Exponential Integrators for Stochastic Partial Differential Equations

Rikard Anton
Abstract

Stochastic partial differential equations (SPDEs) have during the past decades become an important tool for modeling systems which are influenced by randomness. Because of the complex nature of SPDEs, knowledge of efficient numerical methods with good convergence and geometric properties is of considerable importance. Due to this, numerical analysis of SPDEs has become an important and active research field.

The thesis consists of four papers, all dealing with time integration of different SPDEs using exponential integrators. We analyse exponential integrators for the stochastic wave equation, the stochastic heat equation, and the stochastic Schrödinger equation. Our primary focus is to study strong order of convergence of temporal approximations. However, occasionally, we also analyse space approximations such as finite element and finite difference approximations. In addition to this, for some SPDEs, we consider conservation properties of numerical discretizations.

As seen in this thesis, exponential integrators for SPDEs have many benefits over more traditional integrators such as Euler–Maruyama schemes or the Crank–Nicolson–Maruyama scheme. They are explicit and therefore very easy to implement and use in practice. Also, they are excellent at handling stiff problems, which naturally arise from spatial discretizations of SPDEs. While many explicit integrators suffer step size restrictions due to stability issues, exponential integrators do not in general.

In Paper 1 we consider a full discretization of the stochastic wave equation driven by multiplicative noise. We use a finite element method for the spatial discretization, and for the temporal discretization we use a stochastic trigonometric method. In the first part of the paper, we prove mean-square convergence of the full approximation. In the second part, we study the behavior of the total energy, or Hamiltonian, of the wave equation. It is well known that for deterministic (Hamiltonian) wave equations, the total energy remains constant in time. We prove that for stochastic wave equations with additive noise, the expected energy of the exact solution grows linearly with time. We also prove that the numerical approximation produces a small error in this linear drift.

In the second paper, we study an exponential integrator applied to the time discretization of the stochastic Schrödinger equation with a multiplicative potential. We prove strong convergence order 1 and $\frac{1}{2}$ for additive and multiplicative noise, respectively. The deterministic linear Schrödinger equation has several conserved quantities, including the energy, the mass, and the momentum. We first show that for Schrödinger equations driven by additive noise, the expected values of these quantities grow linearly with time. The exponential integrator is shown to preserve these linear drifts for all time in the case of a stochastic Schrödinger equation without potential. For the equation with a multiplicative potential, we obtain a small error in these linear drifts.

The third paper is devoted to studying a full approximation of the one-dimensional stochastic heat equation. For the spatial discretization we use a finite difference method and an exponential integrator is used for the temporal approximation. We prove mean-square convergence and almost sure convergence of the approximation when the
coefficients of the problem are assumed to be Lipschitz continuous. For non-Lipschitz coefficients, we prove convergence in probability.

In Paper 4 we revisit the stochastic Schrödinger equation. We consider this SPDE with a power-law nonlinearity. This nonlinearity is not globally Lipschitz continuous and the exact solution is not assumed to remain bounded for all times. These difficulties are handled by considering a truncated version of the equation and by working with stopping times and random time intervals. We prove almost sure convergence and convergence in probability for the exponential integrator as well as convergence orders of $\frac{1}{2} - \varepsilon$, for all $\varepsilon > 0$, and $\frac{1}{2}$, respectively.
Sammanfattning

Stokastiska partiella differentialekvationer (SPDE) har under de senaste decennierna blivit ett viktigt redskap för att modellera fysikaliska system som påverkas av slumpmässiga störningar. Till följd av SPDEs komplexa natur, har utveckling av effektiva numeriska metoder med bra konvergens- och geometriska egenskaper fått stor betydelse. Detta har lett till stor aktivitet inom forskningsområdet numerisk analys för SPDE.


Som vi kommer se i denna avhandling, så har exponentiella integratorer många fördelar gentemot mer traditionella integratorer som Euler–Maruyama-metoder och Crank–Nicolson-metoden. De är explicita och därför lätt att implementera och applicera i praktiken. De är också väldigt bra på att hantera styva problem, vilka naturligt uppstår från rumsdiskretiseringar av SPDE. Medan många explicita integratorer lider av steglängdsbegränsningar på grund av stabilitetsproblem, så gör exponentiella integratorer i allmänhet inte det.


I den tredje artikeln studerar vi en fullständig approximation av den endimensionella stokastiska värmelandningsekvationen. För rumsdiskretiseringen använder vi en finita differensmetod och för tidsdiskretiseringen använder vi en exponentiell integrator. Vi bevisar konvergens i kvadratiskt medel och konvergens nästan överallt.
för approximationen när koefficienterna antags vara Lipschitzkontinuerliga. För ko-
efficienter som inte är Lipschitzkontinuerliga bevisar vi konvergens i sannolikhet för approximationen.

gensordningar $\frac{1}{2} - \varepsilon$, för alla $\varepsilon > 0$, respektive $\frac{1}{2}$.
Acknowledgements

Most of all, I would like to thank my supervisor David Cohen. I feel very grateful for you having taken the time to support and teach me during these years, and for helping me to achieve this. I would like to thank my co-authors Stig Larsson, Xiaojie Wang, and Lluis Quer-Sardanyons for nice collaborations. Thanks also to Annika Lang for her comments on an earlier draft of the thesis.

Many thanks and much love to my parents Inger and Peter, and the rest of my family. I am truly fortunate to have a family who always supports me and whom I can always rely on. I would also like to express my thanks to my colleagues at the department of mathematics and mathematical statistics for making these years very pleasant.

This thesis was partially supported by the Swedish Research Council (VR) (project nr. 2013 – 4562). The computations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at HPC2N, Umeå University.

Rikard Anton
Umeå, April 2018
List of papers

This thesis consists of an introductory chapter and the following papers∗:


IV. R. Anton, *Convergence of an exponential method for the stochastic Schrödinger equation with power-law nonlinearity*.

∗The published papers have been reformatted to fit the style of the thesis and may contain some very minor differences from the published versions.
## Contents

1 Introduction 1

1.1 Notations ......................................................... 2
1.2 The infinite-dimensional Wiener process .......................... 3
1.3 The stochastic integral for Wiener processes ...................... 7
1.4 Stochastic partial differential equations .......................... 9
1.5 Exponential integrators ........................................... 11

2 Summary of the Papers 13

2.1 Paper I – Full discretization of semilinear stochastic wave equations driven by multiplicative noise ................................ 13
2.2 Paper II – Exponential integrators for stochastic Schrödinger equations driven by Itô noise ...................................... 16
2.3 Paper III – A fully discrete approximation of the one-dimensional stochastic heat equation ...................................... 18
2.4 Paper IV – Convergence of an exponential method for the stochastic Schrödinger equation with power-law nonlinearity ....................... 19

References 21

Papers I-IV
# 1 Introduction

Many systems in the real world are influenced by random perturbations. If the randomness has a large effect on the system, deterministic models might not perform well and it may be a good idea to use stochastic models. Due to this, the analysis of stochastic partial differential equations (SPDEs) has during the past decades become an important field of research. As an example of a physical system that can be modelled by an SPDE, let us consider the position and motion of a strand of DNA in a liquid [13]. There are several forces acting on the strand, including elastic forces, friction forces, and, in particular, a random force. Such random force could model the molecules of the surrounding liquid randomly hitting the strand, causing perturbations in its movement and position. Using Newton’s second law to connect the forces to the acceleration, the author of [13] derives a system of hyperbolic SPDEs. Solutions to systems of SPDEs like these and to SPDEs in general are rarely given in an explicit form. Numerical approximation is thus an important tool for studying solutions to SPDEs. Further motivational examples for using SPDEs to describe real world applications can be found in the introductions of [12,21,24,37], to mention but a few.

The increase in research on SPDEs has pushed the numerical analysis of SPDEs to also become an important and active research field. There are several reasons for this. As mentioned, SPDEs are rarely explicitly solvable. The computations are in general expensive and therefore choosing a good numerical method is vital. Also, since the solutions are stochastic processes, one is often interested in expected values and must therefore do many simulations. It is then desirable to have a numerical method that not only approximates the exact solution well but also computes such approximations fast. In other cases one might be interested in a certain geometric property of the solution or the system and therefore one wants a numerical method that preserves this geometric property.

The main thread of this thesis is the study of exponential integrators for the time integration of partial differential equations (PDEs) driven by Gaussian noise. We focus primarily on strong convergence of exponential integrators. Exponential methods were made popular many years ago when it was noticed that they work particularly well for stiff equations [23]. As stiff equations often arise from the spatial discretization of PDEs and SPDEs, the study of exponential integrators is indeed well motivated. The exponential integrators we analyse exhibit many favourable qualities. They are explicit methods and suffer no step size restrictions, and are therefore easy to implement and to use in practice.

We study strong convergence of the exponential integrators when applied to the stochastic wave equation (Paper I), the stochastic heat equation (Paper III), and the stochastic Schrödinger equation (Papers II and IV). In Papers III and IV we also consider almost sure convergence and convergence in probability. In addition, we study the behaviour of the exponential method on quantities that exhibit linear drift in time. Such quantities include the expected energy for the stochastic wave equation
and the stochastic Schrödinger equation, and the expected $L^2$-norm for the latter equation.

The thesis is structured as follows. We begin with an introductory chapter, where some prerequisite material is presented. This includes the cylindrical Wiener process, the stochastic integral with respect to a Wiener process, and some basics on SPDEs. We also include a short subsection on exponential methods. This follows by summaries of the four papers that constitutes the body of the thesis. The four papers then follow, one by one.

1.1 Notations

We consider SPDEs driven by multiplicative Wiener noise. To make sense of these SPDEs, we need to define both the infinite-dimensional Wiener process and the stochastic integral for Wiener processes. This subsection and the next three subsections about the Wiener process, the stochastic integral, and SPDEs are largely based on material from [12, 21, 29, 37]. Our aim in this introductory part is not to give precise proofs and strict mathematical arguments, but rather to give the overall ideas and motivation for the subject. We do not prove the results stated here, but instead provide the reader with references.

We now introduce some notations. From here on, let $(U, \langle \cdot, \cdot \rangle_U)$ and $(H, \langle \cdot, \cdot \rangle_H)$ be separable Hilbert spaces and let $\mathcal{B}(U)$ denote the Borel $\sigma$-algebra of $U$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $\{\mathcal{F}_t\}_{t \geq 0}$ be a normal filtration on $(\Omega, \mathcal{F}, \mathbb{P})$. That is, $\{\mathcal{F}_t\}_{t \geq 0}$ is a filtration such that $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets and $\mathcal{F}_t = \cap_{s > t} \mathcal{F}_s$, for all $t \geq 0$. A stochastic process $\{X(t)\}_{t \geq 0}$ is said to be adapted to the filtration $\{\mathcal{F}_t\}_{t \geq 0}$ if $X(t)$ is $\mathcal{F}_t$-measurable, for each $t \geq 0$. The expectation of an $H$-valued random variable $Y : \Omega \rightarrow H$ is defined as

$$\mathbb{E}[Y] := \int_{\Omega} Y(\omega) d\mathbb{P}(\omega),$$

where the integral is a Bochner integral (see [37]). Denote by $L^2(H)$ the space of square integrable functions on $H$ and let $\mathcal{L}(U,H)$ be the space of bounded linear operators from $U$ to $H$ (written $\mathcal{L}(U)$ if $U = H$). We define the space of Hilbert–Schmidt operators from $U$ to $H$, denoted by $\mathcal{L}_2(U,H)$, which consists of bounded linear operators $T$ such that

$$\|T\|_{\mathcal{L}_2(U,H)} := \left( \sum_{k=1}^{\infty} \|Te_k\|^2_H \right)^{1/2} < \infty,$$

where $\{e_k\}_{k=1}^{\infty}$ is any orthonormal basis of $U$. The trace of an operator $S \in \mathcal{L}(U)$ is defined as

$$\text{Tr}(S) := \sum_{k=1}^{\infty} \langle Se_k, e_k \rangle_U,$$
where, again, \( \{e_k\}_{k=1}^{\infty} \) is any orthonormal basis of \( U \). Finally, we introduce the space \( L^q(\Omega; H) \), for \( q \geq 1 \), of \( H \)-valued random variables \( Y \) such that
\[
\|Y\|_{L^q(\Omega; H)} := \left( \mathbb{E}[\|Y\|_H^q] \right)^{1/q} < \infty.
\]

1.2 The infinite-dimensional Wiener process

We aim to define and make sense of the cylindrical Wiener process and from it define stochastic integrals in Hilbert spaces. In order to do this, we first need the concept of Gaussian measures in infinite dimensions. We say that a measure \( \mu \) on \( (U, \mathcal{B}(U)) \) is Gaussian if for all \( v \in U \) the bounded linear mapping \( v' : U \to \mathbb{R} \) defined by
\[
u \mapsto \langle u, v \rangle_U, \quad u \in U,
\]
has Gaussian law, i.e. for all \( v \in U \) there exist \( m_v \in \mathbb{R} \) and \( \sigma_v \geq 0 \) such that, if \( \sigma_v > 0 \),
\[
(\mu \circ (v')^{-1})(A) := \mu(\{u \in U : v'(u) \in A\}) = \frac{1}{\sqrt{2\pi\sigma_v^2}} \int_A e^{-\frac{(s-m_v)^2}{2\sigma_v^2}} \, ds,
\]
for all \( A \in \mathcal{B}(\mathbb{R}) \). If \( \sigma_v = 0 \), then we require \( \mu \circ (v')^{-1} = \delta_{m_v} \), where \( \delta_{m_v} := \delta(x - m_v) \) is the Dirac delta function.

We can characterize Gaussian measures in terms of their Fourier transform.

**Theorem 1.2** (Theorem 2.3 in [29]). A finite measure \( \mu \) on \( (U, \mathcal{B}(U)) \) is Gaussian if and only if
\[
\hat{\mu}(u) := \int_U e^{i(u,v)} u \, d\mu(v) = e^{i(m,u)} u - \frac{1}{2} (Qu,u) v, \quad u \in U,
\]
where \( m \in U \) and \( Q \in \mathcal{L}(U) \) is non-negative with finite trace. We write \( \mu = N(m, Q) \), and \( m \) and \( Q \) are called the mean and the covariance operator of \( \mu \). The measure \( \mu \) is uniquely determined by \( m \) and \( Q \).

The concept of Gaussian measures and Gaussian laws is used in the definition of the infinite dimensional Wiener process. Particularly in the last point in the definition below, where we require that the increments should have Gaussian law. The infinite dimensional Wiener process is in large defined analogously to the finite dimensional case.

**Definition 1.3** (Definition 2.1.9 in [37]). Fix \( T > 0 \) and let \( Q \in \mathcal{L}(U) \) be non-negative, symmetric and with finite trace. A \( U \)-valued \( Q \)-Wiener process on a probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) is a stochastic process \( \{W(t)\}_{t \in [0,T]} \) that satisfies
• $W(0) = 0$,
• $W$ has continuous trajectories for almost every $\omega \in \Omega$,
• the increments
  \[ W(t_1), W(t_2) - W(t_1), \ldots, W(t_n) - W(t_{n-1}) \]
  are independent for all $0 \leq t_1 < t_2 < \ldots < t_n \leq T$, $n \in \mathbb{N}$,
• the increments have Gaussian law:
  \[ \mathbb{P} \circ (W(t) - W(s))^{-1} = N(0, (t-s)Q), \quad 0 \leq s \leq t \leq T. \]

Remark 1.4. When the context is clear, we will write $Q$-Wiener process or simply Wiener process.

It can be shown (see Proposition 2.1.10 in [37]) that the $Q$-Wiener process has the following series representation
\[ W(t) = \sum_{k=1}^{\infty} \lambda_k^{1/2} \beta_k(t)e_k, \tag{1.1} \]
where $(\lambda_k, e_k)$, for $k \in \mathbb{N}$, are the eigenpairs of $Q$ and $\{\beta_k\}_{k=1}^{\infty}$ are independent real-valued Brownian motions on $(\Omega, \mathcal{F}, \mathbb{P})$. The series converges in $L^2(\Omega; C([0,T], U))$, for every $T > 0$. To define the stochastic integral for a $Q$-Wiener process, the $Q$-Wiener process $\{W(t)\}_{t \geq 0}$ needs to be adapted to the filtration $\{\mathcal{F}_t\}_{t \geq 0}$ and the difference $W(t) - W(s)$ should be independent of $\mathcal{F}_s$ for all fixed $0 \leq s \leq t$. If these two requirements are met, then we say that $\{W(t)\}_{t \geq 0}$ is a $Q$-Wiener process with respect to the filtration $\{\mathcal{F}_t\}_{t \geq 0}$.

In the next subsection, we would also like to integrate processes whose covariance operator is not necessarily trace-class. To do this we need to define the cylindrical Wiener process. Consider a Hilbert space $(U_1, \langle \cdot, \cdot \rangle_1)$ in which $U$ is continuously embedded. Let $Q \in \mathcal{L}(U)$ be non-negative definite and symmetric. Define the space $U_0 := Q^{1/2}(U)$ with the scalar product
\[ \langle u, v \rangle_0 = \langle Q^{-1/2}u, Q^{-1/2}v \rangle_U, \quad \text{for } u, v \in U_0, \]
where $Q^{1/2}$ is the unique positive square root of $Q$ and $Q^{-1/2}$ is the pseudoinverse of $Q^{1/2}$. Let $J: (U_0, \langle \cdot, \cdot \rangle_0) \to (U_1, \langle \cdot, \cdot \rangle_1)$ be a Hilbert–Schmidt embedding.

Example 1.5. We can always construct $U_1$ and $J$ as described above. Let $Q \in \mathcal{L}(U)$ be non-negative definite and symmetric. Set $U_1 := U$ and define
\[ J: U_0 \to U, \quad u \mapsto \sum_{k=1}^{\infty} \frac{1}{k} \langle u, e_k \rangle_0 e_k, \]
where \( \{e_k\}_{k=1}^{\infty} \) is an orthonormal basis of \( U_0 \). We show that \( J \) is one-to-one and Hilbert–Schmidt. For \( u, v \in U_0 \) and all \( k \in \mathbb{N} \),

\[
\begin{align*}
  u = v & \iff \langle u, e_k \rangle_0 = \langle v, e_k \rangle_0 \\
  & \iff \frac{1}{k} \langle u, e_k \rangle_0 = \frac{1}{k} \langle v, e_k \rangle_0 \\
  & \iff \sum_{k=1}^{\infty} \frac{1}{k} \langle u, e_k \rangle_0 e_k = \sum_{k=1}^{\infty} \frac{1}{k} \langle v, e_k \rangle_0 e_k \\
  & \iff J(u) = J(v).
\end{align*}
\]

To show that \( J \) is Hilbert–Schmidt, we use that

\[
\|Q^{-1/2}e_k\|_U = \|e_k\|_0 = 1, \text{ for all } k \in \mathbb{N}.
\]

We have

\[
\begin{align*}
  \|J\|^2_{L^2(U_0, U)} &= \sum_{k=1}^{\infty} \|Je_k\|^2_U = \sum_{k=1}^{\infty} \left( \sum_{j=1}^{\infty} \frac{1}{j} \langle e_k, e_j \rangle_0 e_j \right)^2_U \\
  &= \sum_{k=1}^{\infty} \|e_k\|^2_U = \sum_{k=1}^{\infty} \frac{1}{k^2} \|Q^{1/2}Q^{-1/2}e_k\|^2_U \\
  &\leq \|Q^{1/2}\|^2_{\mathcal{L}(U)} \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty.
\end{align*}
\]

Here we have also used that \( \|Q^{1/2}Q^{-1/2}e_k\|_U \leq \|Q^{1/2}\|_{\mathcal{L}(U)}\|Q^{-1/2}e_k\|_U \).

Let us now continue with the construction of the cylindrical Wiener process. Define by \( \mathcal{M}^2_T(U) \) the space of all \( U \)-valued continuous, square integrable martingales \( M(t) \), for \( t \in [0, T] \). We have the following result.

**Proposition 1.6** (Proposition 2.5.2 in [37]). Let \( \{e_j\}_{j=1}^{\infty} \) be an orthonormal basis of \( U_0 \) and \( \{\beta_j\}_{j=1}^{\infty} \) a family of independent real-valued Brownian motions. Define \( Q_1 := JJ^* \). Then \( Q_1 \in \mathcal{L}(U_1) \), \( Q_1 \) is non-negative definite and symmetric with finite trace and the series

\[
W(t) = \sum_{j=1}^{\infty} \beta_j(t)e_j, \quad t \in [0, T], \tag{1.2}
\]

converges in \( \mathcal{M}^2_T(U_1) \) and defines a \( Q_1 \)-Wiener process on \( U_1 \). We have that \( Q_1^{1/2}(U_1) = J(U_0) \) and for all \( u_0 \in U_0 \)

\[
\|u_0\|_0 = \|Q_1^{-1/2}Ju_0\|_1 = \|Ju_0\|_{Q_1^{-1/2}(U_1)},
\]

i.e. \( J : U_0 \to Q_1^{1/2}(U_1) \) is an isometry.

**Definition 1.7.** The cylindrical \( Q \)-Wiener process on \( U \) is defined as the \( Q_1 \)-Wiener process on \( U_1 \) given by the series (1.2).
It is important to note that the cylindrical $Q$-Wiener process is in fact not a $U$-valued process. However, let us ignore this fact for the following example.

**Example 1.8.** Let us consider $Q = I$, so that $U_0 = U$. The identity operator is not of trace-class in $U$. If it were, the series representation (1.1) for the Wiener process would be

$$W(t) = \sum_{j=1}^{\infty} \beta_j(t) e_j,$$

where $\{e_j\}_{j=1}^{\infty}$ is an orthonormal basis of $U_0 = U$. Since the cylindrical Wiener process given by (1.2) is a $JJ^*$-Wiener process on $U_1$, we have (see Proposition 2.1.4 in [37])

$$\mathbb{E}[\langle JW(t), u_1 \rangle_1 \langle JW(s), v_1 \rangle_1] = (t \wedge s) \langle JJ^* u_1, v_1 \rangle_1,$$

for any two elements $u_1, v_1 \in U_1$, and times $s, t \geq 0$. Using this and imitating a calculation in [21], we have, for $u, v \in U$ such that $u$ and $v$ are in the range of $J^*$,

$$\mathbb{E}[\langle W(t), u \rangle_U \langle W(s), v \rangle_U] = \mathbb{E}[\langle [W(t), J^* (JJ^*)^{-1} J u]_U, W(s), (JJ^*)^{-1} J v \rangle_U]$$

$$= \mathbb{E}[\langle [JW(t), (JJ^*)^{-1} J u]_1, [JW(t), (JJ^*)^{-1} J v]_1 \rangle_U]$$

$$= (t \wedge s) \langle JJ^* (JJ^*)^{-1} J u, (JJ^*)^{-1} J v \rangle_1$$

$$= (t \wedge s) \langle J u, (JJ^*)^{-1} J v \rangle_1$$

$$= (t \wedge s) \langle u, (JJ^*)^{-1} J v \rangle_U$$

$$= (t \wedge s) \langle u, v \rangle_U.$$

Thus, the series given by (1.3) can intuitively be thought of as a $U$-valued Wiener process with $Q = I$.

We would like to stress again that the Wiener process in Example 1.8 is not an element of $U$ and that the above calculations were purely to improve our understanding of the process, and are not mathematically rigorous (or even correct).

The Hilbert space approach described here to define the $Q$-Wiener process is not the only approach actively used to study SPDEs. For instance, in Paper III we consider the stochastic heat equation, not in terms of a cylindrical $Q$-Wiener process but instead in terms of a Brownian sheet. Before we can define the Brownian sheet, we need the following definitions.

**Definition 1.9** (Definition 7.1 in [34]). A real-valued random field on $\mathcal{D} \subset \mathbb{R}^d$, $d \in \mathbb{N}$, is a set of real-valued random variables $\{X(x, \omega): (x, \omega) \in \mathcal{D} \times \Omega\}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

**Definition 1.10** (Definition 7.3 in [34]). For a set $\mathcal{D} \subset \mathbb{R}^d$, for $d \in \mathbb{N}$, a random field $\{X(x, \omega): (x, \omega) \in \mathcal{D} \times \Omega\}$ is second-order if

$$\mathbb{E}[|X(x, \omega)|^2] < \infty, \quad \text{for every } x \in \mathcal{D}.$$

A second-order random field has covariance function

$$C(x, y) := \mathbb{E}[(X(x) - \mathbb{E}[X(x)])(X(y) - \mathbb{E}[X(y)])], \quad x, y \in \mathcal{D}.$$
Definition 1.11 (Definition 7.5 in [34]). For a set $D \subset \mathbb{R}^d$, for $d \in \mathbb{N}$, a Gaussian random field $\{X(x, \omega) : (x, \omega) \in D \times \Omega\}$ is a second-order random field such that $[X(x_1), X(x_2), \ldots, X(x_M)]^T$ follows the multivariate Gaussian distribution for any $x_1, \ldots, x_M \in D$ and any $M \in \mathbb{N}$.

Let $T > 0$. A Brownian sheet on $[0, T] \times [0, 1]$ is a Gaussian random field $\{B(t, x, \omega) : (t, x, \omega) \in [0, T] \times [0, 1] \times \Omega\}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. It is constructed independently of the $Q$-Wiener process using a martingale approach instead of a Hilbert space approach. In order to not have to introduce new notation and keep with the functional analysis setting, we will construct the Brownian sheet as described in [12]. We would like to mention that it has been shown (see [14]) that the integral for a cylindrical $Q$-Wiener process and the integral for the Brownian sheet are equivalent for a large class of integrands. For a complete presentation of the Brownian sheet we refer to [40].

Let $U = L^2([0, 1])$. If we choose $Q = I$ so that $U_0 = U$, then for every $t \in [0, T]$ and $x \in [0, 1]$, the series

$$B(t, x) = \sum_{j=1}^{\infty} \beta_j(t) \int_0^x \varphi_j(\alpha) \, d\alpha,$$  \hspace{1cm} (1.4)

converges in $L^2(\Omega; L^2([0, 1]))$. Here, $\{\beta_j\}_{j=1}^{\infty}$ is a family of independent real-valued Brownian motions and $\{\varphi_j\}_{j=1}^{\infty}$ is the orthonormal basis of $L^2([0, 1])$ given by $\varphi_j(\alpha) = \sqrt{2} \sin(j \pi \alpha)$ for $j \in \mathbb{N}$. The covariance function of (1.4) is

$$\mathbb{E}[B(t, x)B(s, y)] = \mathbb{E} \left[ \left( \sum_{j=1}^{\infty} \beta_j(t) \int_0^x \varphi_j(\alpha) \, d\alpha \right) \left( \sum_{k=1}^{\infty} \beta_k(s) \int_0^y \varphi_k(\alpha) \, d\alpha \right) \right]$$

$$= \sum_{j=1}^{\infty} \mathbb{E}[\beta_j(t)\beta_j(s)] \int_0^x \varphi_j(\alpha) \, d\alpha \int_0^y \varphi_j(\alpha) \, d\alpha$$

$$= (t \wedge s) \sum_{j=1}^{\infty} \left( \int_0^1 \chi_{[0,x]}(\alpha)\chi_{[0,y]}(\alpha) \varphi_j(\alpha) \, d\alpha \right)^2$$

$$= (t \wedge s) \sum_{j=1}^{\infty} |\langle \chi_{[0,x]}\chi_{[0,y]}, \varphi_j \rangle|^2 = (t \wedge s)(x \wedge y).$$

The series (1.4) has a continuous version called the Brownian sheet. A realization of the Brownian sheet (1.4) with the sum truncated at $N = 200$ can be seen in Figure 1.

1.3 The stochastic integral for Wiener processes

Assume that $\{W(t)\}_{t \in [0, T]}$ is a $U$-valued $Q$-Wiener process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to the filtration $\{\mathcal{F}_t\}_{t \in [0, T]}$ for some fixed $T > 0$. For the moment, let us assume that $\text{Tr}(Q)$ is finite. The stochastic integral for Wiener processes
Figure 1: A realization of the Brownian sheet $B(t,x)$ given by (1.4) for $(t,x) \in [0,1] \times [0,1]$. Left: a MATLAB surf plot. Right: a MATLAB contourf plot. The sum in (1.4) has been truncated at $N = 200$.

is first defined for so-called elementary processes, which are $\mathcal{L}(U,H)$-valued processes $\{\Phi(t)\}_{t \in [0,T]}$ such that there exists $0 = t_0 < t_1 < \ldots < t_N = T$, $N \in \mathbb{N}$, so that

$$
\Phi(t) = \sum_{n=0}^{N-1} \Phi_n 1_{(t_n, t_{n+1}]}(t), \quad t \in [0,T],
$$

where each $\Phi_n : (\Omega, \mathcal{F}) \to \mathcal{L}(U,H)$ is strongly $\mathcal{F}_{t_n}$-measurable and takes only a finite number of values in $\mathcal{L}(U,H)$. Let $\mathcal{E}$ be the set of elementary processes. The integral of processes $\Phi \in \mathcal{E}$ is defined as

$$
\int_0^t \Phi(s) \, dW(s) := \sum_{n=0}^{N-1} \Phi_n \Delta W_n(t),
$$

where $\Delta W_n(t) = W(t_{n+1} \wedge t) - W(t_n \wedge t)$.

One can then show [37] that the integral extends to the completion $\overline{\mathcal{E}}$ of $\mathcal{E}$ by

$$
\int_0^t \Phi(s) \, dW(s) := \lim_{n \to \infty} \int_0^t \Phi_n(s) \, dW(s),
$$

for $\Phi \in \overline{\mathcal{E}}$ and $\{\Phi_n\}_{n=1}^\infty \subset \mathcal{E}$ with $\lim_{n \to \infty} \Phi_n = \Phi$. Let $\Omega_T = [0,T] \times \Omega$, and define the sigma algebra $\mathcal{P}_T$ by

$$
\mathcal{P}_T := \sigma (\{(s,t) \times F \mid 0 \leq s < t \leq T, F \in \mathcal{F}_s\} \cup \{\{0\} \times F \mid F \in \mathcal{F}_0\}).
$$

If $Y : (\Omega_T, \mathcal{P}_T) \to (U,\mathcal{B}(U))$ is measurable, then $Y$ is called $U$-predictable. The completion $\overline{\mathcal{E}}$ has an explicit representation given by (see Theorem 3.11 in [29])

$$
\mathcal{N}_W^2 := \{\Phi : [0,T] \times \Omega \to \mathcal{L}_2(U_0,H) \mid \Phi \text{ is } \mathcal{L}_2(U_0,H)\text{-predictable and } \|\Phi\|_T < \infty\},
$$
where we recall that $U_0 = Q^{1/2}(U)$ and where
\[ \| \Phi \|_T := \left( \mathbb{E} \left[ \int_0^T \| \Phi(s) \|^2_{L^2(U_0,H)} \, ds \right] \right)^{1/2}. \]

We can still integrate processes $\{\Phi(t)\}_{t \in [0,T]}$ which are $L_2(U_0,H)$-predictable when $\text{Tr}(Q) = \infty$. This is accomplished by first observing that since $\text{Tr}(Q_1) < \infty$, we can integrate processes that are $L_2(Q^{1/2}_1(U_1),H)$-predictable. Then we note that $\Phi \in L_2(U_0,H)$ is equivalent to $\Phi J^{-1} \in L_2(Q^{1/2}_1(U_1),H)$. Therefore, the stochastic integral for a cylindrical Wiener process is defined by
\[ \int_0^t \Phi(s) \, dW(s) := \int_0^t \Phi(s) J^{-1} \, dW(s), \]
where $W$ is given by (1.2). The class of integrable processes is the same whether we integrate a cylindrical Wiener process or a Wiener process with trace-class covariance operator. This can be motivated by the following calculation.
\[ \mathbb{E} \left[ \left\| \int_0^t \Phi(s) \, dW(s) \right\|^2_H \right] = \mathbb{E} \left[ \int_0^t \| \Phi(s) J^{-1} \|^2_{L_2(Q^{1/2}_1(U_1),H)} \, ds \right] \]
\[ = \mathbb{E} \left[ \int_0^t \sum_{k=1}^{\infty} \| \Phi(s) J^{-1} J e_k \|^2_U \, ds \right] \]
\[ = \mathbb{E} \left[ \int_0^t \| \Phi(s) J^{-1 \, -1} \|^2_{L_2(U_0,H)} \, ds \right]. \]
Here we have used the Itô isometry (see Proposition 3.5 and Remark 3.12 in [29]) and the fact that $\{e_k\}_{k=1}^\infty$ being an orthonormal basis for $U_0$ is equivalent to $\{J e_k\}_{k=1}^\infty$ being an orthonormal basis for $Q^{1/2}_1(U_1)$. The integral is independent of the embedding $J$, and in the case of $\text{Tr}(Q) < \infty$ we can choose $J = I$ and the definitions coincide.

### 1.4 Stochastic partial differential equations

The types of stochastic partial differential equations that we consider are equations written in the form
\[ dX(t) = (AX(t) + F(X(t))) \, dt + B(X(t)) \, dW(t), \quad 0 < t \leq T \]
\[ X(0) = X_0, \]
where $A: D(A) \subset H \to H$ is a linear operator and the generator of a strongly continuous semigroup $S(t) = e^{At}$ (see comment on semigroups below), and $F: H \to H$ and $B: H \to L_2(U_0,H)$. The process $\{W(t)\}_{t \geq 0}$ is a (possibly cylindrical) $Q$-Wiener process, as previously described. The initial value $X_0$ is an $\mathcal{F}_0$-measurable $H$-valued random variable. The SPDE (1.5) is to be interpreted as the integral equation
\[ X(t) = X_0 + \int_0^t (AX(s) + F(X(s))) \, ds + \int_0^t B(X(s)) \, dW(t), \quad 0 < t \leq T. \]
where the deterministic integral is a Bochner integral and the stochastic integral is defined in Subsection 1.3.

As we are mainly using a semigroup approach in this thesis, some knowledge of semigroups is essential. We take a few moments here to state the definitions of a semigroup and its generator.

**Definition 1.12** (Definition 5.1 in [18]). A family \( \{S(t)\}_{t \geq 0} \) of bounded linear operators on \( H \) is a semigroup if 
\[
S(t+s) = S(t) S(s), \quad \text{for } s, t \geq 0, \text{ and } S(0) = I.
\]
In addition, the semigroup is called a strongly continuous semigroup if the map \( (x, t) \mapsto S(t)x \) is strongly continuous. That is, if
\[
\lim_{t \to 0^+} S(t)x = x, \quad \text{for all } x \in H.
\]

**Definition 1.13** (Definition 1.2 in [18]). The generator \( A : \mathcal{D}(A) \subset H \to H \) of a strongly continuous semigroup \( \{S(t)\}_{t \geq 0} \) on \( H \) is the operator defined by
\[
Ax = \lim_{h \to 0^+} \frac{S(h)x - x}{h},
\]
with \( \mathcal{D}(A) \) being all \( x \in H \) such that the limit exists.

Strong solutions of (1.5) are formulated in terms of the integral equation (1.6), and therefore require the solutions to be in \( \mathcal{D}(A) \). In general this is not the case and hence we do not consider strong solutions of (1.5). Instead we work with mild solutions.

**Definition 1.14.** An \( H \)-valued predictable process \( \{X(t)\}_{t \in [0,T]} \) is a mild solution of (1.5) if, for all \( t \in [0,T] \),
\[
X(t) = S(t)X_0 + \int_0^t S(t-s)F(X(s)) \, ds + \int_0^t S(t-s)B(X(s)) \, dW(s) \quad \mathbb{P}\text{-a.s.,} \quad (1.7)
\]
and such that the integrals are well defined.

If \( F \) and \( B \) satisfy standard Lipschitz and linear growth conditions, then there is a unique global mild solution of (1.5), and the solution has a continuous modification (see Theorem 7.4 in [12]).

We now state two results (and where to find them and their proofs) that are used many times throughout Papers I-IV. The first one is sometimes called Burkholder’s inequality (or Burkholder–Davis–Gundy’s inequality) and the second one is the Itô formula.

**Theorem 1.15** (Lemma 7.2 in [12]). Let \( r \geq 1 \) and let \( \{\Phi(t)\}_{t \in [0,T]} \) be an \( \mathcal{L}_2(U_0, H) \)-valued predictable process. Then, for \( t \in [0,T] \), we have
\[
\mathbb{E} \left[ \sup_{s \in [0,t]} \left\| \int_0^s \Phi(\sigma) \, dW(\sigma) \right\|_{H}^{2r} \right] \leq c_r \mathbb{E} \left[ \left\| \int_0^t \Phi(\sigma) \, dW(\sigma) \right\|_{H}^{2r} \right] \leq C_r \mathbb{E} \left[ \left( \int_0^t \left\| \Phi(\sigma) \right\|_{\mathcal{L}_2(U_0, H)}^2 \, d\sigma \right)^r \right],
\]

where the deterministic integral is a Bochner integral and the stochastic integral is defined in Subsection 1.3.

As we are mainly using a semigroup approach in this thesis, some knowledge of semigroups is essential. We take a few moments here to state the definitions of a semigroup and its generator.
The two constants $c_r$ and $C_r$ are explicitly given by

$$c_r = \left(\frac{2r}{2r-1}\right)^{2r} \quad \text{and} \quad C_r = (r(2r-1))^{r} \left(\frac{2r}{2r-1}\right)^{2r^2}.$$ 

**Theorem 1.16** (Theorem 4.17 in [12]). Let $X(t)$ be given by

$$X(t) = X(0) + \int_0^t \varphi(s) \, ds + \int_0^t \Phi(s) \, dW(s), \quad t \in [0,T],$$

where $X(0)$ is an $\mathcal{F}_0$-measurable $H$-valued random variable, $\varphi$ is an $H$-valued predictable process Bochner integrable on $[0,T]$, $\mathbb{P}$-a.s., and $\Phi$ is an $\mathcal{L}_2(U_0,H)$-valued process such that the stochastic integral is well defined. Assume that a function $G : [0,T] \times H \to \mathbb{R}$ and its partial derivative w.r.t. $t$, $G_t$, and its Fréchet derivatives w.r.t. $x \in H$, $G_x$ and $G_{xx}$, are uniformly continuous on bounded subsets of $[0,T] \times H$. Then, $\mathbb{P}$-a.s. for all $t \in [0,T]$, we have

$$G(t,X(t)) = G(0,X(0)) + \int_0^t (G_x(s,X(s)),\Phi(s)\,dW(s))_H$$

$$+ \int_0^t (G_t(s,X(s)) + (G_x(s,X(s)),\varphi(s))_H) \, ds$$

$$+ \frac{1}{2} \int_0^t \text{Tr} \left[ G_{xx}(s,X(s))(\Phi(s)Q^{1/2})(\Phi(s)Q^{1/2})^* \right] \, ds.$$ 

### 1.5 Exponential integrators

Exponential integrators have been used for many years for the numerical approximation of solutions to ordinary differential equations (ODEs), stochastic differential equations (SDEs), and time integration of PDEs. To motivate exponential integrators, let us consider a system of ODEs

$$u'(t) = Lu(t) + N(u(t)), \quad u(0) = u_0,$$ 

where $t \in [0,T]$, for $T > 0$. Here, $u_0 \in \mathbb{R}^d$, for $d \in \mathbb{N}$, $L$ is a $d \times d$ real matrix, and $N : \mathbb{R}^d \to \mathbb{R}^d$ is a nonlinear function. The system of ODEs could be assumed to be stiff with the stiffness contained in the matrix $L$. Since this is only for motivational purposes, we choose not to specify any further. Note, however, that spatial discretizations of PDEs often lead to systems like (1.8). The basic idea of exponential integrators is to treat the solution of the linear equation $u'(t) = Lu(t)$ exactly, and to use a general linear method for the nonlinear part $N(u)$. By doing so, we integrate the stiff part of the system exactly. Let $N > 0$ be an integer and consider the time step $\Delta t = \frac{T}{N}$ and the grid $0 = t_0 < t_1 < \ldots < t_N = T$ with $t_n = n\Delta t$, $n = 0, \ldots, N$. By the variation-of-constants formula, we have

$$u(t_{n+1}) = e^{\Delta t L} u(t_n) + \int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)L} N(u(s)) \, ds.$$ 

(1.9)
If \( \mathcal{N}(u(t)) \equiv 0 \), then (1.9) is the exact solution of (1.8). In general, however, one needs to discretize the integral on the right-hand side of (1.9). Different ways of doing this leads to different types of exponential methods. In Example 1.17 we state the exponential Euler method and in Example 1.18 we state the more elaborate exponential Runge–Kutta methods. For more information on exponential integrators for ODE and PDE, we refer the reader to [23].

**Example 1.17.** The exponential Euler method is derived by freezing the nonlinear term \( \mathcal{N} \) in (1.9) at the left endpoint and integrating the exponential exactly. It is given by the scheme
\[
 u_{n+1} = e^{\Delta t L} u_n + \Delta t \varphi_1(\Delta t L) \mathcal{N}(u_n),
\]
where
\[
 \varphi_1(z) = e^z - 1.
\]

**Example 1.18.** For \( s \in \mathbb{N} \), consider real coefficient functions \( a_{ij} \), \( b_i \), and \( c_i \) for \( i,j = 1, \ldots, s \). The following class of one-step methods is called exponential Runge–Kutta methods:
\[
 u_{n+1} = e^{\Delta t L} u_n + \Delta t \sum_{i=1}^{s} b_i(\Delta t L) \mathcal{N}(U_{ni})
\]
\[
 U_{ni} = e^{c_i \Delta t L} u_n + \Delta t \sum_{j=1}^{s} a_{ij}(\Delta t L) \mathcal{N}(U_{nj}).
\]

When it comes to SPDEs, exponential integrators are commonly used numerical methods as well (see e.g. [9, 10, 26, 28, 35, 39, 42]). However, due to the relative infancy of the subject of numerical analysis of SPDEs, the study of exponential integrators for SPDEs is far from saturated. Using the notations above, we now derive the exponential scheme that is used throughout the thesis. By freezing the integrands in the mild solution (1.7) at their left endpoint, we obtain the explicit exponential scheme (i.e. an approximation \( X^n \) of \( X(t_n) \) for every \( n = 0, \ldots, N \))
\[
 X^{n+1} = S(\Delta t) X^n + \Delta t S(\Delta t) F(X^n) + S(\Delta t) B(X^n) \Delta W^n, \quad n = 0, \ldots, N-1,
\]
where \( \Delta W^n = W(t_{n+1}) - W(t_n) \) denotes Wiener increments. When considering strong convergence, using a higher order method for the deterministic integral will not increase the order of convergence since the order will be bounded by the stochastic integral. Certainly there are situations where it may be interesting to study different discretizations of the nonlinear term (when studying the evolution of a Hamiltonian for example), but this is not the prime focus of this thesis.

Our main interest is in the strong convergence of \( X^n \) to \( X(t_n) \) in \( L^q(\Omega; H) \) (or in a similar norm), for \( q \geq 2 \). However, in Papers III and IV we also study almost sure convergence and convergence in probability. We state the definitions of the order of convergence for all three types of convergence. For the almost sure convergence and convergence in probability we use the same notation for the orders as in [38].
In Definitions 1.19, 1.20, and 1.21 we assume that \( \{X(t)\}_{t \in [0,T]} \) is the mild solution given by (1.7), and that \( \{X^n\}_{n=0,...,N} \) is any numerical approximation of (1.7).

**Definition 1.19.** The approximation \( \{X^n\}_{n=0,...,N} \) converges to \( \{X(t)\}_{t \in [0,T]} \) in \( L^q(\Omega;H) \) with order \( \gamma > 0 \) if there exists a constant \( C > 0 \), independent of \( \Delta t \), such that

\[
\left( \mathbb{E} \left[ \max_{n=0,...,N} \|X^n - X(t_n)\|_H^q \right] \right)^{1/q} \leq C \cdot (\Delta t)^\gamma,
\]

for \( \Delta t \) small enough.

**Definition 1.20.** The approximation \( \{X^n\}_{n=0,...,N} \) converges to \( \{X(t)\}_{t \in [0,T]} \) almost surely with order \( \gamma > 0 \) if, for almost all \( \omega \in \Omega \), there exists a constant \( C(\omega) > 0 \), independent of \( \Delta t \), such that

\[
\max_{n=0,...,N} \|X^n(\omega) - X(t_n,\omega)\|_H \leq C(\omega)(\Delta t)^\gamma,
\]

for \( \Delta t \) small enough.

**Definition 1.21.** The approximation \( \{X^n\}_{n=0,...,N} \) converges to \( \{X(t)\}_{t \in [0,T]} \) in probability with order \( \gamma > 0 \) if

\[
\lim_{C \to \infty} \mathbb{P} \left( \max_{n=0,...,N} \|X^n - X(t_n)\|_H \geq C \cdot (\Delta t)^\gamma \right) = 0
\]

uniformly with respect to \( \Delta t \).

We will occasionally write convergence order of \( \gamma- \). By this notation we mean convergence order of \( \gamma - \epsilon \) for any \( \epsilon > 0 \).

## 2 Summary of the Papers

In this section we give short summaries of the four papers which the thesis is build on. The published papers (Paper I and II) have been reformatted to fit the style of the thesis and may contain some very minor differences from the published versions.

### 2.1 Paper I – Full discretization of semilinear stochastic wave equations driven by multiplicative noise

In Paper I we use a finite element method and a stochastic trigonometric method to obtain a full approximation of semilinear stochastic wave equations of the form

\[
\begin{align*}
d\dot{u} - \Delta u \, dt &= f(u) \, dt + g(u) \, dW & \text{in } \mathcal{D} \times (0,\infty), \\
u &= 0 & \text{in } \partial \mathcal{D} \times (0,\infty), \\
u(\cdot,0) &= u_0, \quad \dot{u}(\cdot,0) = v_0 & \text{in } \mathcal{D},
\end{align*}
\]

(2.1)
where \( u = u(x,t) \) and \( D \subset \mathbb{R}^d, d = 1, 2, 3, \) is a bounded convex domain with polygonal boundary \( \partial D. \) The “\( \cdot \)\" denotes the time derivative \( \frac{\partial}{\partial t}. \) Let \( H = L^2(D). \) The process \( \{W(t)\}_{t \geq 0} \) is an \( H\)-valued \( Q\)-Wiener process as seen in the introduction. The functions \( f: H \to H \) and \( g: H \to L^2(H_0, H), H_0 := Q^{1/2}(H), \) are assumed to satisfy some regularity and Lipschitz conditions. Paper I generalizes the work [9] in which the same discretization was used for (2.1) with \( f \equiv 0 \) and \( g \equiv I \) to study mean-square convergence and a so-called trace formula. The trace formula considered for the stochastic wave equation (2.1) is the time evolution of the expected value of the energy of the system (see below for further information).

By denoting \( u_2 := \dot{u}_1 := \dot{u}, \) we write (2.1) in the abstract form

\[
\begin{aligned}
    dX(t) &= AX(t) \, dt + F(X(t)) \, dt + G(X(t)) \, dW(t), \quad t > 0, \\
    X(0) &= X_0,
\end{aligned}
\]

where

\[
X := \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \quad A := \begin{bmatrix} 0 & I \\ \Delta & 0 \end{bmatrix}, \quad F(X) := \begin{bmatrix} 0 \\ f(u_1) \end{bmatrix}, \quad G(X) := \begin{bmatrix} 0 \\ g(u_1) \end{bmatrix}, \quad \text{and } X_0 := \begin{bmatrix} u_0 \\ v_0 \end{bmatrix}.
\]

The operator \( A \) is the generator of a strongly continuous semigroup \( E(t) \) given by

\[
E(t) = \begin{bmatrix} C(t) & \Lambda^{-1/2}S(t) \\ -\Lambda^{1/2}S(t) & C(t) \end{bmatrix},
\]

where \( \Lambda = -\Delta, \) \( C(t) = \cos(t\Lambda^{1/2}), \) and \( S(t) = \sin(t\Lambda^{1/2}). \)

To avoid too many technical details in this summary, we simply state that we use a standard piecewise linear finite element method with mesh size \( h. \) It is commonly used, and more details can be found in [1,9,30,31,35,41,43]. The semi-discrete finite element approximation is made fully discrete with the stochastic trigonometric method with step size \( \Delta t, \) and we obtain an approximation

\[
U^n = \begin{bmatrix} U^n_1 \\ U^n_2 \end{bmatrix}
\]

of the exact solution \( X(t_n) \) at the discrete time points \( t_n = n\Delta t, n = 0, 1, \ldots, N. \) The stochastic trigonometric method is the exponential method described by (1.10). This numerical method integrates the semigroup of the operator \( A \) exactly, and, as seen above, one can write this semigroup as a matrix of trigonometric operators. This motivates the name of the considered method.

In the first part of the paper, we prove mean-square convergence of the fully discrete scheme. By assuming that the initial values \( u_0, v_0, \) and the functions \( f \) and \( g \) satisfy some regularity assumptions and Lipschitz conditions, we are able to prove mean-square convergence of the fully discrete scheme of the same order as in the additive noise case:

\[
\left( \mathbb{E}[||U^n_1 - u_1(t_n)||^2_{L^2(D)}]\right)^{1/2} \leq C \cdot (h^{2\beta/3} + (\Delta t)^{\min(\beta,1)}) \quad \text{for } \beta \in [0, 3],
\]
and similarly for the second component (the velocity). The parameter $\beta$ is connected to the regularity of the initial values, the function $f$ and $g$, and to the Wiener process. For more details, see Paper I.

The second part of the paper is dedicated to the study of a trace formula. For this, we consider (2.1) with additive noise ($g \equiv I$) and $f(u) = -V'(u)$ for a smooth potential $V$. The Hamiltonian or energy of the system is defined by (with a slight abuse of notation)

$$H(X(t)) = \frac{1}{2} \int_D (|u_2(x,t)|^2 + |\nabla u_1(x,t)|^2) \, dx + \int_D V(u_1(x,t)) \, dx,$$

where we recall that the notation $X(t)$ denotes

$$X(t) = \begin{bmatrix} u_1(x,t) \\ u_2(x,t) \end{bmatrix}.$$  

It is well known that in the deterministic setting the energy is conserved for all times. Although this does not remain true in the stochastic setting, we still see a nice geometric property. We show, by a use of the Itô formula (Theorem 1.16), that the expected energy of the system grows linearly in time. That is, the exact solution of (2.1) satisfies the trace formula

$$\mathbb{E}[H(X(t))] = \mathbb{E}[H(X_0)] + t \frac{1}{2} \text{Tr}(Q), \quad t \geq 0.$$  

Note that for the above formula to make sense, we require that the covariance operator $Q$ has finite trace. A similar trace formula is also obtained for the finite element solution.

In [9] it was shown that for linear stochastic wave equations, the stochastic trigonometric method exactly preserves the trace formula for all times. When adding a nonlinear term $f(u) = -V'(u)$ to the equation we cannot prove an exact trace formula for this considered numerical discretization. Instead we obtain a small error in the linear drift. We prove the following formula, which we call an almost trace formula.

$$\mathbb{E}[H(U^n)] = \mathbb{E}[H(T^0)] + t_n \frac{1}{2} \text{Tr}(P_h Q P_h) + O((\Delta t)^{\min(2(\beta-1),1)})$$

for $0 \leq t_n \leq T$ and $\beta \in [1,2]$. Here $P_h$ is the orthogonal projection that takes elements in $L^2(D)$ and projects them onto the finite element space. The proof of the almost trace formula includes showing that

$$\left| \mathbb{E}[H(U^n)] - \mathbb{E}[H(X_h(t_n))] \right| = O((\Delta t)^{\min(2(\beta-1),1)}),$$

where $X_h$ is the semi-discrete finite element approximation of $X$. Thus, the almost trace formula can be seen as a weak error estimate.

We conclude Paper I by illustrating our theoretical results through a number of numerical experiments. To demonstrate the results on mean-square convergence, we use
three different examples. First we demonstrate the spatial rate of convergence of the numerical discretization using the one-dimensional hyperbolic Anderson model with multiplicative noise. Next, we illustrate the temporal convergence using two examples: the sine-Gordon equation driven by space-time white noise and the sine-Gordon equation with multiplicative noise. We observe our theoretical orders of convergence for all experiments. For the temporal convergence we also do simulations for other common schemes. Finally, we also illustrate the trace formula through numerical experiments. We find that the stochastic trigonometric method behaves very well, even with a larger time step and longer time interval. It also preserves the trace formula better than the other numerical schemes.

2.2 Paper II – Exponential integrators for stochastic Schrödinger equations driven by Itô noise

In Paper II we are interested in the time discretization of stochastic Schrödinger equations using an explicit exponential method. The equations we consider are of the form

\[
i du = \Delta u \, dt + F(x, u) \, dt + G(u) \, dW \quad \text{in } \mathbb{R}^d \times (0, \infty),
\]

\[u(\cdot, 0) = u_0 \quad \text{in } \mathbb{R}^d,
\]

(2.3)

where \( u = u(x, t) \). Here, \( F(x, u) = V(x)u \) for a real-valued potential \( V(x) \), and \( G(u) = I \) or \( G(u) = u \). That is, we consider both additive and multiplicative noise. For further details on the coefficients and, in particular, \( s \) and the dimension \( d \) we refer to the paper. We study convergence of the explicit exponential scheme to the exact solution in \( L^{2p}(\Omega, H^s) \), where \( p \geq 1 \) and \( H^s := H^s(\mathbb{R}^d) \) is the Sobolev space of order \( s \geq 0 \).

Exponential methods have been used for the time integration of deterministic Schrödinger equations (see [2–4, 6, 8, 17, 22] to mention but a few), yet they have not been used very much in the stochastic versions of this equation. In previous works, it has been more common to use either an implicit Crank–Nicolson method [15, 16], some implicit symplectic method [7,25,27], or a splitting method [11,32,33].

With additive noise in (2.3), we are able to prove order of convergence 1 in \( L^{2p}(\Omega, H^s) \), for \( p \geq 1 \), \( s \geq 0 \), and in dimensions \( d = 1, 2, 3 \). With multiplicative noise we assume that \( s > \frac{d}{2} \), and obtain convergence order of \( \frac{1}{2} \) in \( L^{2p}(\Omega, H^s) \), for \( p \geq 1 \). The same orders of convergence are obtained in [16] for the implicit Crank–Nicolson scheme. However, our regularity assumptions need not be as high since we do not approximate the semigroup of the Schrödinger equation.

With appropriate boundary conditions, such as homogeneous Dirichlet boundary conditions or periodic boundary conditions, the deterministic Schrödinger equation

\[
\frac{i}{\partial} \frac{\partial u(x, t)}{\partial t} = \Delta u(x, t)
\]
has many interesting conserved quantities (see [5]). This includes the mass (we omit
the domain of integration for ease of presentation)

\[ M(u)(t) = \int |u(x,t)|^2 \, dx, \] (2.4)

the energy

\[ H(u)(t) = \frac{1}{2} \int |\nabla u(x,t)|^2 \, dx, \]

and the momentum

\[ p(u)(t) = i \int (u(x,t) \nabla \bar{u}(x,t) - \bar{u}(x,t) \nabla u(x,t)) \, dx. \]

We prove that in the stochastic setting with \( F(x,u) \equiv 0, G(u) \equiv I \) in (2.3), and with
trace-class noise, the expected value of the mass and energy grows linearly in time. With smoother noise, the expected momentum also grows linearly in time. We show
that the same exact linear drifts hold for the numerical approximation given by the
exponential integrator and for all times. In addition, we prove that the forward and
backward Euler–Maruyama schemes and the midpoint scheme do not preserve this
linear drift for the mass. The expected mass of the forward Euler–Maruyama approx-
imation is shown to grow exponentially in time. For the backward Euler–Maruyama
approximation the expected mass grows at a slower rate compared to the rate of the
exact solution. The midpoint approximation will always produce an expected mass
that underestimates the mass of the exact solution.

With \( F(x,u) = V(x)u \) and additive noise, the mass is still given by (2.4). The
energy takes the form

\[ H(u)(t) = \frac{1}{2} \int |\nabla u(x,t)|^2 \, dx - \frac{1}{2} \int V(x)|u(x,t)|^2 \, dx. \]

We prove that for the exact solution to (2.3), the expected mass and energy exhibit lin-
ear growth. However, similarly to the almost trace formula in Paper I, the exponential
scheme produces a small error of size \( O(\Delta t) \), where \( \Delta t \) is the step size.

Paper II includes a vast number of numerical experiments that verify our theo-
retical results. We illustrate the mean-square convergence and the trace formulas by
considering the stochastic Schrödinger equations on the interval \([0,2\pi]\) with periodic
boundary conditions. For the spatial discretization we use a pseudospectral method,
and for the temporal discretization we use the exponential method previously de-
scribed. We plot the mean-square errors for the exponential method for the stochas-
tic Schrödinger equation (2.3) with additive or multiplicative noise. In both cases,
our theoretical orders of convergence can be observed. For the stochastic Schrödinger
equation (2.3) with \( F \equiv 0 \) and additive noise, we observe exact linear drifts for both
the expected mass and energy. Finally, we also illustrate the expected mass when
\( F(x,u) = V(x)u \) in (2.3). Even though the exponential scheme does not preserve this
trace formula exactly, our numerical simulations show that it performs much better
than the Crank–Nicolson scheme and the semi-implicit Euler–Maruyama scheme.
2.3 Paper III – A fully discrete approximation of the one-dimensional stochastic heat equation

In Papers I and II, we considered semilinear equations and treated them as stochastic evolution equations in a Hilbert space. This reduced the equations to problems that resemble stochastic ordinary differential equations of one variable $t$. Error estimates of the approximations were then studied in an appropriate Hilbert space norm such as the $L^2$-norm or some Sobolev norm. Another approach is the martingale approach seen in the introduction, which allows for other error estimates. The two approaches are very different and it can be difficult for someone used to one approach to make use of the literature in the other approach [14]. In Paper III we use the martingale approach to study a full discretization of the one-dimensional stochastic heat equation.

For $(t, x) \in [0, \infty) \times [0, 1]$, the considered stochastic heat equation reads

$$
\begin{align*}
\frac{\partial}{\partial t} u(t,x) &= \frac{\partial^2}{\partial x^2} u(t,x) + f(t,x,u(t,x)) + \sigma(t,x,u(t,x)) \frac{\partial^2}{\partial t \partial x} W(t,x), \\
u(t,0) &= u(t,1) = 0, \\
u(0,x) &= u_0.
\end{align*}
$$

The coefficients $f$ and $\sigma$ are assumed to satisfy a linear growth condition, and the noise is a Brownian sheet. More information on the initial value $u_0$ will be given below.

For the spatial discretization we use a standard finite difference method as in [19], and for the temporal discretization we use an explicit exponential method, as seen in the previous papers. Let $T > 0$ be a fixed final time and let $M, N > 0$ be integers. Denote the semidiscrete finite difference solution by $u^M$ and the full approximation by $u^{M,N}$, where $\Delta x = \frac{1}{M}$ is the mesh size and $\Delta t = \frac{T}{N}$ is the step size in time.

In the first part of the paper we assume that $f$ and $g$ are globally Lipschitz continuous. We prove convergence of the full approximation $u^{M,N}$ to the finite difference solution $u^M$ in the $L^q(\Omega)$-norm

$$
\sup_{t \in [0,T]} \left( \frac{1}{M} \right)^{\frac{1}{2}} \left( \mathbb{E}[|u^{M,N}(t,x) - u^M(t,x)|^q] \right)^{\frac{1}{q}}.
$$

for fixed $t \in [0,T]$ and any $q \geq 2$. In [20], the author considers (2.5) and proves convergence order of $\frac{1}{8}$ when $u_0 \in C([0,1])$ and order $\frac{1}{4}$ when $u_0 \in C^3([0,1])$ for both the explicit and semi-implicit Euler–Maruyama schemes in the $L^q(\Omega)$-norm. Recall that the notation $\frac{1}{8}$ means $\frac{1}{8} - \varepsilon$, for any $\varepsilon > 0$. Using the mild equation and some auxiliary results, we prove convergence order of $\frac{1}{2}$ for the exponential method when $u_0 \in C([0,1])$, and convergence order of $\frac{1}{4}$ when $u_0$ is in the Sobolev space $H^1([0,1])$. In addition, the error bound in the latter case is uniform in both the space and time variable. Combined with error estimates for the finite difference approximation proved in [19], we get estimates for the full approximation. The main reason why we get a higher temporal order of convergence is due to the fact that the exponential method does not approximate the discrete Green’s function in time.
In the second part of the paper, we use the result on convergence in $L^q(\Omega)$, for any $q \geq 2$, together with Markov’s inequality and the Borel–Cantelli lemma to prove almost sure convergence of the full approximation to the exact solution of (2.5).

These results on mean-square convergence and almost sure convergence are accompanied by numerical experiments. A loglog plot of the mean-square error confirms our theoretical order of convergence. By plotting profiles of one realization of the numerical solution for different time steps, we illustrate almost sure convergence of the exponential scheme. In addition to this, we also investigate the computational cost of the exponential scheme, the semi-implicit Euler–Maruyama scheme, and the Crank–Nicholson–Maruyama scheme. By plotting the computational time as a function of the averaged final error for each scheme, we observe that the exponential scheme is far superior to the other two schemes in this regard.

The third part of the paper is concerned with proving convergence in probability of the numerical solution to the exact solution when the coefficients $f$ and $\sigma$ are non-globally Lipschitz continuous. We only assume that $f$ and $\sigma$ satisfy a linear growth condition and are continuous in the $u$ variable. We also assume pathwise uniqueness of the solution of the stochastic heat equation (see Remark 3.1 in Paper III for further information regarding pathwise uniqueness of the solution of (2.5)). With these assumptions, we are able to prove that the full approximation converges in probability to a random field $\{u(t,x) : t \geq 0, x \in [0,1]\}$ and this random field is the unique solution to the stochastic heat equation that we consider. The idea of the proof is inspired by ideas from [19,20,36], in which convergence in probability of the explicit and semi-implicit Euler–Maruyama schemes was proved.

2.4 Paper IV – Convergence of an exponential method for the stochastic Schrödinger equation with power-law nonlinearity

In Paper II we considered a stochastic Schrödinger equation with a real-valued potential. Another common variation of the Schrödinger equation is to have $F(x,u) = \lambda |u|^{2\sigma} u$ in (2.3). Here $\lambda = 1$ or $-1$, which corresponds to the focusing or defocusing case respectively, and $\sigma$ is a positive integer. In Paper IV we study the explicit exponential integrator when applied to (2.3) with $F$ as above and with $G(u) = u$. That is, we consider time discretization of the stochastic Schrödinger equation

$$ idu = \Delta u \, dt + \lambda |u|^{2\sigma} u \, dt + u \, dW \quad \text{in } \mathbb{R}^d \times \mathbb{R}_+, $$

$$ u(\cdot,0) = u_0 \quad \text{in } \mathbb{R}^d, $$

(2.6)

where $u_0 \in H^s(\mathbb{R}^d)$, $s > \frac{d}{2}$, and $d = 1, 2, 3$. The noise $\{W(t)\}_{t \geq 0}$ is an $L^2(\mathbb{R}^d)$-valued $Q$-Wiener process adapted to a normal filtration. For more information on the regularity assumptions for $Q$, we refer to Paper IV. We study (2.6) in its mild form

$$ u(t) = S(t)u_0 - i\lambda \int_0^t S(t-r) |u(r)|^{2\sigma} u(r) \, dr - i \int_0^t S(t-r) u(r) \, dW(r), $$

(2.7)
where $S(t) = e^{-it\Delta}$ is the semigroup of isometries generated by $-i\Delta$.

There are two main problems when working with (2.6). The first one being that the nonlinear term $|u|^{2\sigma} u$ is not globally Lipschitz continuous, and the second being that the solution $u$ is not necessarily bounded on compact time intervals. We say that it may blow up. To handle these problems, we introduce a cut-off function $\theta \in C^\infty(\mathbb{R}_+)$ such that $\theta \geq 0$, $\text{supp}(\theta) \subset [0, 2]$, and $\theta(x) \equiv 1$ for $x \in [0, 1]$. Then we define

$$\theta_R(\cdot) = \theta \left( \frac{\|u\|_s}{R} \right),$$

where $R > 0$, and the norm is the Sobolev norm on $H^s(\mathbb{R}^d)$. We can now define a truncation of the nonlinear term in (2.6), and obtain a truncation of (2.7) by

$$u^R(t) = S(t)u_0 - i\lambda \int_0^t S(t-r)\theta_R(u^R(r))|u^R(r)|^{2\sigma} u^R(r) \, dr \quad - i \int_0^t S(t-r)u^R(r) \, dW(r). \quad (2.8)$$

To deal with the possible blow-up of the solution of (2.6) we restrict our analysis to random time intervals $[0, \tau^\ast]$, where $\tau^\ast$ is a stopping time almost surely less than the blow-up time. In general, the blow-up time is not explicitly known. However, if $\lambda = -1$ or if $\sigma < \frac{2}{d}$, then blow-up will not occur in finite time (see Remark 2.1 in [16]). The results proved in this paper are valid up to a time smaller than the blow-up time.

The numerical approximation is derived as follows. Let $T > 0$ be a fixed final time and $N > 0$ an integer. Let $\Delta t = \frac{T}{N}$ so that $t_n = n\Delta t$, for $n = 0, 1, \ldots, N$. Let $\tau^\ast < \tau(u_0) \wedge T$ a.s. be a stopping time, where $\tau(u_0)$ denotes the blow-up time of the exact solution of (2.6). We define $N_{\tau^\ast} = \left\lceil \frac{\tau^\ast}{\Delta t} \right\rceil$ to be the integer part of $\frac{\tau^\ast}{\Delta t}$. For $n = 1, \ldots, N_{\tau^\ast}$, we approximate the exact solution $u(t_n)$ given by (2.7) by

$$u^n = S(\Delta t)u^{n-1} - i\lambda \Delta t S(\Delta t)|u^{n-1}|^{2\sigma} u^{n-1} - iS(\Delta t)u^{n-1}\Delta W^{n-1}, \quad (2.9)$$

where $\Delta W^{n-1} = W(t_n) - W(t_{n-1})$. We also truncate the approximation (2.9) by

$$u_R^n = S(\Delta t)u_R^{n-1} - i\lambda \Delta t S(\Delta t)\theta_R(u_R^{n-1})|u_R^{n-1}|^{2\sigma} u_R^{n-1} - iS(\Delta t)u_R^{n-1}\Delta W^{n-1}. \quad (2.10)$$

We first prove that the truncated numerical approximation (2.10) converges to the truncated exact solution (2.8) with convergence order $\frac{1}{2}$ in $L^{2p}(\Omega; H^s)$, for any $p \geq 1$. This result is used to also show that the truncated numerical approximation converges to the truncated exact solution almost surely with order $\frac{1}{2}$.

The results for the truncations are used to prove our first main result: almost sure convergence of the numerical solution (2.9) to the exact solution (2.7). We prove that, for $\delta < 1$ and almost all $\omega$, there is a constant $C(\omega)$ such that

$$\max_{n=0,\ldots,N_{\tau^\ast}} \|u^n - u(t_n)\|_s \leq C(\omega) \cdot (\Delta t)^{\delta/2},$$
where we recall that the norm is the $H^s(\mathbb{R}^d)$-norm. This result is obtained via a proof by contradiction argument and other ideas from [16].

The almost sure convergence and other auxiliary results are used to prove our second main result. We prove convergence in probability of the numerical approximation (2.9) to the exact solution (2.7) with order $\frac{1}{2}$:

$$
\lim_{C \to \infty} \lim_{\Delta t \to 0} \mathbb{P}\left( \max_{n=0, \ldots, N_\tau} \| u^n - u(t_n) \|_s \geq C \cdot (\Delta t)^{1/2} \right) = 0. \quad (2.11)
$$

Much like in Paper III, we illustrate the almost sure convergence through phase plots. For two different problems (one with $\lambda = 1$ and one with $\lambda = -1$) we plot one realization of the real and imaginary part of a reference solution and the numerical approximation for different time steps. We observe that as the time step becomes smaller, the numerical approximations approach the reference solution.

We also illustrate the convergence in probability. This is accomplished by calculating the norm of the error for many simulations and counting the number of samples that satisfy

$$
\max_{n=0, \ldots, N_\tau} \| u^n - u(t_n) \|_s \geq C \cdot (\Delta t)^\delta
$$

for different values of $C$, $\Delta t$ and $\delta$. We observe that when $\delta \leq \frac{1}{2}$ this number goes to zero as $\Delta t$ becomes smaller and $C$ becomes larger, which is what we expect should happen according to (2.11).

References


