Graduate thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

Tor Ola Ousdal Solheim

Composition Methods for Stochastic Differential Equations

Graduate thesis in MTFYMA
Supervisor: Anne Kværnø
June 2022

NTNU
Norwegian University of
Science and Technology

Tor Ola Ousdal Solheim

Composition Methods for Stochastic Differential Equations

Graduate thesis in MTFYMA
Supervisor: Anne Kværnø
June 2022

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

Composition methods for stochastic differential equations

Master's Thesis

Tor Ola Solheim

Spring 2022

Abstract

In this Master's thesis, we investigate splitting- and composition methods for solving semi-linear SDEs with additive noise, where the drift satisfies a global one-sided Lipschitz condition, is allowed to grow polynomially at infinity and satisfies a kind of dissipativity condition. We prove that these methods are mean-square convergent of order $p = 1$ and that they can preserve important structural properties of the SDE, such as geometric ergodicity and hypoellipticity. It appears that the composition methods cannot preserve oscillatory dynamics in their entirety, while the splitting methods can. We demonstrate these results by applying our methods to a cubic model problem and the stochastic FitzHugh-Nagumo (FHN) model, and compare our splitting- and composition methods with the drift-implicit Euler-Maruyama method.

Contents

1 Introduction

In this thesis, we investigate the use of splitting- and composition methods for solving semi-linear stochastic differential equations (SDE) with additive noise:

$$d\boldsymbol{X}(t) = \Big(\mathbf{A}\boldsymbol{X}(t) + \boldsymbol{B}\big(\boldsymbol{X}(t)\big)\Big)dt + \boldsymbol{\Sigma} d\boldsymbol{W}(t), \quad t \geq 0, \quad \boldsymbol{X}(0) = \boldsymbol{X}_0 \in \mathbb{R}^d.$$

Here, the drift $F\big(\boldsymbol{X}(t)\big) := \mathbf{A}\boldsymbol{X}(t) + \boldsymbol{B}\big(\boldsymbol{X}(t)\big)$ is semi-linear, i.e. it consists of a linear part $\mathbf{A}\boldsymbol{X}(t)$ and a nonlinear part $\boldsymbol{B}\big(\boldsymbol{X}(t)\big)$. The drift is locally Lipschitz, and is allowed to grow polynomially at infinity. Such SDEs typically do not admit exact solutions, and the solution must be approximated by some numerical method.

In [1], *splitting methods* were investigated as numerical methods for solving these types of SDEs. Their approach was based on splitting the semi-linear SDE into a nonlinear ODE and a linear SDE:

$$\begin{aligned} d\boldsymbol{X}^{(1)}(t) &= \boldsymbol{B}\big(\boldsymbol{X}^{(1)}(t)\big)dt, \quad t \geq 0, \quad \boldsymbol{X}^{(1)}(0) = \boldsymbol{X}_0^{(1)}, \\ d\boldsymbol{X}^{(2)}(t) &= \mathbf{A}\boldsymbol{X}^{(2)}(t)dt + \boldsymbol{\Sigma} d\boldsymbol{W}(t), \quad t \geq 0, \quad \boldsymbol{X}^{(2)}(0) = \boldsymbol{X}_0^{(2)}. \end{aligned}$$

These two sub-equations were then solved exactly, and their exact solutions were composed to yield an approximation $\widetilde{\boldsymbol{X}}(t_n)$ of the overall solution of the SDE. In particular, two approaches of composing the exact solutions of the two sub-equations were considered: The *Strang* approach and the *Lie-Trotter* approach. These two approaches then gives rise to the Lie-Trotter- and Strang splitting methods [2]. It was shown that these splitting methods were mean-square convergent of order 1, meaning that the numerical solution $\widetilde{\boldsymbol{X}}(t_n)$ and the exact solution $\boldsymbol{X}(t_n)$ satisfies

$$\max_{0 \leq t_n \leq T} \sqrt{\mathbb{E}\Big[\Big\| \boldsymbol{X}(t_n) - \widetilde{\boldsymbol{X}}(t_n) \Big\|^2 \Big]} \leq C h^p,$$

for some $C > 0$ and $p = 1$ [3] (p. XXIV). Here, p is referred to as the order of the method, and $h = t_n - t_{n-1}$ is the step-size. It was also shown that the splitting methods could preserve important structural properties of the SDE, such as hypoellipticity, geometric ergodicity and oscillatory dynamics. Moreover, since the exact solutions to both the linear SDE and nonlinear ODE were assumed to be available, the splitting methods were explicit and thereby computationally efficient.

The purpose of the specialization project [16] and this Master's thesis is to generalize and expand on the work of Buckwar et al. [1]. In particular, we investigate the scenario where the nonlinear ODE is solved approximately by some numerical method. It is often the case for nonlinear ODEs that they do not admit exact, explicit solutions, and so relaxing this assumption would allow for a greater variety of SDEs to be considered within the splitting framework. It is commonly understood within this framework that if all sub-equations are solved exactly, then the method is referred to as a splitting method. If one or more sub-equations are solved approximately, then the method is referred to as a *composition method* [4].

The goal of the specialization project and this thesis was then to investigate if our composition methods converge strongly in the mean-square sense, and if they could preserve the same structural properties as their splitting method counterparts. We also wanted to relax an assumption on the linear term $\mathbf{A}\boldsymbol{X}(t)$; in [1], it was assumed that $\left\| e^{\mathbf{A}h} \right\| < 1$. This assumption was mainly used to prove preservation of geometric ergodicity for both splitting methods. If this assumption could be relaxed without sacrificing the preservation of geometric ergodicity, we would have more options in partitioning a given drift term $\boldsymbol{F}\big(\boldsymbol{X}(t)\big)$ into a linear part and a nonlinear part.

In the specialization project, we considered explicit methods to approximate the nonlinear ODE– in particular, the explicit Euler method. We chose the explicit Euler method since this is typically the poorest performing explicit method for solving ODEs, and any results that hold positively for the explicit Euler method is then expected to also hold for more accurate explicit methods. We chose to focus on explicit methods of solving the nonlinear ODE, since we wanted our composition methods to retain the computational efficiency obtained with splitting methods. While mean-square convergence of order $p = 1$ was experimentally observed in the case of an explicit Euler solution of the nonlinear ODE, we were unable to prove this result. The proof of mean-square convergence in [1] required an assumption on linearly bounded growth of the solution to the nonlinear ODE. In order to use the same proof-strategy to prove mean-square convergence of our composition methods, we would need to show that the explicit Euler solution also satisfies the assumption of linearly bounded growth for some step-size $h > 0$. Unfortunately, we were unable to show this for any step-size $h > 0$. We also considered other explicit methods such as the tamed explicit Euler method and higher-order methods in an attempt to regain linearly bounded growth of the solution to the nonlinear ODE. Again, the desired outcome was observed experimentally, but any proof thereof eluded us.

In light of this, we instead turned our focus to composition methods where the nonlinear ODE is solved using an implicit method; specifically the implicit Euler method. In this case, we are able to attain theoretical results on mean-square convergence and on structure preservation. Thus, solving the nonlinear ODE using the implicit Euler method has been the most extensively studied approach in this thesis. Using the implicit Euler method involves, in general, the solution of a system of nonlinear algebraic equations at every time-step. In the best-case scenario, this system of nonlinear algebraic equations admits an explicit solution. If so, our composition method remains explicit, and therefore computationally efficient. In the general case, however, we cannot solve the system of nonlinear algebraic equations explicitly, and the solution must be approximated. This is typically done using some iterative root-finding algorithm, with Newton's method being a standard choice. Unfortunately, such algorithms typically involves repeatedly solving a linear system at every time-step. As a result, our composition method suffers in terms of computational efficiency since it is no longer fully explicit.

This thesis is organized as follows: Section 2 presents the relevant notation and background theory required for this thesis. Section 3 presents the class of SDEs studied in this thesis and investigates its properties. Section 4 develops the numerical methods used in this thesis. Section 5 presents bounds on the solution to the nonlinear ODE, both in the case of an exact solution and in the case of an implicit Euler solution. Section 6 investigates and proves mean-square convergence of our splitting- and composition methods. Section 7 investigates our methods' capabilities for preserving important structural properties of the SDE. Section 8 discusses some practical implementation details, including strategies for speeding up Newton's method. We demonstrate our results by performing numerical experiments on a cubic model problem and the stochastic FitzHugh-Nagumo (FHN) model in sections 9 and 10, respectively. Finally, we report our conclusions in section 11.

2 Preliminaries

This section presents the relevant notation and some required background for the theory presented in this thesis.

2.a Scalars, vectors, matrices and functions

We distinguish between scalars, vectors and matrices by the following typeface conventions:

- *Scalars* are written using italic and light typeface, and may be lower-case or upper-case. Examples: $c, K \in \mathbb{R}$;
- *Vectors* are written using italic and bold typeface, and may be lower-case or upper-case. Examples: $\boldsymbol{b}, \boldsymbol{F} \in \mathbb{R}^d$
- *Matrices* are written using normal (non-italic) and bold typeface, and is restricted to upper-case lettering only. Example: $\mathbf{A} \in \mathbb{R}^{d \times m}$

Let $a \in \mathbb{R}$ be a generic scalar. We denote by $|a|$ the absolute value of a, given by

$$|a| = \begin{cases} a & \text{if } a \geq 0 \\ -a & \text{if } a < 0 \end{cases}$$

Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$ be generic vectors. Then, $x_i \in \mathbb{R}$ denotes the i'th component of $\boldsymbol{x}$ and $\boldsymbol{x}^\top$ denotes the transpose of $\boldsymbol{x}$. In addition, we denote the inner-product of $\boldsymbol{x}$ and $\boldsymbol{y}$ by

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle := x_1 y_1 + x_2 y_2 + ... + x_d y_d = \langle \boldsymbol{y}, \boldsymbol{x} \rangle.$$

Moreover, we denote by $\|\boldsymbol{x}\|$ the Euclidean norm in $\mathbb{R}^d$, given by

$$\|\boldsymbol{x}\| := \sqrt{x_1^2 + x_2^2 + ... + x_d^2} = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}.$$

Let $\mathbf{A} \in \mathbb{R}^{d \times m}$ be a generic matrix. Then, $\mathbf{A}^\top \in \mathbb{R}^{m \times d}$ denotes the transpose of $\mathbf{A}$, and $a_{ij} \in \mathbb{R}$ denotes the scalar element of $\mathbf{A}$ located in row i, column j. The scalar elements of a matrix are, whenever appropriate, expressed using lower-case lettering. If $\mathbf{A}$ is a square matrix, i.e. $\mathbf{A} \in \mathbb{R}^{d \times d}$, we define its trace $\mathrm{Tr}(\mathbf{A})$ as the sum of the diagonal elements of $\mathbf{A}$:

$$\mathrm{Tr}(\mathbf{A}) := \sum_{i=1}^{d} a_{ii}, \quad \mathbf{A} \in \mathbb{R}^{d \times d}.$$

We denote by $\mathbf{0}_d \in \mathbb{R}^d$ the d-dimensional zero vector, whose elements are all equal to zero. Similarly, we denote by $\mathbf{I}_d \in \mathbb{R}^{d \times d}$ the d-by-d-dimensional identity matrix. Note that for $\mathbf{0}_d$ and $\mathbf{I}_d$, we will omit the d-subscript whenever the dimensionality is clear from the context.

We will make repeated use of the matrix exponential $e^{\mathbf{A}t}$. Let $\mathbf{A} \in \mathbb{R}^{d \times d}$ be a square matrix. The matrix exponential $e^{\mathbf{A}t} \in \mathbb{R}^{d \times d}$ is then given by

$$e^{\mathbf{A}t} := \sum_{k=0}^{\infty} \frac{t^k}{k!} \mathbf{A}^k = \mathbf{I} + t\mathbf{A} + \frac{t^2}{2}\mathbf{A}^2 + \mathcal{O}(t^3), \quad t \in \mathbb{R}. \tag{2.1}$$

Let $\mathbf{A} \in \mathbb{R}^{d \times d}$ be a square matrix. Let $\mu = \mu(\mathbf{A})$ be the logarithmic norm of $\mathbf{A}$, defined by

$$\mu = \mu(\mathbf{A}) = \lambda_{\max}\left(\frac{\mathbf{A} + \mathbf{A}^\top}{2}\right), \tag{2.2}$$

where $\lambda_{\max}(\mathbf{M})$ denotes the largest eigenvalue of the matrix $\mathbf{M} \in \mathbb{R}^{d\times d}$. Then, it holds for the matrix exponential that

$$\left\|e^{\mathbf{A}t}\right\| \leq e^{\mu t}. \tag{2.3}$$

Let $\mathbf{G} : \mathbb{R} \mapsto \mathbb{R}^{d\times m}$ be an arbitrary function of $t \in \mathbb{R}$. We say that $\mathbf{G} = \mathbf{G}(t)$ is *of order* p if there exists a constant $c \geq 0$ such that for all $t \in \mathbb{R}$, it holds that

$$\|\mathbf{G}(t)\| \leq c|t|^p.$$

We will frequently write this as $\mathbf{G}(t) = \mathcal{O}(t^p)$.

2.b Stochastic processes

We will here summarize the theory regarding stochastic processes which is relevant for this work. For a more detailed exposition, the reader is referred to e.g. [5] (chapters 2, 3). Let Ω be a given set, and let $\mathcal{F}$ be a family of subsets of Ω with the following properties:

1. $\emptyset \in \mathcal{F}$,
2. $F \in \mathcal{F} \Longrightarrow F^C \in \mathcal{F}, \quad F^C = \Omega\backslash F$,
3. $F_1, F_2, ... \in \mathcal{F} \Longrightarrow F := \bigcup_{i=1}^{\infty} F_i \in \mathcal{F}$.

Then, $\mathcal{F}$ is a σ-algebra defined on Ω, and the tuple $(\Omega, \mathcal{F})$ is referred to as a measurable space. Given any family $\mathcal{U}$ of subsets of Ω, there exists a smallest σ-algebra on Ω containing $\mathcal{U}$, namely the σ-algebra which is the intersection of all σ-algebras containing $\mathcal{U}$. We call this the σ-algebra generated by $\mathcal{U}$. In the special case where $\Omega \subseteq \mathbb{R}^d$ (or some other topological space) and $\mathcal{U}$ is the collection of all open subsets on Ω, the generated σ-algebra is called the Borel σ-algebra, and is denoted $\mathcal{B}(\Omega)$.

A probability measure $\mathbb{P}$ on a measurable space $(\Omega, \mathcal{F})$ is a function $\mathbb{P} : \mathcal{F} \mapsto [0, 1]$ such that

1. $\mathbb{P}(\emptyset) = 0, \quad \mathbb{P}(\Omega) = 1$,
2. If $F_1, F_2, ... \in \mathcal{F}$ and $\{F_i\}_{i=1}^{\infty}$ are disjoint sets (i.e. $F_i \cap F_j = \emptyset$ if $i \neq j$), then

$$\mathbb{P}\Big(\bigcup_{i=1}^{\infty} F_i\Big) = \sum_{i=1}^{\infty} \mathbb{P}(F_i).$$

The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is referred to as a probability space. A probability space is in general not complete, however it can always be made complete. Therefore, whenever we consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ in this work, we will assume that it is a complete probability space. If $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space, then a function $\boldsymbol{Y} : \Omega \mapsto \mathbb{R}^m$ is called $\mathcal{F}$-measurable if

$$\boldsymbol{Y}^{-1}(B) := \{\omega \in \Omega : \boldsymbol{Y}(\omega) \in B\} \in \mathcal{F},$$

for all Borel sets $B \subset \mathbb{R}^m$.

We define a stochastic process as a parameterized collection of random variables $\{\boldsymbol{X}(t)\}_{t\in[\tau,T]}$ for $\tau, T \in \mathbb{R}_{\geq 0}$ with $\tau \leq T$. This collection of random variables are defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The m-dimensional Wiener process $\{\boldsymbol{W}(t)\}_{t\in[\tau,T]}$ (also called Brownian motion) is a stochastic process. We will assume here and throughout that all components $W_i(t)$ of $\boldsymbol{W}(t)$ for $i = 1, ..., m$ are independent. Then, the m-dimensional Wiener process is the stochastic process defined to satisfy the following:

1. $\boldsymbol{W}(0) = \mathbf{0}_m$.

2. $\boldsymbol{W}$ has independent increments; for every $s > 0$, the increments $\boldsymbol{W}(s+t) - \boldsymbol{W}(s)$ are independent from $\boldsymbol{W}(z)$, where $t \geq 0$ and $z \leq s$.

3. The increments $\boldsymbol{W}(s+t) - \boldsymbol{W}(s)$ are normally distributed as

$$\boldsymbol{W}(s+t) - \boldsymbol{W}(s) \sim \mathcal{N}(\mathbf{0}, t\mathbf{I}_m).$$

4. $W_i(t)$ is continuous in t for all $i = 1, ..., m$.

Let $\boldsymbol{W}(t)$ be an m-dimensional Wiener process. Then, we define the filtration $\{\mathcal{F}(t)\}_{t\in[\tau,T]}$ to be the σ-algebra generated by the random variables $\{W_i(s)\}_{1\leq i\leq m, \tau\leq t\leq T}$. Observe that $\mathcal{F}(s) \subset \mathcal{F}(t)$ for $s < t$, i.e. the filtration $\mathcal{F}(t)$ is an increasing family of σ-algebras of subsets of Ω. A process $\boldsymbol{f}(t,\omega) : [\tau, T] \times \Omega \mapsto \mathbb{R}^m$ is called $\mathcal{F}(t)$-adapted if for each $t \in [\tau, T]$, the function $\omega \mapsto \boldsymbol{f}(t,\omega)$ is $\mathcal{F}(t)$-measurable.

Let $\{\boldsymbol{X}(t)\}_{t\in[\tau,T]} \in \mathbb{R}^d$ be a d-dimensional stochastic process on $[\tau, T]$ and let $\{\mathcal{M}(t)\}_{t\in[\tau,T]}$ be the σ-algebra generated by $\{\boldsymbol{X}(s)\}_{s\in[\tau,t]}$. Let $\{\mathcal{F}(t)\}$ be the σ-algebra generated by the m-dimensional Wiener process $\{\boldsymbol{W}(t)\}_{t\in[\tau,T]}$. Then, $\{\boldsymbol{X}(t)\}_{t\in[\tau,T]}$ is a Markov process if

$$\mathbb{E}[\boldsymbol{X}(t)|\mathcal{M}(s)] = \mathbb{E}[\boldsymbol{X}(t)|\mathcal{F}\big(\boldsymbol{X}(s)\big)], \quad s \in [\tau, t].$$

Thus, a Markov process has the property that the future behavior at time $t \geq s \geq \tau$ of the process given what has happened up to time s is the same as the behavior obtained when starting the process at time s; a Markov process "has no memory".

Let $\{\boldsymbol{X}(t)\}_{t\in[\tau,T]} \in \mathbb{R}^d$ be a Markov process. Denoting by $\mathcal{B}(\mathbb{R}^d)$ the Borel σ-algebra on $\mathbb{R}^d$, the process' transition probability is defined as

$$P_{t-\tau}(B, \boldsymbol{x}) := \mathbb{P}\big(\boldsymbol{X}(t) \in B|\boldsymbol{X}(\tau) = \boldsymbol{x}\big), \tag{2.4}$$

where $B \in \mathcal{B}(\mathbb{R}^d)$. The function $P_{t-\tau}(B, \boldsymbol{x})$ gives the probability that the process reaches a Borel set $B \subset \mathbb{R}^d$ at time t, provided that it started in $\boldsymbol{x} \in \mathbb{R}^d$ at time $\tau < t$.

2.c Stochastic integrals

We will in this subsection develop the theory regarding stochastic integration relevant for this work. Let $\boldsymbol{W}(t)$ denote an m-dimensional Wiener process whose components are independent. For a *deterministic* function $\mathbf{G} : [\tau, t] \mapsto \mathbb{R}^{d\times m}$ with $d, m \in \mathbb{N}$, we denote by $\boldsymbol{I} : [\tau, T] \mapsto \mathbb{R}^d$ the stochastic integral defined by

$$\boldsymbol{I}(\tau, t) := \int_\tau^t \mathbf{G}(s) d\boldsymbol{W}(s) \sim \mathcal{N}(\mathbf{0}, \mathbf{C}(t-\tau)), \tag{2.5}$$

where $\tau \leq t \leq T$ and $\boldsymbol{I}(\tau, \tau) := \mathbf{0}$. If $\tau = 0$, we simply write $\boldsymbol{I}(t)$. Since $\mathbf{G}$ is assumed to be a deterministic function, the integral (2.5) represents a particular kind of *Ito integral*, and is sometimes referred to as a *Wiener integral*. For a general Ito-integral, $\mathbf{G}$ need not be a deterministic function. [6]

The expectation of $\boldsymbol{I}(\tau,t)$ is always zero since it is an Ito-integral. Moreover, $\boldsymbol{I}(\tau,t)$ satisfies Ito-isometry [5] (chapter 3):

$$\mathbb{E}\bigg[\Big(\int_\tau^t \mathbf{G}(s)d\boldsymbol{W}(s)\Big)\Big(\int_\tau^t \mathbf{G}(s)d\boldsymbol{W}(s)\Big)^\top\bigg] = \mathbb{E}\Big[\int_\tau^t \mathbf{G}(s)\mathbf{G}^\top(s)ds\Big].$$

Observe that if the stochastic integral $\boldsymbol{I}(\tau,t)$ is given, the covariance matrix $\mathbf{C}(t-\tau)$ of $\boldsymbol{I}(\tau,t)$ may be found directly using Ito-isometry, the fact that the components of the Wiener process are independent and the fact that Ito-integrals are zero in expectation:

$$\begin{aligned}\mathbf{C}(t-\tau) = \mathrm{Cov}[\boldsymbol{I}(\tau,t)] &:= \mathbb{E}\Big[(\boldsymbol{I}(\tau,t)-\mathbb{E}[\boldsymbol{I}(\tau,t)])(\boldsymbol{I}(\tau,t)-\mathbb{E}[\boldsymbol{I}(\tau,t)])^\top\Big] = \mathbb{E}\Big[\boldsymbol{I}(\tau,t)\boldsymbol{I}^\top(\tau,t)\Big] \\ &= \mathbb{E}\bigg[\Big(\int_\tau^t \mathbf{G}(s)d\boldsymbol{W}(s)\Big)\Big(\int_\tau^t \mathbf{G}(s)d\boldsymbol{W}(s)\Big)^\top\bigg] = \int_\tau^t \mathbf{G}(s)\mathbf{G}^\top(s)ds.\end{aligned}$$

Consider the trace of the covariance matrix $\mathbf{C}(t-\tau)$:

$$\mathrm{Tr}\big(\mathbf{C}(t-\tau)\big) = \sum_{i=1}^d \int_\tau^t [\mathbf{G}(s)\mathbf{G}^\top(s)]_{ii}ds = \sum_{i=1}^d \int_\tau^t \sum_{j=1}^m g_{ij}^2(s)ds \tag{2.6}$$

Moreover, consider the mean-square norm of the stochastic integral $\boldsymbol{I}(\tau,t)$:

$$\mathbb{E}\big[\,\|\boldsymbol{I}(\tau,t)\|^2\,\big] = \mathbb{E}\Big[\sum_{i=1}^d I_i^2(\tau,t)\Big] = \mathbb{E}\bigg[\sum_{i=1}^d \Big(\sum_{j=1}^m \int_\tau^t g_{ij}(s)dW_j(s)\Big)^2\bigg]. \tag{2.7}$$

It follows by linearity of expectation and Ito-isometry and that

$$\begin{aligned}\mathbb{E}\bigg[\sum_{i=1}^d \Big(\sum_{j=1}^m \int_\tau^t g_{ij}(s)dW_j(s)\Big)^2\bigg] &= \sum_{i=1}^d \mathbb{E}\bigg[\Big(\sum_{j=1}^m \int_\tau^t g_{ij}(s)dW_j(s)\Big)^2\bigg] \\ &= \sum_{i=1}^d \mathbb{E}\bigg[\Big(\int_\tau^t \sum_{j=1}^m g_{ij}(s)dW_j(s)\Big)^2\bigg] \\ &= \sum_{j=1}^m \int_\tau^t \sum_{j=1}^m g_{ij}^2(s)ds,\end{aligned} \tag{2.8}$$

where the last equality follows from the fact that the components of the Wiener process are mutually independent, such that cross-terms involving $dW_j(s)$ and $dW_k(s)$ for $j \neq k$ vanish in expectation. Thus, by eqs. (2.6), (2.7) and (2.8), it follows that the stochastic integral $\boldsymbol{I}(t-\tau)$ satisfies the following relation:

$$\mathbb{E}\Big[\,\|\boldsymbol{I}(t-\tau)\|^2\,\Big] = \mathrm{Tr}\big(\mathbf{C}(t-\tau)\big). \tag{2.9}$$

For later use, we define the following stochastic integral:

$$\boldsymbol{W}^{(k)}(t) := \int_0^t (t-s)^k d\boldsymbol{W}(s), \quad k = 0, 1, ... \tag{2.10}$$

Observe that $\boldsymbol{W}^{(0)}(t)$ corresponds to the standard Wiener process $\boldsymbol{W}(t)$. Since $\boldsymbol{W}^{(k)}(t)$ is an ito-integral for all $k = 0, 1, ...$, it follows that

$$\boldsymbol{W}^{(k)}(t) \sim \mathcal{N}(\mathbf{0}, \mathbf{C}(t)),$$

where the covariance matrix $\mathbf{C}$ is a diagonal matrix whose entries can be found via Ito-Isometry: Letting $\mathbf{G}(t) = (t-s)^k\mathbf{I}$, we find

$$\mathbf{C}(t) = \mathbb{E}\Big[\int_0^t (t-s)^{2k}\mathbf{I}ds\Big] = \frac{t^{2k+1}}{2k+1}\mathbf{I}.$$

Thus, it follows that

$$\mathbb{E}\Big[\,\Big\|\boldsymbol{W}^{(k)}(t)\Big\|^2\,\Big] = \mathrm{Tr}\big(\mathbf{C}(t)\big) = \mathcal{O}(h^{2k+1}), \quad k = 0, 1, ... \tag{2.11}$$

3 Model and Properties

We will in this section present and develop the model considered in this work. We consider the time-interval $t \in [0, T]$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with filtration $\{\mathcal{F}(t)\}_{t\in[0,T]}$. Moreover, let $\boldsymbol{W}(t)$ be an $\{\mathcal{F}(t)\}$-adapted Wiener process with $m = d$ components defined on $(\Omega, \mathcal{F}, \mathbb{P})$. In the following, we will consider a class of d-dimensional, autonomous, semi-linear stochastic differential equations with additive noise, given below as

$$d\boldsymbol{X}(t) = \boldsymbol{F}\big(\boldsymbol{X}(t)\big)dt + \boldsymbol{\Sigma}d\boldsymbol{W}(t), \quad t \in [0, T], \quad \boldsymbol{X}(0) = \boldsymbol{X}_0 \in \mathbb{R}^d. \tag{3.1}$$

The initial value $\boldsymbol{X}_0$ is an $\mathcal{F}(0)$-measurable random variable which is independent of the Wiener process and has bounded second moment in expectation. The drift $\boldsymbol{F}\big(\boldsymbol{X}(t)\big)$ is semi-linear, and is given by

$$\boldsymbol{F}\big(\boldsymbol{X}(t)\big) = \mathbf{A}\boldsymbol{X}(t) + \boldsymbol{B}\big(\boldsymbol{X}(t)\big). \tag{3.2}$$

Here, $\boldsymbol{X} : [0, T] \mapsto \mathbb{R}^d$, $\boldsymbol{F} : \mathbb{R}^d \mapsto \mathbb{R}^d$, $\mathbf{A} \in \mathbb{R}^{d\times d}$, $\boldsymbol{B} : \mathbb{R}^d \mapsto \mathbb{R}^d$ and $\boldsymbol{\Sigma} \in \mathbb{R}^{d\times d}$. We will assume throughout that the components $\{W_i(t)\}_{i=1}^d$ of the Wiener process are independent.

We suppose that the SDE (3.1) has a unique strong solution, which is regular insofar as it is defined on the entire interval $[0, T]$ such that sample paths do not tend to infinity in finite time. This requires the existence of a stochastic process $\{\boldsymbol{X}(t)\}_{t\in[0,T]}$ which is adapted to the filtration $\{\mathcal{F}(t)\}_{t\in[0,T]}$ and has continuous paths satisfying

$$\boldsymbol{X}(t) = \boldsymbol{X}_0 + \int_0^t \boldsymbol{F}\big(\boldsymbol{X}(s)\big)ds + \int_0^t \boldsymbol{\Sigma}d\boldsymbol{W}(s), \tag{3.3}$$

for all $t \in [0, T]$ $\mathbb{P}$-almost surely. Intuitively, we think of eq. (3.3) as the SDE (3.1) written on integral form. Moreover, $\{\boldsymbol{X}(t)\}_{t\in[0,T]}$ is a Markov process, with transition probability given by eq. (2.4).

As in [1] (Assumption 2.1), we suppose that the drift satisfies a global one-sided Lipschitz condition and is allowed to grow polynomially at infinity. It suffices to place these assumptions on the nonlinear part $\boldsymbol{B}\big(\boldsymbol{X}(t)\big)$ of the drift. We formalize these assumption in Assumptions 3.1 and 3.2, respectively.

Assumption 3.1 *The function $\boldsymbol{B} \in C^1(\mathbb{R}^d)$ is globally one-sided Lipschitz continuous, i.e. there exists a constant $c_1 > 0$ such that*

$$\big\langle \boldsymbol{x}-\boldsymbol{y}, \boldsymbol{B}(\boldsymbol{x})-\boldsymbol{B}(\boldsymbol{y})\big\rangle \leq c_1 \left\|\boldsymbol{x}-\boldsymbol{y}\right\|^2, \quad \forall\ \boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^d.$$

Assumption 3.2 *The function $\boldsymbol{B} \in C^1(\mathbb{R}^d)$ grows at most polynomially, i.e. there exists constants $c_2 > 0$ and $\chi \geq 1$ such that*

$$\left\|\boldsymbol{B}(\boldsymbol{x})-\boldsymbol{B}(\boldsymbol{y})\right\|^2 \leq c_2\Big(1+\left\|\boldsymbol{x}\right\|^{2(\chi-1)}+\left\|\boldsymbol{y}\right\|^{2(\chi-1)}\Big)\left\|\boldsymbol{x}-\boldsymbol{y}\right\|^2, \quad \forall\ \boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^d.$$

Contrary to [1], we also introduce a kind of dissipativity condition on the nonlinear part of the drift $\boldsymbol{B}$. This condition is formalized in Assumption 3.3.

Assumption 3.3 *There exists $K \geq 0$ and $\alpha \in \mathbb{R}$ such that*

$$\big\langle \boldsymbol{x}, \boldsymbol{B}(\boldsymbol{x})\big\rangle + \alpha \left\|\boldsymbol{x}\right\|^2 \leq K, \quad \forall\ \boldsymbol{x} \in \mathbb{R}^d.$$

Note that this condition is a true dissipativity condition only when $K, \alpha > 0$. We introduce Assumption 3.3 at the onset, since it finds applications both in proving mean-square convergence and geometric ergodicity, as we will see in later sections.

Let $\Phi : \mathbb{R}^d \mapsto \mathbb{R}$ be a twice-continuously differentiable function of $\boldsymbol{X}(t)$. We showed in the specialization project [16] using Ito's Lemma that the stochastic process $\big\{\Phi\big(\boldsymbol{X}(t)\big)\big\}_{t\in[0,T]}$ is governed by the SDE

$$d\Phi\big(\boldsymbol{X}(t)\big) = \big(\mathcal{L}^{(0)}\Phi\big)\big(\boldsymbol{X}(t)\big)dt + \sum_{i=1}^{d}\big(\mathcal{L}^{(i)}\Phi\big)\big(\boldsymbol{X}(t)\big)dW_i(t), \quad t \geq 0, \quad \Phi\big(\boldsymbol{X}(0)\big) = \Phi_0, \tag{3.4}$$

where $\mathcal{L}^{(0)}$ and $\mathcal{L}^{(i)}$ are the generators of the SDE (3.1), given below as

$$\mathcal{L}^{(0)} := \sum_{i=1}^{d} F_i \frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i=1} \sigma_{ii}^2 \frac{\partial^2}{\partial x_i^2} \tag{3.5a}$$

$$\mathcal{L}^{(i)} := \sigma_{ii}\frac{\partial}{\partial x_i}, \quad i = 1, ..., d. \tag{3.5b}$$

3.a Noise structure: Ellipticity and hypoellipticity

We proceed with discussing the noise structure of our system. We will assume throughout this work that the diffusion matrix $\boldsymbol{\Sigma}$ is a diagonal matrix whose diagonal elements $\sigma_{ii} \in \mathbb{R}$ are kept constant. Since the diffusion matrix is diagonal, each component $\{X_i(t)\}_{i=1}^d$ of the process $\boldsymbol{X}(t)$ is associated with one and only one component of the Wiener process. As in [1], we will consider two classes of models obtained depending on the noise structure: The first class is called *elliptic* and corresponds to the case when the diffusion matrix $\boldsymbol{\Sigma}$ is of full rank, i.e. $\sigma_{ii} \neq 0$ for all $i = 1, ..., d$. The second class corresponds to the case when $\boldsymbol{\Sigma}$ is a degenerate matrix; in particular, we consider the case when $\sigma_{11} = 0$ and $\sigma_{ii} > 0$ for $i > 1$. Such a degenerate diffusion matrix naturally arises in many applications, and is equivalent to the degenerate diffusion matrix considered in [1] and [8]. In particular, this scenario occurs when transforming a d-dimensional system of second-order SDEs to a $2d$-dimensional system of first-order SDEs. In this second class of models, the first

component of $\{\boldsymbol{X}(t)\}_{t\in[0,T]}$ is called *smooth*, since it is not directly affected by the noise due to its corresponding noise term $\sigma_{11}dW_1(t)$ being equal to zero for all $t \in [0,T]$. By contrast, the remaining $d-1$ components of $\{\boldsymbol{X}(t)\}_{t\in[0,T]}$ are called *rough*, since the noise acts directly on these components. We emphasize that the use of the words *rough* and *smooth* in this work should not be misunderstood to have any connection with the theory of rough paths. [7].

In the scenario of a degenerate diffusion matrix $\boldsymbol{\Sigma}$, the SDE (3.1) is often *hypoelliptic*. This means that the transition probability admits a smooth density, despite the fact that $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top$ is not of full rank. It follows that a sufficient and necessary condition for hypoellipticity of the process $\{\boldsymbol{X}(t)\}_{t\in[0,T]}$ is that at least one of its rough coordinates appear in the first component $F_1\big(\boldsymbol{X}(t)\big)$ of the drift [8]. We summarize these notions in the following assumption:

Assumption 3.4 *Let the diffusion matrix $\boldsymbol{\Sigma} \in \mathbb{R}^{d\times d}$ be a diagonal matrix. In addition, $\boldsymbol{\Sigma}$ satisfies one of the following conditions:*

A.3.2(1) Ellipticity*: The diagonal elements σ_{ii} are nonzero for all $i = 1, ..., d$.*

A.3.2(2) Hypoellipticity*: The first diagonal element σ_{11} equals zero and the remaining diagonal elements $\{\sigma_{ii}\}_{i=2}^d$ are nonzero. Additionally, at least one of the rough components $\{X_i(t)\}_{i=2}^d$ of the process $\{\boldsymbol{X}(t)\}_{t\in[0,T]}$ appears in the first component $F_1\big(\boldsymbol{X}(t)\big)$ of the drift.*

3.b Lyapunov Structure: Geometric ergodicity

The SDE (3.1) is said to be *geometrically ergodic* if the distribution of the Markov process $\{\boldsymbol{X}(t)\}_{t\in[0,T]}$ converges exponentially fast to a unique invariant distribution π satisfying

$$\pi(B) = \int_{\mathbb{R}^d} P_t(B, \boldsymbol{x})\pi(d\boldsymbol{x}) \quad \forall\ B \in \mathcal{B}(\mathbb{R}^d), \quad t \in [0,T].$$

A function $L = L(\boldsymbol{x})$ is a Lyapunov function for the SDE (3.1) if $L(\boldsymbol{x}) \to +\infty$ as $\|\boldsymbol{x}\| \to +\infty$, and there exists constants $\rho, \eta > 0$ such that

$$\big(\mathcal{L}^{(0)}L\big)(\boldsymbol{x}) \le -\rho L(\boldsymbol{x}) + \eta, \quad \forall\ \boldsymbol{x} \in \mathbb{R}^d, \tag{3.6}$$

where $\mathcal{L}^{(0)}$ is the generator from eq. (3.5a). Establishing geometric ergodicity of the SDE (3.1) is, essentially, a matter of finding a Lyapunov function which satisfies eq. (3.6). The process $\{\boldsymbol{X}(t)\}_{t\in[0,T]}$ must also satisfy a minorization condition [9]. It was shown in [9] that if the drift $\boldsymbol{F}\big(\boldsymbol{X}(t)\big)$ satisfies the following dissipativity condition

$$\big\langle \boldsymbol{x}, \boldsymbol{F}(\boldsymbol{x})\big\rangle + c\,\|\boldsymbol{x}\|^2 \le K, \quad \forall\ \boldsymbol{x} \in \mathbb{R}^d, \quad K, c > 0, \tag{3.7}$$

then the function $L : \mathbb{R}^d \mapsto [1, \infty)$, defined by

$$L(\boldsymbol{x}) := 1 + \|\boldsymbol{x}\|^2, \quad \forall\ \boldsymbol{x} \in \mathbb{R}^d, \tag{3.8}$$

is a Lyapunov function for the SDE (3.1). In particular, if the SDE (3.1) is elliptic, condition (3.7) is sufficient to establish geometric ergodicity of the solution of eq. (3.1). If the SDE (3.1) is hypoelliptic, the process is geometrically ergodic if it satisfies (3.7) in addition to satisfying the irreducibility condition $P_t(B, \boldsymbol{x}) > 0$ for all open sets $B \in \mathcal{B}(\mathbb{R}^d)$ and $\boldsymbol{x} \in \mathbb{R}^d$ [1], [9].

We will now show under which conditions Assumption 3.3 implies condition (3.7). We start by proving the following lemma:

Lemma 3.1 *Let $\mathbf{A} \in \mathbb{R}^{d\times d}$ be any square real-valued matrix. Then, it holds that*

$$\langle \boldsymbol{x}, \mathbf{A}\boldsymbol{x}\rangle \leq \mu \left\|\boldsymbol{x}\right\|^2, \quad \forall\ \boldsymbol{x} \in \mathbb{R}^d,$$

where $\mu := \lambda_{\max}\big((\mathbf{A}+\mathbf{A}^\top)/2\big)$ is, by definition, the logarithmic norm of $\mathbf{A}$.

Proof: We start by writing the matrix $\mathbf{A}$ on the form

$$\mathbf{A} = \frac{\mathbf{A}+\mathbf{A}^\top}{2} + \frac{\mathbf{A}-\mathbf{A}^\top}{2}. \tag{3.9}$$

It may easily be shown that $\big(\mathbf{A}+\mathbf{A}^\top\big)/2$ and $\big(\mathbf{A}-\mathbf{A}^\top\big)/2$ are symmetric- and skew-symmetric matrices, respectively. It follows, then, that

$$\langle \boldsymbol{x}, \mathbf{A}\boldsymbol{x}\rangle = \left\langle \boldsymbol{x}, \left(\frac{\mathbf{A}+\mathbf{A}^\top}{2}\right)\boldsymbol{x}\right\rangle + \left\langle \boldsymbol{x}, \left(\frac{\mathbf{A}-\mathbf{A}^\top}{2}\right)\boldsymbol{x}\right\rangle \tag{3.10}$$

Since $(\mathbf{A}+\mathbf{A}^\top)/2$ is symmetric, it follows that its Rayleigh quotient is bounded from above by its largest eigenvalue, i.e. the logarithmic norm μ of $\mathbf{A}$. This immediately yields

$$\left\langle \boldsymbol{x}, \left(\frac{\mathbf{A}+\mathbf{A}^\top}{2}\right)\boldsymbol{x}\right\rangle \leq \mu \left\|\boldsymbol{x}\right\|^2, \quad \forall\ \boldsymbol{x} \in \mathbb{R}^d. \tag{3.11}$$

Moreover, we find that

$$\left\langle \boldsymbol{x}, \left(\frac{\mathbf{A}-\mathbf{A}^\top}{2}\right)\boldsymbol{x}\right\rangle = \frac{1}{2}\langle \boldsymbol{x}, \mathbf{A}\boldsymbol{x}\rangle - \frac{1}{2}\langle \boldsymbol{x}, \mathbf{A}^\top\boldsymbol{x}\rangle = 0 \quad \forall\ \boldsymbol{x} \in \mathbb{R}^d, \tag{3.12}$$

where the above follows from the fact that

$$\langle \boldsymbol{x}, \mathbf{A}\boldsymbol{x}\rangle = \Big(\langle \boldsymbol{x}, \mathbf{A}\boldsymbol{x}\rangle\Big)^\top = \langle \boldsymbol{x}, \mathbf{A}^\top\boldsymbol{x}\rangle \quad \forall\ \boldsymbol{x} \in \mathbb{R}^d.$$

The result then immediately follows from eqs. (3.10) (3.11) and (3.12). ■

Lemma 3.2 *Let $\boldsymbol{B} \in C^1(\mathbb{R}^d)$ be a d-dimensional vector-valued function satisfying Assumption 3.3 for some $K \geq 0$ and $\alpha \in \mathbb{R}$. Let $\boldsymbol{F}(\boldsymbol{x})$ be the function given by*

$$\boldsymbol{F}(\boldsymbol{x}) = \mathbf{A}\boldsymbol{x} + \boldsymbol{B}(\boldsymbol{x}),$$

for some square real-valued matrix $\mathbf{A} \in \mathbb{R}^{d\times d}$. Then it holds that

$$\langle \boldsymbol{x}, \boldsymbol{F}(\boldsymbol{x})\rangle + c\left\|\boldsymbol{x}\right\|^2 \leq K, \quad \forall\ \boldsymbol{x} \in \mathbb{R}^d,$$

where $c := \alpha - \mu \in \mathbb{R}$ and $\mu \in \mathbb{R}$ is the logarithmic norm of $\mathbf{A}$.

Proof: The result is a direct consequence of Assumption 3.3 and Lemma 3.1:

$$\langle \boldsymbol{x}, \boldsymbol{F}(\boldsymbol{x})\rangle = \langle \boldsymbol{x}, \mathbf{A}\boldsymbol{x} + \boldsymbol{B}(\boldsymbol{x})\rangle = \langle \boldsymbol{x}, \mathbf{A}\boldsymbol{x}\rangle + \langle \boldsymbol{x}, \boldsymbol{B}(\boldsymbol{x})\rangle \leq K - (\alpha-\mu)\left\|\boldsymbol{x}\right\|^2.$$

Letting $c = \alpha - \mu$ and rearranging the above yields the result. ■

Lemma 3.2 tells us that if $K > 0$ and $\mu < \alpha$, then the drift $\boldsymbol{F}$ satisfies condition (3.7) with $c = \alpha - \mu > 0$ and the SDE (3.1) is thereby geometrically ergodic. Note that if $\mu > 0$, then the linear part of the drift is unstable. Conversely, if $\alpha > 0$, then the nonlinear part of the drift is stable. Consider the case of $\mu > 0$. The linear part of the drift is unstable, but if $\mu < \alpha$, then *the nonlinear part of the drift is more stable than the linear part is unstable*, and the drift is therefore stable overall. If, on the other hand, $\alpha < 0$, then the nonlinear part of the drift is unstable. However, if $\mu < \alpha$, then *the linear part of the drift is more stable than the nonlinear part is unstable*. Thus, the condition $\mu < \alpha$, which guarantees geometric ergodicity of the SDE (3.1), is also a statement about the stability of the semi-linear drift $\boldsymbol{F}$. In particular, $\mu < \alpha$ tells us that the instability of the linear (nonlinear) term of the drift is compensated for by the stability of the nonlinear (linear) term of the drift, resulting in a stable semi-linear drift overall.

4 Numerical Methods

Consider a discretization $\{t_n\}_{n=0}^N$ of the time-interval $[0, T]$ with uniform step-size $h := t_n - t_{n-1} \in (0, h_0]$ for some $h_0 > 0$, where $t_n := nh \leq Nh = T$. We denote by $\{\widetilde{\boldsymbol{X}}(t_n)\}_{n=0}^N$ a numerical solution of the SDE (3.1) which approximates the process $\{\boldsymbol{X}(t)\}_{t\in[0,T]}$, where $\widetilde{\boldsymbol{X}}(0) = \boldsymbol{X}(0) = \boldsymbol{X}_0$. We start by defining the notions of *compositions* and *flows*.

Definition 4.1 (Composition) *Let $\boldsymbol{f} : U \mapsto V$ and $\boldsymbol{g} : W \mapsto U$ be two well-defined functions with $U, V, W \subseteq \mathbb{R}^d$. The* composition $(\boldsymbol{f} \circ \boldsymbol{g}) : W \mapsto V$ *of $\boldsymbol{f}$ and $\boldsymbol{g}$ is defined by*

$$(\boldsymbol{f} \circ \boldsymbol{g})(\boldsymbol{x}) = \boldsymbol{f}\big(\boldsymbol{g}(\boldsymbol{x})\big), \quad \forall\ \boldsymbol{x} \in W.$$

Definition 4.2 (Flow) *Let $\boldsymbol{X}(t)$ with initial value $\boldsymbol{X}(0) = \boldsymbol{X}_0 \in \mathbb{R}^d$ be a solution to some initial value problem at any time $t \geq 0$. The* flow $\boldsymbol{\varphi}_t(\boldsymbol{X}_0)$ *is then defined as the mapping from $\boldsymbol{X}_0$ to $\boldsymbol{X}(t)$ for all $t \geq 0$, that is*

$$\boldsymbol{\varphi}_t(\boldsymbol{X}_0) = \boldsymbol{X}(t), \quad \boldsymbol{X}(0) = \boldsymbol{X}_0, \quad \forall\ t \geq 0.$$

The flow $\boldsymbol{\varphi}_t(\boldsymbol{X}_0)$ is furthermore defined to satisfy

$$\begin{aligned} \boldsymbol{\varphi}_0(\boldsymbol{X}_0) &= \boldsymbol{X}_0 \quad \text{(Identity)}, \\ \boldsymbol{\varphi}_s\big(\boldsymbol{\varphi}_t(\boldsymbol{X}_0)\big) &= \boldsymbol{\varphi}_{s+t}(\boldsymbol{X}_0), \quad \forall\ t, s \geq 0 \quad \text{(Group law)}. \end{aligned}$$

4.a Splitting- and composition methods

We start by briefly introducing the concept of splitting methods, which in turn will lead us to the concept of composition methods in a straight-forward manner. For a more detailed description of splitting- and composition methods, see e.g. [2], [4]. We present the concept of splitting methods in the context of ODEs, while keeping in mind that the following exposition is valid also for SDEs. Suppose one wants to solve the initial value problem (IVP) given by

$$d\boldsymbol{X}(t) = \boldsymbol{F}\big(\boldsymbol{X}(t)\big)dt, \quad \boldsymbol{X}(0) = \boldsymbol{X}_0, \quad t \geq 0. \tag{4.1}$$

Suppose further that we can decompose the operator $\boldsymbol{F}$ as

$$\boldsymbol{F} = \sum_{j=1}^{q} \boldsymbol{F}^{(j)}, \quad \boldsymbol{F}^{(j)} : \mathbb{R}^d \mapsto \mathbb{R}^d, \quad j = 1, ..., q. \tag{4.2}$$

Using this split, we may construct a collection of q sub-equations on the form

$$d\boldsymbol{X}^{(j)}(t) = \boldsymbol{F}^{(j)}\big(\boldsymbol{X}^{(j)}(t)\big)dt, \quad \boldsymbol{X}^{(j)}(0) = \boldsymbol{X}_0^{(j)}, \quad t \geq 0, \quad j = 1, ..., q, \tag{4.3}$$

whose *exact* solutions at time t are denoted by the exact flows $\{\boldsymbol{\varphi}_t^{(j)}(\boldsymbol{X}_0^{(j)})\}_{j=1}^q$. The solution to our original IVP (4.1) is then constructed by some composition of the flows $\{\boldsymbol{\varphi}_t^{(j)}(\boldsymbol{X}_0^{(j)})\}_{j=1}^q$ of each sub-equation. This strategy is referred to as a *splitting method*. There are in general many ways to decompose a given operator $\boldsymbol{F}$, and the main idea behind splitting methods is to decompose $\boldsymbol{F}$ such that each resulting sub-equation may be solved exactly. The manner in which we then compose the flows $\{\boldsymbol{\varphi}_t^{(j)}(\boldsymbol{X}_0^{(j)})\}_{j=1}^q$ determines the splitting method. Recall that if one or more of the sub-equations are solved approximately by some numerical method, then our method is referred to as a composition method. In this case, we refer to the numerical solution of this sub-equation by the numerical flow $\widetilde{\boldsymbol{\varphi}}_t^{(j)}(\boldsymbol{X}_0^{(j)})$. We proceed with presenting two approaches for composing the (exact or approximate) solutions to our q sub-equations.

Definition 4.3 (Lie-Trotter- and Strang approach) *Let $\overline{\boldsymbol{\varphi}}_t^{(j)}(\overline{\boldsymbol{X}}_0^{(j)})$ represent either the exact or numerical flow for sub-equation $j = 1, ..., q$ given by eq. (4.3). Let $\overline{\boldsymbol{X}}^{LT}(t_n)$ and $\overline{\boldsymbol{X}}^{S}(t_n)$ be the Lie-Trotter- and Strang approaches for composing the flows of each sub-equation $\overline{\boldsymbol{\varphi}}_t^{(j)}(\overline{\boldsymbol{X}}_0^{(j)})$, yielding an approximation to the initial value problem (4.1). Then, $\overline{\boldsymbol{X}}^{LT}(t_n)$ and $\overline{\boldsymbol{X}}^{S}(t_n)$ are defined by the following compositional schemes:*

$$\begin{aligned}
\overline{\boldsymbol{X}}^{LT}(t) &:= \Big(\overline{\boldsymbol{\varphi}}_t^{(q)} \circ ... \circ \overline{\boldsymbol{\varphi}}_t^{(1)}\Big)(\boldsymbol{X}_0), \\
\boldsymbol{X}^{S}(t) &:= \Big(\overline{\boldsymbol{\varphi}}_{t/2}^{(1)} \circ ... \circ \overline{\boldsymbol{\varphi}}_{t/2}^{(q-1)} \circ \overline{\boldsymbol{\varphi}}_t^{(q)} \circ \overline{\boldsymbol{\varphi}}_{t/2}^{(q-1)} \circ ... \circ \overline{\boldsymbol{\varphi}}_{t/2}^{(1)}\Big)(\boldsymbol{X}_0),
\end{aligned}$$

where $\boldsymbol{X}_0$ is the initial value from eq. (4.1).

We proceed with developing the numerical methods used in this work. As in [1], we propose splitting the nonlinear SDE (3.1) into a nonlinear ODE and a linear SDE with additive noise:

$$d\boldsymbol{X}^{(1)}(t) = \boldsymbol{B}\big(\boldsymbol{X}^{(1)}(t)\big)dt, \quad t \geq 0, \quad \boldsymbol{X}^{(1)}(0) = \boldsymbol{X}_0^{(1)}, \tag{4.4a}$$

$$d\boldsymbol{X}^{(2)}(t) = \mathbf{A}\boldsymbol{X}^{(2)}(t)dt + \boldsymbol{\Sigma}d\boldsymbol{W}(t), \quad t \geq 0, \quad \boldsymbol{X}^{(2)}(0) = \boldsymbol{X}_0^{(2)}. \tag{4.4b}$$

We start by discussing solutions to eq. (4.4a). If the nonlinear ODE of eq. (4.4a) can be solved exactly, then its exact solution is given by

$$\boldsymbol{\varphi}_t^{(1)}(\boldsymbol{X}_0^{(1)}) = \boldsymbol{f}(\boldsymbol{X}_0^{(1)}; t), \tag{4.5}$$

where $\boldsymbol{f} : \mathbb{R}^d \times \mathbb{R}_{\geq 0} \mapsto \mathbb{R}^d$ is a function of the initial value $\boldsymbol{X}_0^{(1)}$ and time t. We may thus simulate eq. (4.4a) exactly at each discrete time-point t_n by

$$\boldsymbol{\varphi}_h^{(1)}\big(\boldsymbol{X}^{(1)}(t_{n-1})\big) = \boldsymbol{f}\big(\boldsymbol{X}^{(1)}(t_{n-1}); h\big), \quad n = 1, ..., N, \quad \boldsymbol{X}^{(1)}(0) = \boldsymbol{X}_0^{(1)} = \boldsymbol{X}_0. \tag{4.6}$$

If eq. (4.4a) cannot be solved exactly, we approximate its solution at each discrete time-point t_n using the implicit Euler method. In this case, the numerical solution is given by

$$\widetilde{\boldsymbol{\varphi}}_h^{(1)}\big(\widetilde{\boldsymbol{X}}^{(1)}(t_{n-1})\big) = \boldsymbol{X}^*, \quad n = 1, 2, ..., N, \quad \widetilde{\boldsymbol{X}}^{(1)}(0) = \boldsymbol{X}_0^{(1)} = \boldsymbol{X}_0 \tag{4.7}$$

where $\boldsymbol{X}^*$ satisfies the implicit equation

$$\boldsymbol{X}^* = \widetilde{\boldsymbol{X}}^{(1)}(t_{n-1}) + h\boldsymbol{B}(\boldsymbol{X}^*). \tag{4.8}$$

Obtaining the solution to the implicit equation (4.8) requires, in general, the solving of a nonlinear system of equations. For certain choices of the nonlinear function $\boldsymbol{B}$, however, the implicit equation (4.8) may be solved such that we obtain an expression for $\boldsymbol{X}^*$ which is a function $\widetilde{\boldsymbol{f}}$ that is explicit in $\widetilde{\boldsymbol{X}}^{(1)}(t_{n-1})$ and h:

$$\widetilde{\boldsymbol{\varphi}}_h^{(1)}\big(\widetilde{\boldsymbol{X}}^{(1)}(t_{n-1})\big) = \boldsymbol{X}^* = \widetilde{\boldsymbol{f}}\big(\boldsymbol{X}^{(1)}(t_{n-1}); h\big) \tag{4.9}$$

This is of course highly desirable in terms of computational efficiency, since our implicit Euler method then behaves like an explicit method while still enjoying the unconditional stability of an implicit method. If we cannot solve the implicit equation (4.8) explicitly for $\boldsymbol{X}^*$, we can approximate it to arbitrary precision using Newton's method. This corresponds to iteratively solving the linear system

$$\Big(\mathbf{I} - h\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{z}^k)\Big)\boldsymbol{z}^{(k+1)} = h\boldsymbol{B}(\boldsymbol{z}^{(k)}) - h\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{z}^{(k)})\boldsymbol{z}^{(k)} - \widetilde{\boldsymbol{X}}^{(1)}(t_{n-1}), \quad \boldsymbol{z}^{(0)} = \widetilde{\boldsymbol{X}}^{(1)}(t_{n-1}), \tag{4.10}$$

for $\boldsymbol{z}^{(k+1)}$ at every discrete time-point t_n, with initial value at time $t_0 = 0$ given by $\widetilde{\boldsymbol{X}}^{(1)}(0) = \boldsymbol{X}_0$. Then, we update $k \leftarrow k+1$ and repeat the procedure until $\big\|\boldsymbol{z}^{(k+1)} - \boldsymbol{z}^{(k)}\big\| < \epsilon$ or until a maximum number of iterations has been reached. Here, $\epsilon > 0$ is the error tolerance of our approximation. The last value for $\boldsymbol{z}^{(k+1)}$ is then taken as the approximation of the solution $\boldsymbol{X}^*$ to the implicit equation (4.8). It is well known that Newton's method exhibits local quadratic convergence for initial values sufficiently close to the true solution. For well-posed problems, we can in practice (by C^1-continuity of $\boldsymbol{B}$) guarantee that our initial value is always sufficiently close to the true solution to yield quadratic convergence by choosing the step-size h small enough. Unfortunately, Newton's method requires repeatedly solving a linear system at each time-step, which negatively affects the computational efficiency of our composition methods. We discuss in section 8 two strategies for mitigating this issue.

We proceed by discussing the solution of eq. (4.4b). This linear SDE is linear in the narrow sense, and it can easily be shown (see e.g. [3], chapter 4.2) that its exact solution is given by

$$\boldsymbol{\varphi}_t^{(2)} = e^{\mathbf{A}t}\boldsymbol{X}_0^{(1)} + \boldsymbol{Z}(t), \tag{4.11}$$

where the random variable $\boldsymbol{Z}(t)$ is given by

$$\boldsymbol{Z}(t) := \int_0^t e^{\mathbf{A}(t-s)}\boldsymbol{\Sigma} d\boldsymbol{W}(s) \sim \mathcal{N}\Big(\mathbf{0}, \mathbf{C}(t)\Big). \tag{4.12}$$

We remark that $\boldsymbol{Z}(t)$ is normally distributed with mean zero since it is an Ito integral. Combined with the fact that the Wiener processes are independent, we find its covariance matrix $\mathbf{C}(t)$ via Ito-isometry:

$$\begin{aligned}\mathbf{C}(t) = \mathbb{E}[\boldsymbol{Z}(t)\boldsymbol{Z}^\top(t)] &= \mathbb{E}\left[\left(\int_0^t e^{\mathbf{A}(t-s)}\boldsymbol{\Sigma} d\boldsymbol{W}(s)\right)\left(\int_0^t e^{\mathbf{A}(t-s)}\boldsymbol{\Sigma} d\boldsymbol{W}(s)\right)^\top\right] \\ &= \int_0^t e^{\mathbf{A}(t-s)}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top e^{\mathbf{A}^\top(t-s)} ds = \mathbf{L}(t)\mathbf{L}^\top(t),\end{aligned} \tag{4.13}$$

where $\mathbf{L}(t)$ is the lower-triangular matrix obtained by a Cholesky factorization of $\mathbf{C}(t)$. We can easily simulate the random variable $\boldsymbol{Z}(t)$ by sampling a standard Gaussian $\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, and then constructing $\boldsymbol{Z}(t)$ by $\boldsymbol{Z}(t) = \mathbf{L}(t)\boldsymbol{\zeta}$. We note that when $\boldsymbol{\Sigma}$ is of full rank (i.e. the elliptic case), $\mathbf{C}(t)$ is symmetric and positive-definite (SPD), whereas $\mathbf{C}(t)$ is symmetric and positive semi-definite (SPSD) in the hypoelliptic case. In either case, a Cholesky factorization exists, however in the hypoelliptic case, such a factorization need not be unique.

Thus, we may simulate eq. (4.4b) exactly at the discrete time-points t_n by

$$\boldsymbol{\varphi}_h^{(2)}\big(\boldsymbol{X}^{(2)}(t_{n-1})\big) = e^{\mathbf{A}h}\boldsymbol{X}^{(2)}(t_{n-1}) + \boldsymbol{Z}_{n-1}(h), \quad n = 1, 2, ..., N \tag{4.14}$$

where $\{\boldsymbol{Z}_{n-1}(h)\}_{n=1}^N$ are independent and identically distributed (i.d.d.) random variables given by eq. (4.12) with covariance matrix $\mathbf{C}(h)$ for each $n = 1, 2, ..., N$. Since $\{\boldsymbol{Z}_{n-1}(h)\}_{n=1}^N$ are i.d.d. random variables, we only need to compute $\mathbf{C}(h)$ *once* prior to sampling $\{\boldsymbol{Z}_{n-1}(h)\}_{n=1}^N$. For some systems, we may (often quite laboriously) find closed-form expressions for $\mathbf{C}(h)$, however we may in general always approximate it to arbitrary precision (limited by machine precision, of course) using numerical integration methods, e.g. the trapezoidal rule. To avoid having to derive explicit expressions for $\mathbf{C}(h)$ for each individual problem, we approximate it using the trapezoidal rule in all our implementations.

Having discussed the solutions to each sub-equation (4.4a)-(4.4b), the Lie-Trotter- and Strang approach for composing these solutions are then given by

$$\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_n) = e^{\mathbf{A}h}\overline{\boldsymbol{\varphi}}_h^{(1)}\big(\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_{n-1})\big) + \boldsymbol{Z}_{n-1}(h), \quad n = 1, ..., N, \quad \overline{\boldsymbol{X}}^{\mathrm{LT}}(0) = \boldsymbol{X}_0, \tag{4.15a}$$

$$\overline{\boldsymbol{X}}^{\mathrm{S}}(t_n) = \overline{\boldsymbol{\varphi}}_{h/2}^{(1)}\bigg(e^{\mathbf{A}h}\overline{\boldsymbol{\varphi}}_{h/2}^{(1)}\big(\overline{\boldsymbol{X}}^{\mathrm{S}}(t_{n-1})\big) + \boldsymbol{Z}_{n-1}(h)\bigg), \quad n = 1, ..., N, \quad \overline{\boldsymbol{X}}^{\mathrm{S}}(0) = \boldsymbol{X}_0. \tag{4.15b}$$

Here, $\overline{\boldsymbol{\varphi}}_h^{(1)}$ is a generalized flow, representing both the exact flow $\boldsymbol{\varphi}_h^{(1)}$ and the numerical flow $\widetilde{\boldsymbol{\varphi}}_h^{(1)}$. These two compositional approaches then give rise to two splitting methods and two composition methods, depending on which of the flows $\boldsymbol{\varphi}_h^{(1)}$ or $\widetilde{\boldsymbol{\varphi}}_h^{(1)}$ of eq. (4.4a) is used in eqs. (4.15a) and (4.15b) Specifically, using the flow $\boldsymbol{\varphi}_h^{(1)}$ gives rise to the Lie-Trotter- and Strang splitting methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$. Using the flow $\widetilde{\boldsymbol{\varphi}}_h^{(1)}$ gives rise to the Lie-Trotter Implicit Euler (LTIE) and Strang Implicit Euler (SIE) composition methods $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$.

We emphasize that eqs. (4.15a)-(4.15b) represent the compositional approaches of our methods for one iteration in time. We see that both approaches involve solving multiple sub-equations in succession at every time-step. The first sub-equation to be solved uses the overall solution from the previous time-step as its initial condition. Then, the solution (flow) of one sub-equation is passed as the initial condition to the next in accordance with either eq. (4.15a) or (4.15b). The solution to the final sub-equation is then taken as the overall solution of the method for the current time-step, and the process repeats for the next time-step. Lastly, we note that the Strang approach is expected to be more accurate than the Lie-Trotter approach. This is due to the fact that the Strang approach is equivalent to first performing the Lie-Trotter approach over a half time-step, then performing it again over another half-time step, but reversing the order in which we compose the flows. As such, the Strang approach is symmetric and uses fractional steps, whereas the Lie-Trotter approach does

not. We conclude this section by writing out all four methods:

$$\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n) := e^{\mathbf{A}h}\boldsymbol{\varphi}_h^{(1)}\big(\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_{n-1})\big) + \boldsymbol{Z}_{n-1}(h), \quad \widetilde{\boldsymbol{X}}^{\mathrm{LT}}(0) = \boldsymbol{X}_0, \tag{4.16a}$$

$$\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n) := \boldsymbol{\varphi}_{h/2}^{(1)}\bigg(e^{\mathbf{A}h}\boldsymbol{\varphi}_{h/2}^{(1)}\big(\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_{n-1})\big) + \boldsymbol{Z}_{n-1}(h)\bigg), \quad \widetilde{\boldsymbol{X}}^{\mathrm{S}}(0) = \boldsymbol{X}_0, \tag{4.16b}$$

$$\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n) := e^{\mathbf{A}h}\widetilde{\boldsymbol{\varphi}}_h^{(1)}\big(\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_{n-1})\big) + \boldsymbol{Z}_{n-1}(h), \quad \widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(0) = \boldsymbol{X}_0, \tag{4.16c}$$

$$\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n) := \widetilde{\boldsymbol{\varphi}}_{h/2}^{(1)}\bigg(e^{\mathbf{A}h}\widetilde{\boldsymbol{\varphi}}_{h/2}^{(1)}\big(\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_{n-1})\big) + \boldsymbol{Z}_{n-1}(h)\bigg), \quad \widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(0) = \boldsymbol{X}_0. \tag{4.16d}$$

4.b Drift-Implicit Euler-Maruyama method

In our numerical experiments, we will frequently compare our methods with the Drift-Implicit Euler-Maruyama method $\widetilde{\boldsymbol{X}}^{\mathrm{DIEM}}(t_n)$, defined below as

$$\widetilde{\boldsymbol{X}}^{\mathrm{DIEM}}(t_n) = \widetilde{\boldsymbol{X}}^{\mathrm{DIEM}}(t_{n-1}) + h\boldsymbol{F}\big(\widetilde{\boldsymbol{X}}^{\mathrm{DIEM}}(t_n)\big) + \sqrt{h}\boldsymbol{\Sigma}\boldsymbol{\zeta}_n, \quad \boldsymbol{\zeta}_{n-1} \sim \mathcal{N}\big(\mathbf{0}, \mathbf{I}_d\big), \tag{4.17}$$

with initial condition $\widetilde{\boldsymbol{X}}^{\mathrm{DIEM}}(0) = \boldsymbol{X}_0$. For SDEs with additive noise, it is known that the DIEM method is mean-square convergent of order $p = 1$ [3] (chapter 12). The DIEM method represents a standard implicit Euler-Maruyama-type method for solving SDEs with locally Lipschitz drift, and as such, we use it to benchmark our splitting- and composition methods of eqs. (4.16a)-(4.16d). Since this is an implicit method, we use Newton's method to approximate the solution at every time-step.

5 Nonlinear ODE: Bounds on solutions

Before proving mean-square convergence of our considered methods, we will derive bounds on the exact- and implicit Euler solutions to the nonlinear ODE (4.4a). These bounds are required in our proof of mean-square boundedness, a central ingredient of mean-square convergence. Moreover, these bounds will help establish our results on geometric ergodicity, and will be used for deriving asymptotic bounds in the limit $t_n \to +\infty$. The key to obtaining these bounds lies in Assumption 3.3.

Consider a nonlinear ODE of the form

$$\frac{d\boldsymbol{X}(t)}{dt} = \boldsymbol{B}\big(\boldsymbol{X}(t)\big), \quad t \in [t_{n-1}, t_n], \quad \boldsymbol{X}(t_{n-1}) = \boldsymbol{x} \in \mathbb{R}^d, \tag{5.1}$$

where $h = t_n - t_{n-1}$. Let $\boldsymbol{\varphi}_h(\boldsymbol{x}) = \boldsymbol{X}(t_n)$ denote the exact flow of eq. (5.1). Moreover, let $\widetilde{\boldsymbol{\varphi}}_h(\boldsymbol{x}) = \widetilde{\boldsymbol{X}}(t_n)$ denote the numerical flow when eq. (5.1) is solved using the implicit Euler method. The Lipschitz condition from Assumption 3.1 guarantees existence and uniqueness of both solutions by Picard-Lindelöf's theorem. Moreover, as the next two theorems establish, Assumption 3.3 determines the bounds on the solutions. We begin with the exact case.

Theorem 5.1 *Let $\boldsymbol{\varphi}_h(\boldsymbol{x})$ denote the exact flow of eq. (5.1) at time t_n with initial value $\boldsymbol{X}(t_{n-1}) = \boldsymbol{x} \in \mathbb{R}^d$. Suppose $\boldsymbol{B}$ satisfies Assumptions 3.1 and 3.3. Then, there exists $h_0 > 0$ such that for all*

$h \in (0, h_0]$, the exact flow $\boldsymbol{\varphi}_h(\boldsymbol{x})$ satisfies

$$0 \leq \|\boldsymbol{\varphi}_h(\boldsymbol{x})\|^2 \leq \begin{cases} e^{-2\alpha h}\|\boldsymbol{x}\|^2 + \frac{K}{\alpha}\Big(1 - e^{-2\alpha h}\Big), & \alpha \neq 0, \\ \|\boldsymbol{x}\|^2 + 2Kh, & \alpha = 0, \end{cases}$$

where $K \geq 0$ and $\alpha \in \mathbb{R}$ are the constants from Assumption 3.3.

Proof: The key starting point is the following relation:

$$\frac{d\,\|\boldsymbol{X}(t)\|^2}{dt} = 2\Big\langle \boldsymbol{X}(t), \boldsymbol{B}\big(\boldsymbol{X}(t)\big)\Big\rangle, \tag{5.2}$$

which follows directly via the chain rule and holds for all $\boldsymbol{X}(t) \in \mathbb{R}^d$. We will first consider the case of $\alpha \neq 0$: It follows from Assumption 3.3 that

$$2\Big\langle \boldsymbol{X}(t), \boldsymbol{B}\big(\boldsymbol{X}(t)\big)\Big\rangle \leq 2K - 2\alpha\,\|\boldsymbol{X}(t)\|^2. \tag{5.3}$$

Inserting for the above into eq. (5.2) and using $v(t) := \|\boldsymbol{X}(t)\|^2$, we obtain the following differential inequality:

$$v'(t) \leq -2\alpha v(t) + 2K, \quad \forall\ t \in [t_{n-1}, t_n], \tag{5.4}$$

where $v'(t) = \frac{dv(t)}{dt}$. Consider now the auxiliary function given by

$$u(t) := e^{2\alpha t}\bigg(v(t) - \frac{K}{\alpha}\bigg).$$

It follows by the product rule that

$$\begin{aligned} u'(t) &= \frac{d}{dt}\left[e^{2\alpha t}\bigg(v(t) - \frac{K}{\alpha}\bigg)\right] \\ &= e^{2\alpha t}v'(t) + 2\alpha e^{2\alpha t}\bigg(v(t) - \frac{K}{\alpha}\bigg) \\ &= e^{2\alpha t}\Big(v'(t) + 2\alpha v(t) - 2K\Big) \end{aligned}$$

Observe from the differential inequality (5.4) that

$$v'(t) + 2\alpha v(t) - 2K \leq 0, \quad \forall\ t \in [t_{n-1}, t_n].$$

Since $e^{2\alpha t} > 0$ for all $t \in [t_{n-1}, t_n]$, it follows that $u'(t) \leq 0$ for all $t \in [t_{n-1}, t_n]$. That is, $u(t)$ is a monotonically decreasing function over the interval $t \in [t_{n-1}, t_n]$. Hence, it holds that

$$e^{2\alpha t_n}\bigg(v(t_n) - \frac{K}{\alpha}\bigg) \leq e^{2\alpha t_{n-1}}\bigg(v(t_{n-1}) - \frac{K}{\alpha}\bigg).$$

Expressing the above inequality in terms of $v(t_n)$ yields

$$\begin{aligned} v(t_n) &\leq e^{-2\alpha(t_n - t_{n-1})}v(t_{n-1}) + \frac{K}{\alpha}\Big(1 - e^{-2\alpha(t_n - t_{n-1})}\Big) \\ &= e^{-2\alpha h}v(t_{n-1}) + \frac{K}{\alpha}\Big(1 - e^{-2\alpha h}\Big), \end{aligned}$$

where we used the fact that $h = t_n - t_{n-1}$. Inserting for $v(t_n) = \|\boldsymbol{X}(t_n)\|^2 = \|\boldsymbol{\varphi}_h(\boldsymbol{x})\|^2$ and $v(t_{n-1}) = \|\boldsymbol{x}\|^2$ then yields the result for the case of $\alpha \neq 0$.

In the case of $\alpha = 0$, the proof is considerably easier: In this case, eq. (5.3) reduces to

$$2\Big\langle \boldsymbol{X}(t), \boldsymbol{B}\big(\boldsymbol{X}(t)\big)\Big\rangle \leq 2K.$$

Using this, it follows directly by integrating eq. (5.2) with respect to time from t_{n-1} to t_n that

$$\|\boldsymbol{\varphi}_h(\boldsymbol{x})\|^2 = \|\boldsymbol{X}(t_n)\|^2 \leq \|\boldsymbol{x}\|^2 + 2Kh,$$

which is the result when $\alpha = 0$. We also note that the result for $\alpha \neq 0$ reduces to the result for $\alpha = 0$ in the limit, as expected. When $\alpha \to 0^{\pm}$, we get $e^{-2\alpha h} \to 1$. Moreover, it follows that

$$\begin{aligned}\lim_{\alpha\to 0^{\pm}} \frac{K}{\alpha}\Big(1 - e^{-2\alpha h}\Big) &= \lim_{\alpha\to 0^{\pm}} 2Kh\frac{1 - e^{-2\alpha h}}{2\alpha h} \\ &= 2Kh \lim_{\alpha\to 0^{\pm}} \frac{1 - e^{-2\alpha h}}{2\alpha h} \\ &= 2Kh,\end{aligned}$$

where we used the fact that

$$\lim_{\alpha\to 0^{\pm}} \frac{1 - e^{-2\alpha h}}{2\alpha h} = 1.$$

Thus, the bound on $\|\boldsymbol{\varphi}_h(\boldsymbol{x})\|^2$ for $\alpha \neq 0$ reduces to the case of $\alpha = 0$ in the limit $\alpha \to 0^{\pm}$. ■

We also remark that if $\alpha > 0$, then the flow $\boldsymbol{\varphi}_h(\boldsymbol{x})$ is asymptotically bounded in the limit $h \to +\infty$:

$$\lim_{h\to+\infty} \|\boldsymbol{\varphi}_h(\boldsymbol{x})\|^2 \leq \frac{K}{\alpha}. \tag{5.5}$$

We proceed with proving an analogous theorem, where we consider the numerical approximation $\widetilde{\boldsymbol{\varphi}}_h(\boldsymbol{x}) = \boldsymbol{X}^*$ of eq. (5.1), where $\boldsymbol{X}^*$ satisfies the implicit Euler equation

$$\boldsymbol{X}^* = \boldsymbol{x} + h\boldsymbol{B}(\boldsymbol{X}^*). \tag{5.6}$$

To prove the analogue of Theorem 5.1 for the implicit Euler case, we require that the upper bound h_0 on the stepsize h satisfies the bound

$$1 + 2\alpha h_0 > 0. \tag{5.7}$$

We require this condition to ensure positivity of $\|\boldsymbol{X}^*\|^2$. Observe that as $h_0 \to 0^+$, the condition (5.7) may be satisfied for any value of $\alpha \in \mathbb{R}$.

Theorem 5.2 *Let $\widetilde{\boldsymbol{\varphi}}_h(\boldsymbol{x}) = \boldsymbol{X}^* \in \mathbb{R}^d$ be the solution of eq. (5.1) as computed by the implicit Euler method such that $\boldsymbol{X}^*$ satisfies eq. (5.6). Let $\boldsymbol{B}$ satisfy Assumptions 3.1 and 3.3 for some $K \geq 0$ and $\alpha \in \mathbb{R}$. Then, there exists $h_0 > 0$ satisfying $1 + 2\alpha h_0 > 0$ such that for all $h \in (0, h_0]$, the numerical flow $\widetilde{\boldsymbol{\varphi}}_h(\boldsymbol{x})$ satisfies*

$$0 \leq \|\widetilde{\boldsymbol{\varphi}}_h(\boldsymbol{x})\|^2 \leq e^{-2\widetilde{\alpha}_0 h}\,\|\boldsymbol{x}\|^2 + 2Kh,$$

where $\widetilde{\alpha}_0$ is given by

$$\widetilde{\alpha}_0 := \frac{1}{2h_0}\log\left(1 + 2\alpha h_0\right) \leq \alpha.$$

Proof: Taking the norm-squared of both sides of eq. (5.6) yields

$$\begin{aligned}\|\boldsymbol{X}^*\|^2 &= \|\boldsymbol{x}\|^2 + 2h\Big\langle \boldsymbol{x}, \boldsymbol{B}(\boldsymbol{X}^*)\Big\rangle + h^2\|\boldsymbol{B}(\boldsymbol{X}^*)\|^2 \\ &= \|\boldsymbol{x}\|^2 + 2h\Big\langle \boldsymbol{X}^* - h\boldsymbol{B}(\boldsymbol{X}^*), \boldsymbol{B}(\boldsymbol{X}^*)\Big\rangle + h^2\|\boldsymbol{B}(\boldsymbol{X}^*)\|^2 \\ &= \|\boldsymbol{x}\|^2 + 2h\Big\langle \boldsymbol{X}^*, \boldsymbol{B}(\boldsymbol{X}^*)\Big\rangle - h^2\|\boldsymbol{B}(\boldsymbol{X}^*)\|^2 \\ &\leq \|\boldsymbol{x}\|^2 + 2h\Big(K - \alpha\|\boldsymbol{X}^*\|^2\Big),\end{aligned}$$

where the last inequality follows from Assumption 3.3 and from the fact that $-h^2\|\boldsymbol{B}(\boldsymbol{X}^*)\|^2 \leq 0$ for all $\boldsymbol{X}^* \in \mathbb{R}^d$ and all $h > 0$. Rearranging and solving for $\|\boldsymbol{X}^*\|^2$ then yields

$$\|\boldsymbol{X}^*\|^2 \leq \frac{1}{1+2\alpha h}\|\boldsymbol{x}\|^2 + 2Kh. \tag{5.8}$$

Observe at this point that if $1 + 2\alpha h_0 > 0$ holds, then it holds that $1 + 2\alpha h > 0$ for all $h \in (0, h_0]$: First consider the case of $\alpha \geq 0$. Then, $1 + 2\alpha h > 0$ holds trivially for any $h > 0$. Second, consider the case $\alpha = -|\alpha| < 0$. In this case, it follows from $1 + 2\alpha h = 1 - 2|\alpha|h_0 > 0$ and $h \leq h_0$ that

$$1 > 2|\alpha|h_0 > 2|\alpha|h_0 > 0,$$

which immediately yields $1 - 2|\alpha|h = 1 + 2\alpha h > 0$ for all $h \in (0, h_0]$. Moreover, consider at this point the function $\widetilde{\alpha} = \widetilde{\alpha}(\alpha, h)$, defined by

$$\widetilde{\alpha}(\alpha, h) := \frac{1}{2h}\log(1 + 2\alpha h).$$

Observe that this function is monotonically decreasing in h such that for all $h \in (0, h_0]$, it holds that

$$\widetilde{\alpha}_0 := \widetilde{\alpha}(\alpha, h_0) = \frac{1}{2h_0}\log 1 + 2\alpha h_0 \leq \frac{1}{2h}\log(1 + 2\alpha h) = \widetilde{\alpha}(\alpha, h).$$

Rearranging the above inequality then yields

$$\frac{1}{1+2\alpha h} \leq e^{-2\widetilde{\alpha}_0 h}, \quad \forall\ h \in (0, h_0]. \tag{5.9}$$

Inserting for eq. (5.9) into eq. (5.8) and recalling that $\widetilde{\boldsymbol{\varphi}}_h(\boldsymbol{x}) = \boldsymbol{X}^*$ then yields the result:

$$0 \leq \|\widetilde{\boldsymbol{\varphi}}_h(\boldsymbol{x})\|^2 \leq e^{-2\widetilde{\alpha}_0 h}\|\boldsymbol{x}\|^2 + 2Kh, \quad \forall h \in (0, h_0],$$

where positivity of $\|\widetilde{\boldsymbol{\varphi}}_h(\boldsymbol{x})\|^2$ is guaranteed by the facts that $K \geq 0$ and $e^{-2\widetilde{\alpha}_0 h} > 0$. ■

6 Mean Square Convergence

We will in this section prove mean-square convergence in finite time of both splitting methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$ and both composition methods $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$. Since the solution to an SDE is a stochastic process, we can never guarantee convergence of any single solution.

However, we can show that the solution converges *in expectation*. For this reason, the issue of convergence is considered in the *mean-square* sense.

In proving mean-square convergence, we rely on Theorem 2.1. from [10], which is an extension of Milstein's Fundamental Theorem [11] for globally Lipschitz functions as considered through Assumptions 3.1 and 3.2. We start with establishing some necessary definitions and stating the main mean-square convergence theorem of this work. We will formulate these definitions and the main mean-square convergence theorem as in [1]. To facilitate this, we introduce the following notation: Let $\boldsymbol{X}_{t_{n-1},\boldsymbol{\xi}}(t_n)$ denote the exact solution of eq. (3.1) at time t_n starting from $\boldsymbol{\xi}$ at time t_{n-1}; that is, $\boldsymbol{X}(t_{n-1}) = \boldsymbol{\xi}$. Similarly, let $\widetilde{\boldsymbol{X}}_{t_{n-1},\boldsymbol{\xi}}(t_n)$ denote the one-step approximation starting from $\boldsymbol{\xi}$ at time t_{n-1} used to construct a numerical solution $\widetilde{\boldsymbol{X}}(t_n)$ of eq. (3.1). As before, we use $\widetilde{\boldsymbol{X}}(t_n)$ as a placeholder for all four methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$. It follows from eqs. (4.16a)-(4.16b) that the one-step approximations for the Lie-Trotter- and Strang splitting methods are given by

$$\boldsymbol{X}^{\mathrm{LT}}_{t_{n-1},\boldsymbol{\xi}}(t_n) = e^{\mathbf{A}h}\boldsymbol{\varphi}^{(1)}_h(\boldsymbol{\xi}) + \int_0^h e^{\mathbf{A}(h-s)}\boldsymbol{\Sigma} d\boldsymbol{W}(s), \tag{6.1a}$$

$$\boldsymbol{X}^{\mathrm{S}}_{t_{n-1},\boldsymbol{\xi}}(t_n) = \boldsymbol{\varphi}^{(1)}_{h/2}\Big(e^{\mathbf{A}h}\boldsymbol{\varphi}^{(1)}_{h/2}(\boldsymbol{\xi}) + \int_0^h e^{\mathbf{A}(h-s)}\boldsymbol{\Sigma} d\boldsymbol{W}(s)\Big). \tag{6.1b}$$

Similarly, it follows from eqs. (4.16c)-(4.16d) that the one-step approximations for the LTIE- and SIE composition methods are given by

$$\boldsymbol{X}^{\mathrm{LTIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n) = e^{\mathbf{A}h}\widetilde{\boldsymbol{\varphi}}^{(1)}_h(\boldsymbol{\xi}) + \int_0^h e^{\mathbf{A}(h-s)}\boldsymbol{\Sigma} d\boldsymbol{W}(s), \tag{6.2a}$$

$$\boldsymbol{X}^{\mathrm{SIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n) = \widetilde{\boldsymbol{\varphi}}^{(1)}_{h/2}\Big(e^{\mathbf{A}h}\widetilde{\boldsymbol{\varphi}}^{(1)}_{h/2}(\boldsymbol{\xi}) + \int_0^h e^{\mathbf{A}(h-s)}\boldsymbol{\Sigma} d\boldsymbol{W}(s)\Big). \tag{6.2b}$$

Definition 6.1 (Mean-square consistency) *The one-step approximation* $\widetilde{\boldsymbol{X}}_{t_{n-1},\boldsymbol{\xi}}(t_n)$ *used to construct a numerical solution* $\widetilde{\boldsymbol{X}}(t_n)$ *of eq.* (3.1) *is* mean-square consistent *of order* $p > 0$ *if there exists* $h_0 > 0$ *such that for arbitrary* $\{t_n\}_{n=1}^N$*, for all* $\boldsymbol{\xi} \in \mathbb{R}^d$ *and for all* $h \in (0, h_0]$*, it has the following orders of accuracy:*

$$\Big\|\mathbb{E}\Big[\boldsymbol{X}_{t_{n-1},\boldsymbol{\xi}}(t_n) - \widetilde{\boldsymbol{X}}_{t_{n-1},\boldsymbol{\xi}}(t_n)\Big]\Big\| = \mathcal{O}(h^{p+1}),$$

$$\sqrt{\mathbb{E}\Big[\,\Big\|\boldsymbol{X}_{t_{n-1},\boldsymbol{\xi}}(t_n) - \widetilde{\boldsymbol{X}}_{t_{n-1},\boldsymbol{\xi}}(t_n)\Big\|^2\Big]} = \mathcal{O}(h^{p+\frac{1}{2}}).$$

Definition 6.2 (Mean-square boundedness) *A numerical solution* $\widetilde{\boldsymbol{X}}(t_n)$ *of eq.* (3.1) *is* mean-square bounded *if there exists* $h_0 > 0$ *and a constant* $\widetilde{K} = \widetilde{K}(h_0, T) > 0$ *such that for all step-sizes* $h \in (0, h_0]$ *it holds that*

$$\max_{t_0 \le t_n \le T} \mathbb{E}\Big[\,\Big\|\widetilde{\boldsymbol{X}}(t_n)\Big\|^2\Big] \le \widetilde{K}(h_0, T)\Big(1 + \mathbb{E}\Big[\,\|\boldsymbol{X}_0\|^2\Big]\Big),$$

where $T \in [0, +\infty)$ *is the end-time.*

Definition 6.3 (Mean-square convergence) *Let $\boldsymbol{X}(t_n)$ denote the exact solution of an SDE on the interval $[0,T]$, and let $\widetilde{\boldsymbol{X}}(t_n)$ denote a numerical solution of this SDE. The numerical solution $\widetilde{\boldsymbol{X}}(t_n)$ is mean-square convergent of order $p>0$ if*

$$\max_{0\leq t_n\leq T}\sqrt{\mathbb{E}\Big[\,\Big\|\boldsymbol{X}(t_n)-\widetilde{\boldsymbol{X}}(t_n)\Big\|^2\,\Big]}=\mathcal{O}(h^p).$$

Theorem 6.1 (Mean-square convergence (Thm. 2.1. [10])) *Let $\widetilde{\boldsymbol{X}}(t_n)$ denote a numerical solution of eq. (3.1) at time t_n starting from $\widetilde{\boldsymbol{X}}(t_0)=\widetilde{\boldsymbol{X}}(0)=\boldsymbol{X}(0)=\boldsymbol{X}_0$, constructed using the one-step approximation $\widetilde{\boldsymbol{X}}_{t_{n-1},\boldsymbol{\xi}}(t_n)$. Furthermore, let Assumptions 3.1, 3.2, 3.3 and 3.4 be satisfied. If*

1. *The one-step approximation $\widetilde{\boldsymbol{X}}_{t_{n-1},\boldsymbol{\xi}}(t_n)$ is* mean-square consistent *of order $p>0$ in the sense of Definition 6.1;*
2. *The numerical solution $\widetilde{\boldsymbol{X}}(t_n)$ is* mean-square bounded *in the sense of Definition 6.2;*

Then the numerical method $\widetilde{\boldsymbol{X}}(t_n)$ is mean-square convergent *of order p in the sense of Definition 6.3.*

It is perhaps clear from the preceding definitions and theorem that proving mean-square convergence of a given method involves two tasks: (1) Proving that the method is mean-square *consistent* and (2) proving that the method is mean-square *bounded*. In [1], mean-square convergence of order $p=1$ was proved for both splitting methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$. This proof, however, did not rely on Assumption 3.3. In this thesis, we prove mean-square convergence of order $p=1$ for both splitting methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$ under Assumption 3.3. We include this assumption, because it will allow us to formulate asymptotic bounds on $\mathbb{E}\big[\,\big\|\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)\big\|^2\,\big]$ and $\mathbb{E}\big[\,\big\|\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)\big\|^2\,\big]$ as $T\to+\infty$ even when $\mu\geq 0$, which could not be done using the theory of [1]. We also show that both composition methods $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$ are mean-square convergent of order $p=1$, also under Assumption 3.3.

Proving mean-square *consistency* does not rely on Assumption 3.3. For this reason, we do not prove mean-square consistency of $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ nor $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$, since we may rely fully on the proof given in [1]. Here, it was shown that both splitting methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$ were mean-square consistent of order $p=1$. Thus, we only prove mean-square consistency for our composition methods $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$, and we prove that they are both mean-square consistent of order $p=1$.

In our proof of mean-square *boundedness*, however, we rely heavily on Assumption 3.3, since this assumption yields bounds on the solution of the nonlinear ODE (4.4a), as shown in section 5. Thus, we provide mean-square boundedness proofs for all four methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$.

The mean-square boundedness proof for $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ are rather similar. For this reason, we show mean-square boundedness for the generalized Lie-Trotter method as defined in eq. (4.15a). Recall that the generalized Lie-Trotter method uses the Lie-Trotter approach with a generalized flow $\overline{\boldsymbol{\varphi}}_h^{(1)}$ of eq. (4.4a), which can represent either the exact flow $\boldsymbol{\varphi}_h^{(1)}$ or the numerical flow $\widetilde{\boldsymbol{\varphi}}_h^{(1)}$. The result then holds for the $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ methods upon inserting for their

respective flows $\boldsymbol{\varphi}_h^{(1)}$ or $\widetilde{\boldsymbol{\varphi}}_h^{(1)}$. Similarly, we prove mean-square boundedness for a generalized Strang method defined through eq. (4.15b). Also here do we obtain the result for either $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$ or $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$ when inserting for the appropriate flow.

Once mean-square consistency of order $p = 1$ has been proved for both composition methods $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$, and mean-square boundedness has been proved for all four methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$, it follows immediately from Theorem 6.1 that all four methods are mean-square convergent of order $p = 1$. We state this as separate theorems towards the end of this section.

6.a Mean-square consistency

We will in this section prove the following Lemma:

Lemma 6.2 (Mean-square consistency) *Let Assumptions 3.1, 3.2 and 3.4 be satisfied, and let $\boldsymbol{X}^{LTIE}_{t_{n-1},\boldsymbol{\xi}}(t_n)$ and $\boldsymbol{X}^{SIE}_{t_{n-1},\boldsymbol{\xi}}(t_n)$ be the one-step approximations of the LTIE- and SIE methods defined through eqs.* (6.2a) *and* (6.2b)*, respectively. Then, $\boldsymbol{X}^{LTIE}_{t_{n-1},\boldsymbol{\xi}}(t_n)$ and $\boldsymbol{X}^{SIE}_{t_{n-1},\boldsymbol{\xi}}(t_n)$ are mean-square consistent of order $p = 1$ in the sense of Definition 6.1.*

Proof: The key to proving Lemma 6.2 lies in applying a stochastic Taylor expansion to $\boldsymbol{X}_{t_{n-1},\boldsymbol{\xi}}(t_n)$ around $\boldsymbol{\xi}$. We then compare this expansion with the corresponding expansions of our one-step approximations $\boldsymbol{X}^{\mathrm{LTIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n)$ and $\boldsymbol{X}^{\mathrm{SIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n)$ to find the local error of each composition method. It was shown in the specialization project preceding this work [16] that the one-step approximation of the exact solution is given by

$$\begin{aligned}\boldsymbol{X}_{t_{n-1},\boldsymbol{\xi}}(t_n) = \boldsymbol{\xi} + h\mathbf{A}\boldsymbol{\xi} + h\boldsymbol{B}(\boldsymbol{\xi}) + \frac{h^2}{2}\mathbf{A}^2\boldsymbol{\xi} + \frac{h^2}{2}\mathbf{A}\boldsymbol{B}(\boldsymbol{\xi}) + \frac{h^2}{2}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\mathbf{A}\boldsymbol{\xi} + \frac{h^2}{2}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{B}(\boldsymbol{\xi}) \\ + \frac{h^2}{2}\boldsymbol{D}(\boldsymbol{\xi}) + \boldsymbol{\Sigma}\boldsymbol{W}^{(0)}_{n-1}(h) + \Big(\mathbf{A}\boldsymbol{\Sigma} + \mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{\Sigma}\Big)\boldsymbol{W}^{(1)}_{n-1}(h) + \mathcal{O}(h^{5/2}).\end{aligned} \tag{6.3}$$

Here, $\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})$ is the Jacobian of $\boldsymbol{B}$ evaluated at $\boldsymbol{\xi}$ whose elements are given by

$$\big[\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\big]_{ij} = \partial B_i(\boldsymbol{\xi})/\partial x_j, \quad i,j = 1,...,d.$$

Moreover, $\boldsymbol{D}(\boldsymbol{\xi})$ is the vector whose i'th element is given by

$$D_i(\boldsymbol{\xi}) = \mathrm{Tr}\Big(\boldsymbol{\Sigma}^2\mathbf{H}_{B_i}(\boldsymbol{\xi})\Big) = \sum_{j=1}^{d}\sigma_{jj}^2\frac{\partial^2 B_i(\boldsymbol{\xi})}{\partial x_j^2}, \quad i = 1,...,d,$$

with $\mathbf{H}_{B_i}(\boldsymbol{\xi})$ being the Hessian of B_i evaluated at $\boldsymbol{\xi}$ and $\mathrm{Tr}(\cdot)$ denoting the trace of a matrix. In addition, $\boldsymbol{W}^{(0)}_{n-1}(h)$ and $\boldsymbol{W}^{(1)}_{n-1}(h)$ are stochastic Ito integrals which, for each $n > 0$, are given by eq. (2.10).

We proceed with expanding the one-step approximations for our LTIE- and SIE composition methods from eqs. (6.2a)-(6.2b). Recall from eq. (4.7) that

$$\widetilde{\boldsymbol{\varphi}}_h^{(1)}(\boldsymbol{\xi}) = \boldsymbol{X}^*,$$

where $\boldsymbol{X}^*$ satisfies the equation

$$\boldsymbol{X}^* = \boldsymbol{\xi} + h\boldsymbol{B}(\boldsymbol{X}^*). \tag{6.4}$$

We start by expanding $\boldsymbol{B}(\boldsymbol{X}^*)$ to the first order around $\boldsymbol{\xi}$:

$$\boldsymbol{B}(\boldsymbol{X}^*) = \boldsymbol{B}(\boldsymbol{\xi}) + \mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})(\boldsymbol{X}^* - \boldsymbol{\xi}) + \mathcal{O}(h^2). \tag{6.5}$$

To see why the last term is $\mathcal{O}(h^2)$, observe that the quadratic term in the expansion of $\boldsymbol{B}(\boldsymbol{X}^*)$ around $\boldsymbol{\xi}$ is given by $\boldsymbol{Q}$, where Q_i satisfies

$$Q_i = \frac{1}{2}\Big(\boldsymbol{X}^* - \boldsymbol{\xi}\Big)^\top \mathbf{H}_{B_i}(\boldsymbol{\xi})\Big(\boldsymbol{X}^* - \boldsymbol{\xi}\Big), \quad i = 1, ..., d.$$

From eq. (6.4) it is clear that $\boldsymbol{X}^* - \boldsymbol{\xi} = h\boldsymbol{B}(\boldsymbol{X}^*) = \mathcal{O}(h)$. It then follows directly from the expression for Q_i that $Q_i = \mathcal{O}(h^2)$ for all $i = 1, ..., d$; hence, $\boldsymbol{Q} = \mathcal{O}(h^2)$, which justifies why the last term in eq. (6.5) is $\mathcal{O}(h^2)$. Inserting for eq. (6.5) into eq. (6.4) and rearranging then yields

$$\boldsymbol{X}^* = \boldsymbol{\xi} + h\Big(\mathbf{I} - h\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\Big)^{-1}\boldsymbol{B}(\boldsymbol{\xi}) + \mathcal{O}(h^3). \tag{6.6}$$

Using the Neumann series, we have the following expansion:

$$\Big(\mathbf{I} - h\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\Big)^{-1} = \sum_{k=0}^{\infty} h^k \mathbf{J}_{\boldsymbol{B}}^k(\boldsymbol{\xi}) = \mathbf{I} + h\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi}) + \mathcal{O}(h^2). \tag{6.7}$$

Inserting for (6.7) back into eq. (6.6) then yields

$$\boldsymbol{X}^* = \boldsymbol{\xi} + h\boldsymbol{B}(\boldsymbol{\xi}) + h^2\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{B}(\boldsymbol{\xi}) + \mathcal{O}(h^3). \tag{6.8}$$

Recalling that $\widetilde{\boldsymbol{\varphi}}_h^{(1)}(\boldsymbol{\xi}) = \boldsymbol{X}^*$, we've found

$$\widetilde{\boldsymbol{\varphi}}_h^{(1)}(\boldsymbol{\xi}) = \boldsymbol{\xi} + h\boldsymbol{B}(\boldsymbol{\xi}) + h^2\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{B}(\boldsymbol{\xi}) + \mathcal{O}(h^3). \tag{6.9}$$

We will now insert for the expansion (6.9) and the expansion of the matrix exponential given by eq. (2.1) into the one-step approximations (6.2a)-(6.2b) of our composition methods. The one-step approximation for the LTIE composition method then reads

$$\begin{aligned}
\boldsymbol{X}^{\text{LTIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n) &= e^{\mathbf{A}h}\widetilde{\boldsymbol{\varphi}}_h^{(1)}(\boldsymbol{\xi}) + \int_0^h e^{\mathbf{A}(h-s)}\boldsymbol{\Sigma}d\boldsymbol{W}(s) \\
&= \left(\mathbf{I} + h\mathbf{A} + \frac{h^2}{2}\mathbf{A}^2\right)\Big(\boldsymbol{\xi} + h\boldsymbol{B}(\boldsymbol{\xi}) + h^2\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{B}(\boldsymbol{\xi})\Big) \\
&\quad + \int_0^h \left(\mathbf{I} + (h-s)\mathbf{A} + \frac{(h-s)^2}{2}\mathbf{A}^2\right)\boldsymbol{\Sigma}d\boldsymbol{W}(s) + \mathcal{O}(h^3) \\
&= \boldsymbol{\xi} + h\boldsymbol{B}(\boldsymbol{\xi}) + h^2\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{B}(\boldsymbol{\xi}) + h\mathbf{A}\boldsymbol{\xi} + h^2\mathbf{A}\boldsymbol{B}(\boldsymbol{\xi}) + \frac{h^2}{2}\mathbf{A}^2\boldsymbol{\xi} + \boldsymbol{\Sigma}\int_0^h d\boldsymbol{W}(s) \\
&\quad + \mathbf{A}\boldsymbol{\Sigma}\int_0^h (h-s)d\boldsymbol{W}(s) + \frac{1}{2}\mathbf{A}^2\boldsymbol{\Sigma}\int_0^h (h-s)^2 d\boldsymbol{W}(s) + \mathcal{O}(h^3) \\
&= \boldsymbol{\xi} + h\boldsymbol{B}(\boldsymbol{\xi}) + h^2\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{B}(\boldsymbol{\xi}) + h\mathbf{A}\boldsymbol{\xi} + h^2\mathbf{A}\boldsymbol{B}(\boldsymbol{\xi}) + \frac{h^2}{2}\mathbf{A}^2\boldsymbol{\xi} \\
&\quad + \boldsymbol{\Sigma}\boldsymbol{W}_{n-1}^{(0)}(h) + \mathbf{A}\boldsymbol{\Sigma}\boldsymbol{W}_{n-1}^{(1)}(h) + \frac{1}{2}\mathbf{A}^2\boldsymbol{\Sigma}\boldsymbol{W}_{n-1}^{(2)}(h) + \mathcal{O}(h^3)
\end{aligned} \tag{6.10}$$

Similarly, we find the one-step approximation for the SIE composition method. Let $\boldsymbol{S}$ be given by

$$\begin{aligned}\boldsymbol{S} &:= e^{\mathbf{A}h}\widetilde{\boldsymbol{\varphi}}^{(1)}_{h/2}(\boldsymbol{\xi}) + \int_0^h e^{\mathbf{A}(h-s)}\boldsymbol{\Sigma} d\boldsymbol{W}(s) \\ &= \left(\mathbf{I} + h\mathbf{A} + \frac{h^2}{2}\mathbf{A}^2\right)\left(\boldsymbol{\xi} + \frac{h}{2}\boldsymbol{B}(\boldsymbol{\xi}) + \frac{h^2}{4}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{B}(\boldsymbol{\xi})\right) \\ &\quad + \int_0^h \left(\mathbf{I} + (h-s)\mathbf{A} + \frac{(h-s)^2}{2}\mathbf{A}^2\right)\boldsymbol{\Sigma} d\boldsymbol{W}(s) + \mathcal{O}(h^3) \\ &= \boldsymbol{\xi} + \frac{h}{2}\boldsymbol{B}(\boldsymbol{\xi}) + \frac{h^2}{4}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{B}(\boldsymbol{\xi}) + h\mathbf{A}\boldsymbol{\xi} + \frac{h^2}{2}\mathbf{A}\boldsymbol{B}(\boldsymbol{\xi}) + \frac{h^2}{2}\mathbf{A}^2\boldsymbol{\xi} \\ &\quad + \boldsymbol{\Sigma}\boldsymbol{W}^{(0)}_{n-1}(h) + \mathbf{A}\boldsymbol{\Sigma}\boldsymbol{W}^{(1)}_{n-1}(h) + \frac{1}{2}\mathbf{A}^2\boldsymbol{\Sigma}\boldsymbol{W}^{(2)}_{n-1}(h) + \mathcal{O}(h^3) \\ &= \boldsymbol{\xi} + \boldsymbol{y},\end{aligned} \tag{6.11}$$

where $\boldsymbol{y}$ is given by

$$\begin{aligned}\boldsymbol{y} &:= \frac{h}{2}\boldsymbol{B}(\boldsymbol{\xi}) + \frac{h^2}{4}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{B}(\boldsymbol{\xi}) + h\mathbf{A}\boldsymbol{\xi} + \frac{h^2}{2}\mathbf{A}\boldsymbol{B}(\boldsymbol{\xi}) + \frac{h^2}{2}\mathbf{A}^2\boldsymbol{\xi} \\ &\quad + \boldsymbol{\Sigma}\boldsymbol{W}^{(0)}_{n-1}(h) + \mathbf{A}\boldsymbol{\Sigma}\boldsymbol{W}^{(1)}_{n-1}(h) + \frac{1}{2}\mathbf{A}^2\boldsymbol{\Sigma}\boldsymbol{W}^{(2)}_{n-1}(h) + \mathcal{O}(h^3).\end{aligned} \tag{6.12}$$

Using eqs. (6.11)-(6.12), we find that the one-step approximation for the SIE method reads

$$\begin{aligned}\boldsymbol{X}^{\mathrm{SIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n) &= \widetilde{\boldsymbol{\varphi}}^{(1)}_{h/2}(\boldsymbol{S}) \\ &= \boldsymbol{S} + \frac{h}{2}\boldsymbol{B}(\boldsymbol{S}) + \frac{h^2}{4}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{S})\boldsymbol{B}(\boldsymbol{S}) + \mathcal{O}(h^3) \\ &= \boldsymbol{\xi} + \boldsymbol{y} + \frac{h}{2}\boldsymbol{B}(\boldsymbol{\xi}+\boldsymbol{y}) + \frac{h^2}{4}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi}+\boldsymbol{y})\boldsymbol{B}(\boldsymbol{\xi}+\boldsymbol{y}) + \mathcal{O}(h^3).\end{aligned} \tag{6.13}$$

Expanding $\boldsymbol{B}(\boldsymbol{\xi}+\boldsymbol{y})$ around $\boldsymbol{\xi}$ yields

$$\boldsymbol{B}(\boldsymbol{\xi}+\boldsymbol{y}) = \boldsymbol{B}(\boldsymbol{\xi}) + \mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{y} + \boldsymbol{G}(\boldsymbol{\xi}) + \mathcal{O}(\|\boldsymbol{y}\|^3),$$

where $G_i(\boldsymbol{\xi})$ is given by

$$G_i = \boldsymbol{y}^\top \mathbf{H}_{B_i}(\boldsymbol{\xi})\boldsymbol{y}, \quad i = 1, ..., d,$$

and $\mathcal{O}(\|\boldsymbol{y}\|^3) = \mathcal{O}(h^{3/2})$ since

$$\mathcal{O}(\|\boldsymbol{y}\|) = \mathcal{O}(\|\boldsymbol{W}_{n-1}(h)\|) = \mathcal{O}(h^{1/2}).$$

Thus, we have that

$$\frac{h}{2}\boldsymbol{B}(\boldsymbol{\xi}+\boldsymbol{y}) = \frac{h}{2}\boldsymbol{B}(\boldsymbol{\xi}) + \frac{h}{2}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{y} + \frac{h}{2}\boldsymbol{G}(\boldsymbol{\xi}) + \mathcal{O}(h^{5/2}), \tag{6.14}$$

and

$$\frac{h^2}{4}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi}+\boldsymbol{y})\boldsymbol{B}(\boldsymbol{\xi}+\boldsymbol{y}) = \frac{h^2}{4}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{B}(\boldsymbol{\xi}) + \mathcal{O}(h^{5/2}). \tag{6.15}$$

Inserting for eqs. (6.14) and (6.15) into eq. (6.13) then yields

$$\boldsymbol{X}^{\mathrm{SIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n) = \boldsymbol{\xi} + \boldsymbol{y} + \frac{h}{2}\boldsymbol{B}(\boldsymbol{\xi}) + \frac{h}{2}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{y} + \frac{h}{2}\boldsymbol{G}(\boldsymbol{\xi}) + \frac{h^2}{4}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{B}(\boldsymbol{\xi}) + \mathcal{O}(h^{5/2}) \tag{6.16}$$

Inserting for $\boldsymbol{y}$ from eq. (6.12), we find that

$$\begin{aligned}\boldsymbol{\xi} + \boldsymbol{y} + \frac{h}{2}\boldsymbol{B}(\boldsymbol{\xi}) = \boldsymbol{\xi} + h\boldsymbol{B}(\boldsymbol{\xi}) + \frac{h^2}{4}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{B}(\boldsymbol{\xi}) + h\mathbf{A}\boldsymbol{\xi} + \frac{h^2}{2}\mathbf{A}\boldsymbol{B}(\boldsymbol{\xi}) + \frac{h^2}{2}\mathbf{A}^2\boldsymbol{\xi} \\ + \boldsymbol{\Sigma}\boldsymbol{W}^{(0)}_{n-1}(h) + \mathbf{A}\boldsymbol{\Sigma}\boldsymbol{W}^{(1)}_{n-1}(h) + \frac{1}{2}\mathbf{A}^2\boldsymbol{\Sigma}\boldsymbol{W}^{(2)}_{n-1}(h) + \mathcal{O}(h^3),\end{aligned} \tag{6.17}$$

and

$$\begin{aligned}\frac{h}{2}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{y} = \frac{h^2}{4}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{B}(\boldsymbol{\xi}) + \frac{h^2}{2}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\mathbf{A}\boldsymbol{\xi} + \frac{h}{2}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{\Sigma}\boldsymbol{W}^{(0)}_{n-1}(h) + \frac{h}{2}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\mathbf{A}\boldsymbol{\Sigma}\boldsymbol{W}^{(1)}_{n-1}(h) \\ + \frac{h}{4}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\mathbf{A}^2\boldsymbol{\Sigma}\boldsymbol{W}^{(2)}_{n-1}(h) + \mathcal{O}(h^3).\end{aligned} \tag{6.18}$$

Finally, inserting for eqs. (6.17) and (6.18) into eq. (6.16) yields

$$\begin{aligned}\boldsymbol{X}^{\mathrm{SIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n) = \boldsymbol{\xi} + h\boldsymbol{B}(\boldsymbol{\xi}) + \frac{h^2}{4}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{B}(\boldsymbol{\xi}) + h\mathbf{A}\boldsymbol{\xi} + \frac{h^2}{2}\mathbf{A}\boldsymbol{B}(\boldsymbol{\xi}) + \frac{h^2}{2}\mathbf{A}^2\boldsymbol{\xi} \\ + \boldsymbol{\Sigma}\boldsymbol{W}^{(0)}_{n-1}(h) + \mathbf{A}\boldsymbol{\Sigma}\boldsymbol{W}^{(1)}_{n-1}(h) + \frac{1}{2}\mathbf{A}^2\boldsymbol{\Sigma}\boldsymbol{W}^{(2)}_{n-1}(h) \\ + \frac{h^2}{4}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{B}(\boldsymbol{\xi}) + \frac{h^2}{2}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\mathbf{A}\boldsymbol{\xi} + \frac{h}{2}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{\Sigma}\boldsymbol{W}^{(0)}_{n-1}(h) \\ + \frac{h}{2}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\mathbf{A}\boldsymbol{\Sigma}\boldsymbol{W}^{(1)}_{n-1}(h) + \frac{h}{4}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\mathbf{A}^2\boldsymbol{\Sigma}\boldsymbol{W}^{(2)}_{n-1}(h) + \frac{h}{2}\boldsymbol{G}(\boldsymbol{\xi}) \\ + \frac{h^2}{4}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{B}(\boldsymbol{\xi}) + \mathcal{O}(h^{5/2})\end{aligned} \tag{6.19}$$

We now introduce the local errors $\boldsymbol{\delta}^{\mathrm{LTIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n)$ and $\boldsymbol{\delta}^{\mathrm{SIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n)$, defined by

$$\boldsymbol{\delta}^{\mathrm{LTIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n) := \boldsymbol{X}_{t_{n-1},\boldsymbol{\xi}}(t_n) - \boldsymbol{X}^{\mathrm{LTIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n), \tag{6.20a}$$

$$\boldsymbol{\delta}^{\mathrm{SIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n) := \boldsymbol{X}_{t_{n-1},\boldsymbol{\xi}}(t_n) - \boldsymbol{X}^{\mathrm{SIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n). \tag{6.20b}$$

It follows from eqs. (6.3) and (6.10) that $\boldsymbol{\delta}^{\mathrm{LTIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n)$ is given by

$$\begin{aligned}\boldsymbol{\delta}^{\mathrm{LTIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n) = \frac{h^2}{2}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\mathbf{A}\boldsymbol{\xi} - \frac{h^2}{2}\mathbf{A}\boldsymbol{B}(\boldsymbol{\xi}) - \frac{h^2}{2}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{B}(\boldsymbol{\xi}) + \frac{h^2}{2}\boldsymbol{D}(\boldsymbol{\xi}) \\ + \mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{\Sigma}\boldsymbol{W}^{(1)}_{n-1}(h) - \frac{1}{2}\mathbf{A}^2\boldsymbol{\Sigma}\boldsymbol{W}^{(2)}_{n-1}(h) + \mathcal{O}(h^{5/2})\end{aligned} \tag{6.21}$$

Similarly, it follows from eqs. (6.3) and (6.19) that $\boldsymbol{\delta}^{\mathrm{SIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n)$ is given by

$$\begin{aligned}\boldsymbol{\delta}^{\mathrm{SIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n) = \frac{h^2}{2}\Big(\boldsymbol{D}(\boldsymbol{\xi}) - \boldsymbol{G}(\boldsymbol{\xi})\Big) - \frac{h^2}{4}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{B}(\boldsymbol{\xi}) + \mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{\Sigma}\boldsymbol{W}^{(1)}_{n-1}(h) \\ - \frac{1}{2}\mathbf{A}^2\boldsymbol{\Sigma}\boldsymbol{W}^{(2)}_{n-1}(h) - \frac{h}{2}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\boldsymbol{\Sigma}\boldsymbol{W}^{(0)}_{n-1}(h) - \frac{h}{2}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\mathbf{A}\boldsymbol{\Sigma}\boldsymbol{W}^{(1)}_{n-1}(h) \\ - \frac{h}{4}\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{\xi})\mathbf{A}^2\boldsymbol{\Sigma}\boldsymbol{W}^{(2)}_{n-1}(h) + \mathcal{O}(h^{5/2})\end{aligned} \tag{6.22}$$

Taking expectations of (6.21) and (6.22) yields

$$\left\| \mathbb{E}\left[\boldsymbol{\delta}^{\mathrm{LTIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n) \right] \right\| := \mathcal{O}(h^2), \tag{6.23a}$$

$$\left\| \mathbb{E}\left[\boldsymbol{\delta}^{\mathrm{SIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n) \right] \right\| := \mathcal{O}(h^2), \tag{6.23b}$$

where we used the fact that

$$\mathbb{E}\Big[\boldsymbol{W}^{(k)}_{n-1}(h) \Big] = \mathbf{0} \quad \forall\ k \geq 0, n > 0.$$

Recalling from eq. (2.11) that $\mathbb{E}\Big[\left\| \boldsymbol{W}^{(k)}_{n-1}(h) \right\|^2 \Big] = \mathcal{O}(h^{2k+1})$, it follows that

$$\sqrt{\mathbb{E}\Big[\left\| \boldsymbol{\delta}^{\mathrm{LTIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n) \right\|^2 \Big]} = \mathcal{O}(h^{3/2}), \tag{6.24a}$$

$$\sqrt{\mathbb{E}\Big[\left\| \boldsymbol{\delta}^{\mathrm{SIE}}_{t_{n-1},\boldsymbol{\xi}}(t_n) \right\|^2 \Big]} = \mathcal{O}(h^{3/2}). \tag{6.24b}$$

By eqs. (6.23a)-(6.23b) and (6.24a)-(6.24b), it follows that $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$ are mean-square consistent of order $p = 1$ in the sense of Definition 6.1. ■

6.b Mean-square boundedness

We will now establish mean-square boundedness of all four methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$. This proof requires that the solution (exact or approximate) to the nonlinear ODE (4.4a) is linearly bounded. In this thesis, we prove mean-square boundedness under Assumption 3.3, wherein the assumption of linearly bounded growth on the exact- and implicit-Euler solutions of (4.4a) follows readily through Theorems 5.1 and 5.2. Recall from [1] and the specialization project [16] that linearly bounded solutions of (4.4a) had to be assumed explicitly in lieu of Assumption 3.3. Recall also from the specialization project that we approximated the solution to (4.4a) via the explicit Euler method. This was problematic, since it was discovered that the explicit Euler method could not guarantee linearly bounded solutions in the sense required (thus leading us to consider the case where (4.4a) is solved using the implicit Euler method).

We will prove mean-square boundedness for all four methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$ by proving mean-square boundedness for the generalized methods $\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ and $\overline{\boldsymbol{X}}^{\mathrm{S}}(t_n)$ defined through eqs. (4.15a)-(4.15b). At this point, we introduce the function $\Theta(z)$, defined by

$$\Theta(z) := \frac{1 - e^{-z}}{z}, \quad z \in \mathbb{R}. \tag{6.25}$$

Observe that while $\Theta(z)$ is undefined for $z = 0$, it follows in the limit that

$$\lim_{z \to 0^{\pm}} \Theta(z) = 1.$$

Thus, we define $\Theta(0) := 1$. Observe that since Assumptions 3.1 and 3.3 holds, it follows from Theorems 5.1 and 5.2 that

$$\left\| \boldsymbol{\varphi}^{(1)}_h(\boldsymbol{x}) \right\|^2 \leq e^{-2\alpha h} \left\| \boldsymbol{x} \right\|^2 + \frac{K}{\alpha}\big(1 - e^{-2\alpha h}\big) = e^{-2\alpha h} \left\| \boldsymbol{x} \right\|^2 + 2K\Theta(2\alpha h)h, \quad \forall\ h \in (0, h_0], \tag{6.26a}$$

$$\left\| \widetilde{\boldsymbol{\varphi}}^{(1)}_h(\boldsymbol{x}) \right\|^2 \leq e^{-2\widetilde{\alpha}_0 h} \left\| \boldsymbol{x} \right\|^2 + 2Kh = e^{-2\widetilde{\alpha}_0 h} \left\| \boldsymbol{x} \right\|^2 + 2K\Theta(0)h \quad \forall\ h \in (0, h_0], \tag{6.26b}$$

where, in eq. (6.26b), the upper-bound h_0 on the step-size h must satisfy $1+2\alpha h_0 > 0$. Recall that $\alpha, \widetilde{\alpha}_0 \in \mathbb{R}$ and that $\widetilde{\alpha}_0$ is given by

$$\widetilde{\alpha}_0 = \frac{1}{2h_0}\log\left(1+2\alpha h_0\right). \tag{6.27}$$

It is clear that we may write eqs. (6.26a)-(6.26b) as a single expression using the generalized flow $\overline{\boldsymbol{\varphi}}_h^{(1)}(\boldsymbol{x})$:

$$\left\|\overline{\boldsymbol{\varphi}}_h^{(1)}(\boldsymbol{x})\right\|^2 \leq e^{-2ah}\left\|\boldsymbol{x}\right\|^2 + 2K\Theta(2\beta h)h, \quad \forall\ h \in (0, h_0]. \tag{6.28}$$

Observe that if $a = \beta = \alpha$, then eq. (6.28) reduces to eq. (6.26a). Similarly, if $a = \widetilde{\alpha}_0$ and $\beta = 0$, then eq. (6.28) reduces to eqs. (6.26b).

Having defined the necessary ingredients, the goal of this subsection is to prove the following lemma:

Lemma 6.3 (mean-square boundedness) *Let $\widetilde{\boldsymbol{X}}^{LT}(t_n)$ and $\widetilde{\boldsymbol{X}}^{S}(t_n)$ be the Lie-trotter and Strang splitting methods defined through eqs.* (4.16a) *and* (4.16b)*, respectively. Similarly, let $\widetilde{\boldsymbol{X}}^{LTIE}(t_n)$ and $\widetilde{\boldsymbol{X}}^{SIE}(t_n)$ represent the LTIE- and SIE composition methods defined through eqs.* (4.16c) *and* (4.16d)*, respectively. Let Assumptions 3.1, 3.2, 3.3 and 3.4 hold. Then there exists $h_0 > 0$ such that $\widetilde{\boldsymbol{X}}^{LT}(t_n)$, $\widetilde{\boldsymbol{X}}^{S}(t_n)$, $\widetilde{\boldsymbol{X}}^{LTIE}(t_n)$ and $\widetilde{\boldsymbol{X}}^{SIE}(t_n)$ are mean-square bounded in the sense of Definition 6.2 for all $h \in (0, h_0]$.*

Proof: We start by proving mean-square boundedness for the generalized Lie-Trotter method defined through eq. (4.15a). Taking the expectation of the norm-squared of (4.15a), we find

$$\begin{aligned}
\mathbb{E}\Big[\left\|\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_n)\right\|^2\Big] &= \mathbb{E}\Big[\left\|e^{\mathbf{A}h}\overline{\boldsymbol{\varphi}}_h^{(1)}\big(\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_{n-1})\big) + \boldsymbol{Z}_{n-1}(h)\right\|^2\Big] \\
&= \mathbb{E}\Big[\left\|e^{\mathbf{A}h}\overline{\boldsymbol{\varphi}}_h^{(1)}\big(\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_{n-1})\big)\right\|^2\Big] + \mathbb{E}\Big[\left\|\boldsymbol{Z}_{n-1}(h)\right\|^2\Big] \\
&\quad + 2\mathbb{E}\Big[\Big\langle e^{\mathbf{A}h}\overline{\boldsymbol{\varphi}}_h^{(1)}\big(\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_{n-1})\big), \boldsymbol{Z}_{n-1}(h)\Big\rangle\Big], \quad n = 0, 1, ..., N \in \mathbb{N},
\end{aligned} \tag{6.29}$$

where we recall from eq. (4.12) that

$$\boldsymbol{Z}_{n-1}(h) = \int_0^h e^{\mathbf{A}(h-s)}\boldsymbol{\Sigma}d\boldsymbol{W}(s) \sim \mathcal{N}\big(\mathbf{0}, \mathbf{C}(h)\big), \quad n = 1, ..., N.$$

Moreover, it follows from eq. (2.9) that $\mathbb{E}\Big[\left\|\boldsymbol{Z}_{n-1}(h)\right\|^2\Big] = \mathrm{Tr}\big(\mathbf{C}(h)\big)$. The trace of a matrix satisfies for any matrix $\mathbf{M} \in \mathbb{R}^{d\times d}$ the following relation:

$$\mathrm{Tr}(\mathbf{M}\mathbf{M}^\top) = \mathrm{Tr}(\mathbf{M}^\top\mathbf{M}) = \left\|\mathbf{M}\right\|_{\mathrm{F}}^2, \tag{6.30}$$

with $\left\|\cdot\right\|_{\mathrm{F}}$ denoting the Frobenius norm, which is a submultiplicative norm. Additionally, the Frobenius norm satisfies the following relation for any matrices $\mathbf{M}_1, \mathbf{M}_2 \in \mathbb{R}^{d\times d}$

$$\left\|\mathbf{M}_1\mathbf{M}_2\right\|_{\mathrm{F}}^2 \leq \left\|\mathbf{M}_1\right\|^2\left\|\mathbf{M}_2\right\|_{\mathrm{F}}^2, \tag{6.31}$$

where $\|\cdot\|$ is the usual Euclidean norm. Thus, it follows from eq. (4.13), submultiplicativity of the Frobenius norm and eqs. (6.30)-(6.31) that

$$\begin{aligned}
\operatorname{Tr}\big(\mathbf{C}(h)\big) &= \operatorname{Tr}\left(\int_0^h \Big(e^{\mathbf{A}(h-s)}\mathbf{\Sigma}\Big)\Big(e^{\mathbf{A}(h-s)}\mathbf{\Sigma}\Big)^\top ds\right) \\
&= \int_0^h \operatorname{Tr}\left(\Big(e^{\mathbf{A}(h-s)}\mathbf{\Sigma}\Big)\Big(e^{\mathbf{A}(h-s)}\mathbf{\Sigma}\Big)^\top\right) ds \\
&= \int_0^h \operatorname{Tr}\left(\Big(e^{\mathbf{A}(h-s)}\mathbf{\Sigma}\Big)^\top\Big(e^{\mathbf{A}(h-s)}\mathbf{\Sigma}\Big)\right) ds \\
&= \int_0^h \left\|e^{\mathbf{A}(h-s)}\mathbf{\Sigma}\right\|_{\mathrm{F}}^2 ds \\
&\leq \|\mathbf{\Sigma}\|_{\mathrm{F}}^2 \int_0^h \left\|e^{\mathbf{A}(h-s)}\right\|^2 ds
\end{aligned}$$

Letting $\mu = \mu(\mathbf{A})$ denote the logarithmic norm of $\mathbf{A}$, it follows from eq. (2.3) that

$$\left\|e^{\mathbf{A}h}\right\|^2 \leq e^{2\mu h}. \tag{6.32}$$

Thus, we find that $\operatorname{Tr}\big(\mathbf{C}(h)\big)$ satisfies

$$\begin{aligned}
\operatorname{Tr}\big(\mathbf{C}(h)\big) &\leq \|\mathbf{\Sigma}\|_{\mathrm{F}}^2 \int_0^h \left\|e^{\mathbf{A}(h-s)}\right\|^2 ds \\
&\leq \|\mathbf{\Sigma}\|_{\mathrm{F}}^2 \int_0^h e^{2\mu(h-s)} ds \\
&= \|\mathbf{\Sigma}\|_{\mathrm{F}}^2 \, \frac{e^{2\mu h}-1}{2\mu} \\
&= \|\mathbf{\Sigma}\|_{\mathrm{F}}^2 \, \frac{e^{2\mu h}-1}{2\mu h} h \\
&= \|\mathbf{\Sigma}\|_{\mathrm{F}}^2 \, e^{2\mu h} \Theta(2\mu h) h,
\end{aligned} \tag{6.33}$$

It follows from [5] (Theorem B2 (e)) and the fact that $\boldsymbol{Z}_{n-1}(h)$ is zero in expectation that

$$\mathbb{E}\left[\left\langle e^{\mathbf{A}h}\overline{\boldsymbol{\varphi}}_h^{(1)}\big(\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_{n-1})\big), \boldsymbol{Z}_{n-1}(h)\right\rangle\right] = \left\langle e^{\mathbf{A}h}\overline{\boldsymbol{\varphi}}_h^{(1)}\big(\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_{n-1})\big), \mathbb{E}\Big[\boldsymbol{Z}_{n-1}(h)\Big]\right\rangle = 0. \tag{6.34}$$

Thus, by eqs. (6.32), (6.33) and (6.34), we find that (6.29) reduces to

$$\begin{aligned}
\mathbb{E}\Big[\left\|\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_n)\right\|^2\Big] &\leq \mathbb{E}\Big[\left\|e^{\mathbf{A}h}\overline{\boldsymbol{\varphi}}_h^{(1)}\big(\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_{n-1})\big)\right\|^2\Big] + \|\mathbf{\Sigma}\|_{\mathrm{F}}^2 \, e^{2\mu h}\Theta(2\mu h)h \\
&\leq \mathbb{E}\Big[\left\|e^{\mathbf{A}h}\right\|^2 \left\|\overline{\boldsymbol{\varphi}}_h^{(1)}\big(\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_{n-1})\big)\right\|^2\Big] + \|\mathbf{\Sigma}\|_{\mathrm{F}}^2 \, e^{2\mu h}\Theta(2\mu h)h \\
&\leq e^{2\mu h}\mathbb{E}\Big[\left\|\overline{\boldsymbol{\varphi}}_h^{(1)}\big(\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_{n-1})\big)\right\|^2\Big] + \|\mathbf{\Sigma}\|_{\mathrm{F}}^2 \, e^{2\mu h}\Theta(2\mu h)h
\end{aligned} \tag{6.35}$$

By eq. (6.28), it follows that

$$\left\|\overline{\boldsymbol{\varphi}}_h^{(1)}\big(\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_{n-1})\big)\right\|^2 \leq e^{-2ah}\left\|\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_{n-1})\right\|^2 + 2K\Theta(2\beta h)h \tag{6.36}$$

Inserting for eq. (6.36) into (6.35) yields

$$\begin{aligned}\mathbb{E}\Big[\left\|\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_n)\right\|^2\Big] &\leq e^{2\mu h}\mathbb{E}\Big[e^{-2ah}\left\|\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_{n-1})\right\|^2 + 2K\Theta(2\beta h)h\Big] + \|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2\, e^{2\mu h}\Theta(2\mu h)h \\ &= e^{2(\mu-a)h}\mathbb{E}\Big[\left\|\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_{n-1})\right\|^2\Big] + e^{2\mu h}\Big(2K\Theta(2\beta h) + \|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2\,\Theta(2\mu h)\Big)h \\ &= e^{2(\mu-a)h}\mathbb{E}\Big[\left\|\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_{n-1})\right\|^2\Big] + \overline{C}^{\mathrm{LT}}h,\end{aligned} \tag{6.37}$$

where $\overline{C}^{\mathrm{LT}}$ is given by

$$\overline{C}^{\mathrm{LT}} = \overline{C}^{\mathrm{LT}}(\beta,\mu,h) = e^{2\mu h}\Big(2K\Theta(2\beta h) + \|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2\,\Theta(2\mu h)\Big). \tag{6.38}$$

Note that $\overline{C}^{\mathrm{LT}}$ reduces to constants $\widetilde{C}^{\mathrm{LT}}$ associated with the LT splitting method and $\widetilde{C}^{\mathrm{LTIE}}$ associated with the LTIE composition method by inserting for $\beta = \alpha$ and $\beta = 0$, respectively. Back-iterating eq. (6.37) n times yields

$$\mathbb{E}\Big[\left\|\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_n)\right\|^2\Big] \leq e^{2(\mu-a)nh}\mathbb{E}\big[\,\|\boldsymbol{X}_0\|^2\,\big] + \overline{C}^{\mathrm{LT}}h\sum_{k=0}^{n-1} e^{2(\mu-a)kh}. \tag{6.39}$$

Observe at this point that the sum involved above is a geometric sum, and satisfies

$$h\sum_{k=0}^{n-1} e^{2(\mu-a)kh} = h\frac{e^{2(\mu-a)t_n}-1}{e^{2(\mu-a)h}-1} =: \Psi(t_n), \tag{6.40}$$

where we used the fact that $nh = t_n$. Inserting for eq. (6.40) into eq. (6.39), we find

$$\mathbb{E}\Big[\left\|\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_n)\right\|^2\Big] \leq e^{2(\mu-a)t_n}\mathbb{E}\big[\,\|\boldsymbol{X}_0\|^2\,\big] + \overline{C}^{\mathrm{LT}}\Psi(t_n). \tag{6.41}$$

Recalling that $t_n \leq t_N = T$, we proceed with showing that $\Psi(t_n)$ may always be bounded from above by a constant $\Psi_0(T)$ which is independent of t_n and h. For the case of $\mu - a > 0$, it follows that

$$\frac{h}{e^{2(\mu-a)h}-1} \leq \frac{1}{2(\mu-a)}, \qquad e^{2(\mu-a)t_n} - 1 \leq e^{2(\mu-a)T} - 1, \qquad h > 0, \quad \mu - a > 0,$$

implying that

$$\Psi(t_n) \leq \Psi_0(T) = \frac{e^{2(\mu-a)T}-1}{2(\mu-a)}, \quad \mu - a > 0.$$

For the case $\mu - a < 0$, it follows that $\Psi(t_n)$ is a monotonically increasing function in $t_n \geq 0$ and $h > 0$, yielding

$$\Psi(t_n) \leq \Psi_0(T) = h_0\frac{e^{2(\mu-a)T}-1}{e^{2(\mu-a)h_0}-1}, \quad \mu - a < 0.$$

Lastly, we have that $\Psi(t_n)$ is undefined for $\mu - a = 0$, however we have in the limit that

$$\lim_{\mu - a \to 0^{\pm}} \Psi(t_n) = t_n \leq \Psi_0(T) = T.$$

Thus, we see that $\Psi(t_n)$ may be bounded from above for all $\mu - a \in \mathbb{R}$ by a constant $\Psi_0(T)$ which is independent of t_n and h. Similarly, the constant $\overline{C}^{\mathrm{LT}} = \overline{C}^{\mathrm{LT}}(\beta, \mu, h)$ may be bounded from above by $\overline{C}_0^{\mathrm{LT}} := \overline{C}^{\mathrm{LT}}(\beta, \mu, h_0)$, where $\overline{C}_0^{\mathrm{LT}}$ is independent of h. Note that depending on the sign of β, we have

$$\Theta(2\beta h) \leq \begin{cases} 1, & \beta \geq 0 \\ \Theta(2\beta h_0), & \beta < 0 \end{cases}$$

Thus, depending on the sign of μ and β, it follows that $\overline{C}_0^{\mathrm{LT}}$ is given by

$$\overline{C}_0^{\mathrm{LT}} = \begin{cases} e^{2\mu h_0}\big(2K + \|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2\, \Theta(2\mu h_0)\big), & \mu > 0, \quad \beta \geq 0 \\ e^{2\mu h_0}\Big(2K\Theta(2\beta h_0) + \|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2\, \Theta(2\mu h_0)\Big), & \mu > 0, \quad \beta < 0, \\ 2K + \|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2, & \mu \leq 0, \quad \beta \geq 0, \\ 2K\Theta(2\beta h_0) + \|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2, & \mu \leq 0, \quad \beta < 0. \end{cases} \tag{6.42}$$

Thus, we see that there exists constants $\overline{C}_0^{\mathrm{LT}}$ and $\Psi_0(T)$ independent of h and t_n such that $\overline{C}^{\mathrm{LT}}\Psi(t_n) \leq \overline{C}_0^{\mathrm{LT}}\Psi_0(T)$. Then, it follows from eq. (6.41) that

$$\mathbb{E}\Big[\, \Big\| \overline{\boldsymbol{X}}^{\mathrm{LT}}(t_n) \Big\|^2 \,\Big] \leq e^{2(\mu - a)t_n} \mathbb{E}\big[\, \|\boldsymbol{X}_0\|^2 \,\big] + \overline{C}_0^{\mathrm{LT}} \Psi_0(T) \tag{6.43}$$

Moreover, it holds that

$$e^{2(\mu-a)t_n} \leq \begin{cases} e^{2(\mu-a)T}, & \mu - a > 0, \\ 1, & \mu - a \leq 0. \end{cases} \tag{6.44}$$

Letting $\overline{K}^{\mathrm{LT}}$ be given by

$$\overline{K}^{\mathrm{LT}} = \max\{1, \quad e^{2(\mu-a)T}, \quad \overline{C}_0^{\mathrm{LT}} \Psi_0(T)\}, \tag{6.45}$$

it follows from eqs. (6.43) and (6.44) that

$$\mathbb{E}\Big[\, \Big\| \overline{\boldsymbol{X}}^{\mathrm{LT}}(t_n) \Big\|^2 \,\Big] \leq \overline{K}^{\mathrm{LT}} \Big(1 + \mathbb{E}\big[\, \|\boldsymbol{X}_0\|^2 \,\big]\Big) \tag{6.46}$$

Since eq. (6.46) holds for all $n = 0, 1, ..., N$, it also holds for the value of n which maximizes $\mathbb{E}\Big[\, \Big\| \overline{\boldsymbol{X}}^{\mathrm{LT}}(t_n) \Big\|^2 \,\Big]$. Thus, we've shown that the generalized method $\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ defined through eq. (4.15a) is mean-square bounded in the sense of Definition 6.2. Since this generalized method represents both the Lie-Trotter splitting method $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ (recovered by setting $a = \beta = \alpha$) and the LTIE composition method $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ (recovered by setting $a = \widetilde{\alpha}_0$ and $\beta = 0$), it follows that both $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ are mean-square bounded in the sense of Definition 6.2.

We proceed with proving mean-square boundedness for the generalized Strang method defined through eq. (4.15b). It follows from eq. (6.28) that

$$\begin{aligned}\mathbb{E}\Big[\,\Big\|\overline{\boldsymbol{X}}^{\mathrm{S}}(t_n)\Big\|^2\,\Big] &= \mathbb{E}\Big[\,\Big\|\overline{\boldsymbol{\varphi}}^{(1)}_{h/2}\Big(e^{\mathbf{A}h}\overline{\boldsymbol{\varphi}}^{(1)}_{h/2}\big(\overline{\boldsymbol{X}}^{\mathrm{S}}(t_{n-1})\big)+\boldsymbol{Z}_{n-1}(h)\Big)\Big\|^2\,\Big]\\ &\leq e^{-ah}\mathbb{E}\Big[\,\Big\|e^{\mathbf{A}h}\overline{\boldsymbol{\varphi}}^{(1)}_{h/2}\big(\overline{\boldsymbol{X}}^{\mathrm{S}}(t_{n-1})\big)+\boldsymbol{Z}_{n-1}(h)\Big\|^2\,\Big]+K\Theta(\beta h)h\end{aligned} \tag{6.47}$$

Using the same logic which brought us from eq. (6.29) to eq. (6.35) for the generalized Lie-Trotter method, we find

$$\begin{aligned}\mathbb{E}\Big[\,\Big\|e^{\mathbf{A}h}\overline{\boldsymbol{\varphi}}^{(1)}_{h/2}\big(\overline{\boldsymbol{X}}^{\mathrm{S}}(t_{n-1})\big)+\boldsymbol{Z}_{n-1}(h)\Big\|^2\,\Big] &\leq e^{2\mu h}\mathbb{E}\Big[\,\Big\|\overline{\boldsymbol{\varphi}}^{(1)}_{h/2}\big(\overline{\boldsymbol{X}}^{\mathrm{S}}(t_{n-1})\big)\Big\|^2\,\Big]\\ &\quad+\|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2\,e^{2\mu h}\Theta(2\mu h)h\end{aligned} \tag{6.48}$$

It follows once more from eq. (6.28) that

$$\mathbb{E}\Big[\,\Big\|\overline{\boldsymbol{\varphi}}^{(1)}_{h/2}\big(\overline{\boldsymbol{X}}^{\mathrm{S}}(t_{n-1})\big)\Big\|^2\,\Big]\leq e^{-ah}\mathbb{E}\Big[\,\Big\|\overline{\boldsymbol{X}}^{\mathrm{S}}(t_{n-1})\Big\|^2\,\Big]+K\Theta(\beta h)h \tag{6.49}$$

Inserting for eq. (6.49) back into eq. (6.48) then yields

$$\begin{aligned}\mathbb{E}\Big[\,\Big\|e^{\mathbf{A}h}\overline{\boldsymbol{\varphi}}^{(1)}_{h/2}\big(\overline{\boldsymbol{X}}^{\mathrm{S}}(t_{n-1})\big)+\boldsymbol{Z}_{n-1}\Big\|^2\,\Big] &\leq e^{2\mu h}\Big(e^{-ah}\mathbb{E}\Big[\,\Big\|\overline{\boldsymbol{X}}^{\mathrm{S}}(t_{n-1})\Big\|^2\,\Big]+K\Theta(\beta h)h\Big)\\ &\quad+\|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2\,e^{2\mu h}\Theta(2\mu h)h\end{aligned} \tag{6.50}$$

Moreover, inserting for eq. (6.50) into eq. (6.47) yields

$$\begin{aligned}\mathbb{E}\Big[\,\Big\|\overline{\boldsymbol{X}}^{\mathrm{S}}(t_n)\Big\|^2\,\Big] &\leq e^{-ah}\mathbb{E}\Big[\,\Big\|e^{\mathbf{A}h}\overline{\boldsymbol{\varphi}}^{(1)}_{h/2}\big(\overline{\boldsymbol{X}}^{\mathrm{S}}(t_{n-1})\big)+\boldsymbol{Z}_{n-1}\Big\|^2\,\Big]+K\Theta(\beta h)h\\ &\leq e^{-ah}e^{2\mu h}\Big(e^{-ah}\mathbb{E}\Big[\,\Big\|\overline{\boldsymbol{X}}^{\mathrm{S}}(t_{n-1})\Big\|^2\,\Big]+K\Theta(\beta h)h\Big)\\ &\quad+e^{-ah}\,\|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2\,e^{2\mu h}\Theta(2\mu h)h+K\Theta(\beta h)h\\ &= e^{2(\mu-a)h}\mathbb{E}\Big[\,\Big\|\overline{\boldsymbol{X}}^{\mathrm{S}}(t_{n-1})\Big\|^2\,\Big]+e^{(2\mu-a)h}K\Theta(\beta h)h\\ &\quad+e^{(2\mu-a)h}\,\|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2\,\Theta(2\mu h)h+K\Theta(\beta h)h\\ &= e^{2(\mu-a)h}\mathbb{E}\Big[\,\Big\|\overline{\boldsymbol{X}}^{\mathrm{S}}(t_{n-1})\Big\|^2\,\Big]\\ &\quad+\Big(\big(e^{(2\mu-a)h}+1\big)K\Theta(\beta h)+e^{(2\mu-a)h}\,\|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2\,\Theta(2\mu h)\Big)h\\ &= e^{2(\mu-a)h}\mathbb{E}\Big[\,\Big\|\overline{\boldsymbol{X}}^{\mathrm{S}}(t_{n-1})\Big\|^2\,\Big]+\overline{C}^{\mathrm{S}}h,\end{aligned} \tag{6.51}$$

where $\overline{C}^{\mathrm{S}}$ is given by

$$\overline{C}^{\mathrm{S}}=\overline{C}^{\mathrm{S}}(a,\beta,\mu,h)=\big(e^{(2\mu-a)h}+1\big)K\Theta(\beta h)+e^{(2\mu-a)h}\,\|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2\,\Theta(2\mu h). \tag{6.52}$$

Note that $\overline{C}^{\mathrm{S}}$ reduces to constants $\widetilde{C}^{\mathrm{S}}$ associated with the Strang splitting method and $\widetilde{C}^{\mathrm{SIE}}$ associated with the SIE composition method by inserting for $a = \beta = \alpha$ and $(a = \widetilde{\alpha}_0, \beta = 0)$, respectively. By back-iterating eq. (6.51) n times and inserting for $t_n = nh$, we obtain

$$\mathbb{E}\Big[\,\Big\|\overline{\boldsymbol{X}}^{\mathrm{S}}(t_n)\Big\|^2\,\Big] \leq e^{2(\mu-a)t_n}\mathbb{E}\Big[\,\|\boldsymbol{X}_0\|^2\,\Big] + \overline{C}^{\mathrm{S}}\Psi(t_n), \tag{6.53}$$

where $\Psi(t_n)$ is the same function from eq. (6.40). Comparing eq. (6.53) with the corresponding expression (6.41) for the generalized Lie-Trotter case, it becomes clear that the remainder of the proof is essentially identical to the generalized Lie-Trotter case: It is clear from eq. (6.52) that there must exist a constant $\overline{C}_0^{\mathrm{S}}$ independent of h such that $\overline{C}^{\mathrm{S}} \leq \overline{C}_0^{\mathrm{S}}$. Thus, eq. (6.53) reduces to

$$\mathbb{E}\Big[\,\Big\|\overline{\boldsymbol{X}}^{\mathrm{S}}(t_n)\Big\|^2\,\Big] \leq e^{2(\mu-a)t_n}\mathbb{E}\Big[\,\|\boldsymbol{X}_0\|^2\,\Big] + \overline{C}_0^{\mathrm{S}}\Psi_0(T), \tag{6.54}$$

at which point it immediately follows that

$$\mathbb{E}\Big[\,\Big\|\overline{\boldsymbol{X}}^{\mathrm{S}}(t_n)\Big\|^2\,\Big] \leq \overline{K}^{\mathrm{S}}\bigg(1 + \mathbb{E}\Big[\,\|\boldsymbol{X}_0\|^2\,\Big]\bigg), \tag{6.55}$$

where $\overline{K}^{\mathrm{S}}$ is independent of t_n and h, and is given by

$$\overline{K}^{\mathrm{S}} = \max\{1, \quad e^{2(\mu-a)T}, \quad \widetilde{C}_0^{\mathrm{S}}\Psi_0(T)\}. \tag{6.56}$$

Since eq. (6.55) holds for all $n = 0, 1, ..., N$, it also holds for the value of n which maximizes $\mathbb{E}\Big[\,\Big\|\overline{\boldsymbol{X}}^{\mathrm{S}}(t_n)\Big\|^2\,\Big]$. Thus, we've shown that the generalized Strang method $\overline{\boldsymbol{X}}^{\mathrm{S}}(t_n)$ defined through eq. (4.15b) is mean-square bounded in the sense of Definition 6.2. Since this generalized method represents both the Strang splitting method $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$ (recovered by setting $a = \beta = \alpha$) and the SIE composition method $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$ (recovered by setting $a = \widetilde{\alpha}_0$ and $\beta = 0$), it follows that both $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$ are mean-square bounded in the sense of Definition 6.2. Thus, we've proven that all four methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$ are mean-square bounded in the sense of Definition 6.2. ■

We summarize the results of this section in the following theorems:

Theorem 6.4 (Mean-square convergence of splitting methods) *Let $\widetilde{\boldsymbol{X}}^{LT}(t_n)$ and $\widetilde{\boldsymbol{X}}^{S}(t_n)$ be the Lie-Trotter- and Strang splitting methods defined through eqs.* (4.16a) *and* (4.16b)*, respectively. Let Assumptions 3.1, 3.2, 3.3 and 3.4 hold. Then, there exists $h_0 > 0$ such that for all $h \in (0, h_0]$, $\widetilde{\boldsymbol{X}}^{LT}(t_n)$ and $\widetilde{\boldsymbol{X}}^{S}(t_n)$ are mean-square convergent of order $p = 1$ in the sense of Definition 6.3.*

Proof: The result follows directly from Lemma 4.1. of [1] and Lemma 6.3. ■

Theorem 6.5 (Mean-square convergence of composition methods) *Let $\widetilde{\boldsymbol{X}}^{LTIE}(t_n)$ and $\widetilde{\boldsymbol{X}}^{SIE}(t_n)$ be the LTIE- and SIE composition methods defined through eqs.* (4.16c) *and* (4.16d)*, respectively. Let Assumptions 3.1, 3.2, 3.3 and 3.4 hold. Then, there exists $h_0 > 0$ satisfying $1 + 2\alpha h_0 > 0$ such that for all $h \in (0, h_0]$, $\widetilde{\boldsymbol{X}}^{LTIE}(t_n)$ and $\widetilde{\boldsymbol{X}}^{SIE}(t_n)$ are mean-square convergent of order $p = 1$ in the sense of Definition 6.3*

Proof: The result follows directly from Lemma 6.2 and Lemma 6.3. ■

7 Structure Preservation

In this section, we investigate the conditions under which our considered methods are able to preserve important structural properties; in particular, 1-step hypoellipticity and geometric ergodicity.

7.a Preservation of noise structure and 1-step hypoellipticity

Recall that the SDE (3.1) is hypoelliptic if its transition probability (2.4) admits a smooth density, even though $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top$ is not of full rank. The discrete analogue of this property is termed 1-step hypoellipticity, and it states that the distribution of $\widetilde{\boldsymbol{X}}(t_n)$ given $\widetilde{\boldsymbol{X}}(t_{n-1})$ admits a smooth density. Of particular interest is the case where this smooth density is (multivariate) Gaussian. It was shown in [1] that the Lie-Trotter splitting method is 1-step hypoelliptic, and that the distribution of $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ given $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_{n-1})$ is normally distributed with mean and covariance matrix given by

$$\begin{aligned}\mathbb{E}\Big[\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)|\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_{n-1})\Big] &= e^{\mathbf{A}h}\boldsymbol{\varphi}_h^{(1)}\big(\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_{n-1})\big),\\ \mathrm{Cov}\Big[\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)|\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_{n-1})\Big] &= \mathbf{C}(h),\end{aligned}$$

where $\mathbf{C}(h)$ is the covariance matrix of eq. (4.13). By contrast, the conditional distribution for the Strang splitting method is in general not explicitly available [1]. We prove in Theorem 7.1 that the LTIE method also satisfies 1-step hypoellipticity, and that the distribution of $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ given $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_{n-1})$ also is multivariate normal. We do not prove this for the SIE method since, also here, the conditional distribution is in general not explicitly available. Furthermore, the property of 1-step hypoellipticity implies that, for a hypoelliptic SDE, the numerical method propagates the noise into the smooth component after 1 simulation step. It is known that Euler-Maruyama-type methods don't satisfy this property; rather, they are 2-step hypoelliptic, meaning that the noise is propagated into the smooth component after 2 simulation steps [9].

Theorem 7.1 *Let $\widetilde{\boldsymbol{X}}^{LTIE}(t_n)$ be the LTIE composition method defined through eq.* (4.16c)*. Assume that the SDE* (3.1) *is hypoelliptic in the sense of Assumption 3.4. Suppose further that Assumptions 3.1 and 3.3 holds. Then, there exists $h_0 > 0$ satisfying $1 + 2\alpha h_0 > 0$ such that for all $h \in (0, h_0]$, $\widetilde{\boldsymbol{X}}^{LTIE}(t_n)$ is 1-step hypoelliptic. Moreover, the distribution of $\widetilde{\boldsymbol{X}}^{LTIE}(t_n)$ given $\widetilde{\boldsymbol{X}}^{LTIE}(t_{n-1})$ is multivariate normal with expectation and covariance given by*

$$\begin{aligned}\mathbb{E}\Big[\widetilde{\boldsymbol{X}}^{LTIE}(t_n)|\widetilde{\boldsymbol{X}}^{LTIE}(t_{n-1})\Big] &= e^{\mathbf{A}h}\widetilde{\boldsymbol{\varphi}}_h^{(1)}\big(\widetilde{\boldsymbol{X}}^{LTIE}(t_{n-1})\big),\\ \mathrm{Cov}\Big[\widetilde{\boldsymbol{X}}^{LTIE}(t_n)|\widetilde{\boldsymbol{X}}^{LTIE}(t_{n-1})\Big] &= \mathbf{C}(h),\end{aligned}$$

where $\mathbf{C}(h)$ is the covariance matrix of eq. (4.13)*.*

Proof: By Assumptions 3.1 and 3.3, and since h_0 satisfies $1+2\alpha h_0 > 0$, the implicit-Euler solution of (4.4a) is well-defined for all $h \in (0, h_0]$. It follows immediately from eq. (4.16c) and from the fact that $\mathbb{E}\big[\boldsymbol{Z}_{n-1}(h)\big] = \mathbf{0}$ that

$$\begin{aligned}\mathbb{E}\Big[\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)|\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_{n-1})\Big] &= \mathbb{E}\Big[e^{\mathbf{A}h}\widetilde{\boldsymbol{\varphi}}_h^{(1)}\big(\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_{n-1}) + \boldsymbol{Z}_{n-1}(h)\big)\Big]\\ &= e^{\mathbf{A}h}\widetilde{\boldsymbol{\varphi}}_h^{(1)}\big(\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_{n-1})\big).\end{aligned}$$

The result then follows by

$$\begin{aligned}
&\mathrm{Cov}\Big[\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)|\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_{n-1})\Big] \\
&= \mathbb{E}\left[\left(\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n) - \mathbb{E}\Big[\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)|\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_{n-1})\Big]\right)\left(\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n) - \mathbb{E}\Big[\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)|\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_{n-1})\Big]\right)^{\top}\right] \\
&= \mathbb{E}\Big[\boldsymbol{Z}_{n-1}(h)\boldsymbol{Z}_{n-1}^{\top}(h)\Big] = \mathbf{C}(h),
\end{aligned}$$

where $\mathbf{C}(h)$ is the covariance matrix of eq. (4.13). ■

7.b Preservation of Lyapunov structure and geometric ergodicity

Recall from Lemma 3.2 that if $\mu < \alpha$ and $K > 0$, then the drift $\boldsymbol{F}$ satisfies the following dissipativity condition:

$$\big\langle \boldsymbol{x}, \boldsymbol{F}(\boldsymbol{x})\big\rangle \leq K - c\,\|\boldsymbol{x}\|^2, \quad \forall\ \boldsymbol{x} \in \mathbb{R}^d, \quad c = \alpha - \mu > 0, \quad K > 0. \tag{7.1}$$

Recall further that this condition ensured that the function

$$L(\boldsymbol{x}) := 1 + \|\boldsymbol{x}\|^2, \quad \forall\ \boldsymbol{x} \in \mathbb{R}$$

was a Lyapunov function satisfying condition (3.6) for the SDE (3.1). The existence of such a function then implies that the SDE is geometrically ergodic [9], [1].

Our goal now is to show that our numerical methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$ preserves the geometric ergodicity of the underlying SDE (3.1). The key to proving this result lies in showing that the methods satisfy a discrete analogue of 3.6, which we refer to (as in [1]) as a *discrete Lyapunov condition*. The discrete Lyapunov condition is formally stated in Definition 7.1.

Definition 7.1 (Discrete Lyapunov Condition) *Let L be a Lyapunov function for the SDE* (3.1). *A numerical solution $\widetilde{\boldsymbol{X}}(t_n)$ of* (3.1) *satisfies the discrete Lyapunov condition if there exists $\widetilde{\rho} \in (0,1)$ and $\widetilde{\eta} > 0$ such that*

$$\mathbb{E}\Big[L\big(\widetilde{\boldsymbol{X}}(t_n)\big)|\widetilde{\boldsymbol{X}}(t_{n-1})\Big] \leq \widetilde{\rho} L\big(\widetilde{\boldsymbol{X}}(t_{n-1})\big) + \widetilde{\eta}.$$

It was shown in [1] that both splitting methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$ satisfy this condition, however, their proof required $\left\|e^{\mathbf{A}h}\right\| < 1$. This is guaranteed if $\mu < 0$ (though the converse does not hold). In this work, we prove that all four methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$ satisfy the discrete Lyapunov condition of Definition 7.1, albeit under slightly different conditions. Specifically, we prove that if $\mu < \alpha$, then there exists $h_0 > 0$ such that $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$ preserves geometric ergodicity of the SDE (3.1) for all $h \in (0, h_0]$. Similarly, we prove that if $\mu < \widetilde{\alpha}_0$, then there exists $h_0 > 0$ satisfying $1 + 2\alpha h_0 > 0$ such that $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$ preserves geometric ergodicity of the SDE (3.1) for all $h \in (0, h_0]$. Crucially, and in contrast to [1], there is room for μ to be non-negative.

The proof is nearly identical for all four methods, so we will prove the result for a generalized method $\widetilde{\boldsymbol{X}}(t_n)$, where $\widetilde{\boldsymbol{X}}(t_n)$ represents all four methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$. Recall from eq. (6.37) that

$$\mathbb{E}\Big[\ \Big\|\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_n)\Big\|^2\ \Big] \leq e^{2(\mu-a)h}\mathbb{E}\Big[\ \Big\|\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_{n-1})\Big\|^2\ \Big] + \overline{C}^{\mathrm{LT}}h, \tag{7.2}$$

where $\overline{C}^{\mathrm{LT}} = \overline{C}^{\mathrm{LT}}(\beta, \mu, h)$ is given by eq. (6.38). Here, $\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ represents both methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$. Setting $a = \beta = \alpha$ causes $\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ and $\overline{C}^{\mathrm{LT}}$ to reduce to $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ and $\widetilde{C}^{\mathrm{LT}}$, respectively, where $\widetilde{C}^{\mathrm{LT}} = \overline{C}^{\mathrm{LT}}(\alpha, \mu, h)$. Similarly, setting $a = \widetilde{\alpha}_0$ and $\beta = 0$ causes $\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ and $\overline{C}^{\mathrm{LT}}$ to reduce to $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{C}^{\mathrm{LTIE}}$, respectively, where $\widetilde{C}^{\mathrm{LTIE}} = \overline{C}^{\mathrm{LT}}(0, \mu, h)$. Recall also from (6.51) that

$$\mathbb{E}\Big[\, \Big\| \overline{\boldsymbol{X}}^{\mathrm{S}}(t_n) \Big\|^2 \,\Big] \leq e^{2(\mu - a)h} \mathbb{E}\Big[\, \Big\| \overline{\boldsymbol{X}}^{\mathrm{S}}(t_{n-1}) \Big\|^2 \,\Big] + \overline{C}^{\mathrm{S}} h. \tag{7.3}$$

where $\overline{C}^{\mathrm{S}} = \overline{C}^{\mathrm{S}}(a, \beta, \mu, h)$ is given by eq. (6.52). Here, $\overline{\boldsymbol{X}}^{\mathrm{S}}(t_n)$ represents both methods $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$. Setting $a = \beta = \alpha$ causes $\overline{\boldsymbol{X}}^{\mathrm{S}}(t_n)$ and $\overline{C}^{\mathrm{S}}$ to reduce to $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$ and $\widetilde{C}^{\mathrm{S}}$, respectively, where $\widetilde{C}^{\mathrm{S}} = \overline{C}^{\mathrm{S}}(\alpha, \alpha, \mu, h)$. Similarly, setting $a = \widetilde{\alpha}_0$ and $\beta = 0$ causes $\overline{\boldsymbol{X}}^{\mathrm{S}}(t_n)$ and $\overline{C}^{\mathrm{S}}$ to reduce to $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$ and $\widetilde{C}^{\mathrm{SIE}}$, respectively, where $\widetilde{C}^{\mathrm{SIE}} = \overline{C}^{\mathrm{S}}(\widetilde{\alpha}_0, 0, \mu, h)$.

Using the generalized method $\widetilde{\boldsymbol{X}}(t_n)$, we can represent eqs. (7.2) and (7.3) by the general expression

$$\mathbb{E}\Big[\, \Big\| \widetilde{\boldsymbol{X}}(t_n) \Big\|^2 \,\Big] \leq e^{2(\mu - a)h} \mathbb{E}\Big[\, \Big\| \widetilde{\boldsymbol{X}}(t_{n-1}) \Big\|^2 \,\Big] + \widetilde{C} h. \tag{7.4}$$

We will use eq. (7.4) to prove that the generalized method $\widetilde{\boldsymbol{X}}(t_n)$ satisfies the discrete Lyapunov condition of Definition 7.1 when $\mu < a$. The result for all four methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$ is then obtained by appropriately inserting for $a = \{\alpha, \widetilde{\alpha}_0\}$ and $\widetilde{C} = \{\widetilde{C}^{\mathrm{LT}}, \widetilde{C}^{\mathrm{S}}, \widetilde{C}^{\mathrm{LTIE}}, \widetilde{C}^{\mathrm{SIE}}\}$.

Theorem 7.2 *Let $\widetilde{\boldsymbol{X}}^{LT}(t_n)$ and $\widetilde{\boldsymbol{X}}^{S}(t_n)$ be the Lie-Trotter and Strang splitting methods defined through eqs.* (4.16a) *and* (4.16b)*, respectively. Moreover, let $\widetilde{\boldsymbol{X}}^{LTIE}(t_n)$ and $\widetilde{\boldsymbol{X}}^{SIE}(t_n)$ be the LTIE- and SIE composition methods defined through eqs.* (4.16c) *and* (4.16d)*, respectively. Let*

$$L(\boldsymbol{x}) := 1 + \|\boldsymbol{x}\|^2, \quad \forall\ \boldsymbol{x} \in \mathbb{R}^d,$$

be a Lyapunov function of the SDE (3.1)*. Let $\mu = \mu(\mathbf{A})$ be the logarithmic norm of $\mathbf{A}$. Suppose Assumption 3.3 holds for some $K \geq 0$ and $\alpha \in \mathbb{R}$. If $\mu < \alpha$, then there exists $h_0 > 0$ such that $\widetilde{\boldsymbol{X}}^{LT}(t_n)$ and $\widetilde{\boldsymbol{X}}^{S}(t_n)$ satisfies the discrete Lyapunov condition of Definition 7.1 for $h \in (0, h_0]$. If $\mu < \widetilde{\alpha}_0$, where $\widetilde{\alpha}_0$ is given by*

$$\widetilde{\alpha}_0 := \frac{1}{2h_0} \log(1 + 2\alpha h_0),$$

then there exists $h_0 > 0$ satisfying $1 + 2\alpha h_0 > 0$ such that $\widetilde{\boldsymbol{X}}^{LTIE}(t_n)$ and $\widetilde{\boldsymbol{X}}^{SIE}(t_n)$ satisfy the discrete Lyapunov condition of Definition 7.1 for $h \in (0, h_0]$.

Proof: We will now show that the generalized method $\widetilde{\boldsymbol{X}}(t_n)$ satisfies the discrete Lyapunov condition of Definition 7.1 using eq. (7.4). By assumption, $\mu < a$. If $a = \alpha$, then this proof corresponds to the case of $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ of $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$. If $a = \widetilde{\alpha}_0$, then this proof corresponds to the case of $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$. In either case, $\mu < a$ implies that $e^{2(\mu - a)h} < 1$ for all $h \in (0, h_0]$. Thus,

we find

$$\begin{aligned}\mathbb{E}\Big[L\big(\widetilde{\boldsymbol{X}}(t_n)\big)|\widetilde{\boldsymbol{X}}(t_{n-1})\Big] &= 1 + \mathbb{E}\Big[\widetilde{\boldsymbol{X}}(t_n)|\widetilde{\boldsymbol{X}}(t_{n-1})\Big] \\ &\leq 1 + e^{2(\mu-a)h}\left\|\widetilde{\boldsymbol{X}}(t_{n-1})\right\|^2 + \widetilde{C}h \\ &\leq e^{2(\mu-a)h}\Big(1 + \left\|\widetilde{\boldsymbol{X}}(t_{n-1})\right\|^2\Big) + \widetilde{C}h + 1 \\ &= e^{2(\mu-a)h}L\big(\widetilde{\boldsymbol{X}}(t_{n-1})\big) + \widetilde{C}h + 1.\end{aligned} \tag{7.5}$$

That is, $\widetilde{\boldsymbol{X}}(t_n)$ satisfies the discrete Lyapunov condition of Definition 7.1 with $\widetilde{\rho}$ and $\widetilde{\eta}$ given by

$$\widetilde{\rho} = e^{2(\mu-\alpha)h} \in (0,1), \qquad \widetilde{\eta} = \widetilde{C}h + 1 > 0, \qquad h \in (0, h_0].$$

Appropriate insertion of $a = \{\alpha, \widetilde{\alpha}_0\}$ and $\widetilde{C} = \{\widetilde{C}^{\mathrm{LT}}, \widetilde{C}^{\mathrm{S}}, \widetilde{C}^{\mathrm{LTIE}}, \widetilde{C}^{\mathrm{SIE}}\}$ yields the result for all four methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$. ■

Recall that $\widetilde{\alpha}_0 \leq \alpha$ with equality only when $\alpha = 0$. As such, it is clear from Theorem 7.2 that the splitting methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$ can preserve geometric ergodicity for larger values of μ than the composition methods $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$. It follows in the limit $h_0 \to 0^+$, however, that

$$\lim_{h_0 \to +\infty} \widetilde{\alpha}_0 = \lim_{h_0 \to +\infty} \frac{1}{2h_0} \log\left(1 + 2\alpha h_0\right) = \alpha.$$

Thus, as the step-size decreases, the condition for preserving geometric ergodicity for the splitting methods and composition methods coincide. Figure 1 shows $\widetilde{\alpha}_0$ as a function of α for three different step-sizes h_0. We also include the straight line $y = \alpha$ for reference. Observe that as the step-size decreases, $\widetilde{\alpha}_0$ approaches the line $y = \alpha$.

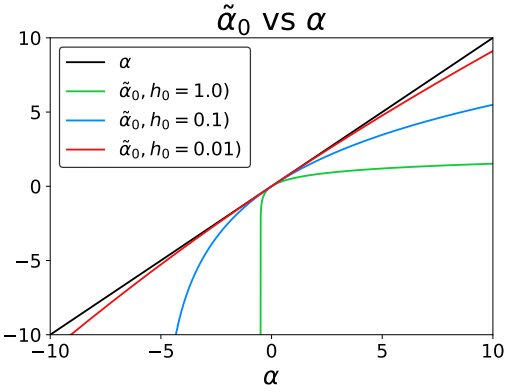


Figure 1: $\widetilde{\alpha}_0$ vs α for different stepsizes h_0.

We proceed with deriving asymptotic bounds for all four methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$ as $t_n \to +\infty$. We will again consider the generalized method $\widetilde{\boldsymbol{X}}(t_n)$ in this proof.

Recall from eq. (7.4) that all splitting- and composition methods satisfy a bound of the type

$$\mathbb{E}\Big[\,\Big\|\widetilde{\boldsymbol{X}}(t_n)\Big\|^2\,\Big] \leq e^{2(\mu-a)h}\mathbb{E}\Big[\,\|\boldsymbol{X}(t_{n-1})\|^2\,\Big] + \widetilde{C}h,$$

where appropriate insertion of $a = \{\alpha, \widetilde{\alpha}_0\}$ and $\widetilde{C} = \{\widetilde{C}^{\mathrm{LT}}, \widetilde{C}^{\mathrm{S}}, \widetilde{C}^{\mathrm{LTIE}}, \widetilde{C}^{\mathrm{SIE}}\}$ yields the particular bound for any of our methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$. Back-iterating eq. (7.4) n times then yields the bound

$$\mathbb{E}\Big[\,\Big\|\widetilde{\boldsymbol{X}}(t_n)\Big\|^2\,\Big] \leq e^{2(\mu-a)t_n}\mathbb{E}\Big[\,\|\boldsymbol{X}_0\|^2\,\Big] + \widetilde{C}\Psi(t_n) =: \widetilde{\kappa}(t_n). \tag{7.6}$$

Here, we obtain four different bounds $\widetilde{\kappa}(t_n) = \{\widetilde{\kappa}^{\mathrm{LT}}(t_n), \widetilde{\kappa}^{\mathrm{S}}(t_n), \widetilde{\kappa}^{\mathrm{LTIE}}(t_n), \widetilde{\kappa}^{\mathrm{SIE}}(t_n)\}$ corresponding to each of our four methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$, depending on our choice of $a = \{\alpha, \widetilde{\alpha}_0\}$ and $\widetilde{C} = \{\widetilde{C}^{\mathrm{LT}}, \widetilde{C}^{\mathrm{S}}, \widetilde{C}^{\mathrm{LTIE}}, \widetilde{C}^{\mathrm{SIE}}\}$.

Corollary 7.2.1 *Let the conditions of Theorem 7.2 hold. If $\mu < \alpha$, then it holds that*

$$\lim_{t_n\to+\infty}\mathbb{E}\Big[\,\Big\|\widetilde{\boldsymbol{X}}^{LT}(t_n)\Big\|^2\,\Big] \leq \frac{h\widetilde{C}^{LT}}{1-e^{2(\mu-\alpha)h}} \leq \frac{h_0\widetilde{C}_0^{LT}}{1-e^{2(\mu-\alpha)h_0}},$$
$$\lim_{t_n\to+\infty}\mathbb{E}\Big[\,\Big\|\widetilde{\boldsymbol{X}}^{S}(t_n)\Big\|^2\,\Big] \leq \frac{h\widetilde{C}^{S}}{1-e^{2(\mu-\alpha)h}} \leq \frac{h_0\widetilde{C}_0^{S}}{1-e^{2(\mu-\alpha)h_0}}.$$

Similarly, if $\mu < \widetilde{\alpha}_0$, then it holds that

$$\lim_{t_n\to+\infty}\mathbb{E}\Big[\,\Big\|\widetilde{\boldsymbol{X}}^{LTIE}(t_n)\Big\|^2\,\Big] \leq \frac{h\widetilde{C}^{LTIE}}{1-e^{2(\mu-\widetilde{\alpha}_0)h}} \leq \frac{h_0\widetilde{C}_0^{LTIE}}{1-e^{2(\mu-\widetilde{\alpha}_0)h_0}},$$
$$\lim_{t_n\to+\infty}\mathbb{E}\Big[\,\Big\|\widetilde{\boldsymbol{X}}^{SIE}(t_n)\Big\|^2\,\Big] \leq \frac{h\widetilde{C}^{SIE}}{1-e^{2(\mu-\widetilde{\alpha}_0)h}} \leq \frac{h_0\widetilde{C}_0^{SIE}}{1-e^{2(\mu-\widetilde{\alpha}_0)h_0}}.$$

Proof: Let $\widetilde{\boldsymbol{X}}(t_n)$ represent all four methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$. Moreover, let a represent both α and $\widetilde{\alpha}_0$, and let $\widetilde{C}$ represent any of the constants $\{\widetilde{C}^{\mathrm{LT}}, \widetilde{C}^{\mathrm{S}}, \widetilde{C}^{\mathrm{LTIE}}, \widetilde{C}^{\mathrm{SIE}}\}$. Suppose now that $\mu < a$. Then, it follows in the limit $t_n \to +\infty$ of eq. (7.6) that

$$\lim_{t_n\to+\infty} e^{2(\mu-a)t_n} = 0,$$
$$\lim_{t_n\to+\infty}\Psi(t_n) = \lim_{t_n\to+\infty} h\frac{e^{2(\mu-a)t_n}-1}{e^{2(\mu-a)h}-1} = \frac{h}{1-e^{2(\mu-a)h}} \leq \frac{h_0}{1-e^{2(\mu-a)h_0}},$$

and the result immediately follows for $\widetilde{\boldsymbol{X}}(t_n)$. Appropriate insertion for $a = \{\alpha, \widetilde{\alpha}_0\}$ and $\widetilde{C} = \{\widetilde{C}^{\mathrm{LT}}, \widetilde{C}^{\mathrm{S}}, \widetilde{C}^{\mathrm{LTIE}}, \widetilde{C}^{\mathrm{SIE}}\}$ then yields the result for all four methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$. ■

As an interesting side-note, we investigate the behavior of our methods in the limit $h_0 \to 0^+$. This also implies that $h \to +\infty$ since $h \in (0, h_0]$. In this limit, we transition from a discretized regime to a continuous regime, whereupon $t_n \to t$. Observe moreover from eq. (6.27) that $\widetilde{\alpha}_0$ approaches α as $h_0 \to 0^+$. We will therefore insert for $a = \alpha$ whenever $h_0 \to 0^+$. Since our

splitting- and composition methods approach the exact solution of (3.1) in this limit, the bound obtained is interpreted as a bound on the exact solution of (3.1). It follows from eq. (6.40)

$$\Psi(t) := \lim_{h\to 0^+} \Psi(t_n) = \lim_{h\to 0^+} h\frac{e^{2(\mu-\alpha)t_n}-1}{e^{2(\mu-\alpha)h}-1} = \frac{e^{2(\mu-\alpha)t}-1}{2(\mu-\alpha)}. \tag{7.7}$$

Moreover, we find from eq. (6.38) that

$$\lim_{h\to 0^+} \overline{C}^{\mathrm{LT}} = \lim_{h\to 0^+} e^{2\mu h}\Big(2K\Theta(2\beta h) + \|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2\,\Theta(2\mu h)\Big) = 2K + \|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2 \tag{7.8}$$

Similarly, we find from eq. (6.52) that

$$\lim_{h\to 0^+} \overline{C}^{\mathrm{S}} = \lim_{h\to 0^+} \Big(\big(e^{(2\mu-a)h}+1\big)K\Theta(\beta h) + e^{(2\mu-a)h}\,\|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2\,\Theta(2\mu h)\Big) = 2K + \|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2 \tag{7.9}$$

Thus, taking the limit $h \to 0^+$ of eqs. (6.41) and (6.53) and inserting for eqs. (7.7), (7.8) and (7.9) yields

$$\lim_{h\to 0^+} \mathbb{E}\Big[\,\Big\|\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_n)\Big\|^2\,\Big] \le e^{2(\mu-\alpha)t}\mathbb{E}\Big[\,\|\boldsymbol{X}_0\|^2\,\Big] + \big(2K + \|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2\big)\frac{e^{2(\mu-\alpha)t}-1}{2(\mu-\alpha)}$$

$$\lim_{h\to 0^+} \mathbb{E}\Big[\,\Big\|\overline{\boldsymbol{X}}^{\mathrm{S}}(t_n)\Big\|^2\,\Big] \le e^{2(\mu-\alpha)t}\mathbb{E}\Big[\,\|\boldsymbol{X}_0\|^2\,\Big] + \big(2K + \|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2\big)\frac{e^{2(\mu-\alpha)t}-1}{2(\mu-\alpha)}$$

Since the upper bounds on $\mathbb{E}\Big[\,\Big\|\overline{\boldsymbol{X}}^{\mathrm{LT}}(t_n)\Big\|^2\,\Big]$ and $\mathbb{E}\Big[\,\Big\|\overline{\boldsymbol{X}}^{\mathrm{S}}(t_n)\Big\|^2\,\Big]$ converge in the limit $h \to 0^+$, we treat this as an upper limit on $\mathbb{E}\Big[\,\|\boldsymbol{X}(t)\|^2\,\Big]$, where $\boldsymbol{X}(t)$ is the exact solution of eq. (3.1). Thus, we find

$$\begin{aligned}\mathbb{E}\Big[\,\|\boldsymbol{X}(t)\|^2\,\Big] &\le e^{2(\mu-\alpha)t}\mathbb{E}\Big[\,\|\boldsymbol{X}_0\|^2\,\Big] + \big(2K + \|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2\big)\frac{e^{2(\mu-\alpha)t}-1}{2(\mu-\alpha)} \\ &= e^{-2(\alpha-\mu)t}\mathbb{E}\Big[\,\|\boldsymbol{X}_0\|^2\,\Big] + \frac{2K + \|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2}{2(\alpha-\mu)}\big(1 - e^{-2(\alpha-\mu)t}\big).\end{aligned} \tag{7.10}$$

The bound (7.10) is only defined for $\alpha - \mu \neq 0$, however in the limit $\alpha - \mu \to 0^{\pm}$ we find

$$\mathbb{E}\Big[\,\|\boldsymbol{X}(t)\|^2\,\Big] \le \mathbb{E}\Big[\,\|\boldsymbol{X}_0\|^2\,\Big] + (2K + \|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2)t \tag{7.11}$$

Moreover, observe that if $\mu < \alpha$, then from eq. (7.10) it holds in the limit $t \to +\infty$ that

$$\lim_{t\to+\infty} \mathbb{E}\Big[\,\|\boldsymbol{X}(t)\|^2\,\Big] \le \frac{2K + \|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2}{2(\alpha-\mu)}, \quad \alpha - \mu > 0. \tag{7.12}$$

Observe the close correspondence between the bounds (7.10) and (7.11), and the bounds from Theorem 5.1. Observe also the similarities between the asymptotic bound for the nonlinear ODE given by eq. (5.5), which we recall here as

$$\lim_{t\to+\infty} \|\boldsymbol{X}(t)\|^2 \le \frac{K}{\alpha}, \quad \alpha > 0,$$

and the bound on the SDE given by eq. (7.12). Note in particular that, for the nonlinear ODE (5.1), asymptotic boundedness is determined by $\alpha > 0$, whereas for the SDE (3.1) asymptotic boundedness is determined by $\alpha - \mu > 0$ (or equivalently, $\mu < \alpha$).

8 Implementation

In this section, we discuss some practical implementation details of our proposed methods. We discuss how to speed up the implementation of Newton's method for the case where the implicit equation (4.8) cannot be solved explicitly. We also discuss our approach for numerically demonstrating mean-square convergence.

8.a Newton's method

In this section we describe two simple strategies for speeding up our implementation of Newton's method when the implicit equation (4.8) cannot be solved explicitly. The first strategy consists in choosing a split of the semi-linear drift $\boldsymbol{F}$ such that the jacobian of the nonlinear part $\boldsymbol{B}$ is diagonal, while still satisfying Assumptions 3.1, 3.2 and 3.3. This may require further splitting of the nonlinear part $\boldsymbol{B}$ into multiple nonlinear parts, increasing the total number of sub-equations to be solved by our composition methods. If the Jacobian is diagonal, then the Newton matrix $\mathbf{I} - h\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{z}^{(k)})$ from eq. (4.10) is trivially invertible, and its inverse is also a diagonal matrix. Thus, we may easily find an explicit expression for the inverse Newton matrix $(\mathbf{I} - h\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{z}^{(k)}))^{-1}$, reducing our problem of solving the linear system (4.10) to the problem of performing the following matrix-vector multiplication at every discrete time-point t_n:

$$\boldsymbol{z}^{(k+1)} = \Big(\mathbf{I} - h\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{z}^{k})\Big)^{-1}\Big[h\boldsymbol{B}(\boldsymbol{z}^{(k)}) - h\mathbf{J}_{\boldsymbol{B}}(\boldsymbol{z}^{(k)})\boldsymbol{z}^{(k)} - \widetilde{\boldsymbol{X}}^{(1)}(t_{n-1})\Big], \quad \boldsymbol{z}^{(0)} = \widetilde{\boldsymbol{X}}^{(1)}(t_{n-1}), \tag{8.1}$$

where $\widetilde{\boldsymbol{X}}^{(1)}(0) = \boldsymbol{X}_0$. Moreover, since the inverse Newton matrix is diagonal, this vector-matrix multiplication can be further reduced by simply multiplying the i'th diagonal element of the inverse Newton matrix with the i'th component of the right-hand-side of eq. (4.10). Then, on an implementation level, Newton's method behaves like a standard fixed-point iteration scheme where $\boldsymbol{z}^{(k+1)} = \boldsymbol{V}(\boldsymbol{z}^{(k)})$ for the function $\boldsymbol{V} : \mathbb{R}^d \mapsto \mathbb{R}^d$ defined by the right-hand-side of eq. (8.1). We note that finding an explicit expression for the inverse Newton matrix may require placing restrictions on the step-size h and/or parameters of the system.

Recall from section 4 that the DIEM method also employs Newton's method at every time-iteration if the corresponding implicit equation (4.17) cannot be solved explicitly. With this method, there is no equation splitting, and so we are forced to use the Jacobian of the full drift, which will in general not be a diagonal matrix. For this case, it would typically be easier to use a standard implementation of Newton's method, where we repeatedly solve a linear system. Lastly, we note that for systems of small dimensionality d, the difference in computational efficiency between a standard implementation of Newton's method, and an implementation which uses the inverse Newton matrix explicitly, will be small. However, for systems of large dimensionality, this difference in computational efficiency could in and of itself be sufficient to prefer the composition methods over the DIEM method.

The second strategy for speeding up our approximation of the solution to the implicit equation involves only performing a single iteration of Newton's method. Letting $k = 0$ and noting that $\boldsymbol{z}^{(0)} = \widetilde{\boldsymbol{X}}^{(1)}(t_{n-1})$, eq. (8.1) reduces to

$$\boldsymbol{z}^{(1)} = \widetilde{\boldsymbol{X}}^{(1)}(t_{n-1}) + h\Big(\mathbf{I} - h\mathbf{J}_{\boldsymbol{B}}\big(\widetilde{\boldsymbol{X}}^{(1)}(t_{n-1})\big)\Big)^{-1}\boldsymbol{B}\big(\widetilde{\boldsymbol{X}}^{(1)}(t_{n-1})\big).$$

The vector $\boldsymbol{z}^{(1)}$ is then taken as the approximation to the implicit Equation of eq. (4.8). Note that with this approach, we are no longer approximating the solution to the nonlinear ODE of eq. (4.4a)

by the implicit Euler method (via Newton's method). Rather, we are employing its linearized cousin; a first-order Rosenbrock method [12]. This method corresponds to approximating the solution to eq. (4.4a) at every time-step t_n by

$$\widehat{\boldsymbol{\varphi}}_h^{(1)}\big(\widehat{\boldsymbol{X}}^{(1)}(t_{n-1})\big) = \widehat{\boldsymbol{X}}^{(1)}(t_{n-1}) + h\Big(\mathbf{I} - h\mathbf{J}_{\boldsymbol{B}}\big(\widehat{\boldsymbol{X}}^{(1)}(t_{n-1})\big)\Big)^{-1}\boldsymbol{B}\big(\widehat{\boldsymbol{X}}^{(1)}(t_{n-1})\big), \tag{8.2}$$

where the initial value is given by $\widehat{\boldsymbol{X}}^{(1)}(0) = \boldsymbol{X}_0$. Combined with a diagonal Jacobian $\mathbf{J}_{\boldsymbol{B}}$, the method of eq. (8.2) can then easily be implemented to behave like an explicit method (in terms of computational efficiency) for solving the nonlinear ODE of eq. (4.4a), while still enjoying increased stability over e.g. the explicit Euler method.

Having introduced another method for approximating the solution to the nonlinear ODE of eq. (4.4a), we remark that the theory of sections 5, 6 and 7, which pertain to the implicit Euler method, no longer applies and would require modification. We will nevertheless include some numerical examples where eq. (4.4a) is solved using eq. (8.2), for comparison purposes. Whenever we do, we will refer to the composition methods obtained by solving the nonlinear ODE using eq. (8.2) as the *Lie-Trotter first-order Rosenbrock* (LTR1) method $\widetilde{\boldsymbol{X}}^{\mathrm{LTR1}}(t_n)$ and the *Strang first-order Rosenbrock* (SR1) method $\widetilde{\boldsymbol{X}}^{\mathrm{SR1}}(t_n)$. These composition methods are given by eq. (4.15a) and 4.15b, respectively, where the generalized flow $\overline{\boldsymbol{\varphi}}_h^{(1)}$ is replaced by the flow $\widehat{\boldsymbol{\varphi}}_h^{(1)}$ of eq. (8.2).

The next two sections will be dedicated to numerical demonstrations of the theory presented in this thesis. First, we consider a simple 1D cubic model problem which will demonstrate several of the theoretic results derived thus far. Second, we will consider the FitzHugh-Nagumo (FHN) model, which was also considered in [1]. Using the theory developed in this thesis, we are able to extend several of the results obtained in [1].

8.b Testing mean-square convergence

We test mean-square convergence by first computing a reference solution using a small step-size h, and then compare the error between this reference solution and the solutions obtained by our numerical method $\widetilde{\boldsymbol{X}}(t_n)$ for successively larger step-sizes h. Since the solution (exact or numerical) of an SDE is itself a stochastic process, we repeat this process across M simulated Brownian paths, recording the error along each path. Our measure for the overall error of the method $\widetilde{\boldsymbol{X}}(t_n)$ as a function of step-size h is the Root Mean-Squared Error (RMSE), defined as

$$\mathrm{RMSE}(h) := \sqrt{\frac{1}{M}\sum_{m=1}^{M}\Big\|\boldsymbol{X}(T) - \widetilde{\boldsymbol{X}}(T)\Big\|^2},$$

where $\widetilde{\boldsymbol{X}}(T)$ represents any of our considered methods. Essentially, the RMSE is the average error of our method across all M simulated paths. Note that all errors are recorded at the end time $t_N = T$. This is because our methods incur a local error at every time-step which accumulate throughout the simulation. Thus, we expect the error to be largest at the end-time T of the simulation.

The issue of reconstructing the Brownian paths at larger step-sizes is a nontrivial matter, and is discussed in the specialization project [16]. Using the terminology of that work, we reconstruct the correct Brownian paths across all step-sizes using an *exact accumulation of fine-grid random variables* when testing mean-square convergence of $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{LTR1}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{SR1}}(t_n)$. For comparison purposes, we also test mean-square convergence of the

DIEM method $\widetilde{\boldsymbol{X}}^{\mathrm{DIEM}}(t_n)$. In this case, we reconstruct the correct Brownian paths using a *linear accumulation of fine-grid random variables*. The RMSE is of order p if it holds for some constant $C > 0$ that

$$\mathrm{RMSE}(h) \leq Ch^p.$$

Note that by taking logs, we may express the above as

$$\log\big(\mathrm{RMSE}(h)\big) \leq \log C + p \log h$$

We may then use a linear least-squares regression to fit a straight line to the data points $\Big(\log h, \log\big(\mathrm{RMSE}(h)\big)\Big)$. The slope of this straight-line is then the experimental order of mean-square convergence.

When computing the reference solution, we would ideally like to use the exact solution of the SDE in question. This is of course rarely available, so we compute our reference solution by using Strang splitting with a small step-size h. We choose Strang splitting since this is presumed to be the most accurate of our methods and will therefore be the best proxy for the exact solution. Nevertheless, the results remain unchanged if we instead compute the reference solution using any of the other splitting- or composition methods, or indeed the DIEM method.

9 1D Cubic model problem

In this section we will investigate the following 1-dimensional cubic model problem:

$$dX(t) = \big(\omega X(t) - X^3(t)\big)dt + \sigma dW(t), \quad t \geq 0, \quad X(0) = X_0 \in \mathbb{R}, \tag{9.1}$$

where $\omega \in \mathbb{R}$ and $\sigma > 0$. The drift $F\big(X(t)\big)$ is on the form

$$F\big(X(t)\big) = AX(t) + B\big(X(t)\big), \quad A = \omega, \quad B\big(X(t)\big) = -X^3(t). \tag{9.2}$$

We start by showing that this choice of $B\big(X(t)\big)$ satisfies Assumptions 3.1, 3.2 and 3.3:

Proposition 9.1 *The function $B\big(X(t)\big) = -X^3(t)$ satisfies Assumption 3.1 for any $c_1 > 0$.*

Proof: Using the relation

$$(x-y)^2(x^2+xy+y^2) = x^4 + y^4 - xy^3 - yx^3, \quad \forall\, x, y \in \mathbb{R},$$

where

$$x^2 + xy + y^2 = \left(x + \frac{y}{2}\right)^2 + \frac{3y^2}{4} \geq 0, \quad \forall\, x, y \in \mathbb{R},$$

the result follows immediately

$$\begin{aligned}\big\langle x - y, B(x) - B(y)\big\rangle &= -x^4 - y^4 + xy^3 + yx^3 \\ &= -(x-y)^2(x^2+xy+y^2) \leq 0 \leq c_1(x-y)^2, \quad \forall x, y \in \mathbb{R}, \quad c_1 > 0. \quad \blacksquare\end{aligned}$$

Proposition 9.2 *The function $B\big(X(t)\big) = -X^3(t)$ satisfies Assumption 3.2 for $c_2 = 8$ and $\chi = 3$.*

Proof: Using the relations

$$\begin{aligned}(y^3 - x^3)^2 &= (x^2 + xy + y^2)^2(x - y)^2, \quad \forall\ x, y \in \mathbb{R},\\ xy &\le x^2 + y^2 \quad \forall\ x, y \in \mathbb{R},\\ 2x^2y^2 &\le x^4 + y^4, \quad \forall\ x, y \in \mathbb{R},\end{aligned}$$

the result follows by

$$\begin{aligned}\big(B(x) - B(y)\big)^2 &= \big(y^3 - x^3\big)^2\\ &= (x^2 + xy + y^2)^2(x - y)^2\\ &\le 4(x^2 + y^2)^2(x - y)^2\\ &= 4(x^4 + y^4 + 2x^2y^2)(x - y)^2\\ &\le 8(x^4 + y^4)(x - y)^2\\ &\le 8(1 + x^4 + y^4)(x - y)^2 \quad \blacksquare\end{aligned}$$

Proposition 9.3 *The function $B\big(X(t)\big)$ satisfies Assumption 3.3 for any $\alpha \in \mathbb{R}$ and $K = \alpha^2/4$.*

Proof: We find

$$xB(x) + \alpha x^2 = -x^4 + \alpha x^2 \tag{9.3}$$

We note that if $\alpha \le 0$, then the above is clearly globally bounded from above by $K = 0$. Assuming that $\alpha > 0$, we differentiate the above, set the derivative equal to zero and solve for x:

$$-4x^3 + 2\alpha x = 0 \quad \Longleftrightarrow \quad x = \{0, \pm\sqrt{\alpha/2}\}.$$

It follows that $x = 0$ corresponds to a local minimum of $xB(x) + \alpha x^2$, whereas $x = \pm\sqrt{\alpha/2}$ correspond to global maxima when $\alpha > 0$. Inserting for $x = \pm\sqrt{\alpha/2}$ into eq. (9.3) then yields the global maxima $\alpha^2/4$ of $xB(x) + \alpha x^2$, which is the result. $\blacksquare$

With the split (9.2), the resulting linear SDE is given by

$$dX^{(2)}(t) = \omega X^{(2)}(t)dt + \sigma dW(t), \quad t \ge 0, \quad X^{(2)}(0) = X_0. \tag{9.4}$$

The exact solution to the linear SDE (9.4) at time t_n starting from $X^{(2)}(t_{n-1})$ is thus given by

$$X^{(2)}(t_n) = e^{\omega h}X^{(2)}(t_{n-1}) + Z_{n-1}(h), \tag{9.5}$$

where $Z_{n-1}(h)$ is given for every $n > 0$ as

$$Z_{n-1}(h) = \sigma \int_0^h e^{\omega(h-s)}dW(s) \sim \mathcal{N}\bigg(0, \frac{\sigma^2}{2\omega}(e^{2\omega h} - 1)\bigg). \tag{9.6}$$

Note that in the limit $\omega \to 0^{\pm}$, the variance of $Z_{n-1}(h)$ reduces to $\sigma^2 h$.

The nonlinear ODE for the split (9.2) is given by

$$dX^{(1)}(t) = -\big(X^{(1)}(t)\big)^3 dt, \quad t \ge 0, \quad X^{(1)}(0) = X_0. \tag{9.7}$$

The exact solution of this ODE at time t_n starting from $X^{(1)}(t_{n-1})$ is given by

$$\varphi_h^{(1)}\big(X^{(1)}(t_{n-1})\big) = \frac{X^{(1)}(t_{n-1})}{\sqrt{1+2\big(X^{(1)}(t_{n-1})\big)^2 h}}, \quad X^{(1)}(0) = X_0. \tag{9.8}$$

The implicit equation (4.8) used to compute the implicit Euler solution of eq. (9.7) is given by

$$X^* = \widetilde{X}^{(1)}(t_{n-1}) - h\big(X^*\big)^3 \tag{9.9}$$

This corresponds to a depressed cubic equation on the form

$$\big(X^*\big)^3 + aX^* = b,$$

where $a = \frac{1}{h}$ and $b = \widetilde{X}^{(1)}(t_{n-1})/h$. Thus, we can find an explicit solution X^* to the implicit equation. Introducing the change-of-variables $3uv = a$ and $u^3 - v^3 = b$, it follows that the solution X^* is given by $X^* = u - v$, where u and v are given by

$$u = \sqrt[3]{\frac{-b + \sqrt{b^2 + 4(a/3)^3}}{2}}, \qquad v = \frac{a}{3u}. \tag{9.10}$$

Thus, the implicit Euler solution is given at every time-step as

$$\widetilde{\varphi}_h^{(1)}\big(\widetilde{X}^{(1)}(t_{n-1})\big) = \sqrt[3]{\frac{-b + \sqrt{b^2 + 4(a/3)^3}}{2}} - \frac{a}{3\sqrt[3]{\frac{-b+\sqrt{b^2+4(a/3)^3}}{2}}}, \quad \widetilde{X}^{(1)}(0) = X_0, \tag{9.11}$$

where $a = 1/h$ and $b = \widetilde{X}^{(1)}(t_{n-1})/h$.

The Jacobian of $B(X(t)) = -X^3(t)$ is simply equal to its derivative since $d = 1$. Thus, the first-order Rosenbrock solution of eq. (9.7) defined through eq. (8.2) is given at every time-step as

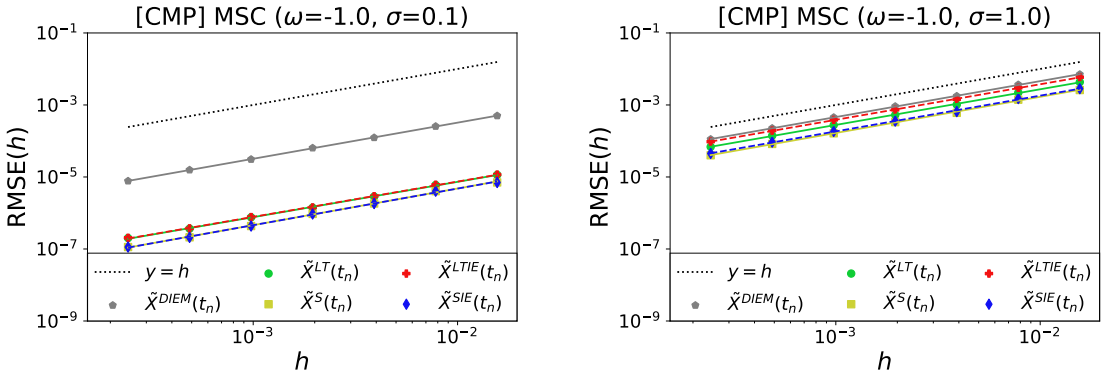
$$\widehat{\varphi}_h^{(1)}\big(\widehat{X}^{(1)}(t_{n-1})\big) = \widehat{X}^{(1)}(t_{n-1}) - \frac{\big(\widehat{X}^{(1)}(t_{n-1})\big)^3 h}{1+3\big(\widehat{X}^{(1)}(t_{n-1})\big)^2 h}, \quad \widehat{X}^{(1)}(0) = X_0. \tag{9.12}$$

We emphasize once more that we have not developed any theory regarding mean-square convergence or structure preservation for the composition methods $\widetilde{X}^{\mathrm{LTR1}}(t_n)$ and $\widetilde{X}^{\mathrm{SR1}}(t_n)$ obtained when the nonlinear ODE is solved by the first-order Rosenbrock method. However, for comparison purposes, we include its results when appropriate.

9.a Mean-square convergence

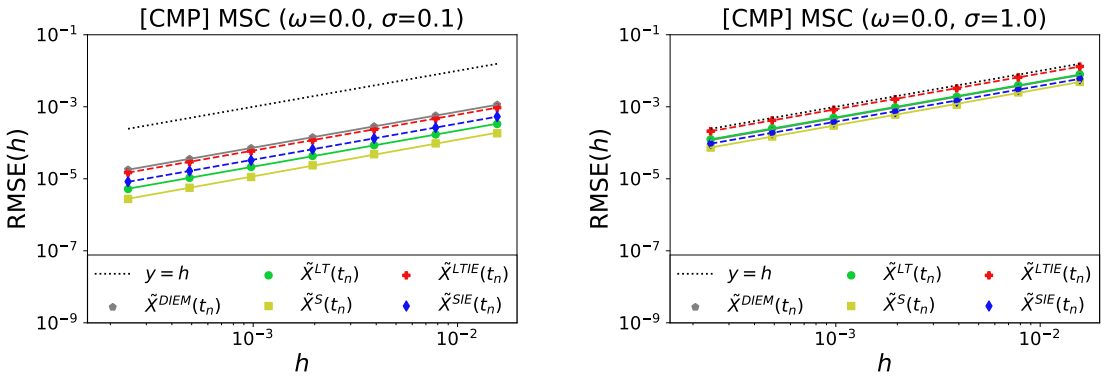
We start by investigating mean-square convergence of our methods when applied to the cubic model problem (9.1). We test mean-square convergence for two noise levels: A moderate noise level given by $\sigma = 0.1$ and a high noise level of $\sigma = 1$. We test mean-square convergence for three distinct choices of the parameter ω, namely $\omega = \{-1, 0, 1\}$. Observe that for the cubic model problem (9.1), the logarithmic norm μ of A is given simply by ω. As such, we can vary the behavior of the linear SDE (4.4b) by simply varying the parameter ω. In particular, $\omega = -1$ yields a stable linear SDE. By contrast, $\omega = 1$ yields an unstable linear SDE, and $\omega = 0$ yields neither a stable nor unstable

SDE. Thus, we refer to the linear SDE as *non-stable* when $\omega = \{0, 1\}$. We perform this test for $\widetilde{X}^{\mathrm{LT}}(t_n)$, $\widetilde{X}^{\mathrm{S}}(t_n)$, $\widetilde{X}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{X}^{\mathrm{SIE}}(t_n)$. For comparison purposes, we also report the results for the DIEM method $\widetilde{X}^{\mathrm{DIEM}}(t_n)$. The reference solution is computed using Strang splitting with step-size $h = 2^{-14}$. We test our methods for step-sizes $h = 2^{-q}$ with $q = \{12, 11, 10, 9, 8, 7, 6\}$. We simulate the solutions across $M = 1000$ Brownian paths. For each path, the initial value is set to $X_0 = 2$ and the end-time is set to $T = 10$. The results for mean-square convergence are reported in Figures 2, 3 and 4.



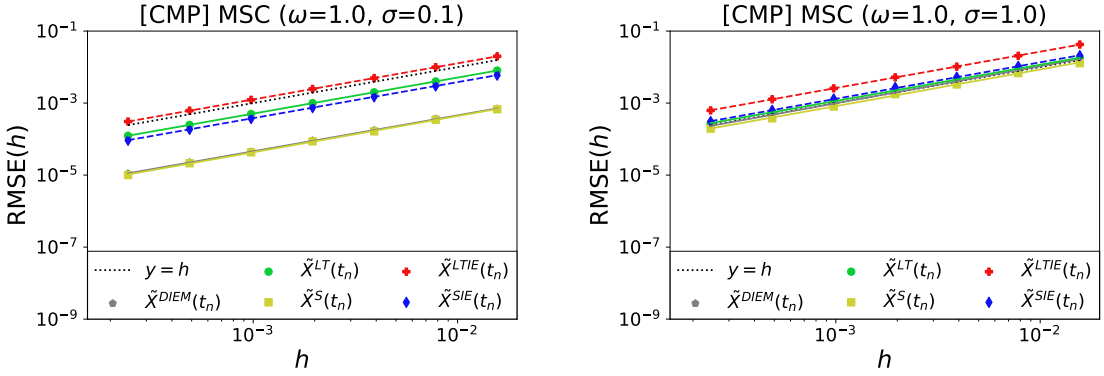
(a) $\sigma = 0.1$ (b) $\sigma = 1$

Figure 2: Cubic model problem; mean-square convergence (msc) test with $\omega = \mu = -1$ using two different noise levels $\sigma = 0.1$ and $\sigma = 1$. Tested for each method across $M = 1000$ simulated Brownian paths with initial value $X_0 = 2$ and end-time $T = 10$ for step-sizes $h = 2^{-q}$ with $q = \{12, 11, 10, 9, 8, 7, 6\}$.



(a) $\sigma = 0.1$ (b) $\sigma = 1$

Figure 3: Cubic model problem; mean-square convergence (msc) test with $\omega = \mu = 0$ using two different noise levels $\sigma = 0.1$ and $\sigma = 1$. Tested for each method across $M = 1000$ simulated Brownian paths with initial value $X_0 = 2$ and end-time $T = 10$ for step-sizes $h = 2^{-q}$ with $q = \{12, 11, 10, 9, 8, 7, 6\}$.



(a) $\sigma = 0.1$ (b) $\sigma = 1$

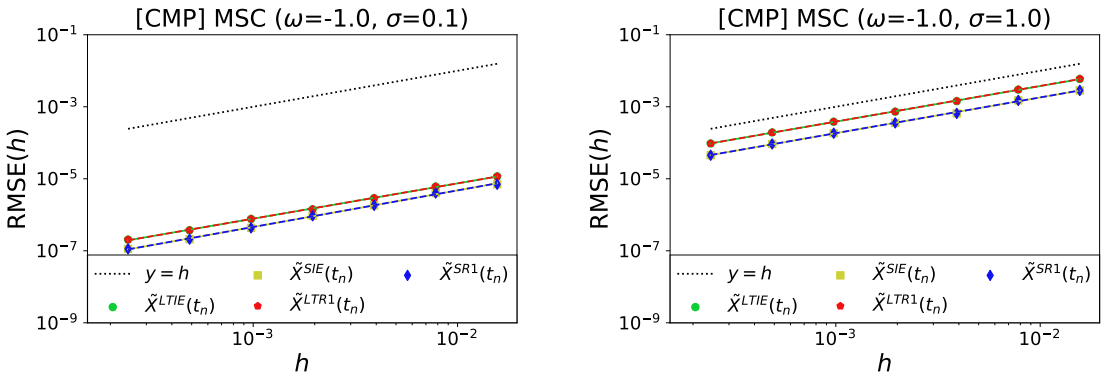
Figure 4: Cubic model problem; mean-square convergence (msc) test with $\omega = \mu = 1$ using two different noise levels $\sigma = 0.1$ and $\sigma = 1$. Tested for each method across $M = 1000$ simulated Brownian paths with initial value $X_0 = 2$ and end-time $T = 10$ for step-sizes $h = 2^{-q}$ with $q = \{12, 11, 10, 9, 8, 7, 6\}$.

Mean-square convergence of order $p = 1$ is observed for all considered method, for all choices of $\omega = \{-1, 0, 1\}$ and noise levels $\sigma = \{0.1, 1\}$. The results for $\widetilde{X}^{\mathrm{DIEM}}(t_n)$ are somewhat difficult to observe in Figures 3b and 4a because they coincide with those of the $\widetilde{X}^{\mathrm{LT}}$ and $\widetilde{X}^{\mathrm{S}}$, respectively. In Figure 4b, the results for $\widetilde{X}^{\mathrm{DIEM}}(t_n)$ lie nestled between those of $\widetilde{X}^{\mathrm{LT}}$ and $\widetilde{X}^{\mathrm{S}}$.

We observe a general trend of the Strang approach outperforming the Lie-Trotter approach, which is expected due to the inherent symmetry and use of fractional steps associated with the Strang approach. Moreover, we observe a general trend of $\widetilde{X}^{\mathrm{S}}(t_n)$ outperforming $\widetilde{X}^{\mathrm{SIE}}(t_n)$ and $\widetilde{X}^{\mathrm{LT}}(t_n)$ outperforming $\widetilde{X}^{\mathrm{LTIE}}(t_n)$, which is also as expected. The difference in error between our splitting- and composition methods for a given compositional approach (Strang or Lie-Trotter) varies across our tests. In Fig. 2a, for instance, it is observed that $\widetilde{X}^{\mathrm{S}}(t_n)$ and $\widetilde{X}^{\mathrm{SIE}}(t_n)$ coincide, and the same is true for $\widetilde{X}^{\mathrm{LT}}(t_n)$ and $\widetilde{X}^{\mathrm{LTIE}}(t_n)$. By contrast, in Fig. 4a for instance, the difference in error between our splitting- and composition methods are rather large for a given compositional approach.

We observe a general trend of error increasing with the noise level σ, which seems reasonable. We also observe that the errors for all splitting- and composition methods tend to increase with ω, however this does not appear to be the case for $\widetilde{X}^{\mathrm{DIEM}}(t_n)$. This is especially interesting considering that, for our splitting- and composition methods, ω only features in the linear SDE, which these methods solve exactly. This is perhaps explained non-rigorously by the fact that when $\omega = \{0, 1\}$, the linear SDE is non-stable. Meanwhile, the nonlinear ODE remains stable. For $\omega = \{0, 1\}$, then, it may be the case that the intra-step dynamics of composing the solutions from a non-stable linear SDE and a stable nonlinear ODE increases error more than solving the linear SDE exactly reduces it. By this logic, it would make sense that the DIEM method is relatively unaffected by the changing values of ω: The DIEM method features no equation splitting, and hence it does not distinguish between the linear and nonlinear parts of the drift; there is just "the drift" $F(X(t)) = \omega X(t) - X^3(t)$. From proposition 9.3, we showed that Assumption 3.3 is satisfied for any $\alpha \in \mathbb{R}$. Hence, we can always guarantee $\omega = \mu < \alpha$, meaning that the drift $F(X(t)) = \omega X(t) - X^3(t)$ is stable for all values of ω. More work is needed to determine if this notion indeed can explain this observed result, and if so, to what extent it generalizes.

Lastly, we investigate mean-square convergence for the LTR1- and SR1 methods in the case of $\omega = -1$, and compare them with the results of the LTIE- and SIE methods. The results are reported in Figure 5



(a) $\sigma = 0.1$ (b) $\sigma = 1$

Figure 5: Cubic model problem; mean-square convergence test comparing the LTIE- and SIE composition methods against the LTR1- and SR1 composition methods with $\omega = \mu = -1$ using two different noise levels $\sigma = 0.1$ and $\sigma = 1$. Tested for each method across $M = 1000$ simulated Brownian paths with initial value $X_0 = 2$ and end-time $T = 10$ for step-sizes $h = 2^{-q}$ with $q = \{12, 11, 10, 9, 8, 7, 6\}$.

We observe that the results for the LTIE- and LTR1 methods essentially coincide, as do the results for the SIE- and SR1 methods. Numerical experiments indicate that this holds true for $\omega = \{-10, 0, 1, 10\}$ as well, for both noise levels. This seems promising, especially since the first-order Rosenbrock method can be made to behave like an explicit method in terms of computational efficiency if we split the drift term $\boldsymbol{F}$ such that all nonlinear parts of the drift has diagonal Jacobians, as discussed in section 8.

9.b Preservation of geometric ergodicity

We proceed by investigating the proposed methods' ability to preserve geometric ergodicity. As shown in Proposition 9.3, $B\big(X(t)\big)$ satisfies Assumption 3.3 for any value of α. Therefore, the SDE is geometrically ergodic for any value of ω since $\omega = \mu < \alpha$ can always be ensured. Consequently, our splitting- and composition methods should be able to preserve geometric ergodicity for any value of ω as well. Of particular interest is the case when $\omega = \mu \geq 0$, since the linear SDE in this case is neither stable nor geometrically ergodic.

We note that for any SDE of the form

$$dX(t) = -\frac{dV(x)}{dx}dt + \sigma dW(t), \quad t \geq 0, \quad X(0) = X_0 \in \mathbb{R}, \tag{9.13}$$

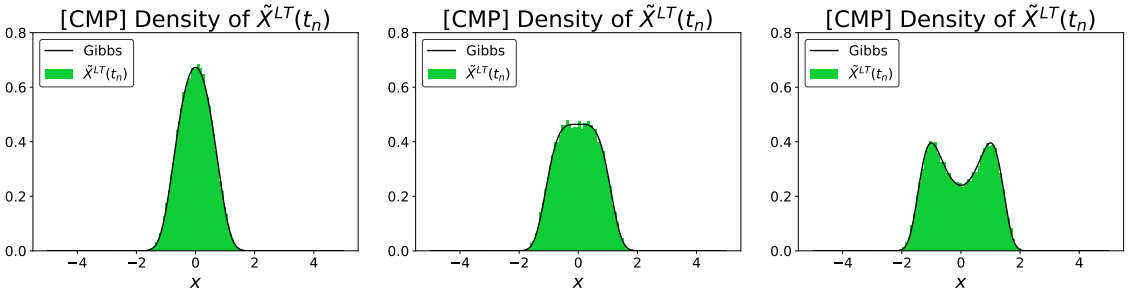
the limiting distribution of the process $X(t)$ is given by the Gibbs distribution [13] (Chapter 4.5)

$$\pi(x) = ze^{-2V(x)/\sigma^2}, \tag{9.14}$$

Observe that may write eq. (9.1) on the form (9.13) by letting $V(x)$ be given by

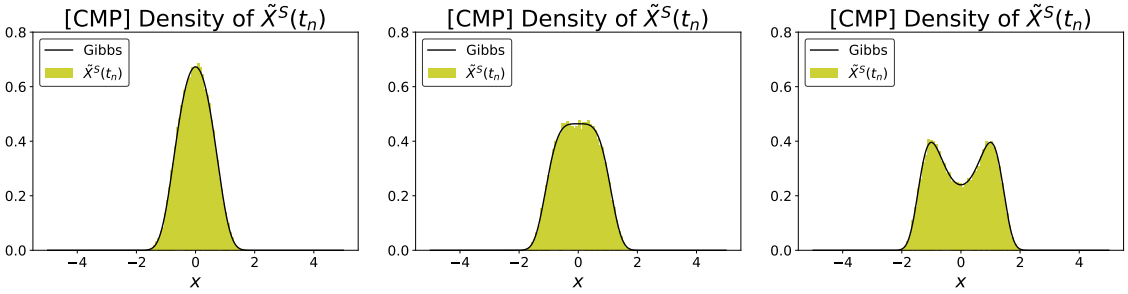
$$V(x) = \frac{\omega x^2}{2} - \frac{x^4}{4}. \tag{9.15}$$

Thus, we may test if our methods preserve geometric ergodicity by whether or not the distribution of their solutions approach, for all initial values, the true limiting distribution given by eq. (9.14) with $V(x)$ as in eq. (9.15). We test this by simulating $M = 50000$ Brownian paths and measuring the value of $\widetilde{X}(t_n)$ at time $t_n = 20$. Moreover, for each simulated Brownian path, the initial value X_0 is drawn from a uniform distribution over the interval $[1, 2]$. That is, $X_0 \sim \mathcal{U}(1, 2)$ for each simulated path. The noise level and step-size are set to $\sigma = 1$ and $h = 0.01$, respectively. The results for $\widetilde{X}^{\mathrm{LT}}(t_n)$, $\widetilde{X}^{\mathrm{S}}(t_n)$, $\widetilde{X}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{X}^{\mathrm{SIE}}(t_n)$ are reported in figures 6, 7, 8 and 9, respectively. For comparison purposes, we also report the results for $\widetilde{X}^{\mathrm{LTR1}}(t_n)$, $\widetilde{X}^{\mathrm{SR1}}(t_n)$ and $\widetilde{X}^{\mathrm{DIEM}}(t_n)$ in Figures 10, 11 and 12, respectively.



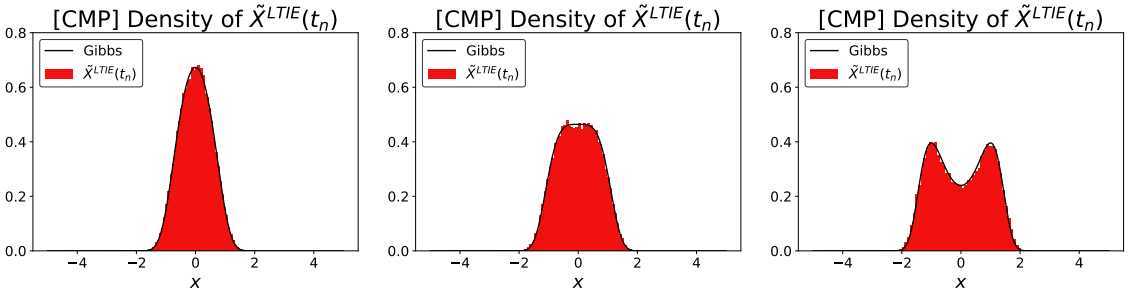
(a) $\omega = -1$ (b) $\omega = 0$ (c) $\omega = 1$

Figure 6: Histogram of $\widetilde{X}^{\mathrm{LT}}(t_n = 20)$ across all $M = 50000$ simulated Brownian paths vs theoretical Gibbs density for $\omega = \{-1, 0, 1\}$. Computed using step-size $h = 0.01$, noise $\sigma = 1$ and initial values $\widetilde{X}^{\mathrm{LT}}(0) = X_0 \sim \mathcal{U}(1, 2)$ for each path.



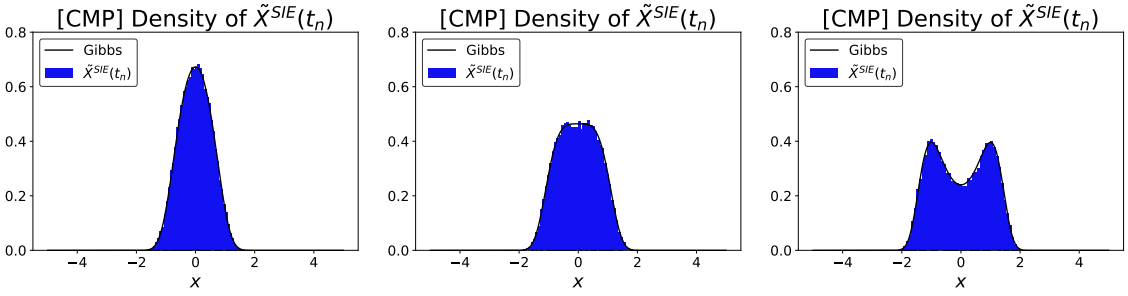
(a) $\omega = -1$ (b) $\omega = 0$ (c) $\omega = 1$

Figure 7: Histogram of $\widetilde{X}^{\mathrm{S}}(t_n = 20)$ across all $M = 50000$ simulated Brownian paths vs theoretical Gibbs density for $\omega = \{-1, 0, 1\}$. Computed using step-size $h = 0.01$, noise $\sigma = 1$ and initial values $\widetilde{X}^{\mathrm{S}}(0) = X_0 \sim \mathcal{U}(1, 2)$ for each path.



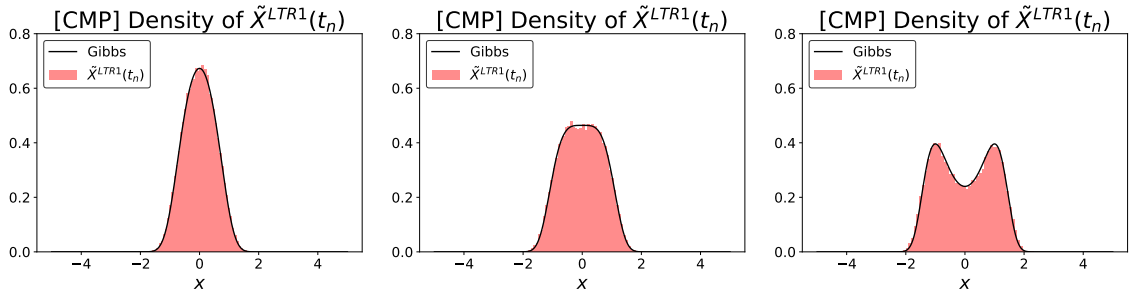
(a) $\omega = -1$ (b) $\omega = 0$ (c) $\omega = 1$

Figure 8: Histogram of $\widetilde{X}^{\mathrm{LTIE}}(t_n = 20)$ across all $M = 50000$ simulated Brownian paths vs theoretical Gibbs density for $\omega = \{-1, 0, 1\}$. Computed using step-size $h = 0.01$, noise $\sigma = 1$ and initial values $\widetilde{X}^{\mathrm{LTIE}}(0) = X_0 \sim \mathcal{U}(1, 2)$ for each path.



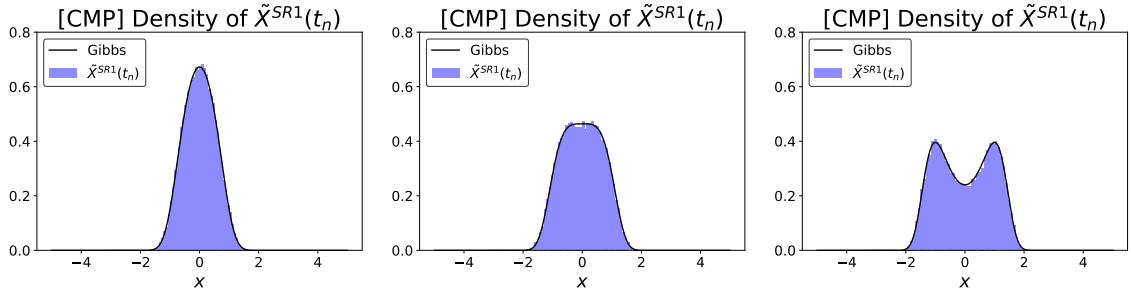
(a) $\omega = -1$ (b) $\omega = 0$ (c) $\omega = 1$

Figure 9: Histogram of $\widetilde{X}^{\mathrm{SIE}}(t_n = 20)$ across all $M = 50000$ simulated Brownian paths vs theoretical Gibbs density for $\omega = \{-1, 0, 1\}$. Computed using step-size $h = 0.01$, noise $\sigma = 1$ and initial values $\widetilde{X}^{\mathrm{SIE}}(0) = X_0 \sim \mathcal{U}(1, 2)$ for each path.



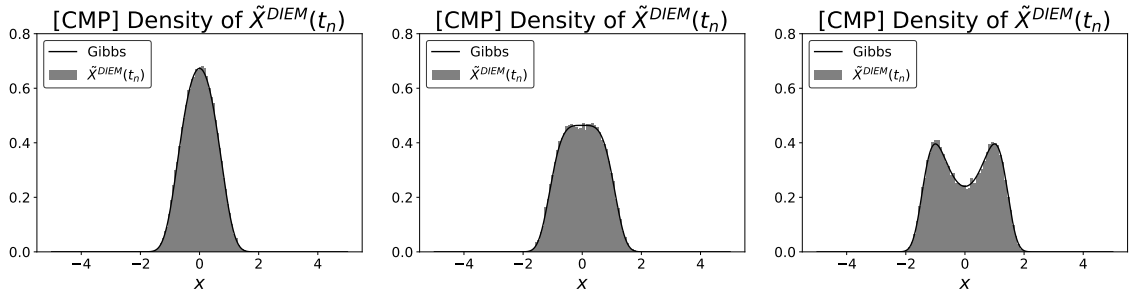
(a) $\omega = -1$ (b) $\omega = 0$ (c) $\omega = 1$

Figure 10: Histogram of $\widetilde{X}^{\mathrm{LTR1}}(t_n = 20)$ across all $M = 50000$ simulated Brownian paths vs theoretical Gibbs density for $\omega = \{-1, 0, 1\}$. Computed using step-size $h = 0.01$, noise $\sigma = 1$ and initial values $\widetilde{X}^{\mathrm{LTR1}}(0) = X_0 \sim \mathcal{U}(1, 2)$ for each path.



(a) $\omega = -1$ (b) $\omega = 0$ (c) $\omega = 1$

Figure 11: Histogram of $\widetilde{X}^{\mathrm{SR1}}(t_n = 20)$ across all $M = 50000$ simulated Brownian paths vs theoretical Gibbs density for $\omega = \{-1, 0, 1\}$. Computed using step-size $h = 0.01$, noise $\sigma = 1$ and initial values $\widetilde{X}^{\mathrm{SR1}}(0) = X_0 \sim \mathcal{U}(1, 2)$ for each path.



(a) $\omega = -1$ (b) $\omega = 0$ (c) $\omega = 1$

Figure 12: Histogram of $\widetilde{X}^{\mathrm{DIEM}}(t_n = 20)$ across all $M = 50000$ simulated Brownian paths vs theoretical Gibbs density for $\omega = \{-1, 0, 1\}$. Computed using step-size $h = 0.01$, noise $\sigma = 1$ and initial values $\widetilde{X}^{\mathrm{DIEM}}(0) = X_0 \sim \mathcal{U}(1, 2)$ for each path.

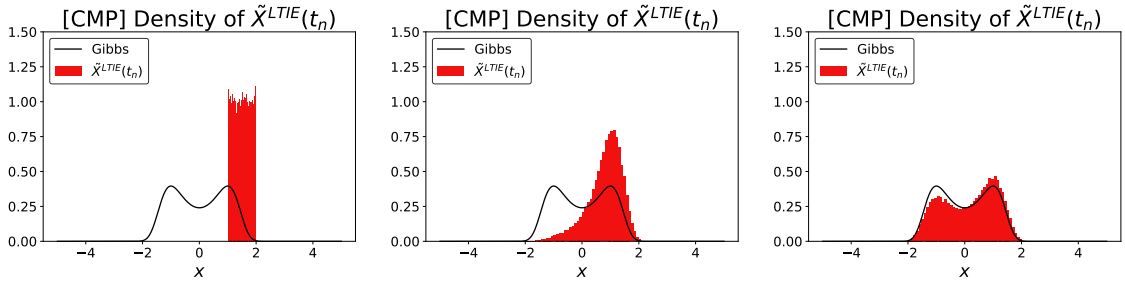
It appears that all methods are able to preserve geometric ergodicity for $\omega = \{-1, 0, 1\}$, since all methods show good correspondence between the predicted Gibbs distribution (9.14) and the distribution of the numerical solution. It is particularly interesting that geometric ergodicity of the SDE (9.1) appears preserved even when $\omega = \mu = \{0, 1\}$, since the linear SDE in this case is not geometrically ergodic nor stable. As we saw in section 7, this is justified for $\widetilde{X}^{\mathrm{LT}}(t_n)$, $\widetilde{X}^{\mathrm{S}}(t_n)$, $\widetilde{X}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{X}^{\mathrm{SIE}}(t_n)$ since the conditions $\mu < \alpha$ and $\mu < \widetilde{\alpha}_0$ can always be guaranteed by Proposition 9.3. It seems reasonable that similar justifications could be derived for $\widetilde{X}^{\mathrm{LTR1}}(t_n)$, $\widetilde{X}^{\mathrm{SR1}}(t_n)$ and $\widetilde{X}^{\mathrm{DIEM}}(t_n)$, given their apparent ability of preserving geometric ergodicity.

When performing this test, the normalization constant z of the Gibbs distribution (9.14) needs to be computed via

$$\int_{-\infty}^{\infty} e^{-2V(x)/\sigma^2} = \frac{1}{z}, \tag{9.16}$$

where $V(x)$ is given by eq. (9.15). Unfortunately, this integral does not admit an analytic expression; hence, we must compute the integral numerically to obtain an approximation of the normalization constant z. It turns out that the Gibbs distribution in eq. (9.14) features very narrow peaks for even moderately small values of σ and moderately large absolute values of ω, leading to an inaccurate approximation of the integral in eq. (9.16). For this reason, we don't test geometric ergodicity for other values of σ or ω.

Lastly, we report the distribution of $\widetilde{X}^{\mathrm{LTIE}}(t_n)$ when $\omega = 1$ for the time-points $t_n = \{0, 1, 5\}$ to see how the distribution evolves over time. The results are reported in Figure 13.



(a) $t_n = 0$ (b) $t_n = 1$ (c) $t_n = 5$

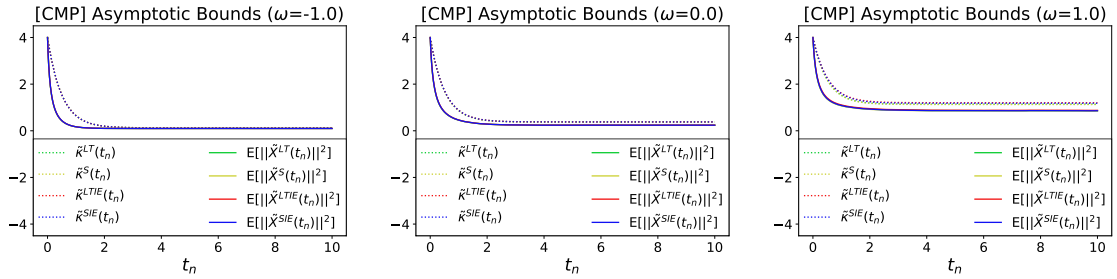
Figure 13: Histogram of $\widetilde{X}^{\mathrm{LTIE}}(t_n)$ across all $M = 50000$ simulated Brownian paths vs theoretical Gibbs density at times $t_n = \{0, 1, 5\}$ for $\omega = 1$. Computed using step-size $h = 0.01$, noise $\sigma = 1$ and initial values $\widetilde{X}^{\mathrm{LT}}(0) = X_0 \sim \mathcal{U}(1, 2)$ for each path.

We also investigate the asymptotic behavior of our methods. Recall from eq (7.6) that our splitting- and composition methods satisfy bounds of the form

$$\mathbb{E}\Big[\,\Big\|\widetilde{\boldsymbol{X}}(t_n)\Big\|^2\,\Big] \leq e^{2(\mu-a)t_n}\mathbb{E}\Big[\,\|\boldsymbol{X}_0\|^2\,\Big] + \widetilde{C}\Psi(t_n) =: \widetilde{\kappa}(t_n). \tag{9.17}$$

Appropriately inserting for $a = \{\alpha, \widetilde{\alpha}_0\}$ and $\widetilde{C} = \{\widetilde{C}^{\mathrm{LT}}, \widetilde{C}^{\mathrm{S}}, \widetilde{C}^{\mathrm{LTIE}}, \widetilde{C}^{\mathrm{SIE}}\}$ then yields bounds $\widetilde{\kappa}(t_n) = \{\widetilde{\kappa}^{\mathrm{LT}}(t_n), \widetilde{\kappa}^{\mathrm{S}}(t_n), \widetilde{\kappa}^{\mathrm{LTIE}}(t_n), \widetilde{\kappa}^{\mathrm{SIE}}(t_n)\}$ for our methods for all $t_n \geq 0$. We perform a numerical experiment to investigate whether or not the bound (9.17) holds for the methods $\widetilde{X}^{\mathrm{LT}}(t_n)$, $\widetilde{X}^{\mathrm{S}}(t_n)$, $\widetilde{X}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{X}^{\mathrm{SIE}}(t_n)$. We do not consider $\widetilde{X}^{\mathrm{LTR1}}(t_n)$, $\widetilde{X}^{\mathrm{SR1}}(t_n)$ nor $\widetilde{X}^{\mathrm{DIEM}}(t_n)$ since we have not derived asymptotic bounds for these methods.

We perform this test by simulating the solution from each considered method $M = 50000$ times, and computing the mean-square norm of the solution of each method as a function of time. We then compare this with the predicted bound from eq. (9.17). We test this for $X_0 = 2, T = 10$, $h = 0.01$, $\sigma = 0.5$ and $\omega = \{-1, 0, 1\}$. Note that for each value of ω, we must choose a value of α such that $\mu < \alpha$ and $\mu < \widetilde{\alpha}_0$ is guaranteed for our choice of step-size $h = 0.01$. It follows that choosing $\alpha = \{0, 1, 2\}$ for $\omega = \{-1, 0, 1\}$, respectively, satisfies $\mu < \alpha$ and $\mu < \widetilde{\alpha}_0$. The results of this test are reported in Figure 14.

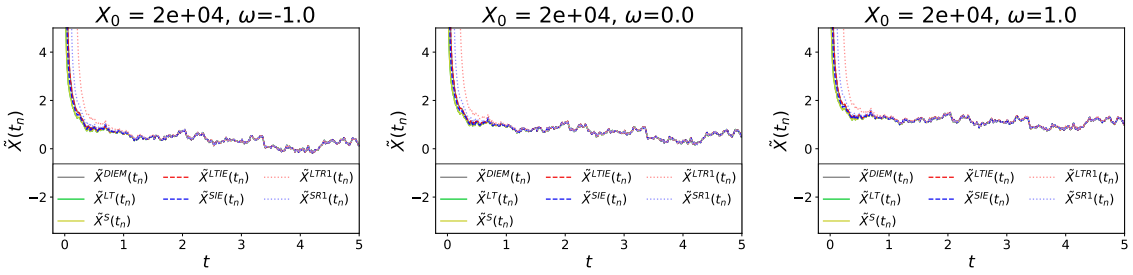


(a) $\omega = -1, \quad \alpha = 0$ (b) $\omega = 0, \quad \alpha = 1$ (c) $\omega = 1, \quad \alpha = 2$

Figure 14: Experimentally observed asymptotic behavior of the mean-square norm of $\widetilde{X}^{\mathrm{LT}}(t_n)$, $\widetilde{X}^{\mathrm{S}}(t_n)$, $\widetilde{X}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{X}^{\mathrm{SIE}}(t_n)$ vs their predicted upper bounds for $\omega = \{-1, 0, 1\}$. For all $M = 50000$ simulated paths we set $X_0 = 2$, $T = 10$, $h = 0.01$ and $\sigma = 0.5$.

We observe good correspondence between the experimentally observed values of the mean-square norm of $\widetilde{X}^{\mathrm{LT}}(t_n)$, $\widetilde{X}^{\mathrm{S}}(t_n)$, $\widetilde{X}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{X}^{\mathrm{SIE}}(t_n)$, and their predicted upper bounds $\widetilde{\kappa}^{\mathrm{LT}}(t_n)$, $\widetilde{\kappa}^{\mathrm{S}}(t_n)$, $\widetilde{\kappa}^{\mathrm{LTIE}}(t_n)$, $\widetilde{\kappa}^{\mathrm{SIE}}(t_n)$, respectively. Note that for $\omega = \{-1, 0\}$, the experimentally observed values and predicted bounds are essentially indistinguishable from one method to the next. Observe also that the bound is particularly tight for $\omega = -1$. When $\omega = 1$, we notice a slight difference between both the experimentally observed values and the predicted bounds from one method to the next. This discrepancy is no longer noticeable when decreasing the step-size to $h = 0.001$.

Lastly, we investigate the effect of a large initial value. For this experiment, we set $X_0 = 2 \cdot 10^4$. We let $h = 0.01$, $T = 5$ and $\sigma = 0.5$, and consider $\omega = \{-1, 0, 1\}$. We perform this test for all methods $\widetilde{X}^{\mathrm{LT}}(t_n)$, $\widetilde{X}^{\mathrm{S}}(t_n)$, $\widetilde{X}^{\mathrm{LTIE}}(t_n)$, $\widetilde{X}^{\mathrm{SIE}}(t_n)$, $\widetilde{X}^{\mathrm{LTR1}}(t_n)$, $\widetilde{X}^{\mathrm{SR1}}(t_n)$ and $\widetilde{X}^{\mathrm{DIEM}}(t_n)$. The results are seen in Figure 15.



(a) $\omega = -1$ (b) $\omega = 0$ (c) $\omega = 1$

Figure 15: Plot of all methods along a single simulated Brownian path with large initial value $X_0 = 2 \cdot 10^4$, where $T = 5$, $\sigma = 0.5$ and $h = 0.01$.

All methods appear to converge towards the same solution for all $\omega = \{-1, 0, 1\}$, even for this very large initial value. There is some discrepancy between our methods in the initial decay of the solution, particularly for the LTR1- and SR1 composition methods, but this discrepancy quickly

resolves. The discrepancy in the initial decay is hardly noticeable when reducing the step-size to $h = 0.001$. It is unsurprising that the LTR1- and SR1 methods are the worst performing methods for this test, since they represent a linearization of the implicit Euler solution. For the rapid initial decay exhibited by this problem when X_0 is large, this linearization has considerable error.

10 FitzHugh-Nagumo model

We will in this section investigate our splitting- and composition methods applied to the stochastic FitzHugh-Nagumo (FHN) model, which is a widely used neuronal model used to model the firing activity of single neurons. For a more in-depth discussion of this model, see e.g. [14] and [15]. This model was also investigated in [1].

The FHN model is given by the following 2-dimensional semi-linear SDE:

$$d\begin{bmatrix} X_1(t) \\ X_2(t) \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{1}{\varepsilon}\big(X_1(t) - X_1^3(t) - X_2(t)\big) \\ \gamma X_1(t) - X_2(t) + \beta \end{bmatrix}}_{:=\boldsymbol{F}\left(\boldsymbol{X}(t)\right)} dt + \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} d\boldsymbol{W}(t), \quad \boldsymbol{X}(0) = \boldsymbol{X}_0, \quad t \in [0,T], \tag{10.1}$$

with solution $\boldsymbol{X}(t) = \big(X_1(t), X_2(t)\big)^\top$ for $t \in [0,T]$. Here $\varepsilon > 0$, $\gamma > 0$ and $\beta \geq 0$ are parameters which describe the neuronal activity. In [1], the following split of $\boldsymbol{F}\big(\boldsymbol{X}(t)\big)$ was considered:

$$\mathbf{A} = \begin{bmatrix} 0 & -\frac{1}{\varepsilon} \\ \gamma & -1 \end{bmatrix}, \qquad \boldsymbol{B}\big(\boldsymbol{X}(t)\big) = \begin{bmatrix} \frac{1}{\varepsilon}\big(X_1(t) - X_1^3(t)\big) \\ \beta \end{bmatrix} \tag{10.2}$$

For this split, mean-square convergence for Lie-Trotter splitting could be proved for all $\beta \geq 0$. For Strang splitting, however, mean-square convergence could only be proved for $\beta = 0$. Numerical demonstrations nevertheless indicated that mean-square boundedness was satisfied in the case of Strang splitting for any $\beta \geq 0$. Moreover, preservation of geometric ergodicity for Strang- and Lie-Trotter splitting could only be proved under $\beta = 0$ and $\gamma = \frac{1}{\varepsilon}$. We remark that, in [1], preservation of geometric ergodicity for the Strang- and Lie-Trotter splitting methods was guaranteed by the condition $\left\|e^{\mathbf{A}h}\right\| < 1$, which holds for the split (10.2) for all $h > 0$ when $\beta = 0$ and $\gamma = 1/\varepsilon$. Lastly, asymptotic bounds on the mean-square norm of the solution could not be derived for this split, since the split (10.2) with $\beta = 0$ and $\gamma = 1/\varepsilon$ yields $\mu = 0$. In the theory of [1], $\mu < 0$ was required to derive such asymptotic bounds. Note that $\mu < 0$ implies $\left\|e^{\mathbf{A}h}\right\| < 1$, but the converse does not hold.

It can be shown that the split (10.2) can only satisfy Assumption 3.3 when $\alpha < 0$. This is a problem, since for this split, $\mu \geq 0$ for all $\varepsilon > 0$ and $\gamma > 0$. Thus, there is no choice of parameters which will yield $\mu < \alpha$ for the split (10.2), and geometric ergodicity for this split cannot be established within the theory of this thesis.

The following split was also briefly discussed in [1]:

$$\mathbf{A} = \begin{bmatrix} 0 & -\frac{1}{\varepsilon} \\ \gamma & 0 \end{bmatrix}, \qquad \boldsymbol{B}\big(\boldsymbol{X}(t)\big) = \begin{bmatrix} \frac{1}{\varepsilon}\big(X_1(t) - X_1^3(t)\big) \\ \beta - X_2(t) \end{bmatrix} \tag{10.3}$$

It was remarked that for this split, mean-square convergence could be established for both splitting methods under $\beta = 0$. However, geometric ergodicity of this split could not be established since $\left\|e^{\mathbf{A}h}\right\| \geq 1$. Employing the theory of this thesis, however, we are able to prove for the split (10.3)

(1) mean-square convergence for both splitting- and composition methods with arbitrary $\beta \geq 0$ and (2) preservation of geometric ergodicity for arbitrary $\beta \geq 0$ and arbitrary $\varepsilon > 0$ with $|\gamma - \frac{1}{\varepsilon}| < 2$.

We start by proving that Assumptions 3.1, 3.2 and 3.3 holds for the function $\boldsymbol{B}$ given by the split (10.3).

Proposition 10.1 *The function* $\boldsymbol{B}$ *from eq.* (10.3) *satisfies Assumption 3.1 for* $c_1 = \frac{1}{\varepsilon}$.

Proof: We find

$$\begin{aligned}
\big\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{B}(\boldsymbol{x}) - \boldsymbol{B}(\boldsymbol{y}) \big\rangle &= (x_1 - y_1)\big(B_1(\boldsymbol{x}) - B_1(\boldsymbol{y})\big) + (x_2 - y_2)\big(B_2(\boldsymbol{x}) - B_2(\boldsymbol{y})\big) \\
&= \frac{1}{\varepsilon}(x_1 - y_1)(x_1 - y_1 + y_1^3 - x_1^3) - (x_2 - y_2)^2 \\
&= \frac{1}{\varepsilon}(x_1 - y_1)^2 - (x_2 - y_2)^2 + \frac{1}{\varepsilon}(x_1 - y_1)(y_1^3 - x_1^3) \\
&\leq \frac{1}{\varepsilon}\Big((x_1 - y_1)^2 + (x_2 - y_2)^2\Big) + \frac{1}{\varepsilon}(x_1 - y_1)(y_1^3 - x_1^3) \\
&= \frac{1}{\varepsilon}\left\|\boldsymbol{x} - \boldsymbol{y}\right\|^2 + \frac{1}{\varepsilon}(x_1 - y_1)(y_1^3 - x_1^3) \\
&= \frac{1}{\varepsilon}\left\|\boldsymbol{x} - \boldsymbol{y}\right\|^2 - \frac{1}{\varepsilon}(x_1 - y_1)^2(x_1^2 + x_1 y_1 + y_1^2) \\
&\leq \frac{1}{\varepsilon}\left\|\boldsymbol{x} - \boldsymbol{y}\right\|^2 . \quad \blacksquare
\end{aligned}$$

Proposition 10.2 *The function* $\boldsymbol{B}$ *from eq.* (10.3) *satisfies Assumption 3.2 for* $c_2 = 1 + 8/\varepsilon^2$ *and* $\chi = 3$.

Proof: We find

$$\begin{aligned}
\|\boldsymbol{B}(\boldsymbol{x}) - \boldsymbol{B}(\boldsymbol{y})\|^2 &= \frac{1}{\varepsilon^2}\Big(x_1 - y_1 + y_1^3 - x_1^3\Big)^2 + (x_2 - y_2)^2 \\
&= \frac{1}{\varepsilon^2}\Big(- (x_1 - y_1)(x_1^2 + x_1 y_1 + y_1^2 - 1)\Big)^2 + (x_2 - y_2)^2 \\
&= \frac{1}{\varepsilon^2}\Big((x_1 - y_1) - (x_1 - y_1)(x_1^2 + x_1 y_1 + y_1^2)\Big)^2 + (x_2 - y_2)^2 \\
&= \frac{1}{\varepsilon^2}\Big((x_1 - y_1)^2 + (x_1 - y_1)^2(x_1^2 + x_1 y_1 + y_1^2)^2 - 2(x_1 - y_1)^2(x_1^2 + x_1 y_1 + y^2)\Big) \\
&\qquad + (x_2 - y_2)^2 \\
&= \frac{1}{\varepsilon^2}(x_1 - y_1)^2\Big(1 + (x_1^2 + x_1 y_1 + y_1^2)^2 - 2(x_1^2 + x_1 y_1 + y_1^2)\Big) + (x_2 - y_2)^2 \\
&\le \frac{1}{\varepsilon^2}(x_1 - y_1)^2\Big(1 + (x_1^2 + x_1 y_1 + y_1^2)^2\Big) + (x_2 - y_2)^2 \\
&\le \frac{1}{\varepsilon^2}(x_1 - y_1)^2(1 + 8x_1^4 + 8y_1^4) + (x_2 - y_2)^2 \\
&\le \frac{1}{\varepsilon^2}(x_1 - y_1)^2(8 + 8x_1^4 + 8x_2^4 + 8y_1^4 + 8y_2^4) + (x_2 - y_2)^2 \\
&\le \frac{8}{\varepsilon^2}(x_1 - y_1)^2(1 + \|\boldsymbol{x}\|^4 + \|\boldsymbol{y}\|^4) + (x_2 - y_2)^2 \\
&\le \left(1 + \frac{8}{\varepsilon^2}\right)(1 + \|\boldsymbol{x}\|^4 + \|\boldsymbol{y}\|^4)\big((x_1 - y_1)^2 + (x_2 - y_2)^2\big) \\
&= \left(1 + \frac{8}{\varepsilon^2}\right)(1 + \|\boldsymbol{x}\|^4 + \|\boldsymbol{y}\|^4)\,\|\boldsymbol{x} - \boldsymbol{y}\|^2 . \quad \blacksquare
\end{aligned}$$

Proposition 10.3 *The function $\boldsymbol{B}$ satisfies Assumptions 3.3 for any $\alpha < 1$ and*

$$K = \frac{1}{4}\left(\frac{(\alpha\varepsilon - 1)^2}{\varepsilon} + \frac{\beta^2}{1 - \alpha}\right), \quad \beta \ge 0, \quad \varepsilon > 0.$$

Proof: We find

$$\big\langle \boldsymbol{x}, \boldsymbol{B}(\boldsymbol{x})\big\rangle + \alpha\,\|\boldsymbol{x}\|^2 = \frac{1}{\varepsilon}x_1^2 - \frac{1}{\varepsilon}x_1^4 + \beta x_2 - x_2^2 + \alpha x_1^2 + \alpha x_2^2. \tag{10.4}$$

We consider the terms involving x_1 and x_2 separately, and investigate their global maxima, should they exist. Defining the two functions

$$f_1(x_1) = \frac{1}{\varepsilon}x_1^2 - \frac{1}{\varepsilon}x_1^4 + \alpha x_1^2, \tag{10.5a}$$

$$f_2(x_2) = \beta x_2 - x_2^2 + \alpha x_2^2, \tag{10.5b}$$

we find their stationary points by solving the equations $d/dx_i f_i(x_i) = 0$ for $i = \{1, 2\}$. Doing this for $f_1(x_1)$ yields stationary points at $x = 0$ and $x = \pm\sqrt{\frac{\alpha\varepsilon+1}{2}}$, whereupon $x = 0$ is revealed to be a local minimum and $x = \pm\sqrt{\frac{\alpha\varepsilon+1}{2}}$ are global maxima. It follows, then, that the maximal value of $f_1(x_1)$ is given by

$$\max_{x_1 \in \mathbb{R}} f_1(x_1) = \frac{(\alpha\varepsilon + 1)^2}{4\varepsilon},$$

attained at both $x = \pm\sqrt{\frac{\alpha\varepsilon+1}{2}}$. Moving on, we find that $f_2(x_2)$ has a stationary point at $x = \frac{\beta}{2(1-\alpha)}$. This stationary point corresponds to a global maxima if and only if $\alpha < 1$; for $\alpha = 1$, $f_2(x_2)$ is a linear function and $\alpha > 1$ yields a convex parabola, neither of which can be bounded from above. Assuming $\alpha < 1$, it follows that the maximum value attained by $f_2(x_2)$ is given by

$$\max_{x_2 \in \mathbb{R}} f_2(x_2) = \frac{\beta^2}{4(1-\alpha)}.$$

Thus, we find that the choice of $\boldsymbol{B}$ in (10.3) satisfies the dissipativity condition 3.3 with

$$K = \frac{1}{4}\left(\frac{(\alpha\varepsilon+1)^2}{\varepsilon} + \frac{\beta^2}{1-\alpha}\right) > 0, \quad \alpha < 1, \quad \varepsilon > 0, \quad \beta \geq 0. \quad \blacksquare \tag{10.6}$$

The exact solution of the nonlinear ODE (4.4a) at time t_n starting from $\boldsymbol{X}^{(1)}(t_{n-1})$, with $\boldsymbol{B}$ given by eq. (10.3), is given by

$$\boldsymbol{\varphi}_h^{(1)}\big(\boldsymbol{X}^{(1)}(t_{n-1})\big) = \begin{bmatrix} \frac{X_1^{(1)}(t_{n-1})}{\sqrt{e^{-2\varepsilon/h} + (1-e^{-2h/\varepsilon})\big(X^{(1)}(t_{n-1})\big)^2}} \\ e^{-h}\big(X_2^{(1)}(t_{n-1}) - \beta\big) + \beta \end{bmatrix}, \quad \boldsymbol{X}^{(1)}(0) = \boldsymbol{X}_0. \tag{10.7}$$

The implicit Euler solution is given at every discrete time-point t_n as

$$\widetilde{\boldsymbol{\varphi}}_h^{(1)}\big(\widetilde{\boldsymbol{X}}^{(1)}(t_{n-1})\big) = \begin{bmatrix} X_1^* \\ X_2^*, \end{bmatrix}, \quad \widetilde{\boldsymbol{X}}^{(1)}(0) = \boldsymbol{X}_0, \tag{10.8}$$

where $\boldsymbol{X}^* = (X_1^*, X_2^*)^\top$ satisfies the implicit equation

$$\begin{bmatrix} X_1^* \\ X_2^* \end{bmatrix} = \begin{bmatrix} \widetilde{X}_1^{(1)}(t_{n-1}) - \frac{h}{\varepsilon}\big(X_1^* - (X_1^*)^3\big) \\ \widetilde{X}_2^{(1)}(t_{n-1}) + h\beta - hX_2^* \end{bmatrix} \tag{10.9}$$

Note that the equations for X_1^* and X_2^* are decoupled, so we can solve them individually. The equation for X_1^* can be solved in the same way as for the cubic model problem, since it can also be written as a depressed cubic equation

$$X_1^* + a(X_1^*)^3 = b,$$

with $a = \frac{\varepsilon}{h} - 1$ and $b = \varepsilon\widetilde{X}_1^{(1)}(t_{n-1})/h$. The equation for X_2^* is easily solved to yield

$$X_2^* = \frac{\widetilde{X}_2^{(1)}(t_{n-1}) + h\beta}{1+h}.$$

Lastly, the Jacobian of $\boldsymbol{B}$ with $\boldsymbol{B}$ defined through eq. (10.3) is diagonal. Thus, the Newton matrix of eq. (8.2) is easily invertible, and the first-order Rosenbrock solution of the nonlinear ODE is given at every discrete time-point by

$$\widehat{\boldsymbol{\varphi}}_h^{(1)}\big(\widehat{\boldsymbol{X}}^{(1)}(t_{n-1})\big) = \begin{bmatrix} \widehat{X}_1^{(1)}(t_{n-1}) + \frac{\frac{h}{\varepsilon}\left(\widehat{X}_1^{(1)}(t_{n-1}) - \left(\widehat{X}_1^{(1)}(t_{n-1})\right)^3\right)}{1 - \frac{h}{\varepsilon}\left(1 - 3\left(\widehat{X}_1^{(1)}(t_{n-1})\right)^2\right)} \\ \widehat{X}_2^{(1)}(t_{n-1}) + \frac{h(\beta - \widehat{X}_2^{(1)}(t_{n-1}))}{1+h}. \end{bmatrix}, \quad \widehat{\boldsymbol{X}}^{(1)}(0) = \boldsymbol{X}_0. \tag{10.10}$$

Note that for the above solution to be defined everywhere, we require $0 < \frac{h}{\varepsilon} < 1$.

10.a Mean-square convergence

We test mean-square convergence of the methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$ applied to the FHN model (10.1) using the split of (10.3). We also include the results for the DIEM method $\widetilde{\boldsymbol{X}}^{\mathrm{DIEM}}(t_n)$ for comparison purposes. We test mean-square convergence in the same manner as described in section 9, using $M = 1000$ simulated paths, computing the reference solution using Strang splitting with step-size $h = 2^{-14}$ and testing our methods for step-sizes 2^{-q} with $q = \{12, 11, 10, 9, 8, 7, 6\}$. We set $\varepsilon = \beta = \gamma = 1$ and let the initial value and end-time be $\boldsymbol{X}_0 = (2, 0)^\top$ and $T = 10$, respectively. The noise is set to $\sigma_1 = \sigma_2 = 0.5$. We report the results in Figure 16.

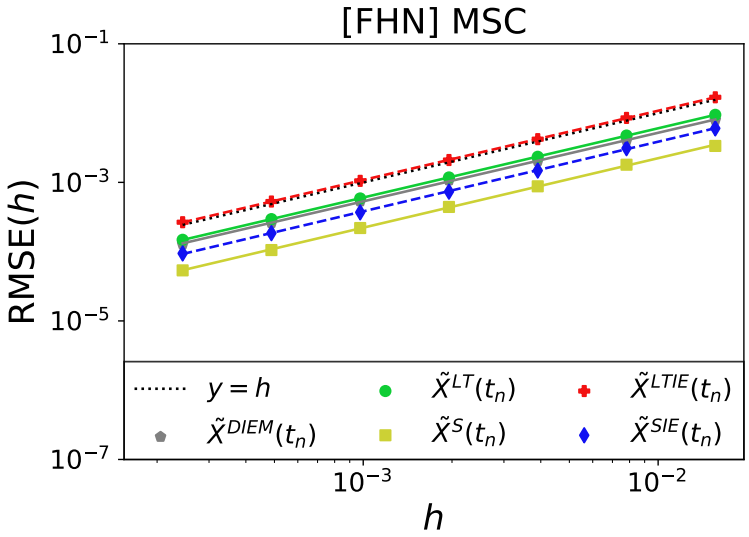


Figure 16: FHN model; mean-square convergence (msc) test using $M = 1000$ simulated Brownian paths with initial value $\boldsymbol{X}_0 = (2, 0)^\top$, end-time $T = 10$ and step-sizes $h = 2^{-q}$ with $q = \{12, 11, 10, 9, 8, 7, 6\}$. We set $\varepsilon = \gamma = \beta = 1$ and $\sigma_1 = \sigma_2 = 0.5$

Mean-square convergence of order $p = 1$ is observed for all considered methods. As for the cubic model problem, it appears that the Strang compositional approach outperforms the Lie-Trotter compositional approach. For a given compositional approach, we observe that the splitting methods outperform the composition methods. Lastly, we observe that the DIEM method performs roughly at the level of the Lie-Trotter splitting method.

10.b Preservation of geometric ergodicity

We proceed by investigating the proposed methods' ability to preserve geometric ergodicity for the FHN model using the split (10.3). Recall from Proposition 10.3 that the function $\boldsymbol{B}$ from eq. (10.3) satisfies Assumption 3.3 for $\alpha < 1$. Thus, our system is geometrically ergodic if $\mu < \alpha < 1$, which in practice reduces to $\mu < 1$ since, for any $\mu < 1$, there exists $\alpha \in (\mu, 1)$. From a simple eigenvalue

computation of $(\mathbf{A}+\mathbf{A}^\top)/2$ with $\mathbf{A}$ given by eq. (10.3), we find that the logarithmic norm of $\mathbf{A}$ is given by

$$\mu = \max\left\{ \pm \frac{\varepsilon\gamma - 1}{2\varepsilon} \right\} = \begin{cases} \frac{\varepsilon\gamma-1}{2\varepsilon}, & \text{if } \gamma \geq \frac{1}{\varepsilon} \\ \frac{1-\varepsilon\gamma}{2\varepsilon}, & \text{if } \gamma < \frac{1}{\varepsilon} \end{cases}$$

Under the constraint that $\mu < 1$, it follows immediately that we obtain a geometrically ergodic system for any $\varepsilon > 0$ and γ satisfying $|\gamma - \frac{1}{\varepsilon}| < 2$.

We test preservation of geometric ergodicity for all methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{LTIE}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{SIE}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{LTR1}}(t_n)$, $\widetilde{\boldsymbol{X}}^{\mathrm{SR1}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{DIEM}}(t_n)$. This test is performed in the same way as for the cubic model problem of section 9: We compute the solution to each method across $M = 50000$ simulated Brownian paths. For each method, we consider the distribution of the $M = 50000$ solutions at time $t_n = 20$, with initial values drawn from a uniform distribution over the interval $(1, 2)$. The step-size is $h = 0.01$, and we set $\varepsilon = 1$, $\gamma = 2$ and $\beta = 2$. Lastly, the noise is set to $\sigma_1 = \sigma_2 = 0.5$. The results for the first component of the solution are reported in Figures 17, 18 and 19. All methods appear to approach the same limiting distribution. We don't report the results for the second component, though this component also appears to approach the same limiting distribution for all methods (a different distribution from that of the first component, though).

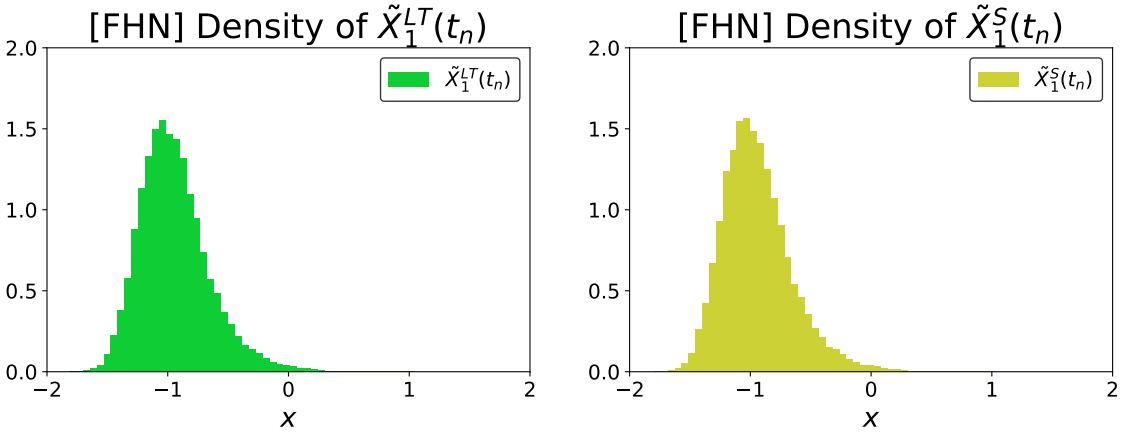


Figure 17: Histograms of $\widetilde{X}_1^{\mathrm{LT}}(t_n = 20)$ and $\widetilde{X}_1^{\mathrm{S}}(t_n = 20)$ across all $M = 50000$ simulated Brownian paths. Computed using step-size $h = 0.01$, noise $\sigma_1 = \sigma_2 = 0.5$ and initial values $X_0 \sim \mathcal{U}(1, 2)$.

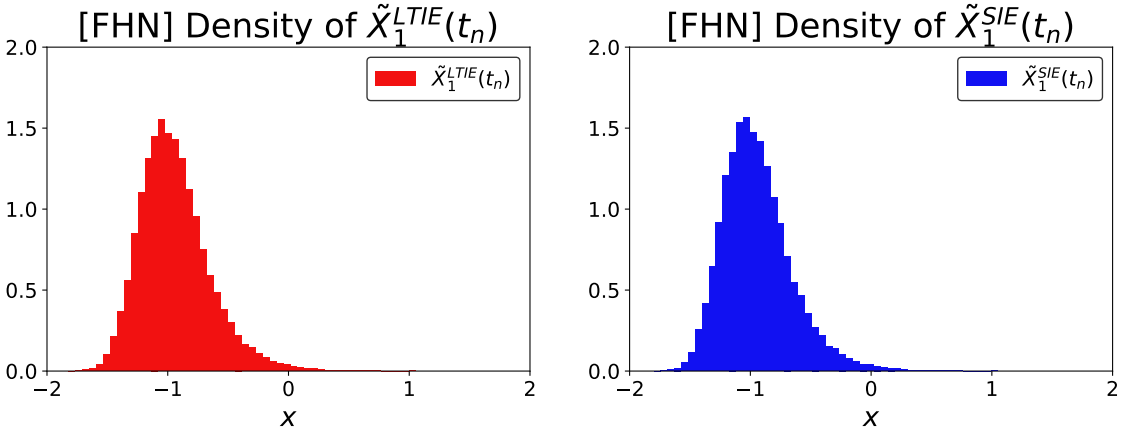


Figure 18: Histograms of $\widetilde{X}_1^{\mathrm{LTIE}}(t_n = 20)$ and $\widetilde{X}_1^{\mathrm{SIE}}(t_n = 20)$ across all $M = 50000$ simulated Brownian paths. Computed using step-size $h = 0.01$, noise $\sigma_1 = \sigma_2 = 0.5$ and initial values $X_0 \sim \mathcal{U}(1, 2)$.

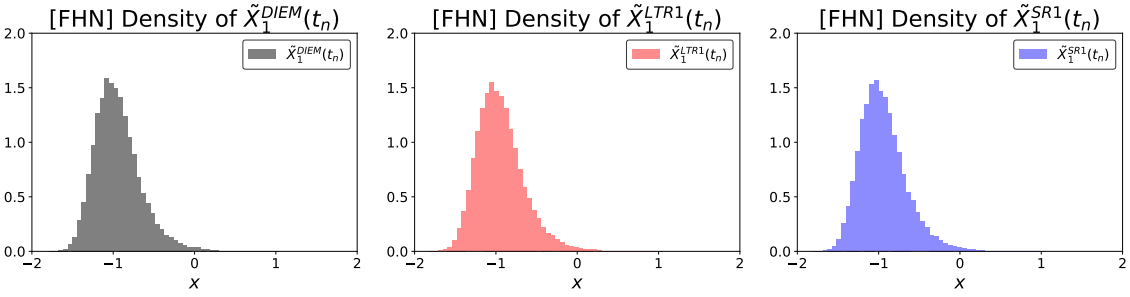
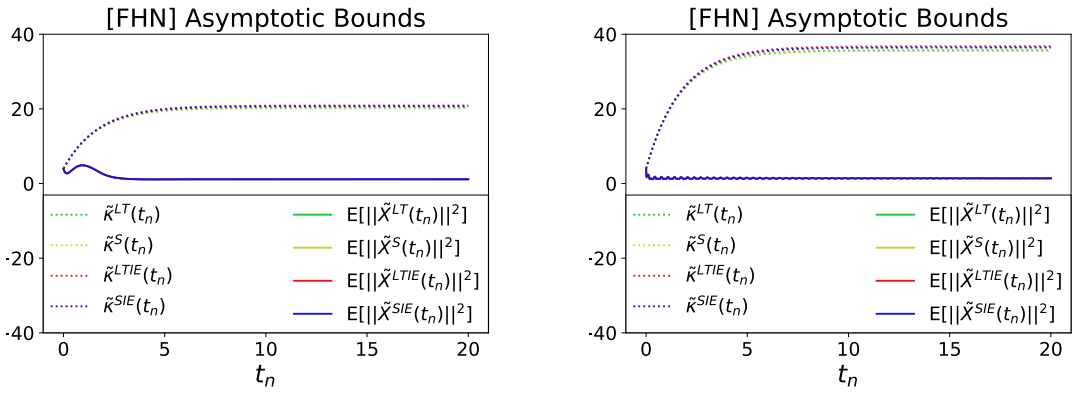


Figure 19: Histograms of $\widetilde{X}_1^{\mathrm{DIEM}}(t_n = 20)$, $\widetilde{X}_1^{\mathrm{LTR1}}(t_n = 20)$ and $\widetilde{X}_1^{\mathrm{SR1}}(t_n = 20)$ across all $M = 50000$ simulated Brownian paths. Computed using step-size $h = 0.01$, noise $\sigma_1 = \sigma_2 = 0.5$ and initial values $X_0 \sim \mathcal{U}(1, 2)$.

We proceed by considering asymptotic bounds of the mean-square norm of the solution for all methods, as we did for the cubic model problem of section 9. Again, the theoretically predicted bounds are given in general by eq. (7.6), which reduce to bounds for each method. We simulate $M = 50000$ Brownian paths, and compute the mean-square norm of the solution across all paths for each method. The initial value, end-time, noise level and step-size are given by $\boldsymbol{X}_0 = (2, 0)^\top$, $T = 20$, $\sigma_1 = \sigma_2 = 0.5$ and $h = 0.01$, respectively. We perform this test for $\varepsilon = \{1, 0.05\}$ with $\gamma = \frac{1}{\varepsilon} + 1$ and $\beta = 2$. For these parameter values, the logarithmic norm μ is given by $\mu = 0.5$. Lastly, we use $\alpha = 0.8$ for this test.



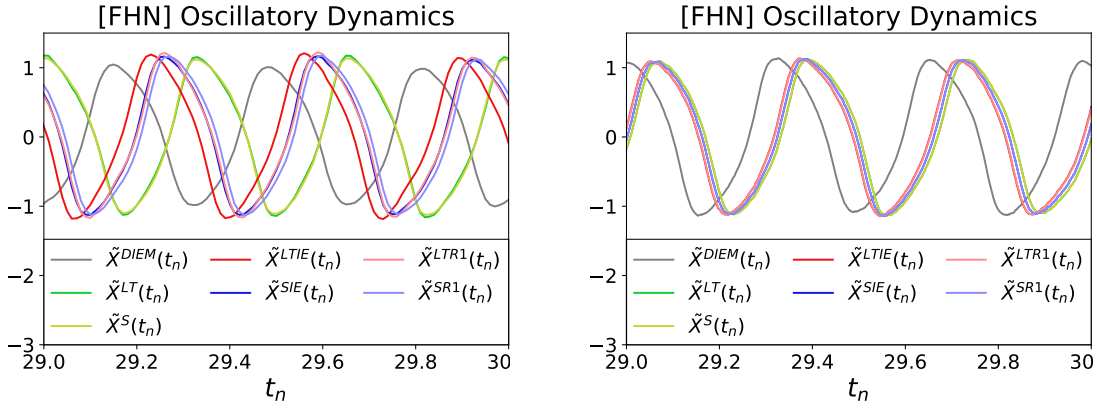
(a) $\varepsilon = 1$ (b) $\varepsilon = 0.05$

Figure 20: Experimentally observed asymptotic behavior of the mean-square norm of $\widetilde{X}^{\mathrm{LT}}(t_n)$, $\widetilde{X}^{\mathrm{S}}(t_n)$, $\widetilde{X}^{\mathrm{LTIE}}(t_n)$ and $\widetilde{X}^{\mathrm{SIE}}(t_n)$ vs their predicted upper bounds for $\varepsilon = \{1, 0.05\}$ with $\gamma = 1/\varepsilon + 1$ and $\beta = 2$. For all $M = 50000$ simulated paths we set $\boldsymbol{X}_0 = (2, 0)^\top$, $T = 20$, $h = 0.01$ and $\sigma_1 = \sigma_2 = 0.5$

The theoretically predicted bounds appear to hold, though they significantly over-estimate the true mean-square norm of the solution to our methods compared to the cubic model problem of section 9. This is due to the fact that the constant K from Assumption 3.3 becomes quite large for the FHN model with the split (10.3) and our chosen values of ε, β and γ, as is evident from eq. (10.6). Moreover, observe that when $\varepsilon = 0.05$, the observed mean-square norm appears to oscillate at first, though this oscillation appears to decay with time.

10.c Oscillatory Dynamics

Lastly, we investigate our proposed methods' ability of preserving oscillatory dynamics. This was also investigated in [1], where the Lie-Trotter- and Strang splitting methods indeed appeared to preserve such dynamics. We test this by setting $\varepsilon = 0.05$, $\gamma = 20$, $\beta = 0.1$, $\sigma_1 = 0.1$ and $\sigma_2 = 0.2$; a parameter choice which in [1] yielded oscillating solutions. We perform this test by computing the solution for all methods across a single simulated Brownian path with end-time $T = 30$ and initial value $X = (-1, 0)$. We then observe how the phase, amplitude and frequency of the first component of the solutions evolve over time. The results are reported in Figure 21 for step-sizes $h = 0.01$ and $h = 0.001$.



(a) $h = 0.01$

(b) $h = 0.001$

Figure 21: Oscillatory dynamics of our methods: We compute the solution for each method along a single simulated path, and consider the first component of each solution. Parameters are set to $\varepsilon = 0.05$, $\gamma = 20$, $\beta = \sigma_1 = 0.1$ and $\sigma_2 = 0.2$ with initial value $\boldsymbol{X}_0 = (-1, 0)^\top$, end-time $T = 30$ and step-sizes $h = \{0.01, 0.001\}$.

In agreement with [1], the splitting methods $\widetilde{\boldsymbol{X}}^{\mathrm{LT}}(t_n)$ and $\widetilde{\boldsymbol{X}}^{\mathrm{S}}(t_n)$ appear to preserve phase, amplitude and frequency, since they coincide at every time-step. By contrast, none of our composition methods appear to preserve phase, though amplitude and frequency appear to be preserved. This is presumably due to the fact that the composition methods only solve the nonlinear ODE approximately, but it is nevertheless interesting that this error appears to only manifest as error in the phase of the solution, and not in the frequency or amplitude. It appears that the DIEM method preserves neither phase nor amplitude, and the error in phase appears larger than for the composition methods. This is especially apparent for the step-size $h = 0.001$.

11 Conclusion

Splitting methods has recently been proven in Buckwar et. al. [1] to be fast and reliable methods for approximating the solutions to semi-linear SDEs with additive noise and globally one-sided Lipschitz-continuous drift, where the drift is allowed polynomial growth at infinity. In this context, the semi-linear SDE is split into a linear SDE and a nonlinear ODE, both of which are solved exactly. The exact solutions are then composed according to the Lie-Trotter and Strang approaches to yield a solution of the overall semi-linear SDE. Depending on the compositional approach used, this strategy gives rise to the Lie-Trotter- and Strang splitting methods, which remained the focus of the work of Buckwar et. al.

Following their success, this thesis set out to generalize their work by considering the case where the nonlinear ODE is solved approximately, which is often required in applications. In such a scenario, our strategy is referred to as a composition method. In particular, we consider the case where the nonlinear ODE is solved using the implicit Euler method. This gives rise to the Lie-Trotter implicit Euler (LTIE) and Strang implicit Euler (SIE) composition methods. The

goal of this thesis was to investigate mean-square convergence and structure preservation of these composition methods.

By introducing an additional condition (Assumption 3.3) on the nonlinear part of the drift, we were able to prove that our composition methods are mean-square convergent of order $p = 1$. This was facilitated by the fact that this additional condition yields global bounds on the exact- and implicit Euler solutions to the nonlinear ODE. For completeness, we also provided a proof of mean-square convergence of the splitting methods under this additional condition. Using this condition, we were further able to prove that our splitting- and composition methods preserve geometric ergodicity, even in cases where the linear SDE itself is not geometrically ergodic, provided that the overall semi-linear SDE is. This result could not be attained using the theory of Buckwar et. al. Lastly, we derived asymptotic bounds on the mean-square norm of the solutions to our splitting- and composition methods, even for cases where the linear SDE is non-stable, provided that the overall semi-linear SDE is. This is also an extension of the work of Buckar et. al.

These results were confirmed by numerical experiments. In particular, we demonstrated our results on a cubic model problem and the FitzHugh-Nagumo (FHN) model. In the cubic model problem, it was also shown that the composition methods remained stable even when the initial value is very large. Under the FHN model, Assumption 3.3 allowed us to show preservation of geometric ergodicity for a larger parameter space than in [1]. Our experiments using the FHN model also showed that our composition methods could not completely preserve oscillatory dynamics of the SDE; in particular, it appears that phase is not preserved. Though frequency and amplitude appears to be preserved, more research is needed to answer this definitively.

If the implicit equation associated with the implicit Euler method can be solved explicitly for the updated value of the solution, then our composition methods retain the computational efficiency of their splitting-method counterparts. If this is not possible, the solution to the implicit equation can be approximated to arbitrary precision using Newton's method. This reduces the computational efficiency of our composition methods, since they are no longer fully explicit. We briefly discussed strategies for mitigating this issue: In particular, by choosing a split of the drift of the semi-linear SDE such that the nonlinear part(s) has only diagonal Jacobians, and by restricting ourselves to a single Newton iteration (corresponding to a Rosenbrock approximation of the solution to the nonlinear ODE), we may devise composition methods which are essentially explicit in terms of computational efficiency. We did not present theoretical results for these composition methods, however numerical demonstrations indicate that these composition methods perform at the same level as the LTIE- and SIE composition methods.

For future work, additional composition methods could be studied by considering alternative numerical methods for solving the nonlinear ODE, e.g. using the first-order Rosenbrock approximation. Future work could also consider other types of semi-linear SDEs than those studied here, e.g. semi-linear SDEs where the diffusion matrix is non-diagonal and/or non-constant in time, or indeed, semi-linear SDEs with multiplicative noise. Lastly, preservation of other structural properties than those considered here is also an interesting topic for future work.

References

[1] E. Buckwar, A. Samson, M. Tamborrino, I. Tubikanec. *Splitting methods for SDEs with locally Lipschitz drift. An illustration on the FitzHugh-Nagumo model.* arXiv:2101.01027 [math.NA].

[2] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration*. Springer, Heidelberg, 2006.

[3] P.E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, Berlin, 1992.

[4] S. Blanes, F. Casas, and A. Murua. *Splitting and composition methods in the numerical integration of differential equations*. Bol. Soc. Esp. Mat. Apl., 45, 2009.

[5] B. Øksendal. *Stochastic Differential Equations*, 6th ed. Berlin, Springer, 2013.

[6] C. Park, D. Skoug. *A Note on Paley-Wiener-Zygmund Stochastic Integrals*. Proceedings of the American Mathematical Society, 103 (2): 591–601, 1988.

[7] T.J. Lyons, M.J. Caruana, T. Lévy. *Differential equations driven by rough paths*. Volume 1908 of Lecture Notes in Mathematics. Springer, Berlin, 2007.

[8] S. Ditlevsen and A. Samson. *Hypoelliptic diffusions: filtering and inference from complete and partial observations*. J. Royal Stat. Soc. B, 81(2):361–384, 2019.

[9] J. C. Mattingly, A. M. Stuart, and D. J. Higham. *Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise.* Stoch. Process. Their Appl., 101(2):185–232, 2002.

[10] M. V. Tretyakov and Z. Zhang. *A fundamental mean-square convergence theorem for SDEs with locally Lipschitz coefficients and its applications*. SIAM J. Numer. Anal., 51(6):3135– 3162, 2013.

[11] G. N. Milstein. *A theorem on the order of convergence of mean-square approximations of solutions of systems of stochastic differential equations*. Theory Probab. Appl., 32(4):738–741, 1988.

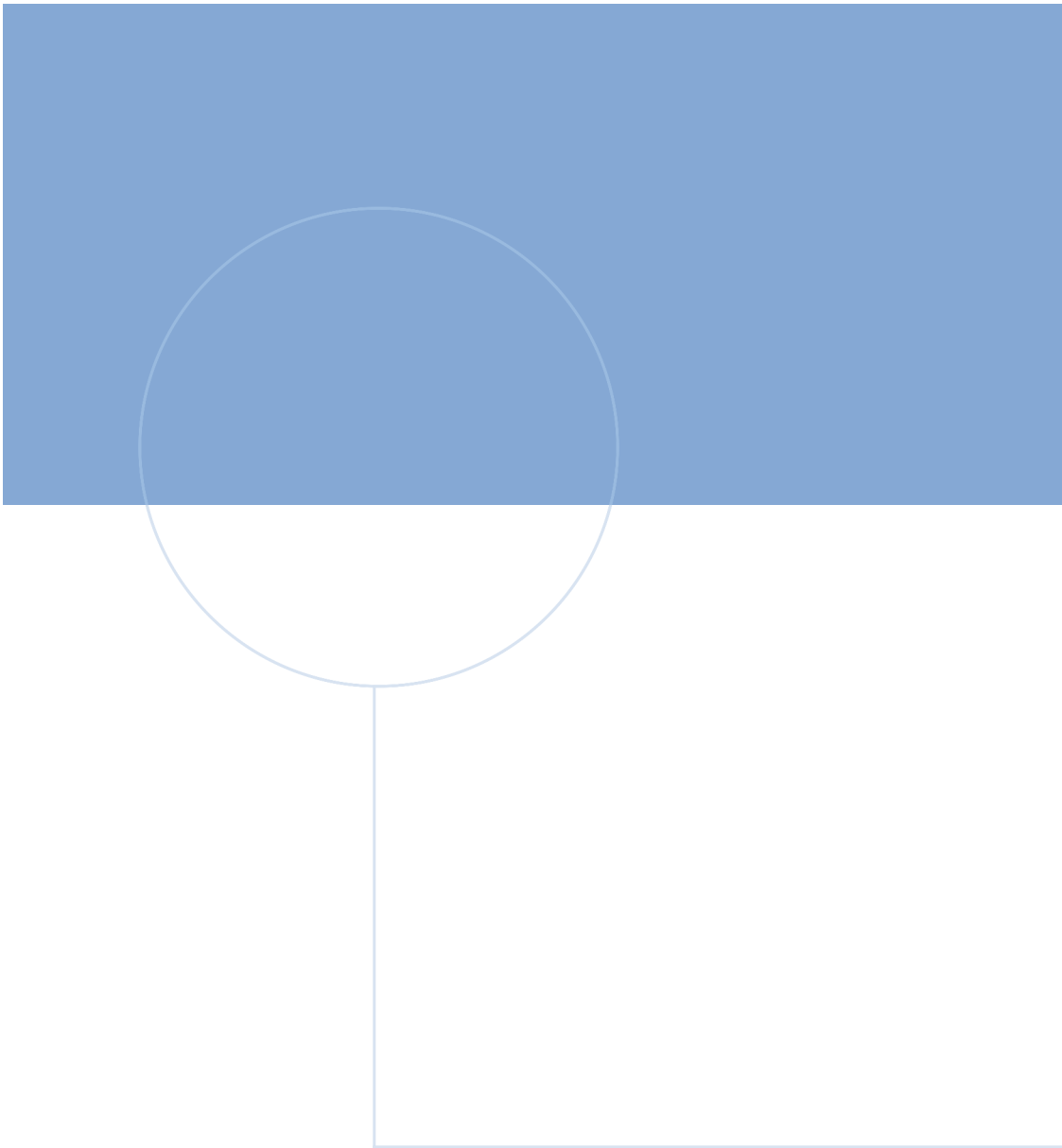
[12] H. H. Rosenbrock. *Some general implicit processes for the numerical solution of differential equations*. The Computer Journal 5(4): 329-330, 1963.

[13] G.A. Pavliotis. *Stochastic Processes and Applications. Diffusion Processes, the Fokker-Planck and Langevin Equations*. Texts in Applied Mathematics, Vol. 60. Springer, Berling, 2014.

[14] R. FitzHugh. *Impulses and physiological states in theoretical models of nerve membrane*. Biophys. J., 1(6):445–466, 1961.

[15] S. Bonaccorsi and E. Mastrogiacomo. *Analysis of the stochastic FitzHugh-Nagumo system*. Infin. Dimens. Anal. Quantum Probab. Relat. Top., 11(03):427–446, 2008.

[16] T.O. Solheim. *Composition Methods for a large class of semi-linear stochastic differential equations*. TMA4500 specialization project, Norwegian University of Science and Technology, 2022 (unpublished).



NTNU

Norwegian University of
Science and Technology