illustrate the effectiveness of the proposed geometric methods, firstly, by simulating the strong convergence order of

the proposed schemes and secondly, by showing the Lie group structure preservation of our methods.

For checking the convergence order, we set  $G = \text{SO}(3)$  and  $\mathfrak{g} = \mathfrak{so}(3)$ . In order to ensure the conditions of Theorem 2.1 we have used the set up of matrices  $K_t$  and  $V_t$  proposed by Muniz et al. [22] for  $d = 1$ . Specifically, we chose the time-dependent functions

$$f_1(t) = \cos(t), \quad f_2(t) = \sin(t), \quad f_3(t) = 1 + t + t^2 + t^3,$$

to compute a skew-symmetric matrix  $V_t$  as a linear combination,

$$V_t = f_1(t)G_1 + f_2(t)G_2 + f_3(t)G_3,$$

where  $G_i, i = 1, 2, 3$  are the following generators of the Lie algebra  $\mathfrak{so}(3)$ ,

$$G_1 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad G_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad G_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Note that the functions  $f_i, i = 1, 2, 3$  can be chosen arbitrarily. We then set the matrix  $K_t$  as the lower triangular matrix of  $V_t^2$  where the diagonal entries of  $K_t$  are 0.5 times the diagonal entries of  $V_t^2$  such that  $K_t + K_t^\top = V_t^2$ .

We simulated  $M = 1000$  different paths of two independent realizations of a standard normally distributed random variable,  $U_1, U_2 \sim \mathcal{N}(0, 1)$ . Then, the random variables used in the numerical method step in Algorithm 3.1 were simulated as  $\widehat{\Delta W} = U_1\sqrt{\Delta}$  and  $\widehat{\Delta Z} = \frac{1}{2}\Delta(\widehat{\Delta W} + U_2\sqrt{\frac{\Delta}{3}})$ . The absolute error as defined in (3.1) was estimated by using the Frobenius norm at  $t_j = T$ , i.e. by

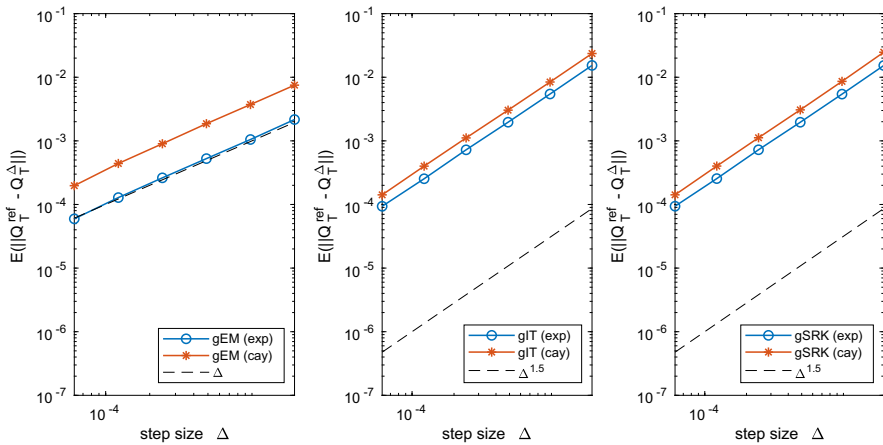
$$\frac{1}{M} \sum_{i=1}^M \left( \| Q_{T,i}^{\text{ref}} - Q_{T,i}^A \|_F \right)$$

where the approximations  $Q_{T,i}^A$  were obtained by using Algorithm 3.1 with step sizes  $\Delta = 2^{-14}, 2^{-13}, 2^{-12}, 2^{-11}, 2^{-10}, 2^{-9}$  and for the reference solution  $Q_{T,i}^{\text{ref}}$  we used the same method with  $\psi = \text{cay}$  and step size  $\Delta = 2^{-16}$ , respectively.

A log–log plot of the estimation of the absolute error against the step sizes can be viewed in Fig. 1. It indicates the strong order of convergence claimed in the sections above for the geometric Euler–Maruyama scheme (3.2), the geometric version of the Itô–Taylor scheme (3.4) and the geometric stochastic Runge–Kutta scheme (3.5).

Examples from financial mathematics and multibody system dynamics verify that the structure-preserving methods derived above can be applied in practice.

In the first example we apply our methods of strong order  $\gamma = 1.5$  to an SDE on  $\text{SO}(2)$  in the context of stochastic correlation modelling. The second example shows how our methods can be used in the modeling of rigid bodies, such as satellites. Although, we have restricted our research for this paper to considering only linear



**Fig. 1** Simulation of the strong convergence order for  $M = 1000$  paths. Left: Geometric Euler–Maruyama scheme (gEM). Center: Geometric version of the Itô–Taylor scheme of strong order 1.5 (gIT). Right: Geometric version of Rößler’s stochastic Runge–Kutta scheme of strong order 1.5 (gSRK)

SDEs on Lie groups, the second example shows that our methods can also be applied to nonlinear SDEs on e.g.  $SO(3)$ .

**4.1 A stochastic correlation model**

Let us assume that a risk manager retrieves from the middle office’s reporting system an initial value of the correlation between two assets and a density function of the considered correlation. Moreover, we assume that the risk manager was given the task to generate correlation matrices that not only approximate the given density function but also respect the stochastic behaviour of correlations.

This problem can be solved by the stochastic correlation model presented in [22]. The main ideas of the approach are outlined below.

For this example we consider historical prices of the S&P 500 index and the Euro/US-Dollar exchange rate and compute moving correlations with a window size of 30 days to obtain correlations from January 03, 2005 to January 06, 2006 (see Fig. 2).

The corresponding initial correlation matrix calculated from this data and imputed to the risk manager is

$$R_0^{\text{hist}} = \begin{pmatrix} 1 & -0.0159 \\ -0.0159 & 1 \end{pmatrix}.$$

Furthermore, we estimate a density function from the historical data using kernel smoothing functions, which is also plotted in Fig. 3. For more details on the density estimation see [2].

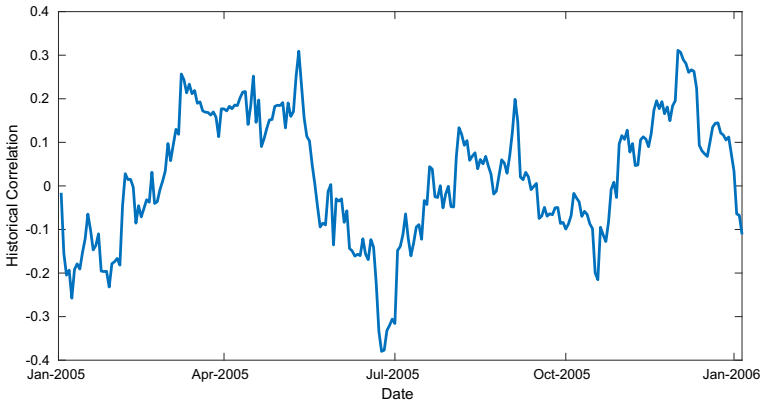


Fig. 2 The 30-day historical correlations between S&P 500 and Euro/US-Dollar exchange rate, source of data: [www.yahoo.com](http://www.yahoo.com)

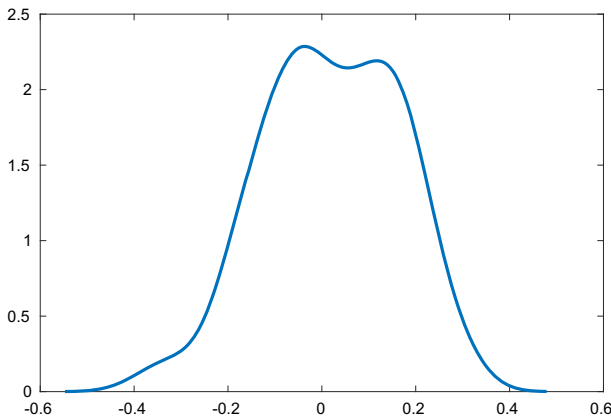


Fig. 3 Empirical density function of the historical correlation between S&P 500 and Euro/US-Dollar exchange rate, computed with the MATLAB function `ksdensity`

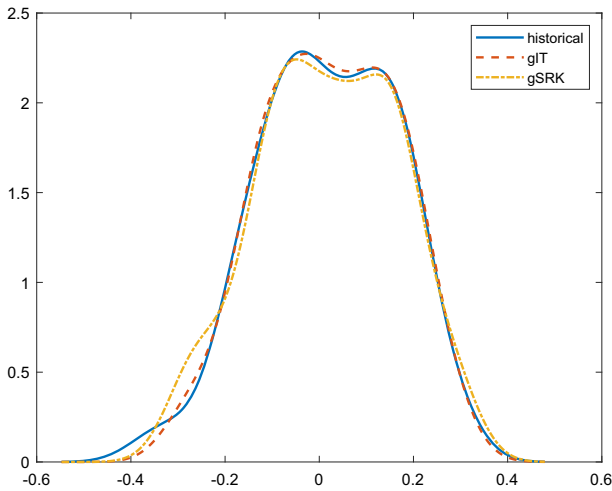
As a first step, we focus on covariance matrices  $P_t, t \geq 0$ . The authors of [29] utilised the principal axis theorem and defined the *covariance flow*

$$P_t = Q_t^\top P_0 Q_t, \quad t \geq 0, \tag{4.1}$$

where  $P_0$  is the initial covariance matrix computed based on  $R_0^{\text{hist}}$  and  $Q_t$  is an orthogonal matrix which without loss of generality can be assumed to have determinant +1, i.e.  $Q_t \in \text{SO}(2)$ . Following the approach in [22] the matrix  $Q_t$  is now assumed to be driven by the SDE (2.3) which can be solved by using Algorithm 3.1. With the resulting matrices approximations of  $P_t$  can be computed with (4.1), which can then be transformed to corresponding correlation matrices

$$R_t = \Sigma_t^{-1} P_t \Sigma_t^{-1}$$





**Fig. 4** Empirical density function of the historical correlation and the correlation flow between S&P 500 and Euro/US-Dollar exchange rate

with  $\Sigma_t = (\text{diag}(P_t))^{1/2}$ .

At last, a density function is estimated from this *correlation flow* and the free parameters involved are calibrated such that the density function matches the density function from the historical data, see [22] for details.

We executed this procedure using the geometric Itô–Taylor scheme (3.4) with  $\psi = \text{cay}$  (gIT) and the geometric Rößler scheme (3.5) with  $\psi = \text{exp}$  and truncation index  $q = 1$  (gSRK) in the second step of Algorithm 3.1, respectively. The results are plotted in Fig. 4, which shows that both density functions approximate the density function of the historical data quite well.

### 4.2 The stochastic rigid body problem

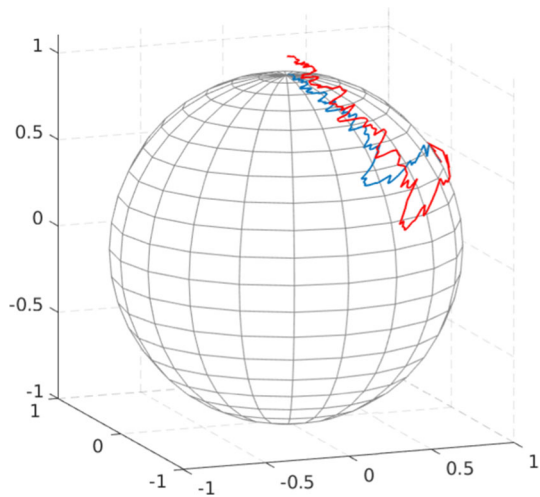
Consider a free rigid body, whose centre of mass is at the origin. Let the vector  $y = (y_1, y_2, y_3)^T$  represent the angular momentum in the body frame and  $I_1, I_2$  and  $I_3$  be the principal moments of inertia [20]. Then the motion of this free rigid body is described by the Euler equations

$$\dot{y} = V(y)y, \quad V(y) = \begin{pmatrix} 0 & y_3/I_3 & -y_2/I_2 \\ -y_3/I_3 & 0 & y_1/I_1 \\ y_2/I_2 & -y_1/I_1 & 0 \end{pmatrix}.$$

We suppose that the rigid body is perturbed by a Wiener process  $W_t$  and compute a matrix  $K(y)$  such that the dynamics are kept on the manifold, i.e. we compute  $K(y)$  from the condition  $K(y) + K^T(y) = V^2(y)$ . Consequently, we regard the Itô SDE

$$dy = K(y)y dt + V(y)y dW_t, \tag{4.2}$$

**Fig. 5** Sample paths of the geometric Euler–Maruyama (blue) and the traditional Euler–Maruyama scheme (red) applied to (4.2)



where the solution evolves on the unit sphere if the initial value  $y_0$  satisfies  $|y_0| = 1$ . Note that stochastic versions of the rigid body problem have already been considered in [16] and [30] but as Stratonovich SDEs.

Since the solution of (4.2) can also be written as  $y = Qy_0$  where  $Q \in \text{SO}(3)$ , we focus on the nonlinear matrix SDE

$$dQ = K(Q)Q dt + V(Q)Q dW_t, \quad Q_0 = I_{3 \times 3}. \tag{4.3}$$

The coefficients of the corresponding SDE in the Lie algebra (2.4) are

$$\begin{aligned} A(\Omega) &= d\psi_\Omega^{-1} \left( K(\psi(\Omega)Q_0) - \frac{1}{2}V^2(\psi(\Omega)Q_0) \right), \\ \Gamma(\Omega) &= d\psi_\Omega^{-1} \left( V(\psi(\Omega)Q_0) \right). \end{aligned} \tag{4.4}$$

Now, SDE (4.2) can be solved by applying Algorithm 3.1 to the SDE (4.3). Note that we deal with right multiplication of the solution  $Q$  on the right hand side of (4.3) instead of left multiplication as in (2.3). As a consequence, the sign of the index of the operator  $d\psi^{-1}$  is changed in (4.4) and the solution of the Projection step in Algorithm 3.1 should be  $Q_{j+1} = \psi(\Omega_{j+1})Q_j$ . We refer to [23] for more details on this matter.

In Fig. 5 we simulated 200 steps of the trajectory of (4.2) with a step size of  $\Delta = 0.03$  by using Algorithm 3.1 with the initial values  $y_0 = (\sin(1.1), 0, \cos(1.1))^T$  and the moments of inertia  $I_1 = 2, I_2 = 1$  and  $I_3 = 2/3$ . For the numerical method step of Algorithm 3.1 we used the Euler–Maruyama scheme with  $\psi = \text{cay}$ . Emphasizing the structure-preserving character of Algorithm 3.1 we also plotted a sample path of the traditional Euler–Maruyama scheme applied directly to (4.2), whose trajectory clearly fails to stay on the manifold. This phenomenon can also be viewed in Fig. 6 where we visualize the distance of the approximate solutions from the manifold.

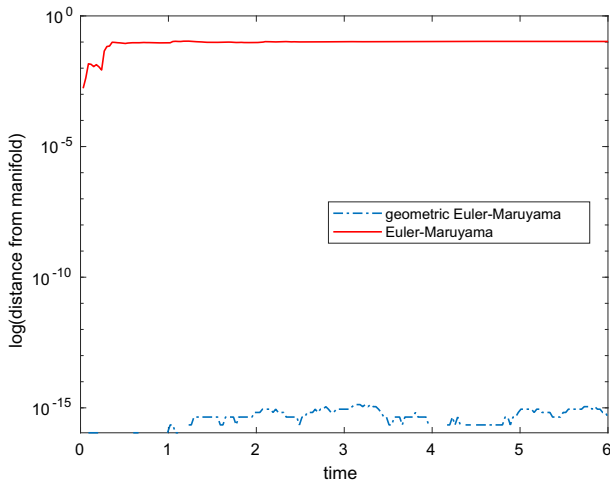


Fig. 6 Log-distance of the numerical solutions to the unit sphere

## 5 Conclusion

We have presented stochastic Lie group methods for linear Itô SDEs on matrix Lie groups that have a higher strong convergence order than the known *geometric Euler–Maruyama* scheme. Based on RKMK methods for ODEs on Lie groups, we have proven a condition on the truncation index of the inverse of  $d \exp_{-\Omega}(H)$  such that the stochastic RKMK method inherits the convergence order of the underlying SRK. This allows us to construct even higher order schemes than the here presented methods of strong order  $\gamma = 1.5$  if one were able to approximate multiple Wiener integrals efficiently. Additionally, we have shown examples for the application of our methods in mechanical engineering and in financial mathematics.

Our methods require further investigations for the application to nonlinear Itô SDEs on matrix Lie groups, which we consider as future work. Moreover, we have restricted our research for this paper to the strong convergence order. In future research, an investigation on the weak convergence order of stochastic Lie group methods will also be conducted.

**Acknowledgements** The authors would like to thank Martin Friesen (Dublin City University) for the in-depth discussions that improved the content of this paper. The work of the authors was partially supported by the bilateral German-Slovakian Project *MATTHIAS—Modelling and Approximation Tools and Techniques for Hamilton-Jacobi-Bellman equations in finance and Innovative Approach to their Solution*, financed by DAAD and the Slovakian Ministry of Education. Further the authors acknowledge partial support from the bilateral German-Portuguese Project *FRAC TAL – FRActional models and CompuTationAL Finance* financed by DAAD and the CRUP - Conselho de Reitores das Universidades Portuguesas.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If

material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## A Proofs

**Theorem A.1** *The solution of (2.3) can be written as  $Q_t = Q_0 \text{cay}(\Omega_t)$ , where  $\Omega_t$  obeys the SDE (2.4) with coefficients given by*

$$A(\Omega_t) = d\text{cay}_{-\Omega_t}^{-1} \left( K_t - \frac{1}{2} \sum_{i=1}^d \left( V_{i,t}^2 + C_i(\Omega_t) \right) \right), \quad \Gamma_i(\Omega_t) = d\text{cay}_{-\Omega_t}^{-1}(V_{i,t})$$

with

$$C_i(\Omega_t) = \left( \frac{d}{d\Omega_t} d\text{cay}_{-\Omega_t}(\Gamma_i(\Omega_t)) \right) \Gamma_i(\Omega_t), \quad i = 1, \dots, d. \quad (\text{A.1})$$

**Proof** Let  $f: \mathfrak{g} \rightarrow G$ ,  $f(\Omega_t) = Q_0 \text{cay}(\Omega_t)$  such that  $Q_t = f(\Omega_t)$ . Due to the Itô rules, i.e.  $dt \cdot dt = dt \cdot dW_t = dW_t \cdot dt = 0$  and  $dW_t \cdot dW_t = dt$ ,  $dQ_t = d(f(\Omega_t))$  is given exactly by the first two terms of the Taylor expansion

$$\begin{aligned} dQ_t &= \frac{d}{d\epsilon} f(\Omega_t + \epsilon d\Omega_t) \Big|_{\epsilon=0} + \frac{1}{2} \frac{d^2}{d\epsilon^2} f(\Omega_t + \epsilon d\Omega_t) \Big|_{\epsilon=0} \\ &= \left( \frac{d}{d\Omega_t} f(\Omega_t) \right) d\Omega_t + \frac{1}{2} \left( \frac{d^2}{d\Omega_t^2} f(\Omega_t) \right) (d\Omega_t)^2 \\ &= \left( \left( \frac{d}{d\Omega_t} f(\Omega_t) \right) A(\Omega_t) + \frac{1}{2} \sum_{i=1}^d \left( \frac{d^2}{d\Omega_t^2} f(\Omega_t) \right) \Gamma_i^2(\Omega_t) \right) dt \\ &\quad + \sum_{i=1}^d \left( \frac{d}{d\Omega_t} f(\Omega_t) \right) \Gamma_i(\Omega_t) dW_t^i \end{aligned}$$

where we have used the fact that

$$(d\Omega_t)^2 = \left( A(\Omega_t) dt + \sum_{i=1}^d \Gamma_i(\Omega_t) dW_t^i \right)^2 = \sum_{i=1}^d \Gamma_i^2(\Omega_t) dt.$$

We use (2.8) to specify the first part of the drift coefficient

$$\begin{aligned} \left( \frac{d}{d\Omega_t} f(\Omega_t) \right) A(\Omega_t) &= Q_0 \left( \frac{d}{d\Omega_t} \text{cay}(\Omega_t) \right) A(\Omega_t) = Q_0 \text{cay}(\Omega_t) d\text{cay}_{-\Omega_t}(A(\Omega_t)) \\ &= Q_t d\text{cay}_{-\Omega_t}(A(\Omega_t)). \end{aligned}$$

Analogously, we have  $\left( \frac{d}{d\Omega_t} f(\Omega_t) \right) \Gamma_i(\Omega_t) = Q_t d\text{cay}_{-\Omega_t}(\Gamma_i(\Omega_t))$ . For the second derivative we obtain

$$\begin{aligned}
 \left(\frac{d^2}{d\Omega_t^2} f(\Omega_t)\right) \Gamma_i^2(\Omega_t) &= \left(\frac{d}{d\Omega} Q_t dcay_{-\Omega_t}(\Gamma_i(\Omega_t))\right) \Gamma_i(\Omega_t) \\
 &= \left(\left(\frac{d}{d\Omega_t} Q_t\right) \Gamma_i(\Omega_t)\right) dcay_{-\Omega_t}(\Gamma_i(\Omega_t)) \\
 &\quad + Q_t \left(\frac{d}{d\Omega_t} dcay_{-\Omega_t}(\Gamma_i(\Omega_t))\right) \Gamma_i(\Omega_t) \\
 &= Q_t \left(dcay_{-\Omega_t}(\Gamma_i(\Omega_t))\right)^2 + Q_t C_i(\Omega_t),
 \end{aligned}$$

where  $C_i(\Omega_t) = \left(\frac{d}{d\Omega_t} dcay_{-\Omega_t}(\Gamma_i(\Omega_t))\right) \Gamma_i(\Omega_t)$ . Comparing these results with SDE (2.3) we get

$$\begin{aligned}
 V_{i,t} &= dcay_{-\Omega_t}(\Gamma_i(\Omega_t)) \quad \text{and} \quad K_t = dcay_{-\Omega_t}(A(\Omega_t)) \\
 &\quad + \frac{1}{2} \sum_{i=1}^d (dcay_{-\Omega_t}(\Gamma_i(\Omega_t)))^2 + \frac{1}{2} \sum_{i=1}^d C_i(\Omega_t)
 \end{aligned}$$

and thus

$$\Gamma_i(\Omega_t) = dcay_{-\Omega_t}^{-1}(V_{i,t}) \quad \text{and} \quad A(\Omega_t) = dcay_{-\Omega_t}^{-1}\left(K_t - \frac{1}{2} \sum_{i=1}^d (V_{i,t}^2 + C_i(\Omega_t))\right).$$

□

**Lemma A.1** *The coefficient (A.1) is given by*

$$C_i(\Omega_t) = V_{i,t} \Omega_t V_{i,t}. \tag{A.2}$$

**Proof**

$$\begin{aligned}
 &\left(\frac{d}{d\Omega_t} dcay_{-\Omega_t}(H)\right) \tilde{H} \\
 &= \left(2 \frac{d}{d\Omega_t} \left((I + \Omega_t)^{-1} H (I - \Omega_t)^{-1}\right)\right) \tilde{H} \\
 &= 2 \left(\left(\frac{d}{d\Omega_t} (I + \Omega_t)^{-1}\right) \tilde{H}\right) H (I - \Omega_t)^{-1} \\
 &\quad + 2(I + \Omega_t)^{-1} H \left(\frac{d}{d\Omega_t} (I - \Omega_t)^{-1}\right) \tilde{H} \\
 &= -2(I + \Omega_t)^{-1} \tilde{H} (I + \Omega_t)^{-1} H (I - \Omega_t)^{-1} \\
 &\quad + 2(I + \Omega_t)^{-1} H (I - \Omega_t)^{-1} \tilde{H} (I - \Omega_t)^{-1}
 \end{aligned}$$

Inserting for both  $H$  and  $\tilde{H}$  the diffusion coefficient  $\Gamma_i(\Omega_t) = d\text{cay}_{-\Omega_t}^{-1}(V_{i,t}) = \frac{1}{2}(I + \Omega_t)V_{i,t}(I - \Omega_t)$  we get

$$\begin{aligned} & \left( \frac{d}{d\Omega_t} d\text{cay}_{-\Omega_t}(\Gamma_i(\Omega_t)) \right) \Gamma_i(\Omega_t) \\ &= -2(I + \Omega_t)^{-1} \Gamma_i(\Omega_t)(I + \Omega_t)^{-1} \Gamma_i(\Omega_t)(I - \Omega_t)^{-1} \\ & \quad + 2(I + \Omega_t)^{-1} \Gamma_i(\Omega_t)(I - \Omega_t)^{-1} \Gamma_i(\Omega_t)(I - \Omega_t)^{-1} \\ &= -\frac{1}{2}(I + \Omega_t)^{-1}(I + \Omega_t)V_{i,t}(I - \Omega_t)(I + \Omega_t)^{-1}(I + \Omega_t)V_{i,t}(I - \Omega_t)(I - \Omega_t)^{-1} \\ & \quad + \frac{1}{2}(I + \Omega_t)^{-1}(I + \Omega_t)V_{i,t}(I - \Omega_t)(I - \Omega_t)^{-1}(I + \Omega_t)V_{i,t}(I - \Omega_t)(I - \Omega_t)^{-1} \\ &= -\frac{1}{2}V_{i,t}(I - \Omega_t)V_{i,t} + \frac{1}{2}V_{i,t}(I + \Omega_t)V_{i,t} \\ &= V_{i,t}\Omega_t V_{i,t}. \end{aligned}$$

□

## B Matrix derivatives

In this section we provide the matrix derivatives that we used in the geometric version of the Itô–Taylor scheme of strong order  $\gamma = 1.5$  (see (3.4)).

### B.1 Derivatives for $\psi = \text{cay}$

Computing the derivative of (2.9) in the direction of an arbitrary matrix  $\tilde{H}$ , we get

$$\begin{aligned} \left( \frac{d}{d\Omega} d\text{cay}_{-\Omega}^{-1}(H) \right) \tilde{H} &= \left( \frac{1}{2} \frac{d}{d\Omega} (H - H\Omega + \Omega H - \Omega H\Omega) \right) \tilde{H} \\ &= \frac{1}{2} \frac{d}{dt} (H - H(\Omega + t\tilde{H}) + (\Omega + t\tilde{H})H \\ & \quad - (\Omega + t\tilde{H})H(\Omega + t\tilde{H})) \Big|_{t=0} \\ &= \frac{1}{2} (-H\tilde{H} + \tilde{H}H - \Omega H\tilde{H} - \tilde{H}H\Omega). \end{aligned} \tag{B.1}$$

Subsequently, the second directional derivative is

$$\left( \frac{d^2}{d\Omega^2} d\text{cay}_{-\Omega}^{-1}(H) \right) \tilde{H}^2 = -\tilde{H}H\tilde{H}. \tag{B.2}$$

Inserting  $H = V$  and  $\tilde{H} = \Gamma(\Omega) = d\text{cay}_{-\Omega}^{-1}(V)$ , we obtain

$$\left( \frac{d}{d\Omega} \Gamma(\Omega) \right) \Gamma(\Omega) = \left( \frac{d}{d\Omega} d\text{cay}_{-\Omega}^{-1}(V) \right) \Gamma(\Omega)$$

$$\begin{aligned}
 &= \frac{1}{2}(-V\Gamma(\Omega) + \Gamma(\Omega)V - \Omega V\Gamma(\Omega) - \Gamma(\Omega)V\Omega) \\
 &= \frac{1}{2}(-V\Omega V + V\Omega V\Omega - \Omega V\Omega V + \Omega V\Omega V\Omega).
 \end{aligned}$$

Similar expressions are obtained for  $(\frac{d}{d\Omega} A(\Omega)) \Gamma(\Omega)$ ,  $(\frac{d}{d\Omega} A(\Omega)) A(\Omega)$  and  $(\frac{d}{d\Omega} \Gamma(\Omega)) A(\Omega)$  by inserting  $H = K - \frac{1}{2}V^2$  and  $\tilde{H} = \Gamma(\Omega)$ ,  $H = K - \frac{1}{2}V^2$  and  $\tilde{H} = A(\Omega)$  and  $H = V$  and  $\tilde{H} = A(\Omega)$  in (B.1), respectively.

Proceed accordingly to compute the second derivatives  $(\frac{d^2}{d\Omega^2} A(\Omega)) \Gamma^2(\Omega)$  and  $(\frac{d^2}{d\Omega^2} \Gamma(\Omega)) \Gamma^2(\Omega)$ .

**B.2 Derivatives for  $\psi = \exp$**

In the following we present derivatives of (2.1) up to  $k = 4$ , i.e. of

$$\begin{aligned}
 \sum_{k=0}^4 \frac{B_k}{k!} \text{ad}_{-\Omega}^k(H) &= H - \frac{1}{2}[-\Omega, H] + \frac{1}{12}[-\Omega, [-\Omega, H]] \\
 &\quad - \frac{1}{720}[-\Omega, [-\Omega[-\Omega, [-\Omega, H]]]] \\
 &= H - \frac{1}{2}(H\Omega - \Omega H) + \frac{1}{12}(\Omega^2 H + H\Omega^2 - 2\Omega H\Omega) \\
 &\quad - \frac{1}{720}(\Omega^4 H - 4\Omega^3 H\Omega + 6\Omega^2 H\Omega^2 - 4\Omega H\Omega^3 + H\Omega^4).
 \end{aligned}$$

Computing the directional derivative we get

$$\begin{aligned}
 &\left(\frac{d}{d\Omega} \sum_{k=0}^4 \frac{B_k}{k!} \text{ad}_{-\Omega}^k(H)\right) \tilde{H} \\
 &= -\frac{1}{2}(H\tilde{H} - \tilde{H}H) \\
 &\quad + \frac{1}{12}(\Omega\tilde{H}H + \tilde{H}\Omega H + H\Omega\tilde{H} + H\tilde{H}\Omega - 2\tilde{H}H\Omega - 2\Omega H\tilde{H}) \\
 &\quad - \frac{1}{720}(\tilde{H}\Omega^3 H + \Omega\tilde{H}\Omega^2 H + \Omega^2\tilde{H}\Omega H + \Omega^3\tilde{H}H \\
 &\quad - 4(\Omega\tilde{H}\Omega H\Omega + \tilde{H}\Omega^2 H\Omega + \Omega^2\tilde{H}H\Omega + \Omega^3\tilde{H}H) \\
 &\quad + 6(\Omega^2 H\Omega\tilde{H} + \Omega^2 H\tilde{H}\Omega + \Omega\tilde{H}H\Omega^2 + \tilde{H}\Omega H\Omega^2) \\
 &\quad - 4(\Omega H\Omega\tilde{H}\Omega + \Omega H\tilde{H}\Omega^2 + \Omega H\Omega^2\tilde{H} + \tilde{H}H\Omega^3) \\
 &\quad + H\tilde{H}\Omega^3 + H\Omega\tilde{H}\Omega^2 + H\Omega^2\tilde{H}\Omega + H\Omega^3\tilde{H}).
 \end{aligned}$$

Whereas the second directional derivative is given by

$$\begin{aligned} & \left( \frac{d^2}{d\Omega^2} \sum_{k=0}^4 \frac{B_k}{k!} \text{ad}_{-\Omega}^k(H) \right) \tilde{H}^2 \\ &= \frac{1}{6} (\tilde{H}^2 H + H \tilde{H}^2 - 2\tilde{H} H \tilde{H}) \\ &\quad - \frac{1}{360} ((\Omega \tilde{H} \Omega \tilde{H} H + \Omega \tilde{H}^2 \Omega H + \tilde{H}^2 \Omega^2 H \\ &\quad + \Omega^2 \tilde{H}^2 H + \tilde{H} \Omega \tilde{H} \Omega H + \tilde{H} \Omega^2 \tilde{H} H) \\ &\quad - 4(\Omega \tilde{H} \Omega H \tilde{H} + \Omega \tilde{H}^2 H \Omega + \tilde{H}^2 \Omega H \Omega + \tilde{H} \Omega^2 H \tilde{H} + \tilde{H} \Omega H \Omega + \Omega^2 \tilde{H} H \tilde{H}) \\ &\quad + 6(\Omega^2 H \tilde{H}^2 + \Omega \tilde{H} H \Omega \tilde{H} + \tilde{H} \Omega H \Omega \tilde{H} + \Omega \tilde{H} H \tilde{H} \Omega + \tilde{H} \Omega H \tilde{H} \Omega + \tilde{H}^2 H \Omega^2) \\ &\quad - 4(\Omega H \Omega \tilde{H}^2 + \Omega H \tilde{H}^2 \Omega + \tilde{H} H \Omega \tilde{H} \Omega + \Omega H \tilde{H} \Omega \tilde{H} + \tilde{H} H \tilde{H} \Omega^2 + \tilde{H} H \Omega^2 \tilde{H}) \\ &\quad + (H \tilde{H} \Omega \tilde{H} \Omega + H \tilde{H}^2 \Omega^2 + H \tilde{H} \Omega^2 \tilde{H} + H \Omega \tilde{H} \Omega \tilde{H} + H \Omega \tilde{H}^2 \Omega + H \Omega^2 \tilde{H}^2)). \end{aligned}$$

Note that evaluating the derivatives at  $\Omega_0 = 0_{n \times n}$  causes many summands to become zero, which makes computing higher summands ( $k > 4$ ) unnecessary.

The needed derivatives for the geometric Itô–Taylor scheme (3.4) can be computed from the formulas above by inserting correspondingly into  $H$  and  $\tilde{H}$  (see B.1 for instructions).

## References

1. Ableidinger, M., Buckwar, E.: Weak stochastic Runge–Kutta Munthe–Kaas methods for finite spin ensembles. *Appl. Numer. Math.* **118**, 50–63 (2017)
2. Bowman, A.W., Azzalini, A.: *Applied Smoothing Techniques for Data Analysis*. Oxford University Press, New York (1997)
3. Burrage, K., Burrage, P.M.: High strong order methods for non-commutative stochastic ordinary differential equation systems and the Magnus formula. *Phys. D* **133**, 34–48 (1999)
4. Celledoni, E., Marthinsen, H., Owren, B.: An introduction to Lie group integrators—basics, new developments and applications. *J. Comput. Phys.* **257**(Part B), 1040–1061 (2014)
5. Chirikjian, G.S.: *Stochastic Models, Information Theory, and Lie Groups, vol. 1: Classical Results and Geometric Methods*. Springer, Boston (2009)
6. Chirikjian, G.S.: *Stochastic Models, Information Theory, and Lie Groups, vol. 2: Analytic Methods and Modern Applications*. Springer, Boston (2011)
7. Crouch, P.E., Grossman, R.: Numerical integration of ordinary differential equations on manifolds. *J. Nonlinear Sci.* **3**, 1–33 (1993)
8. Elworthy, K.D.: *Stochastic Differential Equations on Manifolds, vol. 70*. Cambridge University Press, Cambridge (1982)
9. Hairer, E., Lubich, C., Wanner, G.: *Geometric Numerical Integration. Springer Series in Computational Mathematics, vol. 31, 2nd edn*. Springer, Berlin (2006)
10. Hall, B.C.: *Lie Groups, Lie Algebras, and Representations. Graduate Texts in Mathematics, vol. 222, 2nd edn*. Springer, Heidelberg (2015)
11. Iserles, A., Munthe-Kaas, H.Z., Nørsett, S.P., Zanna, A.: Lie group methods. *Acta Numer.* **9**, 215–365 (2005)
12. Kamm, K., Pagliarani S., Pascucci, A.: On the stochastic Magnus expansion and its application to SPDEs. arXiv preprint [arXiv:2001.01098](https://arxiv.org/abs/2001.01098) (2020)
13. Kloeden, P.E., Platen, E.: *Numerical Solution of Stochastic Differential Equations*. Springer, Berlin (1992)



14. Lim, N., Privault, N.: Analytic bond pricing for short rate dynamics evolving on matrix Lie groups. *Quant. Finance* **16**(1), 119–129 (2016)
15. Magnus, W.: On the exponential solution of differential equations for a linear operator. *Commun. Pure Appl. Math.* **7**, 649–673 (1954)
16. Malham, S.J.A., Wiese, A.: Stochastic Lie group integrators. *SIAM J. Sci. Comput.* **30**(2), 597–617 (2008)
17. Malham, S.J.A., Wiese, A.: Efficient almost-exact Lévy area sampling. *Stat. Probab. Lett.* **88**, 50–55 (2014)
18. Marjanovic, G., Piggott, M.J., Solo, V.: A simple approach to numerical methods for stochastic differential equations in Lie groups. In: *Proceedings of the 54th IEEE Conference on Decision and Control, IEEE, Osaka, Japan*, pp. 7143–7150 (2015)
19. Marjanovic, G., Solo, V.: Numerical methods for stochastic differential equations in matrix Lie groups made simple. *IEEE Trans. Autom. Control* **63**(12), 4035–4050 (2018)
20. Marsden, J.E., Ratiu, T.S.: *Introduction to Mechanics and Symmetry*, 2nd edn. Springer, New York (1999)
21. Misawa, T.: A Lie algebraic approach to numerical integration of stochastic differential equations. *SIAM J. Sci. Comput.* **23**(3), 866–890 (2001)
22. Muniz, M., Ehrhardt, M., Günther, M.: Approximating correlation matrices using stochastic Lie group methods. *Mathematics* **9**(1), 94 (2021)
23. Munthe-Kaas, H.: Runge–Kutta methods on Lie groups. *BIT Numer. Math.* **38**(1), 92–111 (1998)
24. Munthe-Kaas, H.: High order Runge–Kutta methods on manifolds. *Appl. Numer. Math.* **29**, 115–127 (1999)
25. Park, F.C., Chun, C.M., Han, C.W., Webber, N.: Interest rate models on Lie groups. *Quant. Finance* **11**(4), 559–572 (2010)
26. Piggott, M.J., Solo, V.: Geometric Euler–Maruyama schemes for stochastic differential equations in  $SO(n)$  and  $SE(n)$ . *SIAM J. Numer. Anal.* **54**(4), 2490–2516 (2016)
27. Rößler, A.: Explicit order 1.5 schemes for the strong approximation of Itô stochastic differential equations. *PAMM* **5**(1), 817–818 (2005)
28. Soatto, S., Perona, P., Frezza R., Picci, G.: Motion estimation via dynamic vision. In: *Proceedings of the 33rd IEEE Conference on Decision and Control*, vol.4, pp. 3253–3258. IEEE, Piscataway, NJ (1994)
29. Teng, L., Wu, X., Günther, M., Ehrhardt, M.: A new methodology to create valid time-dependent correlation matrices via isospectral flows. *ESAIM Math. Model. Numer. Anal.* **54**(2), 361–371 (2020)
30. Wang, Z., Ma, Q., Yao, Z., Ding, X.: The magnus expansion for stochastic differential equations. *J. Nonlinear Sci.* **30**, 419–447 (2020)
31. Zuyev, A., Vasylieva, I.: Partial stabilization of stochastic systems with application to rotating rigid bodies. *IFAC-PapersOnLine* **52**(16), 162–167 (2019)

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