University of Salerno

Department of Mathematics

Ph.D. in Mathematics, Physics and Applications

Numerical Modeling of Stochastic Differential Problems with Applications

Giuseppe Giordano

Advisor

Prof. Beatrice Paternoster

Ph.D. Director

Prof. Patrizia Longobardi

Cycle XXXV, SSD MAT/08, A.Y. 2021/2022

Abstract

The interest in the study of stochastic differential equations has grown considerably in recent years. The main reason for this interest is that stochastic differential equations are a potential tool for modeling evolutionary problems, especially when the dynamics are affected by stochastic perturbations. Of particular importance is the study of numerics related to these problems since, in the literature, there are a few stochastic differential equations whose solution is known explicitly. Therefore, the spirit of this thesis is to analyze numerical methods for stochastic differential equations and how these can be used for the description of real-life problems. After an introduction to stochastic differential problems and some models in which these are used, we will then move on to some recalls of numerical methods, known in the literature. Subsequently, the focus will be on new research results obtained, divided into three essential parts. In the first part, based on the well-known idea of collocation for Volterra Integral Equations, we obtained continuous numerical methods, which allow us to know the solution not only at the grid points of the numerical discretization but throughout the entire integration interval. Research on this front continued by pointing out how these continuous extensions can be applied to obtain a good estimate of the local truncation error. In fact, as is also known in the deterministic context, this is a first building-block for the development of a variable step-size algorithm, useful especially in the integration of stiff problems. The second part, instead, focused on geometric numerical integration for stochastic Hamiltonian problems. Unlike deterministic Hamiltonian problems, where energy is conserved over time, stochastic Hamiltonian problems of Itô type and driven by the additive Wiener process satisfy the trace equation, i.e. the expected value of the Hamiltonian function grows linearly over time. Interest in this study derives from some results on stochastic Runge-Kutta methods developed by K. Burrage et al. in 2012. These methods, in fact, have a significant error that increases with increasing stochastic noise. Therefore, through a perturbative analysis, the reason for this behaviour was analyzed, concluding that the preservation of the main features of stochastic Hamiltonian problems does not occur directly for any discretization of time. The research was then extended to stochastic Hamiltonian problems with multiplicative noise, first obtaining a characterization of the behaviour of the mean value of the Hamiltonian and then showing that first-order approximations to such systems are unable to maintain such behaviour. In the last part of this thesis, two different models were analyzed in which stochastic differential equations (including stochastic partial differential

equations) occur. Specifically, first, the stochastic FitzHugh-Nagumo system for signal propagation in nerve cells was analyzed, in which the voltage variable is the solution of a one-dimensional partial derivative differential equation of a parabolic type with a cubic nonlinearity driven by additive space-time white noise. Splitting methods for temporal integration will then be developed for that model, showing that such schemes admit strong convergence order 1/4. Next, the analysis shifted to modeling the spread out of fake news through the stochastic SIR model, which is widely used in the epidemiological context for the spread of an epidemic. In particular, interest was placed on the stiffness property of the differential problem by pointing out that in a given population, the more stiff the problem, the faster the transit of fake news. Numerical evidence, which will demonstrate the effectiveness of the theoretical results, will be provided in the development of the thesis.

Acknowledgements

At the conclusion of my Ph.D., I would like to thank all the people who have stood by me during these three wonderful years of my life and especially who have made it all possible.

My first acknowledgment definitely goes to my advisor **Prof. Beatrice Paternoster**, for helping, supporting, and encouraging me throughout my doctoral period. It is really difficult to find the right words to explain my gratitude to her, but I can absolutely say that it is only thanks to her that I was able to have these doctoral experiences. I will never cease to thank her for all the time and willingness shown to me during these three years of doctoral studies and especially for making me realize how beautiful and important research is. I would also like to thank her for all the opportunities she has given me and for always believing in me. Being a young researcher is not an easy art and if I am one today, it is only thanks to her.

Heartfelt thanks can only go to **Prof. Raffaele D'Ambrosio** from University of L'Aquila, a special, generous, helpful person who always knew how to give me invaluable advice to improve my affinity for mathematics. I thank him for all his patience and for all the time he gave me whenever I needed it. Thank you truly, Professor. I will never forget what you did for me!

I want to sincerely thank **Prof Dajana Conte**, for all her advice and suggestions that she has been able to give me in these three years of my Ph.D. I want to thank her for always encouraging me when I felt in trouble and for constantly pushing me to move forward. I also thank **Prof. Angelamaria Cardone** for sharing with me and supporting me in these three years.

I would also like to thank the XXXV cycle Ph.D. coordinator, **Prof. Patrizia Longobardi**, for always being helpful and attentive to all the needs of the PhD program.

I would like to thank **Prof. David Cohen** from the Chalmers University of Technology of Gothenburg for hosting me for three months in Gothenburg. He is a person who is always ready to help you at any time and I am grateful for all the support he gave me during those months. In those months, I was also able to collaborate with him and with **Prof. Charles-Edouard Bréhier**, from University of Pau, on a very interesting project, which is also one of the

topics of this Ph.D. thesis. Thank you both very much for all the suggestions you gave me.

A sincere thank goes to **Prof. Kevin Burrage**, from the Queensland University of Technology, **Prof. Helmut Podhaisky**, from the Institute of Mathematics Martin Luther University Halle-Wittenberg, and **Prof. Jesús Martín Vaquero**, from the University of Salamanca, for your suggestions that improved the quality of this Ph.D. thesis.

A dutiful and heartfelt appreciation goes to my friends and colleagues of the Numerical Analysis Laboratory. Thanks to **Giovanni** with whom I shared the last two years of my PhD. He wasn't always attentive when I talked to him (I know that I am also a bit long-winded) but I always knew he would be there whenever I needed him. Thanks to **Gianluca** for all the valuable suggestions he has given me, especially in the last year of my PhD. Thanks also to **Patricia**, our first "strutturata", for her invaluable help. Thanks guys for all the help you have given me and above all for the many coffees taken at the machines in the building F.

I would like to thank **my family** for always believing in me and being there in the good times but especially in the not-so-good times. I thank them for helping and supporting me in my choices, never letting me lack their moral support.

Special gratitude goes to my girlfriend **Fabiola** for all the love she has given me and been able to pass on to me during these three years as a Ph.D. student. Despite the many problems we had to overcome, she has always stood by me with love and trust. She has always been able to support me in the moments when everything seemed so complicated, always making me make the best decisions. She is an exceptional person!

Finally, special thanks go to all the people who have been close to me for their valuable suggestions and comments over the years.

Contents

Introduction

This Ph.D. Thesis is focused on the study and analysis of numerical methods for Stochastic Differential Equations (SDEs), and Stochastic Partial Differential Equations (SPDEs) interpreted in Itô sense. This study is motivated by the role that SDEs and SPDEs play in modeling evolutionary problems. For many years, Ordinary Differential Equations (ODEs), or more generally partial differential equations, have been widely used to model phenomena deriving from natural science and technology. Although this class of equations constitutes a powerful tool for modeling evolution phenomena, the introduction of stochasticity in such models guarantees a more realistic representation. As a matter of fact, when the system's dynamics are affected by random perturbations, SDEs (or SPDEs) are preferable. In some sense, SDEs can be regarded as differential equations that contain random elements [235]: when the random perturbations are absent, the SDEs become ODEs.

The origin of the study and of the analysis of stochastic differential equations appeared in the literature for the first time in 1930 with the well-known *Ornstein-Uhlenbeck* model of *Brownian motion*, which the solution is known as *Ornstein-Uhlenbeck process*. It is named after Leonard Ornstein and George Eugene Uhlenbeck. The first application dates back to the physical sciences with a model for the velocity of a massive Brownian particle under the influence of friction.

Although the first equation appeared in 1930, the first and true rigorous mathematical theory was born in the mid-20th century by the Japanese mathematician Kiyosi Itô , when in 1942 he published the paper *On stochastic processes (Infinitely divisible laws of probability)* on the *Japanese Journal of Mathematics* [164].

Just to mention a few examples, using stochastic differential equations, one can model several phenomena like the dispersion of a pollutant in the air or in the water, the effect of noise on the transmission of telecommunication signals, the dynamics of chemical reactions or the dynamics of several populations of living beings when random environmental perturbations affect their growth rate.

In the last year, the number of researchers involved in the numerics for SDEs has grown considerably. The interest is given by the fact that in the existing literature, there are only a few equations in which the solution is analytically known. Therefore the development of ad hoc numerical methods, able to solve approximately the problem under consideration, becomes essential.

A natural approach to construct numerical methods is to perturb classical

deterministic methods for ordinary differential equations, such as the Euler-Maruyama method (named after Leonhard Euler and Gisiro Maruyama) or the more general θ- methods. Subsequently, the research shifted to derive numerical methods with a higher order of convergence under numerical discretizations, such as numerical methods based on the well-known Itô-Taylor formula (considered as the stochastic counterpart of Taylor's formula) or stochastic Runge-Kutta method following the same deterministic settings.

Taking into consideration the existing literature on numerical methods for stochastic differential equations, the spirit of this Ph.D. thesis is certainly to illustrate how SDEs are involved in the modeling of evolution problems, thus also considering real-life problems and how it is possible to derive numerical methods for approximating the solution of the problem under consideration. This Ph.D. thesis can be divided into three parts, as seen in the sequel.

The first part will entirely focus on studying continuous numerical methods for stochastic differential equations. These methods will be built following the well-known collocation technique, based on the idea of approximating the exact solution with a suitable approximant belonging to a chosen finite dimensional space, usually a piecewise algebraic polynomial, which exactly satisfies the equation on a specific subset of the integration interval. The collocation methods are very advantageous due to they provide us with an approximation of the solution of the equation, not only in the grid points of the numerical discretization but in the whole integration interval. Furthermore, it is also important that the continuous approximant can be chosen as a linear combination of functions ad hoc for the problem considered with the aim to better reproduce the qualitative behavior of the solution. Another advantage of collocation methods is in the fact that they are a powerful tool for developing a variable step size algorithm used mainly to integrate problems with a high stiffness. It should also be remembered that collocation also has an important theoretical significance: indeed, many numerical methods are difficult to analyze as discrete schemes, whereas their analysis, when reformulated as collocation-based methods, is quite simplified and can be done in a very elegant way. Therefore, following the same approach for the construction of collocation methods for Volterra Integral Equations (VIEs), we will construct continuous extensions of selected numerical methods for stochastic differential equations. In particular, for selected data settings, a continuous extension of the Euler-Maruyama method will be obtained. As happens in the deterministic context, also in the stochastic case, continuous approximants make the analysis of the corresponding numerical methods simpler and easier to use. So a question we asked ourselves is: can continuous extension provide an efficient local truncation error estimation? However, our answer is positive but at a preliminary stage. Indeed, through continuous extensions reported by Higham et. al. in [152], it was found that the estimate obtained is accurate when applied to different test problems with three different numerical methods: Euler-Maruyama, Split-Step Backward Euler and Backward Euler methods.

The second part of this Ph.D. thesis is focused on studying geometric numerical integration, widely used in the deterministic context for ordinary differ-

ential equations (see, for instance, [144]). The spirit is to investigate whether numerical discretizations are able to preserve specific invariant laws when applied to stochastic differential systems that possess a particular characterizing structure or a property of interest. In a nutshell, we want to reproduce the geometric structures of a given continuous problem along with its discretized counterpart. In particular, the purpose of this Ph.D. thesis concerns the analysis of energy-preserving methods applied to stochastic Hamiltonian problems in their formulation according to Itô. Stochastic Hamiltonian problems are the most suitable candidates to conciliate the Hamiltonian nature of classical mechanics, which is closely related to the canonical character of the evolution equations, with the non-differentiable nature of the Wiener process, which describes the continuous innovative character of stochastic effects [21, 22].

Stochastic Hamiltonian problems are used in many applications, such as molecular dynamics, celestial mechanics, population dynamics, weather forecast, immunology, control theory, fluid dynamics, theory of adiabatic invariants, continuum mechanics, nonlinear optics, elastodynamics, oceanography, electromagnetism, cosmology and quantum field theory, just to mention a few examples.

In this Ph.D. thesis, particular interest will be initially placed on Itô-type stochastic Hamiltonian problems with additive noise, which leads us to a system of stochastic differential equations with additive noise. Different from what happens for deterministic Hamiltonian problems, where the Hamiltonian function is a first integral of the system, in the stochastic case, Burrage et. al. in [68, 74] proved that, in the case of additive noise, the Hamiltonian function satisfies the *trace equation*, revealing that the expected value of the Hamiltonian grows linearly over time. The authors also have proved that, if we consider stochastic Runge-Kutta methods involving multiple Wiener processes, they preserve the trace law along its numerical dynamics, at least for quadratic Hamiltonians. However, stochastic Runge-Kutta methods exhibit a remarkable error growth as the parameter ϵ of the diffusive part increases. Through a perturbative analysis, in terms of ϵ expansions, we investigate the reason of this behavior, due to the presence of a secular term $\epsilon\sqrt{t}$ destroying the overall conservation accuracy. In [84], the authors have derived a drift-preserving numerical integrator for stochastic Hamiltonian problems with additive noise, in order to reproduce the trace law along the numerical discretization.
The analysis will then be extended to the case of stochastic Hamiltonian systems with (small) multiplicative noise. First, a characterization of the expected value of the Hamiltonian function will be obtained, emphasizing the separable and quadratic Hamiltonian. Next, we will show that, in general, first-order approximations to such systems cannot retain the same behavior discovered for the exact averaged Hamiltonian. At the conclusion of this second part, the analysis for the specific case of Euler-Maruyama methods is performed.

The last part of this thesis aims to study and analyze evolution models, one involving stochastic partial differential equations and the second stochastic differential equations. In detail:

- **FitzHugh-Nagumo** (FhN) system (named after Richard FitzHugh)

is a model widely used in neuroscience. In the deterministic context, FitzHugh–Nagumo system is a simplified two-dimensional version of the famous Hodgkin–Huxley model, which describes how action potentials propagate along an axon. Historically this was the first model created to describe this process, for which its discoverers, physiologists Alan Lloyd Hodgkin and Andrew Huxley, won the Nobel Prize in Physiology in 1963. Noise is omnipresent in neural systems and arises from different sources: it could be internal noise (such as random synaptic input from other neurons) or external noise. It was noted in [237] that adding an appropriate amount of noise in the model helps to detect weak signals.
This Ph.D. thesis proposes a splitting strategy for the stochastic FitzHugh-Nagumo system. The main idea of a splitting strategy is to decompose the vector field, appearing in the evolution equation, into several parts, in order to exhibit subsystems that can be integrated exactly (or easily). The contributions of non linearity, diffusion operators, and noise can be integrated and combined separately to obtain an explicit, easy-to-implement, and effective numerical scheme. The most crucial result concerning this study is the proof of mean square convergence at a rate of at least $1/4$ of the proposed numerical scheme. We also prove that the scheme satisfies appropriate moment bounds, which is remarkable since we consider schemes that explicitly deal with a non linearity with cubic growth.
The analysis combines some features of the analysis of splitting schemes of the stochastic Allen-Cahn equation [45] and of finite-dimensional versions of the FhN stochastic system; however, we need to develop non-trivial and original arguments to treat the stochastic FitzHugh-Nagumo system.

- The problem of **fake news** represents a real threat to the stability and cohesion of societies, especially in a context like the current one, where news is now accessible from everyday devices such as smartphones, PCs, or smart speakers. In the face of this growing number of information sources, a greater critical sense is needed among users because not all news is founded, and some can even be very dangerous. In fact, although the phenomenon of hoaxes existed even before the Internet, with the web, the spread of misleading news increases, due to the modern channels of transmission, such as Social Media, which can easily reach millions of users worldwide. As a matter of fact, in today's society, fake news is spreading significantly more widely, quickly, deeply and with a long-range, compared to real news, in a dynamic that affects all sectors of information, often exponentially.
  Through the use of epidemiological models, widely used to analyze the spread of a disease in a population, the aim of the last part of this Ph.D. thesis is to be able to model the spread of fake news in such a way as to limit its serious consequences that could occur in the future.
  In the first analysis, the classic SIR model was considered, where the population is divided into three subgroups: susceptible, infected and re-

covered. Each of these categories will be given a meaning with regard to the phenomenon of fake news. In particular, the susceptible will be potentially authoring the spreading of fake news; the infected will be the wide variety of authors highly active in posting fake information; the recovered will be authors who are inactive in the spreading of fake news. This Ph.D. thesis highlights how the stiffness index, which is often used as a measure of stiffness for differential problems, can be employed to model the spread of fake news. In particular, we show that the more the stiffness index of the SIR model is high, the more the transit of fake news in a given population is rapid.
It is clear that the phenomenon of fake news can be affected by uncertainty. Indeed, the dissemination of information is a stochastic process that may depend, for example, on the content of the information, the influence of users and the structure of the network. In this context, it is highlighted how the phenomenon of fake news is best described by a model of stochastic differential equations. In this thesis, we use a stochastic SIR model obtained by perturbing the classic SIR model. The analysis linked to the stiffness index turns out to be more complex this time, as in the stochastic community, a strict definition of a stiffness index is still unclear. Therefore, here this analysis has been carried out by pointing out that the larger the diffusive parameter of the stochastic counterpart is, the more the order reduction phenomenon occurs and therefore the more the problem becomes stiff [30, 31]. Furthermore, it has been experimentally proved that the more the stiff problem, the more the contagion curve tends to lower, a consequence of the fact that stochastic perturbations within the model create noise, which in turn could generate chaos in the population.

This Ph.D. thesis is organized as follows: in the first chapter, the fundamental notions regarding stochastic differential equations will be recalled, emphasizing those that are the main results of the existence and uniqueness of the SDE and the famous formula of Itô. In Chapter 2, instead, the standard aspects of some numerical methods for stochastic differential equations will be reported, which will be useful in the sequel of this thesis. The third chapter aims to obtain continuous methods for stochastic differential equations. First, the collocation technique for the Volterra integral equations will be recalled. Subsequently, however, by exploiting this technique, continuous extensions of some numerical methods for SDE will be obtained. The original results shown in this chapter are an object of [93]. The study then continues with the analysis of stochastic Hamiltonian problems, which will be studied in the fourth chapter of this Ph.D. thesis. In particular, after introducing the importance of Hamiltonian problems in real-life modeling phenomena, we will move on to the study of stochastic Hamiltonian problems with both additive and multiplicative noise. For each of these classes, the perturbative analysis will be carried out through the ϵ expansions, concluding that it is not possible to preserve the trace laws under any numerical discretizations. The novelty enclosed in this chapter is an object of [121]. The Chapter 5 contains new research that

I have conducted in collaboration with Prof. David Cohen (from Chalmers University of Gothenburg and University of Gothenburg) and Prof. Charles-Edouard Brehier (from the University of Pau). It is based on the study of a used model for signal propagation in nerve cells. A splitting method will be analyzed, and the related convergence analyzes will be discussed. The actual result is reported in [44]. The last chapter instead deals with a relevant problem in modern society concerning the spread of fake news. In particular, we will first see how the deterministic stiffness index can be used to model the spread of fake news, and then we will study the stochastic counterpart. Part of this research is reported in [119]. All the results obtained will be confirmed by numerical simulations that will demonstrate the theoretical efficiency. Finally, conclusions and possible future work will be discussed.

Chapter 1

Stochastic Differential Equations

In this first chapter a brief introduction to stochastic differential equations (SDEs) will be provided, highlighting the fundamental results useful for the sequel of this thesis. Some examples of applications in which SDEs are widely used will be introduced.

1.1 Stochastic differential equations

Let $(\Omega, \mathcal{F}, \mathbb{P})$ a be complete probability space with filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ satisfying the usual conditions and let $W(t) = (W_1(t), \dots, W_m(t))$ be an m–dimensional Wiener process (or also called Brownian motion) defined on the space, having independent Wiener process components. If $f : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $G : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ are two functions Borel measurable, the d–dimensional stochastic differential equation of Itô type is given by

$$dX(t) = f(t, X(t))dt + G(t, X(t))dW(t), \quad t \in [t_0, T], \tag{1.1}$$

with initial value $X(t_0) = X_0$ an F_{t_0}–measurable $\mathbb{R}^d$–valued random variable, such that $\mathbb{E}|X_0|^2$ is finite. Usually, in the literature (see [151], for instance), f is called the *drift coefficient* and G the *diffusion coefficient* of the equation (1.1). Different from the deterministic case, in the stochastic setting one is not allowed to write $dW(t)/dt$, since the Brownian motion, usually, is nowhere differentiable in time with probability 1. Therefore, equation (1.1) assumes only a theoretical interest and it is difficult to be managed.
To better understand the considered problem, by integrating equation (1.1) in $[t_0, t]$, the corresponding integral formulation of (1.1) is introduced, and it is given by

$$X(t) = X_0 + \int_{t_0}^{t} f(s, X(s))\, ds + \int_{t_0}^{t} G(s, X(s))\, dW(s), \quad t \in [t_0, T]. \tag{1.2}$$

The first integral appearing in (1.2) is a standard Riemann integral along paths while the second integral of the right–side is intended in the Itô sense, that is, respect to uniform grid $t_0 < t_1 < \ldots < t_N = T$ it reads

$$\int_{t_0}^{t} G(s, X(s))\, dW(s) = \lim_{N\to+\infty} \sum_{j=0}^{N-1} G(t_j, X(t_j))\Delta W_j, \tag{1.3}$$

where $\Delta W_j := W(t_{j+1}) - W(t_j)$ are the so-called Wiener increments, namely normally distributed random variable with mean 0 and variance $\Delta t := t_{j+1} - t_j$.

Remark 1.1.1. *The stochastic differential equations (1.1) is interpreted in Itô sense, since the stochastic integral, appearing in (1.2), is in the Itô formulation. As a matter of fact, there are infinite ways to define the stochastic integral in (1.2). As reported in the Gard's monograph [131], if $H(t, X(t)) = [h_1(t, X), \ldots, h_m(t, X)]^\intercal$ is an $n \times m$ suitable matrix–valued function with components h_j n–vector valued functions, by taking $t_0 < t_1 < \ldots < t_N = T$ an uniform mesh of the interval $[t_0, T]$, the family of stochastic integrals is defined by*

$$\int_{t_0}^{T} H(s, X(s))\, dW(s) = \lim_{N \to +\infty} \sum_{j=0}^{N-1} H\left(t_j, \lambda X(t_{j+1}) + (1-\lambda) X(t_j)\right) \Delta W_j, \tag{1.4}$$

for some real-value $\lambda \in [0, 1]$. In terms of columns of H, equation (1.4) can be written as

$$\int_{t_0}^{T} \sum_{i=1}^{m} h_i(s, X(s))\, dW_i(s) = \lim_{N \to +\infty} \sum_{j=1}^{N-1} \left(\sum_{i=1}^{m} h_i(t, \lambda X(t_{j+1}) + (1-\lambda) X(t_j)) \right) \Delta W_{ij},$$

where $\Delta W_{ij} := W_i(t_{j+1} - W_i(t_j))$.
The reader may note that in the the case $\lambda = 0$, the well-defined Itô integral is obtained, while, the case $\lambda = 1/2$ (the Stratonovich integral) deserves to be treated separately for its physical meaning [37, 135, 190, 248] and for the numerous and several applications in different areas, see for instance [13, 154, 172, 193, 212] and references therein.
The Stratonovich integral is usually denoted by

$$\int_{t_0}^{T} H(s, X(s)) \circ dW(s),$$

and correspondingly, the SDEs in the Stratonovich sense are written as

$$dX(t) = f(t, X(t))dt + G(t, X(t)) \circ dW(t), \quad t \in [0, T]. \tag{1.5}$$

Effortlessly, one can show that the Stratonovich integral coincides with ordinary calculus in the deterministic case. Furthermore, one can prove that the Itô SDEs (1.2) are equivalent to the Stratonovich SDEs

$$dX(t) = f(t, X(t))dt - \frac{1}{2} G'(t, X(t)) G(t, X(t))dt + G(t, X(t)) \circ dW(t),$$

in the sense that if $X(t)$ solves one that, it also solves the other.

1.2 Solution of SDEs

In this section, the concept of solution of SDEs will be presented and some theorems of existence and uniqueness of the solution will be stated. In particular, these results will allow us to state that equation (1.2) is definitely well-defined.

Denote by $\mathcal{L}^p([t_0,T];\mathbb{R}^d)$ the family of $\mathbb{R}^d$-valued $\mathcal{F}_t$-adapted process $\{f(t)\}_{t_0\leqslant t\leqslant T}$ such that $\int_{t_0}^{T}|f(t)|^p\,dt<+\infty$ almost surely (a.s).

Definition 1.2.1. *A $\mathbb{R}^d$–valued stochastic process $\{X(t)\}$, with $t_0\leqslant t\leqslant T$, is a solution of (1.1) if and only if it satisfies the following properties:*

1. *$\{X(t)\}$ is continuous and $\mathcal{F}_t$–adapted;*
2. *$\{f(t,X(t))\}\in\mathcal{L}([t_0,T];\mathbb{R}^d)$ and $\{G(t,X(t))\}\in\mathcal{L}^2([t_0,T];\mathbb{R}^{d\times m})$, for $\in[t_0,T]$;*
3. *equation (1.2) holds for every $t\in[t_0,T]$ with probability 1.*

According to the Definition 1.2.1, the solution of a stochastic differential equation, is a stochastic process, therefore for any $t\in[t_0,T]$, $X(t)$ is a random variable. Furthermore, in the first place one could immediately observe that in order for equation (1.2) to make sense, the functions f and G in (1.2) must satisfy appropriate conditions. The next theorem establishes the existence and uniqueness of the solution to SDEs (1.1), under a very strong condition.

Theorem 1.2.1. *Assume that:*

1. *the functions $f(t,x(t))$ and $G(t,x(t))$ are Borel measurable with respect to $t\in[t_0,T]$ and $x\in\mathbb{R}^d$;*
2. *there exists a constant K such that for all $t\in[t_0,T]$, and $x,y\in\mathbb{R}^d$*
   (a) *(Lipschitz condition)*
   $$|f(t,x)-f(t,y)|+|G(t,x)-G(t,y)|\leqslant K|x-y|;$$
   (b) *(Linear growth condition)*
   $$|f(t,x)|^2+|G(t,x)|^2\leqslant K^2(1+|x|^2);$$
3. *X_0 is independent of $W(t)$, for $t>t_0$, and $\mathbb{E}|X_0|^2<\infty$.*

Then there exists a solution $X(t)$ to (1.1) defined on $[t_0,T]$ which is continuous with probability 1 and such that

$$\sup_{t\in[t_0,T]}\mathbb{E}|X(t)|^2<\infty.$$

Moreover, the solution $X(t)$ is pathwise unique, in the sense that if $X(t)$ and $Y(t)$ are two solutions to SDE (1.1), then

$$\mathbb{P}\left(\sup_{t\in[t_0,T]}|X(t)-Y(t)|=0\right)=1.$$

The proof of the above theorem is technical and requires knowledge of some probabilistic aspects, therefore it is omitted. For details, the interested reader could read, for instance [10, 165].
According to Theorem 1.2.1, the existence and uniqueness of the solution to (1.1) require the functions f and G satisfy the Lipschitz condition and linear growth in the state variable. The condition (2a) is responsable for uniqueness of the solution, while the condition (2b) ensures the existence of the solution. The Lipschitz condition (2a) of Theorem 1.2.1 means that the drift and diffusion coefficients of (1.1) do not change faster than a linear function of x. Clearly, these conditions imply the continuity of $f(t, X(t))$ and $G(t, X(t))$ in $X(t)$ for all $t \in [t_0, T]$. However, these conditions impose much more stringent constraints and, therefore, may be relaxed; in particular, the Lipschitz (uniform) condition (2a) may be substituted with the local Lipschitz condition, as follow:

(2a') For each $N > 0$, there is a constant K_N such that for all $t \in [0, T]$, and for X and Y, with $|X| \leqslant N$, $|Y| \leqslant N$,

$$|f(t, X) - f(t, Y)| + |G(t, X) - G(t, Y)| \leqslant K_N |X - Y|.$$

Furthermore, the linear growth condition is too restricted. Therefore Mao, in his monograph [191], shows and proves the following remarkable theorem:

Theorem 1.2.2. *Assume that the local Lipschitz condition (2a') holds, but the linear growth condition (2b) is replaced with the following monotone condition: there exists a positive constant K such that for all $(t, X(t)) \in [t_0, T] \times \mathbb{R}^d$*

$$X^\intercal f(t, X(t)) + \frac{1}{2}|G(t, X(t))|^2 \leqslant K(1 + |X|^2).$$

Then there exists a unique solution $X(t)$ to equation (1.1).

According with [110], the following definitions can be given.

Definition 1.2.2. *Il the diffusion term G of equation (1.1) does not depend on the variable X, that is $G(t, X) \equiv G(t)$ for all $t \in [t_0, T]$, then the SDE (1.1) is said to have an additive noise. Otherwise it is said to have a multiplicative noise.*

Definition 1.2.3. *If the coefficients f and G of (1.1) do not depend explicitly on time, then the SDE (1.1) is called autonomous.*

Keeping in mind the definitions just considered, an SDE with additive noise, i.e.

$$X(t) = X(t_0) + \int_{t_0}^{t} f(s, X(s))\, ds + \sigma W(t) \tag{1.6}$$

with $\sigma \in \mathbb{R}$, can be seen as the simplest way to perturb an ordinary differential equation (ODE). In fact, let consider an ODE of the form

$$\frac{dx}{dt} = f(t, x(t)), \quad t \in [t_0, T], \tag{1.7}$$

with initial condition $x(t_0) = x_0$. The solution of (1.7) can be written as

$$x(t) = x(t_0) + \int_{t_0}^{t} f(s, x(s))ds \tag{1.8}$$

Therefore, equation (1.6) appears as a stochastic perturbation of the solution (1.8) to ODE (1.7).

1.3 The Itô formula

To complete the basic theory of stochastic differential equations it is important to consider an essential key tool in the stochastic calculus: Itô's Lemma (or also called Itô 's formula). In particular, it can be regarded as a stochastic version of the deterministic chain-rule, widely used to differentiate functions of the solutions to deterministic differential equations. As a matter of fact, Itô's formula gives the stochastic differential of a smooth function of the solution of a stochastic differential equation.
As the reader will recall, if $x(t)$ is the solution of the following ordinary differential equation

$$\frac{d}{dt}x(t) = f(t, x(t)), \quad t \geqslant t_0$$

with a suitable function f, then the chain rule applied to a smooth suitable function $v(x(t))$ gives

$$\frac{d}{dt}v(x(t)) = \frac{d}{dx}v(x(t))\frac{d}{dt}x(t) = v'(x(t))f(x(t)).$$

To better understand the Itô's Lemma, the scalar version will be introduced first, that is applied to a scalar stochastic differential equation, and subsequently it will be possible to state the same Lemma for a system of SDEs.

Theorem 1.3.1 (The one-dimensional Itô formula)**.** *Consider a one-dimensional SDE*

$$\begin{cases} dX(t) = f(t, X(t))\, dt + g(t, X(t))\, dW(t), \quad t \in [t_0, T], \\ X(t_0) = X_0. \end{cases} \tag{1.9}$$

with $X(t) \in \mathbb{R}$ for all $t \in [t_0, T]$. Assume that the functions f and g of (1.9) satisfy the conditions of Theorem 1.2.1 guaranteeing the existence and uniqueness of solution. If the real–valued function $F(t, x)$ has continuous partial derivatives $\partial F/\partial t$, $\partial F/\partial x$, and $\partial^2 F/\partial x^2$ for $t \in [t_0, T]$ and $x \in \mathbb{R}$, and $X(t)$ is a solution of (1.9), then the process $F(t, X(t))$ has the following stochastic differential

$$dF(t, X(t)) = \left[\frac{\partial F}{\partial t} + f\frac{\partial F}{\partial x} + \frac{1}{2}g^2\frac{\partial^2 F}{\partial x^2}\right](t, X(t))dt + \left[\frac{\partial F}{\partial x}g\right](t, X(t))dW(t). \tag{1.10}$$

The integral representation of (1.10) reads

$$F(t,X(t)) = F(t_0,X_0) + \int_{t_0}^{t} \left[\frac{\partial F}{\partial t} + f\frac{\partial F}{\partial x} + \frac{1}{2}g^2\frac{\partial^2 F}{\partial x^2}\right](s,X(s))ds + \int_{t_0}^{t} \left[\frac{\partial F}{\partial x}g\right](s,X(s))dW(s).$$

Itô's Lemma is a powerful result in the stochastic calculus, since the solutions of many stochastic differential equations are computed, the bounds for their moments are provided and numerical methods are obtained for (1.9). A simplifying example will be shown to understand the usefulness of the Itô formula to find the solution of a given SDE.

Example 1.3.1. *Consider the following scalar linear stochastic differential equation*

$$dX(t) = f(t)X(t)dt + g(t)X(t)dW(t), \tag{1.11}$$

where f and g are continuous functions. Letting $F(x) = x^2$, Itô's Lemma can be applied and it asserts that

$$dX^2(t) = \left[2f(t)X^2 + g^2(t)X^2(t)\right] + 2g(t)X^2(t)dW(t).$$

The stochastic process

$$Y(t) = X^2(t)$$

solves the SDE

$$dY(t) = \left[2f(t) + g^2(t)\right]Y(t) + 2g(t)Y(t)dW(t). \tag{1.12}$$

Similarly to the scalar case, one shall extend the one–dimensional Itô's formula to the multidimensional case. For a function $F : [t_0,T]\times\mathbb{R}^d \to \mathbb{R}$ with continuous partial derivatives $\partial F/\partial t$, $\partial F/\partial x_i$ $\partial^2 F/\partial x_i\partial x_j$, putting $X = \{X_i\}$, $W = \{W_j\}$, $f = \{f_i\}$, and $G = \{g_{ij}\}$ for $i = 1,\dots,d$ and $j = 1,\dots,m$, one has, in component form

$$\begin{aligned} dF(t,X(t)) = &\left[\frac{\partial F}{\partial t} + \sum_{i=1}^{d} f_i\frac{\partial F}{\partial x_i} + \sum_{i,j=1}^{d}\sum_{k=1}^{m}\frac{1}{2}g_{ik}g_{jk}\frac{\partial^2 F}{\partial x_i\partial x_j}\right](t,X(t))dt \\ &+ \sum_{i=1}^{d}\frac{\partial F}{\partial x_i}(t,X(t))\sum_{j=1}^{m} g_{ij}(t,X(t))dW_j(t), \end{aligned} \tag{1.13}$$

whose integral form is given by

$$\begin{aligned} F(t,X(t)) = F(t_0,X_0) &+ \int_{t_0}^{t}\left[\frac{\partial F}{\partial t} + \sum_{i=1}^{d} f_i\frac{\partial F}{\partial x_i} + \sum_{i,j=1}^{d}\sum_{k=1}^{m}\frac{1}{2}g_{ik}g_{jk}\frac{\partial^2 F}{\partial x_i\partial x_j}\right](s,X(s))ds \\ &+ \sum_{i=1}^{d}\int_{t_0}^{t}\frac{\partial F}{\partial x_i}(s,X(s))\sum_{j=1}^{m} g_{ij}(s,X(s))dW_j(s), \end{aligned}$$

Equivalently, in vector–matrix form, equation (1.13) can be rearranged in a more compact form as

$$dF = \left[F_t + f^\intercal F_x + \frac{1}{2}\text{trace}(GG^\intercal F_{xx})\right](t, X(t))dt + \left[F_x^\intercal G\right](t, X(t))dW(t), \tag{1.14}$$

where 'trace' denotes the trace operator on the sum of main diagonal entries, and F_x, F_{xx} are respectively the vector gradient of the function F, that is $(F_x)_j = \dfrac{\partial}{\partial x_j}F$, for $j = 1, \dots, d$ and the Hessian matrix of the function F, i.e., $(F_{xx})_{j,i} = \dfrac{\partial}{\partial x_j \partial x_i}F$, for any $j = 1, \dots, d$ and $i = 1, \dots, m$.

Another important result concerning the stochastic calculus is the so-called *Itô isometry*, which for a sufficiently smooth function $h : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}$ is given by

$$\mathbb{E}\left[\left|\int_{t_0}^{t} h(s, X(s))dW(s)\right|^2\right] = \int_{t_0}^{t} \mathbb{E}\left[h(s, X(s))^2\right] ds$$

This above results derive substantially from an application of the Itô's Lemma and on the principle that $dW(t)^2 \equiv dt$ (see [151]).

1.4 Some applications

This subsection will describe some applications in which stochastic differential equations are widely used. Given the random nature of SDEs solution, these can be used in multiple real-life phenomena, especially when the latter are subject to random perturbations. Hence some problems involving SDEs will be listed and described below, starting from dynamic population models to financial models.

1.4.1 Population dynamics

Population dynamics, as reported in [24] is the ecology field that aims to study the variation in time and space of population size and density for one or more species in a crowded environment. In the deterministic case, a classical model to describe population dynamics is given by the logistic linear–quadratic Verhulst equation [19]:

$$\frac{dx}{dt} = r(K - x)x, \tag{1.15}$$

where the constant $K > 0$ is the *carrying capacity* of the environment and the constant $r > 0$ is such that $1/r$ is a characteristic *timescale*. The solution $x(t)$ represents the population density of time t
In a very natural way, by randomizing the constant K and r, equation (1.15) leads to the following scalar autonomous SDE in Itô sense [151]

$$dX(t) = rX(t)(K - X(t))dt + \beta X(t)dW(t), \quad t \in [0, T]. \tag{1.16}$$

Immediately, one can observe that if the stochastic noise β vanishes, the standard logistic ODE (1.15) is obtained. The constant β controls the strength of the stochasticity term. For this model (1.16), Kloeden and Platen in [165] have shown that the exact solution of the SDE (1.16) is

$$X(t) = \frac{X(0)e^{(rK-\frac{1}{2}\beta^2)t+\beta W(t)}}{1+X(0)r\int_0^t e^{(rK-\frac{1}{2}\beta^2)s+\beta W(s)}ds}.$$

Therefore, to maintain the positivity of the solution to the model, it is necessary that the initial value $X(0) \geqslant 0$ with probability 1. In this hypothesis then $X(t) \geqslant 0$ for every $t \in [0,T]$ with probability 1.
Another very studied deterministic model of multi–species interaction is the Lotka-Volterra system [239](or also called Predatory–prey model), that for a d–dimension population it is given by

$$\frac{dx_i}{dt} = x_i\left(b_i + \sum_{j=1}^{d} a_{ij}x_j\right). \tag{1.17}$$

Here x_i denotes the density of i–th species, the coefficient b_i is the intrinsic growth rate for i–species and a_{ij} represents the effect of the j–th species upon the i–th one. The matrix $A = (a_{ij})$ is the so-called *interaction matrix*. As reported in Kloeden and Platen monograph [165], randomizing the growth coefficients, equation (1.17) becomes a system of SDEs as follows

$$dX_i(t) = X_i(t)\left(b_i + \sum_{j=1}^{d} a_{ij}X_j(t)\right)dt + \beta_i X_i(t)dW_i(t), \quad t \in [0,T]. \tag{1.18}$$

Different from the Verhulst model, here the exact solutions are not known; therefore they must be solved numerically, with some numerical methods (see Chapter 2).

1.4.2 Stochastic SIS Model

In the epidemiological field, it is widespread to use deterministic-compartmental models when studying the spread of a new disease. This subsection will pay particular attention to the study of the SIS epidemic model (susceptible-infected-susceptible) and how it is possible to define it in a stochastic context, introducing random perturbations. If $S(t)$ denotes the number of susceptible people and $I(t)$ the number of infected at time t, the differential model is defined as

$$\begin{cases} \dfrac{dS(t)}{dt} = \mu N - \beta S(t)I(t) + \gamma I(t) - \mu S(t), \\ \dfrac{dI(t)}{dt} = \beta S(t)I(t) - (\gamma+\mu)I(t), \end{cases} \tag{1.19}$$

with positive initial condition S_0, I_0 such that $S_0 + I_0 = N$, where N is the total size of the population. Here, the parameter μ represents the per capita death

rate, γ the rate at which infected people became cured, and the coefficient β the disease transmission rate.
If one supposes that random environmental factor takes action simultaneously on each individual in the population [136], then the stochastic SIS model in the Itô sense assumes the form

$$\begin{cases} dS(t) = (\mu N - \beta S(t)I(t) + \gamma I(t) - \mu S(t))\,dt - \sigma S(t)I(t)dW(t), \\ \\ dI(t) = (\beta S(t)I(t) - (\gamma + \mu)I(t))\,dt + \sigma S(t)I(t)dW(t), \end{cases} \tag{1.20}$$

where the parameter μ, γ, β assume the same meaning as the deterministic SIS model, while the parameter σ is the amplitude of the stochastic counterpart. Keeping in mind that $S(t) = N - I(t)$, then the model (1.20) can be rearranged as a single scalar SDE, as follows

$$dI(t) = I(t)(\beta N - \mu - \gamma - \beta I(t))dt + \sigma(N - I(t))dW(t), \tag{1.21}$$

with initial condition $I(0) = I_0$. As shown in [136], stochastic SIS model (1.21) admits a unique and positive solution $I(t) \in (0, N)$, in the sense that

$$\mathbb{P}\{I(t) \in (0, N), \forall t \geqslant 0\} = 1.$$

There are multiple models that describe the spread of an epidemic. For instance, in Chapter 6 the stochastic SIR model (susceptible-infected-recovered) will be introduced, adapting it to the context of the spread of fake news.

1.4.3 Twitter Activity

In recent years, the rapid increase in the use of social networks has changed the way people interact with each other. In fact, the role of social networks is precisely to create a place for discussion within which users want to interact and discuss. In particular, Twitter is a micro-blogging platform where each registered user can send and share "tweets" (considered as small information) that can be private or made public to the user's followers. In [196] a model that analyzes the correlated human activity in massive social organizations is described through SDE. Specifically, the stochastic process for the average waiting time between successive tweets is determined by the following scalar SDE:

$$dX(t) = (1 + f(X(t)))\,dt + \sqrt{X(t)}dW(t),$$

where the deterministic function $f(X(t))$ is chosen so that the process has a non-trivial stationary distribution.

1.4.4 Seismography

The study of earthquakes has been extensively studied over the years using deterministic differential models. Of course, from the intuitive point of view, these phenomena are preferably described by stochastic differential equations, which better acquire the behaviour of the phenomenon studied. As reported

in [165], the first classic model for seismography was introduced by Bolotin, assuming that the acceleration $z''(t)$ of the ground level $z(t)$ at time t:

$$z''(t) = I(t)\xi(t), \tag{1.22}$$

with $\xi(t)$ Gaussian white noise (stationary and ergodic random process with zero mean [192]), and $I(t) = hte^{\alpha t}$, where α, h are positive known constants. Usually, the second order stochastic differential equation may be rewritten as a system of SDEs

$$\begin{cases} dZ_1(t) = Z_2(t)dt \\ \\ dZ_2(t) = I(t)dW(t) \end{cases} \tag{1.23}$$

where $Z_1(t) = z(t)$ is the vertical displacement and $Z_2(t) = z'(t)$ the velocity in the vertical direction.

Kozin has proposed a different variant of (1.22) [165]. In particular, he assumes that the acceleration can be replaced by a time-dependent linear combination of the acceleration and its derivatives. Assuming that $a(t) = z''(t)$, the proposed model is:

$$a''(t) + c_1(t)a'(t) + c_0(t)a(t) = I(t)\xi(t) \tag{1.24}$$

with $c_1(t) = c_{10}$ and $c_0(t) = c_{00} + c_{01}t + c_{02}t^2 + c_{03}t^3$ for known coefficients $c_{10}, c_{00}, c_{01}, c_{02}, c_{03}$ and $\xi(t)$ Gaussian white noise.
In a similar way to the previous model, equation (1.24) can be rearranged as a system of SDEs

$$\begin{cases} dZ_1(t) = Z_1(t)dt \\ \\ dZ_2(t) = Z_2(t)dt \\ \\ dZ_3(t) = Z_3(t)dt \\ \\ dZ_4(t) = (-c_1(t)Z_4(t) - c_0(t)Z_3(t))dt + I(t)dW(t) \end{cases}$$

with $Z_1(t) = z(t)$, $Z_2(t) = z'(t)$, $Z_3(t) = z''(t) = a(t)$ and $Z_4(t) = z'''(t) = a'(t)$.

1.4.5 Investment Finance

One of the most relevant problems in the field of financial mathematics concerns the study of that process that governs the dynamics of an asset, that is, a financial object whose value is known but is subject to change in the future. One of the widely used model is the geometric Brownian motion, used to model stock prices in the Black–Scholes partial differential equation [148]:

$$dS(t) = \lambda S(t)dt + \sigma S(t)dW(t). \tag{1.25}$$

Here, the quantity $S(t)$ is the price of the asset at time t. The constant parameters λ and σ are respectively the measure of the average rate of growth

of the asset price (also known as *drift* term) and the measure of the standard deviation of the returns (also known as *volatility*).

The explicit solution of the linear SDE (1.25) is known and it is given by:

$$S(t) = S_0 \exp\left[\left(\lambda - \frac{1}{2}\sigma^2\right)t + \sigma W(t)\right] \tag{1.26}$$

with initial condition $S(0) = S_0$.

As seen in the following sections, this equation will often occur when one wants to study numerical linear stability for SDE. Straightforwardly, from (1.26), one can find that the expected value of (1.26) is

$$\mathbb{E}[S(t)] = S_0 \exp(\lambda t).$$

Hence, the average price increases exponentially, and is independent of σ.

Another very used model in mathematical finance is the so-called Ornstein-Uhlenbeck process, named after Leonard Ornstein and George Eugene Uhlenbeck. It is used to model interest rate dynamics. The Ornstein–Uhlenbeck process $S(t)$ is the solution of the following SDE

$$dS(t) = \lambda(\mu - S(t))dt + \sigma dW(t). \tag{1.27}$$

with initial condition $S(t_0) = S_0$ and μ constant parameter. The explicit solution of the equation (1.27) is known, and it is given in integral form by

$$S(t) = \mu + e^{\lambda t}(S_0 - \mu) + \sigma e^{-\lambda t}\int_0^t e^{\lambda s} dW(s).$$

It is important to note that the solution $S(t)$ of the model (1.27) could become negative, but if $\mu > 1.5\sigma^2/\lambda$, the probability of $S(t)$ being negative is rather small for any sufficiently large value of t [191].

1.4.6 Double-Well potential

A *potential function* is a sufficiently smooth function V that is bounded below and such that there exists a function x such that

$$\frac{dx}{dt} = -V'(x(t)). \tag{1.28}$$

This function is widely used in quantum mechanics, in quantum field theory and elsewhere for the exploration of various physical phenomena.
The stochastic additive version of (1.28) is given by

$$dX(t) = -V'(X(t))dt + \sigma dW(t) \tag{1.29}$$

where $\sigma \in \mathbb{R}$. The *double-well potential* [151] is defined as

$$V(x) = x^2(x-2)^2. \tag{1.30}$$

Applying (1.30) to (1.29), the SDE is

$$dX(t) = \left(-8X(t) + 12X(t)^2 - 4X(t)^3\right)dt + \sigma dW(t).$$

This problem will be used in the sequel of this thesis.

Chapter 2

Numerics for SDEs

As the reader noted in the previous chapter, there are many phenomena in which stochastic differential equations are involved. However, it seems evident that, as happens in the case of ordinary differential equations, there are very few equations whose explicit solution is known. Therefore, if you want to broaden the set of solvable stochastic differential equations, it is necessary to resort to numerical methods that allow us to find an approximate solution to the problem under consideration. Therefore this chapter will be entirely dedicated to numerics for SDEs, analyzing different numerical methods for SDE and studying he corresponding properties of convergence and numerical stability. Firstly, the Euler-Maruyama method will be remembered, considered as the stochastic counterpart of the classic Euler method for ODEs. Then the reader will see how it is possible to obtain numerical methods with higher order of convergence. Moreover a class of implicit methods, knows as stochastic θ methods, will then be analyzed. Finally a stochastic version of Runge-Kutta methods is given, obtaining methods as a stochastic perturbation of deterministic Runge-Kutta methods.

As the reader will observe in the sequel of this chapter, constructing numerical methods for SDEs with high order of strong convergence is by no means a trivial problem, since the stochastic perturbation of a deterministic ordinary differential equation involves the random variable $W(t)$ which behaves as $\mathcal{O}(\Delta t)$. This led researchers to construct a general theory for constructing higher-order numerical schemes for SDE, making use of the so-called Itô-Taylor stochastic expansion, considered as a stochastic generalization of the classical Taylor expansion for functions of the solution of an ODE [110].

A numerical method for ODEs of the form

$$\begin{cases} \dfrac{dx}{dt} = f(t, x(t)), \quad t \in [t_0, T], \\ \\ x(t_0) = x_0, \end{cases} \tag{2.1}$$

with $X(t) \in \mathbb{R}^d$ consists of discretizing the integration integral $[t_0, T]$ in a mesh of N points of the form $t_n = t_0 + n\Delta t$, with $\Delta t > 0$ and for $n = 0, \dots N-1$, and to generate a sequence of vectors $\{x_n\}$ such that x_n in an approximation of the solution $x(t_n)$. Numerical methods for ODEs have been extensively studied

over the years by many researchers; for example, it is possible to refer (but not exhaustively) to [76, 144, 146, 174, 233, 253] and references therein.

Following the same idea of numerical methods for ODEs, some of the numerical methods existing in the literature for SDEs and their relative properties of convergence and stability will be reviewed in the following subsections.

2.1 Euler-Maruyama method

Euler-Maruyama method (EM) is the simplest numerical method used to approximate a given SDE or a system of SDEs. It is considered as the stochastic counterpart of the explicit Euler method for ODEs [151, 150]. Remember that given a initial value problem of the form

$$\begin{cases} \dfrac{d}{dt}x(t) = f(t, x(t)) \qquad t \in [t_0, T], \\ \\ x(t_0) = x_0, \end{cases} \tag{2.2}$$

with $x(t) \in \mathbb{R}^d$, the explicit Euler method applied to (2.2) is

$$x_{n+1} = x_n + \Delta t\, f(t_n, x_n) \tag{2.3}$$

for $n = 0, \dots N-1$, where $x_n \approx x(t_n)$.

Now, consider a well-posed stochastic initial value problem

$$\begin{cases} dX(t) = f(t, X(t))\, dt + G(t, X(t))\, dW(t), \qquad t \in [t_0, T], \\ \\ X(t_0) = X_0. \end{cases} \tag{2.4}$$

Let $t_i = t_0 + i\Delta t$ be a uniform partition of the integration interval $[t_0, T]$ with step-size $\Delta t = (T - t_0)/N$, for some integer N. The integral formulation of (2.4), computed for $t \in (t_n, t_{n+1}]$, is given by

$$X(t) = X_0 + \int_{t_n}^{t} f(s, X(s))\, ds + \int_{t_n}^{t} G(s, X(s))\, dW(s), \quad t \in [t_0, T]. \tag{2.5}$$

By evaluating (2.5) for $t = t_{n+1}$, one gets

$$X(t_{n+1}) = X(t_n) + \int_{t_n}^{t_{n+1}} f(s, X(s))ds + \int_{t_n}^{t_{n+1}} G(s, X(s))dW(s). \tag{2.6}$$

Now, one can approximate the two integrals by the simple left-endpoint Riemann sums, i.e.,

$$\begin{aligned} &\int_{t_n}^{t_{n+1}} f(s, X(s)ds) \approx \Delta t\, f(t_n, X(t_n)), \\ &\int_{t_n}^{t_{n+1}} g(s, X(s)dW(s)) \approx \Delta W_n\, g(t_n, X(t_n)). \end{aligned} \tag{2.7}$$

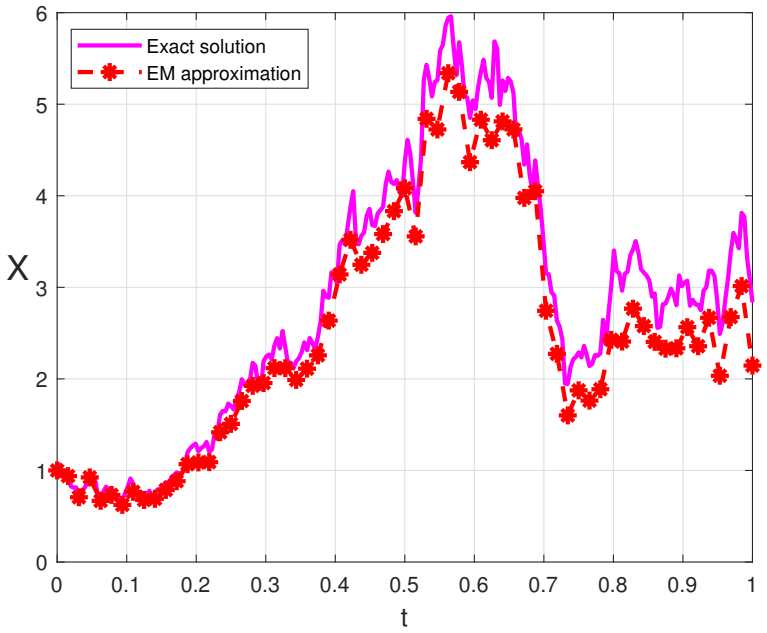


Figure 2.1: True solution and EM approximation of (1.25)

where $\Delta W_n = W(t_{n+1}) - W(t_n)$ is a normal variable with 0 mean and variance Δt. By inserting (2.7) into (2.6), the approximating sequence defined as

$$X_{n+1} = X_n + \Delta t f(t_n, X_n) + \Delta W_n G(t_n, X_n), \tag{2.8}$$

for $n = 0, \ldots, N-1$, defines the well known *Euler-Maruyama* method.
It is worth to observe that in the case $G \equiv 0$, with probability 1, and X_0 constant, then the method (2.8) reduces to the classic Euler method for ODEs (2.3).

Example 2.1.1. *Apply the Euler-Maruyama method (2.8) to equation (1.25) in the interval $[0, 1]$, choosing the parameters $\lambda = 2$, $\sigma = 1$ and initial condition $X_0 = 1$. To simulate a trajectory of the EM method, firstly, the Wiener process is discretized with resolution step-size $\delta t = 2^{-8}$ and then the EM method is applied with a step-size $\Delta t = R\delta t$ with $R = 4$ as follow.*

$$X_{n+1} = X_n + \lambda t X_n \Delta t + \sigma X_n \Delta W_n.$$

Figure 2.1 shows the exact solution (1.26) with a solid line and the approximate one through red asterisks. As noted in [150], the discrepancy between the true solution and the EM approximation at the endpoint $t = 1$ is 0.6907. Reducing the step-size $\Delta t = 2\delta t$ and $\Delta t = \delta t$, the endpoint errors become respectively 0.1595 and 0.0821. This suggests that the EM method's discrepancy error is reduced as the step-size Δt decreases or equivalently increases the number N.

In studying numerical methods for differential equations, it is crucial to establish whether, given a numerical scheme, it turns out to be convergent or

not, and if so, to study the order of convergence. Another important question concerns the study of (linear) stability, which looks at the asymptotic behavior when $T \to +\infty$. For these two reasons, the analysis of convergence and stability for the EM method will be discussed below.

Unlike the study of convergence for ordinary differential equations, in the stochastic case, two different definitions of convergence of numerical methods are distinguished: *weak convergence* and *strong convergence*. In a nutshell, the first will look at the error between the mean of the exact solution and the mean of the numerical one, while the latter will look at the mean of the norm of the error between the true solution and the approximate one. In detail, according to [150, 151], the *weak error* is

$$e_{\Delta t}^{\text{weak}} := \sup_{t_0 \leqslant t_n \leqslant T} \left| \mathbb{E}[\phi(X_n)] - \mathbb{E}[\phi(X(t_n))] \right|, \tag{2.9}$$

where the function $\phi : \mathbb{R}^d \to \mathbb{R}$ must satisfy smoothness and polynomial growth conditions. A numerical method is said *weakly convergent* if

$$\lim_{\Delta t \to 0} e_{\Delta t}^{\text{weak}} = 0. \tag{2.10}$$

Therefore, a numerical method has *weak order of convergence* γ if there exists a constant C (which does not depend on Δt) and a step-size level Δt^* (sufficiently small) such that

$$e_{\Delta t}^{\text{weak}} \leqslant C \Delta t^{\gamma}, \quad \text{for each } \Delta t \in (0, \Delta t^*). \tag{2.11}$$

Note that in the deterministic case (i.e. $G \equiv 0$ and non random initial value), the property (2.10) with $\phi(x) = x$ reduces to the usual deterministic convergence criterion. Under suitable conditions, one can prove that the EM method has weak order of convergence $\gamma = 1$, that is

$$e_{\Delta t}^{\text{weak}} \leqslant C \Delta t.$$

A sketch of the proof is given in [151, Chapter 9], but if the reader is interested in the complete proof, it is possible to refer to [165]. From an experimental point of view, Figure 2.2 illustrates the weak convergence order of the EM method, on a log-log scale, when applied to the linear equation (1.25) on $[0, T]$ with $T = 1$ and parameters $\lambda = 2$, $\mu = 0.1$ and $X_0 = 1$ (constant). Here the function ϕ of the equation (2.9) is the identity map. The asterisks in the Figure 2.2 represent the endpoint error for five different step-sizes $\Delta t = 2^{p-10}$ with $1 \leqslant p \leqslant 5$, computed with $M = 50000$ simulations over $[0, 1]$. Note that the exact solution of (1.25) has expected value

$$\mathbb{E}[X(T)] = e^{\lambda T}.$$

The expected value of the EM endpoint approximation $\mathbb{E}[X_N]$ is computed with a simple average, that is

$$\mathbb{E}[X_N] \sim \frac{1}{M} \sum_{s=1}^{M} X_N^{(s)},$$

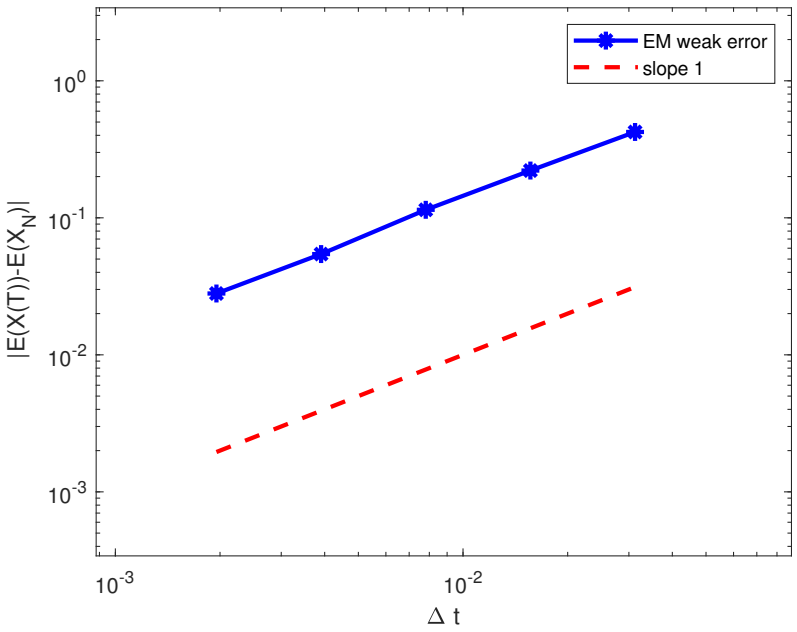


Figure 2.2: Weak endpoint error plot for EM method

where $X_N^{(s)}$ denotes the s-th sample path of the numerical solution at the endpoint $T = 1$.

Similarly to what was done for the study of weak convergence, the following quantity is defined as *strong error*:

$$e_{\Delta t}^{\text{strong}} := \sup_{t_0 \leqslant t_n \leqslant T} \mathbb{E}[|X(t_n) - X_n|]. \tag{2.12}$$

Basically, the strong error considers the expected value of the absolute error $|X(t_n) - X_n|$ at each time point. A numerical method is said to be *strongly convergent* if and only if

$$\lim_{\Delta t \to 0} e_{\Delta t}^{\text{strong}} = 0.$$

Consequentially a numerical method has *strong order* β if there exists a constant K and a step-size level Δt^* such that

$$e_{\Delta t}^{\text{strong}} \leqslant K \Delta t^{\beta}, \quad \text{for each } \Delta t \in (0, \Delta t^*).$$

Assuming f and G are two locally Lipschitz functions and the exact and numerical solution have a bounded p-th moment for some $p > 2$, then Higham et. al. in [152] have proved the EM method converges strongly. Furthermore, it is possible to prove that the strong order of the EM method is $\beta = \frac{1}{2}$, that is

$$e_{\Delta t}^{\text{strong}} \leqslant K \Delta t^{\frac{1}{2}}.$$

In [151, Chapter 10], a sketch of the proof is given.

Figure 2.3 shows the strong endpoint error, on a log-log scale, of EM method when it is applied to linear test problem (1.25) with $\lambda = 2$, $\sigma = 1$ and $X_0 = 1$

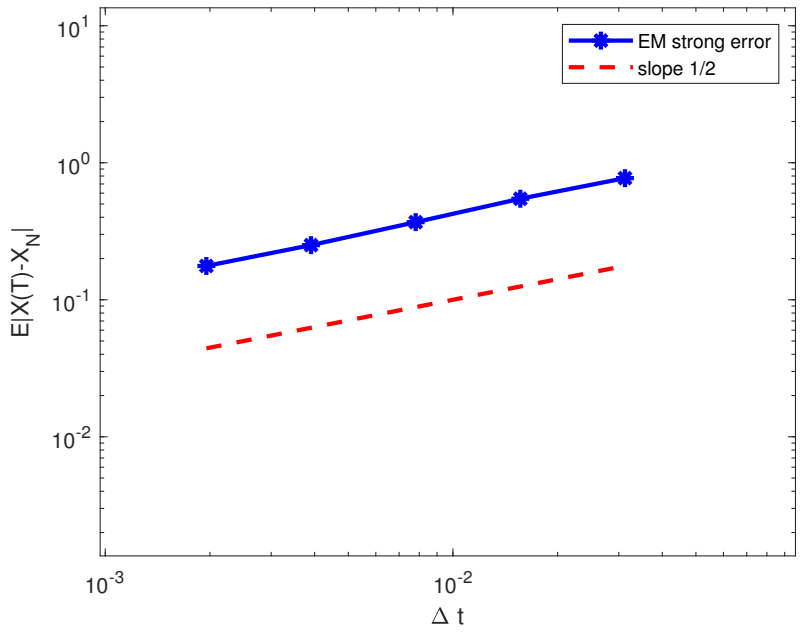


Figure 2.3: Strong endpoint error plot for EM method

(constant). Firstly, $M = 1000$ discretized Wiener process paths are computed with step-size $\delta t = 2^{-9}$ over $[0, 1]$ and for each path, EM method is applied with 5 different step-sizes $\Delta t = 2^{p-1}\delta t$ for $1 \leqslant p \leqslant 5$. The expected value of (2.12) is computed by

$$\mathbb{E}[|X(T) - X_N|] \sim \frac{1}{M}\sum_{s=1}^{M} |X(T) - X_N|^{(s)}$$

where $|X(T) - X_N|^{(s)}$ is the error between the exact solution (1.26) in the final point $T = 1$ and the approximate one, at the s-th sample path. Experimentally, Figure 2.3 confirms that EM method has strong order of convergence $\frac{1}{2}$.
It is also possible to prove that the EM method has strong convergence order $\beta = 1$ if applied to additive noise problems, i.e. when the function $G(t, X(t))$ of (1.1) is constant: see for instance, [151] for details.

However, there are some problems in which the Euler-Maruyama method may not converge in either a weak or a strong sense, as clarified by the following equation

$$dX(t) = -X(t)^3 dt + dW(t). \tag{2.13}$$

The drift coefficient of the problem does not satisfy the global Lipschitz condition. The proof of the existence and uniqueness of the solution related to this problem has been proved in [197]. In [155], the authors showed that there is a small (but not zero) probability that the EM applied to (2.13) produces a point $X_n = \Delta t^{-1}$ that makes the cubic term X_n^3 dominant. This growth can make the expectation uncontrollable and, therefore, the method could not converge in both a strong and a weak sense.

The study of numerics linked to stochastic differential equations goes on to the study of stability. Actually, the study of the convergence of the numerical methods regarding the study of the accuracy over a finite interval $[t_0, T]$. On the other hand, the stability property looks at the behaviour of the SDEs for a long time $T \to +\infty$. To make the analysis simpler, it will focus on the linear test equation (1.25)

$$dX(t) = \lambda X(t)dt + \sigma X(t)dW(t), \tag{2.14}$$

where the parameters λ and σ are allowed to be complex. Recalling that the exact solution to the above problem is

$$X(t) = X_0 \exp\left[\left(\lambda - \frac{1}{2}\sigma^2\right)t + \sigma W(t)\right]. \tag{2.15}$$

For this reason, the conducted analysis will also be called *linear stability*. From (2.15), it is possible to show that

$$\mathbb{E}[X(t)^2] = \mathbb{E}[X_0^2]e^{(2\lambda+\sigma^2)t}. \tag{2.16}$$

Assuming that $X_0 \neq 0$ with probability 1, the *mean-square stability* is defined as

$$\lim_{t\to+\infty} \mathbb{E}[X(t)^2] = 0. \tag{2.17}$$

This property (2.17) is completely characterized by the following condition

$$\mathrm{Re}\{\lambda\} + \frac{1}{2}|\sigma|^2 < 0, \tag{2.18}$$

where $\mathrm{Re}\{\lambda\}$ denotes the real part of the parameter λ. Therefore, for the SDE (2.14) to be mean-square stable, the parameters characterizing it must satisfy the property (2.18). Similarly, from (2.15) follows

$$\lim_{t\to+\infty} |X(t)| = 0, \text{ with prob. } 1 \quad \Leftrightarrow \quad \mathrm{Re}\left\{\lambda - \frac{1}{2}\sigma^2\right\} < 0. \tag{2.19}$$

The property on the left-side of (2.19) is known as *asymptotic stability*. Clearly, from (2.18) and (2.19), it is therefore easy to observe that mean-square stability implies asymptotic stability but not vice versa. Furthermore, both definitions, in the case in which the diffusion term $\sigma = 0$ is reduced to the same condition $\mathrm{Re}\{\lambda\} < 0$, which characterizes the deterministic linear stability.
Correspondingly, a numerical method for SDEs will be said to be:

1. *mean-square stable* if
$$\lim_{n\to+\infty} \mathbb{E}[X_n^2] = 0; \tag{2.20}$$
2. *asymptotic stable* if
$$\lim_{n\to+\infty} |X_n| = 0, \tag{2.21}$$
with probability 1.

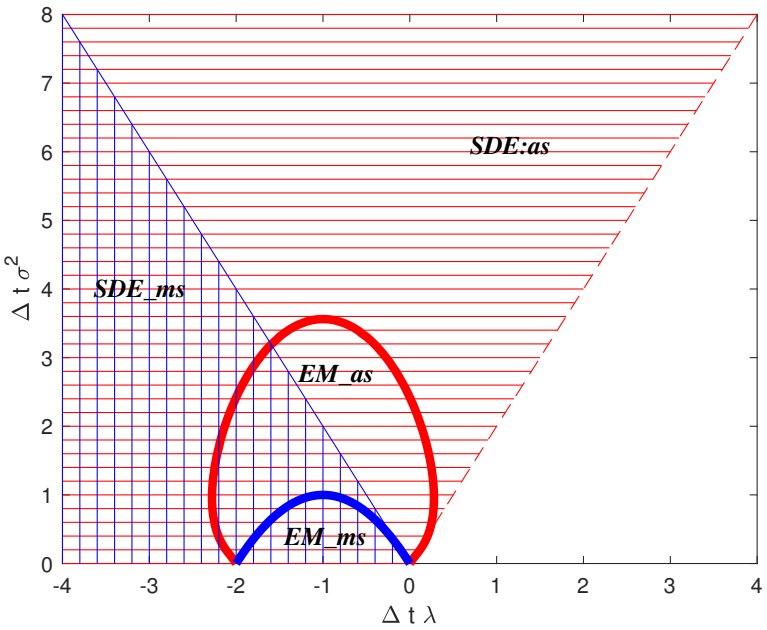


Figure 2.4: Mean-square and Asymptotic stability region to SDE (2.14) and EM method

Regarding the EM method, properties (2.20) and (2.21) can be adapted as follows:

$$\text{EM is } \textit{mean-square stable} \quad \Leftrightarrow \quad |1+\Delta t\lambda|^2 + \Delta t|\sigma|^2 < 1, \tag{2.22}$$

and

$$\text{EM is } \textit{asymptotically stable} \quad \Leftrightarrow \quad \mathbb{E}\left[\log\left|1+\Delta t+\sqrt{\Delta t}\sigma\mathcal{N}(0,1)\right|\right] < 0, \tag{2.23}$$

with $\mathcal{N}(0,1)$ standard normal variable i.e. normally distributed random variable with mean 0 and variance 1. In the case λ and σ real numbers, fixing $x = \Delta t\lambda$ and $y = \Delta t\sigma^2$, follows immediately that the mean-square stability condition (2.18) to SDE (2.14) becomes $y < -2x$ and the asymptotic stability (2.19) as $y > 2x$. The corresponding condition for the mean-square EM method (2.22) becomes $y < -x^2 - 2x$. Regarding the asymptotic stability of the EM method, in [151], the authors have shown the condition (2.23) is equivalent to the following one

$$\log|1+x| + \gamma(c) < 0,$$

where $c := y/(1+x)$ and $\gamma(c) = \mathbb{E}[\log|1+cV|]$, with V a standard normal variable. The explicit expression of the function $\gamma(c)$ is

$$\gamma(c) = \int_{-\infty}^{+\infty} \log|1+cs|e^{-s^2/2}ds. \tag{2.24}$$

According to the current notation, Figure 2.4 shows the mean-square and asymptotic stability regions. In particular vertical blue shading shows the

mean-square stability region of the SDE (2.14) (marked SDE_ms), and correspondingly, the mean-square stability region for EM (marked as EM_ms) is plotted with the solid blue curve. Of course, horizontal red shading shows the asymptotic stability region of the SDE (2.14) (marked as SDE_as), and with the solid red curve, the EM asymptotic stability region is plotted. The function (2.24) can be approximate by using a numerical quadrature formula [63, 95, 97, 207, 232].

2.2 Stochastic two-step methods

In this section a class of stochastic two-step methods is introduced, based on the idea of the deterministic multistep numerical methods, widely used in the deterministic context to numerically approximate the solution of a given ODEs [146, 156].

On a uniform grid, a stochastic two-step linear method for SDE (1.1), has the following form

$$\begin{aligned}\alpha_2 X_{n+1} + \alpha_1 X_n + \alpha_0 X_{n-1} = \\ \Delta t\,[\beta_2\, f(t_{n+1}, X_{n+1}) + \beta_1\, f(t_n, X_n) + \beta_0\, f(t_{n-1}, X_{n-1})] \\ + \gamma_1\, G(t_n, X_n)\Delta W_n + \gamma_0\, G(t_{n-1}, X_{n-1})\Delta W_{n-1}\end{aligned} \tag{2.25}$$

for $n = 1, 2, \ldots, N-2$, $\alpha_2, \alpha_1, \alpha_0, \beta_2, \beta_1, \beta_0, \gamma_1, \gamma_0 \in \mathbb{R}$ are the coefficients of the method and where X_0 and X_1 are assumed to be given. The two Wiener increments $\Delta W_n = W(t_{n+1}) - W(t_n)$ and $\Delta W_{n-1} = W(t_n) - W(t_{n-1})$ are required to be independent at each time step n. Moreover, as the reader can note, if $\beta_2 = 0$ one obtains an explicit numerical method, otherwise, if $\beta \neq 0$, one obtains an implicit numerical method, which requires a computation of a system of non linear equations (if the drift function f and the diffusion term G are non linear). The convergence and the linear stability of such methods have been studied in the literature (see [60]). Exponential non linear stability are, instead, analysed in [59].

2.3 High strong order method: Milstein scheme

So far, the Euler Maruyama method has been described, looking at its convergence and stability properties. In particular, it should be remembered that the EM method has a strong convergence order 1/2 and a weak convergence order equal to 1. This subsection aims to derive a high strong convergence order which will be called the Milstein method. It is named after Grigori N. Milstein, who first published it in 1974 [204]. Consider the scalar stochastic differential equation

$$\begin{cases} dX(t) = f(t, X(t))\, dt + g(t, X(t))\, dW(t), \qquad t \in [t_0, T], \\ X(t_0) = X_0, \end{cases}$$

that in integral form can be written as

$$X(t) = X_0 + \int_{t_0}^{t} f(s, X(s))\, ds + \int_{t_0}^{t} g(s, X(s))\, dW(s), \quad t \in [t_0, T]. \tag{2.26}$$

Keeping in mind that for a scalar function $F(t, X(t))$ of the solution $X(t)$ of (2.26), Itô's Lemma leads to

$$F(t, X(t)) = F(t_0, X(t_0)) + \int_{t_0}^{t} \mathcal{L}_0 F(s, X(s))\, ds + \int_{t_0}^{T} \mathcal{L}_1 F(s, X(s))\, dW(s), \tag{2.27}$$

where the two differential operators are given by

$$\mathcal{L}_0 = \frac{\partial}{\partial t} + f\frac{\partial}{\partial x} + \frac{1}{2}g^2\frac{\partial^2}{\partial x^2},$$

$$\mathcal{L}_1 = g\frac{\partial}{\partial x}.$$

It is easy to observe that:

- If $F(t, X(t)) = X(t)$ then the Itô's formula (2.27) is just the equation (2.26);
- If $F(t, X(t)) = f(t, X(t))$, then

$$f(t, X(t)) = f(t_0, X(t_0)) + \int_{t_0}^{t} \mathcal{L}_0 f(s, X(s))\, ds + \int_{t_0}^{T} \mathcal{L}_1 f(s, X(s))\, dW(s); \tag{2.28}$$

- Similarly, if $F(t, X(t)) = g(t, X(t))$, then

$$g(t, X(t)) = g(t_0, X(t_0)) + \int_{t_0}^{t} \mathcal{L}_0 g(s, X(s))\, ds + \int_{t_0}^{T} \mathcal{L}_1 g(s, X(s))\, dW(s). \tag{2.29}$$

Therefore, substituting (2.28) and (2.29) to (2.26) leads to

$$\begin{aligned} X(t) &= X(t_0) + \int_{t_0}^{t} \left[f(t_0, X(t_0)) + \int_{t_0}^{s} \mathcal{L}_0 f(u, X(u))\, du + \int_{t_0}^{s} \mathcal{L}_1 f(u, X(u))\, du \right] ds \\ &\quad + \int_{t_0}^{t} \left[g(t_0, X(t_0)) + \int_{t_0}^{s} \mathcal{L}_0 g(u, X(u))\, du + \int_{t_0}^{s} \mathcal{L}_1 g(u, X(u))\, du \right] dW(s) \\ &= X(t_0) + f(t_0, X(t_0)) \int_{t_0}^{t} ds + g(t_0, X(t_0)) \int_{t_0}^{t} dW(s) + R_1(t, t_0), \end{aligned}$$

with the remainder

$$\begin{aligned} R_1(t, t_0) &= \int_{t_0}^{t}\int_{t_0}^{s} \mathcal{L}_0 f(u, X(u))\, du\, ds + \int_{t_0}^{t}\int_{t_0}^{s} \mathcal{L}_1 f(u, X(u))\, dW(u)\, ds \\ &\quad + \int_{t_0}^{t}\int_{t_0}^{s} \mathcal{L}_0 g(u, X(u))\, du\, dW(s) + \int_{t_0}^{t}\int_{t_0}^{s} \mathcal{L}_1 g(u, X(u))\, dW(u)\, dW(s). \end{aligned} \tag{2.30}$$

Continuing once more by applying Itô formula to $\mathcal{L}_1 g$ in the fourth double integral of the remainder $R_1(t, t_0)$ in (2.30), the following Itô-Taylor expansion is obtained

$$\begin{aligned} X(t) = X(t_0) + f(t_0, X(t_0)) \int_{t_0}^{t} ds + g(t_0, X(t_0)) \int_{t_0}^{t} dW(s) \\ + \mathcal{L}_1 g(t_0, X(t_0)) \int_{t_0}^{t} \int_{t_0}^{s} dW(u)\, dW(s) + R_2(t, t_0), \end{aligned} \tag{2.31}$$

with the second remainder $R_2(t, t_0)$ given by

$$\begin{aligned} R_2(t, t_0) = &\int_{t_0}^{t} \int_{t_0}^{s} \mathcal{L}_0 f(u, X(u))\, du\, ds + \int_{t_0}^{t} \int_{t_0}^{s} \mathcal{L}_1 f(u, X(u))\, dW(u)\, ds \\ &+ \int_{t_0}^{t} \int_{t_0}^{s} \mathcal{L}_0 g(u, X(u))\, du\, dW(s) \\ &+ \int_{t_0}^{t} \int_{t_0}^{s} \int_{t_0}^{u} \mathcal{L}_0 \mathcal{L}_1 g(z, X(z))\, dz\, dW(u)\, dW(s) \\ &+ \int_{t_0}^{t} \int_{t_0}^{s} \int_{t_0}^{u} \mathcal{L}_1 \mathcal{L}_1 g(z, X(z))\, dW(z)\, dW(u)\, dW(s). \end{aligned}$$

Starting from equation (2.31), replacing t_0 by t_n and t by t_{n+1} and leaving out the remainder, the following numerical scheme is thus obtained

$$X_{n+1} = X_n + \Delta t f(t_n, X_n) + \Delta W_n g(t_n, X_n) + \mathcal{L}_1 g(t_n, X_n) \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} dW(u)\, dW(s). \tag{2.32}$$

It is possible to show that

$$\int_{t_n}^{t_{n+1}} \int_{t_n}^{t} dW(s)\, dW(t) = \frac{1}{2}((\Delta W_n)^2 - \Delta t). \tag{2.33}$$

Replacing (2.33) into (2.32), the *Milstein scheme* [151] is thus obtained:

$$\begin{aligned} X_{n+1} = X_n + \Delta t f(t_n, X_n) + \Delta W_n g(t_n, X_n) \\ + \frac{1}{2} \frac{\partial g}{\partial X}(t_n, X_n) g(t_n, X_n)((\Delta W_n)^2 - \Delta t). \end{aligned} \tag{2.34}$$

The Milstein method thus constructed has strong convergence order 1, but nevertheless, the weak convergence order remains unchanged from that of the Euler-Maruyama method, which is equal to 1. Figure 2.5 shows experimentally how the strong convergence order of the Milstein method is equal to 1. In particular, the Milstein method has been applied to the equation (1.16) by population dynamics by choosing the parameters $r = 2$, $K = 1$, $\beta = 0.25$ and initial condition $X_0 = 1$. The Wiener process is discretized over $[0, 1]$ with step-size $\delta t = 2^{-11}$. Since the exact solution of the problem involved a stochastic integral, then a reference solution can be obtained by applying the Milstein method with step size $\delta t = 2^{-11}$. This solution has been compared with the

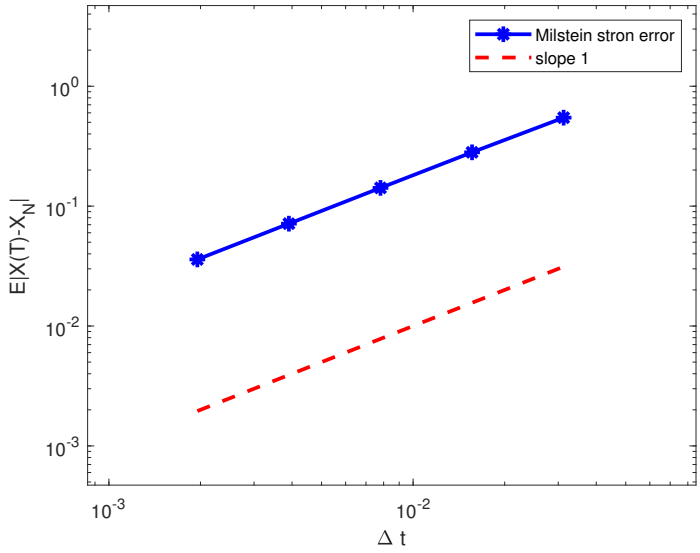


Figure 2.5: Strong endpoint error plot for Milstein scheme

Milstein approximation using step sizes $\Delta t = R\delta t$ with $R = 16, 32, 64, 128$ on 500 sample paths.
Clearly, the Milstein scheme in equation (2.34) refers to the case in which the SDE considered is of the scalar type. In a non-trivial way, it is possible to extend the Milstein method to the multidimensional case.
In particular, let $W(t) = [W_1(t), \ldots, W_m(t)]^\intercal$ be an m-multidimensional Wiener process and the function $G = [g^1(t, X(t)), \ldots, g^d(t, X(t))]^\intercal$, with the j-th component of $g^k = [g^{1,k}(t, X(t)), \ldots, g^{d,k}(t, X(t))]^\intercal$. Then a system of SDEs can be written as

$$dX(t) = f(t, X(t))dt + \sum_{k=1}^{m} g^k(t, X(t))dW^k(t).$$

In this case, the Milstein method becomes:

$$\begin{aligned} X_{n+1} = X_n + \Delta t f(t_n, X_n) + \sum_{k=1}^{m} g^k(t_n, X_n)\Delta W_n^k \\ + \sum_{j_1, j_2=1}^{m} \mathcal{L}^{j_1} g^{j_2}(t_n, X_n) \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} dW^{j_1}(s)\, dW^{j_2}(t). \end{aligned}$$

For details, the reader can refer to [151, 165]. Furthermore, the study of the stability of the Milstein method can be found, for instance, in [62, 149].

2.4 Stochastic θ-methods

The methods studied so far are explicit schemes, that is the solution in the time-step t_{n+1} of the considered uniform grid mesh depends only on the numerical solution at time-step t_n. It is well known from the deterministic case that

there are classes of differential equations (so-called stiff differential equation) that arise in many applications [2, 3, 4, 5, 6, 7, 8, 70, 73, 75, 142, 143, 221], for which it is not possible to apply an explicit numerical method since it would require a very small step size Δt. This fact makes the numerical computation, practically, impossible. For this reason, it is necessary to introduce implicit numerical methods, which, unlike the explicit ones, require the solution of an equation or, more precisely, a system of non linear equations. Although more expensive from a computational point of view, these methods have higher stability properties and, therefore, one can obtain a good approximation of the exact solution of the problem with a smaller step-size. The class of implicit numerical methods considered in this subsection are called stochastic θ-methods.

According to the definition of θ-methods for deterministic ODEs, the stochastic θ-methods for SDEs (1.1) is given by

$$X_{n+1} = X_n + (1-\theta)\Delta t f(t_n, X_n) + \theta \Delta t f(t_{n+1}, X_{n+1}) + \Delta W_n G(t_n, X_n) \quad (2.35)$$

with $\theta \in [0,1]$ and $\Delta W_n = [\Delta W_n^1, \ldots, \Delta W_n^m]^\intercal$. Note, in the case $\theta = 0$, the Euler-Maruyama methods is so obtained. When $\theta \neq 0$, a class of implicit numerical method is derived. In particular, the numerical method, obtained in correspondence of $\theta = 1$, is called the *stochastic backward Euler method*. It is also possible to show that if f satisfies the global Lipschitz condition, then (2.35) admits a unique solution for all sufficiently small Δt. Furthermore, for a definite classes of non linear functions f, a unique solution exists for any $\Delta t > 0$.

The linear stability properties of the θ-method are now studied, first for the mean-square case and subsequently for the asymptotic case. Before this, it is worth remembering that linear stability is studied on the scalar test linear equation (2.14) whose exact solution is given by (2.15). Also remember that conditions that guarantee the mean-square stability and asymptotic stability of SDE (2.14) are discussed in the subsection 2.1 by (2.18) and (2.19).

The stochastic θ-methods applied to linear test equation (2.14) gives

$$(1-\theta\Delta t\lambda)X_{n+1} = X_n(1 + (1-\theta)\Delta t\lambda + \sqrt{\Delta t}\sigma V_n) \quad (2.36)$$

where $V_n \sim \mathcal{N}(0,1)$ are independent and identically distributed (i.i.d.) random variables. By squaring both sides of (2.36) and considering the expected value, one has

$$\mathbb{E}[X_{n+1}^2] = \frac{(1 + (1-\theta)\Delta t\lambda)^2 + \Delta t\sigma^2}{(1-\theta\Delta t\lambda)^2}\mathbb{E}[X_n^2]. \quad (2.37)$$

with the assumption that $(1-\theta\Delta t\lambda) \neq 0$. However, this hypothesis is not restrictive because if $(1-\theta\Delta t\lambda)$ was equal to 0, this would imply that $\mu > 0$, which leads us to a non-stable SDE, in the sense that

$$\lim_{t \to +\infty} \mathbb{E}[X(t)^2] = +\infty.$$

Based on the definition (2.20), the stochastic θ-methods are mean-square stable if and only if

$$\frac{(1 + (1-\theta)\Delta t\lambda)^2 + \Delta t\sigma^2}{(1-\theta\Delta t\lambda)^2} < 1 \quad (2.38)$$

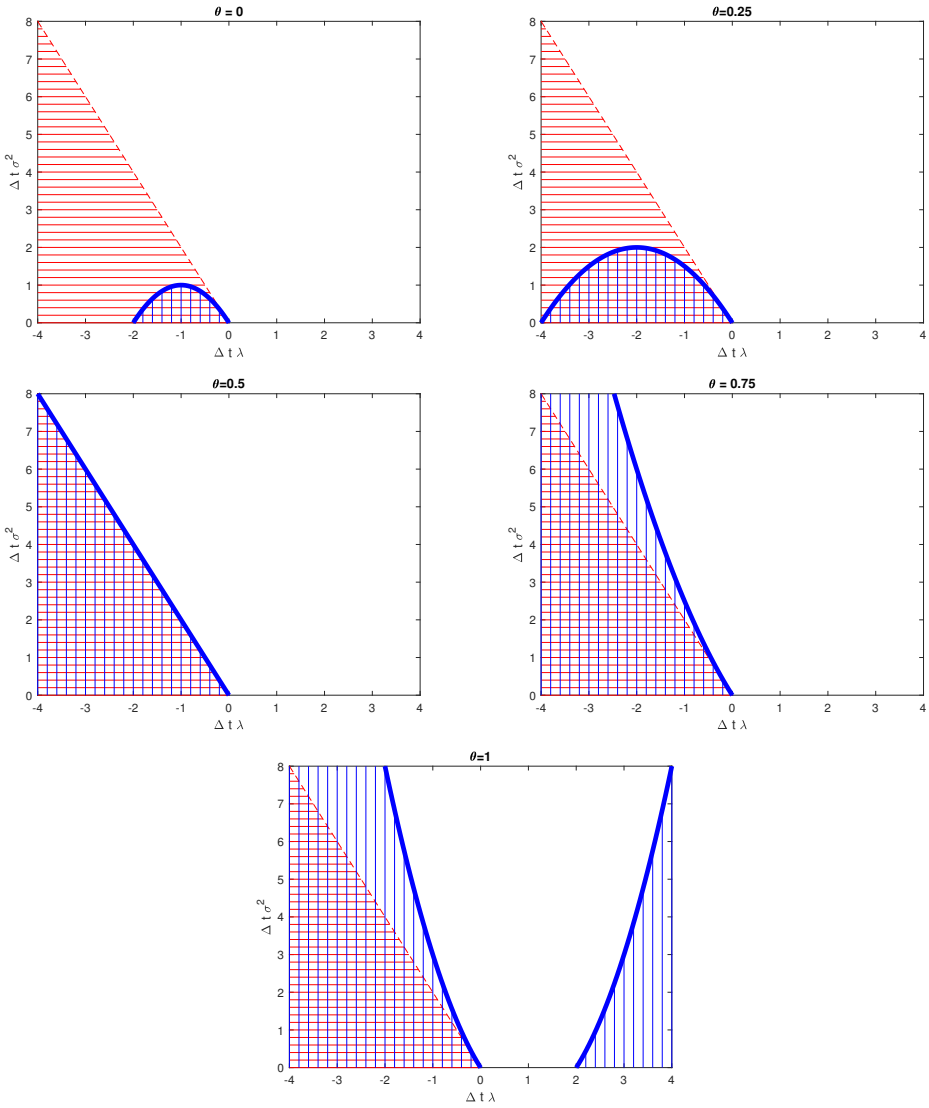


Figure 2.6: Mean-square stability region for the stochastic θ-method and the underlying SDE

holds. Equation (2.38) can also be read as

$$\Delta t(1-2\theta)\lambda^2 < -2\left(\lambda + \frac{1}{2}\sigma^2\right). \tag{2.39}$$

If the parameters λ and σ of the linear test SDE (2.14) are real, by setting $x = \Delta t\lambda$ and $y = \Delta t\sigma^2$, equation (2.39) can be easily rewritten as $y < (2\theta - 1)x^2 - 2x$, while the mean-square stability condition of the SDE (2.18) can be rewritten as $y < -2x$. Figure 2.6 shows the stability regions in a finite portion of the $x - y$ plane for different values of the θ parameter. In particular, for all figures, the horizontal red stripes show the SDE stability region, bounded by the line $y = -2x$, while the vertical blue lines show the stability region for the numerical method. Therefore, from Figure 2.6, one can observe that for $0 \leqslant \theta < \frac{1}{2}$ the method (2.35) has a finite stability region. For $\frac{1}{2} < \theta \leqslant 1$, the

method is *over-stable*, i.e. it is always stable on a stable SDE but can also be stable on a non-stable SDE. For $\theta = \frac{1}{2}$ the stability region of the numerical method coincides with the stability region of the SDE. To summarize, the following theorem, proved by Higham in [149] holds:

Theorem 2.4.1. *Consider the method* (2.35) *applied to the test problem* (2.14) *and let*

$$\Delta t_S := \frac{2|\lambda + \frac{1}{2}\sigma^2|}{(1-2\theta)\lambda^2}.$$

For $0 \leqslant \theta < \frac{1}{2}$,

1. *SDE stable* $\Rightarrow$ θ*-method stable for* $\Delta t < \Delta t_S$*;*
2. *SDE unstable* $\Rightarrow$ θ*-method unstable for ant* $\Delta t > 0$.

For $\frac{1}{2} < \theta \leqslant 1$

1. *SDE stable* $\Rightarrow$ θ*-method stable for all* $\Delta t > 0$*;*
2. *SDE unstable* $\Rightarrow$ θ*-method unstable for* $\Delta t < \Delta t_S$.

For $\theta = \frac{1}{2}$

1. *SDE stable* $\Leftrightarrow$ θ*-method stable for all* $\Delta t > 0$.

Let's focus now on the concept of asymptotic stability, which even if it is more complex to study, it is just as important as the mean-square stability, since it appears in many modelling contexts [191]. For the analysis it is crucial to rewrite (2.36) as

$$X_{n+1} = (a + bV_n)X_n, \tag{2.40}$$

with

$$a := \frac{1 + (1-\theta)x}{1 - \theta x}, \qquad b := \frac{\sqrt{y}}{1 - \theta x},$$

where $x := \Delta t\lambda$ and $y := \Delta t\sigma^2$ with assumption $1 - \theta x \neq 0$, as for mean-square stability analysis. The following Lemma, proved in [57], holds

Lemma 2.4.1. *The stochastic* θ*-methods are asymptotically stable if and only if*

$$\mathbb{E}[\log|a + bV_n|] < 0. \tag{2.41}$$

The condition (2.41) that characterized the asymptotic stability of the stochastic θ-methods can be rearranged as

$$\log|a| + \gamma(c) < 0, \tag{2.42}$$

where $c := b/a$ and $\gamma(c)$ is defined by (2.24). The following theorem gives us an important result about the boundness of the asymptotic stability region.

Theorem 2.4.2. *The stochastic* θ*-methods have unbounded asymptotic stability region if and only if* $\theta \geqslant \frac{1}{1+e^D}$*, where* $D := -\min_{c\in\mathbb{R}} \gamma(c)$.

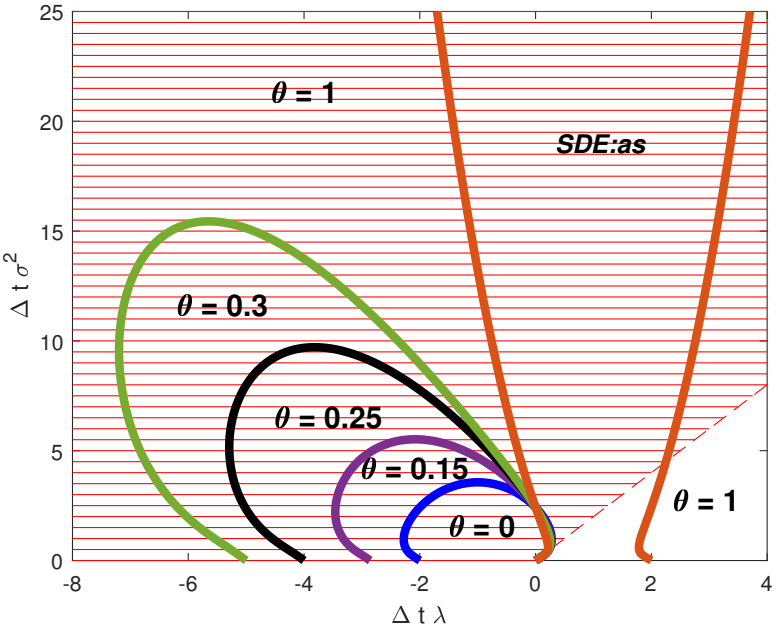


Figure 2.7: Asymptotic stability region for the stochastic θ-method

Numerically, it is possible to find that $D \approx 0.2454$ and hence $1/(1+e^D) \approx 0.4390$. Figure 2.7 plots the asymptotic stability region for different values of parameters θ and the asymptotic stability region of the SDE considered, that in this notation it has bounded by the red line $y = 2x$. Note that Figure 2.7 confirms the theoretical analysis and furthermore the asymptotic stability region increases monotonically with θ, and all are strictly contained in the SDE stability region.

2.5 Stochastic Runge-Kutta methods

In the study of numerics related to deterministic differential calculus, Runge-Kutta methods constitute an important family of numerical methods for solving initial value ordinary differential systems.

For a differential problem of the form

$$\begin{cases} y'(t) = f(t, y(t)), & t \in [t_0, T], \\ y(t_0) = y_0, \end{cases}$$

an s-stage Runge-Kutta method takes the form

$$\begin{aligned} y_{n+1} &= y_n + \Delta t \sum_{j=1}^{s} b_j f(t_n + c_j \Delta t, Y_j) \\ Y_i &= y_n + \Delta t \sum_{j=1}^{s} a_{ij} f(t_n + c_j \Delta t, Y_j), \quad i = 1, \dots, s. \end{aligned} \tag{2.43}$$

Often, the class of Runge-Kutta methods is represented by its Butcher tableau

$$\begin{array}{c|c} c & A \\ \hline & b^{\intercal} \end{array}$$

where $A = (a_{ij})_{i=1,\dots,s}^{j=1\dots,s}$, $b^{\intercal} = (b_1, \dots, b_s)$, $c = Ae$, $e = (1, \dots, 1)^{\intercal}$. Generally, when the matrix A is strictly lower triangular, the method is said to be explicit; otherwise, the method is implicit and may be suitable for integrating stiff problems if the method has appropriate stability properties.

Our interest is turned to the study of stochastic Runge-Kutta methods applied to the d-dimensional Itô SDE of the type

$$dX(t) = f(t, X(t))dt + \sum_{k=1}^{m} g^k(t, X(t))dW^k(t), \quad t \in [t_0, T] \tag{2.44}$$

driven by m-dimensional Wiener process and initial condition $X(t_0) = X_0$. Here $f : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $g^k : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ for $k = 1, \dots, m$. Often, this class of methods derives from stochastic perturbation of the associated deterministic Runge-Kutta method.

First of all, as the reader may recall, the Euler-Maruyama method for numerically solving SDEs (2.44) is the most basic numerical integration scheme. In general, as well as in the deterministic case, the stochastic Heun method produces an order of magnitude improvement in step size compared to the classical Euler-Maruyama method. For the multidimensional SDE (2.44), on the uniform grid

$$I_{\Delta t} = \{t_n = t_0 + n\Delta t,\ n = 0, 1, \dots, n,\ N\Delta t = T\},$$

the Heun scheme [131] is given by

$$\begin{aligned} X_{n+1} = X_n + \frac{1}{2}\Big[f(t_n, X_n) + f(t_{n+1}, \tilde{X}_{n+1})\Big]\Delta t \\ + \frac{1}{2}\Big[G(t_n, X_n) + G(t_{n+1}, \tilde{X}_{n+1})\Big]\Delta W_n, \end{aligned} \tag{2.45}$$

with the Euler predictor

$$\tilde{X}_{n+1} = \tilde{X}_n + f(t_n, \tilde{X}_n)\Delta t + G(t_n, \tilde{X}_n)\Delta W_n.$$

McShane, in [200], proved the scheme (2.45) converges in mean-square sense to the Itô solution of

$$\begin{aligned} dX(t) = \left[f(t, X(t) + \frac{1}{2}\sum_{k=1}^{m} \nabla_X \left(g^k(t, X(t)))g^k(t, X(t))\right)\right] dt \\ + \sum_{k=1}^{m} g^k(t, X(t))dW(t) \end{aligned}$$

where ∇_X denotes the Jacobian operator.
Notice that if one replaces f by $f - \frac{!}{2}\sum_{k=1}^{m} \nabla_X g^k(t, X(t)) g^k(t, X(t)))$ in (2.45), one obtains convergence to the solution of the Itô SDEs (2.44).

As in the case of ordinary differential calculus, one might be interested in obtaining methods with higher-order accuracy via higher-order stochastic Runge-Kutta schemes (SRK), introduced by Rümelin in [226]. The main problem related to these schemes concerns the lack of knowledge of the limit within which these methods converge (see [255, 166]).

Let $G(t, X(t)) = [g^1(t, X(t)), \dots, g^d(t, X(t))]^\intercal \in \mathbb{R}^{d\times m}$, the explicit $(s+1)$-th order SRK scheme is given by

$$X_{n+1} = X_n + \sum_{j=0}^{s} p_j K_j \, \Delta t + \sum_{j=0}^{s} q_j L_j \, \Delta W_n \tag{2.46}$$

where

$$\begin{aligned}
&K_0 = f(t_n + \alpha_0 \Delta t, X_n), \quad L_0 = G(t_n + \alpha_0 \Delta t, X_n)\\
&X_n^{(1)} = X_n + \beta_{10} K_0 \Delta t + \gamma_{10} L_0 \Delta W_n\\
&K_1 = f(t_n + \alpha_1 \Delta t, X_n^{(1)}), \quad L_1 = G(t_n + \alpha_1 \Delta t, X_n^{(1)})\\
&X_n^{(2)} = X_n + [\beta_{20} K_0 + \beta_{21} K_1]\Delta t + [\gamma_{20} L_0 + \gamma_{21} L_1]\Delta W_n\\
&K_2 = f(t_n + \alpha_2 \Delta t, X_n^{(2)}), \quad L_2 = G(t_n + \alpha_2 \Delta t, X_n^{(2)})\\
&\vdots\\
&\vdots\\
&X_n^{(s)} = X_n + \sum_{j=0}^{s-1} \beta_{sj} K_j \, \Delta t + \sum_{j=0}^{s-1} \gamma_{sj} L_j \, \Delta W_n\\
&K_s = f(t_n + \alpha_s \Delta t, X_n^{(s)}), \quad L_s = G(t_n + \alpha_s \Delta t, X_n^{(s)})
\end{aligned} \tag{2.47}$$

and

$$\sum_{j=0}^{s} p_j = \sum_{j=0}^{s} q_j = 1. \tag{2.48}$$

The coefficients of the method are then collected in the following Butcher tableau

$$\begin{array}{c|ccccc|ccccc}
\alpha_0 & & & & & & & & & & \\
\alpha_1 & \beta_{10} & & & & & \gamma_{10} & & & & \\
\alpha_2 & \beta_{20} & \beta_{21} & & & & \gamma_{20} & \gamma_{21} & & & \\
\vdots & \vdots & \vdots & \ddots & & & \vdots & \vdots & \ddots & & \\
\alpha_s & \beta_{s0} & \beta_{s1} & \cdots & \beta_{s,s-1} & & \gamma_s & \gamma_{s1} & \cdots & \gamma_{s,s-1} & \\
\hline
 & p_0 & p_1 & \cdots & p_{s-1} & p_s & q_0 & q_1 & \cdots & q_{s-1} & q_s
\end{array}$$

Rümelin in [226] established under what conditions it is possible to have convergence results (in mean-square sense) for SRK schemes. In particular, the following theorem gives us convergences results for one-dimensional SDE (i.e. SDE (2.44) with $d = m = 1$).

Theorem 2.5.1. *Suppose* $f, \partial f/\partial t, \partial f/\partial X, g, \partial g/\partial t, \partial g/\partial X, \partial^2 g/\partial t^2, \partial g^2/\partial t \partial X, \partial^2 g/\partial X^2$ *are bounded. Then the approximation* X_n *defined by* (2.46)-(2.48) *converges uniformly on* $[t_0, T]$ *in mean-square sense to the Itô solution* $Y(t)$ *of*

$$dY(t) = \left[f(t, X(t)) + \gamma \frac{\partial g}{\partial x}(t, X(t)) g(t, X(t)) \right] dt + g(t, X(t)) dW(t), \quad (2.49)$$

with correction factor $\lambda = 0$ *for* $s = 0$ *and*

$$\lambda = \sum_{i=1}^{m} q_i \sum_{j=0}^{i-1} \gamma_{ij}, \quad m \geqslant 1. \quad (2.50)$$

Note that if one replaces f by $f - \lambda g(\partial g/\partial X)$, then the scheme (2.46) converges uniformly to the solution of the Itô equation

$$dX(t) = f(t, X(t))dt + g(t, X(t))dW(t).$$

The family of SRK methods contains both the Euler Maruyama and Heun methods. In particular, the first one correspond to the choice $s = 0$ with $\alpha_0 = 0$ and $p_0 = q_0 = 1$. The second one correspond to the choice $s = 1$ with $\alpha_0 = 0$, $\alpha_1 = \beta_{10} = \gamma_{10} = 1$, $p_0 = p_1 = q_0 = q_1 = 1/2$ and the corresponding correction factor $\gamma = 1/2$. It is possible to prove that Theorem 2.5.1 can be easily obtained again in the case of a multidimensional SDE (2.44). Clearly the term $\partial g/\partial X g$ in (2.49) has to be replaced by

$$\sum_{k=1}^{m} (\nabla_X g^k) g^k,$$

where again $\nabla_X g^k$ denotes the Jacobian of g^k with elements $(\partial g_i^k/\partial X_j)_{ij}$. The correction factor λ remains the same as (2.50).

However, this class of explicit methods just described presents some problems related to the numerical order they can reach. Indeed Rümelin in [226] proved the following strong error barrier theorem.

Theorem 2.5.2. *Let* f *and* g *be arbitrary scalar functions and consider the corresponding scalar SDE*

$$dX(t) = f(t, X(t))dt + g(t, X(t))dW(t).$$

If f *and* g *have continuous and bounded partial derivatives up to the sixth order, then the one-step strong order (i.e. the order of magnitude of local error) of* (2.46)-(2.48) *cannot exceed 3/2.*

Note that the greatest order of magnitude of the one-step error is obtained by the Heun method. In addition, Rümelin also proved that, in the multidimensional case (2.44), the highest order of magnitude of one step error is 1, and it is obtained directly from the Euler-Maruyama method (see Gard's monograph [131]). Furthermore, the order 3/2 can only be reached if and only if the following commutative condition is satisfied:

$$(\nabla_X g^i) g^j = (\nabla_X g^j) g^i, \quad \text{for all } i, j = 1 \ldots, m.$$

Hence, if higher strong order methods are required, (2.46) must be modified in some way to include other stochastic elements. For example, K. Burrage et. al. in [64, 66, 67, 71, 72] have developed a class of high order stochastic Runge-Kutta methods, either explicit or implicit, thus breaking the barrier Theorem 2.5.2 proposed by Rümelin. In all these works in order to obtain order conditions, the authors generalized Butcher's tree theory [78] to the stochastic case. Other contributions in this area have been given by Rößler, building a new class of stochastic Runge-Kutta methods (either explicit or implicit) with high order of convergence [221, 223, 224]. Further, in the context of strong approximations of SDEs, SRK methods have been studied, for instance, in [16, 65, 70, 123, 165, 198, 203, 236, 242, 251]. Weak convergence of SRK methods have been discussed, e.g, in [16, 167, 168, 169, 187, 198, 203, 219, 220, 222, 244, 245].

The analysis of the numerical stability properties of various types of SRK methods has been extensively addressed by the existing literature; in particular, they focus their attention on determining which conditions the coefficients of the method should satisfy to obtain good stability properties. The interested reader can therefore see, for example [71, 72, 88, 116, 123, 173, 227, 228, 243] and references contained therein.

Chapter 3

Continuous-time numerical methods for SDEs

This section aims to describe some procedures to obtain continuous extensions of determining numerical methods for solving stochastic differential equations. As known in the deterministic case, continuous extensions allow us to get a dense output, this means knowing the approximate solution is not in the grid points of the mesh but in the whole integration interval. This clearly turns out to be essential to build a reasonable error estimate and then move on to constructing an efficient variable step algorithm, which is essential to solve stiff SDEs. Therefore this section will be structured as follows: first, a procedure for obtaining continuous extensions of methods for SDEs will be described based on the idea of numerical collocation for Volterra integral equations [93]. Subsequently, preliminary results about a procedure to obtain an efficient local error estimation will be given associated with selecting numerical methods for SDEs using their possible continuous extension.

3.1 Collocation method for Volterra Integral Equations

Numerical collocation for deterministic problems is a powerful and feasible technique for developing functional equations. In other words, the provided approximant (the *collocation function*) is constructed as a linear combination of selected basis functions, spanning a finite dimensional functional space; usually, these functions are piecewise algebraic polynomials. The collocation function is required to exactly satisfy the given equation at a selected set of points of the integration interval, denoted as *collocation points*. It is worth observing that the collocation function can be chosen as a linear combination of *ad hoc* basis functions, chosen coherently with the qualitative character of the given problem.

Collocation methods allow us to obtain an approximation over the entire integration interval of the solution of the given equation. This feature is mainly, for instance, to provide high continuous order methods not suffering from or-

der reduction when applied to stiff problems [77, 118], as well as to provide an efficient and accurate procedure of error estimation useful in variable step-size implementations [122]. Nonetheless, re-casting discrete numerical methods as collocation methods make their analysis benefit from their continuous formulation.

In this subsection, particular interest will be given to the study of collocation methods for Volterra integral equations (VIEs) because these will be the key point for developing continuous numerical methods for stochastic differential equations. In the literature, many authors have analyzed collocation methods for VIE, see [55, 56] and references therein contained. Consider VIEs of the form

$$y(t) = g(t) + \int_0^t k(t, s, y(s))\, ds, \quad t \in I = [0, T], \tag{3.1}$$

where $k \in C(D \times R)$, with $D = \{(t, s) : 0 \leqslant s \leqslant t \leqslant T\}$ and $g \in C(t)$. Let

$$I_{\Delta t} = \{t_n = n\Delta t,\ n = 0, 1, \ldots, n,\ N\Delta t = T\}, \quad \Delta t > 0, \tag{3.2}$$

be a uniform discretization of the interval I. With respect to (3.2), the problem (3.1) can be rewritten as

$$y(t) = F_n(t) + \phi_n(t), \quad t \in [t_n, t_{n+1}],$$

where the *lag term* F_n, which contains the past history of the phenomena, is

$$F_n(t) := g(t) + \int_0^{t_n} k(t, s, y(s))\, ds,$$

and the *increment function* $\phi_n(t)$ given by

$$\phi_n(t) := \int_{t_n}^t k(t, s, y(s))\, ds.$$

Now, let

$$0 \leqslant c_1 < \ldots < c_m \leqslant 1$$

be m collocation parameters, and denote by $t_{nj} = t_n + c_j\Delta t$ the so-called collocation points. Collocation methods allow us to approximate the exact solution $y(t)$ of (3.1) by a piecewise polynomial

$$P(t) \in S_{m-1}^{(-1)}(I_h) = \{v|_{(t_n, t_{n+1}} \in \Pi_{m-1}, n = 0, 1, \ldots, N-1\},$$

where Π_{m-1} denotes the space of (real) algebraic polynomials of degree not exceeding $m-1$. The collocation polynomial, restricted to the interval $[t_n, t_{n+1}]$, is of the form

$$P_n(t_n + \theta\Delta t) = \sum_{j=1}^m L_j(\theta) Y_{nj}, \quad \theta \in [0, 1], \tag{3.3}$$

for $n = 0, \ldots, N-1$, where $L_j(\theta)$ is the j-th Lagrange fundamental polynomial

$$L_j(\theta) := \prod_{k=1, k\neq j}^{m} \frac{\theta - c_k}{c_j - c_k}$$

and $Y_{nj} = P_n(t_{nj})$. The exact collocation equation is obtained by imposing that the polynomial (3.3) exactly satisfies the integral equation at the collocation points and requires, at each time step, the solution of a system of m nonlinear equations in the m unknowns Y_{ni}, as

$$\begin{cases} Y_{ni} = F_{ni} + \Phi_{ni}, \\ \\ y_{n+1} = \sum_{j=1}^{m} L_j(1) Y_{nj}, \end{cases} \tag{3.4}$$

where

$$F_{ni} = g(t_{ni}) + \Delta t \sum_{v=0}^{n-1} \int_0^1 k\left(t_{ni}, t_v + \theta \Delta t, P_v(t_v + \theta \Delta t)\right) d\theta, \tag{3.5}$$

$$\Phi_{ni} = h \int_0^{c_i} k\left(t_{ni}, t_n + \theta \Delta t, P_n(t_n + \theta \Delta t)\right) d\theta, \tag{3.6}$$

for $i = 1, \ldots, m$.

The maximum attainable order of convergence of such methods is at most $2m$, and it is achieved if Gauss nodes are used as collocation points in the iterated collocation methods. Furthermore, such methods have local order of convergence $2m - 2$ if one uses m Lobatto collocation points or $m - 1$ Gauss points with $c_m = 1$, and order $2m - 1$ if one uses m Radau II points. Unfortunately, the maximum order is gained only at the mesh points. In fact, the classical collocation methods have uniform order $O(\Delta t^m)$ over the entire integration interval.

Different from the ODEs case, the collocation equations are generally not yet in a form amenable to numerical computation due to the presence of the memory term.To make the method (3.4) a fully discretized numerical scheme, another discretization based on suitable quadrature formulas $\tilde{F}_{ni} \approx F_{ni}$ and $\tilde{\phi}_{ni} \approx \phi_{ni}$ for approximating the lag term and the increment function, are needed. This procedure leads to the well-known idea of *discretized collocation*. The discretized collocation polynomial is of the form

$$\tilde{P}_n(t_n + \theta \Delta t) = \sum_{j=1}^{m} L_j(\theta) \tilde{Y}_{nj}, \quad \theta \in [0, 1], \tag{3.7}$$

for $n = 0, \ldots, N-1$, where $\tilde{Y}_{nj} = \tilde{P}_n(t_{nj})$. The corresponding discretized collocation method assumes the form:

$$\begin{cases} \tilde{Y}_{ni} = \tilde{F}_{ni} + \tilde{\Phi}_{ni}, \\ \\ \tilde{y}_{n+1} = \sum_{j=1}^{m} L_j(1) \tilde{Y}_{nj}, \end{cases}$$

where

$$\tilde{F}_{ni} = g(t_{ni}) + \Delta t \sum_{v=0}^{n-1} \sum_{l=1}^{\mu_1} b_l k \left(t_{ni}, t_v + \xi_l \, \Delta t, \tilde{P}_v(t_v + \xi_l \, \Delta t) \right), \tag{3.8}$$

$$\tilde{\Phi}_{ni} = \Delta t \sum_{l=1}^{\mu_2} w_{il} k \left(t_{ni}, t_n + d_{il} \, \Delta t, \tilde{P}_n(t_n + d_{il} \, \Delta t) \right), \tag{3.9}$$

for $i = 1, \ldots, m$, where

$$(\xi_l, b_l)_{l=1}^{\mu_1} \qquad (d_{il}, w_{il})_{l=1}^{\mu_2}$$

are two quadrature formulas with nodes ξ_l and d_{il} such that

$$0 \leqslant \xi_1 < \ldots < \xi_{\mu_1} \leqslant 1$$

and

$$0 \leqslant d_{i1} < \ldots < d_{i\mu_2} \leqslant 1.$$

Here μ_1 and μ_2 are positive integers and b_l and w_{il} are the weights of the formulae.

Brunner in [56] has shown that under suitable conditions on the quadrature formulas (3.8) and (3.9), such methods preserve the same order of the exact collocation methods. Furthermore, it is possible to prove that a collocation method for VIEs is equivalent to an implicit Runge-Kutta method for VIEs if and only if $c_m = 1$. Other continuous extensions of Runge-Kutta methods for VIEs (not necessary with collocation technique) have been introduced in [25].

The most commonly used collocation methods are those based on the zeros of determined orthogonal polynomials, that is Lobatto, Radau and Gauss, which, remember, have respectively order of local convergence order $2m - 2$, $2m - 1$, $2m$, where m is the number of collocation points. The stability properties of such methods can be found in [26, 56, 81, 100] and references therein contained.

To try to increase the uniform convergence order of the one-step collocation methods, first in [180] for the ODEs and then in [96] for the VIEs, multistep collocation methods have been introduced. Generally, multistep collocation methods depend on more parameters than the classical ones without significantly increasing computational cost. Hence, many more freedom degrees are to be spent to obtain high stability properties and a high order of convergence. For this reason, as an immediate consequence, it can be shown that multistep collocation methods generally have a higher order of convergence than one-stage collocation methods with the same number of stages. Moreover, due to their high order of convergence, they do not suffer from the phenomenon of order reduction, which occurs in the integration of stiff problems. For all details, the interested reader can refer to [82] and the references contained therein.

3.2 Continuous extension of selected numerical methods for SDEs

Following the idea of collocation for Volterra integral equations, this section provides a procedure to obtain an approximation of the solution of a given $X(t)$ of (1.2) on the overall integration interval will be described. In particular an approximation $u(t)$ of the solution $X(t)$ of (1.2) will be provided, supposing that its restriction $u_n(t)$ to the interval $[t_n, t_{n+1}]$ is a linear polynomial.

To this purpose, let $c_1, \dots, c_m$ be the collocation parameters such that

$$0 \leqslant c_1 < \dots < c_m \leqslant 1$$

and let

$$t_{nj} = t_n + \Delta t c_j,$$

for $j = 1, \dots, m$, the corresponding collocation points in the interval $[t_n, t_{n+1}]$. The following ansatz is therefore required: in a sufficiently small interval of length Δt, the solution $X(t)$ to (1.2) can be approximated by the linear function

$$u_n(t_n + \theta \Delta t) = \sum_{j=1}^{m} L_j(\theta) U_j^{[n]}, \quad \theta \in [0, 1], \tag{3.10}$$

where $L_j(\theta)$ are the Lagrange fundamental polynomials with respect to the collocation parameters and $U_i^{[n]} = u_n(t_{ni})$, for $i = 1, \dots, m$.

Correspondingly, the following numerical scheme is obtained

$$\begin{cases} U_i^{[n]} = F_n + \Phi_i^{[n]}, \quad i = 1, \dots, m \\ X_{n+1} = \sum_{j=1}^{m} L_j(1) U_j^{[n]}, \end{cases} \tag{3.11}$$

where the lag-term

$$F_n = X_0 + \int_{t_0}^{t_n} f(s, u(s))\, ds + \int_{t_0}^{t_n} g(s, u(s))\, dW(s), \tag{3.12}$$

contains all the past history of the dynamics up to the grid point t_n, while the incremental terms $\Phi_i^{[n]}$, for $i = 1 \dots, m$, are defined by

$$\Phi_i^{[n]} = \int_{t_n}^{t_{ni}} f(s, u_n(s))\, ds + \int_{t_n}^{t_{ni}} g(s, u_n(s))\, dW(s).$$

By employing (3.10) and by a suitable change of variable, the incremental terms assume the form

$$\begin{aligned} \Phi_i^{[n]} = \Delta t \int_0^{c_i} f\left(t_n + \theta \Delta t, \sum_{j=1}^{m} L_j(\theta) U_j^{[n]} \right) d\theta \\ + \int_{t_n}^{t_{ni}} g\left(s, \sum_{j=1}^{m} L_j\left(\frac{s - t_n}{\Delta t} \right) U_j^{[n]} \right) dW(s). \end{aligned} \tag{3.13}$$

Clearly, the scheme thus written (3.11) requires the solution of a system of m-nonlinear equations in the unknowns $U_1^{[n]}, \dots, U_m^{[n]}$.

By setting $c_m = 1$, the scheme (3.11) may be substantively simplified. In fact, in this case, one can note that

$$\Phi_m^{[n]} = \int_{t_n}^{t_{n+1}} f(s, u_n(s))ds + \int_{t_n}^{t_{n+1}} g(s, u_n(s))dW(s), \tag{3.14}$$

therefore, by (3.11), (3.12) and (3.14),

$$U_m^{[n]} = F_n + \Phi_m^{[n]} = F_{n+1}, \quad n = 0, \dots, N-1.$$

As a consequence, by (3.10),

$$F_n = U_m^{[n-1]} = u_{n-1}(t_{n-1} + c_m\Delta t) = u_{n-1}(t_n) = X_n,$$

i.e. for each time step the lag term F_n is equal to the numerical solution X_n at the mesh point t_n. Moreover,

$$X_{n+1} = u_n(t_n + \Delta t) = u_n(t_n + c_m\Delta t) = U_m^{[n]}$$

and the method (3.11) assumes the form

$$\begin{cases} U_i^{[n]} = X_n + \Phi_i^{[n]}, \quad i = 1, \dots, m \\ X_{n+1} = U_m^{[n]}, \end{cases}$$

where the expressions for $\Phi_i^{[n]}$ are the same as (3.13), for $i = 1, \dots, m$.

As with the collocation methods for the VIEs, also in this case, one can observe that the method reported in (3.11) is not fully discretized. In fact, it requires the resolution of integrals both in the deterministic and in the stochastic part. To obtain a full discretization method, another discretization based on suitable quadrature formulas $F_n \approx \tilde{F}_n$ and $\Phi_i^{[n]} \approx \tilde{\Phi}_i^{[n]}$ for approximating the lag-term and the incremental terms, are needed. The discretized linear function (3.10) becomes

$$\tilde{u}_n(t_n + \theta\Delta t) = \sum_{j=1}^{m} L_j(\theta)\tilde{U}_j^{[n]}, \quad \theta \in [0, 1], \tag{3.15}$$

for $n = 0, \dots, N-1$ and $\tilde{U}_j^{[n]} := \tilde{u}_n(t_{nj})$. The m unknowns $U_j^{[n]}$ are determined by imposing that the linear function (3.15) satisfy exactly the given SDEs at the collocation points and by using suitable quadrature formulas for the approximation of the integrals. Therefore the corresponding discretized numerical scheme assumes the form:

$$\begin{cases} \tilde{U}_i^{[n]} = \tilde{F}_n + \tilde{\Phi}_i^{[n]}, \quad i = 1, \dots, m \\ \tilde{X}_{n+1} = \sum_{j=1}^{m} L_j(1)\tilde{U}_j^{[n]}, \end{cases} \tag{3.16}$$

where

$$\tilde{F}_n = X_0 + \sum_{l=1}^{\mu_0} b_l \, f(\xi_l, \tilde{u}_n(\xi_l)) + \sum_{l=1}^{\mu_1} A_l \, g(\omega_l, \tilde{u}_n(\omega_l)), \tag{3.17}$$

and

$$\begin{aligned}\tilde{\Phi}_i^{[n]} = \Delta t \sum_{l=1}^{\mu_2} \tau_{il} f(t_n + d_{il}\Delta t, \tilde{u}_n(t_n + d_{il}\Delta t)) \\ + \sum_{l=1}^{\mu_3} B_{il} g\left(\rho_{il}, \sum_{j=1}^{m} L_j \left(\frac{\rho_{il} - t_n}{\Delta t}\right) \tilde{U}_j^{[n]}\right).\end{aligned} \tag{3.18}$$

The formulas (3.17) and (3.18) are obtained by using deterministic quadrature formulas of the form

$$(\xi_l, b_l)_{l=1}^{\mu_0} \quad (d_{il}, \tau_{il})_{l=1}^{\mu_2}, \;\; i = 1, \ldots, m,$$

and the following stochastic quadrature formulas [140]

$$(\omega_l, A_l)_{l=1}^{\mu_1} \quad (\rho_{il}, B_{il})_{l=0}^{\mu_3}, \;\; i = 1, \ldots, m,$$

where $\xi_l, d_{il}, \omega_l, \rho_{il}$ and $b_l, \tau_{il}, A_l, B_{il}$ are suitable quadrature nodes and weights respectively and $\mu_0, \mu_1, \mu_2, \mu_3$ are positive integers. By setting the last collocation point $c_m = 1$, then, by virtue of the previous observations, $\tilde{F}^{[n]} = \tilde{X}_n$, and choosing as quadrature formulas the rectangular quadrature formula, the numerical scheme can be simplified. In fact, directly, the expression of $\tilde{U}_i^{[n]}$ for $i = 1, \ldots, m$ become

$$\tilde{U}_i^{[n]} = \tilde{X}_n + \Delta t \, c_i f\left(t_n, \sum_{j=1}^{m} p_j \tilde{U}_j^{[n]}\right) + g\left(t_n, \sum_{j=1}^{m} p_j U_j^{[n]}\right) \sqrt{c_i \Delta t} V_i^{[n]} \tag{3.19}$$

where $p_j = L_j(0)$ and $V_j^{[n]}$ are normal variables of mean 0 and variance 1, for any $j = 1, \ldots, m$. Furthermore, from (3.19), if one chooses $c_1 = 0$, then $U_1^{[n]} = \tilde{X}_n$.

Let us now consider a special case: let $m = 2$, $0 \leqslant c_1 < 1$ and $c_2 = 1$. The corresponding discretized method, obtained from (3.16) by approximating the integrals using the rectangular quadrature rule is

$$\begin{cases} \tilde{U}_1^{[n]} = \tilde{X}_n + \Delta t \, c_1 f\left(p_1 \tilde{U}_1^{[n]} + p_2 \tilde{U}_2^{[n]}\right) + g\left(p_1 \tilde{U}_1^{[n]} + p_2 \tilde{U}_2^{[n]}\right) \sqrt{c_1 \, \Delta t} V_1^{[n]}, \\ \tilde{U}_2^{[n]} = \tilde{X}_n + \Delta t \, f\left(p_1 \tilde{U}_1^{[n]} + p_2 \tilde{U}_2^{[n]}\right) + g\left(p_1 \tilde{U}_1^{[n]} + p_2 \tilde{U}_2^{[n]}\right) \sqrt{\Delta t} V_2^{[n]}, \\ \tilde{X}_{n+1} = \tilde{U}_2^{[n]}, \end{cases} \tag{3.20}$$

where, again, $p_1 = L_1(0)$, $p_2 = L_2(0)$ and $V_1^{[n]}$, $V_2^{[n]}$ are normal variables of mean 0 and variance 1.
The continuous approximant, in the interval $[t_n, t_{n+1}]$, assumes the form

$$\tilde{u}_n(t_n + \theta \Delta t) = \frac{1 - \theta}{1 - c_1} \tilde{U}_1^{[n]} + \frac{\theta - c_1}{1 - c_1} \tilde{U}_2^{[n]} \quad \theta \in [0, 1]. \tag{3.21}$$

Remark 3.2.1. *The proposed numerical scheme (3.20) requires the solution of a system of two non linear equations in the unknown variable $U_1^{[n]}$ and $U_2^{[n]}$. In order to discuss its solvability, two different approach are proposed. First of all, the system (3.20) in vector-form becomes*

$$U^{[n]} = \tilde{X}_n e + \Delta t\, f(p^\intercal U) + \sqrt{\Delta t}\, \alpha^{[n]} g(p^\intercal U), \tag{3.22}$$

where $U^{[n]} = [U_1^{[n]}, U_2^{[n]}]^\intercal$, $p = [p1, p2]^\intercal$, $c = [c_1, 1]^\intercal$, $\alpha^{[n]} = [\sqrt{c_1} V_1^{[n]}, V_2^{[n]}]^\intercal$, and $e = [1, 1]^\intercal$.
A first approach is based on the fixed point theorem. Assuming that f and g are globally Lipschitz continuous functions and let us consider the function

$$\varphi(U^{[n]}) := \Delta t\, c f(p^\intercal U^{[n]}) + \sqrt{\Delta t}\, \alpha^{[n]} g(p^\intercal U^{[n]}).$$

The solvability of (3.22) is equivalent to prove the function $\varphi(U^{[n]})$ is a contractive mapping. Therefore,

$$\varphi(U^{[n]}) - \varphi(Z^{[n]}) = \Delta t\, c \left[f(p^\intercal U^{[n]}) - f(p^\intercal Z^{[n]}) \right] + \Delta t\, \alpha^{[n]} \left[g(p^\intercal U^{[n]}) - g(p^\intercal Z^{[n]}) \right],$$

then

$$\begin{aligned} |\varphi(U^{[n]}) - \varphi(Z^{[n]})| &\leqslant \Delta t\, |c|\, |f(p^\intercal U^{[n]}) - f(p^\intercal Z^{[n]})| \\ &\quad + \sqrt{\Delta t}\, |\alpha^{[n]}|\, |g(p^\intercal U^{[n]}) - g(p^\intercal Z^{[n]})| \\ &\leqslant (\Delta t L_f |c| + \sqrt{\Delta t} L_g |\alpha^{[n]}|) |p^\intercal|\, |U^{[n]} - Z^{[n]}|. \end{aligned}$$

Hence, in order for the function $\varphi(U^{[n]})$ to be a contractive mapping, then Δt must be satisfy the following condition

$$\Delta t L_f |c| + \sqrt{\Delta t} L_g |\alpha^{[n]}| < \frac{1}{|p^\intercal|}.$$

A second approach is, instead, based on the implicit function theorem. If f and g are $\mathcal{C}^2$ functions, by defining

$$\psi(U^{[n]}, \tilde{X}_n, \Delta t, c, \alpha^{[n]}) := U^{[n]} - \tilde{X}_n e - \Delta t\, c f(p^\intercal U^{[n]}) - \sqrt{\Delta t}\, \alpha^{[n]} g(p^\intercal U^{[n]}),$$

then the solvability of $U^{[n]}$ in (3.22), for sufficiently small Δt, is equivalent in showing that

$$\left| \nabla_{U^{[n]}} \psi(U^{[n]}, \tilde{X}_n, \Delta t, c, \alpha^{[n]})|_{\Delta t = 0} \right| \neq 0.$$

Since

$$\nabla_{U^{[n]}} \psi(U^{[n]}, \tilde{X}_n, \Delta t, c, \alpha^{[n]}) = I_2 - \Delta t\, c\, p^\intercal f'(p^\intercal U^{[n]}) - \sqrt{\Delta t} \alpha^{[n]} p^\intercal g'(p^\intercal U^{[n]}),$$

then

$$\left| \nabla_{U^{[n]}} \psi(U^{[n]}, \tilde{X}_n, \Delta t, c, \alpha^{[n]})|_{\Delta t = 0} \right| = |I_2| = 1 \neq 0.$$

where I_2 is the identity matrix of dimension two. Hence, by the implicit function theorem, for sufficiently small value of Δt, the function $U^{[n]}$ is solvable.

Table 3.1: Absolute errors in the endpoint $T = 1$ obtained applying (3.11) to problem (3.25), for several number of grid points and for selected values of the parameter c_1.

N	$c_1 = 0$	$c_1 = 1/4$	$c_1 = 1/2$	$c_1 = 3/4$
2^7	$3.0967 \cdot 10^{-3}$	$3.0675 \cdot 10^{-3}$	$3.0095 \cdot 10^{-3}$	$2.8355 \cdot 10^{-3}$
2^6	$6.5909 \cdot 10^{-3}$	$6.5634 \cdot 10^{-3}$	$6.5083 \cdot 10^{-3}$	$6.3435 \cdot 10^{-3}$
2^5	$1.5071 \cdot 10^{-2}$	$1.5044 \cdot 10^{-2}$	$1.5071 \cdot 10^{-2}$	$1.4834 \cdot 10^{-2}$
2^4	$2.8721 \cdot 10^{-2}$	$2.8698 \cdot 10^{-2}$	$2.8653 \cdot 10^{-2}$	$2.8520 \cdot 10^{-2}$

Example 3.2.1. *The derived family of continuous methods (3.20) contains the EM method as a particular case when $c_1 = 0$. As a matter of fact, imposing $c_1 = 0$ in (3.20) and taking into account that $p_1 = 1$ and $p_2 = 0$, the method assumes the form*

$$\begin{cases} \tilde{U}_1^{[n]} = \tilde{X}_n, \\ \tilde{U}_2^{[n]} = \tilde{X}_n + \Delta t\, f(\tilde{X}_n) + g(\tilde{X}_n)\sqrt{\Delta t} V_2^{[n]} \end{cases}$$

and the value of $\tilde{X}_{n+1}$ can be evaluated as

$$\tilde{X}_{n+1} = \tilde{X}_n + \Delta t\, f(\tilde{X}_n) + g(\tilde{X}_n)\sqrt{\Delta t} V_2^{[n]}, \tag{3.23}$$

which corresponds to the EM method (2.8). In the case of the EM method, the continuous approximant assumes the form

$$\tilde{u}_n(t_n + \theta\Delta t) = (1-\theta)\tilde{X}_n + \theta\tilde{X}_{n+1} \quad \theta \in [0,1]. \tag{3.24}$$

From an experimental point of view, one can observe the accuracy of the new class of methods obtained above, emphasising its weak accuracy (2.11).

Let us assume the following SDE as a test problem

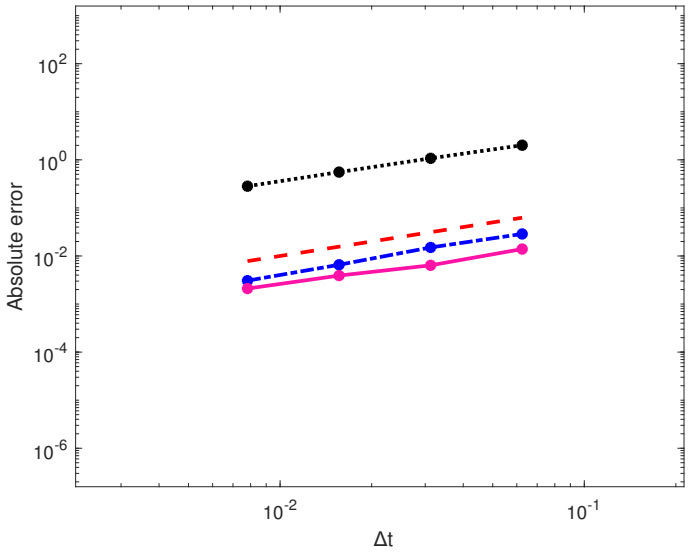
$$dX(t) = t^2 dt + e^{\frac{t}{2}} \cos(X(t)) dW(t), \quad X(0) = 0, \quad t \in [0,1]. \tag{3.25}$$

The expected value of the exact solution is $\mathbb{E}(X(t)) = \dfrac{t^3}{3}$.

Table 6.1 shows the absolute errors corresponding to various numerical methods belonging to the family (3.11), for selected values of c_1 and setting $c_2 = 1$, applied to problem (3.25). All methods exhibit the same weak order of convergence of EM method (2.8), which is equal to 1. In other terms, passing to the continuous extension does not deteriorate the order EM.
Let us now set $c_1 = \frac{1}{4}$ and $c_2 = 1$, and apply the corresponding method to the following additional SDE

$$dX(t) = e^{-\frac{t}{2}} dt + e^{X(t) - \frac{t}{2}} dW(t), \quad X(0) = 0, \quad t \in [0,1], \tag{3.26}$$

Figure 3.1: Absolute weak errors associated to the application of (3.11) with $c_1 = 1/4$ to problem (3.25) (dashed-dotted blue line), problem (3.26) (solid magenta line) and problem (3.27) (dotted black line). The reference slope of order 1 is depicted by the dashed red line.



whose exact solution has expected value $\mathbb{E}(X(t)) = 2 - 2e^{-\frac{t}{2}}$, and to the problem

$$d(X(t)) = (2X(t) + e^{2t})dt + X(t)dW(t), \quad X(0) = 1, \quad t \in [0, 1], \tag{3.27}$$

whose exact solution has expectation given by $\mathbb{E}(X(t)) = -e^{2t}(1 + t)$.

Figure 1 shows the behaviour of the absolute errors for decreasing step size values. One can observe that the method obtained by imposing $c_1 = 1/4$, applied to (3.25), (3.26) and (3.27), exhibits weak order of convergence equal to 1. So, for values of c_1 different from 0, the same weak order of convergence of EM method (2.8) is preserved.

3.3 Are continuous extension able to provide an LTE estimation?

This section aims to show purely preliminary results obtained from numerical experiments regarding an efficient and accurate estimation of the local truncation error (LTE) associated to selected numerical methods for autonomous SDEs of the type

$$\begin{aligned} dX(t) &= f(X(t))\, dt + g(X(t))\, dW(t), \quad t \in [t_0, T] \\ X(t_0) &= X_0, \end{aligned} \tag{3.28}$$

with $X(t) \in \mathbb{R}^d$ for each t, $f : \mathbb{R}^d \to \mathbb{R}^d$, $g : \mathbb{R}^d \to \mathbb{R}^{d\times m}$, and $W(t)$ is a m-dimensional Wiener process.

Specifically, in our analysis, Euler-Maruyama, Split-Step Backward Euler and Backward Euler methods are considered here.

The local truncation error of a numerical method is an estimate of the error introduced in a single iteration of the method, assuming that everything fed into the method was perfectly accurate and there is no round-off error. In a nutshell, to determine the LTE, one has to look at the following difference

$$\mathrm{lte}(t_{k+1}) := X(t_{k+1}) - \tilde{X}_{n+1} \tag{3.29}$$

where $X(\cdot)$ is the exact solution of (1.1) and $\tilde{X}$ is the numerical solution based on the assumption that it is determined from exact information.

Through a continuous extension of the methods mentioned above, in the spirit of [152], an estimation of LTE is therefore numerically obtained, extending the idea effectively used in the deterministic setting via numerical collocation (see, for instance, [118]).

Continuous-time extension of Euler-Maruyama method

Briefly, it is recalled that the Euler-Maruyama method for an autonomous SDE (3.28), respect to the uniform grid

$$I_{\Delta t} = \{t_k = k\Delta t, \quad k = 0, 1, \cdots, N, \quad N\Delta t = T\},$$

is given by

$$X_{k+1} = X_k + \Delta t f(X_k) + g(X_k)\Delta W_k. \tag{3.30}$$

A natural continuous-time extension (see [152]) of the discrete approximation defined in (3.30) is then given by

$$\tilde{X}(t) := X_k + (t - t_k)f(X_k) + g(X_k)(W(t) - W(t_k)), \quad \text{for } t \in [t_{k+1}, t_k), \tag{3.31}$$

or equivalently, by putting $t - t_k = s\Delta t$, (3.31) becomes

$$\tilde{X}(t_k + s\Delta t) := X_k + s\Delta t f(X_k) + \sqrt{s\Delta t} g(X_k)V_k, \quad \text{for } s \in [0, 1),$$

where V_k is a normal standard variable.

Continuous-time extension of Split-Step Backward Euler method

Before defining the continuous extension of the Split-Step Backward Euler method (SSBE), let us impose some useful assumptions on the SDE (3.28).

Assumption 3.3.1. *The functions f and g in (3.28) are C^1, and there exist constants $\mu, c > 0$, such that:*

$$\langle a - b, f(a) - f(b)\rangle \leqslant \mu|a - b|^2, \quad \forall a, b \in \mathbb{R}^d,$$

$$|g(a) - g(b)|^2 \leqslant c|a - b|^2, \quad \forall a, b \in \mathbb{R}^d.$$

where $|\cdot|$ denotes both the Euclidean vector norm and the Frobenius (or trace) matrix norm.

By taking $Y_0 = X_0$, the SSBE is defined as

$$\begin{cases} Y_k^* = Y_k + \Delta t f(Y_k^*), \\ Y_{k+1} = Y_k^* + g(Y_k^*)\Delta W_k, \end{cases} \tag{3.32}$$

where Y_k^* represents an intermediate approximation of the solution, useful to produce the successive approximation Y_{k+1} to the solution of (3.28). In our preliminary analysis, it is essential to remember a key property, which can be proved under hypotheses 3.3.1. Basically, it asserts that the SSBE method can be revised as the EM method applied to a modified SDE of a similar form, as visible from the following Lemma and Corollary, proved in [152].

Lemma 3.3.1. *Let us assume Assumption 3.3.1 and suppose* $\Delta t \in (0, \Delta t_c)$, $\Delta t_c < 1/(2\beta)$, *where* $\beta = \min\{\left(\mu + \frac{1}{2}\right), 2c\}$. *Given* $d \in \mathbb{R}^d$, *the implicit equation*

$$c = d + \Delta t f(c)$$

has an unique solution c. *By defining the functions* $F_{\Delta t}(\cdot)$, $f_{\Delta t}(\cdot)$ *and* $g_{\Delta t}(\cdot)$ *by*

$$F_{\Delta t}(d) = c, \quad f_{\Delta t}(d) = f(F_{\Delta t}(d)), \quad g_{\Delta t}(d) = g(F_{\Delta t}(d)),$$

then $F_{\Delta t}, f_{\Delta t}, g_{\Delta t} \in C^1$, $g_{\Delta t}(\cdot) \to g(\cdot)$ *and* $f_{\Delta t}(\cdot) \to f(\cdot)$ *as* $\Delta t \to 0$ *in* C^1 *uniformly on compact sets and, for any* $a, b \in \mathbb{R}^d$,

$$|f_{\Delta t}(a)| \leqslant \frac{|f(a)|}{1 - \Delta t\mu},$$

$$|F_{\Delta t(d)} - F_{\Delta t}(c)|^2 \leqslant \frac{1}{1 - 2\Delta t\mu}|d - c|^2,$$

$$\langle a - b, f_{\Delta t}(a) - f_{\Delta t}(b)\rangle \leqslant \frac{\mu}{1 - 2\mu\Delta t}|a - b|^2.$$

Further, $g_{\Delta t}$ *is globally Lipschitz, and there exist* $\alpha', \beta' > 0$ *such that*

$$\min\{\langle f_{\Delta t}(a), a\rangle, |g_{\Delta t}(a)|^2\} \leqslant \alpha' + \beta'|a|^2 \quad \forall a \in \mathbb{R}^m.$$

Corollary 3.3.1. *Let us assume Assumptions 3.3.1 and suppose* $\Delta t \in (0, \Delta t_c)$. *Then the SSBE applied to (3.28) is equivalent to the EM method applied to the modified SDE*

$$dy_{\Delta t}(t) = f_{\Delta t}(y_{\Delta t}(t))\, dt + g_{\Delta t}(y_{\Delta t})\, dW(t), \quad 0 \leqslant t \leqslant T, \quad y_{\Delta t}(0) = y_0, \tag{3.33}$$

where $f_{\Delta t}, g_{\Delta t}$ *are defined in Lemma 3.3.1.*

Therefore, Corollary 3.3.1 allows us to rewrite SSBE as the EM method applied to (3.33). Therefore, the continuous-time extension of SSBE can be obtained from (3.31) with f and g replaced by $f_{\Delta t}$ and $g_{\Delta t}$, respectively. Then, it is given by

$$\overline{Y}(t) := Y_k + (t - t_k) f_{\Delta t}(Y_k) + g_{\Delta t}(Y_k)(W(t) - W(t_k)), \quad \text{for } t \in [t_{k+1}, t_k). \tag{3.34}$$

In an equivalent formulation, by taking $t - t_k = s\Delta t$, (3.34) becomes,

$$\overline{Y}(t_k + s\Delta t) := Y_k + s\,\Delta t\, f_{\Delta t}(Y_k) + \sqrt{s\Delta t}\, g_{\Delta t}(Y_k) V_k, \tag{3.35}$$

where V_k is a normal standard variable.

Continuous-time extension of Backward Euler

The stochastic backward Euler (BE) method [152] applied to (3.28) is given by $Z_0 = X_0$ and

$$Z_{k+1} = Z_k + \Delta t\, f(Z_{k+1}) + g(Z_k)\Delta W_k. \tag{3.36}$$

The following result gives us a connection between the solution computed by the SSBE method and that arising from the BE method [152].

Lemma 3.3.2. *Let $\{Y_k\}$ and $\{Z_k\}$ denote SSBE and BE solutions, respectively given by (3.32) and (3.36), respectively. Under Assumption 3.3.1, if $Y_0 = Z_0 - \Delta t\, f(Z_0)$, then*

$$Z_k = Y_k + \Delta t\, f_{\Delta t}(Y_k), \quad \forall k \geqslant 0.$$

By using Lemma 3.3.2, it is natural to define the continuous-time extension to the BE method as

$$\overline{Z}(t) = \overline{Y}(t) + \Delta t\, f_{\Delta t}(\overline{Y}(t)), \tag{3.37}$$

where $\overline{Z}(0) = X_0$. Here $\overline{Y}(t)$ denotes the continuous-time extension to SSBE (3.34) and $f_{\Delta t}$ is defined in the previous subsection.

3.4 Numerical tests

In this section, based on the continuous extensions provided in the previous section, numerical evidences on selected test problems highlight the effectiveness of our preliminary procedure to estimate the local truncation error for the selected one-step stochastic numerical methods. In particular, the Figures 3.2, 3.3, 3.4 show how the estimate tends to approach the true local error, as the value of the step size Δt decreases. Specifically, the following problems are considered:

- the equation of geometric Brownian motion

$$\begin{cases} dX(t) = -2X(t)dt + X(t)dW(t), & t \in [0, 50], \\ X(0) = 1, \end{cases} \tag{3.38}$$

  whose exact solution is

$$X(t) = \exp\left(-\frac{5}{2}t + W(t)\right);$$

- the nonlinear problem

$$\begin{cases} dX(t) = -X(t)(1 - X^2(t))dt + (1 - X^2(t))dW(t), & t \in [0, 10], \\ X(0) = X_0 = 0.25, \end{cases} \tag{3.39}$$

  with the exact solution given by

$$X(t) = \frac{(1 + X_0)\exp(2W(t)) + X_0 - 1}{(1 + X_0)\exp(2W(t)) - X_0 + 1};$$

- the nonlinear SDE

$$\begin{cases} \mathrm{d}X(t) = -\dfrac{1}{900}X(t)(1-X^2(t))\mathrm{d}t + \dfrac{1}{30}(1-X^2(t))\mathrm{d}W(t), & t \in [0,20], \\ X(0) = \dfrac{1}{10}, \end{cases} \tag{3.40}$$

  whose exact solution is given by

$$X(t) = \tanh\left(\frac{1}{30}W(t) + \operatorname{arctanh}\left(X(0)\right)\right).$$

- the non linear problem with two Wiener process

$$\begin{cases} \mathrm{d}X(t) = -\dfrac{1}{900}X(t)(1-X^2(t))\mathrm{d}t + \dfrac{1}{30}(1-X^2(t))\mathrm{d}W_1(t) + 0.01\mathrm{d}W_2(t), \\ X(0) = 0.9, \end{cases} \tag{3.41}$$

  with $t \in [0,5]$. Since the exact solution is not knows, the reference one is given by applying the considered numerical schemes with $N = 2^{16}$.

- the system of SDEs

$$\begin{cases} \mathrm{d}X_1(t) = (-X_1(t) + X_2(t))\mathrm{d}t + 0.1X_1(t)\mathrm{d}W_1(t) \\ \mathrm{d}X_2(t) = (X_1(t) - X_2(t))\mathrm{d}t + 0.1X_2(t)\mathrm{d}W_2(t) \qquad t \in [0,1]. \\ X_1(0) = 1, X_2(0) = 0. \end{cases} \tag{3.42}$$

  The reference solution is given by applying the considered numerical schemes with $N = 2^{16}$.

The results, contained in Figures 3.2, 3.3, 3.4, 3.5 and 3.6 show the plot of the local error and its estimation for Euler-Maruyama, Split step Backward Euler and Backward Euler methods applied to the above problems, with $N = 2^9$ and $N = 2^{13}$ grid points. The expected values of the errors and their estimates are computed over 1000 paths. In particular, these figures confirm the accuracy of the estimation in all considered cases.

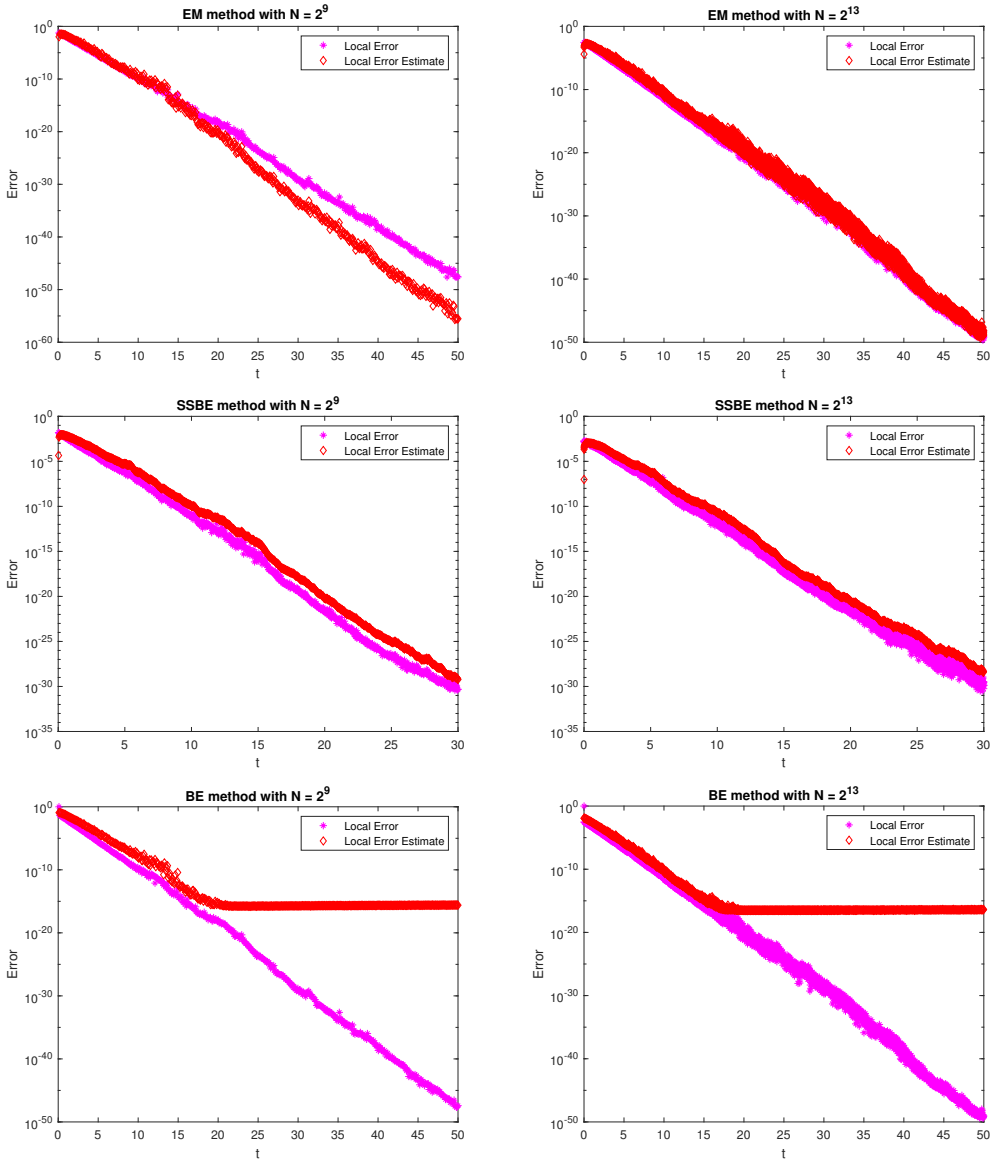


Figure 3.2: **Left:** Local error estimate VS true local error for EM, SSBE and BE methods applied with $N = 2^9$ grid points over 1000 paths on (3.38); **right:** same scenario with $N = 2^{13}$ grid points.

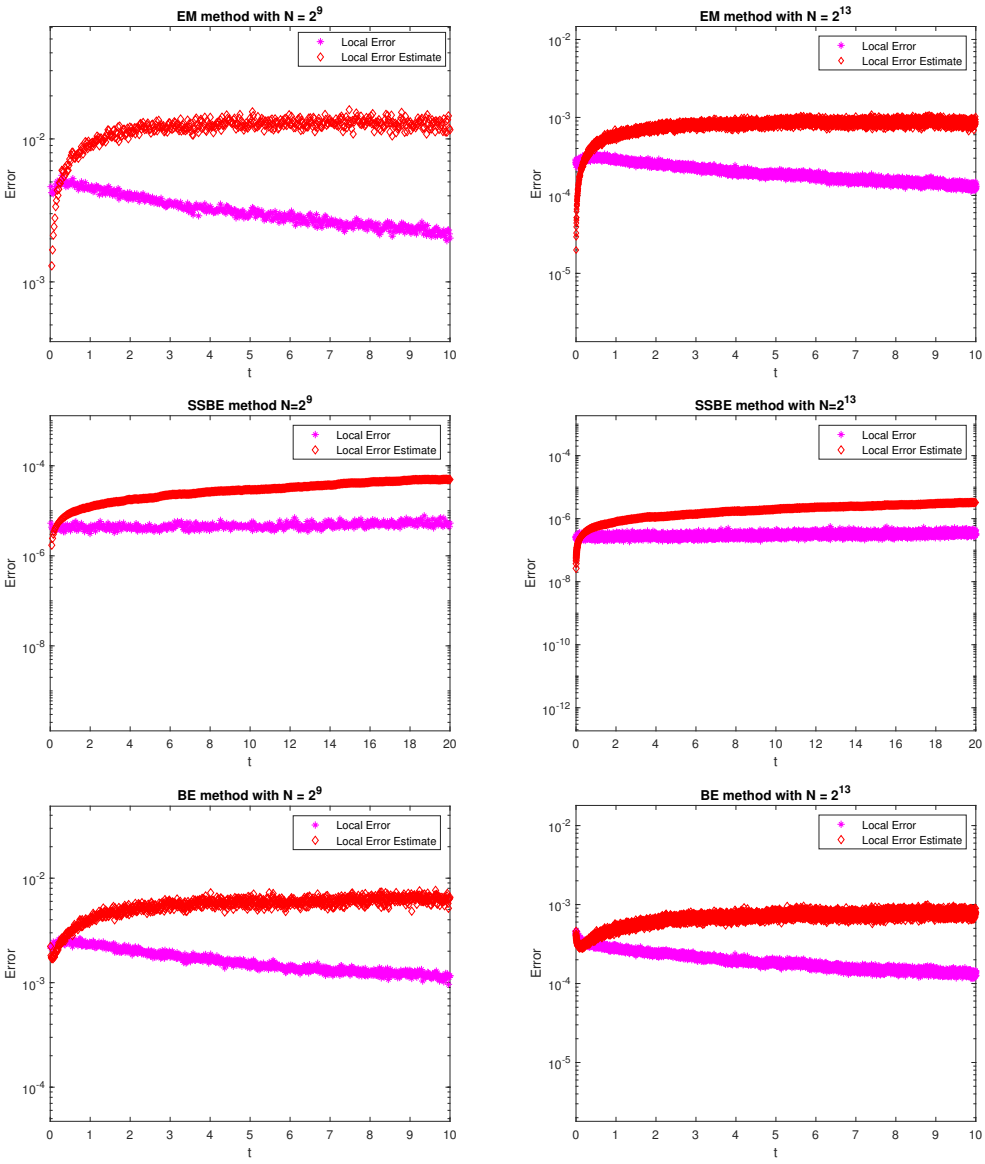


Figure 3.3: **Left:** Local error estimate VS true local error for EM, SSBE and BE methods applied with $N = 2^9$ grid points over 1000 paths on (3.39); **right:** same scenario with $N = 2^{13}$ grid points.

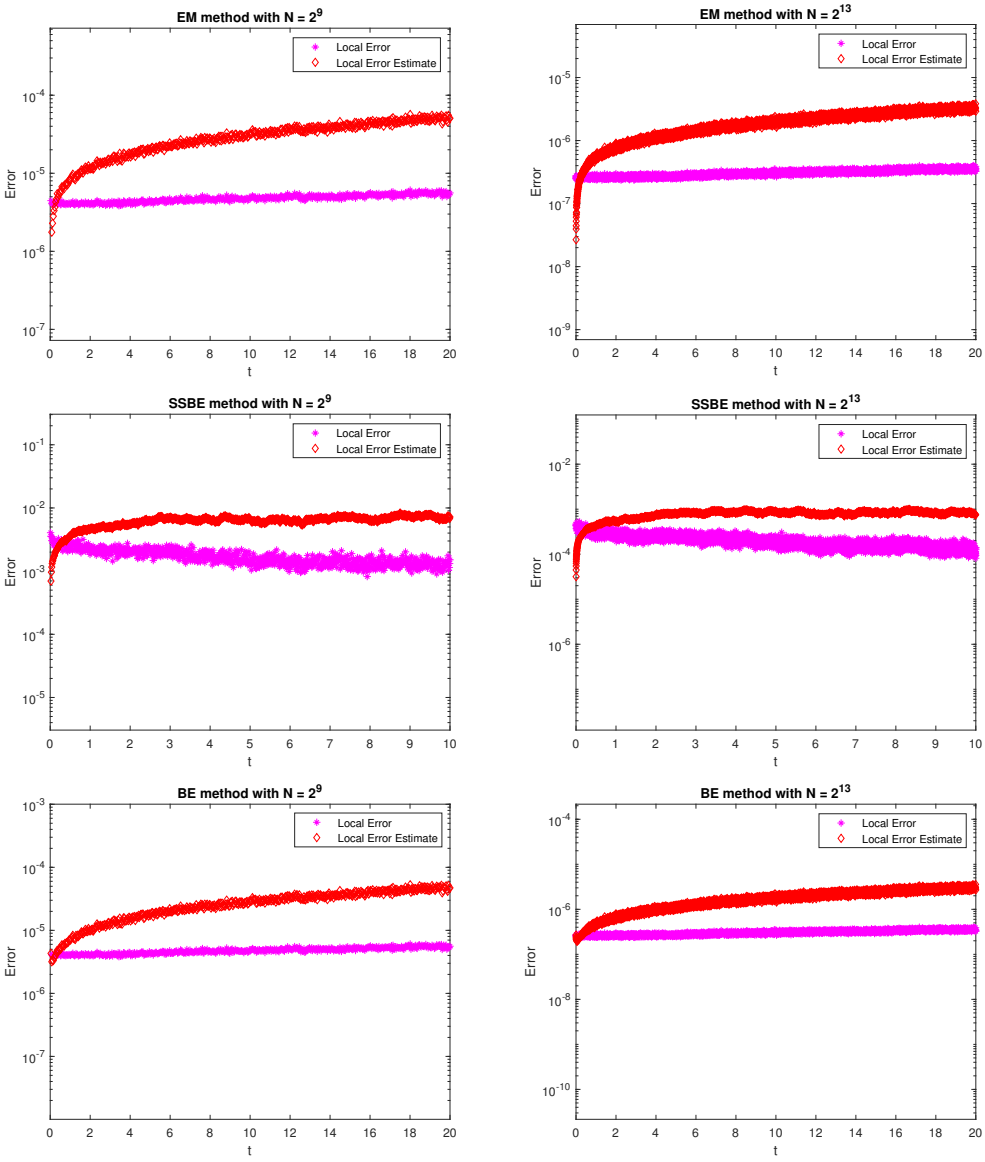


Figure 3.4: **Left:** Local error estimate VS true local error for EM, SSBE and BE methods applied with $N = 2^9$ grid points over 1000 paths on (3.40); **right:** same scenario with $N = 2^{13}$ grid points.

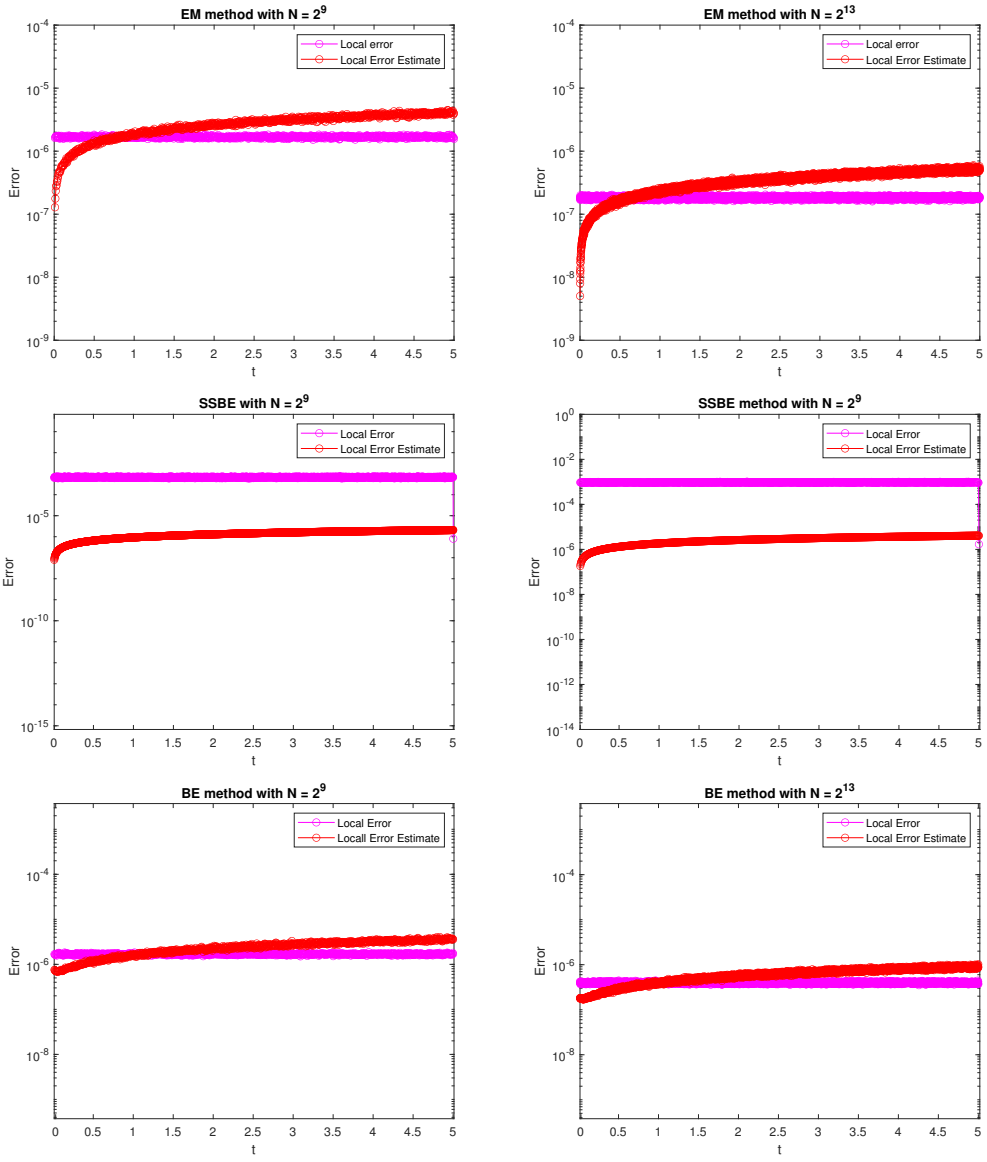


Figure 3.5: **Left:** Local error estimate VS true local error for EM, SSBE and BE methods applied with $N = 2^9$ grid points over 1000 paths on (3.41); **right:** same scenario with $N = 2^{13}$ grid points.

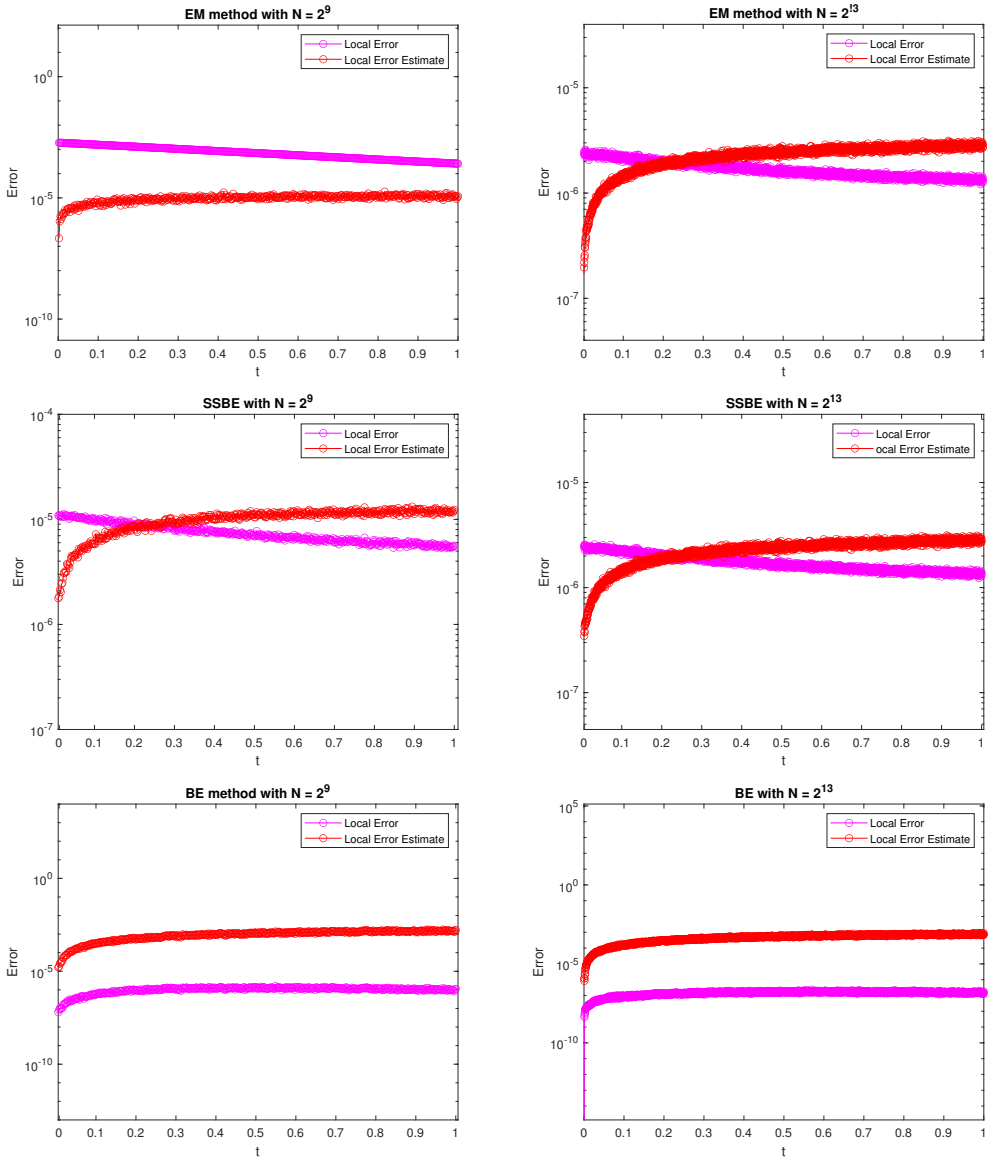


Figure 3.6: **Left:** Local error estimate VS true local error for EM, SSBE and BE methods applied with $N = 2^9$ grid points over 1000 paths on (3.42); **right:** same scenario with $N = 2^{13}$ grid points.

Chapter 4

Stochastic Hamiltonian Problems

A Hamiltonian problem is a dynamic system used to describe the evolution of a physical system. Hamilton's equations first appeared in 1834 by Hamilton, inspired in part by his earlier work in the field of optics. Later, these equations are found to be essential in multiple branches of physics and beyond and range from the nanoscale to molecular dynamics at the macroscale of celestial mechanics. Hamiltonian equations arise as models in weather prediction and meteorology [98, 234], solid mechanics [17, 134, 240] and elastodynamics [138, 238], nonlinear optics [133], oceanography [18, 158], cosmology [215, 250], electromagnetism [27] and quantum field theory [106, 170, 225], for example. Other applications can be found in [1, 112, 141, 210, 256, 258].
Clearly, when the dynamics of the system are affected by random noisy perturbation, deterministic Hamiltonian systems are not well matched, therefore stochastic versions are preferable instead [68, 74, 84, 121, 185, 260].
The irreversible character of stochastic dynamics destroys the idea of isolated systems, since the particles are repeatedly influenced by small unpredictable perturbations of the external environment. In order to combine the laws of evolution for average macroscopic observable with the microscopic dynamics of the constituent particles of a statistical system, it is necessary that the microscopic dynamics have a highly chaotic character and the overall system can be decomposed into almost independent microscopic subsystems. Therefore, stochastic Hamiltonian problems are the most suitable candidates to conciliate the Hamiltonian nature of classical mechanics, which is closely related to the canonical character of the evolution equations, with the non-differentiable nature of the Wiener process, which describes the continuous innovative character of stochastic effects [21, 22]. In particular, in this thesis the behaviour of discretizations to these problems are analyzed, motivated by some results on stochastic Runge-Kutta methods (SRK) developed by Burrage and Burrage in [68]. In fact, SRK methods present a remarkable error that increases with the parameter of the diffusive part of the problem. Through a perturbative analysis, the reason of this behaviour is investigated, leading to a negative answer: retaining the main features of stochastic Hamiltonian problems does not happen straightforwardly for any time discretization.
This section is organized as follows: first of all a new class of stochastic Runge-

Kutta methods are introduced and subsequently our attention will be turned to stochastic Hamiltonian problems first considering an additive type noise and then linear type. Therefore the perturbation analysis will be performed for both cases.

4.1 Hamiltonian problems

Generally, a deterministic Hamiltonian system on the interval $[t_0, T]$ is represented in the form

$$\begin{aligned} p'(t) &= -\frac{\partial}{\partial q}\mathcal{H}(p(t), q(t)), \\ q'(t) &= \frac{\partial}{\partial p}\mathcal{H}(p(t), q(t)) \end{aligned} \tag{4.1}$$

where $\mathcal{H} : \mathbb{R}^{2d} \to \mathbb{R}$ is the Hamiltonian function of the system, while $p(t), q(t) \in \mathbb{R}^d$ respectively denote *generalized momenta* and *generalized coordinates* associated to the mechanical system. In a more compact form, system (4.1) may be rewritten as

$$y'(t) = J\,\nabla\mathcal{H}(y(t)) \tag{4.2}$$

where $y(t) = (p(t), q(t))^\intercal \in \mathbb{R}^{2d}$ and

$$J = \begin{pmatrix} 0 & -I_d \\ I_d & 0 \end{pmatrix} = -J^\intercal = -J^{-1}$$

with I_d the identity matrix of dimension d. Moreover, $\nabla\mathcal{H}$ denotes the gradient of $\mathcal{H}$.

It is well known from the existing literature, $\mathcal{H}(t)$ is a first integral of the system, that is its value remains constant along the solution $y(t)$

$$\mathcal{H}(y(t)) = \mathcal{H}(y_0), \quad \text{for all } t \in [t_0, T].$$

The search for first integrals is interesting since their knowledge allows immediately to obtain qualitative information on the dynamics of the system. In the mechanical systems field, the Hamiltonian is interpreted as the total energy of the system and therefore, it is of interest to derive methods which are able to preserve this property in the discrete solution. Geometric numerical integration is able to perform an excellent long-time conservation of the Hamiltonian along the numerical solution. In particular, the researchers are directed towards employing symplectic Runge-Kutta methods [76, 144, 229], which are meant to exactly preserve quadratic invariants possessed by (4.2) along the numerical solution. Remember that a generic RK method (2.43) has the property of being symplectic if and only if the coefficients that characterize it satisfy the following identities

$$b_i a_{ij} + b_j a_{ji} = b_i b_j$$

for all $i, j = 1, \ldots, s$. Furthermore, symplectic RK methods are capable of preserving any Hamiltonian function over exponentially long times with an

exponentially decreasing error, as proved by Benettin and Giorgilli (see [144]). However, even though symplectic RK methods accurately preserve any Hamiltonian function, they have the limit of being implicit, and thus they require a remarkable computational effort in the integration process. Therefore many contributions in the literature are devoted to the construction and development of semi-implicit numerical methods with a lower computational cost that are able to near preserve these invariant laws [79, 80, 103, 114, 144, 145]. Other numerical methods able to preserve the total energy of the mechanical system are developed by Brugnano et. al. in [51, 52], called *Hamiltonian Boundary Value Methods* (HBVMs). Such type of methods are based on the discretization of a local Fourier expansion of the given ODE problem in which different choices of the basis lead to different classes of methods.

4.2 Stochastic Hamiltonian problems with additive noise

As described at the beginning of this chapter, our interest is devoted to stochastic Hamiltonian problems. For a fixed positive integer m, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with filtration $\{\mathcal{F}_t\}_{t\in[t_0,T]}$ and let $W : [t_0, T] \times \Omega \to \mathbb{R}^m$ a standard $(\mathcal{F}_\sqcup)_{t\in[t_0,T]}$–Wiener process with continuous sample paths on $(\Omega, \mathcal{F}, \mathbb{P})$. The stochastic Hamiltonian systems of Itô type with additive noise have the form

$$\begin{cases} dq(t) = \nabla_p \mathcal{H}(p(t), q(t))\, dt, \\ \\ dp(t) = -\nabla_q \mathcal{H}(p(t), q(t))\, dt + \Sigma\, dW(t) \end{cases} \tag{4.3}$$

for $t \in [t_0, T]$, where $p(t), q(t) \in \mathbb{R}^m$ and $\Sigma \in \mathbb{R}^{m\times m}$ is a diagonal matrix, whose generic diagonal element is denoted by ϵ_i, for $i = 1, \dots, m$. Moreover, let us assume that the initial values (p_0, q_0) are such that $\mathbb{E}(\mathcal{H}(p_0, q_0))$ is finite. One can immediately observe that if the matrix Σ is the zero matrix, the stochastic Hamiltonian system (4.3) recasts the deterministic Hamiltonian system (4.1).

It is well-known that the existing literature (see, for instance, [68, 74, 84] and references therein) has revealed that neither the Hamiltonian function $\mathcal{H}(p(t), q(t))$ nor its expectation are preserved along the dynamics described by Equation (4.3). Through the use of Itô's Lemma (1.14), a stochastic differential equation can be rewritten in such a way that the Hamiltonian function $\mathcal{H}$ is a solution. For this reason, given the $2m$– Itô SDE driven by m-dimensional Wiener process,

$$dX(t) = f(t, X(t))dt + G(t, X(t))dW(t), \quad X(t_0) = X_0 \tag{4.4}$$

where $f : [t_0, T] \times \mathbb{R}^{2m} \to \mathbb{R}^{2m}$ and $G(t, X(t)) = [g^1(t, X(t), \dots, g^m(t, X(t))]^\intercal$ with $g^i : [t_0, T] \times \mathbb{R}^{2m} \to \mathbb{R}^{2m}$ for $i = 1, \dots, m$. Then the SDE for $U = \mathcal{H}(t, X(t))$ is given by

$$dU = \left(\frac{\partial U}{\partial t} + \sum_{i=1}^{2m} f_i \frac{\partial U}{\partial X_i} + \frac{1}{2}\mathrm{trace}(GG^\intercal \nabla^2 U) \right) dt + \sum_{i=1}^{m} g^i \frac{\partial U}{\partial X_{m+i}} dW_i.$$

Therefore from (4.3) by putting $U = \mathcal{H}(q(t), p(t))$ and

- $f = (\nabla_p \mathcal{H}, -\nabla_q \mathcal{H})^\intercal$
- $\nabla U = (\nabla_q \mathcal{H}, \nabla_p \mathcal{H})^\intercal$
- $\nabla^U = \begin{pmatrix} \nabla_{qq}\mathcal{H} & \nabla_{qp}\mathcal{H} \\ \nabla_{pq}\mathcal{H} & \nabla_{pp}\mathcal{H} \end{pmatrix}$
- $g^i = \epsilon_i (0, e_i)^\intercal$, where 0 is the zero vector of dimension m and e_i is the i-th basis vector of $\mathbb{R}^m$

one has

$$d\mathcal{H} = \left(\frac{1}{2} \sum_{i=1}^{m} \epsilon_i^2 \nabla_{pp}^{ii} \mathcal{H} \right) dt + \sum_{i=1}^{m} \epsilon_i \nabla_p^i \mathcal{H} dW_i$$

where $\nabla_p^i \mathcal{H}$ is the i-th component of $\nabla_p \mathcal{H}$ and ∇_{pp}^{ii} is the element in position (i, i) of the Hessian matrix associated to the function $\mathcal{H}$ with respect to p.

Integrating from t_0 to t, gives

$$\mathcal{H}(p(t), q(t)) = \mathcal{H}(p(t_0), q(t_0)) + \frac{1}{2} \sum_{i=1}^{m} \varepsilon_i^2 \int_{t_0}^{t} \nabla_{pp}^{ii} \mathcal{H} \, ds + \sum_{i=1}^{m} \varepsilon_i \int_{t_0}^{t} \nabla_p^i \mathcal{H} \, dW_i(s).$$

The expected value of Hamiltonian is then given by

$$\mathbb{E}\left[\mathcal{H}(p(t), q(t))\right] = \mathbb{E}\left[\mathcal{H}(p(t_0), q(t_0))\right] + \frac{1}{2} \sum_{i=1}^{m} \varepsilon_i^2 \int_{t_0}^{t} \mathbb{E}(\nabla_{pp}^{ii} H) \, ds, \tag{4.5}$$

confirming that the expected value of the Hamiltonian is not preserved for Itô stochastic Hamiltonian problems. In particular, if the Hamiltonian is of the form

$$\mathcal{H}(p(t), q(t)) = \frac{1}{2} \sum_{i=1}^{m} p_i(t)^2 + V(q(t)),$$

i.e. the sum of the kinetic and potential energy of the system, for a suitable smooth potential $V : \mathbb{R}^m \to \mathbb{R}$, the corresponding Hamiltonian system of Itô type reads

$$\begin{cases} dq(t) = p(t) \, dt, \\ \\ dp(t) = -\nabla_q V(q(t)) \, dt + \Sigma \, dW(t). \end{cases} \tag{4.6}$$

Correspondingly, its expectation is then given by

$$\mathbb{E}\left[\mathcal{H}(p(t), q(t))\right] = \mathbb{E}\left[\mathcal{H}(p(t_0), q(t_0))\right] + \frac{1}{2} \sum_{i=1}^{m} \varepsilon_i^2 (t - t_0). \tag{4.7}$$

Therefore, the Hamiltonian function grows linearly in time.

In the existing literature there are several contributions concerning the numerics related to stochastic Hamiltonian problems [14, 74, 68, 84, 121, 202]. They are aimed at determining numerical methods capable of maintaining

equation (4.7) along the discretized dynamics, i.e. methods able to satisfy the following equation

$$\mathbb{E}[\mathcal{H}(q_n, p_n)] = \mathbb{E}[\mathcal{H}(q_0, p_0)] + \frac{1}{2}\sum_{i=1}^{m} \epsilon_i^2 (t_n - t_0)$$

where $t_n = n\Delta t$ for $n = 1, \ldots, N$, $\Delta t = T/N$ and N positive integer. Here q_n and p_n denote the approximation of $q(t_n)$ and $p(t_n)$ respectively.

A natural question that arises is the following:: is it straightforward to accurately retain qualitative features of stochastic Hamiltonian problems under time discretizations? Therefore our attention is paid to investigate a long-term behaviour of numerical methods analyzed in [68, 74] for stochastic Hamiltonian problems (4.6) and in particular on the understanding the validity of (4.7) for large time windows and for large values of the entries of Σ, in line with the well-knows Benettin-Giorgilli theorem, explaining the long-term behaviour of symplectic Runge-Kutta schemes for the deterministic Hamiltonian systems. For this reason, in the next subsection, Runge-Kutta methods proposed by Burrage and Burrage in [68] will be analyzed, showing that the latter are able to maintain linear growth in the expected value of the Hamiltonian function (4.7).

4.2.1 A new version of stochastic Runge-Kutta methods

Let us consider a general problem with additive noise, i.e.,

$$dy(t) = f(y)\, dt + \epsilon\, r\, dW, \tag{4.8}$$

where $y, r \in \mathbb{R}^{2m}$, ε is a positive scalar value and $W(t)$ is a scalar Wiener process. The new formulation of stochastic Runge-Kutta methods is based on the idea of Brugnano et. al. [53, 54]. Then, the integral formulation of (4.8) reads

$$y(t_{n+1}) = y(t_n) + \int_{t_n}^{t_{n+1}} f(y(u))\, du + \epsilon\, r \int_{t_n}^{t_{n+1}} dW(u). \tag{4.9}$$

In particular for the intermediate points $t_n + c_i\Delta t$, equation (4.9) becomes

$$y(t_n + c_i\Delta t) = y(t_n) + \int_{t_n}^{t_n + c_i\Delta t} f(y(u))\, du + \epsilon\, r \int_{t_n}^{t_n + c_i\Delta t} dW(u). \tag{4.10}$$

Denoting Y_i the approximation to (4.10), then a different formulation of stochastic Runge-Kutta methods, proposed by Burrage et. al. in [68], is

$$\begin{aligned} Y_i &= y_n + \Delta t \sum_{j=1}^{s} a_{ij} f(Y_j) + \epsilon\, (W(t_n + c_i\, \Delta t) - W(t_n))\; r, \quad i = 1, \ldots, s \\ y_{n+1} &= y_n + \Delta t \sum_{j=1}^{s} b_j f(Y_j) + \epsilon\, (W(t_n + \Delta t) - W(t_n))\; r \end{aligned} \tag{4.11}$$

Furthermore, it must be taken into account that the sampling $W(t_n + c_i\,\Delta t) - W(t_n)$ can be written as a sum of Wiener increments of the type

$$\sum_{j=1}^{i} \left(W(t_n + c_j\,\Delta t) - W(t_n + c_{j-1}\,\Delta t)\right).$$

Therefore, by putting

$$z_i = W(t_n + c_i\,\Delta t) - W(t_n + c_{i-1}\,\Delta t) = \sqrt{c_i - c_{i-1}}\,\Delta W_i \quad \text{for } i = 1, \ldots, s+1,$$

where $c_0 = 0, c_{s+1} = 1$ and ΔW_i are $s+1$ independent samples of normally distributed variables $\mathcal{N}(0, \Delta t)$, then the stochastic Runge-Kutta methods (4.11) become

$$\begin{aligned} Y_i &= y_n + h \sum_{j=1}^{s} a_{ij}\, f(Y_j) + \varepsilon \sum_{j=1}^{i} z_j\, r, \\ y_{n+1} &= y_n + h \sum_{j=1}^{s} b_j\, f(Y_j) + \varepsilon \sum_{j=1}^{s+1} z_j\, r. \end{aligned} \tag{4.12}$$

As proved by Gard in his monograph [131], the strong convergence of (4.12) is equivalent to the convergence of the underlying Runge-Kutta method with coefficients (A, b, c), i.e., it occurs when $\sum_{i=1}^{s} b_i = 1$. The analysis of SRK arising from the perturbation of deterministic Runge-Kutta methods is also described in [88], where conditions to inherit A-stability are provided. Other issues of stochastic Runge-Kutta methods are analyzed, for instance, in [116, 147, 153, 221] and references therein. Analogous issues for the time integration of stochastic problems with memory are discussed in [94, 115].

Let us consider a linear version of (4.8) as test equation, i.e.,

$$dy(t) = Qy\,dt + \varepsilon\, r\, dW \tag{4.13}$$

where

$$Q = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \qquad r = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

In other terms, the following Hamiltonian problem is considered

$$\begin{aligned} dq(t) &= p(t)\,dt \\ dp(t) &= -q(t)\,dt + \epsilon\, dW(t) \end{aligned} \tag{4.14}$$

as test equation, with Hamiltonian

$$\mathcal{H}(p(t), q(t)) = \frac{1}{2}(p^2 + q^2). \tag{4.15}$$

The performance of the following methods, when applied to the aforementioned test problem, are now evaluated:

- IM, the implicit midpoint rule

$$\begin{array}{c|c} 1/2 & 1/2 \\ \hline & 1 \end{array} \tag{4.16}$$

- the 3-stage Lobatto IIIA (LIIIA)

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ 1/2 & 5/24 & 1/3 & -1/24 \\ 1 & 1/6 & 2/3 & 1/6 \\ \hline & 1/6 & 2/3 & 1/6 \end{array} \tag{4.17}$$

- MQ, able to exactly preserve quartic Hamiltonian [74]

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ 1/2 & 1/12 & 1/3 & 1/12 \\ 1 & 1/6 & 2/3 & 1/6 \\ \hline & 1/6 & 2/3 & 1/6 \end{array} \tag{4.18}$$

Burrage and Burrage in [68] have shown that all three of these above methods (4.16)-(4.18), applied to the test linear problem (4.14), preserve the linear growth of the expected value of the Hamiltonian function (4.7). In particular, to confirm the validity of this result, it is possible to prove the following theorem.

Theorem 4.2.1. *The implicit midpoint rule* (4.16) *preserves the expectation of the Hamiltonian* (4.15) *exactly, that is it preserves* (4.7).

Proof. Consider the IM method in formulation (4.12)

$$\begin{aligned} Y &= y_n + \frac{\Delta t}{2} f(Y) + \frac{1}{\sqrt{2}} \Delta W_1 \, \epsilon \, r \\ y_{n+1} &= y_n + \Delta t \, f(Y) + \frac{1}{\sqrt{2}} \left(\Delta W_1 + \Delta W_2 \right) \epsilon \, r \end{aligned} \tag{4.19}$$

By applying (4.19) to the linear test problem (4.13), the IM method can be rewritten as

$$y_{n+1} = R(Q\,\Delta t) y_n + \epsilon \, S(Q\,\Delta t)\, r$$

where

$$\begin{aligned} R(Q\,\Delta t) &= \left(I - \frac{\Delta t}{2} Q \right)^{-1} \left(I + \frac{\Delta t}{2} Q \right), \\ S(Q\,\Delta t) &= z_2 I + R(Q\,\Delta t) z_1. \end{aligned}$$

with I 2-dimensional identity matrix. By using the expectation properties, one has

$$\mathbb{E}[y_{n+1}] = R(Q\,\Delta t)\mathbb{E}[y_n]$$

and furthermore

$$\mathbb{E}[y_{n+1}^{\intercal} y_{n+1}] = \mathbb{E}[y_n^{\intercal} \, R^{\intercal}(Q\,\Delta t) R(Q\,\Delta t)\, y_n] + \epsilon^2 r^{\intercal} \, \mathbb{E}[S^{\intercal}(Q\,\Delta t) S(Q\,\Delta t)]\, r.$$

Since Q is a skew-symmetric matrix, then the implicit midpoint rule has the following trivial properties

- $R^{\intercal}(Q\,\Delta t)R(Q\,\Delta t) = I$,
- $\mathbb{E}[S^{\intercal}(Q\,\Delta t)S(Q\,\Delta t)] = \mathbb{E}[z_1^2 + z_2^2]\,I = I\Delta t$.

Thus

$$\mathbb{E}[\mathcal{H}_n] = \mathbb{E}[y_{n+1}^{\intercal} y_{n+1}] = \mathbb{E}[\mathcal{H}_n] + \frac{\epsilon^2}{2}\Delta t.$$

This concludes the proof. □

The theorem thus obtained can be proved trivially even if the LIIIA (4.17) and MQ (4.18) methods are used.

Let us show the numerical evidence arising from the application of the three aforementioned methods to the following stochastic Hamiltonian problems:

- the double-well potential, with a Hamiltonian function

$$H(p(t), q(t)) = \frac{1}{2}p^2 + \frac{1}{4}q^4 - \frac{1}{2}q^2, \qquad t \in [0, 40] \tag{4.20}$$

  and initial condition $[q_0, p_0]^{\intercal} = \left[\sqrt{2}, \sqrt{2}\right]_{\intercal}$. The expected Hamiltonian in the endpoint of the integration interval is given by

$$\mathbb{E}\left[H(p(40), q(40)\right] = 1 + 20\varepsilon^2;$$

- the Hénon-Heiles problem, with a Hamiltonian function

$$H(p(t), q(t)) = \frac{1}{2}(p_1^2 + p_2^2) + \frac{1}{2}(q_1^2 + q_2^2) + \frac{1}{16}\left(q_1 q_2^2 - \frac{1}{3}q_1^3\right), \qquad t \in [0, 40] \tag{4.21}$$

  and initial condition $[q_0, p_0]^{\intercal} = \left[\sqrt{3}, 1, 1, 1\right]^{\intercal}$. This problem is characterized by the presence of two sources of noise, both with amplitude ϵ. Hence, in this case, the expected Hamiltonian in the endpoint of the integration interval is

$$\mathbb{E}\left[H(p(40), q(40)\right] = 3 + 40\epsilon^2,$$

  according to (4.7).

4.2.2 Numerical tests and ϵ-expansion

The experiments are carried out for fixed values of the stepsize Δt and for several values of the diffusive parameter ϵ. The results, displayed in Table 4.1, show the increase of the error on the expected Hamiltonian for increasing values of ϵ, which is not mitigated if the stepsize is reduced. Actually, this behaviour is more general: indeed, the linear drift in the expected Hamiltonian is not well preserved by any discretization. In fact, as one can see in Table 4.1 that there is a strong correlation between the value of ϵ and the error of the expected value of Hamiltonian. In particular, for bigger values of ϵ, like 0.2, the error does not decrease by reducing the value of the step size. While this behaviour can partly

ε	Δt	MQ	LIIIA	IM
0	0.05	1.5(-15)	7.9(-7)	4.3(-4)
0	0.2	1.7(-15)	2.1(-4)	6.1(-3)
0.001	0.1	1.2(-5)	3.4(-6)	1.7(-3)
0.01	0.1	4.2(-5)	3.7(-4)	6.7(-4)
0.5	0.05	1.2(-2)	9.3(-2)	1.2(-3)
0.5	0.2	9.6(-2)	4.1(-2)	4.8(-2)

ε	Δt	MQ	LIIIA	IM
0	0.1	1.8(-15)	3.4(-8)	6.9(-5)
0.001	0.1	1.1(-4)	1.1(-4)	1.2(-4)
0.01	0.1	7.8(-4)	2.8(-5)	1.1(-4)
0.01	0.2	8.4(-4)	6.6(-4)	1.4(-3)
0.2	0.1	8.5(-3)	1.3(-2)	3.2(-3)
0.2	0.2	4.0(-3)	6.5(-3)	1.1(-2)

Table 4.1: Hamiltonian errors in the endpoint of the integration interval for the implicit midpoint rule (IM), the 3-stage Lobatto IIIA (LIIIA) and the MQ method, applied to the double well potential (4.20) (top) and to the Hénon-Heiles problem (4.21) (bottom).

be ascribed to the pseudo-random generator used in the implementations, now this issue in terms of perturbation arguments is investigated and widely used in the context of deterministic differential problems [144].

Let us rewrite the linear equation (4.13) as a second-order SDE

$$\ddot{q} = -q + \varepsilon\sqrt{t}, \tag{4.22}$$

where ε is a normal distributed random variable.

Let us assume the following ansatz for the solution $q(t)$ to (4.22)

$$q(t) = \sum_{i=0}^{\infty} q_i(t)\varepsilon^i$$

and replace it into (4.22), leading to

$$\ddot{q_0}(t) + \varepsilon\,\ddot{q_1}(t) + \cdots = -q_0(t) - \varepsilon\,q_1(t) + \cdots + \varepsilon\sqrt{t}.$$

Collecting in terms of powers of ε yields

$$(\ddot{q_0}(t) + q_0(t)) + \varepsilon\left(\ddot{q_1}(t) + q_1(t) - \sqrt{t}\right) + \cdots = 0.$$

Let us focus on the coefficients of ε^0 and ε^1, satisfying the differential equations

$$\ddot{q_0} + q_0 = 0,$$
$$\ddot{q_1} + q_1 = \sqrt{t},$$

respectively. The equation for $q_0(t)$ is equipped by the initial values $q_0(t_0) = \dot{q}_0(t_0) = 1$. Then, its exact solution is easily computable and given by

$$q_0(t) = \cos(t) + \sin(t).$$

Correspondingly, assuming the ansatz

$$p(t) = \sum_{i=0}^{\infty} q_i(t)\varepsilon^i$$

for $p(t) = \dot{q}(t)$, one has $p_0(t) = \cos(t) - \sin(t)$.

Let us now consider the equation for $q_1(t)$, equipped by the initial values $q_1(t_0) = \dot{q}_1(t_0) = 0$. Then, its exact solution can be computed by variation of constants, leading to

$$\begin{aligned} q_1(t) &= -\sqrt{\frac{\pi}{2}}\, C\left(\sqrt{\frac{2}{\pi}}\sqrt{t}\right)\cos(t) - \sqrt{\frac{\pi}{2}}\, S\left(\sqrt{\frac{2}{\pi}}\sqrt{t}\right)\sin(t) + \sqrt{t}, \\ p_1(t) &= \sqrt{\frac{\pi}{2}}\left(C\left(\sqrt{\frac{2}{\pi}}\sqrt{t}\right)\sin(t) - S\left(\sqrt{\frac{2}{\pi}}\sqrt{t}\right)\cos(t)\right), \end{aligned}$$

where S and C are the Fresnel integrals

$$\begin{aligned} S(x) &= \int_0^x \sin\left(\frac{\pi}{2}\, t^2\right) dt, \\ C(x) &= \int_0^x \cos\left(\frac{\pi}{2}\, t^2\right) dt. \end{aligned}$$

Hence, the solution of (4.13) can be represented as

$$\begin{aligned} q(t) &= \cos t + \sin t - \varepsilon\left(\sqrt{\frac{\pi}{2}}\, C\left(\sqrt{\frac{2}{\pi}}\sqrt{t}\right)\cos(t) + \sqrt{\frac{\pi}{2}}\, S\left(\sqrt{\frac{2}{\pi}}\sqrt{t}\right)\sin(t) - \sqrt{t}\right) + \dots \\ p(t) &= \cos t - \sin t + \varepsilon\sqrt{\frac{\pi}{2}}\left(C\left(\sqrt{\frac{2}{\pi}}\sqrt{t}\right)\sin(t) - S\left(\sqrt{\frac{2}{\pi}}\sqrt{t}\right)\cos(t)\right) + \dots \end{aligned}$$

It is worth observing the presence of the secular term $\sqrt{t}$ in the solution of $q(t)$. This term is unbounded as t grows, making an accurate long-time integration unfeasible. This secular term cannot be removed as it is often done in perturbative methods of ODEs with various techniques like that of Poincaré-Lindstedt.

Furthermore, when the value ε is small, i.e. $\varepsilon \ll 1$, the presence of the secular term is mitigated over reasonable short time intervals. When ε is large, the secular term is dominating. This result is also graphically represented in Figure 4.1, where a comparison between the computed expected Hamiltonian vs its exact counterpart is shown for several values of ε.

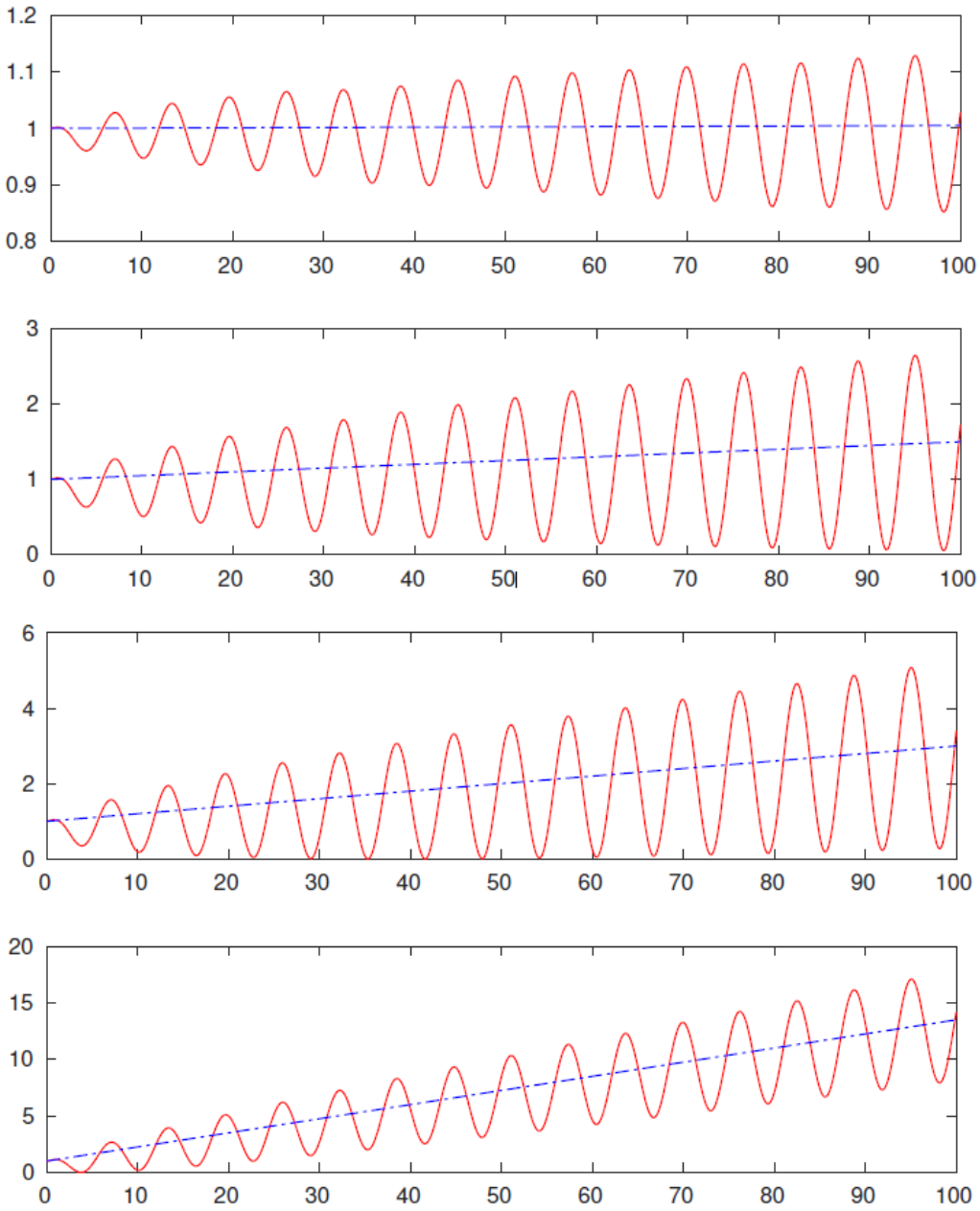


Figure 4.1: Comparison between the computed expected Hamiltonian (red) vs its exact counterpart (blue), for $\varepsilon = 0.01, 0.1, 0.2, 0.5$

4.3 Stochastic Hamiltonian problems with multiplicative noise

In this subsection, our analysis aims to come out as an introductory attempt to introduce multiplicative (small) noise in Hamiltonian systems (4.1), that at least at the best of our knowledge and under a numerical perspective, nowadays still represents an undiscovered field of research. As one may observe from the previous subsection, the numerical investigation of Hamiltonian systems driven by additive noise, indeed, constitutes a full-analyzed topic in the scientific literature, as visible, for example, in [49, 68, 74, 91, 105, 108, 120, 121, 201, 202, 203] and reference therein. It turned out, indeed, that the Hamiltonian is not conserved (neither in expectation) along the exact flow of such systems. Here, a characterization of the averaged Hamiltonian will be provided when the stochastic system is obtained by the employ of a multiplicative noise, that is, the following system will be considered

$$\begin{aligned} \mathrm{d}q(t) =& \nabla_p \mathcal{H}(p(t), q(t))\,\mathrm{d}t, \\ \mathrm{d}p(t) =& -\nabla_q \mathcal{H}(p(t), q(t))\,\mathrm{d}t + q(t)\Sigma^{\intercal}\mathrm{d}W(t), \end{aligned} \tag{4.23}$$

for $t \in [t_0, T]$, where $p(t), q(t) \in \mathbb{R}^m$, $\Sigma \in \mathbb{R}^m$ and $W \in \mathbb{R}^m$ is a vector of m independent Wiener processes. As visible in (4.23), only multiplicative noises are considered, that are linear in $q(t)$.

In the case of multiplicative noises, the following result generalizes the so-called *trace law* provided in (4.7).

Theorem 4.3.1. *For a given Itô Hamiltonian problem (4.23), the expected Hamiltonian satisfies*

$$\begin{aligned} \mathbb{E}\left[\mathcal{H}(q(t), p(t))\right] =& \mathbb{E}\left[\mathcal{H}(q(0), p(0))\right] \\ &+ \frac{1}{2}\mathrm{trace}\left(\Sigma^{\intercal}\Sigma\right) \sum_{i,k=1}^{m} \int_0^t \mathbb{E}\left[q_i(s) q_k(s) \nabla_{pp}^{ki} \mathcal{H}(q(s), p(s))\right] \mathrm{d}s, \end{aligned} \tag{4.24}$$

for $t \geqslant 0$.

Proof. One immediately observes that the proof straightforwardly descends from Itô lemma (1.13) applied to the function $\mathcal{H}(q(t), p(t))$ in (4.23). For this purpose, it is convenient to rewrite (4.23) in a more compact form, i.e., by considering

$$\mathrm{d}y(t) = f(y(t))\mathrm{d}t + g(y(t))\mathrm{d}W(t), \tag{4.25}$$

where $y(t) = [q(t), p(t)]^{\intercal} \in \mathbb{R}^{2m}$, $f(y(t)) = [\nabla_p \mathcal{H}(y(t)), -\nabla_q \mathcal{H}(y(t))]^{\intercal}$ and $G(y(t)) = [O_{m\times m} \;\; q\Sigma_{\intercal}]^{\intercal} \in \mathbb{R}^{2m\times m}$. Indeed, one has

$$\mathrm{d}\mathcal{H}(q(t), p(t)) = \mathcal{L}_0 \mathcal{H}(q(t), p(t))\mathrm{d}t + \mathcal{L}_1 \mathcal{H}(q(t), p(t))\mathrm{d}W(t), \tag{4.26}$$

for $t \geqslant 0$, where

$$\mathcal{L}_0 \mathcal{H}(y(t)) = f(y(t))^{\intercal} \nabla_y \mathcal{H}(y(t)) + \frac{1}{2}\mathrm{trace}\left(g(y(t)) g(y(t))^{\intercal} \nabla_{yy} \mathcal{H}(y(t))\right) \tag{4.27}$$

and

$$\mathcal{L}_1\mathcal{H}(q(t),p(t)) = \nabla_y\mathcal{H}(y(t))^\intercal G(y(t)). \tag{4.28}$$

By (4.23), the first term in the right-hand side of (4.27) vanishes, Hence, one gets

$$\begin{aligned}\mathcal{L}_0\mathcal{H}(q(t),p(t)) &= \frac{1}{2}\text{trace}\left(q(t)\Sigma^\intercal\Sigma q(t)^\intercal\nabla_{pp}\mathcal{H}(q(t),p(t))\right)\\ &= \frac{1}{2}\text{trace}\left(\Sigma^\intercal\Sigma\right)\sum_{i,k=1}^{m} q_i(t)q_k(t)\nabla_{pp}^{ki}\mathcal{H}(q(t),p(t))\end{aligned}$$

and

$$\mathcal{L}_1\mathcal{H}(q(t),p(t)) = \nabla_p\mathcal{H}(q(t),p(t))^\intercal q(t)\cdot\Sigma^\intercal.$$

Then, by passing to the integral representation of (4.26), one gets

$$\begin{aligned}\mathcal{H}(q(t),p(t)) &= \mathcal{H}(q(0),p(0))\\ &\quad+\frac{1}{2}\text{Tr}\left(\Sigma^\intercal\Sigma\right)\sum_{i,k=1}^{m}\int_0^t q_i(s)q_k(s)\nabla_{pp}^{ki}\mathcal{H}(q(s),p(s))\,\mathrm{d}s+\mathcal{I}_1(t),\end{aligned} \tag{4.29}$$

where $\mathcal{I}_1(t)$ here stands for a martingale term with zero initial expectation. Passing to side-by-side expectation concludes the proof. □

It is straightforward to verify that, in the case of additive noise, Equation (4.24) recasts the usual trace law (4.7).

4.3.1 The case of quadratic Hamiltonian

Let consider a separable and quadratic Hamiltonian function $H(q,p)$, i.e.,

$$\mathcal{H}(q,p) = \frac{1}{2}(q^2+p^2).$$

In this case, it is possible to gain some insights on Equation (4.29). For this purpose, as a test problem the linear system arising from (4.23) with quadratic potential will be considered, leading to

$$\mathrm{d}y(t) = Jy(t)\mathrm{d}t+\sigma\widetilde{J}y(t)\mathrm{d}W(t), \tag{4.30}$$

where

$$J=\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix},\qquad \widetilde{J}=\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}.$$

Equation (4.24) can be recasted as follows

$$\mathbb{E}\left[\mathcal{H}(q(t),p(t))\right] = \mathbb{E}\left[\mathcal{H}(q_0,p_0)\right]+\frac{1}{2}\sigma^2\int_0^t\mathbb{E}\left[q(s)^2\right]\,\mathrm{d}s \tag{4.31}$$

and the computation of the integral appearing on the right-hand side can be properly computed by usual SDEs arguments [165], as briefly reported in the remainder. By denoting

$$P(t) = \mathbb{E}\left[y(t)y(t)^{\intercal}\right]$$

the correlation matrix associated with the system (4.30), then $p(t)$ is the solution of the following system of ODEs

$$\frac{\mathrm{d}P}{\mathrm{d}t} = JP(t) - P(t)J + \sigma^2 \widetilde{J}P(t)\widetilde{J}^{\intercal}.$$

Moreover, by setting

$$\hat{p}(t) = \left[\mathbb{E}\left[q(t)^2\right] \;\; \mathbb{E}\left[q(t)p(t)\right] \;\; \mathbb{E}\left[p(t)^2\right]\right]^{\intercal},$$

then, it turns out that $\hat{p}(t)$ solves the linear system of ODEs

$$\frac{\mathrm{d}\hat{p}}{\mathrm{d}t} = S\hat{p}(t), \tag{4.32}$$

where

$$S = \begin{bmatrix} 0 & 2 & 0 \\ -1 & 0 & 1 \\ \sigma^2 & -2 & 0 \end{bmatrix},$$

whose characteristic polynomial is given by

$$\mu(\lambda) = \lambda^3 + 4\lambda - 2\sigma^2.$$

By Cartesio's rule and since $\mathrm{Tr}\,(S) = 0$, it turns out that the above polynomial admits one real positive root and a pair of complex conjugate roots with negative real parts. Let λ_1 be the real eigenvalue of S and λ_2 and $\overline{\lambda}_2$ the two complex conjugate eigenvalues. As a consequence, solving (4.32) leads to the following expression for the second moment of $q(t)$

$$\mathbb{E}\left[q(t)^2\right] = \alpha e^{\lambda_1 t} + e^{\mathrm{Re}(\lambda_2)t}\left[\beta\cos\left(\mathrm{Im}(\lambda_2)t\right) + \gamma\sin\left(\mathrm{Im}(\lambda_2)t\right)\right], \tag{4.33}$$

where $\lambda_1 > 0, \mathrm{Re}(\lambda_2) < 0$ and α, β, γ are constants depending on the initial conditions. Then, one has

$$\begin{aligned}\left|\mathbb{E}\left[H(q(t),p(t))\right] - \mathbb{E}\left[H(q(0),p(0))\right]\right| &= \left|\frac{1}{2}\sigma^2\int_0^t \mathbb{E}\left[q(s)^2\right]\,\mathrm{d}s\right| \\ &\leqslant \frac{1}{2}\sigma^2\int_0^t |\alpha|\, e^{\lambda_1 s} + K\, e^{\mathrm{Re}(\lambda_2)s}\,\mathrm{d}s \\ &\leqslant \frac{1}{2}\sigma^2\left[\frac{\alpha}{\lambda_1}\left(e^{\lambda_1 t} - 1\right) + \frac{K}{|\mathrm{Re}(\lambda_2)|}\right],\end{aligned} \tag{4.34}$$

where $K = \max\{|\beta|, |\gamma|\}$. Hence, it turns out that for sufficiently large $t \geqslant 0$, the averaged deviation of the Hamiltonian function blows up exponentially.

On the other side, one can conclude that, for reasonable values of $\lambda_1(\sigma)t$ such that $e^{\lambda_1 t}$ is well approximated by $1+\lambda_1(\sigma)t$, one finally gets

$$\left|\mathbb{E}\left[H(q(t),p(t))\right]-\mathbb{E}\left[H(q(0),p(0))\right]\right| \leqslant \frac{1}{2}\sigma^2\left[\lambda_1(\sigma)t+\frac{K}{|\mathrm{Re}(\lambda_2)|}\right]. \tag{4.35}$$

Symbolic computations have revealed that λ_1 is a smooth increasing function of σ and, hence, it is possible to have a suitable control on it by properly choosing small values of σ. Hence for suitably small noises, the averaged Hamiltonian deviation, in modulus, is bounded by a linear drift, up to a constant term depending on the initial conditions. Of course, it is worth recalling that, if σ is large enough and/or for large time windows, the quantity on the right-hand side of (4.35) grows exponentially in time.

4.3.2 Perturbative analysis and numerical tests

Following the lines drawn in [121], our attention is turned to analyzing the behaviour of the expected Hamiltonian along a general first-order approximation to system (4.23), in order to understand if it is always possible to retain (4.35) along suitable discretizations to (4.30). With this aim, let us assume as ansatz the possibility to represent $y(t)$ in power series of σ, i.e.,

$$y(t)=\sum_{i\geqslant 0} y_i(t)\sigma^i,$$

where $y_i(t)$, $i\geqslant 0$, are random processes to be determined. By substituting the above assumption in (4.30), one gets

$$\sum_{i\geqslant 0}\mathrm{d}y_i(t)\sigma^i = J\sum_{i\geqslant 0} y_i(t)\sigma^i\mathrm{d}t+\widetilde{J}\sum_{i\geqslant 0} y_i(t)\sigma^{i+1}\mathrm{d}W(t),$$

leading to

$$\begin{aligned}
&\mathrm{d}y_0(t)=Jy_0(t)\mathrm{d}t, \quad y_0(0)=y_0,\\
&\mathrm{d}y_i(t)=Jy_i(t)\mathrm{d}t+\widetilde{J}y_{i-1}(t)\mathrm{d}W(t), \quad y_i(0)=0, \quad i=1,2\ldots.
\end{aligned}$$

Then, by assuming that y_0 is a deterministic value, one obtains

$$\begin{aligned}
&y_0(t)=e^{Jt}y_0,\\
&y_i(t)=\int_0^t e^{J(t-s)}\widetilde{J}y_{i-1}(s)\,\mathrm{d}W(s), \quad i=1,2,\ldots
\end{aligned}$$

Hence, denoting a general first-order approximation to $y(t)$ as $\overline{y}(t)=y_0(t)+\sigma y_1(t)$, one gets

$$\overline{y}(t)=e^{Jt}y_0+\sigma\int_0^t e^{J(t-s)}\widetilde{J}y_0(s)\,\mathrm{d}W(s).$$

Then, by using the properties of Itô integrals [165] and by the orthogonality of the matrix J, one gets

$$\begin{aligned}
\mathbb{E}\left[\|\overline{y}(t)\|^2\right] &= \mathbb{E}\left[\left\|e^{Jt}y_0 + \sigma\int_0^t e^{J(t-s)}\widetilde{J}y_0(s)\mathrm{d}W(s)\right\|^2\right] \\
&= \mathbb{E}\left[\left\|e^{Jt}y_0\right\|^2\right] + \sigma^2\mathbb{E}\left[\left\|\int_0^t e^{J(t-s)}\widetilde{J}y_0(s)\,\mathrm{d}W(s)\right\|^2\right] \\
&+ 2\sigma\mathbb{E}\left[e^{Jt}y_0\int_0^t e^{J(t-s)}\widetilde{J}y_0(s)\mathrm{d}W(s)\right] \\
&= \mathbb{E}\left[\|y_0\|^2\right] + \sigma^2\int_0^t\left\|e^{J(t-s)}\widetilde{J}y_0(s)\right\|^2\mathrm{d}s \\
&= \mathbb{E}\left[\|y_0\|^2\right] + \sigma^2\int_0^t\left\|\widetilde{J}y_0(s)\right\|^2\mathrm{d}s \\
&= \mathbb{E}\left[\|y_0\|^2\right] + \sigma^2\int_0^t\left\|\widetilde{J}e^{Js}y_0\right\|^2\mathrm{d}s.
\end{aligned}$$

Since, for $v = [v_1\ v_2]^\intercal \in \mathbb{R}^2$, one has $\left\|\widetilde{J}v\right\|^2 = v_1^2$, one gets

$$\left\|\widetilde{J}e^{Js}y_0\right\|^2 = \left(e^{Js}y_0\right)_1^2 = \left(\cos(s)q_0 + \sin(s)p_0\right)^2.$$

Therefore, one obtains

$$\begin{aligned}
\mathbb{E}\left[\|\overline{y}(t)\|^2\right] &= \mathbb{E}\left[\|y_0\|^2\right] + \sigma^2\int_0^t\left(\cos(s)q_0 + \sin(s)p_0\right)^2\mathrm{d}s \\
&= \mathbb{E}\left[\|y_0\|^2\right] + \frac{\sigma^2}{2}\mathbb{E}\left[\|y_0\|^2\right]t + \frac{\sigma^2}{2}\left[\alpha_0\cos(t)\sin(t) + \beta_0\sin^2(t)\right],
\end{aligned}$$

where $\alpha_0 = q_0^2 - p_0^2$ and $\beta_0 = 2q_0p_0$. As a consequence,

$$\mathbb{E}\left[\mathcal{H}(\overline{y}(t))\right] = \mathbb{E}\left[\mathcal{H}(y_0)\right] + \frac{\sigma^2}{2}\mathbb{E}\left[\mathcal{H}(y_0)\right]t + \frac{\sigma^2}{4}\left[\alpha_0\cos(t)\sin(t) + \beta_0\sin^2(t)\right].$$

Finally, one can conclude that

$$\begin{aligned}
\left|\mathbb{E}\left[\mathcal{H}(\overline{y}(t))\right] - \mathbb{E}\left[\mathcal{H}(y_0)\right]\right| &= \left|\frac{\sigma^2}{2}\mathbb{E}\left[\mathcal{H}(y_0)\right]t + \frac{\sigma^2}{4}\left[\alpha_0\cos(t)\sin(t) + \beta_0\sin^2(t)\right]\right| \\
&\leqslant \frac{\sigma^2}{2}\left[\mathbb{E}\left[\mathcal{H}(y_0)\right]t + K_0\right],
\end{aligned} \tag{4.36}$$

where $K_0 = \frac{1}{2}\max\{\alpha_0, \beta_0\}$. The comparison between (4.35) and (4.36) reveals that it is not possible to obtain a good behaviour in terms of conservation of (4.36) through any general first-order approximation to (4.30); the accuracy is destroyed for large values of σ and over long time windows.

As a selected example of expected failure in preserving inequality (4.35) along the numerical dynamics generated by approximations to (4.30), now the

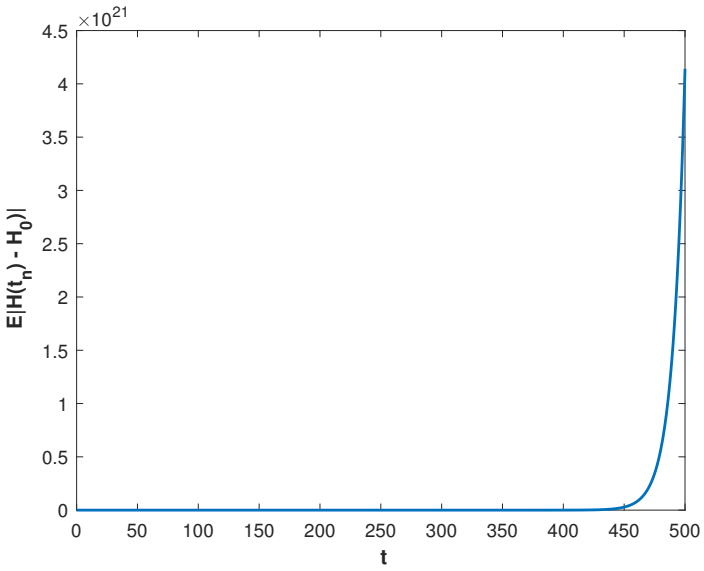


Figure 4.2: Averaged Hamiltonian deviation for Euler-Maruyama scheme (4.37) computed with $\Delta t = 0.1, \sigma = 0.01, M = 10^5$.

behaviour of the Euler-Maruyama scheme [131, 165] when applied to system (4.30) is studied. The method is as follows:

$$y_n = y_{n-1} + \Delta t J y_{n-1} + \sigma \widetilde{J} y_{n-1} \Delta W_n = \left(\mathcal{I} + \Delta t J + \sigma \widetilde{J} \Delta W_{n-1} \right) y_{n-1}. \quad (4.37)$$

Therefore, one has

$$\begin{aligned} \mathbb{E}\left[\|y_n\|^2 \right] &= \mathbb{E}\left[\left\| \left(\prod_{i=1}^{n} \mathcal{I} + \Delta t J + \sigma \widetilde{J} \Delta W_{n-i} \right) y_0 \right\|^2 \right] \\ &\leqslant \prod_{i=1}^{n} \mathbb{E}\left[\left(1 + \Delta t \|J\| + \sigma \Delta W_{n-1} \left\| \widetilde{J} \right\| \right)^2 \right] \|y_0\|^2 \\ &= \prod_{i=1}^{n} \left(1 + \Delta t^2 \|J\|^2 + \sigma^2 \Delta t \left\| \widetilde{J} \right\|^2 + 2\Delta t \|J\| \right) \|y_0\|^2 \\ &= \sum_{k=0}^{n} \binom{n}{k} \Delta t^k \left(\sigma^2 + 2 + \Delta t \right)^k \|y_0\|^2, \end{aligned}$$

that is,

$$\mathbb{E}\left[\mathcal{H}(y_n) \right] - \mathbb{E}\left[\mathcal{H}(y_0) \right] \leqslant \frac{1}{2} \sum_{k=1}^{n} \binom{n}{k} \Delta t^k \left(\sigma^2 + 2 + \Delta t \right)^k. \quad (4.38)$$

According to the inequality in (4.38), extra terms appear in the computation of the deviation of the averaged Hamiltonian destroying any eventual possibility of inheriting (4.35). A numerical evidence is depicted in Figure 4.2, where the growth in time of the averaged Hamiltonian deviation is well visible, according to (4.35).

Chapter 5

SPDEs: Splitting schemes for FitzHugh–Nagumo equation

5.1 Introduction

Ordinary differential equations and the corresponding stochastic version are good models when time is the only variable needed for the description of the studied phenomena. However, very often it happens that the latter is not able to describe the phenomenon under consideration and for this reason, it is necessary to resort to the use of partial differential equations (PDEs) and correspondingly stochastic partial differential equations (SPDEs), which also take into account the conditions of the spatial domain. As a matter of fact, the SPDEs can be seen as the stochastic version of the deterministic PDEs in some Hilbert or Banach space with coefficients (including the inhomogeneous terms) being random functions of space and time. In the literature, there are multiple models described by SPDEs. For example, the reader can find some examples in Pao-Liu Chow's monograph [85], ranging from the problem of turbulent transport to financial models. Clearly, determining the analytical solution of SPDEs is by no means an easy-to-solve problem. In fact, many researchers have gone so far as to determine advanced numerical methods capable of obtaining a good approximation of the solution of the problem, trying to minimize the computational cost. For example the interested reader can see [15, 20, 28, 48, 39, 42, 89, 90, 91, 107, 171, 175, 252] and the references therein.
In this thesis, our interest is focused on the study of the stochastic FitzHugh–Nagumo system and on numerical integrators, which treat the nonlinearity explicitly for the temporal discretization of the system, based on splitting strategies [44].

Splitting schemes have been used extensively in previous years. In particular, the recent work [61] analyses the strong convergence of splitting schemes for a class of semi-linear SDEs as well as preservation of possible structural properties of the problem. The work [246] performs extensive numerical simulations on the FitzHugh–Nagumo equation with space-time white noise in one dimension: a finite difference discretization is used in space, while the

classical Euler–Maruyama is used in time. The article [34] studies numerically the FitzHugh–Nagumo equation with coloured noise in two dimension. More precisely, the authors use a finite element discretization in space and the semi-implicit Euler–Maruyama scheme in time. The two previously mentioned works employ crude explicit discretization for the nonlinearity and therefore may have issues about moment bounds. The work [139] proves convergence (without rates) of a fully-discrete numerical scheme, based on a Galerkin method in space and the tamed Euler scheme in time, for a general SPDE with super-linearly growing operators. This is then applied to the FitzHugh–Nagumo equation with space-time white noise in one dimension. The articles [230] and [231] prove strong convergence rates of a finite difference spatial discretization of the FitzHugh–Nagumo equation with space-time white noise in one dimension.

The deterministic FitzHugh–Nagumo system is a simplified two-dimensional version of the famous Hodgkin–Huxley model that describes how action potentials propagate along an axon. Noise is omnipresent in neural systems and arises from different sources: it could be internal noise (such as random synaptic input from other neurons) or external noise, see for instance [181] for details. It was noted in [237] that the addition of an appropriate amount of noise in the model helps to detect weak signals. All of this has attracted a large body of works on the analysis of the influence of external random perturbations in neurons in the recent years, see for instance [179, 181, 184, 206, 214, 237, 241, 247, 257].

The stochastic FitzHugh–Nagumo system is given by

$$\begin{cases} \dfrac{\partial}{\partial t}u(t,\zeta) = \dfrac{\partial^2}{\partial \zeta^2}u(t,\zeta) + u(t,\zeta) - u^3(t,\zeta) - v(t,\zeta) + \dfrac{\partial^2}{\partial t \partial \zeta}W(t,\zeta), \\ \dfrac{\partial}{\partial t}v(t,\zeta) = \gamma_1 u(t,\zeta) - \gamma_2 v(t,\zeta) + \beta, \\ \dfrac{\partial}{\partial \zeta}u(t,0) = \dfrac{\partial}{\partial \zeta}u(t,1) = 0, \\ u(0,\zeta) = u_0(\zeta), v(0,\zeta) = v_0(\zeta), \end{cases} \tag{5.1}$$

for $\zeta \in (0,1)$ and $t \geqslant 0$. Here, the unknowns $u = \big(u(t)\big)_{t\geqslant 0}$ and $v = \big(v(t)\big)_{t\geqslant 0}$ are $L^2(0,1)$-valued stochastic processes, with initial values $u_0, v_0 \in L^2(0,1)$, see Section 5.2 and the standard monograph [102] on stochastic evolution equations in Hilbert spaces. In addition, $\gamma_1, \gamma_2, \beta \in \mathbb{R}$ are three real-valued parameters, $\Delta = \frac{\partial^2}{\partial \zeta^2}$ is the Laplace operator endowed with homogeneous Neumann boundary conditions, and $\big(W(t)\big)_{t\geqslant 0}$ is a cylindrical Wiener process, meaning that the component u is driven by space-time white noise.

In its physical sense, the component u represents the voltage variable while the component v is the recovery variable. The noise represents random fluctuations of the membrane potential, see [237] for a related model with a scalar noise. Note that in the considered system only the evolution of the voltage variable u is driven by a Wiener process. Having noise for the evolution of the recovery variable v would correspond to modelling different biological phenomena that are not treated in this work.

The major difficulty in the theoretical and numerical analysis of the SPDE system above is the nonlinearity $u - u^3$ appearing in the evolution of the component u: this nonlinearity is not globally Lipschitz continuous and has polynomial growth. As proved in [23], using a standard explicit discretization such as the Euler–Maruyama method (2.8) would yield numerical schemes which usually do not converge. More precisely, moment bounds, uniform with respect to the time step size, would not hold for such methods.

For an efficient numerical simulation of the above SPDE system, a splitting strategy is therefore used to define integrators and furthermore appropriate moment bounds and strong error estimates will be obtained. Splitting schemes have a long history in the numerical analysis of ordinary and partial differential equations, see for instance [32, 144, 178, 199] and references therein. Splitting integrators have recently been applied and analysed in the context of stochastic ordinary and partial differential equations. Without being exhaustive, the interested reader can refer to [9, 11, 33, 58, 84, 177, 205] for the finite-dimensional context and to [29, 41, 43, 45, 47, 99, 101, 113, 137, 182, 194, 195, 211] for the context of SPDEs.

In a nutshell, the key idea of the splitting integrators is to decompose the vector field, appearing in the evolution equation, in several parts, in order to exhibit subsystems that can be integrated exactly (or easily). One then composes the (exact or approximate) flows associated with the subsystems to define integrators applied to the original problem.

The main result obtained in this Thesis is to demonstrate the strong convergence of the aforementioned method, with relative order of convergence 1/4, for easy to implement splitting integrators, see Equation (5.24) in Subsection 5.3.2, for the time discretization of the SPDE defined above, see Theorem 5.3.2 for a precise statement.

The first non-trivial step of the analysis is to obtain suitable moment bounds for the splitting scheme, see Theorem 5.3.1. Note that the proof of the moment bounds of Theorem 5.3.1 is inspired by the work [46] where splitting schemes for the stochastic Allen–Cahn equation

$$\frac{\partial u(t,\zeta)}{\partial t} = \frac{\partial^2 u(t,\zeta)}{\partial \zeta^2} + u(t,\zeta) - u^3(t,\zeta) + \frac{\partial^2}{\partial t \partial \zeta} W(t,\zeta)$$

were studied. The proof of the strong convergence error estimates of Theorem 5.3.2 is inspired by the article [45]. However, one needs a dedicated and detailed analysis since the considered stochastic FitzHugh–Nagumo system is not a parabolic stochastic evolution system, and several arguments are non trivial. Note also that the construction of the splitting scheme is inspired by the recent article [61] that treats a finite dimensional version

$$\begin{cases} \mathrm{d}u(t) = (u(t) - u^3(t) - v(t))\,\mathrm{d}t, \\ \mathrm{d}v(t) = (\gamma_1 u(t) - \gamma_2 v(t) + \beta)\,\mathrm{d}t + \mathrm{d}B(t), \\ u(0) = u_0, v(0) = v_0, \end{cases}$$

of the stochastic FitzHugh–Nagumo system (where the finite-dimensional noise B is in the v-component).

5.2 Background of the problem

This section is devoted to introducing the functional framework, the linear and nonlinear operators, and the Wiener process. This allows us to consider the stochastic FitzHugh–Nagumo SPDE system as a stochastic evolution equation in the classical framework of [102].

5.2.1 Functional framework

Let us first introduce the infinite-dimensional, separable Hilbert space $H = L^2(0,1)$ of square integrable functions from $(0,1)$ to $\mathbb{R}$, equipped with the inner product $\langle \cdot, \cdot \rangle_H$ and the norm $\| \cdot \|_H$ that satisfy

$$\langle u_1, u_2 \rangle_H = \int_0^1 u_1(\zeta) u_2(\zeta) \, \mathrm{d}\zeta, \quad \|u\|_H = \sqrt{\langle u, u \rangle_H},$$

respectively, for all $u_1, u_2, u \in H$. Let us then introduce the product space $\mathcal{H} = H \times H$, which is also an infinite-dimensional, separable Hilbert space, with the inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and the norm $\| \cdot \|_{\mathcal{H}}$ defined by

$$\langle x_1, x_2 \rangle_{\mathcal{H}} = \langle u_1, u_2 \rangle_H + \langle v_1, v_2 \rangle_H, \quad \|x\|_{\mathcal{H}} = \sqrt{\|u\|_H^2 + \|v\|_H^2},$$

for all $x_1 = (u_1, v_1), x_2 = (u_2, v_2), x = (u, v) \in \mathcal{H}$.

Let also $E = \mathcal{C}^0([0,1])$ be the space of continuous functions from $[0,1]$ to $\mathbb{R}$, and $\mathcal{E} = E \times E$ the corresponding product space. One immediately observes that E and $\mathcal{E}$ are both separable Banach spaces, with the following norms

$$\|u\|_E = \max_{\zeta \in [0,1]} |u(\zeta)|, \quad \|x\|_{\mathcal{E}} = \max\bigl(\|u\|_E, \|v\|_E\bigr)$$

for all $u \in E$ and $x = (u, v) \in \mathcal{E}$.

Let us denote the inner product and the norm in the finite-dimensional Euclidean space $\mathbb{R}^2$ by $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$, respectively. If M is a 2×2 real-valued matrix, let $|||M||| = \sup_{x \in \mathbb{R}^2; \|x\| = 1} \|Mx\|$.

5.2.2 Linear operators

This subsection presents the material required to use the semigroup approach for SPDEs, see for instance [102].

For all $j \in \mathbb{N}$, set $\lambda_j = (j\pi)^2$ and $e_j(\zeta) = \sqrt{2} \cos(j\pi\zeta)$ for all $\zeta \in [0,1]$. In addition, set $\lambda_0 = 0$ and $e_0(\zeta) = 1$ for all $\zeta \in [0,1]$. Then $\bigl(e_j\bigr)_{j \geqslant 0}$ is a complete orthonormal system of H, and one has

$$\Delta e_j = -\lambda_j e_j$$

for all $j \geqslant 0$, where Δ denotes the Laplace operator with homogeneous Neumann boundary conditions. For all $u \in H$ and all $t \geqslant 0$, set

$$e^{t\Delta} u = \sum_{j \geqslant 0} e^{-t\lambda_j} \langle u, e_j \rangle_H e_j. \tag{5.2}$$

Then, for any $u_0 \in H$, the mapping $(t,\zeta) \mapsto u(t,\zeta) = e^{t\Delta}u_0(\zeta)$ is the unique solution of the heat equation on $(0,1)$ with homogeneous Neumann boundary conditions and initial value $u(0,\cdot) = u_0$:

$$\begin{cases} \dfrac{\partial u(t,\zeta)}{\partial t} = \Delta u(t,\zeta), \quad t > 0, \zeta \in (0,1), \\ \dfrac{\partial u(t,0)}{\partial \zeta} = \dfrac{\partial u(t,1)}{\partial \zeta} = 0, \quad t > 0, \\ u(0,\zeta) = u_0(\zeta), \quad \zeta \in (0,1). \end{cases}$$

Moreover, for all $\alpha \in [0,2]$, set

$$H^\alpha = \left\{ u \in H; \sum_{j\geqslant 0} \lambda_j^\alpha \langle u, e_j\rangle_H^2 < \infty \right\},$$
$$(-\Delta)^{\frac{\alpha}{2}} u = \sum_{j\geqslant 0} \lambda_j^{\frac{\alpha}{2}} \langle u, e_j\rangle_H e_j, \quad u \in H^\alpha.$$

Observe that $H^0 = H = L^2(0,1)$. The Laplace operator Δ with homogeneous Neumann boundary conditions is a self-adjoint unbounded linear operator on H, with domain $D(\Delta) = H^2$. Also be $\mathcal{H}^\alpha = H^\alpha \times H$ for all $\alpha \in [0,2]$.

Let us now introduce the linear operator Λ, defined as follows: for all $x = (u,v) \in \mathcal{H}^2$, set

$$\Lambda x = \begin{pmatrix} -\Delta u \\ 0 \end{pmatrix}.$$

Then Λ is a self-adjoint unbounded linear operator on $\mathcal{H}$, with domain $D(\Lambda) = \mathcal{H}^2$. For all $x = (u,v) \in \mathcal{H}$ and $t \geqslant 0$, set

$$e^{-t\Lambda} x = \begin{pmatrix} e^{t\Delta}u \\ v \end{pmatrix}. \tag{5.3}$$

Regularity estimates for this operator are presented in Section 5.4 below.

5.2.3 Nonlinear operator

Let $\beta, \gamma_1, \gamma_2 \in \mathbb{R}$ be parameters of the model. Define the mapping $F : \mathbb{R}^2 \to \mathbb{R}^2$ such that for all $x = (u,v) \in \mathbb{R}^2$ one has

$$F(x) = \begin{pmatrix} u - u^3 - v \\ \gamma_1 u - \gamma_2 v + \beta \end{pmatrix}.$$

To define splitting schemes, it is convenient to introduce two auxiliary mappings $F^{\mathrm{NL}} : \mathbb{R}^2 \to \mathbb{R}^2$ and $F^{\mathrm{L}} : \mathbb{R}^2 \to \mathbb{R}^2$ defined as follows: for all $x = (u,v) \in \mathbb{R}^2$ one has

$$F^{\mathrm{NL}}(x) = \begin{pmatrix} u - u^3 \\ \beta \end{pmatrix}$$
$$F^{\mathrm{L}}(x) = \begin{pmatrix} -v \\ \gamma_1 u - \gamma_2 v \end{pmatrix} = Bx,$$

where the matrix B is defined by

$$B = \begin{pmatrix} 0 & -1 \\ \gamma_1 & -\gamma_2 \end{pmatrix}.$$

Therefore, one immediately has

$$F(x) = F^{\mathrm{NL}}(x) + F^{\mathrm{L}}(x) \tag{5.4}$$

for all $x \in \mathbb{R}^2$.

The mapping F^{L} is globally Lipschitz continuous. In fact, for all $x_1, x_2 \in \mathbb{R}^2$ then

$$\|F^{\mathrm{L}}(x_2) - F^{\mathrm{L}}(x_1)\| \leqslant |||B|||\|x_2 - x_1\|.$$

However F and F^{NL} are only locally Lipschitz continuous, and satisfy a one-sided Lipschitz continuity property, that is there exists a positive constant $C \in (0, \infty)$ such that for all $x_1, x_2 \in \mathbb{R}^2$ such that

$$\begin{aligned} &\langle x_2 - x_1, F(x_2) - F(x_1)\rangle \leqslant C\|x_2 - x_1\|^2, \\ &\langle x_2 - x_1, F^{\mathrm{NL}}(x_2) - F^{\mathrm{NL}}(x_1)\rangle \leqslant C\|x_2 - x_1\|^2. \end{aligned} \tag{5.5}$$

In the sequel, an abuse of notation is used for simplicity: the same notation is employed for a mapping $f : \mathbb{R}^2 \to \mathbb{R}^2$ and for the associated Nemytskii operator defined on $\mathcal{H}$ or on $\mathcal{E}$ by $f(u, v) = f(u(\cdot), v(\cdot))$.

5.2.4 Cylindrical Wiener process

It remains to define the noise that drives the stochastic FitzHugh–Nagumo system. Let $\big(W(t)\big)_{t\geqslant 0}$ be a cylindrical Wiener process on H: given a sequence $\big(\beta_j(\cdot)\big)_{j\geqslant 0}$ of independent standard real-valued Wiener processes, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $\big(\mathcal{F}_t\big)_{t\geqslant 0}$ that satisfies the usual conditions, set

$$W(t) = \sum_{j\geqslant 0} \beta_j(t) e_j. \tag{5.6}$$

where $\big(e_j\big)_{j\geqslant 0}$ is a complete orthonormal system of H, defined above. For all $t \geqslant 0$, define

$$\mathcal{W}(t) = \begin{pmatrix} W(t) \\ 0 \end{pmatrix} = \sum_{j\geqslant 0} \beta_j(t) \begin{pmatrix} e_j \\ 0 \end{pmatrix},$$

then $\big(\mathcal{W}(t)\big)_{t\geqslant 0}$ is a generalized $\mathcal{Q}$-Wiener process on $\mathcal{H}$, with the covariance operator

$$\mathcal{Q} = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}.$$

Note that almost surely $W(t) \notin H$ and $\mathcal{W}(t) \notin \mathcal{H}$ for all $t > 0$. However, for all $T \geqslant 0$, the Itô stochastic integrals $\int_0^T L(t)\,\mathrm{d}W(t)$ and $\int_0^T \mathcal{L}(t)\,\mathrm{d}\mathcal{W}(t)$ are well-defined H-valued and $\mathcal{H}$-valued random variables respectively, if $\big(L(t)\big)_{0\leqslant t\leqslant T}$

and $\big(\mathcal{L}(t)\big)_{0\leqslant t\leqslant T}$ are adapted processes that satisfy $\sum_{j\geqslant 0}\int_0^T \mathbb{E}[\|L(t)e_j\|_H^2]\,\mathrm{d}t < \infty$ and $\sum_{j\geqslant 0}\int_0^T \mathbb{E}[\|\mathcal{L}(t)\begin{pmatrix} e_j \\ 0\end{pmatrix}\|_{\mathcal{H}}^2]\,\mathrm{d}t < \infty$, respectively.

Observe that for all $T \geqslant 0$ one has

$$\sum_{j\geqslant 0}\int_0^T \|e^{t\Delta}e_j\|_H^2\,\mathrm{d}t = \sum_{j\geqslant 0}\int_0^T \|e^{-t\Lambda}\begin{pmatrix} e_j \\ 0\end{pmatrix}\|_{\mathcal{H}}^2\,\mathrm{d}t \leqslant T + \sum_{j\geqslant 1}\lambda_j^{-1} < \infty.$$

Therefore, for all $t \geqslant 0$ one can define the H-valued random variable $Z(t)$ and the $\mathcal{H}$-valued random variable $\mathcal{Z}(t)$, called the stochastic convolutions, by

$$\begin{aligned} Z(t) &= \int_0^t e^{(t-s)\Delta}\,\mathrm{d}W(s), \\ \mathcal{Z}(t) &= \int_0^t e^{-(t-s)\Lambda}\,\mathrm{d}\mathcal{W}(s). \end{aligned} \tag{5.7}$$

The processes $\big(Z(t)\big)_{t\geqslant 0}$ and $\big(\mathcal{Z}(t)\big)_{t\geqslant 0}$ are interpreted as the mild solutions of the stochastic evolution equations

$$\begin{aligned} \mathrm{d}Z(t) &= \Delta Z(t)\,\mathrm{d}t + \mathrm{d}W(t), \\ \mathrm{d}\mathcal{Z}(t) &= -\Lambda\mathcal{Z}(t)\,\mathrm{d}t + \mathrm{d}\mathcal{W}(t) \end{aligned}$$

with initial values $Z(0) = 0$ and $\mathcal{Z}(0) = 0$. Note that $\mathcal{Z}(t) = \begin{pmatrix} Z(t) \\ 0\end{pmatrix}$ for all $t \geqslant 0$.

5.2.5 The stochastic FitzHugh–Nagumo SPDE system

The stochastic FhN SPDE system (5.1) can be rewritten as a stochastic evolution system

$$\begin{cases} \mathrm{d}u(t) = \Delta u(t)\,\mathrm{d}t + (u(t) - u^3(t) - v(t))\,\mathrm{d}t + \mathrm{d}W(t), \\ \mathrm{d}v(t) = (\gamma_1 u(t) - \gamma_2 v(t) + \beta)\,\mathrm{d}t, \\ u(0) = u_0, v(0) = v_0, \end{cases} \tag{5.8}$$

where the unknowns $u(\cdot) = \big(u(t)\big)_{t\geqslant 0}$ and $v(\cdot) = \big(v(t)\big)_{t\geqslant 0}$ are H-valued stochastic processes, and with initial values $u_0 \in H$ and $v_0 \in H$. Recall that Neumann boundary conditions are used in the above system. Using the notation introduced above and setting $X(t) = (u(t), v(t))$ for all $t \geqslant 0$, the stochastic evolution system (5.8) is treated in the sequel as the stochastic evolution equation

$$\mathrm{d}X(t) = -\Lambda X(t)\,\mathrm{d}t + F(X(t))\,\mathrm{d}t + \mathrm{d}\mathcal{W}(t),\ X(0) = x_0, \tag{5.9}$$

with the initial value $x_0 = (u_0, v_0) \in \mathcal{H}$. For all $T \in (0,\infty)$, a stochastic process $\big(X(t)\big)_{0\leqslant t\leqslant T}$ is called a mild solution of (5.9) if it has continuous trajectories with values in $\mathcal{H}$, and if for all $t \in [0,T]$ one has

$$X(t) = e^{-t\Lambda}x_0 + \int_0^t e^{-(t-s)\Lambda}F(X(s))\,\mathrm{d}s + \int_0^t e^{-(t-s)\Lambda}\,\mathrm{d}\mathcal{W}(s). \tag{5.10}$$

In the framework presented in this section, the stochastic evolution equation (5.9) admits a unique global mild solution, for any initial value $x_0 \in \mathcal{H}^{2\alpha} \cap \mathcal{E}$ and for $\alpha \in [0, \frac{1}{4})$, see Proposition 5.4.2 below.

For simplicity, the initial values u_0, v_0, resp. x_0, appearing in (5.8), resp. (5.9), are deterministic.

5.3 Splitting numerical schemes

The time-step size of the integrators defined below is denoted by Δt. Without loss of generality, it is assumed that $\Delta t \in (0, \Delta t_0)$, where Δt_0 is an arbitrary positive real number, and that there exists $T \in (0, \infty)$ and $N \in \mathbb{N}$ such that $\Delta t = T/N$. The notation $t_n = n\Delta t$ for $n \in \{0, \dots, N\}$ is used in the sequel. The increments of the Wiener processes $\big(W(t)\big)_{t\geqslant 0}$ and $\big(\mathcal{W}(t)\big)_{t\geqslant 0}$ are denoted by

$$\delta W_n = W(t_{n+1}) - W(t_n), \quad \delta \mathcal{W}_n = \mathcal{W}(t_{n+1}) - \mathcal{W}(t_n) = \begin{pmatrix} \delta W_n \\ 0 \end{pmatrix}.$$

The proposed time integrators for the SPDE (5.9) are based on a splitting strategy. Recall that the main principle of splitting integrators is to decompose the vector field of the evolution problem in several parts, such that the arising subsystems are exactly (or easily) integrated. In the next Subsection 5.3.1-5.3.3 all the details regarding the type of subsystems to be solved and the main results will be provided.

5.3.1 Solutions of auxiliary subsystems

The construction of the proposed splitting schemes is based on the combination of exact or approximate solutions of the three subsystems considered below.

- The nonlinear differential equation (considered on the Euclidean space $\mathbb{R}^2$)

$$\begin{cases} \dfrac{\mathrm{d}x^{\mathrm{NL}}(t)}{\mathrm{d}t} = F^{\mathrm{NL}}(x^{\mathrm{NL}}(t)), \\ x^{\mathrm{NL}}(0) = x_0 \in \mathbb{R}^2 \end{cases} \tag{5.11}$$

admits a unique global solution $\big(x^{\mathrm{NL}}(t)\big)_{t\geqslant 0}$. This solution has the following exact expression, see for instance [46, Equation (3)]: for all $t \geqslant 0$ and $x_0 = (u_0, v_0) \in \mathbb{R}^2$, one has

$$x^{\mathrm{NL}}(t) = \phi_t^{\mathrm{NL}}(x_0) = \begin{pmatrix} \frac{u_0}{\sqrt{u_0^2 + (1-u_0^2)e^{-2t}}} \\ v_0 + \beta t \end{pmatrix}. \tag{5.12}$$

- The linear differential equation (considered on the Euclidean space $\mathbb{R}^2$)

$$\begin{cases} \dfrac{\mathrm{d}x^{\mathrm{L}}(t)}{\mathrm{d}t} = F^{\mathrm{L}}(x^{\mathrm{L}}(t)), \\ x^{\mathrm{L}}(0) = x_0 \in \mathbb{R}^2 \end{cases} \tag{5.13}$$

admits a unique global solution $\big(x^{\mathrm{L}}(t)\big)_{t\geqslant 0}$. This solution has the following expression: for all $t \geqslant 0$ and $x_0 = (u_0, v_0) \in \mathbb{R}^2$, one has

$$x^{\mathrm{L}}(t) = \phi_t^{\mathrm{L}}(x_0) = e^{tB}x_0. \tag{5.14}$$

- The stochastic evolution equation (considered on the Hilbert space $\mathcal{H}$)

$$\begin{cases} \mathrm{d}X^{\mathrm{s}}(t) = -\Lambda X^{\mathrm{s}}(t)\,\mathrm{d}t + \mathrm{d}\mathcal{W}(t) \\ X^{\mathrm{s}}(0) = x_0 \in \mathcal{H} \end{cases} \tag{5.15}$$

admits a unique global solution $\big(X^{\mathrm{s}}(t)\big)_{t\geqslant 0}$. This solution has the following expression: for all $t \geqslant 0$ and $x_0 = (u_0, v_0) \in \mathcal{H}$, one has

$$X^{\mathrm{s}}(t) = e^{-t\Lambda}x_0 + \int_0^t e^{-(t-s)\Lambda}\mathrm{d}\mathcal{W}(s) = \begin{pmatrix} e^{t\Delta}u_0 + \int_0^t e^{(t-s)\Delta}\,\mathrm{d}W(s) \\ v_0 \end{pmatrix}, \tag{5.16}$$

see (5.7) for the expression of the stochastic convolution. For all $n \in \{0, \dots, N-1\}$, set $X_n^{\mathrm{s,exact}} = X^{\mathrm{s}}(t_n)$, then one has the following recursion formula

$$X_{n+1}^{\mathrm{s,exact}} = e^{-\Delta t\Lambda}X_n^{\mathrm{s,exact}} + \int_{t_n}^{t_{n+1}} e^{-(t_{n+1}-s)\Lambda}\,\mathrm{d}\mathcal{W}(s) \tag{5.17}$$

recalling the notation $t_n = n\Delta t$.

Instead of using the exact solution (5.16) of the stochastic convolution (5.15), one can use approximate solutions $\big(X_n^{\mathrm{s,exp}}\big)_{n\geqslant 0} = \big(u_n^{\mathrm{s,exp}}, v_n^{\mathrm{s,exp}}\big)_{n\geqslant 0}$ and $\big(X_n^{\mathrm{s,imp}}\big)_{n\geqslant 0} = \big(u_n^{\mathrm{s,imp}}, v_n^{\mathrm{s,imp}}\big)_{n\geqslant 0}$ defined by an exponential Euler scheme and a linear implicit Euler scheme, respectively:

$$X_{n+1}^{\mathrm{s,exp}} = e^{-\Delta t\Lambda}\Big(X_n^{\mathrm{s,exp}} + \delta\mathcal{W}_n\Big) = \begin{pmatrix} e^{\Delta t\Delta}\big(u_n^{\mathrm{s,exp}} + \delta W_n\big) \\ v_n^{\mathrm{s,exp}} \end{pmatrix}, \tag{5.18}$$

and

$$X_{n+1}^{\mathrm{s,imp}} = \big(I + \Delta t\Lambda\big)^{-1}\Big(X_n^{\mathrm{s,imp}} + \delta\mathcal{W}_n\Big) = \begin{pmatrix} (I - \Delta t\Delta)^{-1}\big(u_n^{\mathrm{s,imp}} + \delta W_n\big) \\ v_n^{\mathrm{s,imp}} \end{pmatrix}, \tag{5.19}$$

with initial values $X_0^{\mathrm{s,exp}} = X_0^{\mathrm{s,imp}} = x_0 = (u_0, v_0) \in \mathcal{H}$, $u_0^{\mathrm{s,exp}} = u_0^{\mathrm{s,imp}} = u_0 \in H$ and $v_0^{\mathrm{s,exp}} = v_0^{\mathrm{s,imp}} = v_0 \in H$.

5.3.2 Definition of the splitting schemes

Now the three splitting schemes are introduced. They are constructed using a Lie–Trotter strategy, where first the subsystems (5.11), (5.13) are solved exactly using the flow maps (5.12) and (5.14) respectively, and where the subsystem (5.15) is either solved exactly using (5.16) or approximately using (5.18) or (5.19).

For the composition of the first two subsystems, define the mapping $\phi_{\Delta t} : \mathbb{R}^2 \to \mathbb{R}^2$ as follows: for all $\Delta t \in (0, \Delta t_0)$, set

$$\phi_{\Delta t} = \phi_{\Delta t}^{\mathrm{L}} \circ \phi_{\Delta t}^{\mathrm{NL}}. \tag{5.20}$$

Using the expression (5.17) for the exact solution (5.16) of (5.15) leads to the definition of the following explicit splitting scheme for the stochastic FitzHugh–Nagumo SPDE system (5.8):

$$X_{n+1}^{\mathrm{LT,exact}} = e^{-\Delta t\Lambda}\phi_{\Delta t}\big(X_n^{\mathrm{LT,exact}}\big) + \int_{t_n}^{t_{n+1}} e^{-(t_{n+1}-s)\Lambda}\,\mathrm{d}\mathcal{W}(s). \tag{5.21}$$

Using the exponential Euler scheme (5.18) to approximate the solution of (5.15) leads to the definition of the following explicit splitting scheme for (5.8):

$$X_{n+1}^{\mathrm{LT,expo}} = e^{-\Delta t\Lambda}\phi_{\Delta t}\big(X_n^{\mathrm{LT,expo}}\big) + e^{-\Delta t\Lambda}\delta\mathcal{W}_n. \tag{5.22}$$

Using the linear implicit Euler scheme (5.19) to approximate the solution of (5.15) leads to the definition of the following splitting scheme for (5.8):

$$X_{n+1}^{\mathrm{LT,imp}} = (I+\Delta t\Lambda)^{-1}\phi_{\Delta t}\big(X_n^{\mathrm{LT,imp}}\big) + (I+\Delta t\Lambda)^{-1}\delta\mathcal{W}_n. \tag{5.23}$$

For these three Lie–Trotter splitting schemes (5.21), (5.22) and (5.23), the same initial value is imposed:

$$X_0^{\mathrm{LT,exact}} = X_0^{\mathrm{LT,expo}} = X_0^{\mathrm{LT,imp}} = x_0 \in \mathcal{H}.$$

Before proceeding with the statements of the main results, let us give several observations and auxiliary tools.

Observe that all these three schemes (5.21), (5.22) and (5.23) can be written using the single formulation

$$X_{n+1} = \mathcal{A}_{\Delta t}\phi_{\Delta t}(X_n) + \int_{t_n}^{t_{n+1}} \mathcal{B}_{t_{n+1}-s}\,\mathrm{d}\mathcal{W}(s) \tag{5.24}$$

which is used in the analysis below. The expressions of the linear operators $\mathcal{A}_{\Delta t}$ and $\mathcal{B}_{t_{n+1}-s}$ for each of the three schemes are given by: $\mathcal{A}_{\Delta t} = e^{-\Delta t\Lambda}$, $\mathcal{B}_{t_{n+1}-s} = e^{-(t_{n+1}-s)\Lambda}$ for the scheme (5.21) $\mathcal{A}_\Delta t = \mathcal{B}_{t_{n+1}-s} = e^{-\Delta t\Lambda}$ for the scheme (5.22), and $\mathcal{A}_{\Delta t} = \mathcal{B}_{t_{n+1}-s} = (I+\Delta t\Lambda)^{-1}$ for the scheme (5.23).

For any value $\Delta t \in (0, \Delta t_0)$ of the time-step size, introduce the mapping $\psi_{\Delta t} : \mathbb{R}^2 \to \mathbb{R}^2$ defined as follows: for all $x \in \mathbb{R}^2$,

$$\psi_{\Delta t}(x) = \frac{\phi_{\Delta t}(x) - x}{\Delta t}. \tag{5.25}$$

The Lie–Trotter splitting scheme (5.24) is then written as

$$X_{n+1} = \mathcal{A}_{\Delta t}X_n + \Delta t\mathcal{A}_{\Delta t}\psi_{\Delta t}(X_n) + \int_{t_n}^{t_{n+1}} \mathcal{B}_{t_{n+1}-s}\,\mathrm{d}\mathcal{W}(s)$$

and can thus be interpreted as a numerical scheme applied to the auxiliary stochastic evolution equation

$$\mathrm{d}X_{\Delta t}(t) = -\Lambda X_{\Delta t}(t)\,\mathrm{d}t + \psi_{\Delta t}(X_{\Delta t}(t))\,\mathrm{d}t + \mathrm{d}\mathcal{W}(t),\ X_{\Delta t}(0) = x_0. \tag{5.26}$$

Note that the SPDE (5.26) is similar to the original problem (5.9), however the nonlinearity F is replaced by the auxiliary mapping $\psi_{\Delta t}$.

5.3.3 Principal results

In this Subsection, the principal results of this chapter will be stated. First, moment bounds for the three splitting schemes (5.24) are given, see Theorem 5.3.1. Then, strong error estimates, with rate of convergence 1/4, for the numerical approximations of the solution of the stochastic FitzHugh–Nagumo SPDE system (5.9), see Theorem 5.3.2 will be given.

Theorem 5.3.1. *For all $T \in (0,\infty)$ and $p \in [1,\infty)$, there exists $C_p(T) \in (0,\infty)$ such that for all $x_0 \in \mathcal{E}$ one has*

$$\sup_{\Delta t\in(0,\Delta t_0)}\sup_{0\leqslant n\leqslant N}\mathbb{E}[\|X_n\|_{\mathcal{E}}^p] \leqslant C_p(T)\big(1+\|x_0\|_{\mathcal{E}}^p\big), \tag{5.27}$$

where $\big(X_n\big)_{n\geqslant 0}$ is given by (5.24) (with initial value $X_0 = x_0$), and where $T = N\Delta t$ with $N \in \mathbb{N}$.

The proof of this theorem is postponed to Section 5.5.

The following theorem give us the main results of this Chapter regarding the strong convergence of the three splitting schemes of type (5.24). Its proof is given in Section 5.5.

Theorem 5.3.2. *For all $T \in (0,\infty)$, $p \in [1,\infty)$ and $\alpha \in [0,\frac{1}{4})$, there exists $C_{\alpha,p}(T) \in (0,\infty)$ such that for all $x_0 = (u_0,v_0) \in \mathcal{H}^{2\alpha}\cap\mathcal{E}$, all $\Delta t \in (0,\Delta t_0)$, one has*

$$\sup_{0\leqslant n\leqslant N}\big(\mathbb{E}[\|X(t_n)-X_n\|_{\mathcal{H}}^p]\big)^{\frac{1}{p}} \leqslant C_{\alpha,p}(T)\Delta t^{\alpha}\big(1+\|(-\Delta)^{\alpha}u_0\|_H^7+\|x_0\|_{\mathcal{E}}^7\big). \tag{5.28}$$

The order of convergence 1/4 obtained in Theorem 5.3.2 is consistent with the temporal Hölder regularity property of the trajectories $t \mapsto X(t) \in \mathcal{H}$. It is also consistent with the strong convergence rate obtained in [45] for the stochastic Allen–Cahn equation. However new arguments are required to study the FitzHugh–Nagumo system which is not a parabolic SPDE problem, and which has a cubic nonlinearity.

Let us state two of the main auxiliary results which are used in the proofs of the main results. These propositions are proved in Section 5.5.

Proposition 5.3.1. *For all $\Delta t \in (0,\Delta t_0)$, the mapping $\phi_{\Delta t}:\mathbb{R}^2\to\mathbb{R}^2$ defined by (5.20) is globally Lipschitz continuous. In addition, for all $\Delta t \in (0,\Delta t_0)$ and all $x_1,x_2\in\mathbb{R}^2$ one has*

$$\|\phi_{\Delta t}(x_2)-\phi_{\Delta t}(x_1)\| \leqslant e^{(1+\|B\|)\Delta t}\|x_2-x_1\|. \tag{5.29}$$

Proposition 5.3.2. *There exists $C(\Delta t_0)\in(0,\infty)$ such that for all $\Delta t \in (0,\Delta t_0)$, the mapping $\psi_{\Delta t}:\mathbb{R}^2\to\mathbb{R}^2$ defined by (5.25) satisfies the following properties: for all $x_1,x_2\in\mathbb{R}^2$, one has*

$$\langle x_2-x_1,\psi_{\Delta t}(x_2)-\psi_{\Delta t}(x_1)\rangle \leqslant C(\Delta t_0)\|x_2-x_1\|^2 \tag{5.30}$$

$$\|\psi_{\Delta t}(x_2)-\psi_{\Delta t}(x_1)\| \leqslant C(\Delta t_0)\big(1+\|x_1\|^3+\|x_2\|^3\big)\|x_2-x_1\|, \tag{5.31}$$

and for all $x \in \mathbb{R}^2$ *one has*

$$\|\psi_{\Delta t}(x) - F(x)\| \leqslant C(\Delta t_0)\Delta t\big(1 + \|x\|^5\big). \tag{5.32}$$

Finally,

$$\sup_{\Delta t \in (0,\Delta t_0)} \|\psi_{\Delta t}(0)\| < \infty. \tag{5.33}$$

The inequality (5.30) states that $\psi_{\Delta t}$ satisfies a one-sided Lipschitz continuity property that is uniform with respect to $\Delta t \in (0, \Delta t_0)$. This is similar to the property (5.5) satisfied by F. It is straightforward to check that $\psi_{\Delta t}$ is in fact globally Lipschitz continuous for any fixed $\Delta t \in (0, \Delta t_0)$, however this property does not hold uniformly with respect to $\Delta t \in (0, \Delta t_0)$. Instead, one has the one-sided Lipschitz continuity property (5.30) and the local Lipschitz continuity property (5.31) that are both uniform with respect to $\Delta t \in (0, \Delta t_0)$.

5.4 Preliminary results

In this section several results will be proven, which are required for the analysis of the three splitting schemes of type (5.24). In particular, properties of the semigroup (Proposition 5.4.1) are obtained and moreover the properties of the auxiliary mappings $\phi_{\Delta t}$ (Proposition 5.3.1) and $\psi_{\Delta t}$ (Proposition 5.3.2) will be given. Furthermore the well-posedness and moment bounds for the mild solution of the considered SPDE will be studied.

5.4.1 Properties of the semigroup

In this subsection, properties of the semigroup generated by the linear operator Λ in the stochastic FitzHugh–Nagumo system (5.9) are studied. In addition, estimates for the operator $(I + \Delta t\Lambda)^{-1}$ used in the semi-linear splitting schemes (5.19) and (5.23) are also provided.

Proposition 5.4.1. *The semigroup* $\big(e^{-t\Lambda}\big)_{t \geqslant 0}$ *defined by* (5.3) *satisfies the following properties:*

1. *For all* $t \geqslant 0$, $e^{-t\Lambda}$ *is a bounded linear operator from* $\mathcal{H}$ *to* $\mathcal{H}$ *and from* $\mathcal{E}$ *to* $\mathcal{E}$. *In addition, for all* $t \geqslant 0$ *one has*

$$\sup_{x \in \mathcal{H}\setminus\{0\}} \frac{\|e^{-t\Lambda}x\|_{\mathcal{H}}}{\|x\|_{\mathcal{H}}} = 1, \qquad \sup_{x \in \mathcal{E}\setminus\{0\}} \frac{\|e^{-t\Lambda}x\|_{\mathcal{E}}}{\|x\|_{\mathcal{E}}} = 1. \tag{5.34}$$

2. *Smoothing property. For all* $\alpha \in [0, \infty)$, *there exists a real number* $C_\alpha \in (0, \infty)$ *such that, for all* $(u, v) \in \mathcal{H}$ *and all* $t \in (0, \infty)$, *one has*

$$\|e^{-t\Lambda}\begin{pmatrix}(-\Delta)^\alpha u \\ v\end{pmatrix}\|_{\mathcal{H}} \leqslant C_\alpha(\min\{1, t\})^{-\alpha}\|(u, v)\|_{\mathcal{H}}. \tag{5.35}$$

3. *Temporal regularity. For all $\mu, \nu \geqslant 0$ with $\mu + \nu \leqslant 1$, there exists a real number $C_{\mu,\nu} \in (0,\infty)$ such that, for all $x = (u,v) \in \mathcal{H}^{2\nu}$ and all $t_1, t_2 \in (0,\infty)$, one has*

$$\|e^{-t_2\Lambda}x - e^{-t_1\Lambda}x\|_{\mathcal{H}} \leqslant C_{\mu,\nu}\frac{|t_2 - t_1|^{\mu+\nu}}{(\min\{t_2,t_1\})^{\mu}}\|(-\Delta)^{\nu}u\|_H. \tag{5.36}$$

Proof. :

1. Since the eigenvalues $\big(\lambda_j\big)_{j\geqslant 0}$ of $-\Delta$ are nonnegative, it is straightforward to see that for all $x = (u,v) \in \mathcal{H}$ and $t \geqslant 0$ one has $e^{-t\Lambda}x \in \mathcal{H}$, and

$$\|e^{-t\Lambda}x\|_{\mathcal{H}}^2 = \|e^{t\Delta}u\|_H^2 + \|v\|_H^2 \leqslant \|u\|_H^2 + \|v\|_H^2 = \|x\|_{\mathcal{H}}^2.$$

   This proves that $e^{-t\Lambda}$ is a bounded linear operator from $\mathcal{H}$ to $\mathcal{H}$ for all $t \geqslant 0$, and that

$$\sup_{x\in\mathcal{H}\setminus\{0\}}\frac{\|e^{-t\Lambda}x\|_{\mathcal{H}}}{\|x\|_{\mathcal{H}}} \leqslant 1.$$

   On the other hand, using the formula for the Green function of the heat equation with homogeneous Neumann boundary conditions, the semigroup $\big(e^{t\Delta}\big)_{t\geqslant 0}$ defined by (5.2) satisfies the following properties: for all $t \geqslant 0$ and $u \in E$, one has $e^{t\Delta}u \in E$ and $\|e^{t\Delta}u\|_E \leqslant \|u\|_E$. As a consequence of the facts, for all $x = (u,v) \in \mathcal{E}$, one has $e^{-t\Lambda}x = (e^{t\Delta}u, v) \in \mathcal{E}$ and

$$\|e^{-t\Lambda}x\|_{\mathcal{E}} = \max\big(\|e^{t\Delta}u\|_E, \|v\|_E\big) \leqslant \max\big(\|u\|_E, \|v\|_E\big) = \|x\|_{\mathcal{E}}.$$

   To conclude the proof of (5.34), it suffices to check that for $x = (0,v)$ and all $t \geqslant 0$ one has $e^{-t\Lambda}x = x$.

2. The smoothing property (5.35) is a straightforward consequence of the smoothing property for the semigroup $\big(e^{t\Delta}\big)_{t\geqslant 0}$: for all $\alpha \in [0,\infty)$, $t \geqslant 0$ and $u \in H$, one has (recall that $\lambda_0 = 0$)

$$\|e^{t\Delta}(-\Delta)^{\alpha}u\|_H^2 = \sum_{j\geqslant 1} e^{-2t\lambda_j}\lambda_j^{2\alpha}\langle u, e_j\rangle_H^2 \leqslant \sup_{\xi\in(0,\infty)}\big(\xi^{2\alpha}e^{-2\xi}\big)t^{-2\alpha}\|u\|_H^2.$$

   As a consequence, for all $\alpha \in [0,\infty)$, $t \geqslant 0$ and $x = (u,v) \in \mathcal{H}$, one has

$$\begin{aligned}\|e^{-t\Lambda}\begin{pmatrix}(-\Delta)^{\alpha}u\\ v\end{pmatrix}\|_{\mathcal{H}}^2 &= \|e^{t\Delta}(-\Delta)^{\alpha}u\|_H^2 + \|v\|_H^2\\ &\leqslant C_{\alpha}^2 t^{-2\alpha}\|u\|_H^2 + \|v\|_H^2\\ &\leqslant C_{\alpha}^2(\min\{1,t\})^{-2\alpha}\|x\|_{\mathcal{H}}^2.\end{aligned}$$

3. The regularity property (5.36) is a straightforward consequence of the following regularity property for the semigroup $\big(e^{t\Delta}\big)_{t\geqslant 0}$: for all $\mu, \nu \in$

$[0,1]$ with $\mu+\nu \leqslant 1$, $0 \leqslant t_1 \leqslant t_2$ and $u \in H^{2\nu}$, one has

$$\begin{aligned}
\|e^{t_2\Delta}u - e^{t_1\Delta}u\|_H^2 &= \|(e^{(t_2-t_1)\Delta} - I)e^{t_1\Delta}u\|_H^2 \\
&= \sum_{j\geqslant 1}\big(e^{-(t_2-t_1)\lambda_j} - 1\big)^2 e^{-2t_1\lambda_j}\langle u, e_j\rangle_H^2 \\
&\leqslant 2^{2(\mu+\nu)}(t_2-t_1)^{2(\mu+\nu)}\sum_{j\geqslant 1}\lambda_j^{2(\mu+\nu)}e^{-2t_1\lambda_j}\langle u, e_j\rangle_H^2 \\
&\leqslant 2^{2(\mu+\nu)}\sup_{\xi\in(0,\infty)}\big(\xi^{2\mu}e^{-2\xi}\big)\frac{(t_2-t_1)^{2(\mu+\nu)}}{t_1^{2\mu}}\sum_{j\geqslant 1}\lambda_j^{2\nu}\langle u, e_j\rangle_H^2 \\
&\leqslant 2^{2(\mu+\nu)}\sup_{\xi\in(0,\infty)}\big(\xi^{2\mu}e^{-2\xi}\big)\frac{(t_2-t_1)^{2(\mu+\nu)}}{t_1^{2\mu}}\|(-\Delta)^\nu u\|_H^2.
\end{aligned}$$

As a consequence, for all $\mu,\nu \in [0,1]$ with $\mu+\nu \leqslant 1$, $0 \leqslant t_1 \leqslant t_2$ and $x = (u,v) \in H^{2\nu}\times H$, one has

$$\|e^{-t_2\Lambda}x - e^{-t_1\Lambda}x\|_{\mathcal{H}} = \|e^{t_2\Delta}u - e^{t_1\Delta}u\|_H \leqslant C_{\mu,\nu}\frac{|t_2-t_1|^{\mu+\nu}}{t_1^{\mu}}\|(-\Delta)^\nu u\|_H.$$

The proof of Proposition 5.4.1 is thus completed. □

In the sequel, the following properties are also used for the analysis of the splitting scheme (5.23) for which a linear implicit Euler method is used for the approximation (5.19) of the stochastic convolution: for all $t \geqslant 0$, $(I+t\Lambda)^{-1}$ is a bounded linear operator from $\mathcal{H}$ to $\mathcal{H}$ and from $\mathcal{E}$ to $\mathcal{E}$, and one has

$$\sup_{x\in\mathcal{H}\setminus\{0\}}\frac{\|(I+t\Lambda)^{-1}x\|_{\mathcal{H}}}{\|x\|_{\mathcal{H}}} = 1, \qquad \sup_{x\in\mathcal{E}\setminus\{0\}}\frac{\|(I+t\Lambda)^{-1}x\|_{\mathcal{E}}}{\|x\|_{\mathcal{E}}} = 1. \tag{5.37}$$

The proof of the inequality (5.37) is straightforward. Indeed, for all $x \in \mathcal{H}$ or $x \in \mathcal{E}$, and all $t \geqslant 0$, one has

$$(I+t\Lambda)^{-1}x = \int_0^\infty e^{-(I+t\Lambda)s}x\,\mathrm{d}s.$$

Using (5.34), one then obtains the inequalities

$$\begin{aligned}
\|(I+t\Lambda)^{-1}x\|_{\mathcal{H}} &\leqslant \int_0^\infty e^{-s}\|e^{-ts\Lambda}x\|_{\mathcal{H}}\,\mathrm{d}s \leqslant \int_0^\infty e^{-s}\,\mathrm{d}s\|x\|_{\mathcal{H}} = \|x\|_{\mathcal{H}} \\
\|(I+t\Lambda)^{-1}x\|_{\mathcal{E}} &\leqslant \int_0^\infty e^{-s}\|e^{-ts\Lambda}x\|_{\mathcal{E}}\,\mathrm{d}s \leqslant \int_0^\infty e^{-s}\,\mathrm{d}s\|x\|_{\mathcal{E}} = \|x\|_{\mathcal{E}}.
\end{aligned}$$

Like in the proof of (5.34), choosing $x = (0,v)$ gives $(I+t\Lambda)^{-1}x = x$ for all $t \geqslant 0$, and thus concludes the proof of (5.37).

In order to prove Propositions 5.3.1 and 5.3.2 that state properties of the mappings $\phi_{\Delta t}: \mathbb{R}^2 \to \mathbb{R}^2$ and $\psi_{\Delta t}: \mathbb{R}^2 \to \mathbb{R}^2$ defined by (5.20) and (5.25), it is convenient to introduce the auxiliary mappings $\phi_t^{\mathrm{AC}}: \mathbb{R} \to \mathbb{R}$ and $\psi_t^{\mathrm{AC}}: \mathbb{R} \to \mathbb{R}$, defined as follows: for all $t \in (0,\infty)$ and $u \in \mathbb{R}$, set

$$\phi_t^{\mathrm{AC}}(u) = \frac{u}{\sqrt{u^2 + (1-u^2)e^{-2t}}}, \qquad \psi_t^{\mathrm{AC}}(u) = \frac{\phi_t^{\mathrm{AC}}(u) - u}{t}. \tag{5.38}$$

The mapping ϕ_t^{AC} is the flow map associated with the nonlinear differential equation, see the subsystem (5.11),

$$\frac{\mathrm{d}u^{\mathrm{AC}}(t)}{\mathrm{d}t} = u^{\mathrm{AC}}(t) - (u^{\mathrm{AC}}(t))^3,$$

meaning that $u^{\mathrm{AC}}(t) = \phi_t^{\mathrm{AC}}(u^{\mathrm{AC}}(0))$ for all $t \geqslant 0$. The properties of the mappings $\phi_{\Delta t}^{\mathrm{AC}}$ and $\psi_{\Delta t}^{\mathrm{AC}}$ stated in Lemma 5.4.1 are given by [46, Lemma 3.1–3.4].

Lemma 5.4.1. *There exists $C(\Delta t_0) \in (0, \infty)$ such that for all $\Delta t \in (0, \Delta t_0)$, the mappings $\phi_{\Delta t}^{\mathrm{AC}} : \mathbb{R} \to \mathbb{R}$ and $\psi_{\Delta t}^{\mathrm{AC}} : \mathbb{R} \to \mathbb{R}$ satisfy the following properties:*

- *For all $\Delta t \in (0, \Delta t_0)$ and $u_1, u_2 \in \mathbb{R}$, one has*

$$|\phi_{\Delta t}^{\mathrm{AC}}(u_2) - \phi_{\Delta t}^{\mathrm{AC}}(u_1)| \leqslant e^{\Delta t}|u_2 - u_1|. \tag{5.39}$$

- *For all $\Delta t \in (0, \Delta t_0)$ and $u_1, u_2 \in \mathbb{R}$, one has*

$$\big(u_2 - u_1\big)\big(\psi_{\Delta t}^{\mathrm{AC}}(u_2) - \psi_{\Delta t}^{\mathrm{AC}}(u_1)\big) \leqslant C(\Delta t_0)|u_2 - u_1|^2, \tag{5.40}$$
$$|\psi_{\Delta t}^{\mathrm{AC}}(u_2) - \psi_{\Delta t}^{\mathrm{AC}}(u_1)| \leqslant C(\Delta t_0)\big(1 + |u_1|^3 + |u_2|^3\big)|u_2 - u_1|, \tag{5.41}$$

  and for all $\Delta t \in (0, \Delta t_0)$ and $u \in \mathbb{R}$, one has

$$|\psi_{\Delta t}(u) - (u - u^3)| \leqslant C(\Delta t_0)\Delta t\big(1 + |u|^5\big). \tag{5.42}$$

Proof of Proposition 5.3.1. Note that for all $\Delta t \in (0, \Delta t_0)$ and $x = (u, v) \in \mathbb{R}^2$ one has

$$\phi_{\Delta t}^{\mathrm{NL}}(x) = \begin{pmatrix} \phi_{\Delta t}^{\mathrm{AC}}(u) \\ v + \beta \Delta t \end{pmatrix}.$$

Using the definition (5.20) and the inequality (5.39) from Lemma 5.4.1, one then obtains the following inequality: for all $\Delta t \in (0, \Delta t_0)$ and all $x_1 = (u_1, v_1), x_2 = (u_2, v_2) \in \mathbb{R}^2$, one has

$$\begin{aligned}
\|\phi_{\Delta t}(x_2) - \phi_{\Delta t}(x_1)\|^2 &= \|\phi_{\Delta t}^{\mathrm{L}}(\phi_{\Delta t}^{\mathrm{NL}}(x_2)) - \phi_{\Delta t}^{\mathrm{L}}(\phi_{\Delta t}^{\mathrm{NL}}(x_1))\|^2 \\
&= \|e^{\Delta t B}\big(\phi_{\Delta t}^{\mathrm{NL}}(x_2) - \phi_{\Delta t}^{\mathrm{NL}}(x_1)\big)\|^2 \\
&\leqslant e^{2\Delta t \|\!| B |\!\|}\|\phi_{\Delta t}^{\mathrm{NL}}(x_2) - \phi_{\Delta t}^{\mathrm{NL}}(x_1)\|^2 \\
&\leqslant e^{2\Delta t \|\!| B |\!\|}\big(|\phi_{\Delta t}^{\mathrm{AC}}(u_2) - \phi_{\Delta t}^{\mathrm{AC}}(u_1)|^2 + |v_2 - v_1|^2\big) \\
&\leqslant e^{2\Delta t \|\!| B |\!\|}\big(e^{2\Delta t}|u_2 - u_1|^2 + |v_2 - v_1|^2\big) \\
&\leqslant e^{2\Delta t(1 + \|\!| B |\!\|)}\|x_2 - x_1\|^2.
\end{aligned}$$

Then the result is straightforward. In fact, $\phi_{\Delta t}$ is the composition of the two globally Lipschitz continuous mappings $\phi_{\Delta t}^{\mathrm{L}}$ and $\phi_{\Delta t}^{\mathrm{NL}}$. The proof is given to exhibit the dependence of the Lipschitz constant with respect to the time-step size $\Delta t \in (0, \Delta t_0)$.
This concludes the proof of Proposition 5.3.1. □

In order to prove Proposition 5.3.2, the main tool is the following expression for the mapping $\psi_{\Delta t}$ defined by (5.25): for all $\Delta t \in (0, \Delta t_0)$ and $x \in \mathbb{R}^2$, one has

$$\psi_{\Delta t}(x) = \psi^{\mathrm{L}}_{\Delta t}(\phi^{\mathrm{NL}}_{\Delta t}(x)) + \psi^{\mathrm{NL}}_{\Delta t}(x), \tag{5.43}$$

where the mappings $\psi^{\mathrm{L}}_{\Delta t}$ and $\psi^{\mathrm{NL}}_{\Delta t}$ are given by

$$\psi^{\mathrm{L}}_{\Delta t}(x) = \frac{\phi^{\mathrm{L}}_{\Delta t}(x) - x}{\Delta t} = \frac{e^{\Delta t B} - I}{\Delta t} x$$
$$\psi^{\mathrm{NL}}_{\Delta t}(x) = \frac{\phi^{\mathrm{NL}}_{\Delta t}(x) - x}{\Delta t} = \begin{pmatrix} \psi^{\mathrm{AC}}_{\Delta t}(u) \\ \beta \end{pmatrix}$$

for all $\Delta t \in (0, \Delta t_0)$ and $x = (u, v) \in \mathbb{R}^2$.

The proof of the equality (5.43) is straightforward: using (5.20), one has

$$\begin{aligned} \psi_{\Delta t}(x) &= \frac{\phi_{\Delta t}(x) - x}{\Delta t} = \frac{\phi^{\mathrm{L}}_{\Delta t}(\phi^{\mathrm{NL}}_{\Delta t}(x)) - \phi^{\mathrm{NL}}_{\Delta t}(x)}{\Delta t} + \frac{\phi^{\mathrm{NL}}_{\Delta t}(x) - x}{\Delta t} \\ &= \psi^{\mathrm{L}}_{\Delta t}(\phi^{\mathrm{NL}}_{\Delta t}(x)) + \psi^{\mathrm{NL}}_{\Delta t}(x). \end{aligned}$$

Thanks to the identity (5.43), Proposition 5.3.2 can be proved.

Proof of Proposition 5.3.2. Note that the mapping $\psi^{\mathrm{L}}_{\Delta t} : \mathbb{R}^2 \to \mathbb{R}^2$ is linear and therefore is globally Lipschitz continuous. In addition, for all $\Delta t \in (0, \Delta t_0)$ and $x_1, x_2 \in \mathbb{R}^2$, one has

$$\|\psi^{\mathrm{L}}_{\Delta t}(x_2) - \psi^{\mathrm{L}}_{\Delta t}(x_1)\| \leqslant |\!|\!| \frac{e^{\Delta t B} - I}{\Delta t} |\!|\!| \, \|x_2 - x_1\| \leqslant \frac{e^{\Delta t_0 |\!|\!| B |\!|\!|} - 1}{\Delta t_0} \|x_2 - x_1\|, \tag{5.44}$$

using the inequalities

$$\begin{aligned} |\!|\!| \frac{e^{\Delta t B} - I}{\Delta t} |\!|\!| &= |\!|\!| \sum_{k=1}^{\infty} \frac{\Delta t^{k-1}}{k!} B^k |\!|\!| \\ &\leqslant \sum_{k=1}^{\infty} \frac{\Delta t^{k-1}}{k!} |\!|\!| B |\!|\!|^k \\ &\leqslant \sum_{k=1}^{\infty} \frac{\Delta t_0^{k-1}}{k!} |\!|\!| B |\!|\!|^k \\ &= \frac{e^{\Delta t_0 |\!|\!| B |\!|\!|} - 1}{\Delta t_0}. \end{aligned}$$

Let us first prove the one-sided Lipschitz continuity property (5.30): for all $\Delta t \in (0, \Delta t_0)$ and $x_1, x_2 \in \mathbb{R}^2$, using the identity (5.43), then the Cauchy–Schwarz inequality and (5.44), one has

$$\begin{aligned} \langle x_2 - x_1, \psi_{\Delta t}(x_2) - \psi_{\Delta t}(x_1) \rangle &= \langle x_2 - x_1, \psi^{\mathrm{L}}_{\Delta t}(\phi^{\mathrm{NL}}_{\Delta t}(x_2)) - \psi^{\mathrm{L}}_{\Delta t}(\phi^{\mathrm{NL}}_{\Delta t}(x_1)) \rangle \\ &\quad + \langle x_2 - x_1, \psi^{\mathrm{NL}}_{\Delta t}(x_2) - \psi^{\mathrm{NL}}_{\Delta t}(x_1) \rangle \\ &\leqslant \frac{e^{\Delta t_0 |\!|\!| B |\!|\!|} - 1}{\Delta t_0} \|x_2 - x_1\| \|\phi^{\mathrm{NL}}_{\Delta t}(x_2) - \phi^{\mathrm{NL}}_{\Delta t}(x_2)\| \\ &\quad + \langle x_2 - x_1, \psi^{\mathrm{NL}}_{\Delta t}(x_2) - \psi^{\mathrm{NL}}_{\Delta t}(x_1) \rangle. \end{aligned}$$

On the one hand, using the same arguments as in the proof of Proposition 5.3.1, one has

$$\|\phi_{\Delta t}^{\mathrm{NL}}(x_2) - \phi_{\Delta t}^{\mathrm{NL}}(x_1)\| \leqslant e^{\Delta} t \|x_2 - x_1\| \leqslant e^{\Delta t_0} \|x_2 - x_1\|.$$

On the other hand, for all $x = (u, v) \in \mathbb{R}^2$ one has

$$\psi_{\Delta t}^{\mathrm{NL}}(x) = \begin{pmatrix} \psi_{\Delta t}^{\mathrm{AC}}(u) \\ \beta \end{pmatrix}.$$

Using the inequality (5.40) from Lemma 5.4.1, one then obtains

$$\langle x_2 - x_1, \psi_{\Delta t}^{\mathrm{NL}}(x_2) - \psi_{\Delta t}^{\mathrm{NL}}(x_1) \rangle \leqslant e^{\Delta} t \|x_2 - x_1\|^2 \leqslant e^{\Delta t_0} \|x_2 - x_1\|^2.$$

Gathering the results then gives

$$\langle x_2 - x_1, \psi_{\Delta t}(x_2) - \psi_{\Delta t}(x_1) \rangle \leqslant \Big(\frac{e^{\Delta t_0 |||B|||} - 1}{\Delta t_0} + 1 \Big) e^{\Delta t_0} \|x_2 - x_1\|^2,$$

which concludes the proof of the inequality (5.30).

Let us now prove the local Lipschitz continuity property (5.31). Using the identity (5.43) and the inequality (5.41), for all $\Delta t \in (0, \Delta t_0)$ and $x_1 = (u_1, v_1), x_2 = (u_2, v_2) \in \mathbb{R}^2$, one has

$$\begin{aligned}
\|\psi_{\Delta t}(x_2) - \psi_{\Delta t}(x_1)\| &\leqslant \|\psi_{\Delta t}^{\mathrm{L}}(\phi_{\Delta t}^{\mathrm{NL}}(x_2)) - \psi_{\Delta t}^{\mathrm{L}}(\phi_{\Delta t}^{\mathrm{NL}}(x_1))\| + \|\psi_{\Delta t}^{\mathrm{NL}}(x_2) - \psi_{\Delta t}^{\mathrm{NL}}(x_1)\| \\
&\leqslant \frac{e^{\Delta t_0 |||B|||} - 1}{\Delta t_0} \|\phi_{\Delta t}^{\mathrm{NL}}(x_2) - \phi_{\Delta t}^{\mathrm{NL}}(x_1)\| \\
&\qquad + C(\Delta t_0) \big(1 + |u_1|^3 + |u_2|^3\big) |u_2 - u_1| \\
&\leqslant \Big(\frac{e^{\Delta t_0 |||B|||} - 1}{\Delta t_0} e^{\Delta t_0} + C(\Delta t_0) \Big) \big(1 + \|x_1\|^3 + \|x_2\|^3\big) \|x_2 - x_1\|.
\end{aligned}$$

Let us now prove the error estimate (5.32). Using the identities (5.4) and (5.43), for all $\Delta t \in (0, \Delta t_0)$ and $x = (u, v) \in \mathbb{R}^2$, one has

$$\|\psi_{\Delta t}(x) - F(x)\| \leqslant \|\psi_{\Delta t}^{\mathrm{L}}(\phi_{\Delta t}^{\mathrm{NL}}(x)) - F^{\mathrm{L}}(x)\| + \|\psi_{\Delta t}^{\mathrm{NL}}(x) - F^{\mathrm{NL}}(x)\|.$$

On the one hand, using the inequality (5.44), the expressions of the linear mappings F^{L} and $\psi_{\Delta t}^{\mathrm{L}}$ and the definition of $\psi_{\Delta t}^{\mathrm{NL}}$, one has

$$\begin{aligned}
\|\psi_{\Delta t}^{\mathrm{L}}(\phi_{\Delta t}^{\mathrm{NL}}(x)) - F^{\mathrm{L}}(x)\| &\leqslant \|\psi_{\Delta t}^{\mathrm{L}}(\phi_{\Delta t}^{\mathrm{NL}}(x)) - F^{\mathrm{L}}(\phi_{\Delta t}^{\mathrm{NL}}(x))\| + \|F^{\mathrm{L}}(\phi_{\Delta t}^{\mathrm{NL}}(x)) - F^{\mathrm{L}}(x)\| \\
&\leqslant |||\frac{e^{\Delta t B} - I - \Delta t B}{\Delta t}||| \|\phi_{\Delta t}^{\mathrm{NL}}(x)\| + \Delta t |||B||| \|\psi_{\Delta t}^{\mathrm{NL}}(x)\|.
\end{aligned}$$

Note that $\phi_{\Delta t}^{\mathrm{NL}}(0) = (\phi_{\Delta t}^{\mathrm{AC}}(0), \beta \Delta t) = (0, \beta \Delta t)$ and $\psi_{\Delta t}^{\mathrm{NL}}(0) = (\psi_{\Delta t}^{\mathrm{AC}}(0), \beta) = (0, \beta)$. In addition, one has

$$\begin{aligned}
|||\frac{e^{\Delta t B} - I - \Delta t B}{\Delta t}||| &\leqslant \sum_{k=2}^{\infty} \frac{\Delta t^{k-1}}{k!} |||B|||^k \\
&\leqslant \Delta t \sum_{k=2}^{\infty} \frac{\Delta t_0^{k-2}}{k!} |||B|||^k \\
&= \Delta t \frac{e^{\Delta t_0 |||B|||} - 1 - \Delta t_0 |||B|||}{\Delta t_0}.
\end{aligned}$$

Therefore, using the inequalities (5.29) from Proposition 5.3.1 and (5.41) from Lemma 5.4.1, one has

$$\|\psi^{\mathrm{L}}_{\Delta t}(\phi^{\mathrm{NL}}_{\Delta t}(x)) - F^{\mathrm{L}}(x)\| \leqslant C(\Delta t_0)\Delta t(1+\|x\|^4).$$

On the other hand, using the inequality (5.42) from Lemma 5.4.1, one has

$$\|\psi^{\mathrm{NL}}_{\Delta t}(x) - F^{\mathrm{NL}}(x)\| = |\psi^{\mathrm{AC}}_{\Delta t}(u) - (u-u^3)| \leqslant C(\Delta t_0)\Delta t\big(1+|u|^5\big).$$

Gathering the estimates then gives the inequality

$$\|\psi_{\Delta t}(x) - F(x)\| \leqslant C(\Delta t_0)\Delta t(1+\|x\|^5),$$

which concludes the proof of (5.32).

It remains to prove the inequality (5.33). The proof is straightforward: using (5.43) and the equalities $\phi^{\mathrm{AC}}_{\Delta t}(0) = \psi^{\mathrm{AC}}_{\Delta t}(0) = 0$, one has

$$\psi_{\Delta t}(0) = \frac{e^{\Delta t B} - I}{\Delta t}\phi^{\mathrm{NL}}_{\Delta t}(0) + \psi^{\mathrm{NL}}_{\Delta t}(0) = e^{\Delta t B}\begin{pmatrix}0\\ \beta\end{pmatrix}.$$

Therefore one gets

$$\sup_{\Delta t\in(0,\Delta t_0)} \|\psi_{\Delta t}(0)\| \leqslant e^{\Delta t_0 |||B|||}|\beta|.$$

The proof of Proposition 5.3.2 is thus completed. □

Let us conclude this subsection with a remark concerning the order of the composition of the two subsystems to define the splitting schemes, see equation (5.20).

Remark 5.4.1. *Let $\hat\phi_{\Delta t} : \mathbb{R}^2 \to \mathbb{R}^2$ be defined as follows: for all $\Delta t \in (0, \Delta t_0)$, set*

$$\hat\phi_{\Delta t} = \phi^{\mathrm{NL}}_{\Delta t} \circ \phi^{\mathrm{L}}_{\Delta t}. \tag{5.45}$$

Compared with the definition (5.20) of $\phi_{\Delta t}$, the order of the composition of the integrators $\phi^{\mathrm{L}}_{\Delta t}$ and $\phi^{\mathrm{NL}}_{\Delta t}$ associated with the subsystems (5.13) and (5.11), respectively is reversed. Define also

$$\hat\psi_{\Delta t}(x) = \frac{\hat\phi_{\Delta t}(x) - x}{\Delta t} \tag{5.46}$$

for all $\Delta t \in (0, \Delta t_0)$ and $x \in \mathbb{R}^2$. Using the mapping $\hat\phi_{\Delta t}$, modifying the definition of the scheme (5.24) gives the alternative splitting scheme

$$\hat X_{n+1} = \mathcal{A}_{\Delta t}\hat\phi_{\Delta t}(\hat X_n) + \int_{t_n}^{t_{n+1}} \mathcal{B}_{t_{n+1}-s}\, d\mathcal{W}(s) \tag{5.47}$$

for the approximation of the stochastic evolution equation (5.9). Precisely, alternatives of the splitting schemes (5.21), (5.22) and (5.23) are obtained from the formulation (5.47). However, the analysis performed in this paper does not encompass the case of the scheme (5.47), due to missing properties for the mapping $\hat\psi_{\Delta t}$, compared with $\psi_{\Delta t}$, as explained below.

Note that the result of Proposition 5.3.1 also holds with $\phi_{\Delta t}$ replaced by $\hat{\phi}_{\Delta t}$. However, it is not clear whether the one-sided Lipschitz continuity property (5.30) from Proposition 5.3.2 holds also with $\psi_{\Delta t}$ replaced by $\hat{\psi}_{\Delta t}$ (uniformly with respect to $\Delta t \in (0, \Delta t_0)$). The proof of the inequality (5.30) exploits the global Lipschitz continuity property (5.44) of the auxiliary mapping $\psi^{\mathrm{L}}_{\Delta t}$, which is a linear mapping from $\mathbb{R}^2$ to $\mathbb{R}^2$. Instead of the identity (5.43), one has

$$\hat{\psi}_{\Delta t}(x) = \psi^{\mathrm{NL}}_{\Delta t}(\phi^{\mathrm{L}}_{\Delta t}(x)) + \psi^{\mathrm{L}}_{\Delta t}(x), \tag{5.48}$$

and since $\psi^{\mathrm{NL}}_{\Delta t}$ is not globally Lipschitz continuous uniformly with respect to $\Delta t \in (0, \Delta t_0)$, the arguments of the proof above cannot be repeated for the splitting scheme (5.47).

Now, moment bounds for the solutions of the stochastic evolution equations (5.9) and (5.26) are obtained.

Let us first state the moment bounds for the stochastic convolution defined by (5.7).

Lemma 5.4.2. *Let $\big(\mathcal{Z}(t)\big)_{t\geqslant 0}$ be defined by (5.7). For all $T \in (0, \infty)$ and $p \in [1, \infty)$, one has*

$$\sup_{0\leqslant t\leqslant T} \mathbb{E}[\|\mathcal{Z}(t)\|^p_{\mathcal{E}}] < \infty.$$

Proof. Let us only provide the sketch of the proof. To deal with homogeneous Neumann boundary conditions, it is convenient to introduce $Z_0(t) = \langle Z(t), e_0\rangle e_0 = \beta_0(t)e_0$ and $Z_\perp(t) = Z(t) - Z_0(t)$ for all $t \geqslant 0$. Let also $\mathcal{Z}_0(t) = \begin{pmatrix} Z_0(t) \\ 0 \end{pmatrix}$ and $\mathcal{Z}_\perp(t) = \mathcal{Z}(t) - \mathcal{Z}_0(t)$. On the one hand, one has

$$\sup_{0\leqslant t\leqslant T} \mathbb{E}[\|\mathcal{Z}_0(t)\|^p_{\mathcal{E}}] = \sup_{0\leqslant t\leqslant T} \mathbb{E}[\|Z_0(t)\|^p_{E}] \leqslant \sup_{0\leqslant t\leqslant T} \mathbb{E}[|\beta_0(t)|^p]\|e_0\|^p_E \leqslant CT^{\frac{p}{2}}.$$

On the other hand, applying the temporal and spatial increment bounds [102, Lemma 5.21] and the Kolmogorov regularity criterion [161, Theorem C.6] gives

$$\sup_{0\leqslant t\leqslant T} \mathbb{E}[\|\mathcal{Z}_\perp(t)\|^p_{\mathcal{E}}] = \sup_{0\leqslant t\leqslant T} \mathbb{E}[\|Z_\perp(t)\|^p_{E}] \leqslant C(T) < \infty.$$

Combining the moment bounds for $\mathcal{Z}_0(t)$ and $\mathcal{Z}_\perp(t)$ then concludes the proof of Lemma 5.4.2. □

Well-posedness and moment bounds properties are provided, first for the solutions to the stochastic FitzHugh–Nagumo SPDE system (5.8), second for the solutions to the auxiliary SPDE (5.9).

Proposition 5.4.2. *For any initial value $x_0 \in \mathcal{H}$, the stochastic evolution equation (5.9) admits a unique global mild solution $\big(X(t)\big)_{t\geqslant 0}$, in the sense that (5.10) is satisfied. Moreover, for all $T \in (0, \infty)$ and all $p \in [1, \infty)$, there exists $C_p(T) \in (0, \infty)$ such that for all $x_0 \in \mathcal{E}$ one has*

$$\sup_{0\leqslant t\leqslant T} \mathbb{E}[\|X(t)\|^p_{\mathcal{E}}] \leqslant C_p(T)\big(1 + \|x_0\|^p_{\mathcal{E}}\big). \tag{5.49}$$

Proposition 5.4.3. *For any initial value $x_0 \in \mathcal{H}$ and for all $\Delta t \in (0, \Delta t_0)$, the stochastic evolution equation (5.26) admits a unique global mild solution $\big(X_{\Delta t}(t)\big)_{t \geqslant 0}$, in the sense that*

$$X_{\Delta t}(t) = e^{-t\Lambda} x_0 + \int_0^t e^{-(t-s)\Lambda} \psi_{\Delta t}(X_{\Delta t}(s))\, ds + \int_0^t e^{-(t-s)\Lambda}\, d\mathcal{W}(s) \tag{5.50}$$

is satisfied for all $t \geqslant 0$. Moreover, for all $T \in (0, \infty)$ and all $p \in [1, \infty)$, there exists $C_p(T, \Delta t_0) \in (0, \infty)$ such that for all $x_0 \in \mathcal{E}$ one has

$$\sup_{\Delta t \in (0, \Delta t_0)} \sup_{0 \leqslant t \leqslant T} \mathbb{E}[\|X_{\Delta t}(t)\|_{\mathcal{E}}^p] \leqslant C_p(T) \big(1 + \|x_0\|_{\mathcal{E}}^p\big). \tag{5.51}$$

The detailed proofs of Propositions 5.4.2 and 5.4.3 are omitted. However let us emphasize that the main arguments used in the proofs are, on the one hand, the one-sided Lipschitz continuity properties (5.5) and (5.30) of F and $\psi_{\Delta t}$, respectively, and on the other hand, the moment bounds on $\mathcal{Z}(t)$ from Lemma 5.4.2. Observe that the mapping $\psi_{\Delta t}$ is globally Lipschitz continuous for any $\Delta t > 0$, therefore the existence and uniqueness of the mild solution $\big(X_{\Delta t}(t)\big)_{t \geqslant 0}$ satisfying (5.50) follows from standard fixed point arguments, see for instance [102, Theorem 7.5]. The proof of the moment bounds (5.51) requires some care: indeed, one needs to obtain upper bounds which are uniform with respect to $\Delta t \in (0, \Delta t_0)$, and applying [102, Theorem 7.5] would not be appropriate since the Lipschitz constant of $\psi_{\Delta t}$ is unbounded for $\Delta t \in (0, \Delta t_0)$. Introducing $Y_{\Delta t}(t) = X_{\Delta t}(t) - \mathcal{Z}(t)$, one obtains the moment bounds (5.51) using the one-sided Lipschitz continuity property (5.30) from Proposition 5.3.2, which is uniform with respect to $\Delta t \in (0, \Delta t_0)$. Similar arguments are used to prove Proposition 5.4.2. Propositions 5.4.2 and 5.4.3 are variants of [46, Propositions 1 and 2] for the analysis of the stochastic Allen–Cahn equation and one refers to [83, Proposition 6.2.2] for a more general version. Some arguments need to be adapted since the considered systems (5.9) and (5.26) are not parabolic systems.

Finally, let us state the following result which is required in Section 5.5 below.

Lemma 5.4.3. *For all $T \in (0, \infty)$, $p \in [1, \infty)$ and $\alpha \in [0, \frac{1}{4})$, there exists $C_{\alpha,p}(T) \in (0, \infty)$ such that for all $x_0 = (u_0, v_0) \in \mathcal{H}^{2\alpha} \cap \mathcal{E}$, all $\Delta t \in (0, \Delta t_0)$ and $t_1, t_2 \in [0, T]$, one has*

$$\big(\mathbb{E}[\|X_{\Delta t}(t_2) - X_{\Delta t}(t_1)\|_{\mathcal{H}}^p]\big)^{\frac{1}{p}} \leqslant C_{\alpha,p}(T) |t_2 - t_1|^{\alpha} \big(1 + \|(-\Delta)^{\alpha} u_0\|_H^4 + \|x_0\|_{\mathcal{E}}^4\big). \tag{5.52}$$

Proof. Let $0 \leqslant t_1 < t_2 \leqslant T$, using the mild form (5.50) of the auxiliary stochastic evolution equation, one obtains the estimate

$$\begin{aligned}
\big(\mathbb{E}[\|X_{\Delta t}(t_2) - X_{\Delta t}(t_1)\|_{\mathcal{H}}^p]\big)^{\frac{1}{p}} &\leqslant \|e^{-t_2\Lambda} x_0 - e^{-t_1\Lambda} x_0\|_{\mathcal{H}} + \big(\mathbb{E}\big[\|\mathcal{Z}(t_2) - \mathcal{Z}(t_1)\|_{\mathcal{H}}^p\big]\big)^{\frac{1}{p}} \\
&\quad + \int_0^{t_1} \big(\mathbb{E}\big[\big\| \big(e^{-(t_2-s)\Lambda} - e^{-(t_1-s)\Lambda}\big) \psi_{\Delta t}(X_{\Delta t}(s))\big\|_{\mathcal{H}}^p\big]\big)^{\frac{1}{p}} \, ds \\
&\quad + \int_{t_1}^{t_2} \big(\mathbb{E}\big[\|e^{-(t_2-s)\Lambda} \psi_{\Delta t}(X_{\Delta t}(s))\|_{\mathcal{H}}^p\big]\big)^{\frac{1}{p}} \, ds,
\end{aligned}$$

where as the reader recalls, $\mathcal{Z}(t)$ denotes the stochastic convolution (5.7).

The first term on the right-hand side is estimated using the inequality (5.36) in order to get

$$\|e^{-t_2\Lambda}x_0 - e^{-t_1\Lambda}x_0\|_{\mathcal{H}} \leqslant |t_2 - t_1|^{\alpha}\|(-\Lambda)^{\alpha}x_0\|_{\mathcal{H}}.$$

The second term corresponds to the temporal regularity of the stochastic convolution

$$\left(\mathbb{E}\left[\|\mathcal{Z}(t_2) - \mathcal{Z}(t_1)\|_{\mathcal{H}}^p\right]\right)^{\frac{1}{p}} \leqslant |t_2 - t_1|^{\alpha}.$$

This is obtained combining the proofs of Lemma 5.4.2 and of [38, Theorem 4.4].

The last two terms are estimated using the polynomial growth

$$\|\psi_{\Delta t}(x)\| \leqslant C(\Delta t_0)\left(1 + \|x\|\right)^4,$$

see equations (5.31) and(5.33) in Proposition 5.3.2. Indeed, one has

$$\begin{aligned}\|\left(e^{-(t_2-s)\Lambda} - e^{-(t_1-s)\Lambda}\right)\psi_{\Delta t}(X_{\Delta t}(s))\|_{\mathcal{H}} &\leqslant C_{\alpha}\frac{|t_2 - t_1|^{\alpha}}{|t_1 - s|^{\alpha}}\|\psi_{\Delta t}(X_{\Delta t}(s))\|_{\mathcal{H}} \\ &\leqslant C_{\alpha}(\Delta t_0)\frac{|t_2 - t_1|^{\alpha}}{|t_1 - s|^{\alpha}}\left(1 + \|X_{\Delta t}(s)\|_{\mathcal{E}}^4\right)\end{aligned}$$

and

$$\|e^{-(t_2-s)\Lambda}\psi_{\Delta t}(X_{\Delta t}(s))\|_{\mathcal{H}} \leqslant \left(1 + \|X_{\Delta t}(s)\|_{\mathcal{E}}^4\right)$$

for the last term. One concludes the proof using the moment bounds of the solution of the auxiliary stochastic evolution equation, see Proposition 5.4.3. □

5.5 Proofs of the main results

In this section, the detailed proofs for the main results of the present work are presented. At first, the moment bounds for the three splitting schemes (Theorem 5.3.1) will be provided and then the strong error estimates with rate of convergence at least 1/4 (Theorem 5.3.2) will be proved.

Proof of Theorem 5.3.1

The proof of the moment bounds (5.27) given below is inspired by the proof of [46, Proposition 3] and requires some auxiliary tools.

Given the time-step size $\Delta t \in (0, \Delta t_0)$, introduce the auxiliary scheme $\left(\mathcal{Z}_n\right)_{n\geqslant 0}$ defined as follows: for all $n \geqslant 0$,

$$\mathcal{Z}_{n+1} = \mathcal{A}_{\Delta t}\mathcal{Z}_n + \int_{t_n}^{t_{n+1}} \mathcal{B}_{t_{n+1}-s}\,\mathrm{d}\mathcal{W}(s) \tag{5.53}$$

with initial value $\mathcal{Z}_0 = 0$, using the same notation as for the general expression (5.24) of the three splitting schemes (5.21), (5.22) and (5.23). One has the following moment bounds for the solution of the scheme (5.53). Recall that one has $T = N\Delta t$ for some integer $N \in \mathbb{N}$.

Lemma 5.5.1. *For all $T \in (0,\infty)$ and $p \in [1,\infty)$, one has*

$$\sup_{\Delta t\in(0,\Delta t_0)} \sup_{0\leqslant n\leqslant N} \mathbb{E}[\|\mathcal{Z}_n\|_{\mathcal{E}}^p] < \infty. \tag{5.54}$$

Lemma 5.5.1 is a variant of [46, Lemma 3.5], using the same arguments as in the sketch of proof of Lemma 5.4.2 above. The proof of Lemma 5.5.1 is therefore omitted.

The proof of Theorem 5.3.1 will now be given.

Proof of Theorem 5.3.1. For all $n \in \{0,\dots,N\}$, set

$$r_n = X_n - \mathcal{Z}_n. \tag{5.55}$$

Using the definitions (5.24) and (5.53) and the definition (5.25) of the mapping $\psi_{\Delta t}$, for all $n \in \{0,\dots,N-1\}$, one has

$$\begin{aligned} r_{n+1} = X_{n+1} - \mathcal{Z}_{n+1} &= \mathcal{A}_{\Delta t}\Big(\phi_{\Delta t}(X_n) - \mathcal{Z}_n\Big) \\ &= \mathcal{A}_{\Delta t}\Big(\phi_{\Delta t}(r_n + \mathcal{Z}_n) - \phi_{\Delta t}(\mathcal{Z}_n)\Big) + \Delta t \mathcal{A}_{\Delta t}\psi_{\Delta t}(\mathcal{Z}_n). \end{aligned}$$

On the one hand, using the inequalities (5.34) and (5.37) and the global Lipschitz continuity property (5.29) of $\phi_{\Delta t}$ (see Proposition 5.3.1), one has

$$\|\mathcal{A}_{\Delta t}\Big(\phi_{\Delta t}(r_n+\mathcal{Z}_n)-\phi_{\Delta t}(\mathcal{Z}_n)\Big)\|_{\mathcal{E}} \leqslant \|\phi_{\Delta t}(r_n+\mathcal{Z}_n)-\phi_{\Delta t}(\mathcal{Z}_n)\|_{\mathcal{E}} \leqslant e^{\Delta t(1+|||B|||)}\|r_n\|_{\mathcal{E}}.$$

On the other hand, using the inequalities (5.34) and (5.37), the local Lipschitz continuity property (5.31) of $\psi_{\Delta t}$ (see Proposition 5.3.2) and the upper bound (5.33), one has

$$\|\mathcal{A}_{\Delta t}\psi_{\Delta t}(\mathcal{Z}_n)\|_{\mathcal{E}} \leqslant C(\Delta t_0)\big(1 + \|\mathcal{Z}_n\|_{\mathcal{E}}^4\big).$$

Therefore one obtains the following inequality

$$\|r_{n+1}\|_{\mathcal{E}} \leqslant e^{\Delta t(1+|||B|||)}\|r_n\|_{\mathcal{E}} + C(\Delta t_0)\big(1 + \|\mathcal{Z}_n\|_{\mathcal{E}}^4\big),$$

and by a straightforward argument, using the fact that $N\Delta t = T$, one has the estimate:

$$\|r_n\|_{\mathcal{E}} \leqslant C(T,\Delta t_0)\Big(\|r_0\|_{\mathcal{E}} + \sum_{k=0}^{n-1}\big(1 + \|\mathcal{Z}_k\|_{\mathcal{E}}^4\big)\Big),$$

for all $n \in \{0,\dots,N\}$.

Finally, for all $p \in [1,\infty)$, using the moment bound (5.54) from Lemma 5.5.1, one obtains for all $n \in \{0,\dots,N\}$

$$\big(\mathbb{E}[\|r_n\|_{\mathcal{E}}^p]\big)^{\frac{1}{p}} \leqslant C(T,\Delta t_0)\Big(\|r_0\|_{\mathcal{E}} + \sum_{k=0}^{n-1}\big(1+\big(\mathbb{E}[\|\mathcal{Z}_k\|_{\mathcal{E}}^{4p}]\big)^{\frac{1}{p}}\big)\Big) \leqslant C_p(T,\Delta t_0)\Big(\|r_0\|_{\mathcal{E}}+1\Big).$$

Since $X_n = r_n + \mathcal{Z}_n$ owing to (5.55), using the moment bound above and the moment bound (5.54) from Lemma 5.5.1 then concludes the proof of the moment bound (5.27). The proof of Theorem 5.3.1 is thus completed. □

Proof of Theorem 5.3.2

Recall that the numerical scheme is given by (5.24). It is straightforward to check that for all $n \geqslant 0$ one has

$$X_n = \mathcal{A}_{\Delta t}^n x_0 + \Delta t \sum_{k=0}^{n-1} \mathcal{A}_{\Delta t}^{n-k} \psi_{\Delta t}(X_k) + \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathcal{A}_{\Delta t}^{n-k-1} \mathcal{B}_{t_{k+1}-s} \, \mathrm{d}\mathcal{W}(s). \quad (5.56)$$

Let us introduce the auxiliary process $\big(X_n^{\mathrm{aux}}\big)_{n\geqslant 0}$ which is defined as follows: for all $n \geqslant 0$ one has

$$X_n^{\mathrm{aux}} = \mathcal{A}_{\Delta t}^n x_0 + \Delta t \sum_{k=0}^{n-1} \mathcal{A}_{\Delta t}^{n-k} \psi_{\Delta t}(X_{\Delta t}(t_k)) + \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathcal{A}_{\Delta t}^{n-k-1} \mathcal{B}_{t_{k+1}-s} \, \mathrm{d}\mathcal{W}(s), \quad (5.57)$$

where, as the reader can recall, $t_k = k\Delta t$ and $\big(X_{\Delta t}(t)\big)_{t\geqslant 0}$ is the unique mild solution of the auxiliary stochastic evolution equation (5.26). Note that for all $n \geqslant 0$ one has

$$X_{n+1}^{\mathrm{aux}} = \mathcal{A}_{\Delta t} X_n^{\mathrm{aux}} + \Delta t \mathcal{A}_{\Delta t} \psi_{\Delta t}(X_{\Delta t}(t_n)) + \int_{t_n}^{t_{n+1}} \mathcal{B}_{t_{n+1}-s} \, \mathrm{d}\mathcal{W}(s). \quad (5.58)$$

Lemma 5.5.2. *For all $T \in (0,\infty)$ and $p \in [1,\infty)$, there exists $C_p(T) \in (0,\infty)$ such that for all $x_0 \in \mathcal{E}$ one has*

$$\sup_{\Delta t\in(0,\Delta t_0)} \sup_{0\leqslant n\leqslant N} \mathbb{E}[\|X_n^{\mathrm{aux}}\|_{\mathcal{E}}^p] \leqslant C_p(T)\big(1 + \|x_0\|_{\mathcal{E}}^p\big). \quad (5.59)$$

Proof of Lemma 5.5.2. Using the discrete mild formulation (5.57) of X_n^{aux}, the inequalities (5.34) and (5.37), the local Lipschitz continuity property (5.31) of $\psi_{\Delta t}$ and the upper bound (5.33) (see Proposition 5.3.2), for all $\Delta t \in (0, \Delta t_0)$ and $n \geqslant 0$ one has

$$\|X_n^{\mathrm{aux}}\|_{\mathcal{E}} \leqslant \|x_0\|_{\mathcal{E}} + C(\Delta t_0)\Delta t \sum_{k=0}^{n-1} \big(1 + \|X_{\Delta t}(t_k)\|_{\mathcal{E}}^4\big) + \|\mathcal{Z}_n\|_{\mathcal{E}}.$$

It suffices to use the moment bounds (5.51) for the auxiliary process $X_{\Delta t}$ from Proposition 5.4.3 and (5.54) for the Gaussian random variables $\mathcal{Z}_n$ from Lemma 5.5.1, and the Minkowskii inequality, to conclude the proof of the moment bounds (5.59). The proof of Lemma 5.5.2 is thus completed. □

Observe that for all $n \in \{0,\dots,N\}$ the error $X(t_n) - X_n$ can be decomposed as follows:

$$X(t_n) - X_n = X(t_n) - X_{\Delta t}(t_n) + X_{\Delta t}(t_n) - X_n^{\mathrm{aux}} + X_n^{\mathrm{aux}} - X_n. \quad (5.60)$$

In order to prove Theorem 5.3.2, it suffices to prove error bounds for the three error terms appearing in the right-hand side of (5.60). They are given in Lemma 5.5.3, Lemma 5.5.4 and Lemma 5.5.5 respectively. The proofs of these technical lemmas are presented at the end of the section.

Lemma 5.5.3. *For all $T \in (0,\infty)$ and $p \in [1,\infty)$, there exists $C_p(T,\Delta t_0) \in (0,\infty)$ such that for all $x_0 \in \mathcal{E}$ and all $\Delta t \in (0,\Delta t_0)$, one has*

$$\sup_{t\in[0,T]} \big(\mathbb{E}[\|X(t) - X_{\Delta t}(t)\|_{\mathcal{H}}^p]\big)^{\frac{1}{p}} \leqslant C_p(T,\Delta t_0)\Delta t\big(1 + \|x_0\|_{\mathcal{E}}^5\big). \tag{5.61}$$

Lemma 5.5.4. *For all $T \in (0,\infty)$, $p \in [1,\infty)$ and $\alpha \in [0,\frac{1}{4})$, there exists $C_{\alpha,p}(T) \in (0,\infty)$ such that for all $x_0 = (u_0,v_0) \in \mathcal{H}^{2\alpha} \cap \mathcal{E}$, all $\Delta t \in (0,\Delta t_0)$, one has*

$$\sup_{0\leqslant n\leqslant N} \big(\mathbb{E}[\|X_{\Delta t}(t_n) - X_n^{\mathrm{aux}}\|_{\mathcal{H}}^p]\big)^{\frac{1}{p}} \leqslant C_{\alpha,p}(T)\Delta t^{\alpha}\big(1 + \|(-\Delta)^{\alpha}u_0\|_H^7 + \|x_0\|_{\mathcal{E}}^7\big). \tag{5.62}$$

Lemma 5.5.5. *For all $T \in (0,\infty)$, $p \in [1,\infty)$ and $\alpha \in [0,\frac{1}{4})$, there exists $C_{\alpha,p}(T) \in (0,\infty)$ such that for all $x_0 = (u_0,v_0) \in \mathcal{H}^{2\alpha} \cap \mathcal{E}$, all $\Delta t \in (0,\Delta t_0)$, one has*

$$\sup_{0\leqslant n\leqslant N} \big(\mathbb{E}[\|X_n^{\mathrm{aux}} - X_n\|_{\mathcal{H}}^p]\big)^{\frac{1}{p}} \leqslant C_{\alpha,p}(T)\Delta t^{\alpha}\big(1 + \|(-\Delta)^{\alpha}u_0\|_H^7 + \|x_0\|_{\mathcal{E}}^7\big). \tag{5.63}$$

With the auxiliary error estimates given above, it is straightforward to give the proof of Theorem 5.3.2.

Proof of Theorem 5.3.2. Using the decomposition of the error (5.60), using the Minkowskii inequality and the error estimates (5.61), (5.62) and (5.63), one obtains the following result: for all $\alpha \in [0,\frac{1}{4})$ and $p \in [1,\infty)$, there exists $C_{\alpha,p} \in (0,\infty)$ such that for all $\Delta t \in (0,\Delta t_0)$ one has

$$\begin{aligned}
\sup_{0\leqslant n\leqslant N} \big(\mathbb{E}[\|X(t_n) - X_n\|_{\mathcal{H}}^p]\big)^{\frac{1}{p}} &\leqslant \sup_{0\leqslant n\leqslant N} \big(\mathbb{E}[\|X(t_n) - X_{\Delta t}(t_n)\|_{\mathcal{H}}^p]\big)^{\frac{1}{p}} \\
&\quad + \sup_{0\leqslant n\leqslant N} \big(\mathbb{E}[\|X_{\Delta t}(t_n) - X_n^{\mathrm{aux}}\|_{\mathcal{H}}^p]\big)^{\frac{1}{p}} \\
&\quad + \sup_{0\leqslant n\leqslant N} \big(\mathbb{E}[\|X_n^{\mathrm{aux}} - X_n\|_{\mathcal{H}}^p]\big)^{\frac{1}{p}} \\
&\leqslant C_p(T,\Delta t_0)\Delta t\big(1 + \|x_0\|_{\mathcal{E}}^5\big) \\
&\quad + C_{\alpha,p}(T)\Delta t^{\alpha}\big(1 + \|(-\Delta)^{\alpha}u_0\|_H^7 + \|x_0\|_{\mathcal{E}}^7\big) \\
&\quad + C_{\alpha,p}(T)\Delta t^{\alpha}\big(1 + \|(-\Delta)^{\alpha}u_0\|_H^7 + \|x_0\|_{\mathcal{E}}^7\big) \\
&\leqslant C_{\alpha,p}(T)\Delta t^{\alpha}\big(1 + \|(-\Delta)^{\alpha}u_0\|_H^7 + \|x_0\|_{\mathcal{E}}^7\big).
\end{aligned}$$

This concludes the proof of the inequality (5.28) and the proof of Theorem 5.3.2 is thus completed. □

Let us now give the proofs of the auxiliary error estimates. Note that the proof of Lemma 5.5.5 requires the error estimate (5.62) from Lemma 5.5.4.

Proof of Lemma 5.5.3. For all $t \geqslant 0$ and $\Delta t \in (0,\Delta t_0)$, set

$$R_{\Delta t}(t) = X_{\Delta t}(t) - X(t).$$

The auxiliary process $\big(R_{\Delta t}(t)\big)_{t\geqslant 0}$ is the unique solution of the evolution equation

$$\frac{\mathrm{d}R_{\Delta t}(t)}{\mathrm{d}t} = -\Lambda R_{\Delta t}(t) + \psi_{\Delta t}(X_{\Delta t}(t)) - \psi_{\Delta t}(X(t)) + \psi_{\Delta t}(X(t)) - F(X(t))$$

with the initial value $R_{\Delta t}(0) = 0$. Therefore one obtains, almost surely, for all $t \geqslant 0$

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}\|R_{\Delta t}(t)\|_{\mathcal{H}}^2}{\mathrm{d}t} &= \langle R_{\Delta t}(t), -\Lambda R_{\Delta t}(t)\rangle_{\mathcal{H}} + \langle R_{\Delta t}(t), \psi_{\Delta t}(X_{\Delta t}(t)) - \psi_{\Delta t}(X(t))\rangle_{\mathcal{H}} \\ &\quad + \langle R_{\Delta t}(t), \psi_{\Delta t}(X(t)) - F(X(t))\rangle_{\mathcal{H}}.\end{aligned}$$

First, one has

$$\langle R_{\Delta t}(t), -\Lambda R_{\Delta t}(t)\rangle_{\mathcal{H}} \leqslant 0.$$

Second, using the one-sided Lipschitz continuity property (5.30) from Proposition 5.3.2 for $\psi_{\Delta t}$ (uniformly with respect to $\Delta t \in (0, \Delta t_0)$), one has

$$\langle R_{\Delta t}(t), \psi_{\Delta t}(X_{\Delta t}(t)) - \psi_{\Delta t}(X(t))\rangle_{\mathcal{H}} \leqslant C(\Delta t_0)\|R_{\Delta t}(t)\|_{\mathcal{H}}^2.$$

Finally, using the Cauchy–Schwarz and Young inequalities and the error estimate (5.32) from Proposition 5.3.2 , one has

$$\begin{aligned}\langle R_{\Delta t}(t), \psi_{\Delta t}(X(t)) - F(X(t))\rangle_{\mathcal{H}} &\leqslant \|R_{\Delta t}(t)\|_{\mathcal{H}}\|\psi_{\Delta t}(X(t)) - F(X(t))\|_{\mathcal{H}} \\ &\leqslant \frac{1}{2}\|R_{\Delta t}(t)\|_{\mathcal{H}}^2 + \frac{1}{2}\|\psi_{\Delta t}(X(t)) - F(X(t))\|_{\mathcal{H}}^2 \\ &\leqslant \frac{1}{2}\|R_{\Delta t}(t)\|_{\mathcal{H}}^2 + C(\Delta t_0)\Delta t^2\big(1 + \|X(t)\|_{\mathcal{E}}^{10}\big).\end{aligned}$$

Gathering the upper bounds above and using Gronwall's lemma, one obtains, almost surely, for all $t \in [0, T]$

$$\|R_{\Delta t}(t)\|_{\mathcal{H}}^2 \leqslant C(T, \Delta t_0)\Delta t^2 \int_0^T \big(1 + \|X(s)\|_{\mathcal{E}}^{10}\big)\,\mathrm{d}s.$$

Using the moment bound (5.49) from Proposition 5.4.2, one then obtains for all $t \in [0, T]$ and all $p \in [2, \infty)$

$$\begin{aligned}\big(\mathbb{E}[\|R_{\Delta t}(t)\|_{\mathcal{H}}^p]\big)^{\frac{2}{p}} &\leqslant C(T, \Delta t_0)\Delta t^2 \int_0^T \big(1 + \mathbb{E}[\|X(s)\|_{\mathcal{E}}^{5p}]^{\frac{2}{p}}\big)\,\mathrm{d}s \\ &\leqslant C(T, \Delta t_0)\Delta t^2\big(1 + \sup_{s\in[0,T]} \mathbb{E}[\|X(s)\|_{\mathcal{E}}^{5p}]^{\frac{2}{p}}\big) \\ &\leqslant C_p(T, \Delta t_0)\Delta t^2\big(1 + \|x_0\|_{\mathcal{E}}^{10}\big).\end{aligned}$$

This estimate has been proved for $p \in [2, \infty)$, however it is also valid for $p \in [1, 2)$. This concludes the proof of the error estimate (5.61) and of Lemma 5.5.3. □

In order to prove Lemma 5.5.4, let us recall the following useful standard inequality:

Proposition 5.5.1. *For all $n \in \mathbb{N}$ and $z \in [0,\infty)$, then*

$$\sup_{n\in\mathbb{N},z\in[0,\infty)} n\Big|\frac{1}{(1+z)^n} - e^{-nz}\Big| + \sup_{n\in\mathbb{N},z\in[0,\infty)} \frac{\big|\frac{1}{(1+z)^n} - e^{-nz}\big|}{\min(1,z)} < \infty. \tag{5.64}$$

Proof. Let us first state two elementary inequalities:

- for all $0 \leqslant a \leqslant b$ and $n \in \mathbb{N}$, one has $0 \leqslant b^n - a^n \leqslant nb^{n-1}(b-a)$,
- for all $z \in [0,\infty)$, one has $0 \leqslant \frac{1}{1+z} - e^{-z} \leqslant C\min(1,z^2)$.

As a consequence, for all $n \in \mathbb{N}$ and $z \in [0,\infty)$ one has

$$0 \leqslant \frac{1}{(1+z)^n} - e^{-nz} \leqslant \frac{n}{(1+z)^{n-1}}\Big(\frac{1}{1+z} - e^{-z}\Big).$$

For all $n \geqslant 3$ and $z \in [0,\infty)$, one has

$$\begin{aligned} n\Big|\frac{1}{(1+z)^n} - e^{-nz}\Big| &\leqslant \frac{Cn^2z^2}{(1+z)^{n-1}} \\ &\leqslant \frac{Cn^2z^2}{1+(n-1)z+\frac{(n-1)(n-2)}{2}z^2} \\ &\leqslant \frac{2Cn^2}{(n-1)(n-2)} \\ &\leqslant C. \end{aligned}$$

The cases $n = 1$ and $n = 2$ are treated separately, one has

$$\sup_{z\in[0,\infty)} \Big|\frac{1}{(1+z)} - e^{-z}\Big| + \sup_{z\in[0,\infty)} 2\Big|\frac{1}{(1+z)^2} - e^{-2z}\Big| < \infty.$$

This concludes the proof of the first inequality. To prove the second inequality, observe first that one has

$$\sup_{n\in\mathbb{N},z\in[0,\infty)} \Big|\frac{1}{(1+z)^n} - e^{-nz}\Big| \leqslant 2.$$

In addition, for all $n \geqslant 2$ and $z \in [0,\infty)$, one has

$$\frac{\big|\frac{1}{(1+z)^n} - e^{-nz}\big|}{z} \leqslant \frac{Cnz}{(1+z)^{n-1}} \leqslant \frac{Cnz}{1+(n-1)z} \leqslant \frac{Cn}{n-1} \leqslant C.$$

The case $n = 1$ is treated separately: using the inequality $\min(1,z^2) \leqslant z$ one has

$$\sup_{z\in[0,\infty)} \frac{\big|\frac{1}{1+z} - e^{-z}\big|}{z} \leqslant C.$$

Gathering the results concludes the proof of the second inequality. □

Furthermore, it is straightforward to prove that for all $\alpha \in [0,1]$, one has $\min(1,z) \leqslant z^\alpha$.

Proof of Lemma 5.5.4. Using the mild formulations (5.50) for $X_{\Delta t}(t_n)$ and (5.57) for X_n^{aux}, one obtains the following decomposition of the error: for all $n \geqslant 0$, one has

$$X_{\Delta t}(t_n) - X_n^{\mathrm{aux}} = E_n^{\Delta t,1} + E_n^{\Delta t,2} + E_n^{\Delta t,3} + E_n^{\Delta t,4} + E_n^{\Delta t,5}, \tag{5.65}$$

where

$$E_n^{\Delta t,1} = (e^{-n\Delta t\Lambda} - \mathcal{A}_{\Delta t}^n)x_0 \tag{5.66}$$

$$E_n^{\Delta t,2} = \mathcal{Z}(t_n) - \mathcal{Z}_n \tag{5.67}$$

$$E_n^{\Delta t,3} = \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} e^{-(t_n-s)\Lambda}\big(\psi_{\Delta t}(X_{\Delta t}(s)) - \psi_{\Delta t}(X_{\Delta t}(t_k))\big)\,\mathrm{d}s \tag{5.68}$$

$$E_n^{\Delta t,4} = \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \big(e^{-(t_n-s)\Lambda} - e^{-(t_n-t_k)\Lambda}\big)\psi_{\Delta t}(X_{\Delta t}(t_k))\,\mathrm{d}s \tag{5.69}$$

$$E_n^{\Delta t,5} = \Delta t \sum_{k=0}^{n-1} \big(e^{-(t_n-t_k)\Lambda} - \mathcal{A}_{\Delta t}^{n-k}\big)\psi_{\Delta t}(X_{\Delta t}(t_k)). \tag{5.70}$$

Let us now give estimates for those five error terms.

- If the splitting schemes (5.21) and (5.22) are considered, one has $\mathcal{A}_{\Delta t} = e^{-\Delta t\Lambda}$ and thus $E_n^{\Delta t,1} = 0$ for all $n \geqslant 0$. If the splitting scheme (5.23) is considered, one has $\mathcal{A}_{\Delta t} = (I + \Delta t\Lambda)^{-1}$, thus using the inequality (5.64), for all $n \in \{0, \dots, N\}$, one has

$$\begin{aligned}
\|E_n^{\Delta t,1}\|_{\mathcal{H}}^2 &= \|\big(e^{n\Delta t\Delta} - ((I - \Delta t\Delta)^{-1})^n\big)u_0\|_H^2 \\
&= \sum_{j=1}^{\infty}\Big(\frac{1}{(1+\Delta t\lambda_j)^n} - e^{-n\Delta t\lambda_j}\Big)^2\langle u_0, e_j\rangle_H^2 \\
&\leqslant C_\alpha \sum_{j=1}^{\infty}(\Delta t\lambda_j)^{2\alpha}\langle u_0, e_j\rangle_H^2 \\
&\leqslant C_\alpha \Delta t^{2\alpha}\|(-\Delta)^\alpha u_0\|_H^2.
\end{aligned}$$

  Therefore one obtains the following upper bound: for all $\alpha \in [0, \frac{1}{4})$, there exists $C_\alpha \in (0, \infty)$ such that for all $\Delta t \in (0, \Delta t_0)$ one has

$$\sup_{0\leqslant n\leqslant N}\big(\mathbb{E}[\|E_n^{\Delta t,1}\|_{\mathcal{H}}^p]\big)^{\frac{1}{p}} \leqslant C_\alpha \Delta t^\alpha \|(-\Delta)^\alpha u_0\|_H. \tag{5.71}$$

- Note that if the splitting scheme (5.21) is considered ($X_n = X_n^{\mathrm{LT,exact}}$ for all $n \geqslant 0$), one has $E_n^{\Delta t,2} = 0$ for all $n \geqslant 0$. If the splitting schemes (5.22) and (5.23) are considered, for all $n \geqslant 0$ one has

$$E_n^{\Delta t,2} = \mathcal{Z}(t_n) - \mathcal{Z}_n = \begin{pmatrix} Z(t_n) - Z_n \\ 0 \end{pmatrix},$$

with $Z_n = Z_n^{\mathrm{LT,expo}}$ (resp. $Z_n = Z_n^{\mathrm{LT,imp}}$) if the scheme (5.22) (resp. the scheme (5.23)) is considered. Here, let us denote

$$Z_{n+1}^{\mathrm{LT,expo}} = e^{\Delta t \Delta}\Big(Z_n^{\mathrm{LT,expo}} + \delta W_n\Big)$$
$$Z_{n+1}^{\mathrm{LT,imp}} = (I - \Delta t \Delta)^{-1}\Big(Z_n^{\mathrm{LT,imp}} + \delta W_n\Big).$$

One has the following mean-square error estimate, which are standard results in the analysis of numerical schemes for parabolic semilinear stochastic partial differential equations, see for instance [213, Theorem 3.2]: for all $\alpha \in [0, \frac{1}{4})$, there exists $C_\alpha \in (0, \infty)$ such that

$$\sup_{n \geqslant 0} \mathbb{E}[\|Z(t_n) - Z_n\|_H^2] \leqslant C_\alpha \Delta t^{2\alpha},$$

if $Z_n = Z_n^{\mathrm{LT,expo}}$ and $Z_n = Z_n^{\mathrm{LT,imp}}$. Since $Z(t_n) - Z_n$ is a H-valued Gaussian random variable, one obtains the following upper bound: for all $\alpha \in [0, \frac{1}{4})$ and $p \in [1, \infty)$, there exists $C_{\alpha,p} \in (0, \infty)$ such that for all $\Delta t \in (0, \Delta t_0)$ one has

$$\sup_{0 \leqslant n \leqslant N} \big(\mathbb{E}[\|E_n^{\Delta t,2}\|_{\mathcal{H}}^p]\big)^{\frac{1}{p}} \leqslant C_{\alpha,p} \Delta t^\alpha. \tag{5.72}$$

- Using the inequality (5.34) and the local Lipschitz continuity property (5.31) of $\psi_{\Delta t}$ (Proposition 5.3.2), one obtains

$$\|E_n^{\Delta t,3}\|_{\mathcal{H}} \leqslant \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \|e^{-(t_n - s)\Lambda}\big(\psi_{\Delta t}(X_{\Delta t}(s)) - \psi_{\Delta t}(X_{\Delta t}(t_k))\big)\|_{\mathcal{H}} \, \mathrm{d}s$$
$$\leqslant \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \|\big(\psi_{\Delta t}(X_{\Delta t}(s)) - \psi_{\Delta t}(X_{\Delta t}(t_k))\big)\|_{\mathcal{H}} \, \mathrm{d}s$$
$$\leqslant C(\Delta t_0) \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \big(1 + \|X_{\Delta t}(s)\|_{\mathcal{E}}^3 + \|X_{\Delta t}(t_k)\|_{\mathcal{E}}^3\big)\|X_{\Delta t}(s) - X_{\Delta t}(t_k)\|_{\mathcal{H}} \, \mathrm{d}s.$$

Using the Minkowskii and Cauchy–Schwarz inequalities, the moment bound (5.51) (Proposition 5.4.3) and the regularity estimate (5.52) (Lemma 5.4.3), one has

$$\big(\mathbb{E}[\|E_n^{\Delta t,3}\|_{\mathcal{H}}^p]\big)^{\frac{1}{p}} \leqslant C(\Delta t_0) \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \Big(1 + \sup_{r \in [t_k, t_{k+1}]} \big(\mathbb{E}[\|X_{\Delta t}(r)\|_{\mathcal{E}}^{6p}]\big)^{\frac{1}{2p}}\Big)$$
$$\times \big(\mathbb{E}[\|X_{\Delta t}(s) - X_{\Delta t}(t_k)\|_{\mathcal{H}}^{2p}]\big)^{\frac{1}{2p}} \, \mathrm{d}s$$
$$\leqslant C_{\alpha,p}(T) \Delta t^\alpha (1 + \|x_0\|_{\mathcal{E}}^3)\big(1 + \|(-\Delta)^\alpha u_0\|_H^4 + \|x_0\|_{\mathcal{E}}^4\big).$$

Therefore one obtains the following upper bound: for all $\alpha \in [0, \frac{1}{4})$, $p \in [1, \infty)$ and $T \in (0, \infty)$, there exists $C_{\alpha,p}(T) \in (0, \infty)$ such that for all $\Delta t \in (0, \Delta t_0)$ one has

$$\sup_{0 \leqslant n \leqslant N} \big(\mathbb{E}[\|E_n^{\Delta t,3}\|_{\mathcal{H}}^p]\big)^{\frac{1}{p}} \leqslant C_{\alpha,p}(T) \Delta t^\alpha \big(1 + \|(-\Delta)^\alpha u_0\|_H^7 + \|x_0\|_{\mathcal{E}}^7\big). \tag{5.73}$$

- Using the inequality (5.36) from Proposition 5.4.1 (with $\mu = \alpha \in [0,1)$ and $\nu = 0$) and the local Lipschitz continuity property (5.31) of $\psi_{\Delta t}$ combined with the bound (5.33) (Proposition 5.3.2), one has for all $s \in [t_k, t_{k+1}]$

$$\begin{aligned}\|\big(e^{-(t_n-s)\Lambda} - e^{-(t_n-t_k)\Lambda}\big)\psi_{\Delta t}(X_{\Delta t}(t_k))\|_{\mathcal{H}} &\leqslant C_\alpha \frac{|s-t_k|^\alpha}{(t_n-s)^\alpha}\|\psi_{\Delta t}(X_{\Delta t}(t_k))\|_{\mathcal{H}} \\ &\leqslant C_\alpha \frac{\Delta t^\alpha}{(t_n-s)^\alpha}\big(1+\|X_{\Delta t}(t_k)\|_{\mathcal{E}}^4\big).\end{aligned}$$

  Using the Minkoswskii inequality, the moment bounds (5.51) from Proposition 5.4.3, and the fact that $\int_0^T s^{-\alpha}\,\mathrm{d}s < \infty$ for $\alpha \in [0,1)$, one obtains the following upper bound: for all $\alpha \in [0, \frac{1}{4})$, $p \in [1,\infty)$ and $T \in (0,\infty)$, there exists $C_{\alpha,p}(T) \in (0,\infty)$ such that for all $\Delta t \in (0, \Delta t_0)$ one has

$$\sup_{0\leqslant n\leqslant N}\big(\mathbb{E}[\|E_n^{\Delta t,4}\|_{\mathcal{H}}^p]\big)^{\frac{1}{p}} \leqslant C_{\alpha,p}(T)\Delta t^\alpha\big(1+\|x_0\|_{\mathcal{E}}^4\big). \tag{5.74}$$

- Note that if the splitting schemes (5.21) and (5.22) are considered, one has $\mathcal{A}_{\Delta t} = e^{-\Delta t\Lambda}$ and thus $E_n^{\Delta t,5} = 0$ for all $n \geqslant 0$. If the splitting scheme (5.23) is considered, one has $\mathcal{A}_{\Delta t} = (I + \Delta t\Lambda)^{-1}$. Using the inequality (5.64), for all $x = (u,v) \in \mathcal{H}$ and all $0 \leqslant k \leqslant n-1$ one has

$$\begin{aligned}\|(e^{-(t_n-t_k)\Lambda}x - \mathcal{A}_{\Delta t}^{n-k}x\|_{\mathcal{H}} &= \|e^{(n-k)\Delta t\Delta}u - ((I-\Delta t\Delta)^{-1})^{n-k}u\|_H \\ &\leqslant \frac{C\|u\|_H}{(n-k)} \leqslant \frac{C\|x\|_{\mathcal{H}}}{(n-k)^\alpha}.\end{aligned}$$

  As a consequence, using the Minkowskii inequality, the local Lipschitz continuity property (5.31) of $\psi_{\Delta t}$ combined with the bound (5.33) (Proposition 5.3.2) and the moment bounds (5.51) from Proposition 5.4.3, one has

$$\begin{aligned}\big(\mathbb{E}[\|E_n^{\Delta t,5}\|_{\mathcal{H}}^p]\big)^{\frac{1}{p}} &\leqslant \Delta t\sum_{k=0}^{n-1}\frac{C}{(n-k)^\alpha}\big(1+\big(\mathbb{E}[\|X_{\Delta t}(t_k)\|_{\mathcal{E}}^{4p}]\big)^{\frac{1}{p}}\big) \\ &\leqslant C_p(T)\Delta t\sum_{\ell=1}^{n}\frac{1}{t_\ell^\alpha}\Delta t^\alpha\big(1+\|x_0\|_{\mathcal{E}}^4\big).\end{aligned}$$

  Using the fact that for all $\alpha \in [0,1)$ one has

$$\sup_{\Delta t\in(0,\Delta t_0)}\Delta t\sum_{\ell=1}^{N}\frac{1}{t_\ell^\alpha} < \infty,$$

  one obtains the following upper bound: for all $\alpha \in [0, \frac{1}{4})$, $p \in [1,\infty)$ and $T \in (0,\infty)$, there exists $C_{\alpha,p}(T) \in (0,\infty)$ such that for all $\Delta t \in (0, \Delta t_0)$ one has

$$\sup_{0\leqslant n\leqslant N}\big(\mathbb{E}[\|E_n^{\Delta t,5}\|_{\mathcal{H}}^p]\big)^{\frac{1}{p}} \leqslant C_{\alpha,p}(T)\Delta t^\alpha\big(1+\|x_0\|_{\mathcal{E}}^4\big). \tag{5.75}$$

To conclude the proof: using the decomposition of the error (5.65) and the upper bounds (5.71), (5.72), (5.73), (5.74) and (5.75), one obtains the following upper bound: for all $\alpha \in [0, \frac{1}{4})$, $p \in [1, \infty)$ and $T \in (0, \infty)$, there exists $C_{\alpha,p}(T) \in (0, \infty)$ such that for all $\Delta t \in (0, \Delta t_0)$ one has

$$\sup_{0 \leqslant n \leqslant N} \big(\mathbb{E}[\|X_{\Delta t}(t_n) - X_n^{\mathrm{aux}}\|_{\mathcal{H}}^p]\big)^{\frac{1}{p}} \leqslant C_{\alpha,p}(T)\Delta t^{\alpha}\big(1 + \|(-\Delta)^{\alpha} u_0\|_H^7 + \|x_0\|_{\mathcal{E}}^7\big).$$

This concludes the proof of the inequality (5.62) and the proof of Lemma 5.5.4 is completed. □

Proof of Lemma 5.5.5. Using the expressions (5.58) and (5.24) for X_n^{aux} and X_n, and the definition (5.25) of the mapping $\psi_{\Delta t}$, for all $n \in \{0, \dots, N-1\}$ one obtains

$$X_{n+1}^{\mathrm{aux}} - X_{n+1} = \mathcal{A}_{\Delta t}\big(X_n^{\mathrm{aux}} - X_n\big) + \Delta t \mathcal{A}_{\Delta t}\big(\psi_{\Delta t}(X_{\Delta t}(t_n)) - \psi_{\Delta t}(X_n)\big).$$

Writing

$$\psi_{\Delta t}(X_{\Delta t}(t_n)) = \psi_{\Delta t}(X_{\Delta t}(t_n)) - \psi_{\Delta t}(X_n^{\mathrm{aux}}) + \psi_{\Delta t}(X_n^{\mathrm{aux}}),$$

and using again the identity (5.25), one obtains

$$X_{n+1}^{\mathrm{aux}} - X_{n+1} = \mathcal{A}_{\Delta t}\big(\phi_{\Delta t}(X_n^{\mathrm{aux}}) - \phi_{\Delta t}(X_n)\big) + \Delta t \mathcal{A}_{\Delta t}\big(\psi_{\Delta t}(X_{\Delta t}(t_n)) - \psi_{\Delta t}(X_n^{\mathrm{aux}})\big). \tag{5.76}$$

On the one hand, using the inequalities (5.34) (Proposition 5.4.1), if $\mathcal{A}_{\Delta t} = e^{-\Delta t \Lambda}$ and (5.37), if $\mathcal{A}_{\Delta t} = (I + \Delta t \Lambda)^{-1}$, and the global Lipschitz continuity property (5.29) of $\phi_{\Delta t}$ (Proposition 5.3.1), one obtains

$$\begin{aligned}\|\mathcal{A}_{\Delta t}\big(\phi_{\Delta t}(X_n^{\mathrm{aux}}) - \phi_{\Delta t}(X_n)\big)\|_{\mathcal{H}} &\leqslant \|\phi_{\Delta t}(X_n^{\mathrm{aux}}) - \phi_{\Delta t}(X_n)\|_{\mathcal{H}} \\ &\leqslant e^{\Delta t(1+\|\|B\|\|)}\|X_n^{\mathrm{aux}} - X_n\|_{\mathcal{H}}.\end{aligned}$$

On the other hand, using the inequalities (5.34) (Proposition 5.4.1), if $\mathcal{A}_{\Delta t} = e^{-\Delta t \Lambda}$ and (5.37), if $\mathcal{A}_{\Delta t} = (I + \Delta t \Lambda)^{-1}$, and the local Lipschitz continuity property (5.31) of $\psi_{\Delta t}$ (Proposition 5.3.2), one obtains

$$\begin{aligned}\|\mathcal{A}_{\Delta t}\big(\psi_{\Delta t}(X_{\Delta t}(t_n)) - \psi_{\Delta t}(X_n^{\mathrm{aux}})\big)\|_{\mathcal{H}} &\leqslant \|\psi_{\Delta t}(X_{\Delta t}(t_n)) - \psi_{\Delta t}(X_n^{\mathrm{aux}})\|_{\mathcal{H}} \\ &\leqslant C(\Delta t_0)\Big(1 + \|X_{\Delta t}(t_n)\|_{\mathcal{E}}^3 + \|X_n^{\mathrm{aux}}\|_{\mathcal{E}}^3\Big) \\ &\quad \times \|X_{\Delta t}(t_n) - X_n^{\mathrm{aux}}\|_{\mathcal{H}}.\end{aligned}$$

By a straightforward argument, since $X_0^{\mathrm{aux}} = X_0 = x_0$, for all $n \in \{0, \dots, N\}$, one has

$$\begin{aligned}&\|X_n^{\mathrm{aux}} - X_n\|_{\mathcal{H}} \leqslant \\ &\quad \leqslant C(\Delta t_0) e^{T(1+\|\|B\|\|)} \Delta t \sum_{k=1}^{N} \Big(1 + \|X_{\Delta t}(t_k)\|_{\mathcal{E}}^3 + \|X_k^{\mathrm{aux}}\|_{\mathcal{E}}^3\Big)\|X_{\Delta t}(t_k) - X_k^{\mathrm{aux}}\|_{\mathcal{H}}.\end{aligned}$$

Using the Minkowskii and Cauchy–Schwarz inequalities, the moment bounds (5.51) and (5.59) from Proposition 5.4.3 and Lemma 5.5.2 respectively, and

the error estimate (5.62) from Lemma 5.5.4, one obtains the following strong error estimate: for all $\alpha \in [0, \frac{1}{4})$, $p \in [1, \infty)$ and $T \in (0, \infty)$, there exists $C_{\alpha,p}(T) \in (0, \infty)$ such that for all $\Delta t \in (0, \Delta t_0)$ one has

$$\begin{aligned}
&\sup_{0 \leqslant n \leqslant N} \big(\mathbb{E}[\|X_n^{\mathrm{aux}} - X_n\|_{\mathcal{H}}^p]\big)^{\frac{1}{p}} \\
&\qquad \leqslant C(T)\Delta t \sum_{k=1}^{N} \Big(1 + \big(\mathbb{E}[\|X_{\Delta t}(t_k)\|_{\mathcal{E}}^{6p}]\big)^{\frac{1}{2p}} + \big(\|X_k^{\mathrm{aux}}\|_{\mathcal{E}}^{6p}]\big)^{\frac{1}{2p}}\Big) \\
&\qquad\quad \times \big(\mathbb{E}[\|X_{\Delta t}(t_k) - X_k^{\mathrm{aux}}\|_{\mathcal{H}}^{2p}\big)^{\frac{1}{2p}} \\
&\qquad \leqslant C_{\alpha,p}(T)\Delta t^{\alpha}\big(1 + \|x_0\|_{\mathcal{E}}^3\big)\big(1 + \|(-\Delta)^{\alpha} u_0\|_H^4 + \|x_0\|_{\mathcal{E}}^4\big).
\end{aligned}$$

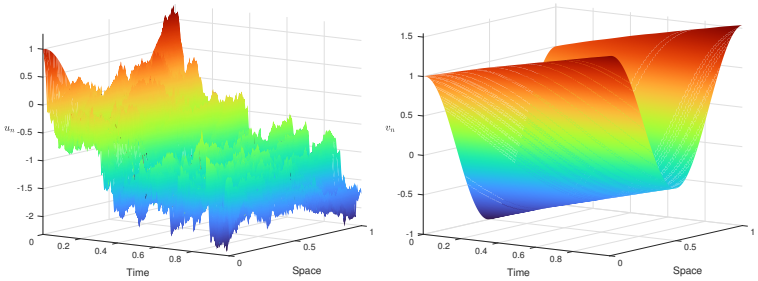
This concludes the proof of the inequality (5.63) and the proof of Lemma 5.5.5 is thus completed. □

5.6 Numerical experiments

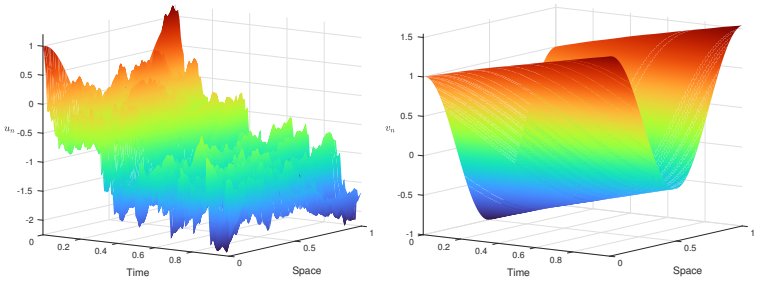
This section presents numerical experiments to support and illustrate the above theoretical results. To perform these numerical experiments, The stochastic FitzHugh–Nagumo SPDE system (5.8) with Neumann boundary conditions on the interval $[0, 1]$ is considered. The spatial discretization is performed using a standard finite difference method with mesh size denoted by h. In order to obtain a linear system with a symmetric matrix, centered differences are used to the numerical discretization of the Laplacian, while first order differences are used for the discretization of the Neumann boundary conditions. The initial values are given by $u_0(\zeta) = \cos(2\pi\zeta)$ and $v_0(\zeta) = \cos(2\pi\zeta)$. For the temporal discretization, the three Lie–Trotter splitting integrators (5.21), (5.22) are used and (5.23) studied in this paper, denoted below by LTexact, LTexpo, LTimp respectively.

5.6.1 Evolution plots

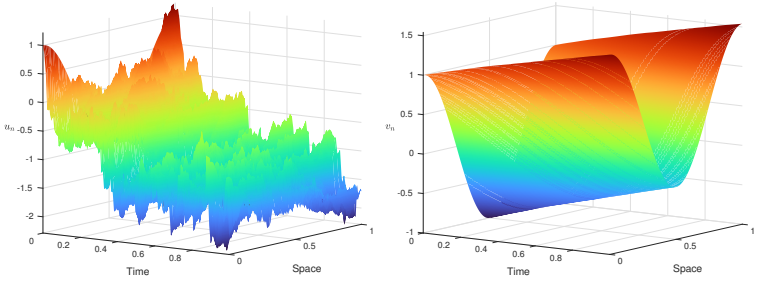
Let us first display one sample of the numerical solutions of the stochastic FitzHugh–Nagumo system (5.8) with the parameters $\gamma_1 = 0.08$, $\gamma_2 = 0.8\gamma_1$ and $\beta = 0.7$. The SPDE is discretized with finite differences with mesh $h = 2^{-10}$. The time interval $[0, T] = [0, 1]$ is considered and the integrators with time step size $\Delta t = 2^{-15}$ are applied. The results are presented in Figure 5.1. The general behaviour of the numerical solutions given by the three splitting schemes is the same. However, one can observe a spatial smoothing effect in the u component of the solution when the schemes LTexpo–(5.21) or to some extent LTimp–(5.23) are applied: for a given time step size, the spatial regularity of the numerical solution is increased compared with the one of the exact solution. On the contrary, the scheme LTexact–(5.21) preserves the spatial regularity of the solution for any value of the time step size. One can refer to the recent preprint [40] for the analysis of this phenomenon for parabolic semilinear SPDEs. Let us emphasize that the phenomenon is due to the way the



(a) Component u for LTEX-ACT (b) Component v for LTEX-ACT



(c) Component u for LTEXPO (d) Component v for LTEXPO



(e) Component u for LTIMP (f) Component v for LTIMP

Figure 5.1: Space-time evolution plots of u and v using the Lie–Trotter splitting schemes LTEXACT, LTEXPO, and LTIMP.

stochastic convolution is computed, exactly for the scheme LTEXACT–(5.21) or approximately for the schemes LTEXPO–(5.21) and LTIMP–(5.23).

5.6.2 Mean-square error plots

To illustrate the rates of strong convergence for the Lie–Trotter splitting schemes stated in Theorem 5.3.2, the stochastic FitzHugh–Nagumo system (5.8) with the parameters $\gamma_1 = \gamma_2 = \beta = 1$ is considered, with $T = 1$ and a finite difference method is applied with $h = 2^{-9}$ for spatial discretization. The Lie–Trotter splitting schemes are applied with time steps ranging from 2^{-10} to 2^{-18}. The reference solution is computed using the scheme LTEXACT–(5.21) with time step size $\Delta t_{\mathrm{ref}} = 2^{-18}$. The expectation is approximated using $M_s = 100$ samples. A plot in logarithmic scales for the mean-square errors

$$\left(\mathbb{E}[\|X(t_N) - X_N\|_{\mathcal{H}}^2]\right)^{\frac{1}{2}}$$

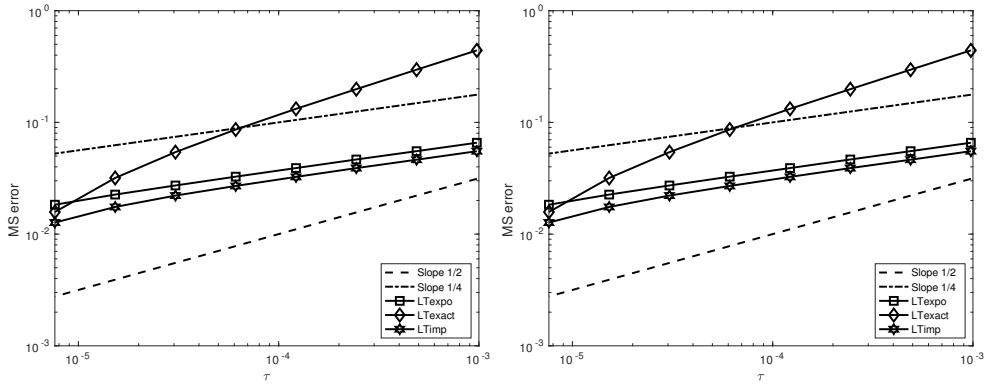


Figure 5.2: Mean-square errors as a function of the time step: Lie–Trotter splitting schemes: left ($\phi_{\Delta t} = \phi^{\mathrm{L}}_{\Delta t} \circ \phi^{\mathrm{NL}}_{\Delta t}$) and right ($\phi_{\Delta t} = \phi^{\mathrm{NL}}_{\Delta t} \circ \phi^{\mathrm{L}}_{\Delta t}$) ($\diamond$ for LTexact, $\square$ for LTexpo, $*$ for LTimp). The dotted lines have slopes 1/2 and 1/4.

is given on the left-hand side of Figure 5.2. One observes that the strong rate of convergence for the three considered Lie–Trotter splitting schemes is at least 1/4, which illustrates the result stated in Theorem 5.3.2. Furthermore, the numerical experiments suggest that for the scheme LTexact–(5.21) the order of convergence is 1/2, which is not covered by Theorem 5.3.2. The fact that using an accelerated exponential Euler scheme where the stochastic convolution is computed exactly yields higher order of convergence is known for parabolic semilinear stochastic PDEs driven by space-time white noise, under appropriate conditions, see for instance [157] or [40, Proposition 7.3]. However, the stochastic FitzHugh–Nagumo equations considered in this article are not parabolic systems therefore it is not known how to prove the observed higher order strong rate of convergence.

The right-hand side of Figure 5.2 shows the errors for the variant (5.47) of the splitting scheme (5.24) introduced in Remark 5.4.1: the mapping $\phi_{\Delta t} = \phi^{\mathrm{L}}_{\Delta t} \circ \phi^{\mathrm{NL}}_{\Delta t}$ given by (5.20) is replaced by $\hat{\phi}_{\Delta t} = \phi^{\mathrm{NL}}_{\Delta t} \circ \phi^{\mathrm{L}}_{\Delta t}$ given by (5.45). As explained in Remark 5.4.1, this type of Lie–Trotter schemes is not covered by the results in Section 5.3.3, more precisely the moment bounds in Theorem 5.3.1 cannot be proved by the techniques used in this article. However, the numerical experiments are similar to those on the left-hand side of Figure 5.2 and suggest that the strong order of convergence for this variant is at least 1/4, and that higher order convergence with rate 1/2 may be obtained for the variant of the scheme LTexact–(5.21).

Chapter 6

Stochastic evolution models: fake news diffusion problem

Fake news is a term usually used to define articles containing invented, misleading information created to misinform and make hoaxes viral through the internet. Although the spread of fake news has recently become of great interest, this problem has always existed. What seems to have clearly changed today compared to the past, is the amount of fake news circulating in information, especially on-line, and the importance that they are taking on. Fake news is created by completely ignoring the editorial rules and processes used to ensure its compliance and truthfulness [132]. Actually, it is created for different purposes; the most common is certainly the electoral one in order to discredit the political opponent by conditioning public opinion. Another important purpose is the profit that is obtained online in proportion to the number of visitors to the article.

In today's society, fake news is spreading significantly faster, compared to real news, because of social media that has radically changed the way in which one transmits malformations. Nowadays, internet is widely used as the main medium of information all over the world, since it is a rapidly accessible and low cost source respect to traditional media such as radio, television, newspapers, etc.

The growing need to stop fake news is highlighted also by the latest COVID-19 pandemic. In fact, the vaccination campaign against this disease has been enormously slowed down due to the spread of fake news and the inability of individuals to discern the authenticity of such information [130]. In order to counter the spread of fake news, a CoronaVirusFacts Alliance was created at the Poynter Institute, where the data database containing fake news is updated daily to insert new publications. For example, Figure 6.1 shows the categories of fact-checks that help the alliance to identify successive waves of misinformation. In particular, the bigger the wave, the more fake news regarding the respective topic have been diffused.

The stronger is the content of the published post, the quicker it will circulate on social platforms. In order to understand the phenomena of fake news it is crucial to analyze it in relation to how social media works. When a fake

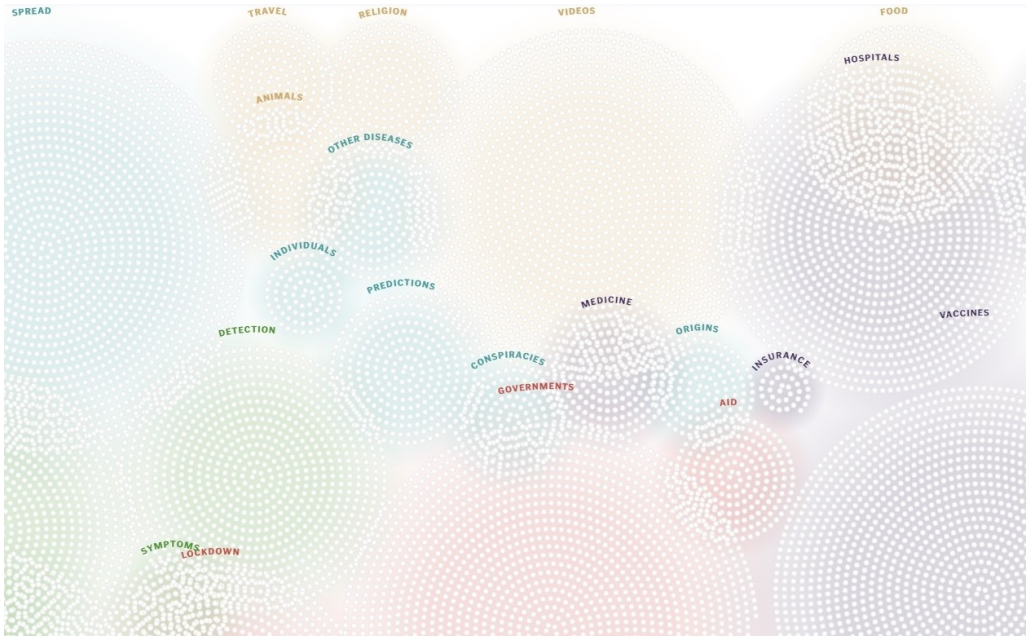


Figure 6.1: Waves of hoaxes of COVID-19 pandemic from `https://www.poynter.org`

news becomes viral, there is no way to stop it. People who create fake news in most cases use powerful images, tough language and harsh words because it has been demonstrated that this kind of contents is most likely to be shared due to its emotional and persuasive power. [12, 50, 129, 176, 208].
It has been demonstrated that fake news usually circulate faster and more broadly than the truth. A study published from the MIT in 2018 revealed that fake news are 70% more likely to be shared on social media than real news, especially if the topic is politics [249]. Also, from an emotional point of view, fake news broad diffusion is related to the feelings they stimulate in the users, which are used to allow themselves to be bullied into a content that simply confirm them their thoughts and ideas, even if it is not the truth [176].

Therefore the problem of fake news is one of the most important that modern society faces us and we need to be ready to respond with determination and precision. An important step therefore concerns the development of models capable of detecting whether a given news is true or false. The interested reader can refer to the review [132] in which recent methods based on the aid of big data and artificial intelligence tools are described. Another important strategy consists in the analysis of the multiple parts that form a fake news, that is, both on the part of the creator and on the part of the user, but also on the linguistics and semantics of the real content and its style, and finally on the social context of the information.
Currently there are some important companies which try to limit the spread of fake news and thus avoid the bad consequences that they can cause. For example, the Kinzen company detects dangerous misinformation and harmful content in audio. Through internal inguistic models, they generate optimized audio transcripts to detect dangerous misinformation in multiple languages.

Advanced machine learning models detect and report harmful languages, statements, narratives, and policy violations within audio. The analysis is further enhanced by the data points and crucial context of human experts, including hashtags, keywords, phrases, slogans and insults.
Furthermore, there are many initiatives that countries around the world take to counter the spread of fake news. In Italy, for example, the National Recovery and Resilience Plan (PNRR), in the section dedicated to cybersecurity, supports the danger of the spread of fake news and the risks they represent for democracy and for the decision-making processes that are the basis of our political system. In particular, as reported in the Plan, an explicit research objective for cybersecurity is to defend democracy and fight fake news through a multidisciplinary, monitoring and "Early Warning" approach based on the flow of information and techniques for predicting human behaviour.
In recent years, many authors have tried to create adequate mathematical models capable of predicting the spread of fake news, in order to limit in advance the effects that the spread of these are having on our society: the interested reader can refer, for instance, to [35, 50, 128, 162, 183, 188, 209, 216, 217, 259] and references therein.
In this thesis a famous model for the spread of a disease, i.e. the SIR model, is used as a model for fake news diffusion. The SIR model, well known in the existing literature of mathematical epidemiology (see, for instance, the very first contributions on compartmental models [159, 160, 218], as well as the monographs [36, 86, 87, 111, 126, 127, 186, 189] and references therein), describes the effects of the spread out of an epidemics in a population ideally divided into three subgroups (susceptible, infected and recovered people). As visible in the literature (see, for instance, [109, 132, 183]), epidemiological models can be profitably used to describe to diffusion of fake information as an infectious disease. Most of these models are given by nonlinear differential equations, whose dynamics can be understood by first looking at the eigenvalues of the Jacobian of the linearized problem and then, linking the stiffness ratio (i.e. the ratio between the largest and the smallest moduli of the eigenvalues of the aforementioned Jacobian matrix) to the evolution of the spreading. The SIR model is the main channel to provide our analysis: the stiffness ratio can help to understand the requested time to recover the truth in a given country exposed to a certain fake news. To some extent, our aim is to provide a novel element to understand the effectiveness of modeling the diffusion of fake news, as well as the re-establishment of the truth: the more the stiffness ratio is high, the faster is the re-establishment of the truth after the diffusion of fake information. These kinds of arguments are not available in the literature, at the best of our knowledge.

6.1 Formulation of the SIR model

The SIR model was first introduced in 1927 by Kermack and McKendrick [160], even if seminal contributions on compartmental models are also given in [159, 218]. It represents an extremely simple mathematical model for describing the

transmission of an infectious disease. It is a special type of model called a compartmental model, because each letter refers to a "compartment" in which an individual can reside. This model describes the mutual interactions of three population of individuals: the population S of susceptible people, i.e. the healthy individuals who can contract the disease; the population I of the infected, i.e. individuals who have contracted the disease and are able to transmit it; the population R of the recovered, that is, individuals who are healed (see Figure 6.2).

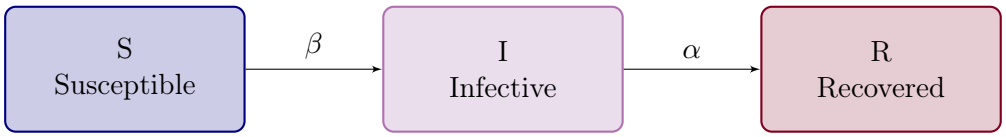


Figure 6.2: Scheme of the classical SIR model.

In our model for the spread out of fake information, above populations are described as follows:

- $S(t)$: potentially authoring the spreading of fake news;
- $I(t)$: the wide variety of authors highly active in posting fake information;
- $R(t)$: authors who are inactive to the spreading of fake news.

The model is based on the continuous interaction between susceptible and infected individuals along time. The corresponding system of ODEs is then given by:

$$\begin{cases} \dfrac{dS(t)}{dt} = -\beta S(t)I(t), \\ \dfrac{dI(t)}{dt} = \beta S(t)I(t) - \alpha I(t), \\ \dfrac{dR(t)}{dt} = \alpha I(t), \end{cases} \tag{6.1}$$

where α is the rate of recovery and β the contact rate. Specifically, since the purpose of our method is to compare the impact of fake news in different countries, the parameters β and α, are related to two important indices, commonly used to describe the social, economic and cultural performance of our society. In particular,

$$\beta = \frac{i}{10}, \quad \alpha = \frac{h}{100},$$

where i is the *internet penetration index* of the country and h is the human development index of the same country. These values are commonly provided in the annual report of United Nations Development Programme. In general, the value of α is smaller than that of β because it is easier to spread a lie than reaffirming the truth. Table 6.1 shows, for instance, the values of α and β for selected countries, i.e., Australia, Brazil, France, India, Italy, Mexico, Mozambique and USA, referring to the year 2019.

Table 6.1: Values of the constants α, β, for France, India, Italy, Mexico and United States, referring to 2019.

	α	β
Australia	0.009	0.087
Brazil	0.008	0.072
France	0.009	0.089
India	0.006	0.035
Italy	0.009	0.061
Mexico	0.008	0.064
Mozambique	0.005	0.021
United States	0.009	0.075

Remark 6.1.1. *It is crucial to highlight the different power of the denial of a fake news and that of the fake news itself [249]. Fake information circulates online much faster than truth one and is more prone to be shared by users who encounter it; on the other hand, the re-assessment of the truth is nowhere viral and reaches far fewer people than those who have read or spread fake information. Due to this intrinsic characteristic of true news compared to false news, the former require a much greater commitment on the part of individual users than that related to the spread of fake news. Above all, the spread out of fake information does not necessarily require a particularly strong human presence: often fake news are circulated by bots created specifically by someone or shared by fake accounts, so they do not correspond to real people. Real people, on the other hand, are necessary and represent the only option to restore the truth. This motivates the choice of the recover rate α linked to human development index per country and the contact rate β to the spread of the internet in the same country.*

6.2 Stiffness analysis

In this section the stiffness ratio of system (6.1) will be analysed. To do this, let us consider an initial value vector

$$\begin{bmatrix} S_0 \\ I_0 \\ R_0 \end{bmatrix} = \begin{bmatrix} S(0) \\ I(0) \\ R(0) \end{bmatrix},$$

and let us linearise model (6.1) around it, leading to

$$\begin{cases} \dfrac{dS(t)}{dt} = \beta S_0 I_0 - \beta I_0 S(t) - \beta S_0 I(t) + \text{high order terms}, \\ \dfrac{dI(t)}{dt} = -\beta S_0 I_0 + \beta I_0 S(t) + (\beta S_0 - \alpha) I(t) + \text{high order terms}, \\ \dfrac{dR(t)}{dt} = \alpha I(t). \end{cases} \tag{6.2}$$

Correspondingly, let us compute the Jacobian matrix of the linear part of the vector field in (6.2), i.e.,

$$J_{\alpha,\beta}(S_0, I_0) = \begin{bmatrix} -\beta I_0 & -\beta S_0 & 0 \\ \beta I_0 & \beta S_0 - \alpha & 0 \\ 0 & \alpha & 0 \end{bmatrix},$$

whose spectrum consists in one eigenvalue equal to 0 and two real eigenvalues $\lambda^{\min}_{\alpha,\beta}(S_0, I_0)$ and $\lambda^{\max}_{\alpha,\beta}(S_0, I_0)$, with $|\lambda^{\min}_{\alpha,\beta}(S_0, I_0)| < |\lambda^{\max}_{\alpha,\beta}(S_0, I_0)|$. Correspondingly, the ratio

$$\sigma_{\alpha,\beta}(S_0, I_0) = \frac{|\lambda^{\max}_{\alpha,\beta}(S_0, I_0)|}{|\lambda^{\min}_{\alpha,\beta}(S_0, I_0)|}, \tag{6.3}$$

meaningful in the analysis of stiff problems [174], provides the so-called *stiffness ratio* of (6.1). Table 6.2 reports the stiffness ratio for each country, related to the initial value

$$\begin{bmatrix} S_0 \\ I_0 \\ R_0 \end{bmatrix} = \begin{bmatrix} 0.7 \\ 0.1 \\ 0 \end{bmatrix}, \tag{6.4}$$

i.e. assuming that 70% of the initial population is susceptible, 10% are infected and there are no recovered people. The results reveal that, the more the internet penetration index i is higher, the more the stiffness ratio is bigger. As a consequence, the corresponding model (6.1) is more stiff and the spread of fake news should be more damped in time. In other words, the more (6.1) is

Table 6.2: Values of the stiffness ratios (6.3) in France, India, Italy, Mexico and United States, referring to 2019, assuming the vector (6.4) as initial point.

	$S_{\alpha,\beta}(S_0, I_0)$
Australia	20.03
Brazil	20.85
France	23.07
India	8.38
Italy	12.35
Mexico	17.13
Mozambique	4.39
United States	17.00

stiff, the more the corresponding country exhibits a faster transit of fake news. Countries with a lower internet penetration index i are characterized by a less stiff model and, as a consequence, the transit of fake information is slower and circulates for much more time.

6.3 Numerical evidences

In this section some numerical tests are performed to confirm our theoretical analysis, i.e., the spread of fake news is closely linked to the stiffness ratio of (6.1). For each listed country, Figures 6.3–6.10 show the solution of problem (6.1) in the interval [0,1000], computed by the standard Matlab built-in function `ode15s`, and the pattern of the ratio $\tau_{\alpha,\beta}(S(t), I(t))$ between the maximum and minimum moduli of the non-zero eigenvalues of the matrix

$$J_{\alpha,\beta}(S(t), I(t)) = \begin{bmatrix} -\beta I(t) & -\beta S(t) & 0 \\ \beta I(t) & \beta S(t) - \alpha & 0 \\ 0 & \alpha & 0 \end{bmatrix},$$

that corresponds to the Jacobian of the problem (6.1), frozen at time t. To some extent, we aim to check the evolution in time of the stiffness ratio

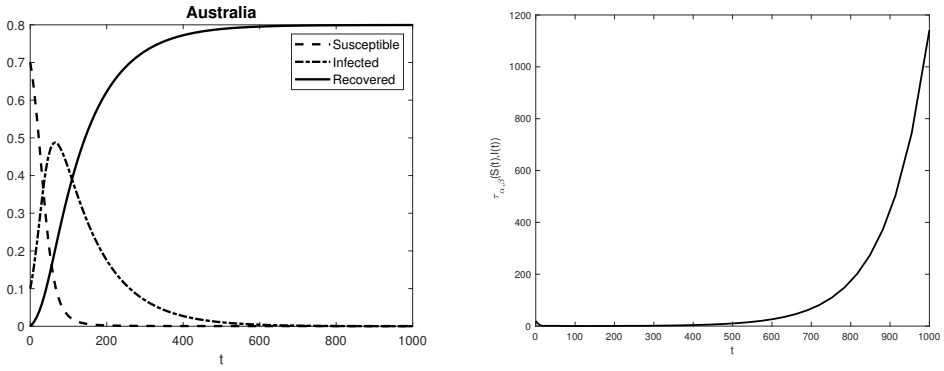


Figure 6.3: Solution to the SIR model (6.1), with initial value given by the vector (6.4), for Australia (top) and corresponding pattern of $\tau_{\alpha,\beta}(S(t), I(t))$ (bottom).

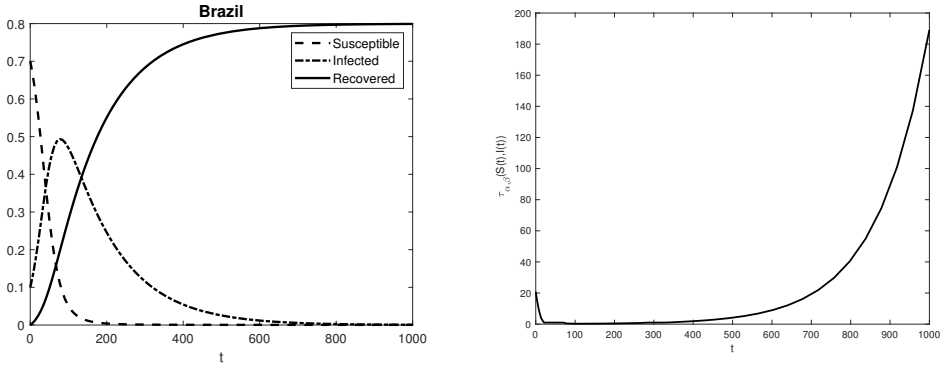


Figure 6.4: Solution to the SIR model (6.1), with initial value given by the vector (6.4), for Brazil (top) and corresponding pattern of $\tau_{\alpha,\beta}(S(t), I(t))$ (bottom).

$\sigma_{\alpha,\beta}(S_0, I_0)$.

Each figure confirms that the higher the stiffness ratio, as listed in Table 6.2, the faster the transit of fake news will be. In some countries, such as India or Mozambique, where the internet penetration index is small, the function $\tau_{\alpha,\beta}(S(t), I(t))$ grows much than in the other cases (corresponding to countries with higher internet penetration indices). As a consequence, smaller values of the stiffness ratio correspond to a slower achievement of the maximum number of infected people and, consequently, to a slower dispersion of fake news. The observed number of time units needed to achieve the maximum number of infected is listed in Table 6.3: one can observe that the number of time units is coherent with the stiffness ratio so, the smallest value is for France, the largest is for Mozambique.

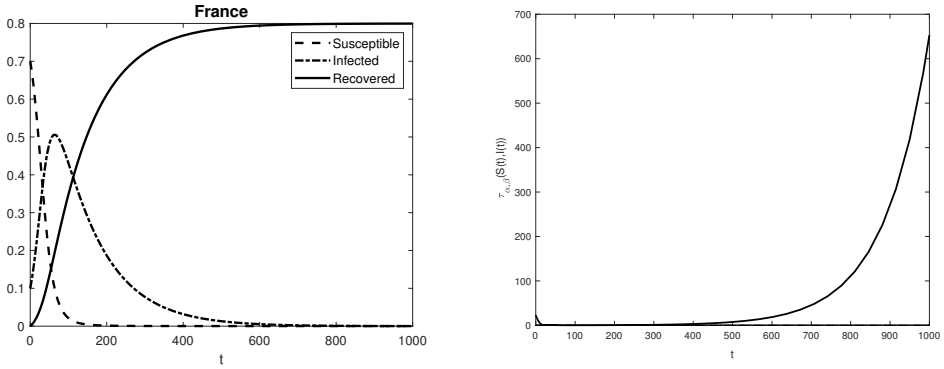


Figure 6.5: Solution to the SIR model (6.1), with initial value given by the vector (6.4), for France (top) aand corresponding pattern of $\tau_{\alpha,\beta}(S(t), I(t))$ (bottom).

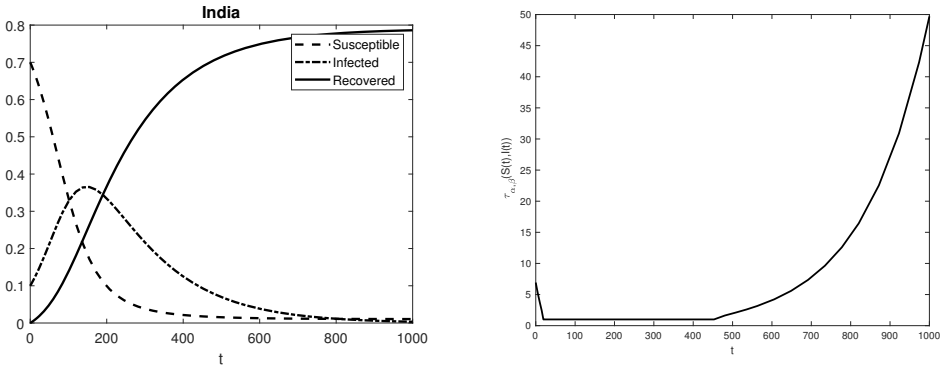


Figure 6.6: Solution to the SIR model (6.1), with initial value given by the vector (6.4), for India (top) and corresponding pattern of $\tau_{\alpha,\beta}(S(t), I(t))$ (bottom).

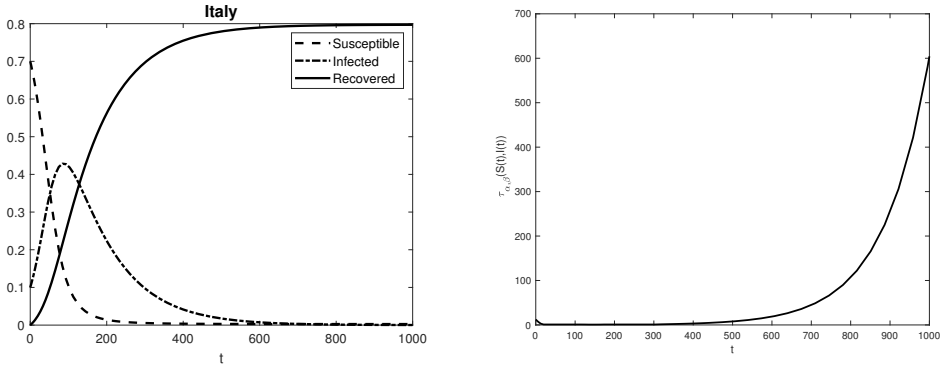


Figure 6.7: Solution to the SIR model (6.1), with initial value given by the vector (6.4), for Italy (top) and corresponding pattern of $\tau_{\alpha,\beta}(S(t), I(t))$ (bottom).

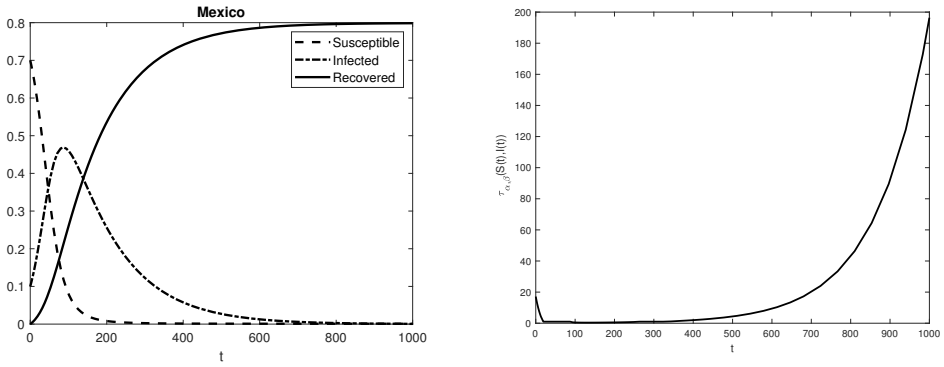


Figure 6.8: Solution to the SIR model (6.1), with initial value given by the vector (6.4), for Mexico (top) and corresponding pattern of $\tau_{\alpha,\beta}(S(t), I(t))$ (bottom).

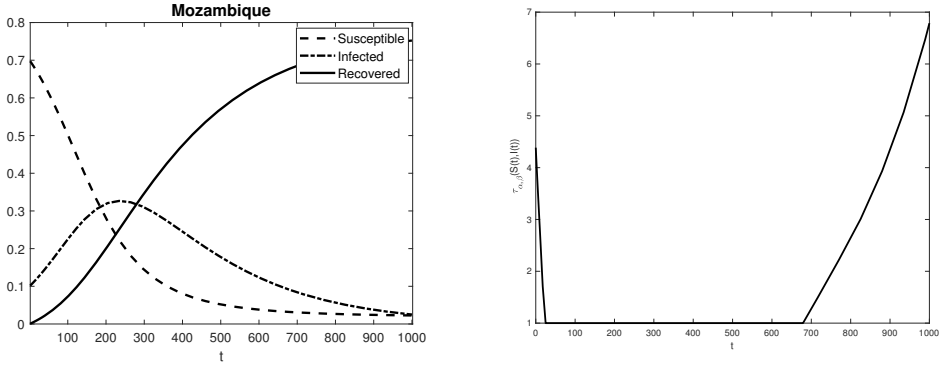


Figure 6.9: Solution to the SIR model (6.1), with initial value given by the vector (6.4), for Mozambique (top) and corresponding pattern of $\tau_{\alpha,\beta}(S(t), I(t))$ (bottom).

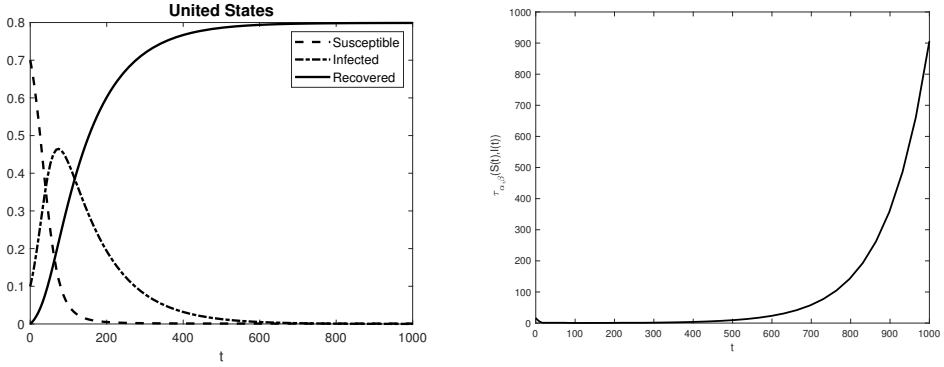


Figure 6.10: Solution to the SIR model (6.1), with initial value given by the vector (6.4), for United States (top) and corresponding pattern of $\tau_{\alpha,\beta}(S(t), I(t))$ (bottom).

Table 6.3: Number of time units required to reach the maximum of infected in France, India, Italy, Mexico and United States, referring to 2019.

	Number of time units
Australia	66.15
Brazil	77.00
France	61.16
India	145.75
Italy	88.90
Mexico	87.02
Mozambique	234.40
United States	74.05

6.4 Stochastic SIR model

As the reader noted in previous chapters, stochastic differential equation models are preferred over deterministic ones when the dynamics are affected by random perturbations. The problem of fake news certainly falls into these phenomena. In fact, the dissemination of information is a stochastic process that may depend, for example, on the content of the information, the influence of users and the structure of the network. Therefore, the number of users exposed to different stories varies greatly and it is needed to consider probabilistic exposure patterns to capture this uncertainty [163].

Let $(\Omega, \mathcal{F}, \mathbb{P})$ to be a complete probability space with filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$, satisfying the usual conditions. Then, by introducing stochastic perturbations in the deterministic SIR model (6.1), the resulting system of SDEs reads

$$\begin{cases} dS(t) = -\beta S(t)I(t)dt - \sigma S(t)I(t)dW(t), \\ dI(t) = (\beta S(t)I(t) - \alpha I(t))dt + \sigma S(t)I(t)dW(t), \\ dR(t) = \alpha I(t)dt, \end{cases} \tag{6.5}$$

where the three different populations $S(t)$, $I(t)$ and $R(t)$ assume the same meaning of the above deterministic model as well as the parameters β and α. The new parameter σ is the so-called *perturbation parameter*, which describes changes in the infection rate changes over time.

Unlike the deterministic case, here the analysis of stiffness is more complex, as a strict definition of the stiffness ratio is not yet dictated in the stochas-

tic community. For example, some authors state that a stochastic differential equation is stiff if by applying a standard explicit integrator, such as the classic Euler-Maruyama method, they face a severe step-size restriction due to numerical stability problems [5]. Other authors instead observe the stiffness property on the stochastic and deterministic component separately. This means that a generic system of SDEs, this could be stiff in the deterministic part and not be so in the stochastic part and vice versa or it could be stiff in both components [69].

In our analysis, following the idea of [30, 31], the stiffness analysis can be conducted by observing that the solution of stiff problems is frequently accompanied by a phenomenon of order reduction. In particular stochastic backward Euler (BE) method is used, that, as the reader remember, for suitable function f and G, it is given by

$$X_{n+1} = X_n + \Delta t\, f(t_{n+1}, X_{n+1}) + G(t_n, X_n)\Delta W_n, \quad \text{for } n = 0, \dots, N. \tag{6.6}$$

Furthermore, one remembers that the strong error of the BE method is $\gamma = 1$, meaning that

$$e_{\Delta t}^{\text{strong}} := \sup_{t_0 \leqslant t_n \leqslant T} \mathbb{E}[|X(t_n) - X_n|] \leqslant K\Delta t \tag{6.7}$$

for some positive constant K.
To apply the BE method (6.6) to problem (6.5), the need to solve a system of nonlinear equations at each step is evident. In the numerical simulations shown in the Figures 6.11-6.18, Newton's iterations were used that involve the evaluation of the Jacobian matrix associated with the drift function, i.e.

$$J(S(t_n), I(t_n)) = \begin{bmatrix} -\beta I(t_n) & -\beta S(t_n) & 0 \\ \beta I(t_n) & \beta S(t_n) - \alpha & 0 \\ 0 & \alpha & 0 \end{bmatrix}.$$

All the numerical simulations shown in the left part of Figures 6.11-6.18 were carried out using the BE method with $N = 2^{18}$ grid points with $M = 1000$ simple paths and thus plotting the expected value of each component, calculated as a simple average on the realizations,i.e.

$$\mathbb{E}[X(t)] = \frac{1}{M} \sum_{i=1}^{M} X^i(t)$$

with $X(t) \in \{S(t), I(t), R(t)\}$. Similarly, the right part of Figures 6.11-6.18 represents the experimental verification of the strong convergence order of the BE method (6.6), where the error is measured with the same reference solution of $N = 2^{16}$ steps.

From the numerical simulations it seems evident that the more the coefficient of the diffusion part of the model (6.5) σ increases, the more the order

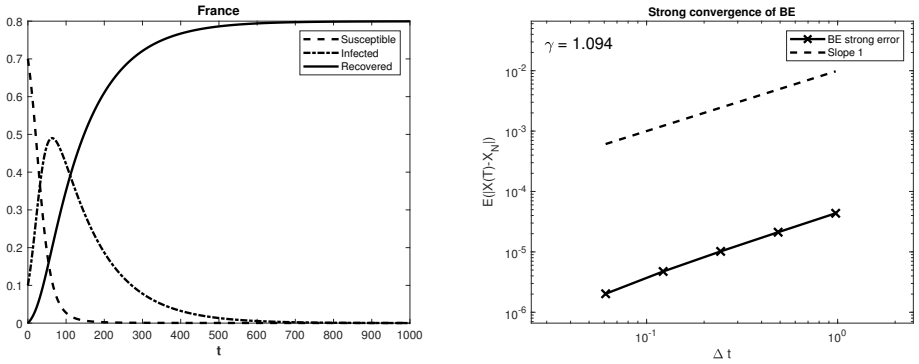


Figure 6.11: Solution to the stochastic SIR model (6.5) for France, with initial value given by the vector (6.4) and $\sigma = 0.01$ (left) and corresponding strong error of BE method applied to (6.5) (right).

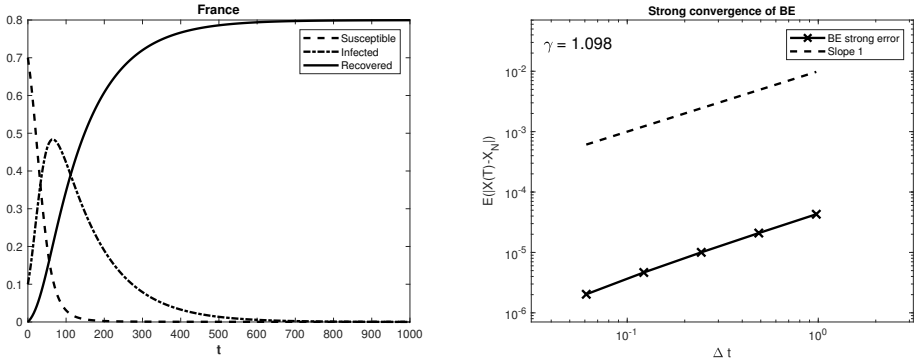


Figure 6.12: Solution to the stochastic SIR model (6.5) for France, with initial value given by the vector (6.4) and $\sigma = 0.1$ (left) and corresponding strong error of BE method applied to (6.5) (right).

of the BE method decreases, both in the case of France (with parameters $\alpha = 0.009$ and $\beta = 0.089$), and in the case of Mozambique (with parameters $\alpha = 0.005$ and $\beta = 0.021$). It is also evident that the more the problem becomes stiff, the more the peak of the infected tends to decrease considerably. In fact, as in the case of France (but in the same way in the case of Mozambique), with the value of $\sigma = 0.01$, the model reaches a peak of infected people of 0.5 (as happens in the deterministic case, see Figure 6.9); while, when σ becomes 0.5, the peak of infectious is considerably reduced until it reaches a value of about 0.3. However, this phenomenon can be explained considering that adding stochastic perturbations inside the model creates noise, which in turn could generate chaos in the population. So this leads us to a decrease in the number of infected. However, this result is in a preliminary version and an extension of this analysis will want to be made in future works.

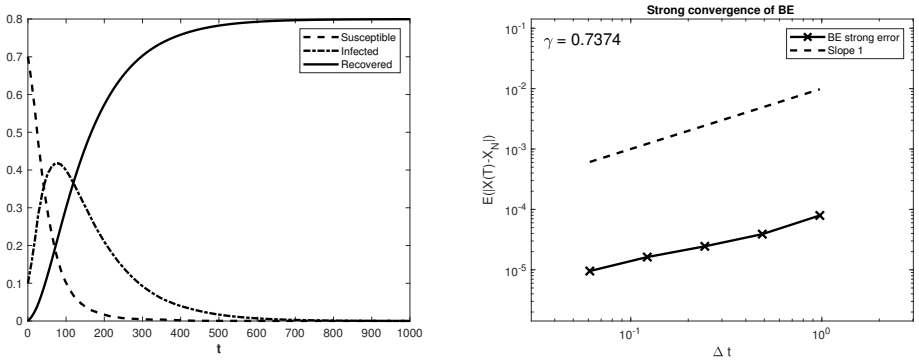


Figure 6.13: Solution to the stochastic SIR model (6.5) for France, with initial value given by the vector (6.4) and $\sigma = 0.3$ (left) and corresponding strong error of BE method applied to (6.5) (right).

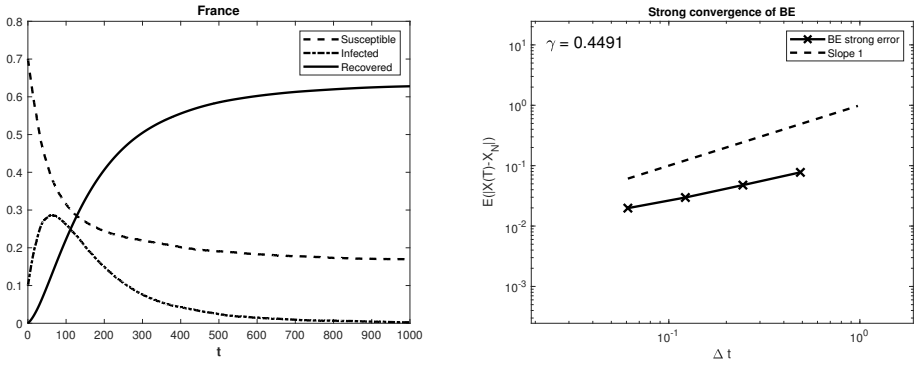


Figure 6.14: Solution to the stochastic SIR model (6.5) for France, with initial value given by the vector (6.4) and $\sigma = 0.5$ (left) and corresponding strong error of BE method applied to (6.5) (right).

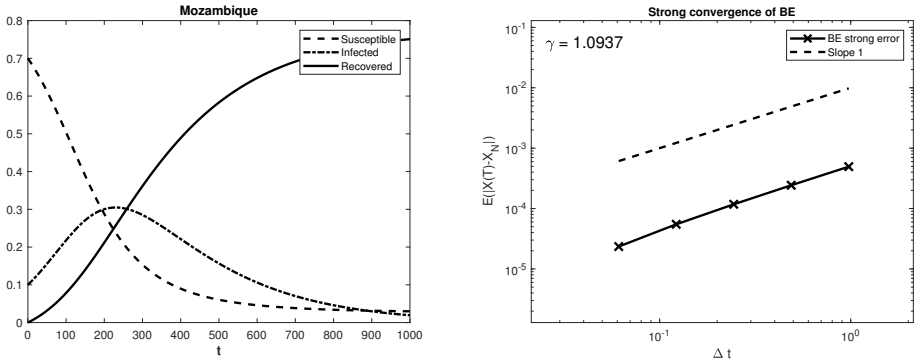


Figure 6.15: Solution to the stochastic SIR model (6.5) for Mozambique, with initial value given by the vector (6.4) and $\sigma = 0.01$ (left) and corresponding strong error of BE method applied to (6.5) (right).

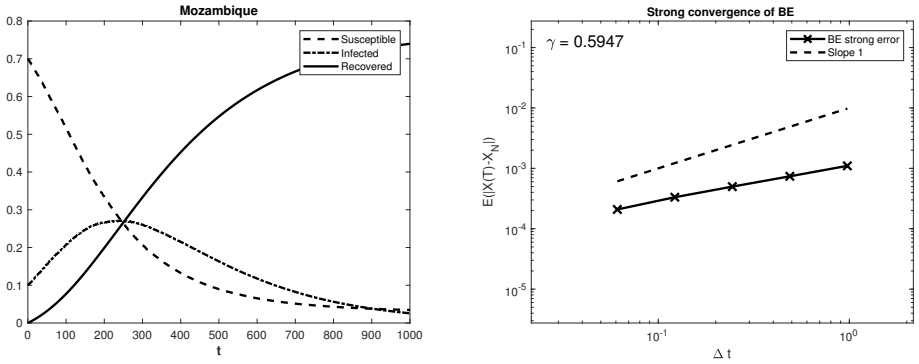


Figure 6.16: Solution to the stochastic SIR model (6.5) for Mozambique, with initial value given by the vector (6.4) and $\sigma = 0.1$ (left) and corresponding strong error of BE method applied to (6.5) (right).

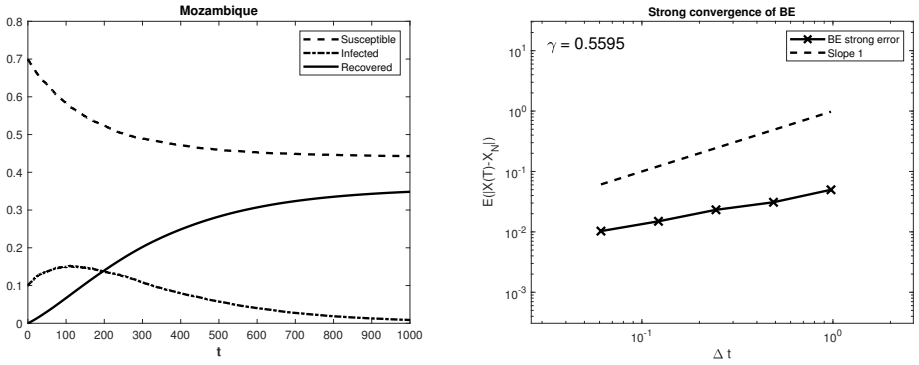


Figure 6.17: Solution to the stochastic SIR model (6.5) for Mozambique, with initial value given by the vector (6.4) and $\sigma = 0.3$ (left) and corresponding strong error of BE method applied to (6.5) (right).

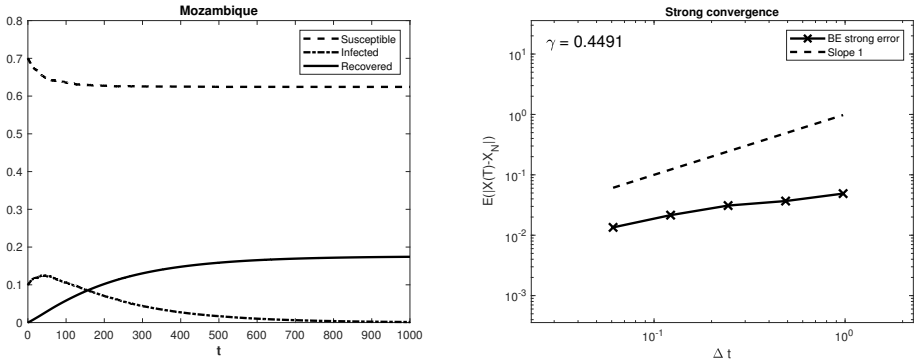


Figure 6.18: Solution to the stochastic SIR model (6.5) for Mozambique, with initial value given by the vector (6.4) and $\sigma = 0.5$ (left) and corresponding strong error of BE method applied to (6.5) (right).

Conclusions and Future Perspectives

This dissertation was devoted to numerics and the analysis of numerical methods for stochastic differential equations, paying particular attention to how these can be used in the context of modeling. Indeed, as we have observed throughout this thesis, stochastic differential equations are powerful tools to model real-life model problems, especially when their dynamics are affected by random perturbations.

Following this spirit, the thesis has been essentially divided into three parts. The first part has been focused on the development of continuous approximants for stochastic differential equations. Following the well-known idea of collocation for deterministic differential equations, particularly for Volterra Integral Equations, we have obtained continuous extensions for specific numerical methods for SDEs. In this family of methods, for particular choices of the parameters that characterize it, a continuous extension of the well-known Euler-Maruyama method is obtained, considered the simplest numerical method for stochastic differential equations. The reader may wonder if continuous extensions for numerical methods for SDE are useful. We know well that in the deterministic context, they constitute a very powerful tool for many aspects. One of these certainly concerns obtaining a good and efficient estimate of the local truncation error and consequently the development of variable step algorithms, which are very useful, especially for the numerical integration of stiff problems. Even in the stochastic context, continuous extensions have found wide use to prove the theorems of existence and uniqueness of solutions in not too stringent hypotheses, such as the global Lipschitz condition: the reader can look at the work, for example [34]. As in the deterministic context, our goal here was to define a procedure capable of efficiently estimating the local truncation error. However, for this estimate, it is necessary to pass through the following identity

$$W(t_n + s\Delta t) - W(t_n) = \sqrt{s\Delta t}V_k, \qquad s \in [0, 1]$$

where V_k is a standard normal variable. We know to be true only punctually, that is, for fixed values of s, and not uniformly. In particular, for a varying $s > 0$ the left side of the identity is only 1/2-Hölder continuous, while the right side is locally Lipschitz continuous. Despite this, the results we have obtained appear to be significant, but the reliability or otherwise of this estimate is, therefore, the object of future works.

Research in this thesis then continued following the spirit of geometric numerical integration. In particular, the attention was placed on stochastic Hamiltonian problems of the Itô type which, in the case of additive noise, we know how to satisfy the so-called trace equation, that is, the expected value of the Hamiltonian function grows linearly over time. Burrage et. al. in [68, 74] demonstrated that not all numerical discretizations are able to preserve this linear growth over time. Therefore the authors derived stochastic Runge-Kutta methods, involving more Wiener processes per step, proving that these exactly preserve the trace equation, at least for quadratic Hamiltonians. In this Ph.D. thesis, we have tried to give a reason for this lack through perturbative analysis via ϵ expansions, being ϵ the amplitude of the stochastic part of the right-hand side. This analysis shows the presence of spurious terms, growing in time and with the parameter ϵ.
The analysis was therefore extended to the case of small multiplicative noise, firstly providing a characterization of the behavior of averaged Hamiltonian arose in such systems, with more emphasis on the separable and quadratic Hamiltonians. Then we showed that, in general, first-order approximations to such systems are not able to retain the same behavior discovered for the exact averaged Hamiltonian. Future perspectives concerning the topic addressed in Chapter 4 concern the analysis of energy preserving methods for Hamiltonian problems with different noises. In particular, relevant interest is the study of numerical methods for stochastic Hamiltonian systems with Levy jumps, widely used especially in physics (for example, the reader can refer to [254]). The goal is to understand if there are invariant laws of the problem and the possible numerical discretizations to preserve these laws that characterize the dynamics in question for a long time.

As has been observed in the development of this thesis, stochastic differential equations are a handy tool for describing phenomena that affect our real life. The last part of this thesis aimed to show models in which stochastic differential equations or partial differential stochastic ones, occur. First, the FitzHugh - Nagumo model was presented, which describes how action potentials propagate along an axon. Starting from this model, a splitting schemes, based on the Lie-Trotter strategy, were developed where the contribution of the nonlinearities of the diffusion operators and of the noise can be integrated separately and combined to give an explicit, easy-to-implement and effective numerical scheme. Starting from this method, a mean-square convergence theorem has been proved, in which it has been outlined that the order of convergence is at most 1/4. Some difficulties we need to overcome are due to the low regularity properties of the solutions, the fact that the system is not a parabolic PDE system, and the cubic growth of the nonlinearity. The choice to use the Lie-Trotter strategy was based on the fact that there is a clear separation between the linear and stochastic parts in the FhN system considered. This is only a first result obtained in this context, and surely a future direction will involve the construction of splitting methods based on Strang's strategy, with a higher order of convergence. There are other directions that this research could be developed. First, through the numerical experiments that have

been conducted, it has been shown that for the scheme LTEXACT–(5.21) the order of convergence is 1/2, which is not covered by Theorem 5.3.2. The fact that using an accelerated exponential Euler scheme where the stochastic convolution is computed exactly yields higher order of convergence is known for parabolic semilinear stochastic PDEs driven by space-time white noise, under appropriate conditions, see for instance [157] or [40]. However, the stochastic FitzHugh–Nagumo equations considered in this article are not parabolic systems; therefore, it is not known how to prove the observed higher order strong rate of convergence. Furthermore, it is well known in the deterministic case that splitting methods, when applied to partial differential equations with particular boundary conditions, suffer from the phenomenon of order reduction (for example, the reader can read [124, 125]). Therefore we want to study whether this same phenomenon also occurs in the case of stochastic partial differential equations.

The analysis carried out in the last part of this Ph.D. thesis is helpful to give a measure, suggested by the stiffness ratio, of the speed of reaffirmation of the truth after the spread out of fake news. In particular, the analysis suggests to use SIR models with high stiffness ratio to describe the diffusion of fake information when the country is exposed to a slower transit of fake news. Less stiff models are particularly suitable when the transit of fake news is slower and its survival time in the exposed population is higher. Following this line, we conducted the same analysis using a stochastic SIR model through a stochastic perturbation of the deterministic SIR model. The research clearly turned out to be more complex deriving from the fact that a precise definition of the stiffness index is not yet well known. In fact, here, this property has been highlighted by the fact that the more stiff the problem was, the more the stochastic Backward Euler method showed a strong order reduction [30, 31]. The employed model is the standard stochastic SIR system of stochastic differential equations, but certainly, more complex models may be used in order to describe the diffusion of fake news as an epidemic phenomenon. For example, one may extend the analysis conducted in [117], in which the authors have proposed a novel and alternative interpretation of the SEIR model to describe the diffusion of fake information on the web and the consequent truth reaffirmation to the stochastic context. Other challenging features of this problem are the study of sentimental analysis, social network analysis and applications of machine learning techniques, see [92, 104].

Bibliography

[1] S. S. Abdullaev. *Construction of mappings for Hamiltonian systems and their applications*, volume 691. Springer, 2006.

[2] A. Abdulle. Explicit methods for stiff stochastic differential equations. In B. Engquist, O. Runborg, and Y.-H. R. Tsai, editors, *Numerical Analysis of Multiscale Computations*, pages 1–22, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[3] A. Abdulle, I. Almuslimani, and G. Vilmart. Optimal explicit stabilized integrator of weak order 1 for stiff and ergodic stochastic differential equations. *SIAM/ASA Journal on Uncertainty Quantification*, 6(2):937–964, 2018.

[4] A. Abdulle and A. Blumenthal. Stabilized multilevel Monte Carlo method for stiff stochastic differential equations. *Journal of Computational Physics*, 251:445–460, 2013.

[5] A. Abdulle and S. Cirilli. Stabilized methods for stiff stochastic systems. *Comptes Rendus Mathematique*, 345(10):593–598, 2007.

[6] A. Abdulle and S. Cirilli. S-ROCK: Chebyshev methods for stiff stochastic differential equations. *SIAM Journal on Scientific Computing*, 30(2):997–1014, 2008.

[7] A. Abdulle, G. Vilmart, and K. C. Zygalakis. Mean-square a-stable diagonally drift-implicit integrators of weak second order for stiff itô stochastic differential equations. *BIT Numerical Mathematics*, 53(4):827–840, 2013.

[8] A. Abdulle, G. Vilmart, and K. C. Zygalakis. Weak second order explicit stabilized methods for stiff stochastic differential equations. *SIAM Journal on Scientific Computing*, 35(4):A1792–A1814, 2013.

[9] M. Ableidinger and E. Buckwar. Splitting integrators for the stochastic Landau-Lifshitz equation. *SIAM J. Sci. Comput.*, 38(3):A1788–A1806, 2016.

[10] R. Ahmad. Introduction to stochastic differential equations, 1988.

[11] A. Alamo and J. M. Sanz-Serna. A technique for studying strong and weak local errors of splitting stochastic integrators. *SIAM J. Numer. Anal.*, 54(6):3239–3257, 2016.

[12] H. Allcott, M. Gentzkow, and C. Yu. Trends in the diffusion of misinformation on social media. *Research & Politics*, 6(2):2053168019848554, 2019.

[13] A. Amirdjanova and S. Chivoret. New method for optimal nonlinear filtering of noisy observations by multiple stochastic fractional integral expansions. *Computers & Mathematics with Applications*, 52(1-2):161–178, 2006.

[14] C. Anton. Weak backward error analysis for stochastic Hamiltonian systems. *BIT Numerical Mathematics*, 59(3):613–646, 2019.

[15] R. Anton, D. Cohen, S. Larsson, and X. Wang. Full discretization of semilinear stochastic wave equations driven by multiplicative noise. *SIAM Journal on Numerical Analysis*, 54(2):1093–1119, 2016.

[16] S. S. Artemiev and T. A. Averina. Numerical analysis of systems of ordinary and stochastic differential equations. In *Numerical Analysis of Systems of Ordinary and Stochastic Differential Equations*. de Gruyter, 2011.

[17] K. Ba and A. Gakwaya. Thermomechanical total Lagrangian SPH formulation for solid mechanics in large deformation problems. *Computer Methods in Applied Mechanics and Engineering*, 342:458–473, 2018.

[18] A. Babiano, J. H. Cartwright, O. Piro, and A. Provenzale. Dynamics of a small neutrally buoyant sphere in a fluid and targeting in Hamiltonian systems. *Physical Review Letters*, 84(25):5764, 2000.

[19] N. Bacaër. *Verhulst and the logistic equation (1838)*, pages 35–39. Springer London, London, 2011.

[20] A. Barth and A. Lang. Simulation of stochastic partial differential equations using finite element methods. *Stochastics An International Journal of Probability and Stochastic Processes*, 84(2-3):217–231, 2012.

[21] A. Bazzani. Hamiltonian systems and stochastic processes. *Lecture Notes, University of Bologna*, 2018.

[22] A. Bazzani, S. Siboni, and G. Turchetti. Diffusion in Hamiltonian systems with a small stochastic perturbation. *Physica D: Nonlinear Phenomena*, 76(1-3):8–21, 1994.

[23] M. Beccari, M. Hutzenthaler, A. Jentzen, R. Kurniawan, F. Lindner, and D. Salimova. Strong and weak divergence of exponential and linear-implicit Euler approximations for stochastic partial differential equations with superlinearly growing nonlinearities, 2019.

[24] M. Begon, J. L. Harper, C. R. Townsend, et al. *Ecology. Individuals, populations and communities.* Blackwell scientific publications, 1986.

[25] A. Bellen, Z. Jackiewicz, R. Vermiglio, and M. Zennaro. Natural continuous extensions of Runge-Kutta methods for Volterra integral equations of the second kind and their applications. *mathematics of computation*, 52(185):49–63, 1989.

[26] A. Bellen, Z. Jackiewicz, R. Vermiglio, and M. Zennaro. Stability analysis of Runge-Kutta methods for Volterra integral equations of the second kind. *IMA journal of numerical analysis*, 10(1):103–118, 1990.

[27] G. Belot. Understanding electromagnetism. *The British Journal for the Philosophy of Science*, 49(4):531–555, 1998.

[28] A. Berg, D. Cohen, and G. Dujardin. Lie–trotter splitting for the nonlinear stochastic Manakov system. *Journal of Scientific Computing*, 88(1):1–31, 2021.

[29] A. Berg, D. Cohen, and G. Dujardin. Lie-Trotter splitting for the nonlinear stochastic Manakov system. *J. Sci. Comput.*, 88(1):Paper No. 6, 31, 2021.

[30] W.-J. Beyn, E. Isaak, and R. Kruse. Stochastic C-stability and B-consistency of explicit and implicit euler-type schemes. *Journal of Scientific Computing*, 67(3):955–987, 2016.

[31] H. B. Bjørkøy. Convergence of stiff, nonlinear stochastic differential equations. Master's thesis, NTNU, 2020.

[32] S. Blanes and F. Casas. *A concise introduction to geometric numerical integration.* Monographs and Research Notes in Mathematics. CRC Press, Boca Raton, FL, 2016.

[33] N. Bou-Rabee and H. Owhadi. Long-run accuracy of variational integrators in the stochastic context. *SIAM J. Numer. Anal.*, 48(1):278–297, 2010.

[34] M. Boulakia, A. Genadot, and M. Thieullen. Simulation of SPDEs for excitable media using finite elements. *J. Sci. Comput.*, 65(1):171–195, 2015.

[35] J. Brainard, P. Hunter, and I. R. Hall. An agent-based model about the effects of fake news on a norovirus outbreak. *Revue d'epidemiologie et de sante publique*, 68(2):99–107, 2020.

[36] F. Brauer and C. Kribs. *Dynamical Systems for Biological Modeling.* CRC Press, 2016.

[37] C. A. Braumann. Itô versus Stratonovich calculus in random population growth. *Mathematical biosciences*, 206(1):81–107, 2007.

[38] C.-E. Bréhier. A short introduction to Stochastic PDEs. Apr. 2014.

[39] C.-E. Bréhier. Approximation of the invariant measure with an Euler scheme for stochastic PDEs driven by space-time white noise. *Potential Analysis*, 40(1):1–40, 2014.

[40] C.-E. Bréhier. Analysis of a modified Euler scheme for parabolic semilinear stochastic PDEs. *Preprint*, 2022.

[41] C.-E. Bréhier and D. Cohen. Analysis of a splitting scheme for a class of nonlinear stochastic Schrödinger equations. *Submitted*, 2020.

[42] C.-E. Bréhier and D. Cohen. Strong rates of convergence of a splitting scheme for Schrödinger equations with nonlocal interaction cubic nonlinearity and white noise dispersion. *SIAM/ASA Journal on Uncertainty Quantification*, 10(1):453–480, 2022.

[43] C.-E. Bréhier and D. Cohen. Strong rates of convergence of a splitting scheme for Schrödinger equations with nonlocal interaction cubic nonlinearity and white noise dispersion. *SIAM/ASA J. Uncertain. Quantif.*, 10(1):453–480, 2022.

[44] C.-E. Bréhier, D. Cohen, and G. Giordano. Splitting schemes for FitzHugh–Nagumo stochastic partial differential equations. *arXiv preprint arXiv:2207.10484*, 2022.

[45] C.-E. Bréhier, J. Cui, and J. Hong. Strong convergence rates of semidiscrete splitting approximations for the stochastic Allen-Cahn equation. *IMA J. Numer. Anal.*, 39(4):2096–2134, 2019.

[46] C.-E. Bréhier and L. Goudenège. Analysis of some splitting schemes for the stochastic Allen-Cahn equation. *Discrete Contin. Dyn. Syst. Ser. B*, 24(8):4169–4190, 2019.

[47] C.-E. Bréhier and L. Goudenège. Weak convergence rates of splitting schemes for the stochastic Allen-Cahn equation. *BIT*, 60(3):543–582, 2020.

[48] C.-E. Bréhier and G. Vilmart. High order integrator for sampling the invariant distribution of a class of parabolic stochastic PDEs with additive space-time noise. *SIAM Journal on Scientific Computing*, 38(4):A2283–A2306, 2016.

[49] T. J. Bridges and S. Reich. Numerical methods for Hamiltonian PDEs. *Journal of Physics A: mathematical and general*, 39(19):5287, 2006.

[50] D. C. Brody and D. M. Meier. How to model fake news. *arXiv preprint arXiv:1809.00964*, 2018.

[51] L. Brugnano, F. Iavernaro, and T. Susca. Hamiltonian BVMs (HBVMs): implementation details and applications. In *AIP Conference Proceedings*, volume 1168, pages 723–726. American Institute of Physics, 2009.

[52] L. Brugnano, F. Iavernaro, and D. Trigiante. Hamiltonian BVMs (HBVMs): a family of "drift free" methods for integrating polynomial Hamiltonian problems. In *AIP Conference Proceedings*, volume 1168, pages 715–718. American Institute of Physics, 2009.

[53] L. Brugnano, F. Iavernaro, and D. Trigiante. Hamiltonian boundary value methods (energy preserving discrete line integral methods). *J. Numer. Anal. Ind. Appl. Math*, 5(1-2):17–37, 2010.

[54] L. Brugnano, F. Iavernaro, and D. Trigiante. The lack of continuity and the role of infinite and infinitesimal in numerical methods for ODEs: the case of symplecticity. *Applied Mathematics and Computation*, 218(16):8056–8063, 2012.

[55] H. Brunner. *Collocation methods for Volterra integral and related functional differential equations*, volume 15. Cambridge university press, 2004.

[56] H. Brunner and P. van deHouwen. The numerical solution of Volterra equations. *CWI monographs*, 1986.

[57] A. Bryden and D. J. Higham. On the boundedness of asymptotic stability regions for the stochastic theta method. *BIT Numerical Mathematics*, 43(1):1–6, 2003.

[58] C.-E. Bréhier, D. Cohen, and T. Jahnke. Splitting integrators for stochastic Lie–Poisson systems. *Submitted*, 2021.

[59] E. Buckwar and R. D'Ambrosio. Exponential mean-square stability properties of stochastic linear multistep methods. *Advances in Computational Mathematics*, 47(4):1–14, 2021.

[60] E. Buckwar, R. Horváth-Bokor, and R. Winkler. Asymptotic mean-square stability of two-step methods for stochastic ordinary differential equations. *BIT Numerical Mathematics*, 46(2):261–282, 2006.

[61] E. Buckwar, A. Samson, M. Tamborrino, and I. Tubikanec. A splitting method for SDEs with locally lipschitz drift: Illustration on the FitzHugh-Nagumo model. *Applied Numerical Mathematics*, 2022.

[62] E. Buckwar and T. Sickenberger. A comparative linear mean-square stability analysis of Maruyama- and Milstein-type methods. *Mathematics and Computers in Simulation*, 81(6):1110–1127, 2011.

[63] A. Bultheel, C. Díaz-Mendoza, P. González-Vera, and R. Orive. On the convergence of certain Gauss-type quadrature formulas for unbounded intervals. *Mathematics of computation*, 69(230):721–747, 2000.

[64] K. Burrage and P. Burrage. General order conditions for stochastic Runge-Kutta methods for both commuting and non-commuting stochastic ordinary differential equation systems. *Applied Numerical Mathematics*, 28(2-4):161–177, 1998.

[65] K. Burrage, P. Burrage, and J. Belward. A bound on the maximum strong order of stochastic Runge-Kutta methods for stochastic ordinary differential equations. *BIT Numerical Mathematics*, 37(4):771–780, 1997.

[66] K. Burrage and P. M. Burrage. High strong order explicit Runge-Kutta methods for stochastic ordinary differential equations. *Applied Numerical Mathematics*, 22(1-3):81–101, 1996.

[67] K. Burrage and P. M. Burrage. Order conditions of stochastic Runge–Kutta methods by b-series. *SIAM Journal on Numerical Analysis*, 38(5):1626–1646, 2000.

[68] K. Burrage and P. M. Burrage. Low rank Runge–Kutta methods, symplecticity and stochastic Hamiltonian problems with additive noise. *Journal of Computational and Applied Mathematics*, 236(16):3920–3930, 2012.

[69] K. Burrage and T. Tian. The composite Euler method for stiff stochastic differential equations. *Journal of computational and applied mathematics*, 131(1-2):407–426, 2001.

[70] K. Burrage and T. Tian. Stiffly accurate Runge–Kutta methods for stiff stochastic differential equations. *Computer physics communications*, 142(1-3):186–190, 2001.

[71] K. Burrage and T. Tian. Implicit stochastic Runge–Kutta methods for stochastic differential equations. *BIT Numerical Mathematics*, 44(1):21–39, 2004.

[72] P. M. Burrage. Runge-Kutta methods for stochastic differential equations. 1999.

[73] P. M. Burrage and K. Burrage. A variable stepsize implementation for stochastic differential equations. *SIAM Journal on Scientific Computing*, 24(3):848–864, 2003.

[74] P. M. Burrage and K. Burrage. Structure-preserving Runge-Kutta methods for stochastic Hamiltonian equations with additive noise. *Numerical Algorithms*, 65(3):519–532, 2014.

[75] P. M. Burrage, R. Herdiana, and K. Burrage. Adaptive stepsize based on control theory for stochastic differential equations. *Journal of Computational and Applied Mathematics*, 170(2):317–336, 2004.

[76] J. C. Butcher. *The numerical analysis of ordinary differential equations: Runge-Kutta and general linear methods*. Wiley-Interscience, 1987.

[77] J. C. Butcher. *Numerical methods for ordinary differential equations*. John Wiley & Sons, 2016.

[78] J. C. Butcher. *B-series: algebraic analysis of numerical methods*, volume 55. Springer Nature, 2021.

[79] J. C. Butcher and R. D'Ambrosio. Partitioned general linear methods for separable Hamiltonian problems. *Applied Numerical Mathematics*, 117:69–86, 2017.

[80] J. C. Butcher, Y. Habib, A. T. Hill, and T. J. Norton. The control of parasitism in g-symplectic methods. *SIAM Journal on Numerical Analysis*, 52(5):2440–2465, 2014.

[81] G. Capobianco, D. Conte, I. Del Prete, and E. Russo. Stability analysis of fast numerical methods for Volterra integral equations. *Electronic Transactions on Numerical Analysis*, 30:305–322, 2008.

[82] A. Cardone, D. Conte, R. D'Ambrosio, and B. Paternoster. Collocation methods for Volterra integral and integro-differential equations: A review. *axioms*, 7(3):45, 2018.

[83] S. Cerrai. *Second order PDE's in finite and infinite dimension*, volume 1762 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2001. A probabilistic approach.

[84] C. Chen, D. Cohen, R. D'Ambrosio, and A. Lang. Drift-preserving numerical integrators for stochastic Hamiltonian systems. *Advances in Computational Mathematics*, 46(2):1–22, 2020.

[85] P.-L. Chow. *Stochastic partial differential equations*. Chapman and Hall/CRC, 2007.

[86] G. Chowell and J. M. Hyman. *Mathematical and statistical modeling for emerging and re-emerging infectious diseases*. Springer, 2016.

[87] G. Chowell, J. M. Hyman, L. M. Bettencourt, C. Castillo-Chavez, and H. Nishiura. *Mathematical and statistical estimation approaches in epidemiology*. Springer, 2009.

[88] V. Citro, R. D'Ambrosio, and S. Di Giovacchino. A-stability preserving perturbation of Runge–Kutta methods for stochastic differential equations. *Applied Mathematics Letters*, 102:106098, 2020.

[89] D. Cohen and G. Dujardin. Exponential integrators for nonlinear Schrödinger equations with white noise dispersion. *Stochastics and Partial Differential Equations: Analysis and Computations*, 5(4):592–613, 2017.

[90] D. Cohen, T. Matsuo, and X. Raynaud. A multi-symplectic numerical integrator for the two-component Camassa–Holm equation. *Journal of Nonlinear Mathematical Physics*, 21(3):442–453, 2014.

[91] D. Cohen and G. Vilmart. Drift-preserving numerical integrators for stochastic Poisson systems. *International Journal of Computer Mathematics*, 99(1):4–20, 2022.

[92] D. Conte, R. D'Ambrosio, and B. Paternoster. Improved theta-methods for stochastic Volterra integral equations. *Communications in Nonlinear Science and Numerical Simulation*, 93:article number 105528, 2021. https://doi.org/10.1016/j.cnsns.2020.105528.

[93] D. Conte, R. D'Ambrosio, G. Giordano, and B. Paternoster. Continuous extension of euler-maruyama method for stochastic differential equations. In *International Conference on Computational Science and Its Applications*, pages 135–145. Springer, 2021.

[94] D. Conte, R. D'Ambrosio, and B. Paternoster. Improved ϑ-methods for stochastic Volterra integral equations. *Communications in Nonlinear Science and Numerical Simulation*, 93:105528, 2021.

[95] D. Conte, L. G. Ixaru, B. Paternoster, and G. Santomauro. Exponentially-fitted Gauss–Laguerre quadrature rule for integrals over an unbounded interval. *Journal of Computational and Applied Mathematics*, 255:725–736, 2014.

[96] D. Conte and B. Paternoster. Multistep collocation methods for Volterra integral equations. *Applied numerical mathematics*, 59(8):1721–1736, 2009.

[97] D. Conte, B. Paternoster, and G. Santomauro. An exponentially fitted quadrature rule over unbounded intervals. In *AIP Conference Proceedings*, volume 1479, pages 1173–1176. American Institute of Physics, 2012.

[98] C. J. Cotter and S. Reich. Adiabatic invariance and applications: from molecular dynamics to numerical weather prediction. *BIT Numerical Mathematics*, 44(3):439–455, 2004.

[99] S. Cox and J. van Neerven. Convergence rates of the splitting scheme for parabolic linear stochastic Cauchy problems. *SIAM J. Numer. Anal.*, 48(2):428–451, 2010.

[100] M. R. Crisci, E. Russo, and A. Vecchio. Stability results for one-step discretized collocation methods in the numerical treatment of Volterra integral equations. *mathematics of computation*, 58(197):119–134, 1992.

[101] J. Cui, J. Hong, Z. Liu, and W. Zhou. Strong convergence rate of splitting schemes for stochastic nonlinear Schrödinger equations. *J. Differential Equations*, 266(9):5625–5663, 2019.

[102] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 152 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 2014.

[103] R. D'Ambrosio, G. De Martino, and B. Paternoster. Construction of nearly conservative multivalue numerical methods for Hamiltonian problems. *Communications in Applied and Industrial Mathematics*, 3(2), 2013.

[104] R. D'Ambrosio, M. Moccaldi, and B. Paternoster. Adapted numerical methods for advection-reaction-diffusion problems generating periodic wavefronts. *Computers & Mathematics with Applications*, 74(5):1029–1042, 2017. https://doi.org/10.1016/j.camwa.2017.04.023.

[105] R. D'Ambrosio, M. Moccaldi, and B. Paternoster. Numerical preservation of long-term dynamics by stochastic two-step methods. *Discrete & Continuous Dynamical Systems-B*, 23(7):2763, 2018.

[106] A. M. O. De Almeida. *Hamiltonian systems: Chaos and quantization*. Cambridge University Press, 1990.

[107] A. Debussche and J. Printems. Weak order for the discretization of the stochastic heat equation. *Mathematics of computation*, 78(266):845–863, 2009.

[108] J. Deng, C. Anton, and Y. S. Wong. High-order symplectic schemes for stochastic Hamiltonian systems. *Communications in Computational Physics*, 16(1):169–200, 2014.

[109] J. Deters, I. P. Aguiar, and J. Feuerborn. The mathematics of gossip. *CODEE Journal*, 12(1):7, 2019.

[110] S. Di Giovacchino. *Structure-preserving numerical approximation of stochastic evolution problems*. PhD thesis, University of L'Aquila, 2021.

[111] O. Diekmann, H. Heesterbeek, and T. Britton. *Mathematical tools for understanding infectious disease dynamics*, volume 7. Princeton University Press, 2013.

[112] P. Droz-Vincent. Hamiltonian systems of relativistic particles. *Reports on Mathematical Physics*, 8(1):79–101, 1975.

[113] R. Duboscq and R. Marty. Analysis of a splitting scheme for a class of random nonlinear partial differential equations. *ESAIM Probab. Stat.*, 20:572–589, 2016.

[114] R. D'Ambrosio, G. De Martino, and B. Paternoster. Numerical integration of Hamiltonian problems by g-symplectic methods. *Advances in Computational Mathematics*, 40(2):553–575, 2014.

[115] R. D'Ambrosio and S. Di Giovacchino. Mean-square contractivity of stochastic ϑ-methods. *Communications in Nonlinear Science and Numerical Simulation*, 96:105671, 2021.

[116] R. D'Ambrosio and S. Di Giovacchino. Nonlinear stability issues for stochastic Runge-Kutta methods. *Communications in Nonlinear Science and Numerical Simulation*, 94:105549, 2021.

[117] R. D'Ambrosio, P. Díaz de Alba, G. Giordano, and B. Paternoster. A modified SEIR model: Stiffness analysis and application to the diffusion of fake news. In *International Conference on Computational Science and Its Applications*, pages 90–103. Springer, 2022.

[118] R. D'Ambrosio, M. Ferro, Z. Jackiewicz, and B. Paternoster. Two-step almost collocation methods for ordinary differential equations. *Numerical Algorithms*, 53(2):195–217, 2010.

[119] R. D'Ambrosio, G. Giordano, S. Mottola, and B. Paternoster. Stiffness analysis to predict the spread out of fake information. *Future Internet*, 13(9):222, 2021.

[120] R. D'Ambrosio, G. Giordano, and B. Paternoster. Numerical conservation issues for stochastic Hamiltonian problems. In *AIP Conference Proceedings*, volume 2425, page 090007. AIP Publishing LLC, 2022.

[121] R. D'Ambrosio, G. Giordano, B. Paternoster, and A. Ventola. Perturbative analysis of stochastic Hamiltonian problems under time discretizations. *Applied Mathematics Letters*, 120:107223, 2021.

[122] R. D'Ambrosio and Z. Jackiewicz. Construction and implementation of highly stable two-step continuous methods for stiff differential systems. *Mathematics and Computers in Simulation*, 81(9):1707–1728, 2011.

[123] R. D'Ambrosio and C. Scalone. Two-step Runge-Kutta methods for stochastic differential equations. *Applied Mathematics and Computation*, 403:125930, 2021.

[124] L. Einkemmer and A. Ostermann. Overcoming order reduction in diffusion-reaction splitting. part 1: Dirichlet boundary conditions. *SIAM Journal on Scientific Computing*, 37(3):A1577–A1592, 2015.

[125] L. Einkemmer and A. Ostermann. Overcoming order reduction in diffusion-reaction splitting. part 2: Oblique boundary conditions. *SIAM Journal on Scientific Computing*, 38(6):A3741–A3757, 2016.

[126] R. C. Elston and W. Johnson. *Basic biostatistics for geneticists and epidemiologists: a practical approach*. John Wiley & Sons, 2008.

[127] J. M. Epstein. *Nonlinear dynamics, mathematical biology, and social science*. CRC Press, 2018.

[128] D. Fan, G.-P. Jiang, Y.-R. Song, and Y.-W. Li. Novel fake news spreading model with similarity on pso-based networks. *Physica A: Statistical Mechanics and Its Applications*, 549:124319, 2020.

[129] N. Fernández-García. Fake news: una oportunidad para la alfabetización mediática. *Nueva sociedad*, (269), 2017.

[130] J. Franceschi and L. Pareschi. Spreading of fake news, competence and learning: kinetic modelling and numerical approximation. *Philosophical Transactions of the Royal Society A*, 380(2224):20210159, 2022.

[131] T. Gard. *Introduction to stochastic differential equations*. Monographs and Textbooks in Pure and Applied Mathematics. New York, NY, 1988.

[132] G. Giordano, S. Mottola, and B. Paternoster. A short review of some mathematical methods to detect fake news. *International Journal of Circuits, Systems and Signal Processing*, 14:255–265, 2020.

[133] A. Goncharenko, Y. A. Logvin, A. Samson, P. Shapovalov, and S. Turovets. Ermakov Hamiltonian systems in nonlinear optics of elliptic gaussian beams. *Physics Letters A*, 160(2):138–142, 1991.

[134] O. Gonzalez and J. C. Simo. On the stability of symplectic and energy-momentum algorithms for non-linear Hamiltonian systems with symmetry. *Computer methods in applied mechanics and engineering*, 134(3-4):197–222, 1996.

[135] J. Gough. Asymptotic stochastic transformations for nonlinear quantum dynamical systems. *Reports on Mathematical Physics*, 44(3):313–338, 1999.

[136] A. Gray, D. Greenhalgh, L. Hu, X. Mao, and J. Pan. A stochastic differential equation SIS epidemic model. *SIAM Journal on Applied Mathematics*, 71(3):876–902, 2011.

[137] W. Grecksch and H. Lisei. Approximation of stochastic nonlinear equations of Schrödinger type by the splitting method. *Stoch. Anal. Appl.*, 31(2):314–335, 2013.

[138] M. Groß. *Conserving time integrators for nonlinear elastodynamics*. PhD thesis, Technische Universität Kaiserslautern, 2004.

[139] I. Gyöngy, S. Sabanis, and D. Šiška. Convergence of tamed Euler schemes for a class of stochastic evolution equations. *Stoch. Partial Differ. Equ. Anal. Comput.*, 4(2):225–245, 2016.

[140] S. Haber. Stochastic quadrature formulas. *Mathematics of Computation*, 23(108):751–764, 1969.

[141] J. D. Hadjidemetriou. Mapping models for Hamiltonian systems with application to resonant asteroid motion. In *Predictability, Stability, and Chaos in N-Body Dynamical Systems*, pages 157–175. Springer, 1991.

[142] A. Haghighi, S. M. Hosseini, and A. Rößler. Diagonally drift-implicit Runge–Kutta methods of strong order one for stiff stochastic differential systems. *Journal of Computational and Applied Mathematics*, 293:82–93, 2016.

[143] A. Haghighi and A. Rößler. Split-step double balanced approximation methods for stiff stochastic differential equations. *International Journal of Computer Mathematics*, 96(5):1030–1047, 2019.

[144] E. Haier, C. Lubich, and G. Wanner. *Geometric Numerical integration: structure-preserving algorithms for ordinary differential equations*. Springer, 2006.

[145] E. Hairer and C. Lubich. Symmetric multistep methods over long times. *Numerische Mathematik*, 97(4):699–723, 2004.

[146] E. Hairer, S. Nørsett, and G. Wanner. Solving ordinary differential equations, i: Nonstiff problems, springer verlag. *Berlin, Germany [Google Scholar]*, 1993.

[147] M. Han, Q. Ma, and X. Ding. High-order stochastic symplectic partitioned Runge-Kutta methods for stochastic hamiltonian systems with additive noise. *Applied Mathematics and Computation*, 346:575–593, 2019.

[148] V. Henderson. Black–Scholes model. *Wiley StatsRef: Statistics Reference Online*, 2014.

[149] D. J. Higham. A-stability and stochastic mean-square stability. *BIT Numerical Mathematics*, 40(2):404–409, 2000.

[150] D. J. Higham. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM review*, 43(3):525–546, 2001.

[151] D. J. Higham and P. E. Kloeden. *An Introduction to the Numerical Simulation of Stochastic Differential Equations*, volume 169. SIAM, 2021.

[152] D. J. Higham, X. Mao, and A. M. Stuart. Strong convergence of Euler-type methods for nonlinear stochastic differential equations. *SIAM Journal on Numerical Analysis*, 40(3):1041–1063, 2002.

[153] J. Hong, D. Xu, and P. Wang. Preservation of quadratic invariants of stochastic differential equations via Runge–Kutta methods. *Applied Numerical Mathematics*, 87:38–52, 2015.

[154] S. Hottovy, G. Volpe, and J. Wehr. Noise-induced drift in stochastic differential equations with arbitrary friction and diffusion in the Smoluchowski-Kramers limit. *Journal of Statistical Physics*, 146(4):762–773, 2012.

[155] M. Hutzenthaler, A. Jentzen, and P. E. Kloeden. Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally lipschitz continuous coefficients. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 467(2130):1563–1576, 2011.

[156] Z. Jackiewicz. *General linear methods for ordinary differential equations*. John Wiley & Sons, 2009.

[157] A. Jentzen and P. E. Kloeden. Overcoming the order barrier in the numerical approximation of stochastic partial differential equations with additive space-time noise. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 465(2102):649–667, 2009.

[158] C. Jones, S. Wiggins, and G. Haller. Dynamical systems and oceanography. Technical report, BROWN UNIV PROVIDENCE RI DIV OF APPLIED MATHEMATICS, 1994.

[159] D. G. Kendall. Deterministic and stochastic epidemics in closed populations. In *Proceedings of the third Berkeley symposium on mathematical statistics and probability*, volume 4, pages 149–165. University of California Press Berkeley, 1956.

[160] W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.

[161] D. Khoshnevisan. *Analysis of stochastic partial differential equations*, volume 119. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2014.

[162] P. Khurana and D. Kumar. Sir model for fake news spreading through whatsapp. In *Proceedings of 3rd International Conference on Internet of Things and Connected Technologies (ICIoTCT)*, pages 26–27, 2018.

[163] J. Kim, B. Tabibian, A. Oh, B. Schölkopf, and M. Gomez-Rodriguez. Leveraging the crowd to detect and reduce the spread of fake news and misinformation. In *Proceedings of the eleventh ACM international conference on web search and data mining*, pages 324–332, 2018.

[164] I. Kiyosi. On stochastic processes (infinitely divisible laws of probability. *Japanese Journal of Mathematics*, 18:261–301, 1942.

[165] P. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2011.

[166] P. E. Kloeden and R. A. Pearson. The numerical solution of stochastic differential equations. *The Journal of the Australian Mathematical Society. Series B. Applied Mathematics*, 20(1):8–12, 1977.

[167] Y. Komori. Stahle ROW-type weak scheme for stochastic differential equations. 1995.

[168] Y. Komori, D. Cohen, and K. Burrage. Weak second order explicit exponential Runge–Kutta methods for stochastic differential equations. *SIAM Journal on Scientific Computing*, 39(6):A2857–A2878, 2017.

[169] Y. Komori, T. Mitsui, and H. Sugiura. Rooted tree analysis of the order conditions of ROW-type scheme for stochastic differential equations. *BIT Numerical Mathematics*, 37(1):43–66, 1997.

[170] A. Kossakowski. On quantum statistical mechanics of non-Hamiltonian systems. *Reports on Mathematical Physics*, 3(4):247–274, 1972.

[171] M. Kovács, S. Larsson, and F. Lindgren. Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise. *BIT Numerical Mathematics*, 52(1):85–108, 2012.

[172] R. Kupferman, G. A. Pavliotis, and A. M. Stuart. Itô versus Stratonovich white-noise limits for systems with inertia and colored multiplicative noise. *Physical Review E*, 70(3):036120, 2004.

[173] D. Küpper, A. Kværnø, and A. Rößler. Stability analysis and classification of Runge–Kutta methods for index 1 stochastic differential-algebraic equations with scalar noise. *Applied Numerical Mathematics*, 96:24–44, 2015.

[174] J. Lambert. *Numerical Methods for Ordinary Systems: The Initial Value Problem*. Chichester, 1991.

[175] A. Lang, A. Petersson, and A. Thalhammer. Mean-square stability analysis of approximations of stochastic differential equations in infinite dimensions. *BIT Numerical Mathematics*, 57(4):963–990, 2017.

[176] D. M. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, et al. The science of fake news. *Science*, 359(6380):1094–1096, 2018.

[177] B. Leimkuhler and C. Matthews. *Molecular dynamics*, volume 39 of *Interdisciplinary Applied Mathematics*. Springer, Cham, 2015. With deterministic and stochastic numerical methods.

[178] B. Leimkuhler and S. Reich. *Simulating Hamiltonian dynamics*, volume 14 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2004.

[179] T. Li, K. Wu, M. Yan, Z. Liu, and H. Zheng. Stochastic dynamic behavior of FitzHugh-Nagumo neurons stimulated by white noise. *Internat. J. Modern Phys. B*, 35(10):Paper No. 2150137, 24, 2021.

[180] I. Lie and S. P. Nørsett. Superconvergence for multistep collocation. *Mathematics of computation*, 52(185):65–79, 1989.

[181] B. Lindner, J. Garcıa-Ojalvo, A. Neiman, and L. Schimansky-Geier. Effects of noise in excitable systems. *Physics reports*, 392(6):321–424, 2004.

[182] J. Liu. A mass-preserving splitting scheme for the stochastic Schrödinger equation with multiplicative noise. *IMA J. Numer. Anal.*, 33(4):1469–1479, 2013.

[183] Q. F. Lotito, D. Zanella, and P. Casari. Realistic aspects of simulation models for fake news epidemics over social networks. *Future Internet*, 13(3):76, 2021.

[184] Y. Lv and W. Wang. Limit dynamics for the stochastic FitzHugh-Nagumo system. *Nonlinear Anal. Real World Appl.*, 11(4):3091–3105, 2010.

[185] Q. Ma, D. Ding, and X. Ding. Symplectic conditions and stochastic generating functions of stochastic Runge–Kutta methods for stochastic Hamiltonian systems with multiplicative noise. *Applied Mathematics and Computation*, 219(2):635–643, 2012.

[186] Z. Ma. *Dynamical modeling and analysis of epidemics*. World Scientific, 2009.

[187] V. Mackevičius and J. Navikas. Second order weak Runge–Kutta type methods for Itô equations. *Mathematics and computers in simulation*, 57(1-2):29–34, 2001.

[188] H. Mahmoud. A model for the spreading of fake news. *Journal of Applied Probability*, 57(1):332–342, 2020.

[189] H. Malchow. *Spatiotemporal patterns in ecology and epidemiology: theory, models, and simulation*. Chapman and Hall/CRC, 2007.

[190] R. Mannella and P. V. McClintock. Itô versus Stratonovich: 30 years later. *The Random and Fluctuating World: Celebrating Two Decades of Fluctuation and Noise Letters*, pages 9–18, 2022.

[191] X. Mao. *Stochastic differential equations and applications*. Elsevier, 2007.

[192] V. Z. Marmarelis. Appendix ii: Gaussian white noise. 2004.

[193] E. Martınez, L. López-Dıaz, L. Torres, and O. Alejos. On the interpretations of Langevin stochastic equation in different coordinate systems. *Physica B: Condensed Matter*, 343(1-4):252–256, 2004.

[194] R. Marty. On a splitting scheme for the nonlinear Schrödinger equation in a random medium. *Commun. Math. Sci.*, 4(4):679–705, 2006.

[195] R. Marty. Local error of a splitting scheme for a nonlinear Schrödinger-type equation with random dispersion. *Commun. Math. Sci.*, 19(4):1051–1069, 2021.

[196] J. Mathiesen, L. Angheluta, P. T. Ahlgren, and M. H. Jensen. Excitable human dynamics driven by extrinsic events in massive communities. *Proceedings of the National Academy of Sciences*, 110(43):17259–17262, 2013.

[197] J. C. Mattingly, A. M. Stuart, and D. J. Higham. Ergodicity for sdes and approximations: locally lipschitz vector fields and degenerate noise. *Stochastic processes and their applications*, 101(2):185–232, 2002.

[198] S. Mauthner. Step size control in the numerical solution of stochastic differential equations. *Journal of computational and applied mathematics*, 100(1):93–109, 1998.

[199] R. I. McLachlan and G. R. W. Quispel. Splitting methods. *Acta Numer.*, 11:341–434, 2002.

[200] E. J. McShane. *Stochastic calculus and stochastic models*, volume 25. Academic Press, 2014.

[201] G. N. Milstein, Y. M. Repin, and M. V. Tretyakov. Numerical methods for stochastic systems preserving symplectic structure. *SIAM Journal on Numerical Analysis*, 40(4):1583–1604, 2002.

[202] G. N. Milstein, Y. M. Repin, and M. V. Tretyakov. Symplectic integration of Hamiltonian systems with additive noise. *SIAM Journal on Numerical Analysis*, 39(6):2066–2088, 2002.

[203] G. N. Milstein and M. V. Tretyakov. *Stochastic numerics for mathematical physics*, volume 456. Springer, 2004.

[204] G. Mil'shtejn. Approximate integration of stochastic differential equations. *Theory of Probability & Its Applications*, 19(3):557–562, 1975.

[205] T. Misawa. A Lie algebraic approach to numerical integration of stochastic differential equations. *SIAM J. Sci. Comput.*, 23(3):866–890, 2001.

[206] D. Mishra, A. Yadav, S. Ray, and P. K. Kalra. Effects of Noise on the Dynamics of Biological Neuron Models. In A. Abraham, Y. Dote, T. Furuhashi, M. Köppen, A. Ohuchi, and Y. Ohsawa, editors, *Soft Computing as Transdisciplinary Science and Technology*, pages 61–69, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.

[207] G. Monegato and L. Scuderi. Quadrature rules for unbounded intervals and their application to integral equations. In *Approximation and Computation*, pages 185–208. Springer, 2010.

[208] S. Mottola. Fake news as social phenomena. *Linguistic analysis and persuasion in hoax in Italian and Spanish, to appear in Discurso & sociedad, Barcelona*.

[209] T. Murayama, S. Wakamiya, E. Aramaki, and R. Kobayashi. Modeling the spread of fake news on twitter. *Plos one*, 16(4):e0250419, 2021.

[210] P. J. Olver. Hamiltonian and non-Hamiltonian models for water waves. In *Trends and applications of pure mathematics to mechanics*, pages 273–290. Springer, 1984.

[211] J. L. Padgett and Q. Sheng. Convergence of an operator splitting scheme for abstract stochastic evolution equations. In *Advances in mathematical methods and high performance computing*, volume 41 of *Adv. Mech. Math.*, pages 163–179. Springer, Cham, 2019.

[212] G. Pesce, A. McDaniel, S. Hottovy, J. Wehr, and G. Volpe. Stratonovich-to-Itô transition in noisy systems with multiplicative feedback. *Nature communications*, 4(1):1–7, 2013.

[213] J. Printems. On the discretization in time of parabolic stochastic partial differential equations. *M2AN Math. Model. Numer. Anal.*, 35(6):1055–1078, 2001.

[214] L. Qin, D. Ma, and J. Shu. Wong-Zakai approximations and long term behavior of stochastic FitzHugh-Nagumo system. *Int. J. Biomath.*, 14(3):Paper No. 2150008, 30, 2021.

[215] J. Ren, X.-H. Meng, and L. Zhao. Hamiltonian formalism in friedmann cosmology and its quantization. *Physical Review D*, 76(4):043521, 2007.

[216] A. Reyes-Menendez, J. R. Saura, and F. Filipe. The importance of behavioral data to identify online fake reviews for tourism businesses: a systematic review. *PeerJ Computer Science*, 5:e219, 2019.

[217] A. Reyes-Menendez, J. R. Saura, and J. G. Martinez-Navalon. The impact of e-wom on hotels management reputation: exploring tripadvisor review credibility with the elm model. *IEEE Access*, 7:68868–68877, 2019.

[218] R. Ross. An application of the theory of probabilities to the study of a priori pathometry.—part i. *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character*, 92(638):204–230, 1916.

[219] A. Rößler. Runge–Kutta methods for stratonovich stochastic differential equation systems with commutative noise. *Journal of computational and applied mathematics*, 164:613–627, 2004.

[220] A. Rößler. Rooted tree analysis for order conditions of stochastic Runge-Kutta methods for the weak approximation of stochastic differential equations. *Stochastic analysis and applications*, 24(1):97–134, 2006.

[221] A. Rössler. Runge–Kutta methods for Itô stochastic differential equations with scalar noise. *BIT Numerical Mathematics*, 46(1):97–110, 2006.

[222] A. Rößler. Second order Runge–Kutta methods for Stratonovich stochastic differential equations. *BIT Numerical Mathematics*, 47(3):657–680, 2007.

[223] A. Rößler. Second order Runge–Kutta methods for Itô stochastic differential equations. *SIAM Journal on Numerical Analysis*, 47(3):1713–1738, 2009.

[224] A. Rößler. Runge–Kutta methods for the strong approximation of solutions of stochastic differential equations. *SIAM Journal on Numerical Analysis*, 48(3):922–952, 2010.

[225] H. J. Rothe and K. D. Rothe. *Classical and quantum dynamics of constrained Hamiltonian systems*, volume 81. World Scientific, 2010.

[226] W. Rüemelin. Numerical treatment of stochastic differential equations. *SIAM Journal on Numerical Analysis*, 19(3):604–613, 1982.

[227] Y. Saito and T. Mitsui. Stability analysis of numerical schemes for stochastic differential equations. *SIAM Journal on Numerical Analysis*, 33(6):2254–2267, 1996.

[228] A. Samsudin, N. Rosli, and A. Ariffin. Stability analysis of explicit and implicit stochastic Runge-Kutta methods for stochastic differential equations. In *Journal of Physics: Conference Series*, volume 890, page 012084. IOP Publishing, 2017.

[229] J.-M. Sanz-Serna and M.-P. Calvo. *Numerical Hhamiltonian problems*. Courier Dover Publications, 2018.

[230] M. Sauer and W. Stannat. Lattice approximation for stochastic reaction diffusion equations with one-sided Lipschitz condition. *Math. Comp.*, 84(292):743–766, 2015.

[231] M. Sauer and W. Stannat. Analysis and approximation of stochastic nerve axon equations. *Math. Comp.*, 85(301):2457–2481, 2016.

[232] L. F. Shampine. Vectorized adaptive quadrature in matlab. *Journal of Computational and Applied Mathematics*, 211(2):131–140, 2008.

[233] L. F. Shampine. *Numerical solution of ordinary differential equations*. Routledge, 2018.

[234] T. G. Shepherd. Hamiltonian geophysical fluid dynamics. 2003.

[235] K. Sobczyk. Introduction: Origin of stochastic differential equations. In *Stochastic Differential Equations*, pages 1–5. Springer, 1991.

[236] A. R. Soheili and M. Namjoo. Strong approximation of stochastic differential equations with Runge-Kutta methods. *World Journal of Modelling and Simulation*, 4(2):83–93, 2008.

[237] B. Spagnolo and E. V. Pankratova. Influence of noise sources on FitzHugh-Nagumo model in suprathreshold regime. In N. G. Stocks, D. Abbott, and R. P. Morse, editors, *Fluctuations and Noise in Biological, Biophysical, and Biomedical Systems III*, volume 5841, pages 174 – 185. SPIE, 2005.

[238] Y. Suzuki and S. Koshizuka. A Hamiltonian particle method for nonlinear elastodynamics. *International Journal for Numerical Methods in Engineering*, 74(8):1344–1373, 2008.

[239] Y. Takeuchi. *Global Dynamical Properties Of Lotka-Volterra Systems*. World Scientific Publishing Company, 1996.

[240] N. Tarnow. *Energy and momentum conserving algorithms for Hamiltonian systems in the nonlinear dynamics of solids*. Stanford University, 1993.

[241] M. Thieullen. Deterministic and stochastic FitzHugh-Nagumo systems. In *Stochastic biomathematical models*, volume 2058 of *Lecture Notes in Math.*, pages 175–186. Springer, Heidelberg, 2013.

[242] T. Tian and K. Burrage. Two-stage stochastic Runge-Kutta methods for stochastic differential equations. *BIT Numerical Mathematics*, 42(3):625–643, 2002.

[243] A. Tocino. Mean-square stability of second-order Runge–Kutta methods for stochastic differential equations. *Journal of computational and applied mathematics*, 175(2):355–367, 2005.

[244] A. Tocino and R. Ardanuy. Runge–Kutta methods for numerical solution of stochastic differential equations. *Journal of Computational and Applied Mathematics*, 138(2):219–241, 2002.

[245] A. Tocino and J. Vigo-Aguiar. Weak second order conditions for stochastic Runge–Kutta methods. *SIAM Journal on Scientific Computing*, 24(2):507–523, 2002.

[246] H. C. Tuckwell. Analytical and simulation results for the stochastic spatial FitzHugh-Nagumo model neuron. *Neural Comput.*, 20(12):3003–3033, 2008.

[247] H. C. Tuckwell. Stochastic partial differential equations in neurobiology: linear and nonlinear models for spiking neurons. In *Stochastic biomathematical models*, volume 2058 of *Lecture Notes in Math.*, pages 149–173. Springer, Heidelberg, 2013.

[248] N. G. Van Kampen. Itô versus Stratonovich. *Journal of Statistical Physics*, 24(1):175–187, 1981.

[249] S. Vosoughi, D. Roy, and S. Aral. The spread of true and false news online. *science*, 359(6380):1146–1151, 2018.

[250] J. Wainwright and G. F. R. Ellis. *Dynamical systems in cosmology*. 1997.

[251] P. Wang. Three-stage stochastic Runge–Kutta methods for stochastic differential equations. *Journal of Computational and Applied Mathematics*, 222(2):324–332, 2008.

[252] X. Wang. Weak error estimates of the exponential euler scheme for semi-linear SPDEs without Malliavin calculus. *arXiv preprint arXiv:1408.0713*, 2014.

[253] G. Wanner and E. Hairer. *Solving ordinary differential equations II*, volume 375. Springer Berlin Heidelberg New York, 1996.

[254] P. Wei, Y. Chao, and J. Duan. Hamiltonian systems with lévy noise: Symplecticity, Hamilton's principle and averaging principle. *Physica D: Nonlinear Phenomena*, 398:69–83, 2019.

[255] D. Wright. The digital simulation of stochastic differential equations. *IEEE Transactions on Automatic Control*, 19(1):75–76, 1974.

[256] Y. Yalcin, L. G. SÜMER, and S. Kurtulan. Discrete-time modeling of Hamiltonian systems. *Turkish Journal of Electrical Engineering and Computer Sciences*, 23(1):149–170, 2015.

[257] M. E. Yamakou, T. D. Tran, L. H. Duc, and J. Jost. The stochastic FitzHugh-Nagumo neuron model in the excitable regime embeds a leaky integrate-and-fire model. *J. Math. Biol.*, 79(2):509–532, 2019.

[258] G. M. Zaslavsky. *The physics of chaos in Hamiltonian systems*. world scientific, 2007.

[259] Z. Zhao, J. Zhao, Y. Sano, O. Levy, H. Takayasu, M. Takayasu, D. Li, J. Wu, and S. Havlin. Fake news propagates differently from real news even at early stages of spreading. *EPJ data science*, 9(1):7, 2020.

[260] K. C. Zygalakis. On the existence and the applications of modified equations for stochastic differential equations. *SIAM Journal on Scientific Computing*, 33(1):102–130, 2011.